Karlsruher Institut für Technologie

Bayesian Optimal Investment and Reinsurance to Maximize Exponential Utility of Terminal Wealth

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

von der KIT-Fakultät für Mathematik des
Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von

Gregor Leimcke, M. Sc.
aus
Karl-Marx-Stadt jetzt Chemnitz

Tag der mündlichen Prüfung: 22. Januar 2020

1. Referentin: Prof. Dr. Nicole Bäuerle
2. Referent: Prof. Dr. Thorsten Schmidt

献给云琦和露伊莎

Acknowledgements

At this point I would like to take the opportunity to express my great gratitude to several people. First of all, my sincere appreciation belongs to Prof. Dr. Nicole Bäuerle for providing me the opportunity to write this thesis at the Institute for Stochastics at Karlsruhe Institute of Technology (KIT). I am very thankful for her excellent guidance and advices whenever I needed it. The numerous fruitful discussions during the three years of my doctoral research push this thesis forward.

Special thanks go to Prof. Dr. Thorsten Schmidt not only for acting as second examiner, but also for allowing me to learn from his excellent expertise during my mathematics studies at the TU Chemnitz. The cooperation was very inspiring and has been of great benefit to me.

Furthermore, I want to thank all members of the Institute for Stochastics. The working environment here was very positive and motivating all the time, which was far beyond anything I would have expected. I will always remember the last three years in the most positive way.

Finally, I want to thank my family. Thanks go to my brother and sisters Jonas, Judith and Sarah and especially to my parents Monika and Wolf-Dietrich, for always being there and giving me their caring encouragement, not only during my PhD studies but also during my entire life. In particular, I want to express my special gratitude to my wife, Yunqi, for her continued and loving support, patience and understanding as well as source of energy and optimistic enthusiasm for life.

Karlsruhe, December 2019 *Gregor Leimcke*

Abstract

We herein discuss the surplus process of an insurance company with various lines of business. The claim arrivals of the lines of business are modelled using multivariate point process with interdependencies between the marginal point processes, which depend only on the choice of thinning probabilities. The insurer's aim is to maximize the expected exponential utility of terminal wealth by choosing an investment-reinsurance strategy, in which the insurer can continuously purchase proportional reinsurance and invest its surplus in a Black-Scholes financial market consisting of a risk-free asset and a risky asset. We separately investigate the resulting stochastic control problem under unknown thinning probabilities, unknown claim arrival intensities and unknown claim size distribution for a univariate case. We overcome the issue of uncertainty for these three partial information control problems using Bayesian approaches that result in reduced control problems, for which we characterize the value functions and optimal strategies with the help of the generalized Hamilton-Jacobi-Bellman equation, in which derivatives are replaced by Clarke's generalized gradients. As a result, we could verify that the proposed investment-reinsurance strategy is indeed optimal. Moreover, we analysed the influence of unobservable parameters on optimal reinsurance strategies by deriving comparative results with the case of complete information, which shows a more risk-averse behaviour under more uncertainty. Finally, we provide numerical examples to illustrate the comparison results.

Contents

List of Figures

List of Tables

Abbreviations

BSDE	backward stochastic differential equation
càdlàg	continu à droite avec limité à gauche
DPP	dynamical programming principle
DSPP	doubly stochastic Poisson process
FTCL	Fundamental Theorem of Calculus for the Lebesgue integral
FV	finite variation
HJB	Hamilton-Jacobi-Bellman
iid	independent and identically distributed
LoB	line of business
MPP	marked point process
mPP	mixed Poisson process
ODE	ordinary differential equation
PDE	partial differential equation
PRM	Poisson random measure
SDE	stochastic differential equation
SPP	simple point process

Basic Notations

Integers and real numbers

$\mathbb{N}_0$	$\mathbb{N} \cup \{0\}$, set of non-negative integers
$\overline{\mathbb{N}}_0$	$\mathbb{N}_0 \cup \{+\infty\}$, set of non-negative integers including infinity
$\mathbb{R}^+$	$[0, \infty)$, set of non-negative real numbers
$\overline{\mathbb{R}}^+$	$[0, \infty) \cup \{+\infty\}$, compactification of non-negative real numbers

Special symbols

$\|\cdot\|$	ℓ_1-norm
$\|\cdot\|_2$	Euclidean norm
$\delta_{(x)}$	Dirac measure located at x, i.e. $\delta_{(x)}(A) := \mathbb{1}_A(x)$
$\mathbb{1}_A$	Indicator function of a set A
$\operatorname{co}\{A\}$	Convex hull of $A \subset \mathbb{R}^n$
$\mathcal{B}(E)$	Borel σ-algebra on E
$\mathcal{B}^+$	Borel σ-algebra on $\mathbb{R}^+$
$\mathcal{P}(\Omega)$	Set of all subsets of a non-empty set Ω
$\mathcal{P}(\mathfrak{F})$	Predictable σ-algebra of the filtration $\mathfrak{F}$
$\sigma(A)$	σ-algebra generated by a set A
$A \vee B$	$\sigma(A \cup B)$, σ-algebra generated by A and B
e_k	k-th unit vector in $\mathbb{R}^m$, $k \in \{1, \ldots, m\}$
$\partial^C f(x)$	Clarke's generalized gradient, cf. Def. 2.6
$f^\circ(x; v)$	Generalized directional derivative of f at x in direction v, cf. Def. 2.4
$(X)^c$	Continuous part of the process X, cf. Def. 2.49
$[X]$	Quadratic variation process of X, cf. Def. 2.50
$[X, Y]$	Quadratic covariation process of X and Y, cf. Def. 2.50
$\mathcal{E}(X)$	Doléans-Dade exponential of X, cf. Def. 2.61

Special sets

E^d	$(0,\infty)^d \times \mathcal{P}(\mathbb{D})$, mark space of the MPP $\Psi = (T_n, (Y_n, Z_n))_{n\in\mathbb{N}}$
Δ_k	$(k-1)$-dimensional probability simplex, $k \geq 2$
$\mathring{\Delta}_k$	Interior of the probability simplex Δ_k, $k \geq 2$
$\mathcal{A}(\mathbb{P}, \mathfrak{F})$	Set of all $\mathfrak{F}$-adapted finite variation processes w.r.t. $\mathbb{P}$, cf. Def. 2.47
$\mathcal{M}(\mathbb{P}, \mathfrak{F})$	Set of all càdlàg $(\mathbb{P}, \mathfrak{F})$-martingales starting at zero, cf. Def. 2.32
$\mathcal{M}_{loc}(\mathbb{P}, \mathfrak{F})$	Set of all càdlàg local $(\mathbb{P}, \mathfrak{F})$-martingales starting at zero

Function spaces

$AC([a,b])$	Set of all absolutely continuous functions on $[a,b]$
$BV([a,b])$	Set of all function of bounded variation on $[a,b]$
$Lip(I)$	Set of all Lipschitz function on $I \subseteq \mathbb{R}^n$
$Lip_{loc}(I)$	Set of all locally Lipschitz function on $I \subseteq \mathbb{R}^n$

Chapter 1

Introduction

1.1 Motivation and literature overview

Many challenges are currently facing the insurance industry. On the one hand, the number and volume of insurance losses are growing as a result of the weather fluctuations due to climate change.[1] On the other hand, the current structural low interest rate environment and higher volatility on financial markets are making it more difficult to achieve profitable investments. Further challenges arise from the transparency of insurance contracts provided by comparison portals, as well as the comprehensive inter-connectedness resulting from the digitalization trend and the data generated from it, which can be used to identify and record risks. For many years, questions inherited from the first two challenges regarding effective strategies for reducing the insurance risk and optimal capital investment have been attracting the attention of researches in actuarial mathematics. In fact, a classical task in risk theory is to deal with optimal risk control and optimal asset allocation for an insurance company.

Generally, the risk of an insurer results from the compensation of insurance claims in exchange for regular premiums, in which an insurance claim is a request to an insurance company for a payment related to the terms of an insurance policy.[2] This risk can be reduced by ceding claims to a reinsurance company in return for relinquishing part of the premium income to the reinsurer. More precisely, the reinsurer covers part of the costs of claims against the insurer. Notice that we refer to the cost of a claim as the claim size, magnitude, loss or amount of damage.

The surplus of the insurance company arises from the premiums left to the insurer after transferring the risk to the reinsurer and from the payments to be made by the insurer. This surplus is deposited in a financial market, which leads to an optimal investment-reinsurance problem in continuous time under the assumption that the insurer can continuously purchase a reinsurance contract and invest in a financial market. These problems have been previously intensively studied in the literature using various optimization criteria, in which maximizing the utility and minimizing the ruin probability are two frequently used optimization criteria.

Schmidli [110], Promislow and Young [103] and Cao and Wan [34] employed a Black-Scholes-type financial market and proportional reinsurance (as will be done in this work) for optimal control problems. While the first two articles provide optimal investment-reinsurance strategies (a closed-form and analytical expression for the reinsurance strategy, respectively) under the criteria of minimizing the ruin probability, the third article offers an explicit expression for the problem of maximizing the exponential utility of terminal wealth. Other articles (with other settings) that are worth mentioning are

[1] See Faust and Rauch [58].
[2] See Cambridge Dict. [33].

those by Zhang and Siu [123], in which an optimization problem was formulated as a stochastic differential game between the insurer and the market; Schmidli [111], who studied several optimization problems in insurance under different frameworks; and Bai and Guo [11], who showed in some special cases the equivalence of optimal strategies for maximizing the expected exponential utility of terminal wealth and minimizing the probability of ruin.

The cited literature deals with the jump part of the surplus process, which describes the *net claim process* (premium income minus claim compensation), in two different ways. The first way is to use the well-known Cramér-Lundberg model from classical risk theory (see Bühlmann [15]) to describe the net claim process, as was done by Schmidli [110]. The Cramér-Lundberg model was first introduced by Filip Lundberg in his work [88] and was also named after Harald Cramér because of his basic findings with that model (see [43]). The second way is to use the diffusion approximation considered by Iglehart [70] for the jump term in the Cramér-Lundberg model, as outlined by Grandell [66, Sec. 1.2]. Hence, with such an approximation approach, the optimization problem was studied by Cao and Wan [34] and Promislow and Young [103]. Both approaches were also examined by Zhang and Siu [123] and Schmidli [111].

In all of the articles mentioned so far, the assumption of full information is used as a common feature, which means that the insurer has complete knowledge of the model parameters. However, in reality, insurance companies operate in a setting with partial information; that is, with regard to the net claim process, only the claim arrival times and magnitudes are directly observable, but the claim intensity, which is required by all net claim models, is not observable by the insurer, as pointed out by Grandell [66, Ch. 2]. In the context of financial markets, partial information means that the terms of drift and volatility are unknown, even though the term of volatility is typically assumed to be known as it can be estimated very well, whereas the appreciation rate is notoriously difficult to estimate (see Rogers [107, Sec. 4.2]).

As mentioned, insurers make decisions solely on the basis of the information at their disposal in practice. Therefore, we herein investigate the optimal investment and reinsurance problem in a partial information framework. We first emphasize that partial information control problems are different from partial observation control problems in that the controls of the latter problems are based on noisy observations of the state process. Di Nunno and Øksendal [52] were the first to study a partial information optimal portfolio problem in the sense that the dealer has access to only some information represented by a filtration, which is generally smaller than the one generated by the financial market. This problem was also investigated by Liang et al. [87] in the presence of both investment and reinsurance, in what partial information refers to the financial market. In this work, we assume that full information is available on the financial market and focus on the insurance risk with an unobservable claim intensity.

On the basis of the suggestion of Albrecher and Asmussen [4, p. 165], Liang and Bayraktar [85] considered the optimal investment and reinsurance problem for maximizing exponential utility under the assumption that the claim intensity and loss distribution depend on the state of a non-observable Markov chain (hidden Markov chain), which describes different environment states, whereby the net claim process is modelled as a compound Poisson process and the fully observable financial market is modelled as a Black-Scholes financial market with one risky and one risk-free asset. In this thesis, we use the same financial market model; moreover, our assumption on the claim intensity can be considered as a special case of one state of the above-mentioned Markovian regime-switching model; namely, we model the intensity as an unobservable random variable, which places us in a Bayesian setting. However, literature with a setting of partial information focuses only on one line of business (LoB) to gain an optimal reinsurance strategy. However, in reality, there is often a dependency between the different risk processes of an insurance company. This results from the fact that customers of a typical insurance company have insurance policies of different types, such as building, private-liability or health insurance contracts.

A simplified example of a possible dependency between several types of risk is that of a storm event accompanied by heavy rainfall, wherein flying roof tiles cause damage to third parties and flooding leads to damage in buildings. In addition to this dependency between private-liability insurance and building insurance, there may even be a dependency on motor-liability insurance and health insurance if a car accident occurs as a result of adverse traffic circumstances due to that heavy rainfall. Therefore, in order to appropriately model the insurance risks of an insurer, we need to capture the dependency structure using a multivariate model.

Thinning is a commonly used approach to impose dependency between several types of insurance risks, which is also the case in this thesis. The idea of this approach is that the occurrence of claims depends on a certain process that generates events that cause damage to LoB i with probability p_i and to LoB j with probability p_j, where all claims occur simultaneously at the trigger arrival time. Therefore, these models are referred to as *common shock risk models*. An example of a shock event is the storm event described above. Typically, the corresponding claim sizes are determined independently of the appearance times.

This typical assumption is also considered to be fulfilled in this work. The thinning approach traces back to Yuen and Wang [122]. Anastasiadis and Chukova [8] provided a literature overview of multivariate insurance models from 1971 to 2008.

Another multivariate model that avoids referencing an external mechanism was given by Bäuerle and Grübel [28], who proposed a multivariate continuous Markov chain of pure birth type with inter-dependency arising from the dependency of the birth rates on the number of claims in other component processes. Scherer and Selch [108] constructed the dependency of the marginal processes of a multiple claim arrival process by introducing a Lévy subordinator serving as a joint stochastic clock, which lead to a multivariate Cox process in the sense that marginal processes are univariate Cox processes.

Another frequently discussed dependency concept based on copula. Cont and Tankov [41, Ch. 5] described the dependency structure of multi-dimensional Lévy processes in terms of Lévy copula. However, as pointed out by Bäuerle and Grübel [27, p. 5], the dependency modelling for Lévy processes is reduced to the choice of thinning properties as a consequence of their defining properties. On the basis of this limitation, Bäuerle and Grübel [27] extended Lévy models by incorporating random shifts in time such that the timings of claims caused by a single trigger event are shifted according to some distribution, where some of these claims are thinned out and do not occur. Section 5.9 discusses why such shifts cannot be incorporated in this work. For further multivariate claim count models, please refer to the literature cited by Scherer and Selch [108, Sec. 1.3].

In connection with optimal reinsurance problems, a Lévy approach was discussed by Bäuerle and Blatter [25]. They showed that constant investment and reinsurance (proportional reinsurance as well as a mixture of proportional and excess-of-loss reinsurance) is the optimal strategy for maximizing the exponential utility of terminal wealth.

In addition to the Lévy model, optimization problems with common shock models have been investigated by Centeno [37], who studied optimal excess-of-loss retention limits for a bivariate compound Poisson risk model in a static setting. The corresponding dynamic model was used by Bai et al. [10] to derive optimal excess-of-loss reinsurance policies (which turned out to be constant) under the criterion of minimizing the ruin probability by making use of a diffusion approximation. For the same model, Liang and Yuen [86] derived a closed-form expression for the optimal proportional reinsurance strategy of the exponential utility maximizing problem both with and without diffusion approximation using the variance premium principle. Bi and Chen [16] also investigated the same problem with the expected value premium principle in the presence of a Black-Scholes financial market. In the case of an insurance company with more than two LoBs, Yuen et al. [121] and Wei et al. [117] sought optimal proportional reinsurance to maximize the exponential utility of terminal wealth and the adjustment coefficient, respectively, in which the strategies are only stated for two LoBs. However, all optimization problems with multivariate insurance models are considered under full information.

In this work, we will describe the dependency structure between different LoBs using the thinning approach while dealing with unobservable thinning probabilities. To our knowledge, this is the first time an optimal reinsurance and investment problem under partial information using a multivariate claim arrival model with possibly dependent marginal processes is studied. In order to solve this optimal control problem, the dynamic programming Hamilton-Jacobi-Bellman (HJB) approach will be applied, which is the most widely used method for stochastic control problems. The HJB equation is a classical tool for deriving optimal strategies for control problems. This equation can be obtained by applying the dynamic programming principle, which was pioneered by Richard Bellman, after whom the HJB equation is named, in the 1950s (see [13, 14]). In classical physics, the diffusion case of this equation can be viewed as an extended Hamilton-Jacobi equation, which was named after William Rowan Hamilton and Carl Gustav Jacob Jacobi.[3]

The HJB equation is a deterministic (integro-) partial differential equation whose solution is the value function of the corresponding stochastic control problem under certain conditions. However, in general, the existence of a solution to this equation is not guaranteed because of smoothness requirements of the solution. In our setting, we will have to deal with the strong assumption of differentiability of the value function, which cannot be guaranteed. Over the past decades, a rich theory has been developed to overcome this difficulty. In the early 1980s, Pierre-Louis Lions and Michael Crandall introduced the currently popular concept known as the *viscosity solution* for non-linear first-order partial differential equations (see [45]), which claims that the value function is the unique viscosity solution to the HJB equation under mild conditions (continuity and, in more general frameworks, even discontinuity).[4] The basic idea behind this concept is to estimate the value function from above and below using smooth test functions. Fleming and Soner [61] applied the viscosity approach to optimise the control of Markov processes.

Another approach is to generalize the HJB equation by including Clarke's generalized gradient, which is a weaker notation of differentiability. Clarke [39] and Davis [47] came up with this idea which the strong assumption of differentiability of the value function can be weakened to local Lipschitz continuity. The generalization concept of the HJB equation has been applied in an article by Liang and Bayraktar [85] and is also used in this work.

As indicated above, we consider the surplus process of an insurance company with various LoBs in which claim arrivals are modelled using a common shock model under incomplete information (i.e. the claim intensity and the thinning probabilities are unknown to the insurer). Our aim is to solve the optimization problem facing insurance

[3]See Section 5.1.2 in Blatter [17] and Section 9.1 in Popp [102].

[4]For a historical survey for the development of the viscosity solution, we refer the reader to Yong and Zhou [120, Sec. 4.1] and Fleming and Soner [61, Sec. II.17].

companies by trying to find investment-reinsurance strategies that maximizes the expected exponential utility of terminal wealth. Using a Bayesian approach, we overcome the issue of uncertainty and obtain a reduced control problem, which is investigated with the help of the dynamic programming principle and a generalized HJB equation. Before entering the world of our insurance model, let us introduce the outline of the thesis and highlight the main results.

1.2 Main results and outline

In this thesis, we deal with a multi-dimensional insurance risk model. The main aim here is to study the impact of partial information regarding the inter-dependencies between marginal risk processes on the optimal investment and proportional reinsurance under the criterion of maximizing the expected exponential utility of terminal wealth. Preparing for the introduction of the risk model and the solution technique of the corresponding control problem, Chapter 2 is dedicated to the fundamentals, starting with the concept of Clarke's generalized gradient in Section 2.1. After providing a brief overview of the basic definitions and properties of stochastic processes in Section 2.2, we recall some important tools of stochastic analysis in Section 2.3. This is then followed by Section 2.4, which is devoted to the (marked) point processes used for modelling the net claim process. Section 2.4.1 includes basic definitions and notations, and Section 2.4.2 deals with the concept of the intensities of point processes, our main object of interest for the characterization of point processes. This concept is generalized in Section 2.4.4 to marked point processes. In Sections 2.4.3 and 2.4.5, we proceed with the study of the innovation method for filtering with point and marked point process observations, respectively, in a simplified setting, which is sufficient for the following chapters.

Following this introductory chapter, we introduce a control problem under partial information in Chapter 3. For this purpose, we specify the multivariate claim arrival model in Section 3.1 as a common shock model, in which the shock generating intensity (background intensity) and the thinning probabilities describing the affected LoBs are unobservable, which is considered by modelling these parameters as random variables. Moreover, we also model the claim size distribution as unknown. In contrast to the cited literature considering optimal reinsurance problems in a multivariate setting, the reinsurance strategies of different LoBs are supposed to be equal in our work, which is equivalent to the assumption of one reinsurance policy for an entire insurance company. In Section 4.10, we discuss the consequences of various reinsurance contracts. Furthermore, as the focus is on the optimal reinsurance strategy, we use a quite simple financial market model with one risk-free and one risky asset modelled as geometric Brownian motion, which is presented in Section 3.2. Subsequently, we introduce investment strategies (see Section 3.3), proportional reinsurance strategies (see Section 4.3) and a reinsurance premium principle (see Section 3.5). This results in the definition of the surplus process in Section 3.6, the basic object of interest in the stochastic control problem under partial information stated in Section 3.7. The next three chapters address the incomplete information problem under different assumptions for unknown parameters.

For the sake of simplicity, we restrict ourselves to the study of the case of an observable background intensity and claim size distribution in Chapter 4. This case is specified in Section 4.1, where we suppose that the vector of thinning probabilities takes values in a finite set. In order to overcome the difficulty of partial information in thinning probabilities, we determine an observable estimator of these probabilities by means of filtering theory for marked point process observations in Section 4.2. With the help of the derived filter, we formulate a reduced control model and problem in Section 4.3, for which we derive the generalized HJB equation in Section 4.4 by replacing the partial derivative with respect to time by the corresponding Clarke's generalized subdifferential, which is introduced in Section 2.1. We then receive candidates for an optimal investment and reinsurance strategy in Sections 4.5 and 4.6. It then turns out that the candidate for the optimal investment strategy is deterministic, in particular independent of the reinsurance strategy, whereas the unique candidate for the optimal reinsurance strategy depends not only on the safety loading parameter of reinsurance and time, but also on the background intensity, claim size distribution and the filter process. In Section 4.7, we continue with the verification procedure by showing that a solution to the HJB equation does indeed coincide with the value function and that the derived candidates for the optimal strategies are indeed optimal. Afterwards, we prove the existence of a solution to the HJB equation. Section 4.8 includes a comparative result of the optimal reinsurance strategy under full information with the one under partial information, in which we suppose identical claim size distributions for every LoB. We find that the optimal reinsurance strategy in the multivariate risk model with known thinning probabilities is always greater than or equal to the one in the risk model with unknown thinning probabilities. The comparison result is illustrated in Section 4.9. We close the chapter with a discussion on the generalizations of the setting used in this chapter, particularly concerning the thinning probabilities.

In Chapter 5 we investigate the partial information problem under the assumptions of observable claim size distribution, unobservable background intensity taking values in a finite set and Dirichlet distributed thinning probabilities (see Section 5.1). Using a filter as an estimator for the background intensity and the conjugated property of the Dirichlet distribution, we proceed as in the previous chapter by stating the reduced control problem (see Section 5.3) and the corresponding generalized HJB equation, in which we need to replace partial derivatives with respect to time and the components of the filter for the background intensity by the corresponding Clarke's generalized gradient. The HJB equation yields the same optimal investment strategy as in Chapter 4, and the optimal reinsurance strategy has a similar structure. The verification step runs as before (see Section 5.6) and shows the optimality of the proposed investment-reinsurance strategy. In Section 5.7, we provide a comparison result similar to the one in the previous chapter, again under the assumption of identical claim size distributions for all insurance classes, which is visualized in Section 5.8. Finally, some generalizations of the setting of this chapter and resulting difficulties regarding the used solution technique are discussed.

Chapter 6 is devoted to the case of an unobservable claim size distribution for the introduced control problem under partial information, in which the framework is quite simple as a result of supposing that there are only a finite number of potential claim size distributions. Section 6.9 establishes the difficulties faced with more general settings. In order to simplify the optimality analysis, we consider working on the insurance model with one LoB and observable background intensity (see Section 6.1), with the result being a similarly optimal investment-reinsurance strategy and correspondingly analogous verification. Moreover, we develop a similar comparative result as before, which yields a very small range of possible optimal reinsurance strategies for some scenarios in the numerical study in Section 6.8.

Appendix A includes auxiliary results for the verification procedure in Chapters 4, 5 and 6, and some useful inequalities are covered in Appendix B.

Chapter 2

Fundamentals

Before getting into the detail of our insurance model and the optimization problem, let us recall some foundations of stochastic processes and filter results with (marked) point process observations to make this work self-contained.

2.1 Clarke's generalized gradient

We start this chapter by briefly introducing a concept of nonsmooth analysis, namely Clarke's generalized gradient, which was introduced by Clarke [39, Ch. 2]. We consider only the case of functions defined on the Euclidean space $(\mathbb{R}^n, \|\cdot\|_2)$ equipped with the Euclidean norm $\|\cdot\|_2$ instead of a general Banach space.

Notation. Let $r > 0$ some scalar and $x \in \mathbb{R}^n$ some vector. We denote the open ball of radius r about x by $B_r(x) := \{y \in \mathbb{R}^n : \|x - y\| < r\}$.

Definition 2.1 ([39], p. 25; [115], Def. 4.6.9; Lipschitz function)**.** Let $I \subset \mathbb{R}^n$ be a subset of $\mathbb{R}^n$ and $f : I \to \mathbb{R}$ be a function defined on I.

(i) We say f is *Lipschitz* on I (of rank K) or *satisfies a Lipschitz condition* on I (of rank K) if there exists $0 < K < \infty$ such that

$$|f(x_1) - f(x_2)| \leq K\|x_1 - x_2\|_2$$

for all $x_1, x_2 \in I$.

(ii) The function f is said to be *Lipschitz* (of rank K) near $x \in I$ if there exists $\varepsilon = \varepsilon(x) > 0$ such that f is Lipschitz on $I \cap B_\varepsilon(x)$. If f is Lipschitz (of rank K) near x for all $x \in I$, then we say f is *locally Lipschitz* on I (of rank K).

We write $Lip(I)$ for the set of all Lipschitz function on I and $Lip_{loc}(I)$ for the collection of all locally Lipschitz function on I.

A useful result is that every convex function defined on an open convex set is locally Lipschitz.

Theorem 2.2 ([106], Thm. A)**.** *Let f be convex on an open convex set $I \subseteq \mathbb{R}^n$. Then $f \in Lip_{loc}(I)$ and, consequently, $f \in Lip(C)$ for all compact sets $C \subset I$.*

Let us mention further result of Lipschitz functions.

Theorem 2.3 ([40], Thm. 3.4.1, Remark 3.4.2; Rademacher's Theorem)**.** *Let $I \subseteq \mathbb{R}^n$ be a subset and $f \in Lip_{loc}(I)$. Then f is differentiable almost everywhere on I in the sense of the Lebesgue measure.*

To define Clarke's generalized gradient, we must first introduce the generalized directional derivation.

Definition 2.4 ([39], p. 25)**.** Let $x \in \mathbb{R}^n$ be a given point and $v \in \mathbb{R}^n$ be some other vector. Moreover, let f be Lipschitz near x. Then the *generalized directional derivative* of f at x in the direction v, denoted by $f^\circ(x;v)$, is defined by

$$f^\circ(x;v) = \limsup_{\substack{y\to x\\ h\downarrow 0}} \frac{f(y+h\,v)-f(y)}{h}.$$

Justification of the definition. Due to the locally Lipschitz property of f, the difference quotient is bounded above by $K\|v\|_2$ for some $0 < K < \infty$ and for y sufficient near x as well as h sufficient near 0. Therefore, $f^\circ(x;v)$ is well-defined since the upper limit is taken from the bounded difference quotient and no limit is presupposed. □

Beside the finite property, the generalized directional derivative admits the following elementary properties.

Proposition 2.5 ([39], Prop. 2.1.1)**.** *Let f be Lipschitz near $x \in \mathbb{R}^n$. Then $v \mapsto f^\circ(x;v)$ is positively homogeneous and subadditive.*

These properties justifies the existence of the next defined generalized gradient. On account of the Hahn-Banach Theorem, we know that any positively homogeneous and subadditive functional on $\mathbb{R}^n$ majorizes some linear functional on $\mathbb{R}^n$. Consequently, the proposition above implies the existence of at least one linear functional $\xi : \mathbb{R}^n \to \mathbb{R}$ with $f^\circ(x;v) \ge \xi(v)$. It therefore follows by $\xi(v) \le K\|v\|_2$ for some $0 < K < \infty$ that ξ belongs to the dual space of $\mathbb{R}^n$ of continuous linear functionals on $\mathbb{R}^n$. Clarke's generalized gradient will be defined as a subset of the continuous linear functionals and thus is non-empty by the explanation above. Since a continuously linear functional ξ on $\mathbb{R}^n$ can be identified by $\xi \in \mathbb{R}^n$, the generalized gradient is a subset of $\mathbb{R}^n$ in our setting.

Definition 2.6 ([39], p. 27)**.** Let f be Lipschitz near $x \in \mathbb{R}^n$. Then *Clarke's generalized gradient* (*generalized gradient* for short) of f at x, denoted by $\partial^C f(x)$, is given by

$$\partial^C f(x) := \left\{\xi \in \mathbb{R}^n : f^\circ(x;v) \ge \xi^\top v \;\forall\, v \in \mathbb{R}^n\right\}.$$

In the univariate case we call $\partial^C f(x)$ *Clarke's generalized subdifferential* (*generalized subdifferential* for short) of f at x.

We continue with properties of the generalized gradient.

Proposition 2.7 ([39], Prop. 2.1.2)**.** *Let f be Lipschitz near $x \in \mathbb{R}^n$. Then $\partial^C f(x)$ is a convex and compact subset of $\mathbb{R}^n$.*

Proposition 2.8 ([39], Prop. 2.2.4)**.** *If f is strictly differentiable at $x \in \mathbb{R}^n$ such that some differential operator D is defined, then f is Lipschitz near x and $\partial^C f(x) = \{Df(x)\}$. Conversely, if f is Lipschitz near x and $\partial^C f(x)$ reduces to a singleton $\{\xi\}$, then f is strictly differentiable at x and $Df(x) = \xi$.*

The next characterization of Clarke's generalized gradient will be needed to show existence of a solution of the generalized HJB equation (cf. Theorems 4.33, 5.26 and 6.17) since it allows us to reduce the case of non-differentiability to the case of differentiability.

Theorem 2.9 ([39], Thm. 2.5.1)**.** *Let f be Lipschitz near $x \in \mathbb{R}^n$, Ω_f the set of points at which the function f is not differentiable and S an arbitrary set of Lebesgue-measure 0 in $\mathbb{R}^n$. Then*

$$\partial^C f(x) = \operatorname{co}\Big\{ \lim_{n\to\infty} \nabla f(x_n) : x_n \to x, x_n \notin S, x_n \notin \Omega_f \Big\},$$

where $\operatorname{co}\{A\}$ denotes the convex hull of $A \subset \mathbb{R}^n$.

2.2 Stochastic processes

We start this chapter with briefly recalling some general definitions and results about stochastic processes cited from Karatzas and Shreve [73, Ch. 1], Protter [104, Ch. 1, 2], Capasso and Bakstein [35, Sec. 2.1], Elstrodt [56, Sec. II.6], Bain and Crisan [12, Sec. A.5], Klebaner [75, Ch. 8] and Chung and Williams [38, Ch. 2], which will be of importance in the following proceedings. Throughout this chapter all stochastic quantities are defined on a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Recall that a *stochastic process* X on $(\Omega, \mathcal{F}, \mathbb{P})$ is a family of $\mathbb{R}^d$-valued random variable $(X_t)_{t\geq 0}$ for $d \in \mathbb{N}$. For the sake of convenience, we will use the shorter term *process* instead of stochastic process. A process can be seen as a function $X : \mathbb{R}^+ \times \Omega \to \mathbb{R}^d$ where $X(t, \cdot) = X_t$ is an $\mathcal{F}$-measurable random variable for all $t \geq 0$. For a fixed $\omega \in \Omega$, the mapping $t \mapsto X_t(\omega)$ from $\mathbb{R}^+$ into $\mathbb{R}$ is called a *sample path* or *trajectory* of X.

For two processes X and Y, the notation $X > Y$ means $X_t(\omega) > Y_t(\omega)$ for all $t \geq 0$ and all $\omega \in \Omega$. In particular, $X \geq 0$ stands for $X_t(\omega) \geq 0$ for all $\omega \in \Omega$ and $t \geq 0$. Similarly, we use the notations $X \geq Y$, $X < Y$, $X \leq Y$ and $X = Y$. Furthermore, we say X and Y are the same if and only if $X = Y$. As we know null sets are normally overlooked in the present of probability measures. Accordingly, we introduce in the following alternative concepts of "equality".

Definition 2.10 ([73], Def. 1.1.2, Def. 1.1.3)**.** Let X and Y be two processes. Then Y is called a *modification* or *version* of X if $\mathbb{P}(X_t = Y_t) = 1$ for all $t \geq 0$. If $\mathbb{P}(X_t = Y_t,\ t \geq 0) = 1$, then X and Y are said to be *indistinguishable*.

Remark 2.11 ([104], p. 4)**.** The sample paths of indistinguishable processes differ only on a $\mathbb{P}$-null set, which does not hold for modifications in general since the uncountable union of null sets can have any probability between 0 and 1, and it can even be non-measurable.

Convention. We say that an equation with functions of processes on both sides (e.g. evolution equations for processes) *holds up to indistinguishability* if the processes described by the both sides of the equality are indistinguishable.

Next, we move on to some regularity properties of sample paths, which are defined for almost all ω since indistinguishable processes are regarded as equal.

Definition 2.12. A process $X = (X_t)_{t\geq 0}$ is called *continuous* if $\lim_{s\to t} X_s(\omega) = X_t(\omega)$ for all $t \geq 0$ and $\mathbb{P}$-almost all $\omega \in \Omega$. Moreover, X is said to be *right-continuous* (*left-continuous*) if $\lim_{s\downarrow t} X_s(\omega) = X_t(\omega)$ ($\lim_{s\uparrow t} X_s(\omega) = X_t(\omega)$) for all $t \geq 0$ ($t > 0$) and $\mathbb{P}$-almost all $\omega \in \Omega$. Furthermore, X is called *càdlàg* if it is right-continuous and the left-hand limit $\lim_{s\uparrow t} X_s(\omega)$ exists for all $t > 0$ and $\mathbb{P}$-almost all $\omega \in \Omega$. If X is càdlàg, then the process $X_- = (X_{t-})_{t\geq 0}$ defined by $X_{0-} := 0$, $X_{t-} := \lim_{s\uparrow t} X_s$ for all $t > 0$ is said to be the *left-hand limit process* of X, and the process $\Delta X = (\Delta X_t)_{t\geq 0}$ defined by $\Delta X_0 := 0$, $\Delta X_t := X_t - X_{t-}$ for all $t > 0$ is called the *jump process* of X.

Lemma 2.13 ([12], Lem. A.14). *Let $X = (X_t)_{t\geq 0}$ be a càdlàg process. Then $\{t \geq 0 : \mathbb{P}(X_{t-} \neq X_t) > 0\}$ contains at most countably many points.*

Proposition 2.14 ([104], Thm. I.2). *Let X and Y be two processes where Y is a modification of X. If X and Y have right-continuous sample paths $\mathbb{P}$-almost surely, then X and Y are indistinguishable.*

Definition 2.15 ([73], Def. 1.1.6; Measurability). A process $X = (X_t)_{t\geq 0}$ is called *measurable w.r.t.* $\mathcal{F}$ if the mapping $(t, \omega) \mapsto X_t(\omega) : \big(\mathbb{R}^+ \times \Omega, \mathcal{B}^+ \otimes \mathcal{F}\big) \to \big(\mathbb{R}, \mathcal{B}(\mathbb{R})\big)$ is measurable, i.e. $\{(t, \omega) \in \mathbb{R}_+ \times \Omega : X_t(\omega) \in B, B \in \mathcal{B}(\mathbb{R})\} \in \mathcal{B}_+ \otimes \mathcal{F}$.

Proposition 2.16. *If the process $X = (X_t)_{t\geq 0}$ is either left- or right-continuous, then X is measurable.*

Proposition 2.17. *If $X = (X_t)_{t\geq 0}$ is a measurable process, then the sample path $X_\cdot(\omega) : \mathbb{R}^+ \to \mathbb{R}$ is $\mathcal{B}^+$-measurable for all $\omega \in \Omega$.*

Another important notion in the context of stochastic processes is the filtration. In insurance mathematics, filtrations are used to model the available information for the insurer. Especially for the models with partial information, the measurability of processes w.r.t. different filtration processes is an important aspect.

Definition 2.18 ([73], p. 3 ff.). A family of σ-algebras $\mathfrak{F} = (\mathcal{F}_t)_{t\geq 0}$ with $\mathcal{F}_t \subset \mathcal{F}$ for all $t \geq 0$ is called a *filtration* if $\mathcal{F}_s \subset \mathcal{F}_t$ for all $0 \leq s \leq t$. We set $\mathcal{F}_\infty := \sigma\big(\bigcup_{t\geq 0} \mathcal{F}_t\big)$. For a filtration $\mathfrak{F} = (F_t)_{t\geq 0}$, we define by $\mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$ the *σ-algebra of events immediately after* $t \geq 0$ and by $\mathcal{F}_{t-} := \sigma\big(\bigcup_{s<t} \mathcal{F}_s\big)$ the *σ-algebra of events strictly prior to* $t \geq 0$, where $\mathcal{F}_{0-} := 0$. We say the filtration $\mathfrak{F} = (\mathcal{F}_t)_{t\geq 0}$ is *right-(left-)continuous* if $\mathcal{F}_t = \mathcal{F}_{t+}$ (resp. $\mathcal{F}_t = \mathcal{F}_{t-}$) for all $t \geq 0$. The probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\mathfrak{F}$, denoted by $(\Omega, \mathcal{F}, \mathfrak{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, is called a *filtrated probability space*.

Notation. To shorten notation we write in the following $\mathfrak{F}$ to denote $\mathfrak{F} = (\mathcal{F}_t)_{t\geq 0}$.

Definition 2.19 ([73], Def. 1.1.9; Adaption). A process $X = (X_t)_{t\geq 0}$ defined on a filtrated probability space $(\Omega, \mathcal{F}, \mathfrak{F}, \mathbb{P})$ is called $\mathfrak{F}$-*adapted* if X_t is $\mathcal{F}_t$-measurable for each $t \geq 0$.

Definition 2.20 ([104], p. 16; Natural filtration). For a process $X = (X_t)_{t\geq 0}$, the filtration $\mathfrak{F}^X = (\mathcal{F}_t^X)_{t\geq 0}$ defined by $\mathcal{F}_t^X := \sigma(X_s : 0 \leq s \leq t)$ is said to be the *natural filtration* of X.

So the natural filtration $\mathfrak{F}^X$ is the smallest filtration making X adapted. It should be noted that the following two statements do not hold in general: 1. $\mathcal{F}_0^X$ contains all $\mathbb{P}$-null sets of $\mathcal{F}$; 2. $\mathfrak{F}^X$ is right-continuous. However these two statements are important technical assumptions for numerous results involving stochastic processes. So filtrations are usually modified as shown below to meet these technical requirements.

Definition 2.21 ([56], Def. 6.1). A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called *complete* if for all $A \subset B \in \mathcal{F}$ with $\mathbb{P}(B) = 0$ implies that $A \in \mathcal{F}$.

The assumption of a complete probability space is not a restriction since for any probability space there exists a unique one which is complete.

Proposition 2.22 ([56], Thm. 6.3). *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and*

$$\begin{aligned}
&\mathcal{N} := \{A \subset N : N \in \mathcal{F}, \mathbb{P}(N) = 0\},\\
&\widetilde{\mathcal{F}} := \{A \cup N : A \in \mathcal{F}, N \in \mathcal{N}\},\\
&\widetilde{\mathbb{P}} : \widetilde{\mathcal{F}} \to [0,1], \quad \widetilde{\mathbb{P}}(A \cup N) := \mathbb{P}(A) \quad \textit{for } A \in \mathcal{F}, N \in \mathcal{N}.
\end{aligned}$$

Then $\widetilde{\mathcal{F}}$ *is a* σ*-algebra,* $\widetilde{\mathbb{P}}$ *is well-defined and* $(\Omega, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ *is a complete probability space. Futhermore,* $\widetilde{\mathcal{F}}$ *is the smallest* σ*-algebra such that* $(\Omega, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ *is complete.*

Definition 2.23 ([56], p. 64)**.** The probability measure $\widetilde{\mathbb{P}}$ in Proposition 2.22 is called a *completion* of $\mathbb{P}$ and the probability space $(\Omega, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ in Proposition 2.22 is called a *completion* of $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 2.24 ([104], p. 3)**.** A filtered probability space $(\Omega, \mathcal{F}, \mathfrak{F}, \mathbb{P})$ is said to satisfy the *usual conditions* if

(i) the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is complete,

(ii) $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets of $\mathcal{F}$,

(iii) $\mathfrak{F}$ is right-continuous.

Remark 2.25. Proposition 2.22 has shown that for any probability space there exists a unique completion. A complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with the filtration $\mathfrak{F}$ can be enlarged to a filtrated probability space $(\Omega, \mathcal{F}, \widetilde{\mathfrak{F}}, \mathbb{P})$ satisfying the usual conditions by

$$\widetilde{\mathcal{F}}_t := \bigcap_{s>t} \sigma(F_s, \mathcal{N}) \quad \text{with } \mathcal{N} := \{A \in \mathcal{F} : \mathbb{P}(A) = 0\}.$$

Obviously, $\widetilde{\mathcal{F}}_t = \widetilde{\mathcal{F}}_{t+}$ for all $t \geq 0$ and $\widetilde{\mathcal{F}}_0$ contains all $\mathbb{P}$-null sets of $\mathcal{F}$. Therefore, for any given filtrated probability space, we can easily find one holding the usual conditions.

In the following, we assume that $(\Omega, \mathcal{F}, \mathfrak{F}, \mathbb{P})$ *is a filtrated probability space satisfying the usual conditions.*

We continue by introducing a stricter concept of measurability than in Definition 2.15.

Definition 2.26 ([73], Def. 1.1.11; Progessiv measurability)**.** A process $X = (X_t)_{t\geq 0}$ is called an $\mathfrak{F}$*-progressive process* or an $\mathfrak{F}$*-progressively measurable* if $(s,\omega) \mapsto X_s(\omega) : ([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}_t) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable for all $t \geq 0$, i.e. $\{(s,\omega) \in [0,t] \times \Omega : X_s(\omega) \in B, B \in \mathcal{B}(\mathbb{R})\} \in \mathcal{B}([0,t]) \otimes \mathcal{F}_t$ for all $t \geq 0$.

Proposition 2.27. *If a process* $X = (X_t)_{t\geq 0}$ *is* $\mathfrak{F}$*-progressively measurable, then* X *is measurable and* $\mathfrak{F}$*-adapted.*

Proof. Fix $t \geq 0$ and $B \in \mathcal{B}(\mathbb{R})$. From Definition 2.26 follows directly that $\{X_s \in B\} \in \mathcal{F}_t$ for all $s \in [0,t]$ and, in particular, $\{X_t \in B\} \in \mathcal{F}_t$, which means that X is $\mathfrak{F}$-adapted. Another consequence of Definition 2.26 is that $\{(s,\omega) \in [0,n] \times \Omega : X_s(\omega) \in B\} \in \mathcal{B}([0,n]) \otimes \mathcal{F}_t \subset \mathcal{B}^+ \otimes \mathcal{F}$ for all $n \in \mathbb{N}$. Hence $X\mathbb{1}_{[0,n]\times\Omega}$ is measurable w.r.t. $\mathcal{F}$ and, consequently, $X = \lim_{n\to\infty} X\mathbb{1}_{\Omega\times[0,n]}$ as well. □

Proposition 2.28 ([73], Prop. 1.1.12)**.** *Let* $X = (X_t)_{t\geq 0}$ *be measurable w.r.t.* $\mathcal{F}$ *and* $\mathfrak{F}$*-adapted. Then* X *has an* $\mathfrak{F}$*-progressively measurable modification.*

Proposition 2.29 ([73], Prop. 1.1.13)**.** *If a process* $X = (X_t)_{t\geq 0}$ *is* $\mathfrak{F}$*-adapted and left- or right-continuous, then* X *is* $\mathfrak{F}$*-progressively measurable.*

Lemma 2.30. *Let $X = (X_t)_{t\geq 0}$ be a non-negative $\mathfrak{F}$-progressive process. Then the process $(\int_0^t X_s \,\mathrm{d}s)_{t\geq 0}$ is $\mathfrak{F}$-adapted.*

Proof. The proof is given in Brémaud [20], solution of Exercise E10 on page 53. □

Proposition 2.31 ([79], Lemma 1.1 (a))**.** *Let $(E, \mathcal{E})$ be a measurable space, where E is a complete separable metric space. Let $X = (X_t)_{t\geq 0}$ be a càdlàg E-valued process defined on a filtrated probability space $(\Omega, \mathcal{F}, \mathfrak{F}, \mathbb{P})$ satisfying the usual conditions. Furthermore, let $\mathfrak{G} = (\mathcal{G}_t)_{t\geq 0}$ be a right-continuous filtration such that $\mathcal{G}_t$ contains all $\mathbb{P}$-null sets of $(\Omega, \mathcal{F})$ and $\mathcal{G}_t \subset \mathcal{F}_t$ for all $t \geq 0$. Then there exists a càdlàg modification of the process $(\mathbb{E}[f(X_t)\,|\,\mathcal{G}_t])_{t\geq 0}$ for all measurable bounded functions f defined on E.*

Let us recall the basic concept of (local) martingals.

Definition 2.32 ([104], p. 7; Martingale)**.** An $\mathfrak{F}$-adapted process $M = (M_t)_{t\geq 0}$ with $\mathbb{E}[|M_t|] < \infty$ for all $t \geq 0$ is called $(\mathbb{P}, \mathfrak{F})$-*martingale* if $\mathbb{E}[M_t|\mathcal{F}_s] = M_s$ for all $0 \leq s < t$. If the equation is weakened to $\mathbb{E}[M_t|\mathcal{F}_s] \geq M_s$ ($\mathbb{E}[M_t|\mathcal{F}_s] \leq M_s$) for every $0 \leq s < t$, then M is said to be a $(\mathbb{P}, \mathfrak{F})$-*submartingale* ($(\mathbb{P}, \mathfrak{F})$-*supermartingale*).

Convention. If it is clear that the underlying probability measure is $\mathbb{P}$, then we omit $\mathbb{P}$ in the appellation $(\mathbb{P}, \mathfrak{F})$-martingale and write only $\mathfrak{F}$-martingale. A similar convention applies to all subsequent definitions, in which a probability measure appears.

Notation. $\mathcal{M}(\mathbb{P}, \mathfrak{F})$ denotes the set of all càdlàg $(\mathbb{P}, \mathfrak{F})$-martingales starting at zero.

Definition 2.33. Let $0 < C < \infty$ be some constant. A process $X = (X_t)_{t\geq 0}$ is called *bounded* by C if $\sup_{t\geq 0} |X_t| < C$ $\mathbb{P}$-a.s.

Proposition 2.34 ([77], Rem. 21.68)**.** *A bounded local martingale is a martingale.*

For the definition of local martingales let us recall the notion of stopping time.

Definition 2.35 (Stopping time)**.** An $\mathcal{F}$-measurable function $\tau : \Omega \to [0, \infty]$ is called an $\mathfrak{F}$-*stopping time* if $\{\tau \leq t\} \in \mathcal{F}_t$ for all $t \in [0, \infty]$.

Proposition 2.36 ([104], Thm. I.1)**.** *Let $\tau : \Omega \to [0, \infty]$ be an $\mathcal{F}$-measurable function. Then τ is an $\mathfrak{F}$-stopping time if and only if $\{\tau < t\} \in \mathcal{F}_t$ for all $t \in [0, \infty]$.*

Definition 2.37 ([104], p. 4)**.** Let $X = (X_t)_{t\geq 0}$ be a real-valued process and $A \in \mathcal{B}(\mathbb{R})$. Then τ defined by $\tau(\omega) = \inf\{t > 0 : X_t \in A\}$ is called the *hitting time* of A for X.

Proposition 2.38 ([104], Thm. I.3, Thm. I.4)**.** *Let X be a real-valued $\mathfrak{F}$-adapted càdlàg process and A an open set or a closed set subset of $\mathbb{R}$. Then a hitting time of A for X is an $\mathfrak{F}$-stopping time.*

Definition 2.39. A stopping time τ is called *finite* if $\mathbb{P}(\tau < \infty) = 1$.

The next defined σ-algebra contains the knowledge included in a filtration up to a stopping time.

Definition 2.40 ([104], p. 5)**.** Let τ be a finite $\mathfrak{F}$-stopping time. Then

$$\mathcal{F}_\tau := \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t, t \geq 0\}$$

is said to be the *stopped time σ-algebra*.

Definition 2.41. Let $\mathcal{C}$ be a class of $\mathfrak{F}$-adapted processes. Then X is called a *local* $\mathcal{C}$-*process* if there exists a non-decreasing sequence of $\mathfrak{F}$-stopping times $(\tau_n)_{n\in\mathbb{N}}$ such that $\mathbb{P}(\tau_n \uparrow \infty) = 1$ as $n \to \infty$ and $X^{\tau_n} = (X_t^{\tau_n})_{t\geq 0} \in \mathcal{C}$ for each $n \in \mathbb{N}$, where $X_t^{\tau_n} := X_{t\wedge\tau_n}$. We write $X \in \mathcal{C}_{loc}$. The sequence $(\tau_n)_{n\in\mathbb{N}}$ is called a *localizing sequence* for X.

In accordance with the definition, the set $\mathcal{M}_{loc}(\mathfrak{F})$ denotes the set of local càdlàg $\mathfrak{F}$-martingales that occurs in the definition of a semimartingale. For this definition we have to recall functions of bounded variation.

Definition 2.42 ([38], p. 75; Partition)**.** A finite ordered set $\pi^n[a,b] := \{t_0, t_1, \dots, t_n\}$ for $n \in \mathbb{N}_0$ such that $a = t_0 < t_1 < \cdots < t_n = b$ is called a *partition* of $[a,b]$.

Definition 2.43 ([80], p. 421, [115], Def. 7.6.1; Bounded variation)**.** Let f be a function defined on $[a,b]$. The *total variation function* of f is defined by

$$V_a^b(f) := \sup_{\pi^n[a,b]} \left\{ \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})| \right\}. \tag{2.1}$$

Then the number $V_a^b(f)$ is called *total variation* of f on $[a,b]$. If $V_a^b(f) < \infty$, we say f is of *bounded variation* on $[a,b]$. The set of all functions of bounded variation on $[a,b]$ is denoted by $BV([a,b])$. Let f now be defined on $\mathbb{R}^+$. If $V_0^t(f) < \infty$ for all $t \geq 0$, then f is said to be of *locally bounded variation* on $\mathbb{R}^+$. In the case $\sup_{t>0} V_0^t(f) < \infty$, f is called *of bounded variation* on $\mathbb{R}^+$.

Proposition 2.44 ([115], Cor. 11.5.10)**.** *Let $f : [a,b] \to \mathbb{R}$ be a function of bounded variation. Then f is differentiable almost everywhere (in the sense of the Lebesgue measure).*

A subclass of functions of bounded variation are absolute continuous functions.

Definition 2.45 ([115], Def. 11.5.12)**.** A function $f : [a,b] \to \mathbb{R}$ is said to be *absolutely continuous* if for every $\varepsilon > 0$ there is a $\delta > 0$ such that given any finite sequence $(I_k)_{i=1,\dots,n}$ of pairwise disjoint open intervals $I_k := (a_k, b_k) \subset [a,b]$, we have

$$\sum_{i=1}^{n}(b_k - a_k) < \delta \Rightarrow \sum_{i=1}^{n} |f(b_k) - f(a_k)| < \varepsilon.$$

We write $AC([a,b])$ for the set of all absolutely continuous functions on $[a,b]$.

Lemma 2.46 ([115], Lemma 11.5.14)**.** *We have $Lip([a,b]) \subset AC([a,b]) \subset BV([a,b])$.*

Next we introduce a class of processes with a strong path regularity property.

Definition 2.47 (Protter [104], p. 101; FV process)**.** An $\mathfrak{F}$-adapted càdlàg process $A = (A_t)_{t\geq 0}$ with $A_0 = 0$ is said to be a *finite variation* process (FV process for short) w.r.t. $\mathbb{P}$ if $t \mapsto A_t(\omega)$ is of locally bounded variation on $\mathbb{R}^+$ for $\mathbb{P}$-almost all $\omega \in \Omega$.

Notation. From now on, we write $\mathcal{A}(\mathbb{P}, \mathfrak{F})$ for the set of all $\mathfrak{F}$-adapted finite variation processes w.r.t. $\mathbb{P}$.

Definition 2.48 ([104], p. 102)**.** An $\mathfrak{F}$-adapted càdlàg process $Y = (Y_t)_{t\geq 0}$ is called a $(\mathbb{P}, \mathfrak{F})$-*semimartingale* if there exists a decomposition of the form

$$Y_t = Y_0 + M_t + A_t, \quad t \geq 0, \tag{2.2}$$

where $(M_t)_{t\geq 0} \in \mathcal{M}_{loc}(\mathbb{P}, \mathfrak{F})$ and $(A_t)_{t\in I} \in \mathcal{A}(\mathbb{P}, \mathfrak{F})$.

Definition 2.49 (Continuous part)**.** Let $X = (X_t)_{t\geq 0}$ be an $\mathfrak{F}$-semimartingale. Then process $(X)^c = ((X)^c_t)_{t\geq 0}$ denotes the *continuous part of the process* X starting at zero, i.e.

$$(X)^c_t = X_t - \sum_{0<s\leq t} \Delta X_s - X_0, \quad t \geq 0.$$

A description of the behaviour of stochastic processes is provided by the quadratic variation.

Definition 2.50 ([75], p. 218)**.** Let X and Y be two $(\mathbb{P}, \mathfrak{F})$-semimartingales. The *quadratic covariation* process of X and Y, denoted by $[X,Y] = ([X,Y]_t)_{t\geq 0}$, is defined by

$$[X,Y]_t := \lim \sum_{i=0}^{n-1} \big(X_{t^n_{i+1}} - X_{t^n_i}\big)\big(Y_{t^n_{i+1}} - Y_{t^n_i}\big), \quad t \geq 0,$$

where the limit is understood as the limit in probability and is taken over shrinking partitions $(t^n_i)_i$ of the interval $[0,t]$ when $\delta_n = \max_i(t^n_{i+1} - t^n_i) \to 0$. The process $[X,Y]$ is also known as the *square bracket process.* The *quadratic variation* process X, denoted by $[X] = ([X]_t)_{t\geq 0}$, is defined by $[X]_t = [X,X]_t$ for all $t \geq 0$.

Notice that the existence of the quadratic covariation can be shown. Other fundamental properties are summarized in the next proposition which are cited from Klebaner [75, p. 218 ff.] as well as Protter [104, p. 66 ff.].

Proposition 2.51. *For two* $(\mathbb{P}, \mathfrak{F})$*-semimartingales* X *and* Y *the following statements are satisfied:*

(i) $[X,Y]$ *is a càdlàg* $\mathfrak{F}$*-adapted FV process.*

(ii) $[X,Y]$ *is bilinear and symmetric.*

(iii) $[X,Y] = \frac{1}{2}\big([X+Y] - [X] - [Y]\big)$ *(polarization identity).*

(iv) If one of the processes X *or* Y *is an FV process:* $[X,Y]_t = \sum_{0<s\leq t} \Delta X_s \Delta Y_s$ *for every* $t \geq 0$.

(v) $[X]^c = [X^c]$.

(vi) For any $t \geq 0$, $[X]_t = [X]^c_t + X_0^2 + \sum_{0<s\leq t}(\Delta X_s)^2$.

(vii) $[X,Y]_t = X_t Y_t - \int_0^t X_{s-}\,\mathrm{d}Y_s - \int_0^t Y_{s-} X_s$ *for all* $t \geq 0$.

(viii) For any $t \geq 0$, $\big[\int_0^\cdot H_s\,\mathrm{d}X_s, \int_0^\cdot K_s\,\mathrm{d}Y_s\big]_t = \int_0^t H_s K_s\,\mathrm{d}[X,Y]_s$ *for all* $\mathfrak{F}$*-predictable processes* $(H_t)_{t\geq 0}$, $(K_t)_{t\geq 0}$*, where the stochastic integrals exist.*

We continue with the notion of predictability, which is an integral part of studying processes in the present of jumps. An process describing the wealth of an insurances company is an example for a process with jumps occurring as a result of insurance payments. Since the wealth process of an insurer is the main objection of our optimization problem, we have to deal with the notion of predictable processes.

Definition 2.52 ([38], p. 25)**.** Let $\mathfrak{F} = (\mathcal{F}_t)_{t\geq 0}$ be a filtration. The family $\mathcal{R}(\mathfrak{F})$ of subsets of $\mathbb{R}^+ \times \Omega$, defined by

$$\mathcal{R}(\mathfrak{F}) := \big\{\{0\} \times F_0 : F_0 \in \mathcal{F}_0\big\} \cup \big\{(s,t] \times F : F \in \mathcal{F}_s, 0 \leq s < t\big\},$$

is called the *class of* $\mathfrak{F}$*-predictable rectangles.*

Definition 2.53 ([38], p. 25)**.** The σ-algebra $\mathcal{P}(\mathfrak{F})$ on $\mathbb{R}^+ \times \Omega$ generated by the class of predictable rectangles $\mathcal{P}(\mathfrak{F}) := \sigma(\mathcal{R}(\mathfrak{F}))$ is called *$\mathfrak{F}$-predictable σ-algebra.* A function $X : \mathbb{R}^+ \times \Omega \to \mathbb{R}$ is said to be *$\mathfrak{F}$-predictable ($\mathfrak{F}$-previsible)* if X is $\mathcal{P}(\mathfrak{F})$-measurable.

Theorem 2.54 ([75], p. 213)**.** *An $\mathfrak{F}$-adapted left-continuous process is $\mathcal{P}(\mathfrak{F})$-measurable.*

Example 2.55. Let X be an $\mathfrak{F}$-adapted and càdlàg process. Clearly, the left-hand limit process X_- is left-continuous as well as $\mathfrak{F}$-adapted by assumption. Hence X_- is $\mathfrak{F}$-predictable by Theorem 2.54.

Proposition 2.56 ([104], p. 103)**.** *Every $\mathfrak{F}$-predictable process X is $\mathfrak{F}$-progressively measurable.*

After this short overview of some basics of stochastic processes, we can now turn to some important tools of stochastic analysis, which will be intensively used in the analysis of optimization problems.

2.3 Tools of stochastic analysis

Throughout this section, we suppose that all stochastic quantities are defined on the filtrated probability space $(\Omega, \mathcal{F}, \mathfrak{F}, \mathbb{P})$ satisfying the usual conditions, see Definition 2.24. We begin this section with probably the best known formula of stochastic calculus, the Itô-Doeblin formula named after the Japanese Kiyoshi Itô, who is well-known as one of the developers of the formula, and the French Wolfgang Doeblin, who as soldier during the World War II developed a comparable formula to the other one from Itô. The famous Itô-Doeblin formula is hereafter formulated in the version for general semimartingals. For a treatment of integration w.r.t. general semimartingales we refer the reader to Protter [104] and Klebaner [75]. It should be noted that we use the following notation.

Notation. Throughout this work, we use the Riemann integral notation for integrals w.r.t. the Lebesgue measure λ, i.e. we write

$$\int_a^b g(s)\,\mathrm{d}s \text{ instead of } \int_{[a,b]} g(s)\,\lambda(\mathrm{d}s), \quad a < b,$$

for any Borel measurable function g.

Theorem 2.57 ([104], Thm. II.32; Itô-Doeblin formula)**.** *Let $d \geq 2$, D be an open subset of $\mathbb{R}^d$ and $f \in C^{1,2}(\mathbb{R}^+ \times D)$ be a real valued function. Furthermore, let $X = (X_t)_{t\geq 0}$ be an D-valued $\mathfrak{F}$-semimartingale. Then $f(t, X) = (f(t, X_t))_{t\geq 0}$ is an $\mathfrak{F}$-semimartingale holding*

$$\begin{aligned} f(t, X_t) - f(0, X_0) = &\int_0^t f_t(s, X_s)\,\mathrm{d}s + \sum_{i=1}^d \int_0^t f_{x_i}(s, X_{s-})\,\mathrm{d}X_s^i \\ &+ \frac{1}{2} \sum_{i,j=1}^d \int_0^t f_{x_i x_j}(s, X_{s-})\,\mathrm{d}[X^i, X^j]_s^c \\ &+ \sum_{0<s\leq t} \Big(f(s, X_s) - f(s, X_{s-}) - \sum_{i=1}^d f_{x_i}(s, X_{s-})\Delta X_s^i \Big), \quad t \geq 0. \end{aligned}$$

In this work the following version of the Itô-Doeblin formula is applied frequently, which follows immediately from the definition of the continuous part of the process X given in Definition 2.48.

Corollary 2.58. *Let the conditions of Theorem 2.57 be satisfied. Then* $f(t,X) = (f(t,X_t))_{t\geq 0}$ *is an* $\mathfrak{F}$*-semimartingale holding*

$$\begin{aligned} f(t,X_t) - f(0,X_0) &= \int_0^t f_t(s,X_s)\,\mathrm{d}s + \sum_{i=1}^d \int_0^t f_{x_i}(s,X_{s-})\,\mathrm{d}(X^i)^c_s \\ &\quad + \frac{1}{2}\sum_{i,j=1}^d \int_0^t f_{x_i x_j}(s,X_{s-})\,\mathrm{d}[X^i,X^j]^c_s \\ &\quad + \sum_{0<s\leq t}\Big(f(s,X_s) - f(s,X_{s-})\Big), \quad t \geq 0. \end{aligned}$$

Theorem 2.59 ([104], p. 83; Integration by Parts, Product Rule)**.** *Let* $X = (X_t)_{t\geq 0}$ *and* $Y = (Y_t)_{t\geq 0}$ *be two* $\mathfrak{F}$*-semimartingales. Then* $(X_t\,Y_t)_{t\geq 0}$ *is an* $\mathfrak{F}$*-semimartingale holding*

$$X_t\,Y_t - X_0\,Y_0 = \int_0^t X_{s-}\,\mathrm{d}Y_s + \int_0^t Y_{s-}\,\mathrm{d}X_s + [X,Y]_t, \quad t \geq 0.$$

Theorem 2.60 ([75], Thm. 8.33, Protter [104], Thm. II.37)**.** *Let* $X = (X_t)_{t\geq 0}$ *be an* $\mathfrak{F}$*-semimartingale. Then there exists an unique (up to indistinguishability)* $\mathfrak{F}$*-semimartingale* $Z = (Z_t)_{t\geq 0}$ *that satisfies the stochastic differential equation (SDE for short)*

$$\mathrm{d}Z_t = Z_{t-}\,\mathrm{d}X_t, \quad Z_0 = 1, \quad t \geq 0,$$

where Z *is given by*

$$Z_t = \mathcal{E}(X)_t := e^{X_t - X_0 - \frac{1}{2}[X]^c_t} \prod_{0<s\leq t}(1+\Delta X_s)e^{-\Delta X_s}, \quad t \geq 0,$$

where the infinite product converges.

Definition 2.61. The process $\mathcal{E}(X) = (\mathcal{E}(X)_t)_{t\geq 0}$ defined in the previous theorem is called the *stochastic exponential* or the *Doléans-Dade exponential* of X.

2.4 Simple point processes and marked point processes

In this section we review some standard facts on (marked) point processes, in particular filter results, based on Brémaud [20], Last and Brandt [80], Jacobsen [71] and Leimcke [82, Ch. 2, 3]. Throughout this section, all stochastic quantities are defined on the filtrated probability space $(\Omega, \mathcal{F}, \mathfrak{F}, \mathbb{P})$ satisfying the usual conditions, see Definition 2.24.

2.4.1 Basic definitions

In this work we only deal with point processes on the non-negative real half-line. In actuarial mathematics, simple point processes are useful to describe the arrivals of insurance claims. A simple model of a claim number process is the Poisson process. We introduce a large class of point processes which contains almost all the point processes of interest in insurance mathematics. The main idea behind this concept is that the "nature" of a point process can be described by an "infinitesimal characterisation". Before introducing this characterisation, we give some basic definitions and properties. We start with the definition of a simple point process verbatim cited from Jacobsen [71, Def. 2.1.1].

Definition 2.62 (Simple point process)**.** A sequence $N = (T_n)_{n\in\mathbb{N}}$ of $\overline{\mathbb{R}}_+$-valued random variables is called a *simple point process* (SPP for short) if

$$\begin{aligned}
&\mathbb{P}(0 < T_1 \le T_2 \le \ldots) = 1, \\
&\mathbb{P}(T_n < T_{n+1}, T_n < \infty) = 1, \quad n \in \mathbb{N}, \\
&\mathbb{P}(\lim_{n\to\infty} T_n = \infty) = 1.
\end{aligned} \tag{2.3}$$

Remark 2.63. (i) We will use point processes to model the arrival of insurance claims or trigger events occurring randomly in time. So we interpret the random variables $(T_n)_{n\in\mathbb{N}}$ as random times. In the following, we also refer to these random times as jump times since a process, which counts the number of claims, jumps upwards of size one at these time points.

(ii) It is also possible to define an SPP $N = (T_n)_{n\in\mathbb{N}}$ without the condition given in the last line of (2.3) such that $\lim_{n\to\infty} T_n < \infty$ is possible. If $\lim_{n\to\infty} T_n < \infty$, then the SPP "explodes" at a certain time. That is, there has to be a finite accumulation point of jump times. As models for claim arrival times, only point processes with $\mathbb{P}(\lim_{n\to\infty} T_n = \infty) = 1$ are of practical interest since it is not a realistic situation that an infinite number of claims occur in a finite time interval.

A simple point process can be interpreted as a special case of a marked point process. A marked point process is a double sequence of random variables. The first sequence is a point process describing times at which certain events occur. The second sequence represents additional information about the events. We say that the second sequence is the mark of the arrival times. For instance, the first sequence describes claims arrivals and the second sequence describes the corresponding amount of claims. The following notation is used to describe the mark of events that never occur.

Notation. Let $(E, \mathcal{E})$ denote a measurable space called the *mark space*. Furthermore, ∇ denotes a singleton which is not a point of the set E and we write

$$\overline{E} := E \cup \{\nabla\} \quad \text{and} \quad \overline{\mathcal{E}} := \mathcal{E} \vee \{\nabla\}.$$

The definition of marked point processes is introduced below following Jacobsen [71, Def. 2.1.2].

Definition 2.64 (Marked point process)**.** A double sequence $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$ of $\overline{\mathbb{R}}^+$-valued random variables (T_n) and $\overline{E}$-valued random variables (Y_n) is called a *marked point process* with a *mark space* E (E-MPP for short) if

$$\begin{aligned}
&(T_n)_{n\in\mathbb{N}} \quad \text{is an SPP}, \\
&\mathbb{P}(Y_n \in E, \quad T_n < \infty) = \mathbb{P}(T_n < \infty),\ n \in \mathbb{N}, \\
&\mathbb{P}(Y_n = \nabla, T_n = \infty) = \mathbb{P}(T_n = \infty),\ n \in \mathbb{N}.
\end{aligned}$$

For every $n \in \mathbb{N}$, the random variable Y_n is said to be the *mark* of T_n. The singleton ∇ is called an *irrelevant mark.*

Remark 2.65. An SPP can be seen as an MPP, which has a mark space of cardinality one. This case is called the *univariate case* or the *unmarked case.*

We will discuss different views of point processes in the following. First of all, a common interpretation of a simple point process is in terms of a counting process.

Definition 2.66 (Counting process)**.** For a point process $N = (T_n)_{n\in\mathbb{N}}$ the associated *counting process*, denoted by $N = (N_t)_{t\geq 0}$, is defined by

$$N_t := \sum_{n\in\mathbb{N}} \mathbb{1}_{\{T_n \leq t\}}, \quad t \geq 0.$$

Remark 2.67. It is easily seen that N_t counts the number of jump times T_n which occur up to time t. Clearly, $(N_t)_{t\geq 0}$ is right-continuous, non-decreasing and piecewise constant with jumps upwards of magnitude one, which justifies the name counting process. We have supposed that N is nonexplosive. Moreover, $N_t < \infty$ $\mathbb{P}$-a.s. for every $t \geq 0$ due to the property $\mathbb{P}(\lim_{n\to\infty} T_n = \infty) = 1$. Hence $(N_t)_{t\geq 0}$ is also càdlàg. Thus $(N_t)_{t\geq 0}$ is an FV process.

It is justified to denote the simple point process $N = (T_n)_{n\in\mathbb{N}}$ and the counting process $N = (N_t)_{t\geq 0}$ both by N since there is a one-to-one correspondence between this two views. Indeed, for a given SPP $N = (T_n)_{n\in\mathbb{N}}$, the associated counting process $N = (N_t)_{t\geq 0}$ can be obtained by its definition. Conversely, suppose that a right-continuous non-decreasing $\overline{\mathbb{N}}_0$-valued process $N = (N_t)_{t\geq 0}$ with jumps upwards of size one is given, then the sequence $N = (T_n)_{n\in\mathbb{N}}$ can be easily recovered by the relationship

$$T_n = \inf\{t \geq 0 : N_t \geq n\}, \quad n \in \mathbb{N},$$

where $\inf \emptyset := \infty$. Moreover, we have

$$\{T_n \leq t\} = \{N_t \geq n\}, \quad n \in \mathbb{N}, \quad t \geq 0.$$

Due to one-to-one correspondence we can make the following convention.

Convention. From now on, a counting process $(N_t)_{t\geq 0}$ associated to a simple point process $N = (T_n)_{n\in\mathbb{N}}$ is called a simple point process, too.

A third view of an SPP is as random element taking values in the following defined sequence space.

Notation ([71], p. 10). Let $\bar{t} = (t_n)_{n\in\mathbb{N}}$ denote a sequence of values $t_n \in (0, \infty]$ for every $n \in \mathbb{N}$. Set

$$K := \{\bar{t} \in (0, \infty]^{\mathbb{N}} : t_1 \leq t_2 \leq \cdots, t_n < t_{n+1} \text{ if } t_n < \infty, n \in \mathbb{N}\},$$

Furthermore, we define

$$\bar{t}(A) = \sum_{n\in\mathbb{N}} \mathbb{1}_{\{t_n \in A\}}, \quad A \in \mathcal{B}^+, \tag{2.4}$$

and the coordinate projections

$$\begin{aligned} &T_n^\circ : K \to (0, \infty], \quad T_n^\circ(\bar{t}) := t_n, \quad n \in \mathbb{N}, \\ &T_\infty^\circ(\bar{t}) := \lim_{n\to\infty} T_n^\circ(\bar{t}) = \lim_{n\to\infty} t_n. \end{aligned}$$

Moreover, $\mathcal{K}$ denotes the smallest σ-algebra of subsets of K such that all coordinate projections T_n°, $n \in \mathbb{N}$, are measurable, i.e.

$$\mathcal{K} := \sigma\big(T_n^\circ, n \in \mathbb{N}\big) = \sigma\left(\bigcup_{n\in\mathbb{N}} T_n^{\circ-1}(\mathcal{B}(0, \infty])\right).$$

Thus an SPP can be regarded as a random element with values in the measurable space $(K, \mathcal{K})$.

Remark 2.68. It can be seen from the analysis above that $\bar{t}$ can also be treated as a counting function, which allows the interpretation K as the set of all counting functions.

The view of an SPP as random element will be used to define the so-called mixed Poisson process, which plays an important role in our claim arrival model introduced in Section 3.1. It is clear that the perspective of simple point processes as random sequences or as random elements taking values in the space of sequences K are equivalent. Furthermore, according to Last and Brandt [80, Lemma 2.2.2], it holds $\sigma(N) = \sigma(N(A) : A \in \mathcal{B}^+) = \sigma(N_t : t \geq 0)$, which means that the σ-algebra generated by the K-valued random element N is equal to the σ-algebra generated by the family of random variables $\{N(A) : A \in \mathcal{B}^+\}$ and $\{N_t : t \geq 0\}$, respectively. This justifies the use of the same symbol N for the random sequence and the corresponding random counting measure. Hence we have three different ways to express an SPP, where the three views carry the same information. The accumulated information of an SPP up to time t is described by the following filtration.

Notation. The natural filtration $\mathfrak{F}^N = (\mathcal{F}_t^N)_{t\geq 0}$ of a point process $N = (N_t)_{t\geq 0}$ is given by $\mathcal{F}_t^N = \sigma(N_s : 0 \leq s \leq t)$ for all $t \geq 0$.

Notice that $\mathcal{F}_t^N$ is equal to the σ-algebras generated by the K-valued stochastic process describing the dynamic evolution of N and by the point process N being viewed as random counting measure restricted on $[0, t]$, respectively, cf. [80, Eq. (2.2.16)].

Theorem 2.69 ([104], Thm. I.25)**.** *The natural filtration $\mathfrak{F}^N$ of a simple point process N is right-continuous.*

The next remark is dedicated to the usual conditions of a filtrated probability space with $\mathfrak{F}^N$ as filtration which has to be in force for a martingale representation theorem used in the proof of the filter result stated in Theorem 2.94.

Remark 2.70. The filtrated probability space $(\Omega, \mathcal{F}_\infty^N, \mathfrak{F}^N, \mathbb{P})$ can be modified such that the usual conditions are satisfied. Recall that for any probability space one can find a unique completion, see Proposition 2.22 and that the natural filtration $\mathfrak{F}^N$ of an SPP N is right-continuous. Defining $\widetilde{\mathfrak{F}}^N = (\widetilde{\mathcal{F}}_t^N)_{t\geq 0}$ by $\widetilde{\mathcal{F}}_t^N := \mathcal{F}_t^N \vee \mathcal{N}$, where $\mathcal{N}$ is the family of $\mathbb{P}$-null sets of $\mathcal{F}_\infty^N$, we obtain that $(\Omega, \widetilde{\mathcal{F}}_\infty^N, \widetilde{\mathfrak{F}}^N, \widetilde{\mathbb{P}})$ holds the usual conditions since if $\mathfrak{F}^N$ is right-continuous then $\widetilde{\mathfrak{F}}^N$ is right-continuous, see Brémaud [20, Thm. A.2.T35].

Now we turn our attention to marked point processes. It is often convenient to regard a marked point process as a random counting measure. For this purpose, we introduce the following definitions.

Definition 2.71 ([77], Def. 8.25; Transition kernel)**.** Let $(\Omega_1, \mathcal{A}_1)$, $(\Omega_2, \mathcal{A}_2)$ be two measurable spaces. A mapping $\kappa : \Omega_1 \times \mathcal{A}_2 \to \overline{\mathbb{R}}^+$ is called a *transition kernel* from $(\Omega_1, \mathcal{A}_1)$ to $(\Omega_2, \mathcal{A}_2)$ (from Ω_1 to Ω_2 for short) if

(i) $\kappa(\cdot, A_2)$ is $\mathcal{A}_1$-measurable for all $A_2 \in \mathcal{A}_2$;

(ii) $\kappa(\omega_1, \cdot)$ is a measure on $(\Omega_2, \mathcal{A}_2)$ for $\mathbb{P}$-a.a. $\omega_1 \in \Omega_1$.

If in (ii) the measure is a probability measure for $\mathbb{P}$-a.a. $\omega_1 \in \Omega_1$, then κ is called a *stochastic kernel*, a *transition probability kernel* or a *Markov kernel* from $(\Omega_1, \mathcal{A}_1)$ to $(\Omega_2, \mathcal{A}_2)$.

A random measure is a certain transition kernel.

Definition 2.72 ([80], p. 74; Random measure)**.** Let $(S,\mathcal{S})$ be a measurable space. A transition kernel ν from $(\Omega,\mathcal{F})$ to $(S,\mathcal{S})$ is called a *random measure* on $(S,\mathcal{S})$ (on S for short) and a stochastic kernel from $(\Omega,\mathcal{F})$ to $(S,\mathcal{S})$ is called a *random probability measure* on $(S,\mathcal{S})$ (on S for short).

For clarity, we use the same convention as in Last and Brandt [80, pages 74–75].

Convention. Let ν be a random measure on $(S,\mathcal{S})$. Sometimes we write $\nu(\omega,\cdot)=\nu(\omega)$ for $\omega\in\Omega$ and $\nu(\omega,A)=\nu(A)$ for $A\in\mathcal{S}$. Unless otherwise stated, an equation with a term $\nu(A)$ means for all $\omega\in\Omega$. By definition, $\nu(\omega)$ is a measure on S for every $\omega\in\Omega$. Let (p) be a property of a measure on S. We say ν *has the property* (p) if $\nu(\omega)$ has the property for all $\omega\in\Omega$.

Now, we define the mentioned random counting measure associated to a marked point process similar to Jacobsen [71, Eq. (2,4)].

Notation. For an E-MPP $\Phi=(T_n,Z_n)_{n\in\mathbb{N}}$ we define

$$\Phi(\mathrm{d}t,\mathrm{d}z) := \sum_{n\in\mathbb{N}:T_n<\infty} \delta_{(T_n,Z_n)}(\mathrm{d}t,\mathrm{d}z), \tag{2.5}$$

where $\delta_{(T_n,Z_n)}(\omega,\mathrm{d}t,\mathrm{d}z)=\delta_{(T_n(\omega),Z_n(\omega))}(\mathrm{d}t,\mathrm{d}z)$ is the Dirac measure at the point $(T_n(\omega),Z_n(\omega))$ on the product space $(\mathbb{R}^+\times E,\mathcal{B}^+\otimes\mathcal{E})$.

It is easy to verify that $\Phi(\mathrm{d}t,\mathrm{d}z)$ given by (2.5) is a random counting measure on $\mathbb{R}^+\times E$. It might be confusing that Φ is used in (2.5) although Φ also denotes the MPP. In the following we will see that this notation is justified since an MPP $\Phi=(T_n,Z_n)_{n\in\mathbb{N}}$ can be identified with a random counting measure given by (2.5). Before the identification, we define a counting process associated to Φ.

Definition 2.73 ([80], p. 5)**.** Let $\Phi=(T_n,Z_n)_{n\in\mathbb{N}}$ be an E-MPP and let $\Phi(\mathrm{d}t,\mathrm{d}z)$ be the associated measure. For any $\emptyset\neq A\in\mathcal{E}$, we define a process $(\Phi(t,A))_{t\geq 0}$ by

$$\Phi(t,A) := \Phi([0,t]\times A) = \sum_{n\in\mathbb{N}} \mathbb{1}_{\{T_n<t\}}\mathbb{1}_{\{Z_n\in A\}}, \quad t\geq 0.$$

The process $(\Phi(t,A))_{t\geq 0}$ is called the *counting process with marks in* A associated to Φ. Furthermore, we define a process $(\Phi(t))_{t\geq 0}$ by

$$\Phi(t) := \Phi(t,E) = \sum_{n\in\mathbb{N}} \mathbb{1}_{\{T_n<t\}}$$

The process $(\Phi(t))_{t\geq 0}$ is called the *counting process* associated to Φ.

Definition 2.74. An E-MPP Φ is said to be $\mathfrak{F}$-*adapted* if $(\Phi(t,B))_{t\geq 0}$ is $\mathfrak{F}$-adapted for all $B\in\mathcal{E}$.

Recall that there is a one-to-one correspondence between an SPP expressed as a sequence and counting process. It is more difficult to verify that there is a one-to-one correspondence between a double sequence $\Phi=(T_n,Z_n)_{n\in\mathbb{N}}$ and the associated random counting measure $\Phi(\mathrm{d}t,\mathrm{d}z)$. Let $\Phi=(T_n,Z_n)_{n\in\mathbb{N}}$ be an E-MPP. Clearly, we obtain

$\Phi(A,B)$ for every $(A,B) \in \mathcal{B}^+ \otimes \mathcal{E}$ from $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$ by definition. It is also possible to recover the sequences $(T_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ from $\Phi(\mathrm{d}t, \mathrm{d}z)$ by

$$T_n = \inf\{t \geq 0 : \Phi(t) \geq n\},\ \{Y_n \in B\} = \{T_n < \infty\} \cap \{\Phi(T_n) \times B > 0\},\ B \in \mathcal{E},\ n \in \mathbb{N},$$

cf. [80, Eq. (2.2.13)]. So we have also a one-to-one correspondence between an MPP stated as a double sequence and a random counting measure, which justifies the following convention.

Convention. From now on we call a random counting measure $\Phi(\mathrm{d}t, \mathrm{d}z)$ associated to a MPP $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$ also MPP.

Next we characterize the history of an MPP as in Last and Brandt [80, p. 9], which expresses the accumulated available information.

Notation. Let $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$ be an E-MPP. We define

$$\mathcal{F}_t^{\Phi} := \sigma\big(\Phi((a,b] \times B) : 0 \leq a < b \leq t, B \in \mathcal{E}\big), \quad t \geq 0. \tag{2.6}$$

It should be noted that $\mathcal{F}_t^{\Phi}$ is equal to the σ-algebra generated by the random counting measure of Φ restricted onto $[0,t] \times E$, see [80, Eq. (2.2.16)]. This equation asserts that all characterizations of the dynamic evolution of Φ carry the same information at every time t. So the filtration $\mathfrak{F}^{\Phi} := (\mathcal{F}_t^{\Phi})_{t\geq 0}$ is the natural filtration of Φ, which gives the history of Φ, i.e. $\mathcal{F}_t^{\Phi}$ includes the available information of Φ up to time t.

Theorem 2.75 ([80], Thm. 2.2.4)**.** *The filtration $\mathfrak{F}^{\Phi}$ is right-continuous.*

Notice that $(\Omega, \mathcal{F}_\infty^{\Phi}, \mathfrak{F}^{\Phi}, \mathbb{P})$ can be altered to a filtrated probability space holding the usual conditions with the same arguments as in Remark 2.70. With this concluding remark on the introduction of marked point processes we can move forward to roll out the mentioned infinitesimal characterisation of simple point processes.

2.4.2 Intensities of simple point processes

In modern point process theory, the behaviour of an SPP is specified by an intensity defined hereafter.

Definition 2.76 ([20], Def. II.D7)**.** Let $N = (N_t)_{t\geq 0}$ be an $\mathfrak{F}$-adapted SPP and $\lambda = (\lambda_t)_{t\geq 0}$ a non-negative $\mathfrak{F}$-progressively measurable process such that

$$\int_0^t \lambda_s \,\mathrm{d}s < \infty \quad \mathbb{P}\text{-a.s.}, \quad t \geq 0.$$

Then we say N admits a $(\mathbb{P}, \mathfrak{F})$-*intensity* λ if

$$\mathbb{E}\left[\int_0^\infty H_s \,\mathrm{d}N_s\right] = \mathbb{E}\left[\int_0^\infty H_s\, \lambda_s \,\mathrm{d}s\right] \tag{2.7}$$

for all non-negative $\mathfrak{F}$-predictable processes $(H_t)_{t\geq 0}$.

If we interpret an SPP with an intensity λ as a claim counting process, then the intensity λ is also called the *claim arrival rate*.

Remark 2.77. Equation (2.7) holds for all non-negative $\mathfrak{F}$-predictable processes $P = (P_t)_{t\geq 0}$, that is to say

$$\mathrm{d}N_t(\omega)\,\mathbb{P}(\mathrm{d}\omega) = \lambda_t(\omega)\,\mathrm{d}t\,\mathbb{P}(\mathrm{d}\omega) \quad \text{on } (\mathbb{R}^+\times\Omega, \mathcal{P}(\mathfrak{F})),$$

where $\mathcal{P}(\mathfrak{F})$ denotes the $\mathfrak{F}$-predictable σ-algebra defined on $\mathbb{R}^+\times\Omega$, see Definition 2.53.

Here are some important properties of point processes with intensities.

Theorem 2.78 ([20], Thm. II.T8)**.** *Let $N = (N_t)_{t\geq 0}$ be an $\mathfrak{F}$-adapted SPP with the $\mathfrak{F}$-intensity $\lambda = (\lambda_t)_{t\geq 0}$ and let $M = (M_t)_{t\geq 0}$ be a process defined by*

$$M_t := N_t - \int_0^t \lambda_s\,\mathrm{d}s, \quad t \geq 0.$$

Then the following assertions hold:

(i) *N is non-explosive, i.e. $N_t < \infty$ $\mathbb{P}$-a.s. for all $t \geq 0$.*

(ii) *N is integrable if and only if $\mathbb{E}\left[\int_0^t \lambda_s\,\mathrm{d}s\right] < \infty$ for all $t \geq 0$.*

(iii) *M is a local $\mathfrak{F}$-martingale.*

(iv) *If N is integrable, then M is an $\mathfrak{F}$-martingale.*

(v) *If X is an $\mathfrak{F}$-predictable process such that $\mathbb{E}\big[\int_0^t |X_s|\lambda_s\,\mathrm{d}s\big] < \infty$ for all $t \geq 0$, then $(\int_0^t X_s\,\mathrm{d}M_s)_{t\geq 0}$ is an $\mathfrak{F}$-martingale.*

(vi) *If X is an $\mathfrak{F}$-predictable process such that $\mathbb{P}\big(\int_0^t |X_s|\lambda_s\,\mathrm{d}s < \infty\big) = 1$ for all $t \geq 0$, then $(\int_0^t X_s\,\mathrm{d}M_s)_{t\geq 0}$ is a local $\mathfrak{F}$-martingale.*

Proof. Statements (i), (iii), (v) and (vi) can be found in Theorem II.T8 of Brémaud [20]. Statements (ii) follows by applying $P_s = \mathbb{1}_{\{s\leq t\}}$ to (2.7), which yields $\mathbb{E}[N_t] = \mathbb{E}\left[\int_0^t \lambda_s\,\mathrm{d}s\right] < \infty$ for all $t \geq 0$. It remains to prove statement (iv) which is a special case of statement (v) by choosing $X \equiv 1$. Indeed, statement (iv) yields, if $\mathbb{E}[\int_0^t \lambda_s\,\mathrm{d}s] < \infty$ for $t \geq 0$, then $(\int_0^t(\mathrm{d}N_s - \lambda_s\,\mathrm{d}s))_{t\geq 0} = (N_t - \int_0^t \lambda_s\,\mathrm{d}s)_{t\geq 0}$ is an $\mathfrak{F}$-martingale. By statement (ii), the condition $\mathbb{E}[\int_0^t \lambda_s\,\mathrm{d}s] < \infty$ for $t \geq 0$ is equivalent to the assertion that N is integrable. This completes the proof. Note that, by similar argumentation, statement (iii) is a special case of statement (v). □

The following example shows that we can not expect the intensity to be unique.

Example 2.79. Let N be an SPP with an $\mathfrak{F}$-intensity $\lambda = (\lambda_t)_{t\geq 0}$ that is càdlàg such that the left-hand limit process $\lambda_- = (\lambda_{t-})_{t\geq 0}$ is well-defined. By Lemma 2.13, the sample paths of λ and λ_- differ at most countably many points. Hence, for any non-negative $\mathfrak{F}$-predictable processes $(H_t)_{t\geq 0}$, we have

$$\mathbb{E}\left[\int_0^\infty H_s\,\mathrm{d}N_s\right] = \mathbb{E}\left[\int_0^\infty H_s\,\lambda_s\,\mathrm{d}s\right] = \mathbb{E}\left[\int_0^\infty H_s\,\lambda_{s-}\,\mathrm{d}s\right],$$

where the second equation holds since we are integrating w.r.t. Lebesgue measure. Therefore, the integrand can be changed at most countably many points without changing the integral. Consequently, λ_- is also an $\mathfrak{F}^N$-intensity of N.

We will see that the intensity is unique under the requirement of predictability.

Theorem 2.80 ([20], Thm. II.T12, Thm. II.T13)**.** *Let N be an SPP with $\mathfrak{F}$-intensity $\lambda = (\lambda_t)_{t\geq 0}$. Then N admits also an $\mathfrak{F}$-intensity $\widetilde{\lambda} = (\widetilde{\lambda}_t)_{t\geq 0}$ which is $\mathfrak{F}$-predictable. Furthermore, $\widetilde{\lambda}$ is unique in the sense that if $\mu = (\mu_t)_{t\geq 0}$ is another $\mathfrak{F}$-predictable intensity of N, then*

$$\widetilde{\lambda}_t(\omega) = \mu_t(\omega) \quad \mathrm{d}N_t(\omega)\,\mathbb{P}(\mathrm{d}\omega)\text{-a.s.}$$

Convention. If the intensity of an SPP is predictable, we will speak of *the* intensity instead of *an* intensity, according to the uniqueness in that case.

The next proposition deals with a change of the history for intensities.

Proposition 2.81. *Let N be an SPP with an $\mathfrak{F}$-intensity $\lambda = (\lambda_t)_{t\geq 0}$ and let $\mathfrak{G} = (\mathcal{G}_t)_{t\geq 0}$ be a smaller filtration than $\mathfrak{F} = (\mathcal{F}_t)_{t\geq 0}$. If $\mu = (\mu_t)_{t\geq 0}$ is a modification of $(\mathbb{E}[\lambda_t \,|\, \mathcal{G}_t])_{t\geq 0}$ such that μ is $\mathfrak{G}$-progressive measurable, then μ is an $\mathfrak{G}$-intensity of N.*

Proof. This proof is inspired by Brémaud [20, p. 32]. Let $(H_t)_{t\geq 0}$ be an arbitrary non-negative $\mathfrak{G}$-predictable process. By definition of λ as $\mathfrak{F}$-intensity of N, we have

$$\mathbb{E}\left[\int_0^\infty H_s\,\mathrm{d}N_s\right] = \mathbb{E}\left[\int_0^\infty H_s\,\lambda_s\,\mathrm{d}s\right], \tag{2.8}$$

where we take into account that every $\mathfrak{G}$-predictable process is also $\mathfrak{F}$-predictable. Since $(H_t\,\lambda_t)_{t\geq 0}$ is non-negative, we can apply Fubini's theorem which yields

$$\begin{aligned} \mathbb{E}\left[\int_0^\infty H_s\,\lambda_s\,\mathrm{d}s\right] &= \int_0^\infty \mathbb{E}[H_s\,\lambda_s]\,\mathrm{d}s = \int_0^\infty \mathbb{E}[\mathbb{E}[H_s\,\lambda_s \,|\, \mathcal{G}_s]]\,\mathrm{d}s \\ &= \int_0^\infty \mathbb{E}[H_s\,\mathbb{E}[\lambda_s \,|\, \mathcal{G}_s]]\,\mathrm{d}s = \mathbb{E}\left[\int_0^\infty H_s\,\mathbb{E}[\lambda_s \,|\, \mathcal{G}_s]\,\mathrm{d}s\right], \end{aligned} \tag{2.9}$$

where we have used the fact that H_s is $\mathcal{G}_s$-measurable for all $s \geq 0$ (which follows from the $\mathfrak{G}$-predictability, cf. Prop. 2.27). Combining (2.8) with (2.9), we obtain

$$\mathbb{E}\left[\int_0^\infty H_s\,\mathrm{d}N_s\right] = \mathbb{E}\left[\int_0^\infty H_s\,\mathbb{E}[\lambda_s \,|\, \mathcal{G}_s]\,\mathrm{d}s\right],$$

i.e. $(\mathbb{E}[\lambda_t \,|\, \mathcal{G}_t])_{t\geq 0}$ satisfies condition (2.7) of the definition of the intensity. The assumption allows $\mu = (\mu_t)_{t\geq 0}$ to be a modification of $(\mathbb{E}[\lambda_t \,|\, \mathcal{G}_t])_{t\geq 0}$, which is $\mathfrak{G}$-progressive measurable. So we can replace $\mathbb{E}[\lambda_s \,|\, \mathcal{G}_s]$ by μ_s in the last line of (2.9). We are left to show that $\int_0^t \mu_s\,\mathrm{d}s < \infty$ $\mathbb{P}$-a.s. for all $t \geq 0$. For this aim let us set for any $n \in \mathbb{N}$,

$$S_n := \begin{cases} \inf\{t > 0 : \int_0^t \lambda_s\,\mathrm{d}s \geq n\} & \text{if } \{\ldots\} \neq \emptyset, \\ \infty & \text{otherwise.} \end{cases}$$

On account of Proposition 2.38, S_n is an $\mathfrak{F}$-stopping time for all $n \in \mathbb{N}$. From $\int_0^t \lambda_s\,\mathrm{d}s < \infty$ $\mathbb{P}$-a.s. for all $t \geq 0$, it follows that $S_n \nearrow \infty$ $\mathbb{P}$-a.s. as $n \to \infty$. Define a nonnegative process $C = (C_t)_{t\geq 0}$ by $C_t = \mathbb{1}_{\{S_n \geq t\}}$ for $t \geq 0$. By Proposition 2.36, $\{S_n < t\} \in \mathcal{F}_t$ and, in consequence, $\{S_n \geq t\} \in \mathcal{F}_t$ for every $t \geq 0$. Hence C is $\mathfrak{F}$-adapted. The left-continuity of C implies that C is $\mathfrak{F}$-predictable, see Theorem 2.54. Writing (2.9) with C_t, we get

$$\mathbb{E}\left[\int_0^{S_n} \mu_s\,\mathrm{d}s\right] = \mathbb{E}\left[\int_0^{S_n} \lambda_s\,\mathrm{d}s\right] \leq n < \infty, \quad n \in \mathbb{N}_0.$$

Therefore, $\int_0^{S_n} \widetilde{\lambda}_s \,\mathrm{d}s < \infty$ $\mathbb{P}$-a.s. for all $n \in \mathbb{N}$. Since $S_n \nearrow \infty$ $\mathbb{P}$-a.s., there exists, for any $t \geq 0$, a number $\tilde{n} \in \mathbb{N}$ such that $\int_0^t \mu_s \,\mathrm{d}s \leq \int_0^{S_{\tilde{n}}} \mu_s \,\mathrm{d}s < \infty$ $\mathbb{P}$-a.s. □

Remark 2.82. It might be not completely clear why the $\mathfrak{G}$-intensity μ can not always be defined as the projection of λ_t on $\mathcal{G}_t$ under the setting of the previous proposition. The reason is the requirement of $\mathfrak{G}$-progressiveness, which is not satisfied by $(\mathbb{E}[\lambda_t \,|\, \mathcal{G}_t])_{t\geq 0}$ in general. Nevertheless, the $\mathfrak{G}$-intensity μ can also be defined (in an abstract way) without further assumptions on μ, compare Theorem II.T14 in Brémaud [20].

The last part of this sub-section is devoted to another class of simple point processes – the so-called doubly stochastic Poisson processes – which turns out to be a sub-class of point processes with intensity under additional assumptions. Loosely speaking, a doubly stochastic Poisson process is an inhomogeneous Poisson process with a random process as intensity, where the sample path of the stochastic intensity is chosen at the beginning. Afterwards, this trajectory is used to generate a Poisson process by acting as its intensity. So the doubly stochastic Poisson process has a conditional Poisson property. There are different definitions of doubly stochastic Poisson processes. We introduce here the definition given in Brémaud [20], Definition II.D1.

Definition 2.83 (Doubly stochastic Poisson process)**.** Let $N = (N_t)_{t\geq 0}$ be an $\mathfrak{F}$-adapted SPP and let $\lambda = (\lambda_t)_{t\geq 0}$ be a non-negative measurable process satisfying, for every $t \geq 0$,

$$\lambda_t \text{ is } \mathcal{F}_0\text{-measurable} \tag{2.10}$$

and

$$\int_0^t \lambda_v \,\mathrm{d}v < \infty \quad \mathbb{P}\text{-a.s.} \tag{2.11}$$

If, for any $u \in \mathbb{R}$ and $0 \leq s \leq t$,

$$\mathbb{E}\left[e^{iu(N_t - N_s)} \,\middle|\, \mathcal{F}_s \right] = \exp\left\{ (e^{iu} - 1) \int_s^t \lambda_v \,\mathrm{d}v \right\}, \tag{2.12}$$

then N is called a $(\mathbb{P}, \mathfrak{F})$-*doubly stochastic Poisson process* (DSPP) with intensity $\lambda = (\lambda_t)_{t\geq 0}$.

Remark 2.84. (i) In literature, the doubly stochastic Poisson process is also called *Cox process* since the process was introduced by Cox [42].

(ii) Equation (2.12) gives the conditional characteristic function of $N_t - N_s$ given $\mathcal{F}_s$. Since the characteristic function of the Poisson distribution is stated on the right-hand side of (2.12), it follows that

$$\mathbb{P}(N_t - N_s = k \,|\, \mathcal{F}_s) = \frac{\left(\int_s^t \lambda_v \,\mathrm{d}v \right)^k}{k!} e^{-\int_s^t \lambda_v \,\mathrm{d}v}, \quad k \in \mathbb{N}_0, \quad 0 \leq s \leq t.$$

The equation indicates that a DSPP with intensity $(\lambda_t)_{t\geq 0}$ is an inhomogeneous Poisson process[1] with intensity measure $\lambda_t \,\mathrm{d}t$ conditioned upon the realization of $(\lambda_t)_{t\geq 0}$.

Proposition 2.85. *Let $N = (N_t)_{t\geq 0}$ be an integrable $\mathfrak{F}$-DSPP with intensity $\lambda = (\lambda_t)_{t\geq 0}$. Then N is an SPP with intensity $\widetilde{\lambda} = (\widetilde{\lambda}_t)_{t\geq 0}$, where $\widetilde{\lambda}$ is a modification of λ. If λ is additional left- or right-continuous, then $\widetilde{\lambda}$ and λ are indistinguishable.*

[1]For the definition of an inhomogeneous Poisson process see Brémaud [21, Def. 4.D5].

Proof. Following the argumentation given at the beginning of Chapter II.2 in Brémaud [20], it yields

$$\mathbb{E}\left[\int_0^\infty H_s \,\mathrm{d}N_s\right] = \mathbb{E}\left[\int_0^\infty H_s \,\lambda_s \,\mathrm{d}s\right]$$

for all non-negative $\mathfrak{F}$-predictable processes $(H_t)_{t\geq 0}$. Due to Eq. (2.10), λ is $\mathfrak{F}$-adapted and hence there exists an $\mathfrak{F}$-progressively measurable modification of λ (cf. Prop. 2.28), which can clearly replace λ_s in the equation above. Moreover, according to Proposition 2.29, λ is $\mathfrak{F}$-progressively measurable if λ is additional left- or right-continuous. □

The next theorem yields a characterization of doubly stochastic Poisson processes.

Theorem 2.86 ([20], Thm. II.T4)**.** *Let $N = (N_t)_{t\geq 0}$ be an $\mathfrak{F}$-adapted SPP and let $\lambda = (\lambda_t)_{t\geq 0}$ be a non-negative measurable process satisfying*

$$\lambda_t \text{ is } \mathcal{F}_0\text{-measurable, and,} \tag{2.13}$$

$$\int_0^t \lambda_s \,\mathrm{d}s < \infty \quad \mathbb{P}\text{-a.s.} \tag{2.14}$$

for all $t \geq 0$. Then N is an $\mathfrak{F}$-DSPP with intensity λ if

$$\mathbb{E}\left[\int_0^\infty H_s \,\mathrm{d}N_s\right] = \mathbb{E}\left[\int_0^\infty H_s \,\lambda_s \,\mathrm{d}s\right], \tag{2.15}$$

for all non-negative $\mathfrak{F}$-predictable processes $(H_t)_{t\geq 0}$.

Examples for doubly stochastic Poisson processes are the so-called mixed Poisson processes, which can be seen as a sub-class of doubly stochastic Poisson processes. A mixed Poisson process will be an essential component of the claim arrival model in Section 3.1. According to Grandell [67, Rem. 4.2] there are two common definitions of mixed Poisson processe: The first one was introduced by Lundberg [89, p. 72] and involves birth processes, while the second one, cf. e.g. Grandell [67, Def. 4.2], is given by mixing of Poisson processes. We will use the second one since this will lead to a Bayesian interpretation of the mixed Poisson process as explained in Section 3.1. However, Albrecht [6] proved the equivalence of both definitions. There are three more definitions being found in the literature according to Lyberopoulos et al. [90].

Before stating the definition, recall the introduced space K on page 18, which can be seen as a set of all counting functions, wherein an SPP N is viewed as a random element takes values. Furthermore, we need the following notation.

Notation. We denote the distribution of N by Π_N, i.e. Π_N is a probability measure on $(K, \mathcal{K})$. Moreover, Π_λ denotes the distribution of a homogeneous Poisson process with intensity $\lambda > 0$.

The measurability of $\lambda \mapsto \Pi_\lambda(B)$ (see Grandell [65, Lemma 1.1]) allows us to introduce the mixed Poisson process as follows.

Definition 2.87 ([67], Def. 4.2; Mixed Poisson process)**.** Let Λ be a positive random variable with distribution Π_Λ. An SPP N is called a *mixed Poisson process* with *mixing distribution* Π_Λ (mPP(Π_Λ)), if its distribution is given by

$$\Pi_N(B) = \int_0^\infty \Pi_\lambda(B)\,\Pi_\Lambda(\mathrm{d}\lambda), \quad B \in \mathcal{K}.$$

The definition states that a mixed Poisson process is the mixture of homogeneous Poisson processes[2] by a given distribution. It can be interpreted that first a realization λ of the random variable Λ is chosen and given that realization N is a homogeneous Poisson process with intensity λ. This interpretation makes it clear that a mixed Poisson process is a special case of a doubly stochastic Poisson process, see Proposition 2.90.

For calculation issues, it is convenient to use the following definition of a mixed Poisson process, where the equivalence to the definition above holds according to Lyberopoulos et al. [90, p. 2].

Definition 2.88 ([67], Def. 4.3)**.** Let Λ be a positive random variable with distribution Π_Λ and $\widetilde{N}$ a (standard) homogeneous Poisson process with intensity 1, which is independent of Λ. Then the SPP $(N_t)_{t\geq 0}$ given by $N_t = \widetilde{N}_{\Lambda t}$ for every $t \geq 0$ is said to be a *mixed Poisson process* with *mixing distribution* Π_Λ.

With this definition, the following elementary property of a mixed Poisson process can be derived.

Proposition 2.89 (Albrecher et al. [5], p. 144–145)**.** *An an mPP*(Π_Λ) *N satisfies, for any* $t \geq 0$*,*

$$\mathbb{P}(N_t = n) = \int_0^\infty \frac{(\lambda t)^n}{n!} e^{-\lambda t}\, \Pi_\Lambda(\mathrm{d}\lambda), \quad \mathbb{E}[N_t] = t\,\mathbb{E}[\Lambda].$$

Proposition 2.90 ([67], p. 86)**.** *Let* Λ *be an* $\mathcal{F}_0$*-measurable positive random variable with distribution* Π_Λ *and let N be an mPP*(Π_Λ)*. Then N is an* $\mathfrak{F}$*-DSPP with constant intensity* $(\Lambda)_{t\geq 0}$*.*

Proof. The conditions (2.10) and (2.11) of Definition 2.83 of a DSPP are obviously satisfied by the $\mathcal{F}_0$-measurability of Λ. Moreover, the Poisson property (2.12) (compare also Remark 2.84 (ii)) follows immediately from Proposition 2.89. □

Proposition 2.91. *Let* Λ *be a positive random distributed according to distribution* Π_Λ *with a finite mean. Then the mixed Poisson process N with the mixing distribution* Π_Λ *is an SPP with* $\mathfrak{F}$*-intensity* $(\Lambda)_{t\geq 0}$ *and* $\mathfrak{F}^N$*-intensity* $(\mathbb{E}[\Lambda \mid \mathcal{F}_t^N])_{t\geq 0}$*.*

Proof. From Proposition 2.89, we can conclude that N is integrable since $\mathbb{E}[\Lambda] < \infty$. Hence N admits the $\mathfrak{F}$-intensity $(\Lambda)_{t\geq 0}$, which follows by combining Proposition 2.90 with Proposition 2.85. The statement that $(\mathbb{E}[\Lambda \mid \mathcal{F}_t^N])_{t\geq 0}$ in an $\mathfrak{F}^N$-intensity of N is given in Proposition 4.1 in Grandell [67]. □

After these introductions about the characterization of point processes with intensity as well as the sub-classes of doubly stochastic Poisson processes and mixed Poisson processes, we continue with the theory of filtering with point process observations, which provides a method to make inferences about unknown intensities of point processes.

