Optimal control and dependence modeling of portfolios with Lévy dynamics

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

der Fakultät für Mathematik,
Universität Karlsruhe (TH)

vorgelegte

DISSERTATION

von

Dipl.-Math. oec. Anja Bettina Blatter
aus Biberach

Tag der mündlichen Prüfung: 4. Februar 2009

Referentin: Prof. Dr. Nicole Bäuerle
Korreferent: Prof. Dr. Alfred Müller

Karlsruhe
2008

Preface

This thesis mainly deals with dependence concepts and their impact on stochastic optimal control in multidimensional Lévy-driven insurance models. It concludes my doctorate research which was carried out at the Institute for Stochastics at Universität Karlsruhe in the period from April 2006 to November 2008. My work was supervised by Professor Dr. Nicole Bäuerle, Universität Karlsruhe.

Acknowledgements

First of all, I would like to express my sincere appreciation to Professor Dr. Nicole Bäuerle for the opportunity to write this thesis at the Institute for Stochastics. I am very grateful for her excellent guidance and valuable suggestions during the three years of my doctoral research. She is an excellent supervisor and teacher.

Special thanks go to Professor Dr. Alfred Müller for his readiness of being the second examiner. He provided valuable assistance and helpful comments whenever asked for advise.

Moreover, I very much enjoyed to contribute to the teaching duties as well as the social life of the institute. Thanks to all my colleagues at the Institute of Stochastics who made the last three years an unforgettable part of my life.

Finally, I want to express my special gratitude to Martin Riesner and my parents for their never ending support and encouragement.

Karlsruhe, November 2008

Anja Blatter

Contents

Chapter 1

Introduction and summary

Itô (cf. Itô (1998)) describes “mathematical beauty” in very nice metaphors.

> *“In precisely built mathematical structures, mathematicians find the same sort of beauty others find in enchanting pieces of music, or in magnificent architecture. There is, however, one great difference between the beauty of mathematical structures and that of great art. Music by Mozart, for instance, impresses greatly even those who do not know musical theory; the cathedral in Cologne overwhelms spectators even if they know nothing about Christianity. The beauty in mathematical structures, however, cannot be appreciated without understanding of a group of numerical formulae that express laws of logic. Only mathematicians can read ”musical scores” containing many numerical formulae, and play that ”music” in their hearts.”*

1.1 Historical overview and existing literature

Financial and actuarial applications often require multivariate models with jumps allowing for dependence between the univariate components. Before going into more detail we consider the historical development. Jump processes have been successfully used in univariate insurance models. The collective risk model of an insurance company has already been introduced at the beginning of the twentieth century by Filip Lundberg. Later, Harald Cramér enhanced Lundberg’s ideas and formulated in an intuitive way a model for the random variation of an insurance company’s surplus which is today known as a compound Poisson process. In recent years the classical Cramér-Lundberg model has been generalized in many directions. One way to generalize the classical risk process is to consider an arbitrary spectrally negative Lévy process. Recently, this model has attracted a lot of research interest and has been successfully used to construct one-dimensional insurance models, see for example Albrecher et al. (2008) and the references given therein. However, multivariate applications are still dominated by Brownian motion. In this thesis we therefore focus on multidimensional insurance models driven by jump processes. To be more precise, we consider in Chapter 2 a multidimensional extension of the classical Cramér-Lundberg model observed at discrete-time points and in Chapters 5 and 6, a multidimensional insurance model driven by a general Lévy process.

As mentioned before, in reality risk processes of an insurance company are often dependent. Let us think of large insurance groups operating in various countries around the world or offering different types of insurance coverage. One might easily think of instances which may generate dependent claims of the single business lines. Simplified examples for such possible dependencies are a car accident or the occurrence of hurricanes. A car accident may cause a claim for the motor liability and health insurance. Hurricanes might cause losses in different countries. To model this appropriately, multidimensional models capturing complex dependency structures are necessary. Bäuerle and Grübel (2005) suggest to use models incorporating thinning and shifts such that events of a background Poisson process trigger later claims in different categories. Lindskog and McNeil (2003) consider dependence of individual jumps of compound Poisson processes assuming that losses result from underlying shock processes. However, this is only possible if there are only a few sources of risk causing jumps. In Pfeifer and Nešlehová (2004) modeling dependence is entirely based on the total number of claims in a finite time interval described by a static multidimensional copula. Whereas Tankov (2003) suggests to model the dependence structure of spectrally positive Lévy processes via Lévy copulas. This concept is extended by Kallsen and Tankov (2006), see also the book of Cont and Tankov (2004). Advantages of modeling dependence of Lévy processes via Lévy copulas instead of ordinary copulas are summarized in the introduction to Chapter 4.

A fundamental problem in actuarial mathematics deals with optimal risk control and optimal investment in a financial market. The surplus process of an insurance company consists of premium income and the payment arising through claims. An insurance company can reduce its risk by ceding claims to a reinsurance company which of course reduces in return the net premium income. To derive an optimal choice on the amount to reinsure and the investments to make, methods from stochastic control theory may be used. Various optimality criteria are used to formulate the optimization problem of risk control such as minimizing the probability of ruin and maximizing expected exponential utility of terminal wealth. Browne (1995) showed that, in a univariate diffusion model, minimizing the probability of ruin is equal to maximizing the expected exponential utility of terminal wealth which is often used as performance criterion in applications. Since the expected exponential utility function is more tractable to work with, it is widely spread.

Another characteristic of multidimensional insurance models is that there are several possibilities to define the ruin probability. In the most general case, the ruin probability is defined as the probability that the multivariate risk process of the insurance company, eventually hits the so-called insolvency region $\mathcal{A}$, some domain of $\mathbb{R}^d$. If the insurance company only consists of one business line the insolvency region is generally chosen to be $\mathcal{A} = (-\infty, 0)$. Unlike the univariate setting, the concept of ruin for multivariate processes could have various interpretations. The least restrictive and at the same time the least realistic one would be to require that all business lines are simultaneously below zero, i.e. $\mathcal{A} = \{y \in \mathbb{R}^d : \max(y^1, \dots, y^d) < 0\}$. A more restrictive choice is to take $\mathcal{A} = \{y \in \mathbb{R}^d : y^1 + \ldots + y^d < 0\}$, which corresponds to the classical univariate ruin problem for the global company. Thus ruin occurs when the sum of net values of the business units are negative. But as capital is not completely versatile between different business lines, ruin may already occur before the aggregated reserves become negative. That is, ruin might occur when at least one business gets ruined and the solvency region is chosen to be $\mathcal{A} = \{y \in \mathbb{R}^d : \min(y^1, \dots, y^d) < 0\}$. More generally, Hult and Lindskog (2006) propose that capital may be transferred between the business lines. They define ruin as the situation when negative positions in one or several lines of business cannot be balanced by capital transfer.

There is quite a lot of literature on optimal control of ruin probabilities as well as exponential utility of terminal wealth for the univariate setting. For models in discrete-time we only mention Schäl (2004) and Schäl (2005) which serves as a basis for Chapter 2. Optimal reinsurance and investment problems in a continuous-time insurance model can be for example found in Højgaard and Taksar (1998), Schmidli (2001), Schmidli (2002), the book of Schmidli (2008) and Fernández et al. (2008). However, literature on multivariate risk processes is relatively sparse. Collamore (1996) considers in discrete time a multidimensional problem of first passage of a process into a general region and Collamore (2002) simulates this probability. Recently, ruin estimation in multivariate models with heavy-tailed claim sizes and capital transfer between single business lines has been considered by Hult and Lindskog (2006) and Bregman and Filipovic (2006). However, methods which allow to reduce the multidimensional problem to a univariate problem are mainly used. For example, Bregman and Klüppelberg (2005) consider the sum of single business line risk reserves and model dependence via a Clayton Lévy copula to estimate the ruin probability. Chan et al. (2003) define several types of ruin probabilities and obtain bounds for the two-dimensional ruin probabilities using univariate model results. A certain bivariate ruin problem is solved in recent publications by Avram et al. (2008b) and Avram et al. (2008a). They allocate claims according to some fixed proportion between the two business lines. However, they again reduce this problem to a one-dimensional setting.

1.2 Summary and outline

The main aim of this thesis is to deal with dependence concepts and their impact on stochastic optimal control in multidimensional insurance models allowing for jumps in the claim processes and in the investment portfolio.

Discrete-time multivariate risk processes

In Chapter 2 we develop a multidimensional insurance model in which the risk processes of individual business lines can be controlled by reinsurance and investment in a financial market taking into account the special features of multidimensional insurance models. That is, we allow for stochastically dependent development of the risk reserves and for the most general ruin definition in assuming that ruin occurs when the corresponding risk reserve first hits the insolvency region in $\mathbb{R}^d$.
After providing a short introduction to the theory of discrete-time stochastic programming in Section 2.1 we consider in Section 2.2 our multidimensional insurance model. In Sections 2.3 and 2.5 we apply the theory of dynamic programming to maximize expected exponential utility of terminal wealth and to minimize the probability of ruin, respectively. We prove in Section 2.3 that the optimal control of the exponential utility maximization of terminal wealth neither depends on the time nor on the present state of the risk reserve. Modeling dependence between the individual business lines by an Archimedean copula we identify structure conditions of the Archimedean generator under which an insurance company certainly reinsures a larger fraction of claims from one business line than from another. Similarly, structure conditions are derived for the investment portfolio. These results can be found in Section 2.4.
Unlike the exponential utility function, the ruin probability does not resemble the well-known cost criterion of control theory at first glance. However, results from stochastic dynamic programming can be applied in the context of our insurance model by interpreting the ruin probability as total cost without discounting which have to be paid entering the insolvency region. This is similar to Schäl (2004) and Schäl (2005). In Section 2.5 the ruin probability can then be described as a fixed point of a contractive operator and can be approximated using an iteration method.

Basic concepts of Lévy processes

In Chapter 3 we discuss essential tools and concepts of the theory of Lévy processes which are beneficial for Lévy driven risk processes in an insurance model. Beginning with the definition of a Lévy process we subsequently discuss the Lévy measure, the famous Lévy-Itô-decomposition that describes the structure of its sample paths, and the distributional properties of Lévy processes such as the Lévy-Khinchin formula and infinite divisibility. Moreover, we classify the paths behavior of Lévy processes distinguishing between finite and infinite activity and variation. Section 3.3 contains a discussion and actuarial interpretation of examples of Lévy processes i.a. a compound Poisson process, a jump diffusion and a spectrally negative Lévy process since they are commonly used in risk process modeling. In Section 3.4 we recall the martingale and Markov property of Lévy processes and treat the infinitesimal generator of Lévy processes in terms of its characteristic triplet which we need in Chapter 5. The Itô-Doeblin formula, until recently known as Itô-formula, is presented in Section 3.5. However, we do not present its most general version. Instead we restrict ourselves to a version for Lévy-type stochastic integrals which is frequently applied in this thesis. Finally, in Section 3.6 we recall conditions under which existence and uniqueness of solutions of Lévy stochastic differential equations can be guaranteed. These results enable us to derive a maximum inequality for a certain power of the solution of a Lévy stochastic differential equation which is beneficial for the correct formulation of the stochastic control problem in Chapter 5. So far, such a maximum inequality only exists for the pure diffusion case.

Dependence concepts for Lévy processes

An intuitive approach to model dependence of Lévy processes might be the use of copulas for random vectors. However, as we will see in Chapter 4 there are a few drawbacks which make this concept not suitable for Lévy processes. Therefore, in order to characterize the dependence structure of a multivariate Lévy measure the concept of Lévy copulas was recently introduced in Cont and Tankov (2004) and further refined in Kallsen and Tankov (2006). However, caution has to be exercised since Lévy copulas can only be used with respect to some dependence properties to characterize dependence among the univariate components of multidimensional Lévy processes (cf. Bäuerle et al. (2008)).
We start this chapter briefly recalling some general definitions which will be important in the following sections i.a. the notion of d-increasing functions is introduced. In the literature different definitions for d-increasing functions exist. However, we show that the different definitions only differ by an additional condition. We then discuss the conceptually simpler case of spectrally positive Lévy processes (cf. Cont and Tankov (2004)) before treating the concept for general Lévy processes introduced in Kallsen and Tankov (2006). A special class of Lévy copulas known as Archimedean Lévy copulas is our main focus. Archimedean Lévy copulas can be constructed quite easily and at the same time possess a lot of nice properties. For that reason

Archimedean Lévy copulas are widely spread in applications. In Subsection 4.2.2 we derive a sufficient and necessary condition for an Archimedean Lévy generator to create a multidimensional positive Lévy copula in arbitrary dimension. So far, this has only been analyzed for a bivariate Lévy copula, while for dimensions larger than two there is only a sufficient condition. It turns out that the necessary and sufficient condition derived in Theorem 4.2.10 contains the existing results as special cases. Finally, we recall the construction of a general Archimedean Lévy copula introduced in Bäuerle et al. (2008).
Describing the dependence structure of a multidimensional Lévy process in terms of its Lévy copula allows us to quantify the effect of dependence on the retention levels and the investment portfolio in our multidimensional Lévy driven insurance model (cf. Chapter 6 and Chapter 5).

Stochastic control of portfolios with Lévy-dynamics

We consider the optimization of proportional reinsurance and investment strategies in a multidimensional Lévy-driven insurance model with dependent claims and dependent investments of the insurance company's single business lines. Before entering the world of our insurance model driven by Lévy dynamics in Section 5.2, we treat Lévy process stochastic control theory in Section 5.1. As far as we know, there is so far only a very short introduction to stochastic control with respect to jump diffusions provided by Øksendal and Sulem (2007). Inspired by Browne (1995) who discovered that in a one-dimensional diffusion model the control which maximizes expected exponential utility of terminal wealth also minimizes the ruin probability, we consider in Section 5.3 the optimization criterion that maximizes exponential utility of terminal wealth. Imbedding this problem in stochastic control theory and solving the Hamilton-Jacobi Bellman equation we can show that it is optimal to keep the retention level and investment portfolio constant regardless of the time and the company's wealth level. This does not only hold for proportional reinsurance but also for general reinsurance as well as for a mixture of proportional and excess of loss reinsurance in a slightly modified model. In the latter case we can even show that there exists a pure excess of loss policy which is always better than any combined reinsurance policy. We conclude in Section 5.4 with a validation of the conjecture that the policy which maximizes utility of terminal wealth also minimizes the ruin probability in our multidimensional reinsurance model. This holds omitting claims caused by jumps while assuming that ruin occurs when the weighted sum of net values of the business lines become negative.

Structural comparison results

This chapter is dedicated to comparison results with respect to jumps in an insurance company's risk reserve. The performance criterion is still the expected exponential utility of terminal wealth. In the first section the optimal control in a compensated jump diffusion model is compared with the correspondent results in a pure diffusion model. Not surprisingly, it turns out that the optimal retention level of the insurance company is larger in the model without jumps than it is in the model containing jumps. In Section 6.2 we weaken the difference between the models and consider only models differing by the weighting of claims caused by jumps. There, we show that the accumulated risk reserve increases in concave order as the weighting factor of the claims decreases. However, the main focus of this chapter is put on the last section. Based on the results of Chapter 5 we devote Section 6.3 to identify structure conditions with respect to the Archimedean generator and the Lévy measure under which an insurance company reinsures a larger fraction of claims from one business line than from another. Similar results can be obtained with respect to investments on a financial market.

1.3 Notation

This section consists of some basic notation, abbreviations and conventions used throughout this thesis.

$x+y$, $x \leq y$	are componentwise operations for $x, y \in \mathbb{R}^d$
(x,y)	$= \sum_{i=1}^d x^i y^i$ inner product in $\mathbb{R}^d$ where $x, y \in \mathbb{R}^d$
$\|x\|$	$= (x,x)^{\frac{1}{2}}$ Euclidean norm where $x \in \mathbb{R}^d$
$\|x\|_p$	$= \left(\sum_{i=1}^d \lvert x^i \rvert^p\right)^{\frac{1}{p}}$ p-norm where $x \in \mathbb{R}^d$
$\|M\|_{sn}$	$= \sum_{i=1}^d \|M_i^i\|$ matrix semi-norm for $M \in \mathbb{R}^d \times \mathbb{R}^d$
$\mathrm{sgn}(x)$	$\mathrm{sgn}(x) = 1$ if $x \geq 0$ and $\mathrm{sgn}(x) = -1$ if $x < 0$
C	set of continuous functions
C^d	set of all d-times continuously differentiable functions
$C^{1,2}$	set of continuous functions which are continuously differentiable with respect to the first variable and twice continuously differentiable with respect to the second variable
C_0	set of continuous functions vanishing at infinity
C_0^2	set of twice continuously differentiable functions, vanishing at infinity
$f_t(t,x)$, $f_{x^i}(t,x)$	partial derivatives with respect to t and x^i where $x \in \mathbb{R}^d$ and $i \in \{1,\dots,d\}$
Ran f	range of a function f
Dom f	domain of a function f
∂B	the boundary of the set B
$\overline{B}$	the closure of the set B
int B	the open set of B
B^c	the complement of the set
l.s.c.	abbreviation for lower semi-continuous
i.i.d.	abbreviation for independently and identically distributed
δ_x	Dirac measure given by $\delta_x(A) = \mathbb{1}_{x \in A}$ for every Borel set A
$\mathcal{B}(X)$	Borel σ-algebra, i.e. the smallest σ-algebra of subsets of X which contains all the open subsets of X
$\overline{\mathbb{R}}$	$= \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$, the extended real line
$\langle X \rangle$	the quadratic variation of X

Chapter 2

Discrete-time multivariate risk processes

In this chapter we develop a multidimensional insurance model in which the risk processes of individual business lines can be controlled by reinsurance and investment in a financial market. We take into account the special features of multidimensional insurance models described in Chapter 1. That is, we allow for dependent risk reserves which are determined by claims, premiums, and the financial market result. There are strong reasons to believe that claims of different types are stochastically dependent. A simplified example for such possible dependencies is a car accident. On the one hand vehicles might be damaged, on the other hand car occupants might be injured. Thus, two business lines have to bear the damage, the motor insurance and the accident-and health insurance. In addition to that, stock returns on the financial market can depend on each other as well.
Another characteristic of multidimensional insurance models is to have several different possibilities to define ruin. We allow for a quite general definition in assuming that ruin occurs when the corresponding risk reserve hits first the so-called insolvency region, some domain of $\mathbb{R}^d$.

After a short introduction to the theory of discrete-time stochastic programming in Section 2.1 we consider our multidimensional insurance model in Section 2.2. In Sections 2.3 and 2.5 we then apply dynamic programming to maximize the expected exponential utility of terminal wealth and to minimize the probability of ruin, respectively. The main result of Section 2.3 shows that the optimal control of terminal wealth utility maximization neither depends on the time nor on the present state of the risk reserve. Modeling dependence between the individual business lines by an Archimedean copula we identify structure conditions of the Archimedean generator under which an insurance company certainly reinsures a larger fraction of claims from one business line than from another. Similarly, structure conditions

are derived for the investment portfolio. These results can be found in Section 2.4.
Unlike the exponential utility function, the ruin probability does not resemble the well-known cost criterion of control theory at first glance. However, results from stochastic dynamic programming can be applied in the context of our insurance model by interpreting the ruin probability as total cost without discounting which has to be paid entering the insolvency region. This is similar to Schäl (2004) and Schäl (2005). In Section 2.5 the ruin probability can then be described as a fixed point of a contractive operator and approximated using an iteration method.

2.1 Stochastic dynamic programming

We give a short introduction to the theory of stochastic discrete-time dynamic programming. The main purpose of this section is to state terms, definitions as well as results used in the sections to come. A more detailed introduction for a finite and an infinite horizon problem can be found in Bertsekas and Shreve (1978) or in Schäl (2004) on which this section is mainly based. A broad list of excellent literature providing a thorough treatment of discrete stochastic dynamic programming including a historical overview can be found in Puterman (1994).
Our setting of treating the ruin probability control in Section 2.5, however, requires to consider some special features. For example, there is a discontinuity of the system function forcing us to use a so-called structure assumption instead of the usual continuity assumption.

2.1.1 The finite horizon model

A (stationary) finite horizon stochastic optimal control model with horizon $N \in \mathbb{N}$ is a tuple $(S, A, D, E, F^{tr}, g, V_0, \alpha)$ consisting of

1. a state space S which is a nonempty subset of some Euclidean space where $(S, \mathcal{B}(S))$ is a measurable space;

2. an action space A which is a nonempty subset of some Euclidean space where $(A, \mathcal{B}(A))$ is a measurable space;

3. a control constraint $D \subset S \times A$ which is $\mathcal{B}(S) \otimes \mathcal{B}(A)$-measurable and

$$D(x) := \{a \in A : (x, a) \in D\} \neq \emptyset \quad \text{for all } x \in S$$

   is the set of feasible control actions in state $x \in S$.
   For a measurable function $\varphi : S \to A$ we assume that $(x, \varphi(x)) \in D$ for all $x \in S$;

4. a disturbance space E which is a nonempty subset of some Euclidean space where $(E, \mathcal{B}(E))$ is a measurable space;
5. a measurable system function $F^{tr} : D \times E \to S$;
6. a measurable function $g : S \times A \to (-\infty, \infty]$ which is bounded from below, called the cost-per-stage function;
7. a measurable function $V_0 : S \to (-\infty, \infty]$ which is bounded from below, called terminal cost function;
8. a discount factor $\alpha \in [0, 1]$.

The system transition in this model is determined by so-called disturbance variables $(W_n)_{1 \leq n \leq N}$, a sequence of independent and identically distributed random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in the disturbance space $(E, \mathfrak{E})$.

Definition 2.1.1. *(a) A decision function is a measurable function $\varphi : S \to A$ such that $\varphi(x) \in D(x)$ for all $x \in S$. The set of all decision functions is denoted by Φ.*

(b) A policy is a sequence $\pi = (\varphi_n)_{0 \leq n \leq N-1}$ of decision functions $\varphi_n \in \Phi$.

A stochastic dynamic programme can be interpreted as follows. A decision maker is faced with the opportunity of influencing the behavior of a probabilistic system as it evolves over time. Observing the system in state $x_n \in S$ and choosing an action $a_n = \varphi_n(x_n)$ from the set of allowable actions $D(x_n)$ in state x_n, the decision maker has to face costs $g(x_n, a_n)$ and the system state moves to state $x_{n+1} = F^{tr}(x_n, a_n, w_{n+1})$ influenced by a disturbance $W_{n+1} = w_{n+1}$. Repeating this procedure up to time N the cost incurred in state x_N is given by $V_0(x_N)$. We seek a finite sequence of control functions $\pi = (\varphi_0, \ldots, \varphi_{N-1})$ which minimizes the total cost over N stages.

Definition 2.1.2. *(a) The total discounted cost over N periods starting in $X_0 = x \in S$ and choosing policy π is*

$$V_{N\pi}(x) = \mathbb{E}\Big[\sum_{n=0}^{N-1} \alpha^n g(X_n, \varphi_n(X_n)) + \alpha^N V_0(X_N) \Big].$$

(b) The minimal total discounted cost over N periods starting in $X_0 = x \in S$ and choosing policy π is

$$V_N(x) = \inf_{\pi} V_{N\pi}(x).$$

(c) A policy π is said to be optimal if

$$V_{N\pi}(x) = V_N(x) \quad \textit{for all } x \in S.$$

The next lemma shows that the total discounted cost can be computed recursively.

Lemma 2.1.3 (Cost iteration). *For* $\pi = (\varphi_0, \varphi_1, \ldots, \varphi_{N-1})$ *and* $\pi^{\leftarrow} = (\varphi_1, \ldots, \varphi_{N-1})$ *it holds*

$$V_{N\pi}(x) = g(x, \varphi_0(x)) + \alpha\, \mathbb{E}\big[V_{N-1\pi^{\leftarrow}}\big(F^{tr}(x, \varphi_0(x), W)\big)\big].$$

Proof. This follows from Definition 2.1.2. □

It is useful to write this recursion formula in terms of operators.

Definition 2.1.4. *For any* $v : S \to (-\infty, \infty]$ *measurable and bounded from below we define*

(a) $Lv(x,a) = g(x,a) + \alpha\, \mathbb{E}\big[v\big(F^{tr}(x,a,W)\big)\big]$ *for* $(x,a) \in D$*;*

(b) $U_\varphi v(x) = Lv(x, \varphi(x))$ *for any decision function* φ*;*

(c) $Uv(x) = \inf_{a \in D(x)} Lv(x,a)$*.*

It can be shown that these operators are order preserving and costs corresponding to a policy $\pi = (\varphi_0, \ldots, \varphi_{N-1})$ can be defined in terms of the composition of operators $U_{\varphi_0} \ldots U_{\varphi_{N-1}}$.

Lemma 2.1.5. (a) *Let* $v, w : S \to (-\infty, \infty]$ *be measurable and bounded from below and* $v(x) \le w(x)$ *for all* $x \in S$*. This implies*

$$\begin{aligned} U_\varphi v(x) &\le U_\varphi w(x) \quad \text{for all } x \in S,\ \varphi(x) \in D(x), \\ Uv(x) &\le Uw(x) \quad \text{for all } x \in S. \end{aligned}$$

(b) *Let* $\pi = (\varphi_0, \varphi_1, \ldots, \varphi_{N-1})$ *and* $\pi^{\leftarrow} = (\varphi_1, \varphi_2, \ldots, \varphi_{N-1})$ *then*

$$V_{N\pi}(x) = U_{\varphi_0} V_{N-1\pi^{\leftarrow}}(x) = U_{\varphi_0} \ldots U_{\varphi_{N-1}} V_0(x).$$

Proof. (a) Let $v(x) \le w(x)$ for all $x \in S$. Then

$$\begin{aligned} U_\varphi v(x) =& g(x, \varphi(x)) + \alpha\, \mathbb{E}\big[v\big(F^{tr}(x, \varphi(x), W)\big)\big] \\ \le& g(x, \varphi(x)) + \alpha\, \mathbb{E}\big[w\big(F^{tr}(x, \varphi(x), W)\big)\big] = U_\varphi w(x). \end{aligned}$$

Taking the infimum on both sides the second statement follows.

(b) The proof can be found in Bertsekas and Shreve (1978), Lemma 8.1. □

Note that it is not clear if Uv is measurable. However, we avoid the measurability questions by the following assumption.

Structure assumption
There exists a set $\mathcal{V}$ of measurable functions $v : S \to (-\infty, \infty]$ which are bounded from below and a set Φ of decision functions such that

(i) $V_0 \in \mathcal{V}$;

(ii) $Uv \in \mathcal{V}$ for all $v \in \mathcal{V}$;

(iii) For all $v \in \mathcal{V}$ there exists a decision function $\varphi \in \Phi$ such that

$$U_\varphi v = Uv.$$

Definition 2.1.6. *Let $v : S \to (-\infty, \infty]$ be measurable and bounded from below. A decision function $\varphi \in \Phi$ is called a minimizer of v if*

$$U_\varphi v(x) = Uv(x) \quad \text{for all } x \in S.$$

The following theorem states the main result of this section. Beginning with the cost V_0 and applying the operator U successively N times the optimal cost function can be generated. Furthermore, the construction of an optimal policy is given.

Theorem 2.1.7 (Main Theorem)**.** *Suppose that the structure assumption holds. We then have*

(a) The value iteration

$$V_n = UV_{n-1} = U^n V_0 \in \mathcal{V} \text{ for all } n = 1, \ldots, N;$$

(b) The optimal policy
If φ_n is a minimizer of V_{n-1} for all $n = 1, \ldots, N$, then the policy $\pi := (\varphi_N, \ldots, \varphi_1)$ is optimal;

(c) Existence
There exists an optimal policy π.

Proof. The structure assumption directly implies that $U^n V_0$ is well-defined. The proof of the value iteration, the existence and optimality of the policy can be found in Bertsekas and Shreve (1978) Proposition 8.2, Corollary 8.2.1 and Proposition 8.5. □

2.1.2 The infinite horizon model

The infinite horizon stochastic optimal control model is as described in Section 2.1.1 differing only by the horizon which is infinite. One approach to solve infinite horizon stochastic dynamic programmes is to approximate them with finite horizon problems. In doing so we have to guarantee that the sequence of value functions generated by the successive application of the finite horizon value iteration converges to the optimal value function of the infinite horizon problem. For this purpose we can, for example, demand a contraction property or impose a monotonicity assumption. Since it is more convenient for our application we stick to the latter case and the following assumption will hold in this section (cf. Bertsekas and Shreve (1978)).

Uniform increase assumption
It holds

$$LV_0(x,a) \geq V_0(x) \quad \text{for all } x \in S \text{ and } a \in D(x).$$

It is common practice to define $V_0 \equiv 0$ in an infinite horizon model. However, in our application to ruin probabilities we do not want 0 to be in the set $\mathcal{V}$ and therefore assume $V_0 = g$ for some $g \neq 0$.

Remark 1. By the uniform increase assumption and Lemma 2.1.5(b) we get

$$V_{n+1\pi} = U_{\varphi_0} \dots U_{\varphi_{n-1}}(U_{\varphi_n} V_0) \geq U_{\varphi_0} \dots U_{\varphi_{n-1}} V_0 = V_{n\pi}$$

implying monotonicity of the value function, that is for all $n \in \mathbb{N}$

$$V_0 \leq V_{n\pi} \leq V_{n+1\pi} \quad \text{and} \quad V_0 \leq V_n \leq V_{n+1}.$$

By the preceding remark the limit of the sequence of finite horizon value functions is well defined and we are now in the position to define the performance criterion.

Definition 2.1.8. *(a) The total expected cost starting in $X_0 = x$ and choosing policy π is*

$$V_\pi(x) = \lim_{n\to\infty} V_{n\pi}(x).$$

(b) The minimal total expected cost starting in $X_0 = x$ is

$$V(x) = \inf_\pi V_\pi(x).$$

$V(x)$ is called the value function of the infinite horizon problem.

(c) A policy π is optimal if $V_\pi(x) = V(x)$ for all $x \in S$.

(d) The policy $(\varphi, \varphi, \dots)$ where $\varphi \in \Phi$ is called stationary and we write $\pi = \varphi^\infty$.

As in the previous section we write $V_n = \inf_\pi V_{n\pi}$ for $n \in \mathbb{N}$ and define the so-called limit function by

$$V_\infty(x) = \lim_{n\to\infty} V_n(x)$$

which is again well-defined by Remark 1.

Lemma 2.1.9. *(a) Let $\pi = \varphi^\infty$ be a stationary policy. Then*

$$U_\varphi V_{\varphi^\infty} = V_{\varphi^\infty} = \lim_{n\to\infty} (U_\varphi)^n V_0.$$

(b) *It holds that*

$$V_0 \leq V_\infty(x) \leq V(x).$$

Proof. (a) We know that

$$V_{n(\varphi,\ldots,\varphi)} = U_\varphi(U_\varphi \ldots U_\varphi V_0) = (U_\varphi)^n V_0.$$

The assertion follows by the monotone convergence theorem.

(b) We have

$$V_\pi(x) = \lim_{n\to\infty} V_{n\pi}(x) \geq \lim_{n\to\infty} V_n(x) = V_\infty(x) \quad \text{for all } \pi.$$

Thus

$$V(x) = \inf_\pi V_\pi(x) \geq V_\infty(x).$$

□

Analogously to the finite horizon model we need a structure assumption ensuring measurability.

Structure assumption
There exists a set $\mathcal{V}$ of measurable functions $v : S \to (-\infty, \infty]$ which are bounded from below such that

(i) $V_0 \in \mathcal{V}$;

(ii) $Uv \in \mathcal{V}$ for all $v \in \mathcal{V}$;

(iii) $D(x)$ is a compact metric space for all $x \in S$;

(iv) $Lv(x, a)$ is lower semi-continuous in a for all $x \in S$ and $v \in \mathcal{V}$.

Remark 2. Brown and Purves (1973) (cf. Corollary 1) show that there exists a decision function φ such that $U_\varphi v = Uv$ for any $v \in \mathcal{V}$. Note that under the structure assumption the conditions required for Corollary 1 in Brown and Purves (1973) are fulfilled. This means that the structure assumption for the infinite horizon model implies the structure assumption for the finite horizon model.

Let us finally turn to the main theorem for the infinite horizon model.

Theorem 2.1.10 (Main theorem)**.** *Suppose that the structure assumption holds. We then have*

(a) *The value iteration*

$$V = \lim_{n\to\infty} U^n V_0;$$

(b) The optimality equation

$$V = UV;$$

(c) The optimality criterion
If φ is a minimizer of V then the stationary policy φ^∞ is optimal;

(d) Existence
There exists a stationary optimal policy.

Proof. The finite horizon structure assumption follows directly from the structure assumption for the infinite horizon case (cf. Remark 2). Thus, by the main theorem for the finite horizon model (cf. Theorem 2.1.7), we have $V_n \in \mathcal{V}$ for all $n \in \mathbb{N}$.
Fix $x \in S$ and define $v_n(a) := LV_{n-1}(x, a)$ for all $a \in D(x)$. By the structure assumption we know that $v_n(a)$ is lower semi-continuous in a. Further by Remark 1 and Lemma 2.1.5(a) we have that $v_n(a)$ is increasing in n. By use of a variant of Dini's theorem (cf. Schäl (1975), Proposition 10.1) we obtain

$$\lim_{n\to\infty} \inf_{a\in D(x)} v_n(a) = \inf_{a\in D(x)} \lim_{n\to\infty} v_n(a).$$

Applying Theorem 2.1.7 it yields

$$\begin{aligned} V_\infty(x) &= \lim_{n\to\infty} V_n(x) = \lim_{n\to\infty} UV_{n-1}(x) = \lim_{n\to\infty} \inf_{a\in D(x)} LV_{n-1}(x,a) \\ &= \inf_{a\in D(x)} \lim_{n\to\infty} LV_{n-1}(x,a) = \inf_{a\in D(x)} LV_\infty(x,a) = UV_\infty(x). \end{aligned} \tag{2.1}$$

Note that $\lim_{n\to\infty} LV_n(x,a) = LV_\infty(x,a)$ follows directly from the monotone convergence theorem since the value function is increasing (cf. Remark 1). We further know that the limit of an increasing sequence of lower semi-continuous functions, $\lim_{n\to\infty} v_n(a) = LV_\infty(x,a)$, is lower semi-continuous. By the selection theorem of Brown and Purves (1973) (cf. Remark 2) there exists a decision function φ such that

$$U_\varphi V_\infty(x) = UV_\infty(x) \tag{2.2}$$

for all $x \in S$. Thus together with (2.1)

$$V_\infty(x) = U_\varphi V_\infty(x) \tag{2.3}$$

for all $x \in S$. This implies

$$V_\infty(x) = U_\varphi^n V_\infty(x) \geq U_\varphi^n V_0(x).$$

Therefore, by Lemma 2.1.9 (a) we have

$$V_\infty(x) \geq \lim_{n\to\infty} U_\varphi^n V_0(x) = V_{\varphi^\infty}(x)$$

for all $x \in S$. On the other hand we have by Lemma 2.1.9 (b)

$$V_\infty(x) \leq V(x) \leq V_{\varphi^\infty}(x)$$

for all $x \in S$. Thus parts (a), (b) and (d) follow immediately. Let us finally turn to part (c). Since φ is a minimizer of V it holds that $U_\varphi V(x) = UV(x)$ for all $x \in S$ and we therefore obtain

$$V(x) = U_\varphi V(x) = U_\varphi^n V(x) \geq U_\varphi^n V_0(x).$$

Hence for $n \to \infty$ it yields

$$V(x) \geq \lim_{n\to\infty} U_\varphi^n V_0(x) = V_{\varphi^\infty}(x)$$

for all $x \in S$ and hence (c) follows. □

The optimality equation in Theorem 2.1.10 is a so-called fixed-point equation which can be solved in our model by iteration without a contraction property. Another approach to obtain the value function V is the Howard improvement theorem.

Theorem 2.1.11 (Howard improvement)**.** *Let φ and $\tilde{\varphi}$ be any decision functions. Define $J := V_{\varphi^\infty}$, $\tilde{J} := V_{\tilde{\varphi}^\infty}$,*

$$A(x,\varphi) := \big\{a \in A : LJ(x,a) < J(x)\big\} \text{ for } x \in S$$

and $S^* := \{x \in S : A(x,\varphi) \neq \emptyset\}$. Let $\tilde{\varphi}(x) \in A(x,\varphi)$ for $x \in S^*$ and $\tilde{\varphi}(x) = \varphi(x)$ for $x \notin S^*$. If

$$\lim_{n\to\infty} \alpha^n \, \mathbb{E}[(J - V_0)(X_n)] = 0 \quad \text{for all } x \in S \tag{2.4}$$

then

$$\begin{aligned} \tilde{J}(x) \leq& J(x) \quad \text{for all } x \in S \text{ and} \\ \tilde{J}(x) <& J(x) \quad \text{for all } x \in S^*. \end{aligned}$$

Note that assumption (2.4) is always satisfied in the discounted case. The next theorem states that if there is no improvement anymore, which means that $A(x,\varphi) = \emptyset$, then we already have the value function and the optimal policy.

Theorem 2.1.12 (Verification theorem)**.** *Let $v : S \to (-\infty,\infty]$ be a measurable function with $v \geq V_0$ and*

$$v = Uv = U_\varphi v.$$

If (2.4) holds then the function v corresponds to the value function V and φ defines a stationary optimal policy φ^∞.

A proof of Theorem 2.1.11 and Theorem 2.1.12 can be found in Schäl (2004).

2.2 Multidimensional insurance model

In this section we introduce our multidimensional insurance model where the surplus process of the single business lines can be controlled by reinsurance and by investment in a financial market. Our construction is based on the one-dimensional insurance model of Schäl (2004) and Schäl (2005). Considering a multidimensional version we can add a special feature allowing for dependent claim occurrences and dependencies in investments of the different business lines.

The claim process as well as the gain process of the financial market are driven by d-dimensional compound Poisson processes. More precisely, the claim size at time T_n' is described by the d-dimensional random vector $Y_n = (Y_n^1, \dots, Y_n^d)$ with values in $[0, \infty)^d$, whereas a Poisson process $N'(t)$ with intensity λ models the occurrence of claims.
Similarly, the return process in the financial market is defined by the sequences of returns $R_n = (R_n^1, \dots, R_n^d)$ at jump times $T_n^{''}$ and a Poisson process $N''(t)$ with intensity $\tilde{\lambda}$ which specifies the occurrence of jumps in the financial market. The intensity $\tilde{\lambda}$ is generally much larger than λ. Returns are described by a d-dimensional stock price process $S_n = (S_n^1, \dots, S_n^d)$ such that $S_n^i = S_{n-1}^i(1 + R_n^i)$ and $1 + R_n^i > 0$ a.s.

The discrete-time process $X_n = (X_n^1, \dots, X_n^d)$ specifies our risk process where X_n^i denotes the risk process of business line i immediately after time T_n. The time epochs T_n result from the superposition of the claim times T_n' and the jump times at the financial market $T_n^{''}$.

We consider now two possible actions for the insurance company.

1. **The process can be controlled by reinsurance**.
   The insurance company chooses a retention level $b = (b^1, \dots, b^d) \in [\underline{b}, \bar{b}]^d$, $\underline{b} \in \mathbb{R}_+^d$, of reinsurance for a period. The kind of reinsurance is fixed in advance. The (measurable) function $h(b^i, y^i)$ denotes the part of the claim y^i in business line i paid by the insurer, where $0 \leq h(b^i, y^i) \leq y^i$. We mainly constrain ourselves to the case of proportional reinsurance. When a claim arises, the share of the reinsurer and the ceding company are paid in proportions of the claim which are fixed in advance, that is

   $$h(b^i, y^i) = b^i \cdot y^i \text{ with retention level } 0 \leq b^i \leq 1, \text{ for all } i \in \{1, \dots, d\}.$$

   In case of an excess of loss reinsurance the reinsurer only responds if the claims suffered by the insurer exceeds a certain amount, that is

   $$h(b, y) = \min(b, y) \text{ with retention level } 0 \leq b \leq \infty.$$

   For each retention level b the insurer pays a reinsurance premium which has to be deducted from the premium $c = (c^1, \dots, c^d) \in \mathbb{R}^d$ being paid by the policy holder. This leads to a so-called net-income $c(b) = (c^1(b^1), \dots, c^d(b^d))$ that may be calculated according to the expected value principle. It holds that

$$0 \le c(b) \le c = c(\bar{b}) \text{ for } \underline{b} \le b \le \bar{b} \tag{2.5}$$

and $c(b)$ is increasing in b. The smallest retention level $\underline{b}$ has to be chosen such that (2.5) is satisfied.

2. **The insurance company can invest in a financial market**. We consider a financial market where d risky assets (stocks) are traded and described by the price process $S_n = (S_n^1, \dots, S_n^d)$ as mentioned before. At any time T_n the insurance company chooses a portfolio vector $\delta_n = (\delta_n^1, \dots, \delta_n^d)$ where the component δ_n^i denotes the amount of capital invested in stock i by business line i.

To sum up, the risk process can be controlled by reinsurance, choosing retention level $b \in [\underline{b}, \bar{b}]^d$ and by investments in a financial market choosing a portfolio vector $\delta \in \mathbb{R}^d$. Thus our control action $a = (b, \delta)$ consists of two components.
We define the Poisson process $N(t)$ by superposition of the independent Poisson processes $N'(t)$ and $N''(t)$. Therefore, $N(t) = N'(t) + N''(t)$ is a Poisson process with intensity $\lambda + \tilde{\lambda}$ and jump times T_n. The period length Z_n is given by the intervals between the jump times of the Poisson process $N(t)$. We write $K_n = 0$ if the jump is caused by a claim at time T_n and $K_n = 1$ if there is a jump at the financial market. Thus $\mathbb{P}(K_n = 1) = \frac{\tilde{\lambda}}{\tilde{\lambda}+\lambda} = 1 - \mathbb{P}(K_n = 0)$.

The surplus process of business line i can be computed according to

$$X_{n+1}^i = X_n^i + c^i(b_n^i)Z_{n+1} + \delta_n^i R_{n+1}^i K_{n+1} - (1 - K_{n+1})h(b_n^i, Y_{n+1}^i)$$

given the surplus X_n^i and the control action (b_n^i, δ_n^i).

We suppose that the disturbance of our model $W_n = (R_n, Y_n, Z_n, K_n)$ satisfies the following assumption.

Model assumption
The random vectors W_n, $1 \le n \le N$ are independent and identically distributed, (R_n, Z_n, K_n) and Y_n are independent, (Z_n, R_n) and K_n are independent and

(i) Z_n is exponentially distributed with parameter $\lambda + \tilde{\lambda}$,

(ii) $Y_n^i > 0$ for all $i \in \{1, \dots, d\}$,

(iii) $\mathbb{P}(R_n^i < 0) > 0$ and $\mathbb{E}|R_n^i| < \infty$ for all $i \in \{1, \dots, d\}$.

Note that we allow Z_n and R_n to be dependent which is a reasonable approach since indeed the gains on a financial market depend on the duration of investment. Moreover, we allow for dependent claims Y_n^i and dependent returns R_n^i on the financial market of the single business lines. Such possible dependencies might be modeled using the concept of copulas as we show in Section 2.4.

2.3 Maximizing expected exponential utility

We are interested in maximizing expected exponential utility of terminal wealth in our multidimensional insurance model. The choice of this performance criterion was motivated by a result of Browne (1995) in a diffusion model. The policy maximizing expected exponential utility of terminal wealth has been shown to be identical to the policy which minimizes the ruin probability at least under certain constraints.

Imbedding our optimization problem in the theory of discrete-time stochastic dynamic programming we choose

1. the state space $S = \mathbb{R}^d$;

2. the action space $A = D(x) = [\underline{b}, \overline{b}]^d \times \mathbb{R}^d$ meaning we are allowed to borrow an unlimited amount of money. Our decision function (b, δ) is composed of two components where $b \in [\underline{b}, \overline{b}]^d$ denotes the retention level of reinsurance and $\delta \in \mathbb{R}^d$ the amount invested in stocks;

3. the transition function

$$F^{tr}(x, a, w) = x + c(b)z + \delta\rho k - h(b, y)(1 - k)$$

   for $a = (b, \delta)$ and $w = (\rho, y, z, k)$;

4. the cost functions $g(x, a) = 0$ and $V_0(x) = \nu_0 \exp(-\theta \sum_{i=1}^{d} x^i)$ for some $\nu_0 \geq 0$. This means the insurance company is imposed a fine which gets larger the smaller the aggregated risk reserve is at the end of the term.

We aim to minimize the expected penalty to pay. That is,

$$\text{minimize } \mathbb{E}\Big[\exp\Big(-\theta \sum_{i=1}^{d} X_N^{i,x,\pi}\Big)\Big]$$

which is equivalent to the maximization of the expected exponential utility of terminal wealth.

In order to apply the main theorem for the finite horizon model we have to check the conditions of the structure assumption. Define

$$\mathcal{V} := \Big\{ v : \mathbb{R}^d \to [0,\infty) : v(x) = \nu \, \exp\Big(-\theta \sum_{i=1}^d x^i\Big), \nu \geq 0 \Big\}$$

and let Φ be the set of all constant decision functions, i.e.

$$\Phi := \{\varphi : S \to A : \varphi(x) = (b,\delta) \,\forall x \in S,\ (b,\delta) \in [\underline{b},\overline{b}]^d \times \mathbb{R}^d\}.$$

Now, let us verify conditions (i)-(iii) of the structure assumption.

(i) $V_0 \in \mathcal{V}$ trivially holds by definition of the terminal cost function V_0.

(ii) We have to show that $Uv \in \mathcal{V}$ for all $v \in \mathcal{V}$.
For $v(x) = \nu \exp(-\theta \sum_{i=1}^d x^i)$, the control $a = (b,\delta) \in [\underline{b},\overline{b}] \times \mathbb{R}^d$ and the disturbance vector $W = (R^1,\dots,R^d,Y^1,\dots,Y^d,Z,K)$ we have

$$\begin{aligned} &Lv(x,a) \\ =&g(x,a) + \mathbb{E}\big[v\big(F^{tr}(x,a,W)\big)\big] \\ =&\nu\,\mathbb{E}\Big[\exp\Big(-\theta\sum_{i=1}^d \big(x^i + c^i(b^i)Z - (1-K)h(b^i,Y^i) + K\delta^i R^i\big)\Big)\Big] \\ =&\nu\, \exp\Big(-\theta \sum_{i=1}^d x^i\Big) \\ &\qquad \mathbb{E}\Big[\exp\Big(-\theta\sum_{i=1}^d \big(c^i(b^i)Z - (1-K)h(b^i,Y^i) + K\delta^i R^i\big)\Big)\Big]. \end{aligned}$$

This yields

$$Uv(x) = \nu^* \,\nu\, \exp\Big(-\theta\sum_{i=1}^d x^i\Big)$$

where

$$\nu^* := \inf_{(b,\delta)\in A} \mathbb{E}\Big[\exp\Big(-\theta\sum_{i=1}^d \big(c^i(b^i)Z - (1-K)h(b^i,Y^i) + K\delta^i R^i\big)\Big)\Big]. \tag{2.6}$$

Thus our model satisfies condition (ii).

