Decomposition of Stochastic Surplus Processes in Life Insurance

Von der Fakultät für Mathematik und Naturwissenschaften
der Carl von Ossietzky Universität Oldenburg
zur Erlangung des Grades und Titels eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

angenommene Dissertation

von

Herrn Julian Jetses

geboren am 05.08.1994 in Emden

1. Gutachter: Prof. Dr. Marcus C. Christiansen, Carl von Ossietzky Universität Oldenburg

2. Gutachter: Prof. Dr. Alexey Chernov, Carl von Ossietzky Universität Oldenburg

Tag der Disputation: 08.05.2025

„Suppose that you want to teach the 'cat' concept to a very young child. Do you explain that a cat is a relatively small, primarily carnivorous mammal with retractible claws, a distinctive sonic output, etc.? I'll bet not. You probably show the kid a lot of different cats, saying 'kitty' each time, until it gets the idea. To put it more generally, generalizations are best made by abstraction from experience.“

- Ralph P. Boas Jr.

Abstract

In life insurance, the uncertain long-term development of economic and demographic factors represents an undiversifiable risk. To address this risk, life insurers use conservative valuation assumptions, which lead to systematic surplus. By statute, a portion of this surplus must be refunded to the policyholder, with the compensation rates usually depending on the source of the surplus. Therefore, a decomposition of the total surplus with respect to the various risk factors is indispensable. Due to their relevance, several decomposition formulas have been presented in the actuarial literature. However, all contributions use heuristic arguments. A comprehensive decomposition principle that allows existing decomposition formulas to be compared and modern risks (e.g. policyholder behaviour) to be added is still missing. The thesis closes that gap by introducing a so-called infinitesimal sequential updating (ISU) decomposition principle.

The ISU decomposition principle enhances the sequential updating (SU) decomposition principle, which is popular in the economics literature but is subject to order effects. By forming the limit of SU decompositions with respect to the update frequency, the ISU decomposition principle eliminates the order effects while retaining the desired additivity. The thesis shows that this approach is also helpful for other prominent decomposition principles. Furthermore, the expediency of the ISU decomposition principle is demonstrated by replicating the various surplus decompositions known from the actuarial literature. In addition to its relevance for the decomposition of traditional life insurance surplus, the application of the ISU decomposition principle to martingales reveals its great potential in risk management. In particular, conditions are presented under which the ISU decomposition coincides with the recently introduced martingale representation theorem (MRT) decomposition.

Furthermore, evidence for the numerical feasibility of the ISU decomposition principle is provided using the example of a fund-linked pension insurance. In a doubly stochastic Markov setup, integral representations for the individual surplus contributions are obtained. The latter are further used to derive convergent estimators for the ISU decomposition by leveraging multilevel Monte Carlo methods to conditional expectations. The thesis concludes with the presentation of the numerical results, focusing on the impact of the chosen time grid and the chosen update order.

Kurzfassung

In der Lebensversicherung stellt die unsichere Entwicklung wirtschaftlicher und demographischer Faktoren ein nicht-diversifizierbares Risiko dar. Diesem Risiko begegnen die Lebensversicherer durch die Verwendung konservativer Bewertungsannahmen, welche zu systematischen Überschüssen führen. Laut Gesetz muss ein Teil dieser Überschüsse an die Versicherungsnehmer zurückgezahlt werden, dabei hängt die jeweilige Überschussbeteiligung in der Regel von der Überschussquelle ab. Dies erfordert die Zerlegung des Gesamtüberschusses nach Quellen. Entsprechend der praktischen Relevanz von Überschusszerlegungen, werden in der versicherungsmathematischen Literatur bereits verschiedene Zerlegungsformeln vorgeschlagen. Die Herleitungen der Zerlegungsformeln basieren jedoch auf heuristischen Argumenten. Ein Zerlegungsprinzip, das es erlaubt, bestehende Zerlegungen miteinander zu vergleichen und diese um moderne Risiken (z.B. das Verhalten der Versicherungsnehmer) zu erweitern, fehlt bisher. Die vorliegende Arbeit schließt diese Lücke, indem sie ein sogenanntes *infinitesimal sequential updating* (ISU) Zerlegungsprinzip einführt.

Das ISU-Zerlegungsprinzip erweitert das in der wirtschaftswissenschaftlichen Literatur bekannte *sequential updating* (SU) Zerlegungsprinzip. Letzteres hat den Nachteil, dass die resultierende Zerlegung maßgeblich von der gewählten Reihenfolge der Risikofaktoren abhängt. Die ISU Zerlegung ergibt sich hingegen als Grenzwert von SU Zerlegungen durch die Verfeinerung des unterliegenden Zeitgitters. Mithilfe dieses infinitesimalen Ansatzes können in Anwendungen oftmals Reihenfolgeneffekte eliminiert werden, gleichzeitig bleibt die wünschenswerte Additivität der SU Zerlegung aber erhalten. Darüberhinaus wird gezeigt, dass der vorgestellte Ansatz auch für andere bekannte Zerlegungsprinzipien nützlich ist. Die Adäquatheit des ISU Zerlegungsprinzips wird mittels der Replikation aus der Literatur bekannter Überschusszerlegungen nachgewiesen. Zusätzlich stellt sich durch die Anwendung des ISU Zerlegungsprinzips auf Martingale heraus, dass das Risikomanagement ein weiteres breites Anwendungsfeld des vorgestellten Zerlegungsprinzips darstellt. Insbesondere werden in dieser Arbeit Bedingungen präsentiert, unter denen die ISU Zerlegung mit der kürzlich eingeführten *martingale representation theorem* (MRT) Zerlegung übereinstimmt.