2.4.3 Filtering with point process observations

Although claim arrivals can be observed, the claim arrival rate of a claim number process is typically not observable by an insurance company. This leads to a filter problems with counting observations. In this section we present the innovations method for filtering with point process observations introduced by Brémaud [20, Sec. IV.1] in a restricted setting which is sufficient for this work. We assume the unobservable state processes to be constant as our restriction. More precisely, we have the following setting.

[2] A definition of the *homogeneous Poisson process* can be found in Brémaud [21, Def. 4.D4].

Assumption 2.92. Let $(\Omega, \mathcal{F}, \mathfrak{F}, \mathbb{P})$ be a filtrated probability space satisfying the usual conditions, Z an $\mathcal{F}_0$-measurable random variable and N an integrable SPP with $\mathfrak{F}$-intensity $\lambda = (\lambda_t)_{t\geq 0}$. Moreover, we make the following assumptions:

(i) The global filtration $\mathfrak{F} = (\mathcal{F}_t)_{t\geq 0}$ records the events outside the observed history $\mathfrak{F}^N = (\mathcal{F}_t^N)_{t\geq 0}$, i.e. $\mathcal{F}_t^N \subsetneq \mathcal{F}_t$ for all $t \geq 0$. Here we even assume that $\mathcal{F}_\infty^N \subsetneq \mathcal{F}_0$.

(ii) The filtrated probability space $(\Omega, \mathcal{F}_\infty^N, \mathfrak{F}^N, \mathbb{P})$ has been modified as shown in Remark 2.70. By abuse notation, we still write $(\Omega, \mathcal{F}_\infty^N, \mathfrak{F}^N, \mathbb{P})$ instead of $(\Omega, \widetilde{\mathcal{F}}_\infty^N, \widetilde{\mathfrak{F}}^N, \widetilde{\mathbb{P}})$ for the modified space.

Remark 2.93. It should be noted that the assumption above defines a Bayesian setting. Therefore, the unknown (random) parameter of a distribution given some observations is described by the posterior distribution of this parameter given the jump times of the observed point process at the jump times. However, the presented filter method will yield the evolution between the jump times and will turn out to be valuable for the stochastic control approach in Section 4.4 and Section 5.4. This issue will be addressed again in Section 4.2.

To state the filter result, we introduce some notations.

Notation. Throughout this section, $(\widehat{\lambda}_t)_{t\geq 0}$ denotes an $\mathfrak{F}^N$-intensity of N and by $\widehat{Z} = (\widehat{Z}_t)_{t\geq 0}$ a càdlàg modification of $(\mathbb{E}[Z \mid \mathcal{F}_t^N])_{t\geq 0}$, i.e.

$$\widehat{Z}_t = \mathbb{E}[Z \mid \mathcal{F}_t^N] \quad \mathbb{P}\text{-a.s.}, \quad t \geq 0.$$

Justification of the notation. Proposition 2.81 ensures that there exists an $\mathfrak{F}^N$-intensity of N and the process $(\mathbb{E}[Z \mid \mathcal{F}_t^N])_{t\geq 0}$ admits a càdlàg modification according to Proposition 2.31. □

The first step of the innovations approach is to project the unobservable random variable Z (process, in general) on the observed filtration, which yields (see Brémaud [20, Thm. IV.T1])

$$\widehat{Z}_t = \mathbb{E}[Z_0] + \widehat{m}_t, \quad t \geq 0,$$

where $(\widehat{m}_t)_{t\geq 0}$ is an $\mathfrak{F}^N$-martingale starting at zero. According to the martingale representation theorem given in Brémaud [20, Thm. III.T17], there exists an $\mathfrak{F}^N$-predictable process $(K_t)_{t\geq 0}$ such that $\widehat{m}_t = \int_0^t K_s\big(\mathrm{d}N_s - \widehat{\lambda}_s\,\mathrm{d}s\big)$ for all $t \geq 0$. Thus

$$\widehat{Z}_t = \mathbb{E}[Z_0] + \int_0^t \widehat{a}_s\,\mathrm{d}s + \int_0^t K_s\,\mathrm{d}\widehat{M}_s, \quad t \geq 0, \tag{2.16}$$

This equation is called a *filter* of Z. The process $\big(\mathrm{d}N_s - \widehat{\lambda}_s\,\mathrm{d}s\big)_{t\geq 0}$ is said to be the *innovations process* and the process $(K_t)_{t\geq 0}$ is called *innovations gain*. The term innovation is justified since $\mathrm{d}N_t$ describes the observation in $[t, t+\mathrm{d}t)$ and $\widehat{\lambda}_t\,\mathrm{d}t$ is what one expect to happen in $[t, t+\mathrm{d}t)$. Therefore, $\mathrm{d}N_t - \widehat{\lambda}_t\,\mathrm{d}t$ gives the new information.

The filter of Z_t given in (2.16) is an abstract existence result since there is no expression for the innovations gain $(K_t)_{t\geq 0}$. Consequently, the next step is to find the innovations gain, which is made in the following theorem.

Theorem 2.94. *Let Assumption 2.92 be in force. Then*

$$\widehat{Z}_t = \mathbb{E}[Z_0] + \int_0^t (A_s - \widehat{Z}_{s-})\,(\mathrm{d}N_s - \widehat{\lambda}_s\,\mathrm{d}s), \quad t \geq 0, \tag{2.17}$$

up to indistinguishability, where $(A_t)_{t\geq 0}$ *is an* $\mathfrak{F}^N$*-predictable processes* $\widehat{\lambda}_t(\omega)\,\mathrm{d}t\,\mathbb{P}(\mathrm{d}\omega)$*-uniquely given by*

$$\mathbb{E}\left[\int_0^t H_s\, Z\, \lambda_s\, \mathrm{d}s\right] = \mathbb{E}\left[\int_0^t H_s\, A_s\, \widehat{\lambda}_s\, \mathrm{d}s\right], \quad t \geq 0, \tag{2.18}$$

for all non-negative bounded $\mathfrak{F}^N$*-predictable processes* $(H_t)_{t\geq 0}$.

Proof. The assertion follows directly from Theorem IV.T2 in Brémaud [20]. Notice that the mentioned theorem is stated for multivariate point processes, while we formulate only a univariate version. To see the direct connection of the stated result and Theorem IV.T2 in Brémaud [20], it should be noted that the process Ψ_3 in Theorem IV.T2 can be chosen to be constant zero if the unobservable state process and N have no common jump times which is the case in the presented setting since we only consider constant state processes. □

Note that $\widehat{\lambda}_t(\omega)\,\mathrm{d}t\,\mathbb{P}(\mathrm{d}\omega)$ is a measure on $(\mathbb{R}^+\times\Omega, \mathcal{P}(\mathfrak{F}^N))$ which is equal to the measure $\mathrm{d}N_t(\omega)\,\mathbb{P}(\mathrm{d}\omega)$ on the same measure space. After this remark, we proceed by extending the filter results to the case of marked point process observations. For this purpose, an infinitesimal characterization for marked point processes is required.

2.4.4 Intensity kernels of marked point processes

The intensity kernel of an MPP is analogue to the intensity of an SPP. Recall that $(E, \mathcal{E})$ is some measurable space and for the E-MPP Φ, $\Phi(\omega, \mathrm{d}t, \mathrm{d}z)$ is a counting measure on E for every $\omega \in \Omega$. Hence, for any measurable function $H : \mathbb{R}^+ \times \Omega \times E \to \mathbb{R}$, we have

$$\int_0^\infty \int_E H(s,z)\,\Phi(\mathrm{d}s, \mathrm{d}z) = \sum_{n\in\mathbb{N}} H(T_n, Z_n)\mathbb{1}_{\{T_n<\infty\}},$$
$$\int_0^t \int_E H(s,z)\,\Phi(\mathrm{d}s, \mathrm{d}z) = \sum_{n\in\mathbb{N}} H(T_n, Z_n)\mathbb{1}_{\{T_n\leq t\}}, \quad t \geq 0,$$

where we use the convention that $\int_a^b$ is interpreted as $\int_{(a,b]}$ if $b < \infty$ and $\int_{(a,b)}$ if $b = \infty$.
For the next definition, let us recall Definition 2.53 of the $\mathfrak{F}$-predictable σ-algebra $\mathcal{P}(\mathfrak{F})$.

Definition 2.95 ([20], p. 235)**.** A function $H : \mathbb{R}^+\times\Omega\times E \to \mathbb{R}$ is called an $\mathfrak{F}$*-predictable function indexed by* E if H is $\mathcal{P}(\mathfrak{F}) \otimes \mathcal{E}$-measurable.

Definition 2.96 ([20], Def. VIII.D2; Intensity kernel)**.** Let $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$ be an $\mathfrak{F}$-adapted E-MPP. Furthermore, let $\lambda : \mathbb{R}^+ \times \Omega \times \mathcal{E} \to \overline{\mathbb{R}}^+$ be a transition kernel from $(\mathbb{R}^+\times\Omega, \mathcal{B}^+\otimes\mathcal{F})$ to $(E, \mathcal{E})$ such that, for any $B \in \mathcal{E}$, $(\lambda(t, B))_{t\geq 0}$ is the $(\mathbb{P}, \mathfrak{F})$-predictable $(\mathbb{P}, \mathfrak{F})$-intensity of $(\Phi(t, B))_{t\geq 0}$. Then the family of random measures $(\lambda(t, \mathrm{d}z))_{t\geq 0}$ is said to be the $(\mathbb{P}, \mathfrak{F})$*-intensity kernel* of Φ.

Notice that it is justified to write "the" intensity kernel because of the uniqueness of predictable intensities, cf. Thm. 2.80. Another important property of intensity kernels is described in the following theorem.

Theorem 2.97 ([20], Thm. VIII.T3; Projection Theorem)**.** *Let* $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$ *be an* E*-MPP with the* $(\mathbb{P}, \mathfrak{F})$*-intensity kernel* $(\lambda(t, \mathrm{d}z))_{t\geq 0}$*. Then, for any non-negative*

$\mathfrak{F}$-predictable function H indexed by E, we have

$$\mathbb{E}\left[\int_0^\infty \int_E H(s,z)\,\Phi(\mathrm{d}t,\mathrm{d}z)\right] = \mathbb{E}\left[\int_0^\infty \int_E H(s,z)\,\lambda(s,\mathrm{d}z)\,\mathrm{d}s\right].$$

Here are some elementary properties of a marked point process with intensity kernel.

Corollary 2.98 ([20], Cor. VIII.C4)**.** *Let $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$ be an E-MPP with the $(\mathbb{P},\mathfrak{F})$-intensity kernel $(\lambda(t,\mathrm{d}z))_{t\geq 0}$. Furthermore, let H be an $\mathfrak{F}$-predictable function indexed by E such that*

$$\int_0^t \int_E |H(s,z)|\,\lambda(s,\mathrm{d}z)\,\mathrm{d}s < \infty \quad \mathbb{P}\text{-a.s.}, \quad t \geq 0.$$

Then the process $M = (M_t)_{t\geq 0}$ defined by

$$M_t := \int_0^t \int_E H(s,z)\,\big(\Phi(\mathrm{d}t,\mathrm{d}z) - \lambda(s,\mathrm{d}z)\,\mathrm{d}s\big)$$

is a local $(\mathbb{P},\mathfrak{F})$-martingale. Moreover, if

$$\mathbb{E}\left[\int_0^t \int_E |H(s,z)|\,\lambda(s,\mathrm{d}z)\,\mathrm{d}s\right] < \infty, \quad t \geq 0,$$

then M is a $(\mathbb{P},\mathfrak{F})$-martingale.

We continue with discussing the theory of filtering with marked point process observations using the innovations method.

2.4.5 Filtering with marked point process observations

In this section we will cite the filter results given in Brémaud [20, Ch. VIII.2], where we again limit the results to the case of constant unobservable processes. To state the filter results, we need intensity kernels with a special structure given in the next definition.

Definition 2.99 (Local characteristic)**.** Let Φ be an E-MPP with the $(\mathbb{P},\mathfrak{F})$-intensity kernel $(\lambda(t,\mathrm{d}z))_{t\geq 0}$ of the form

$$\lambda(t,\mathrm{d}z) = \lambda_t\,\mu(t,\mathrm{d}z),$$

where $(\lambda_t)_{t\geq 0}$ is a non-negative $\mathfrak{F}$-predictable process and μ is a stochastic kernel from $(\mathbb{R}^+\times\Omega, \mathcal{B}^+\otimes\mathcal{F})$ to $(E,\mathcal{E})$. Then the family of pairs $(\lambda_t, \mu(t,\mathrm{d}z))_{t\geq 0}$ is called a *local $(\mathbb{P},\mathfrak{F})$-characteristic of* Φ.

We are now in the position to specify the filter framework.

Assumption 2.100. Let $(\Omega,\mathcal{F},\mathfrak{F},\mathbb{P})$ be a filtrated probability space satisfying the usual conditions, Z an $\mathcal{F}_0$-measurable random variable and Φ an E-MPP with $\mathfrak{F}$-local characteristic $(\lambda_t,\mu(t,\mathrm{d}z))_{t\geq 0}$. Moreover, we make the following assumptions:

(i) The global filtration $\mathfrak{F} = (\mathcal{F}_t)_{t\geq 0}$ records the events outside the observed history $\mathfrak{F}^\Phi = (\mathcal{F}_t^\Phi)_{t\geq 0}$, i.e. $\mathcal{F}_t^\Phi \subsetneq \mathcal{F}_t$ for all $t\geq 0$. Here we even assume that $\mathcal{F}_\infty^\Phi \subsetneq \mathcal{F}_0$.

(ii) The filtrated probability space $(\Omega,\mathcal{F}_\infty^\Phi,\mathfrak{F}^\Phi,\mathbb{P})$ has been modified as shown in Remark 2.70, where we keep the same notation for the modified object.

Similar to Section 2.4.3, we use the following notation.

Notation. Throughout this section, let $(\widehat{\lambda}_t, \widehat{\mu}(t, \mathrm{d}z))_{t\geq 0}$ denote the $\mathfrak{F}^\Phi$-local characteristic of Φ and by $\widehat{Z} = (\widehat{Z}_t)_{t\geq 0}$ a càdlàg modification of $(\mathbb{E}[Z \mid \mathcal{F}_t^\Phi])_{t\geq 0}$, i.e.

$$\widehat{Z}_t = \mathbb{E}[Z \mid \mathcal{F}_t^\Phi] \quad \mathbb{P}\text{-a.s.}, \quad t \geq 0.$$

With this notation the procedure of Section 2.4.3 can be readily extended to the setting of this section, which results in the following filter equation.

Theorem 2.101. *Let the conditions of Assumption 2.100 be prevailed. Then*

$$\widehat{Z}_t = \mathbb{E}[Z_0] + \int_0^t \int_E \big(A(s,z) - \widehat{Z}_{s-}\big)\big(\Phi(\mathrm{d}s, \mathrm{d}z) - \widehat{\lambda}_s\, \widehat{\mu}(s, \mathrm{d}z)\, \mathrm{d}s\big), \quad t \geq 0, \tag{2.19}$$

up to indistinguishability, where A is an $\mathfrak{F}^\Phi$-predictable function indexed by E which is $\widehat{\lambda}_t(\omega)\widehat{\mu}(t,\omega,\mathrm{d}z)\mathbb{P}(\mathrm{d}\omega)$-uniquely given by

$$\mathbb{E}\left[\int_0^t \int_E Z\, H(s,z)\, \lambda_s\, \mu(s, \mathrm{d}z)\, \mathrm{d}s\right] = \mathbb{E}\left[\int_0^t \int_E A(s,z)\, H(s,z)\, \widehat{\Lambda}_s\, \widehat{\mu}(s, \mathrm{d}z)\, \mathrm{d}s\right],\ t \geq 0, \tag{2.20}$$

for all bounded $\mathfrak{F}^\Phi$-predictable functions H indexed by E.

Proof. The theorem is an immediate consequence of Thm. VIII.T9 in Brémaud [20]. □

It is worth noting that $\widehat{\lambda}_t(\omega)\widehat{\mu}(t,\omega,\mathrm{d}z)\mathbb{P}(\mathrm{d}\omega)$ is a measure on $(\mathbb{R}^+ \times \Omega \times E, \mathcal{P}(\mathfrak{F}^\Phi) \otimes \mathcal{E})$. This remark closes the chapter and now we have all the necessary foundations to address the optimization problem of an insurance company described in the next chapter.

Chapter 3

The control problem under partial information

We consider an insurance company with several lines of business (LoBs). The aim is to maximize the expected utility of the terminal surplus of the considered insurance company by choosing optimal investment and reinsurance strategies. For this purpose, we need a model for the surplus process. An important component of the surplus process is the aggregated claim amount process which will be specified in the following section.

3.1 The aggregated claim amount process

In the following, let $d \in \mathbb{N}$ be the number of LoB of the insurer. The arrival times of the insurance claims will be described by a d dimensional counting process, which determines the dependencies between the claim arrival of the various kinds of insurances risks.

The claim arrival process. To impose interdependencies between different types of insurance risks, we use a common shock risk model inspired from the thinning and shift model considered in Bäuerle and Grübel [27], where we suppose that there is no time gap between the shock events and the claim arrival times. To state the claim arrival model, the following notation will be used.

Notation. We set $\mathbb{D} := \{1, \dots, d\}$ and we denote by $\mathcal{P}(\mathbb{D})$ the set of all subsets of $\mathbb{D}$. Furthermore, we write $\ell := 2^d - 1$ for the number of elements of $\mathcal{P}(\mathbb{D})$ minus one. Moreover, $\mathfrak{F} = (\mathcal{F}_t)_{t \geq 0}$ denotes the global filtration. That means, $\mathfrak{F}$ includes all available information, in particular information the insurer does not have.

For the construction of the claim arrival model, let us introduce an $\mathcal{F}_0$-measurable random variable Λ with distribution Π_Λ. The initial point for the claim arrival model is a mixed Poisson process $N = (N_t)_{t \geq 0}$ with mixing distribution Π_Λ, compare Definition 2.87. We interpret the arrival times of the mixed Poisson process N, denoted by $(T_n)_{n \in \mathbb{N}}$, as time points of the events which trigger various kinds of insurance claims. Therefore, N is called the *trigger process* with the *background intensity* Λ and the jump times $(T_n)_{n \in \mathbb{N}}$ are said to be the *arrival times of trigger events* or *of shock events.*

Before describing the connection between the shock events and the insurance claims, let us go deeper into the interpretation of the trigger process. Mixed Poisson processes[1] were introduce by Dubourdieu [53] for an actuarial application, namely as a claim arrival model in health and accident insurance business. A common interpretation of a mixed Poisson processes is as counting processes which consists of various sub-processes

[1]For full treatment of mixed Poisson processes we refer the reader to Grandell [67].

behaving as Poisson processes with certain intensities.[2] However, Definition 2.87 of a mixed Poisson process provides the natural interpretation as the Bayesian version of the Poisson process, where Π_Λ is the prior distribution of the intensity.[3] This interpretation is in accordance with the intention for our use of the mixed Poisson process since, in practice, the background intensity Λ is unknown to the insurance company. As mentioned in Wichelhaus and von Rohrscheidt [118, Ch. 3], it is unreasonable to use a model with fixed but unknown parameter. To avoid this, we have chosen the intensity of the process N as positive $\mathcal{F}_0$-measurable random variable Λ. So we have a two-step stochastic phenomenon. First, the realisation λ of the background intensity Λ is selected, where the realisation is unknown to the insurer since $\mathfrak{F}$ is the global filtration. Secondly, the trigger arrival times are generated with the chosen intensity λ.

The effects of a shock event on the insurance lines will be discussed next. For this purpose, we introduce a sequence $(Z_n)_{n\in\mathbb{N}}$ of categorical variables taking values in a set with ℓ elements. Each value is associated to a category describing in which LoBs the damages by the corresponding shock are caused. We quantify these categories by sets $\emptyset \neq D \subseteq \mathbb{D}$. So, for any $n \in \mathbb{N}$, Z_n is a $\mathcal{P}(\mathbb{D})\setminus\{\emptyset\}$-valued random element, where the trigger event at T_n affects the LoBs $i \in Z_n$.

Notation. Throughout this work, $Z \sim Z_1$ denotes a random element which is identical distributed as Z_1.

We assume that assumption $\mathbb{P}(Z = \emptyset) = 0$, i.e. every shock event leads to at least one insurance damage, which justifies the following abbreviation.

Notation. For notational convenience, we write $D \subset \mathbb{D}$ instead of $\emptyset \neq D \subseteq \mathbb{D}$. Notice that $D \subset \mathbb{D}$ is equal to $D \in \mathcal{P}(\mathbb{D}) \setminus \{\emptyset\}$.

In contrast to the thinning and shift model introduced in Bäuerle and Grübel [27], our claim arrival model has no time shift between the jump times $(T_n)_{n\in\mathbb{N}}$ of N and the claim arrival times. In this perspective, the model in Bäuerle and Grübel [27] is more general. Nevertheless, our model is more comprehensive concerning the randomness of the intensity.

Recall that $\{i \in Z_n\}$ is the set of all events where the LoBs $i \in Z_n$ are affected by the trigger event at T_n. Therefore, the *(multivariate) claim arrival process*, denoted by $(N^1, \dots, N^d) = (N^1_t, \dots, N^d_t)_{t\geq 0}$, is defined by

$$N^i_t := \sum_{n\in\mathbb{N}} \mathbb{1}_{\{T_n \leq t\}} \mathbb{1}_{\{i\in Z_n\}}, \quad t \geq 0, \quad i = 1, \dots, d.$$

So $(N^1, \dots, N^d)$ is a d-dimensional counting process, where N^i_t counts the number of claims of the ith LoB up to time t.

Remark 3.1. The dependency structure of the marginal counting processes is only given via the synchronicity of jump times. Notice further that $(N^1, \dots, N^d)$ can be seen as a special case of the multivariate claim arrival model introduced in Scherer and Selch [108] with (unobservable) constant Lévy subordinator Λ.

Notation. From now on, we set $\Phi := (T_n, Z_n)_{n\in\mathbb{N}}$. That is, Φ is a $\mathcal{P}(\mathbb{D})$-MPP, cf. Definition 2.64. We further write $\mathfrak{F}^{\bar{N}} = (\mathcal{F}^{\bar{N}}_t)_{t\geq 0}$ for the filtration generated by $(N^1, \dots, N^d)$, i.e. $\mathcal{F}^{\bar{N}}_t := \sigma(N^i_s : 0 \leq s \leq t, i = 1, \dots, d)$ for all $t \geq 0$.

[2] See beginning of Sec. 5.2.3 in Albrecher et al. [5].

[3] See page 64 in Grandell [67].

Using the MPP Φ, we can represent the claim arrival process $(N^1, \dots, N^d)$ by

$$N_t^i = \int_0^t \mathbb{1}_z(i)\,\Phi(\mathrm{d}t, \mathrm{d}z), \quad t \geq 0, \quad i = 1, \dots, d.$$

Representations of jump processes by integrals w.r.t. a random counting measure will be proven to be extremely useful for the solution approach of the optimization problem introduced in Section 3.7.

The insurer is able to observe the claim arrival times. In consequence, the arrival times of the trigger events $(T_n)_{n\in\mathbb{N}}$ and the sequence $(Z_n)_{n\in\mathbb{N}}$ are observable since, in the given setting, we can reconstruct the jump times $(T_n)_{n\in\mathbb{N}}$ of the trigger process as well as the $\mathcal{P}(\mathbb{D})$-valued sequence $(Z_n)_{n\in\mathbb{N}}$ as follows: Let $N^i = (T_n^i)_{n\in\mathbb{N}}$, $i \in \mathbb{D}$, be the claim arrival times of the ith LoB. Then the sequence $(T_n)_{n\in\mathbb{N}}$ of the arrival times of the trigger events is iterative given by

$$\begin{aligned} T_1 &= \min\{T_1^i : i = 1, \dots, d\}, \\ T_n &= \min\{T_m^i : T_m^i > T_{n-1}, m \in \mathbb{N}, i = 1, \dots, d\}, \quad n \geq 2. \end{aligned} \tag{3.1}$$

At a jump time T_n, we obtain the corresponding Z_n by checking which point processes N^i, $i = 1, \dots, d$, jumps at time T_n, i.e.

$$Z_n = \big\{i \in \mathbb{D} : N^i(\{T_n\}) > 0\big\}, \quad n \in \mathbb{N}. \tag{3.2}$$

Therefore, the filtration $\mathfrak{F}^\Phi = (\mathcal{F}_t^\Phi)_{t\geq 0}$ of $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$, see (2.6), carries the same information as the filtration $\mathfrak{F}^{\bar{N}}$ of $(N^1, \dots, N^d)$. That is, $\mathcal{F}_t^\Phi = \mathcal{F}_t^{\bar{N}}$ for all $t \geq 0$. In particular, the observable trigger arrival times $(T_n)_{n\in\mathbb{N}}$ of the mixed Poisson process are observable for the insurer. Hence they can be used by insurer to make inferences about the unknown background intensity.

Next we turn our attention to the interdependencies between the LoBs which are also unknown to the insurance company. To take this uncertainty about the interdependencies between the various kinds of insurance risks into account, we introduce the following notation.

Notation. We set

$$\alpha_D := \mathbb{P}(Z = D \,|\, \mathcal{F}_0), \quad D \subset \mathbb{D},$$

and

$$\bar{\alpha} := (\alpha_D)_{D\subset\mathbb{D}}.$$

Furthermore, we write Δ_k for the $(k-1)$-dimensional *probability simplex*, $k \geq 2$. That is,

$$\Delta_k := \{x = (x_1, \dots, x_k) \in [0,1]^k : x_1 + \ldots + x_k = 1\}.$$

We further use the symbol $\mathring{\Delta}_k$ to denote the interior of the probability simplex Δ_k.

That is, $\bar{\alpha}$ is an $\mathcal{F}_0$-measurable vector taking values in Δ_ℓ. The vector $\bar{\alpha}$ determines the conditional probability mass function of the distribution of Z conditioned on $\bar{\alpha}$. That is,

$$\mathbb{P}_{\bar{\alpha}}(Z \in B) = \sum_{D\in B} \alpha_D, \quad B \in \mathcal{P}(\mathcal{P}(\mathbb{D})).$$

The interdependencies between the LoBs are fully determined by $\bar{\alpha}$. We call the components of $\bar{\alpha}$ *thinning probabilities* since they thin the trigger arrival times. So $(N^1, \dots, N^d)$ is a multivariate counting process with a dependency structure that depends only on the choice of thinning probabilities $\bar{\alpha}$. The uncertainty about the interdependencies between the insurance lines is incorporate by the randomness of $\bar{\alpha}$. So we have also a two-step stochastic phenomenon for the claim arrival dependencies. First, the realisation of the thinning probability vector $\bar{\alpha}$ is chosen, which is not observable for the insurer. Secondly, the affected LoBs are generated with the selection thinning distribution. Notice that we obtain information about $\bar{\alpha}$ through the observable sequence $(Z_n)_{n\in\mathbb{N}}$. Besides, we make the following assumptions.

Assumption 3.2. We assume that $(Z_n)_{n\in\mathbb{N}}$ is a sequence of conditional iid random elements given $\bar{\alpha}$ and takes values in $(\mathcal{P}(\mathbb{D}), \mathcal{P}(\mathcal{P}(\mathbb{D})))$ with $\mathbb{P}(Z_1 = \emptyset) = 0$. Moreover, we suppose that the sequences $(T_n)_{n\in\mathbb{N}}$ and $(Z_n)_{n\in\mathbb{N}}$ are independent.

After we have chosen a model for the claim arrival times, we consider the sizes of the insurance claims at the arrival times to obtain the aggregated claim amount process.

Claim sizes. Beside the unobservable background intensity and thinning probabilities, there are more restrictions on the available information to the insurance company. The insurer faces also uncertainty about the claim size distribution. This is taken into account by the following modelling. Let $\{F_\vartheta : \vartheta \in \Theta\}$, $\Theta \subset \mathbb{R}^n$, be a family of distributions on $(0,\infty)^d$, where ϑ in unknown. We view ϑ as a random element taking values in Θ, i.e. we have a parametric Bayesian model for the insurance losses. Moreover, we suppose that F_ϑ is absolutely continuous with density f_ϑ for every given ϑ, where we use the following convention.

Convention. We denote by ϑ both the random element and a realisation of the random element. A similar convention applies for all subsequent definition involving distributions or densities.

The *claim sizes* are given by a d-dimensional sequence $(Y_n)_{n\in\mathbb{N}}$ with $Y_n = (Y_n^1, \ldots, Y_n^d)$ of $(0,\infty)^d$-valued random variables with distribution F_ϑ. It is worth to note that the claims sizes from various LoBs can be dependent.

Assumption 3.3. We assume that $Y_1, Y_2, \ldots$ are conditional independent and identically distributed according to F_ϑ given ϑ. Furthermore, $(Y_n)_{n\in\mathbb{N}}$ is supposed to be independent of the sequences $(T_n)_{n\in\mathbb{N}}$ and $(Z_n)_{n\in\mathbb{N}}$.

Remark 3.4. Due to the assumption above, $Y_1, Y_2, \ldots$ are independent of $(N^1, \ldots, N^d)$.

With the knowledge about the claim sizes, we now move on to develop the aggregated claim amount process.

Representations of the aggregated claim amount process. The sum of the claim sizes of all d insurance classes which appear at the arrival times of the multivariate claim arrival process $(N^1, \ldots, N^d)$ up to time t, gives the aggregated claim amount process. So the *aggregated claim amount process* or *total claim amount process*, denoted by $S = (S_t)_{t\geq 0}$, is given by

$$S_t = \sum_{i=1}^{d} \sum_{n\in\mathbb{N}} Y_n^i \, \mathbb{1}_{\{T_n \leq t\}} \, \mathbb{1}_{\{i\in Z_n\}}, \quad t \geq 0.$$

Notice that Y_n^i does not describe the nth claim of the ith LoB since some of the components of Y_n will be "deleted" if $Z_n \neq \mathbb{D}$. For this reason, $(Z_n)_{n\in\mathbb{N}}$ is also referred to as the *thinning sequence*.

An alternative representation of the aggregated claim amount process can be given with the help of the following marked point process.

Notation. From now on, we set $\Psi := (T_n, (Y_n, Z_n))_{n\in\mathbb{N}}$. That is, Ψ is the $(0,\infty)^d \times \mathcal{P}(\mathbb{D})$-MPP which contains the information of the claim arrival times, the thinning sequence and the claim sizes. To shorten notation, we further set $E^d := (0,\infty)^d \times \mathcal{P}(\mathbb{D})$ and $\mathcal{E}^d := \mathcal{B}((0,\infty)^d) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D}))$.

Using the introduced MPP Ψ, it holds

$$S_t = \int_0^t \int_{E^d} \sum_{i=1}^d y_i \,\mathbb{1}_z(i)\, \Psi(\mathrm{d}s, \mathrm{d}(y,z)), \quad t \ge 0. \tag{3.3}$$

It should be noted that the aggregated claim amount process S is observable for the insurance company and thus the natural filtration of Ψ, denoted by $\mathfrak{F}^\Psi$, is known by the insurer.

The simulation of the surplus process is a two-step procedure. The first step is to simulate the realizations of the parameters Λ, ϑ and $\bar{\alpha}$, which is not observable for the insurer, and secondly, the observable surplus process is generated with theses parameters. We demonstrate the construction of the surplus process by an illustration for $d = 2$. Table 3.1 displays numerical values for the first four shock event under the assumptions that realization of the background intensity is 2, of the thinning probabilities is $(1/3, 1/3, 1/3)$ and of the claim size distribution is the convolution of two independent exponential distributions with parameter 1. These values are graphically illustrated in Figure 3.1.

n	T_n	Z_n	Y_n^1	Y_n^2
1	0.134	$\{1\}$	2.561	0.134
2	0.761	$\{1,2\}$	1.716	2.051
3	1.212	$\{2\}$	0.455	0.680
4	1.510	$\{1\}$	0.583	0.963

Table 3.1: Numerical values for the Figure 3.1.

We conclude this paragraph with a further comment on the introduced model for aggregated losses.

Remark 3.5. (i) The first part of the remark is devoted to the mentioned literatures in the introduction that consider optimization problems concerning insurance companies with several LoBs. For this purpose, let us suppose that $\Lambda \equiv \lambda$ is deterministic and we define the process $N^D = (N_t^D)_{t\ge 0}$, $D \subset \mathbb{D}$, by

$$N_t^D := \sum_{n\in\mathbb{N}} \mathbb{1}_{\{T_n \le t\}} \mathbb{1}_{\{Z_n = D\}}, \quad t \ge 0.$$

Then

$$N_t^i = \sum_{D \ni i} N_t^D, \quad t \ge 0,$$

and, in the case $d = 2$ of two LoBs, the aggregated claim amount process can be

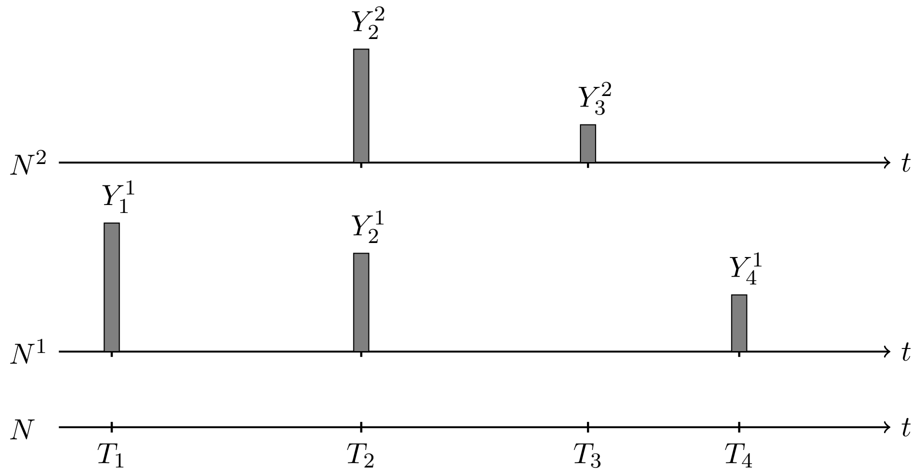

Figure 3.1: An example combining trigger events, thinning and claim sizes for the values of Table 3.1.

written as

$$S_t = \sum_{i=1}^{2} \sum_{j=1}^{N_t^{\{i\}}+N_t^{\{1,2\}}} Y_j^i, \quad t \geq 0.$$

For the aggregated claim amount process with common shock dependency, Yuen et al. [121], Liang and Yuen [86] and Han et al. [68] discuss optimal proportional reinsurance problems under various optimization criterion.

(ii) Due to the fact that we are not considering the aggregated claims of every LoB, S given by (3.3) can be interpreted alternatively as the aggregated claim amount process of an heterogeneous insurance portfolio, where the random elements Z_n yield the information of which type the claim size distribution of the claim at time T_n is. By changing $\{Z_n \in D\}$ to $\{Z_n = i\}$ in (3.3), we obtain a total claim amount process for an inhomogeneous portfolio with d kinds of claim size distribution.

Prior and posterior distributions. According to the explanation above, we have three unknown parameters Λ, $\bar{\alpha}$ and ϑ, where the uncertainty about these parameters is taken into account by modelling these parameters as $\mathcal{F}_0$-measurable random elements. We suppose that the insurance company has a prior belief about these parameters in form of distributions, which are prior distributions[4] from a Bayesian statistical point of view.

Notation. We denote by Π_Λ the prior distribution of Λ, by $\Pi_{\bar{\alpha}}$ the prior distribution of $\bar{\alpha}$ and by Π_ϑ the prior distribution of ϑ.

In general, Π_Λ is a probability measure on $(0, \infty)$, Π_α on Δ_ℓ and Π_ϑ on Θ. These prior distributions could be obtained by estimation based on already existing data using standard estimation procedures (cf. e.g. Czado and Schmidt [46]) or by expert knowledge.

The prior knowledge can be updated with the help of observed claim arrival times and claim sizes. Hence we can calculate posterior distributions[5] at the claim arrival

[4]The distribution of a parameter before observing any data is said to be the *prior distribution* of this parameter, compare DeGroot and Schervish [48, Chap. 7.2].

[5]The *posterior distribution* of a parameter is the conditional distribution of the parameter given observed data, compare DeGroot and Schervish [48, Chap. 7.2].

times by using the information at disposal, where the Bayes rule tells us how to compute them. For example, if we observe the claim sizes $\bar{Y}_n = \bar{y}_n$, where $\bar{Y}_n := (Y_1, \dots, Y_n)$ and $\bar{y}_n := (y_1, \dots, y_n)$, then, according to Klugman et al. [78, Eq. (2.27)], the posterior distribution of ϑ conditioned on $\bar{Y}_n = \bar{y}_n$ is given by

$$\Pi_{\vartheta|\bar{Y}_n=\bar{y}_n}(\mathrm{d}\vartheta) = \frac{L_\vartheta(\bar{y}_n)\,\Pi_\vartheta(\mathrm{d}\vartheta)}{\int_\Theta L_\vartheta(\bar{y}_n)\,\Pi_\vartheta(\mathrm{d}\vartheta)},$$

where $L_\vartheta(\bar{y}_n)$ denotes the likelihood function, i.e.

$$L_\vartheta(\bar{y}_n) = \prod_{i=1}^{n} f_\vartheta(y_i)$$

because of the iid assumption of insurance losses. That is, the posterior distribution is proportional to the Likelihood function times the prior. The mean of the posterior distribution is the *Bayes estimator* which is a minimum mean square error estimator, compare Klugman et al. [78, Thm. 2.17].

We will specify the prior distributions (and thus the posterior distributions) in the following chapters to solve the optimization problem stated in Section 3.7.

3.2 Financial market model

The difference between the aggregated premium income and the aggregated claims is the total wealth or the surplus of the insurance company. This surplus will be invested by the insurer into a financial market, which will be modelled as the classical Black-Scholes market, see e.g. Delbaen and Schachermayer [50, Sec. 4.4]. So it is supposed that there exists one risk-free asset and one risky asset. The price process of the *risk-free asset*, denoted by $B = (B_t)_{t\geq 0}$, is given by

$$\mathrm{d}B_t = rB_t\,\mathrm{d}t, \quad B_0 = 1,$$

where $r \in \mathbb{R}$ denotes the *risk-free interest rate*. That is, $B_t = e^{rt}$ for all $t \geq 0$. The price process of the *risky asset*, denoted by $P = (P_t)_{t\geq 0}$, is given by

$$\mathrm{d}P_t = \mu P_t\,\mathrm{d}t + \sigma P_t\,\mathrm{d}W_t, \quad P_0 = 1,$$

where $\mu \in \mathbb{R}$ and $\sigma > 0$ are constants describing the drift and volatility of the risky asset, respectively, and $W = (W_t)_{t\geq 0}$ is a standard Brownian motion. Therefore

$$P_t = \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma\,W_t\right\}, \quad t \geq 0.$$

Notation. We denote by $\mathfrak{F}^W = (\mathcal{F}_t^W)_{t\geq 0}$ the augmented Brownian filtration of W.

Notice that the σ-algebra generated by the price process of the risk-free asset B is the trivial σ-algebra since the process B is deterministic. So $\mathfrak{F}^W$ represents all the available information about the financial market. We assume that the insurance company can observe the asset prices on the market which means that $\mathfrak{F}^W$ is observable for the insurer. This filtration forms together with the natural filtration of Ψ the observable filtration.

Notation. Throughout this work, $\mathfrak{G} = (\mathcal{G}_t)_{t\geq 0}$ denotes the observable filtration of the insurer which is given by

$$\mathcal{G}_t = \mathcal{F}_t^W \vee \mathcal{F}_t^\Psi, \quad t \geq 0,$$

where $\mathcal{G}_\infty \subsetneq \mathcal{F}_0$. That implies that $\mathcal{G}_t$ provides all the information at disposal of the insurance company up to time t and that Λ, $\bar{\alpha}$ and ϑ are not observable for the insurer at any time.

The later solution procedure requires the following independence assumption of the financial market and the insurance market.

Assumption 3.6. We assume that the Brownian motion W is independent of $(T_n)_{n\in\mathbb{N}}$, $(Y_n)_{n\in\mathbb{N}}$ and $(Z_n)_{n\in\mathbb{N}}$.

Let us conclude this chapter with a brief discussion regarding the simplicity of the introduced financial market.

Remark 3.7. The given financial market is a classical Black-Scholes model with one risky asset. By the Markowitz theory, it is efficient to invest only in a risk-free asset and a particular fund of risky assets, compare Markowitz [91]. In our model, the risky asset can be considered as this particular fund of risky assets. However, we will see that the problem of finding an optimal investments strategy and an optimal strategy for the reinsurance can be separated into two independent problems. Therefore, we can expect that the results may be generalized to more general financial market models, where the corresponding optimal investment problem, maximization exponential utility, is already solved. However, we will focus on choosing an optimal reinsurance strategy and therefore we keep the financial market model simple.

3.3 Investment strategy

We assume that the wealth of the insurance company is invested into the previous described financial market. The insurer can choose the amount of its surplus that is invested at time t into the risky asset P, where we also allow short-sell by the insurer which is represented by a negative volume put into the risky asset. In addition, the insurance company is permit to lend and borrow an infinitesimal amount of money, respectively. That means, the invested capital into the risk free asset can take every value in $\mathbb{R}$. These assumptions are reflected in the following definition of an investment strategy together with some technical conditions.

Definition 3.8. An *investment strategy*, denoted by $\xi = (\xi_t)_{t\geq 0}$, is an $\mathbb{R}$-valued, càdlàg and $\mathfrak{G}$-predictable process with

$$\int_0^t |\xi_s|^2 \,\mathrm{d}s < \infty \quad \mathbb{P}\text{-a.s.} \quad \text{for all } t \geq 0. \tag{3.4}$$

The value ξ_t is the capital put into the risky asset at time t, so $X_t - \xi_t$ stands for the amount of money invested into the risk-free asset at time t.

Remark 3.9. On account of the condition (3.4) the integral $\int_0^t \xi_s\,\sigma\,\mathrm{d}W_s$ in (3.7) is well-defined. It should be noted that (3.7) is also well-defined with the weaker requirement of the progressive measurability of ξ since ξ is only integrated w.r.t. continuous processes. However, the investment strategy might depend on claim arrival times and sizes, which requires the assumed predictability.

Besides allocating the surplus to the two available assets, the insurance company is allowed to share risk with another company.

3.4 Reinsurance strategy

The premium volume is often too small to carry the complete risk, especially the risk resulting from natural catastrophes,[6] which requires the need for risk sharing through a reinsurance contract. There are different forms of reinsurances. A survey of reinsurance types can be found in Albrecher et al. [5, Ch. 2], in which they point out that the proportional reinsurance treaties (quota-share reinsurance) are popular in almost all insurance businesses by reasons of its conceptual and administrative frugality as well as its avoidance of the moral hazard of frowzy claim settlement proceedings.

The insurance company which cedes risk to another insurance company is called the *first-line insurer* or *cedent*[7]. We assume that the first-line insurer has the possibility to take a proportional reinsurance. Therefore, the *part of the insurance claims paid by the insurer*, denoted by $h(b, y)$, satisfies

$$h(b, y) = b \cdot y \tag{3.5}$$

with *retention level* $b \in [0, 1]$ and *insurance claim* $y \in (0, \infty)$. For example, if the insurer chooses a retention level of 0.8 at time t, then the reinsurance company pays 20 percent of the claim y at time t. Here we suppose that the insurer can continuously purchase a reinsurance contract that allows to reinsure a fraction of its claims with retention level $b_t \in [0, 1]$ at every time t. That is, we have a proportional reinsurance depending on time, in which corresponding process satisfies the following conditions.

Definition 3.10. A *reinsurance strategy*, denoted by $b = (b_t)_{t \geq 0}$, is a $[0, 1]$-valued, càdlàg and $\mathfrak{G}$-predictable process.

Remark 3.11. The assumption of predictability implies the reinsurance strategy is fixed in advance. Without this assumption, the insurance company could choose a retention level of zero at time t if there is a claim at time t and otherwise a retention level of one.

Note that the reinsurance strategy is not chosen separately for different LoBs, which means we have only one strategy for the entire insurance company. For this reason the reinsurance strategy is an univariate process. A conceivable supposition would be to suppose that the insurer can choose a reinsurance strategy for each LoB separately. In this case, the solution of an optimal reinsurance strategy of the later considered control problem (see Section 3.7) becomes much more complicated, compare Section 4.10.

3.5 Reinsurance premium model

Of course, sharing risk by ceding proportions of claims to a reinsurer reduces the premium income of the first-line insurer. To discuss this reduction in detail, we first assume that the policyholder's payments to the insurance company are modelled by a fixed *premium (income) rate* $c = (1 + \eta)\kappa$ with a safety loading $\eta > 0$ and a fixed constant $\kappa > 0$, which means that the premiums are calculated by the expected value principle.[8] If the insurer chooses retention levels less than one, then the insurer has to pay premiums to the reinsurer. The *part of the premium rate left to the insurance company* at retention level $b \in [0, 1]$, denoted by $c(b)$, is $c(b) = c - \delta(b)$, where $\delta(b)$ denotes the *reinsurance premium rate*. We say $c(b)$ is the *net income rate*. In the case of no reinsurance (retention level of

[6] See beginning of Sec. 1.7 in Schmidli [112].

[7] See p. 228 in Schmidli [111].

[8] The expected value principle, as well as other principles, are discussed e.g. in Schmidli [112, Sec. 1.10].

1), the net income rate is $c(1) = c$. Moreover, the net income rate $c(b)$ should increase in b, which is fulfilled by setting $\delta(b) := (1-b)(1+\theta)\kappa$ with $\theta > \eta$ which represents the safety loading of the reinsurer. Therefore

$$c(b) = (1+\eta)\kappa - (1-b)(1+\theta)\kappa = (\eta-\theta)\kappa + (1+\theta)\kappa\, b, \tag{3.6}$$

where $\eta - \theta < 0$. This reinsurance premium model is used e.g. in Zhu and Shi [124].

Due to the assumption that the proportional risk load of the reinsurer is greater than that of the cedent, full reinsurance leads to a strictly negative income since $c(0) = (\eta-\theta)\kappa{<}0$ which is a property proposed by Schmidli [111, p. 21]. Notice that other authors suppose that the net income rate is always non-negative, compare e.g. Schäl [109, p. 191].

We are now in the position to describe the wealth of the insurer subscribing a proportional reinsurance contract and investing its surplus in the introduced financial market.

3.6 The surplus process

We restrict ourselves to self-financing strategies in a finite time horizon. So, from now on, we fix some *terminal time* $T > 0$ and we only consider those strategies by which the insurer only invests the wealth obtained from the core business (covering claims in exchange for premiums and reinvesting those premiums into the financial market) and neither adds wealth from other businesses nor distributes part of the profit. To be more precise, the *surplus* of the considered insurance company, denoted by $X^{\xi,b} = (X_t^{\xi,b})_{t\in[0,T]}$, for a self-financing investment and reinsurance strategy is given by

$$\mathrm{d}X_t^{\xi,b} = \frac{X_t^{\xi,b} - \xi_t}{B_t}\,\mathrm{d}B_t + \frac{\xi_t}{P_t}\,\mathrm{d}P_t + c(b_t)\,\mathrm{d}t - b_t\,\mathrm{d}S_t, \quad X_0^{\xi,b} = x_0,$$

where $X_0^{\xi,b} = x_0 > 0$ is the *initial capital* of the insurance company. It is worth to note that $(X_t^{\xi,b} - \xi_t)/B_t$ gives the number of risk-free assets and ξ_t/P_t the number of risky assets held by the insurance company at time t. Therefore, the dynamic of the surplus process $X^{\xi,b}$ can be interpreted as follows: the insurer's current reserve is the initial capital plus the aggregated gain/loss by the investments in risky and risk-free asset plus the net premium income minus the aggregated insurance claims left to the insurer. Notice further that the surplus could be negative, which in practice means that the loan amount of the insurance company is greater than the value of all assets of the insurer.

By the upper indices ξ and b it is taken into account that the surplus process is controlled by the investment and reinsurance strategy. In the following we deal only with admissible strategies defined below.

Definition 3.12. A pair $(\xi, b) = (\xi_t, b_t)_{t\geq 0}$ of an investment strategy $\xi = (\xi_t)_{t\geq 0}$ and a reinsurance strategy $b = (b_t)_{t\geq 0}$ is called an *investment-reinsurance strategy.* For any $t \in [0, T]$, the set of all *admissible investment-reinsurance strategies* in $[t, T]$ is given by

$$\mathcal{U}[t,T] := \big\{(\xi,b) : (\xi,b) = (\xi_s, b_s)_{s\in[t,T]} \text{ is a self-financing investment-reinsurance strategy in } [t,T]\big\}.$$

The set $\mathbb{R} \times [0,1]$ in which the admissible strategies take values is said to be the *control set.*

Using the dynamics B and P of the price process of the risky asset and risk-free asset, respectively, the surplus process $X^{\xi,b} = (X_t^{\xi,b})_{t\in[0,T]}$ under an admissible investment-reinsurance strategy $(\xi, b) \in \mathcal{U}[0,T]$ holds

$$\begin{aligned} \mathrm{d}X_t^{\xi,b} &= (X_t^{\xi,b} - \xi_t) r \,\mathrm{d}t + \xi_t(\mu \,\mathrm{d}t + \sigma \,\mathrm{d}W_t) + c(b_t)\,\mathrm{d}t - b_t \,\mathrm{d}S_t \\ &= \Big(rX_t^{\xi,b} + (\mu - r)\xi_t + c(b_t) \Big) \,\mathrm{d}t + \xi_t \sigma \,\mathrm{d}W_t - b_t \,\mathrm{d}S_t. \end{aligned}$$

Remark 3.13. Recall that the investment strategy ξ gives the absolute value of the wealth invested into the risky asset. As Desmettre [51] pointed out in the introduction of Chapter 2, it is common to optimize the amount of money in an exponential utility set-up. When we optimize the proportion of the wealth put into the risky asset, denoted by π, then the investment strategies π is proportional to $1/X^{\pi,b}$. Therefore, it is possible that an optimal strategy tends to infinity, since the surplus process $X^{\pi,b}$ can attain zero.

An alternative representation of the surplus process with the help of a random measure will be proved to be useful. By using the $(0,\infty)^d \times \mathcal{P}(\mathbb{D})$-MPP $\Psi = (T_n, (Y_n, Z_n))_{n\in\mathbb{N}}$ introduced on page 35, the dynamic of the surplus can be written as

$$\begin{aligned} \mathrm{d}X_t^{\xi,b} = &\Big(rX_s^{\xi,b} + (\mu - r)\xi_s + c(b_s) \Big) \,\mathrm{d}t + \xi_s \sigma \mathrm{d}W_s \\ &- \int_{E^d} b_t \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \Psi(\mathrm{d}t, \mathrm{d}(y,z)), \quad t \in [0,T]. \end{aligned} \tag{3.7}$$

Notice that the surplus process $X^{\xi,b}$ is $\mathfrak{G}$-adapted, i.e. observable for the insurer. Next we consider further properties of the surplus process which will be used in the following procedure.

Proposition 3.14. *The SDE* (3.7) *has a unique strong solution, which is given by*

$$\begin{aligned} X_t^{\xi,b} = x_0 e^{rt} &+ \int_0^t e^{r(t-s)} \Big((\mu - r)\xi_s + c(b_s) \Big) \mathrm{d}s + \int_0^t e^{r(t-s)} \xi_s \sigma \,\mathrm{d}W_s \\ &- \int_0^t \int_{E^d} e^{r(t-s)} b_s \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \Psi(\mathrm{d}s, \mathrm{d}(y,z)), \quad t \in [0,T]. \end{aligned}$$

Proof. Fix $t \in [0,T]$. We will use the general stochastic exponential (see Protter [104, p. 328]) to derive the asserted solution. For this purpose, we observe that the SDE (3.7) can be written as

$$X_t^{\xi,b} = H_t + \int_0^t X_s^{\xi,b} \,\mathrm{d}Z_s, \tag{3.8}$$

where

$$\begin{aligned} H_t &:= x_0 + \int_0^t \big((\mu - r)\xi_s + c(b_s) \big) \mathrm{d}s + \int_0^t \xi_s \sigma \mathrm{d}W_s - \int_0^t \int_{E^d} b_s \sum_{i=1}^d y_i \mathbb{1}_z(i) \Psi(\mathrm{d}s, \mathrm{d}(y,z)), \\ Z_t &:= r\,t. \end{aligned}$$

Notice that $(H_t)_{t\geq 0}$ is a $\mathfrak{G}$-semimartingale and thus (3.7) admits a unique strong solution, which is a $\mathfrak{G}$-semimartingale according to Protter [104, Thm. V.7]. By Protter [104, Thm. V.52], the unique solution of (3.8) is given by

$$X_t^{\xi,b} = \mathcal{E}(Z)_t \left(H_0 + \int_0^t \mathcal{E}(Z)_s^{-1} \,\mathrm{d}(H_s - [H,Z]_s) \right), \tag{3.9}$$

where $\mathcal{E}(Z)$ is the stochastic exponential of Z, i.e. $\mathcal{E}(Z)_t = \exp\{Z_t - \frac{1}{2}[Z]_t\} = e^{rt}$, cf. Protter [104, Thm. II.37]. Furthermore, since Z is a continuous FV process, we have $[H, Z]_t = 0$, which implies

$$X_t^{\xi,b} = e^{rt}\Bigg(x_0 + \int_0^t e^{-rs}\left((\mu - r)\xi_s + c(b_s)\right)\mathrm{d}s + \int_0^t e^{-rs}\xi_s\sigma\,\mathrm{d}W_s \\ - \int_0^t\int_{E^d} e^{-rs}b_s\sum_{i=1}^d y_i\mathbb{1}_z(i)\,\Psi(\mathrm{d}s, \mathrm{d}(y,z))\Bigg).$$

This yields the assertion. □

Proposition 3.15. *Let $X^{\xi,b} = (X_t^{\xi,b})_{t\geq 0}$ be the process given by (3.7). The continuous part $(X^{\xi,b})^c = ((X^{\xi,b})^c_t)_{t\geq 0}$ of $X^{\xi,b}$ is given by*

$$\mathrm{d}(X^{\xi,b})^c_t = \left(rX_t^{\xi,b} + (\mu - r)\xi_t + c(b_t)\right)\mathrm{d}t + \xi_t\sigma\,\mathrm{d}W_t, \quad t \geq 0.$$

The continuous part $[X^{\xi,b}]^c = ([X^{\xi,b}]^c_t)_{t\geq 0}$ of the quadratic variation of $X^{\xi,b}$ is given by

$$\mathrm{d}[X^{\xi,b}]^c_t = \xi_t^2\sigma^2\,\mathrm{d}t, \quad t \geq 0.$$

The process $X^{\xi,b}$ jumps at the times $(T_n)_{n\in\mathbb{N}}$ and satisfies

$$\Delta X^{\xi,b}_{T_n} = -b_{T_n}\sum_{i=1}^d Y_n^i\mathbb{1}_{\{i\in Z_n\}}, \quad n \in \mathbb{N}.$$

Proof. The statements follow immediately from Equation (3.7). □

The surplus process is the main object in the following formulated optimization problem.

3.7 Optimal investment and reinsurance problem under partial information

Clearly, the insurance company is interested in an optimal investment-reinsurance strategy. But there are various optimality criteria to specify optimization of proportional reinsurance and investment strategies. We consider the expected utility of wealth at the terminal time T as a criterion. It is therefore assumed an exponential utility function of the insurer $U : \mathbb{R} \to \mathbb{R}$ with

$$U(x) = -e^{-\alpha x}, \tag{3.10}$$

where the parameter $\alpha > 0$ measures the *degree of risk aversion.*

Remark 3.16. Recall that the surplus is allowed to be negative. In classical ruin theory, an insurance company is bankrupt if the surplus process drops below zero, which is not realistic. Indeed, the loan amount of an insurance company can be greater than the value of all assets of the insurance company. The possible negativity is one reason why we use the exponential utility function because the utility function has to be defined for positive and negative values. This condition rules out the logarithm and power utility functions. Another reason is that the exponential utility function given by (3.10) has a constant absolute risk aversion (CARA):

$$-\frac{U''(x)}{U'(x)} = \frac{\alpha^2\, e^{-\alpha x}}{\alpha\, e^{-\alpha x}} = \alpha.$$

Such a property is important in actuarial mathematics since utility functions with CARA are the only functions among which the so-called zero-utility principle yields a fair premium that does not depend on the surplus level, compare Gerber [63, p. 68].

Before the dynamical version of optimization problem is formulated, we illustrate the feedback control of the surplus process $X^{\xi,b}$ in the block diagram displayed in Figure 3.2[9].

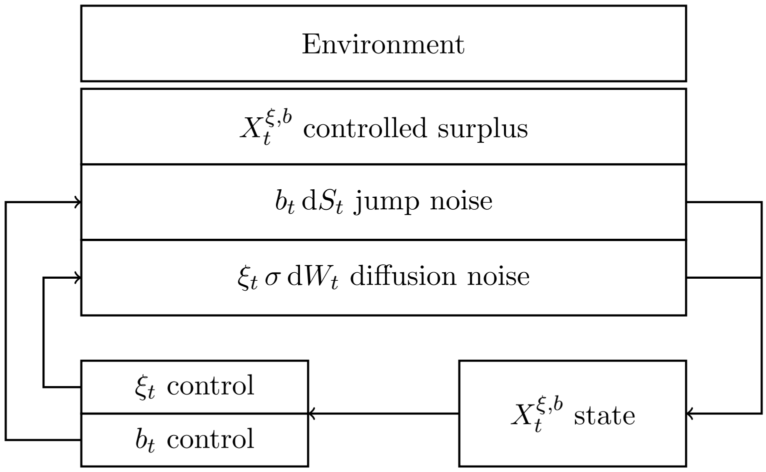

Figure 3.2: The surplus process under feedback control in infinitesimal time intervals.

Next, we are going to formulate the dynamical optimization problem. We define, for any $(t,x) \in [0,T] \times \mathbb{R}$ and $(\xi,b) \in \mathcal{U}[t,T]$,

$$\begin{aligned} \bar{V}^{\xi,b}(t,x) &:= \mathbb{E}^{t,x}\big[U(X_T^{\xi,b}) \,|\, \mathcal{G}_t\big], \\ \bar{V}(t,x) &:= \sup_{(\xi,b)\in\mathcal{U}[t,T]} \bar{V}^{\xi,b}(t,x), \end{aligned} \tag{P}$$

where $\mathbb{E}^{t,x}$ denotes the expectation conditioned $X_t^{\xi,b} = x$. As already mentioned, the insurer wants to choose an optimal strategy. The argument of the supremum, when it exist within the control domain $\mathcal{U}[t,T]$, is the optimal control. That is, assuming the surplus at the time point $t \in [0,T)$ is $x \in \mathbb{R}$, the insurance company is interested in an investment-reinsurance strategy $(\xi^\star, b^\star) \in \mathcal{U}[t,T]$ such that

$$\bar{V}(t,x) = \bar{V}^{\xi^\star,b^\star}(t,x).$$

So our aim is to determine

$$(\xi^\star, b^\star) = \operatorname*{argsup}_{(\xi,b)\in\mathcal{U}[t,T]} \bar{V}^{\xi,b}(t,x).$$

Such a strategy is said to be *optimal*. Note that at time $t = 0$, the optimal strategy is given by

$$(\xi^\star, b^\star) = \operatorname*{argsup}_{(\xi,b)\in\mathcal{U}[0,T]} \bar{V}^{\xi,b}(0,x_0).$$

[9]The figure is adapted from Figure 6.1 in Hanson [69].

An important point to note here is that it makes a difference to solve the problem (P) at time zero, where $X_0^{\xi,b}$ is known, or at time $t > 0$, where $X_t^{\xi,b}$ is known. The reason is that at time zero only the prior distributions of the unobservable parameters Λ, $\bar{\alpha}$ and F are given. In contrast, at time $t > 0$, the observations of claim arrival times and claim sizes (which are included in $\mathcal{G}_t$) yield additional information about the unobservable parameters which can be taken into account for the determination of the optimal strategies. That is, the optimization problem (P) is different for various time points (due to the partial information). In consequence, the dynamic programming method can not be applied directly since the idea of this method is to derive relations between the optimization problems for different initial states and deduce the optimal solution from this relation by solving pointwise optimization problems. That is, the optimal control obtained by the dynamic programming principle cannot incorporate information from the past. To still use the dynamic programming approach, we need to extend the state process by a Markov process which represents the information at disposal. This step is referred as the reduction of the incomplete information problem (P), which requests a characterization of the conditional distributions of Λ, $\bar{\alpha}$ and F using the available information. With regard to the block diagram given in Figure 3.2, we have to extend the state in the bottom right corner by a further one which is affected by the jump diffusion and provides the available information for the insurer about the known parameters. After the reduction step, we can try to solve the reduced control problem with the help from the dynamic programming principle. In the next chapters we address the problem (P) under different assumptions of the prior distribution for Λ, $\bar{\alpha}$ and F.

Chapter 4

Optimal investment and reinsurance with unknown dependency structure between the LoBs

We have introduced the control problem (P) in the last chapter, which is not directly solvable. In this chapter we investigate this problem under partial information focusing on the unknown interdependencies between the lines of business (LoBs), i.e. on the thinning probabilities $\bar{\alpha}$. So the background intensity and the claim size distribution are suppose to be observable for the insurer.

4.1 Setting

First of all, it should be mentioned that the framework from Chapter 3 is valid, in particular the Assumptions 3.2, 3.3 and 3.6 are in force. In addition, as indicated above, we suppose the prior distributions Π_Λ and Π_F are one-point distributions. That means, the background intensity $\Lambda \equiv \lambda$ for some $\lambda > 0$ and the claim size distribution F are observable parameters; only the thinning probabilities $\bar{\alpha}$ are still assumed to be unknown.

Prior distribution for the thinning probabilities. Recall that the prior distribution $\Pi_{\bar{\alpha}}$ of $\bar{\alpha}$ is defined on Δ_ℓ. This may lead to an infinite dimensional stochastic control problem in general. To avoid this, we discretize the probability simplex Δ_ℓ such that the prior distribution $\Pi_{\bar{\alpha}}$ is defined on a finite set.

Assumption 4.1. Let $m \in \mathbb{N}$ be a fixed number. We suppose that $\bar{\alpha} = (\alpha_D)_{D \subset \mathbb{D}}$ is an $\mathcal{F}_0$-measurable random vector taking values in the measure space $(A, \mathcal{A})$, where $A := \{a_1, \dots, a_m\}$ with $a_j = (a_j^D)_{D \subset \mathbb{D}} \in \mathring{\Delta}_\ell$, $j = 1, \dots, m$ and $\mathcal{A} := \mathcal{P}(A)$.

Remark 4.2. (i) The assumption that $A \subset \mathring{\Delta}_\ell$ (and not $A \subset \Delta_\ell$) is required for the filter equation, see Theorem 4.13.

(ii) The $\mathcal{F}_0$-measurability of $\bar{\alpha}$ is owed to the fact that the thinning mechanism does not change in time in this model. Furthermore, $\mathcal{F}_0$ contains information which is not available for the insurer. The knowledge of the insurer at time $t = 0$ about $\bar{\alpha}$ is $\mathcal{F}_0^{\bar{N}} = \{\emptyset, \Omega\}$. But the insurer makes use of expert knowledge about the interdependencies between the LoBs expressed by the prior distribution $\Pi_{\bar{\alpha}}$. Such expert knowledge could be the awareness of high interdependency between the insurance classes "building storm damages" and "building flood damages".

Under the assumption above, the prior distribution $\Pi_{\bar{\alpha}}$ of $\bar{\alpha}$ is a probability measure on $(A, \mathcal{A})$. Hence, the prior distribution $\Pi_{\bar{\alpha}}$ is uniquely determined by the probability mass function

$$\pi_{\bar{\alpha}}(j) := \Pi_{\bar{\alpha}}(a_j) = \mathbb{P}(\bar{\alpha} = a_j), \quad j = 1, \dots, m.$$

Notation. We will use the notation $\bar{\pi} := (\pi_{\bar{\alpha}}(j))_{j=1,\dots,m}$ for the m-dimensional vector in Δ_m, which describes the probability mass function of $\Pi_{\bar{\alpha}}$. Throughout this work we switch between row and column vectors whenever it is typographically convenient.

Claim size distribution. Beside the requirements for the prior distributions, it is necessary to put some restrictions on the claims sizes. We have supposed that the common claim size distribution F is a fixed parameter in this chapter. Insurance claims from different LoBs are allowed to be dependent, but the exponential moments of the sum of losses across all lines must be finite.

Assumption 4.3. We assume that

$$M_F(z) := \mathbb{E}\bigg[\exp\Big\{z\sum_{i=1}^{d} Y_1^i\Big\}\bigg] = \int_{(0,\infty)^d} \exp\Big\{z\sum_{i=1}^{d} y_i\Big\}\, F(\mathrm{d}y) < \infty, \quad z \in \mathbb{R},$$

where $y := (y_1, \dots, y_d)$.

Remark 4.4. The assumption implies that the moment generating function for the marginal distributions of F exists. That means, the losses of each LoB has finite exponential moments. This has the consequence that all moments of the claim sizes exist, due to the relation $M_{Y_1^i}(z) := \mathbb{E}\big[e^{z\,Y_1^i}\big] = \sum_{n=0}^{\infty} \frac{z^n \mathbb{E}[(Y_1^i)^n]}{n!}$, $z \in \mathbb{R}$, cf. Feller [59, p. 285]. Another consequence concerns the variety of possible distributions. The assumption of existence of the moment generation function rules out a lot of distribution, in particular heavy tailed distributions, also the lognormal distributions; all moments of the lognormal distribution exist, but the moment generating function of the lognormal distribution does not exist. It is worth noting that the assumption is satisfied if claim amounts from different LoB are independent and the moment generating functions exist for every marginal loss distribution. A further condition, under which Assumption 4.3 is fulfilled, is boundedness of the claim sizes. This requirement is no restriction in practice since every insurance contract includes an insurance sum which is the maximum value the insurance pays for the insured damages.

Let us mention two consequences of the assumption above.

Lemma 4.5. *Let $z \in \mathbb{R}$ be an arbitrary constant. Then there exist constants $0 < C_1 < \infty$ and $0 < C_2 < \infty$ such that*

(i) $\mathbb{E}\big[Y_1^j \exp\big\{z\sum_{i=1}^{d} Y_1^i\big\}\big] \le C_1, \quad j \in \mathbb{D},$

(ii) $\mathbb{E}\Big[\exp\Big\{z\sum_{i=1}^{d}\sum_{k=1}^{N_t} Y_k^i\Big\}\Big] \le C_2, \quad t \in [0,T].$

Proof. To show statement (i) the Cauchy-Schwarz inequality comes to our assistance. For any $j \in \mathbb{D}$, we have

$$\begin{aligned}\mathbb{E}\bigg[Y_1^j \exp\Big\{z\sum_{i=1}^{d} Y_1^i\Big\}\bigg] &\le \sqrt{\mathbb{E}\big[\big(Y_1^j\big)^2\big]}\sqrt{\mathbb{E}\bigg[\exp\Big\{2z\sum_{i=1}^{d} Y_1^i\Big\}\bigg]} \\ &= \sqrt{\mathbb{E}\big[\big(Y_1^j\big)^2\big]}\sqrt{M_F(2z)} := C_1 < \infty,\end{aligned}$$

where the first expectation is finite according to Remark 4.4 and the finiteness of the second expectation follows from Assumption 4.3. To prove statement (ii), let us fix $t \in [0, T]$. According again to Assumption 4.3 as well as to Assumption 3.3 and the law of total expectation, we obtain

$$\begin{aligned}
\mathbb{E}\Big[\exp\Big\{z\sum_{i=1}^{d}\sum_{k=1}^{N_t} Y_k^i\Big\}\Big] &= \sum_{n=0}^{\infty}\mathbb{P}(N_t = n)\,\mathbb{E}\Big[\exp\Big\{z\sum_{k=1}^{N_t}\sum_{i=1}^{d} Y_k^i\Big\}\,|\,N_t = n\Big] \\
&= \sum_{n=0}^{\infty}\mathbb{P}(N_t = n)\,\prod_{k=1}^{n}\mathbb{E}\Big[\exp\Big\{z\sum_{i=1}^{d} Y_k^i\Big\}\,|\,N_t = n\Big] \\
&= \sum_{n=0}^{\infty}\frac{(\lambda\,t)^n}{n!}e^{-\lambda\,t}\Big(\mathbb{E}\Big[\exp\Big\{z\sum_{i=1}^{d} Y_1^i\Big\}\Big]\Big)^n \\
&= e^{-\lambda\,t}\sum_{n=0}^{\infty}\frac{\big(\lambda\,t\,\mathbb{E}\big[\exp\big\{z\sum_{i=1}^{d} Y_1^i\big\}\big]\big)^n}{n!} \\
&= \exp\Big\{\lambda\,t\Big(\mathbb{E}\Big[\exp\Big\{z\sum_{i=1}^{d} Y_1^i\Big\}\Big] - 1\Big)\Big\} \\
&\le \exp\{\lambda\,T\,M_F(z)\} =: C_2 < \infty,
\end{aligned}$$

which completes the proof. □

Our next target is to reduce the partial information control problem (P) within the introduced framework to one with an extended state process that describes the information at disposal about the unknown interdependencies between the LoBs. To obtain such a control problem, we need to infer the unobservable thinning probabilities $\bar{\alpha}$ by using the relevant observable information. This leads to a stochastic filter problem which will be studied in the following section.

4.2 Filtering and reduction

It has already been mentioned in the previous chapter that we are in a Bayesian setting. Therefore, the reduction problem is to determine the posterior distribution of $\bar{\alpha}$ given the information provided by $\mathcal{G}_t$. It is evident that the available information $\mathfrak{F}^W$ about the financial market contains no information about the interdependencies between the LoBs; only the filtration $\mathfrak{F}^{\bar{N}}$ can be used to make inferences about $\bar{\alpha}$.

The posterior distribution of $\bar{\alpha}$ given $\mathcal{F}_t^{\bar{N}}$ (or equivalent given $Z_1, \dots, Z_{N_t}$) provides all information about $\bar{\alpha}$ which is included in the observed information up to every t. We will see that a characterization of the posterior distribution of $\bar{\alpha}$ can be used to reduce the investment-reinsurance problem under partial information (P) into an equivalent one, which can be solve with the dynamic programming approach. Within the presented Bayesian framework, the posterior distribution of $\bar{\alpha}$ given $(z_1, \dots, z_n) \in \mathcal{P}(\mathbb{D})^n$ is described by

$$\mathbb{P}(\bar{\alpha} = a_j\,|\,z_1, \dots, z_n) = \frac{\prod_{i=1}^{n}\mathbb{P}_{a_j}(Z_i = z_i)\,\mathbb{P}(\bar{\alpha} = a_j)}{\int_A \prod_{i=1}^{n}\mathbb{P}_{\bar{\alpha}}(Z_i = z_i)\,\Pi_{\bar{\alpha}}(\mathrm{d}\bar{\alpha})} = \frac{\prod_{i=1}^{n} a_j^{z_i}\,\pi_{\bar{\alpha}}(j)}{\sum_{k=1}^{m}\prod_{i=1}^{n} a_k^{z_i}\,\pi_{\bar{\alpha}}(k)}$$

for every $j \in \{1, \dots, m\}$. That is,

$$\mathbb{P}(\bar{\alpha} = a_j \mid Z_1, \dots, Z_{N_t}) = \frac{\pi_{\bar{\alpha}}(j) \prod_{i=1}^{N_t} a_j^{Z_i}}{\sum_{k=1}^{m} \pi_{\bar{\alpha}}(k) \prod_{i=1}^{N_t} a_k^{Z_i}}, \quad j = 1, \dots, m. \tag{4.1}$$

However, it has been pointed out in Remark 2.93 that a representation of the posterior distribution by integral processes w.r.t. compensated random counting measures fit with the stochastic control approach. We obtain such a characterization by using filter results.

Classical filter results with multivariate point process observations[1] can not be applied to determine a filter for $\bar{\alpha}$ (see Brémaud [20, Thm. IV.T8]) since point processes $N^1, \dots, N^d$ have common jump times with probability greater than zero if $\alpha_D > 0$ for $D \subset \mathbb{D}$ with $|D| \geq 2$. Fortunately, we have seen in Section 3.1 that the natural filtration of $\Phi = (T_n, Z_n)_{n \in \mathbb{N}}$ carries the same information as the filtration of $(N^1, \dots, N^d)$, i.e. $\mathcal{F}_t^{\Phi} = \mathcal{F}_t^{\bar{N}}$ for all $t \geq 0$. This allows us to use filter results with marked point process observations. For the application of filter results, the usual condition needs to hold for the observed filtration. On this account, we introduce the following notation.

Notation. From now on, $(\Omega, \mathcal{F}_\infty^{\Phi}, \mathfrak{F}^{\Phi}, \mathbb{P})$ denotes the filtrated probability space which is modified as described in Remark 2.70 such that the usual conditions are satisfied.

To state a filter for $\bar{\alpha}$ given by the observed information $\mathfrak{F}^{\Phi}$, we need the local $\mathfrak{F}$- and $\mathfrak{F}^{\Phi}$-characteristic of Φ. To establish these characteristics we start with studying the process $(\Phi(t, D))_{t \geq 0}$, $D \subset \mathbb{D}$. By recalling Definition 2.73, we have

$$\Phi(t, D) = \sum_{n \in \mathbb{N}} \mathbb{1}_{\{T_n \leq t\}} \mathbb{1}_{\{Z_n = D\}}, \quad t \geq 0.$$

Due to the fact that α_D is a random variable, $(\Phi(t, D))_{t \geq 0}$ can not be a Poisson process; $(\Phi(t, D))_{t \geq 0}$ is a Poisson process for deterministic α_D according to the thinning property for Poisson processes, compare e.g. Last and Penrose [81, Cor. 5.9].) It is further known that doubly stochastic Poisson processes (Cox processes) are invariant under thinning (see Karr [74, Lemma 1.1]) and that an SPP is a DSPP if and only if it can be obtained by p-thinning for every $p \in (0, 1)$ (see Mecke [92, Satz 5.1]). This indicates that $(\Phi(t, D))_{t \geq 0}$ may be a doubly stochastic Poisson process.

Lemma 4.6. *For any $D \subset \mathbb{D}$, $(\Phi(t, D))_{t \geq 0}$ is an integrable $\mathfrak{F}$-DSPP with constant intensity $(\lambda\, \alpha_D)_{t \geq 0}$.*

Proof. Fix $D \subset \mathbb{D}$. The conditions (2.10) and (2.11) of Definition 2.83 of a DSPP are obviously satisfied by the $\mathcal{F}_0$-measurability of α_D. Therefore, it remains to show the condition (2.12). For this purpose, we set $X_n = \mathbb{1}_{\{Z_n = D\}}$ for all $n \in \mathbb{N}$. Hence, $\mathbb{P}(X_n = 1 \mid \mathcal{F}_0) = \alpha_D = 1 - \mathbb{P}(X_n = 0 \mid \mathcal{F}_0)$. That is, X_n is conditional Bernoulli distributed given $\mathcal{F}_0$ with parameter α_D and conditionally independent of $(N_t)_{t \geq 0}$ given $\mathcal{F}_0$. Therefore, $S_n := \sum_{i=1}^{n} X_n$ is conditional binominal distributed given $\mathcal{F}_0$ with parameter n and α_D. Thus, by the binomial theorem, for any $z \in [0, 1]$,

$$\mathbb{E}[z^{S_n} \mid \mathcal{F}_0] = \sum_{k=0}^{n} z^k \mathbb{P}(S_n = k \mid \mathcal{F}_0) = \sum_{k=0}^{n} z^k \binom{n}{k} \alpha_D^k (1 - \alpha_D)^{n-k} = \big(z\, \alpha_D + (1 - \alpha_D)\big)^n.$$

[1] Multivariate point process means here that the jump times of the marginal simple point processes are disjunct almost surely. Multivariate point processes can also be seen as special marked point processes, compare the definition given on pages 19–20 in Brémaud [20].

Hence the conditional probability generating function given $\mathcal{F}_0$ of S_{t-s}, where $S_t := S_{N_t}$, $t \geq 0$, holds for every $0 \leq s \leq t$ because of the independence of N_{t-s} and $(X_n)_{n\in\mathbb{N}}$ as well as the Poisson distribution of N_{t-s} with parameter $\lambda(t-s)$,

$$\begin{aligned}\mathbb{E}\big[z^{S_{N_{t-s}}} \,|\, \mathcal{F}_0\big] &= \sum_{n=0}^{\infty} \mathbb{E}[z^{S_n} \,|\, \mathcal{F}_0]\, \mathbb{P}(N_{t-s} = n \,|\, \mathcal{F}_0) \\ &= e^{-\lambda(t-s)} \sum_{n=0}^{\infty} \frac{\big((z\alpha_D + (1-\alpha_D))\lambda(t-s)\big)^n}{n!} = \exp\big\{\lambda\, \alpha_D (t-s)(z-1)\big\}\end{aligned}$$

for all $z \in [0,1]$. Since the distribution of a discrete random variable is uniquely determined through the probability generating function and the probability generating function of a Poisson distributed random variable X with parameter λ is given by $\exp\{\lambda(z-1)\}$, $z \in [0,1]$, we obtain from the equation above that S_{t-s} is conditional Poisson distributed given $\mathcal{F}_0$ with parameter $\lambda\, \alpha_D(t-s)$. Consequently, the conditional characteristic function given $\mathcal{F}_0$ of S_{t-s} is

$$\mathbb{E}\left[e^{iuS_{t-s}} \,|\, \mathcal{F}_0\right] = \exp\left\{(e^{iu}-1)\lambda\, \alpha_D(t-s)\right\}, \quad u \in \mathbb{R}.$$

An easy computation shows

$$S_t - S_s = \sum_{i=N_s+1}^{N_t} X_i = \sum_{n\in\mathbb{N}} \mathbb{1}_{\{T_n \leq t\}} \mathbb{1}_{\{Z_n = D\}} - \sum_{n\in\mathbb{N}} \mathbb{1}_{\{T_n \leq s\}} \mathbb{1}_{\{Z_n = D\}} = \Phi(t,D) - \Phi(s,D).$$

Furthermore, we have, for any $s,t \geq 0$ and $n \in \mathbb{N}_0$,

$$\begin{aligned}\mathbb{P}(S_{t+s} - S_t = n \,|\, \mathcal{F}_0) &= \sum_{i,j\in\mathbb{N}_0, j\geq i} \mathbb{P}\left(\sum_{\ell=i+1}^{j} X_\ell = n, N_s = i, N_{s+t} = j \,|\, \mathcal{F}_0\right) \\ &= \sum_{i,k\in\mathbb{N}_0} \mathbb{P}\left(\sum_{\ell=1}^{k} X_\ell = n \,|\, \mathcal{F}_0\right) \mathbb{P}(N_s = i \,|\, \mathcal{F}_0)\mathbb{P}(N_{s+t} - N_s = k \,|\, \mathcal{F}_0) \\ &= \sum_{k\in\mathbb{N}_0} \mathbb{P}\left(\sum_{\ell=1}^{k} X_\ell = n \,|\, \mathcal{F}_0\right) \mathbb{P}(N_t = k \,|\, \mathcal{F}_0) \sum_{i\in\mathbb{N}_0} \mathbb{P}(N_s = i \,|\, \mathcal{F}_0) \\ &= \sum_{k\in\mathbb{N}_0} \mathbb{P}\left(\sum_{\ell=1}^{k} X_\ell = n, N_t = k \,|\, \mathcal{F}_0\right) = \mathbb{P}\left(\sum_{\ell=1}^{N_t} X_\ell = n \,|\, \mathcal{F}_0\right) = \mathbb{P}(S_t = n \,|\, \mathcal{F}_0).\end{aligned}$$

That is, $(S_t)_{t\geq 0}$ has conditionally stationary increments given $\mathcal{F}_0$ (cf. e.g. Schmidt [113, p. 86] for the definition). Taking this into account, we obtain

$$\begin{aligned}\mathbb{E}\left[e^{iu(\Phi(t,D)-\Phi(s,D))} \,|\, \mathcal{F}_s\right] &= \mathbb{E}\left[e^{iu(S_t - S_s)} \,|\, \mathcal{F}_s\right] = \mathbb{E}\left[e^{iuS_{t-s}} \,|\, \mathcal{F}_0\right] \\ &= \exp\left\{(e^{iu}-1)\lambda\, \alpha_D(t-s)\right\}\end{aligned}$$

for every $u \in \mathbb{R}$ and $0 \leq s \leq t$, which yields condition (2.12). To complete the proof, we show the integrability. Since $(\Phi(t,D))_{t\geq 0}$ is an $\mathfrak{F}$-DSPP with intensity $(\lambda\, \alpha_D)_{t\geq 0}$ and $\bar{\alpha} = (\alpha_D)_{D\subset\mathbb{D}}$ takes values in $\mathring{\Delta}_\ell$ (cf. Assumption 4.1), we get

$$\mathbb{E}[\Phi(t,D)] = \mathbb{E}\left[\mathbb{E}[\Phi(t,D) \,|\, \mathcal{F}_0]\right] = \mathbb{E}\left[\sum_{k=1}^{\infty} k \frac{(\lambda\, \alpha_D\, t)^k}{k!} e^{-\lambda\, \alpha_D\, t}\right]$$

$$= \mathbb{E}\left[\lambda\,\alpha_D t e^{-\lambda\alpha_D t}\sum_{k=1}^{\infty}\frac{(\lambda\alpha_D t)^{k-1}}{(k-1)!}\right] = \lambda t\mathbb{E}\left[\alpha_D\, e^{-\lambda\alpha_D t}\sum_{k=0}^{\infty}\frac{(\lambda\alpha_D t)^{k}}{k!}\right] = \lambda t\mathbb{E}\left[\alpha_D\right] \le \lambda t$$

for every $t \ge 0$. □

Lemma 4.7. *For any $B \in \mathcal{P}(\mathcal{P}(\mathbb{D}))$, the $\mathfrak{F}$-predictable intensity of the SPP $(\Phi(t,B))_{t\ge 0}$ is $\left(\lambda\sum_{D\in B}\alpha_D\right)_{t\ge 0}$.*

Proof. Fix $B \in \mathcal{P}(\mathcal{P}(\mathbb{D}))$. Obviously, the process $\left(\lambda\sum_{D\in B}\alpha_D\right)_{t\ge 0}$ is non-negative and $\mathfrak{F}$-predictable. According to Lemma 4.6 in connection with Proposition 2.85, $(\Phi(t,D))_{t\ge 0}$ is an SPP with intensity $(\lambda\alpha_D)_{t\ge 0}$ for every $D \subset \mathbb{D}$. Consequently, we have, for any non-negative $\mathfrak{F}$-predictable processes $(H_t)_{t\ge 0}$,

$$\begin{aligned}
\mathbb{E}\left[\int_0^{\infty} H_t\,\mathrm{d}\Phi(t,B)\right] &= \mathbb{E}\left[\sum_{n\in\mathbb{N}} H_{T_n}\mathbb{1}_{\{T_n<\infty\}}\mathbb{1}_{\{Z_n\in B\}}\right]\\
&= \mathbb{E}\left[\sum_{n\in\mathbb{N}} H_{T_n}\mathbb{1}_{\{T_n<\infty\}}\sum_{D\in B}\mathbb{1}_{\{Z_n=D\}}\right] = \sum_{D\in B}\mathbb{E}\left[\int_0^{\infty} H_t\,\mathrm{d}\Phi(t,D)\right]\\
&= \sum_{D\in B}\mathbb{E}\left[\int_0^{\infty} H_t\,\lambda\,\alpha_D\,\mathrm{d}t\right] = \mathbb{E}\left[\int_0^{\infty} H_t\,\lambda\sum_{D\in B}\alpha_D\,\mathrm{d}t\right]
\end{aligned}$$

Note that the equation is trivial for $B = \emptyset$ since in this case $(\Phi(t,\emptyset))_{t\ge 0}$ is constant zero. This completes the proof □

Our aim is to apply the filter result for marked-point-process observations (Theorem 2.101), where the observed filtration is $\mathfrak{F}^{\Phi}$. For this purpose, we need local characteristics of $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$ w.r.t. $\mathfrak{F}$ and $\mathfrak{F}^{\Phi}$, respectively (see Definition 2.99). These characteristics are given in the next two propositions using the following notations.

Notation. Let $j \in \{1,\ldots,m\}$. From now on, we set $a_j^{\emptyset} := 0$ and we denote by $p_j = (p_j(t))_{t\ge 0}$ a càdlàg modification of the process $(\mathbb{P}(\bar\alpha = a_j \,|\, \mathcal{F}_t^{\Phi}))_{t\ge 0}$, i.e.

$$p_j(t) = \mathbb{P}(\bar\alpha = a_j \,|\, \mathcal{F}_t^{\Phi}), \quad t \ge 0.$$

Furthermore, for any $D \subset \mathbb{D}$, $p_j^D = (p_j^D(t))_{t\ge 0}$ denotes a càdlàg modification of the process $\left(\mathbb{P}(\alpha_D = a_j^D \,|\, \mathcal{F}_t^{\Phi})\right)_{t\ge 0}$, i.e.

$$p_j^D(t) = \mathbb{P}(\alpha_D = a_j^D \,|\, \mathcal{F}_t^{\Phi}), \quad t \ge 0.$$

Moreover, for any $D \subset \mathbb{D}$, we set

$$p_t^D := \sum_{k=1}^{m} a_k^D\, p_k^D(t), \quad t \ge 0,$$

i.e. p_t^D is a càdlàg modification of $(\mathbb{E}[\alpha_D \,|\, \mathcal{F}_t^{\Phi}])_{t\ge 0}$ since $\sum_{k=1}^m a_k^D\, p_k^D(t) = \mathbb{E}[\alpha_D \,|\, \mathcal{F}_t^{\Phi}]$. We further denote by $(p_t)_{t\ge 0}$ a process which is defined by $p_t := (p_1(t),\ldots,p_m(t))$, $t \ge 0$.

Justification of the notation. It is clear that $\mathbb{P}(\bar\alpha = a_j \,|\, \mathcal{F}_t^{\Phi}) = \mathbb{E}[\mathbb{1}_{\{\bar\alpha=a_j\}} \,|\, \mathcal{F}_t^{\Phi}]$, where $(\mathbb{1}_{\{\bar\alpha=a_j\}})_{t\ge 0}$ is a bounded and càdlàg process. Moreover, the filtration $\mathfrak{F}^{\Phi}$ is right-continuous according to Thm. 2.69. Consequently, by Proposition 2.31, the process $(\mathbb{P}(\bar\alpha = a_j \,|\, \mathcal{F}_t^{\Phi}))_{t\ge 0}$ admits a càdlàg modification. In the same manner we can see that the process $(\mathbb{P}(\alpha_D = a_j^D \,|\, \mathcal{F}_t^{\Phi}))_{t\ge 0}$ has a càdlàg modification. □

The process $(p_t)_{t\geq 0}$ is called the *filter* of $\bar{\alpha}$, which can be seen as the posterior probability mass function of $\bar{\alpha}$ given $Z_1, \ldots, Z_{N_t}$. Therefore, from the insurer's point of view, the filter $(p_t)_{t\geq 0}$ is of main concern. It encapsulates the information gathered so far about the thinning probabilities which can be expected from the claim arrivals.

Remark 4.8. (i) For $t = 0$, $\mathcal{F}_0^\Phi$ is the trivial σ-algebra (i.e. $\mathcal{F}_0^\Phi = \{\emptyset, \Omega\}$) and, consequently, $p_j(0) = \mathbb{P}(\bar{\alpha} = a_j \,|\, \mathcal{F}_0^\Phi) = \mathbb{P}(\bar{p} = a_j) = \pi_{\bar{\alpha}}(j)$, $j = 1, \ldots, m$. Recall that $\bar{\pi} = (\pi_{\bar{\alpha}}(1), \ldots, \pi_{\bar{\alpha}}(m))$ describes the probability mass function of the given prior distribution $\Pi_{\bar{\alpha}}$ of $\bar{\alpha}$. So $p_j(0)$ is known by the insurer for every $j = 1, \ldots, m$.

(ii) Notice further that $p_t^D(\omega) \in (0,1)$ for $\mathbb{P}$-a.a. $\omega \in \Omega$ since $\alpha_D \in (0,1)$ due to the assumption that $\bar{\alpha} = (\alpha_D)_{D\subset\mathbb{D}} \in \mathring{\Delta}_\ell$, compare Assumption 4.1, Moreover, $\sum_{D\subset\mathbb{D}} p_t^D = \mathbb{E}[\sum_{D\subset\mathbb{D}} \alpha_D \,|\, \mathcal{F}_t^\Phi] = 1$ and thus $\sum_{D\subset\mathbb{D}} p_{t-}^D = 1$.

Lemma 4.9. *The $\mathfrak{F}$-intensity kernel of $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$, denoted by $(\mu(t, \mathrm{d}z))_{t\geq 0}$, is given by*

$$\mu(t, B) = \lambda \sum_{D\in B} \alpha_D, \quad t \geq 0, \quad B \in \mathcal{P}(\mathcal{P}(\mathbb{D})).$$

Proof. Fix $B \in \mathcal{P}(\mathcal{P}(\mathbb{D}))$. Firstly, we have to show that λ is a transition kernel from $(\mathbb{R}^+\times\Omega, \mathcal{B}^+\otimes\mathcal{F})$ to $(\mathcal{P}(\mathbb{D}), \mathcal{P}(\mathcal{P}(\mathbb{D})))$, compare Definition 2.96. Since α_D is $\mathcal{F}_0$-measurable (in particular, $\mathcal{F}$-measurable), $(\lambda \sum_{D\in B} \alpha_D)_{t\geq 0}$ is an $\mathfrak{F}$-adapted and continuous process and, consequently, measurable w.r.t. $\mathcal{F}$ (compare Proposition 2.16), i.e. $\lambda \sum_{D\in B} \alpha_D$ is $\mathcal{B}^+ \otimes \mathcal{F}$-measurable. Moreover, it is easily seen that through $\lambda \sum_{D\in B} \alpha_D(\omega)$ a measure on $(\mathcal{P}(\mathbb{D}), \mathcal{P}(\mathcal{P}(\mathbb{D})))$ is defined for every $\omega \in \Omega$ since $\lambda > 0$. Secondly, we have to prove that $(\lambda \sum_{D\in B} \alpha_D)_{t\geq 0}$ is the predictable $\mathfrak{F}$-intensity of $(\Phi(t, B))_{t\geq 0}$ which was already shown in Lemma 4.7. □

Proposition 4.10. *The local $\mathfrak{F}$-characteristic of $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$, denoted by $(\lambda_t, \mu(t, \mathrm{d}z))_{t\geq 0}$, is given by*

$$\lambda_t = \lambda, \quad \mu(t, B) = \sum_{D\in B} \alpha_D, \quad t \geq 0, \quad B \in \mathcal{P}(\mathcal{P}(\mathbb{D})).$$

Proof. The announced characteristic follows immediately from Lemma 4.9 and the obvious fact that a stochastic kernel from $(\mathbb{R}^+ \times \Omega, \mathcal{B}^+ \otimes \mathcal{F})$ to $(\mathcal{P}(\mathbb{D}), \mathcal{P}(\mathcal{P}(\mathbb{D})))$ is defined through $\sum_{D\in B} \alpha_D$. □

Proposition 4.11. *The local $\mathfrak{F}^\Phi$-characteristic of $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$, denoted by $(\widehat{\lambda}_t, \widehat{\mu}(t, \mathrm{d}z))_{t\geq 0}$, is given by*

$$\widehat{\lambda}_t = \lambda, \quad \widehat{\mu}(t, B) = \sum_{D\in B} p_{t-}^D, \quad t \geq 0, \quad B \in \mathcal{P}(\mathcal{P}(\mathbb{D})).$$

Proof. Obviously, $(\lambda)_{t\geq 0}$ is non-negative and it is an $\mathfrak{F}^\Phi$-predictable process. We define $\widehat{\mu}(t, \omega, B) := \sum_{D\in B} p_{t-}^D(\omega)$ for every $t \geq 0$, $B \in \mathcal{P}(\mathcal{P}(\mathbb{D}))$ and $\omega \in \Omega$. According to Remark 4.8 (ii), it holds $\widehat{\mu}(t, \mathcal{P}(\mathbb{D})) = \sum_{D\subset\mathbb{D}} p_{t-}^D = 1$, i.e. $\widehat{\mu}(t, \omega, \mathrm{d}z)$ is a probability measure on $\mathcal{P}(\mathbb{D})$ for $\mathbb{P}$-a.a. $\omega \in \Omega$ and all $t \geq 0$. Furthermore, for any $B \in \mathcal{P}(\mathcal{P}(\mathbb{D}))$, we have

$$\mathbb{R}^+ \times \Omega \ni (t, \omega) \mapsto \widehat{\mu}(t, \omega, B) = \sum_{D\in B} p_{t-}^D(\omega)$$

is $(\mathcal{B}^+ \otimes \mathcal{F}_\infty^\Phi)$-measurable since $(p_{t-}^D)_{t\geq 0}$ is $\mathfrak{F}^\Phi$-predictable and, hence, $\mathfrak{F}^\Phi$-progressively measurable (cf. Prop. 2.56), in particular, measurable w.r.t. $\mathcal{F}_\infty^\Phi$. This establishes that

$\widehat{\mu}(\cdot, \mathrm{d}z)$ is a stochastic kernel from $(\mathbb{R}^+\times\Omega, \mathcal{B}^+\otimes\mathcal{F}^\Phi_\infty)$ to $(\mathcal{P}(\mathbb{D}), \mathcal{P}(\mathcal{P}(\mathbb{D})))$. It remains to prove that $\big(\lambda\sum_{D\in\mathrm{d}z} p^D_{t-}\big)_{t\geq 0}$ is the $\mathfrak{F}^\Phi$-intensity kernel of Φ. Clearly, $\lambda\sum_{D\in\mathrm{d}z} p^D_\cdot(\cdot)$ is a transition kernel from $(\mathbb{R}^+\times\Omega, \mathcal{B}^+\otimes\mathcal{F}^\Phi_\infty)$ to $(\mathcal{P}(\mathbb{D}), \mathcal{P}(\mathcal{P}(\mathbb{D})))$ since $\widehat{\mu}(\cdot, \mathrm{d}z)$ is a stochastic kernel from $(\mathbb{R}^+\times\Omega, \mathcal{B}^+\otimes\mathcal{F}^\Phi_\infty)$ to $(\mathcal{P}(\mathbb{D}), \mathcal{P}(\mathcal{P}(\mathbb{D})))$. Next, we fix $B\in\mathcal{P}(\mathcal{P}(\mathbb{D}))$. The task is now to show that $(\lambda\sum_{D\in B} p^D_{t-})_{t\geq 0}$ is the $\mathfrak{F}^\Phi$-predictable $\mathfrak{F}^\Phi$-intensity of $(\Phi(t,B))_{t\geq 0}$. To do this, let us remind that we have seen in Lemma 4.7 that $\big(\lambda\sum_{D\in B}\alpha_D\big)_{t\geq 0}$ is an $\mathfrak{F}$-intensity of $\big(\Phi(t,B)\big)_{t\geq 0}$. We know already that $\big(\lambda\sum_{D\in B} p^D_t\big)_{t\geq 0}$ is a càdlàg modification of $\big(\lambda\,\mathbb{E}[\sum_{D\in B}\alpha_D\,|\,\mathcal{F}^\Phi_t]\big)_{t\geq 0}$ (see notation above) and, hence, $\mathfrak{F}^\Phi$-progressively measurable. Therefore, by Proposition 2.81, we get that $\big(\lambda\sum_{D\in B} p^D_t\big)_{t\geq 0}$ is an $\mathfrak{F}^\Phi$-intensity of $(\Phi(t,B))_{t\geq 0}$ and, in consequence, $(\lambda\sum_{D\in B} p^D_{t-})_{t\geq 0}$ is the $\mathfrak{F}^\Phi$-predictable $\mathfrak{F}^\Phi$-intensity. □

Notation. We denote by $\widehat{\Phi}(\mathrm{d}t, \mathrm{d}z)$ the compensated random measure of $\Phi(\mathrm{d}t, \mathrm{d}z)$ which defined by

$$\widehat{\Phi}(\mathrm{d}t, \mathrm{d}z) := \Phi(\mathrm{d}t, \mathrm{d}z) - \lambda\,\widehat{\mu}(t, \mathrm{d}z)\,\mathrm{d}t, \tag{4.2}$$

where $\widehat{\mu}$ is defined as in Proposition 4.11.

Remark 4.12. Notice that $\Phi(\mathrm{d}t, \mathrm{d}z)$ and $\widehat{\Phi}(\mathrm{d}t, \mathrm{d}z) + \lambda\widehat{\mu}(t, \mathrm{d}z)\,\mathrm{d}t$ are equal random measures on $\mathbb{R}^+\times\mathcal{P}(\mathbb{D})$. This plain property will ensure that the reduced control problem solves the partially observable problem, compare Section 4.3.

We are now in the position to determine an equation for the dynamic of the process $(p_j(t))_{t\geq 0}$, $j\in\mathbb{D}$, and thus a dynamic for the filter $(p_t)_{t\geq 0}$.

Theorem 4.13. *For any $j\in\{1,\ldots,m\}$, the process $(p_j(t))_{t\geq 0}$ satisfies*

$$p_j(t) = \pi_{\bar{\alpha}}(j) + \int_0^t\int_{\mathcal{P}(\mathbb{D})}\left(\frac{a^z_j\,p_j(s-)}{p^z_{s-}} - p_j(s-)\right)\widehat{\Phi}(\mathrm{d}s, \mathrm{d}z), \quad t\geq 0. \tag{4.3}$$

Proof. The aim is to use the filter results for marked point process observations stated in Theorem 2.101. To do this, let us fix $j\in\{1,\ldots,m\}$ and set $Z_j := \mathbb{1}_{\{\bar{\alpha}=a_j\}}$. Obviously, Z_j is $\mathcal{F}_0$-measurable. Hence Assumption 2.100 is fulfilled and we can apply Theorem 2.101, which yields under consideration of $\mathbb{E}[Z_j] = \mathbb{P}(\bar{\alpha}=a_j) = \pi_{\bar{\alpha}}(j)$ and $\widehat{Z}_j(t) = \mathbb{E}[\mathbb{1}_{\{\bar{\alpha}=a_j\}}\,|\,\mathcal{F}^\Phi_t] = p_j(t)$

$$p_j(t) = \pi_{\bar{\alpha}}(j) + \int_0^t\int_{\mathcal{P}(\mathbb{D})}\big(A_j(t,z) - p_j(t-)\big)\big(\Phi(\mathrm{d}s, \mathrm{d}z) - \lambda\,\widehat{\mu}(s,\,\mathrm{d}z)\,\mathrm{d}s\big), \tag{4.4}$$

where A_j is an $\mathfrak{F}^\Phi$-predictable function indexed by $\mathcal{P}(\mathbb{D})$ satisfying (2.20). It remains to determine A_j. To do this, we denote throughout the proof by H an arbitrary bounded $\mathfrak{F}^\Phi$-predictable function indexed by $\mathcal{P}(\mathcal{P}(\mathbb{D}))$. We know from (2.20) that A_j is computed from

$$\mathbb{E}\left[\int_0^t\int_{\mathcal{P}(\mathbb{D})} Z_j\,H(s,z)\,\lambda\,\mu(s,\mathrm{d}z)\,\mathrm{d}s\right] = \mathbb{E}\left[\int_0^t\int_{\mathcal{P}(\mathbb{D})} A_j(s,z)\,H(s,z)\,\lambda\,\widehat{\mu}(s,\mathrm{d}z)\,\mathrm{d}s\right]. \tag{4.5}$$

Recall that $\mu(t,\mathrm{d}z) = \sum_{D\in\mathrm{d}z}\alpha_D$ and $\widehat{\mu}(t,\mathrm{d}z) = \sum_{D\in\mathrm{d}z} p^D_{t-}$ are probability measures on $\mathcal{P}(\mathbb{D})$ for every $t\geq 0$ and $\omega\in\Omega$ as well as that the integrand of a Lebesgue integral can be changed at countably many points without changing the integral. On account of these facts as well as Fubini's Theorem, we conclude that

$$\mathbb{E}\left[\int_0^t\int_{\mathcal{P}(\mathbb{D})}\frac{a_j^z\,p_j(s-)}{p_{s-}^z}\,H(s,z)\,\lambda\,\widehat{\mu}(s,\mathrm{d}z)\,\mathrm{d}s\right]=\mathbb{E}\left[\int_0^t\int_{\mathcal{P}(\mathbb{D})}\frac{a_j^z\,p_j(s)}{p_s^z}\,H(s,z)\,\lambda\sum_{D\in\mathrm{d}z}p_s^D\,\mathrm{d}s\right]$$
$$=\mathbb{E}\left[\int_0^t\sum_{D\subset\mathbb{D}}H(s,D)\,\lambda\,a_j^D\,p_j(s)\,\mathrm{d}s\right]=\int_0^t\sum_{D\subset\mathbb{D}}\mathbb{E}\left[H(s,D)\,\lambda\,a_j^D\,\mathbb{P}(\bar{\alpha}=a_j\,|\,\mathcal{F}_s^\Phi)\right]\mathrm{d}s$$
$$=\int_0^t\sum_{D\subset\mathbb{D}}\mathbb{E}\left[H(s,D)\,\lambda\sum_{k=1}^m a_k^D\mathbb{1}_{\{a_k=a_j\}}\,\mathbb{P}(\bar{\alpha}=a_k,\alpha_D=a_k^D\,|\,\mathcal{F}_s^\Phi)\right]\mathrm{d}s$$
$$=\int_0^t\sum_{D\subset\mathbb{D}}\mathbb{E}\left[H(s,D)\,\lambda\,\mathbb{E}[\alpha_D\mathbb{1}_{\{\bar{\alpha}=a_j\}}\,|\,\mathcal{F}_s^\Phi]\right]\mathrm{d}s$$
$$=\int_0^t\sum_{D\subset\mathbb{D}}\mathbb{E}\left[\mathbb{E}[H(s,D)\,\lambda\,\alpha_D\,\mathbb{1}_{\{\bar{\alpha}=a_j\}}\,|\,\mathcal{F}_s^\Phi]\right]\mathrm{d}s$$
$$=\int_0^t\sum_{D\subset\mathbb{D}}\mathbb{E}\left[H(s,D)\,\lambda\,\alpha_D\,\mathbb{1}_{\{\bar{\alpha}=a_j\}}\right]\mathrm{d}s$$
$$=\mathbb{E}\left[\int_0^t\sum_{D\subset\mathbb{D}}Z_j\,H(s,D)\,\lambda\,\alpha_D\,\mathrm{d}s\right]=\mathbb{E}\left[\int_0^t\int_{\mathcal{P}(\mathbb{D})}Z_j\,H(s,z)\,\lambda\,\mu(t,\mathrm{d}z)\,\mathrm{d}s\right].$$

In the sixth equality, we have used the $\mathcal{P}(\mathfrak{F}^\Phi)\otimes\mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurability of H, which implies that $H(s,D)$ is $\mathcal{F}_s^\Phi$-measurable for all $s\geq 0$. Recall that, for any $D\subset\mathbb{D}$, $p_t^D\in(0,1)$ $\mathbb{P}$-almost surely, compare Remark 4.8. By setting

$$A_j(t,z):=\frac{a_j^z\,p_j(t-)}{p_{t-}^z},\quad t\geq 0,\quad z\in\mathcal{P}(\mathbb{D}),\tag{4.6}$$

the calculation above has shown that A_j satisfies condition (4.5). Moreover, A_j is an $\mathfrak{F}^\Phi$-predictable function indexed by $\mathcal{P}(\mathbb{D})$. Indeed, in the light of the $\mathfrak{F}^\Phi$-predictability of $(p_j(t-))_{t\geq0}$ and $(p_{t-}^z)_{t\geq0}$, we have that $A_j(\cdot,z)$ is measurable w.r.t. $\mathcal{P}(\mathfrak{F}^\Phi)$ for every $z\in\mathcal{P}(\mathbb{D})$. Furthermore, since $\mathcal{P}(\mathcal{P}(\mathbb{D}))$ is the set of all subset of $\mathcal{P}(\mathbb{D})$, A_j is $\mathcal{P}(\mathfrak{F}^\Phi)\otimes\mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable. Therefore, the proof is completed by inserting the computed A_j in Equation (4.4). □

Remark 4.14. It is worth noting that the process $(p_j(t))_{t\geq0}$ given by (4.3) is an FV process since it is the difference of two Lebesgue-Stieltjes integrals. Notice further that filter p takes values in Δ_m. We verify this by showing that the sum over all $j\in\{1,\dots,m\}$ of right-hand side in (4.3) is one. For any $t\geq0$, we have (since $\bar{\pi}\in\Delta_m$)

$$\sum_{j=1}^m\left(\pi_{\bar{\alpha}}(j)+\int_0^t\int_{\mathcal{P}(\mathbb{D})}\left(\frac{a_j^z\,p_j(s-)}{p_{s-}^z}-p_j(s-)\right)\widehat{\Phi}(\mathrm{d}s,\mathrm{d}z)\right)$$
$$=1+\int_0^t\int_{\mathcal{P}(\mathbb{D})}\left(\frac{\sum_{j=1}^m a_j^z p_j(s-)}{p_{s-}^z}-\sum_{j=1}^m p_j(s-)\right)\widehat{\Phi}(\mathrm{d}s,\mathrm{d}z),$$

where

$$\sum_{j=1}^m p_j(s)=\sum_{j=1}^m\mathbb{P}(\bar{\alpha}=a_j\,|\,\mathcal{F}_t^\Phi)=1$$

and, as above,

$$\frac{\sum_{j=1}^m a_j^z p_j(s)}{p_s^z} = \frac{\sum_{j=1}^m \mathbb{E}[\alpha_z \mathbb{1}_{\{\bar\alpha = a_j\}} \,|\, \mathcal{F}_t^\Phi]}{\mathbb{E}[\alpha_z \,|\, \mathcal{F}_t^\Phi]} = 1, \quad z \in \mathcal{P}(\mathbb{D}).$$

So, by using the left-limit processes of the corresponding càdlàg modifications of the conditional expectations in the equation above, we obtain

$$\frac{\sum_{j=1}^m a_j^z p_j(s-)}{p_{s-}^z} = 1 \quad \text{and} \quad \sum_{j=1}^m p_j(s-) = 1, \quad z \in \mathcal{P}(\mathbb{D}),$$

and thus

$$\int_0^t \int_{\mathcal{P}(\mathbb{D})} \left(\frac{\sum_{j=1}^m a_j^z p_j(s-)}{p_{s-}^D} - \sum_{j=1}^m p_j(s-) \right) \widehat{\Phi}(\mathrm{d}s, \mathrm{d}z) = 1.$$

Let us mention some elementary properties of the filter $(p_t)_{t\geq 0}$.