(iii) We have to find a decision function $\varphi \in \Phi$ such that $U_\varphi v = Uv$ for all $v \in \mathcal{V}$. Here it is sufficient to show that the infimum in (2.6) is attained by some constant policy. The decision function $\varphi \in \Phi$ can then be defined as the constant function $\varphi = (b^*, \delta^*)$.
For this purpose we need some preliminary work which will be done in detail in the following.

In order to verify property (iii) of the structure assumption we first show that

$$(b,\delta) \mapsto \mathbb{E}\Big[\exp\Big(-\theta\sum_{i=1}^{d}\big(x^i + c^i(b^i)Z - (1-K)h(b^i,Y^i) + K\delta^i R^i\big)\Big)\Big] \quad (2.7)$$

is continuous on any compact set $[\underline{b},\bar{b}]^d \times B$. On account of this we need the following assumption.

Continuity assumption

The functions $c^i(b^i)$ and $h(b^i,y^i)$ are continuous in b^i (for each y^i) for all $i \in \{1,\dots,d\}$. Furthermore,

$$\mathbb{E}\Big[\exp\Big(\theta\sum_{i=1}^{d} Y^i\Big)\Big] < \infty \quad \text{and} \quad \mathbb{E}[\exp(\epsilon\|R\|)] < \infty \text{ for all } \epsilon > 0.$$

We are now able to check the continuity of (2.7). Since $0 \le h(b^i,y^i) \le y^i$ for all $i \in \{1,\dots,d\}$, $k \in \{0,1\}$ and the random vectors $(Y^1,\dots,Y^d)$ and $(R^1,\dots,R^d,K)$ are independent by our model assumption, we have

$$\begin{aligned}
&\mathbb{E}\Big[\exp\Big(-\theta\sum_{i=1}^{d}\big(x^i + c^i(b^i)Z - (1-K)h(b^i,Y^i) + K\delta^i R^i\big)\Big)\Big]\\
\le&\mathbb{E}\Big[\exp\Big(\theta\sum_{i=1}^{d} Y^i - \theta K\sum_{i=1}^{d}\delta^i R^i\Big)\Big]\\
=&\mathbb{E}\Big[\exp\Big(\theta\sum_{i=1}^{d} Y^i\Big)\Big]\,\mathbb{E}\Big[\exp\Big(-\theta K\sum_{i=1}^{d}\delta^i R^i\Big)\Big]\\
\le&\mathbb{E}\Big[\exp\Big(\theta\sum_{i=1}^{d} Y^i\Big)\Big]\,\mathbb{E}\Big[\exp\Big|-\theta K\sum_{i=1}^{d}\delta^i R^i\Big|\Big]\\
\le&\mathbb{E}\Big[\exp\Big(\theta\sum_{i=1}^{d} Y^i\Big)\Big]\,\mathbb{E}\big[\exp(\theta\, K\,\|\delta\|\,\|R\|)\big]
\end{aligned}$$

where the last inequality follows from the Cauchy-Schwarz inequality. By the continuity assumption we therefore have,

$$\mathbb{E}\Big[\exp\Big(-\theta\sum_{i=1}^{d}\big(x^i + c^i(b^i)Z - (1-K)h(b^i,Y^i) + K\delta^i R^i\big)\Big)\Big] < \infty$$

for all $(b,\delta) \in [\underline{b},\bar{b}]^d \times B$, where B is compact. Applying Lebesgue's dominated convergence theorem yields the continuity of (2.7).

The control constraint is not compact in our model. But it can be shown that the infimum is attained under the following no arbitrage assumption.

No arbitrage assumption

$$\mathbb{P}\Big(\sum_{i=1}^{d} \delta^i R^i \geq 0\Big) = 1 \text{ implies } \mathbb{P}\Big(\sum_{i=1}^{d} \delta^i R^i = 0\Big) = 1 \text{ for any } \delta \in \mathbb{R}^d.$$

Lemma 2.3.1. *The infimum over* $D(x)$ *of*

$$(b,\delta) \mapsto \mathbb{E}\Big[\exp\Big(-\theta \sum_{i=1}^{d} \big(K\delta^i R^i + c^i(b^i)Z - h(b^i, Y^i)(1-K)\big)\Big)\Big]$$

is attained in (b^*, δ^*)*, where* $\delta^* \in \mathcal{L}$ *and* $\mathcal{L}$ *denotes the smallest linear space in* $\mathbb{R}^d$ *with* $\mathbb{P}(R \in \mathcal{L}) = 1$.

Proof. Let us show that the condition

$$\mathbb{P}\Big(\sum_{i=1}^{d} \delta^i R^i < 0\Big) > 0 \text{ for all } \delta \in \mathcal{L},\ \delta \neq (0,\ldots,0)$$

follows from the no-arbitrage assumption. Let $\delta \in \mathcal{L}$, $\delta \neq (0,\ldots,0)$ and assume that $\mathbb{P}\big(\sum_{i=1}^{d} \delta^i R^i < 0\big) = 0$. That means $\mathbb{P}\big(\sum_{i=1}^{d} \delta^i R^i \geq 0\big) = 1$. By the no-arbitrage assumption

$$\mathbb{P}\Big(\sum_{i=1}^{d} \delta^i R^i \geq 0\Big) = 1 \quad \text{implies} \quad \mathbb{P}\Big(\sum_{i=1}^{d} \delta^i R^i = 0\Big) = 1.$$

That is, $\delta \in \mathcal{L}^{\perp}$ which clearly contradicts the assumption.
By the projection theorem we know for all $\delta \in \mathbb{R}^d$ there is a unique $\check{\delta} \in \mathcal{L}$ and $\breve{\delta} \in \mathcal{L}^{\perp}$ such that $\delta = \check{\delta} + \breve{\delta}$. Thus $\delta \cdot R = \check{\delta} \cdot R$ almost surely for all $\delta \in \mathbb{R}^d$. Therefore, it suffices to consider $\delta \in \mathcal{L}$. Define

$$v(b,\delta) := \mathbb{E}\Big[\exp\Big(-\theta \sum_{i=1}^{d} \big(K\delta^i R^i + c^i(b^i)Z - h(b^i, Y^i)(1-K)\big)\Big)\Big]$$

and

$$F_\lambda := \big\{(b,\delta) \in [\underline{b}, \overline{b}]^d \times \mathcal{L} : \|\delta\| = 1, v(b, \lambda\delta) \leq v(b,0) + 1\big\}.$$

The set F_λ is compact. Furthermore, we know that $v(b,\delta)$ is convex in δ since it is a composition of the exponential and a linear function. Thus, we get that $F_\lambda \subset F_{\lambda'}$ for $0 < \lambda' < \lambda$ since

$$\begin{aligned} v(b,\lambda'\delta) =& v\Big(b, \frac{\lambda'}{\lambda}\lambda\delta + \Big(1 - \frac{\lambda'}{\lambda}\Big)0\Big) \leq \frac{\lambda'}{\lambda} v(b,\lambda\delta) + \Big(1 - \frac{\lambda'}{\lambda}\Big) v(b,0) \\ \leq& \frac{\lambda'}{\lambda}(v(b,0)+1) + \Big(1 - \frac{\lambda'}{\lambda}\Big) v(b,0) < 1 + v(b,0). \end{aligned}$$

From the no-arbitrage assumption and monotone convergence theorem we have for $\delta \in \mathcal{L}$

$$\begin{aligned}\lim_{\lambda\uparrow\infty} v(b,\lambda\delta) \geq& \lim_{\lambda\uparrow\infty} \mathbb{E}\Big[\exp\Big(-\theta\sum_{i=1}^{d}\big(\lambda K\delta^i R^i + Zc^i\big)\Big)\mathbb{1}_{\{\sum_{i=1}^{d}\delta^i R^i<0\}}\Big]\\ =&\mathbb{E}\Big[\lim_{\lambda\uparrow\infty}\exp\Big(-\theta\sum_{i=1}^{d}\big(\lambda K\delta^i R^i + Zc^i\big)\Big)\mathbb{1}_{\{\sum_{i=1}^{d}\delta^i R^i<0\}}\Big]\\ =&\infty.\end{aligned}$$

Thus, there is some $\lambda_0 \in \mathbb{N}$ such that $F_\lambda = \emptyset$ for all $\lambda \geq \lambda_0$. That is, $v(b,\delta) \geq v(b,0)+1$ for all $\delta \in \mathbb{R}^d$ with $\|\delta\| \geq \lambda_0$. Hence,

$$\inf_{(b,\delta)\in D(x)} v(b,\delta) = \min_{(b,\delta)\in[\underline{b},\overline{b}]\times\mathcal{L},\|\delta\|\leq\lambda_0} v(b,\delta).$$

Hence, the infimum over $(b,\delta)\in D(x)$ is attained on the compact set $[\underline{b},\overline{b}]\times\{\delta\in\mathcal{L}:\|\delta\|\leq\lambda_0\}$. □

Therefore, our model satisfies the conditions of the structure assumption with $\mathcal{V}$ and Φ as defined above and we are able to apply the main theorem for the finite horizon problem (cf. Theorem 2.1.7) to obtain the main result of this section.

Theorem 2.3.2. *There exists an optimal policy π^* such that $\pi^* = (b^*,\delta^*) \in [\underline{b},\overline{b}]\times\mathbb{R}^d$ neither depends on the time n nor on the present state x where (b^*,δ^*) is the minimizer of*

$$(b,\delta)\mapsto \mathbb{E}\Big[\exp\Big(-\theta\sum_{i=1}^{d}\big(c^i(b^i)Z-(1-K)h(b^i,Y^i)+K\delta^iR^i\big)\Big)\Big].$$

Remark 3. Let us finally consider a slight modification of the latter model. Unlike in the previously considered model we do only allow for control at the occurrence of claims. The random variable Z_n is the period length between the $(n-1)$-th and n-th claim. Thus R is the return of the period between two claims. Given the surplus X_n and the control action (b_n,δ_n) the surplus process can be computed according to

$$X_{n+1} = X_n + c(b_n)Z_{n+1} + \delta_n R_{n+1} - h(b_n, Y_{n+1}).$$

We get an equivalent result to Theorem 2.3.2. If we additionally assume that R and Z are independent then the optimal portfolio (b^*,δ^*) can be obtained by computing the minimizers of the functions

$$\begin{aligned}b \mapsto&\mathbb{E}\Big[\exp\Big(-\theta\sum_{i=1}^{d}\big(c^i(b^i)Z-h(b^i,Y^i)\big)\Big)\Big]\\ \delta\mapsto&\mathbb{E}\Big[\exp\Big(-\theta\sum_{i=1}^{d}\delta^iR^i\Big)\Big]\end{aligned}$$

separately.

2.4 Structural comparison results

In this section we identify structure conditions with respect to the claim density and the generator of an Archimedean copula under which an insurance company reinsures a larger fraction of claims from one business line than from another. Similarly, structure conditions are derived for the investment portfolio. Since the optimal policy is constant (cf. Theorem 2.3.2) we can apply a similar theory to the one which was developed by Hennessy and Lapan (2002) for a stationary portfolio allocation problem.

We first state Sklar's theorem which is fundamental for copula theory. It illustrates the role of copulas in the relationship between multivariate distribution functions and their univariate margins. A detailed introduction to copulas can be found in Nelsen (2006) or in Joe (1997).

Theorem 2.4.1 (Sklar)**.** *Let F be a d-dimensional distribution function with margins $F^1, \dots, F^d$. Then there exists a d-dimensional copula C such that for all $x \in \overline{\mathbb{R}}^d$,*

$$F(x^1, \dots, x^d) = C\big(F^1(x^1), \dots, F^d(x^d)\big). \tag{2.8}$$

If $F^1, \dots, F^d$ are all continuous then C is unique. Otherwise, C is uniquely determined on Ran $F^1 \times \dots \times$ Ran F^d.
Conversely, if C is a d-dimensional copula and $F^1, \dots, F^d$ are distribution functions, then the function F defined by (2.8) is a d-dimensional distribution function with margins $F^1, \dots, F^d$.

Assuming dependence is modeled via an Archimedean copula, the resulting distribution possesses convenient analytic properties. McNeil and Nešlehová (2008) recently developed a sufficient and necessary condition for Archimedean generators to generate a d-dimensional copula.

Theorem 2.4.2 (Archimedean copula)**.** *Let ϕ be a continuous and strictly decreasing function defined on $[0,1] \to [0,\infty]$ such that $\phi(1) = 0$ and $\phi(0) = \infty$. Then $C : [0,1]^d \to [0,1]$ given by*

$$C(u^1, \dots, u^d) = \phi^{-1}\Big(\sum_{i=1}^{d} \phi(u^i)\Big)$$

defines an Archimedean d-copula if and only if ϕ^{-1} is d-monotone on $[0,\infty)$ i.e. differentiable up to order $d-2$ with derivatives satisfying

$$(-1)^k \big(\phi^{-1}\big)^{(k)}(t) \geq 0 \quad k \in \{1, \dots, d-2\} \tag{2.9}$$

for any $t \in [0,\infty)$ and further $(-1)^{d-2}(\phi^{-1})^{(d-2)}(t)$ is non-increasing and convex in $[0,\infty)$.

Remark 4. In terms of the notation of Nelsen (2006) we have a strict Archimedean copula since we require that $\phi(0) = \infty$. An Archimedean copula can analogously be defined in terms of ϕ^{-1} rather than ϕ, compare Nelsen (2006) or Cont and Tankov (2004).

By Theorem 2.4.1 we know that the pair, copula and marginal law of a d-dimensional random vector gives an alternative description of the law of a random vector. Let us assume in the following that the inverse of the Archimedean generator has derivatives up to order d. Using the Archimedean d-copula we can therefore compute the density function

$$f(x^1, \dots, x^d) = \big(\phi^{-1}\big)^{(d)}\Big(\sum_{i=1}^{d} \phi\big(F^i(x^i)\big)\Big) \prod_{i=1}^{d} s^i(x^i) \tag{2.10}$$

with marginal distribution functions $F^i(x^i)$, marginal densities $f^i(x^i)$, and

$$s^i(\cdot) = \phi^{(1)}\big(F^i(\cdot)\big)\, f^i(\cdot) \quad i \in \{1, \dots, d\}.$$

Thus our optimization problems can be written in terms of Archimedean d-copulas as we will show in Sections 2.4.1 and 2.4.2.

The optimization criterion is still the maximization of expected exponential utility of terminal wealth or equivalently the minimization of

$$\mathbb{E}\Big[\exp\Big(-\theta \sum_{i=1}^{d} X_N^{\pi,i}\Big)\Big]$$

where

$$X_N^{\pi,i} = \sum_{n=0}^{N-1} \big(c(b_n^i)Z_{n+1} - (1-K_{n+1})h(b_n^i, Y_{n+1}^i) + K_{n+1}\delta_n^i R_{n+1}^i\big).$$

By Theorem 2.3.2 we know that the optimal policy $\pi^* = (b^*, \delta^*) \in [\underline{b}, \overline{b}] \times \mathbb{R}^d$ neither depends on the present state of wealth X_n nor on the remaining periods n. Thus, our optimization problem reduces to the minimization of

$$\mathbb{E}\Big[\exp\Big(-\theta \sum_{i=1}^{d}\sum_{n=0}^{N-1} \big(c(b^i)Z_{n+1} - (1-K_{n+1})h(b^i, Y_{n+1}^i) + K_{n+1}\delta^i R_{n+1}^i\big)\Big)\Big]$$

where $(b, \delta) \in [\underline{b}, \overline{b}] \times \mathbb{R}^d$ is some constant policy. Moreover, since the disturbance vector $W_n = (R_n, Y_n, Z_n, K_n)$, $1 \le n \le N$ is independent and identically distributed it suffices to minimize

$$M(b,\delta) := \mathbb{E}\Big[\exp\Big(-\theta \sum_{i=1}^{d} \big(c(b^i)Z - (1-K)h(b^i, Y^i) + K\delta^i R^i\big)\Big)\Big]$$

where $Y^i := Y_1^i$, $R^i := R_1^i$ and $Z := Z_1$. The expectation can be computed using the density in (2.10) given the marginal densities of the claim size vector and the investment portfolio.
In the sequel let us consider the choice of a retention level and a portfolio-vector separately.

2.4.1 Optimal proportional reinsurance

Modeling dependence via an Archimedean copula we establish conditions under which we can be sure that the optimal retention level of one business line is larger than it is from another business line. Suppose we have proportional reinsurance and cheap reinsurance with the same premium income in each business line i.e. $c(b^i) = cb^i$. For notational simplicity we assume that there is no financial market.

Theorem 2.4.3. *Let ϕ be an Archimedean generator and its inverse ϕ^{-1} has derivatives up to order d with alternating signs satisfying*

$$(-1)^k \big(\phi^{-1}\big)^{(k)}(t) \geq 0, \quad k \in \{1, \dots, d\} \tag{2.11}$$

for any $t \in [0, \infty)$. Moreover, let F^i, f^i for $i \in \{1, \dots, d\}$ be the marginal distribution functions, respectively the marginal densities of the negative claim size vector. If

$$s^i(y) \geq s^j(y)$$

on the domain of definition $\mathbb{R}_-^d$, $i, j \in \{1, \dots, d\}$, then

$$b^{*,i} \leq b^{*,j}$$

where $b^* = (b^{1,*}, \dots, b^{d,*})$ is the optimal retention level.

Proof. Let $y \in \mathbb{R}_-^d$ and denote by $f(y)$ the density of the negative claim size vector taking into consideration the dependence as described in (2.10). Our optimization problem is to minimize

$$M(b) = \int_{\mathbb{R}_+} \int_{\mathbb{R}_-^d} \exp(-\lambda z) \exp\Big(-\theta \sum_{i=1}^d \big(cb^i z + b^i y^i\big)\Big) f(y^1, \dots, y^d) \prod_{i=1}^d dy^i \, dz.$$

Note that Z is exponentially distributed with parameter λ. We now aim to identify structure conditions under which $b^{*j} \geq b^{*,i}$. Let us therefore analyze the effect on the expected terminal utility when b^i is permuted with b^j of an arbitrary retention level $b \in \mathbb{R}_+^d$. Without loss of generality consider permutation τ of indices 1 and 2 and keep the retention levels $(b^3, \dots, b^d)$ fixed. We define the permuted policy by $b^\tau = (b^2, b^1, b^3, \dots, b^d)$. Comparing the two expected terminal dis-utility terms we obtain

$$\begin{aligned} &M(b) - M(b^\tau) \\ &= \int_{\mathbb{R}_+} \int_{\mathbb{R}_-^{d-2}} \exp\Big(-\lambda z - \theta c z \sum_{i=1}^d b^i - \theta \sum_{i=3}^d b^i y^i\Big) \mathcal{M}(y^3, \dots, y^d) \prod_{i=3}^d dy^i \, dz \end{aligned}$$

where

$$\mathcal{M}(y^3,\dots,y^d) = \int_{\mathbb{R}^2_-} \exp\Big(-\theta\sum_{i=1}^{2} b^i y^i\Big) f(y^1,\dots,y^d)\,dy^1dy^2 \\ - \int_{\mathbb{R}^2_-} \exp\big(-\theta(b^2y^1+b^1y^2)\big) f(y^1,\dots,y^d)\,dy^1dy^2.$$

Separating the integration area of $\mathcal{M}(y^3,\dots,y^d)$ and changing the integration variables in the second term yields

$$\begin{aligned}
&\mathcal{M}(y^3,\dots,y^d)\\
=&\int_{\{\mathbb{R}^2_-,y_1\geq y_2\}} \exp\Big(-\theta\sum_{i=1}^{2} b^i y^i\Big) f(y^1,\dots,y^d)\,dy^2dy^1\\
&+\int_{\{\mathbb{R}^2_-,y^1\geq y^2\}} \exp\Big(-\theta(b^1y^2+b^2y^1)\Big) f(y^2,y^1,y^3,\dots,y^d)\,dy^2dy^1\\
&-\int_{\{\mathbb{R}^2_-,y_1\geq y_2\}} \exp\Big(-\theta(b^2y^1+b^1y^2)\Big) f(y^1,\dots,y^d)\,dy^2dy^1\\
&-\int_{\{\mathbb{R}^2_-,y^1\geq y^2\}} \exp\Big(-\theta\sum_{i=1}^{2} b^i y^i\Big) f(y^2,y^1,y^3,\dots,y^d)\,dy^2dy^1.
\end{aligned}$$

Plugging in the density in terms of the Archimedean copula as described in (2.10) we have

$$\begin{aligned}
&\mathcal{M}(y^3,\dots,y^d)\\
=&\prod_{i=3}^{d} s^i(y^i)\int_{\{\mathbb{R}^2_-,y^1\geq y^2\}} \big(\mathcal{D}(y^1,y^2)-\mathcal{D}(y^2,y^1)\big)\mathcal{H}(y^1,\dots,y^d)\,dy^2dy^1
\end{aligned}$$

with

$$\mathcal{D}(y^1,y^2) = \exp\Big(-\theta\sum_{i=1}^{2} b^i y^i\Big)$$

and

$$\begin{aligned}
&\mathcal{H}(y^1,\dots,y^d)\\
=&(\phi^{-1})^{(d)}\Big(\sum_{i=1}^{d}\phi\big(F^i(y^i)\big)\Big)s^1(y^1)s^2(y^2)\\
&-(\phi^{-1})^{(d)}\Big(\phi\big(F^1(y^2)\big)+\phi\big(F^2(y^1)\big)+\sum_{i=3}^{d}\phi\big(F^i(y^i)\big)\Big)s^1(y^2)s^2(y^1).
\end{aligned}$$

Define

$$\mathcal{L}(y^3,\ldots,y^d)=\int_{\{\mathbb{R}^2_-,y^1\geq y^2\}}\big(\mathcal{D}(y^1,y^2)-\mathcal{D}(y^2,y^1)\big)\mathcal{H}(y^1,\ldots,y^d)\,dy^2dy^1.$$

Thus, $\mathcal{M}(y^3,\ldots,y^d)$ and therefore $M(b)-M(b^\tau)$ has the same sign as $(-1)^{d-2}\mathcal{L}(y^3,\ldots,y^d)$. In order to deduce an inequality let us integrate $\mathcal{L}(y^3,\ldots,y^d)$ along y^2, that is

$$\begin{aligned}&\mathcal{L}(y^3,\ldots,y^d)\\&=\int_{\mathbb{R}_-}\big(\mathcal{D}(y^1,y^2)-\mathcal{D}(y^2,y^1)\big)\,\mathcal{J}(y^1,\ldots,y^d)\Big|_{-\infty}^{y^1}dy^1\\&\quad-\int_{\{\mathbb{R}^2_-,y^1\geq y^2\}}\big(-b^2\theta\mathcal{D}(y^1,y^2)+b^1\theta\mathcal{D}(y^2,y^1)\big)\mathcal{J}(y^1,\ldots,y^d)\,dy^2dy^1\end{aligned}$$

where

$$\begin{aligned}&\mathcal{J}(y^1,\ldots,y^d)\\&=\big(\phi^{-1}\big)^{(d-1)}\Big(\sum_{i=1}^d\phi\big(F^i(y^i)\big)\Big)s^1(y^1)\\&\quad-\big(\phi^{-1}\big)^{(d-1)}\Big(\phi\big(F^1(y^2)\big)+\phi\big(F^2(y^1)\big)+\sum_{i=3}^d\phi\big(F^i(y^i)\big)\Big)s^2(y^1).\end{aligned}$$

Note that the first summand vanishes through integration-by-parts. More precisely, the first term in the integrand clearly vanishes for $y^2=y^1$ since $\mathcal{D}(y^1,y^2)-\mathcal{D}(y^2,y^1)|_{y^2=y^1}=0$, the second term equals zero since for $y^2=-\infty$ we have $F(-\infty)=0$, $\phi(0)=\infty$ and thus $\mathcal{J}(y^1,y^2,y^3,\ldots,y^d)|_{y^2=-\infty}=0$ (cf. Lemma A.2.2). Thus, it remains

$$\begin{aligned}&\mathcal{L}(y^3,\ldots,y^d)\\&=\theta\int_{\{\mathbb{R}^2_-,y^1\geq y^2\}}\big(b^2\mathcal{D}(y^1,y^2)-b^1\mathcal{D}(y^2,y^1)\big)\mathcal{J}(y^1,\ldots,y^d)\,dy^2dy^1.\end{aligned}$$

Let us now deduce conditions under which $(-1)^{d-2}\mathcal{J}(y^1,\ldots,y^d)\leq 0$. The expression $\mathcal{J}(y^1,\ldots,y^d)$ may be written as

$$\begin{aligned}&\mathcal{J}(y^1,\ldots,y^d)\\&=\big((s^1(y^1)-s^2(y^1)\big)\big(\phi^{-1}\big)^{(d-1)}\Big(\sum_{i=1}^d\phi\big(F^i(y^i)\big)\Big)+s^2(y^1)\mathcal{K}(y^1,\ldots,y^d)\end{aligned}\tag{2.12}$$

where

$$\begin{aligned}\mathcal{K}(y^1,\dots,y^d) =&\big(\phi^{-1}\big)^{(d-1)}\Big(\sum_{i=1}^{d}\phi\big(F^i(y^i)\big)\Big)\\ &-\big(\phi^{-1}\big)^{(d-1)}\Big(\phi\big(F^1(y^2)\big)+\phi\big(F^2(y^1)\big)+\sum_{i=3}^{d}\phi\big(F^i(y^i)\big)\Big).\end{aligned}$$

In determining the sign of $\mathcal{K}(y^1,\dots,y^d)$ we can follow Hennessy and Lapan (2002). Note that $(-1)^{d-1}(\phi^{-1})^{(d-1)}(\cdot)$ is a decreasing function since $(-1)^{d-1}(\phi^{-1})^{(d-2)}(\cdot)\le 0$ by condition (2.11). We therefore obtain

$$(-1)^{d-1}\mathcal{K}(y^1,\dots,y^d)\le 0 \text{ for } y^1\ge y^2$$

if

$$L(y^1,y^2)=\phi\big(F^1(y^1)\big)+\phi\big(F^2(y^2)\big)-\phi\big(F^1(y^2)\big)-\phi\big(F^2(y^1)\big)\ge 0.$$

Fix y^2. We have $L(y^1,y^2)\big|_{y^1=y^2}=0$ and a sufficient and necessary condition for $\frac{\partial L(y^1,y^2)}{\partial y^1}\ge 0$ to hold is that

$$s^1(y^1)-s^2(y^1)\ge 0 \text{ for } y^1\ge y^2. \tag{2.13}$$

Returning to $\mathcal{J}(y^1,\dots,y^d)$, (2.11), (2.13) and $s^2(y^1)\le 0$ imply that

$$(-1)^{d-1}\mathcal{J}(y^1,\dots,y^d)\ge 0 \text{ and } (-1)^{d-2}\mathcal{J}(y^1,\dots,y^d)\le 0.$$

We further now that $(b^1-b^2)(y^1-y^2)\le 0$ for $y^1\ge y^2$ if and only if $b^1\le b^2$. Hence,

$$b^2\mathcal{D}(y^1,y^2)-b^1\mathcal{D}(y^2,y^1)\ge 0$$

for $y^1\ge y^2$ if and only if $b^1\le b^2$. Assuming

$$s^1(y)\ge s^2(y)$$

we have

$$M(b)-M(b^\tau)\le 0 \text{ if and only if } b^1\le b^2.$$

This yields the assumed retention level is preferable to its bivariate permutation. □

2.4.2 Optimal investment

After having derived structure conditions for the optimal retention level let us finally establish conditions under which we can be sure that the optimal investment of one business line is larger than it is from another business line. We exclude short-selling and for notational convenience we disregard the occurrence of claims.

Theorem 2.4.4. *Let ϕ be an Archimedean generator and its inverse ϕ^{-1} has derivatives up to order d with alternating signs satisfying*

$$(-1)^k \big(\phi^{-1}\big)^{(k)}(t) \geq 0, \quad k \in \{1, \dots, d\} \tag{2.14}$$

for any $t \in [0, \infty)$. Moreover, let F^i, f^i for $i \in \{1, \dots, d\}$ be the marginal distribution functions, respectively the marginal densities of the return vector. If

$$s^i(r) \geq s^j(r)$$

on the domain of definition $\mathbb{R}^d$, $i, j \in \{1, \dots, d\}$, then

$$\delta^{*,i} \leq \delta^{*,j}$$

where $\delta^* = (\delta^{*,1}, \dots, \delta^{*,d}) \in \mathbb{R}^d_+$ is the optimal retention level.

Proof. Let $r \in \mathbb{R}^d$ and denote by $f(r)$ the return density taking into consideration dependence as described in (2.10). Our optimization problem is to minimize

$$M(\delta) = \int_{\mathbb{R}_+} \int_{\mathbb{R}^d} \exp\big(-\tilde{\lambda} z\big) \exp\Big(-\theta\Big(cz + \sum_{i=1}^d \delta^i r^i\Big)\Big) f(r^1, \dots, r^d) \prod_{i=1}^d dr^i dz.$$

Note that Z is exponentially distributed with parameter $\tilde{\lambda}$. We now look for conditions under which $\delta^{*,j} \geq \delta^{*,i}$. Let us therefore analyze the effect on the expected terminal utility when δ^i is permuted with δ^j for an arbitrary portfolio vector $\delta \in \mathbb{R}^d_+$. Without loss of generality consider permutation τ of indices 1 and 2 and keep the investment $(\delta^3, \dots, \delta^d)$ fixed. We define the permuted policy by $\delta^\tau = (\delta^2, \delta^1, \delta^3, \dots, \delta^d)$. Comparing the two expected terminal dis-utilities we obtain

$$\begin{aligned} &M(\delta) - M(\delta^\tau) \\ &= \int_{\mathbb{R}_+} \int_{\mathbb{R}^{d-2}} \exp\Big(-\tilde{\lambda} z - \theta\Big(cz + \sum_{i=3}^d \delta^i r^i\Big)\Big) \mathcal{M}(r^3, \dots, r^d) \prod_{i=3}^d dr^i \, dz \end{aligned}$$

where

$$\begin{aligned} \mathcal{M}(r^3, \dots, r^d) = &\int_{\mathbb{R}^2} \exp\Big(-\theta \sum_{i=1}^2 \delta^i r^i\Big) f(r^1, \dots, r^d) \, dr^1 dr^2 \\ &- \int_{\mathbb{R}^2} \exp\big(-\theta(\delta^2 r^1 + \delta^1 r^2)\big) f(r^1, \dots, r^d) \, dr^1 dr^2. \end{aligned}$$

Separating the integration area of $\mathcal{M}(r^3, \dots, r^d)$, changing the integration variables in the second summand and plugging in the return density in terms of the Archimedean copula as described in (2.10) yields

$$\begin{aligned}
&\mathcal{M}(r^3,\dots,r^d)\\
=&\prod_{i=3}^{d} s^i(r^i)\int_{\{\mathbb{R}^2, r^1\geq r^2\}}\big(\mathcal{D}(r^1,r^2)-\mathcal{D}(r^2,r^1)\big)\mathcal{H}(r^1,\dots,r^d)\,dr^1dr^2
\end{aligned}$$

where

$$\mathcal{D}(r_1,r_2)=\exp\Big(-\theta\sum_{i=1}^{2}\delta^i r^i\Big)$$

and

$$\begin{aligned}
&\mathcal{H}(r^1,\dots,r^d)\\
=&\big(\phi^{-1}\big)^{(d)}\Big(\sum_{i=1}^{d}\phi\big(F^i(r^i)\big)\Big)s^1(r^1)s^2(r^2)\\
&-\big(\phi^{-1}\big)^{(d)}\Big(\phi\big(F^1(r^2)\big)+\phi\big(F^2(r^1)\big)+\sum_{i=3}^{d}\phi\big(F^i(r^i)\big)\Big)s^1(r^2)s^2(r^1).
\end{aligned}$$

Define

$$\mathcal{L}(r^3,\dots,r^d)=\int_{\{\mathbb{R}^2, r^1\geq r^2\}}\big(\mathcal{D}(r^1,r^2)-\mathcal{D}(r^2,r^1)\big)\mathcal{H}(r^1,\dots,r^d)\,dr^1dr^2.$$

Thus, $\mathcal{M}(r^3,\dots,r^d)$ and therefore $M(\delta)-M(\delta^\tau)$ has the same sign as $(-1)^{d-2}\mathcal{L}(r^3,\dots,r^d)$. In order to deduce an inequality let us integrate $\mathcal{L}(r^3,\dots,r^d)$ along r^2, that is

$$\begin{aligned}
&\mathcal{L}(r^3,\dots,r^d)\\
=&\int_{\mathbb{R}}\big(\mathcal{D}(r^1,r^2)-\mathcal{D}(r^2,r^1)\big)\mathcal{J}(r^1,r^2)\Big|_{-\infty}^{r^1}dr^1\\
&-\theta\int_{\{\mathbb{R}^2, r^1\geq r^2\}}\big(-\delta^2\mathcal{D}(r^1,r^2)+\delta^1\mathcal{D}(r^2,r^1)\big)\mathcal{J}(r^1,\dots,r^d)\,dr^1dr^2
\end{aligned}$$

where

$$\begin{aligned}
&\mathcal{J}(r^1,\dots,r^d)\\
=&\big(\phi^{-1}\big)^{(d-1)}\Big(\sum_{i=1}^{d}\phi\big(F^i(r^i)\big)\Big)s^1(r^1)\\
&-\big(\phi^{-1}\big)^{(d-1)}\Big(\phi\big(F^1(r^2)\big)+\phi\big(F^2(r^1)\big)+\sum_{i=3}^{d}\phi\big(F^i(r^i)\big)\Big)s^2(r^1).
\end{aligned}$$

Note that the first summand vanishes. More precisely, the first term in the integrand clearly vanishes for $r^2=r^1$ since $\mathcal{D}(r^1,r^2)-\mathcal{D}(r^1,r^1)|_{r^2=r^1}=0$, the second term equals zero since for $r^2 = -\infty$ we have $F(-\infty) = 0$, $\phi(0) = \infty$ and thus $\mathcal{J}(r^1, r^2, r^3, \ldots, r^d)|_{r^2=-\infty} = 0$ (cf. Lemma A.2.2). Thus, it remains

$$\mathcal{L}(r^3, \ldots, r^d) = \theta \int_{\{\mathbb{R}^2, r^1 \geq r^2\}} \big(\delta_2 \mathcal{D}(r^1, r^2) - \delta_1 \mathcal{D}(r^2, r^1)\big) \mathcal{J}(r^1, \ldots, r^d)\, dr^1 dr^2.$$

Similarly to the proof of Theorem 2.4.3 we can show that $(-1)^{d-2}\mathcal{J}(r^1, r^2) \leq 0$ if $s^1(r^1) \geq s^2(r^1)$. Moreover, we know that $\mathcal{D}(r^1, r^2) - \mathcal{D}(r^2, r^1) \geq 0$ for $r^1 \geq r^2$, or equivalently $(\delta^2 - \delta^1)(r^2 - r^1) \leq 0$ for $r^1 \geq r^2$ if and only if $\delta^2 \geq \delta^1$. Further, note that $\delta \in \mathbb{R}^d_+$. Hence, $M(\delta) - M(\delta^\tau) \leq 0$ under the investment policy $\delta = (\delta^1, \ldots, \delta^d)$ where $\delta^1 \leq \delta^2$. This yields that the assumed retention level is preferable to its bivariate permutation. □

2.4.3 Examples

Hennessy and Lapan (2002) show that the reversed hazard dominance on the marginals is sufficient for the stochastic order condition in Theorem 2.4.3 and Theorem 2.4.4. Furthermore, they provide some examples of copulas satisfying this condition. Let us first recall the definition of the reversed hazard order (cf. Müller and Stoyan (2002), Definition 1.3.9).

Definition 2.4.5 (Reversed hazard rate order)**.** *The distribution function $F^2(x)$ is said to dominate the distribution function $F^1(x)$ with respect to the reversed hazard rate order (written $F^1 \leq_{rh} F^2$) if*

$$\mathbb{R} \ni x \mapsto \frac{F^2(x)}{F^1(x)}$$

is increasing.

Proposition 2.4.6. *Let ϕ be an Archimedean generator such that*

$$R_\phi(x) = -x \frac{\phi^{(2)}(x)}{\phi^{(1)}(x)} \geq 1$$

and let $F^j \geq_{rh} F^i$, $i, j \in \{1, \ldots, d\}$, then

$$s^i(x) \geq s^j(x).$$

Proof. For $s^i(x) \geq s^j(x)$ we require

$$f^i(x)\, \phi^{(1)}\big(F^i(x)\big) \geq f^j(x)\, \phi^{(1)}\big(F^j(x)\big).$$

Rewriting this we obtain

$$\frac{f^i(x) F^j(x)}{f^j(x) F^i(x)} \leq \frac{F^j(x)\, \phi^{(1)}\big(F^j(x)\big)}{F^i(x)\, \phi^{(1)}\big(F^i(x)\big)}$$

since $\phi^{(1)}(\cdot) \leq 0$. By definition of the reversed hazard rate order it holds that $F^i \leq_{rh} F^j$ if and only if $\frac{F^j(x)}{F^i(x)}$ is increasing in x or equivalently $f^j(x)F^i(x) \geq f^i(x)F^j(x)$. Thus, it suffices to require

$$\frac{F^j(x)\,\phi^{(1)}\big(F^j(x)\big)}{F^i(x)\,\phi^{(1)}\big(F^i(x)\big)} \geq 1 \quad \text{or} \quad F^j(x)\,\phi^{(1)}\big(F^j(x)\big) \leq F^i(x)\,\phi^{(1)}\big(F^i(x)\big).$$

Since $F^i(x) \geq F^j(x)$ for all x (cf. Müller and Stoyan (2002), Theorem 1.3.14) the condition that $t\phi^{(1)}(t)$ is increasing, is sufficient. That is,

$$R_\phi(t) = -\frac{t\phi^{(2)}(t)}{\phi^{(1)}(t)} \geq 1.$$

□

Indeed there are copulas satisfying the condition in Proposition 2.4.6. For example the Clayton copula with generator $\phi(t) = t^{-\varrho} - 1$, $\varrho > 0$, $t \in [0, \infty)$ satisfies condition $R_\phi \geq 1$.

2.5 Minimizing ruin probability

The concept of ruin for a multivariate risk process could have different interpretations when compared to the univariate risk process. Here, ruin occurs if the multivariate risk process enters some domain of $\mathbb{R}^d$, the so-called insolvency region. In this section we show that the problem of minimizing the probability of ruin for a multi-line insurance company can be imbedded in the framework of discrete-time stochastic dynamic programming. According to Schäl (2004) and Schäl (2005) we show that the probability of ruin can be written as total cost without discounting which has to be paid once entering the insolvency region. As in the one-dimensional model the lack of discounting of the value function and the discontinuity of the system function lead to some special features that have to be considered.

2.5.1 Insolvency region

As mentioned before there are several definitions of ruin for a multidimensional insurance model. In the following, we choose the insolvency region as general as possible, only requiring the insolvency region $\mathcal{A} \subset \mathbb{R}^d$ to be open.

Example 2.5.1. The least restrictive choice of the insolvency region would be

$$\mathcal{A} = \{y \in \mathbb{R}^d : \max(y^1, \ldots, y^d) < 0\},$$

i.e. to require that all business lines are simultaneously below zero. A more restrictive choice is to take

$$\mathcal{A} = \{y \in \mathbb{R}^d : y^1 + \ldots + y^d < 0\},$$

which corresponds to the classical univariate ruin problem for the global company. Thus ruin occurs when the sum of net values of the business units are negative. But as capital is not completely versatile between different business lines, ruin may already occur before the aggregated reserves become negative. That is, ruin might occur when at least one business gets ruined and the solvency region is chosen to be

$$\mathcal{A} = \{y \in \mathbb{R}^d : \min(y^1, \dots, y^d) < 0\}.$$

More generally, Hult and Lindskog (2006) propose that capital may be transferred between the business lines. They define ruin as the situation when negative positions in one or several lines of business cannot be balanced by capital transfer.

2.5.2 Formulation as stochastic dynamic programme

We imbed our insurance model in the theory of discrete-time stochastic dynamic programming. Let us therefore choose

1. the state space as
$$S = \mathbb{R}^d \cup \{-\infty\}^d,$$
that means, if the risk process reaches an insolvency region the system moves to the absorbing state $(-\infty, \dots, -\infty)$.

The risk process of our insurance company can be controlled both by reinsurance and investment in a financial market. Therefore, we choose

2. the control space as
$$A = [\underline{b}, \bar{b}]^d \times \mathbb{R}^d$$
and the set of all admissible controls as
$$D(x) = [\underline{b}, \bar{b}]^d \times \triangle(x),$$
where $\triangle(x)$ denotes the set of all admissible portfolio vectors in state x,
$$\triangle(x) \quad = \quad \{\delta \in \mathbb{R}^d : 0 \le \delta^i \le \alpha x^i, i \in \{1, \dots, d\}\}$$
where $0 \le \alpha \le 1$ is a constant. In order to obtain a compact constraint space which is needed for the structure assumption, we do not allow for negative amounts of δ^i. Thus short selling of stocks is not possible.

The transition equation for the risk process of business line $i \in \{1, \dots, d\}$ is

$$X^i_{n+1} = \begin{cases} X^i_n + H^i(b^i_n, Y^i_{n+1}, Z_{n+1}, K_{n+1}) + \delta^i_n R^i_{n+1} K_{n+1} & \text{for } X_n \in \mathcal{A}^c \\ -\infty & \text{for } X_n \in \mathcal{A}. \end{cases} \tag{2.15}$$

where

$$H^i(b^i, y^i, z, k) = c^i(b^i)z - h(b^i, y^i)(1-k).$$

Choosing $a = (a^1, \dots, a^d)$ where $a^i = (b^i, \delta^i)$ and $w = (w^1, \dots, w^d)$ where $w^i = (\rho^i, y^i, z, k)$

3. the transition function $F^{tr} : D \times E \to S$ is given by

$$F^{tr}(x, a, w) = \begin{cases} \big(F^{tr,1}(x^1, a^1, w^1), \dots, F^{tr,d}(x^d, a^d, w^d)\big) & \text{for } x \in \mathcal{A}^c \\ (-\infty, \dots, -\infty) & \text{for } x \in \mathcal{A}, \end{cases}$$

and $F^{tr}(-\infty, \dots, -\infty, a, w) = (-\infty, \dots, -\infty)$
where

$$F^{tr,i}(x^i, a^i, w^i) = x^i + H^i(b^i, y^i, z, k) + \delta^i \rho^i k$$

is the transition function of business line $i \in \{1, \dots, d\}$.

4. The cost will be defined by

$$g(x, a) = V_0(x) = \mathbb{1}_{\{x \in \mathcal{A}\}}, \quad \alpha = 1$$

and $g(-\infty, \dots, -\infty, a) = V_0(-\infty, \dots, -\infty) = 0$.

We are interested in minimizing the probability of ruin specified by the insolvency region $\mathcal{A}$, that is

$$\Psi^\pi(x) = \mathbb{P}(X_n^{x,\pi} \in \mathcal{A} \text{ for some } n).$$

Therefore, the probability of reaching the insolvency region $\mathcal{A}$ after n periods can be defined by

$$\Psi_n^\pi(x) = \mathbb{P}(X_m^{x,\pi} \in \mathcal{A} \text{ for some } 0 \le m \le n),$$

in particular $\Psi_n^\pi(x) = 1$ for $x \in \mathcal{A}$ and $\Psi_n^\pi(-\infty, \dots, -\infty) := 0$. Let

$$\Psi^\pi(x) = \Psi_\infty^\pi(x).$$

By construction the ruin state is reached only once in our model. Thus, the probability of ruin can be written as total cost without discounting which has to be paid once entering the insolvency region.

$$\Psi_n^\pi(x) = \mathbb{E}\Big[\sum_{0 \le m < n} g\big(X_m, \varphi_m(X_m)\big) + V_0(X_n)\Big].$$

As usual in stochastic optimization it is convenient to introduce the following operators for $\varphi = (b, \delta)$ and any bounded measurable function $v : \mathbb{R}^d \mapsto \mathbb{R}$

$$U_\varphi v(x) = \begin{cases} \dfrac{\tilde\lambda}{\lambda + \tilde\lambda}\mathbb{E}\big[v\big(F^{tr}(x, \varphi, R, Y, Z, 1)\big)\big] & \text{for } x \in \mathcal{A}^c \\ \qquad + \dfrac{\lambda}{\lambda + \tilde\lambda}\mathbb{E}\big[v\big(F^{tr}(x, \varphi, R, Y, Z, 0)\big)\big] & \\ 1 & \text{for } x \in \mathcal{A} \end{cases}$$

and

$$Uv(x)=\inf_{\varphi\in D(x)} U_\varphi v(x).$$

2.5.3 Examining the structure assumption

If our model meets the structure assumption, Theorem 2.1.10 holds which yields more information about the ruin probability. For example, the ruin probability can be described as a fixed point of an operator and moreover, an optimality criterion for the optimal policy can be provided. Let us first state the continuity assumption

Continuity assumption
The functions $c^i(b^i)$ and $h(b^i, y^i)$ are continuous in b^i for all y^i and $i \in \{1, \dots, d\}$.

As in the one-dimensional case there is a discontinuity of the system function $F^{tr}(x, a, w)$ at $x \in \partial\mathcal{A}$. Thus, the usual continuity assumption is not satisfied. As in Schäl (2004) and Schäl (2005) we deal with this problem by choosing a suitable class $\mathcal{V}$ in the structure assumption. We define here

$$\mathcal{V} := \big\{v : [-\infty, \infty)^d \to [0, 1] : v \text{ l.s.c. on } \mathcal{A}^c,\ v(x) = 1 \text{ for } x \in \mathcal{A},\\ v(-\infty, \dots, -\infty) = 0\big\}.$$

We now show that the structure assumption for the infinite horizon model is satisfied and thus the main theorem, Howard's improvement theorem and verification theorem (cf. Theorems 2.1.10, 2.1.11, 2.1.12) can be applied in our context.