Darüber hinaus wird in dieser Arbeit am Beispiel einer fondsgebundenen Rentenversicherung die numerische Umsetzbarkeit des ISU Zerlegungsprinzips untersucht. Genauer werden in einem doppelt stochastischen Markov-Setup Integraldarstellungen für die einzelnen Überschussbeiträge gewonnen. Auf Basis dieser Integraldarstellungen werden konvergente Schätzer für die ISU Zerlegung abgeleitet, indem multilevel Monte Carlo Methoden auf bedingte Erwartungswerte verallgemeinert werden. Die Arbeit schließt mit der Darstellung der numerischen Ergebnisse, wobei sich die Analyse auf den Einfluss des Zeitgitters und der Aktualisierungsreihenfolge fokussiert.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to Marcus C. Christiansen for giving me the opportunity to undertake this PhD journey and for making this an inspiring experience for me. I am very thankful for the many hours of insightful discussion and the unwavering support throughout my thesis. Alexey Chernov, who agreed to be a co-supervisor, also deserves my thanks. The numerical part of this thesis would not have been possible without his valuable advice.

Next, I would like to thank Angelika May, who has not only been a great mentor during my studies, but also a pleasure to work with. Furthermore, I am very grateful to my former office mates and working colleagues from the Institute of Mathematics, who made the time at the university so enjoyable. Many thanks are also due to Henning, Ruben and Stephanie for their support during the completion of the thesis.

Moreover, I am deeply grateful to my friends, who have always been encouraging and understanding, even though the thesis has sometimes robbed me of my free time with them. My heartfelt thanks also go to my family, who never stopped asking about the number of pages written and the date of submission, but more importantly, never stopped believing in me. They have prepared me so lovingly for the future and made me what I am today.

Finally, I am very grateful to my wife for always being there for me and for having my back when I needed it most. She makes even the biggest hurdles seem small, and I am glad to have her by my side.

Declaration of included paper

The present thesis includes the following published manuscript:

> Jetses, J., & Christiansen, M. C. (2022). A general surplus decomposition principle in life insurance. *Scandinavian Actuarial Journal, 2022*(10), 901–925. https://doi.org/10.1080/03461238.2022.2049636

Chapter 2 of the thesis contains parts of Section 2, Section 4 and Section 7 of the manuscript. Chapter 3 of the thesis contains parts of Section 2, Section 3, Section 5 and Section 6 of the manuscript. Further details are provided at the beginning of the respective chapters.

I was responsible for deriving and writing up the main results of the paper. The co-author Marcus C. Christiansen contributed significantly to the conception of the paper and supported me during the further development of the paper.

Contents

1 Introduction

A traditional life insurance policy provides benefits in the event of death or survival, including a guaranteed rate of interest on the premiums paid. Due to the long-term nature of life insurance contracts, the prediction of economic and demographic factors (e.g. interest rates, mortality, and morbidity) harbours a high degree of uncertainty (see Jetses & Christiansen, 2022). To take this non-diversifiable risk into account, conservative assumptions are used for premium calculation. As a consequence, systematic surplus arises during the course of the contract. By virtue of legal requirements or contractual terms ('with-profit life insurance'), a part of this surplus belongs to the policyholder and is therefore refunded to them. Specifications, whether statutory or contractual, usually include compensation rates for the individual sources of surplus (investment surplus, risk surplus, other surplus). For this reason, a decomposition of surplus with respect to the different risk sources is essential.

It is therefore not surprising that surplus decompositions have been repeatedly addressed in actuarial literature since the early 20th century. In time-discrete life insurance setups, the popular contribution formula (*Kontributionsformel*) decomposes the yearly surplus into the mortality surplus, interest surplus, lapse surplus and cost surplus (see e.g. Milbrodt & Helbig, 1999, Section 11.B). In time-continuous life insurance setups, several decomposition formulas in multistate Markov models have been proposed (see e.g. Møller & Steffensen, 2007; Norberg, 1999, 2001; Ramlau-Hansen, 1988, 1991). All these publications derive the individual surplus contribution using heuristic arguments. Accordingly, the different decompositions are hard to compare. Moreover, the model frameworks are rigid, meaning that there is no natural way to add further sources of surplus to the model. However, adding further sources is highly relevant in the age of big data, where trackable policyholder behaviour (e.g. via smartwatches) allows for more accurate risk assessment. Thus, a fundamental decomposition principle that encompasses the existing decompositions on the one hand and allows for flexible expansion of further risk sources on the other hand is missing.