Proposition 4.15. *For any $j \in \{1, \ldots, m\}$, the process $(p_j(t))_{t\geq 0}$ given by* (4.3) *is an $\mathfrak{F}^\Phi$-martingale.*

Proof. Fix $j \in \{1, \ldots, m\}$. Appealing to Corollary 2.98, the process $(p_j(t))_{t\geq 0}$ given by (4.3) is an $\mathfrak{F}^\Phi$-martingale if

$$\mathbb{E}\left[\int_0^t \int_{\mathcal{P}(\mathbb{D})} \left| \frac{a_j^z\, p_j(s-)}{p_{s-}^z} - p_j(s-) \right| \lambda \sum_{D \in \mathrm{d}z} p_{s-}^D \,\mathrm{d}s \right] < \infty, \quad t \geq 0.$$

With the help of Fubini's theorem, the triangle inequality, Remark 4.8 (ii) and the facts that $a_j \in \mathring{\Delta}_\ell$, $j = 1, \ldots, m,$ and $p_j(s), p_s^D > 0$, we get

$$\begin{aligned}
&\mathbb{E}\left[\int_0^t \int_{\mathcal{P}(\mathbb{D})} \left| \frac{a_j^z p_j(s-)}{p_{s-}^z} - p_j(s-) \right| \lambda \sum_{D \in \mathrm{d}z} p_{s-}^D \,\mathrm{d}s \right] = \sum_{D \subset \mathbb{D}} \mathbb{E}\left[\int_0^t \left| \frac{a_j^D p_j(s)}{p_s^D} - p_j(s) \right| \lambda p_s^D \,\mathrm{d}s \right] \\
&\leq \lambda \sum_{D \subset \mathbb{D}} \mathbb{E}\left[\int_0^t a_j^D\, p_j(s) \,\mathrm{d}s \right] + \lambda \sum_{D \subset \mathbb{D}} \mathbb{E}\left[\int_0^t p_j(s)\, p_s^D \,\mathrm{d}s \right] \\
&= \lambda \int_0^t \mathbb{E}\left[p_j(s)\right] \mathrm{d}s \underbrace{\sum_{D \subset \mathbb{D}} a_j^D}_{=1} + \lambda \int_0^t \mathbb{E}\Big[p_j(s) \underbrace{\sum_{D \subset \mathbb{D}} p_s^D}_{=1} \Big] \,\mathrm{d}s \\
&= 2\lambda \int_0^t \mathbb{E}\left[\mathbb{E}[\mathbb{1}_{\{\bar\alpha = a_j\}} \,|\, \mathcal{F}_s^\Phi] \right] \mathrm{d}s = 2\lambda \int_0^t \mathbb{P}(\bar\alpha = a_j) \,\mathrm{d}s = 2\lambda\, \pi_{\bar\alpha}(j)\, t \leq 2\lambda t < \infty,
\end{aligned}$$

for all $t \geq 0$. ∎

Proposition 4.16. *The filter $(p_t)_{t\geq 0}$ is a pure jump process and the new state of $(p_t)_{t\geq 0}$ after the jump times $(T_n)_{n\in\mathbb{N}}$ is given by*

$$p_{T_n} = J\big(p_{T_n-}, Z_n\big), \quad n \in \mathbb{N},$$

where

$$J(p, D) := \left(\frac{a_1^D\, p_1}{\sum_{k=1}^m a_k^D\, p_k}, \ldots, \frac{a_m^D\, p_m}{\sum_{k=1}^m a_k^D\, p_k} \right), \quad p = (p_1, \ldots, p_m) \in \Delta_m \quad D \subset \mathbb{D}. \quad (4.7)$$

Proof. Let us fix $j \in \{1, \dots, m\}$ and $t \geq 0$. First, we show that the continuous part $(p_j^c(t))_{t\geq 0}$ of the process $(p_j(t))_{t\geq 0}$ is constant zero. By definition of $\widehat{\Phi}$ given in (4.2) and (4.3), it follows

$$\begin{aligned} p_j(t) = \pi_{\bar\alpha}(j) &+ \int_0^t \int_{\mathcal{P}(\mathbb{D})} \left(\frac{a_j^z \, p_j(s-)}{p_{s-}^z} - p_j(s-) \right) \Phi(\mathrm{d}s, \mathrm{d}z) \\ &- \sum_{D \subset \mathbb{D}} \int_0^t \int_{\mathcal{P}(\mathbb{D})} \left(\frac{a_j^D p_j(s)}{p_s^D} - p_j(s) \right) \lambda \, p_s^D \, \mathrm{d}s, \end{aligned} \tag{4.8}$$

where the integral w.r.t. Φ is a sum. Hence, due to Remark 4.8 (ii), $(p_j^c(t))_{t\geq 0}$ satisfies

$$p_j^c(t) = -\lambda \sum_{D \subset \mathbb{D}} \int_0^t \left(\frac{a_j^D \, p_j(s)}{p_s^D} - p_j(s) \right) p_s^D \, \mathrm{d}s = -\lambda \int_0^t p_j(s) \Big(\underbrace{\sum_{D \subset \mathbb{D}} a_j^D}_{=1} - \underbrace{\sum_{D \subset \mathbb{D}} p_s^D}_{=1} \Big) \mathrm{d}s = 0.$$

Thus $(p_j(t))_{t\geq 0}$ is constant between the jumps and $p_j(t) = \pi_{\bar\alpha}(j) + \sum_{0<s\leq t} \Delta p_j(s)$. That is, $(p_t)_{t\geq 0}$ is a pure jump process. Moreover, according to Equation (4.8), we have

$$\sum_{0<s\leq t} \Delta p_j(t) = \int_0^t \int_{\mathcal{P}(\mathbb{D})} \left(\frac{a_j^z \, p_j(s-)}{p_{s-}^z} - p_j(s-) \right) \Phi(\mathrm{d}s, \mathrm{d}z).$$

and consequently

$$\Delta p_j(T_n) = \frac{a_j^{Z_n} \, p_j(T_n-)}{p_{T_n-}^{Z_n}} - p_j(T_n-), \quad n \in \mathbb{N}.$$

Therefore, the new state of $(p_j(t))_{t\geq 0}$ at $(T_n)_{n\in\mathbb{N}}$ is

$$p_j(T_n) = p_j(T_n-) + \Delta p_j(T_n) = \frac{a_j^{Z_n} p_j(T_n-)}{p_{T_n-}^{Z_n}}, \quad n \in \mathbb{N}.$$

Notice that

$$\begin{aligned} p_t^D &= \sum_{k=1}^m a_k^D p_k^D(t) = \sum_{k=1}^m a_k^D \, \mathbb{P}(\alpha_D = a_k^D \,|\, \mathcal{F}_t^\Phi) \\ &= \sum_{\substack{k\in\{1,\dots,m\}: \\ a_k^D \neq a_j^D \forall j \in\{1,\dots,m\}}} a_k^D \sum_{\substack{\ell\in\{1,\dots,m\}: \\ a_\ell^D = a_k^D}} \mathbb{P}(\bar\alpha = a_\ell \,|\, \mathcal{F}_t^\Phi) = \sum_{k=1}^m a_k^D \, p_k(t), \end{aligned}$$

since

$$\sum_{\substack{\ell\in\{1,\dots,m\}: \\ a_\ell^D = a_k^D}} \mathbb{P}(\bar\alpha = a_\ell \,|\, \mathcal{F}_t^\Phi), \quad k = 1, \dots, m,$$

is the probability mass function of the conditional distribution of the Dth component of $\bar\alpha$ given $\mathcal{F}_t^\Phi$. Thus

$$p_{T_n-}^{Z_n} = \sum_{k=1}^m a_k^{Z_n} \, p_k(T_n-).$$

This yields $p_{T_n} = J(p_{T_n-}, Z_n)$ for all $n \in \mathbb{N}$, where J is defined by (4.7). $\square$

Remark 4.17. (i) It should be pointed out that the proof of Proposition 4.16 yields

$$p_j(t) = \pi_{\bar{\alpha}}(j) + \int_0^t \int_{\mathcal{P}(\mathbb{D})} \left(\frac{a_j^z\, p_j(s-)}{p_{s-}^z} - p_j(s-) \right) \Phi(\mathrm{d}s, \mathrm{d}z),\ t \geq 0,\ j = 1, \ldots, m.$$

(ii) In the proof we have seen that

$$p_t^D = \sum_{k=1}^m a_k^D p_k(t), \tag{4.9}$$

which is the known relation that the mean of some marginal distribution can be calculated by replacing the marginal distribution with the common distribution in the corresponding integral.

The derived pure jump property of the filter $(p_t)_{t \geq 0}$ is illustrated in Figure 4.1, which shows a sample path of the filter process in the case of $A = \{a_1, a_2, a_3\}$ with $a_1 = (4/9, 4/9, 4/9)$, $a_2 = (5/9, 2/9, 2/9)$, $a_3 = (1/3, 1/3, 1/3)$, where the prior distribution is given by $\bar{\pi}_{\bar{\alpha}} = (2/5, 2/5, 1/5)$. The trajectory of the filter was simulated under the assumption that the (unobservable) realization of $\bar{\alpha}$ is a_2, i.e. $\mathbb{P}(\bar{\alpha} = \alpha_j \,|\, \mathcal{F}_0) = 1$. This is recognized by the filter over time since the probability that the dependencies between the LoBs are given by a_2 is nearly 1 at the end of the consider time interval, compare the red line.

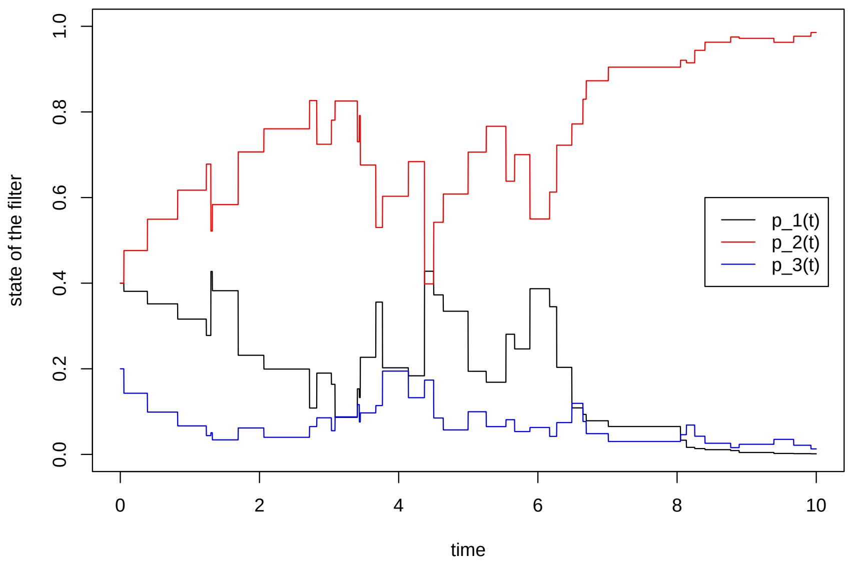

Figure 4.1: A trajectory of the filter process $(p_t)_{t \geq 0}$ under the assumptions $A = \{a_1, a_2, a_3\}$ with $a_1 = (4/9, 4/9, 4/9)$, $a_2 = (5/9, 2/9, 2/9)$, $a_3 = (1/3, 1/3, 1/3)$ and that $\bar{\pi}_{\bar{\alpha}} = (2/5, 2/5, 1/5)$ as well as $\mathbb{P}(\bar{\alpha} = a_2 \,|\, \mathcal{F}_0) = 1$.

The representation of the filter $(p_t)_{t \geq 0}$ given in Theorem 4.13 allows us to reduce the partially observable control problem (P) to one with a state process containing all relevant observable information about the unknown interdependency between the insurance classes whose solution solves the original problem. Before moving on to the reduced control problem, we carry out some useful properties of aggregated claim amount process and the surplus process.

Properties of the aggregated claim amount process and the surplus process. For the stochastic control approach, it will be helpful to use a compensated random measure of the marked point processes $\Psi = (T_n, (Y_n, Z_n))_{n\in\mathbb{N}}$ introduced on page 35 to represent the aggregated claims left to the insurer in the surplus process $X^{\xi,b}$ given by (3.7). The compensator of Ψ is determined by applying the following lemma.

Lemma 4.18. *The $\mathfrak{F}$-intensity kernel of Ψ, denoted by $(\nu(t, \mathrm{d}(y,z)))_{t\geq 0}$, is given by*

$$\nu(t,(A,B)) = \lambda\, F(A) \sum_{D\in B} \alpha_D, \quad t \geq 0, \quad (A,B) \in \mathcal{B}((0,\infty)^d) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D})).$$

Proof. Let us fix $t \geq 0$ and $(A,B) \in \mathcal{B}((0,\infty)^d) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D}))$. The task is now to show that $(\nu(t,(A,B)))_{t\geq 0} = \big(\lambda\, F(A) \sum_{D\in B} \alpha_D\big)_{t\geq 0}$ is the $\mathfrak{F}$-predictable $\mathfrak{F}$-intensity of $(\Psi(t,(A,B)))_{t\geq 0}$, where the required non-negativity and $\mathfrak{F}$-predictability are obviously satisfied. Moreover, it holds $\int_0^t \lambda\, F(A) \sum_{D\in B} \alpha_D \,\mathrm{d}s = \lambda\, F(A) \sum_{D\in B} \alpha_D\, t < \lambda t < \infty$. From Assumption 3.3 and fact that $(\lambda \sum_{D\in B} \alpha_D)_{t\geq 0}$ is an $\mathfrak{F}$-intensity of $(\Phi(t,B))_{t\geq 0}$ (see Lemma 4.9), it follows

$$\begin{aligned}
&\mathbb{E}\left[\int_0^\infty H_s\, \Psi(\mathrm{d}s,(A,B))\right] = \mathbb{E}\left[\sum_{n\in\mathbb{N}} H_{T_n} \mathbb{1}_{\{T_n<\infty\}} \mathbb{1}_{\{Z_n\in B\}} \mathbb{1}_{\{Y_n\in A\}}\right] \\
&= \mathbb{E}\left[\int_0^\infty H_s \Phi(\mathrm{d}s, B)\right] F(A) = \mathbb{E}\left[\int_0^\infty H_s \lambda \sum_{D\in B} \alpha_D \mathrm{d}s\right] F(A) = \mathbb{E}\left[\int_0^\infty H_s \lambda F(A) \sum_{D\in B} \alpha_D \mathrm{d}s\right]
\end{aligned}$$

for all non-negative $\mathfrak{F}^\Phi$-predictable processes $(H_t)_{t\geq 0}$, which finishes the proof. □

Remark 4.19. In this chapter, Ψ is a Poisson random measure (PRM) since the background intensity is deterministic and thus N an homogeneous Poisson process. Therefore, according to Mikosch [94, Prop. 7.3.3], Ψ is an PRM with a mean measure defined on $\mathbb{R}^+ \times (0,\infty)^d \times \mathcal{P}(\mathbb{D})$ which is given by $\nu(\mathrm{d}t, \mathrm{d}(y,z))$, where ν is defined as in Lemma 4.18. This property of Ψ is called independent marking of the PRM N. Notice that in the general setting, if the background intensity Λ of N is randomized, then N has no longer independent increments (cf. e.g. Scherer and Selch [108, p. 141]) and thus Ψ can only be understood as a random counting measure in general, not as PRM.

The following result will be proved in the same way as for Proposition 4.11.

Proposition 4.20. *The $\mathfrak{F}^\Psi$-intensity kernel of Ψ, denoted by $(\widehat{\nu}(t, \mathrm{d}(y,z)))_{t\geq 0}$, is given by*

$$\widehat{\nu}(t,(A,B)) = \lambda\, F(A) \sum_{D\in B} p^D_{t-}, \quad t \geq 0, \quad (A,B) \in \mathcal{B}((0,\infty)^d) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D})).$$

Proof. Fix $t \geq 0$. In the proof of Proposition 4.11 we have seen that

$$\mathbb{R}^+ \times \Omega \times \mathcal{P}(\mathcal{P}(\mathbb{D})) \ni (t,\omega,B) \mapsto \sum_{D\in B} p^D_{t-}(\omega)$$

is a stochastic kernel from $(\mathbb{R}^+ \times \Omega, \mathcal{B}^+ \otimes \mathcal{F}_t^\Phi)$ to $(\mathcal{P}(\mathbb{D}), \mathcal{P}(\mathcal{P}(\mathbb{D})))$. Therefore, since $F(A)$ is deterministic, positive and constant in time for every $A \in \mathcal{B}((0,\infty)^d)$,

$$\mathbb{R}^+ \times \Omega \times \mathcal{B}((0,\infty)^d) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D})) \ni (t, \omega, (A, B)) \mapsto F(A) \sum_{D \in B} p_{t-}^D(\omega)$$

is a transition kernel from $(\mathbb{R}^+ \times \Omega, \mathcal{B}^+ \otimes \mathcal{F}_t^\Phi)$ to $((0,\infty)^d \times \mathcal{P}(\mathbb{D}), \mathcal{B}((0,\infty)^d) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D})))$. The only point remaining concerns the intensity property of $\widehat{\nu}(t, \mathrm{d}(y,z))$. For this purpose let us fix $(A, B) \in \mathcal{B}((0,\infty)^d) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D}))$. The process $(\lambda\, F(A) \sum_{D \in B} p_{t-}^D)_{t \geq 0}$ is obviously non-negative and $\mathfrak{F}^\Psi$-predictable due to the $\mathfrak{F}^\Phi$-predictability of $(p_{t-}^D)_{t \geq 0}$ and satisfies $\int_0^t \lambda\, F(A) \sum_{D \in B} p_{s-}^D \, \mathrm{d}s \leq \lambda\, F(B)\, t < \infty$. Furthermore, using Assumption 3.3, Fubini's Theorem and the fact that $(\lambda \sum_{D \in B} p_{t-}^D)_{t \geq 0}$ is an $\mathfrak{F}^\Phi$-intensity of $(\Phi(t, B))_{t \geq 0}$ (see proof of Proposition 4.11), we obtain by a similar calculation as in proof of Lemma 4.18

$$\begin{aligned}
\mathbb{E}\left[\int_0^\infty H_s\, \Psi(\mathrm{d}s, (A, B))\right] &= \mathbb{E}\left[\int_0^\infty H_s\, \Phi(\mathrm{d}s, B)\right] F(A) \\
&= \mathbb{E}\left[\int_0^\infty H_s\, \lambda \sum_{D \in B} p_{t-}^D \, \mathrm{d}s\right] F(A) \\
&= \mathbb{E}\left[\int_0^\infty H_s\, \lambda\, F(A) \sum_{D \in B} p_{t-}^D \, \mathrm{d}s\right],
\end{aligned}$$

for all non-negative $\mathfrak{F}^\Phi$-predictable processes $(H_t)_{t \geq 0}$, which completes the proof. □

Notation. We denote by $\widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z))$ the compensated random measure of Ψ which is defined by

$$\widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z)) := \Psi(\mathrm{d}t, \mathrm{d}(y,z)) - \widehat{\nu}(t, \mathrm{d}(y,z))\, \mathrm{d}t, \tag{4.10}$$

where $\widehat{\nu}$ is defined as in Proposition 4.20.

Using the introduced compensated random measure $\widehat{\Psi}$, we obtain another way to state the aggregated claims process. Recall the notation of E^d introduced on page 35.

Proposition 4.21. *The aggregated claim process $S = (S_t)_{t \geq 0}$ is given by*

$$S_t = \int_0^t \int_{E^d} \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z)) + \lambda \sum_{D \subset \mathbb{D}} \int_0^t p_s^D \, \mathrm{d}s \sum_{i=1}^d \mathbb{1}_D(i)\, \mathbb{E}[Y_1^i], \quad t \geq 0.$$

Furthermore, S_t satisfies

$$\mathbb{E}[S_t] = \lambda \sum_{D \subset \mathbb{D}} \sum_{k=1}^m a_k^D\, \pi_{\bar{\alpha}}(k) \sum_{i=1}^d \mathbb{1}_D(i)\, \mathbb{E}[Y_1^i]\, t, \quad t \geq 0.$$

Proof. Taking into account (3.3), (4.10) and Proposition 4.20, we get

$$\begin{aligned}
S_t &= \int_0^t \int_{E^d} \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z)) + \sum_{D \subset \mathbb{D}} \int_0^t \int_{(0,\infty)^d} \sum_{i=1}^d y_i \mathbb{1}_D(i) \lambda\, F(\mathrm{d}y)\, p_s^D \, \mathrm{d}s \\
&= \int_0^t \int_{E^d} \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z)) + \lambda \sum_{D \subset \mathbb{D}} \int_0^t p_s^D \, \mathrm{d}s \sum_{i=1}^d \mathbb{1}_D(i)\, \mathbb{E}[Y_1^i], \quad t \geq 0,
\end{aligned}$$

which yields the asserted representation of S_t. Next, we use this representation to infer the expectation of S_t. In order to achieve this value, we will show that the process $(\eta_t)_{t\geq 0}$ given by

$$\eta_t = \int_0^t \int_{E^d} \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z)), \quad t \geq 0,$$

is an $\mathfrak{F}^\Psi$-martingale. By Corollary 2.98, $(\eta_t)_{t\geq 0}$ is an $\mathfrak{F}^\Psi$-martingale if the function $H : \mathbb{R}^+ \times \Omega \times (0,\infty)^d \times \mathcal{P}(\mathbb{D})$ defined by

$$H(t,y,z) := \sum_{i=1}^d y_i \mathbb{1}_z(i)$$

is an $\mathfrak{F}^\Psi$-predictable function indexed by $(0,\infty)^d \times \mathcal{P}(\mathbb{D})$ and satisfies

$$\mathbb{E}\left[\int_0^t \int_{E^d} |H(s,y,z)|\, \widehat{\nu}(s, \mathrm{d}(y,z))\, \mathrm{d}s\right] < \infty, \quad t \geq 0.$$

It is easily seen that H holds the desired predictability property. Having disposed this preliminary step, we can calculate the expectation above by Proposition 4.20 and we obtain that the expectation above is equal to

$$\lambda \sum_{D\subset\mathbb{D}} \mathbb{E}\left[\int_0^t \int_{(0,\infty)^d} \sum_{i=1}^d y_i \mathbb{1}_D(i)\, F(\mathrm{d}y)\, p_s^D\, \mathrm{d}s\right] = \lambda \sum_{i=1}^d \sum_{D\subset\mathbb{D}} \mathbb{1}_D(i) \mathbb{E}[Y_1^i] \mathbb{E}\left[\int_0^t p_s^D\, \mathrm{d}s\right],$$

where, by Fubini's theorem,

$$\begin{aligned} \mathbb{E}\left[\int_0^t p_s^D\right] \mathrm{d}s &= \int_0^t \mathbb{E}\left[p_s^D\right] \mathrm{d}s = \int_0^t \mathbb{E}\left[\mathbb{E}[\alpha_D \,|\, \mathcal{F}_s^\Phi]\right] \mathrm{d}s = \mathbb{E}[\alpha_D]\, t \\ &= \sum_{k=1}^m a_k^D\, \mathbb{P}(\alpha_D = a_k^D)\, t = \sum_{k=1}^m a_k^D\, \pi_{\bar{\alpha}}(k)\, t < \infty \end{aligned}$$

In consequence, the martingale property of $(\eta_t)_{t\geq 0}$ follows since $\mathbb{E}[Y_1^i] < \infty$, $i = 1, \dots, d$, compare Remark 4.4. Moreover, the martingale property of $(\eta_t)_{t\geq 0}$ implies

$$\begin{aligned} \mathbb{E}[S_t] &= \mathbb{E}[\eta_t] + \mathbb{E}\left[\lambda \sum_{i=1}^d \mathbb{E}[Y^i] \sum_{D\subset\mathbb{D}} \mathbb{1}_D(i) \int_0^t p_s^D\, \mathrm{d}s\right] \\ &= \lambda \sum_{D\subset\mathbb{D}} \sum_{k=1}^m a_k^D\, \pi_{\bar{\alpha}}(k) \sum_{i=1}^d \mathbb{1}_D(i)\, \mathbb{E}[Y_1^i]\, t \end{aligned}$$

due to the calculation above. □

The proposition gains in interest for the following reformulation of the surplus process $X^{\xi,b} = (X_t^{\xi,b})_{t\geq 0}$:

$$\begin{aligned} \mathrm{d}X_t^{\xi,b} = &\left(rX_s^{\xi,b} + (\mu - r)\xi_s + c(b_s) - \lambda\, b_t \sum_{D\subset\mathbb{D}} p_t^D \sum_{i=1}^d \mathbb{1}_D(i)\, \mathbb{E}[Y_1^i]\right) \mathrm{d}t \\ &+ \xi_s \sigma \mathrm{d}W_s - \int_{E^d} b_t \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z)), \quad t \geq 0. \end{aligned} \tag{4.11}$$

This representation will be one part of the reduced control model, which will be discussed in the next section.

4.3 The reduced control problem

The state process of the incomplete information problem (P) consists only of the surplus process. To obtain the reduced control model, we have to integrate the additional information about the unknown interdependencies between the various business risks which the insurer receives. For this purpose, we extend the state process by the finite dimensional filter process $(p_t)_{t\geq 0}$ which contains all relevant observable information about the unobservable thinning probabilities. The extended state process is the reduced control model, where the $\mathfrak{G}$-adaptability of all components ensures the observability of this process for the insurer.

After this preliminary consideration we are now in the position to present the control problem in a rigorous way. It should be noted that our aim is to solve the reduced control problem by applying the dynamic programming method. The basic idea of this method is that the solution comes from a family of control problems by varying the initial state values and determining relations between the associated value functions.[2] To state this family of control problems, we fix some initial time $t \in [0,T)$ and consider the state process on $[t,T]$. According to the results above, the complete observable controlled process (state process) $(X_s^{\xi,b}, p_s)_{s\in[t,T]}$ is an $(m+1)$-dimensional process characterized for $(\xi,b) \in \mathcal{U}[t,T]$ by

$$\begin{aligned}\mathrm{d}X_s^{\xi,b} &= \left(rX_s^{\xi,b} + (\mu - r)\xi_s + c(b_s) - \lambda\, b_s \sum_{i=1}^{d} \sum_{D\subset\mathbb{D}} p_s^D \mathbb{1}_D(i)\, \mathbb{E}[Y_1^i] \right) \mathrm{d}s \\ &\qquad + \xi_s\sigma\, \mathrm{d}W_s - \int_{E^d} b_s \sum_{i=1}^{d} y_i \mathbb{1}_z(i)\, \widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z)),\end{aligned} \tag{4.12}$$

$$\mathrm{d}p_j(s) = \int_{\mathcal{P}(\mathbb{D})} \left(\frac{a_j^z\, p_j(s-)}{p_{s-}^z} - p_j(s-) \right) \widehat{\Phi}(\mathrm{d}s, \mathrm{d}z), \quad j = 1,\ldots,m, \tag{4.13}$$

for $s \in [t,T]$, where $X_t^{\xi,b} = x \in \mathbb{R}$ and $p_t = p$ with $p = (p_1,\ldots,p_m) \in \Delta_m$. Recall that $X_0^{\xi,b} = x_0$ and $p_j(0) = \mathbb{P}(\bar{\alpha} = a_j) = \pi_{\bar{\alpha}}(j)$, $j = 1,\ldots,m$. We see that all processes involved in the state process are $\mathfrak{G}$-adapted whereby the reduced model is observable for the insurer. Due to this full observable framework concerning the state process, the reduced control problem stated below is often referred to as the reduced problem under complete information in the literature.

The block diagram displayed in Figure 4.2 illustrates the feedback control in the reduced control model, which shows the additional dependency of the control on the filter process. The block diagram already anticipates that the filter process reasonably only affect the reinsurance strategy but not the investment strategy, as the filter only provides information about the insurance risks.

Now, we can formulate the reduced control problem. For any $(\xi,b) \in \mathcal{U}[t,T]$, the *objective function* is given by

$$\begin{aligned}V^{\xi,b}(t,x,p) &:= \mathbb{E}^{t,x,p}\big[U(X_T^{\xi,b})\big] \\ &:= \mathbb{E}\big[U(X_T^{\xi,b}) \,|\, X_t^{\xi,b} = x, p_t = p\big], \quad (t,x,p) \in [0,T]\times\mathbb{R}\times\Delta_m.\end{aligned}$$

[2]See Section 3.1 in Pham [100].

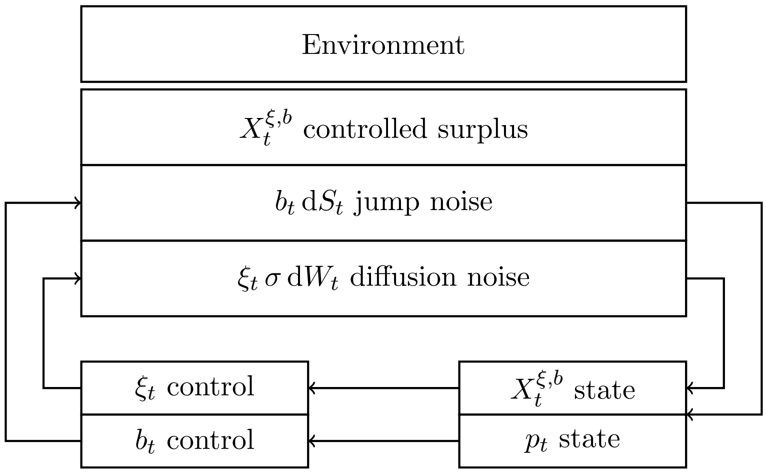

Figure 4.2: The feedback control in the reduced control problem.

With this optimization criterion, the *value function* takes the following form:

$$V(t,x,p) := \sup_{(\xi,b)\in\mathcal{U}[t,T]} V^{\xi,b}(t,x,p), \quad (t,x,p) \in [0,T]\times\mathbb{R}\times\Delta_m. \tag{P1}$$

At this point, it is not trivial a priori whether the value function is measurable since the upper bound of an uncountable set of measurable functions may be a non-measurable function, which is emphasized in Gihman and Skorochod [64, p. 174]. The measurability will turn out later. Similar as before, an investment-reinsurance strategy $(\xi^\star, b^\star)\in\mathcal{U}[t,T]$ is optimal if

$$V(t,x,p) = V^{\xi^\star,b^\star}(t,x,p), \quad (t,x,p)\in[0,T]\times\mathbb{R}\times\Delta_m.$$

The insurance company is interested in the optimal strategies $(\xi^\star, b^\star)\in\mathcal{U}[t,T]$ with

$$(\xi^\star, b^\star) = \operatorname*{argsup}_{(\xi,b)\in\mathcal{U}[t,T]} V^{\xi,b}(t,x,p).$$

Solving the maximization problem directly over all uncountable many strategies is not obvious at all. As already indicated above, the introduced formulation of the problem depending on different initial states allows us to apply the dynamic programming principle which splits the optimization problem into a collection of pointwise maximization separately for every time t. Therefore, an optimal strategy obtained by the dynamic programming principle (more precisely by the HJB equation, which is derived by the dynamic programming principle) is always necessarily Markov, where $(\xi,b)\in\mathcal{U}[t,T]$ is called *Markov* strategy if it is of the form $(\xi_s, b_s) = (v(s, X_s^{\xi,b}, p_s), w(s, X_s^{\xi,b}, p_s))$ for some measurable functions $v:[0,T]\times\mathbb{R}\times\Delta_m\to\mathbb{R}$ and $w:[0,T]\times\mathbb{R}\times\Delta_m\to[0,1]$. The reason for the introduced terminology is that the surplus $X^{\xi,b}$ is a Markov process if (ξ,b) is a Markov strategy w.r.t. the state process $(X_t^{\xi,b}, p_t)_t$. This does not hold in general since the investment-reinsurance strategies may depend on the entire past history. However, since the filter $(p_t)_{t\geq 0}$ contains the information from the past, the strategies depend on the history in the view of the original problem.

Before applying dynamic programming ansatz, we have to realize that solving the reduced control problem (P1) is in fact solving the original problem (P). This is settled by the circumstance that the valued function $\bar{V}(t,x)$ defined in (P) depends on the history $\mathfrak{G}$ only through the filter $(p_t)_{t\geq 0}$. So the filter $(p_t)_{t\geq 0}$ contains the information needed to solve the original control problem (P). Furthermore, the representation of the surplus process in (3.7) and in (4.11) are indistinguishable due to Remark 4.12. Hence, for any $(\xi, b) \in \mathcal{U}[t,T]$,

$$V^{\xi,b}(t,x,p_t) = \bar{V}^{\xi,b}(t,x) \quad \text{and thus} \quad V(t,x,p_t) = \bar{V}(t,x), \quad (t,x) \in [0,T] \times \mathbb{R}.$$

An immediate consequence is that an optimal strategy for the reduced control problem is optimal for the original problem (an vice versa). It should further be noted that the existence of an optimal strategy is not guaranteed. The existence issue will be addressed at the end of Section 4.7.2.

After the verification that an solution of our problem (P) can be obtained by dealing with the problem (P1), we take a closer look at the value function V by the next lemma which states elementary properties of the value function (similar to Lemma 3.3 in Bäuerle and Rieder [31]) using the following notation.

Notation. From now on, we denote by e_k the k-th unit vector in $\mathbb{R}^m$.

Lemma 4.22. *(i) For any $t \in [0,T]$, $(\xi, b) \in \mathcal{U}[t,T]$, $p = (p_1, \ldots, p_m) \in \Delta_m$, and $x \in \mathbb{R}$, it holds*

$$V^{\xi,b}(t,x,p) = \sum_{j=1}^{m} p_j \, V^{\xi,b}(t,x,e_j).$$

(ii) For any $t \in [0,T]$ and $x \in \mathbb{R}$, the function $\Delta_m \ni p \mapsto V(t,x,p)$ is convex.

Proof. (i) The equation follows immediately by conditioning.

(ii) Let us fix $t \in [0,T]$, $p = (p_1, \ldots, p_m) \in \Delta_m$, $q = (q_1, \ldots, q_m) \in \Delta_m$ and $\beta \in [0,1]$. From negativity of the utility function U, it follows immediately that $V^{\xi,b}$ is negative. That is, $\mathcal{U}[t,T] \ni (\xi,b) \mapsto V^{\xi,b}(t,x,p)$ is bounded from above. Therefore, using statement (i) and Proposition B.2, we obtain

$$\begin{aligned}
&V(t,x,\beta p + (1-\beta)q) \\
&= \sup_{(\xi,b)\in\mathcal{U}[t,T]} \sum_{j=1}^{m} (\beta p_j + (1-\beta) q_j) V^{\xi,b}(t,x,e_j) \\
&= \sup_{(\xi,b)\in\mathcal{U}[t,T]} \left(\beta \sum_{j=1}^{m} p_j V^{\xi,b}(t,x,e_j) + (1-\beta) \sum_{j=1}^{m} q_j V^{\xi,b}(t,x,e_j) \right) \\
&\le \beta \sup_{(\xi,b)\in\mathcal{U}[t,T]} \sum_{j=1}^{m} p_j V^{\xi,b}(t,x,e_j) + (1-\beta) \sup_{(\xi,b)\in\mathcal{U}[t,T]} \sum_{j=1}^{m} q_j V^{\xi,b}(t,x,e_j) \\
&= \beta V(t,x,p) + (1-\beta) V(t,x,q),
\end{aligned}$$

for all $x \in \mathbb{R}$. □

Before focusing on the solution of the reduced problem, we conclude the section with a discussion on the used reduction approach.

Remark 4.23. To make inferences about the unknown thinning probabilities we have used a Bayesian approach. Essentially, we make use of the Bayesian estimator for the thinning probabilities. One might wonder if it is possible to use other estimators as well. For the answer let us recall the definition of the objective function $\bar{V}^{\xi,b}(t,x) = \mathbb{E}[\mathbb{E}[U(X_t^{\xi,b}) \mid \mathcal{G}_t] \mid X_t^{\xi,b} = x]$, $(t,x) \in [0,T] \times \mathbb{R}$. The definition of $\bar{V}^{\xi,b}$ uses the conditional expectation, so that the function will not change if we insert a Bayesian estimator. If, on the contrary, we had used another estimator which is not the projection on the observable filtration (e.g. the maximum likelihood estimator), then the objection function $\bar{V}^{\xi,b}$ would change which means that we could not solve the original problem (P) if other estimators than the Bayesian were used for the reduction.

4.4 The Hamilton-Jacobi-Bellman equation

To determine an optimal control of the stochastic control problem stated in (P1), we are going to derive the partial (integro) differential equation of stochastic dynamic programming principle, known as the Hamilton-Jacobi-Bellman equation or simply as the Bellman equation. The HJB equation is derived heuristically with the help of the dynamical programming principle (DPP), which yields, as byproduct, a candidate for an optimal investment-reinsurance strategy. A formal proof of the relationship between the value function and the HJB equation follows later in the verification step in Section 4.7. Due to the heuristic derivation of the HJB equation, no proof of the DPP is required at first.

The starting point for the motivation of the HJB equation is DPP which is also referred as *Bellman principle*: Let $\tau \in [t,T]$ be a $\mathfrak{G}$-stopping time. Then

$$V(t,x,p) = \sup_{(\xi,b)\in\mathcal{U}[t,T]} \mathbb{E}^{t,x,p}\big[V(\tau, X_\tau^{\xi,b}, p_\tau)\big] \tag{4.14}$$

for all $(t,x,p) \in [0,T]\times\mathbb{R}\times\Delta_m$. As explained in Pham [100, p. 41], the Bellman equation can be interpreted as follows: The optimization problem can be split into an optimal control part from time τ given the state of the surplus $X_\tau^{\xi,b}$ and of the filter p_τ (i.e. determine $V(\tau, X_\tau^{\xi,b}, p_\tau)$) and then maximizing the quantity $\mathbb{E}^{t,x,p}\big[V(\tau, X_\tau^{\xi,b}, p_\tau)\big]$ over the investment-reinsurance strategies on $[t,\tau]$.

In the following we choose some fixed time point $t_0 \in (t,T]$. Recall that the filter $(p_t)_{t\geq 0}$ is a pure jump process, compare Proposition 4.16. We suppose that V is sufficient smooth such that the Itô-Doeblin formula can be applied to $V(t_0, X_{t_0}^{\xi,b}, p_{t_0})$ on $[t,t_0]$. It is convenient to use the modified version of the Itô-Doeblin formula given in Corollary 2.58, which yields

$$\begin{aligned}
V(t_0, X_{t_0}^{\xi,b}, p_{t_0}) &= V(t, X_t^{\xi,b}, p_t) + \int_t^{t_0} V_t(s, X_s^{\xi,b}, p_s)\,\mathrm{d}s + \int_t^{t_0} V_x(s, X_{s-}^{\xi,b}, p_{s-})\,\mathrm{d}(X^{\xi,b})_s^c \\
&+ \frac{1}{2}\int_t^{t_0} V_{xx}(s, X_{s-}^{\xi,b}, p_{s-})\,\mathrm{d}[X^{\xi,b}]_s^c + \sum_{t<s\leq t_0}\Big(V(s, X_s^{\xi,b}, p_s) - V(s, X_{s-}^{\xi,b}, p_{s-})\Big),
\end{aligned}$$

where V_t, V_x and V_{xx} denote the partial derivatives of V w.r.t. t, x and xx, respectively. By Propositions 3.15 and 4.16, we get

$$\begin{aligned}
&V(t_0, X_{t_0}^{\xi,b}, p_{t_0}) \\
&= V(t, X_t^{\xi,b}, p_t) + \int_t^{t_0}\Big(V_t(s, X_s^{\xi,b}, p_s) + V_x(s, X_s^{\xi,b}, p_s)\big(rX_s^{\xi,b} + (\mu - r)\xi_s + c(b_s)\big) \\
&\quad + \frac{1}{2}V_{xx}(s, X_s^{\xi,b}, p_s)\xi_s^2\sigma^2\Big)\mathrm{d}s + \int_t^{t_0} V_x(s, X_{s-}^{\xi,b}, p_{s-})\xi_s\sigma\,\mathrm{d}W_s \\
&\quad + \sum_{t<s\leq t_0}\Big(V(s, X_s^{\xi,b}, p_s) - V(s, X_{s-}^{\xi,b}, p_{s-})\Big).
\end{aligned} \tag{4.15}$$

Next we turn our attention to the sum in the equation above. Since $X^{\xi,b}$ and p jump only at trigger arrival times $(T_n)_{n\in\mathbb{N}}$, we obtain

$$\begin{aligned}&\sum_{t<s\le t_0}\big(V(s,X_s^{\xi,b},p_s)-V(s,X_{s-}^{\xi,b},p_{s-})\big)\\&=\sum_{n\in\mathbb{N}}\big(V(T_n,X_{T_n-}^{\xi,b}+\Delta X_{T_n}^{\xi,b},p_{T_n})-V(T_n,X_{T_n-}^{\xi,b},p_{T_n-})\big)\mathbb{1}_{\{T_n\in(t,t_0]\}}.\end{aligned}$$

According to Proposition 3.15 and Proposition 4.16, we have

$$\Delta X_{T_n}^{\xi,b}=-b_{T_n}\sum_{i=1}^d Y_n^i\mathbb{1}_{\{i\in Z_n\}},\quad p_{T_n}=J(p_{T_n-},Z_n),\quad n\in\mathbb{N}.$$

Therefore

$$\begin{aligned}&\sum_{t<s\le t_0}\big(V(s,X_s^{\xi,b},p_s)-V(s,X_{s-}^{\xi,b},p_{s-})\big)\\&=\sum_{n\in\mathbb{N}}\left(V\Big(T_n,X_{T_n-}^{\xi,b}-b_{T_n}\sum_{i=1}^d Y_n^i\mathbb{1}_{\{i\in Z_n\}},J(p_{T_n-},Z_n)\Big)-V(T_n,X_{T_n-}^{\xi,b},p_{T_n-})\right)\mathbb{1}_{\{T_n\in(t,t_0]\}}\\&=\int_t^{t_0}\int_{E^d}\left(V\Big(s,X_{s-}^{\xi,b}-b_s\sum_{i=1}^d y_i\mathbb{1}_z(i),J(p_{s-},z)\Big)-V(s,X_{s-}^{\xi,b},p_{s-})\right)\Psi(\mathrm{d}s,\mathrm{d}(y,z))\end{aligned}$$

Using the compensated random measure $\widehat{\Psi}$ given in (4.10), the last line above is equal to

$$\begin{aligned}&\int_t^{t_0}\int_{E^d}\left(V\Big(s,X_{s-}^{\xi,b}-b_s\sum_{i=1}^d y_i\mathbb{1}_z(i),J(p_{s-},z)\Big)-V(s,X_{s-}^{\xi,b},p_{s-})\right)\widehat{\Psi}(\mathrm{d}s,\mathrm{d}(y,z))\\&+\lambda\sum_{D\subset\mathbb{D}}\int_t^{t_0}p_s^D\int_{(0,\infty)^d}V\Big(s,X_s^{\xi,b}-b_s\sum_{i=1}^d y_i\mathbb{1}_D(i),J(p_s,D)\Big)F(\mathrm{d}y)\,\mathrm{d}s\\&-\lambda\int_t^{t_0}V(s,X_s^{\xi,b},p_s)\underbrace{\sum_{D\subset\mathbb{D}}p_s^D}_{=1}\,\mathrm{d}s.\end{aligned}$$

Inserting this in (4.15) yields

$$\begin{aligned}&V(t_0,X_{t_0}^{\xi,b},p_{t_0})=V(t,X_t^{\xi,b},p_t)\\&+\int_t^{t_0}\Big(V_t(s,X_s^{\xi,b},p_s)+V_x(s,X_s^{\xi,b},p_s)\big(rX_s^{\xi,b}+(\mu-r)\xi_s+c(b_s)\big)\\&+\frac{1}{2}V_{xx}(s,X_s^{\xi,b},p_s)\xi_s^2\sigma^2-\lambda V(s,X_s^{\xi,b},p_s)\\&+\lambda\sum_{D\subset\mathbb{D}}p_s^D\int_{(0,\infty)^d}V\Big(s,X_s^{\xi,b}-b_s\sum_{i=1}^d y_i\mathbb{1}_D(i),J(p_s,D)\Big)F(\mathrm{d}y)\Big)\mathrm{d}s\\&+\eta_{t_0}-\eta_t\end{aligned}\tag{4.16}$$

where, for any $s \in [t, T]$,

$$\begin{aligned}\eta_s := &\int_t^{t_0} V_x(s, X_{s-}^{\xi,b}, p_{s-})\xi_s \sigma \,\mathrm{d}W_s \\ &+ \int_t^{t_0} \int_{E^d} \left(V\Big(s, X_{s-}^{\xi,b} - b_s \sum_{i=1}^d y_i \mathbb{1}_z(i), J(p_{s-}, z)\Big) - V(s, X_{s-}^{\xi,b}, p_{s-}) \right) \widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z)).\end{aligned}$$

We make the assumption that the process $\eta = (\eta_s)_{s \in [t,T]}$ is a $\mathfrak{G}$-martingale. Accordingly, the expectation of η_s is zero for all $s \in [t, T]$, since the process starts at zero. Consequently, by substitution (4.16) back into (4.14), we obtain

$$\begin{aligned}0 = \sup_{(\xi,b) \in \mathcal{U}[t,T]} \mathbb{E}^{t,x,p} \Bigg[\int_t^{t_0} \Bigg(&V_t(s, X_s^{\xi,b}, p_s) - \lambda V(s, X_s^{\xi,b}, p_s) \\ &+ V_x(s, X_s^{\xi,b}, p_s)\big(r X_s^{\xi,b} + (\mu - r)\xi_s + c(b_s)\big) + \frac{1}{2}\sigma^2 V_{xx}(s, X_s^{\xi,b}, p_s)\xi_s^2 \\ &+ \lambda \sum_{D \subset \mathbb{D}} p_s^D \int_{(0,\infty)^d} V\Big(s, X_s^{\xi,b} - b_s \sum_{i=1}^d y_i \mathbb{1}_D(i), J(p_s, D)\Big) F(\mathrm{d}y) \Bigg) \mathrm{d}s \Bigg].\end{aligned}$$

It should be noted that we have subtracted $V(t, x, p)$ from both sides. Next, we divide both sides by $t_0 - t$ and consider $t_0 \downarrow t$. By fundamental theorem of calculus for Lebesgue integrals (FTCL, cf. Sohrab [115, Thm. 11.5.31]) and Equation (4.9), we obtain under the assumption that the limit interchanges with both the supremum and the expectation as well as the Lebesgue integral exists

$$\begin{aligned}0 = \sup_{(\xi,b) \in \mathcal{U}[t,T]} \mathbb{E}^{t,x,p} \Bigg[&V_t(t, X_t^{\xi,b}, p_t) - \lambda V(t, X_t^{\xi,b}, p_t) + \frac{1}{2}\sigma^2 V_{xx}(t, X_t^{\xi,b}, p_t)\xi_t^2 \\ &+ V_x(t, X_t^{\xi,b}, p_t)\big(r\, X_t^{\xi,b} + (\mu - r)\xi_t + c(b_t)\big) \\ &+ \lambda \sum_{D \subset \mathbb{D}} \sum_{k=1}^m a_k^D\, p_k(t) \int_{(0,\infty)^d} V\Big(t, X_t^{\xi,b} - b_t \sum_{i=1}^d y_i \mathbb{1}_D(i), J(p_t, D)\Big) F(\mathrm{d}y) \Bigg].\end{aligned}$$

Using $\xi_t \in \mathbb{R}$ and $b_t \in [0, 1]$, we obtain

$$\begin{aligned}0 = \sup_{(\xi,b) \in \mathbb{R} \times [0,1]} \Bigg\{ &V_t(t, x, p) - \lambda V(t, x, p) + \frac{1}{2}\sigma^2 V_{xx}(t, x, p)\xi^2 \\ &+ V_x(t, x, p)\big(rx + (\mu - r)\xi + c(b)\big) \\ &+ \lambda \sum_{D \subset \mathbb{D}} \sum_{k=1}^m a_k^D p_k \int_{(0,\infty)^d} V\Big(t, x - b\sum_{i=1}^d y_i \mathbb{1}_D(i), J(p, D)\Big) F(\mathrm{d}y) \Bigg\},\end{aligned} \tag{4.17}$$

In case that $t = T$, the value function holds $V(T, x, p) = \mathbb{E}^{T,x,p}\big[U(X_T^{\xi,b})\big] = U(x)$ for all $(x, p) \in [0, T] \times \Delta_m$. Equation (4.17) is the HJB equation for V, which was derived by sending t_0 to t in the DPP given in (4.14) and thus the HJB equation characterises the local behaviour of the value function in that case. Therefore, the HJB equation can be seen as the infinitesimal version of the DPP.

The solution of the HJB equation (4.17) could provide a possible candidate for the value function, which seems difficult to obtain. Therefore, the typical approach is to simplify the equation by a separation approach. We will be able to separate x and p in the sense that the value function is represented as the product of two functions, where one depends on (t,x) and one on (t,p). This may lead to a more easily solvable equation. Thus the following separation approach of the value function – which is typical for an exponential utility function – is important in the solution procedure of the presented optimization problem.

Lemma 4.24. *The value function V holds, for any $(t,x,p) \in [0,T] \times \mathbb{R} \times \Delta_m$,*

$$V(t,x,p) = -e^{-\alpha x e^{r(T-t)}} g(t,p), \tag{4.18}$$

with

$$g(t,p) := \inf_{(\xi,b)\in\mathcal{U}[t,T]} g^{\xi,b}(t,p), \tag{4.19}$$

where

$$\begin{aligned} g^{\xi,b}(t,p) := \mathbb{E}^{t,p}\Bigg[\exp\Bigg\{ -\int_t^T \alpha\, e^{r(T-s)}\big((\mu-r)\xi_s + c(b_s)\big)\,\mathrm{d}s \\ -\int_t^T \alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s + \int_t^T\int_{E^d} \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\,\Psi(\mathrm{d}s,\mathrm{d}(y,z))\Bigg\}\Bigg], \end{aligned} \tag{4.20}$$

where $\mathbb{E}^{t,p}$ denotes the conditional expectation given $p_t = p$.

Proof. From Proposition 3.14 follows

$$\begin{aligned} X_T^{\xi,b} &= X_t^{\xi,b} e^{r(T-t)} + \int_t^T e^{r(T-s)}\Big((\mu-r)\xi_s + c(b_s)\Big)\,\mathrm{d}s + \int_t^T e^{r(T-s)}\xi_s\sigma\,\mathrm{d}W_s \\ &\quad - \int_t^T\int_{E^d} b_s\, e^{r(T-s)}\sum_{i=1}^d y_i\mathbb{1}_z(i)\,\Psi(\mathrm{d}s,\mathrm{d}(y,z)). \end{aligned}$$

Fix $(t,x,p) \in [0,T] \times \mathbb{R} \times \Delta_m$. Then, with the help of the equation above, we get

$$\begin{aligned} &V^{\xi,b}(t,x,p) \\ &= \mathbb{E}^{t,x,p}\big[-\exp\big\{-\alpha X_T^{\xi,b}\big\}\big] \\ &= -e^{-\alpha x e^{r(T-t)}}\mathbb{E}^{t,p}\Bigg[\exp\Bigg\{-\alpha\int_t^T e^{r(T-s)}\Big((\mu-r)\xi_s + c(b_s)\Big)\,\mathrm{d}s \\ &\quad -\alpha\int_t^T e^{r(T-s)}\xi_s\sigma\,\mathrm{d}W_s + \alpha\int_t^T\int_{E^d} b_s\, e^{r(T-s)}\sum_{i=1}^d y_i\mathbb{1}_z(i)\,\Psi(\mathrm{d}s,\mathrm{d}(y,z))\Bigg\}\Bigg]. \end{aligned}$$

Defining $g^{\xi,b}(t,p)$ as in (4.20), we obtain

$$\begin{aligned} V(t,x,p) &= \sup_{(\xi,b)\in\mathcal{U}[t,T]} V^{\xi,b}(t,x,p) \\ &= \sup_{(\xi,b)\in\mathcal{U}[t,T]} \Big\{-e^{-\alpha x e^{r(T-t)}} g^{\xi,b}(t,p)\Big\} \\ &= -e^{-\alpha x e^{r(T-t)}} \inf_{(\xi,b)\in\mathcal{U}[t,T]} g^{\xi,b}(t,p), \end{aligned}$$

which yields the assertion. □

The next presented properties of the function g defined above turn out to be very useful in the following.

Lemma 4.25. *Let g be defined by* (4.19)*. Then the following statements are satisfied*

(i) $g^{\xi,b}(t,p) > 0$ *for all* $(t,p) \in [0,T] \times \Delta_m$ *and* $(\xi,b) \in \mathcal{U}[t,T]$.

(ii) g *is bounded on* $[0,T] \times \Delta_m$.

(iii) $g^{\xi,b}(t,p) = \sum_{j=1}^m p_j \, g^{\xi,b}(t,e_j)$ *for all* $(t,p) \in [0,T] \times \Delta_m$.

(iv) $g^{\xi,b}(t,J(p,D)) = \sum_{j=1}^m \frac{a_j^D \, p_j}{\sum_{k=1}^m a_k^D \, p_k} g^{\xi,b}(t,e_j)$ *for all* $(t,p) \in [0,T] \times \Delta_m$ *and* $D \subset \mathbb{D}$.

(v) $\Delta_m \ni p \mapsto g(t,p)$ *is concave for all* $t \in [0,T]$.

Proof. (i) The statement follows immediately from the definition of $g^{\xi,b}$ given in (4.20).

(ii) Fix $(t,p) \in [0,T] \times \Delta_m$. From the statement above, we have already the lower bound $g(t,p) \geq 0$. It remains to find an upper bound for g. Since $g(t,p) = \inf_{(\xi,b)\in\mathcal{U}[t,T]} g^{\xi,b}(t,p)$, g is bounded from above if there exists a strategy $(\bar{\xi},\bar{b}) \in \mathcal{U}[t,T]$ such that $g^{\bar{\xi},\bar{b}}$ is bounded from above. Let $(\bar{\xi},\bar{b}) \in \mathcal{U}[t,T]$ the strategy which is given by $\bar{\xi}_s \equiv 0$ and $\bar{b}_s \equiv 0$ for all $s \in [t,T]$. Then

$$
\begin{aligned}
g^{\bar{\xi},\bar{b}}(t,p) &= \mathbb{E}^{t,p}\Big[\exp\Big\{-\int_t^T \alpha \, e^{r(T-s)} \, (\eta-\theta) \, \kappa \, \mathrm{d}s\Big\}\Big] \\
&= \exp\Big\{\alpha \, (\theta-\eta) \, \kappa \int_t^T e^{r(T-s)} \, \mathrm{d}s\Big\} \\
&= \exp\Big\{\alpha \, (\theta-\eta) \, \kappa \frac{e^{r(T-t)}-1}{r}\Big\} \\
&\leq \exp\Big\{\alpha \, (\theta-\eta) \, \kappa \frac{e^{rT}-1}{r}\Big\} =: K_0,
\end{aligned}
$$

where $K_0 > 0$ is independent of t and p. Thus K_0 is an upper bound of g which completes the proof.

(iii) Similar to Lemma 4.22 (i), we obtain the equation by conditioning.

(iv) Once again, this assertion follows by conditioning.

(v) The concavity can be proven similar to Lemma 4.22 (iii) by using statement (ii). □

In the following we use the separation approach given in Lemma 4.24 to rearrange the developed HJB eqaution (4.17) for V. Equation (4.18) yields

$$
\begin{aligned}
V_t(t,x,p) &= -e^{-\alpha x e^{r(T-t)}}\big(\alpha \, x \, r \, e^{r(T-t)} g(t,p) + g_t(t,p)\big), \\
V_x(t,x,p) &= -e^{-\alpha x e^{r(T-t)}}\big(-\alpha \, e^{r(T-t)} g(t,p)\big), \\
V_{xx}(t,x,p) &= -e^{-\alpha x e^{r(T-t)}} \, \alpha^2 \, e^{2r(T-t)} g(t,p).
\end{aligned}
$$

Using this partial derivative and the relation

$$
\begin{aligned}
V\Big(t, x - b\sum_{i=1}^d y_i \mathbb{1}_D(i), p\Big) &= -\exp\Big\{-\alpha\Big(x - b\sum_{i=1}^d y_i \mathbb{1}_D(i)\Big)e^{r(T-t)}\Big\} g(t,p) \\
&= -e^{-\alpha x e^{r(T-t)}} \exp\Big\{\alpha \, b \sum_{i=1}^d y_i \mathbb{1}_D(i) e^{r(T-t)}\Big\} g(t,p),
\end{aligned}
$$

Equation (4.17) becomes

$$
\begin{aligned}
0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \Big\{ & \alpha\, x\, r\, e^{r(T-t)} g(t,p) + g_t(t,p) - \lambda\, g(t,p) \\
& - \alpha\, e^{r(T-t)} g(t,p)\big(r\,x + (\mu - r)\xi + c(b)\big) + \frac{1}{2}\sigma^2\,\alpha\, e^{2r(T-t)} g(t,p)\,\xi^2 \\
& + \lambda \sum_{D\subset\mathbb{D}} \sum_{k=1}^{m} a_k^D\, p_k\, g(t, J(p,D)) \int_{(0,\infty)^d} \exp\Big\{\alpha\, b\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y)\Big\},
\end{aligned}
$$

where the integral w.r.t. F is finite according to Assumption 4.3. The equation is equivalent to

$$
\begin{aligned}
0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \Big\{ & g_t(t,p) - \lambda\, g(t,p) \\
& - \alpha\, e^{r(T-t)} g(t,p)\Big((\mu - r)\xi + c(b) - \frac{1}{2}\sigma^2\,\alpha\, e^{r(T-t)}\,\xi^2\Big) \\
& + \lambda \sum_{D\subset\mathbb{D}} \sum_{k=1}^{m} a_k^D\, p_k\, g(t, J(p,D)) \int_{(0,\infty)^d} \exp\Big\{\alpha\, b\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y)\Big\},
\end{aligned}
\tag{4.21}
$$

which is the HJB equation for g. Since the terminal utility conditioned on $X_T^{\xi,b} = x$ yields $-e^{-\alpha x} g(T,p) = V(T,x,p) = U(x) = -e^{-\alpha x}$, a simple final condition follows for g:

$$
g(T,p) = 1, \quad p \in \Delta_m. \tag{4.22}
$$

Now the value function V can be determined by first carrying out the minimization in the HJB equation (4.21) (depending on g), then inserting the obtained minimum, omitting the infimum operator and solving the resulting ordinary differential equation (depending on p) with boundary condition $g(T) = 1$. However, g is probably not differentiable w.r.t. t because of the jump property of the state process. It has already been mentioned in the introduction that this difficulty can be overcome by using a weaker notion of a solution for the HJB equation (viscosity solution) or a weaker notion for differentiability, where we proceed with the second ansatz.

Assuming $t \mapsto g(t,p)$ is Lipschitz on $[0,T]$, we can replace g_t by a weaker notion of differentiability, namely Clarke's generalized subdifferential introduced in Section 2.1, compare Definition 2.6. To define the generalized subdifferential of $g(t,p)$ w.r.t. t, we introduce the following notation.

Notation. For some fixed $p \in \Delta_m$, we write $g_p : [0,T] \to (0,\infty)$ for the function given by

$$
g_p(t) := g(t,p), \quad t \in [0,T].
$$

Using the generalized subdifferential instead of g_t, Equation (4.21) becomes

$$
\begin{aligned}
0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \Big\{ & -\lambda\, g(t,p) - \alpha e^{r(T-t)} g(t,p)\Big((\mu - r)\xi + c(b) - \frac{1}{2}\sigma^2 \alpha e^{r(T-t)}\xi^2\Big) \\
& + \lambda \sum_{D\subset\mathbb{D}} \sum_{k=1}^{m} a_k^D\, p_k\, g(t, J(p,D)) \int_{(0,\infty)^d} \exp\Big\{\alpha\, b\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y)\Big\} \\
& + \inf_{\varphi\in\partial^C g_p(t)} \{\varphi\},
\end{aligned}
\tag{4.23}
$$

which is said to be the *generalized* HJB equation for g. Note that we set $\partial^C g_p(t) = \{g_p'(t)\}$

at the points where the derivative exists. The reason for this convention is that Clarke's generalized gradient reduces only to a singleton at those points where the the function is strictly differentiable (cp. Prop. 2.8); only differentiability is not sufficient. However, in the case of differentiability the corresponding gradient is contained in Clarke's generalized gradient. The reduction of the generalized gradient to a single-point set is required in Section 4.7.

To state the generalized HJB equation above in a compact way, we introduce the following operator.

Notation. Throughout this chapter, let $\mathscr{L}$ denote an operator acting on functions $g : [0,T] \times \Delta_m \to (0,\infty)$ and $(\xi, b) \in \mathbb{R} \times [0,1]$ which is given by

$$\begin{aligned} \mathscr{L} g(t,p;\xi,b) &:= -\lambda\, g(t,p) - \alpha\, e^{r(T-t)} g(t,p)\Big((\mu - r)\xi + c(b) - \frac{1}{2}\sigma^2 \alpha e^{r(T-t)}\xi^2\Big) \\ &\quad + \lambda \sum_{D \subset \mathbb{D}} \sum_{k=1}^{m} a_k^D\, p_k\, g(t, J(p,D)) \int_{(0,\infty)^d} \exp\Big\{\alpha\, b\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y). \end{aligned} \tag{4.24}$$

With the help of the operator $\mathscr{L}$, an equivalent representation of Equation (4.23) is

$$0 = \inf_{(\xi,b) \in \mathbb{R} \times [0,1]} \{\mathscr{L} g(t,p;\xi,b)\} + \inf_{\varphi \in \partial^C g_p(t)} \{\varphi\}. \tag{4.25}$$

Under the assumption that g is positive, the HJB Equation (4.25) can be written as

$$0 = -\lambda\, g(t,p) + \alpha\, e^{r(T-t)} g(t,p) \inf_{\xi \in \mathbb{R}} f_1(t,\xi) + \inf_{b \in [0,1]} f_2(t,p,b) + \inf_{\varphi \in \partial^C g_p(t)} \{\varphi\}, \tag{4.26}$$

where

$$f_1(t,\xi) := -(\mu - r)\xi + \frac{1}{2}\sigma^2\, \alpha\, e^{r(T-t)}\xi^2 \tag{4.27}$$

and

$$\begin{aligned} f_2(t,p,b) :=&\ -\alpha\, e^{r(T-t)}\, c(b)\, g(t,p) + \lambda \sum_{D \subset \mathbb{D}} \sum_{k=1}^{m} a_k^D p_k\, g(t, J(p,D)) \times \\ & \int_{(0,\infty)^d} \exp\Big\{\alpha\, b\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y). \end{aligned} \tag{4.28}$$

Therefore, in order to solve the generalized HJB equation for g, we must first solve the two minimizing problems. The obtained minimums $\xi^\star$ and $b^\star$, if they exist, are the candidates for an optimal investment and reinsurance strategy, respectively, which will be investigated in the next sections. Before determining these candidates, we complete the section with a discussion concerning the HJB equation.

Remark 4.26. If we use (4.21) (HJB equation without generalization of the partial derivative) to derive the value function, then the function g (which can be use to determine the value function V) is the solution of the equation

$$g_t(t,p) = \left(\lambda + \frac{1}{2}\frac{(\mu - r)^2}{\sigma^2} - \lambda\, f_2(b^\star(t,p,g))\right) g(t,p),$$

For a fixed $p \in \Delta_m$, this equation can be regarded as a linear ordinary differential equation (ODE) of first order. A solution of this ODE exists if the right-hand side is continuous in (t, g), compare Peano's existence theorem (see e.g. Ahmad and Ambrosetti

[1, Thm. 2.2.8]). But $b^\star$ depends on g in a very complicated way such that we can neither expect nor prove that $f_2(b^\star(t,p,g))$ is continuous. Therefore, we can not show that there exists a classical solution of the HJB Equation (4.23).

4.5 Candidate for an optimal investment strategy

The generalized HJB equation (4.25) yields as byproduct a candidate for an optimal investment strategy. Namely, due to the representation of the HJB equation given in (4.26), we have to minimize the function f_1 given by (4.27) w.r.t. ξ under the assumption that g is positive which is assumed to be satisfied in this section. First of all, we observe that f_1 does not depend on the parameters x and p. Therefore, the optimal investment strategy is independent of the surplus and the filter process.

In order to derive an optimal investment strategy at some time point, let us fix $t \in [0,T]$. It is plain from definition that $\mathbb{R} \ni \xi \mapsto f_1(t,\xi)$ is twice continuously differentiable w.r.t. ξ and

$$\frac{\partial}{\partial \xi} f_1(t,\xi) = -(\mu - r) + \sigma^2\,\alpha\, e^{r(T-t)}\xi, \quad \frac{\partial^2}{\partial \xi^2} f_1(t,\xi) = \sigma^2\,\alpha\, e^{r(T-t)} > 0.$$

Therefore, by setting the first derivative of f_1 to zero, we get the following unique minimizer $\xi^\star$ of f_1:

$$\xi^\star(t) = \frac{\mu - r}{\sigma^2}\frac{1}{\alpha} e^{-r(T-t)}, \quad t \in [0,T]. \tag{4.29}$$

The value of $f_1(t,\xi)$ at the minimum w.r.t. ξ is

$$\begin{aligned} f_1(t,\xi^\star(t)) &= -\frac{(\mu-r)^2}{\sigma^2}\frac{1}{\alpha}e^{-r(T-t)} + \frac{1}{2}\sigma^2\,\alpha\, e^{r(T-t)}\left(\frac{\mu-r}{\sigma^2}\frac{1}{\alpha}e^{-r(T-t)}\right)^2 \\ &= -\frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{1}{\alpha}e^{-r(T-t)}, \quad t\in[0,T]. \end{aligned} \tag{4.30}$$

Thus the calculated value $\xi^\star(t)$ is the only candidate of an optimal investment strategy at time t. We see that the unobservable thinning probabilities $\bar{\alpha}$ of our claim arrival model have no effect on the strategy because f_1 is independent of p. This was already expectable from the very beginning since the parameters of the financial market are not related to the unobservable parameters. That means, there is no connection between the financial and insurance risks at all, which indicates that we may use every financial market model where the solution of the optimal investment problem of maximizing the exponential utility of terminal wealth is already known. Indeed, the solution of the optimal investment problem in our framework is the well-know solution of the Merton problem with finite time horizon and exponential utility function (cf. e.g. Merton [93, Eq. (49)][3]), where $\frac{\mu-r}{\sigma}$ can be interpreted as the market price of risk.

Due to the simplicity of the financial market model, the optimal investment strategy is quite straightforward, while the candidate for the optimal reinsurance strategy, on the other side, is not straightforward as the insurance risk modelling is more complicated in comparison to the financial risk modelling. The reason for the higher complexity of the insurance market lies in particular in the unobservability of the dependencies between the various business risks, whereby we cannot expect an explicit solution for an optimal

[3]Equation (49) in Merton [93] states the optimal investment strategy for the more general case of the family of HARA (hyperbolic absolute risk-aversion) utility functions, which contains the exponential utility function, compare page 389 in Merton [93].

reinsurance strategy.

4.6 Candidate for an optimal reinsurance strategy

As in the previous section, g is supposed to be positive. Hence, according again to the HJB Equation (4.26), we have to minimize the function f_2 given by (4.28) to obtain a candidate for an optimal reinsurance strategy. The function f_2 includes the term $c(b)$, which depends on the reinsurance premium principle. Recall the assumptions made in Section 3.5 that

$$c(b) = (\eta - \theta)\kappa + (1+\theta)\kappa\, b$$

and thus

$$\begin{aligned} f_2(t,p,b) &= -\alpha\, e^{r(T-t)}\, g(t,p)\,(\eta-\theta)\kappa - \alpha e^{r(T-t)}\, g(t,p)\,(1+\theta)\,\kappa\, b \\ &+ \lambda \sum_{D\subset\mathbb{D}} g(t, J(p,D)) \sum_{k=1}^{m} a_k^D\, p_k \int_{(0,\infty)^d} \exp\left\{\alpha\, b\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_D(i)\right\} F(\mathrm{d}y). \end{aligned} \tag{4.31}$$

The next lemma yields the first order condition for a candidate of an optimal reinsurance strategy.

Lemma 4.27. *For any $(t,p) \in [0,T]\times\Delta_m$, the function $\mathbb{R} \ni b \mapsto f_2(t,p,b)$ is strictly convex and*

$$\begin{aligned} \frac{\partial}{\partial b} f_2(t,p,b) = -\alpha\, e^{r(T-t)}\Bigg(g(t,p)\,(1+\theta)\,\kappa - \lambda \sum_{D\subset\mathbb{D}} g(t,J(p,D)) \sum_{k=1}^{m} a_k^D\, p_k \times \\ \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{\alpha\, b\, e^{r(T-t)} \sum_{j=1}^{d} y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y)\Bigg). \end{aligned}$$

Proof. We first focus on showing that the differentiation w.r.t. b and integration w.r.t. F interchange in (4.31). For this purpose, we fix $t \in [0,T]$ and $D \subset \mathbb{D}$. Moreover, we define a function by

$$h(y,b) := \exp\left\{\alpha\, b\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_D(i)\right\}, \quad y := (y_1,\ldots,y_d) \in (0,\infty)^d, \quad b \in \mathbb{R}.$$

For any $b \in \mathbb{R}$, $\int_{(0,\infty)^d} |h(y,b)|\, F(\mathrm{d}y) < \infty$ according to Assumption 4.3. Furthermore, for every $y \in (0,\infty)^d$, the map $b \mapsto h(y,b)$ is obviously differentiable w.r.t. b and the partial derivative is given by

$$\frac{\partial}{\partial b} h(y,b) = \alpha\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_D(i) \exp\left\{\alpha\, b\, e^{r(T-t)} \sum_{j=1}^{d} y_j \mathbb{1}_D(j)\right\},$$

where

$$\left|\frac{\partial}{\partial b} h(y,b)\right| \le \alpha\, e^{r(T-t)} \sum_{i=1}^{d} y_i \exp\left\{\alpha\, b\, e^{r(T-t)} \sum_{j=1}^{d} y_j\right\},$$

From Lemma 4.5 (i) follows that the right-hand side is integrable w.r.t. F. Therefore, we can apply Theorem 6.28 from Klenke [77], which yields

$$\frac{\partial}{\partial b}\int_{(0,\infty)^d} h(y,b)\,F(\mathrm{d}y) = \int_{(0,\infty)^d} \frac{\partial}{\partial b} h(y,b)\,F(\mathrm{d}y).$$

Hence

$$\begin{aligned}\frac{\partial}{\partial b} f_2(t,p,b) &= -\alpha\, g(t,p)\,(1+\theta)\,\kappa\, e^{r(T-t)} + \alpha\, e^{r(T-t)}\lambda \sum_{D\subset\mathbb{D}} g(t,J(p,D)) \sum_{k=1}^{m} a_k^D p_k \times \\ &\quad \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{\alpha\, b\, e^{r(T-t)} \sum_{j=1}^{d} y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y) \\ &= -\alpha\, e^{r(T-t)} \Bigg(g(t,p)\,(1+\theta)\,\kappa - \lambda \sum_{D\subset\mathbb{D}} g(t,J(p,D)) \sum_{k=1}^{m} a_k^D\, p_k \times \\ &\qquad \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{\alpha\, b\, e^{r(T-t)} \sum_{j=1}^{d} y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y) \Bigg).\end{aligned}$$

The convexity can be seen without calculating the second derivative. Indeed, the integral w.r.t. F is strictly convex since the integrand is strictly convex (both w.r.t. b). Due to the convexity of the sum of linear and convex functions, we obtain that $b \mapsto f_2(t,p,b)$ is convex. □

The previous lemma provides a criterion for a candidate of an optimal reinsurance strategy by setting $\frac{\partial}{\partial b} f_2$ to zero. It is convenient to define the following function to express the first order condition.

Notation. Let $(t,p) \in [0,T] \times \Delta_m$ and $b \in \mathbb{R}$. We define

$$\begin{aligned} h_{\lambda,F}(t,p,b) := \lambda \sum_{D\subset\mathbb{D}} \frac{g(t,J(p,D))}{g(t,p)} \sum_{k=1}^{m} a_k^D p_k \sum_{i=1}^{d} \mathbb{1}_D(i) \times \\ \int_{(0,\infty)^d} y_i \exp\left\{\alpha\, b\, e^{r(T-t)} \sum_{j=1}^{d} y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y). \end{aligned} \tag{4.32}$$

Furthermore, we set

$$\begin{aligned} A_{\lambda,F}(t,p) &:= h_{\lambda,F}(t,p,0), \\ B_{\lambda,F}(t,p) &:= h_{\lambda,F}(t,p,1). \end{aligned} \tag{4.33}$$

We take the dependencies on the fixed background intensity λ and the claim size distribution F by the lower indices (λ, F) into account.

The preceding lemma leads to the following first order condition for the optimal reinsurance strategy:

$$(1+\theta)\,\kappa = h_{\lambda,F}(t,p,b). \tag{4.34}$$

By establishing this equation w.r.t. b we obtain a unique minimizer of f_2 w.r.t. b (if a solution exists) because of the strict convexity of f_2 w.r.t. b.

The next proposition states that this equation is solvable and that the solution takes values in $[0,1]$ depending on the safety loading parameter θ of the reinsurer. Before presenting this crucial statement, let us discuss an alternative reinsurance premium model.

Remark 4.28. We have modelled the reinsurance premium rate by $(1+\theta)\,\kappa$, compare Section 3.5. Using the expected value principle, we obtain by Proposition 4.21 that

$$\kappa = \mathbb{E}[\mathrm{d}S_t] = \lambda \sum_{D\subset\mathbb{D}} \sum_{k=1}^{m} a_k^D \pi_{\bar{\alpha}}(k) \sum_{i=1}^{d} \mathbb{1}_D(i)\mathbb{E}[Y_1^i],$$

where $\sum_{k=1}^m a_k^D \pi_{\bar{\alpha}}(k) = \mathbb{E}[\alpha_D]$. Instead of using $\mathbb{E}[\alpha_D]$ it is thinkable to take the available information about the thinning probabilities into account by replacing $\mathbb{E}[\alpha_D]$ with the left-hand limit process of $\mathbb{E}[\alpha_D \,|\, \mathcal{F}_t^\Phi]$, i.e. with $\sum_{k=1}^m a_k^D p_k(t-)$. That is, κ depends on the filter process:

$$\kappa_{\lambda,F}(p) := \lambda \sum_{D\subset\mathbb{D}} \sum_{k=1}^{m} a_k^D p_k \sum_{i=1}^{d} \mathbb{1}_D(i)\mathbb{E}[Y_1^i].$$

Reinsurance premium principles depending on filter processes are used in Liang and Bayraktar [85] and Brachetta and Ceci [18, Example 2.3]. With this model the reinsurance premium rate is given by $(1+\theta)\,\kappa_{\lambda,F}(p_{t-})$, in which the left-hand limit process of the filter ensures predictability. In fact, the reinsurance premium is time-dependent.[4] Using $\kappa_{\lambda,F}(p)$ instead of κ it can be shown that $A_{\lambda,F}(t,p)/\kappa_{\lambda,F}(p) \geq 1$. The proof of this fact is similar to the procedure in the proof of Theorem 4.41 under use of the order $a_1 \preceq a_2 \preceq \ldots \preceq a_m$, where $\preceq$ is a preorder[5] on the set A defined by

$$a_k \preceq a_\ell \quad :\Longleftrightarrow \quad \sum_{D\subset\mathbb{D}} a_k^D \sum_{i=1}^{d} \mathbb{1}_D(i)\mathbb{E}[Y_1^i] \leq \sum_{D\subset\mathbb{D}} a_\ell^D \sum_{i=1}^{d} \mathbb{1}_D(i)\mathbb{E}[Y_1^i] \tag{4.35}$$

for every $k,\ell \in \{1,\ldots,m\}$, which is no loss of generality. The vector $a_j = (a_j^D)_{D\subset\mathbb{D}} \in \Delta_\ell$ can be seen as weights and thus $\sum_{D\subset\mathbb{D}} a_j^D \sum_{i=1}^d \mathbb{1}_D(i)\mathbb{E}[Y_1^i]$ as the weighted sum of $(\sum_{D\subset\mathbb{D}} \mathbb{E}[Y_1^i]\mathbb{1}_D(i))_{D\subset\mathbb{D}}$ with weights a_j. So (4.35) is an order of the weights $A = \{a_1,\ldots,a_m\}$ such that the weighted sum of the sum of the expected claims of the LoBs $i \in D$ are in increasing order. This can be interpreted as an order from the best to the worst case scenario from the insurer's point of view. Such a order is typically required to derive the comparison result of optimal strategies in the case of partial and full information. (In fact, the order introduced above is too weak to develop a comparison, compare Section 4.8.2.) Using the premium principle introduced above the optimal reinsurance strategy is determined by (4.36) with κ replaced by $\kappa_{\lambda,F}(p)$. Moreover, an analogous comparison result applies as in Corollary 4.42.

We continue with the result which yields the candidate for an optimal reinsurance strategy.

Proposition 4.29. *For any $(t,p) \in [0,T]\times\Delta_m$, Equation (4.34) has a unique root, denoted by $r_{\lambda,F}(t,p)$, which is increasing w.r.t. the safety loading parameter θ. Moreover, it holds*

(i) $r_{\lambda,F}(t,p) \leq 0$ *if* $\theta \leq A_{\lambda,F}(t,p)/\kappa - 1$,

(ii) $0 < r_{\lambda,F}(t,p) < 1$ *if* $A_{\lambda,F}(t,p)/\kappa - 1 < \theta < B_{\lambda,F}(t,p)/\kappa - 1$,

(iii) $r_{\lambda,F}(t,p) \geq 1$ *if* $\theta \geq B_{\lambda,F}(t,p)/\kappa - 1$.

[4] A time-dependent reinsurance premium is used e.g. in Peng and Hu [98].

[5] A *preorder* is a reflexive and transitive binary relation, cf. Bäuerle and Rieder [32, Def. B.3.1].

Proof. The argumentation is inspired by the techniques used in the proofs of Lemma 4.2 and Proposition 4.1 in Liang and Bayraktar [85]. Fix $(t,p) \in [0,T] \times \Delta_m$ and recall $(1+\theta)\,\kappa > 0$ as well as $h_{\lambda,F}(t,p,b) > 0$ for all $b \in \mathbb{R}$. Moreover, the function $\mathbb{R} \ni b \mapsto h_{\lambda,F}(t,p,b)$ is strictly convex (follows easily from the same arguments as for the strict convexity of f_2 w.r.t. b, compare Lemma 4.27) and, consequently, continuous. It is also a simple matter to see that $\mathbb{R} \ni b \mapsto h_{\lambda,F}(t,p,b)$ is strictly increasing, i.e. $\frac{\partial}{\partial b} h_{\lambda,F}(t,p,b) > 0$ for all $b \in \mathbb{R}$. Furthermore, $h_{\lambda,F}$ holds, by the dominated convergence theorem, which can be applied due to Lemma 4.5 (i), and by the convexity of $b \mapsto h_{\lambda,F}(t,p,b)$

$$\lim_{b\to-\infty} h_{\lambda,F}(t,p,b) = 0, \quad \lim_{b\to\infty} h_{\lambda,F}(t,p,b) = \infty.$$

Therefore, $h_{\lambda,F}(t,b)$ and $(1+\theta)\,\kappa$ have a unique point of intersection w.r.t. b, i.e. (4.34) has a unique root. From now on, $r_{\lambda,F}(t,p)$ denotes this unique root of (4.34). Notice that $r_{\lambda,F}(t,p)$ depends on the safety loading parameter θ. By regarding $r_{\lambda,F}(t,p)$ as a function $0 < \theta \mapsto r_{\lambda,F}(t,p,\theta)$, differentiation of both sides of (4.34) w.r.t. θ yields

$$\kappa = \frac{\partial}{\partial b} h_{\lambda,F}(t,p,b)\, \frac{\partial}{\partial \theta} r_{\lambda,F}(t,p,\theta)$$

Thus,

$$\frac{\partial}{\partial \theta} r_{\lambda,F}(t,p,\theta) = \frac{\kappa}{\frac{\partial}{\partial b} h_{\lambda,F}(t,p,b)} > 0.$$

That is, $\theta \mapsto r_{\lambda,F}(t,p,\theta)$ is increasing. Furthermore, we observe that

$$A_{\lambda,F}(t,p) = h_{\lambda,F}(t,p,0) < h_{\lambda,F}(t,p,1) = B_{\lambda,F}(t,p)$$

according to the strict increasing property of $b \mapsto h_{\lambda,F}(t,p,b)$. Thus the cases for θ given in the statement are all cases which can occur. In the case $1+\theta \le A_{\lambda,F}(t,p)/\kappa$, we obtain

$$h_{\lambda,F}(t,p,0) = A_{\lambda,F}(t,p) \ge (1+\theta)\,\kappa.$$

Due to the properties of $h_{\lambda,F}$ (strictly increasing, continuous and $\lim_{b\to-\infty} h_{\lambda,F}(t,p,b) = 0$), assertion (i) follows, compare (4.34). Dealing with the case $A_{\lambda,F}(t,p)/\kappa < 1+\theta < B_{\lambda,F}(t,p)/\kappa$, we have

$$h_{\lambda,F}(t,p,1) = B_{\lambda,F}(t,p) > (1+\theta)\,\kappa,$$

which implies statement (ii). In the same manner, we can see that $r_{\lambda,F}(t,p) \ge 1$, if $\theta \ge B_{\lambda,F}(t,p)/\kappa - 1$, which finalizes the proof. □

Notation. Throughout this chapter, $r_{\lambda,F}(t,p)$ denotes the unique root from Proposition 4.29.

It should be noted that the cases (i) and (ii) for θ can be empty sets if $A_{\lambda,F}(t,p)/\kappa \le 1$ and $B_{\lambda,F}(t,p)/\kappa \le 1$, respectively. It has already been noticed that with an alternative reinsurance premium principle, in which κ depends on the filter process, one can prove the property $A_{\lambda,F}(t,p)/\kappa(t,p) > 1$. However, we will proceed with the introduced reinsurance premium model with constant κ.

Under consideration of the preceding propositions and the $[0,1]$-valuation of a reinsurance strategy, the candidate for an optimal reinsurance strategy can be specified as follows: For any $(t,p) \in [0,T] \times \Delta_m$, we set

$$b_{\lambda,F}(t,p) := \begin{cases} 0, & \theta \le A_{\lambda,F}(t,p)/\kappa - 1, \\ 1, & \theta \ge B_{\lambda,F}(t,p)/\kappa - 1, \\ r_{\lambda,F}(t,p), & \text{otherwise.} \end{cases} \tag{4.36}$$

The candidate for an optimal reinsurance strategy $(b^\star_{\lambda,F}(t))_{t\in[0,T]}$ is thus given by $b^\star_{\lambda,F}(t) := b_{\lambda,F}(t-, p_{t-})$, which is unique because of the uniqueness on $r_{\lambda,F}(t,p)$.

Remark 4.30. According to Proposition 4.29, the value of the optimal reinsurance strategy increases as the value of θ increases since $r_{\lambda,F}(t,p)$ is increasing w.r.t. θ. That is, the cedent expect that the reinsurer pays a greater part of each claim when the reinsurance premium increases. The increasing property of the reinsurance strategy w.r.t. the safety loading parameter θ can also be considered as the consequence of the law of demand which claims that the higher the price the lower the volume demanded, compare Brachetta and Ceci [19]. If $\theta \le A_{\lambda,F}(t,p) - 1$, then the reinsurance premium seems so cheap for the cedent such that she/he chooses a retention level for zero which means that the reinsurer compensates any potential insurance damage at time t. On the other hand, when $\theta \ge B_{\lambda,F}(t,p) - 1$, the reinsurance premium is too expensive for the cedent. As a result the cedent prefers to retain all the risks to itself.

So far we only derived heuristically, under non-trivial assumptions, candidates for a value function as well as an optimal investment-reinsurance strategy. Instead of proceeding by verifying the made assumption and thus showing that the HJB equation follows from the stochastic control problem, we turn the story upside down and start with the HJB equation and assume there exist a solution. This procedure is standard in stochastic control theory and known as verification. Thus in the remainder of this chapter, we will verify that the value function is indeed determined by the HJB equation and that the candidate derived for an investment-reinsurance strategy is indeed optimal.

4.7 Verification

4.7.1 The verification theorem

The following formulated verification theorem collects all necessary assumptions on a function defined on $[0,T] \times \Delta_m$, which have to be satisfied such that the value function can be represented by this function and the developed candidate for an optimal strategy is indeed optimal.

Theorem 4.31. *Suppose there exists a bounded function $h : [0,T] \times \Delta_m \to (0,\infty)$ such that $t \mapsto h(t,p)$ is Lipschitz on $[0,T]$ for all $p \in \Delta_m$, $p \mapsto h(t,p)$ is continuous on Δ_m for all $t \in [0,T]$ and h satisfies the generalized HJB equation*

$$0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \{\mathscr{L}h(t,p;\xi,b)\} + \inf_{\varphi\in\partial^C h_p(t)} \{\varphi\}, \tag{4.37}$$

for all $(t,p) \in [0,T) \times \Delta_m$ with boundary condition

$$h(T,p) = 1, \quad p \in \Delta_m. \tag{4.38}$$

Then

$$V(t,x,p) = -e^{-\alpha x e^{r(T-t)}} h(t,p), \quad (t,x,p) \in [0,T] \times \mathbb{R} \times \Delta_m,$$

and $(\xi^\star, b^\star_{\lambda,F}) = (\xi^\star(s), b^\star_{\lambda,F}(s))_{s\in[t,T]}$ *with* $\xi^\star(s)$ *given by* (4.29) *and* $b^\star_{\lambda,F}(s) := b_{\lambda,F}(s-, p_{s-})$ *given by* (4.36) *(with* g *replaced by* h *in* $A_{\lambda,F}(s,p)$ *and* $B_{\lambda,F}(s,p)$*) is an optimal feedback strategy for the given optimization problem* (P1)*, i.e.* $V(t,x,p) = V^{\xi^\star, b^\star_{\lambda,F}}(t,x,p)$.

One aspect that has to be dealt with in the proof of the verification theorem is a change of measure to show a martingale property. For the definition of the change of measure (see Lemma A.3), we have to restrict the set of admissible strategies such that the investment strategy is bounded and continuous, meanwhile there is no interdependency between the investment and the reinsurance strategy, compare proof of Lemma A.3. But it will become apparent that this is not a restriction.

Notation. Throughout this chapter, we set, for any $t \in [0,T)$,

$$\begin{aligned}\widetilde{\mathcal{U}}[t,T] := \Big\{&(\xi,b) \in \mathcal{U}[t,T] :\ \exists\, K > 0 : |\xi_s| \le K \ \forall\, s \in [t,T],\\ &\xi = (\xi_s)_{s\in[t,T]} \text{ is continuous and } \mathfrak{F}^W\text{-adapted}, b = (b_s)_{s\in[t,T]} \text{ is } \mathfrak{F}^\Psi\text{-predictable}\Big\},\end{aligned} \tag{4.39}$$

i.e. the control set of $(\xi,b) \in \widetilde{\mathcal{U}}[t,T]$ is $[-K,K] \times [0,1]$. Moreover, we set

$$\widetilde{V}(t,x,p) := \sup_{(\xi,b)\in\widetilde{\mathcal{U}}[t,T]} V^{\xi,b}(t,x,p), \quad (t,x,p) \in [0,T] \times \mathbb{R} \times \Delta_m. \tag{4.40}$$

In order to simplify notation we introduce the following operator.

Notation. We define an operator $\mathscr{H}$ acting on functions $v : [0,T] \times \Delta_m \to (0,\infty)$ and $(\xi,b) \in \mathbb{R} \times [0,1]$ by

$$\mathscr{H} v(t,p;\xi,b) := \mathscr{L} v(t,p;\xi,b) + v_t(t,p) \tag{4.41}$$

for all functions $v : [0,T] \times \Delta_m \to (0,\infty)$, where the right-hand side is well-defined.

Using this notation, the HJB equation (4.25) can be written as

$$0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \{\mathscr{H} g(t,p;\xi,b)\}$$

at those points t, where g is differentiable w.r.t. t. Here is used the convention that $\partial^C g_p(t) = \{g_p'(t)\}$ at the points t, where $g_p'(t)$ exists.

Before considering the proof of the preceding theorem, we would like to point out that the preliminary result stated in Lemma A.8 plays an important role in the proof of the verification theorem which is essentially shown by tools from stochastic analysis. Now, we are in the position to prove the verification theorem.

Proof of Theorem 4.31. Let $h : [0,T] \times \Delta_m \to (0,\infty)$ be a function satisfying the conditions stated in the theorem. Notice that every Lipschitz function is also absolutely continuous, compare Lemma 2.46. We set, for any $(t,x,p) \in [0,T] \times \mathbb{R} \times \Delta_m$,

$$f(t,x) := -e^{-\alpha x e^{r(T-t)}} \quad \text{and} \quad G(t,x,p) := f(t,x)\, h(t,p).$$

Let us fix $t \in [0,T]$ and $(\xi,b) \in \widetilde{\mathcal{U}}[t,T]$. From Lemma A.8, it follows

$$G(T, X_T^{\xi,b}, p_T) = G(t, X_t^{\xi,b}, p_t) + \int_t^T f(s, X_s^{\xi,b})\, \mathscr{H} h(s, p_s; \xi_s, b_s)\, \mathrm{d}s + \eta_T^{\xi,b} - \eta_t^{\xi,b}, \tag{4.42}$$

where $(\eta_t^{\xi,b})_{t\in[0,T]}$ is a $\mathfrak{G}$-martingale and we set $\mathscr{H}h(s,p_s;\xi_s,b_s)$ to zero at those points $s\in[t,T]$, where $h_t(s,p_s)$ does not exist. Notice that h is partial differentiable w.r.t. t almost everywhere in the sense of the Lebesgue measure according to the absolute continuity of $t\mapsto h(t,p)$ for all $p\in\Delta_m$. The generalized HJB equation (4.37) implies

$$\mathscr{H}h(s,p_s;\xi_s,b_s)\geq 0,\quad s\in[t,T].$$

In consequence

$$\int_t^T f(s,X_s^{\xi,b})\,\mathscr{H}h(s,p_s;\xi_s,b_s)\,\mathrm{d}s\leq 0,$$

due to the negativity of f. Thus, by (4.42), we get

$$G(T,X_T^{\xi,b},p_T)\leq G(t,X_t^{\xi,b},p_t)+\eta_T^{\xi,b}-\eta_t^{\xi,b}.$$

Using the boundary condition (4.38), we obtain

$$G(T,x,p)=f(T,x)\,h(T,p)=f(T,x)=-e^{-\alpha x}.$$

Hence

$$U(X_T^{\xi,b})=G(T,X_T^{\xi,b},p_T)\leq G(t,X_t^{\xi,b},p_t)+\eta_T^{\xi,b}-\eta_t^{\xi,b}.$$

Now, we take the regular conditional expectation given $X_t^{\xi,b}=x$ and $p_t=p$ on both sides which yields

$$\mathbb{E}^{t,x,p}\big[U(X_T^{\xi,b})\big]\leq G(t,x,p)$$

since $(\eta_s^{\xi,b})_{s\in[t,T]}$ is a $\mathfrak{G}$-martingale. Taking the supremum over all investment-reinsurance strategies $(\xi,b)\in\widetilde{\mathscr{U}}[t,T]$, we obtain

$$\widetilde{V}(t,x,p)\leq G(t,x,p). \tag{4.43}$$

As already seen in (4.26), we can rewrite the generalized HJB equation:

$$\begin{aligned}0&=\inf_{(\xi,b)\in\mathbb{R}\times[0,1]}\{\mathscr{L}h(s,p;\xi,b)\}+\inf_{\varphi\in\partial^C h_p(s)}\{\varphi\}\\&=-\lambda\,h(s,p)+\alpha\,e^{r(T-s)}h(s,p)\inf_{\xi\in\mathbb{R}}f_1(s,\xi)+\inf_{b\in[0,1]}f_2(s,p,b)+\inf_{\varphi\in\partial^C h_p(s)}\{\varphi\},\end{aligned} \tag{4.44}$$

where f_1 is defined by (4.27) and f_2 by (4.28). Moreover, $\xi^\star(s)$ given by (4.29) is the unique minimizer of f_1 on $\mathbb{R}$ and $b_{\lambda,F}(s,p)$ given by (4.36) (with g replaced by h in $A_{\lambda,F}(s,p)$ and $B_{\lambda,F}(s,p)$) is the unique minimizer of f_2 on $[0,1]$. Since $\partial^C h_p(s)$ is a compact subset of $\mathbb{R}$ (cf. Prop. 2.7), we know the $\partial^C h_p(s)$ contains its infimum which we denote by $\varphi^\star(s,p)$. Therefore,

$$\mathscr{L}h(s,p;\xi^\star(s),b_{\lambda,F}(s,p))+\varphi^\star(s,p)=0.$$

We set $b^\star_{\lambda,F}(s):=b_{\lambda,F}(s-,p_{s-})$ for every $s\in[t,T]$. One aspect that has to be dealt with is to show that $(\xi^\star,b^\star)$ is an admissible investment-reinsurance strategy. First, we observe that $(\xi^\star(s))_{s\in[t,T]}$ is a continuous, bounded by $\frac{|\mu-r|}{\sigma^2}\frac{1}{\alpha}$ and deterministic (i.e. $\mathfrak{F}^W$-adapted) process. Secondly, we note that $(t,\omega)\mapsto h(t,p_{t-}(\omega))$ and $(t,\omega)\mapsto h(t,J(p_{t-}(\omega),D))$ are $\mathcal{P}(\mathfrak{F}^\Psi)$-measurable for all $D\subset\mathbb{D}$, compare the arguments in proof of Lemma A.8. This implies that $(t,\omega)\mapsto A_{\lambda,F}(t,p_{t-}(\omega))$ and $(t,\omega)\mapsto B_{\lambda,F}(t,p_{t-}(\omega))$ defined by (4.33) with g replaced by h are $\mathfrak{F}^\Psi$-predictable. Moreover, the $\mathcal{P}(\mathfrak{F}^\Psi)$-measurability of right-hand side of (4.32) has the consequence that the unique root $(t,\omega)\mapsto r_{\lambda,F}(t,p_{t-}(\omega))$ is

$\mathcal{P}(\mathfrak{F}^\Psi)$-measurable. In summary, $(t,\omega) \mapsto b_{\lambda,F}(t, p_{t-}(\omega))$ is $\mathcal{P}(\mathfrak{F}^\Psi)$-measurable. That is, $(b_{\lambda,F}(s))_{s\in[t,T]}$ is an $\mathfrak{F}^\Psi$-predictable, $[0,1]$-valued process and thus $(\xi^\star, b^\star) \in \widetilde{\mathcal{U}}[t,T]$. So we can deduce that

$$\mathscr{H} h(s, p_s; \xi^\star(s), b^\star_{\lambda,F}(s)) = 0, \quad s \in [t,T].$$

This implies

$$\int_t^T f(s, X_s^{\xi^\star,b^\star})\, \mathscr{H} h(s, p_s; \xi^\star(s), b^\star_{\lambda,F}(s))\, \mathrm{d}s = 0.$$

Consequently,

$$U(X_T^{\xi^\star,b^\star}) = G(T, X_T^{\xi^\star,b^\star}, p_T) = G(t, X_t^{\xi^\star,b^\star}, p_t) + \eta_T^{\xi^\star,b^\star} - \eta_t^{\xi^\star,b^\star}.$$

Again, taking the regular conditional expectation given $X_t^{\xi^\star,b^\star} = x$ and $p_t = p$ on both sides then yields

$$\mathbb{E}^{t,x,p}\big[U(X_T^{\xi^\star,b^\star})\big] = G(t,x,p).$$

That is,

$$V^{\xi^\star,b^\star}(t,x,p) = G(t,x,p) = -e^{-\alpha x e^{r(T-t)}} h(t,p),$$

which implies

$$\widetilde{V}(t,x,p) = \sup_{(\xi,b)\in\widetilde{\mathcal{U}}[t,T]} V^{\xi,b}(t,x,p) = V^{\xi^\star,b^\star}(t,x,p) = -e^{-\alpha x e^{r(T-t)}} h(t,p).$$

From the representation of the generalized HJB equation given in (4.44), it follows directly that the optimal investment-reinsurance strategy is an element of $\widetilde{\mathcal{U}}[t,T]$ (compare the arguments above). Therefore,

$$\sup_{(\xi,b)\in\widetilde{\mathcal{U}}[t,T]} V^{\xi,b}(t,x,p) = \sup_{(\xi,b)\in\mathcal{U}[t,T]} V^{\xi,b}(t,x,p)$$

and thus

$$V(t,x,p) = -e^{-\alpha x e^{r(T-t)}} h(t,p)$$

and the proof is complete. □

The verification theorem has validated the optimality of the candidate solution of the generalized HJB equation and thus the optimality of the candidate investment-reinsurance strategy. Moreover the theorem shows that the solution of the generalized HJB equation (if a solution exists) is unique since the solution determines the value function (in a unique way), which is unique.

While the preceding verification step follows a fairly standardized procedure, the proof of existence of a solution to the HJB equation given in the next section is quite individual for each problem.

4.7.2 Existence result for the value function

This section is devoted to the existence of the solution of the generalized HJB equation and thus of the optimal strategy. More precisely, we will show the existence of a function $h : [0,T] \times \Delta_m \to (0,\infty)$ satisfying conditions stated in Theorem 4.31. To see this, let us introduce the following function $\widetilde{g}$.

Notation. We set

$$\widetilde{g}(t,p) := \inf_{(\xi,b)\in\widetilde{\mathcal{U}}[t,T]} g^{\xi,b}(t,p), \quad (t,p)\in[0,T]\times\Delta_m, \tag{4.45}$$

where $g^{\xi,b}$ is given by (4.20) and $\widetilde{\mathcal{U}}[t,T]$ by (4.39).

The next lemma yields useful properties of $\widetilde{g}$.

Lemma 4.32. *The function $\widetilde{g}$ defined by* (4.45) *has the following properties:*

(i) $\widetilde{g}(t,p) > 0$ *for all* $(t,p)\in[0,T]\times\Delta_m$.

(ii) $\widetilde{\mathcal{U}}[0,T]\ni(\xi,b)\mapsto g^{\xi,b}(0,p)$ *is bounded for all* $p\in\Delta_m$.

(iii) There exists a constant $0 < K_3 < \infty$ *such that* $|\widetilde{g}(t,p)|\le K_3$ *for all* $(t,p)\in[0,T]\times\Delta_m$.

(iv) $\Delta_m\ni p\mapsto\widetilde{g}(t,p)$ *is concave on for all* $t\in[0,T]$.