Proposition 2.5.2. *The structure assumption holds under the continuity assumption.*

Proof. Clearly, we have $V_0 \in \mathcal{V}$ and the space of admissible actions in state x, $D(x)$, is compact by construction of our model since we do not allow for short-selling. Thus, it remains to show that $Uv \in \mathcal{V}$ for all $v \in \mathcal{V}$ and that $Lv(x, a)$ is lower semi-continuous in (x, a) for all $x \in S$ and $v \in \mathcal{V}$.
Let $v \in \mathcal{V}$. Then v is lower semi-continuous on $\mathcal{A}^c$ and since the set $\mathcal{A}$ is open, $v(x) = \mathbb{1}_{\{x\in\mathcal{A}\}}$ is lower semi-continuous on $\mathcal{A}$. Therefore, v is lower semi-continuous on $\mathbb{R}^d$. The continuity of $F^{tr}(x, a, w)$ in $(x, a) \in \mathcal{A}^c \times D(x)$ yields $v(F^{tr}(x, a, w))$ is lower semi-continuous in $(x, a) \in \mathcal{A}^c \times D(x)$. Applying Fatou's lemma we obtain by definition of lower semi-continuity

$$\liminf_{(\tilde{x},\tilde{a})\to(x,a)} \mathbb{E}\big[v\big(F^{tr}(\tilde{x}, \tilde{a}, w)\big)\big] \geq \mathbb{E}\big[\liminf_{(\tilde{x},\tilde{a})\to(x,a)} v\big(F^{tr}(\tilde{x}, \tilde{a}, w)\big)\big]\\ \geq \mathbb{E}\big[v\big(F^{tr}(x, a, w)\big)\big],$$

i.e. $Lv(x,a) = \mathbb{E}[v(F^{tr}(x,a,w)]$ is lower semi-continuous in $(x,a) \in \mathcal{A}^c \times D(x)$. From the compactness and continuity of $D(x)$ we conclude the lower semi-continuity of $Uv(x) = \min_{a \in D(x)} Lv(x,a)$ in $x \in \mathcal{A}^c$ (cf. Hernández-Lerma and Lasserre (1996), Proposition D.5). Furthermore, we know that $Lv(x,a) = 1$ for $x \in \mathcal{A}$ and $Lv(x,a) = 0$ for $x = (-\infty, \dots, -\infty)$. Thus, the assertion follows. □

Remark 5. (a) The uniform increase assumption ($LV_0(x,a) \geq V_0(x)$ for all $a \in D(x)$ and for all $x \in S$) is clearly fulfilled by construction of our model.

(b) The structure assumption implies the existence of a measurable function φ such that $U_\varphi v(x) = Uv(x)$ for all $v \in \mathcal{V}$. This is a result of the selection theorem by Brown and Purves (1973) (cf. Corollary 1).

2.5.4 Howard's improvement theorem

In this subsection we show that there is a contraction property for the ruin probability which is strong enough for the application of Howard's improvement theorem and verification theorem (cf. Theorems 2.1.11, 2.1.12) in spite of the lack of discounting. However, we have to specify the insolvency region, i.e.

$$\mathcal{A} := \{y \in S : \min(A^1 y, \dots, A^r y) < 0, A^j \in \mathbb{R}_+^{e \times d}, j \in \{1, \dots, r\}\}.$$

Note that Example 2.5.3 still holds in our context. We assume the following net profit condition to hold.

Net profit condition:

$$(c - \lambda \mathbb{E} Y) \geq 0.$$

For notational convenience we set

$$H(b,y,z,k) = \begin{pmatrix} H^1(b^1, y^1, z, k) \\ \dots \\ H^d(b^d, y^d, z, k) \end{pmatrix}$$

and

$$G(b,\delta,y,r,z,k) = \begin{pmatrix} G^1(b^1, \delta^1, y^1, r^1, z, k) \\ \dots \\ G^d(b^d, \delta^d, y^d, r^d, z, k) \end{pmatrix}$$

where

$$G^i(b^i, \delta^i, y^i, r^i, z, k) = \delta^i r^i k + H^i(b^i, y^i, z, k).$$

In the sequel we exclude cheap reinsurance and we assume the following natural contraction property to hold.

Contraction property (C):

$$\mathbb{P}\big(H(b,Y,Z,0)\in\mathcal{A}\big)>0 \text{ for all } b\in[\underline{b},\overline{b}]^d.$$

Let us give an example under which the contraction condition is fulfilled.

Example 2.5.3. We consider an insurance company consisting of two dependent lines of business allowing for proportional reinsurance. Claims of the two branches are exponentially distributed with parameter λ^1, λ^2 and dependence is modeled via a Clayton copula with parameter $\varrho=1$. If it is not possible to cancel all negative positions the insurance company is insolvent and ruin occurs. That is, we assume the insolvency region is

$$\mathcal{A}=\{x=(x^1,x^2)\in\mathbb{R}^2 : x^1+x^2<0\}.$$

In order to prove that the contraction property holds even in the case the occurrence of claims in the single business lines are dependent, we have to show that

$$\mathbb{P}\big((c^1(b^1)+c^2(b^2))Z-b^1Y^1-b^2Y^2<0\big)>0.$$

First note that by Sklar's theorem and modeling dependence via a Clayton copula with parameter $\varrho=1$ it holds

$$\begin{aligned}&\mathbb{P}\big(Y^1\le y^1,\,Y^2\le y^2\big)=C\big(F^1(y^1),F^2(y^2)\big)\\ =&\big((1-\exp(-\lambda^1y^1))^{-1}+(1-\exp(-\lambda^2y^2))^{-1}-1\big)^{-1}\\ =&\frac{1-\exp(-\lambda^2y^2)-\exp(-\lambda^1y^1)+\exp(-\lambda^2y^2-\lambda^1y^1)}{1-\exp(-\lambda^1y^1-\lambda^2y^2)}\end{aligned}\tag{2.16}$$

assuming that claims are exponentially distributed with parameter λ^1, λ^2. Since the period length Z is exponentially distributed we have

$$\begin{aligned}&\mathbb{P}\big(b^1Y^1+b^2Y^2>(c^1(b^1)+c^2(b^2))Z\big)\\ =&\int_0^\infty \lambda e^{-\lambda z}\mathbb{P}\big(b^1Y^1+b^2Y^2>(c^1(b^1)+c^2(b^2))z\big)\,dz.\end{aligned}$$

Moreover,

$$\mathbb{P}\big(b^1Y^1+b^2Y^2>(c^1(b^1)+c^2(b^2))z\big)\ge\mathbb{P}\big(b^1Y^1>c^1(b^1)z\,,b^2Y^2>c^2(b^2)z\big).$$

For notational simplicity we define $y^1=\frac{c^1(b^1)z}{b^1}$ and $y^2=\frac{c^2(b^2)z}{b^2}$. Note that $b>0$ since we exclude cheap reinsurance. It remains to show that

$$\begin{aligned}&\mathbb{P}(Y^1>y^1,Y^2>y^2)\\ =&1-\mathbb{P}(Y^1\le y^1)-\mathbb{P}(Y^2\le y^2)+\mathbb{P}(Y^1\le y^1,Y^2\le y^2)>0.\end{aligned}$$

Using (2.16) and the fact that claims are exponentially distributed we know that this holds if and only if

$$\exp(-\lambda^1 y^1) - 1 + \exp(-\lambda^2 y^2) \\ + \frac{1 - \exp(-\lambda^2 y^2) - \exp(-\lambda^1 y^1) + \exp(-\lambda^2 y^2 - \lambda^1 y^1)}{1 - \exp(-\lambda^1 y^1 - \lambda^2 y^2)} > 0$$

which is equivalent to

$$\begin{aligned} &\big(\exp(-\lambda^1 y^1) - 1 + \exp(-\lambda^2 y^2)\big)\big(1 - \exp(-\lambda^1 y^1 - \lambda^2 y^2)\big) \\ &+ 1 - \exp(-\lambda^2 y^2) - \exp(-\lambda^1 y^1) + \exp(-\lambda^2 y^2 - \lambda^1 y^1) \\ =& \exp(-\lambda^2 y^2 - \lambda^1 y^1)\big(2 - \exp(-\lambda^1 y^1) - \exp(-\lambda^2 y^2)\big) > 0 \end{aligned}$$

and the contraction property holds in this example.

For notational convenience we set $I_J := \{1, \dots, r\}$, $I_K := \{1, \dots, e\}$ and $M_{KJ} := \{j : I_K \to I_J\}$. The k-th entry of the vector $(\cdot)$ is denoted by $(\cdot)^k$.

Lemma 2.5.4. *Let $\xi : (0, \infty) \to (0, 1]$ be a measurable and integrable function. Then it holds*

(a) The function

$$\begin{aligned} d(b, x) :=& \mathbb{E}\big[\xi(Z)\mathbb{1}_{\{\cap_{k\in I_K}(A^{j(k)}(H(b,Y,Z,0)-x))^k<0\}}\big] \\ =& (\lambda + \tilde{\lambda}) \int \xi(z)\, \mathbb{P}\Big(\bigcap_{k\in I_K} \big(A^{j(k)}(H(b, Y, z, 0) - x)\big)^k < 0\Big)\, dz \end{aligned}$$

is lower semi-continuous in $b \in \mathbb{R}^d$ for all $x \in \mathbb{R}^d$ and some $j \in M_{KJ}$.

(b) There is some $\epsilon \in \mathbb{R}^d_+$ such that

$$\inf_{b\in[\underline{b},\overline{b}]^d} d(b, -\epsilon) > 0.$$

Proof. (a) Since Z and $Y = (Y^1, \dots, Y^d)$ are independent by construction of our model the first equality holds. The function $H^i(b^i, y^i, z, 0) = c^i(b^i)z - h^i(b^i, y^i)$ is continuous in b^i by the continuity assumption. Using this and the fact that the intersection of finite open sets is open, we know that the indicator function $\mathbb{1}_{\{\cap_{k\in I_K}(A^{j(k)}(H(b,y,z,0)-x))^k<0\}}$ is lower semi-continuous in $b \in \mathbb{R}^d$ for all y, z. The assertion follows from Fatou's lemma.

(b) Since $\xi \in (0, 1]$ the function $d(b, x)$ is increasing in x. Using the lower semi-continuity of $d(b, x)$ in $b \in [\underline{b}, \overline{b}]^d$, the compactness of $[\underline{b}, \overline{b}]^d$ and the fact that $d(b,x)$ is increasing in x we can apply a variant of Dini's theorem (cf. Schäl (1975), Proposition 10.1) and obtain

$$\lim_{\|\epsilon\|\to 0} \inf_{b\in[\underline{b},\overline{b}]^d} d(b,-\epsilon) = \inf_{b\in[\underline{b},\overline{b}]^d} d(b,0) = d(b_0,0) \text{ for an } b_0 \in [\underline{b},\overline{b}]^d.$$

By the contraction property (C) the last expression is strictly positive for all $b_0 \in [\underline{b},\overline{b}]^d$.

□

Furthermore, let us state an assumption related to the usual no-arbitrage assumption which excludes portfolios that make profit without risk.

Weak no-arbitrage assumption (NA):
For all $\delta \in \mathbb{R}^d$ and for all $z \in \mathbb{R}_+$ it holds

$$\mathbb{P}\big((\delta^1 R^1,\ldots,\delta^d R^d) \le 0 | Z = z\big) > 0.$$

Lemma 2.5.5. *(a) For all $z \in \mathbb{R}_+$ we have*

$$\inf_{\delta\in\triangle(x)} \mathbb{P}\big((\delta^1 R^1,\ldots,\delta^d R^d) \le 0 | Z = z\big) > 0.$$

(b) There is some $\epsilon \in \mathbb{R}_+^d$ and some $j \in M_{KJ}$ such that

$$\inf_{(b,\delta)\in[\underline{b},\overline{b}]^d\times\triangle(x)} \mathbb{P}\Big(\bigcap_{k\in I_K} \big(A^{j(k)}\big(G(b,\delta,Y,R,Z,K)+\epsilon\big)\big)^k < 0\Big) > 0$$

Proof. (a) By the weak no-arbitrage assumption we have for all $\delta \in \mathbb{R}^d$

$$\mathbb{P}\big((\delta^1 R^1,\ldots,\delta^d R^d) \le 0 | Z = z\big) > 0.$$

As in the preceding lemma it can be shown that the function

$$\delta \mapsto \mathbb{P}\big(\delta^1 R^1 < 0,\ldots,\delta^d R^d < 0 | Z = z\big)$$

is lower semi-continuous. Since $\triangle(x)$ is compact, the infimum is attained on $\triangle(x)$.

(b) With part (a) we have for all $z \in \mathbb{R}_+$

$$\xi(z) := \inf_{\delta\in\triangle(x)} \mathbb{P}\big((\delta^1 R^1,\ldots,\delta^d R^d) \le 0 | Z = z\big) > 0.$$

Set

$$\delta R = (\delta^1 R^1,\ldots,\delta^d R^d).$$

Thus it holds

$$\begin{aligned}
&\mathbb{P}\big(\bigcap_{k\in I_K} \big(A^{j(k)}(G(b,\delta,Y,R,Z,K)+\epsilon)\big)^k < 0\big) \\
\geq &\mathbb{P}\Big(\delta RK \leq 0, \bigcap_{k\in I_K} \big(A^{j(k)}(G(b,\delta,Y,R,Z,K)+\epsilon)\big)^k < 0\Big).
\end{aligned}$$

Since $A^j \in \mathbb{R}_+^{e\times d}$ for all $j \in \{1,\ldots,d\}$ we have

$$\begin{aligned}
&\mathbb{P}\Big(\delta RK \leq 0, \bigcap_{k\in I_K} \big(A^{j(k)}(G(b,\delta,Y,R,Z,K)+\epsilon)\big)^k < 0\Big) \\
\geq &\mathbb{P}\Big(\delta RK \leq 0, \bigcap_{k\in I_K} \big(A^{j(k)}(H(b,Y,Z,K)+\epsilon)\big)^k < 0\Big) \\
\geq &\mathbb{P}\Big(\delta R \leq 0, \bigcap_{k\in I_K} \big(A^{j(k)}(H(b,Y,Z,K)+\epsilon)\big)^k < 0\Big).
\end{aligned}$$

By the independence assumptions of R, Y and K we obtain

$$\begin{aligned}
&\mathbb{P}\Big(\delta R \leq 0, \bigcap_{k\in I_K} \big(A^{j(k)}(H(b,Y,Z,K)+\epsilon)\big)^k < 0\Big) \\
=&(\lambda+\tilde\lambda)\int_0^\infty \mathbb{P}\big(\delta R \leq 0 | Z = z\big) e^{-(\lambda+\tilde\lambda)z} \\
&\qquad\qquad \mathbb{P}\Big(\bigcap_{k\in I_K} \big(A^{j(k)}(H(b,Y,z,K)+\epsilon)\big)^k < 0\Big)\, dz \\
\geq&(\lambda+\tilde\lambda)\int_0^\infty \xi(z)\, e^{-(\lambda+\tilde\lambda)z} \\
&\qquad\qquad \mathbb{P}\Big(K=0, \bigcap_{k\in I_K} \big(A^{j(k)}(H(b,Y,z,0)+\epsilon)\big)^k < 0\Big)\, dz \\
=&(\lambda+\tilde\lambda)\mathbb{P}(K=0) \\
&\quad \int_0^\infty \xi(z)\, e^{-(\lambda+\tilde\lambda)z}\mathbb{P}\Big(\bigcap_{k\in I_K} \big(A^{j(k)}(H(b,Y,z,0)+\epsilon)\big)^k < 0\Big)\, dz.
\end{aligned}$$

The assertion follows from Lemma 2.5.4 for an $\epsilon \in \mathbb{R}_+^d$.

□

For notational convenience we define

$$\overline{\mathcal{A}}_M := \{y \in S : \forall k \in I_K\ \exists j \in M_{KJ} \text{ such that } (A^{j(k)}y)^k \leq (M^{j(k)})^k\}.$$

Proposition 2.5.6. *Let $M \in \mathbb{R}_+^{e\times r}$ be arbitrary.*

(a) There is an $n \in \mathbb{N}$ such that

$$\sup\big\{\mathbb{P}\big(X_n^{x,\pi} \in \mathcal{A}^c\big) : x \in \overline{\mathcal{A}}_M,\, \pi\big\} < 1;$$

(b) It holds

$$\mathbb{P}\big(X_m^{x,\pi} \in \mathcal{A}^c \text{ and } X_m^{x,\pi} \in \overline{\mathcal{A}}_M, \text{ for infinitely many } m\big) = 0.$$

Proof. (a) Choose $\epsilon \in \mathbb{R}_+^d$ as in Lemma 2.5.5 and an arbitrary policy π. Consider in the following the event $\{X_m \in \mathcal{A}^c, 0 \leq m < n\}$. By the transition equation (2.15) it holds

$$X_n^i = X_0^i + \sum_{m=0}^{n-1} \delta_m^i R_{m+1}^i K_{m+1} + H^i(b_m^i, Y_{m+1}^i, Z_{m+1}, K_{m+1}).$$

Further define

$$G_m^i := \delta_m^i R_{m+1}^i K_{m+1} + H^i(b_m^i, Y_{m+1}^i, Z_{m+1}, K_{m+1}).$$

Recall that by Lemma 2.5.5

$$\inf_{(b,\delta)\in[\underline{b},\overline{b}]^d \times \triangle(x)} \mathbb{P}\Big(\bigcap_{k\in I_K} \big(A^{j(k)}(G_m^1 + \epsilon^1, \dots, G_m^d + \epsilon^d)^T\big)^k < 0\Big) =: \nu > 0. \tag{2.17}$$

By induction we obtain

$$\begin{aligned}&\mathbb{P}\Big(\bigcap_{k\in I_K} \Big(A^{j(k)}\Big(\sum_{m=0}^{n-1} G_m^1 + n\epsilon^1, \dots, \sum_{m=0}^{n-1} G_m^d + n\epsilon^d\Big)^T\Big)^k < 0\Big)\\ \geq &\mathbb{P}\Big(\bigcap_{m=0}^{n-1}\bigcap_{k\in I_K} \big(A^{j(k)}(G_m^1 + \epsilon^1, \dots, G_m^d + \epsilon^d)^T\big)^k < 0\Big).\end{aligned}$$

By (2.17) it therefore holds

$$\mathbb{P}\Big(\bigcap_{k\in I_K} \Big(A^{j(k)}\Big(\sum_{m=0}^{n-1} G_m^1 + n\epsilon^1, \dots, \sum_{m=0}^{n-1} G_m^d + n\epsilon^d\Big)^T\Big)^k < 0\Big) \geq \nu^n. \tag{2.18}$$

Taking into consideration the definition of the insolvency region $\mathcal{A}$ we can rewrite the ruin probability as follows

$$
\begin{aligned}
&\mathbb{P}(X_n \in \mathcal{A})\\
&= \mathbb{P}\Big(\Big(\sum_{m=0}^{n-1} G_m^1 + x^1, \dots, \sum_{m=0}^{n-1} G_m^d + x^d\Big)^T \in \mathcal{A}\Big)\\
&=\mathbb{P}\Big(\min\Big(A^1\Big(\sum_{m=0}^{n-1} G_m^1 + x^1, \dots, \sum_{m=0}^{n-1} G_m^d + x^d\Big)^T, \dots,\\
&\qquad A^r\Big(\sum_{m=0}^{n-1} G_m^1 + x^1, \dots, \sum_{m=0}^{n-1} G_m^d + x^d\Big)^T\Big) < 0\Big)\\
&=\mathbb{P}\Big(\bigcap_{k\in I_K} \min\Big(\Big(A^1\Big(\sum_{m=0}^{n-1} G_m^1 + x^1, \dots, \sum_{m=0}^{n-1} G_m^d + x^d\Big)^T\Big)^k, \dots,\\
&\qquad \Big(A^r\Big(\sum_{m=0}^{n-1} G_m^1 + x^1, \dots, \sum_{m=0}^{n-1} G_m^d + x^d\Big)^T\Big)^k\Big) < 0\Big),
\end{aligned}
\tag{2.19}
$$

where the last equality holds since we have to consider the minimum componentwise. Since for all $k \in I_K$ it exists $j \in M_{KJ}$ such that $(A^{j(k)}x)^k \leq (M^{j(k)})^k \leq n(A^{j(k)}\epsilon)^k$ we get

$$
\begin{aligned}
&\min\Big(\dots, \Big(A^{j(k)}\Big(\sum_{m=0}^{n-1} G_m^1 + x^1, \dots, \sum_{m=0}^{n-1} G_m^d + x^d\Big)^T\Big)^k, \dots,\\
&\qquad \Big(A^r\Big(\sum_{m=0}^{n-1} G_m^1 + x^1, \dots, \sum_{m=0}^{n-1} G_m^d + x^d\Big)^T\Big)^k\Big)\\
&\leq \min\Big(\dots, \Big(A^{j(k)}\Big(\sum_{m=0}^{n-1} G_m^1 + n\epsilon^1, \dots, \sum_{m=0}^{n-1} G_m^d + n\epsilon^d\Big)^T\Big)^k, \dots,\\
&\qquad \Big(A^r\Big(\sum_{m=0}^{n-1} G_m^1 + x^1, \dots, \sum_{m=0}^{n-1} G_m^d + x^d\Big)^T\Big)^k\Big).
\end{aligned}
\tag{2.20}
$$

Clearly, (2.20) is smaller than

$$
\Big(A^{j(k)}\Big(\sum_{m=0}^{n-1} G_m^1 + n\epsilon^1, \dots, \sum_{m=0}^{n-1} G_m^d + n\epsilon^d\Big)^T\Big)^k.
$$

In (2.19) we therefore get

$$
\begin{aligned}
&\mathbb{P}(X_n \in \mathcal{A})\\
&\geq \mathbb{P}\Big(\bigcap_{k\in I_K} \Big(A^{j(k)}\Big(\sum_{m=0}^{n-1} G_m^1 + n\epsilon^1, \dots, \sum_{m=0}^{n-1} G_m^d + n\epsilon^d\Big)^T\Big)^k < 0\Big) \geq \nu^n
\end{aligned}
$$

where the last inequality follows from (2.18). Hence, $\mathbb{P}(X_n \in \mathcal{A}^c) \leq 1-\nu^n$ where for all $k \in I_K$ there exists $j(k) \in I_J$ such that $(A^{j(k)}x)^k \leq (M^{j(k)})^k \leq n(A^j\epsilon)^k$.

(b) Applying part (a) we obtain that

$$\eta_{n,M} := \sup_{X_m^{x,\pi} \in \overline{\mathcal{A}}_M,\, \pi} \mathbb{P}(X_{m+n}^{x,\pi} \in \mathcal{A}^c) < 1$$

where n is defined as in part (a). Defining the stopping times

$$\begin{aligned}\tau_0 :=&0,\\ \tau_{k+1} :=&\inf\big\{m \geq \tau_k + n | X_m^{x,\pi} \in \mathcal{A}^c \text{ and } X_m^{x,\pi} \in \overline{\mathcal{A}}_M\big\}\end{aligned}$$

and $\inf \emptyset := +\infty$ we have

$$\begin{aligned}&\mathbb{P}(\tau_1 < \infty, \dots, \tau_{k+1} < \infty)\\ =&\mathbb{P}(\tau_1 < \infty, \dots, \tau_k < \infty)\, \mathbb{P}(\tau_{k+1} < \infty | \tau_1 < \infty, \dots, \tau_k < \infty)\\ \leq&\mathbb{P}(\tau_1 < \infty, \dots, \tau_k < \infty)\, \mathbb{P}(X_{\tau_k+n}^{x,\pi} \in \mathcal{A}^c | \tau_1 < \infty, \dots, \tau_k < \infty)\\ \leq&\mathbb{P}(\tau_1 < \infty, \dots, \tau_k < \infty)\, \eta_{n,M}.\end{aligned}$$

Thus,

$$\begin{aligned}&\mathbb{P}\big(X_m^{x,\pi} \in \mathcal{A}^c \text{ and } X_m^{x,\pi} \in \overline{\mathcal{A}}_M \text{ for infinitely many } m\big)\\ =&\lim_{k\to\infty} \mathbb{P}(\tau_1 < \infty, \dots, \tau_k < \infty)\\ \leq&\mathbb{P}\big(X_m^{x,\pi} \in \mathcal{A}^c \text{ and } X_m^{x,\pi} \in \overline{\mathcal{A}}_M \text{ for infinitely many } m\big)\, \eta_{n,M}\end{aligned}$$

and the assertion follows.

□

From Proposition 2.5.6 we conclude by the continuity of the probability measure that for all $x, M > 0$ and policies π

$$\lim_{n\to\infty} \mathbb{P}\big(X_n^{x,\pi} \in \mathcal{A}^c, X_n^{x,\pi} \in \overline{\mathcal{A}}_M\big) = 0. \tag{2.21}$$

We are finally able to derive a weak contraction property in order to treat models without discounting.

Lemma 2.5.7 (Contraction theorem)**.** *If $\xi : [-\infty, \infty)^d \mapsto [0, \infty)$ is a bounded measurable function such that $\xi(x) = 0$ for $x \in \mathcal{A}$ and*

$$\lim_{n\to\infty} \xi(x^{(n)}) = 0$$

where $(x^{(n)})_{n\in\mathbb{N}}$ is an arbitrary sequence such that

$$\lim_{n\to\infty} ((A^1x^{(n)})^k, \dots, (A^r x^{(n)})^k) = (\infty, \dots, \infty),$$

for some $k \in \{1, \dots, e\}$ then

$$\lim_{n\to\infty} \mathbb{E}\big[\xi(X_n^{1,x_1,\pi}, \dots, X_n^{d,x_d,\pi})\big] = 0 \quad \textit{for all } x, \pi.$$

Proof. For every $\epsilon > 0$ we can choose a $M \in \mathbb{R}_+^{e\times r}$ such that $\xi(x) \leq \epsilon$ where it exists $k \in I_K$, $(A^j x)^k > (M^j)^k$ for all $j \in I_J$. Denote by $\|\xi\|$ the upper bound of ξ. Thus we have

$$\begin{aligned}
&\mathbb{E}\big[\xi(X_n^1,\dots,X_n^d)\big]\\
=&\mathbb{E}\big[\mathbb{1}_{\{X_n\in\mathcal{A}^c\}}\mathbb{1}_{\{\exists k\in I_K,\forall j\in I_J:\,(A^jX_n)^k>(M^j)^k\}}\xi(X_n^1,\dots,X_n^d)\big]\\
&+\mathbb{E}\big[\mathbb{1}_{\{X_n\in\mathcal{A}^c\}}\mathbb{1}_{\{\forall k\in I_K,\exists j(k)\in I_J:\,(A^{j(k)}X_n)^k\leq(M^{j(k)})^k\}}\xi(X_n^1,\dots,X_n^d)\big]\\
\leq&\epsilon+\|\xi\|\mathbb{E}\big[\mathbb{1}_{\{X_n\in\mathcal{A}^c\}}\mathbb{1}_{\{\forall k\in I_K,\exists j(k)\in I_J:\,(A^{j(k)}X_n)^k\leq(M^{j(k)})^k\}}\big].
\end{aligned}$$

The assertion follows from (2.21). □

We are now in the position to show that the contraction property is strong enough for the application of Howard's improvement theorem and the verification theorem (cf. Theorem 2.1.11, Theorem 2.1.12). Denote by φ^∞ the stationary policy under which the decision maker does not invest in the financial market and does not choose to reinsure its claims, i.e.

$$\varphi(x) = (\bar{b}, 0) \quad \text{for all } x \in \mathbb{R}^d.$$

The ruin probability is then denoted by $\Psi := \Psi^{\varphi^\infty}$.

Lemma 2.5.8. *Let $\varphi^\infty = (\bar{b}, 0)$ and let $(x^{(n)})_{n\in\mathbb{N}}$ be an arbitrary sequence such that $\lim_{n\to\infty}((A^1x^{(n)})^k,\dots,(A^rx^{(n)})^k) = (\infty,\dots,\infty)$ for some $k \in I_K$. Then*

(a) $\lim_{n\to\infty}\Psi(x^{(n)}) = 0$.

(b) $\Psi(x) - V_0(x) = 0$ *for* $x \in \mathcal{A}$ *and* $\lim_{n\to\infty}(\Psi(x^{(n)}) - V_0(x^{(n)})) = 0$.

Proof. (a) We have

$$\begin{aligned}
\Psi(x) =&\mathbb{P}\big(\{\min(A^1X_n,\dots,A^rX_n) < 0 \text{ for an } n\in\mathbb{N}\}\big)\\
=&\mathbb{P}\Big(\bigcap_{k=1}^{e}\{\min((A^1X_n)^k,\dots,(A^rX_n)^k) < 0 \text{ for an } n\in\mathbb{N}\}\Big)\\
\leq&\mathbb{P}\big(\{\min((A^1X_n)^k,\dots,(A^rX_n)^k) < 0 \text{ for an } n\in\mathbb{N}\}\big)\\
=&\mathbb{P}\Big(\bigcup_{j=1}^{r}\{(A^jX_n)^k < 0 \text{ for an } n\in\mathbb{N}\}\Big)\\
\leq&\sum_{j=1}^{r}\mathbb{P}(\{(A^jX_n)^k < 0 \text{ for an } n\in\mathbb{N}\})
\end{aligned}$$

for some $k \in I_K$ where $\Psi^{jk} := \mathbb{P}(\{(A^jX_n)^k < 0 \text{ for an } n\in\mathbb{N}\})$ corresponds to the one dimensional ruin probability of the risk reserve $(A^jX_n)^k$ with initial risk reserve $(A^jx)^k$, i.e.

$$(A^jX_n)^k = (A^jx)^k + \sum_{m=0}^{n-1}(A^jc)^kZ_{m+1} - (A^jY_{m+1})^k.$$

The net profit condition implies that

$$A^j(c - \lambda \mathbb{E}Y) \geq 0 \text{ for all } j \in I_J.$$

Therefore, the net profit condition for the one dimensional risk reserves $(A^j X_n)^k$ hold which ensures that $(A^j X_n)^k \to \infty$ as $n \to \infty$ (cf. Rolski et al. (1999), Theorem 6.3.1). Thus, $\Psi^{jk} \to 0$ as the initial risk reserve $(A^j x)^k \to \infty$ and the assertion follows.

(b) It is $V_0(x^1, \ldots, x^d) = 0$ for $(x^1, \ldots, x^d) \in \mathcal{A}^c$ and $\Psi(x^1, \ldots, x^d) = V_0(x^1, \ldots, x^d)$ for $(x^1, \ldots, x^d) \in \mathcal{A}$. The assertion follows from part (a).

□

Hence, the contraction property holds in our model and we may apply Howard's improvement theorem and the verification theorem. We can now check whether an insurance company can operate more successfully than to keep its risk reserve. More precisely, we aim to improve φ^∞ and find conditions such that it is the optimal policy.

Theorem 2.5.9. *Let φ and $\tilde{\varphi}$ be any decision function. Set $\Psi := \Psi^{\varphi^\infty}$, $\tilde{\Psi} := \Psi^{\tilde{\varphi}^\infty}$ and*

$$A(x, \varphi) := \big\{a \in D(x) : L\Psi(x, a) < \Psi(x)\big\} \quad \text{for all } x \in S.$$

Under the assumptions of this section we have

(a) Howard improvement
Let $\tilde{\varphi} \in A(x, \varphi)$ for some $x \in S$ and $\tilde{\varphi} = \varphi$ for other states $x \in S$ then

$$\begin{aligned} \tilde{\Psi}(x) \leq& \Psi(x) \quad \text{for all } x \in S \text{ and} \\ \tilde{\Psi}(x) <& \Psi(x) \quad \text{if } \tilde{\varphi} \in A(x, \varphi). \end{aligned}$$

(b) Verification theorem
If $\Psi = U\Psi = U_\varphi \Psi$ then φ defines a stationary, optimal policy φ^∞.

Chapter 3

Basic concepts of Lévy processes

Lévy processes are named in honors of the French mathematician Paul Lévy (1886 – 1971). Paul Lévy explored a major part of modern stochastic process theory parallel with, but independently from, the Russian mathematicians Andrei N. Kolmogorov (1903 – 1987) and Aleksandr Yakovlevich Khinchin (1894 – 1959). His student Loève (cf. Loève (1973)) gives a vivid description of Lévy's work:

> *"Paul Lévy was a painter in the probabilistic world. Like the very great painting geniuses, his palette was his own and his paintings transmuted forever our vision of reality [...]."*

In this chapter we present an introduction to the theory of Lévy processes and discuss those properties of a Lévy process which are required for Lévy driven risk processes in an insurance model. Moreover, we introduce two fundamental tools, the Lévy-Khinchin formula allowing to study distributional properties of a Lévy process and the Lévy-Itô decomposition describing the structure of its sample paths. However, a more thorough treatment can be found for example in Applebaum (2004), Protter (1990), Sato (1999) or in a more recent book by Cont and Tankov (2004). Finally, we recall conditions under which existence and uniqueness of solutions of Lévy stochastic differential equations can be guaranteed. These results enable us to derive a maximum inequality for a certain power of the solution of a Lévy stochastic differential equation which is beneficial for the correct formulation of the stochastic control problem in Chapter 5. So far, such a maximum inequality only exists for the pure diffusion case. Throughout this thesis we consider risk processes of an insurance company only, thus it suffices to treat the case of a finite time horizon $T \in [0, \infty)$.

3.1 Definitions and basic properties

Let $(\Omega, \mathbb{F}, (\mathcal{F}_t)_{0\le t\le T}, \mathbb{P})$ be a filtered probability space satisfying the usual hypotheses of completeness and right-continuity (cf. Protter (1990), Chapter I). In the following we will always assume this probability space as given. As we frequently need the notion of a martingale we briefly recall its definition. The term martingale roots in French language and originally describes a tack being used on horses to control their head carriage. It shall avoid the horse raising too high its head. Later, the term martingale also referred to betting strategies which were popular in 18th century France. In probability theory, the notion martingale was introduced by Paul Lévy and further developed by Joseph L. Doob (1910 – 2004).

Definition 3.1.1. *A càdlàg $\mathcal{F}_t$-adapted process $(X_t)_{t\ge 0}$ satisfying the integrability condition $\mathbb{E}\|X_t\| < \infty$ for all $t \ge 0$ is said to be a martingale if for all $0 \le s < t < \infty$,*

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s \quad a.s.$$

As in Protter (1990), Chapter I and Cont and Tankov (2004), Chapter 3 we define a Lévy process as follows.

Definition 3.1.2. *An adapted (càdlàg) stochastic process $X = (X_t)_{0\le t\le T}$ with values in $\mathbb{R}^d$ such that $X_0 = 0$ is called a Lévy process if it possesses the following properties:*

(i) Independent increments:

$$X_t - X_s \text{ is independent of } \mathcal{F}_s, \quad 0 \le s < t \le T.$$

(ii) Stationary increments:

$$X_{t+h} - X_t \stackrel{d}{=} X_h, \quad 0 \le t < t+h \le T,$$

that is, the law of $X_{t+h} - X_t$ does not depend on t.

(iii) Stochastic continuity:

$$\forall \epsilon > 0, \lim_{h\to 0} \mathbb{P}(|X_{t+h} - X_t| \ge \epsilon) = 0,$$

that is, processes with jumps at a fixed (non-random) time t occur only with zero probability.

The property "càdlàg" (continu à droite, limites à gauche) means that a stochastic process has a.s. sample paths which are right continuous and have left limits. This property in the definition of a Lévy process is not necessary since it can be shown that every Lévy process has a unique modification which is càdlàg (cf. Protter (1990), Theorem I.30).

Jumps of a Lévy process

Let us turn our attention to the discontinuities of a Lévy process. Since a Lévy process is càdlàg the only type of discontinuities are those caused by a jump which at time t are defined by

$$\triangle X_t = X_t - X_{t-} \text{ where } X_{t-} = \lim_{s \uparrow t} X_s \text{ is the left limit at } t.$$

The main challenge of a Lévy process arises from the fact that there may occur an infinite number of small jumps in every finite time interval, that is $\sum_{0 \leq s \leq t} \|\triangle X(s)\| = \infty$ a.s, $t > 0$.

Rather than analyzing the jumps themselves, we count jumps of specified size. The jump counter of a Lévy process can be described by

$$N_t(B) := \sum_{0<s\leq t} \mathbb{1}(\triangle X_s \neq 0)\mathbb{1}(\triangle X_s \in B), \quad t > 0$$

for every Borel-measurable set $B \in \mathbb{R}^d$. Thus for any Borel-measurable set $B \in \mathbb{R}^d$, $N_t(B)$ is the sum of all jumps taking values in the set B up to the time t. Since the paths of X are càdlàg, this is almost surely a finite sum whenever $0 \notin \overline{B}$ (the closure of B). It can be shown that $N_t(\cdot)$ and $\mathbb{E}[N_1(\cdot)]$ define σ-finite measures on $\mathbb{R}^d \setminus \{0\}$ (cf. Protter (1990), Theorem I.35). Let us now define the Lévy measure.

Definition 3.1.3. *The measure ν on $\mathbb{R}^d$ defined by*

$$\nu(B) = \mathbb{E}[N_1(B)], \quad B \in \mathcal{B}(\mathbb{R}^d)$$

is called the Lévy measure of the Lévy process X, $\nu(B)$ is the expected number of jumps, per unit of time, whose size belongs to $B \in \mathcal{B}(\mathbb{R}^d)$ where $\mathcal{B}(\mathbb{R}^d)$ denotes the Borel-σ-algebra over $\mathbb{R}^d$.

For a fixed Borel-measurable set B in $\mathbb{R}^d$ such that $0 \notin \overline{B}$, it can be shown that $(N_t(B))_{0\leq t\leq T}$ is a Poisson process (cf. Protter (1990), p.26 or Applebaum (2004) Theorem 2.3.5). Therefore, we know that $N_t(B)$ has intensity $t\nu(B)$ and $(N_t(B) - t\nu(B))_{0\leq t\leq T}$ is a martingale.
More generally, one considers a Poisson random measure $N(dt, dx)$ on $[0, T] \times \mathbb{R}^d$ (cf. Cont and Tankov (2004), Chapter 2.6). The connection between the Poisson process N_t and the random measure $N(dt, dx)$ is as follows

$$N_t(B) = \int\int_{[0,t]\times B} N(ds, dx) \qquad \text{for any set } B \in \mathcal{B}(\mathbb{R}^d) \text{ and } t \in [0, T].$$

The measure $dt\nu(dx)$ is called intensity measure or compensator of $N(dt, dx)$ and

$$\tilde{N}(dt, dx) := N(dt, dx) - dt\nu(dx)$$

is called the compensated measure.
We close this section with some properties used extensively later in this thesis. The next lemma is adapted from Applebaum (2004) Theorem 2.3.8, Sato (1999) Proposition 19.5 or (more elementary) from Cont and Tankov (2004) Chapter 2.6.

Lemma 3.1.4. *For a Borel-measurable set $B \in \mathbb{R}^d$ with $0 \notin \overline{B}$ and for a Borel-measurable function $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ such that*

$$\iint_{[0,T]\times B} |f(t,x)|\, dt\nu(dx) < \infty \tag{3.1}$$

it holds

$$\mathbb{E}[\iint_{[0,T]\times B} f(t,x)\, N(dt,dx)] = \iint_{[0,T]\times B} f(t,x)\, dt\nu(dx).$$

Under the assumptions of Lemma 3.1.4 one can show that integrals of the form

$$\iint_{[0,t]\times B} f(s,x)\tilde{N}(ds,dx) \quad 0 \le t \le T$$

are martingales. Observe that

$$\iint_{(s,t]\times B} f(s,x)\, N(ds,dx) \text{ is independent of } \mathcal{F}_s$$

for $0 \le s < t \le T$. Together with Lemma 3.1.4 this immediately implies the martingale property.
Let us finally discuss the relationship between infinitely divisible distributions and Lévy processes. We first cite the following definition from Kyprianou (2006).

Definition 3.1.5. *A $\mathbb{R}^d$-valued random variable Y with law μ is infinitely divisible if for all $n \in \mathbb{N}$ there exist i.i.d. random variables Y_{in}, $i \in \{1,\dots,n\}$, such that*

$$Y \overset{d}{=} \sum_{i=1}^{n} Y_{in}.$$

Or equivalently, the law μ of a $\mathbb{R}^d$- valued random variable is infinitely divisible if for all $n \in \mathbb{N}$, there exists a law μ_n such that

$$\mu = \mu_n^{*n},$$

where μ_n^* denotes the n-fold convolution of μ_n.

If (X_t) is a $\mathbb{R}^d$-valued Lévy process, then X_t is for any t infinitely divisible. To verify that we may write for any fixed t and $n \in \mathbb{N}$

$$X_t = (X_{t_n} - X_{t_{n-1}}) + \ldots + (X_{t_1} - X_{t_0}) \text{ for } t_k = \frac{kt}{n}.$$

Because of the stationarity and independence of the increments of a Lévy process this is the sum of n independent identically distributed random variables.

Theorem 3.1.6 (Infinite divisibility and Lévy processes)**.** *If $(X_t)_{0 \le t \le T}$ is a $\mathbb{R}^d$-valued Lévy process then for any fixed t, X_t has an infinitely divisible distribution.*
Conversely, if μ is an infinitely divisible distribution on $\mathbb{R}^d$ then there exists a Lévy process $(X_t)_{0 \le t \le T}$ such that the distribution of X_1 is given by μ.

Proof. The first part is shown above. For the converse statement we refer to Sato (1999), Theorem 7.10. □

3.2 Lévy-Itô decomposition

After having discussed the discontinuities of a Lévy process in the first section we now turn our attention to the formulation of the so-called Lévy-Itô decomposition. The Lévy-Itô decomposition probably is the most meaningful property of a Lévy process and certainly provides more structure than the definition since it reveals its building blocks. To be more precise, the structure of a Lévy process is described in terms of three independent Lévy processes with different behavior of their sample paths.

Theorem 3.2.1 (Lévy-Itô decomposition)**.** *Let X be a $\mathbb{R}^d$-valued Lévy process and ν its Lévy measure. Then the Lévy measure ν satisfies the integrability condition*

$$\int \min(\|x\|^2, 1)\, \nu(dx) < \infty.$$

Moreover, there exist a Brownian motion B with covariance matrix A and a (constant) vector $\mu \in \mathbb{R}^d$ such that

$$X_t = \mu t + B_t + X_t^l + \tilde{X}_t,$$

where

$$X_t^l = \int_0^t \int_{\{\mathbb{R}^d,\, \|x\|>1\}} x\, N(ds, dx), \qquad \text{and} \qquad \tilde{X}_t = \int_0^t \int_{\{\mathbb{R}^d,\, \|x\| \le 1\}} x\, \tilde{N}(ds, dx).$$

The terms B, X^l and $\tilde{X}^l$ are independent.

Proof. The original proof of Lévy, which has been completed by Itô, can be found in Lévy (1934) and Itô (1942). An outline of the proof can be found in Cont and Tankov (2004), Chapter 3.4. □

We call (μ, A, ν) in Theorem 3.2.1 the generating (or characteristic) triplet of a Lévy process. A is called Gaussian covariance matrix and ν the Lévy measure. The characteristics (μ, A, ν) of a Lévy process are uniquely determined by a Lévy process (cf. Applebaum (2004), Corollary 2.4.21).

Let us follow Cont and Tankov (2004) and some results in Protter (1990) for a brief comment on the jump processes of this decomposition.
Given a Lévy process X with Lévy measure ν, we know that $\nu(B)$ must be finite for every compact set B such that $0 \notin B$ (cf. Protter (1990), Theorem I.35). Thus the sum of the jumps with magnitude larger than one is a finite series and does not invoke any convergence problems. It can be shown that

$$X_t^l = \int_0^t \int_{\{\mathbb{R}^d,\, \|x\|>1\}} x\, N(ds, dx)$$

is a compound Poisson process (cf. Applebaum (2004), Theorem 2.3.10). There is nothing special about the threshold of jump amplitude being equal to one. Further note that for every $\epsilon > 0$, the process

$$\int_0^t \int_{\{\mathbb{R}^d,\, \epsilon<\|x\|<1\}} x\, N(ds, dx) = \sum_{0\leq s\leq t} \triangle X_s \mathbb{1}(\epsilon < |\triangle X_s| < 1)$$

is a well-defined compound Poisson process. However, ν can have infinitely many small jumps close to zero. Therefore, as $\epsilon \to 0$, this sum does not necessarily converge. As we have seen before, replacing the Poisson random measure $N(dt, dx)$ by its compensated measure $\tilde{N}(dt, dx) = N(dt, dx) - t\nu(dx)$ generates a martingale,

$$\tilde{X}_t = \int_0^t \int_{\{\mathbb{R}^d,\, \|x\|<1\}} x\, \tilde{N}(ds, dx),$$

which allows to apply Kolmogorov's three series Theorem (cf. Kallenberg (2002), Theorem 4.18) in order to show the desired convergence.
The process X^l (respectively $\tilde{X}$) could be interpreted as an infinite superposition of independent (compensated) Poisson processes.

The Lévy-Itô decomposition shows that every Lévy process is the sum of a Brownian motion with drift and a possibly infinite sum of independent compound Poisson processes. Therefore, every Lévy process can be approximated by a jump diffusion process, that is, by the sum of a Brownian motion with drift and a compound Poisson process.
Moreover, the Lévy-Itô decomposition shows that any Lévy process can be

decomposed into a martingale $B_t + \tilde{X}_t$ and a process $\mu t + X_t^l$ with paths of finite variation on compacts. Thus we can conclude that every Lévy process is a semimartingale (cf. Protter (1990), Theorems II.7/8/9).

Remark 6. Other choices than $\mathbb{1}_{\{\|x\| \leq 1\}}$ as truncation function can be used. Jumps larger than an arbitrary c might be truncated or one might even allow for a more general truncation function $g : \mathbb{R}^d \to \mathbb{R}$ which behaves as $1 + o(|x|)$ for $x \to 0$ and $O(\frac{1}{|x|})$ for $x \to \infty$. The covariance matrix A as well as the Lévy measure ν remain unaffected only μ changes with different choices of the truncation function

An important implication of the Lévy-Itô decomposition is the Lévy-Khinchin formula for the characteristic function of a Lévy process.

Theorem 3.2.2 (Lévy-Khinchin)**.** *Let X be a $\mathbb{R}^d$-valued Lévy process with characteristic triplet (μ, A, ν). Then it holds*

$$\mathbb{E}[e^{iuX_t}] = e^{t\psi(u)}$$

for $u \in \mathbb{R}^d$ where

$$\psi(u) = -\frac{1}{2}uAu + i\mu u + \int_{\mathbb{R}^d} \left(e^{iux} - 1 - iux\mathbb{1}_{\{\|x\| \leq 1\}}\right) \nu(dx).$$

Proof. A proof can be found in Cont and Tankov (2004), Theorem 3.1. □

In applications the additional assumption,

$$\int_{\{\mathbb{R}^d,\, \|x\| > 1\}} \|x\| \, \nu(dx) < \infty,$$

is often seen. Since there is no need to truncate the magnitude of large jumps, this assumption allows us to simplify the Lévy-Itô decomposition and the Lévy-Khinchin formula. We have $\mathbb{E}[X_1^l] < \infty$ and thus we can represent the Lévy-Itô decomposition in the following way

$$X_t = B_t + \tilde{M}_t + at, \quad 0 \leq t \leq T$$

where

$$\tilde{M}_t = \int_0^t \int_{\mathbb{R}^d} x\tilde{N}(ds, dx) \quad \text{and} \quad a = \mu + \int_{\|x\| \geq 1} x\nu(dx).$$

This decomposition of a Lévy process into a predictable process at and a martingale $B_t + \tilde{M}_t$ is called Doob-Meyer decomposition (cf. Kallenberg (2002), Chapter 25). Analogously, the characteristic exponent simplifies to

$$\psi(u) = -\frac{1}{2}uAu + iau + \int_{\mathbb{R}^d} \left(e^{iux} - 1 - iux\right) \nu(dx).$$

Path properties of a Lévy process

We finally state path properties of a Lévy process in terms of its characteristic triplet. Let X be a $\mathbb{R}^d$-valued Lévy process with generating triplet (μ, A, ν).