This thesis closes the outlined gap by presenting a so-called infinitesimal sequential updating (ISU) decomposition principle that gathers all the existing surplus decompositions under one common roof. The ISU decomposition principle is an extension of the sequential updating (SU) decomposition principle, which decomposes the surplus into the changes resulting from the sequential updating of risk factor information. Though the SU decomposition principle is well-known in economics literature (see e.g. Blinder, 1973; DiNardo et al., 1996; Oaxaca, 1973) it has the disadvantage of depending on the formal order of surplus sources (see Biewen, 2014; Fortin et al., 2011). One can address this shortcoming by using the averaged sequential updating (ASU) decomposition principle, which is also known as the Shapley-Shubik decomposition (see Shubik, 1962). This approach is subject to the curse of dimensionality (see Junike et al., 2024), nonetheless it is frequently used for attributing the prediction score of a machine learning model to its features (see Sundararajan and Najmi, 2020). Another idea discussed in the economics literature is the one-at-a-time (OAT) decomposition principle (see Biewen, 2014). Although the OAT decomposition principle avoids order effects, it involves interaction effects that cannot be attributed to a single source of risk. Due to the stated shortcomings of alternative decomposition principles, the ISU decomposition principle is introduced, which eliminates order effects by pushing the refinement of the valuation intervals to the limit. It is shown that this refinement approach also helps with the ASU and OAT decomposition principles, leading to the averaged infinitesimal SU (AISU) and infinitesimal OAT (IOAT) decomposition principles. The latter has recently also been discussed by Frei (2020). It turns out that all decomposition principles, ISU, AISU and IOAT, result in the same surplus decomposition, whenever the ISU decomposition does not depend on the update order. In particular, this applies to the traditional decompositions of life insurance surplus referred to above.

Though the traditional surplus decompositions has been the focus so far, it is worth noting that the scope of the ISU decomposition principle goes far beyond it. Even if the different surplus sources contribute to the total surplus in a non-linear manner, the ISU decomposition principle provides individual surplus contributions that add up to the total surplus. On the one hand, the need for such a decomposition arises from regulatory requirements for insurance companies (see Flaig & Junike, 2024). For example, Article 123 of the Solvency II Directive (European Parliament and the Council, 2009) requires insurers to perform an annual profit and loss attribution. Additionally, the revised Market Consistent Embedded Value (MCEV) reporting principles from 2016 entail the reconciliation of the opening MCEV and closing MCEV in a change analysis (CFO Forum, 2016). On the other hand, the ISU decomposition principle might be a very useful tool for risk management. The additivity of the resulting decomposition allows for an immediate application of the Euler allocation principle (see Frei, 2020; Karabey et al., 2014).

Taking additivity into account, Schilling et al. (2020) have recently published a comprehensive list of desirable properties of a risk decomposition. In addition, Schilling et al. (2020) have introduced the so-called martingale representation theorem (MRT) decomposition, which, unlike other risk decompositions in the literature, fulfils all desirable properties. The MRT decomposition breaks down the total surplus into individual surplus contributions by attributing the integrals of the martingale representation to the respective risk factors. In particular, this approach requires the total surplus process to be a martingale. The ISU decomposition principle is applicable to martingales, which again emphasises the scope of the decomposition principle presented. This immediately raises the question of the relationship between the MRT decomposition and the ISU decomposition. In this thesis, it is shown that under certain assumptions both decomposition approaches coincide. This result further illustrates the plausibility of the ISU decomposition principle.

The ISU decomposition principle is motivated by the various use cases in practice and its application to well-known examples from the actuarial literature. However, the question of whether the ISU decomposition principle will find a way into practice also depends on its numerical feasibility. Therefore, the numerical feasibility of the ISU decomposition principle is investigated using the example of a fund-linked pension insurance. In a doubly stochastic Markov setup, the ISU decomposition is computed in two steps. Firstly, the ISU decomposition is approached by SU decompositions applying the above-mentioned convergence results for martingale surplus processes. Secondly, the SU decompositions are approximated with the help of multilevel Monte Carlo (MLMC) methods. To develop a theoretical foundation for the second step, the MLMC approach by Giles (2008) is extended to conditional expectations and a systematic notion of (integral) MLMC convergence is introduced. The implementation of the derived estimators in R 4.4.2 (R Core Team, 2024) and the presentation of the numerical results then serve as a proof of concept.

This thesis is structured as follows:

In Chapter 2, the ISU decomposition principle is presented as a refinement of the SU decomposition principle (see e.g. Blinder, 1973). Additionally, further prevalent decomposition approaches are discussed. Section 2.1 introduces the ISU decomposition principle in a very general framework that allows for wide range of applications. By taking up the infinitesimal approach of the ISU decomposition principle, Section 2.2 proposes infinitesimal versions of the widely used ASU and OAT decomposition principles and analyses their relationship to the ISU decomposition principle. Section 2.3 concludes the chapter with an introductory example.

In Chapter 3, traditional surplus decomposition formulas known from the literature are embedded in the framework of the ISU decomposition principle presented in Chapter 2. This provides further evidence to the existing formulas, but also opens the door to the inclusion of further risks. In Section 3.1 the framework of the ISU decomposition principle is underpinned with the notions of life insurance modelling. In Section 3.2, a general multistate life insurance setup is introduced, in which rigorous definitions of surplus are established following Norberg (1999). The application of the ISU decomposition principle leads to the main theorems in Section 3.3, which provide the basis to deduce the traditional surplus decomposition formulas in Section 3.4.