(v) $[0,T]\ni t\mapsto\widetilde{g}(t,p)$ *is Lipschitz on* $[0,T]$ *for all* $p\in\Delta_m$.

(vi) Let M be the set of all points $(t,p)\in[0,T]\times\Delta_m$*, where the partial derivatives of $\widetilde{g}$ w.r.t. t exist. Then there exists a constant* $0<K_4<\infty$ *such that* $|\widetilde{g}_t(t,p)|\le K_4$ *for all* $(t,p)\in M$.

(vii) There exists a constant $0<K_5<\infty$ *such that* $\big|\mathscr{L}\widetilde{g}(t,p;\xi,b)\big|\le K_5$ *for all* $(t,p)\in[0,T]\times\Delta_m$ *and* $(\xi,b)\in[-K,K]\times[0,1]$.

(viii) There exists a constant $0<K_6<\infty$ *such that* $\big|\inf_{(\xi,b)\in[-K,K]\times[0,1]}\mathscr{L}\widetilde{g}(t,p;\xi,b)\big|\le K_6$ *for all* $(t,p)\in[0,T]\times\Delta_m$.

Proof. (i) Fix $(t,p)\in[0,T]\times\Delta_m$ and $(\xi,b)\in\widetilde{\mathcal{U}}[t,T]$. Using the change of measure introduced in Lemma A.3, it follows from the definition of $g^{\xi,b}$ given in (4.20) that

$$\begin{aligned} g^{\xi,b}(t,p) = \mathbb{E}^{t,p}_{\mathbb{Q}^{\xi,b}_t}\Bigg[\exp\Bigg\{\int_t^T\Bigg(&-\alpha\, e^{r(T-t)}\Big((\mu-r)\xi_s + c(b_s) - \frac{1}{2}\alpha\,\sigma^2\, e^{r(T-s)}\xi_s^2\Big)\\ &+\lambda\sum_{D\subset\mathbb{D}} p_s^D\int_{(0,\infty)^d}\exp\Big\{\alpha\, b_s\, e^{r(T-t)}\sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\}F(\mathrm{d}y)-\lambda\Bigg)\mathrm{d}s\Bigg\}\Bigg]. \end{aligned}$$

Since

$$\begin{aligned} \alpha\, e^{r(T-s)}(\mu-r)\xi_s &\le \alpha\, e^{rT}|\mu-r|\,K,\\ \alpha\, e^{r(T-s)}c(b_s) &\le \alpha\, e^{rT}\, c(1) = \alpha\, e^{rT}\,(1+\eta)\,\kappa, \end{aligned} \tag{4.46}$$

for all $s\in[t,T]$, we obtain

$$g^{\xi,b}(t,p)\ge\exp\big\{-\alpha\, T\, e^{rT}\big(|\mu-r|\,K+(1+\eta)\,\kappa\big)-\lambda\, T\big\} =: C > 0.$$

Hence, due to the arbitrariness of $(\xi,b)\in\widetilde{\mathcal{U}}[t,T]$, the infimum of $g^{\xi,b}(t,p)$ over all $(\xi,b)\in\widetilde{\mathcal{U}}[t,T]$ is greater than or equal to C which yields the statement by definition of $\widetilde{g}$ given in (4.45).

(ii) Fix $p\in\Delta_m$. Following Bäuerle and Rieder [31, Proof of Lemma 6.1 c)] and Liang and Bayraktar [85, Proof of Lemma 4.4 (b)], respectively, we make use of a measurement change to show the announced statement similar to the proof of

Lemma A.1. We use the probability measure $\mathbb{Q}_T^{\xi,b}$ introduced in Lemma A.3, which yields

$$\begin{aligned}|g^{\xi,b}(0,p)| &= \mathbb{E}^{0,p}\Bigg[\exp\Bigg\{-\int_0^T \alpha e^{r(T-s)}((\mu-r)\xi_s + c(b_s))\,\mathrm{d}s - \int_0^T \alpha\sigma e^{r(T-s)}\xi_s\,\mathrm{d}W_s \\ &\qquad + \int_0^T\int_{(0,\infty)^d} \alpha\, b_s\, e^{r(T-s)}\sum_{i=1}^d y_i \mathbb{1}_z(i)\,\Psi(\mathrm{d}s,\mathrm{d}(y,z))\Bigg\}\Bigg] \\ &= \mathbb{E}^{0,p}_{\mathbb{Q}_T^{\xi,b}}\Bigg[\exp\Bigg\{\int_0^T\Bigg(-\alpha\, e^{r(T-s)}((\mu-r)\xi_s + c(b_s)) + \frac{1}{2}\alpha^2\,\sigma^2\, e^{2r(T-s)}\xi_s^2 \\ &\qquad + \lambda\sum_{D\subset\mathbb{D}} p_s^D \int_{(0,\infty)^d}\exp\Bigg\{\alpha\, b_s\, e^{r(T-s)}\sum_{i=1}^d y_i\mathbb{1}_D(i)\Bigg\}F(\mathrm{d}y)\Bigg)\mathrm{d}s - \lambda T\Bigg\}\Bigg].\end{aligned}$$

By similar arguments as in the proof of Lemma A.4, we obtain that the expectation above is bounded by a finite constant independent of (ξ,b).

(iii) The statement follows immediately from Lemma 4.25 (iii) since $\widetilde{\mathcal{U}}[t,T]\subset\mathcal{U}[t,T]$.

(iv) For the proof of the concavity, let us fix $t\in[0,T]$. Using the boundedness of $\widetilde{\mathcal{U}}[t,T]\ni(\xi,b)\mapsto g^{\xi,b}(0,p)$ from below by zero and Proposition B.4, we can show the concavity of $\Delta_m\ni p\mapsto\widetilde{g}(t,p)$ similar to the proof of Lemma 4.22 (ii) for all $t\in[0,T]$.

(v) The Lipschitz condition is proven in much the same way as in Bäuerle and Rieder [31, Lemma 6.1 d)]. Let us fix $p\in\Delta_m$ and $t\in[0,T]$. Due to the dependency of $g^{\xi,b}$ on the time horizon T, we use the notation $g_T^{\xi,b}(t,p)$. For any $(\xi,b)\in\widetilde{\mathcal{U}}[t,T]$, we define $(\hat\xi,\hat b)=(\hat\xi_s,\hat b_s)_{s\in[0,T-t]}$ by $(\hat\xi_s,\hat b_s)=(\xi_{t+s},b_{t+s})$ for all $s\in[0,T-t]$. Using this notation, it follows $g_t^{\xi,b}(t,p)=g_{T-t}^{\hat\xi,\hat b}(0,p)$. Now let $0\le t_1<t_2\le T$. Appealing to Proposition B.5 (which can be applied since $\widetilde{\mathcal{U}}[0,T]\ni(\xi,b)\mapsto g^{\xi,b}(0,p)$ is bounded for every $T>0$, compare statement (ii)), we get

$$\begin{aligned}\big|\widetilde{g}(t_1,p)-\widetilde{g}(t_2,p)\big| &= \Big|\inf_{(\xi,b)\in\widetilde{\mathcal{U}}[t_1,T]} g_T^{\xi,b}(t_1,p) - \inf_{(\xi,b)\in\widetilde{\mathcal{U}}[t_2,T]} g_T^{\xi,b}(t_2,p)\Big| \\ &= \Big|\inf_{(\hat\xi,\hat b)\in\widetilde{\mathcal{U}}[0,T-t_1]} g_{T-t_1}^{\hat\xi,\hat b}(0,p) - \inf_{(\hat\xi,\hat b)\in\widetilde{\mathcal{U}}[0,T-t_1]} g_{T-t_2}^{\hat\xi,\hat b}(0,p)\Big| \\ &\le \sup_{(\hat\xi,\hat b)\in\widetilde{\mathcal{U}}[0,T-t_1]}\big|g_{T-t_1}^{\hat\xi,\hat b}(0,p)-g_{T-t_2}^{\hat\xi,\hat b}(0,p)\big|.\end{aligned}$$

Notice that, by $T-t_2\le T-t_1$, we have

$$\inf_{(\hat\xi,\hat b)\in\widetilde{\mathcal{U}}[0,T-t_2]} g_{T-t_2}^{\hat\xi,\hat b}(0,p) = \inf_{(\hat\xi,\hat b)\in\widetilde{\mathcal{U}}[0,T-t_1]} g_{T-t_2}^{\hat\xi,\hat b}(0,p).$$

This is used in the second equality above. Appealing again to Lemma A.3, for any $\varepsilon>0$, there exists a strategy $(\bar\xi,\bar b)\in\widetilde{\mathcal{U}}[0,T-t_1]$ such that

$$\begin{aligned}&\sup_{(\hat\xi,\hat b)\in\widetilde{\mathcal{U}}[0,T-t_1]}\big|g_{T-t_1}^{\hat\xi,\hat b}(0,p)-g_{T-t_2}^{\hat\xi,\hat b}(0,p)\big| \\ &\le \big|g_{T-t_1}^{\bar\xi,\bar b}(0,p)-g_{T-t_2}^{\bar\xi,\bar b}(0,p)\big|+\varepsilon\end{aligned}$$

$$
\begin{aligned}
&= \Bigg| \mathbb{E}^{0,p}_{\mathbb{Q}^{\bar\xi,\bar b}_{T-t_1}} \bigg[\exp\bigg\{ \int_0^{T-t_1} \Big(-\alpha\, e^{r(T-s)}((\mu-r)\bar\xi_s + c(\bar b_s)) + \frac{1}{2}\alpha^2\,\sigma^2\, e^{2r(T-s)}\bar\xi_s^2 \\
&\quad + \lambda \sum_{D\subset\mathbb{D}} p_s^D \int_{(0,\infty)^d} \exp\Big\{ \alpha\,\bar b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i) \Big\} F(\mathrm{d}y) - \lambda \Big) \mathrm{d}s \bigg\} \bigg] \\
&\quad - \mathbb{E}^{0,p}_{\mathbb{Q}^{\bar\xi,\bar b}_{T-t_2}} \bigg[\exp\bigg\{ \int_0^{T-t_2} \Big(-\alpha\, e^{r(T-s)}((\mu-r)\bar\xi_s + c(\bar b_s)) + \frac{1}{2}\alpha^2\,\sigma^2\, e^{2r(T-s)}\bar\xi_s^2 \\
&\quad + \lambda \sum_{D\subset\mathbb{D}} p_s^D \int_{(0,\infty)^d} \exp\Big\{ \alpha\,\bar b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i) \Big\} F(\mathrm{d}y) - \lambda \Big) \mathrm{d}s \bigg\} \bigg] \Bigg| + \varepsilon \\
&= \Bigg| \mathbb{E}^{0,p}_{\mathbb{Q}^{\bar\xi,\bar b}_{T-t_1}} \bigg[\exp\bigg\{ \int_0^{T-t_2} \Big(-\alpha\, e^{r(T-s)}((\mu-r)\bar\xi_s + c(\bar b_s)) + \frac{1}{2}\alpha^2\,\sigma^2\, e^{2r(T-s)}\bar\xi_s^2 \\
&\quad + \lambda \sum_{D\subset\mathbb{D}} p_s^D \int_{(0,\infty)^d} \exp\Big\{ \alpha\,\bar b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i) \Big\} F(\mathrm{d}y) - \lambda \Big) \mathrm{d}s \bigg\} \times \\
&\quad \bigg(\exp\bigg\{ \int_{T-t_2}^{T-t_1} \Big(-\alpha\, e^{r(T-s)}((\mu-r)\bar\xi_s + c(\bar b_s)) + \frac{1}{2}\alpha^2\,\sigma^2\, e^{2r(T-s)}\bar\xi_s^2 \\
&\quad + \lambda \sum_{D\subset\mathbb{D}} p_s^D \int_{(0,\infty)^d} \exp\Big\{ \alpha\,\bar b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i) \Big\} F(\mathrm{d}y) - \lambda \Big) \mathrm{d}s \bigg\} - 1 \bigg) \bigg] \Bigg| + \varepsilon.
\end{aligned}
$$

Notice that the family of probability measures $\big(\mathbb{Q}^{\bar\xi,\bar b}_t\big)_{t\in[0,T-t_1]}$ is consistent in the sense that, for any $A \in \mathcal{G}_t$, $t \in [0, T-t_1]$,

$$
\mathbb{Q}^{\bar\xi,\bar b}_{T-t_1}(A) = \mathbb{E}\Big[\mathbb{1}_A L^{\bar\xi,\bar b}_{T-t_1}\Big] = \mathbb{E}\Big[\mathbb{1}_A \mathbb{E}\Big[L^{\bar\xi,\bar b}_{T-t_1} \,|\, \mathcal{G}_t\Big]\Big] = \mathbb{E}\Big[\mathbb{1}_A L^{\bar\xi,\bar b}_t\Big] = \mathbb{Q}^{\bar\xi,\bar b}_t(A).
$$

By the same arguments as in the proof of Lemma A.4, we obtain that the first exponential function in the expectation above is bounded by a constant $0 < K_0 < \infty$ which is independent of $(\bar\xi, \bar b)$. Therefore, by reapplying the above mentioned argumentation to the second exponential function, we obtain

$$
\begin{aligned}
&\big|\widetilde g(t_1,p) - \widetilde g(t_2,p)\big| \\
&\le K_0\, \mathbb{E}^{0,p}_{\mathbb{Q}^{\bar\xi,\bar b}_{T-t_2}} \bigg[\bigg| \exp\bigg\{ \int_{T-t_2}^{T-t_1} \Big(\alpha\, e^{rT} \Big(|\mu-r|\, K + (2+\eta+\theta)\,\kappa + \frac{1}{2}\alpha\,\sigma^2\, e^{rT}\, K^2 \Big) \\
&\qquad\qquad + \lambda\, M_F\big(\alpha\, e^{rT}\big) \Big) \mathrm{d}s \bigg\} - 1 \bigg| \bigg] + \varepsilon.
\end{aligned}
$$

On account of the Lipschitz condition for exponential function stated in Proposition B.1, we get

$$
\begin{aligned}
&\big|\widetilde g(t_1,p) - \widetilde g(t_2,p)\big| \\
&\le K_0\, \mathbb{E}^{0,p}_{\mathbb{Q}^{\bar\xi,\bar b}_{T-t_2}} \bigg[\bigg| \exp\bigg\{ \Big(\alpha\, e^{rT} \Big(|\mu-r|\, K + (2+\eta+\theta)\,\kappa + \frac{1}{2}\alpha\,\sigma^2\, e^{rT}\, K^2 \Big) \\
&\quad + \lambda\, M_F\big(\alpha\, e^{rT}\big) \Big) T \bigg\} (e-1) \Big(\alpha\, e^{rT} \Big(|\mu-r|\, K + (2+\eta+\theta)\,\kappa + \frac{1}{2}\alpha\,\sigma^2\, e^{rT}\, K^2 \Big)
\end{aligned}
$$

$$
\begin{aligned}
&\quad + \lambda\, M_F\big(\alpha\, e^{rT}\big)\Big)(t_2 - t_1)\Big|\Big] + \varepsilon \\
&= K_0\, K_1\, |t_2 - t_1| + \varepsilon,
\end{aligned}
$$

where $0 < K_1 < \infty$ is the constant given by

$$
\begin{aligned}
K_1 &:= \exp\bigg\{\bigg(\alpha\, e^{rT}\Big(|\mu - r|\, K + (2 + \eta + \theta)\,\kappa + \frac{1}{2}\alpha\,\sigma^2\, e^{rT}\, K^2\Big) \\
&+ \lambda\, M_F\big(\alpha\, e^{rT}\big)\bigg)T\bigg\}(e-1)\bigg(\alpha\, e^{rT}\Big(|\mu - r|\, K + (2 + \eta + \theta)\,\kappa + \frac{1}{2}\alpha\,\sigma^2\, e^{rT}\, K^2\Big) \\
&+ \lambda\, M_F\big(\alpha\, e^{rT}\big)\bigg),
\end{aligned}
$$

which is independent of $(\bar{\xi}, \bar{b})$. Letting $\varepsilon \downarrow 0$ yields the assertion.

(vi) Fix $(t,p) \in M$. Using the same arguments and notation as in the proof of previous statement, we obtain

$$
\begin{aligned}
\widetilde{g}_t(t,p) &= \frac{\partial}{\partial t} \inf_{(\xi,b)\in\widetilde{\mathcal{U}}[t,T]} g^{\xi,b}(t,p) \\
&= \lim_{h\downarrow 0} \frac{1}{h}\bigg(\inf_{(\xi,b)\in\widetilde{\mathcal{U}}[t+h,T]} g^{\xi,b}(t+h,p) - \inf_{(\xi,b)\in\widetilde{\mathcal{U}}[t,T]} g^{\xi,b}(t,p)\bigg) \\
&\le \lim_{h\downarrow 0} \frac{1}{h}\bigg| \inf_{(\xi,b)\in\widetilde{\mathcal{U}}[t,T]} g^{\xi,b}(t,p) - \inf_{(\xi,b)\in\widetilde{\mathcal{U}}[t+h,T]} g^{\xi,b}(t+h,p)\bigg| \\
&\le \lim_{h\downarrow 0} \frac{1}{h} \sup_{(\hat{\xi},\hat{b})\in\widetilde{\mathcal{U}}[0,T-t]} \Big| g^{\hat{\xi},\hat{b}}_{T-t}(0,p) - g^{\hat{\xi},\hat{b}}_{T-(t+h)}(0,p)\Big|.
\end{aligned}
$$

We proceed further as in the proof of the previous statement. For any $\varepsilon > 0$, there exists a strategy $(\bar{\xi}, \bar{b}) \in \widetilde{\mathcal{U}}[0, T-t]$ such that

$$
\begin{aligned}
&\sup_{(\bar{\xi},\bar{b})\in\widetilde{\mathcal{U}}[0,T-t]} \Big| g^{\hat{\xi},\hat{b}}_{T-t}(0,p) - g^{\hat{\xi},\hat{b}}_{T-(t+h)}(0,p)\Big| + \varepsilon \\
&\le \Big| g^{\bar{\xi},\bar{b}}_{T-t}(0,p) - g^{\bar{\xi},\bar{b}}_{T-(t+h)}(0,p)\Big| + \varepsilon \le K_0 |h| + \varepsilon,
\end{aligned}
$$

where $0 < K_0 < \infty$ is a constant independent of $(\bar{\xi}, \bar{b})$. Hence

$$
|\widetilde{g}_t(t,p)| \le \bigg| \lim_{h\downarrow 0} \frac{K_0 |h|}{h} + \varepsilon \bigg| = K_0 + \varepsilon,
$$

which yields statement (vi) for $\varepsilon \downarrow 0$.

(vii) Fix $(t,p) \in [0,T] \times \Delta_m$ and $(\xi, b) \in [-K, K] \times [0,1]$. It follows from the definition of $\mathscr{L}$ given in (4.24) and statement (iii) that

$$
\begin{aligned}
\big|\mathscr{L}\widetilde{g}(t,p;\xi,b)\big| \le K_3 \bigg(\lambda + \alpha\, e^{rT}\Big(|\mu - r| K + (2 + \eta + \theta)\,\kappa + \frac{1}{2}\alpha\,\sigma^2 e^{rT} K^2\Big) \\
+ \lambda\, M_F\big(\alpha\, e^{rT}\big)\bigg) =: K_5.
\end{aligned}
$$

(viii) Fix $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$. In the same way as in the proof the previous statement, the following results arise by taking account of (4.26), (4.30) and (4.31):

$$\begin{aligned}&\Big| \inf_{(\xi,b)\in[-K,K]\times[0,1]} \mathscr{L}\widetilde{g}(t,p,q;\xi,b)\Big| \\ &\le K_3\bigg(\lambda + \alpha\, e^{rT}\frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{1}{\alpha} + \alpha\, e^{rT}(2+\eta+\theta)\,\kappa + \lambda\, M_F\big(\alpha\, e^{rT}\big)\bigg) =: K_6. \quad \square\end{aligned}$$

The preceding lemma is a key ingredient to show the following existence result for a solution of the generalized HJB equation.

Theorem 4.33. *The value function of the investment-reinsurance problem stated in* (P1) *is given by*

$$V(t,x,p) = -e^{-\alpha x e^{r(T-t)}}\, g(t,p),$$

where g is defined by (4.19) *and g satisfies the generalized HJB equation*

$$0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \{\mathscr{L} g(t,p;\xi,b)\} + \inf_{\varphi\in\partial^C g_p(t)} \{\varphi\}, \quad (t,p)\in[0,T]\times\Delta_m,$$

with boundary condition $g(T,p) = 1$ for all $p \in \Delta_m$. Furthermore, $(\xi^\star, b^\star_{\lambda,F}) = (\xi^\star(s), b^\star_{\lambda,F}(s))_{s\in[t,T]}$ with $\xi^\star(s)$ given by (4.29) *and $b^\star_{\lambda,F}(s) := b_{\lambda,F}(s-,p_{s-})$ given by* (4.36) *is an optimal investment-reinsurance strategy for the Problem* (P1).

Proof. Our first concern is to prove that the function $\widetilde{g}$ defined by (4.45) satisfies the generalized HJB equation. The following argumentation is taken from the proof of Theorem 5.2 in Bäuerle and Rieder [31]. Fix $(t,p,x) \in [0,T]\times\mathbb{R}\times\Delta_m$ and $(\xi,b) \in \widetilde{\mathcal{U}}[t,T]$. Let τ be the first jump time of $X^{\xi,b}$ after t. Notice that τ is a $\mathfrak{G}$-stopping time since $X^{\xi,b}$ jumps at the arrival times of the trigger events $(T_n)_{n\in\mathbb{N}}$ which are observable. Hence $\tau \wedge t'$ is a $\mathfrak{G}$-stopping time taking values in $[t,T]$, where $t' \in (t,T]$ is some fixed time point. Using the argumentation of proof of Lemma 4.24, we obtain

$$\widetilde{V}(t,x,p) = f(t,x)\,\widetilde{g}(t,p), \quad (t,x,p)\in[0,T]\times\mathbb{R}\times\Delta_m,$$

where

$$f(t,x) := -e^{-\alpha x e^{r(T-t)}}.$$

Furthermore, from Lemma 4.32 (iii), we know that $\Delta_m \ni p \mapsto \widetilde{g}(t,p)$ is concave and, in consequence, $\Delta_m \ni p \mapsto \widetilde{g}(t,p)$ is continuous. Moreover, by Lemma 4.32 (iv), $[0,T] \ni t \mapsto \widetilde{g}(t,p)$ is Lipschitz on $[0,T]$ for all $p\in\Delta_m$ and hence $\widetilde{g}(t,p)$ is differentiable w.r.t. t almost everywhere on $[0,T]$ in the sense of the Lebesgue measure for all $p\in\Delta_m$, compare Theorem 2.3. Therefore, we can conclude from Lemma A.8 that

$$\begin{aligned}&\widetilde{V}(\tau\wedge t', X^{\xi,b}_{\tau\wedge t'}, p_{\tau\wedge t'}) \\ &= \widetilde{V}(t, X^{\xi,b}_t, p_t) + \int_t^{\tau\wedge t'} f(s, X^{\xi,b}_s)\,\mathscr{H}\widetilde{g}(s,p_s;\xi_s,b_s)\,\mathrm{d}s + \eta^{\xi,b}_{\tau\wedge t'} - \eta^{\xi,b}_t,\end{aligned} \tag{4.47}$$

where $(\eta^{\xi,b}_t)_{t\in[0,T]}$ is a $\mathfrak{G}$-martingale and we set $\mathscr{H}\widetilde{g}(s,p_s;\xi_s,b_s)$ to zero at those $s\in[t,T]$, where the partial derivative of $\widetilde{g}$ w.r.t. t does not exist. For any $\varepsilon > 0$ we can construct a strategy $(\xi^\varepsilon, b^\varepsilon) \in \widetilde{\mathcal{U}}[t,T]$ with $(\xi^\varepsilon_s, b^\varepsilon_s) = (\xi_s, b_s)$ for all $s\in[t,\tau\wedge t']$ from the continuity of V (cf. e.g. proof of Prop. 4.1 in Leobacher et al. [83] and proof of Prop. 3.1 in Azcue and Muler [9]) such that

$$\begin{aligned}\mathbb{E}^{t,x,p}\Big[\widetilde{V}(\tau\wedge t', X^{\xi,b}_{\tau\wedge t'}, p_{\tau\wedge t'})\Big] &\le \mathbb{E}^{t,x,p}\Big[\mathbb{E}^{\tau\wedge t', X^{\xi,b}_{\tau\wedge t'}, p_{\tau\wedge t'}}\Big[U(X^{\xi^\varepsilon,b^\varepsilon}_T)\Big]\Big]+\varepsilon\\ &\le \mathbb{E}^{t,x,p}\Big[U(X^{\xi^\varepsilon,b^\varepsilon}_T)\Big]+\varepsilon \le \widetilde{V}(t,x,p)+\varepsilon.\end{aligned}$$

From the arbitrariness of $\varepsilon > 0$ we conclude

$$\widetilde{V}(t,x,p) \ge \mathbb{E}^{t,x,p}\Big[\widetilde{V}(\tau\wedge t', X^{\xi,b}_{\tau\wedge t'}, p_{\tau\wedge t'})\Big].$$

Inserting (4.47) into the previous equation leads to

$$\begin{aligned}0 &\ge \mathbb{E}^{t,x,p}\Big[\int_t^{\tau\wedge t'} f(s,X^{\xi,b}_s)\,\mathscr{H}\widetilde{g}(s,p_s;\xi_s,b_s)\,\mathrm{d}s\Big]\\ &= \mathbb{E}^{t,x,p}\Big[\int_t^{t'} f(s,X^{\xi,b}_s)\,\mathscr{H}\widetilde{g}(s,p_s;\xi_s,b_s)\,\mathrm{d}s\mathbb{1}_{\{t'<\tau\}}\Big]\\ &\quad+\mathbb{E}^{t,x,p}\Big[\int_t^{\tau} f(s,X^{\xi,b}_s)\,\mathscr{H}\widetilde{g}(s,p_s;\xi_s,b_s)\,\mathrm{d}s\mathbb{1}_{\{t'\ge\tau\}}\Big],\end{aligned}$$

where we have subtracted $V(t,x,p)$ from both sides. By the law of total variation, the inequality is equivalent to

$$\begin{aligned}0 &\ge \mathbb{E}^{t,x,p}\Big[\int_t^{t'} f(s,X^{\xi,b}_s)\,\mathscr{H}\widetilde{g}(s,p_s;\xi_s,b_s)\,\mathrm{d}s\,|\,t'<\tau\Big]\mathbb{P}^{t,x,p}(t'<\tau)\\ &\quad+\mathbb{E}^{t,x,p}\Big[\int_t^{\tau} f(s,X^{\xi,b}_s)\,\mathscr{H}\widetilde{g}(s,p_s;\xi_s,b_s)\,\mathrm{d}s\,|\,t'\ge\tau\Big]\mathbb{P}^{t,x,p}(t'\ge\tau).\end{aligned}$$

Next, we divide both sides in the inequality above by $t'-t$ and consider $t'\downarrow t$, which results in

$$\begin{aligned}0 &\ge \lim_{t'\downarrow t}\mathbb{E}^{t,x,p}\Big[\frac{1}{t'-t}\int_t^{t'} f(s,X^{\xi,b}_s)\,\mathscr{H}\widetilde{g}(s,p_s;\xi_s,b_s)\,\mathrm{d}s\,|\,t'<\tau\Big]\mathbb{P}^{t,x,p}(t'<\tau)\\ &\quad+\lim_{t'\downarrow t}\mathbb{E}^{t,x,p}\Big[\frac{1}{t'-t}\int_t^{\tau} f(s,X^{\xi,b}_s)\,\mathscr{H}\widetilde{g}(s,p_s;\xi_s,b_s)\,\mathrm{d}s\,|\,t'\ge\tau\Big]\mathbb{P}^{t,x,p}(t'\ge\tau).\end{aligned}\tag{4.48}$$

The next aim is to determine the limits in the inequality above. We first determine the probability $\mathbb{P}^{t,x,p}(t'\ge\tau)$. For this purpose, we denote the last jump time of $X^{\xi,b}$ before τ by τ'. Since $X^{\xi,b}$ jumps at $N=(T_n)_{n\in\mathbb{N}}$, which is a Poisson process with intensity λ, we know that $\tau-\tau'$ is exponential distributed with parameter λ. Taking this into account and the memoryless property of the exponential distribution as well as $\tau > t$, we obtain

$$\begin{aligned}\mathbb{P}^{t,x,p}(\tau>t') &= \mathbb{P}(\tau>t') = \mathbb{P}(\tau-\tau'>t'-\tau') = \mathbb{P}(\tau-\tau'>(t'-t)+(t-\tau'))\\ &= \mathbb{P}(\tau-\tau'>t'-t)\,\mathbb{P}(\tau-\tau'>t-\tau') = \mathbb{P}(\tau-\tau'>t'-t)\,\mathbb{P}(\tau>t)\\ &= \mathbb{P}(\tau-\tau'>t'-t) = e^{-\lambda(t'-t)}.\end{aligned}$$

Thus

$$\lim_{t'\downarrow t}\mathbb{P}^{t,x,p}(\tau\le t') = 1-\lim_{t'\downarrow t}e^{-\lambda(t'-t)} = 0.$$

Consequently,

$$0 \geq \lim_{t' \downarrow t} \mathbb{E}^{t,x,p}\left[\frac{1}{t'-t}\int_t^{t'} f(s, X_s^{\xi,b})\,\mathscr{H}\widetilde{g}(s,p_s;\xi_s,b_s)\,\mathrm{d}s \mathbb{1}_{\{t'<\tau\}}\right].$$

Our next concern will be the interchange of the limit and the expectations. Due to Lemma 4.32 (vi), (vii), Lemma A.3 and Lemma A.4, we have

$$\begin{aligned}
&\mathbb{E}^{t,x,p}\left[\left|\frac{1}{t'-t}\int_t^{t'} f(s, X_s^{\xi,b})\,\mathscr{H}\widetilde{g}(s,p_s;\xi_s,b_s)\,\mathrm{d}s\right|\right] \\
&\leq \mathbb{E}^{t,x,p}\left[\frac{1}{t'-t}\int_t^{t'} |f(s, X_s^{\xi,b})|(K_5+K_6)\,\mathrm{d}s\right] \\
&\leq \frac{K_5+K_6}{t'-t}\int_t^{t'} \mathbb{E}^{t,x,p}_{\mathbb{Q}_s^{\xi,b}}\left[\frac{|f(s,X_s^{\xi,b})|}{L_s^{\xi,b}}\right]\mathrm{d}s \leq (K_5+K_6)\,K_1.
\end{aligned}$$

Therefore, by dominated convergence theorem, we obtain

$$0 \geq \mathbb{E}^{t,x,p}\left[\lim_{t' \downarrow t}\frac{1}{t'-t}\int_t^{t'} f(s, X_s^{\xi,b})\,\mathscr{H}\widetilde{g}(s,p_s;\xi_s,b_s)\,\mathrm{d}s \mathbb{1}_{\{t'<\tau\}}\right].$$

Next, we want to apply the FTCL, see Sohrab [115, Thm. 11.5.23, 11.5.31]. For this purpose, we define the function $h : [t,T] \to \mathbb{R}$ by

$$h(s) := f(s, X_s^{\xi,b})\,\mathscr{H}\widetilde{g}(s,p_s;\xi_s,b_s).$$

Notice that, due to the Lipschitz property of $\widetilde{g}$, $s \mapsto \widetilde{g}_t(s,\cdot)$ is bounded on $[t,T]$ almost everywhere and càdlàg functions on compact sets are also bounded almost everywhere. Hence, $s \mapsto h(s)$ is Lebesgue integrable on $[t,T]$. Therefore, we can apply the FTCL and we obtain that the function $H : [t,T] \to \mathbb{R}$ given by

$$H(s) := H(t) + \int_t^s f(u, X_u^{\xi,b})\,\mathscr{H}\widetilde{g}(u,p_u;\xi_u,b_u)\,\mathrm{d}u, \quad s \in [t,T],$$

is absolutely continuous and $H'(s) = h(s)$ for almost all $s \in [t,T]$. Thus,

$$\lim_{t' \downarrow t}\frac{1}{t'-t}\int_t^{t'} h(s)\,\mathrm{d}s = \lim_{t' \downarrow t}\frac{H(t')-H(t)}{t'-t} = H'(t) = h(t).$$

Consequently, by the fact that $\mathbb{1}_{\{t'<\tau\}} \to 1$ $\mathbb{P}$-a.s. for $t' \downarrow t$ since $t < \tau$, we deduce that

$$\begin{aligned}
&\mathbb{E}^{t,x,p}\left[\lim_{t' \downarrow t}\frac{1}{t'-t}\int_t^{t'} f(s, X_s^{\xi,b})\,\mathscr{H}\widetilde{g}(s,p_s;\xi_s,b_s)\,\mathrm{d}s \mathbb{1}_{\{t'<\tau\}}\right] \\
&= \mathbb{E}^{t,x,p}\left[f(t, X_t^{\xi,b})\,\mathscr{H}\widetilde{g}(t,p_t;\xi_t,b_t)\right].
\end{aligned}$$

From now on, let $(\xi,b) \in [-K,K] \times [0,1]$ and $\varepsilon > 0$ be arbitrary constants as well as $(\bar{\xi},\bar{b}) \in \widetilde{\mathcal{U}}[t,T]$ be a fixed strategy with $(\bar{\xi}_s,\bar{b}_s) \equiv (\xi,b)$ for $s \in [t,t+\varepsilon)$. Hence

$$\mathbb{E}^{t,x,p}\left[f(t, X_t^{\xi,b})\,\mathscr{H}\widetilde{g}(t,p_t;\bar{\xi}_t,\bar{b}_t)\right] = f(t,x)\,\mathscr{H}\widetilde{g}(t,p;\xi,b)$$

at those points (t,p), where $\widetilde{g}$ is differentiable w.r.t. t. Consequently,

$$0 \geq f(t,x)\, \mathscr{H}\widetilde{g}(t,p;\xi,b).$$

On account of the negativity of $f(t,x)$, we get

$$0 \leq \mathscr{H}\widetilde{g}(t,p;\xi,b).$$

In the light of the arbitrariness of (ξ,b), we obtain

$$0 \leq \inf_{(\xi,b)\in[-K,K]\times[0,1]} \big\{\mathscr{H}\widetilde{g}(t,p;\xi,b)\big\}.$$

We show next the inequality above if $\widetilde{g}$ is not differentiable w.r.t. t. To see this, we define $\widetilde{g}_p(t) := \widetilde{g}(t,p)$ and we denote set of point at which $\widetilde{g}_p$ is differentiable by $D_{\widetilde{g}_p} \subset [0,T]$. Appealing to Theorem 2.9, we have

$$\partial^C \widetilde{g}_p(t) = \operatorname{co}\Big\{\lim_{n\to\infty} \widetilde{g}_p'(t_n) : t_n \to t, t_n \in D_{\widetilde{g}_p}\Big\}.$$

That is, for every $\varphi \in \partial^C\widetilde{g}_p(t)$, there exist $u \in \mathbb{N}$ and $(\gamma_1,\dots,\gamma_u) \in \Delta_u$ such that $\varphi = \sum_{i=1}^u \gamma_i \varphi^i$, where $\varphi^i = \lim_{n\to\infty} \widetilde{g}_p'(t_n^i)$ for sequences $(t_n^i)_{n\in\mathbb{N}} \subset D_{\widetilde{g}_p}$ with $\lim_{n\to\infty} t_n^i = t$. From what has already been proved, it can be concluded that

$$0 \leq \gamma_i \mathscr{L}\widetilde{g}(t_n^i,p;\xi,b) + \gamma_i \widetilde{g}_p'(t_n^i), \quad i = 1,\dots,u,$$

and thus, by the continuity of $t \mapsto \widetilde{g}(t,p)$,

$$0 \leq \gamma_i \mathscr{L}\widetilde{g}(t,p;\xi,b) + \gamma_i \lim_{n\to\infty} \widetilde{g}_p'(t_n^i), \quad i = 1,\dots,u,$$

which yields

$$0 \leq \mathscr{L}\widetilde{g}(t,p;\xi,b)\sum_{i=1}^u \gamma_i + \sum_{i=1}^u \gamma_i \lim_{n\to\infty} \widetilde{g}_p'(t_n^i) = \mathscr{L}\widetilde{g}(t,p;\xi,b) + \varphi.$$

Due to the arbitrariness of $\varphi \in \partial^C\widetilde{g}_p(t)$ and $(\xi,b) \in [-K,K]\times[0,1]$, we get

$$0 \leq \inf_{(\xi,b)\in[-K,K]\times[0,1]} \big\{\mathscr{L}\widetilde{g}(t,p;\xi,b)\big\} + \inf_{\varphi\in\partial^C\widetilde{g}_p(t)} \{\varphi\}.$$

Our next objective is to evaluate the reverse inequality above. For any $\varepsilon > 0$ and $0 < t < t' \leq T$, there exists a strategy $(\xi^{\varepsilon,t'}, b^{\varepsilon,t'}) \in \widetilde{\mathcal{U}}[t,T]$ such that

$$\widetilde{V}(t,x,p) - \varepsilon\,(t'-t) \leq \mathbb{E}^{t,x,p}\Big[U\big(X_T^{\xi^{\varepsilon,t'},b^{\varepsilon,t'}}\big)\Big] \leq \mathbb{E}^{t,x,p}\Big[\widetilde{V}\big(\tau\wedge t', X_{\tau\wedge t'}^{\xi^{\varepsilon,t'},b^{\varepsilon,t'}}, p_{\tau\wedge t'}\big)\Big].$$

Again on the basis of (4.47), it can be deduced that

$$-\varepsilon\,(t'-t) \leq \mathbb{E}^{t,x,p}\bigg[\int_t^{\tau\wedge t'} f\big(s, X_s^{\xi^{\varepsilon,t'},b^{\varepsilon,t'}}\big)\, \mathscr{H}\widetilde{g}(s,p_s;\xi_s^{\varepsilon,t'},b_s^{\varepsilon,t'})\,\mathrm{d}s\bigg],$$

which is equivalent to

$$-\varepsilon \leq \mathbb{E}^{t,x,p}\bigg[\frac{1}{t'-t}\int_t^{\tau\wedge t'} f\big(s, X_s^{\xi^{\varepsilon,t'},b^{\varepsilon,t'}}\big)\, \mathscr{H}\widetilde{g}(s,p_s;\xi_s^{\varepsilon,t'},b_s^{\varepsilon,t'})\,\mathrm{d}s\bigg]$$

$$\leq \mathbb{E}^{t,x,p}\left[\frac{1}{t'-t}\int_t^{\tau\wedge t'} f\big(s, X_s^{\xi^{\varepsilon,t'},b^{\varepsilon,t'}}\big) \inf_{(\xi,b)\in[-K,K]\times[0,1]} \big\{\mathscr{H}\widetilde{g}(s,p_s;\xi,b)\big\}\,\mathrm{d}s\right].$$

In the same manner as before, we get

$$-\varepsilon \leq \lim_{t'\downarrow t} \mathbb{E}^{t,x,p}\left[\frac{1}{t'-t}\int_t^{t'} f\big(s, X_s^{\xi^{\varepsilon,t'},b^{\varepsilon,t'}}\big) \inf_{(\xi,b)\in[-K,K]\times[0,1]} \big\{\mathscr{H}\widetilde{g}(s,p_s;\xi,b)\big\}\,\mathrm{d}s\mathbb{1}_{\{t'<\tau\}}\right].$$

Making use of Lemma 4.32 (vi), (viii), we obtain as before

$$\begin{aligned}
&\mathbb{E}^{t,x,p}\left[\frac{1}{t'-t}\int_t^{t'} f\big(s, X_s^{\xi^{\varepsilon,t'},b^{\varepsilon,t'}}\big) \inf_{(\xi,b)\in[-K,K]\times[0,1]} \big\{\mathscr{H}\widetilde{g}(s,p_s;\xi,b)\big\}\,\mathrm{d}s\mathbb{1}_{\{t'<\tau\}}\right]\\
&\leq \mathbb{E}^{t,x,p}\left[\frac{1}{t'-t}\int_t^{t'} \Big|f\big(s, X_s^{\xi^{\varepsilon,t'},b^{\varepsilon,t'}}\big)\Big|(K_6+K_4)\,\mathrm{d}s\right] \leq K_1\,(K_6+K_4) < \infty.
\end{aligned}$$

Hence, we can interchange the limit and the infimum again, which yields

$$\begin{aligned}
&\lim_{t'\downarrow t} \mathbb{E}^{t,x,p}\left[\frac{1}{t'-t}\int_t^{\tau\wedge t'} f\big(s, X_s^{\xi^{\varepsilon,t'},b^{\varepsilon,t'}}\big) \inf_{(\xi,b)\in[-K,K]\times[0,1]} \big\{\mathscr{H}\widetilde{g}(s,p_s;\xi,b)\big\}\,\mathrm{d}s\right]\\
&= \mathbb{E}^{t,x,p}\left[f\big(t, X_t^{\xi^{\varepsilon,t'},b^{\varepsilon,t'}}\big) \inf_{(\xi,b)\in[-K,K]\times[0,1]} \big\{\mathscr{H}\widetilde{g}(t,p_t;\xi,b)\big\}\,\mathrm{d}s\right]\\
&= f(t,x) \inf_{(\xi,b)\in[-K,K]\times[0,1]} \big\{\mathscr{H}\widetilde{g}(t,p;\xi,b)\big\}
\end{aligned}$$

at those points (t,p), where $\widetilde{g}_p$ is differentiable. In consequence

$$-\varepsilon \leq f(t,x) \inf_{(\xi,b)\in[-K,K]\times[0,1]} \big\{\mathscr{H}\widetilde{g}(t,p;\xi,b)\big\}.$$

According to $f(t,x)<0$ and the arbitrariness of $\varepsilon>0$, we get by $\varepsilon\downarrow 0$,

$$0 \geq \inf_{(\xi,b)\in[-K,K]\times[0,1]} \big\{\mathscr{H}\widetilde{g}(t,p;\xi,b)\big\}.$$

By the same method as before, we obtain that in the case of no differentiability of $\widetilde{g}$ w.r.t. t

$$0 \geq \inf_{(\xi,b)\in[-K,K]\times[0,1]} \big\{\mathscr{L}\widetilde{g}(t,p;\xi,b)\big\} + \inf_{\varphi\in\partial^C\widetilde{g}_p(t)} \{\varphi\}.$$

In summary, we have

$$0 = \inf_{(\xi,b)\in[-K,K]\times[0,1]} \big\{\mathscr{L}\widetilde{g}(t,p;\xi,b)\big\} + \inf_{\varphi\in\partial^C\widetilde{g}_p(t)} \{\varphi\}$$

for all $t\in[0,T]$ and $p\in\Delta_m$. Notice that, similar to (4.26), it holds

$$\begin{aligned}
&\inf_{(\xi,b)\in[-K,K]\times[0,1]} \big\{\mathscr{L}\widetilde{g}(t,p;\xi,b)\big\}\\
&= -\lambda\,\widetilde{g}(t,p) + \alpha\, e^{r(T-t)}\,\widetilde{g}(t,p) \inf_{\xi\in[-K,K]} f_1(t,\xi) + \inf_{b\in[0,1]} f_2(t,p,b),
\end{aligned}$$

where f_1 is defined by (4.27) and f_2 by (4.28). Furthermore,

$$|\xi^\star(t)| = \Big|\arg\inf_{\xi\in\mathbb{R}} f_1(t,\xi)\Big| = \left|\frac{\mu-r}{\sigma^2}\frac{1}{\alpha}e^{-r(T-t)}\right| \leq \frac{|\mu-r|}{\sigma^2}\frac{1}{\alpha},$$

compare Section 4.5. That is, the optimal investment strategy is continuous, bounded as well as deterministic (in particular $\mathfrak{F}^W$-adapted) and it holds

$$\inf_{(\xi,b)\in[-K,K]\times[0,1]} \left\{\mathscr{L}\widetilde{g}(t,p;\xi,b)\right\} = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \left\{\mathscr{L}\widetilde{g}(t,p;\xi,b)\right\}$$

with $K := \frac{|\mu-r|}{\sigma^2}\frac{1}{\alpha}$. Therefore, $\widetilde{g}$ satisfies the generalized HJB equation with boundary condition $\widetilde{g}(T,p) = 1$, where the boundary condition follows immediately from the definition of $g^{\xi,b}$ given in (4.20). Since, by Lemma 4.32, $\widetilde{g} : [0,T]\times\Delta_m \to (0,\infty)$ is bounded, $p \mapsto \widetilde{g}(t,p)$ is continuous on Δ_m for all $t \in [0,T]$ and $t \mapsto \widetilde{g}(t,p)$ is Lipschitz on $[0,T]$ for all $p \in \Delta_m$, it follows from Theorem 4.31 that

$$V(t,x,p) = -e^{-\alpha x e^{r(T-t)}}\widetilde{g}(t,p), \quad (t,x,p) \in [0,T]\times\mathbb{R}\times\Delta_m,$$

and that $(\xi^\star, b^\star_{\lambda,F}) = (\xi^\star(s), b^\star_{\lambda,F}(s))_{s\in[t,T]}$ from Theorem 4.31 is an optimal investment-reinsurance strategy, where g is replaced by $\widetilde{g}$ in $A_{\lambda,F}(s,p)$ and $B_{\lambda,F}(s,p)$. From what has already been shown, we know that the optimal strategy $\xi^\star = (\xi^\star(s))_{s\in[t,T]}$ is a bounded, continuous and $\mathfrak{F}^W$-adapted process. Moreover, $b^\star_{\lambda,F} = (b^\star_{\lambda,F}(s))_{s\in[t,T]}$ is an $\mathfrak{F}^\Psi$-predictable $[0,1]$-valued process, compare end of the proof of Theorem 4.31. Hence, we have

$$\widetilde{g}(t,p) = \inf_{(\xi,b)\in\widetilde{\mathcal{U}}[t,T]} g^{\xi,b}(t,p) = \inf_{(\xi,b)\in\mathcal{U}[t,T]} g^{\xi,b}(t,p) = g(t,p),$$

and, in consequence,

$$V(t,x,p) = -e^{-\alpha x e^{r(T-t)}} g(t,p), \quad (t,x,p) \in [0,T]\times\mathbb{R}\times\Delta_m,$$

which finishes the proof. □

We close this section with a discussion on an alternative solution technique.

Remark 4.34. We have seen that the problem of choosing optimal investment-reinsurance strategies is decomposed in two separate problems: determination of an optimal investment strategy and of an optimal reinsurance strategy, which are independent from each other. So it is thinkable to consider the optimal reinsurance problem to be isolated. The behaviour of the derived optimal reinsurance strategy between the trigger arrival times $(T_n)_{n\in\mathbb{N}}$ is deterministic. This is expectable since there appears to be no new information about the unknown thinning probabilities between the jump times $(T_n)_{n\in\mathbb{N}}$. The deterministic evolution between the jump times $(T_n)_{n\in\mathbb{N}}$ indicates that we can transform the optimal reinsurance problem in a discrete one, in which at every claim arrival time a deterministic reinsurance strategy until the next claim is chosen. Indeed for this problem the state of the optimal reinsurance problem (without financial market) has a piecewise-deterministic behaviour such that we can define a time-discrete Markov Decision Process (MDP) with a value function coinciding with the one of the original optimal reinsurance problem and, in which every control is a function of the state after the last claim arrival time as well as the time elapsed since the last claim arrival. Making use of the MDP-theory it should be possible to establish an optimal control. Typically, this problem is solvable if the corresponding problem in continuous time is solved. But it is not expectable to obtain a "more explicit" solution in the discrete approach. For a thorough treatment of Markov Decision Processes we refer to Bäuerle and Rieder [32] and Davis [47]. A work considering both the generalized HJB-approach and the MDP-approach for optimal control problems with Markovian jump processes under incomplete information is Winter [119].

4.8 Comparison to the case with complete information

This section covers the study of the influence of the uncertainty about the dependencies between the LoBs on the optimal reinsurance strategy. For this aim it is initially necessary to determine the optimal reinsurance strategy in case of full information.

4.8.1 Solution to the optimal investment and reinsurance under full information

In this section we investigate the optimal investment and reinsurance problem within the setting of Section 4.1 under the additional assumption that $\bar{\alpha} \equiv \bar{c} = (c_D)_{D\subset\mathbb{D}} \in \Delta_\ell$ is deterministic. So we are in a framework with known parameters λ, $\bar{\alpha}$ and F. The setup is tantamount to assuming that $\mathcal{F}_0$ is known for the insurance company.

Due to the deterministic thinning probabilities, the compensated random measure $\widehat{\Psi}$ of Ψ defined in (4.10) is

$$\widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z)) = \Psi(\mathrm{d}t, \mathrm{d}(y,z)) - \lambda\, F(\mathrm{d}y) \sum_{D\in \mathrm{d}z} c_D\, \mathrm{d}t. \tag{4.49}$$

The state process in the full observable case is $(X_s^{\xi,b})_{s\in[t,T]}$, $(\xi,b)\in\mathcal{U}[t,T]$, which evolves as

$$\begin{aligned} \mathrm{d}X_s^{\xi,b} = \Big(& rX_s^{\xi,b} + (\mu - r)\xi_s + c(b_s) - \lambda\, b_s \sum_{D\subset\mathbb{D}} c_D \sum_{i=1}^d \mathbb{E}[Y_1^i]\, \mathbb{1}_D(i)\Big)\mathrm{d}s \\ & + \xi_s\sigma\,\mathrm{d}W_s - \int_{E^d} b_s \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z)), \end{aligned} \tag{4.50}$$

for $s\in[t,T]$ with initial time $t\in[0,T)$, where $X_t^{\xi,b} = x \in \mathbb{R}$. In contrast to the state process in the partial observable in Section 4.3, no further process is required to describe the information at disposal of the thinning probabilities since they are known. Then the value functions are given by, for any $(t,x)\in[0,T]\times\mathbb{R}$ and $(\xi,b)\in\mathcal{U}[t,T]$,

$$\begin{aligned} V^{\xi,b}(t,x) &:= \mathbb{E}^{t,x}\big[U(X_T^{\xi,b})\big], \\ V(t,x) &:= \sup_{(\xi,b)\in\mathcal{U}[t,T]} V^{\xi,b}(t,x), \end{aligned} \tag{O}$$

where $\mathbb{E}^{t,x}$ denotes the conditional expectation given $X_t^{\xi,b} = x$. So (O) is the corresponding control problem with full information to the partially observable problem (P).

The HJB equation for the value function V can be established by the same method as used in Section 4.4 and we obtain

$$\begin{aligned} 0 = \sup_{(\xi,b)\in\mathbb{R}\times[0,1]} \Big\{ & V_t(t,x) - \lambda V(t,x) + \frac{1}{2}\sigma^2 V_{xx}(t,x)\xi^2 + V_x(t,x)\big(rx + (\mu-r)\xi + c(b)\big) \\ & + \lambda \sum_{D\subset\mathbb{D}} c_D \int_{(0,\infty)^d} V\Big(t, x - b\sum_{i=1}^d y_i \mathbb{1}_D(i)\Big) F(\mathrm{d}y)\Big\}, \end{aligned} \tag{4.51}$$

where $V(T,x) = \mathbb{E}^{T,x}\big[U(X_T^{\xi,b})\big] = U(x)$ for all $(x,p)\in[0,T]\times\Delta_m$.

Following the proof of separation approach stated in Lemma 4.24, we obtain

$$V(t,x) = -e^{-\alpha x e^{r(T-t)}} g(t), \quad (t,x) \in [0,T] \times \mathbb{R}, \tag{4.52}$$

with

$$g(t) = \inf_{(\xi,b)\in\mathcal{U}[t,T]} g^{\xi,b}(t),$$

where

$$\begin{aligned} g^{\xi,b}(t) := \mathbb{E}\bigg[\exp\bigg\{ -\int_t^T \alpha\, e^{r(T-s)}\big((\mu-r)\xi_s + c(b_s)\big)\,\mathrm{d}s \\ - \int_t^T \alpha\,\sigma\, e^{r(T-s)} \xi_s \,\mathrm{d}W_s + \int_t^T \int_{E^d} \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \Psi(\mathrm{d}s, \mathrm{d}(y,z)) \bigg\}\bigg]. \end{aligned}$$

It therefore follows

$$\begin{aligned} V_t(t,x) &= -e^{-\alpha x e^{r(T-t)}} \Big(\alpha\, x\, r\, e^{r(T-t)} g(t) + g'(t) \Big), \\ V_x(t,x) &= -e^{-\alpha x e^{r(T-t)}} \Big(-\alpha e^{r(T-t)} g(t) \Big), \\ V_{xx}(t,x) &= -e^{-\alpha x e^{r(T-t)}} \alpha^2 e^{2r(T-t)} g(t), \\ V\Big(t, x - b\sum_{i=1}^d y_i \mathbb{1}_D(i), p\Big) &= -e^{-\alpha x e^{r(T-t)}} \exp\Big\{ \alpha\, b\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_D(i) \Big\} g(t), \end{aligned}$$

which yields the following HJB equation for g:

$$\begin{aligned} 0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \bigg\{ g'(t) - \lambda g(t) - \alpha e^{r(T-t)} g(t) \Big((\mu - r)\xi + c(b) - \frac{1}{2}\alpha\sigma^2 e^{r(T-t)} \xi^2 \Big) \\ + \lambda\, g(t) \sum_{D\subset\mathbb{D}} c_D \int_{(0,\infty)^d} \exp\Big\{ \alpha\, b\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_D(i) \Big\} F(\mathrm{d}y) \bigg\} \end{aligned} \tag{4.53}$$

with $g(T) = 1$. Equation (4.53) is equivalent to

$$0 = g'(t) - \lambda\, g(t) + \alpha e^{r(T-t)} g(t) \inf_{\xi\in\mathbb{R}} f_1(t,\xi) + g(t) \inf_{b\in[0,1]} f_2(t,b), \tag{4.54}$$

where f_1 is defined by (4.27) and

$$\begin{aligned} f_2(t,b) &:= -\alpha\, e^{r(T-t)}\, c(b) + \lambda \sum_{D\subset\mathbb{D}} c_D \int_{(0,\infty)^d} \exp\Big\{ \alpha\, b\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_D(i) \Big\} F(\mathrm{d}y) \\ &= -\alpha\, e^{r(T-t)} (\eta - \theta)\kappa - \alpha\, e^{r(T-t)} (1+\theta)\kappa \\ &\quad + \lambda \sum_{D\subset\mathbb{D}} c_D \int_{(0,\infty)^d} \exp\Big\{ \alpha\, b\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_D(i) \Big\} F(\mathrm{d}y) \end{aligned}$$

From Section 4.5, it is known that the candidate of an optimal investment strategy $(\xi^\star(t))_{t\in[0,T]}$ is again given by

$$\xi^\star(t) = \frac{\mu - r}{\sigma^2} \frac{1}{\alpha} e^{-r(T-t)}, \quad t \in [0,T],$$

and

$$\inf_{\xi\in\mathbb{R}} f_1(t,\xi) = f_1(t,\xi_t^\star) = -\frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{1}{\alpha}e^{-r(T-t)}.$$

An illustration of the investment strategy $\xi^\star$ is provided by Figure 4.7 in Section 4.9.

According to the proof of Lemma 4.27, $\mathbb{R} \ni b \mapsto f_2(t,b)$ is strictly convex and

$$\begin{aligned}\frac{\partial}{\partial b} f_2(t,b) = &-\alpha\, e^{r(T-t)}\Bigg((1+\theta)\,\kappa \\ &-\lambda\sum_{D\subset\mathbb{D}} c_D \sum_{i=1}^d \mathbb{1}_D(i)\int_{(0,\infty)^d} y_i \exp\left\{\alpha\, b\, e^{r(T-t)}\sum_{j=1}^d y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y)\Bigg).\end{aligned}$$

For declaring the first order condition of the optimal reinsurance strategy, we use the following notation.

Notation. For any $t \in [0,T]$ and $b \in \mathbb{R}$, we set

$$h_{\lambda,\bar{c},F}(t,b) := \lambda\sum_{D\subset\mathbb{D}} c_D \sum_{i=1}^d \mathbb{1}_D(i)\int_{(0,\infty)^d} y_i \exp\left\{\alpha\, b\, e^{r(T-t)}\sum_{j=1}^d y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y). \quad (4.55)$$

Furthermore, we define

$$\begin{aligned} A_{\lambda,\bar{c},F}(t) &:= h_{\lambda,\bar{c},F}(t,0), \\ B_{\lambda,\bar{c},F}(t) &:= h_{\lambda,\bar{c},F}(t,1). \end{aligned}$$

The explanation above results in the first order condition:

$$(1+\theta)\kappa = h_{\lambda,\bar{c},F}(t,b). \quad (4.56)$$

It should be noted that the previous equation has a unique solution w.r.t. b, compare proof of Proposition 4.29, which justifies the next notation.

Notation. From now on, $r_{\lambda,\bar{c},F}(t)$ denotes the unique root of Equation (4.56) w.r.t. b.

By the same line of arguments as in Proposition 4.29, we obtain that the optimal reinsurance strategy $\tilde{b}^\star_{\lambda,\bar{c},F} = (\tilde{b}^\star_{\lambda,\bar{c},F}(t))_{t\in[0,T]}$ is given by

$$\tilde{b}^\star_{\lambda,\bar{c},F}(t) := \begin{cases} 0, & \theta \le A_{\lambda,\bar{c},F}(t)/\kappa - 1, \\ 1, & \theta \ge B_{\lambda,\bar{c},F}(t)/\kappa - 1, \\ r_{\lambda,\bar{c},F}(t), & \text{otherwise}. \end{cases} \quad (4.57)$$

Notice that $r_{\lambda,\bar{c},F}(t)$, $A_{\lambda,\bar{c},F}(t)$ and $B_{\lambda,\bar{c},F}(t)$ are continuous in t. Consequently, the optimal reinsurance strategy $\tilde{b}^\star_{\lambda,\bar{c},F}$ is continuous. Moreover, $\tilde{b}^\star_{\lambda,\bar{c},F}$ is deterministic and can be calculated easily. Figure 4.4 in Section 4.9 illustrates the evolution of the optimal reinsurance strategy under complete information.

After solving the minimizing problems in (4.54), we can write the HJB equation for g as

$$g'(t) = \Bigg(-\lambda - \frac{1}{2}\frac{\mu-r}{\sigma^2}\frac{1}{\alpha}e^{-r(T-t)} + f_2(t, b^\star_{\lambda,\bar{c},F}(t))\Bigg) g(t).$$

The solution of this ODE of first order with boundary condition $g(T) = 1$ is

$$g(t) = \exp\left\{ \int_t^T \left(\lambda + \frac{1}{2}\frac{\mu - r}{\sigma^2}\frac{1}{\alpha}e^{-r(T-s)} - f_2(s, b^\star_{\lambda,\bar{c},F}(s))\right)\mathrm{d}s\right\}, \quad t \in [0,T], \tag{4.58}$$

cf. e.g. Polyanin and Zaitsev [101, 1.1.4]. Using (4.52), this results in the following representation of the value function: For any $(t,x) \in [0,T] \times \mathbb{R}$

$$V(t,x) = -\exp\left\{ -\alpha\, x\, e^{r(T-t)} \int_t^T \left(\lambda + \frac{1}{2}\frac{\mu - r}{\sigma^2}\frac{1}{\alpha}e^{-r(T-s)} - f_2(s, b^\star_{\lambda,\bar{c},F}(s))\right)\mathrm{d}s\right\}.$$

In the case of full information and one LoB (i.e. $d = 1$), we are in the same framework as in Section 3 of Liang et al. [84] and our optimal strategies are consistent with the results in the given reference, cf. Eq. (19) and Thm. 3.1 in Liang et al. [84]. In the mentioned sources it is shown that the corresponding HJB equation is a classical solution. The same method can be used to validate the optimality of the classical candidate solution (4.58) to the HJB Equation (4.53) in the full information case and thus the announced strategies are indeed optimal. This verification can be seen as a special case of the verification of the partial information case in Section 4.7.

4.8.2 Comparison results

In this section a comparison result of the optimal reinsurance strategy under partial information given in Theorem 4.33 and the optimal strategy under full information are stated. But first we will deduce a priori bounds to the reinsurance optimal strategy, i.e. bounds which can be calculated at time 0 and thus independent of the observed accidents. For this purpose, we introduce the following terms.

Notation. Let $t \in [0,T]$ and $b \in \mathbb{R}$. Throughout this section, we set

$$h^{\min}_{\lambda,F}(t,b) := \lambda \sum_{D\subset\mathbb{D}} \min_{j\in\{1,\dots,m\}}\{a^D_j\} \int_{(0,\infty)^d} y_i \exp\left\{\alpha b e^{r(T-t)} \sum_{j=1}^d y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y), \tag{4.59}$$

$$h^{\max}_{\lambda,F}(t,b) := \lambda \sum_{D\subset\mathbb{D}} \max_{j\in\{1,\dots,m\}}\{a^D_j\} \int_{(0,\infty)^d} y_i \exp\left\{\alpha b e^{r(T-t)} \sum_{j=1}^d y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y). \tag{4.60}$$

Proposition 4.35. *Let $t \in [0,T]$. Then $\mathbb{R} \ni b \mapsto h^{\min}_{\lambda,F}(t,b)$ and $\mathbb{R} \ni b \mapsto h^{\max}_{\lambda,F}(t,b)$ are strictly increasing and strictly convex. Furthermore, it holds*

$$\lim_{b\to-\infty} h^{\min}_{\lambda,F}(t,b) = \lim_{b\to-\infty} h^{\max}_{\lambda,F}(t,b) = 0, \quad \lim_{b\to\infty} h^{\min}_{\lambda,F}(t,b) = \lim_{b\to\infty} h^{\max}_{\lambda,F}(t,b) = \infty.$$

Proof. This follows by the same analysis as in the proof of Proposition 4.29. □

This proposition ensures the existence of the following notation.

Notation. For some fixed $t \in [0,T]$, we denote the unique root of the equation $(1+\theta)\,\kappa = h^{\max}_{\lambda,F}(t,b)$ w.r.t. b and the unique root of the equation $(1+\theta)\,\kappa = h^{\min}_{\lambda,F}(t,b)$ w.r.t. b by $r^{\max}_{\lambda,F}(t)$ and $r^{\min}_{\lambda,F}(t)$, respectively.

The announced a priori bounds follow immediately from the next result in connection with Proposition 4.35, which make use of the function $h_{\lambda,F}$ given in (4.32).

Proposition 4.36. *For any $(t,p) \in [0,T] \times \Delta_m$ and $b \in \mathbb{R}$, it holds*

$$h^{\min}_{\lambda,F}(t,b) \le h_{\lambda,F}(t,p,b) \le h^{\max}_{\lambda,F}(t,b).$$

Proof. Fix $(t,p) \in [0,T] \times \Delta_m$ and $b \in \mathbb{R}$. For any $(\xi,b) \in \mathcal{U}[t,T]$ and $D \subset \mathbb{D}$, the use of Lemma 4.25 (iii) and (iv) results in

$$g^{\xi,b}(t,J(p,D)) \sum_{k=1}^m a_k^D p_k = \sum_{j=1}^m a_j^D \, p_j \, g^{\xi,b}(t,e_j)$$
$$\le \max_{j\in\{1,\dots,m\}} \{a_j^D\} \sum_{j=1}^m p_j \, g^{\xi,b}(t,e_j) = \max_{j\in\{1,\dots,m\}} \{a_j^D\} \, g^{\xi,b}(t,p).$$

Taking the infimum over all $(\xi,b) \in \mathcal{U}[t,T]$ on both sides, it yields

$$\frac{g(t,J(p,D))}{g(t,p)} \sum_{k=1}^m a_k^D p_k \le \max_{j\in\{1,\dots,m\}} \{a_j^D\}, \quad D \subset \mathbb{D},$$

which is equivalent to

$$\frac{g(t,J(p,D))}{g(t,p)} \sum_{k=1}^m a_k^D p_k \int_{(0,\infty)^d} y_i \exp\left\{\alpha b e^{r(T-t)} \sum_{j=1}^d y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y)$$
$$\le \max_{j\in\{1,\dots,m\}} \{a_j^D\} \int_{(0,\infty)^d} y_i \exp\left\{\alpha b e^{r(T-t)} \sum_{j=1}^d y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y), \quad D \subset \mathbb{D}.$$

Summing over $D \subset \mathbb{D}$ and multiplying with λ gives $h_{\lambda,F}(t,p,b) \le h^{\max}_{\lambda,F}(t,b)$. A passage similar to the above implies $h^{\min}_{\lambda,F}(t,b) \le h_{\lambda,F}(t,p,b)$. □

Corollary 4.37. *The optimal reinsurance strategy $(b^\star_{\lambda,F}(t))_{t\in[0,T]}$ from Theorem 4.33 has the following bounds:*

$$\max\{0, r^{\max}_{\lambda,F}(t)\} \le b^\star_{\lambda,F}(t) \le \min\{1, r^{\min}_{\lambda,F}(t)\}, \quad t \in [0,T].$$

It is simply seen that the bounds for the optimal reinsurance strategy do not have to be in force since $\operatorname{argmin}_{j\in\{1,\dots,m\}}\{a_j^D\}$ varies for different $D \subset \mathbb{D}$ in general. So the lower bound is only in force if, for any $D \subset \mathbb{D}$, $\min_{j\in\{1,\dots,m\}}\{a_j^D\}$ is taken for the same scenario $j \in \{1,\dots,m\}$. The same works for the upper bound. Accordingly, it is to be expected that the range for a possible optimal reinsurance strategy described by these bounds is rather large. In fact, this is illustrated by the Figure 4.3 showing the a priori bounds for the parameters from Section 4.9. The a prior upper bound provides only a useful bound in period from 0 to 2.5 starting by roughly 0.8 with a sharp positive slope. The a prior lower bound has a convex and increasing shapely from around 0.15 to 0.35, i.e. it is never optimal for the insurer to take a full reinsurance.

More interesting bounds involving the optimal reinsurance strategy of the full observable case will be established next. Typically, more uncertainty leads to a more risk-averse behaviour. This is known from literature with comparative studies of partial and full observable settings, compare Liang and Bayraktar [85, Prop. 4.3], Bäuerle and Rieder [31, Thm. 5.6], Bäuerle and Rieder [30, Thm. 6] as well as Bäuerle and Chen [26, Sec. 2.4]. In the presented framework, more risk averseness means that the insurer takes less retention, i.e. the insurer cedes a larger proportion of possible claims to the reinsurer. Such

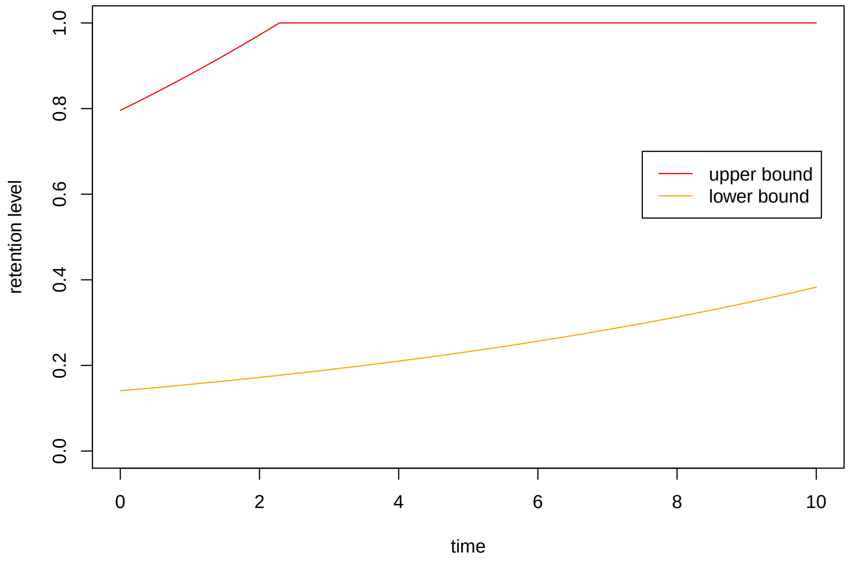

Figure 4.3: The a priori upper bound (red line) and lower bound (orange line) for the optimal reinsurance strategy for parameters from Section 4.9.

a result is also valid in our setting.

As already indicated in Remark 4.28, an order of $A = \{a_1, \dots, a_m\}$ is essential for the comparison result, but the order given in (4.35) is too weak for this purpose. For the definition of a stronger order, it should be noted that a_j, $j = 1, \dots, m$, can be identified with a probability measure on $\mathcal{P}(\mathbb{D})$. From this point of view it is difficult to define an order of A since there is no natural order of the elements of $\mathcal{P}(\mathbb{D})$. However, under the assumption of identical claim size distributions in every LoB, it is easily seen that the order defined in (4.35) holds

$$a_k \preceq a_\ell \quad \Longleftrightarrow \quad \sum_{D \subset \mathbb{D}} a_k^D |D| \leq \sum_{D \subset \mathbb{D}} a_\ell^D |D| \quad \Longleftrightarrow \quad \sum_{i=1}^{d} i \sum_{\substack{D \subset \mathbb{D}: \\ |D|=i}} a_k^D = \sum_{i=1}^{d} i \sum_{\substack{D \subset \mathbb{D}: \\ |D|=i}} a_\ell^D, \quad (4.61)$$

for every $k, \ell \in \{1, \dots, m\}$, where the sums on right-hand side can be interpreted as expectations w.r.t. measures on the set $\mathbb{D} = \{1, \dots, d\}$, which has a natural order. It will turn out that we actually get a useful order under the following assumptions.

Assumption 4.38. Throughout this section, we suppose that

$$F(\mathrm{d}y) = \bar{F}(\mathrm{d}y_1) \otimes \bar{F}(\mathrm{d}y_2) \otimes \cdots \otimes \bar{F}(\mathrm{d}y_m),$$

where $\bar{F}$ is a distribution on $(0, \infty)$ with existing moment generation function and $\otimes$ is the product measure.

Another way of stating the assumption is to say that the claim sizes of every insurance class are independent and identical distributed with distribution $\bar{F}$. Notice that Assumption 4.3 is satisfied under the assumption above, compare Remark 4.4.

The explanation above motivates the following notation.

Notation. For any $k \in \{1, \dots, m\}$, we define

$$\tilde{a}_k(i) := \sum_{\substack{D \subset \mathbb{D}: \\ |D| = i}} a_k^D, \quad i = 1, \dots, d,$$

and

$$\tilde{a}_k := (\tilde{a}_k(1), \dots, \tilde{a}_k(d)).$$

Since a_k^D describes the probability mass of the set D (w.r.t. the distribution which is determined by a_k), $\tilde{a}_k(i)$ is the aggregated mass on all subsets of $\mathbb{D}$ with $i \in \mathbb{D}$ elements. Hence any $\tilde{a}_k$, $k = 1, \dots, m$, characterizes a probability measure on the set $\mathbb{D}$, which is specified in the next notation.

Notation. For any $k \in \{1, \dots, m\}$, we denote the probability measure on $\mathbb{D}$ by $\widetilde{F}_k$, which is defined by

$$\widetilde{F}_k(B) := \sum_{i \in B} \tilde{a}_k(i), \quad B \in \mathcal{P}(\mathbb{D}).$$

In contrast to the set $\mathcal{P}(\mathbb{D})$, the set $\mathbb{D} = \{1, \dots, d\}$ has a natural order, so we will define an order for the above defined probability measures on $\mathbb{D}$, which represents equivalence classes of the set A.

Notation. We denote the equivalence relation on A by $\sim$, which is defined by

$$a_k \sim a_\ell \quad :\Longleftrightarrow \quad \tilde{a}_k = \tilde{a}_\ell,$$

for every $k, \ell \in \{1, \dots, m\}$. Furthermore, for any $k \in \{1, \dots, m\}$, $[a_k]$ is written for the equivalence class of $a_k \in A$ under $\sim$, i.e. $[a_k] := \{a \in A : \tilde{a}_k = \tilde{a}\}$. Moreover, we set $\widetilde{A} = \{[a_1], \dots, [a_m]\}$, where $[a_k]$ and $[a_\ell]$ are either equal or disjoint for all $k, \ell \in \{1, \dots, m\}$.

Justification of the notation. A trivial verification shows that the defined binary relation $\sim$ on A is reflexive, symmetric, transitive and thus an equivalence relation on A. □

Equivalence classes are assumed to be ordered as follows

Assumption 4.39. We suppose that $[a_1] \preceq_{\text{st}} [a_2] \preceq_{\text{st}} \dots \preceq_{\text{st}} [a_m]$, where $\preceq_{\text{st}}$ is an order on the set $\widetilde{A}$ defined by, for any $k, j \in \{1, \dots, m\}$,

$$[a_k] \preceq_{\text{st}} [a_j] \quad :\Longleftrightarrow \quad \widetilde{F}_k(x) \geq \widetilde{F}_j(x), \quad x \in \mathbb{R}. \tag{4.62}$$

Remark 4.40. The defined order can be regarded as the *usual stochastic order*.[6] If $X \sim \widetilde{F}_k$ and $Y \sim \widetilde{F}_j$ and $k \leq j$, then the introduced order is equivalent to $X \preceq_{\text{st}} Y$, where $\preceq_{\text{st}}$ denotes the usual stochastic order. Therefore, the order $\preceq_{\text{st}}$ given by (4.62) is equivalent to

$$\int_{\mathbb{D}} f(x)\, \widetilde{F}_k(\mathrm{d}x) \leq \int_{\mathbb{D}} f(x)\, \widetilde{F}_j(\mathrm{d}x)$$

for all increasing functions $f : \mathbb{D} \to \mathbb{R}$, for which both integrals exist, compare Müller and Stoyan [96, Thm. 1.2.8]. It is also worth noting that the order $\preceq_{\text{st}}$ is consistent with the order defined in (4.35) by choosing f in the equation above as identity.

With the help of the Assumptions 4.38 and 4.39, we can prove the next result which implies the desired comparison result.

[6] For a deeper discussion of the usual stochastic order we refer the reader to Müller and Stoyan [96, Sec. 1.2].

Theorem 4.41. *Let $b_{\lambda,F}$ be the function given by* (4.36) *and $\tilde{b}^{\star}_{\lambda,\bar{c},F}$ be the function given by* (4.57)*. Then, for any $(t,p) \in [0,T] \times \Delta_m$,*

$$b_{\lambda,F}(t,p) \le \tilde{b}^{\star}_{\lambda,w(p),F}(t) \quad t \in [0,T],$$

with

$$w(p) := \left(\sum_{k=1}^{m} a_k^D \, p_k \right)_{D \subset \mathbb{D}}.$$

Proof. Fix $(t,p) \in [0,T] \times \Delta_m$, $\bar{b} \in \mathbb{R}$ as well as $(\xi, b) \in \mathcal{U}[t,T]$. We begin with the observation that the left-hand sides in (4.34) and (4.56) are equal. According to Equations (4.34) and (4.56), we can see that it is sufficient to compare $h_{\lambda,F}(t,p,\bar{b})$ and $h_{\lambda,w(p),F}(t,\bar{b})$ for the comparison of $b_{\lambda,F}$ and $\tilde{b}^{\star}_{\lambda,w(p),F}$. We first observe that a_j is the thinning probability under the condition $p_t = e_j$ and that

$$\begin{aligned} g^{\xi,b}(t,e_j) := \mathbb{E}^{t,e_j}\Bigg[\exp\bigg\{ &- \int_t^T \alpha\, e^{r(T-s)} \big((\mu - r)\xi_s + c(b_s)\big) \mathrm{d}s \\ &- \int_t^T \alpha\, \sigma\, e^{r(T-s)} \xi_s \, \mathrm{d}W_s + \sum_{n=1}^{N_{T-t}} \alpha\, b_{T_n}\, e^{r(T-T_n)} \sum_{\ell=1}^{|Z_n|} Y_n^{\ell} \bigg\} \Bigg], \end{aligned}$$

compare (4.20). Since the integrand above considered to be a function of $|Z_n|$ is increasing and a_j describes the distribution of $|Z_n|$, it follows from Assumption 4.39 in connection with Remark 4.40 that

$$g^{\xi,b}(t,e_1) \le g^{\xi,b}(t,e_2) \le \ldots \le g^{\xi,b}(t,e_m).$$

Using this we conclude with the help of Lemma 4.25 (iii), (iv) and Lemma B.6, for any $D \subset \mathbb{D}$,

$$\begin{aligned} & g^{\xi,b}(t,J(p,D)) \sum_{k=1}^{m} a_k^D p_k \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{ \alpha\, \bar{b}\, e^{r(T-t)} \sum_{j=1}^{d} y_j \mathbb{1}_D(j) \right\} F(\mathrm{d}y) \\ &= \sum_{\ell=1}^{m} p_\ell\, a_\ell^D\, g^{\xi,b}(t,e_\ell) \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{ \alpha\, \bar{b}\, e^{r(T-t)} \sum_{j=1}^{d} y_j \mathbb{1}_D(j) \right\} F(\mathrm{d}y) \\ &\ge \sum_{k=1}^{m} a_k^D p_k \sum_{\ell=1}^{m} p_\ell\, g^{\xi,b}(t,e_\ell) \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{ \alpha\, \bar{b}\, e^{r(T-t)} \sum_{j=1}^{d} y_j \mathbb{1}_D(j) \right\} F(\mathrm{d}y) \\ &= g^{\xi,b}(t,p) \sum_{k=1}^{m} a_k^D p_k \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{ \alpha\, \bar{b}\, e^{r(T-t)} \sum_{j=1}^{d} y_j \mathbb{1}_D(j) \right\} F(\mathrm{d}y), \end{aligned}$$

which yields

$$\begin{aligned} & \frac{g(t,J(p,D))}{g(t,p)} \sum_{k=1}^{m} a_k^D p_k \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{ \alpha\, \bar{b}\, e^{r(T-t)} \sum_{j=1}^{d} y_j \mathbb{1}_D(j) \right\} F(\mathrm{d}y) \\ &\ge \sum_{k=1}^{m} a_k^D p_k \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{ \alpha\, \bar{b}\, e^{r(T-t)} \sum_{j=1}^{d} y_j \mathbb{1}_D(j) \right\} F(\mathrm{d}y), \quad D \subset \mathbb{D}. \end{aligned}$$

by taking the infimum over all $(\xi, b) \in \mathcal{U}[t,T]$. Summing over all $D \subset \mathbb{D}$ and multiplying by λ, we obtain

$$h_{\lambda,F}(t,p,\bar{b}) \geq h_{\lambda,w(p),F}(t,\bar{b}),$$

which completes the proof. □

Corollary 4.42. *Let $\tilde{b}^\star_{\lambda,\bar{c},F}$ be the function given by* (4.57)*. Then the optimal reinsurance strategy under partial information $(b^\star_{\lambda,F}(t))_{t\in[0,T]}$ from Theorem 4.33 satisfies*

$$b^\star_{\lambda,F}(t) \leq \tilde{b}^\star_{\lambda,w(p_{t-}),F}(t), \quad t \in [0,T].$$

It should be noted that $(\tilde{b}^\star_{\lambda,w(p_{t-}),F}(t))_{t\in[0,T]}$ is $\mathfrak{F}^\Phi$-predictable, so that it is an admissible reinsurance strategy. Notice further that $w(p_t) = (\mathbb{E}[\alpha_D \,|\, \mathcal{F}^\Phi_t])_{D\subset\mathbb{D}}$. Therefore $w(p_{t-})$ is the known conditional average thinning probabilities given the available information strict before time t. Therefore Theorem 4.41 makes it legitimate to say that more uncertainty leads to a less or equal retention level since the optimal reinsurance strategy with unknown thinning probabilities is less than or equal to the strategy in the model with known conditional average thinning probabilities given the available information. The comparison result is illustrated in the next section as well as further comparative statistics to provide a deep understanding of the optimal reinsurance strategy under partial information.

4.9 Numerical analyses

In the following some numerical simulations are performed to examine how partial information affects the insurer's optimal reinsurance strategy and to obtain sensitivity analyses of the optimal strategy (under incomplete information). Recall that we have used Clarke's generalized subdifferential to overcome the smoothness assumption on the value function such that a calculation of the solution of the generalized HJB Equation (4.25) is the first step to attain the optimal reinsurance strategy. Using the solution, the optimal feedback control can be designed as the second step. However, the generalized integro PDE 4.25 hardly allows an explicit solution. Therefore we need to rely on numerical procedures. Here we only draw conclusions about the behaviour of the optimal reinsurance strategy by means of the derived comparative result. In order to do this, we assume that the considered insurance company has two LoBs (i.e. $d = 2$) and the insurer's prior belief is given by

$$a_1 = \begin{pmatrix} 4/9 \\ 4/9 \\ 1/9 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 5/9 \\ 2/9 \\ 2/9 \end{pmatrix}, \quad a_3 = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix},$$

where

$$a_j = \left(a_j^{\{1\}}, a_j^{\{2\}}, a_j^{\{1,2\}}\right), \quad i = 1,\ldots,3.$$

The prior probability mass function of $\bar{\alpha}$ is supposed to be

$$\bar{\pi}_{\bar{\alpha}} = \left(\frac{2}{5}, \frac{2}{5}, \frac{1}{5}\right).$$

Since

$$\tilde{a}_1 = \begin{pmatrix} 8/9 \\ 1/9 \end{pmatrix}, \quad \tilde{a}_2 = \begin{pmatrix} 7/9 \\ 2/9 \end{pmatrix}, \quad \tilde{a}_3 = \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix},$$

it holds

$$\widetilde{F}_1(x) \geq \widetilde{F}_2(x) \geq \widetilde{F}_3(x), \quad x \in \mathbb{R}$$

and thus $[a_1] \preceq_{\mathrm{st}} [a_2] \preceq_{\mathrm{st}} [a_3]$, compare Remark 4.40.

Furthermore, it is supposed that the claim sizes from the two insurances classes are independent and identically right-truncated exponential distributed[7] with rate 1 and the truncation is at 10. We denote this distribution by $\bar{F}$. That means, the density function of $\bar{F}$, denoted by $\bar{f}$, is given

$$\bar{f}(y) = \frac{e^{-y}}{1 - e^{-10}}, \quad 0 \leq y \leq 10,$$

and

$$\mathbb{E}[Y_1^1] = \mathbb{E}[Y_1^2] = \frac{1}{1 - e^{-10}}.$$

Hence Assumption 4.3 is fulfilled, compare Remark 4.4, and, due to the identically distributed claim size of each LoB, we can perform the comparison result given in Corollary 4.42 which requested this assumption. Further parameters are fixed in Table 4.1. We are left to specify the parameter κ of the premium principle. We choose $\kappa = \mathbb{E}[\mathrm{d}S_t]$.

parameter	value
x_0	10
T	10
λ	3
r	0.1
μ	0.15
σ	3
α	0.15
θ	0.6
η	0.3

Table 4.1: Simulation parameters for Section 4.9.

That is, by Proposition 4.21, we have

$$\begin{aligned}\kappa &= \lambda \sum_{k=1}^{m} \pi_{\bar{\alpha}}(k) \sum_{D \subset \mathbb{D}} a_k^D \sum_{i=1}^{d} \mathbb{1}_D(i) \mathbb{E}\big[Y_1^i\big] = \lambda\, \mathbb{E}\big[Y_1^1\big] \sum_{k=1}^{m} \pi_{\bar{\alpha}}(k) \sum_{D \subset \mathbb{D}} a_k^D\, |D| \\ &= \lambda\, \mathbb{E}\big[Y_1^1\big] \sum_{k=1}^{m} \pi_{\bar{\alpha}}(k) \sum_{i=1}^{2} \tilde{a}_k(i)\, i = \frac{6}{5} \frac{4}{1 - e^{-10}}.\end{aligned}$$

In the following we suppose that the (unobservable) realization of $\bar{\alpha}$ is a_2. Before we turn our attention to the optimal reinsurance strategy under partial information, let us perform the full information case, where the optimal reinsurance strategy is given by (4.57). In order to be consistent with the incomplete information case, we recalculate κ for the fully observable case by using the same approach and we use the notation $\bar{\kappa}$

[7] The general definition of truncated distributions can be found in Cramér [44, Sec. 19.3].

instead of κ. With the help of Wald's equation, we obtain

$$\mathbb{E}[S_t] = \mathbb{E}\Bigg[\sum_{i=1}^{d}\sum_{n=1}^{N_t^i} Y_n^i\Bigg] = \sum_{i=1}^{d}\mathbb{E}\big[N_t^i\big]\,\mathbb{E}[Y_1^i] = \mathbb{E}[Y_1^1]\sum_{i=1}^{d}\mathbb{E}\big[N_t^i\big] = \mathbb{E}[Y_1^1]\sum_{i=1}^{d}\lambda\sum_{\substack{D\subset\mathbb{D}:\\ D\ni i}}\alpha_D\, t$$
$$= \mathbb{E}[Y_1^1]\,\lambda\, t, \quad t\geq 0,$$

and thus

$$\bar{\kappa} = \lambda\,\mathbb{E}[Y_1^1] = \frac{3}{1-e^{-10}}.$$

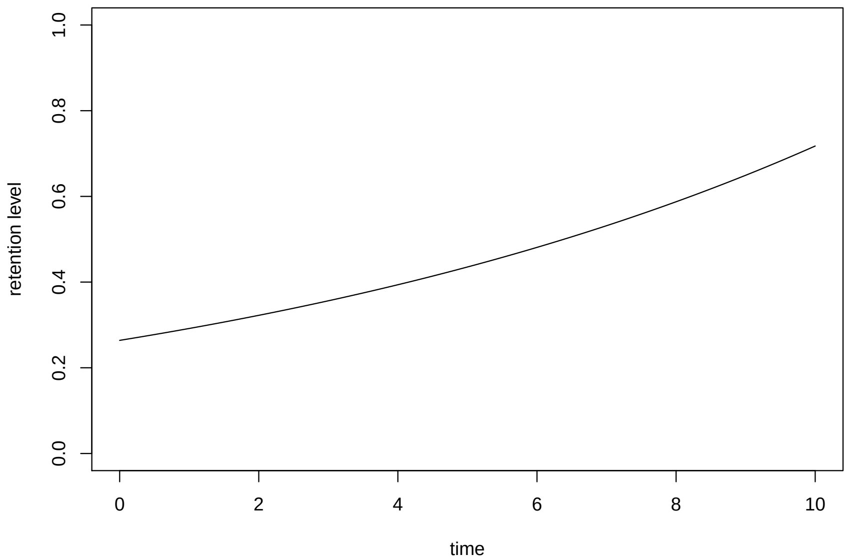

Figure 4.4: The development of the optimal reinsurance strategy in time under full information for the parameter selection of Section 4.9.

The optimal reinsurance strategy under complete information is displayed in Figure 4.4 which shows a rising optimal reinsurance strategy from approximately 0.25 to 0.7 at time 10. The (exponential) increase of the strategy is explained by the exponential utility function. In the given parameter choice, the surplus rises (compare Figure 4.9) in most scenarios and a loss is valued less strongly for a high surplus than for a low surplus. Therefore the insurer behaves more risky at a high surplus which explains the more risky reinsurance strategy at the end of the considered time interval. Since the risk aversion depends on the parameter α, we study graphically the effect of α on the optimal reinsurance strategy at time 5 in Figure 4.5. It can be seen that it is optimal for the insurer to retain all the risk to itself for a very small α which is associated with a less risk aversion. With increasing α, the optimal reinsurance strategy decreases exponential and converged to zero, where the convergence follows immediately from the first order condition given in (4.56).

It is also worth considering the effect of the safety loading parameter θ of the reinsurer on the optimal strategy. This effect is illustrated in Figure 4.6 which displays the expected behaviour that for a small θ (which means a cheap reinsurance premium) it is optimal to transfer the entire risk to the reinsurer. After that, the optimal strategy in-

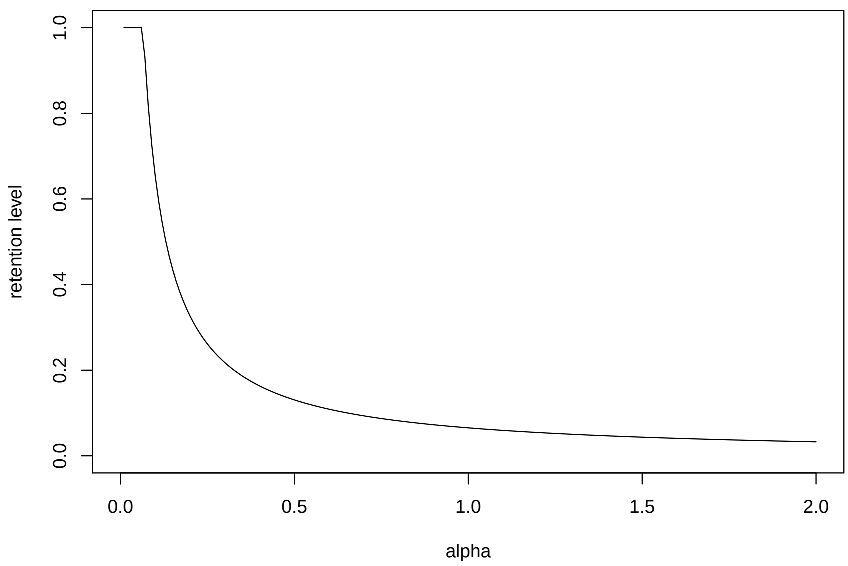

Figure 4.5: The effect of α on the optimal reinsurence strategy at time $t = 5$ in the case of complete information for the parameter selection of Section 4.9.

creases until the premium is so expensive that it is optimal not to purchase a reinsurance contract.

After having analyzed the optimal reinsurance strategy in the fully information case, we examine the optimal investment strategy given by (4.29), which is the same as that under full and partial information. This strategy is illustrated in Figure 4.7 for the given parameter choice of this section. The increasing property can be explained by the exponential utility function again as for the optimal reinsurance strategy.

Now we turn to the comparison result from Section 4.8.2. In Figure 4.3 we have already pictured the a priori bounds for the parameters used in this section. In Figure 4.8, we show these bounds together with two trajectories (black and blue lines) of the reinsurance strategy $(\tilde{b}^{\star}_{\lambda,w(p_{t-}),F}(t))_{t\in[0,T]}$ with $w(p) = (\sum_{k=1}^{m} a_k^D\, p_k)_{D\subset\mathbb{D}}$, which provide an upper bound for the optimal reinsurance strategy for each scenario according to Corollary 4.42. In both scenarios (black and blue lines) the upper bounds, which take the observed data into account, generate a much better upper bound than the a priori one. In combination with the a priori lower bound, the insurer obtain a quite small range of possible optimal reinsurance strategies up to time 7. Afterwards the range becomes bigger.

To conclude the numerical analysis, we consider the path of the surplus process in an insurance loss scenario for three different reinsurance strategies in Figure 4.9. The red line displays the path of the surplus process in the case of full reinsurance (i.e. retention level of 0), which is evident from the fact that this path contains no jumps because the reinsurer covers all losses. The full reinsurance is purchased through a negative premium rate which explains the downward trend. For a constant reinsurance strategy of 0.5, the trajectory of the surplus process is plotted by the blue line. This path is similar to the path in the case of the reinsurance strategy $(\tilde{b}^{\star}_{\lambda,w(p_{t-}),F}(t))_{t\in[0,T]}$ with $w(p) = (\sum_{k=1}^{m} a_k^D\, p_k)_{D\subset\mathbb{D}}$ (black line), which was illustrated in Figure 4.8. The increasing property of this strategy can be recognized by the circumstance that the jump sizes (part of the losses the insurer has to pay) are smaller at the beginning of the

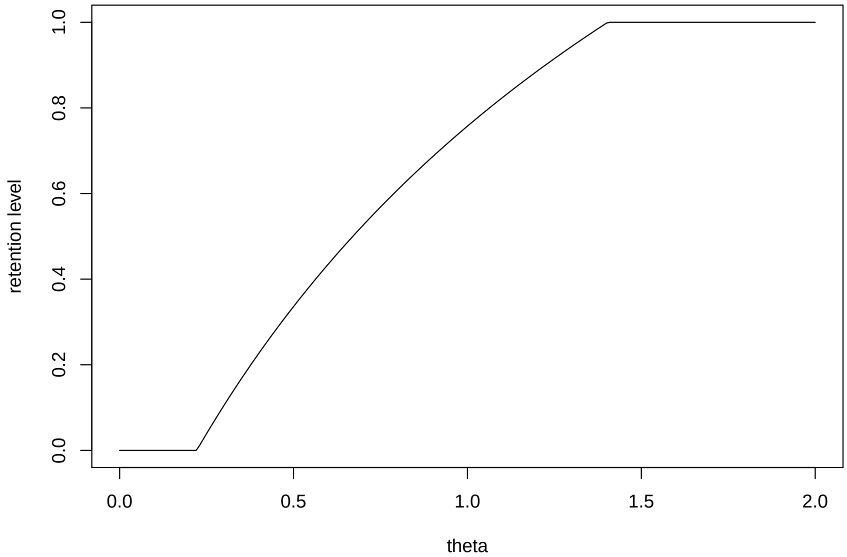

Figure 4.6: The effect of θ on the optimal reinsurence strategy at time $t = 5$ in the case of complete information for the parameter selection of Section 4.9.

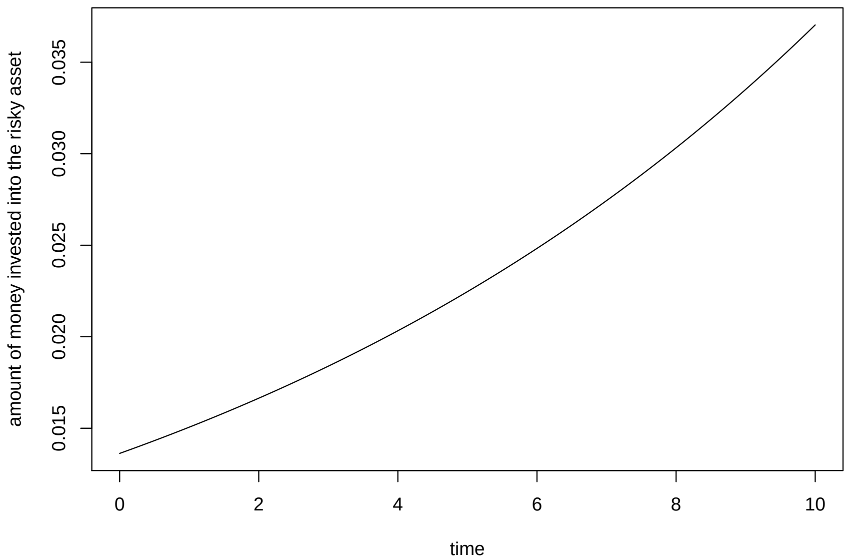

Figure 4.7: The Development of the optimal investment strategy in time for the parameter selection of Section 4.9.

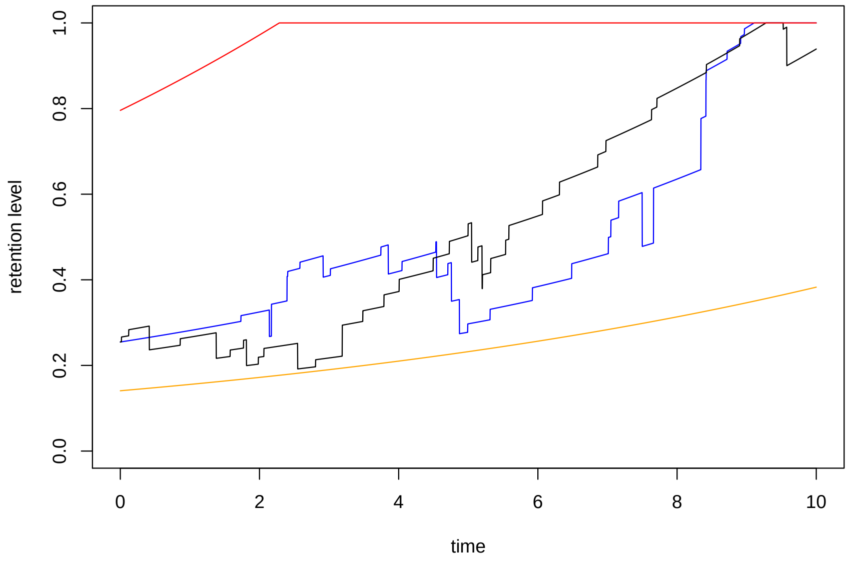

Figure 4.8: The a priori upper bound (red line) and lower bound (orange line) for the optimal reinsurance strategy and two paths of the reinsurance strategy $(\tilde{b}^{\star}_{\lambda,w(p_{t-}),F}(t))_{t\in[0,T]}$ with $w(p) = (\sum_{k=1}^{m} a_k^D\, p_k)_{D\subset\mathbb{D}}$.

observed time interval and larger at the end compared to the constant strategy.

4.10 Comments on generalizations

We close the chapter with a discussion about generalizations of the presented setting.

Multivariate reinsurance strategy. It was already pointed out in the introduction that the reinsurance strategy is multivariate in the literature with multi-dimensional risk models. However, in the literature with common shock models, the optimal reinsurance strategies are only stated in the case of two insurance classes since it is necessary to considered different cases where the number of cases may increase geometrically, compare Yuen et al. [121, Remark 4.3]. A similar effect can be presumed in our setting, which is explained below in detail.

To choose a candidate for an optimal multivariate reinsurance strategy $(b^1_{\lambda,F},\dots,b^d_{\lambda,F})$ with $b^\ell_{\lambda,F}(t)\in[0,1]$ for $t\geq 0$ and $\ell\in\{1,\dots,d\}$ (i.e. a reinsurance strategy for each LoB), we have to solve the following system of equations w.r.t. $(b_1,\dots,b_d)$:

$$(1+\theta_\ell)\kappa_\ell = \lambda \sum_{D\subset\mathbb{D}} \frac{g(t,J(p,D))}{g(t,p)} \sum_{k=1}^{m} a_k^D p_k \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\Big\{\alpha\, e^{r(T-t)} \sum_{j=1}^{d} b_j\, y_j \mathbb{1}_D(j)\Big\} F(\mathrm{d}y), \quad \ell = 1,\dots,d,$$

where θ_ℓ denotes the safety risk load parameter of the reinsurer for the ℓth LoB and κ_ℓ the corresponding parameter of the reinsurance premium principle. Similar to Yuen et al. [121, p. 5], the uniqueness and existence of the solution to the equation above can

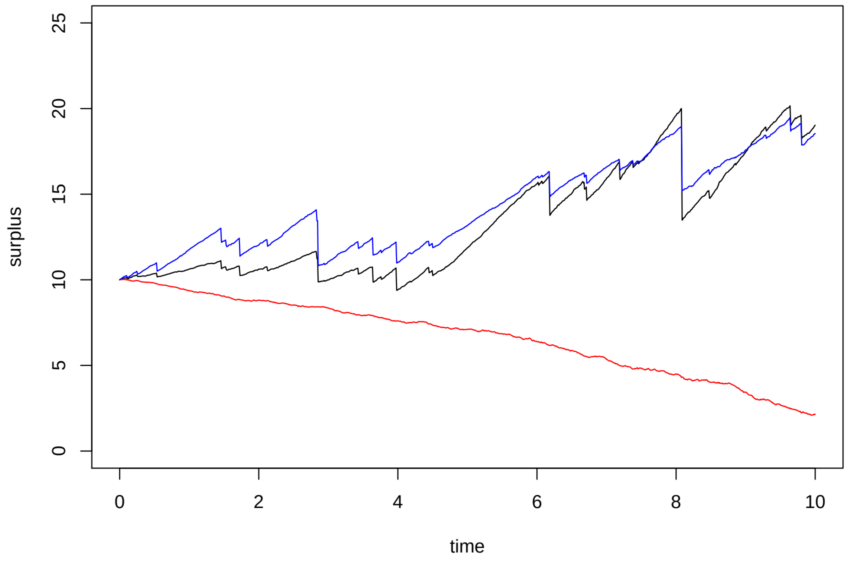

Figure 4.9: Trajectories of the surplus process for an insurance loss scenario in the cases of full reinsurance (red line), constant retention level of 0.5 (blue line) and the reinsurance strategy $(\tilde{b}^{\star}_{\lambda,w(p_{t-}),F}(t))_{t\in[0,T]}$ with $w(p) = (\sum_{k=1}^{m} a_k^D\, p_k)_{D\subset\mathbb{D}}$ (black line).

be shown. But to make sure that the retention level $(b^1_{\lambda,F},\dots,b^d_{\lambda,F})$ takes values in $[0,1]^d$, we need to discuss various cases w.r.t. the order of the unique roots (i.e. $d!$ cases, cf. Yuen et al. [121, p. 5]) which results in a much more complicated optimality analysis.

Regime-switching model. In the present framework, the underlying environment, which determines the interdependencies between the LoBs, does not change. This assumption can be weakened by introducing an unobservable Markov chain with finite state space and supposing that the thinning probabilities change over time according to the state of the hidden Markov model[8]. Such regime-switching models (also called Markov-modulated models) have been intensively studied in actuarial mathematics literature, where the claim arrival intensity and claim size distribution depends on the state of the chain (cf. e.g. Bäuerle [24]) or the aggregated claim rate of a diffusion risk process (cf. e.g. Elliott et al. [55]). Extending the Bayesian setting of this chapter to a hidden Markov model would lead to a different filter equation which can be determined by using the filter result for marked point process observations given in Brémaud [20, Thm. VIII.T9]. It is expectable that the presented solution procedure can be applied analogously. However, in the next chapter we will use an alternative approach to deal with unobservable thinning probabilities which does not allow an extension to a regime-switching model.

[8]For a general treatment of hidden Markov models see e.g. Elliott et al. [54].

Chapter 5

Optimal investment and reinsurance with unknown claim arrival intensities

In Chapter 4 we have studied the control problem (P) under the assumption of observability of the background intensity and the claim size distribution. In this chapter we relax a part of this assumption, namely we suppose that the background intensity is unobservable; the claim size distribution is still assumed to be known. Moreover, we deal with the unknown thinning probabilities in an alternative way.

5.1 Setting

As indicated in the introduction of this chapter, we suppose that the prior distribution Π_ϑ is a one-point distribution such that the claim size distribution is observable for the insurer. In the following, we denote the observable loss distribution by F. Furthermore, we suppose that the Assumptions 3.2, 3.3 and 3.6 are in force.

Prior distribution for the thinning probabilities. Recall the approach in Section 4.1 with the approximation of the $(\ell-1)$-dimensional probability simplex Δ_ℓ (the unknown thinning probabilities $\bar{\alpha}$ take values in Δ_ℓ) by a finite number of points. So the prior distribution $\Pi_{\bar{\alpha}}$ of $\bar{\alpha}$ has been chosen to be defined on a finite set. Under this assumption we have derived a filter process for the probability mass function using the available information of the insurer, in which the filter process was finite dimensional because of the finite discretization. This was the key to obtain a reduced control problem whose optimal strategy is also optimal for the original incomplete information. In this chapter we could do the same procedure as in the previous regarding the thinning probabilities. But the reduction step of the partially observable problem (P) is also possible without the discretization of the probability simplex Δ_ℓ made in Assumption 4.1. This would lead to a stochastic control problem in infinite dimension in general. However, in this section, we are going to use a parametric Bayesian approach which avoids the discretization of the probability simplex. From this point of view, this approach is more general since the thinning probabilities can take every value in the interior of the probability simplex.