Proposition 3.2.3. *(i) (Continuity) Almost all paths of X are continuous if and only if $\nu \equiv 0$.*

(ii) (Piecewise constancy) Almost all paths of X are piecewise constant (it is a compound Poisson process) if and only if $A = 0$, $\nu(\mathbb{R}^d) < \infty$ and $\mu = \int_{\{\|x\| \leq 1\}} x\nu(dx)$.

Proof. A proof can be found in Sato (1999) Theorems 21.1/2. □

Moreover, Lévy processes can be distinguished by their jump activity.

Proposition 3.2.4. *(i) (Finite activity) If $\nu(\mathbb{R}^d) < \infty$ then almost all path of X have only finitely many jumps on any compact interval.*

(ii) (Infinite activity) If $\nu(\mathbb{R}^d) = \infty$ then almost all path of X have infinitely many jumps on any compact interval.

Proof. A proof can be found in Riesner (2006), Theorem 2.2.6. □

The Lévy-Itô decomposition implies that the path of a Lévy process is almost surely of infinite variation in the case that a Brownian motion is present. Let us therefore consider the case $A = 0$. Since the second term in the Lévy-Itô decomposition is a compound Poisson process and therefore of finite variation the third term decides whether the Lévy process is of finite or of infinite variation.

Proposition 3.2.5 (Finite variation)**.** *A Lévy process with generating triplet (μ, A, ν) is of finite variation if and only if*

$$A = 0 \quad \text{and} \quad \int_{\{\|x\| \leq 1\}} \|x\| \, \nu(dx) < \infty.$$

Proof. For a proof we refer to Cont and Tankov (2004) Proposition 3.9. □

3.3 Examples and actuarial interpretation

To conclude the short introduction to Lévy processes we consider some examples. The Brownian motion and the Poisson process are two fundamental examples of Lévy processes. The Brownian motion possesses almost sure continuous paths and normal distributed increments. It can be shown (cf. Proposition 3.2.3) that a Lévy process has continuous sample paths if and only if it is a Brownian motion (with drift). The Poisson process which is a counting process is an example for a Lévy process with discontinuous paths. It is used for the construction of more complex jump processes, for example the compound Poisson process. A compound Poisson process can be written as

$$X_t = \sum_{k=1}^{N_t} Y_k, \quad 0 \le t \le T,$$

where the jump sizes $Y_k \in \mathbb{R}^d$ are an i.i.d. sequence of random vectors with distribution having a density $f(x)$ with respect to the Lebesgue measure and (N_t) is a Poisson process with intensity λ independent from $(Y_k)_{k\geq 1}$. A compound Poisson process is the only Lévy process with almost sure piecewise constant sample paths (cf. Cont and Tankov (2004), Proposition 3.3). The Lévy measure of a compound Poisson process has a simple structure.

Proposition 3.3.1. *Let X be a compound Poisson process with intensity λ and claim size density $f(x)$. Then its Lévy measure is given by*

$$\nu(dx) = \lambda f(x)dx.$$

Proof. For a proof we refer to Cont and Tankov (2004), Proposition 3.5. □

This shows that a compound Poisson process is of finite activity. Moreover, every compound Poisson process can be written in terms of the Poisson random measure, that is

$$X_t = \int_0^t \int_{\mathbb{R}^d} x\, N(ds, dx). \tag{3.2}$$

where N is a Poisson random measure with intensity measure $\nu(dx)dt$.
The exponent in the Lévy-Khintchin formula for compound Poisson processes reduces to

$$\Psi(u) = \lambda \int_{\mathbb{R}^d} (e^{iux} - 1)\, f(dx).$$

In the classical Cramér-Lundberg model the risk process of an insurance company is modeled using the compound Poisson process. The Poisson process (N_t) describes the number of claims during the interval $(0, t]$. At each jump of (N_t) the insurance company has to pay out claims which are assumed to be independent and identically distributed. Then the total payment in the interval $(0, t]$ by the company is a compound Poisson process. For details on the classical risk model we refer to Grandell (1991).

Together with the Brownian motion compound Poisson processes constitute a jump diffusion. A Lévy process of jump diffusion type has the following form.

$$X_t = \mu t + \sigma W_t + \sum_{k=1}^{N_t} Y_k, \quad 0 \le t \le T,$$

where $\mu \in \mathbb{R}^d$, $\sigma > 0$, $(W_t)_{0\leq t\leq T}$ is a standard Brownian motion and $\sum_{k=1}^{N_t} Y_k$ is a compound Poisson process. Note that the jump diffusion process is a special case of the Lévy-Itô decomposition, compare Theorem 3.2.1 and the compound Poisson process representation (3.2). However, it is important to note that not every jump part of a Lévy process can be represented in this form since ν can have a singularity at zero meaning the Lévy process may have an infinite number of small jumps close to zero such that their sum does not necessarily converge.

So far, we restricted the examples mentioned to finite activity Lévy processes, that is $\nu(\mathbb{R}^d) < \infty$ (cf. Proposition 3.2.4). The most well-known examples of infinite activity Lévy processes are generated by Brownian subordination. To be more precise, a Brownian motion $(W_t)_{t\geq 0}$ with possible drift μ is evaluated on a stochastic time scale which is given by a subordinator $(S_t)_{t\geq 0}$. A subordinator is a Lévy process having trajectories which are almost surely increasing. It can therefore be interpreted as time change. A process generated by subordination is still a Lévy process (cf. Cont and Tankov (2004), Theorem 4.2), that is $X_t = \sigma W_{S_t} + \mu S_t$ is again a Lévy process. Examples of infinite activity processes generated by Brownian subordination include the inverse Gaussian model of Barndorff-Nielsen (1998) (infinite variation) and the variance gamma process (finite variation). Details to the variance gamma process can be found in Madan (2001).

We finally follow Kyprianou (2006), Section 2.7.1 to offer an actuarial interpretation of a spectrally negative Lévy process, that is $\nu(0,\infty) = 0$. Rearranging the terms in the Lévy-Khintchin exponent we have

$$\begin{aligned}\psi(u) =&\Big(-\frac{1}{2}uAu\Big) + \Big(i\mu u + \int_{\{\mathbb{R}^d_-,\,\|x\|>1\}} (e^{iux} - 1)\,\nu(dx)\Big)\\ &+ \Big(\int_{\{\mathbb{R}^d_-,\,\|x\|\leq 1\}} (e^{iux} - 1 - iux)\,\nu(dx)\Big).\end{aligned}$$

Let us focus on the infinite activity case in which case the characteristics are definitely different from the Cramér-Lundberg model. The second part of the Lévy-Khintchin exponent represents large claims meaning claims of size larger than one. Analogously to the Cramér-Lundberg model these claims are counterbalanced by some constant premium income $\mu > 0$. The first part might be interpreted as perturbation of claims and incoming premium payments. Finally, the third term represents a countably infinite number of arbitrarily small claims which are counterbalanced by a positive drift which can be interpreted as accumulated premiums over an infinite number of insurance contracts. Note that the first and third parts of the corresponding Lévy process are martingales. Therefore, the long term behavior of the risk process depends on the second term.

3.4 Martingale and Markov property

Directly from the Lévy-Itô-decomposition and the Lévy-Khinchin formula we get conditions under which a Lévy process or its exponential is a martingale.

Proposition 3.4.1. *Let X be a $\mathbb{R}^d$-valued Lévy process with characteristic triplet (μ, A, ν).*

(a) X is a martingale if and only if

$$\int_{\{\|x\|\geq 1\}} \|x\|\,\nu(dx) < \infty \quad \text{and} \quad \mu + \int_{\{\|x\|\geq 1\}} x\,\nu(dx) = 0.$$

(b) The exponential of a Lévy process is a martingale if and only if

$$\int_{\{\|x\|\geq 1\}} \exp(x)\,\nu(dx) < \infty \quad \text{and}$$

$$\frac{A}{2} + \mu + \int_{\mathbb{R}^d} \left(\exp(x) - 1 - x\mathbb{1}_{\{\|x\|\leq 1\}}\right)\nu(dx) = 0.$$

Proof. The assertions follow from Lemma 3.1.4 and Theorems 3.2.1, 3.2.2. □

Another important property of a Lévy process is its Markov property. That means using the whole past history of the process to predict its future behavior is equal to a prediction based on the knowledge of the present only.

Definition 3.4.2. *An adapted process $(X_t)_{t\geq 0}$ defined on the filtered probability space $(\Omega, \mathbb{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is a Markov process if for every $B \in \mathcal{B}(\mathbb{R}^d)$ and for $0 \leq s \leq t < \infty$*

$$\mathbb{E}[X_t \in B|\mathcal{F}_s] = \mathbb{E}[X_t \in B|\sigma(X_s)] \quad a.s.$$

For a Markov process X_t we can define the transition probability for $0 \leq s < t < \infty$, $B \in \mathcal{B}(\mathbb{R}^d)$, $x \in \mathbb{R}^d$ by

$$P_{s,t}(x, B) = \mathbb{P}(X_t \in B|X_s = x).$$

That is, the mappings $P_{s,t}$ describe the probabilities that the process fades from point x at time s to the set B at time t. Let us now take into consideration the special structure of Lévy processes. Lévy processes can be completely characterized as temporally homogeneous Markov processes with spatially homogeneous transition functions (cf. Sato (1999), Theorem 10.5). That means the transition probability is

$$P_{s,t}(x, B) = P_{0,t-s}(0, B - x)$$

for any $0 \leq s \leq t$, $x \in \mathbb{R}^d$ and $B \in \mathcal{B}(\mathbb{R}^d)$, where $B - x = \{y - x : y \in B\}$. Therefore, the Markov property is weaker than the independence and stationarity property of Lévy process increments.

Infinitesimal generator of Lévy processes

The characteristic triplet of a Lévy process can also be used to describe its infinitesimal generator. Let us first define the transition operator for a Lévy process X and $f : \mathbb{R}^d \to \mathbb{R}^d$ by

$$(T_t f)(x) = \int_{\mathbb{R}^d} f(y) P_t(x, dy) = \mathbb{E}[f(X_t + x)].$$

For homogeneous Markov processes, we always write the operators $T_{0,t}$ as T_t and the transition probabilities $P_{0,t}$ as P_t. Applying the Chapman-Kolmogorov equation (cf. Applebaum (2004), Theorem 3.15) we have the following relation between transition operators

$$T_{s+t} = T_s T_t. \tag{3.3}$$

Any family of linear operators on a Banach space that satisfies (3.3) is called a semigroup. There is an extensive theory of semigroups which can be found e.g. in Applebaum (2004), Section 3.2. Moreover, let us introduce the notation of a Feller process based on Applebaum (2004). A homogeneous Markov process X is said to be a Feller process if $T_t f \in C_0(\mathbb{R}^d)$ for $f \in C_0(\mathbb{R}^d)$ and for all $f \in C_0(\mathbb{R}^d)$ it holds

$$\lim_{t \to 0} \|T_t f - f\| = 0.$$

It can be shown that every Lévy process is a Feller process (Applebaum (2004), Theorem 3.1.9). A semigroup T_t verifying the Feller property can be described via its infinitesimal generator which is a linear operator defined by

$$(\mathcal{A}f)(x) = \lim_{t \to 0} \left\| \frac{T_t f(x) - f(x)}{t} \right\| = 0,$$

where f should be chosen such that the limit exists.
While the infinitesimal generator is hard to characterize in general, its behavior on the subset of smooth functions with compact support is well defined in terms of the Lévy characteristics. The following theorem can be found e.g. in Sato (1999), Theorem 31 or in Applebaum (2004), Theorem 3.3.3.

Theorem 3.4.3. *Let X be a $\mathbb{R}^d$-valued Lévy process with characteristic triplet (μ, A, ν) and let $\mathcal{A}$ be its infinitesimal generator. For $f \in C_0^2(\mathbb{R}^d)$ it holds*

$$\begin{aligned}\mathcal{A}f(x) =& \frac{1}{2} \sum_{i,j=1}^{d} A^{ij} \frac{\partial^2 f(x)}{\partial x^i \partial x^j} + \sum_{i=1}^{d} \mu^i \frac{\partial f(x)}{\partial x^i} \\ & \int_{\mathbb{R}^d} \Big(f(x+z) - f(x) - \sum_{i=1}^{d} z^i \frac{\partial f(x)}{\partial x^i} \mathbb{1}_{\{\|z\| \le 1\}} \Big) \nu(dz).\end{aligned}$$

Example 3.4.4. (i) The Brownian motion with drift
Let $(\mu, A, 0)$ be the characteristic triplet of a d-dimensional Brownian motion with drift. The generator is

$$\mathcal{A}f(x) = \sum_{i=1}^{d} \mu^i \frac{\partial f(x)}{\partial x^i} + \frac{1}{2} \sum_{i,k=1}^{d} A^{ik} \frac{\partial^2 f(x)}{\partial x^i \partial x^k}.$$

(ii) The Poisson process
Let $(0, 0, \lambda\delta_0)$ be the characteristic triplet of a Poisson process with intensity λ. The generator is

$$\mathcal{A}f(x) = \lambda(f(x+1) - f(x)).$$

(iii) The compound Poisson process
Let $(0, 0, \lambda g)$ be the characteristic triplet of a compound Poisson process with intensity λ and jump density $g(x)$. Then the generator is

$$\mathcal{A}f(x) = \lambda \int_{\mathbb{R}^d} (f(x+y) - f(x))g(y)\, dy.$$

In the next section, we show that the infinitesimal generator is the term that has to be subtracted from the Lévy process to get a martingale.

3.5 Itô-Doeblin formula for Lévy-type stochastic integrals

In this section we recall the famous Itô-Doeblin formula, until recently known as Itô formula. Since this topic has recently hit the press let us briefly comment on the naming history of this formula. In the year 2000 a sealed letter which had been stored at the Academy of Sciences in Paris for 60 years revealed a sensation. The letter contained a mathematical manuscript titled "On Kolmogorov's equation". It has been written by Wolfgang Doeblin, a young French mathematician of German origin, during his deployment as simple soldier near the Franco-German border in 1939/1940. A few months later he committed suicide to escape being captured by the German Wehrmacht. The letter proves that Wolfgang Doeblin developed a formula comparable to the famous formula of the Japanese mathematician Kiyoshi Itô.

Proving the Itô-Doeblin formula is non-trivial. The reader my consult Applebaum (2004) or Protter (1990) who treats integration with respect to a general semimartingale. Cont and Tankov (2004) also give a brief review.

Theorem 3.5.1 (Itô-Doeblin formula for multidimensional Lévy processes)**.** *Let X be a $\mathbb{R}^d$-valued Lévy process with characteristic triplet (μ, A, ν) and $f \in C^{1,2}([0,\infty) \times \mathbb{R}^d)$. Then we have*

$$\begin{aligned}
&f(t, X_t)\\
=&f(0, X_0) + \int_0^t \frac{\partial f(s, X_s)}{\partial s}\, ds + \int_0^t \sum_{i=1}^d \frac{\partial f(s,x)}{\partial x^i}\Big|_{x=X_{s-}} dX_s^i\\
&+ \frac{1}{2} \int_0^t \sum_{i,j=1}^d A^{ij} \frac{f(s,x)}{\partial x^i \partial x^j}\Big|_{x=X_s} ds\\
&+ \int_0^t \int_{\mathbb{R}^d} \Big(f(s, X_{s-} + z) - f(s, X_{s-}) - \sum_{i=1}^d z^i \frac{\partial f(s,x)}{\partial x^i}\Big|_{x=X_{s-}} \Big) N(ds, dz).
\end{aligned}$$

It is often useful to separate the martingale term in the equation of Theorem 3.5.1 when it comes to the computation of the mean.

Proposition 3.5.2 (Martingale decomposition)**.** *Let X be a $\mathbb{R}^d$-valued Lévy process with characteristic triplet (μ, A, ν) where $A = \sigma\sigma^T$ and $f \in C^2(\mathbb{R}^d)$. Then*

$$f(X_t) = M_t + V_t$$

where M_t is the martingale part given by

$$\begin{aligned}
M_t =& f(X_0) + \int_0^t \sum_{i,k=1}^d \frac{\partial f(x)}{\partial x^i}\Big|_{x=X_s} \sigma^{ik}\, dW_s^k\\
&+ \int_0^t \int_{\mathbb{R}^d} \big(f(X_{s-} + z) - f(X_{s-}) \big)\, \tilde{N}(ds, dz).
\end{aligned}$$

and V_t a continuous finite variation process

$$\begin{aligned}
V_t =& \frac{1}{2} \int_0^t \sum_{i,j=1}^d A^{ij} \frac{\partial^2 f(x)}{\partial x^i \partial x^j}\Big|_{x=X_s} ds + \int_0^t \sum_{i=1}^d \frac{\partial f(x)}{\partial x^i}\Big|_{x=X_s} \mu^i\, ds\\
&+ \int_0^t \int_{\mathbb{R}^d} \Big(f(X_{s-} + z) - f(X_{s-}) - \sum_{i=1}^d z^i \frac{\partial f(x)}{\partial x^i}\Big|_{x=X_{s-}} \mathbb{1}_{\|z\|\le 1} \Big) \nu(dz) ds.
\end{aligned}$$

Proof. Plug the Lévy-Itô decomposition into the Itô-Doeblin formula. □

Let us close this section citing Applebaum (2004)

> *"Itô's fromula is the key to the wonderful world of stochastic calculus."*

3.6 Existence and uniqueness of Lévy stochastic differential equations

In the following theorems on the existence and uniqueness of solutions of Lévy stochastic differential equations (SDE) we do not present the most general theory. However, we restrict ourselves to conditions being adequate for the vast majority of applications. For more general results we refer to Protter (1990), Chapter V.3. and the references given therein.

Recall that we operate on a filtered probability space $(\Omega, \mathbb{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$ satisfying the usual hypotheses of completeness and right-continuity. Let $W = (W_t)_{t\geq 0}$ be a d-dimensional standard Brownian motion, N an independent Poisson random measure on $\mathbb{R}_+ \times (\mathbb{R}\setminus\{0\})^d$ with associated compensated process $\tilde{N}$ and intensity measure ν where we assume that ν is a Lévy measure.

Consider the following Lévy stochastic differential equation in $\mathbb{R}^d$

$$\begin{aligned} dX_t =& \mu(X_t)\,dt + \sigma(X_t)\,dW_t + \int_{\{\mathbb{R}^d, \|z\|<1\}} \gamma_1(X_{t-}, z)\,\tilde{N}(dt,dz) \\ &+ \int_{\{\mathbb{R}^d, \|z\|\geq 1\}} \gamma_2(X_{t-}, z)\,N(dt,dz) \end{aligned} \tag{3.4}$$

with initial condition $X_0 = x \in \mathbb{R}^d$. Written out in detail component number i of X_t in equation (3.4) is

$$\begin{aligned} dX_t^i =& \mu^i(X_t)\,dt + \sum_{k=1}^d \sigma^{ik}(X_t)\,dW_t^k \\ &+ \int_{\{\mathbb{R}^d, \|z\|<1\}} \gamma_1^i(X_{t-}, z)\,\tilde{N}(dt,dz) + \int_{\{\mathbb{R}^d, \|z\|\geq 1\}} \gamma_2^i(X_{t-}, z)\,N(dt,dz) \end{aligned}$$

where $1 \leq i \leq d$. We assume all $\mu^i : \mathbb{R}^d \to \mathbb{R}$, $\sigma^{ik} : \mathbb{R}^d \to \mathbb{R}$, $\gamma_1^i : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ and $\gamma_2^i : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ to be measurable for $1 \leq i, k \leq d$.

The solution to (3.4), if it exists, will be an $\mathbb{R}^d$-valued stochastic process $X_t = (X_t^1, \ldots, X_t^d)$, $t \geq 0$. In related literature such a solution is frequently called a strong solution since W and N are specified in advance. A discussion about the notion of weak solutions can be found in Applebaum (2004), Section 6.7.3. Moreover, we require our solutions to be unique meaning pathwise uniqueness a.s. To be more precise, if X_t and $\tilde{X}_t$ are solutions to (3.4) then $\mathbb{P}(X_t = \tilde{X}_t, \forall t \geq 0) = 1$.
We now impose the following conditions on the coefficients of a Lévy process in order to ensure that (3.4) has a unique strong solution. Note that $\|\cdot\|_{sn}$ is the matrix semi-norm as defined in Section 1.3.

(C1) **Growth condition**
There exists a constant $C_1 > 0$ such that for all $x \in \mathbb{R}^d$,

$$\|\mu(x)\|^2 + \|\sigma(x)\sigma(x)^T\|_{sn} + \int_{\{\mathbb{R}^d, \|z\|<1\}} \|\gamma_1(x,z)\|^2 \, \nu(dz) \leq C_1(1 + \|x\|^2).$$

(C2) **Lipschitz condition**
There exists a constant $C_2 > 0$ such that for all x, $y \in \mathbb{R}^d$,

$$\|\mu(x) - \mu(y)\|^2 + \|\sigma(x)\sigma(x)^T - 2\sigma(x)\sigma(y)^T + \sigma(y)\sigma(y)^T\|_{sn} + \int_{\{\mathbb{R}^d, \|z\|<1\}} \|\gamma_1(x,z) - \gamma_1(y,z)\|^2 \, \nu(dz) \leq C_2\|x - y\|^2.$$

(C3) **Continuity**
The mapping $x \to \gamma_2(x,z)$ is continuous for all $z \geq 1$.

Let us introduce the modified Lévy stochastic differential equation

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dW_t + \int_{\{\mathbb{R}^d, \|z\|<1\}} \gamma_1(X_{t-}, z)\,\tilde{N}(dt,dz). \qquad (3.5)$$

Theorem 3.6.1. *Suppose* (C1) *and* (C2) *hold. Then there exists a unique càdlàg adapted strong solution* $X = (X_t)_{t \geq 0}$ *to the modified stochastic differential equation (3.5) with* $X_0 = x$.

Proof. For a proof we refer to Applebaum (2004), Section 6.2. □

Using the results of the modified equation one can apply a standard interlacing procedure to construct the solution of the original equation. The continuity assumption ensures that the integrands in the Poisson integrals are predictable.

Theorem 3.6.2. *Suppose* (C1), (C2) *and* (C3) *hold. Then there exists a unique càdlàg adapted strong solution* $X = (X_t)_{t \geq 0}$ *to the stochastic differential equation (3.4) with* $X_0 = x$.

Proof. A proof can be found e.g. in Applebaum (2004), Section 6.2. □

For the next theorem we need a stronger assumption than condition (C1) including also the large jump term.

(C1') **Adjusted growth condition**
There exists a constant $C_1' > 0$ such that for all $x \in \mathbb{R}^d$,

$$\|\mu(x)\|^2 + \|\sigma(x)\sigma(x)^T\|_{sn} + \int_{\{\mathbb{R}^d, \|z\|<1\}} \|\gamma_1(x,z)\|^2 \, \nu(dz) + \int_{\{\mathbb{R}^d, \|z\|\geq 1\}} \|\gamma_2(x,z)\|^2 \, \nu(dz) \leq C_1'(1 + \|x\|^2).$$

The following theorem will be helpful in Chapter 5. The corresponding result for diffusion processes can be found for example in Yong and Zhou (1999), Theorem 6.16 or in Karatzas and Shreve (1991), Chapter 5 Problem 3.15. We show that this upper bound does not only hold for diffusion processes but also for general Lévy processes.

Theorem 3.6.3. *Suppose* $(C1')$, $(C2)$ *and* $C3$ *hold. If* $\mathbb{E}\|X_0\|^{2m} < \infty$ *and* $\int_{\{\|z\|\geq 1\}} \|z\|^{2m}\,\nu(dz) < \infty$ *for some* $m \geq 1$ *then*

$$\mathbb{E}\big[\max_{0\leq t\leq T} \|X_t\|^{2m}\big] \leq Ce^{CT}(1+\mathbb{E}\|X_0\|^{2m})$$

where X *is the solution to (3.4) and* C *is some positive constant depending only on* m, T *and* d.

Proof. In this proof we use the following inequality (cf. Karatzas and Shreve (1991), Chapter 5, inequality (9.3)) for $m \in \mathbb{N}$ and $x_1, \ldots, x_d \in \mathbb{R}$

$$|x_1|^m + \ldots + |x_d|^m \leq d(|x_1| + \ldots + |x_d|)^m \leq d^{m+1}(|x_1|^m + \ldots + |x_d|^m). \quad (3.6)$$

Let us consider the stochastic differential equation in (3.4) and let the constant $C(m,d) \geq 0$ depend on m and d but it is not necessarily the same throughout the proof. Applying inequality (3.6) we have almost surely for all $0 \leq t \leq T$

$$\begin{aligned}\|X_t\|^{2m} \leq & C(m,d)\Big[\|X_0\|^{2m} + \Big\| \int_0^t \mu(X_s)\,ds \Big\|^{2m} + \Big\| \int_0^t \sigma(X_s)\,dW_s \Big\|^{2m} \\ & + \Big\| \int_0^t \int_{\{\mathbb{R}^d, \|z\|<1\}} \gamma_1(X_{s-},z)\,\tilde{N}(ds,dz) \Big\|^{2m} \\ & + \Big\| \int_0^t \int_{\{\mathbb{R}^d, \|z\|\geq 1\}} \gamma_2(X_{s-},z)\,N(ds,dz) \Big\|^{2m}\Big]. \end{aligned} \quad (3.7)$$

The upper bound for the drift term can be obtained using Hölder's inequality

$$\Big\| \int_0^t \mu(X_s)\,ds \Big\|^{2m} = \Big[\sum_{i=1}^d \Big(\int_0^t \mu^i(X_s)\,ds\Big)^2\Big]^m \leq t^m \Big[\int_0^t \big\|\mu(X_s)\big\|^2\,ds\Big]^m. \quad (3.8)$$

We compensate the last term in inequality (3.7), apply the triangle inequality and inequality (3.6). Thus,

$$\begin{aligned} & \Big\| \int_0^t \int_{\{\mathbb{R}^d, \|z\|\geq 1\}} \gamma_2(X_{s-},z)\,N(ds,dz) \Big\|^{2m} \\ = & \Big\| \int_0^t \int_{\{\mathbb{R}^d, \|z\|\geq 1\}} \gamma_2(X_{s-},z)\,\tilde{N}(ds,dz) - \int_0^t \int_{\{\mathbb{R}^d, \|z\|\geq 1\}} \gamma_2(X_{s-},z)\,\nu(dz)ds \Big\|^{2m} \\ \leq & 2^{2m}\Big(\Big\| \int_0^t \int_{\{\mathbb{R}^d, \|z\|\geq 1\}} \gamma_2(X_{s-},z)\,\tilde{N}(ds,dz) \Big\|^{2m} \\ & + \Big\| \int_0^t \int_{\{\mathbb{R}^d, \|z\|\geq 1\}} \gamma_2(X_{s-},z)\,\nu(dz)ds \Big\|^{2m}\Big). \end{aligned} \quad (3.9)$$

The bound for the finite variation part of (3.9) can be obtained using Hölder's inequality

$$\begin{aligned}
&\Big\| \int_0^t \int_{\{\mathbb{R}^d, \|z\| \geq 1\}} \gamma_2(X_{s-}, z)\, \nu(dz) ds \Big\|^{2m} \\
=&\Big[\sum_{i=1}^d \Big(\int_0^t \int_{\{\mathbb{R}^d, \|z\| \geq 1\}} \gamma_2^i(X_{s-}, z)\, \nu(dz) ds \Big)^2 \Big]^m \\
\leq&\Big(\int_0^t \int_{\{\mathbb{R}^d, \|z\| \geq 1\}} 1\, \nu(dz) ds \Big)^m \Big(\sum_{i=1}^d \int_0^t \int_{\{\mathbb{R}^d, \|z\| \geq 1\}} |\gamma_2^i(X_{s-}, z)|^2\, \nu(dz) ds \Big)^m .
\end{aligned} \tag{3.10}$$

Using the property $\int_{\{\|z\| \geq 1\}} \nu(dz) < \infty$ of the Lévy measure (cf. Theorem 3.2.1) implies

$$\begin{aligned}
&\Big\| \int_0^t \int_{\{\mathbb{R}^d, \|z\| \geq 1\}} \gamma_2(X_{s-}, z)\, \nu(dz) ds \Big\|^{2m} \\
\leq& K^m \Big(\int_0^t \int_{\{\mathbb{R}^d, \|z\| \geq 1\}} \|\gamma_2(X_{s-}, z)\|^2\, \nu(dz) ds \Big)^m .
\end{aligned} \tag{3.11}$$

In order to get a bound for the martingale terms we apply the Burkholder-Davis-Gundy inequality (cf. Theorem A.1.1). Let T be the stopping time needed therein. Taking into account the bounds obtained in (3.8),(3.9), (3.11), truncating at the stopping time T, taking the maximum and the expectation, inequality (3.7) gives

$$\begin{aligned}
&\mathbb{E}\big[\max_{0 \leq t \leq T} \|X_t\|^{2m} \big] \\
\leq& C(m, d) \Big(\mathbb{E}\|X_0\|^{2m} + T^m \mathbb{E}\Big[\int_0^T \|\mu(X_s)\|^2\, ds \Big]^m \\
&\quad + \mathbb{E}\Big[\int_0^T \int_{\{\|z\| \geq 1\}} \|\gamma_2(X_{s-}, z)\|^2\, \nu(dz) ds \Big]^m \\
&\quad + \mathbb{E}\Big[\max_{0 \leq t \leq T} \Big\| \int_0^t \sigma(X_s)\, dW_s \Big\|^{2m} \Big] \\
&\quad + \mathbb{E}\Big[\max_{0 \leq t \leq T} \Big\| \int_0^t \int_{\{\|z\| \leq 1\}} \gamma_1(X_{s-}, z) \tilde{N}(ds, dz) \Big\|^{2m} \Big] \\
&\quad + \mathbb{E}\Big[\max_{0 \leq t \leq T} \Big\| \int_0^t \int_{\{\|z\| \geq 1\}} \gamma_2(X_{s-}, z) \tilde{N}(ds, dz) \Big\|^{2m} \Big] \Big).
\end{aligned}$$

Applying Corollary A.1.1 we get bounds for

$$\mathbb{E}\Big[\max_{0 \leq t \leq T} \Big\| \int_0^t \sigma(X_s)\, dW_s \Big\|^{2m} \Big] \leq K_m \mathbb{E}\Big[\int_0^T \|\sigma(X_s)\|^2\, ds \Big]^m ,$$

$$
\begin{aligned}
&\mathbb{E}\Big[\max_{0\le t\le T}\Big\|\int_0^t\int_{\{\|z\|\le 1\}}\gamma_1(X_{s-},z)\,\tilde{N}(ds,dz)\Big\|^{2m}\Big]\\
\le &K_m\mathbb{E}\Big[\int_0^T\int_{\{\|z\|\le 1\}}\big\|\gamma_1(X_{s-},z)\big\|^2\,\nu(dz)ds\Big]^m.
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}\Big[\max_{0\le t\le T}\Big\|\int_0^t\int_{\{\|z\|\ge 1\}}\gamma_2(X_{s-},z)\,\tilde{N}(ds,dz)\Big\|^{2m}\Big]\\
\le &K_m\mathbb{E}\Big[\int_0^T\int_{\{\|z\|\ge 1\}}\big\|\gamma_2(X_{s-},z)\big\|^2\,\nu(dz)ds\Big]^m.
\end{aligned}
$$

Note that $(C1')$ is required in order to obtain these bounds (cf. Cont and Tankov (2004), Proposition 8.8). So far, we therefore have

$$
\begin{aligned}
&\mathbb{E}\big[\max_{0\le t\le T}\|X_t\|^{2m}\big]\\
\le &C(m,d)\Big(\mathbb{E}\|X_0\|^{2m}+T^m\mathbb{E}\Big[\int_0^T\big\|\mu(X_s)\big\|^2\,ds\Big]^m\\
&\qquad+\mathbb{E}\Big[\int_0^T\int_{\{\|z\|\ge 1\}}\|\gamma_2(X_{s-},z)\|^2\,\nu(dz)ds\Big]^m\\
&\qquad+\mathbb{E}\Big[\int_0^T\big\|\sigma(X_s)\big\|^2\,ds\Big]^m\\
&\qquad+\mathbb{E}\Big[\int_0^T\int_{\{\|z\|\le 1\}}\big\|\gamma_1(X_{s-},z)\big\|^2\,\nu(dz)ds\Big]^m\Big).
\end{aligned}
$$

Applying (C1') yields

$$
\mathbb{E}\big[\max_{0\le t\le T}\|X_t\|^{2m}\big]\le C(T,m,d)\Big(\mathbb{E}\|X_0\|^{2m}+\mathbb{E}\Big[\int_0^T\big(1+\|X_s\|^2\big)\,ds\Big]^m\Big). \tag{3.12}
$$

Let us have a closer look at the last term in (3.12). Hölder's inequality and inequality (3.6) imply

$$
\begin{aligned}
\mathbb{E}\Big[\int_0^T(1+\|X_s\|)^2\,ds\Big]^m \le &T^{m-1}\mathbb{E}\Big[\int_0^T\big(1+\|X_s\|^2\big)^m\,ds\Big]\\
\le &T^{m-1}2^m\mathbb{E}\Big[\int_0^T(1+\|X_s\|^{2m})\,ds\Big].
\end{aligned} \tag{3.13}
$$

Finally, we have

$$
\mathbb{E}\big[\max_{0\le t\le T}\|X_t\|^{2m}\big]\le \quad C(T,m,d)\Big(\mathbb{E}\|X_0\|^{2m}+1+\mathbb{E}\Big[\int_0^T\|X_s\|^{2m}\,ds\Big]\Big).
$$

Using the assumption that the Lévy measure $\nu(dz)$ of the Lévy process X satisfies $\int_{\{\|z\|\geq 1\}} \|z\|^{2m}\,\nu(dz) < \infty$ we know by Riesner (2006), Lemma 2.2.5 that $\mathbb{E}[\|X_t\|^{2m}] < \infty$ for all t. Therefore, we have with the latter inequality by an interchange of integral and expectation

$$\psi(T) := \mathbb{E}\big[\max_{0\leq t\leq T} \|X_t\|^{2m}\big] < \infty.$$

We apply Gronwall's inequality (e.g. Amann (1983)) to

$$\psi(T) \leq C(T,m,d)\Big(\mathbb{E}\|X_0\|^{2m} + 1 + \int_0^T \psi(s)\,ds\Big)$$

and obtain

$$\begin{aligned}\psi(T) \leq& C(T,m,d)(\mathbb{E}\|X_0\|^{2m}+1)\Big(1+\int_0^T C(s,m,d)e^{C(T,m,d)(T-s)}\,ds\Big)\\ \leq& C(T,m,d)(\mathbb{E}\|X_0\|^{2m}+1)\Big(1+\int_0^T C(T,m,d)e^{C(T,m,d)(T-s)}\,ds\Big).\end{aligned}$$

Simplifying the right hand side yields

$$\begin{aligned}\psi(T) \leq& C(T,m,d)(\mathbb{E}\|X_0\|^{2m}+1)\Big(1-e^{C(T,m,d)(T-s)}\big|_0^T\Big)\\ =& C(T,m,d)(\mathbb{E}\|X_0\|^{2m}+1)e^{C(T,m,d)T}.\end{aligned}$$

As mentioned before $C(T,m,d)$ is not necessarily the same throughout the proof. Finally, we have

$$\mathbb{E}\big[\max_{0\leq t\leq T} \|X_t\|^{2m}\big] \leq C(\mathbb{E}\|X_0\|^{2m}+1)\,e^{Ct}$$

where $C = C(T,m,d)$. This yields the assertion. □

Chapter 4

Dependence concepts for Lévy processes

Financial and actuarial applications often require multidimensional models with jumps allowing for dependence between the univariate components. Jumps can be easily modeled via Lévy processes. An intuitive approach to model dependence of such Lévy processes might be the use of copulas for random vectors. However, there are a few drawbacks which make this concept not suitable for Lévy processes e.g. the copula C_t of $(X_t^1, \ldots, X_t^d)$ may depend on t. An example for a Lévy process with a copula that is not constant can be found in Tankov (2004). Moreover, it is unknown which copula yields a multidimensional infinitely divisible distribution for given infinitely divisible distributions of $X_t^1, \ldots, X_t^d$. Last but not least it would be inconvenient to model dependence via the copula of the probability distribution whereas the behavior of a multidimensional Lévy process is characterized by the Lévy measure. Therefore, the concept of Lévy copulas was recently introduced in Tankov (2003) and further refined in Kallsen and Tankov (2006). However, caution has to be exercised since Lévy copulas can only be used in some cases to characterize dependence among the components of multidimensional Lévy processes (cf. Bäuerle et al. (2008)). It is possible with respect to the dependence property of association, but the Lévy copula fails to characterize some other dependence properties like multivariate total positivity of order 2 or conditionally increasing in sequence.

We start this chapter briefly recalling some general definitions which will be important in the following sections i.a. the notion of d-increasing functions is introduced. In the literature different definitions for d-increasing functions exist. However, we show that the different definitions only differ by an additional condition. We then discuss the conceptually simpler case of spectrally positive Lévy processes (cf. Cont and Tankov (2004)) before treating the concept for general Lévy processes introduced in Kallsen and

Tankov (2006). A special class of Lévy copulas known as Archimedean Lévy copulas is our main focus. Archimedean Lévy copulas can be constructed quite easily and at the same time possess a lot of nice properties. For that reason they are widely spread in applications. In Subsection 4.2.2 we derive a sufficient and necessary condition for an Archimedean Lévy generator to create a multidimensional positive Lévy copula in arbitrary dimension. So far, this has only been analyzed for a bivariate Lévy copula, while for dimensions larger than two there is only a sufficient condition. It turns out that the necessary and sufficient condition derived in Theorem 4.2.10 contains the existing results as special cases. These considerations were inspired by McNeil and Nešlehová (2008) considering ordinary copulas for random vectors. Finally, we recall the construction of a general Archimedean Lévy copula introduced in Bäuerle et al. (2008).
Describing the dependence structure of a multidimensional Lévy process in terms of its Archimedean Lévy copula allows us to quantify the effect of dependence on the retention levels and the investment portfolio in our multidimensional Lévy driven insurance model (cf. Chapter 5 and Chapter 6).

4.1 Basic definitions

Let us start with some general definitions which will be important in the sections to come.

Definition 4.1.1. *(a) Difference operator*
For functions $f : \mathbb{R}^d \to \mathbb{R}$ the multivariate difference operator is defined by

$$\triangle_x^y f = \sum_{(\epsilon^1,\dots,\epsilon^d)\in\{0,1\}^d} (-1)^{\sum_{i=1}^d \epsilon^i} f\big(\epsilon^1 x^1 + (1-\epsilon^1)y^1, \dots, \epsilon^d x^d + (1-\epsilon^d)y^d\big)$$

where $x = (x^1, \dots, x^d)$, $y = (y^1, \dots, y^d) \in \mathbb{R}^d$.

(b) d-increasing
A function $f : \mathbb{R}^d \to \mathbb{R}$ is called d-increasing if $\triangle_x^y f \geq 0$ for all $x < y$.

In the literature the notion of d-increasing functions may have a slightly different meaning and sometimes they are referred to as quasi-monotone functions or $\triangle$-monotone functions. To make the definition of d-increasing in our context clear we show that the different definitions only differ by an additional condition.

Lemma 4.1.2. *Let $f : \mathbb{R}^d \to \mathbb{R}$ and $f(x) = 0$ if $x^j = 0$ for at least one $j \in \{1, \dots, d\}$. It holds that f is d-increasing if and only if*

$$\triangle_{i_1}^{\tilde{\epsilon}1} \dots \triangle_{i_k}^{\tilde{\epsilon}k} f(x) \geq 0 \quad \text{for all } x \in \mathbb{R}^d, \tag{4.1}$$

for every subset $J := \{i_1, \ldots, i_k\} \subset \{1, \ldots, d\}$ *and* $\tilde{\epsilon}^1, \ldots, \tilde{\epsilon}^k > 0$ *where* $\triangle_{i_1}^{\tilde{\epsilon}^1} \ldots \triangle_{i_k}^{\tilde{\epsilon}^k}$ *denotes a sequential application of the first order difference operator* $\triangle_{i_j}^{\tilde{\epsilon}^j} f(x) = f(\ldots, x_{i_j} + \tilde{\epsilon}_j, \ldots) - f(\ldots, x_{i_j}, \ldots)$, $j \in \{1, \ldots, k\}$.

Remark 7. The definition of d-increasing is sometimes given as in (4.1).

Proof. We first make a general observation. Set $J = \{i_1, \ldots, i_k\} \subset \{1, \ldots, d\}$ for $1 \le k \le d$. Let $x \in \mathbb{R}^d$ and $x^j = 0$ for all $j \notin J$ then $f(x) = 0$. W.l.o.g. we assume $i_j = j$ for all $j \in \{1, \ldots, k\}$. Moreover let $y \in \mathbb{R}^d$ such that $x < y$. The multivariate difference operator can therefore be written as

$$\begin{aligned}
&\sum_{(\epsilon^1, \ldots, \epsilon^d) \in \{0,1\}^d} (-1)^{\sum_{i=1}^d \epsilon^i} f(\epsilon^1 x^1 + (1-\epsilon^1) y^1, \ldots, \epsilon^d x^d + (1-\epsilon^d) y^d) \\
=& \sum_{\substack{\epsilon^j \in \{0,1\},\, \forall j \in J \\ \epsilon^j = 0,\, \forall j \notin J}} (-1)^{\sum_{i=1}^d \epsilon^i} f(\epsilon^1 x^1 + (1-\epsilon^1) y^1, \ldots, \epsilon^d x^d + (1-\epsilon^d) y^d) \\
=& \sum_{(\epsilon^1, \ldots, \epsilon^k) \in \{0,1\}^k} (-1)^{\sum_{j=1}^k \epsilon^j} f(\epsilon^1 x^1 + (1-\epsilon^1) y^1, \ldots, \epsilon^k x^k + (1-\epsilon^k) y^k, \\
&\qquad y^{k+1}, \ldots, y^d)
\end{aligned}$$

for all $(x^1, \ldots, x^k, y^{k+1}, \ldots, y^d) \in \mathbb{R}^d$. Writing the i_l-th summand in detail, w.l.o.g. we assume $l = 1$ and obtain

$$\begin{aligned}
&\sum_{(\epsilon^1, \ldots, \epsilon^k) \in \{0,1\}^k} (-1)^{\sum_{j=1}^k \epsilon^j} f(\epsilon^1 x^1 + (1-\epsilon^1) y^1, \ldots, \epsilon^k x^k + (1-\epsilon^k) y^k, \\
&\qquad y^{k+1}, \ldots, y^d) \\
=& \sum_{(\epsilon^2, \ldots, \epsilon^k) \in \{0,1\}^{k-1}} (-1)^{\sum_{j=2}^k \epsilon^j} \big(f(y^1, \epsilon^2 x^2 + (1-\epsilon^2) y^2, \ldots, \\
&\qquad \epsilon^k x^k + (1-\epsilon^k) y^k, y^{k+1}, \ldots, y^d) \\
&\qquad - f(x^1, \epsilon^2 x^2 + (1-\epsilon^2) y^2, \ldots, \\
&\qquad \epsilon^k x^k + (1-\epsilon^k) y^k, y^{k+1}, \ldots, y^d) \big) \\
=& \sum_{(\epsilon^2, \ldots, \epsilon^k) \in \{0,1\}^{k-1}} (-1)^{\sum_{j=2}^k \epsilon^j} \triangle_1^{\tilde{\epsilon}^1} f(x^1, \epsilon^2 x^2 + (1-\epsilon^2) y^2, \ldots, \\
&\qquad \epsilon^k x^k + (1-\epsilon^k) y^k, y^{k+1}, \ldots, y^d) \\
=& \ldots \\
=& \triangle_k^{\tilde{\epsilon}^k} \ldots \triangle_1^{\tilde{\epsilon}^1} f(x^1, \ldots, x^k, y^{k+1}, \ldots, y^d) \qquad (4.2)
\end{aligned}$$

where $\tilde{\epsilon}^j > 0$ is chosen such that

$$\tilde{\epsilon}^j = y^j - x^j \quad \text{for all } j \in \{1, \ldots, k\}.$$

(i) Let f be now d-increasing and let $J = \{i_1, \dots, i_k\} \subset \{1, \dots, d\}$ for $1 \leq k \leq d$. W.l.o.g. we assume $i_j = j$ for all $j \in \{1, \dots, k\}$. By (4.2) we then have

$$\triangle_k^{\tilde{\epsilon}^k} \dots \triangle_1^{\tilde{\epsilon}^1} f(x^1, \dots, x^k, y^{k+1}, \dots, y^d) \geq 0$$

for all $(x^1, \dots, x^k, y^{k+1}, \dots, y^d) \in \mathbb{R}^d$. Therefore,

$$\triangle_k^{\tilde{\epsilon}^k} \dots \triangle_1^{\tilde{\epsilon}^1} f(x^1, \dots, x^k, x^{k+1}, \dots, x^d) \geq 0$$

for all $(x^1, \dots, x^k, x^{k+1}, \dots, x^d) \in \mathbb{R}^d$.

(ii) Now, let (4.1) hold. Choose $J = \{i_1, \dots, i_k\} \subset \{1, \dots, d\}$ for $k \in \{1, \dots, d\}$ and $\tilde{\epsilon}^1, \dots, \tilde{\epsilon}^k > 0$ such that $\tilde{\epsilon}^j = y^{i_j} - x^{i_j}$. W.l.o.g. we assume $i_j = j$ for all $j \in \{1, \dots, k\}$. We have that

$$\triangle_1^{\tilde{\epsilon}^1}, \dots, \triangle_k^{\tilde{\epsilon}^k} f(x^1, \dots, x^k, x^{k+1}, \dots, x^d) \geq 0$$

implies

$$\triangle_1^{\tilde{\epsilon}^1}, \dots, \triangle_k^{\tilde{\epsilon}^k} f(x^1, \dots, x^k, y^{k+1}, \dots, y^d) \geq 0.$$

This yields the assertion. □

The following observation will be convenient in the sequel.