In Chapter 4, the ISU decomposition principle is related to the recently introduced MRT decomposition of Schilling et al. (2020), and a property of the risk factors is established under which both decompositions are equivalent. The underlying model framework, which assumes the surplus process to be a martingale, is introduced in Section 4.1. After defining the above-mentioned property of the surplus sources, it is shown in Section 4.2, that a number of commonly used stochastic processes satisfy this property. In Section 4.3, the ISU decomposition of martingales is derived and its relationship to the MRT decomposition is discussed.

In Chapter 5, the numerical feasibility of the ISU decomposition principle is investigated using the example of a fund-linked pension insurance. After introducing an appropriate model framework in Section 5.1, an analysis of the approximating SU decompositions is considered in Section 5.2. Furthermore, in Section 5.2 the multilevel Monte Carlo approach studied by Giles (2008) is generalised to conditional expectations, which helps with the derivation of convergent estimators for the SU contributions. Section 5.3 contains details of the implementation and a presententation of the results.

In Chapter 6, the results obtained are reflected upon, and further discussion is given to the open questions for future research.

2 A general surplus decomposition principle

This chapter is derived in part from an article published in the Scandinavian Actuarial Journal in 2022 (copyright Taylor & Francis), available online: http://www.tandfonline.com/10.1080/03461238.2022.2049636. More precisely, (parts of) Sections 2, 4 and 7 of the article Jetses and Christiansen (2022) are included in Sections 2.1 and 2.2 of this chapter. The Section 2.3 has been prepared specifically for this thesis. In order to improve readability and standardise the notation in this dissertation, minor changes have been made compared to the original article.

Insurance companies are subject to an increasing number of regulatory requirements. The main objectives of regulation are to protect policyholders and to provide greater transparency of an insurer's financial position to all stakeholders. The latter necessitates comprehensive knowledge of the risks to which the insurer is exposed. In that regard, reconciling the balance sheets from two different valuation dates and linking the changes to the various sources leads to a better awareness for the relevant risks (see Candland & Lotz, 2014). This aspect is also reflected by the Solvency II Directive (European Parliament and the Council, 2009), where Article 123 requires insurers to carry out a profit and loss attribution at least once a year. A similar idea is followed by the analysis of change in liabilities, which is part of the recently introduced reporting standard IFRS17 (IASB, 2017, Article 100 ff.). Furthermore, a change analysis is required by the revised MCEV reporting principles from 2016 (CFO Forum, 2016) in order to reconcile the opening MCEV with the closing MCEV.

The stated examples from regulation underline the need for a systematic approach to profit and loss attribution or change analysis, where both terms are often used interchangeably (see Christiansen, 2022). Allocating a value change between two valuation dates to the different sources can be a very challenging task, especially if the risk factors contribute to the respective key figure in a non-linear way. Usually, the value change is caused by the newly gained information on the risk factors that leads to a reconciliation of expected and actual values for the past period, as well as to changes in estimates of future values. Therefore, a straightforward idea is to update the information from the different risk factors *sequentially* and assign the resulting change of the total to the respective risk factor. This approach, which is called the sequential updating (SU) decomposition principle, leads to additive decompositions and is widely used in various fields of economics (see Jetses & Christiansen, 2022). Examples are provided by Blinder (1973) and Oaxaca (1973) ('Blinder-Oaxaca-Decomposition') in the field of labour economics, see also DiNardo et al. (1996). In insurance economics, Candland and Lotz (2014) present the *Waterfall approach* for profit and loss attribution, which follows the idea of the SU decomposition principle. Further economic applications of the SU decomposition principle can be found in Fortin et al. (2011).

For the application of the SU decomposition principle, it is necessary to decide on a time grid in which the risk drivers will be updated. This can lead to different decompositions depending on the selected time grid. In addition to this drawback, another downside of the SU decomposition principle has been pointed out by Biewen (2014) and Fortin et al. (2011), namely its dependence on the updating order of the risk factors. Fortunately, both disadvantages can be eliminated by refining the time grid of the SU decomposition to the limit, while retaining the desired additivity. As this leads to infinitesimal valuation periods, this approach is referred to as the infinitesimal SU (ISU) decomposition principle.

It is worth mentioning that in the existing literature alternative decomposition principles are discussed that address the drawback of the order dependency. By averaging the SU decompositions over all possible update orders, the order dependency can be removed. Nonetheless, this alternative which is called the averaged sequential updating (ASU) decomposition principle is subject to the curse of dimensionality (see Junike et al., 2024). In economics literature, this approach is also known as the Shapley-Shubik decomposition, which was introduced by Shubik (1962) as a generalisation of the Shapley decomposition (see Shapley, 1953). A further alternative avoiding the ordering problem is the one-at-a-time (OAT) decomposition principle (see Biewen, 2014). However, this decomposition principle involves interaction effects, which cannot be assigned to a single risk driver. Additionally, similar to the SU decomposition principle, the ASU and the OAT decomposition principles depend on the selected time grid. As an advancement of the ASU and OAT decomposition principles, both the averaged infinitesimal SU (AISU) and the infinitesimal OAT (IOAT) decomposition principles are proposed, which are derived by incorporating the above-mentioned infinitesimal approach. It turns out that all three decomposition principles are closely related.