Before stating the assumptions of this section, let us recall that $(Z_n)_{n\in\mathbb{N}}$ denotes a sequence of conditional iid random elements (conditioned on $\bar{\alpha}$) taking values in $(\mathcal{P}(\mathbb{D}), \mathcal{P}(\mathcal{P}(\mathbb{D})))$, where Z_n describes the affected LoBs by the trigger event at T_n, namely the insurance classes $i \in Z_n$ are affected, compare Section 3.1. For the Bayesian approach we have to count the realisations of every Z_n that occurs up to time t, which is accomplished with the help of the following notation.

Notation. For any $n \in \mathbb{N}$, $D \subset \mathbb{D}$ and $t \geq 0$, we define $C_0(D) := 0$ and

$$C_n(D) := \sum_{i=1}^{n} \mathbb{1}_{\{Z_i = D\}}, \qquad C_n := \big(C_n(D)\big)_{D \subset \mathbb{D}},$$
$$q_D(t) := C_{N_t}(D), \qquad q_t := \big(q_D(t)\big)_{D \subset \mathbb{D}}.$$

Thus $q = (q_t)_{t \geq 0}$ is a $\mathbb{N}_0^\ell$-valued process and $q_D(t)$ counts how many times the realizations of $(Z_n)_{n \in \mathbb{N}}$ are D up to time t. Note that we can use $(C_n)_{n \in \mathbb{N}}$ instead of $(Z_n)_{n \in \mathbb{N}}$ to describe the multivariate claim arrival process N since

$$\{i \in Z_n\} = \big\{i \in D \text{ with } C_n(D) > C_{n-1}(D)\big\}.$$

According to the notation above, for any $t \geq 0$, the ℓ-dimensional random vector C_n is multinomial distributed with parameters n and $\bar{\alpha} = (\alpha_D)_{D \subset \mathbb{D}} \in \Delta_\ell$. The probability mass function of C_n is given by

$$f(x \,|\, n, \bar{\alpha}) = \frac{n!}{\prod_{D \subset \mathbb{D}} x_D!} \prod_{D \subset \mathbb{D}} \alpha_D^{x_D}, \quad x = (x_D)_{D \subset \mathbb{D}} \in \mathbb{N}_0^\ell \text{ with } \sum_{D \subset \mathbb{D}} x_D = n,$$

see e.g. DeGroot [49, Sec. 5.2]. We write shortly

$$C_n \,|\, n, \bar{\alpha} \sim \mathrm{Mult}(n, \bar{\alpha}).$$

As indicated at the beginning of this Section, we apply a parametric Bayesian approach in this chapter. More precisely, we choose the Dirichlet distribution as the prior for the thinning probabilities $\bar{\alpha} = (\alpha_D)_{D \subset \mathbb{D}}$.

Definition 5.1 (Dirichlet distribution; [49], p. 49)**.** A random vector $X = (X_1, \ldots, X_k)$ has a *Dirichlet distribution* with parameter vector $\bar{\beta} = (\beta_1, \ldots, \beta_k) \in (0, \infty)^k$, if the probability density function $f_{\bar{\beta}}(\cdot)$ of X is given by

$$f_{\bar{\beta}}(x) = \frac{\Gamma(\beta_1 + \ldots + \beta_k)}{\Gamma(\beta_1) \cdot \ldots \cdot \Gamma(\beta_n)} \prod_{i=1}^{k} x_i^{\beta_i - 1}, \quad x = (x_1, \ldots, x_k) \in \mathring{\Delta}_k,$$

where Γ denotes the gamma function, i.e.

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} \, \mathrm{d}x$$

for all complex numbers z with positive real part. We write shortly

$$X \,|\, \bar{\beta} \sim \mathrm{Dir}(\bar{\beta}).$$

The assumption about the dependence mechanism of the insurance classes $\bar{\alpha}$ is summarized in the next assumption.

Assumption 5.2. We suppose that the $\mathcal{F}_0$-measurable random vector $\bar{\alpha} = (\alpha_D)_{D \subset \mathbb{D}}$ is Dirichlet distributed with parameter vector $\bar{\beta} = (\beta_D)_{D \subset \mathbb{D}} \in (0, \infty)^\ell$, i.e.

$$\bar{\alpha} \,|\, \bar{\beta} \sim \mathrm{Dir}(\bar{\beta}).$$

Remark 5.3. Notice that a Dirichlet distributed random vector takes values in the interior of probability simplex. That is, the probability for every possible dependency between the LoBs is positive. The parameter $\bar{\beta}$ of the Dirichlet distribution can be chosen with already existing date by using for example the maximum likelihood method. In the case of non-existence of pre-information, an uninformed prior can be selected which is $\bar{\beta} = (1, \dots, 1)$ for the Dirichlet distribution such that the thinning probabilities are uniform distributed on the probability simplex.

The reason for the choice of the Dirichlet distribution as the prior is the conjugated property of the Dirichlet prior, which is stated next. If the posterior distribution is of the same type as the prior distribution, then we call the prior *conjugated.* That is, a family of distributions which a conjugated prior belongs to, is closed under sampling.

Theorem 5.4 ([49], Thm 9.8.1). *The posterior distribution of $\bar{\alpha}$ given $C_n = c$ with $c = (c_D)_{D\subset\mathbb{D}} \in \mathbb{N}_0^\ell$ is a Dirichlet distribution with parameter vector $\bar{\beta}+c = (\beta_D+c_D)_{D\subset\mathbb{D}}$, i.e. the posterior density of $\bar{\alpha}$ is*

$$f_{\bar{\beta}}(\bar{\alpha}\,|\,c) = \frac{\Gamma\big(\sum_{D\subset\mathbb{D}}(\beta_D + c_D)\big)}{\prod_{D\subset\mathbb{D}}\Gamma\big(\beta_D + c_D\big)} \prod_{D\subset\mathbb{D}} \alpha_D^{\beta_D+c_D-1}, \quad \bar{\alpha} = (\alpha_D)_{D\subset\mathbb{D}} \in \mathring{\Delta}_\ell.$$

It should be noticed that the marginal distribution of the jth component of a $\mathrm{Dir}(\bar{\beta})$-distributed random vector $(X_1, \dots, X_k)$, $\bar{\beta} = (\beta_1, \dots, \beta_k) \in (0,\infty)^k$, is Beta distributed[1] with parameters β_j and $\sum_{i=1}^k \beta_i - \beta_j$, compare DeGroot [49, p. 50]. This fact implies immediately the following result.

Corollary 5.5. *The posterior distribution of α_D given $C_n = c$ with $c = (c_D)_{D\subset\mathbb{D}} \in \mathbb{N}_0^\ell$ is a Beta distribution with parameters $\beta_D + c_D$ and $\sum_{E\subset\mathbb{D}\setminus\{D\}}(\beta_D + c_D)$, i.e. the posterior density of α_D is*

$$f_{\bar{\beta}}(\alpha_D\,|\,c) = \frac{\Gamma\big(\sum_{D\subset\mathbb{D}}(\beta_D + c_D)\big)}{\Gamma\big(\beta_D + c_D\big)\Gamma\Bigg(\sum\limits_{E\subset\mathbb{D}\setminus\{D\}}(\beta_E + c_E)\Bigg)} \alpha_D^{\beta_D+c_D-1}(1-\alpha_D)^{\sum_{E\subset\mathbb{D}\setminus\{D\}}(\beta_E+c_E)-1}$$

for $\alpha_D \in (0,1)$ and 0 otherwise.

Notation. Throughout this chapter, $\|\cdot\|$ denotes the ℓ_1-norm, i.e. $\|x\| = \sum_{i=1}^n |x_i|$ for some $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

According to the theorem and corollary above, we have

$$\begin{aligned} \bar{\alpha}\,|\,\bar{\beta}, Z_1, \dots, Z_{N_t} &\sim \mathrm{Dir}(\bar{\beta} + q_t), \\ \alpha_D\,|\,\bar{\beta}, Z_1, \dots, Z_{N_t} &\sim \mathrm{Beta}(\beta_D + q_D(t), \|\bar{\beta} + q_t\| - \beta_D - q_D(t)). \end{aligned}$$

The next result yields the predictive distribution[2] of the categorical variable Z given observed categories, in which the proof incorporates the following notations.

Notation. From now on, $\mathbb{P}_x$ denotes the conditional probability measure $\mathbb{P}$ given x, where x is either a random element or a single event, and $\mathbb{E}^x$ denotes the corresponding conditional expectation w.r.t. $\mathbb{P}_x$.

The notation plays an important role in the determination of Hamilton-Jacobi-Bellman equation in Section 5.4 and in the next proof.

[1] For the definition of the Beta distribution we refer the reader to DeGroot [49, Sec. 4.9].

[2] The predictive distribution is the conditional distribution of a new observation given some date, compare Klugman et al. [78, Def. 2.43].

Notation. Throughout this chapter, let $v : \mathbb{N}_0^\ell \times \mathcal{P}(\mathbb{D}) \to \mathbb{N}_0^\ell$ denote a function given by

$$v(q, D) := \big(q_E + \mathbb{1}_{\{E=D\}}\big)_{E\subset\mathbb{D}}, \tag{5.1}$$

where $q = (q_D)_{D\subset\mathbb{D}}$. Moreover, we write $v(q, D)_{D'}$ for the D'th component of the sequence $\big(q_E + \mathbb{1}_{\{E=D\}}\big)_{E\subset\mathbb{D}}$.

Proposition 5.6. *For any $t \geq 0$, we have*

$$\mathbb{P}(Z_{N_t+1} = D \,|\, Z_1, \dots, Z_{N_t}) = \frac{\beta_D + q_D(t)}{\|\bar\beta + q_t\|}.$$

Proof. Fix $D \subset \mathbb{D}$ and $t \geq 0$. First of all, let $\bar z_n := (z_1, \dots, z_n) \in \mathcal{P}(\mathbb{D})^n$ be a realization of $\bar Z_n := (Z_1, \dots, Z_n)$. Due to the Bayes' rule, Assumption 3.2 and Theorem 5.4, we have

$$\mathbb{P}(Z_{n+1} = D \,|\, \bar Z_n = \bar z_n) = \int_{\Delta_\ell} \alpha_D\, f_{\bar\beta}(\bar\alpha \,|\, \bar Z_n = \bar z_n)\, \mathrm{d}\bar\alpha.$$

Consequently, using the fact that $\Gamma(n+1) = n\,\Gamma(n)$ for all $n \in \mathbb{N}$, we obtain

$$\begin{aligned}
&\mathbb{P}(Z_{n+1} = D \,|\, Z_1, \dots, Z_{N_t}) = \int_{\Delta_\ell} \alpha_D\, f_{\bar\beta}(\bar\alpha \,|\, q_t)\, \mathrm{d}\bar\alpha \\
&= \frac{\Gamma\big(\sum_{E\subset\mathbb{D}}(\beta_E + q_E(t))\big)}{\prod_{E\subset\mathbb{D}} \Gamma(\beta_E + q_E(t))} \int_{\Delta_\ell} \prod_{E\subset\mathbb{D}} \alpha_E^{\mathbb{1}_{\{E=D\}}+\beta_E+q_E(t)-1}\, \mathrm{d}\bar\alpha \\
&= \frac{\Gamma\big(\sum_{E\subset\mathbb{D}}(\beta_E + q_E(t)) + 1\big)}{\sum_{E\subset\mathbb{D}}(\beta_E + q_E(t))} \frac{\beta_D + q_D(t)}{\Gamma(\beta_D + q_D(t) + 1)\prod_{E\subset\mathbb{D}\setminus\{D\}} \Gamma(\beta_E + q_E(t))} \times \\
&\quad \int_{\Delta_\ell} \alpha_D^{\beta_D+q_D(t)} \prod_{E\subset\mathbb{D}\setminus\{D\}} \alpha_E^{\beta_E+q_E(t)-1}\, \mathrm{d}\bar\alpha \\
&= \frac{\beta_D + q_D(t)}{\|\bar\beta + q_t\|} \times \\
&\quad \int_{\Delta_\ell} \frac{\Gamma\big(\sum_{E\subset\mathbb{D}}(\beta_E + q_E(t)) + 1\big)}{\Gamma(\beta_D + q_D(t) + 1)\prod_{E\subset\mathbb{D}\setminus\{D\}} \Gamma(\beta_E + q_E(t))} \alpha_D^{\beta_D+q_D(t)} \prod_{E\subset\mathbb{D}\setminus\{D\}} \alpha_E^{\beta_E+q_E(t)-1}\, \mathrm{d}\bar\alpha,
\end{aligned}$$

where the integral is 1 since the integrand is the density of the Dirichlet distribution with parameter vector $v(\bar\beta + q_t, D)$. □

The proposition provides the distribution of a new thinning Z given the appeared categories up to time t. This distribution gains in interest in the proof of Lemma 5.20 (vi).

Due to the conjugation property of the prior for $\bar\alpha$, the posterior distribution of $\bar\alpha$ as well as the parameter of this distribution, which is described by $(p_t)_{t\geq0}$, are known. Thus $(q_t)_{t\geq0}$ provides all available information about $\bar\alpha$ which is included in the observable filtration $\mathfrak{G}$. Therefore, instead of a filter equation, we can use the process $(q_t)_{t\geq0}$, which describes the parameter of the posterior distribution, for the reduction of the control problem (P) under partial information to one with complete information.

After the reduction we use the stochastic control approach to solve the reduced problem. For this purpose, we need the compensated process of $(q_t)_{t\geq0}$. To determine this process, it is necessary to know the dynamics of the projection of the unobservable background intensity Λ to the observable filtration. So firstly, we have to specify the assumption about Λ.

Prior distribution for the background intensity. In contrast to Chapter 4, we regard the trigger process N as mixed Poisson process with mixing distribution Π_Λ, i.e. Λ is a positive $\mathcal{F}_0$-measurable random variable. We suppose that Π_Λ is defined on a finite set such that we can derive a finite dimensional filter equation which describes the background intensity Λ given the available information up to any time t. This ansatz is similar to the model for the thinning probabilities $\bar{\alpha}$ in Section 4.1.

Assumption 5.7. Let $m \in \mathbb{N}$ be fixed. We assume that Λ is an $\mathcal{F}_0$-measurable random variable taking values in the measure space $(A, \mathcal{A})$, where $A := \{\lambda_1, \ldots, \lambda_m\}$ with $\lambda_j \in (0, \infty)$, $j = 1, \ldots, m$, and $\mathcal{A} := \mathcal{P}(A)$. Without restriction to generality, we suppose that $\lambda_1 < \lambda_2 < \ldots < \lambda_m$. Furthermore, Λ and $\bar{\alpha}$ are assumed to be independent.

Later, bounds for the optimal strategy (Corollary 5.29) and a comparative result with optimal strategy in the fully observable case (Corollary 5.32) are derived under use of the order $\lambda_1 < \lambda_2 < \ldots < \lambda_m$, which is an order of the possible background intensities from the best to the worst case scenario from the insurer's point of view. As already mentioned in Section 4.8.2, such orders are typically the basis for determining comparison results. Beside the order of the background intensities, an order is also required for the thinning probabilities which is not as obvious as for the background intensities since the harm of a scenario for the thinning probabilities also depends on the claim size distribution. It will turn out that a proper order can only be given under further assumptions to the claim size distribution, compare Section 5.7.

According to the assumption above, the prior distribution Π_Λ of Λ is defined on A. That means, the insurer is aware that the background intensity is one of the values in A, but it is not known which one. The insurance company has only a prior guess about the distribution of Λ which represents expert knowledge about the unknown background intensity. That is,

$$\Pi_\Lambda(B) = \sum_{j \in \{1,\ldots,m\}: \lambda_j \in B} \pi_\Lambda(j), \quad B \in \mathcal{A},$$

where

$$\pi_\Lambda(j) := \mathbb{P}(\Lambda = \lambda_j), \quad j = 1, \ldots, m.$$

Notation. We write $\bar{\pi}_\Lambda := (\pi_\Lambda(1), \ldots, \pi_\Lambda(m)) \in \Delta_m$ for the m-dimensional vector which describes the probability mass function of Π_λ.

In contrast to the prior for the thinning probabilities, the prior for the background intensity is a distribution on a finite set. Due to the finiteness, we can describe the update of the prior guess of the distribution of Λ with a finite dimensional filter equation which will be determined in the following section. Before we turn to the filter problem, a further assumption must be made regarding the distribution of the claim sizes.

Claim size distribution. We need the same requirements on the claim sizes as in Chapter 4. Thus we suppose that Assumption 4.3 is satisfied, i.e.

$$M_F(z) := \mathbb{E}\left[\exp\left\{z \sum_{i=1}^{d} Y_1^i\right\}\right] = \int_{(0,\infty)^d} \exp\left\{z \sum_{i=1}^{d} y_i\right\} F(\mathrm{d}y) < \infty, \quad z \in \mathbb{R}, \tag{5.2}$$

Due to this assumption, we can state the following properties which will be used for the verification of the solution of the later announced reduced control problem, see Section 5.3.

Lemma 5.8. *Let $z \in \mathbb{R}$ be an arbitrary constant. Then there exists constants $0 < C_1 < \infty$ and $0 < C_2 < \infty$ such that*

(i) $\mathbb{E}\big[Y_1^j \exp\big\{z \sum_{i=1}^d Y_1^i\big\}\big] \le C_1, \quad j = \{1, \dots, d\}$,

(ii) $\mathbb{E}\Big[\exp\Big\{z \sum_{i=1}^d \sum_{k=1}^{N_t} Y_k^i\Big\}\Big] \le C_2, \quad t \in [0, T]$.

Proof. The first statement matches Lemma 4.5 (i). Regarding the second statement, we obtain, by following the same line of arguments as in the proof of Lemma 4.5 with $\mathbb{P}$ replaced by $\mathbb{P}_\lambda$ in the application of law of total variation,

$$\mathbb{E}\Bigg[\exp\Bigg\{z \sum_{i=1}^d \sum_{k=1}^{N_t} Y_k^i\Bigg\}\Bigg] = \mathbb{E}\Bigg[\exp\Bigg\{\lambda\, t\Bigg(\mathbb{E}\Bigg[\exp\Bigg\{z \sum_{i=1}^d Y_1^i\Bigg\}\Bigg] - 1\Bigg)\Bigg\}\Bigg],$$

where, by definition of the prior distribution Π_Λ,

$$\begin{aligned}\mathbb{E}\Bigg[\exp\Bigg\{\lambda\, t\Bigg(\mathbb{E}\Bigg[\exp\Bigg\{z \sum_{i=1}^d Y_1^i\Bigg\}\Bigg] - 1\Bigg)\Bigg\}\Bigg] &= \sum_{j=1}^m \pi_\Lambda(j) \exp\Bigg\{\lambda_j\, t\Bigg(\mathbb{E}\Bigg[\exp\Bigg\{z \sum_{i=1}^d Y_1^i\Bigg\}\Bigg] - 1\Bigg)\Bigg\} \\ &= \sum_{j=1}^m \pi_\Lambda(j) \exp\big\{\lambda_j\, T\, M_F(z)\big\} =: C_2\end{aligned}$$

which is finite according to (5.2) for all $t \in [0, T]$. □

5.2 Filtering and reduction

The task is to reduce the partially observable control problem (P) within the introduced framework to one with a state process that describes the available information about the unknown background intensity and interdependencies between the LoBs. For this purpose, we proceed for the background intensity as in Section 4.2, i.e. we determine a filter equation for the background intensity.

Filtering of the background intensity. We want to describe the conditional distribution of λ given the available information up to time t. For this purpose, we are going to apply the filter result for point-process observations stated in Theorem 2.94, where the observed filtration is the natural filtration of the background process N denoted by $\mathfrak{F}^N = (\mathcal{F}_t^N)_{t \ge 0}$.

Recall that, by Theorem 2.69, the natural filtration of an SPP is right-continuous. This justifies the following notation according to Proposition 2.31.

Notation. Throughout this chapter, we denote the càdlàg modification of the process $(\mathbb{E}[\Lambda | \mathcal{F}_t^N])_{t \ge 0}$ by $(\widehat{\Lambda}_t)_{t \ge 0}$, i.e.

$$\widehat{\Lambda}_t = \mathbb{E}[\Lambda \,|\, \mathcal{F}_t^N], \quad t \ge 0.$$

From Proposition 2.91, we know that $(\widehat{\Lambda}_t)_{t \ge 0}$ is an $\mathfrak{F}^N$-intensity of the mixed Poisson process N since the mixing distribution Π_Λ has a finite mean. The finite mean implies further that N is integral, compare Theorem 2.78 (ii).

Notation. From now on, $(\Omega, \mathcal{F}_\infty^N, \mathfrak{F}^N, \mathbb{P})$ denotes the filtrated probability space which is modified as described in Remark 2.70 such that the usual conditions are satisfied.

Furthermore, let $\widehat{N} = (\widehat{N}_t)_{t\geq 0}$ denote the compensated process of N defined by

$$\widehat{N}_t := N_t - \int_0^t \widehat{\Lambda}_s \,\mathrm{d}s, \quad t \geq 0, \tag{5.3}$$

and we write $(p_j(t))_{t\geq 0}$, $j = 1, \ldots, m$, for the càdlàg modification of the processs $(\mathbb{P}(\Lambda = \lambda_j \,|\, \mathcal{F}_t^N))_{t\geq 0}$, i.e.

$$p_j(t) = \mathbb{P}(\Lambda = \lambda_j \,|\, \mathcal{F}_t^N), \quad t \geq 0.$$

Moreover, let $p = (p_t)_{t\geq 0}$ denote the m-dimensional process defined by

$$p_t := (p_1(t), \ldots, p_m(t)), \quad t \geq 0.$$

Remark 5.9. Notice that $\widehat{\Lambda}_t \leq \lambda_m$ for $t \geq 0$. Furthermore, it is clear that $p_j(0) = \pi_\Lambda(j)$ for every $j \in \{1, \ldots, m\}$.

The following result provides the dynamic of the *filter process* $(p_t)_{t\geq 0}$.

Theorem 5.10. *For any $j \in \{1, \ldots, m\}$, the process $(p_j(t))_{t\geq 0}$ satisfies*

$$p_j(t) = \pi_\Lambda(j) + \int_0^t \Big(\frac{\lambda_j \, p_j(s-)}{\widehat{\Lambda}_{s-}} - p_j(s-) \Big) \mathrm{d}\widehat{N}_s, \quad t \geq 0. \tag{5.4}$$

Proof. Fix $j \in \{1, \ldots, m\}$ and set $X_j := \mathbb{1}_{\{\Lambda = \lambda_j\}}$. Clearly, X_j is $\mathcal{F}_0$-measurable. Therefore, Assumption 2.92 is fulfilled and we can apply Theorem 2.94 which yields, under consideration of $\mathbb{E}[X_j] = \mathbb{P}(\Lambda = \lambda_j) = \pi_\Lambda(j)$ and $\widehat{X}_j(t) = \mathbb{E}[\mathbb{1}_{\{\Lambda=\lambda_j\}} \,|\, \mathcal{F}_t^N] = p_j(t)$,

$$p_j(t) = \pi_\Lambda(j) + \int_0^t \big(A_s - p_j(s-)\big) \,\mathrm{d}\widehat{N}_s, \quad t \geq 0,$$

where $(A_t)_{t\geq 0}$ is an $\mathfrak{F}^N$-predictable process satisfying

$$\mathbb{E}\Big[\int_0^t H_s \, X_j \, \Lambda \,\mathrm{d}s\Big] = \mathbb{E}\Big[\int_0^t H_s \, A_s \, \widehat{\Lambda}_s \,\mathrm{d}s\Big], \quad t \geq 0,$$

for all non-negative bounded $\mathfrak{F}^N$-predictable processes $(H_t)_{t\geq 0}$. For some fixed non-negative bounded $\mathfrak{F}^N$-predictable processes $(H_t)_{t\geq 0}$ we obtain, by Fubini's Theorem and the $\mathcal{F}_s^N$-measurability of H_s and with the fact that càdlàg processes has only countable many jumps, that

$$\begin{aligned}\mathbb{E}\Big[\int_0^t H_s \frac{\mathbb{E}\big[\Lambda \mathbb{1}_{\{\Lambda=\lambda_j\}} \,|\, \mathcal{F}_s^N\big]}{\widehat{\Lambda}_s} \widehat{\Lambda}_s \,\mathrm{d}s\Big] &= \int_0^t \mathbb{E}\big[\mathbb{E}\big[H_s \, \Lambda \mathbb{1}_{\{\Lambda=\lambda_j\}} \,|\, \mathcal{F}_s^N\big]\big] \,\mathrm{d}s \\ = \int_0^t \mathbb{E}\big[H_s \, \Lambda \mathbb{1}_{\{\Lambda=\lambda_j\}}\big] \,\mathrm{d}t &= \mathbb{E}\Big[\int_0^t H_s \, Z_j \, \lambda \,\mathrm{d}s\Big], \quad t \geq 0.\end{aligned}$$

Furthermore,

$$\mathbb{E}\big[\Lambda \mathbb{1}_{\{\Lambda=\lambda_j\}} \,|\, \mathcal{F}_s^N\big] = \sum_{k=1}^m \lambda_k \mathbb{1}_{\{\lambda_k=\lambda_j\}} p_k(t) = \lambda_j \, p_j(t), \quad t \geq 0.$$

Therefore, we can choose

$$A_t := \frac{\lambda_j \, p_j(t-)}{\widehat{\Lambda}_{t-}}, \quad t \geq 0,$$

where $(A_t)_{t\geq 0}$ is obviously $\mathfrak{F}^N$-predictable. □

The filter process $(p_t)_{t\geq 0}$ carries all available information about the background intensity which is encapsulated in the observable filtration. Let us mention further elementary properties of the filter $(p_t)_{t\geq 0}$.

Proposition 5.11. *Let $j \in \{1,\dots,m\}$. The continuous part $(p_j^c(t))_{t\geq 0}$ of $(p_j(t))_{t\geq 0}$ satisfies*

$$p_j^c(t) = \int_0^t p_j(s)\big(\widehat{\Lambda}_s - \lambda_j\big)\,\mathrm{d}s, \quad t \geq 0,$$

and the new state of the filter p at jump times $(T_n)_{n\in\mathbb{N}}$ is

$$p_{T_n} = J\big(p_{T_n-}\big), \quad n \in \mathbb{N},$$

where

$$J\big(p\big) := \left(\frac{\lambda_1\, p_1}{\sum_{k=1}^m \lambda_k\, p_k}, \dots, \frac{\lambda_m\, p_m}{\sum_{k=1}^m \lambda_k\, p_k}\right) \tag{5.5}$$

for $p = (p_1,\dots,p_m) \in \Delta_m$.

Proof. Fix $j \in \{1,\dots,m\}$. Due to the definition of $\widehat{N}$ given in (5.3) and Theorem 5.10, we have

$$p_j(t) = \pi_\Lambda(j) + \int_0^t \left(\frac{\lambda_j\, p_j(s-)}{\widehat{\Lambda}_{s-}} - p_j(s-)\right)\mathrm{d}N_s - \int_0^t p_j(s)\big(\lambda_j - \widehat{\Lambda}_s\big)\,\mathrm{d}s, \quad t \geq 0.$$

Thus the continuous part $(p_j^c(t))_{t\geq 0}$ of $(p_j(t))_{t\geq 0}$ satisfies

$$p_j^c(t) = \int_0^t p_j(s)\big(\widehat{\Lambda}_s - \lambda_j\big)\,\mathrm{d}s, \quad t \geq 0,$$

and

$$\sum_{0<s\leq t} \Delta p_j(s) = \int_0^t \left(\frac{\lambda_j\, p_j(s-)}{\widehat{\Lambda}_{s-}} - p_j(s-)\right)\mathrm{d}N_s, \quad t \geq 0.$$

That is

$$\Delta p_j(T_n) = \frac{\lambda_j\, p_j(T_n-)}{\widehat{\Lambda}_{T_n-}} - p_j(T_n-), \quad n \in \mathbb{N}.$$

Therefore, the new state of the filter p at the jump times $(T_n)_{n\in\mathbb{N}}$ is

$$p_j(T_n) = p_j(T_n-) + \Delta p_j(T_n) = \frac{\lambda_j\, p_j(T_n-)}{\widehat{\Lambda}_{T_n-}}, \quad n \in \mathbb{N}.$$

Consequently, $p_{T_n} = J(p_{T_n-})$ for every $n \in \mathbb{N}$, where J is defined by (5.5). □

Next we investigate the dynamic of the filter process $(p_t)_{t\geq 0}$ between the jump times.

Proposition 5.12. *Let $n \in \mathbb{N}_0$. Assume $p_{T_n} = p$, then the evolution of $(p_t)_{t\geq 0}$ up to the next jump time T_{n+1} is the solution, denoted by $\phi(t) = (\phi_j(t))_{j=1,\dots,m}$, of the following* *system of ordinary differential equations*

$$\begin{cases} \dot{\phi}_1 = \phi_1\big(\sum_{k=1}^m \lambda_k\,\phi_k - \lambda_1\big),\\ \quad\vdots \\ \dot{\phi}_m = \phi_m\big(\sum_{k=1}^m \lambda_k\,\phi_k - \lambda_m\big),\\ \phi(0) = p \in \mathring{\Delta}_m. \end{cases} \tag{5.6}$$

Proof. Fix $n \in \mathbb{N}_0$. The dynamic of the continuous part $(p_j^c(t))_{t\geq 0}$ of $(p_j(t))_{t\geq 0}$ calculated in Proposition 5.11 yields the dynamic of $(p_j(t))_{t\geq 0}$ for any $t \in [T_n, T_{n+1}]$,

$$p_j(t) = p_j(T_{n-1}) + \int_{T_{n-1}}^t p_j(s)\left(\sum_{k=1}^m \lambda_k\,p_k(s) - \lambda_j\right)\mathrm{d}s.$$

For any $j \in \{1,\dots,m\}$ and $s \geq 0$, we have

$$\left|p_j(s)\left(\sum_{k=1}^m \lambda_k\,p_k(s) - \lambda_j\right)\right| \leq \sum_{k=1}^m \lambda_k + \lambda_j < (m+1)\,\lambda_m,$$

since $p_j(s) \leq 1$ and the order $\lambda_1 < \lambda_2 < \ldots < \lambda_m$, compare Assumption 5.7. That is, the integrand above is bounded on $[T_n, T_{n+1}]$. Therefore, by the càdlàg property, the integrand is Riemann integrable according to the Lebesgue's criterion for Riemann integrability (cf. e.g. Sohrab [115, Prop. 11.1.3]). Thus the Lebesgue integral above coincides with the Riemann integral with the same integrand (cf. e.g. Klenke [77, Thm. 4.23]). In consequence, the second fundamental theorem of calculus (cf. e.g. Sohrab [115, Thm. 7.5.8]) implies the announced system of ordinary differential equations (5.6) since the integrand above is continuous on $[T_n, T_{n+1})$. □

Remark 5.13. Due to the previous proposition, the filter process $(p_t)_{t\geq 0}$ is a piecewise deterministic Markov process.

Here are some elementary properties of the evolution of the filter process between the jumps.

Proposition 5.14. *(i) Let $n \in \mathbb{N}_0$. For $t \in [T_n, T_{n+1})$ it holds $p_t = \phi(t - T_n)$.*

(ii) For any $p \in \Delta_m$, the map $t \mapsto \phi(t)$ with $\phi(0) = p$ is Lipschitz of a rank independent of p.

Proof. (i) This statement is an immediate consequence of the representation of the continuous part of $(p_t)_{t\geq 0}$ given in Proposition 5.11.

(ii) Let $\phi(0) = p$ for some $p \in \Delta_m$. From the first statement, we know that $\dot{\phi}_j(t)$, $j = 1,\dots,m$, is independent of p. Furthermore, from (5.6) follows that $t \mapsto \dot{\phi}_j(t)$, $j = 1,\dots,m$, is continuous and thus bounded on the compact set $[0,T]$. Hence

$$\|(\dot{\phi}_1(t),\dots,\dot{\phi}_m(t))\|_2 \leq K,$$

where $0 < K < \infty$ is independent of p and $\|\cdot\|_2$ denotes the Euclidean norm. Consequently, on account of the mean value theorem for vector-valued functions (see Akcoglu et al. [2, Thm. 5.1.13]), we get

$$\|\phi(t_1) - \phi(t_2)\|_2 \leq K\,|t_1 - t_2|$$

for all $t_1, t_2 \in [0, T]$. □

The first statement of the previous proposition expresses the equivalence of the evolution of $(p_t)_{t\geq 0}$ after every jump at $(T_n)_{n\in\mathbb{N}}$. This property as well as the jumps can be seen in Figure 5.1 which displays a path of each component of the filter $(p_t)_{t\geq 0}$ under the assumption that $A = \{2, 4, 5\}$, $\bar{\pi}_\Lambda = (2/5, 2/5, 1/5)$ and $\mathbb{P}(\Lambda = 4 \,|\, \mathcal{F}_0) = 1$. The figure further shows that no trigger event occurred approximately in the time between 2 and 3, which increases the probability for the smallest possible parameter 2. This is expressed by the strongly increasing black line in this period. Subsequently more events occur (especially in the time between 5 and 6), whereby the probability that the true intensity is 2 decreases strongly and the filter rank the probability for the correct parameter (here 4) up in the course of time which is represented by the red line.

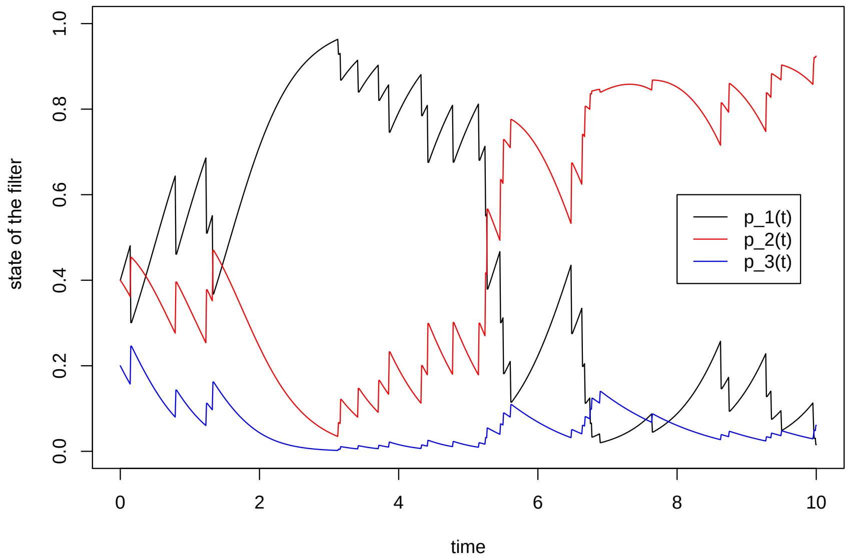

Figure 5.1: A trajectory of the filter process $(p_t)_{t\geq 0}$ under the assumptions that $A = \{2, 4, 5\}$, $\bar{\pi}_\Lambda = (2/5, 2/5, 1/5)$ and $\mathbb{P}(\Lambda = 4 \,|\, \mathcal{F}_0) = 1$, where $p_t = (p_1(t), p_2(t), p_3(t))$ with $p_1(t) = \mathbb{P}(\Lambda = 2 \,|\, \mathcal{G}_t)$, $p_2(t) = \mathbb{P}(\Lambda = 4 \,|\, \mathcal{G}_t)$ and $p_3(t) = \mathbb{P}(\Lambda = 5 \,|\, \mathcal{G}_t)$.

Dynamics for the parameters of the posterior distribution of the thinning probabilities. It has already been mentioned at the beginning of Section 5.1 that we do not have to a solve filter problem for the thinning probabilities since the posterior distribution of $\bar{\alpha}$ given the available information is known, namely through the process $(q_t)_{t\geq 0}$. Therefore, the property of conjugation plays a key role for the reduction in this chapter. Nevertheless we need to compensate the process $(q_t)_{t\geq 0}$ for the stochastic control approach. For this, the next notation and results are required. Before stating these, let us recall that the MPP $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$ carries the information about the claim arrival times, i.e. $\mathfrak{F}^\Phi$ is observable for the insurer, see Section 3.1.

Notation. For any $D \subset \mathbb{D}$, we denote the càdlàg modification of $(\mathbb{E}[\alpha_D \,|\, \mathcal{F}_t^\Phi])_{t\geq 0}$ by $(\widehat{\alpha}_D(t))_{t\geq 0}$, i.e.

$$\widehat{\alpha}_D(t) = \mathbb{E}[\alpha_D \,|\, \mathcal{F}_t^\Phi], \quad t \geq 0.$$

Lemma 5.15. *The* $\mathfrak{F}$*-intensity kernel of* $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$*, denoted by* $(\lambda(t, \mathrm{d}z))_{t\geq 0}$*, is given by*

$$\lambda(t, B) = \Lambda \sum_{D\in B} \alpha_D, \quad t \geq 0, \quad B \in \mathcal{P}(\mathcal{P}(\mathbb{D})).$$

Proof. It is easily seen that λ is a transition kernel from $(\mathbb{R}^+ \times \Omega, \mathcal{B}^+ \otimes \mathcal{F})$ to $(\mathcal{P}(\mathbb{D}), \mathcal{P}(\mathcal{P}(\mathbb{D})))$, compare the arguments in the proof of Lemma 4.9. Moreover, we can show that $(\Lambda \sum_{D\in B} \alpha_D)_{t\geq 0}$ is the predictable $\mathfrak{F}$-intensity of $(\Phi(t, B))_{t\geq 0}$ for some $B \in \mathcal{P}(\mathcal{P}(\mathbb{D}))$ with the same method as in the proof of Lemma 4.7, which completes the proof. □

Proposition 5.16. *The* $\mathfrak{F}^\Phi$*-intensity kernel of* $\Phi = (T_n, Z_n)_{n\in\mathbb{N}}$*, denoted by* $(\widehat{\mu}(t, \mathrm{d}z))_{t\geq 0}$*, is given by*

$$\mu(t, B) = \widehat{\Lambda}_{t-} \sum_{D\in B} \frac{\beta_D + q_D(t-)}{\|\bar{\beta} + q_{t-}\|}, \quad t \geq 0, \quad B \in \mathcal{P}(\mathcal{P}(\mathbb{D})).$$

Proof. The definition of μ clearly forces that μ is a transition kernel from $(\mathbb{R}^+\times \Omega, \mathcal{B}^+\otimes \mathcal{F}^\Phi_\infty)$ to $(\mathcal{P}(\mathbb{D}), \mathcal{P}(\mathcal{P}(\mathbb{D})))$. Now, the procedure is to show that $(\mu(t, B))_{t\geq 0}$ is the $\mathfrak{F}^\Phi$-predictable intensity of $(\Phi(t, B))_{t\geq 0}$ for some fixed $B \in \mathcal{P}(\mathcal{P}(\mathbb{D}))$. To do this, let us remind that $(\Lambda \sum_{D\in B} \alpha_D)_{t\geq 0}$ is an $\mathfrak{F}$-intensity of $(\Phi(t, B))_{t\geq 0}$, compare Lemma 5.15. Since $(\widehat{\Lambda}_t \sum_{D\in B} \widehat{\alpha}_D(t))_{t\geq 0}$ is a càdlàg modification of $(\mathbb{E}[\alpha_D \,|\, \mathcal{F}^\Phi_t])_{t\geq 0}$ (see notation above) and thus $\mathfrak{F}^\Phi$-progressively measurable, it follows, by Proposition 2.81, that $(\widehat{\Lambda}_t \sum_{D\in B} \widehat{\alpha}_D(t))_{t\geq 0}$ is an $\mathfrak{F}^\Phi$-intensity of $(\Phi(t, B))_{t\geq 0}$. Notice that

$$\widehat{\alpha}_D(t) = \mathbb{E}[\alpha_D \,|\, \mathcal{F}^\Phi_t] = \mathbb{E}[\alpha_D \,|\, Z_1, \ldots, Z_{N_t}] = \int_{(0,1)} \alpha_D\, f_{\bar{\beta}}(\alpha_D \,|\, Z_1, \ldots, Z_{N_t})\, \mathrm{d}\alpha_D,$$

where, by Corollary 5.5, the posterior density $f_{\bar{\beta}}$ of α_D given $(Z_1, \ldots, Z_{N_t})$ is the density of the Beta distribution with parameters $\beta_D + q_D(t)$ and $\sum_{E\subset\mathbb{D}\setminus\{D\}}(\beta_E + q_E(t))$. Since the mean of $X \sim \mathrm{Beta}(a, b)$ is $\mathbb{E}[X] = \frac{a}{a+b}$ (cf. e.g. DeGroot [49, p. 50]), we obtain

$$\widehat{\alpha}_D(t) = \frac{\beta_D + q_D(t)}{\|\bar{\beta} + q_t\|},$$

and, in consequence, we have that $\big(\widehat{\Lambda}_{t-} \sum_{D\in B} \frac{\beta_D+q_D(t-)}{\|\bar{\beta}+q_{t-}\|}\big)_{t\geq 0}$ is the $\mathfrak{F}^\Phi$-predictable $\mathfrak{F}^\Phi$-intensity kernel. □

The proof above gives the Bayesian estimator of α_D given $(Z_1, \ldots, Z_{N_t})$, namely

$$\mathbb{E}[\alpha_D \,|\, Z_1, \ldots, Z_{N_t}] = \frac{\beta_D + q_D(t)}{\|\bar{\beta} + q_t\|}, \quad t \geq 0. \tag{5.7}$$

Notation. Let $\widehat{\Phi}(\mathrm{d}t, \mathrm{d}z)$ denote the compensated random measure given by

$$\widehat{\Phi}(\mathrm{d}t, \mathrm{d}z) := \Phi(\mathrm{d}t, \mathrm{d}z) - \mu(t, \mathrm{d}z)\, \mathrm{d}t, \tag{5.8}$$

where μ is given as in Proposition 5.16.

Taking this notation into account, we get for any $D \subset \mathbb{D}$

$$\begin{aligned} q_D(t) &= \int_0^t \int_{\mathcal{P}(\mathbb{D})} \mathbb{1}_{\{z=D\}} \,\widehat{\Phi}(\mathrm{d}s, \mathrm{d}z) + \sum_{E \subset \mathbb{D}} \int_0^t \mathbb{1}_{\{E=D\}} \widehat{\Lambda}_s \frac{\beta_E + q_E(s)}{\|\bar{\beta} + q_s\|} \,\mathrm{d}s \\ &= \int_0^t \int_{\mathcal{P}(\mathbb{D})} \mathbb{1}_{\{z=D\}} \,\widehat{\Phi}(\mathrm{d}s, \mathrm{d}z) + \int_0^t \widehat{\Lambda}_s \frac{\beta_D + q_D(s)}{\|\bar{\beta} + q_s\|} \,\mathrm{d}s, \quad t \geq 0. \end{aligned}$$

Properties of the aggregated claim amount process and the surplus process. The solution method requires a representation of the aggregated claim amount process and the surplus process w.r.t. the compensated random measure of Ψ which is derived next.

Proposition 5.17. *The $\mathfrak{F}^{\Psi}$-intensity kernel of $\Psi = (T_n, (Y_n, Z_n))_{n \in \mathbb{N}}$, denoted by $(\nu(t, \mathrm{d}(y,z)))_{t \geq 0}$, is given by*

$$\nu(t, (A, B)) = \widehat{\Lambda}_{t-} \, F(A) \sum_{D \in B} \frac{\beta_D + q_D(t)}{\|\bar{\beta} + q_t\|}, \quad t \geq 0,\ A \in \mathcal{B}((0,\infty)^d),\ B \in \mathcal{P}(\mathcal{P}(\mathbb{D})).$$

Proof. The assertion follows by using Proposition 5.16 and the same line of arguments as in the proof of Proposition 4.20. □

Notation. Let $\widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z))$ denote the compensated random measure given by

$$\widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z)) := \Psi(\mathrm{d}t, \mathrm{d}(y,z)) - \nu(t, \mathrm{d}(y,z)) \,\mathrm{d}t, \tag{5.9}$$

where ν is defined as in Proposition 5.17.

With assistance of the measure introduced above, we obtain the following characteristics of the aggregated claim amount process.

Proposition 5.18. *The aggregated claim amount process $S = (S_t)_{t \geq 0}$ is given by*

$$S_t = \int_0^t \int_{E^d} \sum_{i=1}^d y_i \mathbb{1}_z(i) \,\widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z)) + \sum_{D \subset \mathbb{D}} \int_0^t \widehat{\Lambda}_s \frac{\beta_D + q_D(s)}{\|\bar{\beta} + q_s\|} \,\mathrm{d}s \sum_{i=1}^d \mathbb{1}_D(i) \mathbb{E}[Y_1^i]$$

and satisfies

$$\mathbb{E}[S_t] = \sum_{k=1}^m \lambda_k \, \pi_\Lambda(k) \sum_{D \subset \mathbb{D}} \frac{\beta_D}{\|\bar{\beta}\|} \sum_{i=1}^d \mathbb{1}_D(i) \mathbb{E}\big[Y_1^i\big] \, t$$

for all $t \geq 0$.

Proof. Combining (3.3) and (4.10), we get for any $t \geq 0$

$$\begin{aligned} S_t &= \int_0^t \int_{E^d} \sum_{i=1}^d y_i \mathbb{1}_z(i) \,\widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z)) \\ &\quad + \sum_{D \subset \mathbb{D}} \int_0^t \int_{(0,\infty)^d} \sum_{i=1}^d y_i \mathbb{1}_D(i) \widehat{\Lambda}_s \frac{\beta_D + q_D(s)}{\|\bar{\beta} + q_s\|} \, F(\mathrm{d}y) \,\mathrm{d}s \\ &= \int_0^t \int_{E^d} \sum_{i=1}^d y_i \mathbb{1}_z(i) \,\widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z)) + \sum_{D \subset \mathbb{D}} \int_0^t \widehat{\Lambda}_s \frac{\beta_D + q_D(s)}{\|\bar{\beta} + q_s\|} \,\mathrm{d}s \sum_{i=1}^d \mathbb{1}_D(i) \mathbb{E}[Y_1^i]. \end{aligned}$$

By Corollary 2.98, the process $(\eta_t)_{t\geq 0}$ defined by

$$\eta_t := \int_0^t \int_{E^d} \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z)), \quad t \geq 0,$$

is an $\mathfrak{F}^\Psi$-martingale if

$$\begin{aligned}
&\mathbb{E}\bigg[\int_0^t \int_{E^d} \bigg|\sum_{i=1}^d y_i \mathbb{1}_z(i)\bigg| \nu(s, \mathrm{d}(y,z))\,\mathrm{d}s\bigg] \\
&= \mathbb{E}\bigg[\int_0^t \widehat{\Lambda}_s \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar{\beta} + q_s\|} \sum_{i=1}^d \mathbb{1}_D(i)\mathbb{E}\big[Y_1^i\big]\,\mathrm{d}s\bigg] < \infty, \quad t \geq 0.
\end{aligned}$$

Notice that the function $H(t,y,z) := \sum_{i=1}^d y_i \mathbb{1}_z(i)$ is obviously an $\mathfrak{F}^\Psi$-predictable function indexed by E^d. Moreover, recall that Λ and $\bar{\alpha}$ are independent (compare Assumption 5.7). Therefore, by Fubini's Theorem and the fact that for any $s \geq 0$ and $D \subset \mathbb{D}$

$$\begin{aligned}
\mathbb{E}\bigg[\widehat{\Lambda}_s \frac{\beta_D + q_D(s)}{\|\bar{\beta} + q_s\|}\bigg] &= \mathbb{E}\big[\mathbb{E}[\Lambda \,|\, \mathcal{F}_s^{\bar{N}}]\mathbb{E}[\alpha_D \,|\, \mathcal{F}_s^{\Phi}]\big] = \mathbb{E}\big[\mathbb{E}[\Lambda\, \alpha_D \,|\, \mathcal{F}_s^{\Psi}]\big] \\
&= \mathbb{E}[\Lambda\, \alpha_D] = \mathbb{E}[\Lambda]\,\mathbb{E}[\alpha_D] = \sum_{k=1}^m \lambda_k\, \pi_\Lambda(k) \frac{\beta_D}{\|\bar{\beta}\|}, \quad s \geq 0,\ D \subset \mathbb{D},
\end{aligned}$$

since $\alpha_D \sim \text{Beta}\left(\beta_D, \sum_{E\subset\mathbb{D}\setminus\{D\}} \beta_E\right)$, compare Section 5.1. Therefore,

$$\begin{aligned}
&\mathbb{E}\bigg[\int_0^t \widehat{\Lambda}_s \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar{\beta} + q_s\|} \sum_{i=1}^d \mathbb{1}_D(i)\mathbb{E}\big[Y_1^i\big]\,\mathrm{d}s\bigg] \\
&= \sum_{D\subset\mathbb{D}} \frac{\beta_D}{\|\bar{\beta}\|} \sum_{i=1}^d \mathbb{1}_D(i)\mathbb{E}\big[Y_1^i\big] \sum_{k=1}^m \lambda_k\, \pi_\Lambda(k)\, t < \infty, \quad t \geq 0,
\end{aligned}$$

and the proposition follows. □

The proposition yields the following indistinguishable representation of the surplus process $X^{\xi,b} = (X_t^{\xi,b})_{t\geq 0}$:

$$\begin{aligned}
\mathrm{d}X_t^{\xi,b} &= \bigg(rX_s^{\xi,b} + (\mu - r)\xi_s + c(b_s) - \widehat{\Lambda}_t\, b_t \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(t)}{\|\bar{\beta} + q_t\|} \sum_{i=1}^d \mathbb{1}_D(i)\, \mathbb{E}[Y_1^i]\bigg)\mathrm{d}t \\
&\quad + \xi_s \sigma \mathrm{d}W_s - \int_{E^d} b_t \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z)), \quad t \geq 0.
\end{aligned}$$

This dynamic will be one part of the reduced control model discussed in the next section.

5.3 The reduced control problem

Recall that the processes $(p_t)_{t\geq 0}$ and $(q_t)_{t\geq 0}$ carry all relevant information about the unknown parameters λ and $\bar{\alpha}$ contained in the observable filtration $\mathfrak{G}$ of the insurer. That is, it is sufficient to know the processes $(p_t)_{t\geq 0}$ and $(q_t)_{t\geq 0}$ instead of the whole history $\mathfrak{G}$. As in Section 4.3, we consider a family of problems by varying the initial time to achieve a relationship among the corresponding value functions. Therefore, the state process of the reduced control problem with complete observation is the $(\ell+m+1)$-dimensional process

$$(X_s^{\xi,b}, p_s, q_s)_{s\in[t,T]}$$

for some fixed initial time $t\in[0,T)$ and $(\xi,b)\in\mathcal{U}[t,T]$, where

$$\begin{aligned}
\mathrm{d}X_s^{\xi,b} &= \left(rX_s^{\xi,b} + (\mu-r)\xi_s + c(b_s) - \widehat{\Lambda}_s\, b_s \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar\beta + q_s\|} \sum_{i=1}^d \mathbb{1}_D(i)\mathbb{E}[Y_1^i]\right)\mathrm{d}s \\
&\qquad + \xi_s\sigma\,\mathrm{d}W_s - \int_{E^d} b_s \sum_{i=1}^d y_i \mathbb{1}_z(i)\,\widehat{\Psi}(\mathrm{d}s,\mathrm{d}(y,z)), && (5.10)\\
\mathrm{d}p_j(s) &= \left(\frac{\lambda_j\, p_j(s-)}{\widehat{\Lambda}_{s-}} - p_j(s-)\right)\mathrm{d}\widehat{N}_s, \quad j=1,\dots,d, && (5.11)\\
\mathrm{d}q_D(s) &= \int_{\mathcal{P}(\mathbb{D})} \mathbb{1}_{\{z=D\}}\,\widehat{\Phi}(\mathrm{d}s,\mathrm{d}z) + \widehat{\Lambda}_s \frac{\beta_D + q_D(s)}{\|\bar\beta + q_s\|}\,\mathrm{d}s, \quad D\subset\mathbb{D}. && (5.12)
\end{aligned}$$

for $s\in[t,T]$, with

$$\big(X_t^{\xi,b}, p_t, q_t\big) = (x,p,q)$$

with $x\in\mathbb{R}$, $p=(p_1,\dots,p_m)\in\Delta_m$ and $q=(q_D)_{D\subset\mathbb{D}}\in\mathbb{N}_0^\ell$. Using this reduced model, we can formulate the reduced control problem. For any $(\xi,b)\in\mathcal{U}[t,T]$, the *objective function* is given by

$$V^{\xi,b}(t,x,p,q) := \mathbb{E}^{t,x,p,q}\big[U(X_T^{\xi,b})\big] := \mathbb{E}\big[U(X_T^{\xi,b}) \,|\, X_t^{\xi,b}=x, p_t=p, q_t=q\big]$$

and the *value function* is defined by

$$V(t,x,p,q) := \sup_{(\xi,b)\in\mathcal{U}[t,T]} V^{\xi,b}(t,x,p,q), \tag{P2}$$

for all $(t,x,p,q)\in[0,T]\times\mathbb{R}\times\Delta_m\times\mathbb{N}_0^\ell$. As before, an investment-reinsurance strategy $(\xi^\star,b^\star)\in\mathcal{U}[t,T]$ is optimal if

$$V(t,x,p,q) = V^{\xi^\star,b^\star}(t,x,p,q),$$

and the insurer is interested in optimal strategies $(\xi^\star,b^\star)\in\mathcal{U}[t,T]$, i.e. in strategies

$$(\xi^\star,b^\star) = \operatorname*{argsup}_{(\xi,b)\in\mathcal{U}[t,T]} V^{\xi,b}(t,x,p,q).$$

Applying the arguments from Section 4.3, we get for any $(\xi,b)\in\mathcal{U}[t,T]$

$$V^{\xi,b}(t,x,p_t,q_t) = \widetilde{V}^{\xi,b}(t,x) \quad\text{and thus}\quad V(t,x,p_t,q_t) = \widetilde{V}(t,x), \quad (t,x)\in[0,T]\times\mathbb{R}.$$

That is, if we solve the reduced control problem (P2), the original problem (P) is also solved (under setting of Section 5.1).

The following properties of the value function are analogous to those in Section 4.3, compare Lemma 4.22.

Lemma 5.19. *(i) For any $(\xi,b)\in\mathcal{U}[t,T]$ and $(t,x,p,q)\in[0,T]\times\mathbb{R}\times\Delta_m\times\mathbb{N}_0^\ell$, it* *holds*

$$V^{\xi,b}(t,x,p,q) = \sum_{j=1}^{m} p_j \, V^{\xi,b}(t,x,e_j,q).$$

(ii) For any $(t,x,q) \in [0,T] \times \mathbb{R} \times \mathbb{N}_0^\ell$*, the function* $\Delta_m \ni p \mapsto V(t,x,p,q)$ *is convex.*

Proof. The assertions follow from the same arguments as in the proof of Lemma 4.22. □

5.4 The Hamilton-Jacobi-Bellman equation

In this section we proceed as in Section 4.4 by developing heuristically the generalized HJB equation, of which a byproduct is a candidate for an optimal investment-reinsurance strategy. Thus the starting point is the assumption that the DPP holds, i.e. the value function V satisfies

$$V(t,x,p,q) = \sup_{(\xi,b)\in\mathcal{U}[t,t_0]} \mathbb{E}^{t,x,p,q}\Big[V\big(t_0, X_{t_0}^{\xi,b}, p_{t_0}, q_{t_0}\big)\Big]$$

for all $(t,x,p,q) \in [0,T] \times \mathbb{R} \times \Delta_m \times \mathbb{N}_0^\ell$ and for some $t_0 \in [t,T]$. Recall that $(q_t)_{t\geq 0}$ is a pure jump process and that $(p_t)_{t\geq 0}$ is an FV process. Therefore, assuming that V is sufficient smooth, we obtain by Itô-Doeblin's formula

$$\begin{aligned}
&V\big(t_0, X_{t_0}^{\xi,b}, p_{t_0}, q_{t_0}\big) = V\big(t, X_t^{\xi,b}, p_t, q_t\big) \\
&+ \int_t^{t_0} V_t\big(s, X_s^{\xi,b}, p_s, q_s\big)\,\mathrm{d}s + \int_t^{t_0} V_x\big(s, X_{s-}^{\xi,b}, p_{s-}, q_{s-}\big)\,\mathrm{d}(X^{\xi,b})_s^c \\
&+ \sum_{j=1}^{m} \int_t^{t_0} V_{p_j}\big(s, X_{s-}^{\xi,b}, p_{s-}, q_{s-}\big)\,\mathrm{d}p_j^c(s) + \frac{1}{2}\int_t^{t_0} V_{xx}\big(s, X_{s-}^{\xi,b}, p_{s-}, q_{s-}\big)\,\mathrm{d}[X^{\xi,b}]_s^c \\
&+ \sum_{0<s\leq t} \Big(V\big(s, X_s^{\xi,b}, p_s, q_s\big) - V\big(s, X_{s-}^{\xi,b}, p_{s-}, q_{s-}\big)\Big).
\end{aligned}$$

According to Proposition 3.15 and Proposition 5.11, we get

$$\begin{aligned}
&V\big(t_0, X_{t_0}^{\xi,b}, p_{t_0}, q_{t_0}\big) = V\big(t, X_t^{\xi,b}, p_t, q_t\big) \\
&+ \int_t^{t_0} \bigg(V_t\big(s, X_s^{\xi,b}, p_s, q_s\big) + V_x\big(s, X_s^{\xi,b}, p_s, q_s\big)\big(rX_s^{\xi,b} + (\mu - r)\xi_s + c(b_s)\big) \\
&+ \sum_{j=1}^{m} V_{p_j}\big(s, X_s^{\xi,b}, p_s, q_s\big)p_j(s)\big(\widehat{\Lambda}_s - \lambda_j\big) + \frac{1}{2}V_{xx}\big(s, X_s^{\xi,b}, p_s, q_s\big)\sigma^2\xi_s^2\bigg)\mathrm{d}s \\
&+ \int_t^{t_0} V_x\big(s, X_{s-}^{\xi,b}, p_{s-}, q_{s-}\big)\sigma\xi_s\,\mathrm{d}W_s + \sum_{0<s\leq t} \Big(V\big(s, X_s^{\xi,b}, p_s, q_s\big) - V\big(s, X_{s-}^{\xi,b}, p_{s-}, q_{s-}\big)\Big).
\end{aligned}$$

To calculate the sum in the equation above, recall the notation made on page 108 of the function v given in (5.1). Once again with the help of Proposition 3.15 and Proposition 5.11 as well as the definition of $\widehat{\Psi}$ given in (5.9), it follows

$$\begin{aligned}
&\sum_{0<s\leq t} \Big(V\big(s, X_s^{\xi,b}, p_s, q_s\big) - V\big(s, X_{s-}^{\xi,b}, p_{s-}, q_{s-}\big)\Big) \\
&= \int_0^t \int_{E^d} \bigg(V\Big(s, X_{s-}^{\xi,b} - b_s \sum_{i=1}^{d} y_i \mathbb{1}_z(i), J(p_{s-}), v(q_{s-}, z)\Big)
\end{aligned}$$

$$
\begin{aligned}
&\qquad - V\big(s, X^{\xi,b}_{s-}, p_{s-}, q_{s-}\big)\Big)\widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z))\\
&+ \int_0^t \widehat{\Lambda}_s \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar\beta + q_s\|} \int_{(0,\infty)^d} V\Big(s, X^{\xi,b}_s - b_s \sum_{i=1}^d y_i \mathbb{1}_D(i), J(p_s), v(q_s, D)\Big) F(\mathrm{d}y)\,\mathrm{d}s\\
&- \int_0^t \widehat{\Lambda}_s\, V\big(s, X^{\xi,b}_s, p_s, q_s\big)\,\mathrm{d}s.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&V\big(t_0, X^{\xi,b}_{t_0}, p_{t_0}, q_{t_0}\big) = V\big(t, X^{\xi,b}_t, p_t, q_t\big)\\
&+ \int_t^{t_0} \bigg(V_t\big(s, X^{\xi,b}_s, p_s, q_s\big) + V_x\big(s, X^{\xi,b}_s, p_s, q_s\big)\big(rX^{\xi,b}_s + (\mu - r)\xi_s + c(b_s)\big)\\
&+ \sum_{j=1}^m V_{p_j}\big(s, X^{\xi,b}_s, p_s, q_s\big) p_j(s) \bigg(\sum_{k=1}^m \lambda_k\, p_k(s) - \lambda_j \bigg) + \frac{1}{2} V_{xx}\big(s, X^{\xi,b}_s, p_s, q_s\big)\sigma^2 \xi_s^2\\
&+ \sum_{k=1}^m \lambda_k p_k(s) \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar\beta + q_s\|} \int_{(0,\infty)^d} V\Big(s, X^{\xi,b}_s - b_s \sum_{i=1}^d y_i \mathbb{1}_D(i), J(p_s), v(q_s, D)\Big) F(\mathrm{d}y)\\
&- \sum_{k=1}^m \lambda_k\, p_k(s) V\big(s, X^{\xi,b}_s, p_s, q_s\big)\bigg)\mathrm{d}s + \int_t^{t_0} V_x\big(s, X^{\xi,b}_{s-}, p_{s-}, q_{s-}\big)\sigma\, \xi_s\, \mathrm{d}W_s\\
&+ \int_0^t \int_{E^d} \bigg(V\Big(s, X^{\xi,b}_{s-} - b_s \sum_{i=1}^d y_i \mathbb{1}_z(i), J(p_{s-}), v(q_{s-}, z)\Big)\\
&\qquad - V\big(s, X^{\xi,b}_{s-}, p_{s-}, q_{s-}\big)\bigg)\widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z)).
\end{aligned}
$$

Using the same arguments as in Section 4.4, we obtain

$$
\begin{aligned}
0 = \sup_{(\xi,b)\in\mathbb{R}\times[0,1]} \bigg\{ &V_t(t,x,p,q) - \sum_{k=1}^m \lambda_k\, p_k V(t,x,p,q) + \frac{1}{2}\sigma^2 V_{xx}(t,x,p,q)\xi^2\\
&+ V_x(t,x,p,q)\big(rx + (\mu - r)\xi + c(b)\big) + \sum_{j=1}^m V_{p_j}(t,x,p,q) p_j \bigg(\sum_{k=1}^m \lambda_k\, p_k - \lambda_j \bigg)\\
&+ \sum_{k=1}^m \lambda_k\, p_k \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar\beta + q\|} \int\limits_{(0,\infty)^d} V\Big(t, x - b\sum_{i=1}^d y_i \mathbb{1}_D(i), J(p), v(q,D)\Big) F(\mathrm{d}y) \bigg\}.
\end{aligned}
\tag{5.13}
$$

Again, we have the following separation approach which follows by similar argumentations as in the proof of Lemma 4.24: For any $(t,x,p,q) \in [0,T] \times \mathbb{R} \times \Delta_m \times \mathbb{N}_0^\ell$, we have

$$
V(t,x,p,q) = -e^{-\alpha x e^{r(T-t)}} g(t,p,q) \tag{5.14}
$$

with

$$
g(t,p,q) := \inf_{(\xi,b)\in\mathcal{U}[t,T]} g^{\xi,b}(t,p,q), \tag{5.15}
$$

where

$$g^{\xi,b}(t,p,q) := \mathbb{E}^{t,p,q}\Big[\exp\Big\{-\int_t^T \alpha\, e^{r(T-s)}\big((\mu-r)\,\xi_s + c(b_s)\big)\,\mathrm{d}s \\ - \int_t^T \alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s + \int_t^T\int_{E^d} \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\,\Psi(\mathrm{d}s,\mathrm{d}(y,z))\Big\}\Big], \tag{5.16}$$

where $\mathbb{E}^{t,p,q}$ denotes the conditional expectation given $p_t = p$, $q_t = q$. Before we use the separation approach to rearrange Equation (5.13), let us mention some useful properties of the introduced function g.

Lemma 5.20. *The function g defined by* (5.15) *has the following properties:*

(i) g *is bounded on* $[0,T]\times\Delta_m\times\mathbb{N}_0^\ell$.

(ii) $g^{\xi,b}(t,p,q) > 0$ *for all* $(t,p,q)\in[0,T]\times\Delta_m\times\mathbb{N}_0^\ell$ *and* $(\xi,b)\in\mathcal{U}[t,T]$.

(iii) $g^{\xi,b}(t,p,q) = \sum_{j=1}^m p_j\, g^{\xi,b}(t,e_j,q)$ *for all* $(t,p,q)\in[0,T]\times\Delta_m\times\mathbb{N}_0^\ell$ *and* $(\xi,b)\in\mathcal{U}[t,T]$.

(iv) $g^{\xi,b}(t,J(p),q) = \sum_{j=1}^m \frac{\lambda_j\, p_j}{\sum_{k=1}^m \lambda_k\, p_k} g^{\xi,b}(t,e_j,q)$ *for all* $(t,p,q)\in[0,T]\times\Delta_m\times\mathbb{N}_0^\ell$ *and* $(\xi,b)\in\mathcal{U}[t,T]$.

(v) $g^{\xi,b}(t,p,q) = \sum_{D\subset\mathbb{D}} \frac{\beta_D+q_D}{\|\beta+q\|} g^{\xi,b}(t,p,v(q,D))$ *for all* $(t,p,q)\in[0,T]\times\Delta_m\times\mathbb{N}_0^\ell$ *and* $(\xi,b)\in\mathcal{U}[t,T]$.

(vi) $\Delta_m\ni p\mapsto g(t,p,q)$ *is concave for all* $(t,q)\in[0,T]\times\mathbb{N}_0^\ell$.

Proof. (i) This can be shown by the same method as in the proof of Lemma 4.25 (iii).

(ii) This statement follows immediately from the definition of $g^{\xi,b}$ given in (5.16).

(iii) As in the proof of Lemma 5.19 (ii), the announced assertions follow by conditioning.

(iv) Once again, the statement follows by conditioning.

(v) Fix $(t,p,q)\in[0,T]\times\Delta_m\times\mathbb{N}_0^\ell$ and $(\xi,b)\in\mathcal{U}[t,T]$. Notice that $g^{\xi,b}$ can be written as

$$g^{\xi,b}(t,p,q) := \mathbb{E}^{t,p,q}\Big[\exp\Big\{-\int_t^T \alpha\, e^{r(T-s)}\big((\mu-r)\,\xi_s + c(b_s)\big)\,\mathrm{d}s \\ - \int_t^T \alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s + \alpha\sum_{n=1}^{N_{T-t}} b_{T_n}\, e^{r(T-T_n)} \sum_{i=1}^d Y_n \mathbb{1}_{Z_n}(i)\Big\}\Big],$$

compare (5.16), where p describes the intensity of the of the trigger process $(N_t)_{t\geq 0}$ and q the distribution of the thinning probabilities, both given the relevant information up to time t. We observe that

$$g^{\xi,b}(t,p,q) = \int_{A\times\mathring{\Delta}_\ell} h(\lambda,\tilde\alpha)\,\mathbb{P}^{t,p,q}(\Lambda\in\mathrm{d}\lambda,\bar\alpha\in\mathrm{d}\tilde\alpha)$$

with

$$h(\lambda,\tilde{\alpha}) := \mathbb{E}^{t,p,q}\Big[\exp\Big\{-\int_t^T \alpha\, e^{r(T-s)}\big((\mu-r)\,\xi_s + c(b_s)\big)\,\mathrm{d}s \\ - \int_t^T \alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s + \alpha\sum_{n=1}^{N_{T-t}} b_{T_n}\,e^{r(T-T_n)}\sum_{i=1}^{d} Y_n \mathbb{1}_{Z_n}(i)\Big\}\,\Big|\,\Lambda=\lambda,\bar{\alpha}=\tilde{\alpha}\Big],$$

where, by Assumptions 3.3 and 5.7, Λ and $\bar{\alpha}$ are conditional independent given $p_t = p$ and $q_t = q$, i.e.

$$\mathbb{P}^{t,p,q}(\Lambda\in\mathrm{d}\lambda,\bar{\alpha}\in\mathrm{d}\tilde{\alpha}) = \mathbb{P}^{t,p,q}(\Lambda\in\mathrm{d}\lambda)\,\mathbb{P}^{t,p,q}(\bar{\alpha}\in\mathrm{d}\tilde{\alpha}).$$

Hence

$$g^{\xi,b}(t,p,q) = \sum_{k=1}^{m}\mathbb{P}^{t,p,q}(\Lambda=\lambda_k)\int_{\mathring{\Delta}_\ell} h(\lambda_k,\tilde{\alpha})\,\mathbb{P}^{t,p,q}(\bar{\alpha}\in\mathrm{d}\tilde{\alpha})$$

with $\mathbb{P}^{t,p,q}(\Lambda=\lambda_k) = p_k$ and $\mathbb{P}^{t,p,q}(\bar{\alpha}\in\mathrm{d}\tilde{\alpha}) = f_{\bar{\beta}}(\tilde{\alpha}\,|\,q)\mathrm{d}\tilde{\alpha}$, where $f_{\bar{\beta}}(\tilde{\alpha}\,|\,q)$ denotes the posterior density function of $\tilde{\alpha}$ given $q_t = q$, compare Theorem 5.4. That is,

$$f_{\bar{\beta}}(\tilde{\alpha}\,|\,q) = \frac{\Gamma\big(\sum_{E\subset\mathbb{D}}(\beta_E+q_E)\big)}{\prod_{E\subset\mathbb{D}}\Gamma\big(\beta_E+q_E\big)}\prod_{E\subset\mathbb{D}}\alpha_E^{\beta_E+q_E-1},\quad \tilde{\alpha}=(\alpha_E)_{E\subset\mathbb{D}}\in\mathring{\Delta}_\ell.$$

Consequently, the statement holds if

$$\sum_{D\subset\mathbb{D}}\frac{\beta_D+q_D}{\|\bar{\beta}+q\|}f_{\bar{\beta}}(\tilde{\alpha}\,|\,v(q,D)) = f_{\bar{\beta}}(\tilde{\alpha}\,|\,q).$$

Indeed, using $\Gamma(n+1) = n\,\Gamma(n)$ for all $n\in\mathbb{N}$, we have for any $\tilde{\alpha}=(\alpha_E)_{E\subset\mathbb{D}}\in\mathring{\Delta}_\ell$

$$\begin{aligned}
&\sum_{D\subset\mathbb{D}}\frac{\beta_D+q_D}{\|\bar{\beta}+q\|}f_{\bar{\beta}}(\tilde{\alpha}\,|\,v(q,D))\\
&=\sum_{D\subset\mathbb{D}}\frac{\beta_D+q_D}{\|\bar{\beta}+q\|}\frac{\Gamma\big(\sum_{E\subset\mathbb{D}}(\beta_E+q_E)+1\big)}{\Gamma\big(\beta_D+q_D+1\big)\prod_{E\subset\mathbb{D}\setminus\{D\}}\Gamma\big(\beta_E+q_E\big)}\alpha_D^{\beta_D+q_D}\prod_{E\subset\mathbb{D}\setminus\{D\}}\alpha_E^{\beta_E+q_E-1}\\
&=\sum_{D\subset\mathbb{D}}\frac{\beta_D+q_D}{\sum_{E\subset\mathbb{D}}(\beta_E+q_E)}\frac{\big(\sum_{E\subset\mathbb{D}}(\beta_E+q_E)\big)\Gamma\big(\sum_{E\subset\mathbb{D}}(\beta_E+q_E)\big)}{(\beta_D+q_D)\prod_{E\subset\mathbb{D}}\Gamma\big(\beta_E+q_E\big)}\alpha_D\prod_{E\subset\mathbb{D}}\alpha_E^{\beta_E+q_E-1}\\
&=\frac{\Gamma\big(\sum_{E\subset\mathbb{D}}(\beta_E+q_E)\big)}{\prod_{E\subset\mathbb{D}}\Gamma\big(\beta_E+q_E\big)}\prod_{E\subset\mathbb{D}}\alpha_E^{\beta_E+q_E-1}\sum_{D\subset\mathbb{D}}\alpha_D\\
&=f_{\bar{\beta}}(\tilde{\alpha}\,|\,q),
\end{aligned}$$

since $\sum_{D\subset\mathbb{D}}\alpha_D = 1$.

(vi) We can conclude the assertion by using (5.14) and the convexitiy of V w.r.t. p, compare Lemma 5.19 (ii). □

The separation approach (5.14) implies

$$\begin{aligned}
V_t(t,x,p,q) &= -e^{-\alpha x e^{r(T-t)}}\Big(\alpha\,x\,r\,e^{r(T-t)}g(t,p,q) + g_t(t,p,q)\Big),\\
V_x(t,x,p,q) &= -e^{-\alpha x e^{r(T-t)}}\Big(-\alpha\,e^{r(T-t)}g(t,p,q)\Big),
\end{aligned}$$

$$V_{xx}(t,x,p,q) = -e^{-\alpha x e^{r(T-t)}} \alpha^2 \, e^{2r(T-t)} g(t,p,q),$$
$$V_{p_j}(t,x,p,q) = -e^{-\alpha x e^{r(T-t)}} g_{p_j}(t,p,q), \quad j = 1,\dots,m.$$

The partial derivative w.r.t. t and p_j, $j = 1,\dots,m$, are only defined on the open sets $(0,T)$ and $(0,1)$, respectively. However, we will generalize this partial derivatives later. Using the relations derived above as well as

$$V\Big(t, x - b\sum_{i=1}^{d} y_i \mathbb{1}_D(i), p, q\Big) = -e^{-\alpha x e^{r(T-t)}} \exp\Big\{\alpha \, b \, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_D(i)\Big\} g(t,p,q),$$

we conclude from (5.13)

$$\begin{aligned} 0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \Big\{ & g_t(t,p,q) - \sum_{k=1}^{m} \lambda_k \, p_k \, g(t,p,q) + \sum_{j=1}^{m} g_{p_j}(t,p,q) p_j \Big(\sum_{k=1}^{m} \lambda_k \, p_k - \lambda_j \Big) \\ & - \alpha \, e^{r(T-t)} g(t,p,q) \Big((\mu - r)\xi + c(b) - \frac{1}{2} \alpha \, \sigma^2 \, e^{r(T-t)} \xi^2 \Big) \\ & + \sum_{k=1}^{m} \lambda_k \, p_k \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar{\beta} + q\|} g\big(t, J(p), v(q,D)\big) \int_{(0,\infty)^d} \exp\Big\{\alpha \, b \, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y) \Big\}. \end{aligned} \tag{5.17}$$

Notice that

$$-e^{-\alpha x} g(T,p,q) = V(T,x,p,q) = \mathbb{E}^{T,x,p,q}\big[U(X_T^{\xi,b})\big] = -e^{-\alpha x}$$

for all $(x,p,q) \in \mathbb{R} \times \Delta_m \times \mathbb{N}_0^\ell$, i.e.

$$g(T,p,q) = 1, \quad (p,q) \in \Delta_m \times \mathbb{N}_0^\ell.$$

However, g is probably not differentiable w.r.t. t and p_j, $j = 1,\dots,m$, since every component of the state process jumps. Assuming $(t,p) \mapsto g(t,p,q)$ is Lipschitz on $[0,T] \times \Delta_m$ for all $q \in \mathbb{N}_0^\ell$, we can replace the partial derivatives of g w.r.t. t and p_j, $j = 1,\dots,m$, by Clarke's generalized gradient, compare Section 2.1. For this purpose, we introduce the following notation.

Notation. For fixed $q \in \mathbb{N}_0^\ell$, we write $g_q(t,p) : [0,T] \times \Delta_m \to (0,\infty)$ for the function which is given by

$$g_q(t,p) := g(t,p,q), \quad (t,p) \in [0,T] \times \Delta_m.$$

Furthermore, we denote the components of an $(m+1)$-dimensional vector $\varphi \in \partial^C g_q(t,p)$ by $\varphi = (\varphi_0, \varphi_1, \dots, \varphi_m)$.

We further introduce the following operator.

Notation. Throughout this chapter, let $\mathscr{L}$ denote an operator acting on functions $g : [0,T] \times \Delta_m \times \mathbb{N}_0^\ell \to (0,\infty)$ and $(\xi,b) \in \mathbb{R} \times [0,1]$ which is defined by

$$\begin{aligned} & \mathscr{L} g(t,p,q;\xi,b) \\ & := -\sum_{k=1}^{m} \lambda_k \, p_k \, g(t,p,q) - \alpha \, e^{r(T-t)} g(t,p,q) \Big((\mu - r)\xi + c(b) - \frac{1}{2} \alpha \, \sigma^2 \, e^{r(T-t)} \xi^2 \Big) \\ & + \sum_{k=1}^{m} \lambda_k p_k \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar{\beta} + q\|} g\big(t, J(p), v(q,D)\big) \int_{(0,\infty)^d} \exp\Big\{\alpha b e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y). \end{aligned} \tag{5.18}$$

Using this operator and replacing the partial derivatives of g w.r.t. t and p_j, $j = 1, \ldots, m$, in (5.17) by the generalized gradient, we get the generalized HJB equation for g:

$$0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \{\mathscr{L} g(t,p,q;\xi,b)\} + \inf_{\varphi\in\partial^C g_q(t,p)} \Big\{\varphi_0 + \sum_{j=1}^m \varphi_j\, p_j \Big(\sum_{k=1}^m \lambda_k\, p_k - \lambda_j\Big)\Big\} \tag{5.19}$$

for all $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$ with the boundary condition

$$g(T,p,q) = 1, \quad (p,q) \in \Delta_m \times \mathbb{N}_0^\ell. \tag{5.20}$$

Note that we set $\partial^C g_q(t,p) = \{\nabla g_q(t,p)\}$ at the points (t,p) where the gradient exists. In the next section we continue to determine a candidate for an optimal strategy.

5.5 Candidate for an optimal strategy

To obtain candidates for an optimal strategy, we rewrite the generalized HJB equation (5.19). But first, it should be noted that $g(t,p,q) \geq 0$ for all $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$ which is an immediate consequence of Lemma 5.20 (i). However, in the following we need positiveness of the function g, which is assumed from now on throughout this section. Then we obtain from (5.19)

$$\begin{aligned} 0 = &-\sum_{k=1}^m \lambda_k\, p_k\, g(t,p,q) + \alpha\, e^{r(T-t)} g(t,p,q) \inf_{\xi\in\mathbb{R}} f_1(t,\xi) + \inf_{b\in[0,1]} f_2(t,p,q,b) \\ &+ \inf_{\varphi\in\partial^C g_q(t,p)} \Big\{\varphi_0 + \sum_{j=1}^m \varphi_j\, p_j \Big(\sum_{k=1}^m \lambda_k\, p_k - \lambda_j\Big)\Big\}, \end{aligned} \tag{5.21}$$

where f_1 is defined by (4.27) and

$$\begin{aligned} f_2(t,p,q,b) := -\alpha\, e^{r(T-t)}\, c(b)\, g(t,p,q) + \sum_{k=1}^m \lambda_k\, p_k \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar\beta + q\|} g\big(t, J(p), v(q,D)\big) \times \\ \int_{(0,\infty)^d} \exp\Big\{\alpha\, b\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y). \end{aligned}$$

Hence we can conclude from Section 4.5 that the unique candidate of an optimal investment strategy $\xi^\star = (\xi^\star(t))_{t\in[0,T]}$ is given by

$$\xi^\star(t) = \frac{\mu - r}{\sigma^2} \frac{1}{\alpha} e^{-r(T-t)}, \quad t \in [0,T]. \tag{5.22}$$

We next proceed similar to Section 4.6 to obtain a candidate for an optimal reinsurance strategy. Using the reinsurance premium model given in (3.6), we get

$$\begin{aligned} f_2(t,p,q,b) = &-\alpha\, e^{r(T-t)}\, g(t,p,q)(\eta - \theta)\, \kappa - \alpha\, e^{r(T-t)}\, g(t,p,q)(1+\theta)\, \kappa\, b \\ &+ \sum_{k=1}^m \lambda_k\, p_k \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar\beta + q\|} g\big(t, J(p), v(q,D)\big) \times \\ &\int_{(0,\infty)^d} \exp\Big\{\alpha\, b\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y). \end{aligned} \tag{5.23}$$

The following lemma yields the first order condition for a candidate of an optimal reinsurance strategy.

Lemma 5.21. *For any* $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$, *the function* $\mathbb{R} \ni b \mapsto f_2(t,p,q,b)$ *is strictly convex and*

$$\frac{\partial}{\partial b} f_2(t,p,q,b) = -\alpha\, e^{r(T-t)} \Bigg(g(t,p,q)\,(1+\theta)\,\kappa - \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar\beta + q\|} g\big(t, J(p), v(q,D)\big) \times \\ \sum_{i=1}^d \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\Big\{\alpha\, b\, e^{r(T-t)} \sum_{j=1}^d y_j \mathbb{1}_D(j)\Big\} F(\mathrm{d}y) \sum_{k=1}^m \lambda_k\, p_k \Bigg).$$

Proof. The lemma can be proven with the same arguments as Lemma 4.27. □

The previous lemma provides a criterion for a candidate of an optimal reinsurance strategy as well as the uniqueness of the candidate. The criterion is expressed with help of the following notation.

Notation. For any $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$ and $b \in \mathbb{R}$, we define

$$h_F(t,p,q,b) := \sum_{k=1}^m \lambda_k\, p_k \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar\beta + q\|} \frac{g\big(t, J(p), v(q,D)\big)}{g(t,p,q)} \times \\ \sum_{i=1}^d \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\Big\{\alpha\, b\, e^{r(T-t)} \sum_{j=1}^d y_j \mathbb{1}_D(j)\Big\} F(\mathrm{d}y). \tag{5.24}$$

Furthermore, we set

$$A_F(t,p,q) := h_F(t,p,q,0), \\ B_F(t,p,q) := h_F(t,p,q,1).$$

Before we state the announced first order condition for the optimal reinsurance strategy, let us mention an alternative reinsurance premium model.

Remark 5.22. As in Remark 4.28, we discuss shortly a time-dependent premium calculation principle. Recall that by Proposition 5.18

$$\mathbb{E}[\mathrm{d}S_t] = \sum_{k=1}^m \lambda_k\, \pi_\Lambda(k) \sum_{D\subset\mathbb{D}} \frac{\beta_D}{\|\bar\beta\|} \sum_{i=1}^d \mathbb{1}_D(i) \mathbb{E}\big[Y_1^i\big] = \widehat\Lambda_0 \sum_{D\subset\mathbb{D}} \mathbb{E}[\alpha_D] \sum_{i=1}^d \mathbb{1}_D(i)\, \mathbb{E}\big[Y_1^i\big].$$

It has already been explained in Remark 4.28 that it is reasonable to replace the a priori estimators of unknown parameters in the expression above by the posterior estimators given the available information such that κ depends on the processes p and q, which include all relevant available information:

$$\kappa_F(p,q) = \sum_{k=1}^m \lambda_k\, p_k \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar\beta + q\|} \sum_{i=1}^d \mathbb{1}_D(i)\, \mathbb{E}\big[Y_1^i\big].$$

Then it can be shown by similar arguments as in the proof of Theorem 5.31 that

$$\frac{A_F(t,p,q)}{\kappa_F(p,q)} = \frac{\sum\limits_{D\subset\mathbb{D}} \frac{\beta_D+q_D}{\|\bar\beta+q\|}\frac{g(t,J(p),v(q,D))}{g(t,p,q)}\sum\limits_{i=1}^{d}\mathbb{1}_D(i)\,\mathbb{E}[Y_1^i]}{\sum\limits_{D\subset\mathbb{D}} \frac{\beta_D+q_D}{\|\bar\beta+q\|}\sum\limits_{i=1}^{d}\mathbb{1}_D(i)\mathbb{E}[Y_1^i]} \geq 1,$$

under the assumption (which is no loss of generality) of an increasing order of $(v(q,D))_{D\subset\mathbb{D}}$ w.r.t. $\preceq$, where $\preceq$ is the order of $\mathbb{N}_0^\ell$ defined by

$$q \preceq q' \quad :\Longleftrightarrow \sum_{D\subset\mathbb{D}} \frac{\beta_D+q_D}{\|\bar\beta+q\|}\sum_{i=1}^{d}\mathbb{1}_D(i)\mathbb{E}[Y_1^i] \leq \sum_{D\subset\mathbb{D}} \frac{\beta_D+q'_D}{\|\bar\beta+q'\|}\sum_{i=1}^{d}\mathbb{1}_D(i)\mathbb{E}[Y_1^i] \tag{5.25}$$

for every $q=(q_D)_{D\subset\mathbb{D}}, q'=(q'_D)_{D\subset\mathbb{D}}\in\mathbb{N}_0^\ell$. Using the premium model introduced above, it may be shown that the optimal strategy is described by (5.27) with κ replaced by $\kappa(p,q)$. Furthermore, an analogous comparison result applies as in Corollary 5.32.

Now we move on to the aforementioned condition for the optimal reinsurance strategy further under the assumption of constant κ. Setting $\frac{\partial}{\partial b}f_2$ to zero (cf. Lemma 5.21), we obtain the first order condition

$$(1+\theta)\,\kappa = h_F(t,p,q,a). \tag{5.26}$$

By establishing this equation w.r.t. a we obtain a minimizer of f_2 w.r.t. b. If such a minimizer exists then the minimizer is unique because of the strict convexity property of f_2 w.r.t. a. The next proposition states that this equation is solvable and the solution takes values in $[0,1]$ depending on the safety loading parameter θ of the reinsurer.

Proposition 5.23. *For any $(t,p,q)\in[0,T]\times\Delta_m\times\mathbb{N}_0^\ell$, Equation (5.26) has the unique root w.r.t. a, denoted by $r_F(t,p,q)$, which is increasing w.r.t. the safety loading parameter of the reinsurer θ. Moreover, it holds,*

(i) $r_F(t,p,q)\leq 0$ *if* $\theta\leq A_F(t,p,q)/\kappa-1$,

(ii) $0<r_F(t,p,q)<1$ *if* $A_F(t,p,q)/\kappa-1<\theta<B_F(t,p,q)/\kappa-1$,

(iii) $r_F(t,p,q)\geq 1$ *if* $\theta\geq B_F(t,p,q)/\kappa-1$.

Proof. This follows by the same method as in the proof of Proposition 4.29. □

Notation. Throughout this chapter, $r_F(t,p,q)$ denotes the unique root from Proposition 5.23.

Notice that the cases (i) and (ii) in the proposition above could be empty sets. These cases would certainly be possible by using the setting given in Remark 5.22 with the modified $\kappa_F(p,q)$ depending on p and q. However, we carry on with the constant $\kappa>0$. Therefore, the proposition above provides the candidate for an optimal reinsurance strategy with the same structure as in Section 4.6. For any $(t,p,q)\in[0,T]\times\Delta_m\times\mathbb{N}_0^\ell$, we set

$$b_F(t,p,q) := \begin{cases} 0, & \theta\leq A_F(t,p,q)/\kappa-1,\\ 1, & \theta\geq B_F(t,p,q)/\kappa-1,\\ r_F(t,p,q), & \text{otherwise.}\end{cases} \tag{5.27}$$

Then the candidate for an optimal reinsurance strategy $(b_F^\star(t))_{t\in[0,T]}$ is given by $b_F^\star(t) := b_F(t-, p_{t-}, q_{t-})$. It is worth noting that the interpretation about the optimal reinsurance strategy given in Remark 4.30 applies here as well.

5.6 Verification

This section is devoted to a verification theorem to ensure that the solution of the stated generalized HJB equation yields the value function (see Theorem 5.24). We also demonstrate an existence theorem of a solution of the HJB equation (see Theorem 5.26), where we adapt the method from Section 4.7.

5.6.1 The verification theorem

Theorem 5.24. *Suppose there exists a bounded function $h : [0,T] \times \Delta_m \times \mathbb{N}_0^\ell \to (0,\infty)$ such that $t \mapsto h(t,p,q)$ and $t \mapsto h(t,\phi(t),q)$ with $\phi(0) = p$ are Lipschitz on $[0,T]$ for all $(p,q) \in \Delta_m \times \mathbb{N}_0^\ell$ as well as $p \mapsto h(t,p,q)$ is concave for all $(t,q) \in [0,T] \times \mathbb{N}_0^\ell$. Furthermore, h satisfies the generalized HJB equation*

$$0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \{\mathscr{L} h(t,p,q,\xi,b)\} + \inf_{\varphi\in\partial^C h_q(t,p)} \left\{\varphi_0 + \sum_{j=1}^m \varphi_j\, p_j \left(\sum_{k=1}^m \lambda_k\, p_k - \lambda_j\right)\right\}, \quad (5.28)$$

for all $(t,p,q) \in [0,T) \times \Delta_m \times \mathbb{N}_0^\ell$ with boundary condition

$$h(T,p,q) = 1, \quad (p,q) \in \Delta_m \times [0,T]. \quad (5.29)$$

Then

$$V(t,x,p,q) = -e^{-\alpha x e^{r(T-t)}} h(t,p,q), \quad (t,x,p,q) \in [0,T] \times \mathbb{R} \times \Delta_m \times \mathbb{N}_0^\ell,$$

and $(\xi^\star, b_F^\star) = (\xi^\star(s), b_F^\star(s))_{s\in[t,T]}$ with $\xi^\star(s)$ given by (4.29) and $b_F^\star(s) := b_F(s-, p_{s-}, q_{s-})$ given by (5.27) (with g replaced by h in $A_F(s,p,q)$ and $B_F(s,p,q)$) is an optimal feedback strategy for the given optimization problem (P2), i.e. $V(t,x,p,q) = V^{\xi^\star, b_F^\star}(t,x,p,q)$.

Same as in the previous chapter, an auxiliary result is essential to prove the theorem above, which can be shown with the aid of a measurement change introduced in Lemma A.10. This requires again a restriction of the admissible strategy and further notation.

Notation. Throughout this chapter, we set, for any $t \in [0,T)$,

$$\begin{aligned} \widetilde{\mathcal{U}}[t,T] &:= \{(\xi,b) \in \mathcal{U}[t,T] :\ \exists\, 0 < K < \infty : |\xi_s| \le K\ \forall\, s \in [t,T], \\ &\xi = (\xi_s)_{s\in[t,T]} \text{ is continuous and } \mathfrak{F}^W\text{-adapted}, b = (b_s)_{s\in[t,T]} \text{ is } \mathfrak{F}^\Psi\text{-predictable}\}. \end{aligned} \quad (5.30)$$

Moreover, we define

$$\widetilde{V}(t,x,p,q) := \sup_{(\xi,b)\in\widetilde{\mathcal{U}}[t,T]} V^{\xi,b}(t,x,p,q) \quad (5.31)$$

for all $(t,p,q) \in [0,T] \times \mathbb{R} \times \Delta_m \times \mathbb{N}_0^\ell$ and the operator D acting on function $h : [0,T] \times \Delta_m \times \mathbb{N}_0^\ell \to (0,\infty)$ by

$$Dh(t,p,q) := h_t(t,p,q) + \sum_{j=1}^m h_{p_j}(t,p,q)\, p_j \left(\sum_{k=1}^m \lambda_k\, p_k - \lambda_j\right) \quad (5.32)$$

for all functions h, where the right-hand side exists. Furthermore, we define an operator $\mathscr{H}$ acting on functions $v : [0,T] \times \Delta_m \times \mathbb{N}_0^\ell \to (0,\infty)$ and $(\xi, b) \in \mathbb{R} \times [0,1]$ by

$$\mathscr{H} v(t,p,q;\xi,b) := \mathscr{L} v(t,p,q;\xi,b) + D\, v(t,p,q) \tag{5.33}$$

for all functions $v : [0,T] \times \Delta_m \times \mathbb{N}_0^\ell \to (0,\infty)$, where the right-hand side is well-defined.

Using this notation, the generalized HJB equation (5.19) can be written as

$$0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \{\mathscr{H} g(t,p,q;\xi,b)\}$$

at those points (t,p,q) with existing $Dg(t,p,q)$. With this in mind, Lemma A.15 can be proven which is of decisive importance in the proof of the Verification Theorem 5.24.

Proof of Theorem 5.24. Using the Lemmata A.13, A.14 and A.15, the proof takes place complete analogously to the proof of Theorem 4.31. □

5.6.2 Existence result for the value function

We proceed as in Section 4.7.2 and show that there exists a function $h : [0,T] \times \Delta_m \times \mathbb{N}_0^\ell \to (0,\infty)$ satisfying the conditions stated in Theorem 5.24. For this purpose, let us define

$$\widetilde{g}(t,p,q) := \inf_{(\xi,b)\in\widetilde{\mathscr{U}}[t,T]} g^{\xi,b}(t,p,q), \quad (t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell, \tag{5.34}$$

where $g^{\xi,b}$ is given by (5.15) and $\widetilde{\mathscr{U}}[t,T]$ by (5.30). We begin with some properties of this function proved in a similar manner as Lemma 4.32.

Lemma 5.25. *The function $\widetilde{g}$ defined by* (5.34) *has the following properties:*

(i) $\widetilde{g}(t,p,q) > 0$ *for all* $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$.

(ii) $\widetilde{\mathscr{U}}[0,T] \ni (\xi,b) \mapsto g^{\xi,b}(0,p,q)$ *is bounded for all* $(p,q) \in \Delta_m \times \mathbb{N}_0^\ell$.

(iii) There exists a constant $0 < K_3 < \infty$ *such that* $\left|\widetilde{g}(t,p,q)\right| \le K_3$ *for all* $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$.

(iv) $\Delta_m \ni p \mapsto \widetilde{g}(t,p,q)$ *is concave for all* $(t,q) \in [0,T] \times \mathbb{N}_0^\ell$.

(v) $[0,T] \ni t \mapsto \widetilde{g}(t,p,q)$ *is Lipschitz on* $[0,T]$ *for all* $(p,q) \in \Delta_m \times \mathbb{N}_0^\ell$.

(vi) $[0,T] \ni t \mapsto \widetilde{g}(t,\phi(t),q)$ *with* $\phi(0) = p$ *is Lipschitz on* $[0,T]$ *for all* $(p,q) \in \Delta_m \times \mathbb{N}_0^\ell$.

(vii) Let M be the set of all points $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$, *where* $D\,\widetilde{g}$ *exists. Then there exists a constant* $0 < K_4 < \infty$ *such that* $|D\,\widetilde{g}(t,p,q)| \le K_4$ *for all* $(t,p,q) \in M$.

(viii) There exists a constant $0 < K_5 < \infty$ *such that* $\left|\mathscr{L}\widetilde{g}(t,p,q;\xi,b)\right| \le K_5$ *for all* $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$ *and* $(\xi,b) \in [-K,K] \times [0,1]$.

(ix) There exists a constant $0 < K_6 < \infty$ *such that* $\left|\inf_{(\xi,b)\in[-K,K]\times[0,1]} \mathscr{L}\widetilde{g}(t,p,q;\xi,b)\right| \le K_6$ *for all* $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$.

Proof. (i) Let us fix $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$ as well as $(\xi,b) \in \widetilde{\mathcal{U}}[t,T]$. Using the change of measure defined in Lemma A.10, it follows from the definition of $g^{\xi,b}$ given in (5.16) that

$$g^{\xi,b}(t,p,q) = \mathbb{E}^{t,p,q}_{\mathbb{Q}^{\xi,b}_t}\Bigg[\exp\bigg\{\int_t^T \bigg(-\alpha\, e^{r(T-s)}\Big((\mu-r)\xi_s + c(b_s) - \frac{1}{2}\alpha\,\sigma^2\, e^{r(T-s)}\xi_s^2\Big)$$
$$-\widehat{\Lambda}_s + \widehat{\Lambda}_s \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar\beta + q_s\|}\int_{(0,\infty)^d} \exp\Big\{\alpha\, b_s\, e^{r(T-s)}\sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y)\bigg)\mathrm{d}s\bigg\}\Bigg].$$

Taking (4.46) into account, we obtain

$$g^{\xi,b}(t,p,q) \geq \exp\big\{-\alpha\, T\, e^{rT}\big(|\mu-r|\, K + (1+\eta)\,\kappa\big) - m\,\lambda_m\, T\big\} =: C > 0.$$

Hence, due to the arbitrariness of $(\xi,b) \in \widetilde{\mathcal{U}}[t,T]$, the infimum of $g^{\xi,b}(t,p,q)$ over all $(\xi,b) \in \widetilde{\mathcal{U}}[t,T]$ is greater than or equal to C which yields the statement by definition of $\widetilde{g}$ given in (5.34).

(ii) Fix $(p,q) \in \Delta_m \times \mathbb{N}_0^\ell$. From the definition of $g^{\xi,b}$ given in (5.16) and Lemma A.10, we get

$$\big|g^{\xi,b}(0,p,q)\big| = \mathbb{E}^{0,p,q}_{\mathbb{Q}^{\xi,b}_T}\Bigg[\exp\bigg\{\int_0^T \bigg(-\alpha e^{r(T-s)}\Big((\mu-r)\xi_s + c(b_s) - \frac{1}{2}\alpha\sigma^2 e^{r(T-s)}\xi_s^2\Big)$$
$$+\widehat{\Lambda}_s \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar\beta + q_s\|}\int_{(0,\infty)^d} \exp\Big\{\alpha\, b_s\, e^{r(T-s)}\sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y) - \widehat{\Lambda}_s\bigg)\mathrm{d}s\bigg\}\Bigg]$$
$$\leq \exp\Big\{\alpha e^{rT}\Big(|\mu-r|K + (2+\eta+\theta)\,\kappa + \frac{1}{2}\alpha\sigma^2 e^{rT}K^2\Big)T + m\lambda_m\, M_F(\alpha e^{rT})T\Big\},$$

which yields statement (i), where we have used the same arguments as in the proof of Lemma A.11 for the last inequality above which is possible since $\widetilde{\mathcal{U}}[t,T] \subset \mathcal{U}[t,T]$.

(iii) The statement follows immediately from Lemma 5.20 (i).

(iv) The announced assertion is directly implied by Lemma 5.20 (vi).

(v) Let us fix $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$. Following the proof of Lemma 4.32 (iv) and using Lemma A.10, for any $0 \leq t_1 < t_2 \leq T$ and $\varepsilon > 0$, there exists a strategy $(\bar\xi,\bar b) \in \widetilde{\mathcal{U}}[0,T-t_1]$ such that

$$\big|\widetilde{g}(t_1,p,q) - \widetilde{g}(t_2,p,q)\big|$$
$$\leq \Bigg|\mathbb{E}^{0,p,q}_{\mathbb{Q}^{\bar\xi,\bar b}_{T-t_1}}\Bigg[\exp\bigg\{\int_0^{T-t_2} F_s\,\mathrm{d}s\bigg\}\bigg(\exp\bigg\{\int_{T-t_2}^{T-t_1} F_s\,\mathrm{d}s\bigg\} - 1\bigg)\Bigg]\Bigg| + \varepsilon$$
$$\leq \Bigg|\mathbb{E}^{0,p,q}_{\mathbb{Q}^{\bar\xi,\bar b}_{T-t_1}}\Bigg[\exp\bigg\{\int_0^{T} |F_s|\,\mathrm{d}s\bigg\}\bigg(\exp\bigg\{\int_{T-t_2}^{T-t_1} |F_s|\,\mathrm{d}s\bigg\} - e^0\bigg)\Bigg]\Bigg| + \varepsilon$$

where

$$F_s := -\alpha e^{r(T-s)}\Big((\mu-r)\xi_s + c(b_s) - \frac{1}{2}\alpha\sigma^2 e^{r(T-s)}\xi_s^2\Big)$$
$$+\widehat{\Lambda}_s \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar\beta + q_s\|}\int_{(0,\infty)^d} \exp\Big\{\alpha b_s e^{r(T-s)}\sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y) - \widehat{\Lambda}_s.$$

Same as in the proof of statement (ii), we have $|F_s| \le C$ for every $s \in [0,T]$, where

$$C := \alpha e^{rT}\Big(|\mu - r|\,K + (2+\eta+\theta)\kappa + \frac{1}{2}\alpha\sigma^2 e^{rT} K^2\Big) + m\lambda_m\, M_F(\alpha e^{rT}).$$

Therefore, by Proposition B.1,

$$\begin{aligned}
&\big|\widetilde{g}(t_1,p,q) - \widetilde{g}(t_2,p,q)\big| \\
&\le \left|\mathbb{E}^{0,p,q}_{\mathbb{Q}^{\bar{\xi},\bar{b}}_{T-t_1}}\left[e^{CT}e^{CT}(e-1)\Big(\int_{T-t_2}^{T-t_1}|F_s|\,\mathrm{d}s - 0\Big)\right]\right| + \varepsilon \\
&\le \big|e^{2CT}(e-1)C(T-t_1-T+t_2)\big| + \varepsilon \le e^{2CT}(e-1)C|t_2-t_1| + \varepsilon.
\end{aligned}$$

By $\varepsilon \downarrow 0$, we get that $t \mapsto \widetilde{g}(t,p,q)$ is Lipschitz on $[0,T]$ of rank $L := e^{2CT}(e-1)C$.

(vi) The approach of this proof is taken from the proof of Lemma 6.1 e) in Bäuerle and Rieder [31]. Fix $(t,p,q) \in [0,T]\times\Delta_m\times\mathbb{N}_0^\ell$ and $0 \le t_1 < t_2 \le T$. We first observe that

$$\begin{aligned}
&\big|\widetilde{g}(t_2,\phi(t_2),q) - \widetilde{g}(t_1,\phi(t_1),q)\big| \\
&\le \big| - \widetilde{g}(t_2,\phi(t_2),q) - (-\widetilde{g}(t_2,\phi(t_1),q))\big| + \big|\widetilde{g}(t_2,\phi(t_1),q) - \widetilde{g}(t_1,\phi(t_1),q)\big|.
\end{aligned}$$

According to statement (iii), we have

$$\big|\widetilde{g}(t_2,\phi(t_1),q) - \widetilde{g}(t_1,\phi(t_1),q)\big| \le K_3\,|t_2-t_1|.$$

Moreover, in the case that p is located on the boundary of Δ_m, we have $\phi(t_1) = \phi(t_2) = p$ and thus $\big| - \widetilde{g}(t_2,\phi(t_2),q) - (-\widetilde{g}(t_2,\phi(t_1),q))\big| = 0$. In the following, we consider the case $p \in \mathring{\Delta}_m$. By statement (iv), $\widetilde{g}(t_2,\cdot,q)$ is concave on $\mathring{\Delta}_m$ and thus it follows from Theorem 2.2 that $\widetilde{g}(t_2,\cdot,q) \in Lip_{loc}(\mathring{\Delta}_m)$. Due to the continuity of ϕ, $\phi_{\min} := \min_{t\in[0,T]}\phi(t)$ and $\phi_{\max} := \max_{t\in[0,T]}\phi(t) \in \mathring{\Delta}_m$ exist and thus $\phi : [0,T] \to [\phi_{\min},\phi_{\max}] \subset \mathring{\Delta}_m$, i.e. ϕ maps on a compact subset of $\mathring{\Delta}_m$. Hence, $\widetilde{g}(t_2,\phi(\cdot),q) \in Lip([\phi_{\min},\phi_{\max}])$, and, in consequence, there exists a constant $0 < K_7 < \infty$ such that

$$\big| - \widetilde{g}(t_2,\phi(t_2),q) - (-\widetilde{g}(t_2,\phi(t_1),q))\big| \le K_7|t_2-t_1|.$$

From Proposition 5.14 (ii), we know that $t \mapsto \phi(t)$ is Lipschitz of some rank $0 < K_8 < \infty$, where K_8 is independent of p. In summary, we obtain

$$\big|\widetilde{g}(t_2,\phi(t_2,p),q) - \widetilde{g}(t_1,\phi(t_1,p),q)\big| \le (K_7\,K_8 + K_3)\,|t_2-t_1|.$$

(vii) Fix $(t,p,q) \in M$ and thus $p \in \mathring{\Delta}_m$. Using statement (v) as well as the local Lipschitz property of convex functions on convex open sets, here $\mathring{\Delta}_m$ (see Theorem 2.2), we obtain

$$\begin{aligned}
\big|D\,\widetilde{g}(t,p,q)\big| &\le \lim_{h\downarrow 0}\frac{1}{h}\Big|\widetilde{g}(t+h,p,q) - \widetilde{g}(t,p,q)\Big| \\
&\quad + \sum_{j=1}^m \lim_{h\downarrow 0}\frac{1}{h}\Big| - \widetilde{g}(t,p+he_j,q) - (-\widetilde{g}(t,p,q))\Big)\Big|\,p_j\,\Big|\sum_{k=1}^m p_k\,\lambda_k - \lambda_j\Big|
\end{aligned}$$

$$\le \lim_{h\downarrow 0} \frac{1}{h} L_1\, h + \lim_{h\downarrow 0} \frac{1}{h} \sum_{j=1}^m L_2^j\, h \left| \sum_{k=1}^m p_k\, \lambda_k - \lambda_j \right| \le L_1 + \sum_{j=1}^m L_2^j\,(m+1)\lambda_m =: K_4$$

for some constants $0 < L_1 < \infty$ and $0 < L_2^1, \ldots, L_2^m < \infty$.