Proposition 4.1.3. *Let $f : \mathbb{R}_+^d \to \mathbb{R}_+$. It holds that $x \mapsto (-1)^d f(x)$ is d-increasing for $x \geq 0$ if and only if $x \mapsto f(-x)$ is d-increasing for $x \leq 0$.*

Proof. Define $\tilde{f}(x) := f(-x)$ for $x \leq 0$. By Definition 4.1.1 we know that $x \mapsto \tilde{f}(x)$ is d-increasing if and only if $\triangle_x^y \tilde{f} \geq 0$ for $x < y$ where

$$\begin{aligned} &\triangle_x^y \tilde{f} \\ =& \sum_{(\epsilon^1, \dots, \epsilon^d) \in \{0,1\}^d} (-1)^{\sum_{i=1}^d \epsilon^i} \tilde{f}\big(\epsilon^1 x^1 + (1-\epsilon^1) y^1, \dots, \epsilon^d x^d + (1-\epsilon^d) y^d\big) \\ =& \sum_{(\epsilon^1, \dots, \epsilon^d) \in \{0,1\}^d} (-1)^{\sum_{i=1}^d \epsilon^i} f\big(-\epsilon^1 x^1 - (1-\epsilon^1) y^1, \dots, -\epsilon^d x^d - (1-\epsilon^d) y^d\big). \end{aligned}$$

Define $\tilde{x}^i := -x^i$, $\tilde{y}^i := -y^i$ and $\tilde{\epsilon}^i := 1 - \epsilon^i$. Then

$$\begin{aligned} &\triangle_x^y \tilde{f} \\ =& \sum_{(\tilde{\epsilon}^1, \dots, \tilde{\epsilon}^d) \in \{0,1\}^d} (-1)^{\sum_{i=1}^d (1-\tilde{\epsilon}^i)} f\big((1-\tilde{\epsilon}^1)\tilde{x}^1 + \tilde{\epsilon}^1 \tilde{y}^1, \dots, (1-\tilde{\epsilon}^d)\tilde{x}^d + \tilde{\epsilon}^d \tilde{y}^d\big) \\ =& (-1)^d \triangle_{\tilde{y}}^{\tilde{x}} f \end{aligned}$$

for $\tilde{y} < \tilde{x}$. Note that

$$(-1)^{d - \sum_{i=1}^d \epsilon^i} = (-1)^{\sum_{i=1}^d \epsilon^i - d}.$$

This yields the assertion. □

The following definition of margins can be found in Kallsen and Tankov (2006).

Definition 4.1.4. *Let $f : (-\infty, \infty]^d \to (-\infty, \infty]$ be a d-increasing function such that $f(u^1, \dots, u^d) = 0$ if $u^i = 0$ for at least one $i \in \{1, \dots, d\}$. For any non-empty index set $I \subset \{1, \dots, d\}$, the I-margin of f is the function $f^I : (-\infty, \infty]^I \to (-\infty, \infty]$ defined by*

$$f^I\big((u^i)_{i\in I}\big) := \lim_{c\to\infty} \sum_{(u^j)_{j\in I^c}\in\{-c,\infty\}^{|I^c|}} f(u^1, \dots, u^d) \prod_{j\in I^c} sgn\ u^j,$$

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

For notational simplicity, the I-margin of F for $I = \{i\}$ is denoted by F^i.

Remark 8. Let $f : [0, \infty]^d \to [0, \infty]$ satisfy the conditions of Definition 4.1.4. Then the I-margin of f is given by

$$f^I\big((u^i)_{i\in I}\big) := f(u^1, \dots, u^d)\big|_{u^i=\infty,\, i\in I^c}.$$

4.2 Lévy copula for spectrally positive Lévy processes

The term copula originates from the Latin noun for link, tie or bond. It is originally used in linguistic to link the subject and predicate of a sentence. It first appears in statistics in a theorem by Sklar to characterize the function which links univariate distribution functions to construct a multivariate distribution function and which is named after him today. A detailed introduction to the theory of ordinary copulas, mostly in the bivariate case, can be found in Nelsen (2006). For the multivariate case Joe (1997) is a good reference.
In the context of Lévy processes the idea is to establish an analogue of Sklar's theorem for Lévy measures. Since, with the exception of a compound Poisson process, Lévy measures have a non-integrable singularity at zero, we follow Kallsen and Tankov (2006) and introduce a Lévy copula on the tail integral which is defined for each quadrant separately. We therefore first treat Lévy copulas for Lévy processes with only positive jumps in each component. A more detailed discussion about positive Lévy copulas can be found e.g. in Cont and Tankov (2004).

4.2.1 Definitions and Sklar's theorem

We only consider Lévy measures and Lévy copulas living on $\mathbb{R}_+^d$. Unlike copulas for random vectors we have to take into account the possibility of the Lévy measure to be infinite.

Definition 4.2.1 (Positive Lévy copula). *A d-dimensional positive Lévy copula is a function $F : [0,\infty]^d \to [0,\infty]$ satisfying*

(i) $F(u^1,\dots,u^d)$ *is d-increasing*

(ii) $F(u^1,\dots,u^d) = 0$ *whenever* $u^i = 0$ *for at least one* $i \in \{1,\dots,d\}$

(iii) F *has uniform margins, i.e.* $F^i(u) = u$ *for all* $i \in \{1,\dots,d\}$ *and* $u \in [0,\infty]$.

We state the definition of a tail integral which is adapted from Tankov (2003). The link between Lévy copulas and tail integrals will be illustrated later in this section. The tail integral can be interpreted as the tail integral of a Lévy measure and it characterizes the Lévy measure like the distribution function characterizes the law of a random vector. As before we only consider Lévy processes with positive jumps in every component.

Definition 4.2.2 (Tail integral of a Lévy measure). *Let* X *be a d-dimensional Lévy process with positive jumps and Lévy measure* ν *on* $(\mathbb{R}_+ \setminus \{0\})^d$. *Its tail integral is a function* $U : (\mathbb{R}_+ \setminus \{0\})^d \to [0,\infty]$ *such that for* $(x^1,\dots,x^d) \in (\mathbb{R}_+ \setminus \{0\})^d$

$$U(x^1,\dots,x^d) := \nu\big([x^1,\infty) \times \ldots \times [x^d,\infty)\big).$$

Remark 9. We know that (cf. Rüschendorf (1980)) f satisfies (4.1) and is right-continuous if, and only if, f determines a measure μ on $\mathbb{R}^d$ by $\mu((x,y]) = \triangle_x^y f$. Therefore, Definition 4.2.2 corresponds to the abstract definition of a tail integral in Cont and Tankov (2004) (cf. Definition 5.7). However, it possesses two additional properties:

(i) U is left-continuous in each variable.

(ii) U integrates $\|x\|^2$ near 0.

$$\int_{\|x\|\le 1} \|x\|^2 \nu(dx) = \int_{\|x\|\le 1} \|x\|^2 dU < \infty. \tag{4.3}$$

Property (i) is equivalent to the right-continuity of $U(-x)$ for $x < 0$. Moreover the definition of abstract tail integrals demands $x \mapsto (-1)^d U(x)$ to be d-increasing where $x > 0$ which is equivalent to $x \mapsto U(-x)$ d-increasing for $x < 0$ (cf. Proposition 4.1.3). Therefore, $U(-x)$ is a measure defining function for $x < 0$. Furthermore, property (ii) entails that ν satisfies the properties of a Lévy measure.

Definition 4.2.3 (Marginal tail integral). *For* $I \subset \{1,\dots,d\}$ *non-empty, the I-marginal tail integral* U^I *of* X *is the tail integral of the process* $X^I = (X^i)_{i\in I}$. *To simplify notation, we denote one-dimensional margins by* $U^i := U^{\{i\}}$.

Tankov (2004) shows that there is a reformulation of Sklar's theorem for tail integrals. In analogy to survival copulas, multidimensional tail integrals are coupled to their margins via a Lévy copula. The following theorem known as Sklar's theorem for Lévy processes can be found in Kallsen and Tankov (2006).

Theorem 4.2.4 (Sklar's theorem for Lévy processes). *(a) Let X be a $\mathbb{R}_+^d$-valued Lévy process. Then there exists a Lévy copula F on $[0,\infty]^d$ such that the tail integral of X is*

$$U^I\big((x^i)_{i\in I}\big) = F^I\big((U^i(x^i))_{i\in I}\big) \tag{4.4}$$

for all $I \subset \{1,\dots,d\}$ and all $x \in (\mathbb{R}_+ \setminus \{0\})^{|I|}$. The function F is uniquely determined on $\prod_{i=1}^d \overline{\text{Ran } U^i}$.

(b) Conversely, let $X^1,\dots,X^d$ be $\mathbb{R}_+$-valued Lévy processes with tail integrals $U^1,\dots,U^d$ and let F be a Lévy copula on $[0,\infty]^d$. Then equation (4.4) defines for all $I \subset \{1,\dots,d\}$ and all $x \in (\mathbb{R}_+ \setminus \{0\})^{|I|}$ the marginal tail integrals of a $\mathbb{R}_+^d$-valued Lévy process whose components have tail integrals $U^1,\dots,U^d$.

In the case that the marginal Lévy measures $\nu^1,\dots,\nu^d$ are infinite and have no atoms, i.e. $\overline{\text{Ran } U^i} = [0,\infty]$ for all i, the Lévy copula is unique.

On the one hand, Sklar's theorem states that Lévy copulas characterize the tail integral of a Lévy process given fixed marginal tail integrals. On the other hand, for given Lévy copula and one-dimensional tail integrals a multidimensional tail integral can be constructed. That is, the knowledge of the marginal tail integrals does not provide the entire knowledge for the tail integral of a multidimensional Lévy process.

For Lévy measures having Lebesgue densities, these densities can be computed by differentiation. The following result does not only hold in the positive Lévy copula case but also for general Lévy copulas and can be found e.g. in Cont and Tankov (2004), Proposition 5.8.

Proposition 4.2.5. *Let F be a positive, d-dimensional Lévy copula continuous on $[0,\infty]^d$ such that*

$$\frac{\partial^d F(y^1,\dots,y^d)}{\partial y^1 \dots \partial y^d} \text{ exists on } (0,\infty)^d$$

and let $U^1,\dots,U^d$ be univariate tail integrals with densities $\nu^1,\dots,\nu^d$. Then

$$\nu(x^1,\dots,x^d) = \frac{\partial^d F(y^1,\dots,y^d)}{\partial y^1 \dots \partial y^d}\Big|_{y^i=U^i(x^i),i\in\{1,\dots,d\}} \prod_{i=1}^d \nu^i(x^i)$$

is the Lévy density of a Lévy measure with marginal Lévy densities $\nu^1,\dots,\nu^d$.

Further properties of positive Lévy copulas as well as examples can be found in the literature cited in this section as well as in Barndorff-Nielsen and Lindner (2005).

4.2.2 Positive Archimedean Lévy copula

We focus on Archimedean Lévy copulas, a special class of Lévy copulas. The term "Archimedean" originates from ordinary Archimedean copulas satisfying a version of the Archimedean axiom, (cf. Nelsen (2006), Theorem 4.3.1).
We derive a necessary and sufficient condition for an Archimedean Lévy generator to define a multidimensional positive Lévy copula. Inspired by McNeil and Nešlehová (2008) we show that such a generator creates an Archimedean Lévy copula if and only if it is a d-monotone function. So far, a sufficient and necessary condition has only been analyzed for a bivariate Lévy copula. Cont and Tankov (2004) (cf. Proposition 5.6) prove that a bivariate Lévy copula is defined by an Archimedean Lévy generator if and only if this generator is strictly decreasing and convex. While for a d-dimensional Lévy copula Tankov (2003) requires the Archimedean Lévy generator to have derivatives up to the order d with alternating signs. However, this condition is only sufficient and not necessary for dimensions greater than two.

Definition 4.2.6 (Archimedean Lévy generator and copula)**.**

(a) A strictly decreasing and continuous function $\psi : [0,\infty] \to [0,\infty]$ which satisfies the conditions $\psi(0) = \infty$ and $\psi(\infty) = 0$ is called an Archimedean Lévy generator.

(b) A d-dimensional Lévy copula F is called Archimedean if it permits the presentation

$$F(u^1,\dots,u^d) = \psi\big(\psi^{-1}(u^1) + \ldots + \psi^{-1}(u^d)\big), \quad u \in [0,\infty]^d$$

for some Archimedean Lévy generator ψ.

Remark 10. We use the more familiar additive generator approach in contrast to the multiplicative generator approach used in Kallsen and Tankov (2006), but they differ only by a logarithmic transformation (cf. Nelsen (2006) for the case of Archimedean copulas of probability measures).

The following definition describes the requirement on the Archimedean Lévy generator to create a positive Lévy copula as we will show later in this section.

Definition 4.2.7. *(a) d-monotonicity*
A real function f is called d-monotone in (a,b), where $a,b \in \overline{\mathbb{R}}$ and

$d \geq 2$, if it is differentiable in (a, b) up to order $d-2$ and the derivatives satisfy

$$(-1)^k f^{(k)}(x) \geq 0, \quad k = 0, 1, \ldots, d-2$$

for any $x \in (a, b)$ and further if $(-1)^{d-2} f^{(d-2)}$ is non-increasing and convex in (a, b).
For $d = 1$, f is called 1-monotone in (a, b) if it is non-negative and non-increasing on (a, b).

(b) Complete monotonicity
If f has derivatives of all orders in (a, b) and if $(-1)^k f^{(k)} \geq 0$ for any x in (a, b), then f is called completely monotone.

Before we state the main theorem of this section some preliminary work has to be done. The following proposition can be found in McNeil and Nešlehová (2008), Proposition 15.

Proposition 4.2.8. *Let f be a real function on (a, b), $a, b \in \overline{\mathbb{R}}$, let $d \geq 1$ and let further $\tilde{f}$ denote a function on $(-b, -a)$ given by $\tilde{f}(x) = f(-x)$. Then the following statements are equivalent*

(i) f is d-monotone on (a, b).

(ii) f is non-negative and satisfies for any $k = 1, \ldots, d$, any $x \in (-b, -a)$ and any $h_i > 0$, $i = 1, \ldots, k$ such that $(x + h_1 + \ldots + h_k) \in (-b, -a)$

$$\triangle_{h^k} \ldots \triangle_{h^1} \tilde{f}(x) \geq 0$$

where $\triangle_{h^k} \ldots \triangle_{h^1}$ denotes a sequential application of the first order difference operator $\triangle_x^{x+h}$ whenever x and $(x + h) \in (-b, -a)$.

(iii) f is non-negative and satisfies for any $k = 1, \ldots, d$, any $x \in (a, b)$ and any $h > 0$ such that $(x + kh) \in (-b, -a)$

$$(\triangle_h)^k \tilde{f}(x) \geq 0$$

where $(\triangle_h)^k$ denotes the k-fold sequential use of the operator $\triangle_x^{x+h}$.

There is a crucial relation between d-monotonicity and d-increasing functions which is essential for the proof of the main theorem of this section.

Proposition 4.2.9. *Let $f : [0, \infty] \to [0, \infty]$ be a continuous, strictly decreasing real function satisfying $f(\infty) = 0$ and $G : [0, \infty]^d \to [0, \infty]$ be a function specified by*

$$G(f(x^1), \ldots, f(x^d)) = f(\|x\|_1).$$

Then it holds G is d-increasing if and only if f is a d-monotone function.

Proof. First note that the boundary condition $f(\infty) = 0$ forces the function G to be grounded, i.e.

$$G(0, f(x^2), \ldots, f(x^d)) = f(\infty) = 0.$$

Define $\tilde{f}(x) := f(-x)$ for any $x \in (-\infty, 0)$, $J \subset \{1, \ldots, k\}$ and

$$\epsilon^i = \begin{cases} 1, & \text{for } i \in J \\ 0, & \text{for } i \in \{1, \ldots, k\} \setminus J. \end{cases}$$

For $k \in \{1, \ldots, d\}$ and $h > 0$ such that $x + kh < 0$ it holds that

$$\begin{aligned} &(\triangle_h)^k \tilde{f}(x) \\ &= \sum_{(\epsilon^1, \ldots, \epsilon^k) \in \{0,1\}^k} (-1)^{\sum_{i=1}^k \epsilon^i} \tilde{f}\Big(\frac{x}{k} \sum_{i=1}^k \epsilon^i + \Big(\frac{x}{k} + h\Big)\Big(k - \sum_{i=1}^k \epsilon^i\Big)\Big) \\ &= \sum_{(\epsilon^1, \ldots, \epsilon^k) \in \{0,1\}^k} (-1)^{\sum_{i=1}^k \epsilon^i} f\Big(-\frac{x}{k} \sum_{i=1}^k \epsilon^i - \Big(\frac{x}{k} + h\Big)\Big(k - \sum_{i=1}^k \epsilon^i\Big)\Big) \\ &= \sum_{J \subset \{1, \ldots, k\}} (-1)^{|J|} G\Big(f\Big(-\frac{x}{k}\Big)_{j \in J}, f\Big(-\frac{x}{k} - h\Big)_{j \in \{1, \ldots, k\} \setminus J}, f(0), \ldots, f(0)\Big) \end{aligned}$$

First, let G be d-increasing. Together with the property that G is grounded we obtain by Lemma 4.1.2 that

$$\sum_{J \subset \{1, \ldots, k\}} (-1)^{|J|} G\Big(f\Big(-\frac{x}{k}\Big)_{j \in J}, f\Big(-\frac{x}{k} - h\Big)_{j \in \{1, \ldots, k\} \setminus J}, f(0), \ldots, f(0)\Big) \geq 0 \tag{4.5}$$

since f is decreasing. Proposition 4.2.8 implies that f is d-monotone on $(0, \infty)$.
Now, let us assume that f is d-monotone on $(0, \infty)^d$ which is according to Proposition 4.2.8 equivalent to $(\triangle_h)^k \tilde{f}(x) \geq 0$. Therefore the inequality (4.5) holds and Lemma 4.1.2 implies that G is d-increasing. □

We are now in the position to finally show that d-monotonicity is a sufficient and necessary condition for an Archimedean Lévy generator to create a d-dimensional Lévy copula.

Theorem 4.2.10. *Let ψ be an Archimedean Lévy generator. Then $F : [0, \infty]^d \to [0, \infty]$ given by*

$$F(u^1, \ldots, u^d) = \psi\big(\psi^{-1}(u^1) + \ldots + \psi^{-1}(u^d)\big), \quad u \in [0, \infty]^d$$

is a d-dimensional Lévy copula if and only if ψ is d-monotone on $(0, \infty)$.

Proof. F always satisfies (ii) and (iii) of Definition 4.2.1. It remains to show that F is d-increasing if and only if ψ is d-monotone on $(0,\infty)$. We obtain

$$F(\psi(x^1),\dots,\psi(x^d)) = \psi(\|x\|_1)$$

for any $x \in (0,\infty)^d$ and $\psi(x^i) = u^i$. By Proposition 4.2.9 it holds that $F(\psi(x^1),\dots,\psi(x^d))$ is d-increasing if and only if ψ is d-monotone. This yields the assertion. □

This Theorem entails the following two corollaries as special case (cf. Cont and Tankov (2004), Propositions 5.6 and Proposition 5.7 for the if part).

Corollary 4.2.11. *Let ψ be an Archimedean Lévy generator. Then ψ generates a bidimensional positive Lévy copula if and only if ψ is convex.*

Corollary 4.2.12. *Let ψ be an Archimedean Lévy generator with derivatives up to order d. Then ψ generates a d-dimensional positive Lévy copula if and only if ψ has alternating signs of derivatives up to order d, i.e.*

$$(-1)^k\psi^{(k)}(x) \geq 0 \quad \text{for } k \in \{1,\dots,d\}.$$

Let us consider a few examples to illustrate Theorem 4.2.10.

Example 4.2.13 (Clayton Lévy copula)**.** The generator $\psi(t) = t^{-\frac{1}{\delta}}$ for $\delta \in (0,\infty)$ generates the Clayton copula given by

$$F_\delta(u^1,\dots u^d) = \left((u^1)^{-\delta} + \dots + (u^d)^{-\delta}\right)^{-\frac{1}{\delta}}.$$

The limiting cases include the complete dependence and independence Lévy copulas, i.e.

$$\begin{aligned}\lim_{\delta\to\infty} F_\delta(u^1,\dots,u^d) &= \min(u^1,\dots,u^d),\\ \lim_{\delta\to 0} F_\delta(u^1,\dots,u^d) &= u^1\mathbb{1}_{\{u^2=\dots=u^d=0\}} + \dots + u^d\mathbb{1}_{\{u^1=\dots=u^{d-1}=0\}}.\end{aligned}$$

Therefore, δ allows to adjust the degree of dependence. We refer to Tankov (2004) for a characterization of independence and complete dependence Lévy copulas.

Example 4.2.14 (Gumbel Lévy copula)**.** The generator $\psi(t) = \exp(t^{-\frac{1}{\delta}}) - 1$, $\delta \in (0,\infty)$ generates the Gumbel Lévy copula given by

$$F_\delta(u^1,\dots,u^d) = \exp\Big(\Big(\sum_{i=1}^d (\log(1+u^i))^{-\delta}\Big)^{-\frac{1}{\delta}}\Big) - 1.$$

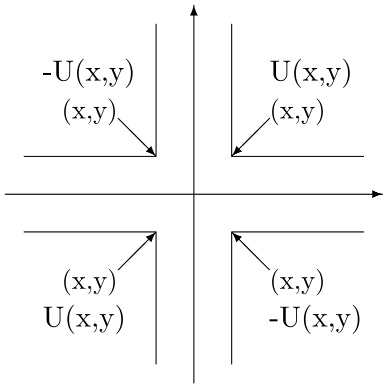

Figure 4.1: Bidimensional tail integral, $U : (\mathbb{R} \setminus \{0\}) \times (\mathbb{R} \setminus \{0\}) \to \mathbb{R}$, $x > 0$ $y > 0 \mapsto \nu\big([x, \infty) \times [y, \infty)\big)$, $x > 0$ $y < 0 \mapsto \nu\big([x, \infty) \times (-\infty, y)\big)$, $x < 0$ $y > 0 \mapsto \nu\big((-\infty, x) \times [y, \infty)\big)$, $x < 0$ $y < 0 \mapsto \nu\big((-\infty, x) \times (-\infty, y)\big)$.

Example 4.2.15 (Ali-Mikhail-Haq Lévy-copula)**.** The generator $\psi(t) = \frac{1-\delta}{e^t - 1}$, $\delta \in [-1, 1)$ generates

$$F_\delta(u^1, \ldots, u^d) = \frac{1 - \delta}{\prod_{i=1}^d \left(\frac{1-\delta}{u^i} + 1\right) - 1}.$$

Inspired by Nelsen (2006) who chooses a similar generator for a copula for random vectors we call this Lévy copula Ali-Mikhail-Haq Lévy-copula.

4.3 Lévy copula for general Lévy processes

Next we give a brief review of Lévy copulas for general Lévy processes. As mentioned before, the main difficulty of Lévy processes is that Lévy measures may have a singularity at zero. Therefore, each corner of the Lévy measure has to be considered separately and we have to define tail integrals for each orthant, compare Figure 4.1. For a more detailed approach and details concerning the following definitions see Kallsen and Tankov (2006). For $I = \{i_1, \ldots, i_k\} \subset \{1, \ldots, d\}$ and $x = (x^1, \ldots, x^d) \in \mathbb{R}^d$ we define $x^I := (x^{i_1}, \ldots, x^{i_k})$.

4.3.1 Definitions and Sklar's theorem

Analogously to the case of spectrally positive Lévy processes, the Lévy measure for the general case is characterized by its tail integral which is given by the following definition.

Definition 4.3.1. *Let X be a $\mathbb{R}^d$-valued Lévy process with Lévy measure ν. The tail integral of X is the function $U : (\mathbb{R}\backslash\{0\})^d \to \mathbb{R}$ defined by*

$$U(x^1, \dots, x^d) := \prod_{i=1}^{d} sgn(x^i)\, \nu\Big(\prod_{j=1}^{d} \mathcal{I}(x^j)\Big),$$

where

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

This definition differs from the positive Lévy copula case in that the Lévy measure is not characterized by a single tail integral anymore. Instead, the Lévy measure is characterized by all marginal tail integrals, each defined on $(\mathbb{R} \setminus \{0\})^I$. Note that the tail integral does not determine the Lévy measure uniquely. However, the Lévy measure is completely determined by the set of its marginal tail integrals (cf. Tankov (2003), Lemma 4.7).
The abstract definition for general Lévy copulas is similar to the spectrally positive case but it is defined on a different domain. Lévy copulas for general Lévy processes are functions from $(-\infty, \infty]^d$ to $(-\infty, \infty]$ instead of functions from $[0, \infty]^d$ to $[0, \infty]$ in the positive Lévy copula case.

Definition 4.3.2 (Lévy copula)**.** *A d-dimensional Lévy copula is a function $F : [-\infty, \infty]^d \to [-\infty, \infty]$ satisfying*

(i) $F(u^1, \dots, u^d)$ is d-increasing

(ii) $F(u^1, \dots, u^d) = 0$ whenever $u^i = 0$ for at least one $i \in \{1, \dots, d\}$

(iii) F has uniform margins, i.e. $F^i(u) = u$ for all $i \in \{1, \dots, d\}$ and $u \in [-\infty, \infty]$.

Finally, let us state a version of Sklar's theorem for general Lévy processes.

Theorem 4.3.3 (Sklar)**.** *(a) Let X be $\mathbb{R}^d$-valued Lévy process. Then there exists a Lévy copula F on $(-\infty, \infty]^d$ such that the tail integral of X is*

$$U^I\big((x^i)_{i\in I}\big) = F^I\big((U^i(x^i))_{i\in I}\big) \tag{4.6}$$

for all $I \subset \{1, \dots, d\}$ and all $x \in (\mathbb{R} \setminus \{0\})^{|I|}$. The function F is uniquely determined on $\prod_{i=1}^{d} \overline{Ran\ U^1}$.

(b) Conversely, let $(X^1), \dots, (X^d)$ be $\mathbb{R}$-valued Lévy processes with tail integrals $U^1, \dots, U^d$ and let F be a Lévy copula on $(-\infty, \infty]^d$. Then equation (4.6) defines for all $I \subset \{1, \dots, d\}$ and all $x \in (\mathbb{R} \setminus \{0\})^{|I|}$ the marginal tail integrals of a $\mathbb{R}^d$-valued Lévy process whose components have tail integrals $U^1, \dots, U^d$.

Here $\prod_{i=1}^{d} \overline{\text{Ran} U^i} = [-\infty, \infty]^d$ if as before ν^i, $i = 1, \dots, d$ has no atoms and $\nu^i\big((-\infty, 0)\big) = \infty$, $\nu^i\big((0, \infty)\big) = \infty$. In this case the Lévy copula is unique.

4.3.2 General Archimedean Lévy copula

To complete this chapter, let us finally turn to general Archimedean Lévy copulas on $(-\infty,\infty]^d$. As mentioned before, the construction described above is not valid if there are jumps in several directions. In this case the definition of positive Lévy copulas has to be extended to the whole Euclidean space. This can be done by constructing an Archimedean Lévy copula for each orthant according to Subsection 4.2.2 and combining them in a certain way. The following is adapted from Bäuerle et al. (2008).

Let $I = \{-1, 1\}^d$ and define for each $i = (i^1, \ldots, i^d) \in I$ an orthant

$$O^i = \{x \in \mathbb{R}^d : \text{ sgn } (x^j) = i^j, j \in \{1, \ldots, d\}\}.$$

Given a set of functions F_{ϕ_i}, $i \in I$, and a weight function $\eta : I \to [0, 1]$ having the property that for each $k \in \{1, \ldots, d\}$ it holds

$$\sum_{i:\, i^k=-1} \eta(i) = \sum_{i:\, i^k=1} \eta(i) = 1,$$

we can define an Archimedean Lévy copula on $\mathbb{R}^d$ by

$$F(u^1, \ldots, u^d) = \sum_{i \in I} \big(\eta(i) F_{\phi_i}(|u^1|, \ldots, |u^d|) \mathbb{1}_{u \in O^i} \prod_{j=1}^{d} \text{ sgn } (u^j)\big)$$

if $|u^j| > 0, j \in \{1, \ldots, d\}$ and $F(u^1, \ldots, u^d) = 0$ else.

Remark 11. Kallsen and Tankov (2006) have introduced a family of Archimedean Lévy copulas on $(-\infty,\infty]^d$. However, it can be checked that the class of Lévy processes generated by their family of Archimedean Lévy copulas does not include positive dependent Lévy processes in terms of association.

Chapter 5

Stochastic control of portfolios with Lévy-dynamics

We investigate now optimization of reinsurance and investment strategies in a multidimensional Lévy-driven insurance model. A special feature of our construction allows for dependent claims and dependent investments of the insurance company's single business lines. Before considering our insurance model driven by Lévy dynamics in Section 5.2, we treat Lévy process stochastic control theory in Section 5.1. As far as we know, there is by now only a very short introduction to stochastic control with respect to jump diffusions provided by Øksendal and Sulem (2007). Specific conditions are required to get existence of a solution of the Lévy stochastic differential equations and to have a well-defined mean in the reward functional.
Inspired by Browne (1995) who discovered that in a one-dimensional diffusion model the control which maximizes expected exponential utility of terminal wealth also minimizes the ruin probability, we consider in Section 5.3 the optimization criterion that maximizes exponential utility of terminal wealth. Imbedding this problem in stochastic control theory and solving the Hamilton-Jacobi Bellman equation we can show that it is optimal to keep the retention level and investment portfolio constant regardless of the time and the company's wealth level. This does not only hold for proportional reinsurance but also for general reinsurance as well as for a mixture of proportional and excess of loss reinsurance in a slightly modified model. In the latter case we can even show that there exists a pure excess of loss strategy which is always better than any combined reinsurance policy. We conclude in Section 5.4 with a validation of the conjecture that the policy which maximizes utility of terminal wealth also minimizes the ruin probability in our multidimensional reinsurance model. This holds omitting claims caused by jumps while assuming that ruin occurs when the weighted sum of net values

of the business units become negative.

5.1 Stochastic control of Lévy processes

There is quite a bit of literature on continuous time stochastic control theory for processes having continuous paths. We only mention a few e.g. Fleming and Rishel (1975), Yong and Zhou (1999), Schmidli (2008) in actuarial context and Korn (1997) with respect to portfolio optimization. However, literature on stochastic control for jump processes is quite sparse. As far as we know, only Øksendal and Sulem (2007) provide a short introduction to stochastic control of jump diffusions. For this reason we provide a thorough introduction to stochastic control of general Lévy processes.

5.1.1 Terminology of classical control theory

Let $(\Omega, \mathbb{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$ be a filtered probability space satisfying the usual hypotheses of completeness and right-continuity (cf. Protter (1990), Chapter I) and let $\mathcal{U}$ be a given set. We refer to $X = (X_s)_{s\in I}$ defined in Section 3.6 as state process in the time interval $I \subset \mathbb{R}_+$. We are free to choose a control process $u = (u_s)_{s\in I}$, $u : I \times \Omega \to \mathcal{U}$ affecting the dynamics of the system. The control process $u = (u_s)_{s\in I}$ is assumed to be predictable, i.e. measurable with respect to the σ-algebra generated by all adapted left-continuous processes. In the finite horizon case the state process is only controlled in a finite interval $I = [t, T]$ whereas in the infinite horizon case it is controlled in the interval $I = [0, \infty)$. Choosing a control u the dynamics of the system are governed by the following differential equation

$$\begin{aligned} dX_s =&\mu(X_s, u_s)\, ds + \sigma(X_s, u_s)\, dW_s + \int_{\mathbb{R}^d} \gamma(X_{s-}, u_s, z)\, \bar{N}(ds, dz) \\ X_t =&x \end{aligned} \tag{5.1}$$

where

$$\bar{N}(ds, dz) = N(ds, dz) - \mathbb{1}_{\{\|z\|<1\}}\nu(dz)ds,$$

$$\mu : \mathbb{R}^d \times \mathcal{U} \to \mathbb{R}^d, \quad \sigma : \mathbb{R}^d \times \mathcal{U} \to \mathbb{R}^{d\times r} \quad \text{and} \quad \gamma : \mathbb{R}^d \times \mathcal{U} \times \mathbb{R}^d \to \mathbb{R}^d.$$

If it exists, the process $X = (X_s)_{s\in I}$ is the solution of the controlled Lévy stochastic differential equation (5.1).

Remark 12. If (X_s) is a càdlàg adapted process, then (X_{s-}) is a predictable process (cf. Jacod and Shiryaev (2003), Proposition 2.6). Therefore our stochastic integral in (5.1) is well-defined.

Exercising control does not only affect the dynamics of the system but also results in rewards. We therefore associate a reward functional with each control u which is for the

(a) finite horizon

$$J(t,x;u) = \mathbb{E}_{t,x}\Big[\int_t^T g(s,X_s,u_s)\,ds + h(X_T)\Big]. \tag{5.2}$$

The function $g(s,x,u)$ is called running reward function and $h(x)$ terminal reward function. Both, the running reward and the terminal reward function are assumed to be continuous. The functional $J(t,x;u)$ is therefore the total expected reward of using control u on the interval $[t,T]$ when starting at time t in position x. This is also indicated by the indices at the conditional expectation.

(b) infinite horizon

$$J(x;u) = \mathbb{E}_x\Big[\int_0^\infty e^{-rs} g(X_s,u_s)\,ds\Big]. \tag{5.3}$$

The running reward $g(x,u)$ is assumed to be continuous and it does not depend on s. The functional $J(x;u)$ is the discounted reward of using control u when starting at time $t=0$ in position x. The parameter $r>0$ is called discount rate.

We denote by $\mathcal{U}(t,x) \subset \mathcal{U}$ the set of all admissible controls u where the controlled process starts in $(t,x) \in I \times \mathbb{R}^d$ if the stochastic differential equation (5.1) with initial condition $X_t = x$ admits a unique solution X for all $x \in \mathbb{R}^d$ and if the mean in the reward functional (5.2) (respectively in (5.3)) is well-defined.

The aim is to choose the control process which maximizes the reward function. That is, in order to get a solution we have to find an optimal control $u^* \in \mathcal{U}(t,x)$ and the value function $V(t,x)$ or $V(x)$ defined by

$$V(t,x) = \sup_{u\in\mathcal{U}(t,x)} J(t,x;u) = J(t,x;u^*),$$

and

$$V(x) = \sup_{u\in\mathcal{U}(x)} J(x;u) = J(x;u^*)$$

respectively.

In the following we only focus on the finite horizon problem. Let us introduce some necessary assumptions. As we have seen in the previous section, even the solution of an uncontrolled stochastic differential equation may not exist unless the drift, the diffusion coefficient and the jump coefficient satisfy certain regularity conditions. Furthermore, the reward functional could be infinite for each admissible control if the reward functions' rate of growth is too large. We assume that the functions $\mu(\cdot,\cdot)$, $\sigma(\cdot,\cdot)$, $\gamma(\cdot,\cdot)$ and the functions $g(\cdot,\cdot,\cdot)$ and $h(\cdot,\cdot)$ satisfy the following conditions:

(A1) **Adjusted linear growth condition**
There exists a constant C_1 such that,

$$\begin{aligned}&\|\mu(x,u)\|^2 + \|\sigma(x,u)\sigma(x,u)^T\|_{sn} + \int_{\mathbb{R}^d} \|\gamma(x,u,z)\|^2 \,\nu(dz)\\ \le& C_1(1+\|x\|^2+\|u\|^2).\end{aligned}$$

(A2) **Lipschitz condition**
There exists a constant C_2 such that

$$\begin{aligned}&\|\mu(x,u)-\mu(y,v)\|^2 + \int_{\{\mathbb{R}^d,\|z\|<1\}} \|\gamma(x,u,z)-\gamma(y,v,z)\|^2 \,\nu(dz)\\ &+\|\sigma(x,u)\sigma(x,u)^T - 2\sigma(x,u)\sigma(y,v)^T + \sigma(y,v)\sigma(y,v)^T\|_{sn}\\ \le& C_2(\|x-y\|^2+\|u-v\|^2).\end{aligned}$$

(A3) **Continuity**
The mapping $(x,u)\to\gamma(x,u,z)$ is continuous for all $z \ge 1$.

(A4) **Polynomial growth condition**
There exist constants C_3 and $k<\infty$ such that

$$\begin{aligned}|g(t,x,u)| \le& C_3(1+\|x\|^k+\|u\|^k)\\ |h(t,x)| \le& C_3(1+\|x\|^k).\end{aligned}$$

(A5) **Moment condition**
For some $k \ge 1$ the Lévy measure $\nu(dz)$ satisfies

$$\int_{\{\mathbb{R}^d,\|z\|\ge 1\}} \|z\|^k \,\nu(dz) < \infty.$$

By Theorem 3.6.2 existence and uniqueness of the solution to (5.1) follow from $(A1)$, $(A2)$ and $(A3)$. The next proposition shows that these assumptions also ensure that the mean in the reward functional is well-defined.

Proposition 5.1.1. *Suppose (A1)-(A5) and* $\mathbb{E}[\int_t^T \|u_s\|^k \,ds] < \infty$ *for all* $k<\infty$ *hold. Then the reward functional* $J(t,x;u)$ *and the value function* $V(t,x)$ *satisfy*

$$|J(t,x;u)| \le K(1+\|x\|^k) \quad \text{for all } u \in \mathcal{U} \tag{5.4}$$

$$|V(t,x)| \le K(1+\|x\|^k) \quad \text{for all } (t,x)\in[0,T]\times\mathbb{R}^d. \tag{5.5}$$

Proof. Let $(t,x)\in[0,T)\times\mathbb{R}^d$ be fixed. By Theorem 3.6.3 we have

$$\mathbb{E}\big[\max_{t\le s\le T}\|X_s\|^k\big] \le C(1+\|x\|^k) \tag{5.6}$$

where X is the solution of (5.1). Together with (A4) we get an upper bound

$$|J(t,x;u)| \leq K(1+\|x\|^k) \quad \forall u \in \mathcal{U}(t,x)$$

where K depends on $C = C(T,m,d)$, C_3, T as well as the upper bound of $\mathbb{E}\big[\int_t^T \|u_s\|^k\,ds\big]$. This implies (5.5). □

Under the assumptions of this section it therefore suffices to define the set of admissible controls $\mathcal{U}(t,x)$ for fixed $(t,x) \in I \times \mathbb{R}^d$ such that

$$\mathcal{U}(t,x) = \Big\{u \in \mathcal{U} : (u_s)_{s\in[t,T]} \text{ is predictable and } \mathbb{E}\Big[\int_t^T \|u_s\|^k\,ds\Big] < \infty \text{ for all } k < \infty\Big\}.$$

5.1.2 Hamilton-Jacobi-Bellman equation

The Hamilton-Jacobi-Bellman (HJB) equation serves a method to derive the solution of stochastic control problems. This famous equation is named after Richard Bellman, who pioneered the theory of dynamic programming in the 1950s and after whom the corresponding discrete-time equation is named. It is further named after William Rowan Hamilton and Carl Gustav Jacob Jacobi since the Hamilton-Jacobi-Bellman equation for the diffusion case can be seen as an extension of the Hamilton-Jacobi equation in classical physics.

Let us now derive the Hamilton-Jacobi-Bellman equation for the finite horizon problem where we extend the classical Hamilton-Jacobi-Bellman equation to the case of the controlled process being a general Lévy process. The main tool is the Bellman equation which is sometimes also called Dynamic Programming principle. The following citation is due to Bellman (cf. Bellman (1957), p. 83).

> *"An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."*

Theorem 5.1.2 (Bellman equation)**.** *The optimal reward function $V(t,x)$ satisfies the Bellman equation, i.e. for all $(t,x) \in [0,T) \times \mathbb{R}^d$,*

$$V(t,x) = \sup_{u\in\mathcal{U}(t,x)} \mathbb{E}_{t,x}\Big[\int_t^{\tilde{t}} g(s,X_s,u_s)\,ds + V(\tilde{t},X_{\tilde{t}})\Big], \qquad \textit{for all } \tilde{t} \in [t,T],$$

where X denotes the solution of the stochastic differential equation (5.1).

Proof. We omit this rather technical proof and refer to Yong and Zhou (1999) (cf. Theorem 3.3) or Højgaard and Taksar (2007) (cf. Theorem B.1). □

The Bellman equation states that the maximal expected value $V(t,x)$ in the interval $[t,T]$ can be derived by taking the supremum over the following strategy: Choose control u in $[t,\tilde{t}]$ and act optimal from $X_{\tilde{t}}$ onwards which leads to a reward $V(\tilde{t}, X_{\tilde{t}})$.

Corollary 5.1.3 (Bellman's principle of optimality). *If (X^*, u^*) is optimal on $[t,T]$ then*

$$V(\tilde{t}, X^*_{\tilde{t}}) = J(\tilde{t}, X^*_{\tilde{t}}; u^*) \quad \text{for } \tilde{t} \geq t.$$

That means the restriction of (X^*, u^*) on the interval $[\tilde{t}, T]$ with $\tilde{t} \geq t$ is optimal.

Proof. For all $t \leq \tilde{t}$ we have

$$\begin{aligned} V(t,x) =& J(t,x;u^*) = \mathbb{E}_{t,x}\Big[\int_t^{\tilde{t}} g(s, X_s, u^*_s)\, ds + J(\tilde{t}, X^*_{\tilde{t}}; u^*)\Big] \\ \leq& \mathbb{E}_{t,x}\Big[\int_t^{\tilde{t}} g(s, X_s, u^*_s)\, ds + V(\tilde{t}, X^*_{\tilde{t}})\Big] \leq V(t,x) \end{aligned}$$

where the last inequality follows from the Bellman equation (cf. Theorem 5.1.2). Therefore,

$$\mathbb{E}_{t,x}\big[J(\tilde{t}, X^*_{\tilde{t}}; u^*)\big] = \mathbb{E}_{t,x}\big[V(\tilde{t}, X^*_{\tilde{t}})\big]$$

and thus

$$J(\tilde{t}, X^*_{\tilde{t}}; u^*) = V(\tilde{t}, X^*_{\tilde{t}}) \text{ a.s.}$$

which yields the assertion. □

Let us now define the Hamiltonian generator $\mathcal{A}^u f(t,x)$ with respect to general Lévy processes for any $C^{1,2}$ function $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ such that

$$\begin{aligned} &\mathcal{A}^u f(t,x) \\ =& \frac{\partial f(t,x)}{\partial t} + \sum_{i=1}^d \mu^i(x,u) \frac{\partial f(t,x)}{\partial x^i} + \frac{1}{2} \sum_{i,j=1}^d A^{ij}(x,u) \frac{\partial^2 f(t,x)}{\partial x^i \partial x^j} \\ &+ \int_{\mathbb{R}^d} \Big(f(t, x + \gamma(x,u,z)) - f(t,x) - \sum_{i=1}^d \gamma^i(x,u,z) \frac{\partial f(t,x)}{\partial x^i} \mathbb{1}_{\{\|z\| \leq 1\}} \Big) \nu(dz) \end{aligned}$$

where $A(x,u) = \sigma(x,u)\sigma(x,u)^T$ denotes the covariance matrix of our state process. Before we state the Hamilton-Jacobi Bellman equation based on Lévy processes, we need some more assumptions to ensure the martingale preserving property with respect to general Lévy processes.

(M1) For $V \in C^1$,

$$\mathbb{E}\Big[\int_0^T \Big\| \frac{\partial V(t,x)}{\partial x}\Big|_{x=X_t} \sigma(X_t, u_t) \Big\|^2 dt\Big] < \infty.$$

(M2) $\mathbb{E}\Big[\int_0^T\int_{\mathbb{R}^d}\big\|V(t,X_{t-}+\gamma(X_{t-},u_t,z))-V(t,X_{t-})\big\|^2\,\nu(dz)dt\Big]<\infty.$

We are now able to derive the Hamilton-Jacobi Bellman equation.