After providing the basic notations, Section 2.1 introduces the ISU decomposition principle. Section 2.2 examines alternative decomposition principles and their relation to the ISU decomposition principle. This chapter is concluded with an introductory example in Section 2.3.

2.1 The ISU decomposition principle

We generally assume that we have a complete probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with a right-continuous and complete filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geqslant 0}$. Let the so-called *risk basis*

$$X = (X_1, \dots, X_m),$$

be given by a multivariate adapted process composed of so-called *risk factors* $X_1, \dots, X_m$. Furthermore, let $R = (R(t))_{t \geqslant 0}$ be a stochastic process that rests on the *risk basis* X, i.e. R is adapted to the right-continuous and complete filtration generated by X. The right-continuous and complete filtration generated by X does not necessarily coincide with $\mathbb{F}$ but may be a strict sub-filtration of $\mathbb{F}$ in such a way that at least R is adapted. We interpret $R(t)$ as a proxy in t of a random variable that depends on the risk factors' future development. For example, one could define $R(t)$ as the conditional expectation of the (discounted) total cashflow given the available information at time $t \geqslant 0$. In this sense, R increases or decreases due to revaluation gains or losses. Therefore, we call R the *revaluation surplus process.* The information provided by X at time t can be represented by the stopped process X^t, formally defined by

$$X^t(s) = \mathbb{1}_{s \leqslant t} X(s) + \mathbb{1}_{s > t} X(t). \tag{2.1}$$

Thus, at each time t, the proxy $R(t)$ can be interpreted as the value of a mapping

$$(t, X^t) \mapsto R(t)$$

that assigns at each time t to the current information X^t the random variable $R(t)$. In this thesis, we assume that there even exists a mapping ϱ such that

$$\varrho(X^t) = R(t), \quad t \geqslant 0.$$

In the latter equation, the time parameter t itself is not an argument of ϱ and only appears as stopping parameter in X^t. That means that the dynamics of R are solely driven by the increase of information through X^t.

The central aim of this thesis is to decompose R as

$$R(t) = R(0) + D_1(t) + \cdots + D_m(t), \quad t \geqslant 0, \tag{2.2}$$

where $D_1, \dots, D_m$ are adapted processes that start at zero and describe the contributions of each risk factor $X_1, \dots, X_m$ to the dynamics of R. The first addend $R(0)$ represents initial surplus, which is not decomposed here.

Suppose that the information updates of the risk factors $X_1, \dots, X_m$ are asynchronously delayed with $t_1, \dots, t_m \leqslant t$ being the current update statuses of each risk factor. Then

$$U(t_1, \dots, t_m) := \varrho((X_1^{t_1}, \dots, X_m^{t_m})) \tag{2.3}$$

is the value of the delayed revaluation process at time points $t_1, \dots, t_m$. Furthermore, we denote $U = (U(t_1, \dots, t_m))_{t_1, \dots, t_m \geqslant 0}$ as the *revaluation surplus surface* with respect to X. We can recover the revaluation surplus process R from the revaluation surplus surface U as

$$R(t) = U(t, \dots, t), \quad t \geqslant 0.$$

For any partition $\mathcal{T}(t) = \{0 = t_0 < t_1 < \cdots < t_l = t\}$ of the interval $[0, t]$, we can build the telescoping series

$$\begin{aligned} R(t) - R(0) &= U(t, \dots, t) - U(0, \dots, 0) \\ &= \sum_{k=0}^{l-1} \Big(U(t_{k+1}, t_k, \dots, t_k) - U(t_k, \dots, t_k) \Big) \\ &\quad + \sum_{k=0}^{l-1} \Big(U(t_{k+1}, t_{k+1}, t_k, \dots, t_k) - U(t_{k+1}, t_k, \dots, t_k) \Big) \\ &\quad + \cdots \\ &\quad + \sum_{k=0}^{l-1} \Big(U(t_{k+1}, \dots, t_{k+1}) - U(t_{k+1}, \dots, t_{k+1}, t_k) \Big). \end{aligned}$$

It is natural to interpret the m different sums on the right hand side as an additive decomposition $R(t) - R(0) = D_1(t) + \cdots + D_m(t)$, since the i-th sum collects exactly the information updates for the i-th risk factor.

Definition 2.1. The random vector $D(t) = (D_1(t), \dots, D_m(t))$ defined by

$$\begin{aligned} D_1(t) &= \sum_{k=0}^{l-1} \Big(U(t_{k+1}, t_k, \dots, t_k) - U(t_k, \dots, t_k) \Big), \\ &\cdots \\ D_m(t) &= \sum_{k=0}^{l-1} \Big(U(t_{k+1}, \dots, t_{k+1}) - U(t_{k+1}, \dots, t_{k+1}, t_k) \Big), \end{aligned} \tag{2.4}$$

is called the *sequential updating (SU) decomposition* of $R(t) = \varrho(X^t)$ with respect to $\mathcal{T}(t)$.