(viii) Fix $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$ and $(\xi, b) \in [-K,K] \times [0,1]$. It follows from the definition of $\mathscr{L}$ given in (5.18) statement (iii) that

$$\big|\mathscr{L}\widetilde{g}(t,p,q;\xi,b)\big| \le K_3 \bigg(m\lambda_m + \alpha\, e^{rT} \Big(|\mu - r| K + (2+\eta+\theta)\,\kappa + \frac{1}{2}\alpha\,\sigma^2 e^{rT} K^2 \Big) + m\lambda_m\, M_F\big(\alpha\, e^{rT}\big) \bigg) =: K_5$$

(ix) Fix $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$. In the same way as in the proof the previous statement, the following results arise by taking account of (4.30), (5.21) and (5.23):

$$\Big| \inf_{(\xi,b)\in[-K,K]\times[0,1]} \mathscr{L}\widetilde{g}(t,p,q;\xi,b) \Big| \le K_3 \bigg(m\lambda_m + \alpha e^{rT} \frac{1}{2} \frac{(\mu-r)^2}{\sigma^2} \frac{1}{\alpha} + \alpha e^{rT}(2+\eta+\theta)\kappa + m\lambda_m M_F(\alpha e^{rT}) \bigg) =: K_6. \ \square$$

The proof of the next existence result of a solution of the generalized HJB equation is quite similar to the proof of Theorem 4.33, for which reason some analogous argumentation is omitted.

Theorem 5.26. *The value function of the investment-reinsurance problem stated in* (P2) *is given by*

$$V(t,x,p,q) = -e^{-\alpha x e^{r(T-t)}} g(t,p,q), \quad (t,x,p,q) \in [0,T] \times \mathbb{R} \times \Delta_m \times \mathbb{N}_0^\ell,$$

where g is defined by (5.34) *and satisfies the generalized HJB equation*

$$0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \{\mathscr{L}g(t,p,q;\xi,b)\} + \inf_{\varphi \in \partial^C g_q(t,p)} \bigg\{ \varphi_0 + \sum_{j=1}^m \varphi_j\, p_j \Big(\sum_{k=1}^m \lambda_k\, p_k - \lambda_j \Big) \bigg\}$$

for all $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$ with boundary condition $g(T,p,q) = 1$ for all $(p,q) \in \Delta_m \times \mathbb{N}_0^\ell$. Furthermore, $(\xi^\star, b_F^\star) = (\xi^\star(s), b_F^\star(s))_{s\in[t,T]}$ with $\xi^\star(s)$ given by (4.29) *and $b_F^\star(s) = b_F(s-, p_{s-}, q_{s-})$ given by* (5.27) *is the optimal investment-reinsurance strategy of the Problem* (P2).

Proof. We follow the proof of Theorem 4.33. Fix $t \in [0,T)$ and $(\xi,b) \in \widetilde{\mathcal{U}}[t,T]$ and set

$$f(t,x) := -e^{-\alpha x e^{r(T-t)}}, \quad (t,x) \in [0,T] \times \mathbb{R}.$$

Let τ be the first jump time of $X^{\xi,b}$ after t, τ' the last jump of $X^{\xi,b}$ before τ and $t' \in (t,T]$. Since $\widetilde{\mathcal{U}}[t,T] \subset \mathcal{U}[t,T]$, it follows from Lemma A.15 that

$$\begin{aligned} &\widetilde{V}(\tau \wedge t', X_{\tau\wedge t'}^{\xi,b}, p_{\tau\wedge t'}, q_{\tau\wedge t'}) \\ &= \widetilde{V}(t, X_t^{\xi,b}, p_t, q_t) + \int_t^{\tau\wedge t'} f(s, X_s^{\xi,b})\, \mathscr{H}\widetilde{g}(s,p_s,q_s;\xi_s,b_s)\, \mathrm{d}s + \eta_{\tau\wedge t'}^{\xi,b} - \eta_t^{\xi,b}, \end{aligned} \tag{5.35}$$

where $(\eta_t^{\xi,b})_{t\in[0,T]}$ is a $\mathfrak{G}$-martingale starting at zero and we set $\mathscr{H}\widetilde{g}(s,p_s,q_s;\xi_s,b_s)$ to zero at those $s \in [t,T]$ where the $D\widetilde{g}(s,p_s,q_s)$ does not exist. Using the same arguments as in the proof of Theorem 4.33, we obtain

$$\begin{aligned}0 \geq \lim_{t' \downarrow t} \mathbb{E}^{t,x,p,q}\left[\frac{1}{t'-t}\int_t^{t'} f(s,X_s^{\xi,b})\,\mathscr{H}\widetilde{g}(s,p_s,q_s;\xi_s,b_s)\,\mathrm{d}s \,|\, t'<\tau\right]\mathbb{P}^{t,x,p,q}(t'<\tau)\\ + \lim_{t' \downarrow t} \mathbb{E}^{t,x,p,q}\left[\frac{1}{t'-t}\int_t^{\tau} f(s,X_s^{\xi,b})\,\mathscr{H}\widetilde{g}(s,p_s,q_s;\xi_s,b_s)\,\mathrm{d}s \,|\, t'\geq\tau\right]\mathbb{P}^{t,x,p,q}(t'\geq\tau).\end{aligned}$$

Again according to the proof of Theorem 4.33, we have $\mathbb{P}_\lambda^{t,x,p,q}(\tau > t') = e^{-\lambda(t'-t)}$ $\mathbb{P}$-a.s. and, in consequence,

$$\mathbb{P}^{t,x,p,q}(\tau \leq t') = \int_A \mathbb{P}_\lambda(\tau \leq t')\,\Pi_\Lambda(\mathrm{d}\lambda) = \sum_{j=1}^m \left(1-e^{-\lambda(t'-t)}\right)\pi_\Lambda(j).$$

Thus

$$\lim_{t'\downarrow t} \mathbb{P}^{t,x,p,q}(\tau \leq t') = \sum_{j=1}^m \left(1-\lim_{t'\downarrow t} e^{-\lambda(t'-t)}\right)\pi_\Lambda(j) = 0.$$

Consequently,

$$0 \geq \lim_{t' \downarrow t} \mathbb{E}^{t,x,p,q}\left[\frac{1}{t'-t}\int_t^{t'} f(s,X_s^{\xi,b})\,\mathscr{H}\widetilde{g}(s,p_s,q_s;\xi_s,b_s)\,\mathrm{d}s\mathbb{1}_{\{t'<\tau\}}\right].$$

Notice that, due to Lemma 5.25 (vii) and (viii) as well as Lemma A.11

$$\begin{aligned}&\mathbb{E}^{t,x,p,q}\left[\frac{1}{t'-t}\int_t^{t'} f(s,X_s^{\xi,b})\,\mathscr{H}\widetilde{g}(s,p_s,q_s;\xi_s,b_s)\,\mathrm{d}s\mathbb{1}_{\{t'<\tau\}}\right]\\ &\leq \mathbb{E}^{t,x,p,q}\left[\frac{1}{t'-t}\int_t^{t'} |f(s,X_s^{\xi,b})|(K_5+K_4)\,\mathrm{d}s\right]\\ &\leq \frac{K_4+K_5}{t'-t}\int_t^{t'} \mathbb{E}^{t,x,p,q}_{\mathbb{Q}_s^{\xi,b}}\left[\frac{|f(s,X_s^{\xi,b})|}{L_s^{\xi,b}}\right]\mathrm{d}s\\ &\leq \frac{K_4+K_5}{t'-t}K_1(t'-t) = (K_4+K_5)\,K_1 < \infty.\end{aligned}$$

Therefore, by the dominated convergence theorem, we can interchange the limit and the expectation. That is,

$$0 \geq \mathbb{E}^{t,x,p,q}\left[\lim_{t' \downarrow t}\frac{1}{t'-t}\int_t^{t'} f(s,X_s^{\xi,b})\,\mathscr{H}\widetilde{g}(s,p_s,q_s;\xi_s,b_s)\,\mathrm{d}s\mathbb{1}_{\{t'<\tau\}}\right].$$

As in the proof of Theorem 4.33, we obtain by the FTLC and $\mathbb{1}_{\{t'<\tau\}} \to 1$ $\mathbb{P}$-a.s. for $t' \downarrow t$.

$$0 \geq \mathbb{E}^{t,x,p,q}\left[f(t,X_t^{\xi,b})\,\mathscr{H}\widetilde{g}(t,p_t,q_t;\xi_t,b_t)\,\mathrm{d}s\right].$$

From now on, let $(\xi,b) \in [-K,K]\times[0,1]$ and $\varepsilon > 0$ as well as $(\bar{\xi},\bar{b}) \in \widetilde{\mathcal{U}}[t,T]$ be a fixed strategy with $(\bar{\xi}_s,\bar{b}_s) \equiv (\xi,b)$ for $s \in [t,t+\varepsilon)$. Then

$$0 \geq \mathbb{E}^{t,x,p,q}\left[f(t,X_t^{\bar{\xi},\bar{b}})\,\mathscr{H}\widetilde{g}(t,p_t,q_t;\bar{\xi}_t,\bar{b}_t)\,\mathrm{d}s\right] = f(t,x)\mathscr{H}\widetilde{g}(t,p,q;\xi,b)$$

at those points (t,p,q), where $D\widetilde{g}(t,p,q)$ exists. On account of the negativity of f, we get

$$0 \le \mathscr{H}\widetilde{g}(t,p,q;\xi,b).$$

In the light of the arbitrariness of (ξ,b), we obtain

$$0 \le \inf_{(\xi,b)\in[-K,K]\times[0,1]} \big\{\mathscr{H}\widetilde{g}(t,p,q;\xi,b)\big\}.$$

We show next the inequality above if $D\widetilde{g}$ does not exist. For this purpose, let $M_q \subset [0,T]\times\Delta_m$ denote the set of points at which $\nabla\widetilde{g}_q(t,p)$ exists for any $q\in\mathbb{N}_0^\ell$. On the basis of Theorem 2.9, we have, for any $q\in\mathbb{N}_0^\ell$,

$$\partial^C\widetilde{g}_q(t,p) = \operatorname{co}\Big\{\lim_{n\to\infty}\nabla\widetilde{g}_q(t_n,p_n) : (t_n,p_n)\to(t,p), (t_n,p_n)\in M_q\Big\}.$$

That is, for every $\varphi\in\partial^C\widetilde{g}_q(t,p)\subset[0,T]\times\Delta_m$, there exists $u\in\mathbb{N}$ and $(\beta_1,\dots,\beta_u)\in\Delta_u$ such that $\varphi=\sum_{i=1}^u\beta_i\,\varphi^i$, where $\varphi^i=\lim_{n\to\infty}\nabla\widetilde{g}_q(t_n^i,p_n^i)$ for sequences $(t_n^i,p_n^i)_{n\in\mathbb{N}}$ with $\lim_{n\to\infty}(t_n^i,p_n^i)=(t,p)$ along with existing $\nabla\widetilde{g}_q$. From what has already been proved, it can be concluded that, for any $i=1,\dots,u$

$$0 \le \mathscr{L}\widetilde{g}(t_n^i,p_n^i,q;\xi,b)+\widetilde{g}_t(t_n^i,q_n^i,q)+\sum_{j=1}^m\widetilde{g}_{p_j}(t_n^i,p_n^i,q)(p_n^i)_j\bigg(\sum_{k=1}^m\lambda_k(p_n^i)_k-\lambda_j\bigg),$$

where $(p_n^i)_j$, $j=1,\dots,m$, denotes the jth component of the m-dimensional vector p_n^i. Thus, by the continuity of $t\mapsto\widetilde{g}(t,p,q)$, $p\mapsto\widetilde{g}(t,p,q)$ and $p\mapsto J(p)$, we get

$$\begin{aligned}0 \le{}& \beta_i\mathscr{L}\widetilde{g}(t,p,q;\xi,b)+\beta_i\lim_{n\to\infty}\widetilde{g}_t(t_n^i,q_n^i,q)\\&+\sum_{j=1}^m\beta_i\lim_{n\to\infty}\widetilde{g}_{p_j}(t_n^i,p_n^i,q)p_j\bigg(\sum_{k=1}^m\lambda_k\,p_k-\lambda_j\bigg),\quad i=1,\dots,u,\end{aligned}$$

which yields

$$\begin{aligned}0 \le{}& \mathscr{L}\widetilde{g}(t,p,q;\xi,b)\sum_{i=1}^u\beta_i+\sum_{i=1}^u\beta_i\lim_{n\to\infty}\widetilde{g}_t(t_n^i,q_n^i,q)\\&+\sum_{j=1}^m\sum_{i=1}^u\beta_i\lim_{n\to\infty}\widetilde{g}_{p_j}(t_n^i,p_n^i,q)p_j\bigg(\sum_{k=1}^m\lambda_k\,p_k-\lambda_j\bigg)\\={}&\mathscr{L}\widetilde{g}(t,p,q;\xi,b)+\varphi_0+\sum_{j=1}^m\varphi_j\,p_j\bigg(\sum_{k=1}^m\lambda_k\,p_k-\lambda_j\bigg).\end{aligned}$$

Due to the arbitrariness of $\varphi\in\partial^C\widetilde{g}_q(t,p)$ and $(\xi,b)\in[-K,K]\times[0,1]$, we obtain

$$0 \le \inf_{(\xi,b)\in[-K,K]\times[0,1]}\big\{\mathscr{L}\widetilde{g}(t,p,q;\xi,b)\big\}+\inf_{\varphi\in\partial^C g_q(t,p)}\bigg\{\varphi_0+\sum_{j=1}^m\varphi_j\,p_j\bigg(\sum_{k=1}^m\lambda_k\,p_k-\lambda_j\bigg)\bigg\}.$$

As next objective we can establish the reverse inequality with the same arguments as in the proof of Theorem 4.33. To summarize, we have

$$0 = \inf_{(\xi,b)\in[-K,K]\times[0,1]} \left\{\mathscr{L}\widetilde{g}(t,p,q;\xi,b)\right\} + \inf_{\varphi\in\partial^C g_q(t,p)} \left\{\varphi_0 + \sum_{j=1}^m \varphi_j\, p_j \left(\sum_{k=1}^m \lambda_k\, p_k - \lambda_j\right)\right\}.$$

With the same arguments as in the proof of Theorem 4.33, we get

$$\inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \left\{\mathscr{L}\widetilde{g}(t,p,q;\xi,b)\right\} = \inf_{(\xi,b)\in[-K,K]\times[0,1]} \left\{\mathscr{L}\widetilde{g}(t,p,q;\xi,b)\right\}$$

and that

$$\widetilde{g}(t,p,q) = \inf_{(\xi,b)\in\widetilde{\mathcal{U}}[t,T]} g^{\xi,b}(t,p,q) = \inf_{(\xi,b)\in\mathcal{U}[t,T]} g^{\xi,b}(t,p,q) = g(t,p,q).$$

Therefore, it follows from Lemma 5.25 (i), (iii), (iv), (v), (vi) and Theorem 5.24 that

$$V(t,x,p,q) = -e^{-\alpha x e^{r(T-t)}} g(t,p,q), \quad (t,x,p,q) \in [0,T]\times\mathbb{R}\times\Delta_m\times\mathbb{N}_0^\ell,$$

and that $(\xi^\star, b_F^\star) = (\xi^\star(s), b_F^\star(s))_{s\in[t,T]}$ with $\xi^\star(t)$ given by (4.29). Moreover, we obtain that $b_F^\star(s) := b_F(s-, p_{s-}, q_{s-})$ given by (5.27) is the optimal investment-reinsurance strategy of the optimization problem (P2). □

We have seen in the previous Theorem that g is the solution of the generalized HJB equation. Consequently, the generalized HJB equation (5.19) has a unique solution due to the uniqueness of g by definition.

5.7 Comparison results with the complete information case

We will present in this section a comparison result of the optimal reinsurance strategy given in Theorem 5.26 and the one in the case with full information given by (4.57). But first of all we derive bounds to the optimal strategy which can be calculated a priori, i.e. independent of the filter process $(p_t)_{t\geq 0}$ and the process $(q_t)_{t\geq 0}$. For this determination, we introduce the following terms.

Notation. Let $t \in [0,T]$ and $b \in \mathbb{R}$. Throughout this section, we set

$$h_F^{\min}(t,b) := \lambda_1 \min_{D\subset\mathbb{D}} \left\{ \sum_{i=1}^d \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{\alpha\, b\, e^{r(T-t)} \sum_{j=1}^m y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y) \right\},$$
$$h_F^{\max}(t,b) := \lambda_m \max_{D\subset\mathbb{D}} \left\{ \sum_{i=1}^d \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{\alpha\, b\, e^{r(T-t)} \sum_{j=1}^m y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y) \right\}.$$

Proposition 5.27. *Let $t \in [0,T]$. Then $\mathbb{R} \ni b \mapsto h_F^{\min}(t,b)$ and $\mathbb{R} \ni b \mapsto h_F^{\max}(t,b)$ are strictly increasing and strictly convex. Furthermore, it holds*

$$\lim_{b\to-\infty} h_F^{\min}(t,b) = \lim_{b\to-\infty} h_F^{\max}(t,b) = 0, \quad \lim_{b\to\infty} h_F^{\min}(t,b) = \lim_{b\to\infty} h_F^{\max}(t,b) = \infty.$$

Proof. This follows by the same analysis as in the proof of Proposition 4.29. □

The proposition justifies the next notation.

Notation. For some fixed $t \in [0,T]$, we denote the unique root of the equation $(1+\theta)\,\kappa = h_F^{\min}(t,b)$ w.r.t. b, and the unique root of the equation $(1+\theta)\,\kappa = h_F^{\max}(t,b)$ w.r.t. b by $r_F^{\min}(t)$ and $r_F^{\max}(t)$, respectively.

The announced a priori bounds are a direct consequence of the following theorem in connection with Proposition 5.27. Recall the definition of the function h_F given in (5.24).

Proposition 5.28. *For any $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$ and $b \in \mathbb{R}$, we have*

$$h_F^{\min}(t,b) \le h_F(t,p,q,b) \le h_F^{\max}(t,b).$$

Proof. Choose some $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$ and $\bar{b} \in \mathbb{R}$. Recall that $\lambda_1 < \lambda_2 < \ldots < \lambda_m$, compare Assumption 5.7. For any $(\xi,b) \in \mathcal{U}[t,T]$, an application of Lemma 5.20 (iv), (v) and (vi) yields

$$\begin{aligned}
&\sum_{k=1}^m \lambda_k p_k \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar{\beta}+q\|} g^{\xi,b}(t,J(p),v(q,D)) \times \\
&\qquad \sum_{i=1}^d \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{\alpha\,\bar{b}\,e^{r(T-t)} \sum_{j=1}^m y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y) \\
&= \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar{\beta}+q\|} \sum_{j=1}^m \lambda_j\,p_j\,g^{\xi,b}(t,e_j,v(q,D)) \times \\
&\qquad \sum_{i=1}^d \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{\alpha\,\bar{b}\,e^{r(T-t)} \sum_{j=1}^m y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y) \\
&\le h_F^{\max}(t,\bar{b}) \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar{\beta}+q\|} \sum_{j=1}^m p_j g^{\xi,b}(t,e_j,v(q,D)) \\
&= h_F^{\max}(t,\bar{b}) \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar{\beta}+q\|} g^{\xi,b}(t,p,v(q,D)) \\
&= h_F^{\max}(t,\bar{b})\,g^{\xi,b}(t,p,q).
\end{aligned}$$

Hence, by taking the infimum over all $(\xi,b) \in \mathcal{U}[t,T]$ on both sides, we get $h_F(t,p,q,\bar{b}) \le h_F^{\max}(t,\bar{b})$. The other announced inequality is obtained by analogue arguments. ☐

Corollary 5.29. *The optimal reinsurance strategy $b_F^\star = (b_F^\star(t))_{t\in[0,T]}$ from Theorem 5.26 has the following bounds:*

$$\max\{0, r_F^{\max}(t)\} \le b_F^\star(t) \le \min\{1, r_F^{\min}(t)\}, \quad t \in [0,T].$$

The bounds provide only a very large range of optimality due to the rough estimation in the proposition above. This is illustrated in Figure 5.2, where the bounds are calculated for the parameter selection from Section 5.7. For this setting, the graphic shows that the insurer should not take a full reinsurance (i.e. retention level of zero) since the lower bound (orange line) is always greater than zero. The upper bound (red line) is trivial in this example. It should be noted again that these bounds can be calculated by the insurer already at time zero for the entire time horizon, which is why only rough bounds can be expected.

Let us move on to the comparison result of the partial and the full observable case which should provide a tighter bound because of the dependence on observations. As

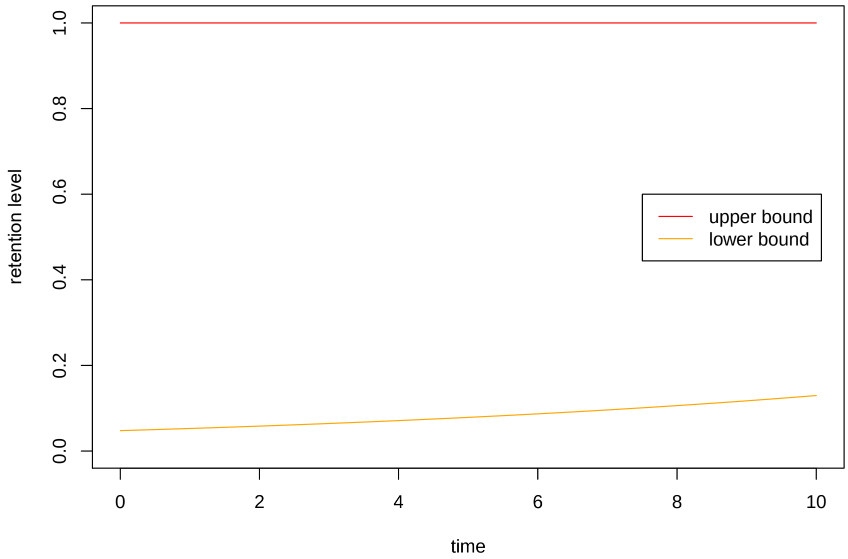

Figure 5.2: A priori upper bound (red line) and lower bound (orange line) for the optimal reinsurance strategy.

already explained in Section 4.8.2, a comparison result requires an order of the thinning probabilities w.r.t. the degree of harm for the insurer. In the given setting the $\mathbb{N}_0^\ell$-valued process $(q_t)_{t\geq 0}$ encapsulates the gathered information about the thinning probabilities expected from the claim arrivals. Therefore, an order on $\mathbb{N}_0^\ell$ is a necessity. Such an order is given by (5.25), which can be interpreted as an order of the weighted sum of expected claims of the LoBs $i \in D$ from the best to the worst case scenario from the insurer's perspective. However a stronger order than those in (5.25) is required, which is difficult to define since there is no natural order of the elements of $\mathcal{P}(\mathbb{D})$. Under the assumption of identical claim size distributions in every LoB, it is easily seen that the order defined in (5.25) satisfies

$$\begin{aligned} q \preceq q' \quad &\iff \quad \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar\beta + q\|}|D| \leq \sum_{D\subset\mathbb{D}} \frac{\beta_D + q'_D}{\|\bar\beta + q'\|}|D| \\ &\iff \quad \sum_{i=1}^d i \sum_{\substack{D\subset\mathbb{D}:\\ |D|=i}} \frac{\beta_D + q_D}{\|\bar\beta + q\|} \leq \sum_{i=1}^d i \sum_{\substack{D\subset\mathbb{D}:\\ |D|=i}} \frac{\beta_D + q'_D}{\|\bar\beta + q'\|}, \end{aligned}$$

for every $q = (q_D)_{D\subset\mathbb{D}} \in \mathbb{N}_0^\ell$ and $q' = (q'_D)_{D\subset\mathbb{D}} \in \mathbb{N}_0^\ell$, where the sums in the last line can be interpreted as expectations w.r.t. measures on the set $\mathbb{D} = \{1,\dots,d\}$. This observation is the initial point for the development of an order, which is useful for a comparison result. So we deal with the comparison under the assumption of identical claim size distributions for every LoB.

Assumption 5.30. Throughout this section, we suppose that

$$F(\mathrm{d}y) = \bar F(\mathrm{d}y_1) \otimes \bar F(\mathrm{d}y_2) \otimes \cdots \otimes \bar F(\mathrm{d}y_m),$$

where $\bar F$ is a distribution on $(0,\infty)$ with existing moment generation function.

It should be noted that Assumption 4.3 is satisfied under the assumption above, compare Remark 4.4. Due to the assumption, the order given in (5.25) depends only on the number of elements of $D \subset \mathbb{D}$. This motivates the next notation.

Notation. For any $q = (q_D)_{D\subset\mathbb{D}} \in \mathbb{N}_0^\ell$, we define $\tilde{q} = (\tilde{q}(1), \dots, \tilde{q}(d)) \in \mathbb{N}_0^d$ by

$$\tilde{q}(i) := \sum_{\substack{D\subset\mathbb{D}:\\ |D|=i}} q_D.$$

Using this notation, we define an equivalence relation on $\mathbb{N}_0^\ell$.

Notation. Throughout this section, let $\sim$ denote the equivalence relation on $\mathbb{N}_0^\ell$ which is defined by

$$q \sim q' \quad :\Longleftrightarrow \quad \tilde{q}' = \tilde{q}. \tag{5.36}$$

for every $q, q' \in \mathbb{N}_0^\ell$. Furthermore, $[q]$ is written for the equivalence class of $q \in \mathbb{N}_0^\ell$.

Justification of the notation. A trivial verification shows that the defined binary relation $\sim$ on $\mathbb{N}_0^\ell$ is reflexive, symmetric and transitive and thus an equivalence relation. □

Under consideration of this notation, we can identify every sequence $((\beta_D + q_D)/\|\bar{\beta} + q\|)_{D\subset\mathbb{D}} \in \Delta_\ell$ with a probability measure on $\mathbb{D}$.

Notation. Let $q = (q(i))_{i=1,\dots,d} \in \mathbb{N}_0^d$. Throughout this chapter, we denote the probability measure on $\mathbb{D}$ by F_q, which is defined by

$$F_q(B) := \sum_{i\in B} \frac{\tilde{\beta}(i) + q(i)}{\|\bar{\beta} + q\|}, \quad B \in \mathcal{P}(\mathbb{D}),$$

where $\tilde{\beta} = (\tilde{\beta}(1), \dots, \tilde{\beta}(d))$ with $\tilde{\beta}(i) = \sum_{D\subset\mathbb{D}:|D|=i} \beta_D$, $i = 1, \dots, d$.

Notice that $\|\bar{\beta}\| = \|\tilde{\beta}\|$. We are now in the position to define an order which turns out to be useful for the comparison result.

Notation. Let $q, q' \in \mathbb{N}_0^d$. Throughout this chapter, we define an order $\preceq_{\text{st}}$ on $\mathbb{N}_0^d$ by

$$q \preceq_{\text{st}} q' \quad :\Longleftrightarrow \quad F_q(x) \geq F_{q'}(x), \quad x \in \mathbb{R}, \tag{5.37}$$

where $F_a(x)$ denotes the distribution function of F_a at x for some $a \in \mathbb{N}_0^d$.

The defined order can be regarded as the usual stochastic order since, if $X \sim F_q$ and $Y \sim F_{q'}$, then the introduced order is equivalent to $X \preceq_{\text{st}} Y$, where $\preceq_{\text{st}}$ denotes the usual stochastic order.

The announced comparison result for the optimal reinsurance strategies under partial and full information is an immediate consequence of the next theorem. The proof of the theorem makes use of an order for the equivalence classes $[v(q, D)]$, $D \subset \mathbb{D}$. Since $[v(q, D)] = [v(q, D')]$ if $|D| = |D'|$, there are d equivalence classes of the set $\{[v(q, D)] : D \subset \mathbb{D}\}$. For the sake of simplicity, we introduce the following notation for these classes.

Notation. Throughout this section, we write $\tilde{v}_i \in \mathbb{N}_0^\ell$ for a represent of the equivalence class $\{v(q, D) : D \subset \mathbb{D} \text{ with } |D| = i\}$.

Using the introduced notations, it holds for $\tilde{v}_i = (\tilde{v}_i(1), \dots, \tilde{v}_i(d))$

$$\tilde{v}_i(j) = \begin{cases} \tilde{q}(j), & i \neq j, \\ \tilde{q}(j) + 1, & i = j, \end{cases}$$

for all $i \in \{1, \dots, d\}$. Consequently,

$$\sum_{j=1}^{n} \frac{\tilde{\beta}(j) + \tilde{v}_1(j)}{\|\bar{\beta} + q\| + 1} \geq \sum_{j=1}^{n} \frac{\tilde{\beta}(j) + \tilde{v}_2(j)}{\|\bar{\beta} + q\| + 1} \geq \dots \geq \sum_{j=1}^{n} \frac{\tilde{\beta}(j) + \tilde{v}_d(j)}{\|\bar{\beta} + q\| + 1}, \quad n = 1, \dots, d,$$

and thus

$$F_{\tilde{v}_1}(x) \geq F_{\tilde{v}_2}(x) \geq \dots \geq F_{\tilde{v}_d}(x), \quad x \in \mathbb{R}. \tag{5.38}$$

That is,

$$\tilde{v}_1 \preceq_{\mathrm{st}} \tilde{v}_2 \preceq_{\mathrm{st}} \dots \preceq_{\mathrm{st}} \tilde{v}_d. \tag{5.39}$$

Notice that the order above is equivalent to

$$\int_{\mathbb{D}} f(x)\, F_{\tilde{v}_1}(x) \leq \int_{\mathbb{D}} f(x)\, F_{\tilde{v}_2}(x) \leq \dots \leq \int_{\mathbb{D}} f(x)\, F_{\tilde{v}_d}(\mathrm{d}x), \tag{5.40}$$

for all increasing functions $f : \mathbb{D} \to \mathbb{R}$, for which both expectations exist, compare Müller and Stoyan [96, Thm. 1.2.8]. This order is crucial for the proof of the next theorem.

Theorem 5.31. *Let b_F be the function given by* (5.27) *and $\tilde{b}^{\star}_{\lambda,\bar{c},F}$ the function given by* (4.57). *Then, for any $(t, p, q) \in [0, T] \times \Delta_m \times \mathbb{N}_0^{\ell}$,*

$$b_F(t, p, q) \leq \tilde{b}^{\star}_{u(p),w(q),F}(t)$$

with

$$u(p) := \sum_{k=1}^{m} \lambda_k p_k, \quad w(q) := \left(\frac{\beta_D + q_D}{\|\bar{\beta} + q\|} \right)_{D \subset \mathbb{D}}.$$

Proof. Fix $(t, p, q) \in [0, T] \times \Delta_m \times \mathbb{N}_0^{\ell}$, $\bar{b} \in \mathbb{R}^+$ and $(\xi, b) \in \mathcal{U}[t, T]$. As in the proof of Theorem 4.41, it is sufficient to compare $h_F(t, p, q, \bar{b})$ given by (5.24) and $h_{u(p),w(q),F}(t, \bar{b})$ given by (4.55) due to the first order conditions (5.26) and (4.56). To draw this comparison, we first observe that, by Assumption 5.30,

$$\begin{aligned} g^{\xi,b}(t,p,q) := \mathbb{E}^{t,p,q}\Bigg[\exp\Bigg\{ &- \int_t^T \alpha\, e^{r(T-s)} \big((\mu - r)\xi_s + c(b_s) \big)\, \mathrm{d}s \\ &- \int_t^T \alpha\, \sigma\, e^{r(T-s)} \xi_s \,\mathrm{d}W_s + \alpha \sum_{n=1}^{N_{T-t}} b_{T_n} e^{r(T-T_n)} \sum_{\ell=1}^{|Z_n|} Y_n^{\ell} \Bigg\} \Bigg], \end{aligned} \tag{5.41}$$

compare the definition of $g^{\xi,b}$ given in (5.16). Notice that q determines the distribution of $\bar{\alpha} = (\alpha_D)_{D \subset \mathbb{D}}$ and the distribution of $|Z|$ is described by $\tilde{\alpha} = (\tilde{\alpha}(i))_{i=1,\dots,d}$ with

$$\tilde{\alpha}_i := \sum_{\substack{D \subset \mathbb{D}: \\ |D| = i}} \alpha_D,$$

where $\tilde{\alpha} \sim \mathrm{Dir}(\tilde{\beta} + \tilde{q})$ since $\bar{\alpha} \sim \mathrm{Dir}(\bar{\beta} + q)$, see DeGroot [49, p. 50]. Furthermore, by the independence of $(T_n)_{n \in \mathbb{N}}$ and $(Z_n)_{n \in \mathbb{N}}$ (see Assumption 3.2), we have for any

$i \in \{1, \dots, d\}$

$$g^{\xi,b}(t,p,\tilde{v}_i) = \sum_{m=0}^{\infty} \mathbb{P}^{t,p,\tilde{v}_i}(N_{T-t} = m) \int_{\mathbb{D}^m} \mathbb{E}^{t,p,\tilde{v}_i}\Big[\exp\Big\{ - \int_t^T \alpha\, e^{r(T-s)} \big((\mu - r)\,\xi_s + c(b_s)\big)\,\mathrm{d}s - \int_t^T \alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s + \alpha \sum_{n=1}^{m} b_{T_n}\, e^{r(T-T_n)} \sum_{\ell=1}^{|Z_n|} Y_n^\ell \Big\} \Big| N_{T-t} = m, (|Z_1|, \dots, |Z_m|) = (z_1, \dots, z_m)\Big] \mathbb{P}^{t,p,\tilde{v}_i}(|Z_1| \in \mathrm{d}z_1) \dots \mathbb{P}^{t,p,\tilde{v}_i}(|Z_m| \in \mathrm{d}z_m),$$

where, by Proposition 5.6, the distribution of $|Z_1|$ given $q_t = \tilde{v}_i$ (which represents the relevant information about the thinning probabilities up to time t) is

$$\mathbb{P}^{t,p,\tilde{v}_i}(|Z_1| = k) = \mathbb{P}^{t,\tilde{v}_i}(|Z_1| = k) = \mathbb{P}^{t,\tilde{v}_i}(Z_1 = E \text{ for all } E \subset \mathbb{D} \text{ with } |E| = k) = \sum_{\substack{E \subset \mathbb{D}:\\ |E| = k}} \mathbb{P}^{t,\tilde{v}_i}(Z_1 = E) = \frac{\tilde{\beta}(k) + \tilde{v}_i(k)}{\|\bar{\beta} + q\| + 1}, \quad k = 1, \dots, d.$$

That is, $\mathbb{P}^{t,\tilde{v}_i}(|Z_1| \in \mathrm{d}z_1) = F_{\tilde{v}_i}(\mathrm{d}z_1)$. Combining this with the fact that the integrand of the expectation of $g^{\xi,b}(t,p,\tilde{v}_i)$ is increasing in $|Z_n|$ for every $n = 1, \dots, m$, it follows from (5.40) that

$$g^{\xi,b}(t,p,\tilde{v}_1) \le g^{\xi,b}(t,p,\tilde{v}_2) \le \dots \le g^{\xi,b}(t,p,\tilde{v}_d). \tag{5.42}$$

Moreover, on account of Assumption 5.30, it holds

$$\begin{aligned}
&\sum_{i=1}^{d} \frac{\tilde{\beta}(i) + \tilde{q}(i)}{\|\bar{\beta} + q\|}\, i \Big(\int_{(0,\infty)} e^{\alpha \bar{b} e^{r(T-t)} y_1} \bar{F}(\mathrm{d}y_1) \Big)^{i-1} \int_{(0,\infty)} y_1 e^{\alpha \bar{b} e^{r(T-t)} y_1} \bar{F}(\mathrm{d}y_1) \\
&= \sum_{i=1}^{d} \sum_{\substack{D \subset \mathbb{D}:\\ |D| = i}} \frac{\beta_D + q_D}{\|\bar{\beta} + q\|} |D| \Big(\int_{(0,\infty)} e^{\alpha \bar{b} e^{r(T-t)} y_1} \bar{F}(\mathrm{d}y_1) \Big)^{|D|-1} \int_{(0,\infty)} y_1 e^{\alpha \bar{b} e^{r(T-t)} y_1} \bar{F}(\mathrm{d}y_1) \\
&= \sum_{D \subset \mathbb{D}} \frac{\beta_D + q_D}{\|\bar{\beta} + q\|} \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)} e^{\alpha \bar{b} e^{r(T-t)} y_1 \mathbb{1}_D(1)} \bar{F}(\mathrm{d}y_1) \times \dots \\
&\quad \int_{(0,\infty)} e^{\alpha \bar{b} e^{r(T-t)} y_1 \mathbb{1}_D(i-1)} \bar{F}(\mathrm{d}y_1) \int_{(0,\infty)} y_1 e^{\alpha \bar{b} e^{r(T-t)} y_1 \mathbb{1}_D(i)} \bar{F}(\mathrm{d}y_1) \times \\
&\quad \int_{(0,\infty)} e^{\alpha \bar{b} e^{r(T-t)} y_1 \mathbb{1}_D(i+1)} \bar{F}(\mathrm{d}y_1) \dots \int_{(0,\infty)} e^{\alpha \bar{b} e^{r(T-t)} y_1 \mathbb{1}_D(m)} \bar{F}(\mathrm{d}y_1) \\
&= \sum_{D \subset \mathbb{D}} \frac{\beta_D + q_D}{\|\bar{\beta} + q\|} \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\Big\{ \alpha\, \bar{b}\, e^{r(T-t)} \sum_{j=1}^{m} y_j \mathbb{1}_D(j) \Big\} F(\mathrm{d}y),
\end{aligned}$$

Furthermore, since $\int_{(0,\infty)} e^{\alpha \bar{b} e^{r(T-t)} y_1} \bar{F}(\mathrm{d}y_1) \ge 1$, the sequence

$$\Big(i \Big(\int_{(0,\infty)} e^{\alpha \bar{b} e^{r(T-t)} y_1} \bar{F}(\mathrm{d}y_1) \Big)^{i-1} \int_{(0,\infty)} y_1 e^{\alpha \bar{b} e^{r(T-t)} y_1} \bar{F}(\mathrm{d}y_1) \Big)_{i=1,\dots,d} \tag{5.43}$$

is orderly increasing. Notice that $((\tilde{\beta}_i + \tilde{q}_i)/\|\bar{\beta} + q\|)_{i=1,\dots,d} \in \Delta_d$. Consequently, applying

Lemma 5.20 (iv), (v) and Lemma B.6 (in connection with the increasing orders in (5.42) and (5.43)) yields

$$
\begin{aligned}
&\sum_{k=1}^{m} \lambda_k p_k \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar\beta + q\|} g^{\xi,b}(t, J(p), v(q,D)) \times \\
&\quad \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{\alpha\,\bar b\, e^{r(T-t)} \sum_{j=1}^{m} y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y) \\
&= \sum_{j=1}^{m} p_j \lambda_j \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar\beta + q\|} g^{\xi,b}(t, e_j, v(q,D)) \times \\
&\quad \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{\alpha\,\bar b\, e^{r(T-t)} \sum_{j=1}^{m} y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y) \\
&= \sum_{j=1}^{m} p_j \lambda_j \sum_{i=1}^{d} \frac{\tilde\beta(i) + \tilde q(i)}{\|\bar\beta + q\|}\, g^{\xi,b}(t, e_j, \tilde v_i)\, i \left(\int_{(0,\infty)} e^{\alpha\bar b e^{r(T-t)} y_1} \bar F(\mathrm{d}y_1)\right)^{i-1} \times \\
&\quad \int_{(0,\infty)} y_1 e^{\alpha\bar b e^{r(T-t)} y_1} \bar F(\mathrm{d}y_1) \\
&\geq \sum_{j=1}^{m} p_j \lambda_j \sum_{i=1}^{d} \frac{\tilde\beta(i) + \tilde q(i)}{\|\bar\beta + q\|} g^{\xi,b}(t, e_j, \tilde v_i) \sum_{k=1}^{d} \frac{\tilde\beta(k) + \tilde q(k)}{\|\bar\beta + q\|} k \left(\int_{(0,\infty)} e^{\alpha\bar b e^{r(T-t)} y_1} \bar F(\mathrm{d}y_1)\right)^{k-1} \times \\
&\quad \int_{(0,\infty)} y_1 e^{\alpha\bar b e^{r(T-t)} y_1} \bar F(\mathrm{d}y_1) \\
&= \sum_{j=1}^{m} p_j \lambda_j \sum_{E\subset\mathbb{D}} \frac{\beta_E + q_E}{\|\bar\beta + q\|} g^{\xi,b}(t, e_j, v(q,E)) \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar\beta + q\|} \times \\
&\quad \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{\alpha\,\bar b\, e^{r(T-t)} \sum_{j=1}^{m} y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y) \\
&= \sum_{j=1}^{m} p_j \lambda_j g^{\xi,b}(t, e_j, q) \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar\beta + q\|} \sum_{i=1}^{d} \mathbb{1}_D(i) \int\limits_{(0,\infty)^d} y_i \exp\left\{\alpha\bar b e^{r(T-t)} \sum_{j=1}^{m} y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y),
\end{aligned}
$$

where we have used the fact that every representation of the equivalence class $[v(q,D)]$ carries the same information about the distribution of $|Z_1|$ in the second equality. Recall that $\lambda_1 < \lambda_2 < \ldots < \lambda_m$ and notice that the background intensity of $(N_t)_{t\geq 0}$ in the definition of $g^{\xi,b}(t, e_j, q)$, $j = 1, \ldots, m$, is λ_j due to the choice of $p = e_j$. It is well-established that a Poisson process $(\widetilde N_t)_{t\geq 0}$ with intensity λ_k pathwise stochastically dominates a Poisson process $(\bar N_t)_{t\geq 0}$ with intensity λ_j, if $\lambda_k \geq \lambda_j$, $k, j \in \{1, \ldots, m\}$, compare Müller and Stoyan [96, Sec. 4.3.3], and thus $\sum_{n=1}^{\widetilde N_{T-t}} b_{T_n} e^{r(T-T_n)} \sum_{\ell=1}^{|Z_n|} Y_n^\ell$ pathwise stochastically dominates $\sum_{n=1}^{\bar N_{T-t}} b_{T_n} e^{r(T-T_n)} \sum_{\ell=1}^{|Z_n|} Y_n^\ell$. Consequently, $g^{\xi,b}(t, e_1, q) \leq g^{\xi,b}(t, e_2, q) \leq \ldots \leq g^{\xi,b}(t, e_m, q)$, compare (5.41). Therefore, a repeated application of Lemma B.6 as well as Lemma 5.20 (iii) leads to

$$
\sum_{j=1}^{m} p_j \lambda_j g^{\xi,b}(t, e_j, q) \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D}{\|\bar\beta + q\|} \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\left\{\alpha\bar b e^{r(T-t)} \sum_{j=1}^{m} y_j \mathbb{1}_D(j)\right\} F(\mathrm{d}y)
$$

$$
\begin{aligned}
&\geq \sum_{k=1}^{m} p_k\lambda_k \sum_{j=1}^{m} p_j g^{\xi,b}(t,e_j,q) \sum_{D\subset\mathbb{D}} \frac{\beta_D+q_D}{\|\bar\beta+q\|}\times \\
&\quad \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\Big\{\alpha\,\bar b\, e^{r(T-t)} \sum_{j=1}^{m} y_j \mathbb{1}_D(j)\Big\} F(\mathrm{d}y) \\
&= g^{\xi,b}(t,p,q) \sum_{k=1}^{m} p_k\lambda_k \sum_{D\subset\mathbb{D}} \frac{\beta_D+q_D}{\|\bar\beta+q\|} \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\Big\{\alpha\bar b e^{r(T-t)} \sum_{j=1}^{m} y_j \mathbb{1}_D(j)\Big\} F(\mathrm{d}y).
\end{aligned}
$$

In summary, we have

$$
\begin{aligned}
&\sum_{k=1}^{m} \lambda_k p_k \sum_{D\subset\mathbb{D}} \frac{\beta_D+q_D}{\|\bar\beta+q\|} g^{\xi,b}(t,J(p),v(q,D))\times \\
&\quad \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\Big\{\alpha\,\bar b\, e^{r(T-t)} \sum_{j=1}^{m} y_j \mathbb{1}_D(j)\Big\} F(\mathrm{d}y) \\
&\geq g^{\xi,b}(t,p,q) \sum_{k=1}^{m} \lambda_k p_k \sum_{D\subset\mathbb{D}} \frac{\beta_D+q_D}{\|\bar\beta+q\|} \sum_{i=1}^{d} \mathbb{1}_D(i) \int_{(0,\infty)^d} y_i \exp\Big\{\alpha\bar b e^{r(T-t)} \sum_{j=1}^{m} y_j \mathbb{1}_D(j)\Big\} F(\mathrm{d}y),
\end{aligned}
$$

which yields $h_F(t,p,q,\bar b) \geq h_{u(p),w(q),F}(t,\bar b)$ by taking the infimum over all $(\xi,b)\in\mathcal{U}[t,T]$ on both sides and the proof is complete. □

Corollary 5.32. *Let $\tilde b^\star_{\lambda,\bar c,F}$ be the function given by (4.57). Then the optimal reinsurance strategy under partial information $(b^\star_F(t))_{t\in[0,T]}$ from Theorem 5.26 satisfies*

$$
b^\star_F(t) \leq \tilde b^\star_{u(p_{t-}),w(q_{t-}),F}(t), \quad t\in[0,T].
$$

It should be noted that $(\tilde b^\star_{u(p_{t-}),w(q_{t-}),F}(t))_{t\in[0,T]}$ is $\mathfrak{G}$-predictable. Furthermore $u(p_{t-}) = \widehat\Lambda_{t-}$, which indicates that $u(p_{t-})$ is the known conditional average background intensity given the available information strict before time t. Moreover $w(q_t) = (\mathbb{E}[\alpha_D \mid \mathcal{F}^\Phi_t])_{D\subset\mathbb{D}}$ and thus $w(q_{t-})$ can be seen as the known conditional average thinning probabilities given again the information strict before time t. Consequently, the comparison result above has the same interpretation as the comparison result in the previous chapter (see Corollary 4.42); namely, more uncertainty leads to a more cautious optimal reinsurance strategy (that means, lesser or equal retention level). The comparison result will be graphically illustrated in the next section.

5.8 Numerical analyses

In this section we illustrate some numerical results in the case of two LoBs (i.e. $d=2$). The set of possible background intensities is $A=\{2,4,5\}$ (i.e. Λ takes values in A) and the prior probability mass function of Λ is supposed to be

$$
\bar\pi_\Lambda = \Big(\frac{2}{5},\frac{2}{5},\frac{1}{5}\Big).
$$

Furthermore, we assume that the prior parameter of the Dirichlet distribution of the thinning probabilities $\bar{\alpha}$ is

$$\bar{\beta} = (\beta_{\{1\}}, \beta_{\{2\}}, \beta_{\{1,2\}}) = (8, 7, 5).$$

Next we specify the claim size distribution F. Since we want to present the comparison result graphically and this result was derived for the assumption of independently and identically distributed claim sizes for every insurance class, we choose the same claim size distribution for both LoBs, namely a right-truncated exponential distribution with rate 1 and truncation at 3, i.e.

$$\mathbb{E}[Y_1^1] = \mathbb{E}[Y_1^2] = \frac{1}{1 - e^{-3}}.$$

Notice that Assumption 4.3 is fulfilled according to Remark 4.4. For the parameter κ of the premium principle, we select $\mathbb{E}[\mathrm{d}S_t]$, which yields, by Proposition 5.18,

$$\begin{aligned}\kappa &= \sum_{k=1}^{m} \lambda_k \, \pi_\Lambda(k) \sum_{D \subset \mathbb{D}} \frac{\beta_D}{\|\bar{\beta}\|} \sum_{i=1}^{d} \mathbb{1}_D(i) \mathbb{E}\big[Y_1^i\big] \\ &= \sum_{k=1}^{m} \lambda_k \, \pi_\Lambda(k) \mathbb{E}\big[Y_1^1\big] \frac{\tilde{\beta}(1) + 2\tilde{\beta}(2)}{\|\bar{\beta}\|} = \frac{17}{4 - 4e^{-3}}.\end{aligned}$$

The remaining parameters are chosen as in Table 5.1.

parameter	value
x_0	10
T	10
r	0.1
μ	0.2
σ	3
α	0.2
θ	0.6
η	0.2

Table 5.1: Simulation parameters for Section 5.8.

The following simulations are generated under the assumption that the realization of $\bar{\alpha}$ is $(0.38, 0.48, 0.14)$ and that the realization of Λ is 4 (i.e. $\mathbb{P}(\Lambda = 4 \,|\, \mathcal{F}_0) = 1$). Beside the a priori bounds (red and orange lines), we also illustrated in Figure 5.3 two trajectories (black and blue lines) of the reinsurance strategy $(\tilde{b}^{\star}_{u(p_{t-}),w(q_{t-}),F}(t))_{t\in[0,T]}$ with $u(p) = \sum_{k=1}^{m} \lambda_k p_k$ and $w(q) = ((\beta_D + q_D)/\|\bar{\beta} + q\|)_{D\subset\mathbb{D}}$, which provide for each scenario an upper bound for corresponding optimal reinsurance strategy according to Corollary 5.32. So the black and blue lines depend on the realized trigger arrival times and the affected LoBs. In both scenarios, the upper bounds (black and blue line) obtained from the comparison result are only useful up to approximately time 8. Before this, a strong dependence on the realizations can be seen.

Concluding the numerical illustration, we show the path of the surplus process in an insurance loss scenario for three different insurance strategies in Figure 5.4. In the case of full reinsurance (i.e. retention level of 0) the trajectory of the surplus process tends downward (red line), which is purchased through a negative premium rate. The blue line

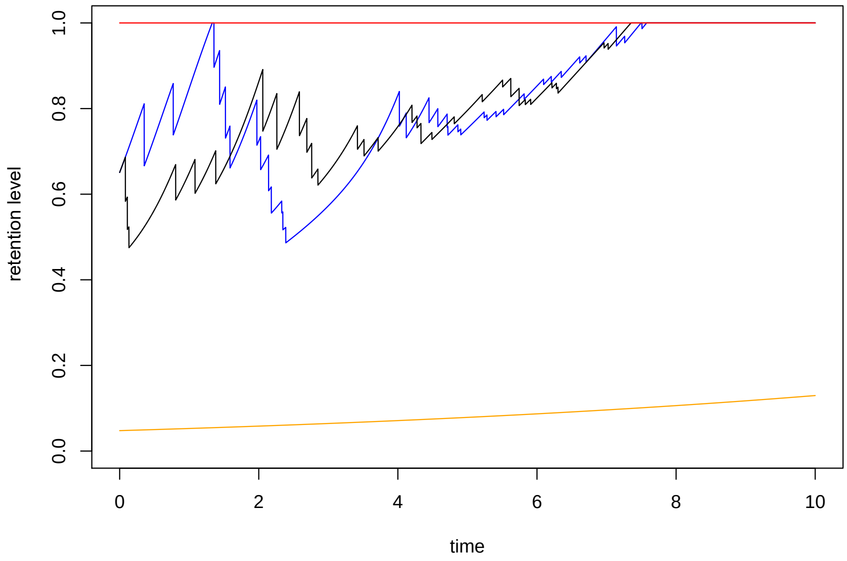

Figure 5.3: The a priori upper bound (red line) and lower bound (orange line) for the optimal reinsurance strategy and two paths of the reinsurance strategy $(\tilde{b}^{\star}_{u(p_{t-}),w(q_{t-}),F}(t))_{t\in[0,T]}$ with $u(p) := \sum_{k=1}^{m} \lambda_k p_k$ and $w(q) := ((\beta_D + q_D)/\|\bar{\beta} + q\|)_{D\subset\mathbb{D}}$.

displays a trajectory of the surplus for a constant reinsurance strategy of 0.5 and the black line for the reinsurance strategy $(\tilde{b}^{\star}_{u(p_{t-}),w(q_{t-}),F}(t))_{t\in[0,T]}$ with $u(p) = \sum_{k=1}^{m} \lambda_k p_k$ and $w(q) = ((\beta_D + q_D)/\|\bar{\beta} + q\|)_{D\subset\mathbb{D}}$. From Figure 5.3, we known that the latter reinsurance strategy tends upwards, which is evident in Figure 5.4 since the jump sizes of the black line are higher at the end of the considered time interval than those of the blue line. But because of the lower level of reinsurance, the surplus between losses rises stronger (as the premium rate is higher) than in the case of the constant reinsurance strategy.

5.9 Comments on generalizations

In this section we discuss generalizations of the setting and resulting difficulties with the used solution technique.

Conjugated prior for background intensity. The first generalization we look at concerns the prior distribution Π_Λ for the background intensity Λ. It was assumed that Π_Λ is defined on the finite set A. This assumption can be generalized in the sense that Λ is a $(0,\infty)$-valued random variable. To obtain a finite dimensional control problem, the prior distribution Π_Λ has to be a conjugated prior. In Reiss and Thomas [105, Eq. (3.52)], it is shown that the Gamma distribution is a conjugated prior for Λ, which essentially follows from the fact that the Gamma distribution is a conjugated prior for the exponential distribution. More precisely, we have

$$\begin{aligned} &\Lambda \,|\, \gamma, \zeta \sim \Gamma(\gamma, \zeta), \quad \gamma, \zeta > 0, \\ &\Lambda \,|\, \gamma, \zeta, T_1, \ldots, T_{N_t} \sim \Gamma(\gamma + N_t, \zeta + t). \end{aligned}$$

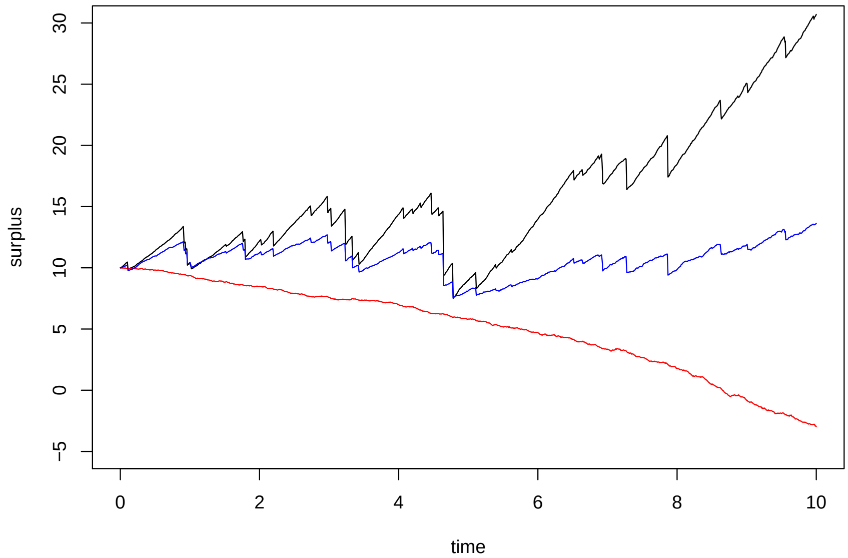

Figure 5.4: Trajectories of the surplus process for an insurance loss scenario in the cases of full reinsurance (red line), constant retention level of 0.5 (blue line) and the reinsurance strategy $(\tilde{b}^{\star}_{u(p_{t-}),w(q_{t-}),F}(t))_{t\in[0,T]}$ with $u(p) := \sum_{k=1}^{m}\lambda_k p_k$ and $w(q) := ((\beta_D + q_D)/\|\bar{\beta}+q\|)_{D\subset\mathbb{D}}$ (black line).

So the trigger process N is a mixed Poisson process with mixing $\Gamma(\gamma,\zeta)$-distribution. Such a process is called *Pólya process* which is a popular case of the mixed Poisson process in insurance mathematics, cf. e.g. Example (b) on p. 146 in Albrecher et al. [5].

Due to the conjugated property stated above, all information about the unknown parameter Λ, which is included in the observable filtration $\mathfrak{G}$ up to time t, is described by the processes $(N_t)_{t\geq 0}$ and time. In consequence, the state process of the reduced control problem with complete observation is the $(\ell+2)$-dimensional process $(X_s^{\xi,b}, N_s, q_s)_{s\in[t,T]}$ for some fixed initial time $t \in [0,T)$ and $(\xi,b) \in \mathcal{U}[t,T]$. The reduced control problem is given by

$$\begin{aligned} V^{\xi,b}(t,x,n,q) &:= \mathbb{E}^{t,x,n,q}\big[U(X_T^{\xi,b})\big] := \mathbb{E}\big[U(X_T^{\xi,b}) \,|\, X_t^{\xi,b} = x, N_t = n, q_t = q\big], \\ V(t,x,n,q) &:= \sup_{(\xi,b)\in\mathcal{U}[t,T]} V^{\xi,b}(t,x,n,q), \end{aligned}$$

for all $(t,x,n,q) \in [0,T] \times \mathbb{R} \times \mathbb{N}_0 \times \mathbb{N}_0^{\ell}$. Note that the processes $(N_t)_{t\geq 0}$ and $(q_t)_{t\geq 0}$ are pure jump processes. For this reason, analogue to Chapter 4, the following generalized HJB equation results:

$$\begin{aligned} 0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \Big\{ &-\frac{\gamma+n}{\zeta+u} g(t,n,q) \\ &- \alpha\, e^{r(T-t)} g(t,n,q)\Big((\mu-r)\xi + c(b) - \frac{1}{2}\alpha\sigma^2 e^{r(T-t)}\xi^2\Big) \\ &+ \frac{\gamma+n}{\zeta+t} \sum_{D\subset\mathbb{D}} \frac{\beta_D+q_D}{\|\bar{\beta}+q\|} g(t,n+1,v(q,D)) \int_{(0,\infty)^d} \exp\Big\{\alpha\, b\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y)\Big\} \end{aligned}$$

$$+ \inf_{\varphi \in \partial^C g_{n,q}(t)} \{\varphi\}$$

with boundary condition $g(T,n,q) = 1$, where

$$V(t,x,n,q) = -e^{-\alpha x e^{r(T-t)}} g(t,n,q),$$
$$g(t,n,q) := \inf_{(\xi,b)\in\mathcal{U}[t,T]} g^{\xi,b}(t,n,q)$$

and

$$g^{\xi,b}(t,n,q) := \mathbb{E}^{t,n,q}\bigg[\exp\bigg\{-\int_t^T \alpha\, e^{r(T-s)}((\mu-r)\xi_s + c(b_s))\,\mathrm{d}s - \int_t^T \alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s + \int_t^T\int_{E^d} \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\,\Psi(\mathrm{d}s,\mathrm{d}(y,z))\bigg\}\bigg].$$

With the arguments of Section 4.5 and 4.6, this arises the candidate for an optimal investment strategy, which is $(\xi^\star_t)_{t\in[0,T]}$ given by

$$\xi^\star_t = \frac{\mu-r}{\sigma^2}\frac{1}{\alpha}e^{-r(T-t)},$$

and the candidate for an optimal reinsurance strategy, which is $(b^\star_F(t))_{t\in[0,T]}$ given by $b_F(t-,N_{t-},q_{t-})$ with

$$b_F(t,n,q) = \begin{cases} 0, & \theta \le A_F(t,n,q)/\kappa - 1,\\ 1, & \theta \ge B_F(t,n,q)/\kappa - 1,\\ r_F(t,n,q), & \text{otherwise,}\end{cases}$$

where, for any $(t,n,q,b) \in [0,T]\times\mathbb{N}_0\times\mathbb{N}_0^\ell$,

$$\begin{aligned} h_F(t,n,q,b) &:= \frac{\gamma+n}{\zeta+t}\sum_{D\subset\mathbb{D}} \frac{\beta_D+q_D}{\|\bar\beta+q\|}\frac{g(t,n+1,v(q,D))}{g(t,n,q)}\times\\ &\qquad \sum_{i=1}^d \mathbb{1}_D(i)\int_{(0,\infty)^d} y_i \exp\bigg\{\alpha\, b\, e^{r(T-t)}\sum_{j=1}^d y_j \mathbb{1}_D(j)\bigg\}F(\mathrm{d}y)\\ A_F(t,n,q) &:= h_F(t,n,q,0),\\ B_F(t,n,q) &:= h_F(t,n,q,1),\end{aligned}$$

and $r_F(t,n,p)$ is the unique root w.r.t. b of

$$(1+\theta)\,\kappa = h_F(t,n,q,b).$$

The verification goes through similar to Section 4.7 under the condition that

$$\mathbb{E}\big[\exp\big\{\Lambda\big(M_F\big(2\alpha\, e^{rT}\big)-1\big)T\big\}\big] < \infty,$$

which is necessary to define an equivalent change of measure analogue to Lemma A.3, cf. proof of Lemma A.2. Unfortunately, the expectation above is only finite if the reciprocal scale parameter ζ of the prior distribution Π_Λ fulfils

$$\zeta > \big(M_F\big(2\alpha\, e^{rT}\big) - 1\big)T.$$

But in this case the $\Gamma(\gamma, \zeta)$-distribution is nearly a dirac distribution depending on the shape parameter γ. So if γ is chosen, Λ is nearly a known deterministic value. That means, we are in the case of an observable background intensity as in Chapter 4.

This approach with a conjugated prior for Λ works only for a conjugated prior distribution with existing moment generating function. But there does not exist such a prior distribution by the best knowledge of the author.

Additional time shift. Another generalization concerns the in insurance industry well-known acronym IBNR (Incurred But Not Reported) as well as a time gap between the shock event and the claims. In order to take these aspects into account, the model must be extended by a further random component which describes an additional time shift between the time of the trigger event and the occurrence of damages in some insurance lines. Such a model is developed in Bäuerle and Grübel [27, Ch. 3]. An additional time shift raises numerous problems for the optimization problem under partial information since then it is not clear which claim belongs to which trigger event and thus the category random elements $(Z_n)_{n\in\mathbb{N}}$ are no longer observable. This problem can be solved by assuming that the causation of the damage and the trigger events are known. But even then we cannot proceed as before since at every time point there can arise another damage as a result of a past trigger event. So only if the time shift is bounded by some constant $0 < K < \infty$, we have full information about the sequence Z_n at time $T_n + K$. Thus we could proceed as described in this chapter with the difference that the processes $(\tilde{p}_t)_{t\geq 0}$ and $(\tilde{q}_t)_{t\geq 0}$ containing the information about the unknown parameters at disposal are given by $\tilde{p}_t = p_{t-K}$ and $\tilde{q}_t = q_{t-K}$.

It is worth noting that there exists an estimation procedure for the background intensity λ in the model with shift without the assumptions described above. The estimation method is developed in Brown [23] for an analogous problem in the queuing theory. However, it is not readily possible to use other estimators than Bayesian estimators for the reduction, compare Remark 4.23.

More general trigger process. In the actuarial literature, more general models than the mixed Poisson process are discussed, which describe claim arrival processes. Schmidt [114] as well as Albrecher and Asmussen [3] suggest the use of a shot-noise driven claim arrival model, in particular for catastrophe modelling. An even more general process is the so-called dynamic contagion process which combines a shot-noise and a self-exciting property that can be observed for LoBs included claims as a result of abrasion. Analogous to our setting, the stochastic intensities of such claim arrival models are not observable, which in consequence requires stochastic filter technique. As seen, for example, in Leimcke [82, Thm. 3.29], the filter equation describing all information about the intensity process of the dynamic contagion process contained in the observed filtration has infinite dimension. This would lead to an infinite dimensional filter problem if such a process were used as the claim arrival model (or trigger process) and therefore cannot be solved with the approach presented in this paper. However, the dynamic programming HJB method is already extended to infinite-dimensional Hilbert space (compare Fabbri et al. [57] for an overview) which may be used to solve such kind of problem. Moreover, Brachetta and Ceci [18] provided recently an approach with backward stochastic differential equations (BSDEs) for optimal reinsurance problems under partial information with infinite dimensional filter equations, in which value process and the optimal reinsurance strategy are characterized as the unique solution of a BSDE driven by a marked point process.

Inter-dependency between financial and insurance risks. The last part of this section concerns the independence of the financial and insurance risks assumed in this work which results in an optimal investment and reinsurance strategy without interdependencies. But, according to Wang et al. [116, p. 114], there exist at least two reasons to believe that the insurance and financial risk should be dependent. First, (re)insurance companies transfer their insurance risks to the capital market by using insurance-linked securities, like catastrophe bonds, for instance. As a result, an insurer invested in the financial market is exposed to the insurance risks exported by another insurance company to the financial market, and there may be dependencies among these risks and the insurance risks of the insurance company invested in the financial market, for example through natural catastrophes. A second interconnectedness among financial and insurance risks in insurance contracts for financial guarantees, which can cause systemic risk. One way to establish a dependency between the financial and insurance market would be by a financial market with partial information described next. Portfolio optimization problems in a Bayesian financial market with one risk-free and one risky-asset, in which the drift rate of the risky asset is modelled as a random variable, whose outcome is unknown to the investor, have been investigated extensively, cf. e.g. Bäuerle and Grether [29] and the references given therein. This approach can be generalized by making the drift dependent on the state of an unobservable Markov chain. A portfolio optimization problem with such an unobservable Markov-modulated drift process is considered in Bäuerle and Rieder [30]. As already described in introduction, it is a common approach to modulate the claim arrival process by a Markov chain as well. The financial and insurance markets become dependent when both the drift rate of the risky asset and the intensity of the claim arrival process are modulated by the same unobservable Markov chain. The reduction of an optimization problem with such a model requires filter results with continuous and point process observations as provided by Ceci and Colaneri [36] since the insurer observes continuously the price process of the risky asset on the one hand and insurance claims on the other hand. However, even with such filter results, the solution approach of this paper cannot be applied without further thought, because we have shown the changes of measure in Lemma A.3 and A.10 and A.17 assuming independence between insurance and financial risks.

Chapter 6

Optimal investment and reinsurance for the univariate case with unknown claim size distribution

So far, we have always analyzed the partially observable problem (P) for a given loss distribution, but not for an unobservable one yet. This gap will be closed in this chapter, whereby we assume the background intensity to be observable and consider only one LoB for simplification.

6.1 Setting

This chapter only deals with one insurance line, i.e. we set $d = 1$. Therefore, the unobservable parameter $\bar{\alpha}$, which describes the dependencies between the lines, is trivially $\bar{\alpha} \equiv 1$ (and the sequence $(Z_n)_{n\in\mathbb{N}}$ is deterministic) and thus does not have to be taken into account further. In addition, we suppose that the prior distribution of Π_Λ is a one-point distribution. This makes the background intensity Λ directly observable for the insurance company and we set $\Lambda \equiv \lambda$ (i.e. $\Pi_\Lambda = \delta_\lambda$) for some $\lambda > 0$. Furthermore, Assumption 3.6 is considered to be in force.

Claim size distribution. Notice that in this section $(Y_n)_{n\in\mathbb{N}}$ is the sequence of $(0,\infty)$-valued random variable describing the amount of losses at the claim arrival times $(T_n)_{n\in\mathbb{N}}$, in which we assume that the claim size distribution is not observable and that Assumption 3.3 is satisfied (i.e. $(Y_n)_{n\in\mathbb{N}}$ and $(T_n)_{n\in\mathbb{N}}$ are independent). Next, we precise the prior distribution Π_ϑ.

Assumption 6.1. Let $m \in \mathbb{N}$ be a fixed number. We assume ϑ is an $\mathcal{F}_0$-measurable random variable taking values in the measure space $(\Theta, \mathcal{P}(\Theta))$, where $\Theta := \{1,\dots,m\}$. We further suppose that $F_1,\dots,F_m$ be absolutely continuous distributions on $(0,\infty)$ with densities $f_1,\dots,f_m$ such that

$$M_j(z) := \int_{(0,\infty)} e^{zy} f_j(y)\,\mathrm{d}y < \infty, \quad z \in \mathbb{R}, \quad j = 1,\dots,m.$$

Based on this assumption, the unknownness of the claim size distribution is modelled similarly to the background intensity in the previous chapter, namely by assuming that there are m potential loss distributions but the true distribution is unknown to the insurer. Thus, the entire information about the claim size distribution is encapsulated in the unknown parameter ϑ which results in the following characterization of the prior distribution for this parameter.

Notation. We set

$$\pi_\vartheta(j) := \mathbb{P}(\vartheta = j), \quad j = 1, \dots, m,$$

and we write $\bar{\pi}_\vartheta := (\pi_\vartheta(1), \dots, \pi_\vartheta(m))$ for the m-dimensional vector which describes the probability mass function of the distribution of ϑ.

Note that, like in the previous chapters, the subsequent considerations require exponential moments of the claim sizes which are ensured by the assumption above. The first moments of the distributions are also frequently demanded in the following, for that we introduce the next abbreviation.

Notation. For any $j \in \{1, \dots, m\}$, let $\mu_j := \int_0^\infty y\, f_j(y)\, \mathrm{d}y = \mathbb{E}_j[Y]$ denote the mean of the jth distribution, where $Y \,|\, \vartheta \sim Y_1 \,|\, \vartheta$ and $\mathbb{E}_j$ denotes the expectation w.r.t. to the distribution F_j.

Assumption 6.2. Without loss of generality we suppose that $\mu_1 \le \mu_2 \le \ldots \le \mu_m$.

Before we turn our attention to the reduction of the partially observable problem, let us consider some properties of the claim sizes.

Lemma 6.3. *Let $z \in \mathbb{R}$ be an arbitrary constant. Then there exist constants $0 < C_1 < \infty$ and $0 < C_2 < \infty$ such that*

(i) $\mathbb{E}\big[Y \exp\{z\, Y\}\big] \le C_1$,

(ii) $\mathbb{E}\Big[\exp\Big\{z \sum_{k=1}^{N_t} Y_k\Big\}\Big] \le C_2, \quad t \in [0, T]$.

Proof. As in the proof of Lemma 4.5 (i), we get

$$\mathbb{E}_k\big[Y \exp\{z\, Y\}\big] \le c_k, \quad k = \{1, \dots, m\},$$

where $0 < c_k < \infty$. Hence, on account of

$$\begin{aligned} \int_{(0,\infty)} g(y)\, \mathbb{P}(Y \in \mathrm{d}y) &= \int_\Theta \int_{(0,\infty)} g(y)\, \mathbb{P}(Y \in \mathrm{d}y, \vartheta \in \mathrm{d}\vartheta) \\ &= \int_\Theta \int_{(0,\infty)} g(y)\, \mathbb{P}(Y \in \mathrm{d}y \,|\, \vartheta = \vartheta')\, \Pi_\vartheta(\mathrm{d}\vartheta') \\ &= \sum_{k=1}^m \int_{(0,\infty)} g(y)\, \mathbb{P}(Y \in \mathrm{d}y \,|\, \vartheta = k)\, \pi_\vartheta(k) \\ &= \sum_{k=1}^m \int_{(0,\infty)} g(y)\, f_k(y)\, \mathrm{d}y\, \pi_\vartheta(k) \end{aligned} \tag{6.1}$$

for all $\mathcal{B}((0,\infty))$-measurable functions g, it follows

$$\mathbb{E}\big[Y \exp\{z\, Y\}\big] = \int_{(0,\infty)} y\, e^{zy}\, \mathbb{P}(Y \in \mathrm{d}y) = \sum_{k=1}^m \mathbb{E}_k\big[Y \exp\{z\, Y\}\big] \pi_\vartheta(k) \le \sum_{k=1}^m c_k\, \pi_\vartheta(k),$$

which is finite since the number of summands is finite. For the second statement we refer to the proof of Lemma 4.5 (ii) which yields

$$\mathbb{E}_k\Bigg[\exp\Bigg\{z \sum_{k=1}^{N_t} Y_k\Bigg\}\Bigg] = \exp\big\{\lambda\, t\big(\mathbb{E}_k\big[\exp\{z\, Y\}\big] - 1\big)\big\}$$

for every $k \in \{1, \dots, m\}$. Similar as above, this implies

$$\mathbb{E}\bigg[\exp\bigg\{z\sum_{k=1}^{N_t} Y_k\bigg\}\bigg] \le \sum_{k=1}^{m} \exp\{\lambda\, T\, c_k\}\, \pi_\vartheta(k) =: C_2,$$

where $0 < C_2 < \infty$ for all $t \in [0, T]$. □

6.2 Filtering and reduction

The aim of this section is to reduce the partially observable control problem (P) within the given framework in the previous section such that the state process of the reduced model takes the accident realizations into account, which yield information about the distribution of the claim sizes. Thus the reduction requires a filter process again. Notice that in the case of one LoB, the MPP $\Psi = (T_n, (Y_n, Z_n))_{n\in\mathbb{N}}$ can be identified with the $(0,\infty)$-MPP $\Psi = (T_n, Y_n)_{n\in\mathbb{N}}$, which justifies the use of Ψ for both marked point processes.

Filter for the claim size distribution. By the Bayes rule, the posterior probability mass function of ϑ given the observation $\bar{Y}_n = \bar{y}_n$ with $\bar{Y}_n := (Y_1, \dots, Y_n)$ and $\bar{y}_n := (y_1, \dots, y_n)$ is

$$\mathbb{P}(\vartheta = j \,|\, \bar{Y}_n = \bar{y}_n) = \frac{f_{\bar{Y}_n}(\bar{y}_n \,|\, \vartheta = j)\, \pi_\vartheta(j)}{\sum_{k=1}^{m} f_{\bar{Y}_n}(\bar{y}_n \,|\, \vartheta = k)\, \pi_\vartheta(k)} = \frac{\pi_\vartheta(j) \prod_{i=1}^{n} f_j(y_i)}{\sum_{k=1}^{m} \pi_\vartheta(k) \prod_{i=1}^{n} f_k(y_i)} \tag{6.2}$$

for all $j \in \{1, \dots, m\}$. However, the solution method necessitates a dynamic representation of this posterior probability distribution given the information up to any time t. To achieve that, let us introduce the following notation.

Notation. Throughout this chapter, we denote by $(p_j(t))_{t\ge 0}$ the càdlàg modification of the process $(\mathbb{P}(\vartheta = j \,|\, \mathcal{F}_t^\Psi))_{t\ge 0}$ for every $j \in \{1, \dots, m\}$, i.e.

$$p_j(t) = \mathbb{P}(\vartheta = j \,|\, \mathcal{F}_t^\Psi), \quad t \ge 0.$$

Moreover, let $(p_t)_{t\ge 0}$ denote the m-dimensional process defined by

$$p_t := (p_1(t), \dots, p_m(t)), \quad t \ge 0.$$

Remark 6.4. It is worth noting that $\sum_{k=1}^{m} p_k(t)\, \mu_k \le \sum_{k=1}^{m} \mu_k \le m\, \mu_m$ for all $t \ge 0$, due to Assumption 6.2. Furthermore, it is readily seen that $p_j(0) = \pi_\vartheta(j)$ for every $j \in \{1, \dots, m\}$.

We are interested in a representation of the process $(p_t)_{t\ge 0}$.

Proposition 6.5. *For any $j \in \{1, \dots, m\}$, the process $(p_j(t))_{t\ge 0}$ satisfies*

$$p_j(t) = \pi_\vartheta(j) + \int_0^t \int_{(0,\infty)} \bigg(\frac{p_j(s-)\, f_j(y)}{\sum_{k=1}^{m} p_k(s-)\, f_k(y)} - p_j(s-)\bigg) \Psi(\mathrm{d}s, \mathrm{d}y), \quad t \ge 0.$$

Proof. Fix $j \in \{1, \dots, m\}$. From (6.2) follows that the posterior probability mass function of ϑ given the observation $\bar{Y}_{n+1} = \bar{y}_{n+1}$ has the following recursive structure:

$$\begin{aligned}\mathbb{P}(\vartheta = j \,|\, \bar{Y}_{n+1} = \bar{y}_{n+1}) &= \frac{\pi_\vartheta(j) \prod_{i=1}^n f_j(y_i)\, f_j(y_{n+1})}{\sum_{k=1}^m \pi_\vartheta(k) \prod_{i=1}^n f_k(y_i)\, f_k(y_{n+1})} \\ &= \frac{f_j(y_{n+1})\, \mathbb{P}(\vartheta = j \,|\, \bar{Y}_n = \bar{y}_n)}{\sum_{k=1}^m f_k(y_{n+1})\, \mathbb{P}(\vartheta = k \,|\, \bar{Y}_n = \bar{y}_n)}.\end{aligned}$$

Thus, using the above notation, we get

$$p_j(t) = \mathbb{P}(\vartheta = j \,|\, Y_1, \ldots, Y_{N_t}) = \frac{f_j(Y_{N_t})\, p_j(t-)}{\sum_{k=1}^m f_k(Y_{N_t})\, p_k(t-)},$$

since $p_j(t-) = \mathbb{P}(\vartheta = j \,|\, Y_1, \ldots, Y_{N_{t-}}) = \mathbb{P}(\vartheta = j \,|\, Y_1, \ldots, Y_{N_t - 1})$. Clearly, the process $(p_j(t))_{t \geq 0}$ jumps only at the claim arrival times $(T_n)_{n \in \mathbb{N}}$ (at these time points the insurer gains new information about the unknown claim size distribution), where the jump size is

$$\Delta p_j(T_n) = \frac{f_j(Y_n)\, p_j(T_n-)}{\sum_{k=1}^m f_k(Y_n)\, p_k(T_n-)} - p_j(T_n-), \quad n \in \mathbb{N}.$$

Therefore,

$$\begin{aligned}p_j(t) &= p_j(0) + \sum_{0 < s \leq t} \Delta p_j(s) \\ &= \pi_\vartheta(j) + \int_0^t \int_{(0,\infty)} \left(\frac{p_j(s-)\, f_j(y)}{\sum_{k=1}^m p_k(s-)\, f_k(y)} - p_j(s-) \right) \Psi(\mathrm{d}s, \mathrm{d}y).\end{aligned}$$

for all $t \geq 0$. □

Let us state some elementary properties of the filter process $(p_t)_{t \geq 0}$, which follow immediately from the representation of the filter given in the previous proposition.

Corollary 6.6. *The filter $(p_t)_{t \geq 0}$ is a pure jump process and the new state of $(p_t)_{t \geq 0}$ after the jump times $(T_n)_{n \in \mathbb{N}}$ with jump size $(Y_n)_{n \in \mathbb{N}}$ is given by*

$$p_{T_n} = J\big(p_{T_n-}, Y_n\big), \quad n \in \mathbb{N},$$

where

$$J(p, y) := \left(\frac{f_1(y)\, p_1}{\sum_{k=1}^m f_k(y)\, p_k}, \ldots, \frac{f_m(y)\, p_m}{\sum_{k=1}^m f_k(y)\, p_k} \right), \tag{6.3}$$

for every $p = (p_1, \ldots, p_m) \in \Delta_m$ and $y \in (0, \infty)$.

Figure 6.1 shows a path of the filter process $(p_t)_{t \geq 0}$ under the conditions from Section 6.8 with three possible loss distributions, in which the second is the true one and the prior is given by $\bar{\pi}_\vartheta = (1/3, 1/3, 1/3)$. We see that the filter tends to increase the probability of the second (true) claim size distribution over time.

Properties of the aggregated claim amount process and the surplus process. It is essential to have representations of the aggregated claim amount process and the surplus process w.r.t. the compensated random measure Ψ for the development of the HJB equation of the reduced control problem stated in the next section. This measure is provided by the posterior predictive distribution of Y given the observed claim sizes up to time t.

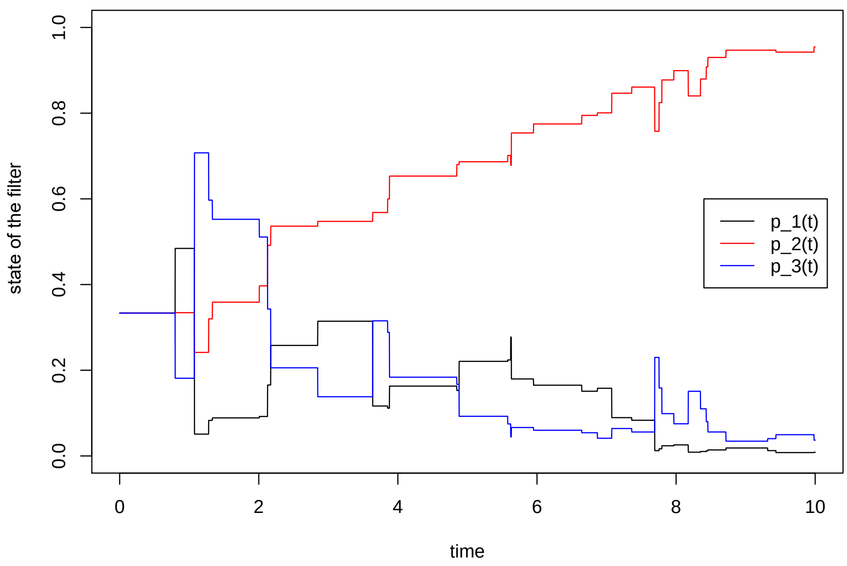

Figure 6.1: A trajectory of the filter process $(p_t)_{t\geq 0}$ for the setting from Section 6.8 with three potential claim size distributions under the assumptions that $\bar{\pi}_\vartheta = (1/3, 1/3, 1/3)$ and $\mathbb{P}(\vartheta = 2 \,|\, \mathcal{F}_0) = 1$.

Lemma 6.7. *For any $t \geq 0$, it holds*

$$\mathbb{P}(Y \in B \,|\, Y_1, \ldots, Y_{N_t}) = \sum_{k=1}^m p_k(t)\, F_k(B), \quad B \in \mathcal{B}((0,\infty)).$$

Proof. With the same arguments as in (6.1) follows for any $n \in \mathbb{N}$

$$\begin{aligned} f_{Y|\bar{Y}_n=\bar{y}_n}(y) &= \frac{f_{Y,\bar{Y}_n}(y,\bar{y}_n)}{f_{\bar{Y}_n}(\bar{y}_n)} = \sum_{k=1}^m \frac{f_{Y,\bar{Y}_n|\vartheta=k}(y,\bar{y}_n)}{f_{\bar{Y}_n}(\bar{y}_n)}\, \mathbb{P}(\vartheta = k) \\ &= \sum_{k=1}^m \frac{f_k(y)\, f_{\bar{Y}_n|\vartheta=k}(\bar{y}_n)}{f_{\bar{Y}_n}(\bar{y}_n)}\, \mathbb{P}(\vartheta = k) = \sum_{k=1}^m f_k(y)\, \mathbb{P}(\vartheta = k \,|\, \bar{Y}_n = \bar{y}_n), \end{aligned}$$

where we have used the conditional independence of the claims sizes given ϑ (cf. Assumption 6.1) in the third equality. Hence

$$f_{Y|Y_1,\ldots,Y_{N_t}}(y) = \sum_{k=1}^m f_k(y)\, p_k(t), \quad t \geq 0,$$

which implies the announced statement. □

Proposition 6.8. *The $\mathfrak{F}^\Psi$-intensity kernel of $\Psi = (T_n, Y_n)_{n\in\mathbb{N}}$, denoted by $(\nu(t, \mathrm{d}y))_{t\geq 0}$, is given by*

$$\nu(t,B) = \lambda \sum_{k=1}^m p_k(t-) \int_B f_k(y)\, \mathrm{d}y = \lambda \sum_{k=1}^m p_k(t-)\, F_k(B), \quad B \in \mathcal{B}((0,\infty)).$$

Proof. Fix $B \in \mathcal{B}((0,\infty))$. We first show that ν is a transition kernel from $(\mathbb{R}^+\times\Omega, \mathcal{B}^+\otimes\mathcal{F}^\Psi_\infty)$ to $((0,\infty), \mathcal{B}((0,\infty)))$. It is clear that $(t,\omega)\mapsto \lambda\sum_{k=1}^m p_k(\omega,t-)\,F_k(B)$ is $\mathcal{B}^+\otimes\mathcal{F}^\Psi_\infty$-measurable due to the measurability w.r.t. $\mathcal{F}^\Psi_\infty$ of the $\mathfrak{F}^\Psi$-predictable process $(p_k(t-))_{t\geq 0}$, $k=1,\dots,m$, compare Proposition 2.56 and Proposition 2.27. Moreover, it is easily seen that through $\lambda\sum_{k=1}^m p_k(t-)\,F_k(\mathrm{d}y)$, a measure on $((0,\infty),\mathcal{B}((0,\infty)))$ is defined for all $(t,\omega)\in\mathbb{R}^+\times\Omega$ since $F_k(\mathrm{d}y)$ is a distribution on $((0,\infty),\mathcal{B}((0,\infty)))$ for every $k\in\{1,\dots,m\}$. Now the procedure is to show that $(\nu(t,B))_{t\geq0} = (\lambda\sum_{k=1}^m p_k(t-)\,F_k(B))_{t\geq 0}$ is the predictable $\mathfrak{F}^\Psi$-intensity of $(\Psi(t,B))_{t\geq0}$, where the predictability follows immediately from the $\mathfrak{F}^\Psi$-predictability of $(p_k(t-))_{t\geq0}$ for every $k\in\{1,\dots,m\}$. To do this let $(H_t)_{t\geq0}$ be some non-negative $\mathfrak{F}^\Psi$-predictable process. It follows from Fubini's theorem and Lemma 6.7 that

$$
\begin{aligned}
&\mathbb{E}\Big[\int_0^\infty H_t\,\nu(t,B)\,\mathrm{d}t\Big] = \lambda\int_0^\infty \mathbb{E}\Big[H_t\sum_{k=1}^m p_k(t)\,F_k(B)\Big]\,\mathrm{d}t = \lambda\int_0^\infty \mathbb{E}\big[H_t\mathbb{P}(Y\in B\,|\,\mathcal{F}^\Psi_t)\big]\,\mathrm{d}t\\
&= \lambda\int_0^\infty \mathbb{E}\big[\mathbb{E}\big[H_t\,\mathbb{1}_{\{Y\in B\}}\,|\,\mathcal{F}^\Psi_t\big]\big]\mathrm{d}t = \lambda\int_0^\infty \mathbb{E}\big[\mathbb{E}\big[H_t\,\mathbb{1}_{\{Y\in B\}}\,|\,\vartheta\big]\big]\mathrm{d}t\\
&= \mathbb{E}\Big[\int_0^\infty \mathbb{E}\big[H_t\,\mathbb{1}_{\{Y\in B\}}\,|\,\vartheta\big]\,\lambda\,\mathrm{d}t\Big].
\end{aligned}
$$

Notice that the process $(\mathbb{E}\big[H_t\,\mathbb{1}_{\{Y\in B\}}\,|\,\vartheta\big])_{t\geq0}$ is a non-negative $\mathfrak{F}$-predictable process due to the $\mathcal{F}_0$-measurability of ϑ. Therefore, Brémaud [20, Eq. (2.3)] implies

$$
\mathbb{E}\Big[\int_0^\infty \mathbb{E}\big[H_t\,\mathbb{1}_{\{Y\in B\}}\,|\,\vartheta\big]\,\lambda\,\mathrm{d}t\Big] = \mathbb{E}\Big[\int_0^\infty \mathbb{E}\big[H_t\,\mathbb{1}_{\{Y\in B\}}\,|\,\vartheta\big]\,\mathrm{d}N_t\Big].
$$

It is worth noting that this arguments would also hold if λ is unobservable (i.e. $\mathcal{F}_0$-measurable), which means that we can use the same arguments to determine the $\mathfrak{F}^\Psi$-intensity for a setting with unobservable claim size distribution. By Assumption 3.3 we further have

$$
\begin{aligned}
&\mathbb{E}\Big[\int_0^\infty \mathbb{E}\big[H_t\,\mathbb{1}_{\{Y\in B\}}\,|\,\vartheta\big]\,\mathrm{d}N_t\Big] = \sum_{n\in\mathbb{N}}\mathbb{E}\big[\mathbb{E}\big[H_{T_n}\,\mathbb{1}_{\{T_n<\infty\}}\,\mathbb{1}_{\{Y\in B\}}\,|\,\vartheta\big]\big]\\
&= \sum_{n\in\mathbb{N}}\mathbb{E}\big[\mathbb{E}\big[H_{T_n}\,\mathbb{1}_{\{T_n<\infty\}}\,\mathbb{1}_{\{Y_n\in B\}}\,|\,\vartheta\big]\big] = \mathbb{E}\Big[\sum_{n\in\mathbb{N}} H_{T_n}\mathbb{1}_{\{T_n<\infty\}}\,\mathbb{1}_{\{Y_n\in B\}}\Big]\\
&= \mathbb{E}\Big[\int_0^\infty H_t\,\Psi(\mathrm{d}t,B)\Big].
\end{aligned}
$$

In summary, we have

$$
\mathbb{E}\Big[\int_0^\infty H_t\,\nu(t,B)\,\mathrm{d}t\Big] = \mathbb{E}\Big[\int_0^\infty H_t\,\Psi(\mathrm{d}t,B)\Big],
$$

i.e. $(\nu(t,\mathrm{d}y))_{t\geq0}$ is the $\mathfrak{F}^\Psi$-intensity kernel of $\Psi=(T_n,Y_n)_{n\in\mathbb{N}}$. □

Notation. Let $\widehat{\Psi}(\mathrm{d}t,\mathrm{d}y)$ denote the compensated random measure given by

$$
\widehat{\Psi}(\mathrm{d}t,\mathrm{d}y) := \Psi(\mathrm{d}t,\mathrm{d}y) - \nu(t,\mathrm{d}y)\,\mathrm{d}t, \tag{6.4}
$$

where ν is defined as in Proposition 6.8.

This notation leads to the following representation of aggregated claim amount process.

Proposition 6.9. *The aggregated claim amount process $S = (S_t)_{t\geq 0}$ is given by*

$$S_t = \int_0^t \int_{(0,\infty)} y\, \widehat{\Psi}(\mathrm{d}s, \mathrm{d}y) + \lambda \int_0^t \sum_{k=1}^m p_k(s)\, \mu_k \,\mathrm{d}s, \quad t \geq 0.$$

and satisfies

$$\mathbb{E}[S_t] = \lambda \sum_{k=1}^m \pi_\vartheta(k)\, \mu_k\, t, \quad t \geq 0.$$

Proof. An easy verification gives the stated representation of S by combining (3.7) and (6.4). By Corollary 2.98, the process $(\eta_t)_{t\geq 0}$ defined by

$$\eta_t := \int_0^t \int_{(0,\infty)} y\, \widehat{\Psi}(\mathrm{d}s, \mathrm{d}y), \quad t \geq 0,$$

is an $\mathfrak{F}^\Psi$-martingale since

$$\mathbb{E}\Big[\int_0^t \int_{(0,\infty)} y \sum_{k=1}^m p_k(s)\, f_k(y)\,\mathrm{d}y\,\mathrm{d}s\Big] = \mathbb{E}\Big[\int_0^t \sum_{k=1}^m p_k(s)\, \mu_k \,\mathrm{d}s\Big] = \lambda \sum_{k=1}^m \int_0^t \mathbb{E}[p_k(s)]\, \mu_k \,\mathrm{d}s$$
$$= \lambda \sum_{k=1}^m \pi_\vartheta(k)\, \mu_k\, t < \lambda\, \mu_m\, t < \infty, \quad t \geq 0,$$

which also provides the announced expected value. □

Using the proposition, we obtain the following representation of the surplus process, which is indistinguishable from the one given in (3.7) (with $d = 1$):

$$\mathrm{d}X_t^{\xi,b} = \Big(rX_s^{\xi,b} + (\mu - r)\xi_s + c(b_s) - \lambda\, b_t \sum_{k=1}^m p_k(s)\mu_k\Big)\mathrm{d}t + \xi_s\sigma\mathrm{d}W_s - \int\limits_{(0,\infty)} b_t\, y\, \widehat{\Psi}(\mathrm{d}t, \mathrm{d}y)$$

for all $t \geq 0$. Notice further that

$$\begin{aligned}
&\int_0^t \int_{(0,\infty)} \Big(\frac{p_j(s-)\, f_j(y)}{\sum_{k=1}^m p_k(s-) f_k(y)} - p_j(s-)\Big)\nu(s, \mathrm{d}y)\,\mathrm{d}s \\
&= \lambda \int_0^t p_j(s) \int_{(0,\infty)} \Big(f_j(y) - \sum_{k=1}^m p_k(s) f_k(y)\Big)\mathrm{d}y\,\mathrm{d}s \\
&= \lambda \int_0^t p_j(s) \Big(1 - \sum_{k=1}^m p_k(s)\Big)\,\mathrm{d}s = 0,
\end{aligned}$$

since $\sum_{k=1}^m p_k(s) = 1$ for all $s \geq 0$. Hence

$$\mathrm{d}p_j(t) = \int_{(0,\infty)} \Big(\frac{p_j(s-)\, f_j(y)}{\sum_{k=1}^m p_k(s-)\, f_k(y)} - p_j(t-)\Big)\widehat{\Psi}(\mathrm{d}s, \mathrm{d}y), \quad t \geq 0.$$

This result can also be obtained by directly applying the filter result for marked point process observations from Theorem 2.101.

6.3 The reduced control problem

In the reduced control model we incorporate accident realization by extending the state process with the filter process $(p_t)_{t\geq 0}$. Thus the state process of the reduced control problem is the $(m+1)$-dimensional process

$$(X_s^{\xi,b}, p_s)_{s\in[t,T]}$$

for some fixed initial time $t \in [0,T)$ and $(\xi,b) \in \mathcal{U}[t,T]$, where

$$\begin{aligned}
\mathrm{d}X_s^{\xi,b} &= \left(rX_s^{\xi,b} + (\mu - r)\xi_s + c(b_s) - \lambda\, b_s \sum_{k=1}^m p_k(s)\, \mu_k \right) \mathrm{d}s + \xi_s \sigma\, \mathrm{d}W_s \\
&\quad - \int_{(0,\infty)} b_s\, y\, \widehat{\Psi}(\mathrm{d}s, \mathrm{d}y), \\
\mathrm{d}p_j(s) &= \left(\frac{f_j(y)\, p_j(s-)}{\sum_{k=1}^m f_k(y)\, p_k(s-)} - p_j(s-) \right) \widehat{\Psi}(\mathrm{d}s, \mathrm{d}y), \quad j = 1, \ldots, m,
\end{aligned}$$

for $s \in [t,T]$, with

$$\big(X_t^{\xi,b}, p_t\big) = (x,p), \quad x \in \mathbb{R}, \quad p = (p_1, \ldots, p_m) \in \Delta_m.$$

Using this reduced model, we can formulate the reduced control problem. For any $(\xi,b) \in \mathcal{U}[t,T]$, the *objective function* is given by

$$V^{\xi,b}(t,x,p) := \mathbb{E}^{t,x,p}\big[U(X_T^{\xi,b})\big] := \mathbb{E}\big[U(X_T^{\xi,b}) \,|\, X_t^{\xi,b} = x, p_t = p\big],$$

and the *value function* is defined by

$$V(t,x,p) := \sup_{(\xi,b)\in\mathcal{U}[t,T]} V^{\xi,b}(t,x,p), \tag{P3}$$

for all $(t,x,p) \in [0,T]\times\mathbb{R}\times\Delta_m$. As before, an investment-reinsurance strategy $(\xi^\star, b^\star) \in \mathcal{U}[t,T]$ is optimal if

$$V(t,x,p) = V^{\xi^\star,b^\star}(t,x,p),$$

and the insurer is interested in optimal strategies $(\xi^\star, b^\star) \in \mathcal{U}[t,T]$, i.e. in strategies

$$(\xi^\star, b^\star) = \underset{(\xi,b)\in\mathcal{U}[t,T]}{\operatorname{argsup}}\; V^{\xi,b}(t,x,p).$$

The same line of arguments as in Section 4.3 yields, for any $(\xi,b) \in \mathcal{U}[t,T]$,

$$V^{\xi,b}(t,x,p_t) = \widetilde{V}^{\xi,b}(t,x) \quad \text{and thus} \quad V(t,x,p_t) = \widetilde{V}(t,x), \quad (t,x) \in [0,T]\times\mathbb{R}.$$

That is, solving the reduced control problem (P3) gives a solution of the original problem (P) under the framework given in Section 6.1. The solution approach runs as before with the generalized Hamilton-Jacobi-Bellman equation. Before we get into that, it is a good point to briefly discuss an extension of the assumptions of this chapter to those in the previous one.

Remark 6.10. Suppose the framework of Chapter 5 is in force (cf. Sec. 5.1) and the claim sizes of every insurance line are independent. Then the setting of Chapter 5 can be readily extended to one with unobservable claim size distribution as presented in this chapter. That is, we have for every LoB a finite number of possible loss distributions, namely $m_i \in \mathbb{N}$ for the LoB $i \in \{1, \dots, d\}$. Hence the state process of the reduced control model has $(\ell + m + 1 + \sum_{i=1}^{d} m_i)$ components. Accordingly, the generalized HJB equation and the optimal reinsurance strategy derived from it become very complex while the verification procedure is still analogue to the one in the previous and this section.

6.4 The Hamilton-Jacobi-Bellman equation

The heuristic development of the generalized Hamilton-Jacobi-Bellman equation is similar to Section 4.4 and is therefore omitted. Using sufficient assumptions, we obtain

$$\begin{aligned} 0 = \sup_{(\xi,b)\in\mathbb{R}\times[0,1]} \bigg\{ & V_t(t,x,p) - \lambda V(t,x,p) + \frac{1}{2}\sigma^2 V_{xx}(t,x,p)\xi^2 \\ & + V_x(t,x,p)\big(rx + (\mu - r)\xi + c(b)\big) + \lambda \sum_{k=1}^{m} p_k \int_{(0,\infty)} V\Big(t, x - b\,y, J(p,y)\Big) f_k(y)\mathrm{d}y \bigg\}. \end{aligned} \tag{6.5}$$

By applying the arguments from the proof of Lemma 4.24, we obtain for any $(t,x,p) \in [0,T] \times \mathbb{R} \times \Delta_m$

$$V(t,x,p) = -e^{-\alpha x e^{r(T-t)}} g(t,p) \tag{6.6}$$

with

$$g(t,p) := \inf_{(\xi,b)\in\mathcal{U}[t,T]} g^{\xi,b}(t,p), \tag{6.7}$$

where

$$\begin{aligned} g^{\xi,b}(t,p) := \mathbb{E}^{t,p}\bigg[\exp\bigg\{ & - \int_t^T \alpha\, e^{r(T-s)} \big((\mu - r)\,\xi_s + c(b_s)\big)\,\mathrm{d}s \\ & - \int_t^T \alpha\,\sigma\, e^{r(T-s)} \xi_s \,\mathrm{d}W_s + \int_t^T \int_{(0,\infty)} \alpha\, b_s\, y\, e^{r(T-s)}\, \Psi(\mathrm{d}s,\mathrm{d}y) \bigg\}\bigg], \end{aligned} \tag{6.8}$$

where $\mathbb{E}^{t,p}$ denotes the conditional expectation given $p_t = p$. We reformulate the Equation (6.5) by using the separation approach above. First, useful properties of the introduced function g should be mentioned.

Lemma 6.11. *Let g be defined by* (6.7)*. Then the following statements are satisfied:*

(i) $g^{\xi,b}(t,p) > 0$ *for all* $t \in [0,T]$, $p \in \Delta_m$ *and* $(\xi,b) \in \mathcal{U}[t,T]$.

(ii) g *is bounded on* $[0,T] \times \Delta_m$.

(iii) $g^{\xi,b}(t,p) = \sum_{j=1}^{m} p_j\, g^{\xi,b}(t,e_j)$ *for all* $t \in [0,T]$ *and* $p \in \Delta_m$.

(iv) $g^{\xi,b}(t,J(p,y)) = \sum_{j=1}^{m} \frac{f_j(y)\,p_j}{\sum_{k=1}^{m} f_k(y)\,p_k} g^{\xi,b}(t,e_j)$ *for all* $(t,p,y) \in [0,T] \times \Delta_m \times (0,\infty)$.

(v) $\Delta_m \ni p \mapsto g(t,p)$ *is concave for all* $t \in [0,T]$.

Proof. A passage similar to the proof of Lemma 4.25 implies the statements. □

The separation approach (6.6) implies

$$\begin{aligned} V_t(t,x,p) &= -e^{-\alpha x e^{r(T-t)}} \big(\alpha\, x\, r\, e^{r(T-t)} g(t,p) + g_t(t,p)\big), \\ V_x(t,x,p) &= -e^{-\alpha x e^{r(T-t)}} \big(- \alpha\, e^{r(T-t)} g(t,p)\big), \\ V_{xx}(t,x,p) &= -e^{-\alpha x e^{r(T-t)}}\, \alpha^2\, e^{2r(T-t)} g(t,p), \\ V(t,x-b\,y,p) &= -e^{-\alpha x e^{r(T-t)}} \exp\big\{\alpha\, b\, y\, e^{r(T-t)}\big\}\, g(t,p), \end{aligned}$$

where the partial derivative w.r.t. t is only defined on $(0,T)$. However, since g is probably not differentiable w.r.t. t due ot the jumps of the state process, we replace the partial derivative g_t by Clarke's generalized subdifferential again. Recall further the notation g_p introduced on page 68.

Using the relations stated above as well as the generalized subdifferential $\partial^C g_p$ instead of g_t, we conclude from (6.5)

$$\begin{aligned} 0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \bigg\{ &- \lambda\, g(t,p) - \alpha e^{r(T-t)} g(t,p) \Big((\mu - r)\xi + c(b) - \frac{1}{2}\sigma^2 \alpha e^{r(T-t)} \xi^2 \Big) \\ &+ \lambda \sum_{k=1}^{m} p_k \int_{(0,\infty)} g(t, J(p,y)) \exp\big\{\alpha\, b\, y\, e^{r(T-t)}\big\}\, f_k(y)\,\mathrm{d}y \bigg\} \\ + \inf_{\varphi\in\partial^C g_p(t)} \{\varphi\} & \end{aligned}$$

for all $(t,p) \in [0,T) \times \Delta_m$ with boundary condition $g(T,p) = 1$ for all $p \in \Delta_m$.