Theorem 5.1.4. *Suppose (A1)-(A5), (M1)-(M2) and* $V(t,x)\in C^{1,2}([0,T]\times\mathbb{R}^d)$ *hold. Then* $V(t,x)$ *satisfies the Hamilton-Jacobi-Bellman (HJB) equation*

$$\sup_{u\in\mathcal{U}}\big\{\mathcal{A}^u V(t,x)+g(t,x,u)\big\}=0,\quad (t,x)\in[0,T)\times\mathbb{R}^d, \tag{5.7}$$

with terminal condition

$$V(T,x)=h(x),\quad x\in\mathbb{R}^d. \tag{5.8}$$

Proof. Fix $(t,x)\in[0,T)\times\mathbb{R}^d$ as well as an arbitrary constant control $u\in\mathcal{U}$. Let X be the corresponding state trajectory to $u\in\mathcal{U}(t,x)$ with $u_s\equiv u$ for all s. For $\epsilon>0$ let $\eta_u^\epsilon=(\epsilon+t)\wedge\inf\{s\geq t:\|X_s-x\|\geq\epsilon\}$. Then $\eta_u^\epsilon<\infty$ a.s. and $\eta_u^\epsilon\to t$ as $\epsilon\to 0$. Applying the Itô-Doeblin formula for multidimensional Lévy processes (cf. Theorem 3.5.1) on the value function we have for $\eta_u^\epsilon\geq t$ in terms of the Hamiltonian generator

$$\begin{aligned}V(\eta_u^\epsilon,X_{\eta_u^\epsilon})=&V(t,x)+\int_t^{\eta_u^\epsilon}\mathcal{A}^uV(s,X_s)\,ds\\&+\int_t^{\eta_u^\epsilon}\sum_{i,k=1}^d\frac{V(s,x)}{\partial x^i}\Big|_{x=X_s}\sigma^{ik}(s,X_s,u)\,dW_s^k\\&+\int_t^{\eta_u^\epsilon}\int_{\mathbb{R}^d}\big(V(s,X_{s-}+\gamma(X_{s-},u,z))-V(s,X_{s-})\big)\,\tilde{N}(ds,dz).\end{aligned} \tag{5.9}$$

The finite variation process is contained in the Hamiltonian generator, the remaining term of (5.9) is a martingale term (cf. Theorem 3.5.2). Note that for the application of Itô-Doeblin's formula the assumption $V(t,x)\in C^{1,2}([0,T]\times\mathbb{R}^d)$ is required. Applying Bellman's equation (cf. Theorem 5.1.2) yields

$$0\geq\mathbb{E}\big[V(\eta_u^\epsilon,X_{\eta_u^\epsilon})-V(t,x)\big]+\mathbb{E}\Big[\int_t^{\eta_u^\epsilon}g(s,X_s,u)\,ds\Big].$$

Dividing by $\mathbb{E}[\eta_u^\epsilon-t]$ and pasting in (5.9) gives us

$$\begin{aligned}0\geq&\frac{1}{\mathbb{E}[\eta_u^\epsilon-t]}\,\mathbb{E}\Big[V(\eta_u^\epsilon,X_{\eta_u^\epsilon})-V(t,x)+\int_t^{\eta_u^\epsilon}g(s,X_s,u)\,ds\Big]\\=&\frac{1}{\mathbb{E}[\eta_u^\epsilon-t]}\,\mathbb{E}\Big[\int_t^{\eta_u^\epsilon}\big(\mathcal{A}^uV(s,X_s)+g(s,X_s,u)\big)\,ds\\&\qquad+\int_t^{\eta_u^\epsilon}\sum_{i,k=1}^d\frac{V(s,x)}{\partial x^i}\Big|_{x=X_s}\sigma^{ik}(s,X_s,u)\,dW_s^k\end{aligned}$$

$$+\int_t^{\eta_u^\epsilon}\int_{\mathbb{R}^d}\big(V(s,X_{s-}+\gamma(X_{s-},u,z)) \\ -V(s,X_{s-})\big)\,\tilde{N}(ds,dz)\Big].$$

Because of assumptions (M1)-(M2) and the martingale preserving property for Brownian integrals and compensated Poisson integrals (cf. Cont and Tankov (2004) p.259 ff.) it holds that

$$\mathbb{E}\Big[\int_t^{\eta_u^\epsilon}\sum_{i,k=1}^d \frac{V(s,x)}{\partial x^i}\Big|_{x=X_s}\sigma^{ik}(s,X_s,u)\,dW_s^k\Big]=0 \tag{5.10}$$

as well as

$$\mathbb{E}\Big[\int_t^{\eta_u^\epsilon}\int_{\mathbb{R}^d}\big(V(s,X_{s-}+\gamma(X_{s-},u_{s-},z))-V(s,X_{s-})\big)\,\tilde{N}(ds,dz)\Big]=0. \tag{5.11}$$

Thus it remains

$$0\geq\frac{1}{\mathbb{E}[\eta_u^\epsilon-t]}\,\mathbb{E}\Big[\int_t^{\eta_u^\epsilon}\big(\mathcal{A}^uV(s,X_s)+g(s,X_s,u)\big)\,ds\Big].$$

By construction of η_u^ϵ we have that $\eta_u^\epsilon-t\leq\epsilon$ and $\|X_s-x\|<\epsilon$ for all $s\in[t,\eta_u^\epsilon)$. Therefore,

$$g(s,X_s,u)+\mathcal{A}^uV(s,X_s)=g(t,x,u)+\mathcal{A}^uV(t,x)+\phi(s)$$

where

$$|\phi(s)|\leq\sup_{\|x-y\|<\epsilon,|t-s|<\epsilon}|g(t,x,u)-g(s,y,u)|+|\mathcal{A}^uV(t,x)-\mathcal{A}^uV(s,y)|\to 0$$

as $\epsilon\to 0$ by the continuity of g, the Lipschitz condition (A2) and $V\in C^{12}$. Letting η_u^ϵ tend to t we obtain

$$0\geq\mathcal{A}^uV(t,x)+g(t,x,u)$$

for all $u\in\mathcal{U}$. Arbitrariness of u yields

$$\sup_{u\in\mathcal{U}}\big\{\mathcal{A}^uV(t,x)+g(t,x,u)\big\}\leq 0.$$

Now, let $\tilde{\epsilon}>0$, η_u^ϵ as defined above and $0\leq t<\eta_u^\epsilon\leq T$ with $\mathbb{E}[\eta_u^\epsilon-t]$ small enough. Then there exists $u\equiv u_{\tilde{\epsilon},\eta_u^\epsilon}\in\mathcal{U}(t,x)$ with

$$V(t,x)-\tilde{\epsilon}\,\mathbb{E}[\eta_u^\epsilon-t]\leq\mathbb{E}\Big[\int_t^{\eta_u^\epsilon}g(s,X_s,u_s)\,ds+V(\eta_u^\epsilon,X_{\eta_u^\epsilon})\Big].$$

Thus

$$-\tilde{\epsilon} \leq \frac{1}{\mathbb{E}[\eta_u^\epsilon - t]}\, \mathbb{E}\Big[V(\eta_u^\epsilon, X_{\eta_u^\epsilon}) - V(t,x) + \int_t^{\eta_u^\epsilon} g(s, X_s, u_s)\, ds\Big].$$

Therefore by (5.9) and the same arguments as above we have

$$\begin{aligned} -\tilde{\epsilon} \leq & \frac{1}{\mathbb{E}[\eta_u^\epsilon - t]}\, \mathbb{E}\Big[\int_t^{\eta_u^\epsilon} \big(\mathcal{A}^u V(s, X_s)) + g(s, X_s, u_s)\big)\, ds\Big] \\ \leq & \frac{1}{\mathbb{E}[\eta_u^\epsilon - t]}\, \mathbb{E}\Big[\int_t^{\eta_u^\epsilon} \sup_{u \in \mathcal{U}} \big\{\mathcal{A}^u V(s, X_s) + g(s, X_s, u)\big\}\, ds\Big] \end{aligned}$$

and letting $\epsilon \to 0$ yields

$$-\tilde{\epsilon} \leq \sup_{u \in \mathcal{U}} \big\{\mathcal{A}^u V(t,x) + g(t,x,u)\big\}.$$

We obtain

$$0 \leq \sup_{u \in \mathcal{U}} \big\{\mathcal{A}^u V(t,x) + g(t,x,u)\big\}$$

and the assertion follows. □

We proved that the Hamilton-Jacobi-Bellman equation is satisfied by the optimal reward function $V(t,x)$ under certain regularity conditions. However, it is hardly possible to prove the smoothness of the optimal reward function a priori. But the following theorem states that if there exists a solution to the HJB-equation and certain assumptions are satisfied then this solution coincides with the optimal reward function. In addition, the proof provides a method for the construction of an optimal control.

Theorem 5.1.5 (Verification theorem)**.** *Let $\tilde{V}(t,x) \in C^{1,2}([0,T] \times \mathbb{R}^d)$ be a solution of the HJB-equation (5.7) with terminal condition (5.8). Suppose (A1)-(A5) and (M1)-(M2) for $\tilde{V}$ hold. Then*

(a) $\tilde{V}(t,x) \geq V(t,x)$ for all $(t,x) \in [0,T] \times \mathbb{R}^d$.

(b) If for all $(t,x) \in [0,T] \times \mathbb{R}^d$ there exists an $\tilde{u}(\cdot,\cdot) \in \mathcal{U}(t,x)$ such that $\tilde{u}(s,y)$ is a maximum point of

$$u \mapsto \mathcal{A}^u \tilde{V}(s,y) + g(s,y,u) \quad \text{for all } s \in [t,T] \tag{5.12}$$

and X_s^* the corresponding state process, i.e. X_s^* is the solution to the following stochastic differential equation

$$\begin{aligned} dX_s =& \mu\big(s, X_s, \tilde{u}(s, X_s)\big)\, ds + \sigma\big(s, X_s, \tilde{u}(s, X_s)\big)\, dW_s \\ &+ \int_{\mathbb{R}^d} \gamma\big(X_{t-}, \tilde{u}(t, X_{t-}), z\big)\, \bar{N}(dt, dz) \\ X_t =& x. \end{aligned}$$

Then

$$u_s^* = \tilde{u}(s, X_s^*) \quad \textit{for all } s \in [t,T]$$

is an optimal control and

$$\tilde{V}(t,x) = V(t,x) \qquad \textit{for all } (t,x) \in [0,T) \times \mathbb{R}^d.$$

Proof. (a) Fix $(t,x) \in [0,T) \times \mathbb{R}^d$ as well as an arbitrary control. Let X be the corresponding state trajectory to $u \in \mathcal{U}(t,x)$. Applying Itô-Doeblin's formula for multidimensional Lévy processes (cf. Theorem 3.5.1) we get under assumption (M1)-(M2)

$$\begin{aligned}\mathbb{E}_{t,x}\big[\tilde{V}(T,X_T) - \tilde{V}(t,x)\big] =& \mathbb{E}_{t,x}\Big[\int_t^T \mathcal{A}^u \tilde{V}(s,X_s)\, ds\Big] \\ \leq& \mathbb{E}_{t,x}\Big[-\int_t^T g\big(s,X_s,u_s\big)\, ds\Big]\end{aligned}$$

where the last step follows from the Hamilton-Jacobi-Bellman equation (5.7). Since (t,x) is fixed, $\tilde{V}(t,x)$ is deterministic. Thus

$$\begin{aligned}\tilde{V}(t,x) \geq& \mathbb{E}_{t,x}\Big[\int_t^T g(s,X_s,u_s)\, ds + \tilde{V}(T,X_T)\Big] \\ =& \mathbb{E}_{t,x}\Big[\int_t^T g(s,X_s,u_s)\, ds + h(X_T)\Big] = J(t,x;u) \qquad (5.13)\end{aligned}$$

for all $u \in \mathcal{U}$. Arbitrariness of u yields

$$\tilde{V}(t,x) \geq V(t,x).$$

(b) Now, apply the above argument to $u_s^* = \tilde{u}(s, X_s^*)$ for all $s \in [t,T]$. Then we get equality in (5.13) and u^* is the optimal solution, since

$$\tilde{V}(t,x) = V(t,x) = J(t,x;u^*).$$

□

The verification theorem (cf. Theorem 5.1.5) offers a technique to solve our stochastic control problem. It is described by the following steps.

Verification technique

Step 1 Determine $\tilde{u}(s,y)$ as a maximum point of (5.12) depending on the unknown function $\tilde{V}$.

Step 2: Determine $\tilde{V}(t,x)$ as a solution of

$$\begin{aligned}A^{\tilde{u}(t,x)}\tilde{V}(t,x) + g(t,x,\tilde{u}(t,x)) =& 0, \\ V(T,x) =& h(x) \quad (t,x) \in [0,T] \times \mathbb{R}^d.\end{aligned}$$

Step 3: Compute X_s^* as a solution of

$$\begin{aligned} dX_s =& \mu\big(s, X_s, \tilde{u}(s, X_s)\big)\, dt + \sigma\big(s, X_s, u^*(s, X_s)\big)\, dW_s \\ &+ \int_{\mathbb{R}^d} \gamma\big(X_{s-}, u^*(s-, X_{s-}), z\big)\, \bar{N}(ds, dz) \\ X_t =& x. \end{aligned}$$

Define $u^*(s) = \tilde{u}(s, X_s^*)$.

Step 4: Verify optimality, i.e. check all assumptions of Theorem 5.1.5.

There is a simple method to confirm that the candidate for the optimal control is indeed optimal.

Theorem 5.1.6 (Martingale optimality principle)**.** *Let $u \in \mathcal{U}(t,x)$ and define the following reward process by*

$$M_\vartheta^{u,t,x} := \int_t^\vartheta g\big(s, X_s^u, u_s\big)\, ds + V(\vartheta, X_\vartheta^u) \quad \textit{for all } \vartheta \in [t, T].$$

Then

(a) $\big(M_\vartheta^{u,t,x}\big)$ is a supermartingale for all $u \in \mathcal{U}(t,x)$.

(b) $\big(M_\vartheta^{u^*,t,x}\big)$ is a martingale if and only if u^* is the optimal control.

Proof. First note by assumptions (A1)-(A5) it holds that $\mathbb{E}|M_\vartheta^{u,t,x}| < \infty$ for all $\vartheta \in [t, T]$ (cf. Proposition 5.1.1).

(a) For $\tilde{\vartheta} \leq \vartheta$ it holds

$$\begin{aligned} &\mathbb{E}[M_\vartheta^{u,t,x} | \mathcal{F}_{\tilde{\vartheta}}] \\ =& \mathbb{E}\Big[\int_t^{\tilde{\vartheta}} g(s, X_s^u, u_s)\, ds + \int_{\tilde{\vartheta}}^\vartheta g(s, X_s^u, u_s)\, ds + V(\vartheta, X_\vartheta^u) \Big| \mathcal{F}_{\tilde{\vartheta}}\Big] \\ =& \int_t^{\tilde{\vartheta}} g(s, X_s^u, u_s)\, ds + \mathbb{E}\Big[\int_{\tilde{\vartheta}}^\vartheta g(s, X_s^u, u_s)\, ds + V(\vartheta, X_\vartheta^u) \Big| \mathcal{F}_{\tilde{\vartheta}}\Big] \\ \leq& \int_t^{\tilde{\vartheta}} g(s, X_s^u, u_s)\, ds + V(\tilde{\vartheta}, X_{\tilde{\vartheta}}^u) = M_{\tilde{\vartheta}}^{u,t,x}. \end{aligned}$$

The inequality results from the Bellman equation (cf. Theorem 5.1.2). Hence, $\big(M_\vartheta^{u,t,x}\big)$ is a supermartingale.

(b) First, let $\big(M_\vartheta^{u^*,t,x}\big)$ be a martingale. Then

$$V(t, x) = M_t^{u^*,t,x} = \mathbb{E}[M_T^{u^*,t,x} | \mathcal{F}_t] = J(t, x; u^*).$$

Therefore, u^* is optimal. Second, let u^* be optimal on $[t,T]$. We have by Bellman's optimality principle (cf. Corollary 5.1.3)

$$V(\tilde{\vartheta}, X^*_{\tilde{\vartheta}}) = J(\tilde{\vartheta}, X^*_{\tilde{\vartheta}}; u^*) \quad \text{for all } \tilde{\vartheta} \in [t,T].$$

Hence for all $\vartheta \geq \tilde{\vartheta}$ it yields

$$\begin{aligned} V(\tilde{\vartheta}, X^*_{\tilde{\vartheta}}) =& J(\tilde{\vartheta}, X^*_{\tilde{\vartheta}}; u^*) \\ =& \mathbb{E}\Big[\int_{\tilde{\vartheta}}^{\vartheta} g(s, X^{u^*}_s, u^*_s)\, ds + V(\vartheta, X^*_{\vartheta}) \big| \mathcal{F}_{\tilde{\vartheta}}\Big]. \end{aligned}$$

Therefore, we have equality in (a) and

$$M^{u^*,t,x}_{\tilde{\vartheta}} = \mathbb{E}[M^{u^*,t,x}_{\vartheta} | \mathcal{F}_{\tilde{\vartheta}}],$$

i.e. $\big(M^{u^*,t,x}_{\vartheta}\big)$ is a martingale.

□

Intuitively, the submartingale inequality of Theorem 5.1.6 yields that the difference $\mathbb{E}[M^{u,t,x}_{\vartheta}|\mathcal{F}_{\tilde{\vartheta}}] - M^{u,t,x}_{\tilde{\vartheta}}$ is the expected cost caused by using the non-optimal control over the time interval $[\tilde{\vartheta}, \vartheta]$ rather than switching to an optimal control already at time $\tilde{\vartheta}$.

Remark 13. The Hamilton-Jacobi Bellman verification theorem (cf. Theorem 5.1.5) is a special case of the martingale optimality principle (cf. Theorem 5.1.6) since by the Itô-Doeblin formula (cf. Theorem 3.5.1) we have

$$\begin{aligned} dM^{u,t,x}_{\vartheta} =& g(\vartheta, X_{\vartheta}, u_{\vartheta})\, d\vartheta + dV(\vartheta, X^u_{\vartheta}) \\ =& g(\vartheta, X_{\vartheta}, u_{\vartheta})\, d\vartheta + \mathcal{A}^u V(\vartheta, X_{\vartheta}) + \tilde{M}, \end{aligned}$$

where $\tilde{M}$ is the martingale term (cf. Theorem 3.5.2). Therefore, part (a) of Theorem 5.1.6 holds if and only if

$$g(\vartheta, X_{\vartheta}, u_{\vartheta})\, d\vartheta + \mathcal{A}^u V(\vartheta, X_{\vartheta}) \leq 0$$

for all $u \in \mathcal{U}(t,x)$, respectively part (b) if and only if

$$g(\vartheta, X_{\vartheta}, u^*_{\vartheta})\, d\vartheta + \mathcal{A}^{u^*} V(\vartheta, X_{\vartheta}) = 0$$

for the optimal control u^*.

5.2 Insurance model driven by Lévy dynamics

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space endowed with filtration $(\mathcal{F}_t)_{t\geq 0}$ that satisfies the usual hypotheses of completeness and right-continuity (cf. Protter (1990), Chapter I). The σ-algebra $\mathcal{F}_t$ represents the information available at time t and any decision is made upon this information. Let $(W_t)_{t\geq 0}$ be a d-dimensional standard Brownian motion and N an independent Poisson random measure with associated compensated Poisson random measure $\tilde{N}$ and intensity measure ν where we assume that ν is a Lévy measure. For notational convenience let us again define $\overline{N}(dt, dz) := N(dt, dz) - \mathbb{1}_{\{\|z\|<1\}}\nu(dz)dt$.

We consider an insurance company consisting of d lines of business. Let X_t^i be a stochastic process that denotes the current reserve of business line i at any time t where $i \in \{1, \ldots, d\}$. In such case that no reinsurance is realized and no investments on the financial market are made, the risk process evolves according to

$$X_t^i = x^i + c^i t + \sum_{k=1}^{r} \sigma^{C,ik} W_t^{C,k} - \int_0^t \int_{\mathbb{R}^d} y^i \, \overline{N}^C(ds, dy)$$

where $x^i > 0$ is the initial reserve of business line i, $c^i > 0$ is the (constant) premium income over time and $\sigma^{C,ik} \in \mathbb{R}$. The index C denotes the relation to the claim process.

A special feature of our construction allows the risk reserves to be dependent. One might be interested in insurance losses occurring in different lines of business supposing that different types of casualties are dependent. A simplified example for such possible dependencies is that of natural disasters which might affect several lines of business. Let us think of a storm tide. On the one hand a storm tide may cause claims on buildings. On the other hand it may contaminate drinking water which results in infections of people. Thus two lines of business of the insurance company are affected, the property as well as the health insurance. For a mathematical formulation of dependencies among Lévy processes we refer to Chapter 4 of this thesis.

In the following, we deal with the evolution of the reserves of an insurance company with $d \geq 1$ lines of business facing dependent risks where the risk reserves of the single business lines can be controlled by reinsurance and investments on a financial market.
That is, on the one hand our d-dimensional stochastic process can be controlled by a proportional reinsurance policy $b_t = (b_t^1, \ldots, b_t^d)$ which reinsures the fraction $1 - b_t^i$ of the incoming claims of business line i at any time t. In other words $b_t = (b_t^1, \ldots, b_t^d)$ denotes the fraction of the incoming claims that the company insures itself, the so-called retention level. We require

$b_t \in \mathbb{R}_+^d$, where $b^i > 1$ indicates that additional business should be signed to line of business i. Thus the risk reserve is now constructed by a dynamic choice of $b_t \in \mathbb{R}_+^d$. The corresponding equation for the risk reserve including proportional reinsurance is

$$\begin{aligned} X_t^{b,i} =& x^i + c^i \int_0^t b_s^i \, ds + \int_0^t b_s^i \sum_{k=1}^r \sigma^{C,ik} \, dW_s^{C,k} \\ & - \int_0^t \int_{\mathbb{R}^d} b_s^i y^i \, \overline{N}^C(ds, dy). \end{aligned}$$

This is based on the assumption, that reinsurance companies have the same safety loading as the insurance company itself. This is usually called cheap reinsurance. Of course, one might also think of other choices of reinsurance policies including mixtures (cf. Section 5.3.3). However, we restrict ourselves in the majority of cases to proportional reinsurance since it is the most tractable one.

On the other hand, the surplus of the insurance company can be invested in a Lévy-driven financial market. Suppose we have a market with d risky investments with price dynamics

$$dS_t^i = S_{t_-}^i \Big(\mu^i \, dt + \sum_{k=1}^r \sigma^{S,ik} \, dW_t^{S,k} + \int_{(-1,\infty)^d} z^i \, \overline{N}^S(ds, dz) \Big)$$

where $S_0^i = s^i > 0$, $\mu > 0$ and $\sigma^{S,ik} \in \mathbb{R}$ for $i \in \{1, \dots, d\}$. The index S denotes the relation to the stock price. The assumption that jumps in the financial market are larger than -1 is necessary for the stock prices to be positive (cf. Cont and Tankov (2004), Proposition 8.21, Doléans-Dade exponential). Our construction allows the evolution of the price processes $(S^1, \dots, S^d)$ to be dependent. Let $\delta_t = (\delta_t^1, \dots, \delta_t^d) \in \mathbb{R}^d$ be the amounts of money invested in the stock at any time t. We allow δ_t to become negative which represents short selling of stocks.

Thus our control action $a_t = (b_t, \delta_t)$ consists of two components where b_t specifies the retention level and δ_t the portfolio vector at any time t. We require $a_t = (b_t, \delta_t)$ to be predictable.

To sum up, the risk process of business line i controlled by reinsurance and investment on a financial market evolves according to

$$\begin{aligned}X_t^{a,i} =& x^i + \int_0^t \left(c^i b_s^i + \mu^i \delta_s^i\right) ds + \int_0^t \sum_{k=1}^r b_s^i \sigma^{C,ik}\, dW_s^{C,k} \\ &+ \int_0^t \sum_{k=1}^r \delta_s^i \sigma^{S,ik}\, dW_s^{S,k} \\ &+ \int_0^t \int_{\mathbb{R}^d} \int_{(-1,\infty)^d} \left(\delta_s^i z^i - b_s^i y^i\right) \overline{N}(ds, dz, dy)\end{aligned}$$

where $\overline{N} = \overline{N}^S \otimes \overline{N}^C$. We assume the claim process to be independent from the stock price process, which means that the stock price process and the claim process never jump together. In other words, the support of the Lévy measure ν is contained in the set $\{(y, z) : y \cdot z = 0\}$ (cf. Cont and Tankov (2004), Proposition 5.3).

5.3 Maximizing exponential utility

In this section we consider an insurance company that is interested in maximizing its utility from terminal wealth at time T. We further suppose that the insurance company has an exponential utility function

$$u(x^1, \ldots, x^d) = \lambda - \frac{\gamma}{\theta} \exp\Big(-\theta \sum_{i=1}^d x^i\Big) \tag{5.14}$$

where $\gamma > 0$ and $\theta > 0$. The function $V(t, x)$ describes the maximal expected exponential utility starting at time t in state x. That is,

$$V(t, x) = \sup_a \mathbb{E}[u(X_T^a) | X_t^a = x].$$

5.3.1 Proportional reinsurance and investment

We imbed the problem of maximizing exponential utility of terminal wealth in the framework of stochastic dynamic programming. Solving the Hamilton-Jacobi-Bellman equation, we can show that it is optimal to keep the retention level and investments constant, regardless of time and the company's level of wealth.

We need some integrability conditions. Let

$$\int_{\{\mathbb{R}^d, \|y\|>1\}} \|y\| \exp\left(\theta\Gamma\|y\|\right) \nu^C(dy) < \infty \tag{5.15a}$$

and

$$\int_{\{\mathbb{R}^d, \|z\|>1\}} \|z\| \exp\left(\theta\Lambda\|z\|\right) \nu^S(dz) < \infty \tag{5.15b}$$

for some constants $0 < \Gamma, \Lambda < \infty$.

Theorem 5.3.1. *The optimal strategy to maximize expected exponential utility at a terminal time T is to keep the retention level constantly equal to the solution $b^* = (b^{*,1}, \ldots, b^{*,d})$ of the equation*

$$0 = c^i - \theta \sum_{k=1}^{r} \sum_{j=1}^{d} b^{*,j} \sigma^{C,ik} \sigma^{C,jk} - \int_{\mathbb{R}^d} y^i \Big(\exp\Big(\theta \sum_{j=1}^{d} b^{*,j} y^j \Big) - \mathbb{1}_{\{\|y\|<1\}} \Big) \nu^C(dy)$$

over the interval $[0, \infty)^d$ and to invest at each time $t \leq T$ the constant amount equal to the solution $\delta^* = (\delta^{*,1}, \ldots, \delta^{*,d})$ of the equation

$$\begin{aligned} 0 =& \mu^i - \theta \sum_{k=1}^{r} \sum_{j=1}^{d} \delta^{*,j} \sigma^{S,ik} \sigma^{S,jk} \\ &+ \int_{(-1,\infty)^d} z^i \Big(\exp\Big(-\theta \sum_{j=1}^{d} \delta^{*,j} z^j \Big) - \mathbb{1}_{\{\|z\|<1\}} \Big) \nu^S(dz) \end{aligned}$$

over $\mathbb{R}^d$. Then the value function is

$$V(t,x) = \lambda - \frac{\gamma}{\theta} \exp\Big(-\theta \sum_{i=1}^{d} x^i + (T-t) h_c^* \Big)$$

where h_c^* is a constant given by

$$\begin{aligned} h_c^* =& -\theta \sum_{i=1}^{d} \big(c^i b^{*,i} + \mu^i \delta^{*,i} \big) \\ &+ \frac{1}{2} \theta^2 \sum_{i,j=1}^{d} \sum_{k=1}^{r} \big(b^{*,i} b^{*,j} \sigma^{C,ik} \sigma^{C,jk} + \delta^{*,i} \delta^{*,j} \sigma^{S,ik} \sigma^{S,jk} \big) \\ &+ \int_{\mathbb{R}^d} \Big(\exp\Big(\theta \sum_{i=1}^{d} b^{*,i} y^i \Big) - 1 - \theta \mathbb{1}_{\{\|y\|<1\}} \sum_{i=1}^{d} b^{*,i} y^i \Big) \nu^C(dy) \\ &+ \int_{(-1,\infty)^d} \Big(\exp\Big(-\theta \sum_{i=1}^{d} \delta^{*,i} z^i \Big) - 1 + \theta \mathbb{1}_{\{\|z\|<1\}} \sum_{i=1}^{d} \delta^{*,i} z^i \Big) \nu^S(dz). \end{aligned}$$

Proof. Suppose first that we search for the optimal control on a compact set $\|b\| \leq \Gamma < \infty$ and $\|\delta\| \leq \Lambda < \infty$. For the problem of maximizing exponential utility of terminal wealth at a fixed terminal time T, the Hamilton-Jacobi-Bellman equation becomes

$$\begin{aligned} \sup_{a \in \mathbb{R}_+^d \times \mathbb{R}^d} \mathcal{A}^a V(t,x) =& 0 \\ V(T,x) =& u(x). \end{aligned} \tag{5.16}$$

That means, we have to solve (5.16) for all (t,x) and derive $a(t,x)$ which maximizes the generator with respect to our wealth process given by

$$\begin{aligned}
&\mathcal{A}^a V(t,x)\\
=&\frac{\partial V(t,x)}{\partial t}+\sum_{i=1}^d \frac{\partial V(t,x)}{\partial x^i}\big(c^i b^i+\mu^i\delta^i\big)\\
&+\frac{1}{2}\sum_{i,j=1}^d\sum_{k=1}^r \frac{\partial^2 V(t,x)}{\partial x^i\partial x^j}\big(b^i b^j\sigma^{C,ik}\sigma^{C,jk}+\delta^i\delta^j\sigma^{S,ik}\sigma^{S,jk}\big)\\
&+\int_{\mathbb{R}^d}\int_{(-1,\infty)^d}\Big(V(t,x-by+\delta z)-V(t,x)\\
&\qquad\qquad -\sum_{i=1}^d \frac{\partial V(t,x)}{\partial x^i}(\delta^i z^i-b^i y^i)\mathbb{1}_{\{\|y\|^2+\|z\|^2<1\}}\Big)\,\nu(dz,dy)
\end{aligned}$$

where $x,y,z\in\mathbb{R}^d$. The Lévy measure ν is supported by the set $\{(y,z): y\cdot z=0\}$ which means that in the integrand y is always zero whenever $z\neq 0$ and vice versa. Therefore, it can represented in the form $\nu(B)=\nu^C(B_Y)+\nu^S(B_Z)$ where $B_Y=\{y:(y,0)\in B\}$ and $B_Z=\{z:(0,z)\in B\}$ for all $B\in\mathcal{B}(\mathbb{R}^d)$ (cf. Cont and Tankov (2004), Proposition 5.3). Therefore, the generator can be rewritten as

$$\begin{aligned}
&\mathcal{A}^a V(t,x)\\
=&\frac{\partial V(t,x)}{\partial t}+\sum_{i=1}^d \frac{\partial V(t,x)}{\partial x^i}\big(c^i b^i+\mu^i\delta^i\big)\\
&+\frac{1}{2}\sum_{i,j=1}^d\sum_{k=1}^r \frac{\partial^2 V(t,x)}{\partial x^i\partial x^j}\big(b^i b^j\sigma^{C,ik}\sigma^{C,jk}+\delta^i\delta^j\sigma^{S,ik}\sigma^{S,jk}\big)\\
&+\int_{\mathbb{R}^d}\Big(V(t,x-by)-V(t,x)+\sum_{i=1}^d \frac{\partial V(t,x)}{\partial x^i}b^i y^i\mathbb{1}_{\{\|y\|<1\}}\Big)\,\nu^C(dy)\\
&+\int_{(-1,\infty)^d}\Big(V(t,x+\delta z)-V(t,x)-\sum_{i=1}^d \frac{\partial V(t,x)}{\partial x^i}\delta^i z^i\mathbb{1}_{\{\|z\|<1\}}\Big)\,\nu^S(dz).
\end{aligned}$$

We choose the Ansatz (cf. Browne (1995))

$$\tilde{V}(t,x)=\lambda-\frac{\gamma}{\theta}\exp\Big(-\theta\sum_{i=1}^d x^i+h_c(T-t)\Big)$$

for a suitable constant h_c. Note that the boundary condition is $\tilde{V}(T,x)=$

$\lambda - \frac{\gamma}{\theta}\exp(-\theta\sum_{i=1}^{d} x^i) = u(x)$. For this trial solution we have

$$\frac{\partial \tilde{V}(t,x)}{\partial t} = -\, h_c(\tilde{V}(t,x) - \lambda) \tag{5.17}$$

$$\frac{\partial \tilde{V}(t,x)}{\partial x^i} = -\, \theta(\tilde{V}(t,x) - \lambda) \tag{5.18}$$

$$\frac{\partial^2 \tilde{V}(t,x)}{\partial x^i \partial x^j} = \theta^2(\tilde{V}(t,x) - \lambda). \tag{5.19}$$

Substituting (5.17), (5.18) and (5.19) into the generator gives

$$\begin{aligned}
&\mathcal{A}^a \tilde{V}(t,x) \\
=& -h_c(\tilde{V}(t,x) - \lambda) - \theta(\tilde{V}(t,x) - \lambda)\sum_{i=1}^{d}(c^i b^i + \mu^i \delta^i) \\
&+ \frac{1}{2}\theta^2(\tilde{V}(t,x) - \lambda)\sum_{i,j=1}^{d}\sum_{k=1}^{r}\big(b^i b^j \sigma^{C,ik}\sigma^{C,jk} + \delta^i \delta^j \sigma^{S,ik}\sigma^{S,jk}\big) \\
&+ \int_{\mathbb{R}^d}(\tilde{V}(t,x) - \lambda)\Big(\exp\Big(\theta\sum_{i=1}^{d} b^i y^i\Big) - 1 - \theta\sum_{i=1}^{d} b^i y^i \mathbb{1}_{\{\|y\|<1\}}\Big)\,\nu^C(dy) \\
&+ \int_{(-1,\infty)^d}(\tilde{V}(t,x) - \lambda)\Big(\exp\Big(-\theta\sum_{i=1}^{d}\delta^i z^i\Big) - 1 + \theta\sum_{i=1}^{d}\delta^i z^i \mathbb{1}_{\{\|z\|<1\}}\Big)\,\nu^S(dz).
\end{aligned}$$

Since $\tilde{V}(t,x) - \lambda < 0$ the Hamilton-Jacobi-Bellman equation gives

$$\sup_{(b,\delta)\in\mathbb{R}^d_+\times\mathbb{R}^d} \mathcal{L}(b,\delta) = 0$$

where

$$\begin{aligned}
\mathcal{L}(b,\delta) =& h_c + \theta\sum_{i=1}^{d}(c^i b^i + \mu^i \delta^i) \\
&- \frac{1}{2}\theta^2\sum_{i,j=1}^{d}\sum_{k=1}^{r}\big(b^i b^j \sigma^{C,ik}\sigma^{C,jk} + \delta^i \delta^j \sigma^{S,ik}\sigma^{S,jk}\big) \\
&- \int_{\mathbb{R}^d}\Big(\exp\Big(\theta\sum_{i=1}^{d} b^i y^i\Big) - 1 - \theta\mathbb{1}_{\{\|y\|<1\}}\sum_{i=1}^{d} b^i y^i\Big)\,\nu^C(dy) \\
&- \int_{(-1,\infty)^d}\Big(\exp\Big(-\theta\sum_{i=1}^{d}\delta^i z^i\Big) - 1 + \theta\mathbb{1}_{\{\|z\|<1\}}\sum_{i=1}^{d}\delta^i z^i\Big)\,\nu^S(dz)
\end{aligned}$$

which is concave in $(b,\delta) \in \mathbb{R}^d_+ \times \mathbb{R}^d$. Moreover, we notice that the maximum point $(\tilde{b},\tilde{\delta})$ is constant, both with respect to t and x. Let us now find the value $\tilde{a} = (\tilde{b}^1, \ldots, \tilde{b}^d, \tilde{\delta}^1, \ldots, \tilde{\delta}^d)$ which maximizes $\mathcal{L}(b, \delta)$. The first order condition for maximality of $\tilde{b}^i$ is

$$0 = c^i - \theta \sum_{k=1}^{r} \sum_{j=1}^{d} \tilde{b}^j \sigma^{C,ik} \sigma^{C,jk} - \int_{\mathbb{R}^d} y^i \Big(\exp\Big(\theta \sum_{j=1}^{d} \tilde{b}^j y^j \Big) - \mathbb{1}_{\{\|y\|<1\}} \Big) \nu^C(dy)$$

for all $i \in \{1, \ldots d\}$. Likewise the first order condition for maximality of $\tilde{\delta}^i$ is

$$\begin{aligned} 0 =& \mu^i - \theta \sum_{k=1}^{r} \sum_{j=1}^{d} \tilde{\delta}^j \sigma^{S,ik} \sigma^{S,jk} \\ &+ \int_{(-1,\infty)^d} z^i \Big(\exp\Big(-\theta \sum_{j=1}^{d} \tilde{\delta}^j z^j \Big) - \mathbb{1}_{\{\|z\|<1\}} \Big) \nu_S(dz) \end{aligned}$$

for all $i \in \{1, \ldots, d\}$. Let us denote the optimal policy by $a^* = (a^{*,1}, \ldots, a^{*,d})$ where $a^{*,i} = (b^{*,i}, \delta^{*,i})$. Inserting $a^* = (a^{*,1}, \ldots, a^{*,d})$ in the Hamilton-Jacobi-Bellman equation yields

$$\begin{aligned} h_c^* =& -\theta \sum_{i=1}^{d} \big(c^i b^{*,i} + \mu^i \delta^{*,i} \big) \\ &+ \frac{1}{2} \theta^2 \sum_{i,j=1}^{d} \sum_{k=1}^{r} \big(b^{*,i} b^{*,j} \sigma^{C,ik} \sigma^{C,jk} + \delta^{*,i} \delta^{*,j} \sigma^{S,ik} \sigma^{S,jk} \big) \\ &+ \int_{\mathbb{R}^d} \Big(\exp\Big(\theta \sum_{i=1}^{d} b^{*,i} y^i \Big) - 1 - \theta \mathbb{1}_{\{\|y\|<1\}} \sum_{i=1}^{d} b^{*,i} y^i \Big) \nu^C(dy) \\ &+ \int_{(-1,\infty)^d} \Big(\exp\Big(-\theta \sum_{i=1}^{d} \delta^{*,i} z^i \Big) - 1 + \theta \mathbb{1}_{\{\|z\|<1\}} \sum_{i=1}^{d} \delta^{*,i} z^i \Big) \nu^S(dz). \end{aligned} \tag{5.20}$$

In order to verify that both the control and the value function are in fact optimal let us finally apply the martingale optimality principle (cf. Theorem 5.1.6). For notational simplicity define

$$X_t^{a^*} = \sum_{i=1}^{d} X_t^{a^*,i}.$$

By the independence and stationarity of the increments it holds that

$$\mathbb{E}[\exp(-\theta(X_{t+s}^{a^*} - X_t^{a^*}))|\mathcal{F}_t] = \mathbb{E}[\exp(-\theta(X_{t+s}^{a^*} - X_t^{a^*}))] = \mathbb{E}[\exp(-\theta X_s^{a^*})]$$

which can be computed in terms of the characteristic triplet using the Lévy-Khinchin representation (cf. Theorem 3.2.2)

$$\mathbb{E}[\exp(-\theta X_s^{a^*})] = \exp(s h_c^*)$$

with h_c^* is as in (5.20). Thus under policy a^* the value function $\tilde{V}(t,x)$ is a martingale for the optimally controlled process $(X_t^{a^*})$. Note that

$$\mathbb{E}[\tilde{V}(t+s,x)|\mathcal{F}_t] = \tilde{V}(t,x)$$

if and only if

$$\mathbb{E}[\exp(-\theta(X_{t+s}^{a^*} - X_t^{a^*}))|\mathcal{F}_t] = \exp(sh_c^*).$$

Furthermore, it can be shown that $\tilde{V}(t, X_t^a)$ is a supermartingale under any other admissible policy $a \in \mathbb{R}_+^d \times \mathbb{R}^d$, which establishes optimality. Alternatively, we may check the conditions of the verification theorem (cf. Theorem 5.1.5).

Finally, we show that the moment-generating functions and its first derivatives used in this proof indeed exist. We only show the existence for the reinsurance term, for the financial market term one can proceed analogously. We require that

$$\left| \int_{\mathbb{R}^d} \left(\exp\left(\theta \sum_{i=1}^d b^i y^i\right) - 1 - \theta \mathbb{1}_{\{\|y\|<1\}} \sum_{i=1}^d b^i y^i \right) \nu^C(dy) \right| < \infty. \tag{5.21}$$

By Hölder's inequality we have for $m \geq 1$

$$\left(\sum_{i=1}^d |b^i y^i| \right)^m \leq \left(\sum_{i=1}^d |b^i|^2 \right)^{\frac{m}{2}} \left(\sum_{i=1}^d |y^i|^2 \right)^{\frac{m}{2}} = \|b\|^m \|y\|^m. \tag{5.22}$$

For $\|y\| \leq 1$ and $m \geq 2$ it holds that $\|y\|^m \leq \|y\|^2$. Condition (5.21) always holds for $\|y\| \leq 1$ since

$$\begin{aligned}
&\left| \int_{\{\mathbb{R}^d,\, \|y\|\leq 1\}} \left(\exp\left(\theta \sum_{i=1}^d b^i y^i\right) - 1 - \theta \sum_{i=1}^d b^i y^i \right) \nu^C(dy) \right| \\
=&\left| \int_{\{\mathbb{R}^d,\, \|y\|\leq 1\}} \sum_{m=2}^{\infty} \frac{\left(\theta \sum_{i=1}^d b^i y^i\right)^m}{m!} \nu^C(dy) \right| \\
\leq& \int_{\{\mathbb{R}^d,\, \|y\|\leq 1\}} \sum_{m=2}^{\infty} \frac{\theta^m \left(\sum_{i=1}^d |b^i y^i| \right)^m}{m!} \nu^C(dy) \\
\leq& \sum_{m=2}^{\infty} \frac{\theta^m \|b\|^m}{m!} \int_{\{\mathbb{R}^d,\, \|y\|\leq 1\}} \|y\|^2 \, \nu^C(dy) \\
=&(\exp\left(\theta\|b\|\right) - 1 - \theta\|b\|) \int_{\{\mathbb{R}^d,\, \|y\|\leq 1\}} \|y\|^2 \, \nu^C(dy) < \infty.
\end{aligned}$$

Note that $\|b\| \leq \Gamma$ and ν^C is a Lévy measure and therefore satisfies the integrability condition

$$\int_{\mathbb{R}^d} (\|y\|^2 \wedge 1) \, \nu^C(dy) < \infty.$$

Proceeding as in the case of $\|y\| \leq 1$ we obtain for $\|y\| \geq 1$

$$\begin{aligned}&\left|\int_{\{\mathbb{R}^d, \|y\| \geq 1\}} \left(\exp\left(\theta \sum_{i=1}^{d} b^i y^i\right) - 1\right) \nu^C(dy)\right| \\ &\leq \int_{\{\mathbb{R}^d, \|y\| \geq 1\}} \exp\left(\theta \|b\| \, \|y\|\right) \nu^C(dy) + \int_{\{\mathbb{R}^d, \|y\| \geq 1\}} \nu^C(dy).\end{aligned}$$

By the integrability condition (5.15a) and the properties of a Lévy measure it holds

$$\int_{\{\mathbb{R}^d, \|y\| \geq 1\}} \exp\left(\theta \|b\| \, \|y\|\right) \nu^C(dy) + \int_{\{\mathbb{R}^d, \|y\| \geq 1\}} \nu^C(dy) < \infty.$$

Note that $\|b\| < \infty$ on our compact set. For the first derivative we require

$$\left|\int_{\mathbb{R}^d} y^i \left(\exp\left(\theta \sum_{i=1}^{d} b^i y^i\right) - \mathbb{1}_{\{\|y\| < 1\}}\right) \nu^C(dy)\right| < \infty \tag{5.23}$$

for all $i \in \{1, \ldots, d\}$. Again we start with the case of $\|y\| \leq 1$. It holds that $\|y\|^{m+1} \leq \|y\|^2$ for $m \geq 1$. Taking into account (5.22) for $m \geq 1$ we obtain

$$\begin{aligned}&\left|\int_{\{\mathbb{R}^d, \|y\| \leq 1\}} y^i \left(\exp\left(\theta \sum_{i=1}^{d} b^i y^i\right) - 1\right) \nu^C(dy)\right| \\ &\leq \int_{\{\mathbb{R}^d, \|y\| \leq 1\}} \|y\|^2 \left|\sum_{m=1}^{\infty} \frac{\theta^m \left(\sum_{i=1}^{d} b^i y^i\right)^m}{m!}\right| \nu^C(dy) \\ &\leq \int_{\{\mathbb{R}^d, \|y\| \leq 1\}} \sum_{m=1}^{\infty} \frac{\theta^m \|b\|^m \|y\|^{m+1}}{m!} \nu^C(dy) \\ &\leq \left(\exp\left(\theta \|b\|\right) - 1\right) \int_{\{\mathbb{R}^d, \|y\| \leq 1\}} \|y\|^2 \nu^C(dy) < \infty.\end{aligned}$$

It remains to consider the case of $\|y\| \geq 1$. It holds that

$$\begin{aligned}&\left|\int_{\{\mathbb{R}^d, \|y\| \geq 1\}} y^i \exp\left(\theta \sum_{i=1}^{d} b^i y^i\right) \nu^C(dy)\right| \leq \int_{\{\mathbb{R}^d, \|y\| \geq 1\}} |y^i| \exp\left(\theta \|b\| \, \|y\|\right) \nu^C(dy) \\ &\leq \int_{\{\mathbb{R}^d, \|y\| \geq 1\}} \|y\| \exp\left(\theta \|b\| \, \|y\|\right) \nu^C(dy) < \infty\end{aligned}$$

for all $i \in \{1, \ldots, d\}$ by the integrability assumption (5.15a) and $\|b\| \leq \Gamma$. Since for any b we can find Γ such that $\|b\| \leq \Gamma$ and (5.15a) holds, this finally yields the assertion. □

Remark 14 (Existence of a solution).
Existence of a solution might be shown using a generalization of Brouwer's fixed point theorem which is known as the Bolzano-Poincaré-Miranda theorem (cf. Schäfer (2007)):

> Let $\mathcal{R} = \{x \in \mathbb{R}^d : |x^i| \leq L, i \in \{1, \dots, d\}\}$ and let $f : \mathcal{R} \to \mathbb{R}^d$ be continuous satisfying
>
> $$\begin{aligned} f^i(x^1, x^2, \dots, x^{i-1}, -L, x^{i+1}, \dots, x^d) &\geq 0 \\ f^i(x^1, x^2, \dots, x^{i-1}, +L, x^{i+1}, \dots, x^d) &\leq 0 \end{aligned}$$
>
> for all $i \in \{1, \dots, d\}$. Then $f(x) = 0$ has a solution in $\mathcal{R}$. For $d = 1$ this reduces to the well-known intermediate-value theorem.

Define for $i \in \{1, \dots, d\}$

$$\begin{aligned} f^i(b^1, \dots, b^d) =& c^i - \theta \sum_{j=1}^{d} \sum_{k=1}^{r} b^j \sigma^{C,ik} \sigma^{C,jk} \\ & - \int_{\mathbb{R}^d} y^i \Big(\exp \Big(\theta \sum_{j=1}^{d} b^j y^j \Big) - \mathbb{1}_{\{\|y\| < 1\}} \Big) \, \nu^C(dy). \end{aligned}$$

If the model parameters are chosen such that

$$\begin{aligned} f^i(b^1, b^2, \dots, b^{i-1}, -\Gamma, b^{i+1}, \dots, b^d) \geq 0 \\ \text{and } f^i(b^1, b^2, \dots, b^{i-1}, \Gamma, b^{i+1}, \dots, b^d) \leq 0 \end{aligned}$$

for all $i \in \{1, \dots, d\}$ then there exists a solution in $\mathcal{R} = \{b \in \mathbb{R}^d_+ : |b^i| \leq \Gamma, i \in \{1, \dots, d\}\}$ by Bolzano-Poincaré-Miranda theorem. A negative retention level indicates that no claims are retained. However, explicit conditions for the model parameters can not be derived. Conditions for the existence of a solution of (5.44) might be stated analogously.

Let us finish this subsection with an explicit computation of a proportional reinsurance policy which is possible in the pure diffusion case. For notational convenience let us give the result for $d = r = 2$ and let us assume that there are no investments on a financial market. Set $\sigma^C := \sigma$.