The SU decomposition principle is used in various fields of economics (see e.g. Biewen, 2014; Fortin et al., 2011). In (2.4) we update the information on X in a specific order, starting with risk factor X_1, then updating X_2, and so on. Unfortunately, the decomposition is not invariant with respect to this update order, which is a major drawback of the SU concept. We can reduce the impact of the update order by increasing the number of updating steps, i.e. refining the partition $\mathcal{T}_n(t)$. In a next step we push such refinements to the limit.

Let $\mathcal{T}_n(t) = \{0 = t_0^n < t_1^n < \cdots < t_{l_n}^n = t\}$, $n \in \mathbb{N}$, be a sequence of partitions of $[0, t]$ with vanishing step lengths (i.e. $\lim_{n\to\infty} \max_{1\leqslant k\leqslant l_n} |t_k^n - t_{k-1}^n| = 0$). For each $n \in \mathbb{N}$ let $D^n(t) = (D_1^n(t), \dots, D_m^n(t))$ be the SU decomposition of $R(t) = \varrho(X^t)$ with respect to $\mathcal{T}_n(t)$. We are looking for a random vector $D(t)$ that satisfies

$$D_i(t) = \operatorname*{plim}_{n\to\infty} D_i^n(t), \quad i \in \{1, \dots, m\}, \tag{2.5}$$

where $\mathrm{plim}_{n\to\infty}$ describes the convergence in probability.

Definition 2.2. Let $(\mathcal{T}_n(t))_{n\in\mathbb{N}}$ be a sequence of partitions of $[0,t]$ with vanishing step lengths. If $D(t)$ satisfies (2.5), then we call $D(t)$ the *infinitesimal sequential updating (ISU) decomposition* of $R(t) = \varrho(X^t)$ with respect to $(\mathcal{T}_n(t))_{n\in\mathbb{N}}$.

An axiomatic approach to the ISU decomposition can be found in Christiansen (2022).

2.2 Alternative decomposition principles

As we will see in the next chapters, moving forward to the limit of SU decompositions by pushing the step lengths to zero is an effective way to eliminate order dependencies. However, this is not the only way to approach this issue. Therefore, we want to elaborate on this point by discussing two alternative decomposition principles in this section. Instead of updating the sources of risk sequentially, we could also update only one source of risk at a time and quantify its impact on total revaluation surplus $R(t) - R(0)$, which will lead us to the OAT decomposition principle.

Recall that $U(t_1, \ldots, t_m) = \varrho((X_1^{t_1}, \ldots, X_m^{t_m}))$ is the value of the delayed revaluation process at time points $t_1, \ldots, t_m$ (see (2.3)). For any partition

$$\mathcal{T}(t) = \{0 = t_0 < t_1 < \cdots < t_l = t\}$$

of the interval $[0,t]$ we can decompose

$$\begin{aligned}
&R(t) - R(0)\\
&= U(t, \ldots, t) - U(0, \ldots, 0)\\
&= \sum_{k=0}^{l-1} \Big(U(t_{k+1}, t_k, \ldots, t_k) - U(t_k, \ldots, t_k) \Big)\\
&\quad + \sum_{k=0}^{l-1} \Big(U(t_k, t_{k+1}, t_k, \ldots, t_k) - U(t_k, \ldots, t_k) \Big)\\
&\quad + \ldots\\
&\quad + \sum_{k=0}^{l-1} \Big(U(t_k, \ldots, t_k, t_{k+1}) - U(t_k, \ldots, , t_k) \Big)\\
&\quad + \sum_{k=0}^{l-1} \Big(U(t_{k+1}, \ldots, t_{k+1}) - U(t_k, \ldots, t_k) \Big)\\
&\quad - \sum_{k=0}^{l-1} \Big(U(t_{k+1}, t_k, \ldots, t_k) - U(t_k, \ldots, t_k) + \ldots + U(t_k, \ldots, t_k, t_{k+1}) - U(t_k, \ldots, t_k) \Big).
\end{aligned}$$

Here, the first m sums quantify the single effect of the corresponding source of risk. Following Biewen (2014), we call them the *ceteris paribus effects*. Since the ceteris paribus effects do not necessarily add up to the total revaluation surplus $R(t) - R(0)$, we get an extra term in the last two lines, which is called the *interaction effect* (see Biewen, 2014). Based on this construction, we get a decomposition principle with a joint risk factor.

Definition 2.3. The random vector $D(t) = (D_1(t), \dots, D_m(t), \overline{D}(t))$ defined by

$$\begin{aligned}
D_1(t) &= \sum_{k=0}^{l-1} \Big(U(t_{k+1}, t_k, \dots, t_k) - U(t_k, t_k, \dots, t_k) \Big), \\
&\dots \\
D_m(t) &= \sum_{k=0}^{l-1} \Big(U(t_k, \dots, t_k, t_{k+1}) - U(t_k, \dots, t_k) \Big), \\
\overline{D}(t) &= R(t) - R(0) - \sum_{j=1}^{m} D_j(t)
\end{aligned} \tag{2.6}$$

is called the *one-at-a-time (OAT) decomposition* of $R(t) = \varrho(X^t)$ with respect to $\mathcal{T}(t)$.