Notation. Throughout this chapter, let $\mathscr{L}$ denote an operator acting on functions $g : [0,T] \times \Delta_m \to (0,\infty)$ and $(\xi,b) \in \mathbb{R} \times [0,1]$, which is defined by

$$\begin{aligned} \mathscr{L} g(t,p;\xi,b) &:= -\lambda\, g(t,p) - \alpha\, e^{r(T-t)} g(t,p) \Big((\mu - r)\xi + c(b) \\ &- \frac{1}{2}\alpha\, \sigma^2\, e^{r(T-t)} \xi^2 \Big) + \lambda \sum_{k=1}^{m} p_k \int_{(0,\infty)} g(t, J(p,y)) \exp\big\{\alpha\, b\, y\, e^{r(T-t)}\big\}\, f_k(y)\,\mathrm{d}y. \end{aligned} \tag{6.9}$$

Using this notation the generalized HJB equation for g is given by

$$0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \big\{\mathscr{L} g(t,p;\xi,b)\big\} + \inf_{\varphi\in\partial^C g_p(t)} \{\varphi\} \tag{6.10}$$

for all $(t,p) \in [0,T) \times \Delta_m$ with boundary condition

$$g(T,p) = 1, \quad p \in \Delta_m. \tag{6.11}$$

Note that we set $\partial^C g_p(t) = \{g_p'(t)\}$ at the points t, where the derivative exists. In the next section we continue to determine a candidate for an optimal strategy.

6.5 Candidate for an optimal strategy

We obtain a candidate for an optimal strategy as byproduct of the generalized HJB equation (6.10) by rewriting this equation as follows:

$$0 = -\lambda g(t,p) + \alpha e^{r(T-t)} g(t,p) \inf_{\xi\in\mathbb{R}} f_1(t,\xi) + \inf_{b\in[0,1]} f_2(t,p,b) + \inf_{\varphi\in\partial^C g_p(t)} \{\varphi\}, \tag{6.12}$$

where f_1 is defined by (4.27) and

$$f_2(t,p,q) := -\alpha\, e^{r(T-t)}\, g(t,p)\,(\eta-\theta)\,\kappa - \alpha\, e^{r(T-t)}\, g(t,p)\,(1+\theta)\,\kappa\, b \\ - \lambda \sum_{k=1}^{m} p_k \int_{(0,\infty)} g(t,J(p,y))\, \exp\left\{\alpha\, b\, y\, e^{r(T-t)}\right\} f_k(y)\,\mathrm{d}y.$$

Notice that we have used the reinsurance premium model given in (3.6). Hence, it follows from Section 4.5 that the unique candidate of an optimal investment strategy $\xi^\star = (\xi^\star(t))_{t\in[0,T]}$ is given by

$$\xi^\star(t) = \frac{\mu-r}{\sigma^2}\frac{1}{\alpha}e^{-r(T-t)}, \quad t\in[0,T]. \tag{6.13}$$

The first order condition for the candidate of an optimal reinsurance strategy is provided by the next result.

Lemma 6.12. *For any $(t,p)\in[0,T]\times\Delta_m$, the function $\mathbb{R}\ni b\mapsto f_2(t,p,b)$ is strictly convex and*

$$\frac{\partial}{\partial b} f_2(t,p,b) = -\alpha\, e^{r(T-t)}\Bigg(g(t,p)\,(1+\theta)\,\kappa \\ - \lambda\sum_{k=1}^{m} p_k \int_{(0,\infty)} g(t,J(p,y))\, y\, \exp\left\{\alpha\, b\, y\, e^{r(T-t)}\right\} f_k(y)\,\mathrm{d}y\Bigg).$$

Proof. The lemma can be proven with the same arguments as Lemma 4.27. □

The lemma yields a criterion for a candidate of an optimal reinsurance strategy as well as the uniqueness of the candidate. We express the criterion by using the following notation, in which we suppose that g is positive (Lemma 6.11 (i) yields only the nonnegativity of g) throughout this section.

Notation. For any $(t,p)\in[0,T]\times\Delta_m$ and $b\in\mathbb{R}$, we define

$$h_\lambda(t,p,b) := \lambda\sum_{k=1}^{m} p_k \int_{(0,\infty)} \frac{g(t,J(p,y))}{g(t,p)}\, y\, \exp\left\{\alpha\, b\, y\, e^{r(T-t)}\right\} f_k(y)\,\mathrm{d}y. \tag{6.14}$$

Furthermore, we set

$$\begin{aligned} A_\lambda(t,p) &:= h_\lambda(t,p,0),\\ B_\lambda(t,p) &:= h_\lambda(t,p,1). \end{aligned} \tag{6.15}$$

Remark 6.13. This remark is devoted to an alternative reinsurance premium model similar to Remark 4.28. According to Proposition 6.9, we have

$$\mathbb{E}[\mathrm{d}S_t] = \lambda\sum_{k=1}^{m}\pi_\vartheta(k)\,\mu_k, \quad t\geq 0,$$

which is a reasonable choice of κ. As already discussed in Remark 4.28, it makes sense to replace the prior estimator Π_ϑ with the posterior estimator $p_k(t)$, which takes the available information into account. Therefore, it makes sense to modify κ so that it depends on the filter $(p_t)_{t\geq 0}$ as follows:

$$\kappa_\lambda(p) = \lambda \sum_{k=1}^{m} p_k \, \mu_k.$$

In this case it can be shown (similar as Theorem 6.23) that $A_\lambda(t,p)/\kappa_\lambda(p) \geq 1$ for all $(t,p) \in [0,T] \times \Delta_m$ and that the optimal strategy is given by (6.17) with κ replaced by $\kappa_\lambda(p)$ under Assumption 6.18. In addition, a comparison result in analogy to Corollary 6.24 holds.

We continue the discussion with a constant κ and obtain the following first order condition for the optimal reinsurance strategy by setting $\frac{\partial}{\partial b} f_2$ to zero (cf. Lemma 6.12):

$$(1+\theta)\,\kappa = h_\lambda(t,p,b). \tag{6.16}$$

By establishing this equation w.r.t. b, we obtain a minimizer of f_2 w.r.t. b which is unique due to the strict convexity of f_2 w.r.t. b (if such a minimizer exists). The first order condition is solvable and the solution takes values in $[0,1]$ depending on the safety loading parameter θ of the reinsurer.

Proposition 6.14. *For any $(t,p) \in [0,T] \times \Delta_m$, Equation (6.16) has a unique root, denoted by $r_\lambda(t,p)$, which is increasing w.r.t. the safety loading parameter θ. Moreover, it holds,*

(i) $r_\lambda(t,p) \leq 0$ *if* $\theta \leq A_\lambda(t,p)/\kappa - 1$,

(ii) $0 < r_\lambda(t,p) < 1$ *if* $A_\lambda(t,p)/\kappa - 1 < \theta < B_\lambda(t,p)/\kappa - 1$,

(iii) $r_\lambda(t,p) \geq 1$ *if* $\theta \geq B_\lambda(t,p)/\kappa - 1$.

Proof. This follows by the same method as in the proof of Proposition 4.29. □

Notation. In this chapter, $r_\lambda(t,p)$ denotes the unique root from Proposition 6.14.

There is a possibility that cases (i) and (ii) are empty sets, which can not occur with the modified $\kappa_\lambda(p)$ described in Remark 6.13. However, using the constant κ, Proposition 6.14 implies

$$b_\lambda(t,p) := \begin{cases} 0, & \theta \leq A_\lambda(t,p)/\kappa - 1, \\ 1, & \theta \geq B_\lambda(t,p)/\kappa - 1, \\ r_\lambda(t,p), & \text{otherwise}, \end{cases} \tag{6.17}$$

for every $(t,p) \in [0,T] \times \Delta_m$, the candidate for an optimal reinsurance strategy $(b^\star_\lambda(t))_{t\in[0,T]}$ is given by $b^\star_\lambda(t) := b_\lambda(t-, p_{t-})$. It is worth noting that the interpretation about the optimal reinsurance strategy given in Remark 4.30 applies here as well.

6.6 Verification

This section shows that the solution of the generalized HJB equation, from which the optimal strategy is derived, is the determining equation for the value function (Theorem 6.15). Furthermore, we focus on the existence of a solution for the generalized HJB equation (Theorem 6.17) and thus on the optimality of the given candidates for an optimal investment-reinsurance strategy. We use the same procedure as in the Section 4.7.

6.6.1 The verification theorem

Theorem 6.15. *Suppose there exists a bounded function* $h : [0,T] \times \Delta_m \to (0,\infty)$ *such that* $t \mapsto h(t,p)$ *is Lipschitz on* $[0,T]$ *for all* $p \in \Delta_m$, $p \mapsto h(t,p)$ *is continuous on* Δ_m *for all* $t \in [0,T]$ *and* h *satisfies the generalized HJB equation*

$$0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \{\mathscr{L}h(t,p;\xi,b)\} + \inf_{\varphi\in\partial^C h_p(t)} \{\varphi\},$$

for all $(t,p) \in [0,T) \times \Delta_m$ *with boundary condition*

$$h(T,p) = 1, \quad p \in \Delta_m.$$

Then

$$V(t,x,p) = -e^{-\alpha x e^{r(T-t)}} h(t,p), \quad (t,x,p) \in [0,T] \times \mathbb{R} \times \Delta_m,$$

and $(\xi^\star, b^\star_\lambda) = (\xi^\star(s), b^\star_\lambda(s))_{s\in[t,T]}$ *with* $\xi^\star(s)$ *given by* (6.13) *and* $b^\star_\lambda(s) := b_\lambda(s-, p_{s-})$ *given by* (6.17) *(with* g *replaced by* h *in* $A_\lambda(s,p)$ *and* $B_\lambda(s,p)$*) is an optimal feedback strategy for the given optimization problem* (P3)*, i.e.* $V(t,x,p) = V^{\xi^\star,b^\star_\lambda}(t,x,p)$.

For the proof of the verification theorem a measure change is applied, which is introduced in Lemma A.17. To do this, the set of admissible strategies must be constrained as follows.

Notation. Throughout this chapter, we set for any $t \in [0,T)$

$$\begin{aligned} \widetilde{\mathcal{U}}[t,T] := \{&(\xi,b) \in \mathcal{U}[t,T] :\ \exists\, K > 0 : |\xi_s| \le K\ \forall\, s \in [t,T], \\ &\xi = (\xi_s)_{s\in[t,T]} \text{ is continuous and } \mathfrak{F}^W\text{-adapted}, b = (b_s)_{s\in[t,T]} \text{ is } \mathfrak{F}^\Psi\text{-predictable}\}. \end{aligned} \tag{6.18}$$

Moreover, we set

$$\widetilde{V}(t,x,p) := \sup_{(\xi,b)\in\widetilde{\mathcal{U}}[t,T]} V^{\xi,b}(t,x,p), \quad (t,x,p) \in [0,T] \times \mathbb{R} \times \Delta_m. \tag{6.19}$$

Notice that Lemma A.22 is crucial to prove the verification theorem, which makes use of the following operator.

Notation. We define an operator $\mathscr{H}$ acting on functions $v : [0,T] \times \Delta_m \to (0,\infty)$ and $(\xi,b) \in \mathbb{R} \times [0,1]$ by

$$\mathscr{H}v(t,p;\xi,b) := \mathscr{L}v(t,p;\xi,b) + v_t(t,p) \tag{6.20}$$

for all functions $v : [0,T] \times \Delta_m \to (0,\infty)$, where the right-hand side is well-defined.

Using this notation, the HJB equation (6.10) can be written as

$$0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \{\mathscr{H}g(t,p;\xi,b)\}$$

at those points t, where g is differentiable w.r.t. t.

Proof of Theorem 6.15. Using the Lemmata A.20, A.21 and A.22, analysis similar to the proof of Theorem 4.31 yields the announced assertion. □

6.6.2 Existence result for the value function

We demonstrate below that there exists a function $h : [0,T] \times \Delta_m \to (0,\infty)$ which fulfils the conditions stated in Theorem 6.15. In order to do this, we introduce the following function.

Notation. We set

$$\widetilde{g}(t,p) := \inf_{(\xi,b)\in\widetilde{\mathcal{U}}[t,T]} g^{\xi,b}(t,p), \quad (t,p) \in [0,T] \times \Delta_m, \tag{6.21}$$

where $g^{\xi,b}$ is given by (6.8) and $\widetilde{\mathcal{U}}[t,T]$ by (6.18).

Properties of the function introduced above are given below.

Lemma 6.16. *The function $\widetilde{g}$ defined by* (6.21) *has the following properties:*

(i) $\widetilde{g}(t,p) > 0$ *for all* $(t,p) \in [0,T] \times \Delta_m$.

(ii) $\widetilde{\mathcal{U}}[0,T] \ni (\xi,b) \mapsto g^{\xi,b}(0,p)$ *is bounded for all* $p \in \Delta_m$.

(iii) There exists a constant $0 < K_3 < \infty$ *such that* $|\widetilde{g}(t,p)| \le K_3$ *for all* $(t,p) \in [0,T] \times \Delta_m$.

(iv) $\Delta_m \ni p \mapsto \widetilde{g}(t,p)$ *is concave for all* $t \in [0,T]$.

(v) $[0,T] \ni t \mapsto \widetilde{g}(t,p)$ *is Lipschitz on* $[0,T]$ *for all* $p \in \Delta_m$.

(vi) Let M be the set of all points $(t,p) \in [0,T] \times \Delta_m$, where the partial derivative of $\widetilde{g}$ w.r.t. t exists. Then there exists a constant $0 < K_4 < \infty$ such that $|\widetilde{g}_t(t,p)| \le K_4$ for all $(t,p) \in M$.

(vii) There exists a constant $0 < K_5 < \infty$ such that $\left|\mathscr{L}\widetilde{g}(t,p;\xi,b)\right| \le K_5$ for all $(t,p) \in [0,T] \times \Delta_m$ and $(\xi,b) \in [-K,K] \times [0,1]$.

(viii) There exists a constant $0 < K_6 < \infty$ such that $\left|\inf_{(\xi,b)\in[-K,K]\times[0,1]} \mathscr{L}\widetilde{g}(t,p;\xi,b)\right| \le K_6$ for all $(t,p) \in [0,T] \times \Delta_m$.

Proof. The proof runs in the same manner as the proof of Lemma 4.32. □

Now we can state the result of the existence of the HJB equation, whose proof works analogously to Theorem 4.33 in Section 4.7.2 by aid of Lemma 6.16.

Theorem 6.17. *The value function of our investment-reinsurance problem* (P3) *is given by*

$$V(t,x,p) = -e^{-\alpha x e^{r(T-t)}} g(t,p),$$

where g is defined by (4.19) *and satisfies the generalized HJB equation*

$$0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \{\mathscr{L}g(t,p;\xi,b)\} + \inf_{\varphi\in\partial^C g_p(t)} \{\varphi\}, \quad (t,p) \in [0,T) \times \Delta_m,$$

with boundary condition $g(T,p) = 1$ for all $p \in \Delta_m$. Furthermore, $(\xi^\star, b^\star_\lambda) = (\xi^\star(s), b^\star_\lambda(s))_{s\in[t,T]}$ with $\xi^\star(s)$ given by (6.13) *and $b^\star_\lambda(s) = b_\lambda(s-, p_{s-})$, where b_λ is given by* (6.17)*, is an optimal investment-reinsurance strategy for the optimization problem* (P3).

Proof. Applying the arguments from the proof of Theorem 4.33 again, we obtain the announced assertion. □

6.7 Comparison results with the complete information case

For the considered setting with an unknown claim size distribution, a comparison of the optimal reinsurance strategy given in Theorem 6.17 and the one in the complete information case given in (4.57) is derived as in the previous chapters. In addition, we calculate bounds for the optimal strategy, which are independent of the filter and thus of the observed claim amounts. Both the comparison result and the a priori bounds can be only applied under the following assumption.

Assumption 6.18. Throughout this section, we assume that

$$F_1(x) \geq F_2(x) \geq \ldots \geq F_m(x)$$

for all $x \in \mathbb{R}$.

Remark 6.19. A useful equivalent formulation of the assumption above is

$$\int_{(0,\infty)} g(x)\, F_1(\mathrm{d}x) \leq \ldots \leq \int_{(0,\infty)} g(x)\, F_m(\mathrm{d}x)$$

for all increasing functions g, for which both expectations exist, compare Müller and Stoyan [96, Thm. 1.2.8]. Hence Assumption 6.2 about the order of the means of the potential claim size distributions fits with Assumption 6.18 by choosing g as identity above.

Another way of stating Assumption 6.18 is to say that the claim sizes given ϑ are stochastically ordered:

$$Y \mid \vartheta = 1 \preceq_{\mathrm{st}} Y \mid \vartheta = 2 \preceq_{\mathrm{st}} \ldots \preceq_{\mathrm{st}} Y \mid \vartheta = m,$$

where $\preceq_{\mathrm{st}}$ denotes the usual stochastic order. It is readily that this order is an order from the best to the worst case scenario from the perspective of the insurer. The following notations and results are prerequisites for the representation of the announced a priori bounds

Notation. Let $t \in [0,T]$ and $b \in \mathbb{R}$. Throughout this section, we set

$$h_\lambda^{\min}(t,b) := \lambda \int_{(0,\infty)} y \left\{\alpha\, b\, y\, e^{r(T-t)}\right\} f_1(y)\, \mathrm{d}y,$$
$$h_\lambda^{\max}(t,b) := \lambda \int_{(0,\infty)} y \left\{\alpha\, b\, y\, e^{r(T-t)}\right\} f_m(y)\, \mathrm{d}y.$$

Proposition 6.20. *Let $t \in [0,T]$. Then $\mathbb{R} \ni b \mapsto h_\lambda^{\min}(t,b)$ and $\mathbb{R} \ni b \mapsto h_\lambda^{\max}(t,b)$ are strictly increasing and strictly convex. Furthermore, it holds*

$$\lim_{b\to-\infty} h_\lambda^{\min}(t,b) = \lim_{b\to-\infty} h_\lambda^{\max}(t,b) = 0, \quad \lim_{b\to\infty} h_\lambda^{\min}(t,b) = \lim_{b\to\infty} h_\lambda^{\max}(t,b) = \infty.$$

Proof. This follows by the same analysis as in the proof of Proposition 4.29. □

The proposition justifies the next notation, where we refer again to the proof of Proposition 4.29 for details.

Notation. For some fixed $t \in [0,T]$, we denote the unique root of the equation $(1+\theta)\,\kappa = h_\lambda^{\min}(t,b)$ w.r.t. b, and the unique root of the equation $(1+\theta)\,\kappa = h_\lambda^{\max}(t,b)$ w.r.t. b by $r_\lambda^{\min}(t)$ and $r_\lambda^{\max}(t)$, respectively.

The announced a priori bounds are a direct consequence of the following theorem in connection with Proposition 6.20.

Proposition 6.21. *For any* $(t,p) \in [0,T] \times \Delta_m$ *and* $b \in \mathbb{R}^+$, *we have*

$$h_\lambda^{\min}(t,b) \le h_\lambda(t,p,b) \le h_\lambda^{\max}(t,b).$$

Proof. Choose some $(t,p) \in [0,T] \times \Delta_m$ and $\bar b \in \mathbb{R}^+$. Due to the increasing property of $(0,\infty) \ni y \mapsto y \exp\{\alpha \bar b\, y\, e^{r(T-t)}\}$, it follows from Remark 6.19 that

$$\int_{(0,\infty)} y \exp\left\{\alpha \bar b\, y\, e^{r(T-t)}\right\} f_1(y)\,\mathrm{d}y \le \ldots \le \int_{(0,\infty)} y \exp\left\{\alpha \bar b\, y\, e^{r(T-t)}\right\} f_m(y)\,\mathrm{d}y. \quad (6.22)$$

Taking this order as well as Lemma 6.11 (iii), (iv) into account, we get

$$\begin{aligned}
&\lambda \sum_{k=1}^m p_k \int_{(0,\infty)} g^{\xi,b}(t, J(p,y))\, y \exp\left\{\alpha \bar b\, y\, e^{r(T-t)}\right\} f_k(y)\,\mathrm{d}y \\
&= \lambda \sum_{k=1}^m p_k \int_{(0,\infty)} \sum_{j=1}^m \frac{p_j\, f_j(y)}{\sum_{\ell=1}^m p_\ell\, f_\ell(y)} g^{\xi,b}(t,e_j)\, y \exp\left\{\alpha \bar b\, y\, e^{r(T-t)}\right\} f_k(y)\,\mathrm{d}y \\
&= \lambda \sum_{j=1}^m p_j\, g^{\xi,b}(t,e_j) \int_{(0,\infty)} \frac{\sum_{k=1}^m p_k\, f_k(y)}{\sum_{\ell=1}^m p_\ell\, f_\ell(y)}\, y \exp\left\{\alpha \bar b\, y\, e^{r(T-t)}\right\} f_j(y)\,\mathrm{d}y \\
&= \lambda \sum_{j=1}^m p_j\, g^{\xi,b}(t,e_j) \int_{(0,\infty)} y \exp\left\{\alpha \bar b\, y\, e^{r(T-t)}\right\} f_j(y)\,\mathrm{d}y \\
&\le h_\lambda^{\max}(t,\bar b) \sum_{j=1}^m p_j\, g^{\xi,b}(t,e_j) = h_\lambda^{\max}(t,\bar b)\, g^{\xi,b}(t,p)
\end{aligned}$$

for every $(\xi,b) \in \mathcal{U}[t,T]$, which yields $h_\lambda(t,p,\bar b) \le h_\lambda^{\max}(t,\bar b)$ by taking the infimum over all $(\xi,b) \in \mathcal{U}[t,T]$ on both sides. The other inequality is obtained by an analogue argumentation. □

Corollary 6.22. *The optimal reinsurance strategy* $b_\lambda^\star = (b_\lambda^\star(t))_{t\in[0,T]}$ *from Theorem 6.17 has the following boundaries:*

$$\max\{0, r_\lambda^{\max}(t)\} \le b_\lambda^\star(t) \le \min\{1, r_\lambda^{\min}(t)\}, \quad t \in [0,T].$$

The range of optimality for the reinsurance strategy provided by the corollary above increases in m (the number of possible claim size distributions) because of the definition of $h_\lambda^{\min}$ and $h_\lambda^{\max}$. In the case of three potential loss distributions, chosen as stated Section 6.7 (as well as other parameters), we can see in Figure 6.2 that only the lower bound (orange line) is useful; the upper one (red line) yields a trivial bound. Due to the computability of these bounds at time zero for the entire time horizon, they provide the insurer the a priori information that it is never optimal to transfer the entire risk to the reinsurer (retention level zero). One reason for the (exponential) rise of the lower bound from approximately 0.2 to 0.6 lies in the choice of the exponential utility function: As the level of the surplus process rises, a loss is valued less strongly, which leads to a more risky behaviour of the insurer when the surplus rises over time.

A tighter upper bound for the optimal strategy is provided by the comparison result of the optimal reinsurance strategy under partial information and full information, which is an immediate consequence of the next theorem.

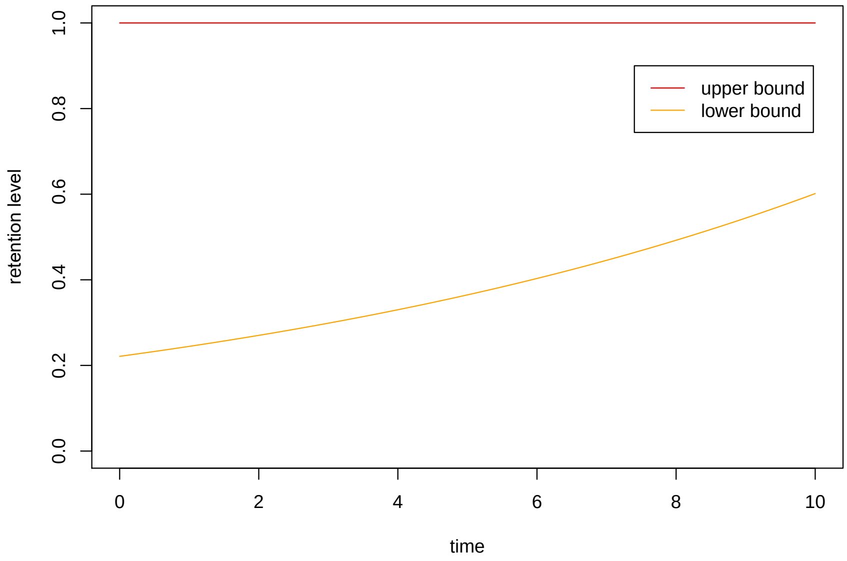

Figure 6.2: The a priori upper bound (red line) and lower bound (orange line) for the optimal reinsurance strategy for the setting from Section 6.8.

Theorem 6.23. *Let b_λ be the function given by* (6.17) *and $\tilde{b}^\star_{\lambda,\bar{c},F}$ the function given by* (4.57)*. Then, for any $(t,p) \in [0,T] \times \Delta_m$,*

$$b^\star_\lambda(t,p) \le \tilde{b}^\star_{\lambda,1,\bar{F}_p}(t)$$

with

$$\bar{F}_p(\mathrm{d}y) := \sum_{k=1}^m p_k\, f_k(y)\, \mathrm{d}y.$$

Proof. Let us fix $(t,p) \in [0,T] \times \Delta_m$ and $\bar{b} \in \mathbb{R}^+$. As in the proof of Theorem 4.41 it is sufficient to compare h_λ given by (6.14) and $h_{\lambda,1,\bar{F}_p}$ given by (4.55) due to the first order conditions (6.16) and (4.56). From the proof of Proposition 6.21, we know already that

$$\begin{aligned}&\lambda \sum_{k=1}^m p_k \int_{(0,\infty)} g^{\xi,b}(t, J(p,y))\, y\, \exp\left\{\alpha\, \bar{b}\, y\, e^{r(T-t)}\right\} f_k(y)\, \mathrm{d}y \\ &= \lambda \sum_{j=1}^m p_j\, g^{\xi,b}(t,e_j) \int_{(0,\infty)} y\, \exp\left\{\alpha\, \bar{b}\, y\, e^{r(T-t)}\right\} f_j(y)\, \mathrm{d}y\end{aligned}$$

for every $(\xi,b) \in \mathcal{U}[t,T]$. Remark 6.19 implies that $g^{\xi,b}(t,e_1) \le \ldots \le g^{\xi,b}(t,e_m)$ because the integrand of

$$\begin{aligned} g^{\xi,b}(t,p) = \mathbb{E}^{t,p}\Bigg[\exp\Bigg\{ -\int_t^T \alpha e^{r(T-s)}\big((\mu-r)\xi_s + c(b_s)\big)\, \mathrm{d}s - \int_t^T \alpha\, e^{r(T-s)}\xi\, \mathrm{d}W_s \\ + \sum_{n=1}^{N_{T-t}} \alpha\, b_{T_n} Y_n\, e^{r(T-T_n)}\Bigg\}\Bigg], \end{aligned}$$

is increasing in Y_n, whose distribution is given by F_j under the condition $p = e_j$, $j =$

$1,\dots,m$. By considering this order, the order given in (6.22), Lemma 6.11 (iii) as well as Lemma B.6, we obtain

$$\begin{aligned}
&\lambda \sum_{j=1}^{m} p_j\, g^{\xi,b}(t,e_j) \int_{(0,\infty)} y\, \exp\left\{\alpha\, \bar{b}\, y\, e^{r(T-t)}\right\} f_j(y)\,\mathrm{d}y \\
&\geq \lambda \sum_{j=1}^{m} p_j\, g^{\xi,b}(t,e_j) \sum_{k=1}^{m} p_k \int_{(0,\infty)} y\, \exp\left\{\alpha\, \bar{b}\, y\, e^{r(T-t)}\right\} f_k(y)\,\mathrm{d}y \\
&= \lambda\, g^{\xi,b}(t,p) \int_{(0,\infty)} y\, \exp\left\{\alpha\, \bar{b}\, y\, e^{r(T-t)}\right\} \sum_{k=1}^{m} p_k\, f_k(y)\,\mathrm{d}y
\end{aligned}$$

for every $(\xi,b)\in\mathcal{U}[t,T]$. In summary, we have for any

$$\begin{aligned}
&\lambda \sum_{k=1}^{m} p_k \int_{(0,\infty)} g^{\xi,b}(t,J(p,y))\, y\, \exp\left\{\alpha\, \bar{b}\, y\, e^{r(T-t)}\right\} f_k(y)\,\mathrm{d}y \\
&\geq \lambda\, g^{\xi,b}(t,p) \int_{(0,\infty)} y\, \exp\left\{\alpha\, \bar{b}\, y\, e^{r(T-t)}\right\} \sum_{k=1}^{m} p_k\, f_k(y)\,\mathrm{d}y
\end{aligned}$$

for all $(\xi,b)\in\mathcal{U}[t,T]$, which gives $h_\lambda(t,p,\bar{b}) \geq h_{\lambda,1,\bar{F}}(t,\bar{b})$ by taking the infimum over all $(\xi,b)\in\mathcal{U}[t,T]$ on both sides. □

Corollary 6.24. *Let $\tilde{b}^\star_{\lambda,\bar{c},F}$ be the function given by* (4.57). *Then the optimal reinsurance strategy under partial information $(b^\star_\lambda(t))_{t\in[0,T]}$ from Theorem 6.17 satisfies*

$$b^\star_\lambda(t) \leq \tilde{b}^\star_{\lambda,1,\bar{F}_{p_{t-}}}(t), \quad t\in[0,T].$$

Since $(\tilde{b}^\star_{\lambda,1,\bar{F}_{p_{t-}}}(t))_{t\in[0,T]}$ is $\mathfrak{G}$-predictable, it is an admissible reinsurance strategy. Furthermore, $\bar{F}_{p_{t-}}(\mathrm{d}y)$ can be seen as the known conditional average claim size distribution given the available information strict before time t. Consequently, the comparison result above has the same interpretation as the comparison result in Section 4.8.2 and 5.7; namely, more uncertainty leads to a higher level of protection, i.e. to a lesser or equal retention level. The comparison result will be illustrated in the next section.

6.8 Numerical analyses

The results of the numerical experiments in this section have the purpose to support the analytic results of the optimal reinsurance strategy, in particular of the comparison result from the previous section.

We suppose the $\Theta=\{1,2,3\}$ and f_1, f_2, f_3 are the density functions of right-truncated exponential distributions with rate 3, 2 and 1, respectively, all three truncated at 3. That is, for any $y\in\mathbb{R}$

$$f_1(y)=\frac{3\,e^{-3y}}{1-e^{-9}}\mathbb{1}_{[0,3]}(y),\ f_2(y)=\frac{2\,e^{-2y}}{1-e^{-6}}\mathbb{1}_{[0,3]}(y),\ f_3(y)=\frac{e^{-y}}{1-e^{-3}}\mathbb{1}_{[0,3]}(y),$$

and

$$\mu_1=\frac{1}{3\,(1-e^{-9})},\quad \mu_2=\frac{1}{2\,(1-e^{-6})},\quad \mu_3=\frac{1}{1-e^{-3}}.$$

Due to this setting, we have for any $0 < x \le y \le 3$

$$\frac{f_1(y)}{f_2(y)} \le \frac{f_1(x)}{f_2(x)} \iff \frac{3\,e^{-3y}}{2\,e^{-2y}}\frac{1-e^{-6}}{1-e^{-9}} \le \frac{3\,e^{-3x}}{2\,e^{-2x}}\frac{1-e^{-6}}{1-e^{-9}} \iff e^{-y} \le e^{-x},$$

where the last inequality is obviously satisfied. Therefore $f_1(y)\,f_2(x) \le f_1(x)\,f_2(y)$ for all $x, y \in \mathbb{R}$ with $x \le y$. Hence $Y_1 \,|\, \vartheta = 1 \preceq_{\mathrm{lr}} Y_1 \,|\, \vartheta = 2$, while $\preceq_{\mathrm{lr}}$ denotes the *likelihood ration order*, see e.g. Müller and Stoyan [96, Def. 1.4.1]. Consequently, $Y_1 \,|\, \vartheta = 1 \preceq_{\mathrm{st}} Y_1 \,|\, \vartheta = 2$ since the likelihood ration order implies the stochastic order according to Müller and Stoyan [96, Thm. 1.4.5]. The same conclusion can be drawn for f_2 and f_3, i.e. $Y_1 \,|\, \vartheta = 2 \preceq_{\mathrm{st}} Y_1 \,|\, \vartheta = 3$. Thus Assumption 6.18 is fulfilled. The prior probability mass function of ϑ is supposed to be

$$\bar{\pi}_\vartheta = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right).$$

Further parameters are specified in the Table 6.1. The parameter κ of the premium

parameter	value
x_0	10
T	10
λ	3
r	0.1
μ	0.2
σ	3
α	0.2
θ	0.6
η	0.2

Table 6.1: Simulation parameters for Section 6.8.

principle is choose as $\mathbb{E}[\mathrm{d}S_t]$, i.e.

$$\kappa = \lambda \sum_{k=1}^{3} \pi_\vartheta(k)\,\mu_k = \sum_{k=1}^{3} \mu_k,$$

compare Proposition 6.9.

Now all parameters are fixed and we can visualize the comparison result from Corollary 6.24 graphically. The following simulation results have been generated under the assumption that the realization of ϑ is 2, which means that the underlying loss distribution is F_2. In Figure 6.2 we have already illustrated the a priori bounds for the parameter selection of this section. Figure 6.3 shows these bounds together with two paths (black and blue lines) of the reinsurance strategy $(\tilde{b}^\star_{\lambda,1,\bar{F}_{p_{t-}}}(t))_{t\in[0,T]}$ with $\bar{F}_p(\mathrm{d}y) := \sum_{k=1}^{m} p_k\, f_k(y)\,\mathrm{d}y$, $p \in \Delta_m$, which provide an upper bound of the corresponding optimal reinsurance strategy for each scenario according to Corollary 6.24. That means, the black line and blue line depend on the realized claim arrival times and the corresponding losses. In the scenario of the black line, the insurer receives rarely useful information about the choice of the optimal reinsurance strategy since the path of the strategy serving as upper bound is 1 for almost the entire period. But in the other scenario of the blue line, the range of optimality for the reinsurance strategy (area between the blue line and orange line) in the time between 1 and 6 is very small. In summary, we can conclude that the quality of the upper bound provided by the comparison result depends strongly on the realized loss amounts.

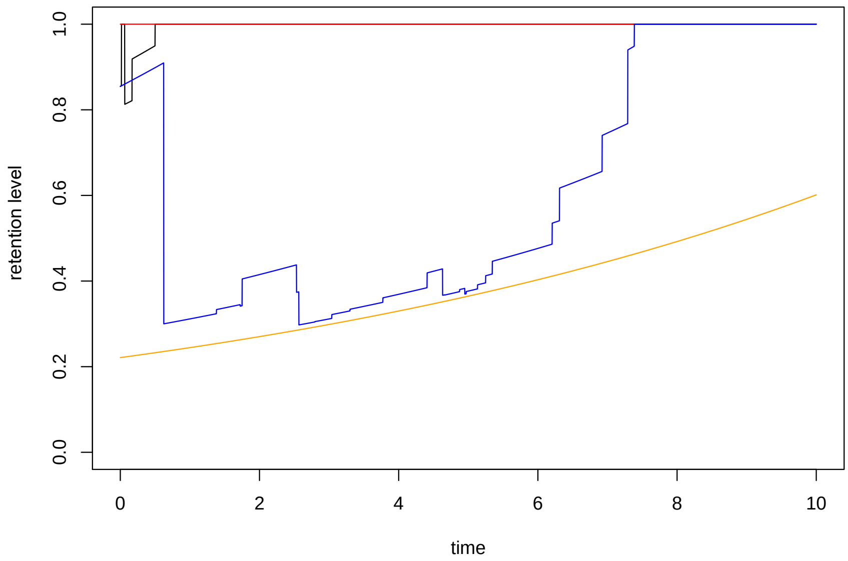

Figure 6.3: The a priori upper bound (red line) and lower bound (orange line) for the optimal reinsurance strategy as well as two trajectories (black and blue lines) of the reinsurance strategy $(\tilde{b}^{\star}_{\lambda,1,\bar{F}_{p_{t-}}}(t))_{t\in[0,T]}$ with $\bar{F}_p(dy) := \sum_{k=1}^{m} p_k\, f_k(y)\, dy$.

We conclude the numerical analysis with Figure 6.4, which shows the path of the surplus process in an insurance loss scenario for three different insurance strategies. The red line displays the trajectory of the surplus process in the case of full reinsurance (i.e. retention level of 0), which tends upwards in contrast to the corresponding paths in Sections 4.9 and 5.8. The reason is that the parameter κ selected by means of the expected value premium principles is smaller than in the other sections under the consideration of a single LoB. For a constant reinsurance strategy of 0.5, the trajectory of the surplus process is plotted by the blue line. Up to time 4, this path is similar to the one in the case of the reinsurance strategy $(\tilde{b}^{\star}_{\lambda,1,\bar{F}_{p_{t-}}}(t))_{t\in[0,T]}$ with $\bar{F}_p(dy) := \sum_{k=1}^{m} p_k\, f_k(y)\, dy$ (black line). Since the later reinsurance strategy tends upwards (cf. Figure 5.3) at the end of the considered time horizon, the (negative) jump sizes become higher in comparison to the blue line, but at the same time the path between the claims rises more strongly because of the lower reinsurance premium.

6.9 Comments on generalizations

The framework of this chapter is one of the simplest conceivable settings for an unobservable claim size distribution. Therefore it is desirable to solve the optimization problem in a more general setting.

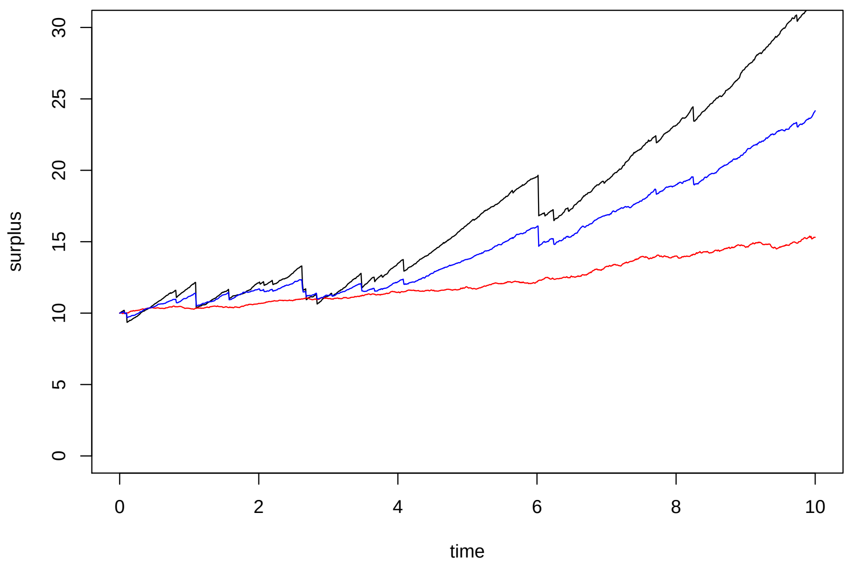

Figure 6.4: Trajectories of the surplus process for an insurance loss scenario in the cases of full reinsurance (red line), constant retention level of 0.5 (blue line) and the reinsurance strategy $(\tilde{b}^{\star}_{\lambda,1,\bar{F}_{p_{t-}}}(t))_{t\in[0,T]}$ with $\bar{F}_p(dy) := \sum_{k=1}^{m} p_k\, f_k(y)\, dy$ (black line).

Finite mixture model. A generalization is to model the unknown distribution as a finite mixture distribution[1] with an unknown allocation. In general, this would lead to an infinite dimension control problem, compare the explanation in the paragraph "More general trigger process" in Section 5.9. Moreover, there are no natural conjugate priors available for finite mixture models with unknown allocation (see Frühwirth-Schnatter [62, p. 53]). Therefore, it is not possible to solve the optimization problem within a finite mixture framework using the methods presented in this paper.

Dirichlet process. Another approach for considering the unknownness of the claim size distribution is to use the Dirichlet process as the model for the loss distribution, which was introduced in the breakthrough paper from Ferguson [60]. If α is some finite measure on $(0,\infty)$, then the Dirichlet process with parameter α, written as $\mathrm{DP}(\alpha)$, chooses a discrete claim size distribution on $(0,\infty)$. However, the Dirichlet process is rich in the sense that there is a positive probability that any fixed distribution, which is absolutely continuous w.r.t. α, is approximate as closely as desired by a sample function of the Dirichlet process, compare Phadia [99, p. 29 f.]. Therefore the Dirichlet process approximates every relevant loss distribution, which justifies the use of the Dirichlet process as a distribution for the claim size distribution. So it is reasonable to assume that $F \mid \alpha \sim \mathrm{DP}(\alpha)$, where α is a finite measure on $(0,\infty)$ and the claim sizes are conditional iid with $F \mid \alpha$. To obtain the reduced control problem, we have to characterize the random claim size distribution by using the observed claim sizes. For this purpose, the

[1] An introduction to finite mixture modelling can be found in Frühwirth-Schnatter [62, Ch. 1].

following conjugated property of the Dirichlet process can be used

$$F \mid \alpha, Y_1, \dots, Y_{N_t} \sim \mathrm{DP}\left(\alpha + \sum_{i=1}^{N_t} \delta_{Y_i}\right),$$

compare Ferguson [60, Thm. 1], where δ_x denotes the Dirac measure at x. Due to the conjugated property, the family of random measures

$$C_t(\mathrm{d}x) := \sum_{i=1}^{N_t} \delta_{Y_i}(\mathrm{d}x) = \int_0^t \int_{(0,\infty)} \mathbb{1}_{\{y \in \mathrm{d}x\}} \Phi(\mathrm{d}s, \mathrm{d}y), \quad t \in [0,T],$$

encapsulates all available information about F, which is included in the observable filtration $\mathfrak{G}$. The random counting measure $C_t(\mathrm{d}x)$ is uniquely determined by

$$\int_{(0,\infty)} f(x)\, C_t(\mathrm{d}x) = \sum_{i=1}^{N_t} f(Y_i) = \int_0^t \int_{(0,\infty)} f(y)\, \Phi(\mathrm{d}t, \mathrm{d}y), \quad t \in [0,T],$$

for all functions $f \in B(0,\infty)$, $B(0,\infty)$ denotes the set of all measurable bounded functions defined on $(0,\infty)$. Therefore the process $\pi(f) = (\pi_t(f))_{t\in[0,T]}$ defined by

$$\pi_t(f) := \int_0^t \int_{(0,\infty)} f(y)\, \Phi(\mathrm{d}t, \mathrm{d}y), \quad t \in [0,T], \quad f \in B(0,\infty),$$

characterises the relevant information about F. Therefore the state process of the reduced control problem is of infinite dimension, which requires a different solution approach than the one presented here.

A parametric Bayesian model. A further concept for the claim size modelling is to choose a parametric Bayesian model with a conjugated prior. Such a model is the exponential distributions with Gamma distributed rate. So $\{F_\vartheta : \vartheta \in \Theta\}$ is a family of conditional exponential distribution given ϑ, where ϑ is an $\mathcal{F}_0$-measurable random variable taking values in $\Theta := (0,\infty)$, which is Gamma distributed. That is, the prior knowledge of the insurer about the unknown rate of the claim size distribution is expressed by a Gamma distribution. It is well-known that the Gamma distribution is a conjugated prior for the rate of an exponential distribution, cf. e.g. DeGroot [49, Thm. 9.4.3]. More precisely, we have the following setting:

$$\begin{aligned}
&Y_1, Y_2, \dots \mid \vartheta \overset{iid}{\sim} Y \\
&Y \mid \vartheta \sim \mathrm{Exp}(\vartheta), \\
&\vartheta \mid \gamma, \zeta \sim \Gamma(\gamma, \zeta), \quad \gamma, \zeta > 0, \\
&\vartheta \mid \gamma, \zeta, Y_1, \dots, Y_{N_t} \sim \Gamma\big(\gamma + N_t, \zeta + q_t\big) \quad \text{with } q_t := \textstyle\sum_{i=1}^{N_t} Y_i.
\end{aligned}$$

The distribution for the claim amounts is an exponential-Gamma mixture model, which can also be interpreted as the average of individual exponential distributed claim sizes, where the heterogeneity of the individual losses is taken into account by the mixing, see Pacáková and Zapletal [97, Sec. III]. However, an easy calculation shows that the unconditional density of Y is

$$f_Y(y) = \int_\Theta f_{Y,\vartheta}(y,\vartheta')\,\mathrm{d}\vartheta' = \int_0^\infty f_{Y|\vartheta=\vartheta'}(y)\,h_{\gamma,\zeta}(\vartheta')\,\mathrm{d}\vartheta', \quad y \in (0,\infty),$$

where $f_{Y|\vartheta=\vartheta'} = f_{\vartheta'}$ is the density of the exponential distribution with parameter ϑ' and $h_{\gamma,\zeta}$ denotes the density of the Gamma distribution with parameter γ and ζ. That is,

$$\begin{aligned} f_Y(y) &= \int_0^\infty \vartheta\, e^{-\vartheta y}\frac{\zeta^\gamma}{\Gamma(\gamma)}\vartheta^{\gamma-1}e^{-\zeta\,\vartheta}\,\mathrm{d}\vartheta = \frac{\gamma^\zeta}{\Gamma(\gamma)}\frac{\Gamma(\gamma+1)}{(y+\zeta)^{\gamma+1}}\int_0^\infty \frac{(y+\zeta)^{\gamma+1}}{\Gamma(\gamma+1)}\vartheta^\gamma\, e^{-(\zeta+y)\vartheta}\,\mathrm{d}\vartheta \\ &= \frac{\gamma\,\zeta^\gamma}{(y+\zeta)^{\gamma+1}}, \quad y \in (0,\infty), \end{aligned}$$

and thus

$$F_Y(y) = 1 - \left(1+\frac{y}{\zeta}\right)^{-\gamma}, \quad y \in (0,\infty),$$

which is the distribution function of the Lomax-distribution with parameters γ and ζ (also known as Pareto (II) distribution with location parameter of zero), compare Kleiber and Kotz [76, Sec. 6.4.2]. That is, the losses are heavy tailed and $\mathbb{E}[Y\exp\{\alpha\, e^{rT}\,Y\}] = \infty$. But the existence of the expectation is necessary for the proof of a change of measure (similar to Lemma A.17), which is an integral part of the verification. In consequence, the solution procedure of this chapter can not be applied to an exponential Gamma-mixture model for the claim sizes. A possible way out is an approximation of this distribution by a right-truncated Lomax-distribution. Since such an approximation no longer solves the original control problem, we must proceed as follows. First of all, it should be noted that the predictive distribution of an insurance loss given the information at disposal at time t is Lomax-distributed with parameter $\gamma+N_t$ and $\zeta+q_t$, which follows by the same calculation as above because of the conjugated property of the Gamma prior. Assuming that the predictive distribution of Y given $\bar{Y}_n = \bar{y}_n$ is

$$\hat{f}_{Y|\bar{Y}_n=\bar{y}_n}(y) = \frac{(\gamma+n)\big((\gamma+\sum_{i=1}^n y_i)\,K + (\gamma+\sum_{i=1}^n y_i)^2\big)^{\gamma+n}}{(y+\gamma+\sum_{i=1}^n y_i)^{\gamma+n+1}\big((\zeta+\sum_{i=1}^n y_i+K)^{\gamma+n} - (\zeta+\sum_{i=1}^n y_i)^{\gamma+n}\big)},$$

for $0 \le y \le K$, where $K > 0$ is some upper bound for the claim sizes. This is the density function of Lomax-distribution with parameter $\gamma+n$ and $\zeta+\sum_{i=1}^n y_i$ right-truncated at K. With the notation $f_{\gamma,\zeta,K}$ as density function of Lomax-distribution with parameter γ and ζ right-truncated at K, the $\mathfrak{F}^\Psi$-intensity kernel of the $(0,K)$-MPP $\Psi = (T_n,Y_n)_{n\in\mathbb{N}}$ is given by $\nu(t,\mathrm{d}y) = \lambda\, f_{\gamma+N_t,\zeta+q_t,K}(y)\,\mathrm{d}y\,\mathrm{d}t$, $t \ge 0$, which follows similar to Proposition 6.8. The processes $(N_t)_{t\ge0}$ and $(q_t)_{t\ge0}$ provides the information at disposal about the claim size distribution. Therefore, the state process of the reduced control problem is the 3-dimensional process $(X_s^{\xi,b}, N_s, q_s)_{s\in[t,T]}$ for some fixed initial time $t \in [0,T)$ and $(\xi,b) \in \mathcal{U}[t,T]$ with

$$\mathrm{d}X_s^{\xi,b} = \Big(rX_s^{\xi,b} + (\mu-r)\xi_s + c(b_s) - \lambda\, b_s\,\mu_s\Big)\mathrm{d}s + \xi_s\sigma\,\mathrm{d}W_s - \int_{(0,K)} b_s\,y\,\widehat{\Psi}(\mathrm{d}s,\mathrm{d}y),$$

where

$$\mu_t := \int_0^K y\, f_{\gamma+N_t,\zeta+q_t,K}(y)\,\mathrm{d}y, \quad t \ge 0,$$

is a process describing the conditional mean of the loss distribution given the available information up to time t. The reduced control problem is given by

$$\begin{aligned} V^{\xi,b}(t,x,n,q) &:= \mathbb{E}^{t,x,n,q}\big[U(X_T^{\xi,b})\big] := \mathbb{E}\big[U(X_T^{\xi,b}) \,|\, X_t^{\xi,b} = x, N_t = n, q_t = q\big], \\ V(t,x,n,q) &:= \sup_{(\xi,b)\in\mathcal{U}[t,T]} V^{\xi,b}(t,x,n,q), \end{aligned}$$

for every $(t,x,n,q) \in [0,T] \times \mathbb{R} \times \mathbb{N}_0 \times (0,\infty)$. With the same arguments as before, we obtain the generalized HJB equation

$$\begin{aligned} 0 = \inf_{(\xi,b)\in\mathbb{R}\times[0,1]} \bigg\{ &-\lambda\, g(t,n,q) - \alpha e^{r(T-t)} g(t,n,q)\Big((\mu-r)\xi + c(b) - \frac{1}{2}\sigma^2\alpha e^{r(T-t)}\xi^2\Big) \\ &+ \lambda \int_0^K g(t,n+1,q+y)\, \exp\big\{\alpha\, b\, y\, e^{r(T-t)}\big\}\, f_{\gamma+n,\zeta+q,K}(y)\,\mathrm{d}y \bigg\} \\ + \inf_{\varphi\in\partial^C g_{n,q}(t)} &\{\varphi\}. \end{aligned}$$

An analogous procedure as in the previous chapters provides that the optimal reinsurance strategy $(b^\star_\lambda(t))_{t\in[0,T]}$ is given by $b_\lambda(t-,N_{t-},q_{t-})$ with

$$b_\lambda(t,n,q) := \begin{cases} 0, & \theta \le A_\lambda(t,n,q)/\kappa - 1, \\ 1, & \theta \ge B_\lambda(t,n,q)/\kappa - 1, \\ r_\lambda(t,n,q), & \text{otherwise}, \end{cases}$$

where

$$\begin{aligned} h_\lambda(t,n,q,b) &:= \lambda \int_0^K \frac{g(t,n+1,q+y)}{g(t,n,q)} y\, \exp\big\{\alpha\, b\, y\, e^{r(T-t)}\big\}\, f_{\gamma+n,\zeta+q,K}(y)\,\mathrm{d}y, \\ A_\lambda(t,n,q) &:= h_\lambda(t,n,q,0), \\ B_\lambda(t,n,q) &:= h_\lambda(t,n,q,1), \end{aligned}$$

and $r_\lambda(t,n,p)$ is the unique root w.r.t. b of

$$(1+\theta)\,\kappa = h_\lambda(t,n,q,b).$$

Appendix A

Auxiliary Results

A.1 Auxiliary results to Section 4.7

The following two results will be used to provide a change of measure in Lemma A.3 Recall the definition of $\widetilde{\mathcal{U}}[t,T]$ given in (4.39).

Lemma A.1. *Let $\xi = (\xi_t)_{t\in[0,T]}$ be some continuous, bounded and $\mathfrak{F}^W$-adapted investment strategy. Furthermore, let $A = (A_t)_{t\in[0,T]}$ be the process which is given by*

$$A_t^\xi := -\int_0^t \alpha\,\sigma\,e^{r(T-s)}\xi_s\,\mathrm{d}W_s, \quad t\in[0,T]. \tag{A.1}$$

Then the stochastic exponential $\mathcal{E}(A^\xi) = (\mathcal{E}(A^\xi)_t)_{t\in[0,T]}$ of A is an $\mathfrak{F}^W$-martingale on $[0,T]$.

Proof. Fix some continuous, bounded and $\mathfrak{F}^W$-adapted investment strategy $\xi = (\xi_t)_{t\in[0,T]}$. Applying Theorem 5.2 in Klebaner [75], we obtain

$$\mathcal{E}(A)_t = \exp\left\{-\int_0^t \alpha\,\sigma\,e^{r(T-s)}\xi_s\,\mathrm{d}W_s - \frac{1}{2}\int_0^t \alpha^2\,\sigma^2\,e^{2r(T-s)}\xi_s\,\mathrm{d}s\right\}, \quad t\in[0,T].$$

The process $\mathcal{E}(A) = (\mathcal{E}(A)_t)_{t\in[0,T]}$ is obviously $\mathfrak{F}^W$-adapted. From the boundedness of ξ follows, by the Novikov condition (cf. e.g. Corollary 3.5.13 in Karatzas and Shreve [73]), the announced martingale property of $\mathcal{E}(A)$. □

Lemma A.2. *Let $b = (b_t)_{t\in[0,T]}$ be some $\mathfrak{F}^\Psi$-predictable reinsurance strategy. Furthermore, let $B^b = (B_t^b)_{t\in[0,T]}$ be the process which is defined by*

$$B_t^b := \int_0^t\int_{E^d}\left(\exp\left\{\alpha\,b_s\,e^{r(T-s)}\sum_{i=1}^d y_i\mathbb{1}_z(i)\right\}-1\right)\widehat{\Psi}(\mathrm{d}s,\mathrm{d}(y,z)). \tag{A.2}$$

Then the stochastic exponential $\mathcal{E}(B^b) = (\mathcal{E}(B^b)_t)_{t\in[0,T]}$ of B^b is given by

$$\begin{aligned}\mathcal{E}(B^b)_t = \exp\Bigg\{&\int_0^t\int_{E^d}\alpha\,b_s\,e^{r(T-s)}\sum_{i=1}^d y_i\mathbb{1}_z(i)\Psi(\mathrm{d}s,\mathrm{d}(y,z)) + \lambda\,t\\ &-\int_0^t\lambda\sum_{D\subset\mathbb{D}}p_s^D\int_{(0,\infty)^d}\exp\left\{\alpha\,b_s\,e^{r(T-s)}\sum_{i=1}^d y_i\mathbb{1}_D(i)\right\}F(\mathrm{d}y)\mathrm{d}s\Bigg\}.\end{aligned}$$

Furthermore, $\mathcal{E}(B^b)$ is an $\mathfrak{F}^\Psi$-martingale on $[0,T]$.

Proof. Fix some $\mathfrak{F}^\Psi$-predictable reinsurance strategy $b = (b_t)_{t\in[0,T]}$. According to Theorem 2.60, we have

$$\mathcal{E}(B^b)_t = e^{B^b_t - B^b_0 - \frac{1}{2}[B^b]^c_t} \prod_{0<s\le t} \big(1 + \Delta B^b_s\big) e^{-\Delta B^b_s},$$

where, by Proposition 2.51 (v) and the definition of $\widehat{\Psi}$ given in (4.10),

$$[B^b]^c = [(B^b)^c] = \left[\int_0^\cdot \int_{E^d} \left(\exp\left\{ \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i) \right\} - 1 \right) \widehat{\nu}(s, \mathrm{d}(y,z))\, \mathrm{d}s \right] \equiv 0$$

and

$$\prod_{0<s\le t} \big(1 + \Delta B^b_s\big) e^{-\Delta B^b_s} = \exp\left\{ - \sum_{0<s\le t} \Delta B^b_s \right\} \prod_{0<s\le t} \big(1 + \Delta B^b_s\big)$$

with

$$\begin{aligned}
\prod_{0<s\le t} (1 + \Delta B_s) &= \prod_{n\in\mathbb{N}} \left(1 + \left(\exp\left\{ \alpha\, b_{T_n}\, e^{r(T-T_n)} \sum_{i=1}^d Y_n^i \mathbb{1}_{Z_n}(i) \right\} - 1 \right) \mathbb{1}_{\{T_n\le t\}} \right) \\
&= \prod_{n\in\mathbb{N}} \left(\exp\left\{ \alpha\, b_{T_n}\, e^{r(T-T_n)} \sum_{i=1}^d Y_n^i \mathbb{1}_{Z_n}(i) \right\} \mathbb{1}_{\{T_n\le t\}} + \mathbb{1}_{\{T_n > t\}} \right) \\
&= \exp\left\{ \sum_{n\in\mathbb{N}} \alpha\, b_{T_n}\, e^{r(T-T_n)} \sum_{i=1}^d Y_n^i \mathbb{1}_{Z_n}(i) \mathbb{1}_{\{T_n\le t\}} \right\} \\
&= \exp\left\{ \int_0^t \int_{E^d} \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i) \Psi(\mathrm{d}s, \mathrm{d}(y,z)) \right\}
\end{aligned}$$

and

$$\begin{aligned}
&\exp\left\{ - \sum_{0<s\le t} \Delta B^b_s \right\} \\
&= \exp\left\{ - \sum_{n\in\mathbb{N}} \left(\exp\left\{ \alpha\, b_{T_n}\, e^{r(T-T_n)} \sum_{i=1}^d Y_n^i \mathbb{1}_{Z_n}(i) \right\} - 1 \right) \mathbb{1}_{\{T_n\le t\}} \right\} \\
&= \exp\left\{ - \int_0^t \int_{E^d} \left(\exp\left\{ \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i) \right\} - 1 \right) \Psi(\mathrm{d}s, \mathrm{d}(y,z)) \right\}.
\end{aligned}$$

Therefore, again in accordance with (4.10), we obtain

$$\begin{aligned}
\mathcal{E}(B^b)_t = \exp\Bigg\{ &\int_0^t \int_{E^d} \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i) \Psi(\mathrm{d}s, \mathrm{d}(y,z)) \\
&- \int_0^t \int_{E^d} \left(\exp\left\{ \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i) \right\} - 1 \right) \widehat{\nu}(s, \mathrm{d}(y,z))\, \mathrm{d}s \Bigg\},
\end{aligned}$$

which reduces to, by Proposition 4.20 and $\sum_{D\subset\mathbb{D}} p_t^D = 1$,

$$\mathcal{E}(B^b)_t = \exp\bigg\{\int_0^t\int_{E^d}\alpha\, b_s\, e^{r(T-s)}\sum_{i=1}^d y_i \mathbb{1}_z(i)\Psi(\mathrm{d}s,\mathrm{d}(y,z)) + \lambda\, t \\ - \int_0^t \lambda \sum_{D\subset\mathbb{D}} p_s^D \int_{(0,\infty)^d}\exp\Big\{\alpha\, b_s\, e^{r(T-s)}\sum_{i=1}^d y_i\mathbb{1}_D(i)\Big\}F(\mathrm{d}y)\mathrm{d}s\bigg\}.$$

We continue by analysing the process $\mathcal{E}(B^b)$ for the desired martingale property. The process $\mathcal{E}(B^b) = (\mathcal{E}(B^b)_t)_{t\in[0,T]}$ is obviously $\mathfrak{F}^\Psi$-adapted. By definition of the stochastic exponential, we have

$$\begin{aligned}\mathcal{E}(B^b)_t &= \int_0^t \mathcal{E}(B^b)_{s-}\,\mathrm{d}B_s^b \\ &= \int_0^t\int_{E^d}\mathcal{E}(B^b)_{s-}\Big(\exp\Big\{\alpha\, b_s\, e^{r(T-s)}\sum_{i=1}^d y_i\mathbb{1}_z(i)\Big\}-1\Big)\widehat{\Psi}(\mathrm{d}s,\mathrm{d}(y,z)), \quad t\in[0,T].\end{aligned}$$

Therefore, according to Corollary 2.98, $\mathcal{E}(B^b)$ is an $\mathfrak{F}^\Psi$-martingale on $[0,T]$ if

$$\mathbb{E}\bigg[\int_0^T\int_{E^d}\Big|\mathcal{E}(B^b)_{t-}\Big(\exp\Big\{\alpha\, b_t\, e^{r(T-t)}\sum_{i=1}^d y_i\mathbb{1}_z(i)\Big\}-1\Big)\Big|\widehat{\nu}(\mathrm{d}t,\mathrm{d}(y,z))\bigg] < \infty.$$

Notice that the integrand process above is obviously $\mathfrak{F}^\Psi$-predictable due to the $\mathfrak{F}^\Psi$-predictability of $(b_t)_{t\geq 0}$. By the triangle inequality, Assumption 4.3, $b_t \leq 1$ and $\sum_{D\subset\mathbb{D}} p_t^D = 1$, we obtain that the expectation above is less or equal to

$$\begin{aligned}&\mathbb{E}\bigg[\int_0^T\int_{E^d}\exp\Big\{\int_0^t\int_{E^d}\alpha\, b_s\, e^{r(T-s)}\sum_{i=1}^d y_i\mathbb{1}_z(i)\,\Psi(\mathrm{d}s,\mathrm{d}(y,z))+\lambda\, t\Big\}\times \\ &\qquad\Big(\exp\Big\{\alpha\, b_t\, e^{r(T-t)}\sum_{i=1}^d y_i\mathbb{1}_z(i)\Big\}+1\Big)\lambda\sum_{D\in\mathrm{d}z} p_s^D\, F(\mathrm{d}y)\,\mathrm{d}t\bigg] \\ &\leq \mathbb{E}\bigg[\int_0^T\exp\Big\{\alpha\, e^{|r|T}\sum_{i=1}^d\sum_{n=1}^{N_t} Y_n^i + \lambda\, T\Big\}\times \\ &\qquad \lambda\sum_{D\subset\mathbb{D}} p_s^D\Big(\int_{(0,\infty)^d}\exp\Big\{\alpha\, e^{|r|T}\sum_{i=1}^d y_i\Big\}F(\mathrm{d}y)+1\Big)\,\mathrm{d}t\bigg] \\ &= \Big(M_F\big(\alpha e^{|r|T}\big)+1\Big)\lambda\, e^{\lambda T}\int_0^T\mathbb{E}\bigg[\exp\Big\{\alpha e^{|r|T}\sum_{i=1}^d\sum_{k=1}^{N_t} Y_k^i\Big\}\bigg]\mathrm{d}t,\end{aligned}$$

where, by Lemma 4.5 (ii), the expectation above is finite as well as the other term. □

Now the aforementioned change of measure will be introduced.

Lemma A.3. *Let $t\in[0,T]$ and let $(\xi,b)\in\widetilde{\mathcal{U}}[0,T]$ be an arbitrary admissible strategy. We set*

$$L_t^{\xi,b} := \exp\bigg\{-\int_0^t \alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s - \frac{1}{2}\int_0^t \alpha^2\,\sigma^2\, e^{2r(T-s)}\xi_s^2\,\mathrm{d}s$$

$$+ \int_0^t \int_E \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \Psi(\mathrm{d}s, \mathrm{d}(y,z)) + \lambda\, t$$
$$- \int_0^t \lambda \sum_{D \subset \mathbb{D}} p_s^D \int_{(0,\infty)^d} \exp\left\{ \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i) \right\} F(\mathrm{d}y)\, \mathrm{d}s \Bigg\}.$$

Then, a probability measure on $(\Omega, \mathcal{G}_t)$ is defined by $\mathbb{Q}_t^{\xi,b}(A) := \int_A L_t^{\xi,b}\, \mathrm{d}\mathbb{P}$, $A \in \mathcal{G}_t$, for every $t \in [0,T]$, i.e. $\frac{\mathrm{d}\mathbb{Q}_t^{\xi,b}}{\mathrm{d}\mathbb{P}} := L_t^{\xi,b}$. The probability measures $\mathbb{Q}_t^{\xi,b}$ and $\mathbb{P}$ are equivalent.

Proof. Fix $(\xi, b) \in \widetilde{\mathcal{U}}[0,T]$. We define a process $L^{\xi,b} = (L_t^{\xi,b})_{t\in[0,T]}$ as solution of the stochastic differential equation

$$\mathrm{d}L_t^{\xi,b} = L_{t-}^{\xi,b}\, \mathrm{d}Z_t^{\xi,b}, \quad L_0^{\xi,b} = 1,$$

with $Z_t^{\xi,b} := A_t^\xi + B_t^b$, $t \in [0,T]$, where A_t^ξ is defined by (A.1) and B_t^b by (A.2). That is, $L^{\xi,b}$ is the Doléans-Dade exponential of $Z^{\xi,b}$ which is denoted by $\mathcal{E}(Z^{\xi,b})$. Let us fix some $t \in [0,T]$. From Theorem II.38 in Protter [104], it follows

$$L_t^{\xi,b} = \mathcal{E}(Z^{\xi,b})_t = \mathcal{E}(A^\xi)_t\, \mathcal{E}(B^b)_t,$$

where $\mathcal{E}(A^\xi) = (\mathcal{E}(A^\xi)_t)_{t\in[0,T]}$ and $\mathcal{E}(B^b) = (\mathcal{E}(B^b)_t)_{t\in[0,T]}$ are the Doléans-Dade exponential of $A^\xi = (A_t^\xi)_{t\in[0,T]}$ and $B^b = (B_t^b)_{t\in[0,T]}$, respectively. Therefore, Lemma A.1 and Lemma A.2 imply

$$L_t^{\xi,b} = \exp\Bigg\{ - \int_0^t \alpha\, \sigma\, e^{r(T-s)} \xi_s\, \mathrm{d}W_s - \frac{1}{2} \int_0^t \alpha^2\, \sigma^2\, e^{2r(T-s)} \xi_s^2\, \mathrm{d}s$$
$$+ \int_0^t \int_{E^d} \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i) \Psi(\mathrm{d}s, \mathrm{d}(y,z)) + \lambda\, t$$
$$- \int_0^t \lambda \sum_{D \subset \mathbb{D}} p_s^D \int_{E^d} \exp\left\{ \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i) \right\} F(\mathrm{d}y) \mathrm{d}s \Bigg\}.$$

We are left with the task of showing that $L^{\xi,b}$ is a $(\mathbb{P}, \mathfrak{G})$-martingale. Recall that $\mathcal{E}(A^\xi)$ is a $(\mathbb{P}, \mathfrak{F}^W)$-martingale on $[0,T]$ and that $\mathcal{E}(B^b)$ is a $(\mathbb{P}, \mathfrak{F}^\Psi)$-martingale on $[0,T]$, compare Lemma A.1 and Lemma A.9. Recall that the Brownian motion $(W_t)_{t\geq 0}$ is independent of $(T_n)_{n\in\mathbb{N}}$, $(Y_n)_{n\in\mathbb{N}}$ and $(Z_n)_{n\in\mathbb{N}}$ according to Assumption 3.6. Therefore, since $\mathcal{G}_t = \mathcal{F}_t^W \vee \mathcal{F}_t^\Psi$, $t \geq 0$, and the product of two independent martingales (each with respect to its natural filtration) is a martingale (with respect to the natural filtration of the product), compare Theorem 2.1 of Chapter "Some particular Problems of Martingale Theory" in Kabanov et al. [72], it follows that $L^{\xi,b}$ is a $(\mathbb{P}, \mathfrak{G})$-martingale on $[0,T]$. Thus $\mathbb{E}[L_t^{\xi,b}] = 1$. Therefore, we can define a new measure $\mathbb{Q}_t^{\xi,b}$ on $(\Omega, \mathcal{G}_t)$ by $\mathbb{Q}_t^{\xi,b}(A) = \int_A L_t^{\xi,b}\, \mathrm{d}\mathbb{P}$ for every $A \in \mathcal{G}_t$, where it is easily seen that $\mathbb{Q}_t^{\xi,b}$ and $\mathbb{P}$ are equivalent. □

The point of the lemma is that it allows one to prove that $\mathbb{E}[\exp\{-\alpha\, e^{r(T-t)} X_t^{\xi,b}\}]$ is bounded for all $t \in [0,T]$. This condition will be needed in the proof of the important Lemma A.8. It is clear that $\exp\{-\alpha\, e^{r(T-t)} X_t^{\xi,b}\} < \infty$ $\mathbb{P}$-a.s. due to the càdlàg property of $X^{\xi,b}$ (since a càdlàg function is bounded on a compact set). To show the boundedness of the expectation, the basic idea is to change the measure $\mathbb{P}$ to an equivalent probability measure such that the expectation only includes Lebesgue integrals bounded on $[0,T]$.

And this is where the lemma above comes into picture.

Lemma A.4. *Let $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a function defined by* (A.3)*. Furthermore, let $(\xi, b) \in \widetilde{\mathcal{U}}[0,T]$ and let $L^{\xi,b} = (L_t^{\xi,b})_{t\in[0,T]}$ be the density process of Lemma A.3. Then there exists a constant $0 < K_1 < \infty$ such that*

$$\frac{\big|f(t, X_t^{\xi,b})\big|}{L_t^{\xi,b}} \leq K_1 \quad \mathbb{P}\text{-}a.s.$$

for all $t \in [0,T]$.

Proof. Fix $t \in [0,T]$ and $(\xi, b) \in \widetilde{\mathcal{U}}[0,t]$. From Proposition 3.14, we know

$$\big|f(t, X_t^{\xi,b})\big| = \exp\bigg\{ -\alpha\, x_0\, e^{r(T-t)} e^{rt} - \alpha \int_0^t e^{r(T-t)} e^{r(t-s)} \big((\mu - r)\,\xi_s + c(b_s)\big)\, \mathrm{d}s$$
$$- \alpha \int_0^t \sigma\, e^{r(T-t)} e^{r(t-s)} \xi_s\, \mathrm{d}W_s + \alpha \int_0^t \int_{E^d} b_s\, e^{r(T-t)} e^{r(t-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \Psi(\mathrm{d}s, \mathrm{d}(y,z))\bigg\}.$$

Consequently

$$\frac{\big|f(t, X_t^{\xi,b})\big|}{L_t^{\xi,b}} = \exp\bigg\{ -\alpha x_0 e^{rT} + \int_0^t \bigg(-\alpha\, e^{r(T-s)} \big((\mu - r)\,\xi_s + c(b_s)\big) + \frac{1}{2}\alpha^2\, \sigma^2\, e^{2r(T-s)} \xi_s^2$$
$$+ \lambda \sum_{D\subset\mathbb{D}} p_s^D \int_{(0,\infty)^d} \exp\bigg\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\bigg\} F(\mathrm{d}y)\bigg)\, \mathrm{d}s - \lambda t\bigg\}$$
$$\leq \exp\bigg\{ \int_0^t \bigg(-\alpha\, e^{r(T-s)} \big((\mu - r)\,\xi_s + c(b_s)\big) + \frac{1}{2}\alpha^2\, \sigma^2\, e^{2r(T-s)} \xi_s^2$$
$$+ \lambda \sum_{D\subset\mathbb{D}} p_s^D \int_{(0,\infty)^d} \exp\bigg\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\bigg\} F(\mathrm{d}y)\bigg)\, \mathrm{d}s\bigg\}.$$

Using Assumption 4.3, $|\xi_s| \leq K$, $\sum_{D\subset\mathbb{D}} p_s^D = 1$, $c(b_t) = (1+\eta)\kappa - (1-b_t)(1+\theta)\kappa$ and $e^x \leq e^{|x|}$ for all $x \in \mathbb{R}$, we obtain

$$\exp\bigg\{ \int_0^t \bigg(-\alpha\, e^{r(T-s)} \big((\mu - r)\,\xi_s + c(b_s)\big) + \frac{1}{2}\alpha^2\, \sigma^2\, e^{2r(T-s)} \xi_s^2$$
$$+ \lambda \sum_{D\subset\mathbb{D}} p_s^D \int_{(0,\infty)^d} \exp\bigg\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\bigg\} F(\mathrm{d}y)\bigg)\, \mathrm{d}s\bigg\}$$
$$\leq \exp\bigg\{\bigg(\alpha e^{|r|T}\Big(|\mu - r|K + (2+\eta+\theta)\kappa + \frac{1}{2}\alpha\, \sigma^2 e^{|r|T} K^2\Big) + \lambda M_F\big(\alpha e^{|r|T}\big)\bigg)T\bigg\} =: K_1,$$

where $0 < K_1 < \infty$ is independent of $t \in [0,T]$ as well as (ξ, b). □

Corollary A.5. *Let $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a function defined by* (A.3)*. Furthermore, let $(\xi, b) \in \widetilde{\mathcal{U}}[0,T]$ and let $\widetilde{L}^{\xi,b} = (\widetilde{L}_t^{\xi,b})_{t\in[0,T]}$ be the density process of Lemma A.3 with α replaced by 2α. Then there exists a constant $0 < K_2 < \infty$ such that*

$$\frac{\big(f(t, X_t^{\xi,b})\big)^2}{\widetilde{L}_t^{\xi,b}} \leq K_2 \quad \mathbb{P}\text{-}a.s.$$

for all $t \in [0,T]$.

Proof. Let $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ be defined by (A.3). Since, by Proposition 3.14,

$$\big(f(t, X_t^{\xi,b})\big)^2 = \exp\bigg\{ -2\alpha\, x_0\, e^{r(T-t)} e^{rt} - 2\alpha \int_0^t e^{r(T-t)} e^{r(t-s)} \big((\mu - r)\, \xi_s + c(b_s)\big)\, \mathrm{d}s$$
$$- 2\alpha \int_0^t \sigma\, e^{r(T-t)} e^{r(t-s)} \xi_s\, \mathrm{d}W_s + 2\alpha \int_0^t \int_{E^d} b_s\, e^{r(T-t)} e^{r(t-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \Psi(\mathrm{d}s, \mathrm{d}(y,z))\bigg\},$$

the assertion follows as in the proof of Lemma A.4. □

The next two lemmata are used to prove Lemma A.8.

Lemma A.6. *The function* $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ *given by*

$$f(t,x) := -e^{-\alpha x e^{r(T-t)}}, \quad (t,x) \in [0,T] \times \mathbb{R}, \tag{A.3}$$

satisfies

$$f(t, X_t^{\xi,b}) = f(0, X_0^{\xi,b}) + \int_0^t \bigg(f(s, X_s^{\xi,b}) \alpha e^{r(T-s)} \Big(\frac{1}{2} \alpha\, \sigma^2 e^{r(T-s)} \xi_s^2 - (\mu - r)\xi_s - c(b_s)\Big)$$
$$+ \lambda \sum_{D \subset \mathbb{D}} p_s^D\, f(s, X_s^{\xi,b}) \int_{(0,\infty)^d} \exp\bigg\{ \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i) \bigg\}\, F(\mathrm{d}y) - \lambda\, f(s, X_s^{\xi,b}) \bigg) \mathrm{d}s$$
$$- \int_0^t f(s, X_{s-}^{\xi,b})\, \alpha\, \sigma e^{r(T-s)} \xi_s\, \mathrm{d}W_s$$
$$+ \int_0^t \int_{E^d} f(s, X_{s-}^{\xi,b}) \bigg(\exp\bigg\{ \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i) \bigg\} - 1 \bigg) \widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z)),$$

for all $t \in [0,T]$.

Proof. The proof is a straightforward application of Itô-Doeblin's formula. Set

$$f(t,x) := -e^{-\alpha x e^{r(T-t)}}, \quad t \in [0,T],\ x \in \mathbb{R}.$$

An easy verification shows that $f \in C^{1,2}((0,T) \times \mathbb{R})$ and, for any $(t,x) \in (0,T) \times \mathbb{R}$,

$$f_t(t,x) = -\alpha\, x\, r\, e^{r(T-t)}\, e^{-\alpha x e^{r(T-t)}} = \alpha\, x\, r\, e^{r(T-t)} f(t,x),$$
$$f_x(t,x) = \alpha\, e^{r(T-t)} e^{-\alpha x e^{r(T-t)}} = -\alpha\, e^{r(T-t)} f(t,x),$$
$$f_{xx}(t,x) = -\alpha^2\, e^{2r(T-t)} e^{-\alpha x e^{r(T-t)}} = \alpha^2\, e^{2r(T-t)} f(t,x),$$

where $f_t(t,x)$, $f_x(t,x)$ and $f_{xx}(t,x)$ denote the partial derivatives of f at (t,x). Fix $t \in [0,T]$. By Corollary 2.58 (Itô-Doeblin's formula), we have

$$f(t, X_t^{\xi,b}) = f(0, X_0^{\xi,b}) + \int_0^t f_t(s, X_s^{\xi,b})\, \mathrm{d}s + \int_0^t f_x(s, X_{s-}^{\xi,b})\, \mathrm{d}(X^{\xi,b})_s^c$$
$$+ \frac{1}{2} \int_0^t f_{xx}(s, X_{s-}^{\xi,b})\, \mathrm{d}[X^{\xi,b}]_s^c + \sum_{0<s\le t} \big(f(s, X_s^{\xi,b}) - f(s-, X_{s-}^{\xi,b})\big).$$

Therefore, by Proposition 3.15,

$$\begin{aligned} f(t, X_t^{\xi,b}) &= f(0, X_0^{\xi,b}) + \int_0^t f(s, X_s^{\xi,b})\, \alpha\, e^{r(T-t)} \Big(r X_s^{\xi,b} - \big(r X_s^{\xi,b} + (\mu - r)\xi_s + c(b_s) \big) \\ &\quad + \frac{1}{2} \alpha\, \sigma^2\, e^{r(T-s)} \xi_s^2 \Big)\, \mathrm{d}s - \int_0^t f(s, X_{s-}^{\xi,b}) \alpha\, \sigma\, e^{r(T-s)} \xi_s\, \mathrm{d}W_s \\ &\quad + \sum_{0<s\le t} \big(f(s, X_s^{\xi,b}) - f(s, X_{s-}^{\xi,b}) \big). \end{aligned}$$

Furthermore, since $X^{\xi,b}$ jumps only at the arrival times of the trigger events $N = (T_n)_{n\in\mathbb{N}}$,

$$\begin{aligned} &\sum_{0<s\le t} \big(f(s, X_s^{\xi,b}) - f(s, X_{s-}^{\xi,b}) \big) \\ &= \sum_{n\in\mathbb{N}} \bigg(f\Big(T_n, X_{T_n-}^{\xi,b} - b_{T_n} \sum_{i=1}^d y_i \mathbb{1}_{\{i\in Z_n\}} \Big) - f\big(T_n, X_{T_n-}^{\xi,b} \big) \bigg) \mathbb{1}_{\{T_n \le t\}} \\ &= \int_0^t \int_{E^d} \bigg(f\Big(s, X_{s-}^{\xi,b} - b_s \sum_{i=1}^d y_i \mathbb{1}_z(i) \Big) - f(s, X_{s-}^{\xi,b}) \bigg) \Psi(\mathrm{d}s, \mathrm{d}(y,z)) \end{aligned}$$

Using the compensated random measure $\widehat{\Psi}$ defined in (4.10), we obtain

$$\begin{aligned} &\sum_{0<s\le t} \big(f(s, X_s^{\xi,b}) - f(s, X_{s-}^{\xi,b}) \big) \\ &= \int_0^T \int_{E^d} \bigg(f\Big(s, X_{s-}^{\xi,b} - b_s \sum_{i=1}^d y_i \mathbb{1}_z(i) \Big) - f(s, X_{s-}^{\xi,b}) \bigg) \widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z)) \\ &\quad + \lambda \sum_{D\subset\mathbb{D}} \int_0^t \int_{(0,\infty)^d} \bigg(f\Big(s, X_{s-}^{\xi,b} - b_s \sum_{i=1}^d y_i \mathbb{1}_z(i) \Big) - f(s, X_{s-}^{\xi,b}) \bigg) \lambda\, p_s^D\, F(\mathrm{d}y)\, \mathrm{d}s. \end{aligned}$$

Due to the relation

$$\begin{aligned} f\Big(t, x - b \sum_{i=1}^d y_i \mathbb{1}_D(i) \Big) &= -\exp\Big\{ -\alpha \Big(x - b \sum_{i=1}^d y_i \mathbb{1}_D(i) \Big) e^{r(T-t)} \Big\} \\ &= -\exp\Big\{ -\alpha\, x\, e^{r(T-t)} \Big\} \exp\Big\{ \alpha\, b \sum_{i=1}^d y_i \mathbb{1}_D(i) e^{r(T-t)} \Big\} \\ &= f(t,x) \exp\Big\{ \alpha\, b \sum_{i=1}^d y_i \mathbb{1}_D(i) e^{r(T-t)} \Big\}, \end{aligned} \tag{A.4}$$

we get, by $\sum_{D\subset\mathbb{D}} p_s^D = 1$,

$$\begin{aligned} &\sum_{0<s\le t} \big(f(s, X_s^{\xi,b}) - f(s, X_{s-}^{\xi,b}) \big) \\ &= \int_0^t \int_{E^d} f(s, X_{s-}^{\xi,b}) \Big(\exp\Big\{ \alpha b_s \sum_{i=1}^d y_i \mathbb{1}_z(i) e^{r(T-s)} \Big\} - 1 \Big) \widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z)) \end{aligned}$$

$$+\int_0^t \lambda \sum_{D\subset\mathbb{D}} p_s^D\, f(s,X_s^{\xi,b}) \int_{(0,\infty)^d} \exp\left\{\alpha b_s \sum_{i=1}^d y_i \mathbb{1}_D(i) e^{r(T-s)}\right\} F(\mathrm{d}y)\,\mathrm{d}s$$
$$-\int_0^t \lambda\, f(s,X_s^{\xi,b})\,\mathrm{d}s.$$

This yields the assertion. □

Lemma A.7. *Let $h : [0,T]\times\Delta_m \to (0,\infty)$ be a function such that $t\mapsto h(t,p)$ is absolutely continuous on $[0,T]$ for all $p\in\Delta_m$. Then*

$$h(t,p_t) = h(0,p_0) + \int_0^t \Big(h_t(s,p_s) - \lambda h(s,p_s) + \lambda \sum_{D\subset\mathbb{D}} p_s^D h\big(s, J(p_s,D)\big)\Big)\,\mathrm{d}s$$
$$+\int_0^t \int_{\mathcal{P}(\mathbb{D})} \Big(h\big(s,J(p_{s-},z)\big) - h(s,p_{s-})\Big)\,\widehat{\Phi}(\mathrm{d}s,\mathrm{d}z), \quad t\in[0,T].$$

Proof. According to the assumption of absolutely continuity of $[0,T]\ni t\mapsto h(t,p)$ for all $p\in\Delta_m$ and the piecewise constancy of the filter $(p_t)_{t\geq 0}$ between the jump times $(T_n)_{n\in\mathbb{N}}$ (see Proposition 4.16, it follows that the function $F(t) := h(t,p_t)$ is absolutely continuous on $[T_{n-1},T_n]$ for every $n\in\mathbb{N}$. Hence the FTCL (cf. Sohrab [115, Thm. 11.5.23, 11.5.31]) implies

$$F(t) = F(T_{n-1}) + \int_{T_{n-1}}^t F'(s)\,\mathrm{d}s, \quad n\in\mathbb{N},$$

where F' is Lebesgue integrable. The absolutely continuity of $t\mapsto h(t,p)$ also implies that the derivative of $h(t,p)$ w.r.t. t exists almost everywhere on $[0,T]$ in the sense of the Lebesgue measure, compare Proposition 2.44 in connection with Lemma 2.46. Hence

$$h(t,p_t) = h(T_{n-1},p_{T_{n-1}}) + \int_{T_{n-1}}^{T_n-} h_t(s,p_s)\,\mathrm{d}s, \quad t\in[T_{n-1},T_n],\ n\in\mathbb{N},$$

and, in consequence,

$$h(t,p_t) = h(0,p_0) + \int_0^t h_t(s,p_s)\,\mathrm{d}s + \sum_{0<s\leq t}\big(h(s,p_s) - h(s,p_{s-})\big), \quad t\in[0,T],$$

where, by Proposition 4.16 and the definition compensated random measure $\widehat{\Phi}$ in (4.2),

$$\sum_{0<s\leq t}\big(h(s,p_s) - h(s-,p_{s-})\big) = \sum_{n\in\mathbb{N}} \Big(h\big(T_n, J(p_{T_n-},Z_n)\big) - h(T_n,p_{T_n-})\Big)\mathbb{1}_{\{T_n\leq t\}}$$
$$= \int_0^t \int_{\mathcal{P}(\mathbb{D})} \big(h(s,J(p_{s-},z)) - h(s,p_{s-})\big)\,\widehat{\Phi}(\mathrm{d}s,\mathrm{d}z)$$
$$+\lambda \sum_{D\subset\mathbb{D}} \int_0^t h\big(s,J(p_s,D)\big)\,p_s^D\,\mathrm{d}s - \lambda\int_0^t h(s,p_s)\,\mathrm{d}s,$$

which yields the assertion. □

The next result is crucial for the proof of the Verification Theorem 4.31. It makes use of the notation of the operator $\mathscr{H}$ given by 4.41.

Lemma A.8. *Suppose that $(\xi,b)\in\widetilde{\mathcal{U}}[0,T]$ is an arbitrary strategy and $h : [0,T]\times\Delta_m \to (0,\infty)$ is a bounded function such that $t\mapsto h(t,p)$ is absolutely continuous on $[0,T]$ for* *all $p \in \Delta_m$ and $p \mapsto h(t,p)$ is continuous on Δ_m for all $t \in [0,T]$. Then, the function $G : [0,T] \times \mathbb{R} \times \Delta_m \to \mathbb{R}$ defined by*

$$G(t,x,p) := -e^{-\alpha x e^{r(T-t)}}\, h(t,p)$$

satisfies

$$\mathrm{d}G(t, X_t^{\xi,b}, p_t) = -e^{-\alpha X_t^{\xi,b} e^{r(T-t)}}\, \mathscr{H} h(t, p_t; \xi_t, b_t)\,\mathrm{d}t + \mathrm{d}\eta_t^{\xi,b}, \quad t \in [0,T],$$

where $(\eta_t^{\xi,b})_{t\in[0,T]}$ is a $\mathfrak{G}$-martingale starting at zero and we set $\mathscr{H} h(t,p;\xi,b)$ to zero at those points (t,p) where the partial derivative of h w.r.t. t does not exist.

Proof. Let $(\xi, b) \in \widetilde{\mathcal{U}}[0,T]$ be an arbitrary strategy and let $h : [0,T] \times \Delta_m \to (0,\infty)$ be a function satisfying the conditions stated in the lemma, where $0 < K_0 < \infty$ is some constant which bounds h, i.e. $|h(t,p)| \le K_0$ for all $(t,p) \in [0,T] \times \Delta_m$. Let us fix $(t,x,p) \in [0,T] \times \mathbb{R} \times \Delta_m$ and set

$$G(t,x,p) := -e^{-\alpha x e^{r(T-t)}} h(t,p) \quad \text{and} \quad f(t,x) := -e^{-\alpha x e^{r(T-t)}}.$$

From Lemma A.6, we get

$$\begin{aligned}
\mathrm{d}f(t,X_t^{\xi,b}) = &\Bigg[f(t,X_t^{\xi,b})\alpha e^{r(T-t)}\Big(\frac{1}{2}\sigma^2\alpha e^{r(T-t)}\xi_t^2 - (\mu - r)\xi_t - c(b_t)\Big) \\
&+ \lambda \sum_{D\subset\mathbb{D}} p_t^D\, f(t,X_t^{\xi,b}) \int_{(0,\infty)^d} \exp\Big\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y) - \lambda\, f(t,X_t^{\xi,b}) \Bigg] \mathrm{d}t \\
&- f(t,X_{t-}^{\xi,b})\alpha\sigma e^{r(T-t)}\xi_t\,\mathrm{d}W_t \\
&+ \int_{E^d} f(t,X_{t-}^{\xi,b}) \Big(\exp\Big\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big\} - 1\Big) \widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z)),
\end{aligned} \tag{A.5}$$

Moreover, Lemma A.7 yields

$$\begin{aligned}
\mathrm{d}h(t,p_t) = &\Big(h_t(t,p_t) - \lambda h(t,p_t) + \lambda \sum_{D\subset\mathbb{D}} p_t^D h\big(t, J(p_t,D)\big)\Big)\,\mathrm{d}t \\
&+ \int_{\mathcal{P}(\mathbb{D})} \Big(h\big(t, J(p_{t-},D)\big) - h(t-,p_{t-})\Big)\, \widehat{\Phi}(\mathrm{d}t,\mathrm{d}z).
\end{aligned} \tag{A.6}$$

Since, $G(t,X_t^{\xi,b},p_t) = f(t,X_t^{\xi,b})\, h(t,p_t)$, the product rule (cf. Thm. 2.59) implies

$$\mathrm{d}G(t,X_t^{\xi,b},p_t) = h(t-,p_{t-})\,\mathrm{d}f(t,X_t^{\xi,b}) + f(t,X_{t-}^{\xi,b})\,\mathrm{d}h(t,p_t) + \mathrm{d}\big[f(\cdot,X_\cdot^{\xi,b}), h(\cdot,p_\cdot)\big]_t. \tag{A.7}$$

From Proposition 2.51 (iii), (v) and (vi) follows

$$\begin{aligned}
&\big[f(\cdot,X_\cdot^{\xi,b}), h(\cdot,p_\cdot)\big]_t \\
&= \frac{1}{2}\Big([f(\cdot,X_\cdot^{\xi,b}) + h(\cdot,p_\cdot)]_t - [f(\cdot,X_\cdot^{\xi,b})]_t - [h(\cdot,p_\cdot)]_t \Big) \\
&= \frac{1}{2}\Big([f(\cdot,X_\cdot^{\xi,b}) + h(\cdot,p_\cdot)]_t^c + \big(f(0,X_0^{\xi,b}) + h(0,X_0^{\xi,b})\big)^2 \\
&\quad + \sum_{0<s\le t} \big(\Delta f(s,X_s^{\xi,b}) + \Delta h(s,p_s)\big)^2 - [f(\cdot,X_\cdot^{\xi,b})]_t^c - f(0,X_0^{\xi,b})^2
\end{aligned}$$

$$- \sum_{0<s\le t} \big(\Delta f(s, X_s^{\xi,b})\big)^2 - [h(\cdot, p.)]_t^c - h(0,p_0)^2 - \sum_{0<s\le t} \big(\Delta h(s,p_s)\big)^2\Big).$$

Since $h(\cdot, p.)$ is an FV process, we have

$$[f(\cdot, X_\cdot^{\xi,b}) + h(\cdot, p.)]_t^c = \int_0^t f(s, X_s^{\xi,b})^2 \alpha^2 \sigma^2 e^{2r(T-s)} \xi_s^2 \,\mathrm{d}s = [f(\cdot, X_\cdot^{\xi,b})]_t^c$$

and

$$[h(\cdot, p.)]_t^c = 0.$$

Thus,

$$\begin{aligned}
&\big[f(\cdot, X_\cdot^{\xi,b}), h(\cdot, p.)\big]_t \\
&= \frac{1}{2}\Big(2\, f(0, X_0^{\xi,b})\, h(0,p_0) + 2 \sum_{0<s\le t} \Delta f(s, X_s^{\xi,b})\, \Delta h(s,p_s)\Big) \\
&= f(0, X_0^{\xi,b})\, h(0,p_0) + \sum_{0<s\le t} \big(f(s, X_s^{\xi,b}) - f(s, X_{s-}^{\xi,b})\big)\big(h(s,p_s) - h(s-, p_{s-})\big) \\
&= f(0, X_0^{\xi,b})\, h(0,p_0) + \sum_{0<s\le t} \bigg(f(s, X_s^{\xi,b})\, \big(h(s,p_s) - h(s-, p_{s-})\big) \\
&\quad - f(s, X_{s-}^{\xi,b})\, \big(h(s,p_s) - h(s-, p_{s-})\big)\bigg).
\end{aligned}$$

Due to the jumps of $X^{\xi,b}$ and p at the arrival times of the trigger events $(T_n)_{n\in\mathbb{N}}$, we obtain, by using Proposition 4.16,

$$\begin{aligned}
&\big[f(\cdot, X_\cdot^{\xi,b}), h(\cdot, p.)\big]_t = f(0, X_0^{\xi,b})\, h(0,p_0) \\
&+ \sum_{n\in\mathbb{N}} \bigg(f\Big(T_n, X_{T_n-}^{\xi,b} - b_{T_n} \sum_{i=1}^d y_i \mathbb{1}_{\{i\in Z_n\}}\Big) \Big(h\big(T_n, J(p_{T_n-}, Z_n)\big) - h(T_n-, p_{T_n-})\Big) \\
&\quad - f(T_n, X_{T_n-}^{\xi,b})\, \big(h(T_n, J(p_{T_n-}, Z_n)) - h(T_n-, p_{T_n-})\big)\bigg) \mathbb{1}_{\{T_n \le t\}} \\
&= f(0, X_0^{\xi,b})\, h(0,p_0) \\
&+ \int_0^t \int_{E^d} f\Big(s, X_{s-}^{\xi,b} - b_s \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big) \big(h(s, J(p_{s-}, z)) - h(s-, p_{s-})\big)\, \Psi(\mathrm{d}s, \mathrm{d}(y,z)) \\
&- \int_0^t \int_{\mathcal{P}(\mathbb{D})} f(s, X_{s-}^{\xi,b}) \Big(h(s, J(p_{s-}, z)) - h(s-, p_{s-})\Big)\, \Phi(\mathrm{d}s, \mathrm{d}z)
\end{aligned}$$

Once again, we use the compensated random measures $\widehat{\Phi}$ and $\widehat{\Psi}$ (compare (4.2) and (4.10), respectively) and we obtain

$$\begin{aligned}
&\mathrm{d}\big[f(\cdot, X_\cdot^{\xi,b}), h(\cdot, p.)\big]_t = \\
&\int_{E^d} f\Big(t, X_{t-}^{\xi,b} - b_t \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big) \Big(h\big(t, J(p_{t-}, z)\big) - h(t-, p_{t-})\Big)\, \widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z)) \\
&+ \lambda \sum_{D\subset\mathbb{D}} p_t^D \Big(h\big(t, J(p_t, D)\big) - h(t, p_t)\Big) \int_{(0,\infty)^d} f\Big(t, X_t^{\xi,b} - b_t \sum_{i=1}^d y_i \mathbb{1}_D(i)\Big)\, F(\mathrm{d}y)\,\mathrm{d}t
\end{aligned}$$

$$- \int_{\mathcal{P}(\mathbb{D})} f(t, X_{t-}^{\xi,b}) \Big(h\big(t, J(p_{t-}, z)\big) - h(t-, p_{t-})\Big)\, \widehat{\Phi}(\mathrm{d}t, \mathrm{d}z)$$
$$- \lambda \sum_{D \subset \mathbb{D}} p_t^D f(t, X_t^{\xi,b}) \Big(h\big(t, J(p_t, D)\big) - h(t, p_t)\Big)\, \mathrm{d}t.$$

Using the relation (A.4), we get

$$\begin{aligned}
\mathrm{d}\big[f(\cdot, X_\cdot^{\xi,b}), h(\cdot, p_\cdot)\big]_t = \Bigg[&\lambda \sum_{D \subset \mathbb{D}} p_t^D\, f(t, X_t^{\xi,b})\Big(h\big(t, J(a^D, p_t)\big) - h(t, p_t)\Big)\times \\
&\int_{(0,\infty)^d} \exp\Big\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y) \\
&+ \lambda\, f(t, X_t^{\xi,b})\, h(t,p) - \lambda \sum_{D \subset \mathbb{D}} p_t^D\, f(t, X_t^{\xi,b})\, h\big(t, J(p_t, D)\big)\Bigg]\, \mathrm{d}t \\
&+ \int_{E^d} f(t, X_{t-}^{\xi,b})\, h\big(t, J(p_{t-}, z)\big)\, \exp\Big\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big\} \widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z)) \\
&- \int_{E^d} f(t, X_{t-}^{\xi,b})\, h(t-, p_{t-})\, \exp\Big\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big\} \widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z)) \\
&- \int_{\mathcal{P}(\mathbb{D})} f(t, X_{t-}^{\xi,b}) \Big(h\big(t, J(p_{t-}, z)\big) - h(t-, p_{t-})\Big)\, \widehat{\Phi}(\mathrm{d}t, \mathrm{d}z).
\end{aligned} \tag{A.8}$$

Inserting (A.5), (A.6) and (A.8) into (A.7), we obtain

$$\begin{aligned}
\mathrm{d}G(t, X_t^{\xi,b}, p_t) = \Bigg[&h(t,p_t)\, f(t, X_t^{\xi,b})\, \alpha\, e^{r(T-t)} \Big(\frac{1}{2}\sigma^2 \alpha e^{r(T-t)} \xi_t^2 - (\mu - r)\xi_t - c(b_t)\Big) \\
&+ h(t,p_t)\lambda \sum_{D \subset \mathbb{D}} p_t^D\, f(t, X_t^{\xi,b}) \int_{(0,\infty)^d} \exp\Big\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\}\, F(\mathrm{d}y) \\
&- \lambda\, h(t,p_t)\, f(t, X_t^{\xi,b})\Bigg]\, \mathrm{d}t - h(t-, p_{t-})\, f(t, X_{t-}^{\xi,b})\, \sigma\, \alpha\, e^{r(T-t)} \xi_t\, \mathrm{d}W_t \\
&+ \int_{E^d} h(t-, p_{t-})\, f(t, X_{t-}^{\xi,b}) \Big(\exp\Big\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big\} - 1\Big) \widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z)) \\
&+ f(t, X_t^{\xi,b}) \Big(h_t(t,p_t) - \lambda\, h(t,p_t) + \lambda \sum_{D \subset \mathbb{D}} p_t^D h\big(t, J(p_t, D)\big)\Big)\, \mathrm{d}t \\
&+ \int_{\mathcal{P}(\mathbb{D})} f(t, X_{t-}^{\xi,b}) \Big(h\big(t, J(p_{t-}, z)\big) - h(t-, p_{t-})\Big)\, \widehat{\Phi}(\mathrm{d}t, \mathrm{d}z) \\
&+ \Bigg[\lambda \sum_{D \subset \mathbb{D}} p_t^D\, f(t, X_t^{\xi,b})\Big(h\big(t, J(p_t, D)\big) - h(t, p_t)\Big)\times \\
&\int_{(0,\infty)^d} \exp\Big\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y) \\
&+ \lambda\, f(t, X_t^{\xi,b})\, h(t,p_t) - \lambda \sum_{D \subset \mathbb{D}} p_t^D\, f(t, X_t^{\xi,b})\, h\big(t, J(a^D, p_t)\big)\Bigg]\, \mathrm{d}t
\end{aligned}$$

$$+\int_{E^d} f(t, X_{t-}^{\xi,b})\, h\big(t, J(p_{t-}, z)\big)\ \exp\left\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_z(i)\right\} \widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z))$$
$$-\int_{E^d} f(t, X_{t-}^{\xi,b})\, h(t-, p_{t-})\ \exp\left\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_z(i)\right\} \widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z))$$
$$-\int_{\mathcal{P}(\mathbb{D})} f(t, X_{t-}^{\xi,b})\, \Big(h\big(t, J(p_{t-}, z)\big) - h(t-, p_{t-})\Big)\, \widehat{\Phi}(\mathrm{d}t, \mathrm{d}z),$$

which reduces to

$$\begin{aligned}
\mathrm{d}G(t, X_t^{\xi,b}, p_t) = &\Bigg[h(t,p_t)\, f(t, X_t^{\xi,b})\, \alpha\, e^{r(T-t)} \Big(\frac{1}{2}\sigma^2 \alpha e^{r(T-t)} \xi_t^2 - (\mu - r)\xi_t - c(b_t)\Big)\\
&+ \lambda \sum_{D\subset\mathbb{D}} p_t^D\, f(t, X_t^{\xi,b})\, h\big(t, J(p_t, D)\big) \int_{(0,\infty)^d} \exp\left\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_D(i)\right\} F(\mathrm{d}y)\\
&- \lambda\, h(t,p_t)\, f(t, X_t^{\xi,b}) + f(t, X_t^{\xi,b})\, h_t(t,p_t)\Bigg]\, \mathrm{d}t\\
&- h(t-, p_{t-})\, f(t, X_{t-}^{\xi,b})\, \sigma\, \alpha\, e^{r(T-t)} \xi_t\, \mathrm{d}W_t\\
&+ \int_{E^d} f(t, X_{t-}^{\xi,b}) \Bigg(h\big(t, J(p_{t-}, z)\big)\ \exp\left\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_z(i)\right\}\\
&\qquad - h(t-, p_{t-})\Bigg) \widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z)).
\end{aligned}$$

Thus

$$\begin{aligned}
&\mathrm{d}G(t, X_t^{\xi,b}, p_t)\\
&= f(t, X_t^{\xi,b}) \Bigg[- \lambda\, h(t,p_t) - \alpha\, e^{r(T-t)} h(t,p_t) \Big((\mu - r)\xi_t + c(b_t) - \frac{1}{2}\sigma^2 \alpha e^{r(T-t)} \xi_t^2\Big)\\
&\quad + \lambda \sum_{D\subset\mathbb{D}} p_t^D\, h\big(t, J(p_t, D)\big) \int_{(0,\infty)^d} \exp\left\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_D(i)\right\} F(\mathrm{d}y) + h_t(t,p_t)\Bigg] \mathrm{d}t\\
&\quad + \mathrm{d}\eta_t^{\xi,b},
\end{aligned}$$

where

$$\eta_t^{\xi,b} := \hat{\eta}_t^{\xi,b} - \tilde{\eta}_t^{\xi,b}, \quad t \in [0,T], \tag{A.9}$$

with

$$\begin{aligned}
\hat{\eta}_t^{\xi,b} := \int_0^t \int_{E^d} f(s, X_{s-}^{\xi,b}) \Bigg(h\big(s, J(p_{s-}, z)\big)\ \exp\left\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^{d} y_i \mathbb{1}_z(i)\right\}&\\
- h(s-, p_{s-})\Bigg) \widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z)), \quad t \in [0,T],&
\end{aligned} \tag{A.10}$$

and

$$\tilde{\eta}_t^{\xi,b} := \int_0^t h(s-, p_{s-})\, f(s, X_{s-}^{\xi,b})\, \sigma\, \alpha\, e^{r(T-s)} \xi_s\, \mathrm{d}W_s, \quad t \in [0,T]. \tag{A.11}$$

By the absolute continuity of $t \mapsto h(t,p)$ for all $p \in \Delta_m$, the partial derivative of h w.r.t. t exists almost everywhere in the sense of the Lebesgue measure. Therefore, due to the definition of the operator $\mathscr{H}$ given by (4.41) and relation (4.9), we obtain

$$\mathrm{d}G(t, X_t^{\xi,b}, p_t) + f(t, X_t^{\xi,b})\, \mathscr{H}h(t, p_t; \xi_t, b_t)\, \mathrm{d}t + \mathrm{d}\eta_t^{\xi,b},$$

where we set $\mathscr{H}h(t, p_t, \xi_t, b_t)$ to zero at those points $t \in [0, T]$ where the partial derivative $h_t(t, p_t)$ does not exist. The only point remaining concerns the martingale behaviour of $(\eta_t^{\xi,b})_{t\in[0,T]}$. Appealing to Corollary 2.98, the process $(\hat{\eta}_t^{\xi,b})_{t\in[0,T]}$ is a $\mathfrak{G}$-martingale if the function

$$F : [0, T] \times \Omega \times (0, \infty)^d \times \mathcal{P}(\mathbb{D}) \to \mathbb{R}$$

defined by

$$F(t, \omega, y, z) := f(t, X_{t-}^{\xi,b}(\omega))\bigg(h\big(t, J(p_{t-}(\omega), z)\big) \exp\bigg\{\alpha\, b_t(\omega)\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_z(i)\bigg\} \\ - h(t-, p_{t-}(\omega))\bigg),$$

is a $\mathfrak{G}$-predictable function indexed by $E^d = (0, \infty)^d \times \mathcal{P}(\mathbb{D})$ and holds

$$\mathbb{E}\bigg[\int_0^T \int_{E^d} \big|F(s, y, z)\big|\, \lambda \sum_{D\in \mathrm{d}z} p_s^D\, F(\mathrm{d}y)\, \mathrm{d}s\bigg] < \infty.$$

We begin with verifying that F is a $\mathfrak{G}$-predictable function indexed by $(0, \infty)^d \times \mathcal{P}(\mathbb{D})$. First we observe that $\Delta_m \ni p \mapsto h(t, p)$ is $\mathcal{B}(\Delta_m)$-measurable due to the assumed continuity of $p \mapsto h(t, p)$. Since $t \mapsto h(t, p)$ is also continuous, h is a Carathéodory function and thus $\mathcal{B}([0, T]) \otimes \mathcal{B}(\Delta_m)$-measurable, cf. Def. 4.50 and Lemma 4.51 in Aliprantis and Border [7]. Next, we define a function

$$g : [0, T] \times \Omega \times \mathcal{P}(\mathbb{D}) \to [0, T] \times \Delta_m$$

by

$$g(t, \omega, z) = \begin{pmatrix} t \\ J(p_{t-}(\omega), z) \end{pmatrix}.$$

Notice that the function $\Delta_m \ni p \mapsto J(p, z)$ is continuous in every component. Hence $J : \Omega \times \Delta_m \times \mathcal{P}(\mathbb{D}) \to \Delta_m$ is $\mathcal{B}(\Delta_m) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable and thus $(t, \omega, z) \mapsto J(p_{t-}(\omega), z)$ is $\mathcal{P}(\mathfrak{G}) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable due to the $\mathfrak{G}$-predictability of $(p_t)_{t\geq 0}$. Obviously, the first component of the vector above is $\mathcal{B}([0, T])$-measurable, in particular $\mathcal{P}(\mathfrak{G}) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable. Hence g is $\mathcal{P}(\mathfrak{G}) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable, compare Klenke [77, Thm. 1.90]. Since $(h \circ g)(t, \omega, z) = h(t, J(p_{s-}(\omega), z))$ and h is $\mathcal{B}([0, T]) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable, it follows that $(t, \omega, z) \mapsto h(t, J(p_{t-}(\omega), z))$ is $\mathcal{P}(\mathfrak{G}) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable, in particular $\mathcal{P}(\mathfrak{G}) \otimes \mathcal{B}((0, \infty)^d) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable. Moreover, it is easily seen that $(0, \infty)^d \ni y \mapsto F(t, \omega, y, z)$ is continuous in every component for all $(t, \omega, z) \in [0, T] \times \Omega \times \mathcal{P}(\mathbb{D})$. Hence, F is a Carathéodory function. Thus F is $\mathcal{P}(\mathfrak{G}) \otimes \mathcal{B}((0, \infty)^d) \otimes \mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable, i.e. F is a $\mathfrak{G}$-predictable process indexed by $(0, \infty)^d \times \mathcal{P}(\mathbb{D}) = E^d$. Next we can turn our attention to the finiteness of the expectation above. Using the triangle inequality, Assumption 4.3, the boundedness of h with constant K_0 and Remark 4.8 (ii), we see that

$$\mathbb{E}\bigg[\int_0^T \int_{E^d} |F(s, y, z)|\, \lambda \sum_{D\in \mathrm{d}z} p_s^D\, F(\mathrm{d}y)\, \mathrm{d}s\bigg]$$

$$= \lambda \sum_{D\subset\mathbb{D}} \mathbb{E}\bigg[\int_0^T \int_{(0,\infty)^d} \Big| f(s,X_s^{\xi,b})\Big(h\big(s,J(p_s,D)\big) \exp\Big\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\} - h(s,p_s)\Big)\Big|\, p_s^D\, F(\mathrm{d}y)\,\mathrm{d}s\bigg]$$

$$\le \lambda \int_{(0,\infty)^d} \exp\Big\{\alpha\, e^{|r|T} \sum_{i=1}^d y_i\Big\} F(\mathrm{d}y)\, \mathbb{E}\bigg[\int_0^T \sum_{D\subset\mathbb{D}} p_s^D \Big| f(s,X_s^{\xi,b})\, h\big(s,J(p_s,D)\big)\Big|\mathrm{d}s\bigg]$$

$$+ \lambda\, \mathbb{E}\bigg[\int_0^T \Big| f(s,X_s^{\xi,b})\, h(s,p_s)\Big|\,\mathrm{d}s\bigg]$$

$$\le \lambda\, M_F\big(\alpha e^{|r|T}\big) K_0 \mathbb{E}\bigg[\int_0^T \big|f(s,X_s^{\xi,b})\big|\,\mathrm{d}s\bigg] + \lambda K_0 \mathbb{E}\bigg[\int_0^T \big|f(s,X_s^{\xi,b})\big|\,\mathrm{d}s\bigg]$$

$$= \lambda\Big(M_F\big(\alpha e^{|r|T}\big) + 1\Big)\lambda\, K_0\, \mathbb{E}\bigg[\int_0^T \big|f(s,X_s^{\xi,b})\big|\,\mathrm{d}s\bigg]. \tag{A.12}$$

Due to Lemma A.3, Lemma A.4 and Fubini's theorem, we obtain

$$\mathbb{E}\bigg[\int_0^T \big|f(s,X_s^{\xi,b})\big|\,\mathrm{d}s\bigg] = \int_0^T \mathbb{E}_{\mathbb{Q}_s^{\xi,b}}\bigg[\frac{\big|f(s,X_s^{\xi,b})\big|}{L_s^{\xi,b}}\bigg]\mathrm{d}s \le K_1\, T,$$

which yields the desired finiteness of (A.12). To see the martingale property of $(\tilde{\eta}_t^{\xi,b})_{t\in[0,T]}$, we have to show that the process $H = (H_t)_{t\in[0,T]}$ defined by

$$H_t := h(t-,p_{t-})\, f(t,X_{t-}^{\xi,b})\, \sigma\, \alpha e^{r(T-t)} \xi_t, \quad t\in[0,T],$$

is $\mathfrak{G}$-progressively measurable and satisfies $\int_0^T \mathbb{E}[H_s^2]\,\mathrm{d}s < \infty$, cf. Theorem 4.7 in Klebaner [75]. It is easily seen that H is $\mathfrak{G}$-progressively measurable due to the càdlàg property and adaptedness of ξ as well as the $\mathfrak{G}$-predictability of $(h(t-,p_{t-})f(t,X_{t-}^{\xi,b}))_{t\ge 0}$ (which follows from the continuity of h and f in both components). Moreover, by the boundedness of h with K_0 and ξ with K as well as Corollary A.5, we have

$$\mathbb{E}[H_s^2] \le \alpha^2 e^{2|r|T} K_0^2 K^2 K_2.$$

In summary, we can make the desired conclusion that $(\eta_t^{\xi,b})_{t\in[0,T]}$ is a $\mathfrak{G}$-martingale which starts obviously at zero. □

A.2 Auxiliary results to Section 5.6

The following result will be used to provide a change of measure in Lemma A.10

Lemma A.9. *Let $b = (b_t)_{t\in[0,T]}$ be some $\mathfrak{F}^\Psi$-predictable reinsurance strategy. Furthermore, let $B^b = (B_t^b)_{t\in[0,T]}$ be the process which is defined by*

$$B_t^b := \int_0^t \int_{E^d} \Big(\exp\Big\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big\} - 1\Big)\widehat{\Psi}(\mathrm{d}s,\mathrm{d}(y,z)). \tag{A.13}$$

Then the stochastic exponential $\mathcal{E}(B^b) = (\mathcal{E}(B^b)_t)_{t\in[0,T]}$ *of* B^b *is given by*

$$\begin{aligned}\mathcal{E}(B^b)_t =& \exp\bigg\{\int_0^t\int_{E^d} \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \Psi(\mathrm{d}s,\mathrm{d}(y,z)) + \int_0^t \widehat{\Lambda}_s\,\mathrm{d}s \\ &- \int_0^t \widehat{\Lambda}_s \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar\beta + q_s\|} \int_{(0,\infty)^d} \exp\bigg\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\bigg\}\, F(\mathrm{d}y)\,\mathrm{d}s\bigg\}.\end{aligned}$$

Furthermore, $\mathcal{E}(B^b)$ *is an* $\mathfrak{F}^\Psi$*-martingale on* $[0,T]$.