Corollary 5.3.2. *The optimal policy to maximize expected utility at a terminal time T is to choose, at each time $t \leq T$, constant retention levels*

$$b^{*,1} = \max \Big(0, \frac{c^1 \rho^2 - \rho^{12} c^2}{\theta(\rho^2 \rho^1 - (\rho^{12})^2)} \Big), \qquad b^{*,2} = \max \Big(0, \frac{c^2 \rho^1 - \rho^{12} c^1}{\theta(\rho^1 \rho^2 - (\rho^{12})^2)} \Big).$$

where $\rho^1 = (\sigma^{11})^2 + (\sigma^{12})^2$, $\rho^2 = (\sigma^{21})^2 + (\sigma^{22})^2$ and $\rho^{12} = \sigma^{11}\sigma^{21} + \sigma^{12}\sigma^{2,2}$. The optimal value function is

$$V(t, x) = \lambda - \frac{\gamma}{\theta} \exp \big(-\theta(x^1 + x^2) + h^*(T - t) \big),$$

where

$$h_c^* = -\theta(c^1 b^{*,1} + c^2 b^{*,2}) + \frac{1}{2}\theta^2((b^{*,1})^2\rho^1 + (b^{*,2})^2\rho^2 + 2b^{*,1}b^{*,2}\rho^{12}).$$

Proof. By Theorem 5.3.1 the first order conditions for maximality in case of no jumps are

$$\begin{aligned} 0 =& c^1 - \theta(b^1\rho^1 + b^2\rho^{12}) \\ 0 =& c^2 - \theta(b^2\rho^2 + b^1\rho^{12}). \end{aligned}$$

Solving this system of equations we get the optimal retention level. □

5.3.2 General reinsurance policy

Let us glance at the control by a general reinsurance policy. To be more precise, the insurance company can choose a retention level $b_t = (b_t^1, \ldots, b_t^d) \in [\underline{b}, \overline{b}]^d$. The function $h(b, y)$ describes the share of claim y paid by the insurance company itself depending on the retention level. Natural assumptions are $|h(b, y)| \leq |y|$ and $h(b, y)$ is increasing in b. For example, in the case of proportional reinsurance we have $h(b, y) = by$ with retention level $b \in [0, 1]^d$ or in the case of excess of loss reinsurance $h(b, y) = \min(b, y)$ with retention level $b \in [0, \infty]^d$. For each retention level b the insurer pays a premium rate to the reinsurer which has to be deducted from the premium rate $c = (c^1, \ldots, c^d) \in \mathbb{R}_+^d$ the policy holder pays to the insurance company. This leads to a so-called net-income rate $c(b) = (c^1(b^1), \ldots, c^d(b^d))$ that may be calculated according to the expected value principle. It holds that $0 \leq c(b) \leq c = c(\overline{b})$.

Thus the risk reserve is again constructed by a dynamical choice of the retention level $b_t \in [\underline{b}, \overline{b}]^d$. The corresponding risk reserve including general reinsurance is

$$X_t^{b,i} = x^i + \int_0^t c^i(b_s^i)\, ds + \int_0^t \sum_{k=1}^r \sigma^{i,k}\, dW_s^k - \int_0^t\int_{\mathbb{R}^d} h(b_s^i, y^i)\, \overline{N}(ds, dy).$$

We need the following integrability condition

$$\int_{\{\mathbb{R}^d, \|y\| \geq 1\}} \exp\left(\theta\sqrt{d}\|y\|\right) \nu(dy) < \infty. \tag{5.24}$$

Note that since there is no investment on a financial market we skip the claim index C. As in the preceding section we are interested in maximizing expected exponential utility of terminal wealth.

Theorem 5.3.3. *The optimal policy to maximize expected exponential utility at a terminal time T is to keep the retention level fixed, regardless of the time t and the insurance company's level of wealth x.*

Proof. For the problem of maximizing exponential utility of terminal wealth at a fixed terminal time T, the Hamilton-Jacobi-Bellman equation becomes

$$\begin{aligned}\sup_{b\in[\underline{b},\overline{b}]^d} \mathcal{A}^b V(t,x) =& 0 \\ V(T,x) =& u(x).\end{aligned} \tag{5.25}$$

That means, we have to solve (5.25) for all (t,x) and derive $b(t,x)$ which maximizes the generator with respect to our wealth process given by

$$\begin{aligned}&\mathcal{A}^b V(t,x)\\ =&\frac{\partial V(t,x)}{\partial t} + \sum_{i=1}^d \frac{\partial V(t,x)}{\partial x^i} c^i(b^i) + \frac{1}{2}\sum_{i,j=1}^d\sum_{k=1}^r \frac{\partial^2 V(t,x)}{\partial x^i \partial x^j}\sigma^{ik}\sigma^{jk}\\ &+ \int_{\mathbb{R}^d}\Big(V(t,x-h(b,y)) - V(t,x) + \sum_{i=1}^d \frac{\partial V(t,x)}{\partial x^i} h(b^i,y^i)\mathbb{1}_{\{\|y\|<1\}}\Big)\,\nu(dy).\end{aligned}$$

As before we choose the Ansatz

$$\tilde{V}(t,x) = \lambda - \frac{\gamma}{\theta}\exp\Big(-\theta\sum_{i=1}^d x^i + h_c(T-t)\Big)$$

for a suitable constant h_c. As required the boundary condition is $\tilde{V}(T,x) = \lambda - \frac{\gamma}{\theta}\exp(-\theta\sum_{i=1}^d x^i) = u(x)$. For our trial solution we have

$$\frac{\partial \tilde{V}(t,x)}{\partial t} = - h_c(\tilde{V}(t,x)-\lambda) \tag{5.26}$$

$$\frac{\partial \tilde{V}(t,x)}{\partial x^i} = - \theta(\tilde{V}(t,x)-\lambda) \tag{5.27}$$

$$\frac{\partial^2 \tilde{V}(t,x)}{\partial x^i \partial x^j} = \theta^2(\tilde{V}(t,x)-\lambda). \tag{5.28}$$

Substituting (5.26), (5.27) and (5.28) into the generator gives

$$\begin{aligned}&\mathcal{A}^b \tilde{V}(t,x)\\ =&(\tilde{V}(t,x)-\lambda)\Big(-h_c - \theta\sum_{i=1}^d c^i(b^i) + \frac{1}{2}\theta^2\sum_{i,j=1}^d\sum_{k=1}^r \sigma^{ik}\sigma^{jk}\Big)\\ &+\int_{\mathbb{R}^d}(\tilde{V}(t,x)-\lambda)\Big(\exp\Big(\theta\sum_{i=1}^d h(b^i,y^i)\Big) - 1 - \theta\sum_{i=1}^d h(b^i,y^i)\mathbb{1}_{\{\|y\|<1\}}\Big)\nu(dy).\end{aligned}$$

Since $\tilde{V}(t,x)-\lambda<0$ the Hamilton-Jacobi-Bellman equation gives

$$\sup_{b\in[\underline{b},\overline{b}]^d}\mathcal{L}(b) = 0 \tag{5.29}$$

where

$$\mathcal{L}(b) = h_c + \theta \sum_{i=1}^{d} c^i(b^i) - \frac{1}{2}\theta^2 \sum_{i,j=1}^{d} \sum_{k=1}^{r} \sigma^{ik}\sigma^{jk} - \int_{\mathbb{R}^d} \Big(\exp\Big(\theta \sum_{i=1}^{d} h(b^i, y^i)\Big) - 1 - \theta \mathbb{1}_{\{\|y\|<1\}} \sum_{i=1}^{d} h(b^i, y^i)\Big) \nu(dy).$$

Thus we notice that the optimal reinsurance policy is constant, both with respect to t and x.
In order to verify that the control which satisfies (5.29) is in fact optimal let us finally apply the martingale optimality principle (cf. Theorem 5.1.6). For notational simplicity define

$$X_t^{b^*} = \sum_{i=1}^{d} X_t^{b^*,i}.$$

By the independence and stationarity of the increments it holds that

$$\mathbb{E}[\exp(-\theta(X_{t+s}^{b^*} - X_t^{b^*}))|\mathcal{F}_t] = \mathbb{E}[\exp(-\theta(X_{t+s}^{b^*} - X_t^{b^*}))] = \mathbb{E}[\exp(-\theta X_s^{b^*})]$$

which can be computed in terms of the characteristic triplet using the Lévy-Khinchin representation (cf. Theorem 3.2.2). We have

$$\mathbb{E}[e^{-\theta X_s^{b^*}}] = \exp(s h_c^*)$$

where

$$h_c^* = -\theta \sum_{i=1}^{d} c^i(b^{*,i}) + \frac{1}{2}\theta^2 \sum_{i,j=1}^{d} \sum_{k=1}^{r} \sigma^{ik}\sigma^{jk} + \int_{\mathbb{R}^d} \Big(\exp\Big(\theta \sum_{i=1}^{d} h(b^{*,i}, y^i)\Big) - 1 - \theta \mathbb{1}_{\{\|y\|<1\}} \sum_{i=1}^{d} h(b^{*,i}, y^i)\Big) \nu(dy).$$

Thus under policy b^* the value function $\tilde{V}(t,x)$ is a martingale for the optimally controlled process $(X_t^{b^*})$. Furthermore, it can be shown that $\tilde{V}(t, X_t^b)$ is a supermartingale under any other admissible policy $b \in [\underline{b}, \overline{b}]^d$, which establishes optimality. Existence of the moment-generating function can be shown analogously to the proportional reinsurance case using the integrability condition (5.31) and $|h(b^i, y^i)| \leq |y^i|$. Note that for all $m \geq 1$

$$\Big(\sum_{i=1}^{d} \big|h(b^i, y^i)\big|\Big)^m \leq \Big(\sum_{i=1}^{d} \big|y^i\big|\Big)^m \leq (\sqrt{d})^m \|y\|^m.$$

□

5.3.3 Mixture of proportional and excess of loss reinsurance

In this subsection we allow the insurance company to control its risk reserve by a combination of proportional and excess of loss (XL) reinsurance.
Let $b_t = (b_t^1, \ldots, b_t^d)$ denote the retention level of the proportional reinsurance policy and $m_t = (m_t^1, \ldots, m_t^d)$ the retention level of the excess of loss reinsurance. We require $b_t \in [0,1]^d$ and $m_t \in \mathbb{R}_+^d$. Thus the risk reserve is now constructed by a dynamical choice of b_t and m_t. The risk reserve including proportional and excess of loss reinsurance of business line i, $i \in \{1, \ldots, d\}$ evolves according to

$$\begin{aligned} X_t^{(b,m),i} =& x^i + \int_0^t c^i(b_s^i, m_s^i)\, ds + \int_0^t \sum_{k=1}^r \sigma^{ik}\, dW_s^k \\ &- \int_0^t \int_{\mathbb{R}_+^d} (b_s^i y^i \wedge m_s^i)\, \overline{N}(ds, dy). \end{aligned}$$

The premium is calculated according to the expected value principle.

$$\begin{aligned} c^i(b_s^i, m_s^i) =& (1+\tilde{\eta})\mathbb{E}\Big(\int_{\mathbb{R}^d} y^i\, \overline{N}(ds, dy) \Big) \\ &- (1+\eta)\mathbb{E}\Big(\int_{\mathbb{R}^d} \big(y^i - (b_s^i y^i \wedge m_s^i)\big)\, \overline{N}(ds, dy) \Big) \end{aligned}$$

where $\eta > 0$ and $\tilde{\eta} > 0$ are the safety loadings of the insurance company and reinsurance company. For simplicity let us assume, that the reinsurance company has the same safety loading as the insurance company (i.e. cheap reinsurance). Thus we have

$$c^i(b_s^i, m_s^i) = (1+\eta) \int_{\{\mathbb{R}_+^d, \|y\| \geq 1\}} (b_s^i y^i \wedge m_s^i)\, \nu(dy). \tag{5.30}$$

Again, we imbed the problem of maximizing exponential utility of terminal wealth in the framework of stochastic dynamic programming. Solving the Hamilton-Jacobi-Bellman equation, we can show that there exists a pure excess of loss reinsurance strategy which is better than any combination of an excess of loss and proportional reinsurance strategy where $0 < b^i < 1$, $i \in \{1, \ldots, d\}$. Moreover, we get that the optimal retention level for the excess of loss reinsurance is to keep a fixed constant amount of the claims, regardless of the time and the company's level of wealth.
We need an integrability condition. Let

$$\int_{\{\mathbb{R}^d, \|y\| > 1\}} \|y\| \exp\big(\theta \|y\|\big)\, \nu(dy) < \infty. \tag{5.31}$$

Theorem 5.3.4. *The optimal policy to maximize expected exponential utility at a terminal time T is $(1, m^*) = (1, \dots, 1, m^{*,1}, \dots, m^{*,d})$ where the excess of loss retention level $m^{*,i}$ is given by the constant solution of*

$$0 = \int_{m^i}^{\infty} \int_{\mathbb{R}_+^{d-1}} \Big(\eta \mathbb{1}_{\{\|y\| \geq 1\}} + 1 - \exp\Big(\theta m^i + \theta \sum_{j \neq i}^{d} (y^j \wedge m^j) \Big) \Big) \, \nu(dy)$$

over the interval $[0, \infty)^d$. Then the value function is

$$V(t, x) = \lambda - \frac{\gamma}{\theta} \exp\Big(-\theta \sum_{i=1}^{d} x^i + (T - t) h_c^* \Big)$$

where h_c^* is a constant given by

$$\begin{aligned} h_c^* = & -\theta \sum_{i=1}^{d} c^i(1, m^{*,i}) + \frac{1}{2}\theta^2 \sum_{i,j=1}^{d} \sum_{k=1}^{r} \sigma^{ik} \sigma^{jk} \\ & + \int_{\mathbb{R}_+^d} \Big(\exp\Big(\theta \sum_{i=1}^{d} y^i \wedge m^{*,i} \Big) - 1 - \theta \mathbb{1}_{\{\|y\| < 1\}} \sum_{i=1}^{d} (y^i \wedge m^{*,i}) \Big) \, \nu(dy). \end{aligned}$$

Proof. For the problem of maximizing exponential utility of terminal wealth at a fixed terminal time T, the Hamilton-Jacobi-Bellman equation becomes

$$\begin{aligned} \sup_{(b,m) \in [0,1]^d \times \mathbb{R}_+^d} \mathcal{A}^{(b,m)} V(t, x) &= 0 \\ V(T, x) &= u(x). \end{aligned} \tag{5.32}$$

That means, we have to solve (5.32) for all (t, x) and derive $(b(t, x), m(t, x))$ which maximizes the generator with respect to our wealth process given by

$$\begin{aligned} & \mathcal{A}^{(b,m)} V(t, x) \\ = & \frac{\partial V(t, x)}{\partial t} + \sum_{i=1}^{d} \frac{\partial V(t, x)}{\partial x^i} c^i(b^i, m^i) + \frac{1}{2} \sum_{i,j=1}^{d} \sum_{k=1}^{r} \frac{\partial^2 V(t, x)}{\partial x^i \partial x^j} \sigma^{ik} \sigma^{jk} \\ & + \int_{\mathbb{R}_+^d} \Big(V\Big(t, x - b\Big(y \wedge \frac{m}{b}\Big)\Big) - V(t, x) \\ & \qquad + \sum_{i=1}^{d} \frac{\partial V(t, x)}{\partial x^i} \Big(b^i \Big(y^i \wedge \frac{m^i}{b^i} \Big) \Big) \mathbb{1}_{\{\|y\| < 1\}} \Big) \, \nu(dy). \end{aligned}$$

We choose the Ansatz

$$\tilde{V}(t, x) = \lambda - \frac{\gamma}{\theta} \exp\Big(-\theta \sum_{i=1}^{d} x^i + h_c(T - t) \Big)$$

for a suitable constant h_c. As required the boundary condition is $\tilde{V}(T,x) = \lambda - \frac{\gamma}{\theta}\exp(-\theta\sum_{i=1}^d x^i) = u(x)$. For our trial solution we have

$$\frac{\partial \tilde{V}(t,x)}{\partial t} = -\, h_c(\tilde{V}(t,x) - \lambda) \tag{5.33}$$

$$\frac{\partial \tilde{V}(t,x)}{\partial x^i} = -\, \theta(\tilde{V}(t,x) - \lambda) \tag{5.34}$$

$$\frac{\partial^2 \tilde{V}(t,x)}{\partial x^i \partial x^j} = \theta^2(\tilde{V}(t,x) - \lambda). \tag{5.35}$$

Substituting (5.33), (5.34) and (5.35) into the generator we obtain

$$\begin{aligned}
&\mathcal{A}^{(b,m)}\tilde{V}(t,x)\\
=&(\tilde{V}(t,x)-\lambda)\Big(-h_c - \theta\sum_{i=1}^d c^i(b^i,m^i) + \frac{1}{2}\theta^2\sum_{i,j=1}^d\sum_{k=1}^r \sigma^{ik}\sigma^{jk}\Big)\\
&+\int_{\mathbb{R}_+^d}(\tilde{V}(t,x)-\lambda)\Big(\exp\Big(\theta\sum_{i=1}^d b^i\Big(y^i\wedge\frac{m^i}{b^i}\Big)\Big)-1\\
&\qquad -\theta\sum_{i=1}^d b^i\Big(y^i\wedge\frac{m^i}{b^i}\Big)\mathbb{1}_{\{\|y\|<1\}}\Big)\,\nu(dy).
\end{aligned}$$

Since $\tilde{V}(t,x) - \lambda < 0$ the Hamilton-Jacobi-Bellman equation yields

$$\sup_{(b,m)\in[0,1]^d\times\mathbb{R}_+^d} \mathcal{L}(b,m) = 0$$

where

$$\begin{aligned}
&\mathcal{L}(b,m)\\
=&h_c + \theta\sum_{i=1}^d c^i(b^i,m^i) - \frac{1}{2}\theta^2\sum_{i,j=1}^d\sum_{k=1}^r\sigma^{ik}\sigma^{jk}\\
&-\int_{\mathbb{R}_+^d}\Big(\exp\Big(\theta\sum_{i=1}^d b^i\Big(y^i\wedge\frac{m^i}{b^i}\Big)\Big)-1-\theta\mathbb{1}_{\{\|y\|<1\}}\sum_{i=1}^d b^i\Big(y^i\wedge\frac{m^i}{b^i}\Big)\Big)\,\nu(dy).
\end{aligned}$$

Plugging in the premiums (5.30) we get

$$\begin{aligned}
&\mathcal{L}(b,m)\\
=&h_c + \theta(1+\eta)\sum_{i=1}^d\int_{\mathbb{R}_+^d} b^i\Big(y^i\wedge\frac{m^i}{b^i}\Big)\mathbb{1}_{\{\|y\|\geq 1\}}\,\nu(dy) - \frac{1}{2}\theta^2\sum_{i,j=1}^d\sum_{k=1}^r\sigma^{ik}\sigma^{jk}\\
&-\int_{\mathbb{R}_+^d}\Big(\exp\Big(\theta\sum_{i=1}^d b^i\Big(y^i\wedge\frac{m^i}{b^i}\Big)\Big)-1-\theta\mathbb{1}_{\{\|y\|<1\}}\sum_{i=1}^d b^i\Big(y^i\wedge\frac{m^i}{b^i}\Big)\Big)\,\nu(dy).
\end{aligned}$$

Thus we notice that the optimal policy is constant, both with respect to t and x. Differentiating with respect to m^i we get by Lemma A.2.1

$$\frac{\partial \mathcal{L}(b,m)}{\partial m^i} = \int_{\frac{m^i}{b^i}}^{\infty} \int_{\mathbb{R}_+^{d-1}} \left(\eta \mathbb{1}_{\{\|y\| \geq 1\}} + 1 - \exp\left(\theta m^i + \theta \sum_{j \neq i}^{d} b^j \left(y^j \wedge \frac{m^j}{b^j} \right) \right) \right) \nu(dy). \tag{5.36}$$

The first order condition for maximality of $\tilde{m}^i$ is

$$0 = \int_{\frac{m^i}{b^i}}^{\infty} \int_{\mathbb{R}_+^{d-1}} \left(\eta \mathbb{1}_{\{\|y\| \geq 1\}} + 1 - \exp\left(\theta m^i + \theta \sum_{j \neq i}^{d} b^j \left(y^j \wedge \frac{m^j}{b^j} \right) \right) \right) \nu(dy)$$

for all $i \in \{1, \dots d\}$. If

$$\tilde{m}^i(y^1, \dots, y^d) = \frac{\log\left(\eta \mathbb{1}_{\{\|y\| \geq 1\}} + 1 \right)}{\theta} - \sum_{j \neq i}^{d} b^j \left(y^j \wedge \frac{m^j}{b^j} \right) \tag{5.37}$$

the first order condition for maximality holds. Moreover, we have

$$\frac{\partial \mathcal{L}(b,m)}{\partial m^i} > 0$$

for all $m^i \in (0, \tilde{m}^i)$. Since $\mathcal{L}(b,m)$ is an increasing function with respect to $m^i \in (0, \tilde{m}^i)$ we obtain

$$m^{*,i} = \tilde{m}^i.$$

Differentiating with respect to b^i and plugging in $m^{*,i}$ yields (cf. Lemma A.2.1)

$$\begin{aligned} \frac{\partial \mathcal{L}(b,m)}{\partial b^i} = \int_0^{\frac{m^{*,i}}{b^i}} \int_{\mathbb{R}_+^{d-1}} y^i \Big(\eta \mathbb{1}_{\{\|y\| \geq 1\}} + 1 \qquad\qquad (5.38) \\ - \exp\left(\theta b^i y^i + \theta \sum_{j \neq i}^{d} b^j \left(y^j \wedge \frac{m^j}{b^j} \right) \right) \Big) \nu(dy). \end{aligned}$$

For $y^i \in [0, \frac{m^{*,i}}{b^i}]$ we have

$$\begin{aligned} &\eta \mathbb{1}_{\{\|y\| \geq 1\}} + 1 - \exp\left(\theta b^i y^i + \theta \sum_{j \neq i}^{d} b^j \left(y^j \wedge \frac{m^j}{b^j} \right) \right) \\ \geq &\eta \mathbb{1}_{\{\|y\| \geq 1\}} + 1 - \exp\left(\theta m^{*,i} + \theta \sum_{j \neq i}^{d} b^j \left(y^j \wedge \frac{m^j}{b^j} \right) \right) = 0 \end{aligned}$$

since $m^{*,i}$ equals (5.37). Therefore, it yields

$$\frac{\partial \mathcal{L}(b,m)}{\partial b^i} \geq 0$$

for all $b^i \in [0,1]$, $i \in \{1,\dots,d\}$. Thus, the optimal proportional reinsurance policy is $b^{*,i} = 1$ for all $i \in \{1,\dots,d\}$. We will denote the optimal policy by $a^* = (a^{*,1},\dots,a^{*,d})$ where $a^{*,i} = (1, m^{*,i})$.
Inserting $a^* = (a^{*,1},\dots,a^{*,d})$ in the Hamilton-Jacobi-Bellman equation yields

$$\begin{aligned} h_c^* = &- \theta \sum_{i=1}^{d} c^i(1, m^{*,i}) + \frac{1}{2}\theta^2 \sum_{i,j=1}^{d} \sum_{k=1}^{r} \sigma^{ik}\sigma^{jk} \\ &+ \int_{\mathbb{R}_+^d} \left(\exp\left(\theta \sum_{i=1}^{d} (y^i \wedge m^{*,i}) \right) - 1 - \theta \mathbb{1}_{\{\|y\|<1\}} \sum_{i=1}^{d} (y^i \wedge m^{*,i}) \right) \nu(dy). \end{aligned} \tag{5.39}$$

In order to verify that the control and the value function are in fact optimal let us finally apply the martingale optimality principle (cf. Theorem 5.1.6). For notational simplicity define

$$X_t^{a^*} = \sum_{i=1}^{d} X_t^{a^*,i}.$$

By the independence and stationarity of the increments it holds that

$$\mathbb{E}[\exp(-\theta(X_{t+s}^{a^*} - X_t^{a^*}))|\mathcal{F}_t] = \mathbb{E}[\exp(-\theta(X_{t+s}^{a^*} - X_t^{a^*}))] = \mathbb{E}[\exp(-\theta X_s^{a^*})]$$

which can be computed in terms of the characteristic triplet

$$\mathbb{E}[\exp(-\theta X_s^{a^*})] = \exp(s h_c^*)$$

where h_c^* is as in (5.39). Thus under policy a^* the value function $\tilde{V}(t,x)$ is a martingale for the optimally controlled process $(X_t^{a^*})$. Furthermore, it can be shown that $\tilde{V}(t, X_t^a)$ is a supermartingale under any other admissible policy $a \in [0,1]^d \times \mathbb{R}_+^d$, which establishes optimality.
For existence of the moment-generating function and the first derivative we can follow the lines of the proportional reinsurance case since $m^j \leq b^j y^j$ for $y^j \in (\frac{m^j}{b^j}, \infty)$, $b^j(y^j \wedge \frac{m^j}{b^j}) \leq b^j y^j$ for all $j \in \{1,\dots,d\}$ and $\|b\| \leq 1$ using the integrability condition (5.31). □

Theorem 5.3.4 shows that there exists a pure excess of loss policy which is always better than any combined reinsurance policy (b,m) where $0 < b^i < 1$ for all $i \in \{1,\dots,d\}$.

Remark 15. Choosing $m^i = \infty$ for all $i \in \{1, \dots, d\}$, the first order maximality condition

$$\frac{\partial \mathcal{L}(b, m)}{\partial b^i} = 0$$

for b^i, $i \in \{1, \dots, d\}$ (cf. equation (5.38)) corresponds to the maximality condition in the pure proportional reinsurance case (cf. Theorem 5.3.1) disregarding the diffusion term. To be more precise, the net premium income is

$$c^i(b^i, m^i) = b^i(1+\eta) \int_{\{\mathbb{R}^d_+, \|y\| \geq 1\}} y^i \, \nu(dy).$$

Therefore, the premium income $c = (c^1, \dots, c^d)$ corresponds to

$$c^i = (1+\eta) \int_{\{\mathbb{R}^d_+, \|y\| \geq 1\}} y^i \, \nu(dy). \tag{5.40}$$

for $i \in \{1, \dots, d\}$. Substituting (5.40) in (5.38), letting m^i tend to ∞ and rearranging the terms we obtain

$$\begin{aligned}
0 =& \int_{\mathbb{R}^d_+} y^i \Big(\eta \mathbb{1}_{\{\|y\| \geq 1\}} + 1 - \exp\Big(\theta \sum_{j=i}^d b^j y^j \Big) \Big) \, \nu(dy) \\
=& \int_{\mathbb{R}^d_+} y^i (\eta \mathbb{1}_{\{\|y\| \geq 1\}} + 1 - \mathbb{1}_{\{\|y\| \leq 1\}}) \, \nu(dy) \\
&- \int_{\mathbb{R}^d_+} y^i \Big(\exp\Big(\theta \sum_{j=i}^d b^j y^j \Big) - \mathbb{1}_{\{\|y\| \leq 1\}} \Big) \, \nu(dy) \\
=& c - \int_{\mathbb{R}^d_+} y^i \Big(\exp\Big(\theta \sum_{j=i}^d b^j y^j \Big) - \mathbb{1}_{\{\|y\| \leq 1\}} \Big) \, \nu(dy)
\end{aligned}$$

which corresponds to the maximality condition in the pure proportional reinsurance case (cf. Theorem 5.3.1) disregarding the diffusion term.

Example 5.3.5. We assume that our insurance company only consists of one business line and that claims are modeled by a compound Poisson process with jump size distribution F and intensity λ. The optimality equation for the optimal excess of loss retention level (cf. Theorem 5.3.4) reduces to

$$\begin{aligned}
0 =& \int_m^\infty \big(\eta - \exp(\theta m) + 1 \big) \, \nu(dy) = \lambda \big(\eta - \exp(\theta m) + 1 \big) \int_m^\infty F(dy) \\
=& \lambda \big(\eta + 1 - \exp(\theta m) \big) \big(1 - F(m) \big).
\end{aligned}$$

The optimal excess of loss retention level can be computed explicitly, that is

$$m^* = \frac{\ln(1+\eta)}{\theta}.$$

5.4 Minimizing probability of ruin in a pure diffusion model

In this section we focus on obtaining a retention level as well as an investment strategy which are optimal with respect to minimizing the probability of ruin. Inspired by Browne (1995) we then validate the conjecture that maximizing exponential utility of terminal wealth and minimizing probability of ruin are exactly equivalent in our multidimensional insurance model, at least in the case of no jumps.

Let us assume that ruin occurs when the weighted sum of net values of the business units are negative. Therefore, the ruin time is

$$\tau_{\text{ruin}} = \inf\Big\{t \geq 0 : X_t^a := \sum_{i=1}^d w^i X_t^{a,i} < 0, w^i \in \mathbb{R}_+\Big\}.$$

If the weights $w = (w^1, \ldots, w^d)$ are equal to one this corresponds to the ruin problem of the global company, i.e. ruin occurs when the aggregate sum of risk reserves of the single lines of business become negative. The risk process of business line i, $i \in \{1, \ldots, d\}$ controlled by proportional reinsurance and investments on a financial market evolves according to

$$\begin{aligned} X_t^{a,i} =& x^i + \int_0^t \big(c^i b_s^i + \mu^i \delta_s^i\big)\, ds \\ &+ \sum_{k=1}^r \Big(\int_0^t b_s^i \sigma^{C,ik}\, dW_s^{C,k} + \int_0^t \delta_s^i \sigma^{S,ik}\, dW_s^{S,k}\Big). \end{aligned}$$

In order to get the optimal policy in the sense of minimizing the ruin probability let us focus on maximizing the probability of reaching a given upper wealth barrier before hitting a given lower ruin barrier. To be more precise, let τ_z^a denote the first excess time of z under policy $a = (b, \delta) \in \mathbb{R}_+^d \times \mathbb{R}^d$, i.e.

$$\tau_z^a = \inf\{t > 0 : X_t^a = z\},$$

and for the particular two numbers α and β, let

$$\tau^a = \min\{\tau_\alpha^a, \tau_\beta^a\}.$$

We aim to find a control which maximizes

$$\mathbb{P}(X_{\tau^a}^a \geq \beta | X_0 = x)$$

where

$$\alpha < x := \sum_{i=1}^d w^i x^i < \beta.$$

Now, we start at $X_0 = x$ with $0 \leq \alpha < x < \beta \leq \infty$ and our goal is to maximize the probability of reaching the wealth level $\beta \in \mathbb{R}_+$ before exceeding the given lower ruin barrier $\alpha \in \mathbb{R}_+$. Denote by $V(x)$ the optimal value function, i.e.

$$V(x) = \sup_a \mathbb{P}(X^a_\tau \geq \beta | X_0 = x).$$

Theorem 5.4.1. *The optimal policy to maximize the probability of reaching point $\beta \in \mathbb{R}_+$ before $\alpha \in \mathbb{R}_+$ is to keep the retention levels constantly equal to the solution $b^* = (b^{*,1}, \ldots, b^{*,d})$ of the equation*

$$0 = c^i - \theta \sum_{j=1}^{d} \sum_{k=1}^{r} w^j b^{*,j} \sigma^{C,ik} \sigma^{C,jk} \tag{5.41}$$

over the interval $[0, \infty)^d$ and to invest the constant amount equal to the solution $\delta^* = (\delta^{*,1}, \ldots, \delta^{*,d})$ of the equation

$$0 = \mu^i - \theta \sum_{j=1}^{d} \sum_{k=1}^{r} w^j \delta^{*,j} \sigma^{S,ik} \sigma^{S,jk}. \tag{5.42}$$

The optimal value function is

$$V(x) = \frac{e^{-\theta\alpha} - e^{-\theta \sum_{i=1}^{d} w^i x^i}}{e^{-\theta\alpha} - e^{-\theta\beta}}.$$

Note that the optimal control $a^* = (a^{*,1}, \ldots, a^{*,d})$ does not depend on the barriers α and β. Therefore, letting β tend to infinity and α tend to zero we know that this is also the control which minimizes the ruin probability. The following corollary is a direct consequence of Theorem 5.4.1.

Corollary 5.4.2. *The optimal policy to minimize the probability of ruin is to keep the retention levels and the investment portfolio constantly equal to the solutions $b = (b^1, \ldots, b^d)$ and $\delta = (\delta^1, \ldots, \delta^d)$ of the equations (5.41) and (5.42).*

Remark 16. (a) The solution-equations (5.41) and (5.42) for the optimal controls with respect to ruin probability minimization coincide with the corresponding solution-equations for the optimal controls in the case that the performance criterion is the expected exponential utility of terminal wealth (cf. Theorem 5.4.1 and Theorem 5.3.1), choosing weights $w^i = 1$, $i \in \{1, \ldots, d\}$ and disregarding the jump part. This validates the conjecture that the policy that maximizes expected exponential utility of terminal wealth at a fixed terminal time coincides with the policy that minimizes the probability of ruin in case of no jumps.

(b) However, the situation is awkward as soon as jumps are involved. Boundary conditions have to be imposed not only at the boundary but as well outside the boundary. In fact they have to be extended to any point that the process can jump to from inside the domain which makes an explicit computation impossible.

Proof of Theorem 5.4.1. The Hamilton-Jacobi-Bellman equation is

$$\sup_{a\in\mathbb{R}_+^d\times\mathbb{R}^d} \mathcal{A}^a V(x) = 0$$

with the boundary conditions $V(\alpha) = 0$ and $V(\beta) = 1$ where $\alpha, \beta \in \mathbb{R}_+$. The quadratic variation of our wealth process satisfies

$$d\langle X\rangle_t = dt \sum_{i,j=1}^{d}\sum_{k=1}^{r} w^i w^j \left(b_t^i b_t^j \sigma^{C,ik}\sigma^{C,jk} + \delta_t^i\delta_t^j\sigma^{S,ik}\sigma^{S,jk}\right).$$

Applying the Itô-Doeblin formula (cf. Theorem 3.5.1) we obtain the generator with respect to our wealth process

$$\begin{aligned}\mathcal{A}^a V(x) =& \frac{\partial V(x)}{\partial x}\sum_{i=1}^{d} w^i\left(c^i b^i + \mu^i\delta^i\right)\\ &+ \frac{1}{2}\frac{\partial^2 V(x)}{\partial^2 x}\sum_{i,j=1}^{d}\sum_{k=1}^{r} w^i w^j\left(b^i b^j\sigma^{C,ik}\sigma^{C,jk} + \delta^i\delta^j\sigma^{S,ik}\sigma^{S,jk}\right)\end{aligned}$$

for $\alpha < x = \sum_{i=1}^d w^i x^i < \beta$. Let us choose the Ansatz

$$\tilde{V}(x) = \lambda - \frac{\gamma}{\theta}e^{-\theta x}.$$

For our trial solution we have

$$\begin{aligned}\frac{\partial\tilde{V}(x)}{\partial x} =& -\theta(\tilde{V}(x)-\lambda)\\ \frac{\partial^2\tilde{V}(x)}{\partial^2 x} =& \theta^2(\tilde{V}(x)-\lambda).\end{aligned}$$

Since $\tilde{V}(t,x) - \lambda < 0$ the Hamilton-Jacobi-Bellman equation gives

$$\sup_{(b,\delta)\in\mathbb{R}_+^d\times\mathbb{R}^d} \mathcal{L}(b,\delta) = 0$$

where

$$\begin{aligned}\mathcal{L}(b,\delta) =& \theta\sum_{i=1}^{d} w^i\left(c^i b^i + \mu^i\delta^i\right)\\ &- \frac{1}{2}\theta^2\sum_{i,j=1}^{d}\sum_{k=1}^{r} w^i w^j\left(b^i b^j\sigma^{C,ik}\sigma^{C,jk} + \delta^i\delta^j\sigma^{S,ik}\sigma^{S,jk}\right)\end{aligned}$$

which is concave in $(b^1, \ldots, b^d, \delta^1, \ldots, \delta^d)$. Thus we notice that the optimal policy is constant with respect to x and the hitting points α and β.
Let us now find the value $\tilde{a} = (\tilde{b}^1, \ldots, \tilde{b}^d, \tilde{\delta}^1, \ldots, \tilde{\delta}^d)$ which maximizes this function. The first order condition for maximality of $\tilde{b}^i$ is

$$0 = c^i - \theta \sum_{j=1}^{d} \sum_{k=1}^{r} w^j \tilde{b}^j \sigma^{C,ik} \sigma^{C,jk} \tag{5.43}$$

for all $i \in \{1, \ldots d\}$. And the first order condition for maximality of $\tilde{\delta}^i$ is

$$0 = \mu^i - \theta \sum_{j=1}^{d} \sum_{k=1}^{r} w^j \tilde{\delta}^j \sigma^{S,ik} \sigma^{S,jk}. \tag{5.44}$$

for all $i \in \{1, \ldots d\}$. Let us denote the optimal policy by $a^* = (a_1^*, \ldots, a_d^*)$ where $a_i^* = (b_i^*, \delta_i^*)$. In order to verify that the control and the value function are in fact optimal we apply the martingale optimality principle (cf. Theorem 5.1.6). By the independence and stationarity of the increments of our diffusion process and the Lévy-Khinchin representation (cf. Theorem 3.2.2) we obtain

$$\begin{aligned}
&\mathbb{E}[\exp(-\theta(X_{t+s}^{a^*} - X_t^{a^*}))|\mathcal{F}_t] = \mathbb{E}[\exp(-\theta(X_{t+s}^{a^*} - X_t^{a^*}))] = \mathbb{E}[\exp(-\theta X_s^{a^*})] \\
=& \exp\Big(s\Big(-\theta \sum_{i=1}^{d} w^i \big(c^i b^{*,i} + \mu^i \delta^{*,i}\big) \\
&\quad + \frac{1}{2}\theta^2 \sum_{i,j=1}^{d} \sum_{k=1}^{r} w^i w^j \big(b^{*,i} b^{*,j} \sigma^{C,ik} \sigma^{C,jk} + \delta^{*,i} \delta^{*,j} \sigma^{S,ik} \sigma^{S,jk}\big)\Big)\Big) \\
=& 1.
\end{aligned}$$

The last equality holds since b^* is a solution of (5.43) and δ^* is a solution of (5.44). This implies $\mathcal{L}(b^*, \delta^*) = 0$ which corresponds to the exponent. Thus under policy a^* the value function $\tilde{V}(x)$ is a martingale for the optimally controlled process $\big(X_t^{a^*}\big)$. Furthermore, it can be shown that $\tilde{V}(X_t^a)$ is a supermartingale under any other admissible policy $a \in \mathbb{R}_+^d \times \mathbb{R}^d$, which establishes optimality. The boundary condition $V(\alpha) = 0$ determines the constant λ as

$$\lambda = \frac{\gamma}{\theta} e^{-\theta\alpha},$$

and the boundary condition $V(\beta) = 1$ determines the constant γ as

$$\gamma = \frac{\theta}{e^{-\theta\alpha} - e^{-\theta\beta}}.$$

Therefore, the optimal value function is

$$V(x) = \frac{e^{-\theta\alpha} - e^{-\theta x}}{e^{-\theta\alpha} - e^{-\theta\beta}}.$$

□

Chapter 6

Structural comparison results

This chapter is dedicated to comparison results with respect to jumps in our insurance company's risk reserve. The performance criterion is still the expected exponential utility of terminal wealth. In the first section the optimal control in a compensated jump diffusion model is compared with the correspondent results in a pure diffusion model. Not surprisingly, it turns out that the optimal retention level of the insurance company is larger in the model without jumps than it is in the model containing jumps. In Section 6.2 we weaken the difference between the models and consider only models differing by the weighting of claims caused by jumps. There, we show that the accumulated risk reserve increases in concave order as the weighting factor of the claims decreases. However, the main focus of this chapter is put on the last section. Based on the results of Chapter 5 we devote Section 6.3 to identify structure conditions with respect to the Archimedean generator and the Lévy measure under which an insurance company reinsures a larger fraction of claims from one business line than from another. Similar results can be obtained with respect to investments on a financial market.

6.1 Comparison of the pure diffusion and jump diffusion case

We compare the optimal control in a compensated jump diffusion model with the correspondent results in a pure diffusion model. The optimal risk reserve of business line i, $i \in \{1, \ldots, d\}$, in the compensated jump diffusion model evolves according to

$$\begin{aligned}X_t^{a^*,i} =& x^i + (c^i b^{*,i} + \mu^i \delta^{*,i})t + \int_0^t \sum_{k=1}^r b^{*,i} \sigma^{C,ik}\, dW_s^{C,k} \\ &+ \int_0^t \sum_{k=1}^r \delta^{*,i} \sigma^{S,ik}\, dW_s^{S,k} \\ &+ \int_0^t \int_{\mathbb{R}_+^d} \int_{(-1,0]^d} (\delta^{*,i} z^i - b^{*,i} y^i)\, \tilde{N}(ds, dz, dy)\end{aligned}$$

where $a^* = (b^*, \delta^*) \in \mathbb{R}_+^{d\times d}$ denotes the optimal control which can be computed according to Theorem 5.3.1. Whereas the optimal risk reserve of business line i in a pure diffusion model follows

$$\begin{aligned}X_t^{a^{o*},i} =& x^i + (c^i b^{o*,i} + \mu^i \delta^{o*,i})t + \int_0^t \sum_{k=1}^r b^{o*,i} \sigma^{C,ik}\, dW_s^{C,k} \\ &+ \int_0^t \sum_{k=1}^r \delta^{o*,i} \sigma^{S,ik}\, dW_s^{S,k}\end{aligned}$$

where $a^{o*} = (b^{o*}, \delta^{o*}) \in \mathbb{R}_+^{d\times d}$ denotes the optimal control. It is important that we compare a pure diffusion process with a compensated jump diffusion process instead of a general Lévy process since a meaningful comparison is only possible if the means coincide.

Corollary 6.1.1. *Let (b^*, δ^*) be the optimal control in the compensated jump diffusion model and (b^{o*}, δ^{o*}) in the pure diffusion model respectively. Then we have*

$$b^{o*} \geq b^* \qquad \text{componentwise.}$$

Proof. Let $\mathcal{L}(b, \delta)$ and $\mathcal{L}^o(b, \delta)$ denote the Hamiltonian functions as defined in Theorem 5.3.1 for the compensated jump diffusion model and the pure diffusion model respectively. In the jump diffusion case there is only a finite number of jumps in each interval. Hence, $\mathcal{L}(b, \delta)$ reduces to

$$\begin{aligned}\mathcal{L}(b, \delta) =& h_c + \theta \sum_{i=1}^d (c^i b^i + \mu^i \delta^i) \\ &- \frac{1}{2}\theta^2 \sum_{i,j=1}^d \sum_{k=1}^r (b^i b^j \sigma^{C,ik} \sigma^{C,jk} + \delta^i \delta^j \sigma^{S,ik} \sigma^{S,jk}) \\ &- \int_{\mathbb{R}_+^d} \Big(\exp\Big(\theta \sum_{i=1}^d b^i y^i\Big) - 1\Big)\, \nu^C(dy) \\ &- \int_{(-1,0]^d} \Big(\exp\Big(-\theta \sum_{i=1}^d \delta^i z^i\Big) - 1\Big)\, \nu^S(dz).\end{aligned}$$

Let $\mathcal{L}^o(b,\delta)$ be the corresponding function in case of no jumps. Note that $\mathcal{L}$ and $\mathcal{L}^o$ are both concave in b and δ. Furthermore, we have for all $b^i \geq 0$, $i \in \{1,\dots,d\}$

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial b^i}(b,\delta) =& \theta c^i - \theta^2 \sum_{j=1}^d \sum_{k=1}^r b^j \sigma^{C,ik}\sigma^{C,jk} - \int_{\mathbb{R}_+^d} \theta y^i \exp\Big(\theta \sum_{i=1}^d b^i y^i\Big) \nu^C(dy) \\ \leq& \frac{\partial \mathcal{L}^o}{\partial b^i}(b,\delta)\end{aligned}$$

and for all $\delta^i \geq 0$, $i \in \{1,\dots,d\}$

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \delta^i}(b,\delta) =& \theta \mu^i - \theta^2 \sum_{j=1}^d \sum_{k=1}^r \delta^j \sigma^{S,ik}\sigma^{S,jk} \\ &+ \int_{(-1,0]} \theta z^i \exp\Big(-\theta \sum_{i=1}^d \delta^i z^i\Big) \nu^S(dz) \leq \frac{\partial \mathcal{L}^o}{\partial \delta^i}(b,\delta).\end{aligned}$$

Under the model assumptions of Chapter 5 the derivatives can be computed using Lebesgue's dominated convergence theorem (cf. the proof of Theorem 5.3.1). Since $\mathcal{L}$ and $\mathcal{L}^o$ are concave in $(b^1,\dots,b^d)$ and $(\delta^1,\dots,\delta^d)$ we have

$$b^{*,i} \leq b^{o*,i} \quad \text{and} \quad \delta^{*,i} \leq \delta^{o*,i}$$

for all $i \in \{1,\dots,d\}$. □

6.2 Comparison with respect to weighting of claims caused by jumps

We compare two insurance portfolios with different weighting of claims caused by jumps. More precisely, we consider the accumulated terminal risk reserves of the insurance company and insert weights $\beta, \tilde{\beta} \in \mathbb{R}_+$ with $\beta \leq \tilde{\beta}$ such that

$$X_T^{\beta} = \sum_{i=1}^d \big(x^i + c^i b^i T\big) + \sum_{i=1}^d \sum_{k=1}^r \sigma^{ik} W_T^k - \beta \sum_{i=1}^d \int_0^T \int_{\mathbb{R}^d} b^i y^i \, \overline{N}(ds,dy) \quad (6.1)$$

$$X_T^{\tilde{\beta}} = \sum_{i=1}^d \big(x^i + c^i b^i T\big) + \sum_{i=1}^d \sum_{k=1}^r \sigma^{ik} W_T^k - \tilde{\beta} \sum_{i=1}^d \int_0^T \int_{\mathbb{R}^d} b^i y^i \, \overline{N}(ds,dy). \quad (6.2)$$

We show that the accumulated risk reserve increases in concave order as the weighting factor of the jumps decreases. The increasing concave order is explained in the following definition adapted from Definition 1.5.1 in Müller and Stoyan (2002). The reason for choosing this order is that our performance criterion is still the expected utility of terminal wealth.

Definition 6.2.1 (Increasing concave order)**.** *Let X and $\tilde{X}$ be random variables with finite means. Then we say that $\tilde{X}$ is less than X in increasing concave order (written $\tilde{X} \leq_{icv} X$) if $\mathbb{E}f(\tilde{X}) \leq \mathbb{E}f(X)$ for all increasing concave functions f such that the expectations exist.*

Because of the relation $\tilde{X} \leq_{\text{icv}} X$ if and only if $-\tilde{X} \geq_{\text{icx}} -X$ the results of Müller and Stoyan (2002) for increasing convex functions can be applied in our context. Let us cite a sufficient condition for increasing convex order from Müller and Stoyan (2002). For that purpose we need another order definition which corresponds to Definition 1.5.16 in Müller and Stoyan (2002).