The OAT decomposition principle is also known in economics (see Biewen, 2014). In contrast to the SU decomposition, the OAT decomposition is order-invariant, i.e. it does not depend on the order of the risk basis (see Schilling et al., 2020). Nevertheless, we get a joint risk factor that cannot be assigned to any source of risk. In Section 2.1, we faced the order dependence of the SU decomposition by considering increasing sequences of partitions of $[0, t]$. Similarly, we address the unassignable interaction effect in the OAT decomposition.

Let $\mathcal{T}_n(t) = \{0 = t_0^n < t_1^n < \dots < t_{l_n}^n = t\}$, $n \in \mathbb{N}$, be a sequence of partitions of $[0, t]$ with vanishing step lengths (i.e. $\lim_{n\to\infty} \max_{1\leqslant k\leqslant l_n} |t_k^n - t_{k-1}^n| = 0$). For each $n \in \mathbb{N}$ let $D^n(t) = (D_1^n(t), \dots, D_m^n(t), \overline{D}^n(t))$ be the OAT decomposition of $R(t) = \varrho(X^t)$ with respect to $\mathcal{T}_n(t)$. We are looking for a random vector $D(t) = (D_1(t), \dots, D_m(t), \overline{D}(t))$ that satisfies

$$\begin{aligned}
D_i(t) &= \operatorname*{plim}_{n\to\infty} D_i^n(t), \quad i \in \{1, \dots, m\}, \\
\overline{D}(t) &= \operatorname*{plim}_{n\to\infty} \overline{D}^n(t).
\end{aligned} \tag{2.7}$$

Definition 2.4. Let $(\mathcal{T}_n(t))_{n\in\mathbb{N}}$ be a sequence of partitions of $[0, t]$ with vanishing step lengths. If $D(t) = (D_1(t), \dots, D_m(t), \overline{D}(t))$ satisfies (2.7), then we call $D(t)$ the *infinitesimal one-at-a-time (IOAT) decomposition* of $R(t) = \varrho(X^t)$ with respect to $(\mathcal{T}_n(t))_{n\in\mathbb{N}}$.

The next theorem characterizes the relation between the ISU decomposition and the IOAT decomposition.

Theorem 2.5. *The following statements are equivalent:*

a) The ISU decomposition is independent of update order.

b) For each update order, the ISU decomposition is equal to the ceteris paribus effects of the IOAT decomposition.

In both cases, the interaction effect is zero.

Proof. The proof follows Biewen (2014). Let us fix a source of risk ($i = 1, \dots, m$). Choosing an update order, such that this source of risk is updated first, the corresponding risk factor of the ISU decomposition coincides per definition with the ceteris paribus effect of the IOAT decomposition. If the ISU decomposition is independent of update order, the risk factor, corresponding to the fixed source of risk, equals the ceteris paribus effect of the IOAT decomposition for each update order.

Apart from that, the statement in b) directly implies, that the ISU decomposition is independent of update order. Furthermore, if the ISU decomposition equals the IOAT decomposition, then the ceteris paribus effects sum up to total risk $R(t) - R(0)$, therefore the interaction effect is zero. $\square$

By subdividing the interaction effect into different groups of interaction effects (depending on the number of involved risk factors), Biewen (2014) even shows that the particular interaction effects are zero if and only if the ISU decomposition is independent of update order.

If the interaction effect is non-zero, neither the ISU decomposition nor the IOAT decomposition yields an order-invariant decomposition satisfying (2.2). One possible solution for this problem is to build a decomposition principle based on the ISU decomposition principle that is symmetric with respect to the sources of risk. For that, let $\pi\colon \{1, \dots, m\} \to \{1, \dots, m\}$ be a permutation that represents an update order for the ISU decomposition. The set of all possible permutations on $\{1, \dots, m\}$ is denoted by σ_m.

Definition 2.6. Let $(\mathcal{T}_n(t))_{n\in\mathbb{N}}$ be an increasing sequence of partitions of $[0, t]$ with vanishing step lengths and let $\pi \in \sigma_m$. Further, let $D^\pi(t) = (D_1^\pi(t), \dots, D_m^\pi(t))$ denote the ISU decomposition of $R(t) = \varrho(X^t)$ with respect to π and with respect to $(\mathcal{T}_n)_n$. The random vector $D(t) = (D_1(t), \dots, D_m(t))$ defined by

$$\begin{aligned} D_1(t) &= \frac{1}{m!} \sum_{\pi\in\sigma_m} D^\pi_{\pi(1)}(t), \\ &\cdots \\ D_m(t) &= \frac{1}{m!} \sum_{\pi\in\sigma_m} D^\pi_{\pi(m)}(t), \end{aligned} \tag{2.8}$$

is called the *averaged infinitesimal sequential updating (AISU) decomposition* of $R(t) = \varrho(X^t)$ with respect to $(\mathcal{T}_n(t))_{n\in\mathbb{N}}$.