Proof. Fix some $\mathfrak{F}^\Psi$-predictable reinsurance strategy $b = (b_t)_{t\in[0,T]}$. As in the proof of Lemma A.2, we obtain, by definition of $\widehat{\Psi}$ given in (5.9),

$$\begin{aligned}\mathcal{E}(B^b)_t = \exp\bigg\{&\int_0^t\int_{E^d} \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \Psi(\mathrm{d}s,\mathrm{d}(y,z)) \\ &- \int_0^t\int_{E^d} \bigg(\exp\bigg\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\bigg\} - 1\bigg)\nu(\mathrm{d}s,\mathrm{d}(y,z))\bigg\},\end{aligned}$$

which reduces to, by Proposition 5.17,

$$\begin{aligned}\mathcal{E}(B^b)_t =& \exp\bigg\{\int_0^t\int_{E^d} \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \Psi(\mathrm{d}s,\mathrm{d}(y,z)) + \int_0^t \widehat{\Lambda}_s\,\mathrm{d}s \\ &- \int_0^t \widehat{\Lambda}_s \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar\beta + q_s\|} \int_{(0,\infty)^d} \exp\bigg\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\bigg\}\, F(\mathrm{d}y)\,\mathrm{d}s\bigg\}.\end{aligned}$$

The process $\mathcal{E}(B^b) = (\mathcal{E}(B^b)_t)_{t\in[0,T]}$ is obviously $\mathfrak{F}^\Psi$-adapted. By definition of the stochastic exponential, we have

$$\begin{aligned}\mathcal{E}(B^b)_t &= \int_0^t \mathcal{E}(B^b)_{s-}\,\mathrm{d}B^b_s \\ &= \int_0^t\int_{E^d} \mathcal{E}(B^b)_{s-}\bigg(\exp\bigg\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\bigg\} - 1\bigg)\widehat{\Psi}(\mathrm{d}s,\mathrm{d}(y,z)), \quad t\in[0,T].\end{aligned}$$

Therefore, according to Corollary 2.98, $\mathcal{E}(B^b)$ is an $\mathfrak{F}^\Psi$-martingale on $[0,T]$ if

$$\mathbb{E}\bigg[\int_0^T\int_{E^d} \bigg|\mathcal{E}(B^b)_{t-}\bigg(\exp\bigg\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\bigg\} - 1\bigg)\bigg|\nu(\mathrm{d}t,\mathrm{d}(y,z))\bigg] < \infty.$$

Notice that the integrand process above is obviously $\mathfrak{F}^\Psi$-predictable due to the $\mathfrak{F}^\Psi$-predictability of $(b_t)_{t\geq 0}$. By the triangle inequality, we obtain that the expectation above is less or equal to

$$\begin{aligned}\mathbb{E}\bigg[\int_0^T\int_{E^d} &\exp\bigg\{\int_0^t\int_{E^d} \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\, \Psi(\mathrm{d}s,\mathrm{d}(y,z)) + \int_0^t \widehat{\Lambda}_s\,\mathrm{d}s \\ &- \int_0^t \widehat{\Lambda}_s \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar\beta + q_s\|} \int_{(0,\infty)^d} \exp\bigg\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\bigg\}\, F(\mathrm{d}y)\,\mathrm{d}s\bigg\}\times\end{aligned}$$

$$\left(\exp\left\{\alpha\, b_t\, e^{r(T-t)}\sum_{i=1}^d y_i \mathbb{1}_z(i)\right\}+1\right)\widehat{\Lambda}_t \sum_{D\in \mathrm{d}z} \frac{\beta_D+q_D(t)}{\|\bar{\beta}+q_t\|}\, F(\mathrm{d}y)\,\mathrm{d}t\Bigg]$$

$$\leq \mathbb{E}\Bigg[\int_0^T \widehat{\Lambda}_t \exp\left\{\int_0^t\int_{E^d}\alpha\, b_s\, e^{r(T-s)}\sum_{i=1}^d y_i\mathbb{1}_z(i)\,\Psi(\mathrm{d}s,\mathrm{d}(y,z))+\int_0^t \widehat{\Lambda}_s\,\mathrm{d}s\right\}\times$$

$$\sum_{D\subset\mathbb{D}}\frac{\beta_D+q_D(t)}{\|\bar{\beta}+q_t\|}\left(\int_{(0,\infty)^d}\exp\left\{\alpha\, b_t\, e^{r(T-t)}\sum_{i=1}^d y_i\mathbb{1}_z(i)\right\}+1\right)\mathrm{d}t\Bigg].$$

By (5.2), $b_t \leq 1$ and Remark 5.9, we get the following finite upper bound for the expectation above

$$\lambda_m\left(M_F(\alpha e^{|r|T})+1\right)\mathbb{E}\Bigg[\int_0^T \exp\left\{\alpha e^{|r|T}\sum_{i=1}^d\sum_{k=1}^{N_t} Y_k^i+\lambda_m T\right\}\mathrm{d}t\Bigg]$$

$$\leq \lambda_m\left(M_F(\alpha e^{|r|T})+1\right)e^{\lambda_m T}\int_0^T \mathbb{E}\Bigg[\exp\left\{\alpha e^{|r|T}\sum_{i=1}^d\sum_{k=1}^{N_t} Y_k^i\right\}\Bigg]\mathrm{d}t<\infty,$$

where the finiteness follows from Lemma 5.8 (ii). □

Recall the definition of $\widetilde{\mathcal{U}}[t,T]$ given in (5.30).

Lemma A.10. *Let $t\in[0,T]$ and let $(\xi,b)\in\widetilde{\mathcal{U}}[0,T]$ be an arbitrary admissible strategy. We set*

$$L_t^{\xi,b} := \exp\Bigg\{-\int_0^t \alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s-\frac{1}{2}\int_0^t \alpha^2\,\sigma^2\, e^{2r(T-s)}\xi_s^2\,\mathrm{d}s$$

$$+\int_0^t\int_{E^d}\alpha\, b_s\, e^{r(T-s)}\sum_{i=1}^d y_i\mathbb{1}_z(i)\,\Psi(\mathrm{d}s,\mathrm{d}(y,z))+\int_0^t\widehat{\Lambda}_s\,\mathrm{d}s$$

$$-\int_0^t \widehat{\Lambda}_s\sum_{D\subset\mathbb{D}}\frac{\beta_D+q_D(s)}{\|\bar{\beta}+q_s\|}\int_{(0,\infty)^d}\exp\left\{\alpha\, b_s\, e^{r(T-s)}\sum_{i=1}^d y_i\mathbb{1}_D(i)\right\}F(\mathrm{d}y)\,\mathrm{d}s\Bigg\}.$$

Then, a probability measure on $(\Omega,\mathcal{G}_t)$ is defined by $\mathbb{Q}_t^{\xi,b}(A):=\int_A L_t^{\xi,b}\,\mathrm{d}\mathbb{P}$, $A\in\mathcal{G}_t$, for every $t\in[0,T]$, i.e. $\frac{\mathrm{d}\mathbb{Q}_t^{\xi,b}}{\mathrm{d}\mathbb{P}}:=L_t^{\xi,b}$. The probability measures $\mathbb{Q}_t^{\xi,b}$ and $\mathbb{P}$ are equivalent.

Proof. Fix $(\xi,b)\in\widetilde{\mathcal{U}}[0,T]$. We define a process $L^{\xi,b}=(L_t^{\xi,b})_{t\in[0,T]}$ as stochastic exponential of the process $Z^{\xi,b}=(Z_t^{\xi,b})_{t\in[0,T]}$ given by $Z_t^{\xi,b}=A_t^\xi+B_t^b$ with

$$A_t^\xi := -\int_0^t \alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s,$$

$$B_t^b := \int_0^t\int_{E^d}\left(\exp\left\{\alpha\, b_s\, e^{r(T-s)}\sum_{i=1}^d y_i\mathbb{1}_z(i)\right\}-1\right)\widehat{\Psi}(\mathrm{d}s,\mathrm{d}(y,z)).$$

According to Theorem II.38 in Protter [104], we have

$$L_t^{\xi,b}=\mathcal{E}(Z^{\xi,b})_t=\mathcal{E}(A)_t\,\mathcal{E}(B)_t,\quad t\in[0,T],$$

where $\mathcal{E}(A^\xi)=(\mathcal{E}(A^\xi)_t)_{t\in[0,T]}$ and $\mathcal{E}(B^b)=(\mathcal{E}(B^b)_t)_{t\in[0,T]}$ are the Doléans-Dade exponential of $A^\xi=(A_t^\xi)_{t\in[0,T]}$ and $B^b=(B_t^b)_{t\in[0,T]}$, respectively. Therefore, Lemma A.1

and Lemma A.9 imply

$$L_t^{\xi,b} = \exp\bigg\{ -\int_0^t \alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s - \frac{1}{2}\int_0^t \alpha^2\,\sigma^2\, e^{2r(T-s)}\xi_s^2\,\mathrm{d}s + \int_0^t\int_{E^d} \alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\,\Psi(\mathrm{d}s,\mathrm{d}(y,z)) + \int_0^t \widehat{\Lambda}_s\,\mathrm{d}s - \int_0^t \widehat{\Lambda}_s \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar\beta + q_s\|}\int_{(0,\infty)^d} \exp\Big\{\alpha\, b_s\, e^{r(T-s)}\sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\}\,F(\mathrm{d}y)\,\mathrm{d}s\bigg\}.$$

Recall that $\mathcal{E}(A^\xi)$ is a $(\mathbb{P},\mathfrak{F}^W)$-martingale on $[0,T]$ and that $\mathcal{E}(B^b)$ is a $(\mathbb{P},\mathfrak{F}^\Psi)$-martingale on $[0,T]$, compare Lemma A.1 and Lemma A.9, respectively. Recall further that the Brownian motion $(W_t)_{t\ge 0}$ is independent of $(T_n)_{n\in\mathbb{N}}$, $(Y_n)_{n\in\mathbb{N}}$ and $(Z_n)_{n\in\mathbb{N}}$ according to Assumption 3.6. Consequently, $\mathcal{E}(A^\xi)$ and $\mathcal{E}(B^b)$ are independent processes. Therefore, since $\mathcal{G}_t = \mathcal{F}_t^W \vee \mathcal{F}_t^\Psi$, $t\ge 0$, and the product of two independent martingales (each with respect to its natural filtration) is a martingale (with respect to the natural filtration of the product), compare Theorem 2.1 of Chapter "Some particular Problems of Martingale Theory" in Kabanov et al. [72], it follows that $L^{\xi,b}$ is a $(\mathbb{P},\mathfrak{G})$-martingale on $[0,T]$. Thus $\mathbb{E}[L_t^{\xi,b}] = 1$. Therefore, we can define a new measure $\mathbb{Q}_t^{\xi,b}$ on $(\Omega,\mathcal{G}_t)$ by $\mathbb{Q}_t^{\xi,b}(A) = \int_A L_t^{\xi,b}\,\mathrm{d}\mathbb{P}$ for every $A\in\mathcal{G}_t$, where it is easily seen that $\mathbb{Q}_t^{\xi,b}$ and $\mathbb{P}$ are equivalent. □

Lemma A.11. *Let $f:[0,T]\times\mathbb{R}\to\mathbb{R}$ be the function defined by* (A.3)*. Furthermore, let $(\xi,b)\in\widetilde{\mathcal{U}}[0,T]$ and let $L^{\xi,b} = (L_t^{\xi,b})_{t\in[0,T]}$ be the density process of Lemma A.10. Then there exists a constant $0<K_1<\infty$ such that*

$$\frac{\big|f(t,X_t^{\xi,b})\big|}{L_t^{\xi,b}} \le K_1 \quad \mathbb{P}\text{-a.s.}$$

for all $t\in[0,T]$.

Proof. Fix $t\in[0,T]$ and $(\xi,b)\in\widetilde{\mathcal{U}}[0,t]$. A line of arguments as in the proof of Lemma A.4 in connection with $\sum_{D\subset\mathbb{D}}\frac{\beta_D+q_D(s)}{\|\bar\beta+q_s\|} = 1$ and Remark 5.9 yields

$$\begin{aligned}
&\frac{\big|f(t,X_t^{\xi,b})\big|}{L_t^{\xi,b}} \\
&= \exp\bigg\{ -\alpha x_0 e^{rT} - \int_0^t \alpha e^{r(T-s)}\Big((\mu-r)\xi_s + c(b_s) - \frac{1}{2}\alpha\sigma^2 e^{r(T-s)}\xi_s^2\Big)\,\mathrm{d}s \\
&\quad + \int_0^t \widehat{\Lambda}_s \sum_{D\subset\mathbb{D}}\frac{\beta_D+q_D(s)}{\|\bar\beta+q_s\|}\int_{(0,\infty)^d}\exp\Big\{\alpha\, b_s\, e^{r(T-s)}\sum_{i=1}^d y_i\mathbb{1}_D(i)\Big\}\,F(\mathrm{d}y)\,\mathrm{d}s - \int_0^t\widehat{\Lambda}_s\,\mathrm{d}s\bigg\} \\
&\le \exp\bigg\{\Big(\alpha e^{|r|T}\big(|\mu-r|K + (2+\eta+\theta)\kappa + \frac{1}{2}\alpha\sigma^2 e^{|r|T}K^2\big) + \lambda_m M_F\big(\alpha e^{|r|T}\big)\Big)T\bigg\} =: K_1,
\end{aligned}$$

where $0<K_1<\infty$ is independent of $t\in[0,T]$ as well as (ξ,b). □

Corollary A.12. *Let $f:[0,T]\times\mathbb{R}\to\mathbb{R}$ be a function defined by* (A.3)*. Furthermore, let $(\xi,b)\in\widetilde{\mathcal{U}}[0,T]$ and let $\widetilde{L}^{\xi,b} = (\widetilde{L}_t^{\xi,b})_{t\in[0,T]}$ be the density process of Lemma A.10 with*

α replaced by 2α. Then there exists a constant $0 < K_2 < \infty$ such that

$$\frac{\big(f(t, X_t^{\xi,b})\big)^2}{\widetilde{L}_t^{\xi,b}} \leq K_2 \quad \mathbb{P}\text{-a.s.}$$

for all $t \in [0, T]$.

Proof. The assertion follows directly from the proof of Lemma A.11 with that same argument as in the proof of Corollary A.5. □

Lemma A.13. *The function $f : [0, T] \times \mathbb{R} \to \mathbb{R}$ given by* (A.3) *satisfies*

$$\begin{aligned}
f(t, X_t^{\xi,b}) &= f(0, X_0^{\xi,b}) + \int_0^t f\big(s, X_s^{\xi,b}\big)\Big(\alpha e^{r(T-s)}\Big(\frac{1}{2}\alpha\,\sigma^2\, e^{r(T-s)}\xi_s^2 - (\mu - r)\xi_s - c(b_s)\Big) \\
&+ \widehat{\Lambda}_s \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar{\beta} + q_s\|} \int_{(0,\infty)^d} \exp\Big\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y) - \widehat{\Lambda}_s\Big)\,\mathrm{d}s \\
&- \int_0^t f(s, X_{s-}^{\xi,b})\,\alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s \\
&+ \int_0^t \int_{E^d} f(s, X_{s-}^{\xi,b})\Big(\exp\Big\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big\} - 1\Big)\widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y, z)),
\end{aligned}$$

for all $t \in [0, T]$.

Proof. Fix $t \in [0, T]$. Notice that the surplus process $X^{\xi,b}$ given in (5.10) satisfies the properties stated in Proposition 3.15. Therefore, a similar argumentation as in the proof of Lemma A.6 yields

$$\begin{aligned}
f(t, X_t^{\xi,b}) &= f(0, X_0^{\xi,b}) \\
&\quad + \int_0^t f(s, X_s^{\xi,b})\,\alpha\, e^{r(T-t)}\Big(-(\mu - r)\xi_s - c(b_s) + \frac{1}{2}\alpha\,\sigma^2\, e^{r(T-s)}\xi_s^2\Big)\,\mathrm{d}s \\
&\quad - \int_0^t f(s, X_{s-}^{\xi,b})\alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s + \sum_{0<s\leq t}\big(f(s, X_s^{\xi,b}) - f(s, X_{s-}^{\xi,b})\big),
\end{aligned}$$

where, by the definition of compensated random counting measure $\widehat{\Psi}$ given in (5.9),

$$\begin{aligned}
&\sum_{0<s\leq t}\big(f(s, X_s^{\xi,b}) - f(s, X_{s-}^{\xi,b})\big) \\
&= \int_0^t \int_{E^d} \Big(f\Big(s, X_{s-}^{\xi,b} - b_s \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big) - f(s, X_{s-}^{\xi,b})\Big)\widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y, z)) \\
&\quad + \int_0^t \int_{E^d} \Big(f\Big(s, X_{s-}^{\xi,b} - b_s \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big) - f(s, X_{s-}^{\xi,b})\Big)\nu(\mathrm{d}s, \mathrm{d}(y, z)).
\end{aligned}$$

By Proposition 5.17 and (A.4), the last line above is equal to

$$\sum_{D\subset\mathbb{D}} \int_0^t \widehat{\Lambda}_s\Big(\int_{(0,\infty)^d} f\Big(s, X_s^{\xi,b} - b_s \sum_{i=1}^d y_i \mathbb{1}_D(i)\Big) F(\mathrm{d}y) - f(s, X_s^{\xi,b})\Big)\frac{\beta_D + q_D(s)}{\|\bar{\beta} + q_s\|}\,\mathrm{d}s$$

$$= \int_0^t \widehat{\Lambda}_s\, f(s, X_s^{\xi,b}) \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar{\beta} + q_s\|} \int_{(0,\infty)^d} \exp\Big\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y)\,\mathrm{d}s$$
$$- \int_0^t \widehat{\Lambda}_s\, f\big(s, X_s^{\xi,b}\big)\,\mathrm{d}s,$$

which implies the statement of the lemma. □

The next result make use of the functions J and v introduced in (5.5) and (5.1), respectively.

Lemma A.14. *Let $h : [0,T] \times \Delta_m \times \mathbb{N}_0^\ell \to (0,\infty)$ be a function such that $t \mapsto h(t,p,q)$ and $t \mapsto h(t,\phi(t),q)$ with $\phi(0) = p$ are absolutely continuous on $[0,T]$ for all $(p,q) \in \Delta_m \times \mathbb{N}_0^\ell$, and $p \to h(t,p,q)$ is concave for all $(t,q) \in [0,T]\times\mathbb{N}_0^\ell$. Then, for any $t \in [0,T]$,*

$$h(t,p_t,q_t) = h(0,p_0,q_0) + \int_0^t \Big(Dh(s,p_s,q_s) - \widehat{\Lambda}_s\, h(s,p_s,q_s)$$
$$+ \widehat{\Lambda}_s \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar{\beta} + q_s\|} h\big(s, J(p_s), v(q_s,D)\big)\Big)\mathrm{d}s$$
$$+ \int_0^t \int_{\mathcal{P}(\mathbb{D})} \Big(h\big(s, J(p_{s-}), v(q_{s-},z)\big) - h(s,p_{s-},q_{s-})\Big)\, \widehat{\Phi}(\mathrm{d}s,\mathrm{d}z).$$

Proof. Recall that the processes $(p_t)_{t\geq 0}$ and $(q_t)_{t\geq0}$ jump at the arrival times $(T_n)_{n\in\mathbb{N}}$ of the trigger events. In the light of the absolutely continuity of $s \mapsto h(s,\phi(s),q)$ with $\phi(0) = p$ for all $(p,q) \in \Delta_m \times \mathbb{N}_0^\ell$ and the fact that $(q_t)_{t\geq0}$ is a pure jump process, the function $F(t) := h(t,p_t,q_t)$ is absolutely continuous on $[T_{n-1},T_n]$ for every $n \in \mathbb{N}$. Hence, by the FTCL (cf. Sohrab [115, Thm. 11.5.23, 11.5.31]), we have

$$F(t) = F(T_{n-1}) + \int_{T_{n-1}}^t F'(s)\,\mathrm{d}s, \quad t \in [T_{n-1},T_n],$$

where F' is Lebesgue integrable. Since $t \mapsto h(t,p,q)$ is absolutely continuous for all $(p,q) \in \Delta_m \times \mathbb{N}_0^\ell$ and $p \mapsto h(t,p,q)$ is concave for all $(t,q) \in [0,T] \times \mathbb{N}_0^\ell$, the partial derivatives h_t and h_{p_j}, $j = 1,\ldots,m$, exist almost everywhere, compare Theorem 2.2 and Theorem 2.3. Hence

$$F'(t) = h_t(t,p_t,q_t) + \sum_{j=1}^m h_{p_j}(t,p_t,q_t)\dot{\phi}_j(t) \quad \text{for a.a. } t \in (T_{n-1},T_n),\ n \in \mathbb{N}$$

with $\phi(0) = p_{T_{n-1}}$. Using the operator D introduced in (5.32), the representation of $\dot{\phi}_j$ given in (5.6) as well as Proposition 5.14 (i), we obtain

$$F'(t) = Dh(t,p_t,q_t) \quad \text{for a.a. } t \in (T_{n-1},T_n),\ n \in \mathbb{N}.$$

That is,

$$h(t,p_t,q_t) = h(T_{n-1},p_{T_{n-1}},q_{T_{n-1}}) + \int_{T_{n-1}}^{T_n-} Dh(s,p_s,q_s)\,\mathrm{d}s, \quad t \in [T_{n-1},T_n],\ n \in \mathbb{N}.$$

Therefore, for any $t \in [0,T]$

$$h(t,p_t,q_t) = h(0,p_0,q_0) + \int_0^t Dh(s,p_s,q_s)\,\mathrm{d}s + \sum_{0<s\le t}\Big(h(s,p_s,q_s) - h(s,p_{s-},q_{s-})\Big),$$

where, the definition of the compensated random measure $\widehat{\Phi}$ given in (5.8),

$$\begin{aligned}
&\sum_{0<s\le t}\Big(h(s,p_s,q_s) - h(s,p_{s-},q_{s-})\Big)\\
&= \int_0^t \int_{\mathcal{P}(\mathbb{D})} \Big(h\big(s,J(p_{s-}),v(q_{s-},z)\big) - h(s,p_{s-},q_{s-})\Big)\,\widehat{\Phi}(\mathrm{d}s,\mathrm{d}z)\\
&\quad + \int_0^t \widehat{\Lambda}_s \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar{\beta}+q_s\|} h\big(s,J(p_s),v(q_s,D)\big)\,\mathrm{d}s - \int_0^t \widehat{\Lambda}_s\, h(s,p_s,q_s)\,\mathrm{d}s,
\end{aligned}$$

which finishes the proof. □

The next result is crucial for the proof of the Verification Theorem 5.24. It makes use of the notation of the operator $\mathscr{H}$ given by 5.33.

Lemma A.15. *Suppose that $(\xi,b) \in \widetilde{\mathcal{U}}[0,T]$ is an arbitrary strategy and $h : [0,T] \times \Delta_m \times \mathbb{N}_0^\ell \to (0,\infty)$ is a bounded function such that $t \mapsto h(t,p,q)$ and $t \mapsto h(t,\phi(t),q)$ with $\phi(t) = p$ are absolutely continuous on $[0,T]$ for all $(p,q) \in \Delta_m \times \mathbb{N}_0^\ell$ as well as $p \mapsto h(t,p,q)$ is concave for all $(t,q) \in [0,T] \times \mathbb{N}_0^\ell$. Then, the function $G : [0,T] \times \mathbb{R} \times \Delta_m \times \mathbb{N}_0^\ell \to \mathbb{R}$ defined by*

$$G(t,x,p,q) := -e^{-\alpha x e^{r(T-t)}}\, h(t,p,q)$$

satisfies

$$\mathrm{d}G(t,X_t^{\xi,b},p_t,q_t) = -e^{-\alpha X_t^{\xi,b} e^{r(T-t)}}\, \mathscr{H}h(t,p_t,q_t;\xi_t,b_t)\,\mathrm{d}t + \mathrm{d}\eta_t^{\xi,b}, \quad t \in [0,T],$$

where $(\eta_t^{\xi,b})_{t\in[0,T]}$ is a $\mathfrak{G}$-martingale and we set $\mathscr{H}h(t,p,q;\xi,b)$ zero at those points (t,p,q) where Dh does not exist.

Proof. The proof follows closely the proof of Lemma A.8. Let $(\xi,b) \in \widetilde{\mathcal{U}}[0,T]$ and $h : [0,T] \times \Delta_m \times \mathbb{N}_0^\ell \to (0,\infty)$ be some function satisfying the conditions stated in the lemma, where $0 < K_0 < \infty$ is some constant which bounds h, i.e. $|h(t,p,q)| \le K_0$ for all $(t,p,q) \in [0,T] \times \Delta_m \times \mathbb{N}_0^\ell$. Furthermore, we set

$$G(t,x,p,q) := -e^{-\alpha x e^{r(T-t)}}\, h(t,p,q) \quad \text{and} \quad f(t,x) := -e^{-\alpha x e^{r(T-t)}},$$

for any $(t,x,p,q) \in [0,T] \times \mathbb{R} \times \Delta_m \times \mathbb{N}_0^\ell$. Let us fix $t \in [0,T]$. Applying the product rule (compare Theorem 2.59) to $G\big(t,X_t^{\xi,b},p_t,q_t\big) = f\big(t,X_t^{\xi,b}\big)h(t,p_t,q_t)$, we get

$$\begin{aligned}
\mathrm{d}G\big(t,X_t^{\xi,b},p_t,q_t\big) &= h(t,p_{t-},q_{t-})\,\mathrm{d}f\big(t,X_t^{\xi,b}\big) + f\big(t,X_{t-}^{\xi,b}\big)\,\mathrm{d}h(t,p_t,q_t)\\
&\quad + \mathrm{d}\big[f\big(\cdot,X_\cdot^{\xi,b}\big),h(\cdot,p_\cdot,q_\cdot)\big]_t
\end{aligned}$$

and hence, by Lemmata A.13 and A.14,

$$\begin{aligned}
&\mathrm{d}G\big(t, X_t^{\xi,b}, p_t, q_t\big) \\
&= f\big(t, X_t^{\xi,b}\big) h(t, p_t, q_t)\bigg(\alpha e^{r(T-t)}\Big(\frac{1}{2}\alpha\sigma^2 e^{r(T-t)}\xi_t^2 - (\mu - r)\xi_t - c(b_t)\Big) \\
&\quad + \widehat{\Lambda}_t \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(t)}{\|\bar{\beta} + q_t\|} \int_{(0,\infty)^d} \exp\Big\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y) - \widehat{\Lambda}_t\bigg)\mathrm{d}t \\
&\quad - f\big(t, X_{t-}^{\xi,b}\big) h(t, p_{t-}, q_{t-}) \alpha\, \sigma\, e^{r(T-t)} \xi_t \,\mathrm{d}W_t \\
&\quad + \int_{E^d} f\big(t, X_{t-}^{\xi,b}\big) h(t, p_{t-}, q_{t-})\bigg(\exp\Big\{\alpha b_t e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big\} - 1\bigg)\widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y,z)) \qquad \text{(A.14)} \\
&\quad + f\big(t, X_t^{\xi,b}\big)\bigg(Dh(t, p_t, q_t) - \widehat{\Lambda}_t h(t, p_t, q_t) + \\
&\qquad\quad \widehat{\Lambda}_t \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(t)}{\|\bar{\beta} + q_t\|} h\big(t, J(p_t), v(q_t, D)\big)\bigg)\mathrm{d}t \\
&\quad + \int_{\mathcal{P}(\mathbb{D})} f\big(t, X_{t-}^{\xi,b}\big)\Big(h\big(t, J(p_{t-}), v(q_{t-}, z)\big) - h(t, p_{t-}, q_{t-})\Big)\widehat{\Phi}(\mathrm{d}t, \mathrm{d}z) \\
&\quad + \mathrm{d}\big[f\big(\cdot, X_\cdot^{\xi,b}\big), h(\cdot, p_\cdot, q_\cdot)\big]_t.
\end{aligned}$$

Since $h(\cdot, p_\cdot, q_\cdot)$ is an FV process, it holds

$$\big[f\big(\cdot, X_\cdot^{\xi,b}\big) + h(\cdot, p_\cdot, q_\cdot)\big]^c \equiv \big[f\big(\cdot, X_\cdot^{\xi,b}\big)\big]^c \quad \text{und} \quad \big[h(\cdot, p_\cdot, q_\cdot)\big]^c \equiv 0,$$

and, consequently, Proposition 2.51 (iii), (v) and (vi) (compare proof of Lemma A.8 for details) yields

$$\begin{aligned}
&\big[f\big(\cdot, X_\cdot^{\xi,b}\big), h(\cdot, p_\cdot, q_\cdot)\big]_t \\
&= f\big(0, X_0^{\xi,b}\big) h(0, p_0, q_0) + \sum_{0<s\le t} f\big(s, X_s^{\xi,b}\big)\big(h(s, p_s, q_s) - h(s, p_{s-}, q_{s-})\big) \\
&\quad - \sum_{0<s\le t} f\big(s, X_{s-}^{\xi,b}\big)\big(h(s, p_s, q_s) - h(s, p_{s-}, q_{s-})\big) \\
&= f\big(0, X_0^{\xi,b}\big) h(0, p_0, q_0) + \int_0^t \int_{E^d} f\Big(s, X_{s-}^{\xi,b} - b_s \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big) \times \\
&\qquad \Big(h\big(s, J(p_{s-}), v(q_{s-}, z)\big) - h(s, p_{s-}, q_{s-})\Big)\Psi(\mathrm{d}s, \mathrm{d}(y,z)) \\
&\quad - \int_0^t \int_{\mathcal{P}(\mathbb{D})} f\big(s, X_{s-}^{\xi,b}\big)\Big(h(s, J(p_{s-}), v(q_{s-}, z)) - h(s, p_{s-}, q_{s-})\Big)\Phi(\mathrm{d}s, \mathrm{d}z).
\end{aligned}$$

Using the introduced compensated random measures $\widehat{\Phi}$ and $\widehat{\Psi}$ given in (5.8) and (5.9), respectively, as well as Equation (A.4), the variation becomes

$$\begin{aligned}
&\big[f\big(\cdot, X_\cdot^{\xi,b}\big), h(\cdot, p_\cdot, q_\cdot)\big]_t \\
&= f\big(0, X_0^{\xi,b}\big) h(0, p_0, q_0) + \int_0^t \int_{E^d} f\big(s, X_{s-}^{\xi,b}\big) \exp\Big\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big\} \times \\
&\qquad \Big(h\big(s, J(p_{s-}), v(q_{s-}, z)\big) - h(s, p_{s-}, q_{s-})\Big)\widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y,z))
\end{aligned}$$

$$
\begin{aligned}
&-\int_0^t \int_{\mathcal{P}(\mathbb{D})} f\big(s, X_{s-}^{\xi,b}\big)\Big(h\big((s, J(p_{s-}), v(q_{s-}, z)\big) - h(s, p_{s-}, q_{s-})\Big)\widehat{\Phi}(\mathrm{d}s, \mathrm{d}z) \\
&+\int_0^t \widehat{\Lambda}_s\, f\big(s, X_s^{\xi,b}\big) \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar{\beta} + q_s\|} \int_{(0,\infty)^d} \exp\Big\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y)\times \\
&\Big(h\big(s, J(p_s), v(q_s, D)\big) - h(s, p_{s-}, q_{s-})\Big)\,\mathrm{d}s \\
&-\int_0^t \widehat{\Lambda}_s\, f\big(s, X_s^{\xi,b}\big) \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(s)}{\|\bar{\beta} + q_s\|} h(s, J(p_s), v(q_s, D))\,\mathrm{d}s \\
&+\int_0^t \widehat{\Lambda}_s\, f\big(s, X_s^{\xi,b}\big) h(s, p_s, q_s)\,\mathrm{d}s.
\end{aligned}
$$

Substituting this into (A.14), we obtain

$$
\begin{aligned}
&\mathrm{d}G\big(t, X_t^{\xi,b}, p_t, q_t\big) \\
&= f\big(t, X_t^{\xi,b}\big)\bigg(-\alpha\, e^{r(T-t)} h(t, p_t, q_t)\Big((\mu - r)\,\xi_t + c(b_t) - \frac{1}{2}\alpha\,\sigma^2\, e^{r(T-t)} \xi_t^2\Big) \\
&+ \widehat{\Lambda}_t \sum_{D\subset\mathbb{D}} \frac{\beta_D + q_D(t)}{\|\bar{\beta} + q_t\|} h(t, J(p_t), v(q_t, D)) \int_{(0,\infty)^d} \exp\Big\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_D(i)\Big\} F(\mathrm{d}y) \\
&- \widehat{\Lambda}_t\, h(t, p_t, q_t) + Dh(t, p_t, q_t)\bigg)\mathrm{d}t \\
&- f\big(t, X_{t-}^{\xi,b}\big)\, h(t, p_{t-}, q_{t-})\, \alpha\, \sigma\, e^{r(T-t)} \xi_t\, \mathrm{d}W_t - f\big(t, X_{t-}^{\xi,b}\big)\, h(t, p_{t-}, q_{t-})\, \mathrm{d}\widehat{N}_t \\
&+ \int_{E^d} f\big(t, X_{t-}^{\xi,b}\big) \exp\Big\{\alpha\, b_t\, e^{r(T-t)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big\} h(t, J(p_{t-}), v(q_{t-}, z))\widehat{\Psi}(\mathrm{d}t, \mathrm{d}(y, z)).
\end{aligned}
$$

Therefore, by definition of the operator $\mathscr{H}$ given in (5.33), we have

$$
\mathrm{d}G\big(t, X_t^{\xi,b}, p_t, q_t\big) = f\big(t, X_t^{\xi,b}\big)\mathscr{H} h(t, p_t, q_t; \xi_t, b_t)\,\mathrm{d}t + \mathrm{d}\eta_t^{\xi,b},
$$

where

$$
\eta_t^{\xi,b} := \hat{\eta}_t^{\xi,b} - \bar{\eta}_t^{\xi,b} - \tilde{\eta}_t^{\xi,b}
$$

with

$$
\begin{aligned}
\hat{\eta}_t^{\xi,b} &:= \int_0^t \int_{E^d} f\big(s, X_{s-}^{\xi,b}\big) \exp\Big\{\alpha\, b_s\, e^{r(T-s)} \sum_{i=1}^d y_i \mathbb{1}_z(i)\Big\}\times \\
&\qquad h\big(s, J(p_{s-}), v(q_{s-}, z)\big)\widehat{\Psi}(\mathrm{d}s, \mathrm{d}(y, z)), \\
\bar{\eta}_t^{\xi,b} &:= \int_0^t f\big(s, X_{s-}^{\xi,b}\big)\, h(s, p_{s-}, q_{s-})\,\mathrm{d}\widehat{N}_s, \\
\tilde{\eta}_t^{\xi,b} &:= \int_0^t f\big(s, X_{s-}^{\xi,b}\big) h(s, p_{s-}, q_{s-})\, \alpha\, \sigma\, e^{r(T-s)} \xi_s\, \mathrm{d}W_s.
\end{aligned}
$$

To complete the proof we need to show that $(\eta_t^{\xi,b})_{t\in[0,T]}$ is a $\mathfrak{G}$-martingale on $[0, T]$ starting at zero, where the last property is obviously satisfied. According to Corollary 2.98 $(\hat{\eta}_t^{\xi,b})_{t\in[0,T]}$ holds the stated martingale property if the function

$$
F : [0, T] \times \Omega \times (0, \infty)^d \times \mathcal{P}(\mathbb{D}) \to \mathbb{R}
$$

defined by

$$F(t,\omega,y,z) := f\big(t, X^{\xi,b}_{t-}(\omega)\big) \exp\Big\{\alpha\, b_t(\omega)\, e^{r(T-t)} \sum_{i=1}^{d} y_i \mathbb{1}_z(i)\Big\}\times \\ h\big(s, J(p_{t-}(\omega)), v(q_{t-}(\omega),z)\big)$$

is a $\mathfrak{G}$-predictable function indexed by E^d and satisfies

$$\mathbb{E}\Big[\int_0^T \int_{E^d} |F(t,y,z)|\nu(t,\mathrm{d}(y,z))\,\mathrm{d}t\Big] < \infty. \tag{A.15}$$

In order to show the measurability property of F, we first observe that $q \mapsto h(t,p,q)$ is $\mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable and $p \mapsto h(t,p,q)$ is $\mathcal{B}(\Delta_m)$-measurable due to the assumed continuity. Hence $(p,q) \mapsto h(t,p,q)$ is $\mathcal{B}(\Delta_m)\otimes\mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable for all $t \in [0,T]$. We further observe that, by assumption, $t \mapsto h(t,p,q)$ is continuous for all $(p,q) \in \Delta_m \times \mathbb{N}_0^\ell$. That is, h is a Carathéodory function and thus $\mathcal{B}([0,T])\otimes\mathcal{B}(\Delta_m)\otimes\mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable, compare Def. 4.50 and Lemma 4.51 in Aliprantis and Border [7]. Next, we define a function

$$g : [0,T]\times\Omega\times\mathcal{P}(\mathbb{D}) \to [0,T]\times\Delta_m\times\mathbb{N}_0^\ell$$

by

$$g(t,\omega,z) = \begin{pmatrix} t \\ J(p_{t-}(\omega)) \\ v(q_{t-}(\omega),z) \end{pmatrix}.$$

Notice that the function $J : \Delta_m \to \Delta_m$ defined by (5.5) is continuous. Hence $J \circ p_{\cdot-}(\cdot) : [0,T]\times\Omega \to \Delta_m$ is $\mathcal{P}(\mathfrak{G})$-measurable. That is, the second component of the vector above is $\mathcal{P}(\mathfrak{G})$-measurable. Furthermore, it is easily seen that the first component of the vector above is $\mathcal{B}([0,T])$-measurable and the third one $\mathcal{P}(\mathfrak{G})\otimes\mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable. In particular, all components are $\mathcal{P}(\mathfrak{G})\otimes\mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable. Hence g is $\mathcal{P}(\mathfrak{G})\otimes\mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable, compare e.g. Klenke [77, Thm. 1.90]. Since $(h\circ g)(t,\omega,z) = h(t, J(p_{s-}(\omega)), v(q_{t-}(\omega),z))$ and h is $\mathcal{B}([0,T])\otimes\mathcal{B}(\Delta_m)\otimes\mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable, it follows that $(t,\omega,z) \mapsto h(t, J(p_{t-}(\omega)), v(q_{t-}(\omega),z))$ is $\mathcal{P}(\mathfrak{G})\otimes\mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable, in particular $\mathcal{P}(\mathfrak{G})\otimes\mathcal{B}((0,\infty)^d)\otimes\mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable. At the same time, the other terms of the function F are $\mathcal{P}(\mathfrak{G})\otimes\mathcal{B}((0,\infty)^d)\otimes\mathcal{P}(\mathcal{P}(\mathbb{D}))$-measurable as well, according to its continuity. That is, F is a $\mathfrak{G}$-predictable process indexed by $(0,\infty)^d\times\mathcal{P}(\mathbb{D}) = E^d$. Next we show the finiteness of the expectation in (A.15). According to the combined information of Proposition 5.17, the boundedness of h, Remark 5.9, Equation (5.2) and Lemma A.11, an upper bound for the expectation in Equation (A.15) is

$$K_0\mathbb{E}\Big[\int_0^T \widehat{\Lambda}_t\big|f\big(t,X^{\xi,b}_t\big)\big| \int_{(0,\infty)^d} \exp\Big\{\alpha e^{|r|T}\sum_{i=1}^d y_i\Big\}F(\mathrm{d}y)\mathrm{d}t\Big] \\ \le K_0\, M_F(\alpha e^{|r|T})\,\lambda_m \int_0^T \mathbb{E}_{\mathbb{Q}^{\xi,b}_t}\Bigg[\frac{\big|f\big(t,X^{\xi,b}_t\big)\big|}{L^{\xi,b}_t}\Bigg]\mathrm{d}t \le K_0\, M_F(\alpha e^{|r|T})\,\lambda_m\, K_1\, T < \infty.$$

Now we turn our attention to the process $(\bar{\eta}^{\xi,b}_t)_{t\in[0,T]}$ which is a $\mathfrak{G}$-martingale on $[0,T]$ if the process $f(\cdot, X^{\xi,b}_{\cdot-})h(\cdot, p_{\cdot-}, C_{\cdot-})$ is $\mathfrak{G}$-predictable and

$$\mathbb{E}\Big[\int_0^T \big|f(t,X^{\xi,b}_{t-})\,h(t,p_{t-},q_{t-})\big|\,\widehat{\Lambda}_t\,\mathrm{d}t\Big] < \infty.$$

We first show the predictability. Since $(p_{t-})_{t\in[0,T]}$ and $(q_{t-})_{t\in[0,T]}$ are $\mathcal{P}(\mathfrak{G})$-measurable, the process $h(\cdot, p_{\cdot-}, q_{\cdot-})$ is $\mathfrak{G}$-predictable due to the above shown measurability of h. Moreover, by its continuity, $f(\cdot, X^{\xi,b}_{\cdot-})$ is $\mathfrak{G}$-predictable as well which yields the required predictability. Furthermore, the same arguments as above imply

$$\mathbb{E}\Big[\int_0^T \big|f(t, X^{\xi,b}_{t-})\, h(t, p_{t-}, q_{t-})\big|\, \widehat{\Lambda}_t\, \mathrm{d}t\Big] \le K_0\, \lambda_m\, K_1\, T < \infty.$$

The last part of the proof is devoted to the martingale property of $(\tilde{\eta}^{\xi,b}_t)_{t\in[0,T]}$. According to Klebaner [75, Thm. 4.7], $(\tilde{\eta}^{\xi,b}_t)_{t\in[0,T]}$ is a $\mathfrak{G}$-martingale on $[0,T]$ if the process $H = (H_t)_{t\in[0,T]}$ defined by

$$H_t := f\big(t, X^{\xi,b}_{t-}\big) h(t, p_{t-}, q_{t-}) \alpha\sigma e^{r(T-t)} \xi_t$$

is $\mathfrak{G}$-progressively measurable and satisfies $\int_0^T \mathbb{E}[H_t^2]\, \mathrm{d}t < \infty$. From the already shown and the $\mathfrak{G}$-progressive measurability of ξ follows the requested $\mathfrak{G}$-progressive measurability of H. Using Corollary A.12 as well as the boundedness of h and of ξ, we conclude $\int_0^T \mathbb{E}[H_t^2]\, \mathrm{d}t \le \alpha^2\sigma^2 e^{2|r|T} K^2 K_0^2 K_2^2 T < \infty$. In summary, we obtain the desired martingale property of $(\eta^{\xi,b}_t)_{t\ge 0}$ and the proof is complete. □

A.3 Auxiliary results to Section 6.6

The following result will be used to provide a change of measure in Lemma A.17

Lemma A.16. *Let $b = (b_t)_{t\in[0,T]}$ be some $\mathfrak{F}^\Psi$-predictable reinsurance strategy. Furthermore, let $B^b = (B^b_t)_{t\in[0,T]}$ be the process which is defined by*

$$B^b_t := \int_0^t \int_{(0,\infty)} \big(\exp\big\{\alpha\, b_s\, y\, e^{r(T-s)}\big\} - 1\big) \widehat{\Psi}(\mathrm{d}s, \mathrm{d}y).$$

Then the stochastic exponential $\mathcal{E}(B^b) = (\mathcal{E}(B^b)_t)_{t\in[0,T]}$ of B^b is given by

$$\begin{aligned}\mathcal{E}(B^b)_t = \exp\bigg\{ &\int_0^t \int_{(0,\infty)} \alpha\, b_s\, y\, e^{r(T-s)}\, \Psi(\mathrm{d}s, \mathrm{d}y) + \lambda \int_0^t \sum_{k=1}^m p_k(s)\, \mu_k\, \mathrm{d}s \\ &- \lambda \int_0^t \sum_{k=1}^m p_k(s) \int_{(0,\infty)} \exp\big\{\alpha\, b_s\, y\, e^{r(T-s)}\big\}\, f_k(y)\, \mathrm{d}y\, \mathrm{d}s\bigg\}.\end{aligned}$$

Furthermore, $\mathcal{E}(B^b)$ is an $\mathfrak{F}^\Psi$-martingale on $[0,T]$.

Proof. Fix some $\mathfrak{F}^\Psi$-predictable reinsurance strategy $b = (b_t)_{t\in[0,T]}$. As in the proof of Lemma A.2, we obtain, by definition of $\widehat{\Psi}$ given in (6.4),

$$\begin{aligned}\mathcal{E}(B^b)_t = \exp\bigg\{ &\int_0^t \int_{(0,\infty)} \alpha\, b_s\, y\, e^{r(T-s)}\, \Psi(\mathrm{d}s, \mathrm{d}y) \\ &- \int_0^t \int_{(0,\infty)} \big(\exp\big\{\alpha\, b_s\, y\, e^{r(T-s)}\big\} - 1\big) \nu(\mathrm{d}s, \mathrm{d}y)\bigg\},\end{aligned}$$

which reduces to, by Proposition 6.8,

$$\mathcal{E}(B^b)_t = \exp\bigg\{ \int_0^t \int_{(0,\infty)} \alpha\, b_s\, y\, e^{r(T-s)}\, \Psi(\mathrm{d}s, \mathrm{d}y) + \lambda \int_0^t \sum_{k=1}^m p_k(s)\, \mu_k\, \mathrm{d}s - \lambda \int_0^t \sum_{k=1}^m p_k(s) \int_{(0,\infty)} \exp\big\{\alpha\, b_s\, y\, e^{r(T-s)}\big\}\, f_k(y)\, \mathrm{d}y\, \mathrm{d}s \bigg\}.$$

The process $\mathcal{E}(B^b) = (\mathcal{E}(B^b)_t)_{t\in[0,T]}$ is obviously $\mathfrak{F}^{\Psi}$-adapted. By definition of the stochastic exponential, we have

$$\begin{aligned}\mathcal{E}(B^b)_t &= \int_0^t \mathcal{E}(B^b)_{s-}\, \mathrm{d}B^b_s \\ &= \int_0^t \int_{(0,\infty)} \mathcal{E}(B^b)_{s-}\big(\exp\big\{\alpha\, b_s\, y\, e^{r(T-s)}\big\} - 1\big)\widehat{\Psi}(\mathrm{d}s, \mathrm{d}y), \quad t \in [0,T].\end{aligned}$$

Therefore, according to Corollary 2.98, $\mathcal{E}(B^b)$ is an $\mathfrak{F}^{\Psi}$-martingale on $[0,T]$ if

$$\mathbb{E}\bigg[\int_0^T \int_{(0,\infty)} \big|\mathcal{E}(B^b)_{t-}\big(\exp\big\{\alpha\, b_t\, y\, e^{r(T-t)}\big\} - 1\big)\big|\nu(\mathrm{d}t, \mathrm{d}y)\bigg] < \infty.$$

Notice that the integrand process above is obviously $\mathfrak{F}^{\Psi}$-predictable due to the $\mathfrak{F}^{\Psi}$-predictability of $(b_t)_{t\geq 0}$. By the triangle inequality, we obtain that the expectation above is less or equal to

$$\mathbb{E}\bigg[\int_0^T \sum_{k=1}^m p_k(s) \exp\bigg\{\int_0^t \int_{(0,\infty)} \alpha\, b_s\, y\, e^{r(T-s)}\, \Psi(\mathrm{d}s, \mathrm{d}y) + \lambda \int_0^t \sum_{k=1}^m p_k(s)\, \mu_k\, \mathrm{d}s\bigg\}\times \int_{(0,\infty)} \big(\exp\big\{\alpha\, b_t\, e^{r(T-t)}\big\} + 1\big)\, f_k(y)\, \mathrm{d}y\, \mathrm{d}t\bigg].$$

Due to $b_t \leq 1$, $p_k(s) \leq 1$, $k = 1, \ldots, m$, as well as Assumption 6.1 and Remark 6.4, we get the following finite upper bound for the expectation above

$$\sum_{k=1}^m \big(M_{F_k}(\alpha e^{|r|T}) + \mu_k\big)\, e^{\lambda\, m\, \mu_k\, T} \int_0^T \mathbb{E}\bigg[\exp\bigg\{\alpha e^{|r|T} \sum_{k=1}^{N_t} Y_i\bigg\}\bigg]\, \mathrm{d}t < \infty,$$

where the finiteness follows from Lemma 6.3 (ii). □

Recall the definition of $\widetilde{\mathcal{U}}[t,T]$ given in (5.30).

Lemma A.17. *Let $t \in [0,T]$ and let $(\xi, b) \in \widetilde{\mathcal{U}}[0,T]$ be an arbitrary admissible strategy. We set*

$$\begin{aligned} L^{\xi,b}_t := \exp\bigg\{ &- \int_0^t \alpha\, \sigma\, e^{r(T-s)}\xi_s\, \mathrm{d}W_s - \frac{1}{2}\int_0^t \alpha^2\, \sigma^2\, e^{2r(T-s)}\xi_s^2\, \mathrm{d}s \\ &+ \int_0^t \int_{(0,\infty)} \alpha\, b_s\, y\, e^{r(T-s)}\, \Psi(\mathrm{d}s, \mathrm{d}y) + \lambda \int_0^t \sum_{k=1}^m p_k(s)\, \mu_k\, \mathrm{d}s \\ &- \lambda \int_0^t \sum_{k=1}^m \int_{(0,\infty)} \exp\big\{\alpha\, b_s\, y\, e^{r(T-s)}\big\}\, f_k(y)\, \mathrm{d}y\, \mathrm{d}s\bigg\}.\end{aligned}$$

Then, a probability measure on $(\Omega, \mathcal{G}_t)$ *is defined by* $\mathbb{Q}_t^{\xi,b}(A) := \int_A L_t^{\xi,b}\,\mathrm{d}\mathbb{P}$, $A \in \mathcal{G}_t$, *for every* $t \in [0,T]$, *i.e.* $\frac{\mathrm{d}\mathbb{Q}_t^{\xi,b}}{\mathrm{d}\mathbb{P}} := L_t^{\xi,b}$. *The probability measures* $\mathbb{Q}_t^{\xi,b}$ *and* $\mathbb{P}$ *are equivalent.*

Proof. Using Lemma A.1 and Lemma A.16, the proof runs as for Lemma A.10. □

Lemma A.18. *Let* $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ *be the function defined by* (A.3). *Furthermore, let* $(\xi, b) \in \widetilde{\mathcal{U}}[0,T]$ *and let* $L^{\xi,b} = (L_t^{\xi,b})_{t\in[0,T]}$ *be the density process of Lemma A.17. Then there exists a constant* $0 < K_1 < \infty$ *such that*

$$\frac{\big|f(t, X_t^{\xi,b})\big|}{L_t^{\xi,b}} \le K_1 \quad \mathbb{P}\text{-}a.s.$$

for all $t \in [0,T]$.

Proof. Fix $t \in [0,T]$ and $(\xi, b) \in \widetilde{\mathcal{U}}[0,t]$. In the case of one LoB ($d = 1$), Proposition 3.14 yields

$$\begin{aligned}\big|f(t, X_t^{\xi,b})\big| = \exp\bigg\{ &- \alpha\, x_0\, e^{r(T-t)} e^{rt} - \alpha\, e^{r(T-t)} \int_0^t e^{r(t-s)} \big((\mu - r)\,\xi_s + c(b_s)\big)\,\mathrm{d}s \\ &- \alpha\, e^{r(T-t)} \int_0^t \sigma\, e^{r(t-s)}\, \xi_s\, \mathrm{d}W_s + \alpha\, e^{r(T-t)} \int_0^t \int_{(0,\infty)} e^{r(t-s)}\, b_s\, y\, \Psi(\mathrm{d}s, \mathrm{d}y)\bigg\}.\end{aligned}$$

Consequently, we obtain with the help of Lemma A.17, Assumption 6.1 and Remark 6.4

$$\begin{aligned}&\frac{\big|f(t, X_t^{\xi,b})\big|}{L_t^{\xi,b}} \\ &= \exp\bigg\{ - \alpha x_0 e^{rT} + \int_0^t \bigg(- \alpha e^{r(T-s)} \Big((\mu - r)\,\xi_s + c(b_s) - \frac{1}{2}\alpha\,\sigma^2\, e^{r(T-s)} \xi_s^2\Big) \\ &\quad - \lambda \sum_{k=1}^m p_k(s)\, \mu_k + \lambda \sum_{k=1}^m p_k(s) \int_{(0,\infty)} \exp\big\{\alpha\, b_s\, y\, e^{r(T-s)}\big\}\, f_k(y)\,\mathrm{d}y\bigg)\,\mathrm{d}s\bigg\} \\ &\le \exp\bigg\{\bigg(\alpha e^{|r|T} \big(|\mu - r|K + (2{+}\eta{+}\theta)\kappa + \frac{1}{2}\alpha\sigma^2 e^{|r|T} K^2\big) + \lambda \sum_{k=1}^m M_{F_k}\big(\alpha e^{|r|T}\big)\bigg) T\bigg\} =: K_1,\end{aligned}$$

where $0 < K_1 < \infty$ is independent of $t \in [0,T]$ as well as (ξ, b). □

Corollary A.19. *Let* $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ *be a function defined by* (A.3). *Furthermore, let* $(\xi, b) \in \widetilde{\mathcal{U}}[0,T]$ *and let* $\widetilde{L}^{\xi,b} = (\widetilde{L}_t^{\xi,b})_{t\in[0,T]}$ *be the density process of Lemma A.17 with* α *replaced by* 2α. *Then there exists a constant* $0 < K_2 < \infty$ *such that*

$$\frac{\big(f(t, X_t^{\xi,b})\big)^2}{\widetilde{L}_t^{\xi,b}} \le K_2 \quad \mathbb{P}\text{-}a.s.$$

for all $t \in [0,T]$.

Proof. The assertion follows directly from the proof of Lemma A.18 with that same argument as in the proof of Corollary A.5. □

Lemma A.20. *The function* $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ *given by* (A.3) *satisfies*

$$f(t, X_t^{\xi,b}) = f(0, X_0^{\xi,b}) + \int_0^t f\big(s, X_s^{\xi,b}\big)\Big(\alpha e^{r(T-s)}\Big(\frac{1}{2}\alpha\,\sigma^2\, e^{r(T-s)}\xi_s^2 - (\mu - r)\xi_s - c(b_s)\Big)$$
$$+ \lambda \sum_{k=1}^m p_k(s) \int_{(0,\infty)} \exp\big\{\alpha\, b_s\, y\, e^{r(T-s)}\big\}\, f_k(y)\,\mathrm{d}y - \lambda \sum_{k=1}^m p_k(s)\,\mu_k\Big)\,\mathrm{d}s$$
$$- \int_0^t f(s, X_{s-}^{\xi,b})\,\alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s$$
$$+ \int_0^t \int_{(0,\infty)} f(s, X_{s-}^{\xi,b})\big(\exp\big\{\alpha\, b_s\, y\, e^{r(T-s)}\big\} - 1\big)\widehat{\Psi}(\mathrm{d}s, \mathrm{d}y),$$

for all $t \in [0,T]$.

Proof. Fix $t \in [0,T]$. As in the proof of Lemma A.6, we obtain yields

$$f(t, X_t^{\xi,b}) = f(0, X_0^{\xi,b})$$
$$+ \int_0^t f(s, X_s^{\xi,b})\,\alpha\, e^{r(T-t)}\Big(- (\mu - r)\xi_s - c(b_s) + \frac{1}{2}\alpha\,\sigma^2\, e^{r(T-s)}\xi_s^2\Big)\,\mathrm{d}s$$
$$- \int_0^t f(s, X_{s-}^{\xi,b})\alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s + \sum_{0<s\le t}\big(f(s, X_s^{\xi,b}) - f(s, X_{s-}^{\xi,b})\big),$$

where, by the definition of compensated random counting measure $\widehat{\Psi}$ given in (6.4),

$$\sum_{0<s\le t}\big(f(s, X_s^{\xi,b}) - f(s, X_{s-}^{\xi,b})\big)$$
$$= \int_0^t \int_{(0,\infty)} \big(f(s, X_{s-}^{\xi,b} - b_s\, y) - f(s, X_{s-}^{\xi,b})\big)\widehat{\Psi}(\mathrm{d}s, \mathrm{d}y)$$
$$+ \lambda \int_0^t \sum_{k=1}^m p_k(s) \int_{(0,\infty)} f(s, X_s^{\xi,b} - b_s\, y)\, f_k(y)\,\mathrm{d}y\,\mathrm{d}s - \lambda \int_0^t f(s, X_s^{\xi,b}) \sum_{k=1}^m p_k(s)\,\mu_k\,\mathrm{d}s$$
$$= \int_0^t \int_{(0,\infty)} f(s, X_{s-}^{\xi,b})\big(\exp\big\{\alpha\, b_s\, y\, e^{r(T-s)}\big\} - 1\big)\widehat{\Psi}(\mathrm{d}s, \mathrm{d}y)$$
$$+ \lambda \int_0^t f(s, X_s^{\xi,b}) \sum_{k=1}^m p_k(s) \int_{(0,\infty)} \exp\big\{\alpha\, b_s\, y\, e^{r(T-s)}\big\}\, f_k(y)\,\mathrm{d}y\,\mathrm{d}s$$
$$- \lambda \int_0^t f(s, X_s^{\xi,b}) \sum_{k=1}^m p_k(s)\,\mu_k\,\mathrm{d}s,$$

since $f(t, x - b\,y) = f(t,x)\,\exp\{\alpha\, b\, y\, e^{r(T-t)}\}$. □

The next result make use of the function J introduced in (6.3).

Lemma A.21. *Let* $h : [0,T] \times \Delta_m \to (0,\infty)$ *be a function such that* $t \mapsto h(t,p)$ *is absolutely continuous on* $[0,T]$ *for all* $p \in \Delta_m$. *Then*

$$h(t, p_t) = h(0, p_0) + \int_0^t \Big(h_t(s, p_s) - \lambda\, h(s, p_s) \sum_{k=1}^m p_k(s)\,\mu_k$$

$$+\lambda\sum_{k=1}^{m}p_k(s)\int_{(0,\infty)}h(s,J(p_s,y))\,f_k(y)\,\mathrm{d}y\Big)\,\mathrm{d}s$$
$$+\int_0^t\int_{(0,\infty)}\big(h(s,J(p_{s-},y))-h(s,p_{s-})\big)\,\widehat{\Psi}(\mathrm{d}s,\mathrm{d}y),\quad t\in[0,T].$$

Proof. Applying the arguments from the proof of Lemma A.7, we get

$$h(t,p_t)=h(0,p_0)+\int_0^t h_t(s,p_s)\,\mathrm{d}s+\sum_{0<s\le t}\big(h(s,p_s)-h(s,p_{s-})\big),\quad t\in[0,T],$$

where, by the definition compensated random measure $\widehat{\Psi}$ in (6.4),

$$\sum_{0<s\le t}\big(h(s,p_s)-h(s-,p_{s-})\big)$$
$$=\int_0^t\int_{(0,\infty)}\Big(h(s,J(p_{s-},y))-h(s,p_{s-})\Big)\,\widehat{\Psi}(\mathrm{d}s,\mathrm{d}y)$$
$$+\lambda\int_0^t\sum_{k=1}^{m}p_k(s)\int_{(0,\infty)}h(s,J(p_s,y))\,f_k(y)\,\mathrm{d}y\,\mathrm{d}s-\lambda\int_0^t h(s,p_s)\sum_{k=1}^{m}p_k(s)\,\mu_k\,\mathrm{d}s,$$

which yields the assertion. □

The next result is crucial for the proof of the Verification Theorem 6.15. It makes use of the notation of the operator $\mathscr{H}$ given by 6.20.

Lemma A.22. *Suppose that $(\xi,b)\in\widetilde{\mathcal{U}}[0,T]$ is an arbitrary strategy and $h:[0,T]\times\Delta_m\to(0,\infty)$ is a bounded function such that $t\mapsto h(t,p)$ is absolutely continuous on $[0,T]$ for all $p\in\Delta_m$ and $p\mapsto h(t,p)$ is continuous on Δ_m for all $t\in[0,T]$. Then, the function $G:[0,T]\times\mathbb{R}\times\Delta_m\to\mathbb{R}$ defined by*

$$G(t,x,p):=-e^{-\alpha x e^{r(T-t)}}\,h(t,p)$$

satisfies

$$\mathrm{d}G(t,X_t^{\xi,b},p_t)=-e^{-\alpha X_t^{\xi,b}e^{r(T-t)}}\,\mathscr{H}h(t,p_t;\xi_t,b_t)\,\mathrm{d}t+\mathrm{d}\eta_t^{\xi,b},\quad t\in[0,T],$$

where $(\eta_t^{\xi,b})_{t\in[0,T]}$ is a $\mathfrak{G}$-martingale and we set $\mathscr{H}h(t,p;\xi,b)$ to zero at those points (t,p) where the partial derivative of h w.r.t. t does not exist.

Proof. The proof follows closely the proof of Lemma A.8. Fix $(\xi,b)\in\widetilde{\mathcal{U}}[0,T]$ and let $0<K_0<\infty$ be some constant such that $|h(t,p,q)|\le K_0$ for all $(t,p)\in[0,T]\times\Delta_m$, where K_0 exists by assumption. Furthermore, we set $f(t,x):=-e^{-\alpha x e^{r(T-t)}}$ for all $(t,x)\in[0,T]\times\mathbb{R}$. Let us fix $t\in[0,T]$ and apply the product rule (cf. Thm. 2.59) to $G\big(t,X_t^{\xi,b},p_t\big)=f\big(t,X_t^{\xi,b}\big)h(t,p_t)$, which yields

$$\mathrm{d}G\big(t,X_t^{\xi,b},p_t\big)=h(t,p_{t-})\,\mathrm{d}f\big(t,X_t^{\xi,b}\big)+f\big(t,X_{t-}^{\xi,b}\big)\,\mathrm{d}h(t,p_t)+\mathrm{d}\big[f\big(\cdot,X_\cdot^{\xi,b}\big),h(\cdot,p_\cdot)\big]_t$$

Hence, with the help of Lemmata A.20 and A.21, we get

$$
\begin{aligned}
&\mathrm{d}G\big(t, X_t^{\xi,b}, p_t\big) \\
&= f\big(t, X_t^{\xi,b}\big) h(t,p_t) \bigg(\alpha e^{r(T-t)} \Big(\frac{1}{2} \alpha \sigma^2 e^{r(T-t)} \xi_t^2 - (\mu - r)\xi_t - c(b_t) \Big) \\
&\quad + \lambda \sum_{k=1}^m p_k(t) \int_{(0,\infty)} \exp\big\{\alpha\, b_t\, y\, e^{r(T-t)}\big\} f_k(y)\,\mathrm{d}y - \lambda \sum_{k=1}^m p_k(t)\,\mu_k \bigg)\mathrm{d}t \\
&\quad - f\big(t, X_{t-}^{\xi,b}\big) h(t,p_{t-}) \alpha\, \sigma\, e^{r(T-t)} \xi_t \,\mathrm{d}W_t \\
&\quad + \int_{(0,\infty)} f\big(t, X_{t-}^{\xi,b}\big) h(t,p_{t-}) \big(\exp\big\{\alpha\, b_t\, y\, e^{r(T-t)}\big\} - 1\big)\, \widehat{\Psi}(\mathrm{d}t, \mathrm{d}y) \\
&\quad + f\big(t, X_t^{\xi,b}\big) \bigg(h_t(t,p_t) - \lambda h(t,p_t) \sum_{k=1}^m p_k(t)\,\mu_k \\
&\qquad\qquad + \lambda \sum_{k=1}^m p_k(t) \int_{(0,\infty)} h(t, J(p_t,y))\, f_k(y)\,\mathrm{d}y \bigg)\mathrm{d}t \\
&\quad + \int_{(0,\infty)} f\big(t, X_{t-}^{\xi,b}\big) \Big(h(t, J(p_{t-},y)) - h(t,p_{t-}) \Big) \widehat{\Psi}(\mathrm{d}t, \mathrm{d}y) \\
&\quad + \mathrm{d}\big[f\big(\cdot, X_\cdot^{\xi,b}\big), h(\cdot, p_\cdot)\big]_t .
\end{aligned}
\tag{A.16}
$$

Notice that $h(\cdot, p_\cdot, q_\cdot)$ is an FV process and thus

$$
\big[f\big(\cdot, X_\cdot^{\xi,b}\big) + h(\cdot,p_\cdot)\big]^c \equiv \big[f\big(\cdot, X_\cdot^{\xi,b}\big)\big]^c \quad \text{und} \quad \big[h(\cdot,p_\cdot)\big]^c \equiv 0,
$$

Consequently, Proposition 2.51 (iii), (v) and (vi) (compare proof of Lemma A.8 for details) implies

$$
\begin{aligned}
&\big[f\big(\cdot, X_\cdot^{\xi,b}\big), h(\cdot, p_\cdot, q_\cdot)\big]_t \\
&= f\big(0, X_0^{\xi,b}\big) h(0,p_0) + \sum_{0<s\le t} f\big(s, X_s^{\xi,b}\big)\big(h(s,p_s) - h(s,p_{s-})\big) \\
&\quad - \sum_{0<s\le t} f\big(s, X_{s-}^{\xi,b}\big)\big(h(s,p_s) - h(s,p_{s-})\big) \\
&= f\big(0, X_0^{\xi,b}\big) h(0,p_0) + \int_0^t \int_{(0,\infty)} f\Big(s, X_{s-}^{\xi,b} - b_s\, y\Big) \Big(h(s, J(p_{s-},y)) - h(s,p_{s-}) \Big) \Psi(\mathrm{d}s, \mathrm{d}y) \\
&\quad - \int_0^t \int_{(0,\infty)} f\big(s, X_{s-}^{\xi,b}\big) \Big(h(s, J(p_{s-},y)) - h(s,p_{s-}) \Big) \Psi(\mathrm{d}s, \mathrm{d}y).
\end{aligned}
$$

Using the introduced compensated random measures $\widehat{\Psi}$ given in (6.4) as well as the relation $f(t, x - b\,y) = f(t,x)\exp\{\alpha\, b\, y\, e^{r(T-t)}\}$, the variation becomes

$$
\begin{aligned}
&\mathrm{d}\big[f\big(\cdot, X_\cdot^{\xi,b}\big), h(\cdot, p_\cdot, q_\cdot)\big]_t \\
&= \int_{(0,\infty)} f\big(t, X_{t-}^{\xi,b}\big) \Big(h(t, J(p_{t-},y)) - h(t,p_{t-}) \Big) \Big(\exp\big\{\alpha\, b_t\, y\, e^{r(T-t)}\big\} - 1 \Big) \widehat{\Psi}(\mathrm{d}t, \mathrm{d}y) \\
&\quad + \lambda\, f\big(t, X_t^{\xi,b}\big) \sum_{k=1}^m p_k(t) \int_{(0,\infty)} h(t, J(p_t,y))\, \exp\big\{\alpha\, b_t\, y\, e^{r(T-t)}\big\}\, f_k(y)\,\mathrm{d}y\,\mathrm{d}t
\end{aligned}
$$

$$
\begin{aligned}
&- \lambda\, f\big(t, X_t^{\xi,b}\big)\, h(t,p_t) \sum_{k=1}^m p_k(t) \int_{(0,\infty)} \exp\big\{\alpha\, b_t\, y\, e^{r(T-t)}\big\}\, f_k(y)\,\mathrm{d}y\,\mathrm{d}t \\
&- \lambda\, f\big(t, X_t^{\xi,b}\big) \sum_{k=1}^m p_k(t) \int_{(0,\infty)} h(t, J(p_t,y))\, f_k(y)\,\mathrm{d}y\,\mathrm{d}t \\
&+ \lambda\, f\big(t, X_t^{\xi,b}\big)\, h(t,p_t) \sum_{k=1}^m p_k(t)\, \mu_k\,\mathrm{d}t.
\end{aligned}
$$

Substituting this into (A.16), we obtain

$$
\begin{aligned}
&\mathrm{d}G\big(t, X_t^{\xi,b}, p_t, q_t\big) \\
&= f\big(t, X_t^{\xi,b}\big)\bigg(- \alpha\, e^{r(T-t)} h(t,p_t)\Big((\mu - r)\,\xi_t + c(b_t) - \frac{1}{2}\alpha\,\sigma^2\, e^{r(T-t)}\xi_t^2\Big) \\
&+ \lambda \sum_{k=1}^m p_k(t) \int_{(0,\infty)} h(t, J(p_t,y))\, \exp\big\{\alpha\, b_t\, y\, e^{r(T-t)}\big\}\, f_k(y)\,\mathrm{d}y \\
&- \lambda\, h(t,p_t) \sum_{k=1}^m p_k(t)\,\mu_k + h_t(t,p_t)\bigg)\mathrm{d}t \\
&- f\big(t, X_{t-}^{\xi,b}\big)\, h(t,p_{t-})\,\alpha\,\sigma\, e^{r(T-t)}\xi_t\,\mathrm{d}W_t - f\big(t, X_{t-}^{\xi,b}\big)\, h(t,p_{t-})\,\mathrm{d}\widehat{N}_t \\
&+ \int_{(0,\infty)} f\big(t, X_{t-}^{\xi,b}\big)\, h(t, J(p_{t-},y))\, \exp\big\{\alpha\, b_t\, y\, e^{r(T-t)}\big\}\,\widehat{\Psi}(\mathrm{d}t,\mathrm{d}y).
\end{aligned}
$$

Therefore, by definition of the operator $\mathscr{H}$ given in (6.20), we have

$$
\mathrm{d}G\big(t, X_t^{\xi,b}, p_t, q_t\big) = f\big(t, X_t^{\xi,b}\big)\mathscr{H} h(t,p_t,q_t;\xi_t,b_t)\,\mathrm{d}t + \mathrm{d}\eta_t^{\xi,b},
$$

where

$$
\eta_t^{\xi,b} := \hat{\eta}_t^{\xi,b} - \bar{\eta}_t^{\xi,b} - \tilde{\eta}_t^{\xi,b}
$$

with

$$
\begin{aligned}
\hat{\eta}_t^{\xi,b} &:= \int_0^t \int_{(0,\infty)} f\big(s, X_{s-}^{\xi,b}\big)\, h(s, J(p_{s-},y))\, \exp\big\{\alpha\, b_s\, y\, e^{r(T-s)}\big\}\,\widehat{\Psi}(\mathrm{d}s,\mathrm{d}y), \\
\bar{\eta}_t^{\xi,b} &:= \int_0^t f\big(s, X_{s-}^{\xi,b}\big)\, h(s,p_{s-})\,\mathrm{d}\widehat{N}_s, \\
\tilde{\eta}_t^{\xi,b} &:= \int_0^t f\big(s, X_{s-}^{\xi,b}\big)\, h(s,p_{s-})\,\alpha\,\sigma\, e^{r(T-s)}\xi_s\,\mathrm{d}W_s.
\end{aligned}
$$

To finish the proof we need to show that $(\eta_t^{\xi,b})_{t\in[0,T]}$ is a $\mathfrak{G}$-martingale on $[0,T]$, which can be obtained as in the proof of Lemma A.8 and the proof is complete. □

Appendix B

Useful inequalities

The first inequality of this chapter is a Lipschitz condition for exponential function.

Proposition B.1 ([22], Prop. 2.3.19)**.** *Let $A > 0$ and $c \leq x \leq y \leq d$. Then*

$$A^c\left(1 - \frac{1}{A}\right)(y - x) \leq A^y - A^x \leq A^d(A - 1)(y - x).$$

For the following inequalities, let E denote some non-empty set.

Proposition B.2. *Let $f, g : E \to \mathbb{R}$ be bounded functions from above. Then*

$$\sup_{x \in E} \big(f(x) + g(x)\big) \leq \sup_{x \in E} f(x) + \sup_{x \in E} g(x).$$

Proof. The statement follows directly from $f(x) + g(x) \leq \sup_{x \in E} f(x) + \sup_{x \in E} g(x)$ for all $x \in E$. □

Remark B.3. The boundedness from above ensures that we obtain a useful inequality. Unless, we would have the trivial equality $\sup_{x \in E}(f(x) + g(x)) \leq \infty$.

Proposition B.4. *Let $f, g : E \to \mathbb{R}$ be bounded functions from below. Then*

$$\inf_{x \in E} \big(f(x) + g(x)\big) \geq \inf_{x \in E} f(x) + \inf_{x \in E} g(x).$$

Proof. We obtain the stated assertion by applying Proposition B.2 to the functions $-f$ and $-g$. □

Proposition B.5. *Let $f, g : E \to \mathbb{R}$ be bounded functions. Then*

$$\left|\sup_{x \in E} f(x) - \sup_{x \in E} g(x)\right| \leq \sup_{x \in E} |f(x) - g(x)|, \qquad \left|\inf_{x \in E} f(x) - \inf_{x \in E} g(x)\right| \leq \sup_{x \in E} |f(x) - g(x)|.$$

Proof. From $f = f - g + g$ and $f - g \leq |f - g|$ on E, it follows by Proposition B.2 that

$$\sup_{x \in E} f(x) \leq \sup_{x \in E} \big(f(x) - g(x)\big) + \sup_{x \in E} g(x) \leq \sup_{x \in E} \big|f(x) - g(x)\big| + \sup_{x \in E} g(x).$$

Thus

$$\sup_{x \in E} f(x) - \sup_{x \in E} g(x) \leq \sup_{x \in E} \big|f(x) - g(x)\big|.$$

By exchanging f and g in the inequality above, we obtain the first stated inequality in the proposition. Using this inequality and $-\inf_{x \in E} f(x) = \sup_{x \in E}(-f(x))$ leads to the second announced inequality. □

The following result can be found in Mitrinovic et al. [95] or the proof of Lemma 4.1 in Liang and Bayraktar [85].

Lemma B.6. *Let* $\alpha_1 \leq \ldots \leq \alpha_n$ *and* $\beta_1 \leq \ldots \leq \beta_n$ *be real numbers and* $(p_1, \ldots, p_n) \in \Delta_n$. *Then*

$$\sum_{j=1}^{n} p_j \alpha_j \beta_j \geq \sum_{j=1}^{n} p_j \alpha_j \sum_{k=1}^{n} p_k \beta_k.$$

Bibliography

[1] Ahmad, S. and Ambrosetti, A. (2015). *A Textbook on Ordinary Differential Equations*, volume 88 of *Unitext*. Springer, Cham, 2nd edition.

[2] Akcoglu, M., Bartha, P., and Ha, D. (2009). *Analysis in Vector Spaces*. Wiley, Hoboken.

[3] Albrecher, H. and Asmussen, S. (2006). Ruin probabilities and aggregrate claims distributions for shot noise Cox processes. *Scandinavian Actuarial Journal*, 2006(2):86–110.

[4] Albrecher, H. and Asmussen, S. (2010). *Ruin Probabilities*, volume 14 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific, Singapore, 2nd edition.

[5] Albrecher, H., Beirlant, J., and Teugels, J. L. (2017). *Reinsurance: Actuarial and Statistical Aspects*. Wiley Series in Probability and Statistics. Wiley, Oxford.

[6] Albrecht, P. (1981). Über einige Eigenschaften des gemischten Poissonprozesses. *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*, 1:241–250.

[7] Aliprantis, C. D. and Border, K. C. (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, Berlin Heidelberg New York, 3rd edition.

[8] Anastasiadis, S. and Chukova, S. (2012). Multivariate insurance models: An overview. *Insurance: Mathematics and Economics*, 51(1):222–227.

[9] Azcue, P. and Muler, N. (2005). Optimal reinsurance and dividend distribution policies in the Cramér-Lundberg model. *Mathematical Finance*, 15(2):261–308.

[10] Bai, L., Cai, J., and Zhou, M. (2013). Optimal reinsurance policies for an insurer with a bivariate reserve risk process in a dynamic setting. *Insurance: Mathematics and Economics*, 53(3):664–670.

[11] Bai, L. and Guo, J. (2008). Optimal proportional reinsurance and investment with multiple risky assets and no-shorting constraint. *Insurance: Mathematics and Economics*, 42(3):968–975.

[12] Bain, A. and Crisan, D. (2009). *Fundamentals of Stochastic Filtering*, volume 60 of *Stochastic Modelling and Applied Probability*. Springer, New York.

[13] Bellman, R. (1954). The theory of dynamic programming. *Bull. Amer. Math. Soc.*, 60(6):503–515.

[14] Bellman, R. (1957). *Dynamic Programming*. Princeton University Press, Princeton, NJ.

[15] Bühlmann, H. (1970). *Mathematical Methods in Risk Theory*, volume 172 of *A Series of Comprehensive Studies in Mathematics*. Springer, Berlin. 2nd printing 2005.

[16] Bi, J. and Chen, K. (2019). Optimal investment-reinsurance problems with common shock dependent risks under two kinds of premium principles. *RAIRO Operations Research*, 53(1):179–206.

[17] Blatter, A. B. (2008). *Optimal control and dependence modeling of portfolios with Lévy dynamics*. PhD dissertation, University Karlsruhe, Department of Mathematics.

[18] Brachetta, M. and Ceci, C. (2019a). A BSDE-based approach for the optimal reinsurance problem under partial information. Paper 1910.05999, arXiv.org.

[19] Brachetta, M. and Ceci, C. (2019b). Optimal proportional reinsurance and investment for stochastic factor models. *Insurance: Mathematics and Economics*, 87:15–33.

[20] Brémaud, P. (1981). *Point Processes and Queues: Martingale Dynamics*. Springer Series in Statistics. Springer, Berlin.

[21] Brémaud, P. (1988). *An Introduction to Probabilistic Modeling*. Springer Undergraduate Texts in Mathematics and Technology. Springer, Berlin.

[22] Bridger, M. (2007). *Real Analysis: A Constructive Approach*. Wiley, Hoboken, New Jersey.

[23] Brown, M. (1970). An $m/g/\infty$ estimation problem. *Ann. Math. Statist.*, 41(2):651–654.

[24] Bäuerle, N. (1996). Some results about the expected ruin time in Markov-modulated risk models. *Insurance: Mathematics and Economics*, 18(2):119–127.

[25] Bäuerle, N. and Blatter, A. (2011). Optimal control and dependence modeling of insurance portfolios with Lévy dynamics. *Insurance: Mathematics and Economics*, 48(3):398–405.

[26] Bäuerle, N. and Chen, A. (2019). Optimal retirement planning under partial information. *Statistics & Risk Modeling*, 36:37–55.

[27] Bäuerle, N. and Grübel, R. (2005). Multivariate counting processes: Copulas and beyond. *ASTIN Bulletin*, 35(2):379–408.

[28] Bäuerle, N. and Grübel, R. (2008). Multivariate risk processes with interacting intensities. *Advances in Applied Probability*, 40(2):578–601.

[29] Bäuerle, N. and Grether, S. (2017). Extremal behavior of long-term investors with power utility. *International Journal of Theoretical and Applied Finance*, 20(05).

[30] Bäuerle, N. and Rieder, U. (2005). Portfolio optimization with unobservable Markov-modulated drift process. *Journal of Applied Probability*, 42(2):362—-378.

[31] Bäuerle, N. and Rieder, U. (2007). Portfolio optimization with jumps and unobservable intensity process. *Mathematical Finance*, 17(2):205–224.

[32] Bäuerle, N. and Rieder, U. (2011). *Markov Decision Processes with Applications to Finance*. Universitext. Springer, Berlin.

[33] Cambridge Business English Dictionary (n.d.). Definition: Insurance claim. Cambridge University Press, Available online: http://www.dictionary.cambridge.org/dictionary/english/insurance-claim (accessed on 18 December 2019).

[34] Cao, Y. and Wan, N. (2009). Optimal proportional reinsurance and investment based on Hamilton-Jacobi-Bellman equation. *Insurance: Mathematics and Economics*, 45(2):157–162.

[35] Capasso, V. and Bakstein, D. (2005). *An Introduction to Continuous-Time Stochastic Processes: Theory, Models, and Applications to Finance, Biology, and Medicine*. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser, Boston.

[36] Ceci, C. and Colaneri, K. (2012). Nonlinear filtering for jump diffusion observations. *Advances in Applied Probability*, 44(3):678–701.

[37] Centeno, M. d. L. (2005). Dependent risks and excess of loss reinsurance. *Insurance: Mathematics and Economics*, 37(2):229–238. Papers presented at the 8th IME Conference, Rome, 14-16 June 2004.

[38] Chung, K. L. and Williams, R. J. (1990). *Introduction to Stochastic Integration*. Probability and its Applications. Birkhäuser, Boston, 2nd edition.

[39] Clarke, F. H. (1983). *Optimization and nonsmooth analysis*. Canadian Mathematical Society series of monographs and advanced texts - A Wiley-Interscience publication. Wiley, New York.

[40] Cobzaş, Ş., Miculescu, R., and Nicolae, A. (2019). *Lipschitz Functions*, volume 2241 of *Lecture Notes in Mathematics*. Springer, Cham.

[41] Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton.

[42] Cox, D. R. (1955). Some statistical methods connected with series of events. *Journal of the Royal Statistical Society. Series B (Methodological)*, 17(2):129–164.

[43] Cramér, H. (1930). *On the Mathematical Theory of Risk*. Skandia Jubilee Volume, Stockholm.

[44] Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press, Princeton.

[45] Crandall, M. G. and Lions, P.-L. (1983). Viscosity solutions of Hamilton-Jacobi equations. *Transactions of the American Mathematical Society*, 277:1–42.

[46] Czado, C. and Schmidt, T. (2011). *Mathematische Statistik*. Statistik und ihre Anwendungen. Springer, Heidelberg.

[47] Davis, M. H. A. (1993). *Markov Models and Optimization*, volume 49 of *Monographs on statistics and applied probability*. Chapman & Hall, London, 1st edition.

[48] DeGroot, M. and Schervish, M. (2012). *Probability and Statistics*. Addison-Wesley, Boston, 4th edition.

[49] DeGroot, M. H. (1970). *Optimal Statistical Decisions*. McGraw-Hill, New York.

[50] Delbaen, F. and Schachermayer, W. (2006). *The Mathematics of Arbitrage*. Springer Finance. Springer, Berlin, corrected 2nd printing edition.

[51] Desmettre, S. (2010). *Optimal Portfolios for Executive Stockholders*. PhD dissertation, Technische Universität Kaiserslautern, Department of Mathematics.

[52] Di Nunno, G. and Øksendal, B. (2009). Optimal portfolio, partial information and malliavin calculus. *Stochastics: An International Journal of Probability and Stochastics Processes*, 81(3–4):303–322.

[53] Dubourdieu, J. (1938). Remarques relatives à la théorie mathématique de l'assurance-accidents. *Bulletin de l'Institut des Actuaires Français*, 49:76–126.

[54] Elliott, R. J., Lakhdar, A., and Moore, J. B. (1997). *Hidden Markov Models: Estimation and Control*, volume 29 of *Applications of Mathematics*. Springer, New York, corr. print., 2nd edition.

[55] Elliott, R. J., Siu, T. K., and Yang, H. (2011). Ruin theory in a hidden Markov-modulated risk model. *Stochastic Models*, 27(3):474–489.

[56] Elstrodt, J. (2011). *Maß- und Integrationstheorie*. Springer, Berlin, 7th corrected and updated edition.

[57] Fabbri, G., Gozzi, F., and Święch, A. (2017). *Stochastic Optimal Control in Infinite Dimension: Dynamic Programming and HJB Equations*, volume 82 of *Probability Theory and Stochastic Modelling*. with Chapter 6 by M. Fuhrman and G. Tessitore. Springer, Cham.

[58] Faust, E. and Rauch, E. (2019). Series of hot years and more extreme weather. Available online: https://www.munichre.com/topics-online/en/climate-change-and-natural-disasters/climate-change/climate-change-heat-records-and-extreme-weather.html (accessed on 18 December 2019).

[59] Feller, W. (1970). *An Introduction to Probability Theory and its Applications*, volume 1 of *Wiley Series in Probability and Mathematical Statistics*. Wiley, New York, revised printing, 3rd edition.

[60] Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*, 1(2):209–230.

[61] Fleming, W. H. and Soner, H. M. (1993). *Controlled Markov Processes and Viscosity Solutions*, volume 25 of *Applications of mathematics*. Springer, New York.

[62] Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. Springer Series in Statistics. Springer, New York.

[63] Gerber, H. U. (1979). *An Introduction to Mathematical Risk Theory*. In: Huebner Foundation Monograph, vol. 8, Philadelphia.

[64] Gihman, I. I. and Skorochod, A. V. (1979). *Controlled stochastic processes*. Springer, Heidelberg.

[65] Grandell, J. (1976). *Doubly Stochastic Poisson Processes*, volume 529 of *Lecture Notes in Mathematics*. Springer, Berlin.

[66] Grandell, J. (1991). *Aspects of Risk Theory*. Springer Series in Statistics. Springer, New York.

[67] Grandell, J. (1997). *Mixed Poisson Processes*, volume 77 of *Monographs on statistics and applied probability*. Chapman & Hall, London, 1st edition.

[68] Han, X., Liang, Z., and Zhang, C. (2018). Optimal proportional reinsurance with common shock dependence to minimise the probability of drawdown. *Annals of Actuarial Science*, pages 1–27.

[69] Hanson, F. B. (2007). *Applied Stochastic Processes and Control for Jump-Diffsuions: Modeling, Analysis, and Computation*. Advances in Design and Control. Society for Industrial and Applied Mathematics, Philadelphia.

[70] Iglehart, D. L. (1969). Diffusion approximations in collective risk theory. *Journal of Applied Probability*, 6(2):285–292.

[71] Jacobsen, M. (2006). *Point process theory and applications : Marked point and piecewise deterministic processes*. Probability and its Applications. Birkhäuser, Boston.

[72] Kabanov, J. M., Lipster, R., and Stoyanov, J., editors (2006). *From Stochastic Calculus to Mathematical Finance: The Shiryaev Festschrift*. Springer, Berlin.

[73] Karatzas, I. and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer, New York, 2nd edition.

[74] Karr, A. F. (1985). Inference for thinned point processes, with application to cox processes. *Journal of Multivariate Analysis*, 16(3):368–392.

[75] Klebaner, F. C. (2005). *Introduction to stochastic calculus with applications*. Imperial College Pr., London, 2nd edition.

[76] Kleiber, C. and Kotz, S. (2003). *Statistical Size Distributions in Economics and Actuarial Sciences*. Wiley series in probability and statistics. Wiley, Hoboken, NJ.

[77] Klenke, A. (2014). *Probability Theory: A Comprehensive Course*. Universitext. Springer, London, 2nd edition.

[78] Klugman, S. A., Panjer, H. H., and Willmot, G. E. (1998). *Loss Models: From Data to Decisions*. Wiley series in probability and statistics. Wiley, New York, NY.

[79] Kurtz, T. G. and Ocone, D. L. (1988). Unique characterization of conditional distributions in nonlinear filtering. *The Annals of Probability*, 16(1):80–107.

[80] Last, G. and Brandt, A. (1995). *Marked Point Processes on the Real Line: The Dynamical Approach*. Probability and its Applications. Springer, New York.

[81] Last, G. and Penrose, M. (2017). *Lectures on the Poisson Process*. Institute of Mathematical Statistics Textbooks. Cambridge University Press.

[82] Leimcke, G. (2016). Stochastic filtering in non-life insurance. Master thesis, Technische Universität Chemnitz, Faculty of Mathematics.

[83] Leobacher, G., Szölgyenyi, M., and Thonhauser, S. (2014). Bayesian dividend optimization and finite time ruin probabilities. *Stochastic Models*, 30(2):216–249.

[84] Liang, Z., Bai, L., and Guo, J. (2011a). Optimal investment and proportional reinsurance with constrained control variables. *Optimal Control Applications and Methods*, 32(5):587–608.

[85] Liang, Z. and Bayraktar, E. (2014). Optimal reinsurance and investment with unobservable claim size and intensity. *Insurance: Mathematics and Economics*, 55:156 – 166.

[86] Liang, Z. and Yuen, K. C. (2016). Optimal dynamic reinsurance with dependent risks: variance premium principle. *Scandinavian Actuarial Journal*, 2016(1):18–36.

[87] Liang, Z., Yuen, K. C., and Guo, J. (2011b). Optimal proportional reinsurance and investment in a stock market with Ornstein–Uhlenbeck process. *Insurance: Mathematics and Economics*, 49(2):207–215.

[88] Lundberg, F. (1903). *Approximerad framställning av sannolikhetsfunktionen. Återförsäkrin av kollektivrisker.* Almqvist & Wiksell, Uppsala.

[89] Lundberg, O. (1964). *On Random Processes and Their Application to Sickness and Accident Statistics.* Almqvist & Wiksells, Uppsala. 1st ed. 1940.

[90] Lyberopoulos, D. P., Macheras, N. D., and Tzaninis, S. M. (2019). On the equivalence of various definitions of mixed poisson processes. *Mathematica Slovaca*, 69(2):453–468.

[91] Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

[92] Mecke, J. (1968). Eine charakteristische Eigenschaft der doppelt stochastischen Poissonschen Prozesse. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 11(1):74–81.

[93] Merton, R. C. (1971). Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3(4):373 – 413.

[94] Mikosch, T. (2009). *Non-Life Insurance Mathematics: An Introduction with the Poisson Process.* Universitext. Springer, Berlin, 2nd edition.

[95] Mitrinovic, D., Pecaric, J., and Fink, A. (1993). *Classical and New Inequalities in Analysis.* Mathematics and its Applications. Kluwer Academic Publishers, Dordrecht.

[96] Müller, A. and Stoyan, D. (2002). *Comparison Methods for Stochastic Models and Risks.* Wiley, New York.

[97] Pacáková, V. and Zapletal, D. (2013). Mixture distributions in modelling of insurance losses. *In: Proceedings of the 2013 International Conference on Applied Mathematics and Computational Methods in Engineering*, pages 112–125. Rhodes Island, Greece.

[98] Peng, X. and Hu, Y. (2013). Optimal proportional reinsurance and investment under partial information. *Insurance: Mathematics and Economics*, 53(2):416–428.

[99] Phadia, E. G. (2016). *Prior Processes and Their Applications: Nonparametric Bayesian Estimation.* Springer Series in Statistics. Springer, 2nd ed. edition.

[100] Pham, H. (2009). *Continuous-time Stochastic Control and Optimization with Financial Applications*, volume 61 of *Stochastic Modelling and Applied Probability*. Springer, Berlin.

[101] Polyanin, A. D. and Zaitsev, V. F. (1995). *Handbook of Exact Solutions for Ordinary Differential Equations*. CRC Press, Boca Raton.

[102] Popp, A. (2016). *Risk-Sensitive Stopping Problems for Continuous-Time Markov Chains*. PhD dissertation, Karlsruhe Institut of Technology, Department of Mathematics.

[103] Promislow, S. and Young, V. (2005). Minimizing the probability of ruin when claims follow brownian motion with drift. *North American Actuarial Journal*, 9:109–128.

[104] Protter, P. (2005). *Stochastic Integration and Differential Equations: Version 2.1*, volume 21 of *Applications of mathematics*. Springer, Berlin, 2nd edition. Corrected 3rd printing.

[105] Reiss, R.-D. and Thomas, M. (2007). *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*. Birkhäuser, Basel, 3rd edition.

[106] Roberts, A. W. and Varberg, D. E. (1974). Another proof that convex functions are locally lipschitz. *The American Mathematical Monthly*, 81(9):1014–1016.

[107] Rogers, L. C. G. (2013). *Optimal Investment*. SpringerBriefs in Quantitative Finance. Springer, Berlin.

[108] Scherer, M. and Selch, D. (2018). *A Multivariate Claim Count Model for Applications in Insurance*. Springer Actuarial. Springer International Publishing.

[109] Schäl, M. (2004). On discrete-time dynamic programming in insurance: Exponential utility and minimizing the ruin probability. *Scandinavian Actuarial Journal*, 2004(3):189–210.

[110] Schmidli, H. (2002). On minimizing the ruin probability by investment and reinsurance. *The Annals of Applied Probability*, 12(3):890–907.

[111] Schmidli, H. (2007). *Stochastic Control in Insurance*. Probability and Its Applications. Springer, London.

[112] Schmidli, H. (2017). *Risk Theory*. Springer Actuarial Lecure Notes. Springer Nature, Cham.

[113] Schmidt, K. D. (1996). *Lectures on Risk Theory*. Teubner Skripten zur mathematischen Stochastik. Teubner, Stuttgart.

[114] Schmidt, T. (2014). Catastrophe insurance modeled by shot-noise processes. *Risks*, 2(1):3–24.

[115] Sohrab, H. H. (2014). *Basic Real Analysis*. SpringerLink. Birkhäuser, New York, 2nd edition.

[116] Wang, K., Gao, M., Yang, Y., and Chen, Y. (2018). Asymptotics for the finite-time ruin probability in a discrete-time risk model with dependent insurance and financial risks. *Lithuanian Mathematical Journal*, 58(1):113–125.

[117] Wei, W., Liang, Z., and Yuen, K. (2017). Optimal reinsurance in a compound poisson risk model with dependence. *Journal of Applied Mathematics and Computing*, 58:1–24.

[118] Wichelhaus, C. and von Rohrscheidt, M. (2017). Bayesian nonparametric inference for the M/G/1 queueing systems based on the marked departure process.

[119] Winter, J. T. (2008). *Optimal Control of Markovian Jump Processes with Different Information Structures*. PhD dissertation, University Ulm, Department of Mathematics and Economics.

[120] Yong, J. and Zhou, X. Y. (1999). *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43 of *Applications of mathematics*. Springer, New York.

[121] Yuen, K. C., Liang, Z., and Zhou, M. (2015). Optimal proportional reinsurance with common shock dependence. *Insurance: Mathematics and Economics*, 64:1–13.

[122] Yuen, K. C. and Wang, G. (2001). Comparing two models with dependent classes of business. *In: Proceedings of the 36th Actuarial Research Conference*. ARCH (Society of Actuaries), 22 pages.

[123] Zhang, X. and Siu, T. (2009). Optimal investment and reinsurance of an insurer with model uncertainty. *Insurance: Mathematics and Economics*, 45:81–88.

[124] Zhu, S. and Shi, J. (2019). Optimal reinsurance and investment strategies under mean-variance criteria: Partial and full information. Papers 1906.08410, arXiv.org.