Definition 6.2.2 (Dangerous order)**.** *Let X and $\tilde{X}$ be random variables with distribution functions F_X and $F_{\tilde{X}}$. Then X is said to be less dangerous than $\tilde{X}$ (written $X \leq_D \tilde{X}$) if there is some $t_0 \in \mathbb{R}$ such that $F_X(t) \leq F_{\tilde{X}}(t)$ for all $t < t_0$ and $F_X(t) \geq F_{\tilde{X}}(t)$ for all $t \geq t_0$ and if in addition $\mathbb{E}X \leq \mathbb{E}\tilde{X}$.*

Theorem 6.2.3. *$X \leq_D \tilde{X}$ implies $X \leq_{icx} \tilde{X}$.*

Proof. For a proof we refer to Müller and Stoyan (2002), Theorem 1.5.17. □

We are now in the position to compare the two models (6.1) and (6.2) with respect to increasing concave order.

Theorem 6.2.4. *Let the risk reserves be as described in (6.1) and (6.2). Then*

$$X_T^{\beta} \geq_{icv} X_T^{\tilde{\beta}}.$$

Proof. For notational convenience define

$$J_T^{\beta} := \beta \int_0^T \int_{\mathbb{R}^d} b^i y^i \, \overline{N}(ds, dy) \quad \text{and} \quad J_T^{\tilde{\beta}} := \tilde{\beta} \int_0^T \int_{\mathbb{R}^d} b^i y^i \, \overline{N}(ds, dy).$$

By the independence of the diffusion and jump term of Lévy processes (cf. Sato (1999), Theorem 19.2) and since the increasing convex order is preserved under convolution (cf. Müller and Stoyan (2002), Theorem 1.5.5) it suffices to show that

$$J_T^{\tilde{\beta}} \geq_{icx} J_T^{\beta}.$$

It holds

$$\mathbb{P}(J_T^{\tilde{\beta}} \leq t) = \mathbb{P}\Big(\frac{\tilde{\beta}}{\beta} J_T^{\beta} \leq t\Big) = \mathbb{P}\Big(J_T^{\beta} \leq \frac{\beta}{\tilde{\beta}} t\Big).$$

Therefore

$$\mathbb{P}(J_T^{\beta} \leq t) - \mathbb{P}(J_T^{\tilde{\beta}} \leq t) \leq 0 \quad \text{for } t < 0$$
$$\mathbb{P}(J_T^{\beta} \leq t) - \mathbb{P}(J_T^{\tilde{\beta}} \leq t) \geq 0 \quad \text{for } t \geq 0.$$

That is,

$$J_T^{\tilde{\beta}} \geq_D J_T^{\beta}$$

and the assertion follows by Theorem 6.2.3. □

Corollary 6.2.5. *Let u be an utility function and let the risk reserves be as described in (6.1) and (6.2). Then*

$$\mathbb{E}u(X_T^{\beta}) \geq \mathbb{E}u(X_T^{\tilde{\beta}}).$$

6.3 Structure conditions for the Lévy measure

We identify structure conditions with respect to the Lévy measure under which an insurance company certainly reinsures a larger fraction of claims from one business line than from another. Intuitively, such a decision depends on the expected number of jumps in each business line as well as on the dependence structure between the single business lines. The dependence structure of a multivariate Lévy process can be characterized completely by the Lévy measure and the covariance matrix of the Brownian motion. The continuous term and the jump term of a Lévy process are independent (cf. Sato (1999), Theorem 19.2). Therefore, it suffices to consider the jump and diffusion part separately. In this section, let us focus on the jump part only, since the Brownian part is easy to handle.

In order to get meaningful comparisons we have to assume that there are only jumps in one direction. Moreover, we assume that the underlying Lévy process is of finite activity (cf. Proposition 3.2.4). Our performance criterion is still the expected exponential utility of terminal wealth (cf. (5.14)). By Theorem 5.3.1 we know that it is optimal to keep the control constant regardless of the time and the company's level of wealth. That means our dynamic optimization problem is now stationary. Define

$$\begin{aligned} X_T^* = \sum_{i=1}^d \Big(x^i + \int_0^t \left(c^i b^{*,i} + \mu^i \delta^{*,i} \right) ds \\ - \int_0^t \int_{\mathbb{R}_+^d} \int_{[0,1)^d} \left(\delta^{*,i} z^i + b^{*,i} y^i \right) N(ds, dz, dy) \Big). \end{aligned}$$

The value function can be computed in terms of the characteristic triplet of the accumulated Lévy process

$$\mathbb{E}[u(X_T^*)] = \lambda - \frac{\gamma}{\theta} \exp\left(T\psi(b^*, \delta^*) \right)$$

where

$$\psi(b,\delta) = -\,\theta\sum_{i=1}^{d}\big(cb^i+\mu\delta^i\big) + \int_{\mathbb{R}_+^d}\int_{[0,1)^d}\Big(\exp\Big(\theta\sum_{i=1}^{d}(b^iy^i+\delta^iz^i)\Big)-1\Big)\,\nu(dz,dy).$$

The Lévy measure ν is supported by the set $\{(y,z) : yz = 0\}$ which means that in the integrand y is always zero whenever $z \neq 0$ and vice versa. Therefore, the generator can be rewritten in terms of the Lévy measures $\nu^C(A) = \nu(A_Y)$ and $\nu^S(A) = \nu(A_Z)$ for all $A \in \mathcal{B}(\mathbb{R}^d)$ where $A_Y = \{y : (y,0) \in A\}$ and $A_Z = \{z : (0,z) \in A\}$. We can therefore analyze the retention level separately from the investment portfolio. Define

$$M(b) = \int_{\mathbb{R}_+^d}\Big(\exp\Big(\theta\sum_{i=1}^{d}b^iy^i\Big)-1\Big)\,\nu^C(dy).$$

We assume that the dependence structure between the business lines is specified via an Archimedean Lévy copula (cf. Definition 4.2.6) generated by a completely monotone Lévy generator (cf. Definition 4.2.7). We chose here $\phi := \psi^{-1}$ since this leads to simpler expressions in this context. Therefore, the inverse inverse ϕ^{-1} has derivatives up to order d with alternating signs satisfying

$$(-1)^k\big(\phi^{-1}\big)^{(k)}(t) \geq 0, \quad k \in \{1,\dots,d\} \tag{6.3}$$

for any $t \in [0,\infty)$. Moreover, suppose that the one-dimensional tail integrals are sufficiently smooth. Then we know by Proposition 4.2.5 that the Lévy density ν^C can be constructed by specifying the dependence structure and the marginal Lévy measures separately, that is

$$\nu^C(dy^1,\dots,dy^d) = \big((\phi^C)^{-1}\big)^{(d)}\Big(\sum_{i=1}^{d}\phi^C\big(U^{C,i}(y^i)\big)\Big)\prod_{i=1}^{d}s^{C,i}(y^i)\,\nu^{C,i}(dy^i),$$

where

$$s^{C,i}(\cdot) = (\phi^C)^{(1)}(U^{C,i}(\cdot)).$$

Thus our optimization problems can be written in terms of an Archimedean Lévy copula. We determine structure conditions with respect to the Lévy measure under which we can be sure that $b^{*,j} \geq b^{*,k}$ where $b^* = (b^{*,i},\dots,b^{*,d})$ is the optimal retention level. Inspired by the approach of Hennessy and Lapan (2002) we establish an inequality on the level of the expected terminal exponential utility function if $b^{*,j}$ is permuted with $b^{*,k}$ for $j,k \in \{1,\dots,d\}$.

Theorem 6.3.1. *If dependence is modeled via an Archimedean Lévy copula generated by a completely monotone Lévy generator and*

$$s^{C,j}(x)\nu^{C,j}(x) \geq s^{C,k}(x)\nu^{C,k}(x)$$

on the domain of definition, $j, k \in \{1, \dots, d\}$. *Then*

$$b^{*,j} \geq b^{*,k}$$

where $b^* = (b^{*,i}, \dots, b^{*,d})$ *is the optimal retention level.*

Proof. For notational convenience we skip the index C. In the sequel we assume the Lévy measure $\nu(dy)$ to be equivalent to the Lebesgue measure, that is $\nu(dy) = \nu(y)dy$, where $\nu(y)$ is the Lévy density of the Lévy measure $\nu(dy)$. Without loss of generality, consider permutation τ of indices 1 and 2 of an arbitrary retention vector $b \in \mathbb{R}^d_+$. Note that the retentions $(b^3, \dots, b^d)$ are held fixed. We write $b^\tau = (b^2, b^1, b^3, \dots, b^d)$ for the permuted policy and assume that $b^2 \geq b^1$. Comparing the two evaluations we have

$$\begin{aligned} &M(b^\tau) - M(b) \\ &= \int_{\mathbb{R}^d_+} \exp\left(\theta\Big(b^1y^2 + b^2y^1 + \sum_{i=3}^d b^iy^i\Big)\right)\nu(dy^1, \dots, dy^d) \\ &\quad - \int_{\mathbb{R}^d_+} \exp\left(\theta\sum_{i=1}^d b^iy^i\right)\nu(dy^1, \dots, dy^d). \end{aligned}$$

Separating the integration area and changing the notation of the variables in the second summand we obtain

$$\begin{aligned} &M(b^\tau) - M(b) \\ &= \int_{\{\mathbb{R}^d_+, y^1 \geq y^2\}} \exp\left(\theta\Big(b^1y^2 + b^2y^1 + \sum_{i=3}^d b^iy^i\Big)\right)\nu(dy^1, \dots, dy^d) \\ &\quad + \int_{\{\mathbb{R}^d_+, y^1 \geq y^2\}} \exp\left(\theta\sum_{i=1}^d b^iy^i\right)\nu(dy^2, dy^1, dy^3, \dots, dy^d) \\ &\quad - \int_{\{\mathbb{R}^d_+, y^1 \geq y^2\}} \exp\left(\theta\sum_{i=1}^d b^iy^i\right)\nu(dy^1, \dots, dy^d) \\ &\quad - \int_{\{\mathbb{R}^d_+, y^1 \geq y^2\}} \exp\left(\theta\Big(b^1y^2 + b^2y^1 + \sum_{i=3}^d b^iy^i\Big)\right)\nu(dy^2, dy^1, \dots, dy^d). \end{aligned}$$

Defining

$$\begin{aligned}&\mathcal{H}(y^1,\dots,y^d)\\=&\left(\phi^{-1}\right)^{(d)}\Big(\sum_{i=1}^d \phi\big(U^i(y^i)\big)\Big)\prod_{i=1}^d \nu^i(y^i)s^i(y^i)\\&-\left(\phi^{-1}\right)^{(d)}\Big(\phi\big(U^1(y^2)\big)+\phi\big(U^2(y^1)\big)+\sum_{i=3}^d \phi\big(U^i(y^i)\big)\Big)\\&\qquad\qquad \nu^1(y^2)s^1(y^2)\nu^2(y^1)s^2(y^1)\prod_{i=3}^d \nu^i(y^i)s^i(y^i),\end{aligned}$$

and

$$\mathcal{D}(y^1,y^2)=\exp\Big(\theta\sum_{i=1}^d b^i y^i\Big)$$

the above equation can be rewritten as

$$\begin{aligned}&M(b^\tau)-M(b)\\=&\int_{\{\mathbb{R}_+^d,y^1\geq y^2\}}\big(\mathcal{D}(y^2,y^1)-\mathcal{D}(y^1,y^2)\big)\mathcal{H}(y^1,\dots,y^d)\,dy^1\dots dy^d.\end{aligned}$$

In order to obtain an inequality for $M(b^\tau)-M(b)$ we integrate along y^2. That is,

$$\begin{aligned}&M(b^\tau)-M(b)\\=&\int_{\mathbb{R}_+^{d-1}}\big(\mathcal{D}(y^2,y^1)-\mathcal{D}(y^1,y^2)\big)\,\mathcal{J}(y^1,\dots,y^d)\Big|_0^{y^1}\,dy^1dy^3\dots dy^d\\&-\int_{\{\mathbb{R}_+^d,y^1\geq y^2\}}\theta\big(b^1\mathcal{D}(y^2,y^1)-b^2\mathcal{D}(y^1,y^2)\big)\,\mathcal{J}(y^1,\dots,y^d)\,dy^1\dots dy^d\end{aligned}$$

where

$$\begin{aligned}&\mathcal{J}(y^1,\dots,y^d)\\=&\left(\phi^{-1}\right)^{(d-1)}\Big(\sum_{i=1}^d \phi\big(U^i(y^i)\big)\Big)s^1(y^1)\nu^1(y^1)\prod_{i=3}^d s^i(y^i)\nu^i(y^i)\\&-\left(\phi^{-1}\right)^{(d-1)}\Big(\phi\big(U^1(y^2)\big)+\phi\big(U^2(y^1)\big)+\sum_{i=3}^d \phi\big(U^i(y^i)\big)\Big)\\&\qquad\qquad s^2(y^1)\nu^2(y^1)\prod_{i=3}^d s^i(y^i)\nu^i(y^i).\end{aligned}$$

The first term vanishes for $y^1 = y^2$. Upon integration along y^1 in the second term we have

$$\begin{aligned}
&M(b^\tau) - M(b) \\
&= -\int_{\mathbb{R}_+^{d-1}} \big(\mathcal{D}(0,y^1) - \mathcal{D}(y^1,0)\big)\,\mathcal{J}(y^1,0,y^3,\dots,y^d)\,dy^1dy^3\dots dy^d \\
&\quad - \int_{\mathbb{R}_+^{d-1}} \theta\big(b^1\mathcal{D}(y^2,y^1) - b^2\mathcal{D}(y^1,y^2)\big)\,\mathcal{K}(y^1,\dots,y^d)\Big|_{y^2}^{\infty} dy^2\dots dy^d \\
&\quad + \int_{\{\mathbb{R}_+^d, y^1\ge y^2\}} \theta^2 b^2 b^1 \big(\mathcal{D}(y^2,y^1) - \mathcal{D}(y^1,y^2)\big)\mathcal{K}(y^1,\dots,y^d)\,dy_1\dots dy_d
\end{aligned}$$

where

$$\begin{aligned}
&\mathcal{K}(y^1,\dots,y^d) \\
&=\big(\phi^{-1}\big)^{(d-2)}\Big(\sum_{i=1}^d \phi\big(U^i(y^i)\big)\Big)\prod_{i=3}^d s^i(y^i)\nu^i(y^i) \\
&\quad - \big(\phi^{-1}\big)^{(d-2)}\Big(\phi\big(U^1(y^2)\big) + \phi\big(U^2(y^1)\big) + \sum_{i=3}^d \phi\big(U^i(y^i)\big)\Big)\prod_{i=3}^d s^i(y^i)\nu^i(y^i).
\end{aligned}$$

The second term vanishes since $\mathcal{K}(y^1,y^2,y^3,\dots,y^d) = 0$ for $y^1 = y^2$ and $\mathcal{K}(y^1,y^2,y^3,\dots,y^d) = 0$ for $y^1 = \infty$ by Lemma A.2.2 keeping in mind that $U(\infty) = 0$ and $\phi(0) = \infty$. Integrating the first term along y^1 finally yields

$$\begin{aligned}
&M(b^\tau) - M(b) \\
&= -\int_{\mathbb{R}_+^{d-2}} \big(\mathcal{D}(0,y^1) - \mathcal{D}(y^1,0)\big)\,\mathcal{K}(y^1,0,y^3,\dots,y^d)\Big|_0^{\infty} dy^3\dots dy^d \\
&\quad + \int_{\mathbb{R}_+^{d-1}} \theta\big(b^2\mathcal{D}(0,y^1) - b^1\mathcal{D}(y^1,0)\big)\,\mathcal{K}(y^1,0,y^3,\dots,y^d)\,dy^1dy^3\dots dy^d \\
&\quad + \int_{\{\mathbb{R}_+^d, y^1\ge y^2\}} \theta^2 b^1 b^2 \big(\mathcal{D}(y^2,y^1) - \mathcal{D}(y^1,y^2)\big)\,\mathcal{K}(y^1,\dots,y^d)\,dy^1dy^2\dots dy^d.
\end{aligned}$$

Since $\mathcal{K}(y^1,0,y^3,\dots,y^d) = 0$ for $y^1 = \infty$ and $\mathcal{K}(y^1,0,y^3,\dots,y^d) = 0$ for $y^1 = 0$ it remains

$$\begin{aligned}
&M(b^\tau) - M(b) \\
&= \int_{\mathbb{R}_+^{d-1}} \theta\big(b^2\mathcal{D}(0,y^1) - b^1\mathcal{D}(y^1,0)\big)\,\mathcal{K}(y^1,0,y^3,\dots,y^d)\,dy^1dy^3\dots dy^d \\
&\quad + \int_{\{\mathbb{R}_+^d, y^1\ge y^2\}} \theta^2 b^1 b^2 \big(\mathcal{D}(y^2,y^1) - \mathcal{D}(y^1,y^2)\big)\,\mathcal{K}(y^1,\dots,y^d)\,dy^1dy^2\dots dy^d.
\end{aligned} \tag{6.4}$$

We establish conditions under which we can determine the sign of $\mathcal{K}(y^1, \ldots, y^d)$. Note that

$$\operatorname{sgn}\Big(\prod_{i=3}^{d} s^i(y^i)\nu^i(y^i)\Big) = (-1)^{d-2}.$$

By (6.3) we know that $(-1)^{d-2}(\phi^{-1})^{(d-2)}(\cdot)$ is a decreasing function since $(-1)^{d-2}(\phi^{-1})^{(d-1)}(\cdot) \leq 0$. Therefore, it holds that

$$\mathcal{K}(y^1, \ldots, y^d) \geq 0 \text{ for } y^1 \geq y^2$$

if

$$L(y^1, y^2) = \phi\big(U^1(y^1)\big) + \phi\big(U^2(y^2)\big) - \phi\big(U^1(y^2)\big) - \phi\big(U^2(y^1)\big) \leq 0.$$

Fix y^2. We have $y^1 \geq y^2$, $L(y^1, y^2)\big|_{y^1=y^2} = 0$ and a sufficient and necessary condition for $\frac{\partial L(y^1,y^2)}{\partial y^1} \leq 0$ to hold is that

$$s^1(y^1)\nu^1(y^1) \leq s^2(y^1)\nu^2(y^1).$$

We analogously get that $\mathcal{K}(y^1, 0, y^3, \ldots, y^d) \geq 0$ for $y^1 \geq 0$ if and only if $s^1(y)\nu^1(y) \leq s^2(y)\nu^2(y)$.
For $b^2 \geq b^1$ and $y^1 \in \mathbb{R}_+$ it holds that

$$b^2 \exp\Big(\theta\Big(b^2 y^1 + \sum_{i=3}^{d} b^i y^i\Big)\Big) \geq b^1 \exp\Big(\theta\Big(b^1 y^1 + \sum_{i=3}^{d} b^i y^i\Big)\Big).$$

With $K(y^1, 0, y^3, \ldots, y^d) \geq 0$ we therefore know that the first term in (6.4) is positive, so is the second term since for $b^2 \geq b^1$ and $y^1 \geq y^2$ it holds that

$$\exp\Big(\theta\Big(b^1 y^2 + b^2 y^1 + \sum_{i=3}^{d} b^i y^i\Big)\Big) \geq \exp\Big(\theta \sum_{i=1}^{d} b^i y^i\Big).$$

Note that we have $(b^2 - b^1)(y^1 - y^2) \geq 0$ for $y^1 \geq y^2$ if and only if $b^2 \geq b^1$. Hence

$$M(b^\tau) - M(b) \geq 0$$

for

$$s^1(y)\nu^1(y) \leq s^2(y)\nu^2(y) \text{ and } b^2 \geq b^1 \geq 0.$$

Clearly, the assumed allocation is preferable to its bivariate permutation. □

Remark 17. (a) Considering a more general model with different premium incomes in the single business lines c^i for $i \in \{1, \ldots, d\}$ we can easily show by Theorem 6.3.1: If

$$c^j \geq c^k \text{ and } s^{C,j}(y)\nu^{C,j}(y) \geq s^{C,k}(y)\nu^{C,k}(y)$$

on the domain of definition. Then

$$b^{*,j} \geq b^{*,k}.$$

(b) This approach is not valid to identify regularity conditions if there are jumps in several directions. In this case the definition of positive Lévy copulas has to be extended to the whole Euclidean space. The construction of such an extended Archimedean Lévy copula can be found in Chapter 4 of this thesis.

Our structure condition can be simplified to a condition which can be checked more easily

Proposition 6.3.2. *Let $U^k(x) \le U^j(x)$ for all x and let $\frac{U^k(x)}{U^j(x)}$ be decreasing in x, $j,k \in \{1,\dots,d\}$ fixed. For an Archimedean Lévy copula with generator ϕ such that*

$$R_\phi(u) = -u\frac{\phi^{(2)}(u)}{\phi^{(1)}(u)} \ge 1,$$

the structure condition of Theorem 6.3.1 is satisfied.

Proof. We can rewrite condition $s^j(x)\nu^j(x) \ge s^k(x)\nu^k(x)$ as

$$\frac{\nu^j(x)U^k(x)}{\nu^k(x)U^j(x)} \le \frac{\phi^{(1)}\big(U^k(x)\big)}{\phi^{(1)}\big(U^j(x)\big)}\frac{U^k(x)}{U^j(x)},$$

note that $s^i(x) = \phi^{(1)}\big(U^i(x)\big) < 0$.
Since $\frac{U^k(x)}{U^j(x)}$ is decreasing in x that is $\frac{\nu^j(x)U^k(x)}{\nu^k(x)U^j(x)} \le 1$ it suffices to show that

$$\frac{\phi^{(1)}\big(U^k(x)\big)}{\phi^{(1)}\big(U^j(x)\big)}\frac{U^k(x)}{U^j(x)} \ge 1$$

or equivalently

$$\phi^{(1)}\big(U^k(x)\big)U^k(x) \le \phi^{(1)}\big(U^j(x)\big)U^j(x).$$

Since $U^k(x) \le U^j(x)$ for all x it suffices to establish that $u\phi^{(1)}(u)$ is increasing in u. This condition is satisfied if and only if

$$\phi^{(1)}(u) + u\phi^{(2)}(u) \ge 0.$$

Recall that $\phi(u)$ is decreasing in u. Therefore, this condition is satisfied if and only if $-u\frac{\phi^{(2)}(u)}{\phi^{(1)}(u)} \ge 1$. □

Let us finally verify that there are indeed Archimedean Lévy copulas satisfying the condition of Theorem 6.3.1.

Example 6.3.3 (Clayton Lévy copula)**.**
Consider the positive Clayton Lévy copula with generator $\phi(u) = u^{-\frac{1}{\varrho}}$, $\varrho > 0$, $u \ge 0$. Therefore, $\phi^{(1)}(u) = -\frac{1}{\varrho}u^{-\frac{1}{\varrho}-1}$, $\phi^{(2)}(u) = \frac{1}{\varrho}(\frac{1}{\varrho}+1)u^{-\frac{1}{\varrho}-2}$ and

$$R_\phi(u) = -u\frac{\phi^{(2)}(u)}{\phi^{(1)}(u)} = \frac{-u\frac{1}{\varrho}(\frac{1}{\varrho}+1)u^{-\frac{1}{\varrho}-2}}{-\frac{1}{\varrho}u^{-\frac{1}{\varrho}-1}} = 1 + \frac{1}{\varrho} > 1.$$

Assuming that $\frac{U^2(x)}{U^1(x)}$ is decreasing in x and $U^2(x) \leq U^1(x)$ for all x the structure condition is satisfied by Proposition 6.3.2.

Example 6.3.4 (Gumbel Lévy copula)**.**
Consider the positive Gumbel Lévy copula with generator $\phi(u) = \big(\log(u+1)\big)^{-\varrho}$, $\varrho > 0$, $u > 0$. It yields

$$\phi^{(1)}(u) = \frac{-\varrho\big(\log(u+1)\big)^{-\varrho-1}}{u+1} < 0,$$
$$\phi^{(2)}(u) = \frac{\varrho(\varrho+1)\big(\log(u+1)\big)^{-\varrho-2}(u+1) + \varrho\big(\log(u+1)\big)^{-\varrho-1}}{(u+1)^2} > 0.$$

Thus

$$R_\phi(u) = u\frac{(\varrho+1)\big(\log(u+1)\big)^{-1}(u+1)+1}{(u+1)} \geq 1.$$

We therefore have to show that

$$f(u) := u(u+1)(\varrho+1) - \log(u+1) \geq 0.$$

This clearly holds since $f(0) = 0$ and $f'(u) = (\varrho+1)(2u+1) - \frac{1}{1+u} \geq \varrho > 0$. Assuming that $\frac{U^2(x)}{U^1(x)}$ is decreasing in x and $U^2(x) \leq U^1(x)$ for all x the structure condition is satisfied by Proposition 6.3.2.

Appendix A

Supplementary Material

A.1 The Burkholder-Davis-Gundy inequality

The following inequality can be found in Karatzas and Shreve (1991) for continuous local martingales. However, this inequality still holds in the case of an arbitrary martingale, so jumps are included. Details can be found in Dellacherie and Meyer (1982), Theorem VII.92. In the following we use the convention

$$M_t^* = \max_{0 \le s \le t} \|M_s\|.$$

Theorem A.1.1 (The Burkholder-Davis-Gundy inequality)**.** *Let M be an arbitrary local martingale. For every $m > 0$ there exists a constant K_m (depending only on m) such that for every stopping time τ*

$$\mathbb{E}[(M_\tau^*)^{2m}] \le K_m \, \mathbb{E}[\langle M \rangle_\tau^m].$$

Corollary A.1.1. *Let M a d-dimensional local martingale. There exists a constant K_m such that for all $m > 0$ and every stopping time τ*

$$\mathbb{E}[(M_\tau^*)^{2m}] \le K_m \, \mathbb{E}[A_\tau^m] \quad \textit{where } A_t = \sum_{i=1}^d \langle M_i \rangle_t \quad 0 \le t < \infty.$$

Proof. With inequality (3.6) we have

$$\|M_t\|^{2m} = \Big[\sum_{i=1}^d (M_t^i)^2\Big]^m \le d^m \sum_{i=1}^d |M_t^i|^{2m}$$

and

$$\sum_{i=1}^d \langle M_i \rangle_\tau^m \le d\Big(\sum_{i=1}^d \langle M_i \rangle_\tau\Big)^m = dA_\tau^m.$$

Taking maxima and expectation in the first inequality and maxima in the second inequality we obtain by the Burkholder-Davis-Gundy-inequality (cf. Theorem A.1.1)

$$\mathbb{E}(\|M\|_\tau^*)^{2m} \leq d^m \sum_{i=1}^d \mathbb{E}[(M_\tau^i)^*]^{2m} \leq d^m \sum_{i=1}^d K_m \mathbb{E}[\langle M^i \rangle_\tau^m] \leq K_m d^{m+1} \mathbb{E}(A_\tau^m).$$

□

A.2 Selected proofs

The following lemma is needed for the proof of Theorem 5.3.1.

Lemma A.2.1. *Let $\mathcal{L}(b,m)$ be as derived in the proof of Theorem 5.3.1. Then*

$$\frac{\partial \mathcal{L}(b,m)}{\partial m^i} = \int_{\frac{m^i}{b^i}}^\infty \int_{\mathbb{R}_+^{d-1}} \Big(\eta \mathbb{1}_{\{\|y\|\geq 1\}} + 1 - \exp\Big(\theta m^i + \theta \sum_{j\neq i}^d b^j \Big(y^j \wedge \frac{m^j}{b^j}\Big)\Big)\Big)\nu(dy)$$

$$\frac{\partial \mathcal{L}(b,m)}{\partial b^i} = \int_0^{\frac{m^i}{b^i}} \int_{\mathbb{R}_+^{d-1}} y^i \Big(\eta \mathbb{1}_{\{\|y\|\geq 1\}} + 1 - \exp\Big(\theta b^i y^i + \theta \sum_{j\neq i}^d b^j \Big(y^j \wedge \frac{m^j}{b^j}\Big)\Big)\Big)\,\nu(dy).$$

Remark 18 (Differentiation rule for parameter integrals).
Let $f \in C^{1,\cdot}$ and $\varphi \in C^1$ then

$$\frac{\partial}{\partial x}\int_a^{\varphi(x)} f(x,t)\,dt = \int_a^{\varphi(x)} \frac{\partial f(x,t)}{\partial x}\,dt + f(x,\varphi(x))\varphi'(x).$$

Therefore,

$$\frac{\partial}{\partial x}\int_{\varphi(x)}^{a} f(x,t)\,dt = -\frac{\partial}{\partial x}\int_a^{\varphi(x)} f(x,t)\,dt = \int_{\varphi(x)}^{a} \frac{\partial f(x,t)}{\partial x}\,dt - f(x,\varphi(x))\varphi'(x).$$

Proof. Let $\mathcal{L}(b,m)$ be as derived in the proof of Theorem 5.3.1. Defining

$$\mathcal{E}(y) = \theta \sum_{j\neq i}^d b^j \Big(y^j \wedge \frac{m^j}{b^j}\Big)$$

and separating the integrating area of $\mathcal{L}(b,m)$ we have

$$\begin{aligned}
&\mathcal{L}(b,m)\\
=&\theta(1+\eta)\sum_{i=1}^{d}\Big(\int_0^{\frac{m^i}{b^i}}\int_{\mathbb{R}_+^{d-1}} b^i y^i \mathbb{1}_{\{\|y\|\geq 1\}}\,\nu(dy)+\int_{\frac{m^i}{b^i}}^{\infty}\int_{\mathbb{R}_+^{d-1}} m^i \mathbb{1}_{\{\|y\|\geq 1\}}\,\nu(dy)\Big)\\
&+h_c-\frac{1}{2}\theta^2\sum_{i,j=1}^{d}\sum_{k=1}^{r}\sigma^{ik}\sigma^{jk}\\
&-\int_0^{\frac{m^i}{b^i}}\int_{\mathbb{R}_+^{d-1}}\big(\exp\big(\theta b^i y^i+\mathcal{E}(y)\big)-1-\theta\mathbb{1}_{\{\|y\|<1\}}\big(b^i y^i+\mathcal{E}(y)\big)\big)\,\nu(dy)\\
&-\int_{\frac{m^i}{b^i}}^{\infty}\int_{\mathbb{R}_+^{d-1}}\big(\exp\big(\theta m^i+\mathcal{E}(y)\big)-1-\theta\mathbb{1}_{\{\|y\|<1\}}\big(m^i+\mathcal{E}(y)\big)\big)\,\nu(dy).
\end{aligned}$$

In the sequel we assume the Lévy measure has a density with respect to the Lebesgue measure, that is

$$\nu(dy)=\nu(y)\,dy.$$

For notational simplicity we define

$$\mathcal{I}(y)=\sqrt{\sum_{\substack{j=1\\ j\neq i}}^{d}(y^j)^2+\Big(\frac{m^i}{b^i}\Big)^2}.$$

Differentiation with respect to m^i yields

$$\begin{aligned}
&\frac{\partial\mathcal{L}(b,m)}{\partial m^i}\\
=&\theta(1+\eta)\Big(\frac{1}{b^i}\int_{\mathbb{R}_+^{d-1}} m^i\mathbb{1}_{\{\mathcal{I}(y)\geq 1\}}\nu(y)\Big|_{y^i=\frac{m^i}{b^i}}\prod_{\substack{j=1\\ j\neq i}}^{d}dy^j+\int_{\frac{m^i}{b^i}}^{\infty}\int_{\mathbb{R}_+^{d-1}}\mathbb{1}_{\{\|y\|\geq 1\}}\nu(y)dy\\
&\qquad-\frac{1}{b^i}\int_{\mathbb{R}_+^{d-1}} m^i\mathbb{1}_{\{\mathcal{I}(y)\geq 1\}}\nu(y)\Big|_{y^i=\frac{m^i}{b^i}}\prod_{\substack{j=1\\ j\neq i}}^{d}dy^j\Big)\\
&-\frac{1}{b^i}\int_{\mathbb{R}_+^{d-1}}\big(\exp(\theta m^i+\mathcal{E}(y))-1-\theta\mathbb{1}_{\{\mathcal{I}(y)<1\}}(m^i+\mathcal{E}(y))\big)\nu(y)\Big|_{y^i=\frac{m^i}{b^i}}\prod_{\substack{j=1\\ j\neq i}}^{d}dy^j\\
&-\int_{\frac{m^i}{b^i}}^{\infty}\int_{\mathbb{R}_+^{d-1}}\big(\theta\exp(\theta m^i+\mathcal{E}(y))-\theta\mathbb{1}_{\{\|y\|<1\}}\big)\nu(y)\,dy\\
&+\frac{1}{b^i}\int_{\mathbb{R}_+^{d-1}}\big(\exp(\theta m^i+\mathcal{E}(y))-1-\theta\mathbb{1}_{\{\mathcal{I}(y)<1\}}(m^i+\mathcal{E}(y))\big)\nu(y)\Big|_{y^i=\frac{m^i}{b^i}}\prod_{\substack{j=1\\ j\neq i}}^{d}dy^j.
\end{aligned}$$

Note that some terms cancel out. We finally obtain

$$\begin{aligned}\frac{\partial \mathcal{L}(b,m)}{\partial m^i} =&\theta(1+\eta)\int_{\frac{m^i}{b^i}}^{\infty}\int_{\mathbb{R}_+^{d-1}} \mathbb{1}_{\{\|y\|\geq 1\}}\nu(y)\,dy\\ &-\int_{\frac{m^i}{b^i}}^{\infty}\int_{\mathbb{R}_+^{d-1}}\big(\theta\exp(\theta(m^i+\mathcal{E}(y)) - \theta\mathbb{1}_{\{\|y\|<1\}}\big)\nu(y)\,dy\\ =&\theta\int_{\frac{m^i}{b^i}}^{\infty}\int_{\mathbb{R}_+^{d-1}}\big(\eta\mathbb{1}_{\{\|y\|\geq 1\}} - \exp(\theta(m^i+\mathcal{E}(y))+1\big)\nu(y)\,dy.\end{aligned}$$

Differentiation with respect to b^i yields

$$\begin{aligned}&\frac{\partial \mathcal{L}(b,m)}{\partial b^i}\\ =&\theta(1+\eta)\Big(\int_0^{\frac{m^i}{b^i}}\int_{\mathbb{R}_+^{d-1}} y^i\mathbb{1}_{\{\|y\|\geq 1\}}\nu(y)\,dy\\ &\qquad - \frac{m^i}{(b^i)^2}\int_{\mathbb{R}_+^{d-1}} b^i\frac{m^i}{b^i}\mathbb{1}_{\{\mathcal{I}(y)\geq 1\}}\nu(y)\big|_{y^i=\frac{m^i}{b^i}}\prod_{\substack{j=1\\ j\neq i}}^{d} dy^j\\ &\qquad + \frac{m^i}{(b^i)^2}\int_{\mathbb{R}_+^{d-1}} m^i\mathbb{1}_{\{\mathcal{I}(y)\geq 1\}}\nu(y)\big|_{y^i=\frac{m^i}{b^i}}\prod_{\substack{j=1\\ j\neq i}}^{d} dy^j\Big)\\ &-\int_0^{\frac{m^i}{b^i}}\int_{\mathbb{R}_+^{d-1}}\theta y^i\big(\exp\big(\theta(b^iy^i+\mathcal{E}(y))\big) - \mathbb{1}_{\{\|y\|<1\}}\big)\nu(y)\,dy\\ &+\frac{m^i}{(b^i)^2}\int_{\mathbb{R}_+^{d-1}}\big(\exp\big(\theta(m^i+\mathcal{E}(y))\big)-1\\ &\qquad\qquad - \theta\mathbb{1}_{\{\mathcal{I}(y)<1\}}(m^i+\mathcal{E}(y))\big)\nu(y)\big|_{y^i=\frac{m^i}{b^i}}\prod_{\substack{j=1\\ j\neq i}}^{d} dy^j\\ &-\frac{m^i}{(b^i)^2}\int_{\mathbb{R}_+^{d-1}}\big(\exp(\theta(m^i+\mathcal{E}(y))-1\\ &\qquad\qquad - \theta\mathbb{1}_{\{\mathcal{I}(y)<1\}}(m^i+\mathcal{E}(y))\big)\nu(y)\big|_{y^i=\frac{m^i}{b^i}}\prod_{\substack{j=1\\ j\neq i}}^{d} dy^j.\end{aligned}$$

Note that some terms cancel out. We finally obtain

$$
\begin{aligned}
&\frac{\partial \mathcal{L}(b,m)}{\partial b^i}\\
=&\theta(1+\eta)\int_0^{\frac{m^i}{b^i}}\int_{\mathbb{R}_+^{d-1}} y^i \mathbb{1}_{\{\|y\|\geq 1\}}\nu(y)\,dy\\
&-\int_0^{\frac{m^i}{b^i}}\int_{\mathbb{R}_+^{d-1}} \big(\theta y^i \exp\big(\theta(b^i y^i + \mathcal{E}(y))\big) - \theta \mathbb{1}_{\{\|y\|<1\}} y_i\big)\nu(y)\,dy\\
=&\theta\int_0^{\frac{m^i}{b^i}}\int_{\mathbb{R}_+^{d-1}} y^i\big(\eta\mathbb{1}_{\{\|y\|\geq 1\}} + 1 - \exp\big(\theta(b^i y^i + \mathcal{E}(y))\big)\big)\,\nu(dy).
\end{aligned}
$$

□

The following lemma is part of Chapter 6.

Lemma A.2.2. *Let* $f:\mathbb{R}_+\to\mathbb{R}_+$ *and* $f\in C^k$. *Assuming* $(-1)^k f^{(k)}\geq 0$ *and* $\lim_{x\to\infty} f(x)=0$ *it holds that*

$$\lim_{x\to\infty} f^{(k)}(x)=0.$$

Proof. Let $f(x)=-\int_x^\infty f'(y)dy$. For $\epsilon>0$ there exists a x_0 such that for all $x>x_0$: $f(x)\leq\epsilon$. Therefore, it holds for all $x>x_0+1$ that

$$\epsilon\geq\int_{x-1}^{\infty} -f'(y)\,dy > \int_{x-1}^{x} -f'(x)\,dy = -f'(x)$$

since $-f'(y)$ is decreasing in y. Thus $-f'(x)\leq\epsilon$. Assuming that this holds for the k-th derivative and proceeding in the same manner as above for $k+1$ yields the assertion. □

Bibliography

Albrecher, H., Renaud, J.-F., and Zhou, X. (2008) A Lévy insurance risk process with tax. *J. Appl. Probab.* **45**, 363–375.

Amann, H. (1983) *Gewöhnliche Differentialgleichungen.* Walter de Gruyter & Co., Berlin.

Applebaum, D. (2004) *Lévy processes and stochastic calculus.* Cambridge University Press, Cambridge.

Avram, F., Palmowski, Z., and Pistorius, M. (2008a) Exit problem of a two-dimensional risk process from the quadrant: exact and asymptotic results. *Ann. Appl. Probab.* to appear.

Avram, F., Palmowski, Z., and Pistorius, M. (2008b) A two-dimensional ruin problem on the positive quadrant. *Insurance Math. Econom.* **42**, 227–234.

Barndorff-Nielsen, O. E. (1998) Processes of normal inverse Gaussian type. *Finance Stoch.* **2**, 41–68.

Barndorff-Nielsen, O. E. and Lindner, A. (2005) Some aspects of Lévy copulas URL `www.imf.au.dk/publs?id=584`.

Bäuerle, N., Blatter, A., and Müller, A. (2008) Dependence properties and comparison results for Lévy processes. *Math. Methods Oper. Res.* **67**, 161–186.

Bäuerle, N. and Grübel, R. (2005) Multivariate counting processes: copulas and beyond. *Astin Bull.* **35**, 379–408.

Bellman, R. (1957) *Dynamic programming.* Princeton University Press, Princeton, N. J.

Bertsekas, D. P. and Shreve, S. E. (1978) *Stochastic optimal control: The discrete time case.* Academic Press, New York.

Bregman, Y. and Filipovic, D. (2006) Ruin probability estimation in multivariate regularly varying models using Lévy copulas .

Bregman, Y. and Klüppelberg, C. (2005) Ruin estimation in multivariate models with Clayton dependence structure. *Scand. Actuar. J.* 462–480.

Brown, L. and Purves, R. (1973) Measurable selections of extrema. *Ann. Stat.* **1**, 902–912.

Browne, S. (1995) Optimal investment policies for a firm with a random risk process: exponential utility and minimizing the probability of ruin. *Math. Oper. Res.* **20**, 937–958.

Chan, W.-S., Yang, H., and Zhang, L. (2003) Some results on ruin probabilities in a two-dimensional risk model. *Insurance Math. Econom.* **32**, 345–358.

Collamore, J. F. (1996) Hitting probabilities and large deviations. *Ann. Probab.* **24**, 2065–2078.

Collamore, J. F. (2002) Importance sampling techniques for the multidimensional ruin problem for general Markov additive sequences of random vectors. *Ann. Appl. Probab.* **12**, 382–421.

Cont, R. and Tankov, P. (2004) *Financial modelling with jump processes.* Chapman & Hall/CRC, Boca Raton.

Dellacherie, C. and Meyer, P.-A. (1982) *Probabilities and potential. Theory of martingales.* North-Holland Publishing Co., Amsterdam.

Fernández, B., Hernández-Hernández, D., Meda, A., and Saavedra, P. (2008) An optimal investment strategy with maximal risk aversion and its ruin probability. *Math. Methods Oper. Res.* **68**, 159–179.

Fleming, W. H. and Rishel, R. W. (1975) *Deterministic and stochastic optimal control.* Springer-Verlag, Berlin.

Grandell, J. (1991) *Aspects of risk theory.* Springer-Verlag, New York.

Hennessy, D. A. and Lapan, H. E. (2002) The use of Archimedean copulas to model portfolio allocations. *Math. Finance* **12**, 143–154.

Hernández-Lerma, O. and Lasserre, J. B. (1996) *Discrete-time Markov control processes*, volume 30 of *Applications of Mathematics (New York).* Springer-Verlag, New York. Basic optimality criteria.

Højgaard, B. and Taksar, M. (1998) Optimal proportional reinsurance policies for diffusion models. *Scand. Actuar. J.* 166–180.

Højgaard, B. and Taksar, M. (2007) *Diffusion optimization models in insurance and finance.* Early draft.

HULT, H. and LINDSKOG, F. (2006) Heavy-Tailed Insurance Portfolios: Buffer Capital and Ruin Probabilities URL `http://hdl.handle.net/1813/9313`.

ITÔ, K. (1942) On stochastic processes. I. (Infinitely divisible laws of probability). *Jap. J. Math.* **18**, 261–301.

ITÔ, K. (1998) My Sixty Years in Studies of Probability Theory. Acceptance speech of the Kyoto Prize in Basic Sciences.

JACOD, J. and SHIRYAEV, A. N. (2003) *Limit theorems for stochastic processes.* Springer-Verlag, Berlin.

JOE, H. (1997) *Multivariate Models and Dependence Concepts.* Chapman & Hall, London.

KALLENBERG, O. (2002) *Foundations of modern probability.* Springer-Verlag, New York.

KALLSEN, J. and TANKOV, P. (2006) Characterization of dependence of multidimensional Lévy processes using Lévy copulas. *Journal of Multivariate Analysis* **97**, 1551–1572.

KARATZAS, I. and SHREVE, S. E. (1991) *Brownian motion and stochastic calculus.* Springer-Verlag, New York.

KORN, R. (1997) *Optimal portfolios: stochastic models for optimal investment and risk management in continuous time.* Singapore: World Scientific.

KYPRIANOU, A. E. (2006) *Introductory lectures on fluctuations of Lévy processes with applications.* Springer-Verlag, Berlin.

LÉVY, P. (1934) Sur les intégrales dont les éléments sont des variables aléatoires indépendantes. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (2)* **3**, 337–366.

LINDSKOG, F. and MCNEIL, A. J. (2003) Common Poisson shock models: applications to insurance and credit risk modelling. *Astin Bull.* **33**, 209–238.

LOÈVE, M. (1973) Paul Lévy, 1886-1971. *Ann. Probab.* **1**, 1–18.

MADAN, D. B. (2001) Purely discontinuous asset price processes. In *Option pricing, interest rates and risk management.* Cambridge Univ. Press, Cambridge.

MCNEIL, A. J. and NEŠLEHOVÁ, J. (2008) Multivariate Archimedean Copulas, d-monotone Functions and l1 norm Symmetric Distributions. *Annals of Statistics* to appear.

Müller, A. and Stoyan, D. (2002) *Comparison Methods for Stochastic Models and Risks.* John Wiley & Sons Ltd., Chichester.

Nelsen, R. B. (2006) *An introduction to copulas.* Springer Series in Statistics. Springer, New York.

Øksendal, B. and Sulem, A. (2007) *Applied stochastic control of jump diffusions.* Springer, Berlin.

Pfeifer, D. and Nešlehová, J. (2004) Modeling and generating dependent risk processes for IRM and DFA. *Astin Bull.* **34**, 333–360.

Protter, P. (1990) *Stochastic integration and differential equations.* Springer-Verlag, Berlin.

Puterman, M. L. (1994) *Markov decision processes: discrete stochastic dynamic programming.* John Wiley & Sons Inc., New York.

Riesner, M. (2006) *Unit-linked life insurance in Lévy process financial markets.* Dissertation. URL: http://vts.uni-ulm.de/doc.asp?id=5713, Ulm.

Rolski, T., Schmidli, H., Schmidt, V., and Teugels, J. (1999) *Stochastic processes for insurance and finance.* Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester.

Rüschendorf, L. (1980) Inequalities for the expectation of Δ-monotone functions. *Z. Wahrsch. Verw. Gebiete* **54**, 341–349.

Sato, K.-i. (1999) *Lévy processes and infinitely divisible distributions.* Cambridge University Press, Cambridge.

Schäfer, U. (2007) A fixed point theorem based on Miranda. *Fixed Point Theory Appl.* Art. ID 78706, 5.

Schäl, M. (1975) Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **32**, 179–196.

Schäl, M. (2004) On discrete-time dynamic programming in insurance: exponential utility and minimizing the ruin probability. *Scand. Actuar. J.* 189–210.

Schäl, M. (2005) Control of ruin probabilities by discrete-time investments. *Math. Methods Oper. Res.* **62**, 141–158.

Schmidli, H. (2001) Optimal proportional reinsurance policies in a dynamic setting. *Scand. Actuar. J.* 55–68.

Schmidli, H. (2002) On minimizing the ruin probability by investment and reinsurance. *Ann. Appl. Probab.* **12**, 890–907.

SCHMIDLI, H. (2008) *Stochastic control in insurance.* Springer-Verlag, London.

TANKOV, P. (2003) Dependence structure for spectrally positive multidimensional Lévy processes, unpublished.

TANKOV, P. (2004) *Lévy Processes in Finance: Inverse Problems and Dependence Modelling.* PhD thesis. École Polytechnique.

YONG, J. and ZHOU, X. Y. (1999) *Stochastic controls*, volume 43 of *Applications of Mathematics (New York).* Springer-Verlag, New York. Hamiltonian systems and HJB equations.