In a similar manner, Shorrocks (2013) proposes the averaged SU decomposition (without taking limits) for the distributional analysis of poverty in economics literature. A recent contribution by Godin et al. (2023) uses averaged SU decompositions for risk allocation. Axiomatic approaches to the averaged SU decomposition can further be found in Friedman and Moulin (1999) and Sprumont (1998). By construction, the AISU decomposition principle is symmetric with respect to the risk basis and therefore gives an order-invariant surplus decomposition satisfying (2.2) even if the interaction effect is non-zero. Furthermore, the averaged ISU decomposition is in line with the previously proposed decomposition principles as the next theorem shows.

Theorem 2.7. *If the ISU decomposition is independent of update order, then ISU (for each update order), IOAT and averaged ISU yield the same decomposition.*

Proof. Assume that the ISU decomposition principle yields a decomposition

$$D(t) = (D_1(t), \ldots, D_m(t))$$

for each update order. Then, by Theorem 2.5, the ISU decomposition is equal to the IOAT decomposition for each update order. Furthermore, it holds $D^{\pi}_{\pi(i)}(t) = D_i(t)$, $i = 1, \ldots, m$, for every permutation π. Since $\#\sigma_m = m!$, the averaged ISU decomposition is also given by $(D_1(t), \ldots, D_m(t))$. □

Having introduced the decomposition principles, we can consider a first example.

2.3 An introductory example

Before presenting an introductory example for the previously presented decomposition principles, we establish some notation that will be used throughout the thesis. If we calculate a sum across a partition

$$\mathcal{T}(t) = \{0 = t_0 < \ldots < t_l = t\},$$

like we do in (2.4) and (2.6), we write $\sum_{t_k, t_{k+1} \in \mathcal{T}(t)}$ instead of $\sum_{k=0}^{l-1}$. For the grid width of the partition $\mathcal{T}(t)$, we introduce $|\mathcal{T}(t)| := \max_{1 \leqslant k \leqslant l} |t_k - t_{k-1}|$. Furthermore, as already touched in the previous sections, $(\mathcal{T}_n(t))_n$ always refers to a sequence of partitions on $[0, t]$ with vanishing step lengths, i.e. $\lim_{n \to \infty} |\mathcal{T}_n(t)| = 0$. To achieve a better readability, we solely write $t_k, t_{k+1} \in \mathcal{T}_n(t)$ instead of $t^n_k, t^n_{k+1} \in \mathcal{T}_n(t)$. In addition, we will encounter many integrals in this thesis. Unless otherwise stated, the integrals are understood as stochastic integrals. With this clarification of the notation, we can now move on to the example.

We suppose that the risk basis

$$X = (X_1, \ldots, X_m), \ m \in \mathbb{N},$$

consists of $\mathbb{F}$-semimartingales $X_1, \ldots, X_m$, such that the quadratic covariation between the different risk factors is zero, i.e. $[X_i, X_j] = 0$, $i \neq j$. Furthermore, let $\mathcal{C}^2_c(\mathbb{R}^m)$ denote the space of twice continuously differentiable, real-valued functions from $\mathbb{R}^m$ to $\mathbb{R}$ with a compact support.

Additionally, we assume that the revaluation surplus process $R = (R(t))_{t\geqslant 0}$ is given by

$$R(t) = \varrho(X^t) = f(X_1(t), \ldots, X_m(t))$$

for some $f \in \mathcal{C}_c^2(\mathbb{R}^m)$. Thus, the proxy $R(t)$ of the total surplus does not depend on the whole past of the risk sources but only on the current value $X(t)$, which equals a Markovian structure. Applying Itô's formula (Protter, 2005, Chapter II, Theorem 33) we directly get the additive representation

$$\begin{aligned} R(t) - R(0) = &\sum_{i=1}^{m} \int_0^t f_i'(X(s-))\mathrm{d}X_i(s) + \frac{1}{2}\sum_{i=1}^{m}\int_0^t f_{ii}''(X(s-))\mathrm{d}[X_i, X_i]^c(s) \\ &+ \sum_{0<s\leqslant t}\left(f(X(s)) - f(X(s-)) - \sum_{i=1}^{m} f_i'(X(s-))\Delta X_i(s)\right), \end{aligned}$$

where $X(t) = (X_1(t), \ldots, X_m(t))$.

As $[X_i, X_j] = 0$, $i \neq j$, implies $\Delta X_i \Delta X_j = \Delta[X_i, X_j] = 0$ (see e.g. Protter, 2005, Chapter II, Theorem 23), a natural guess for the i-th surplus contribution, $i = 1, \ldots, m$ is

$$\begin{aligned} D_i(t) = &\int_0^t f_i'(X(s-))\mathrm{d}X_i(s) + \frac{1}{2}\int_0^t f_{ii}''(X(s-))\mathrm{d}[X_i, X_i]^c(s) \\ &+ \sum_{\substack{0<s\leqslant t \\ \Delta X_i(s)\neq 0}} \big(f(X(s)) - f(X(s-)) - f_i'(X(s-))\Delta X_i(s)\big). \end{aligned} \tag{2.9}$$

The next theorem shows that the ISU decomposition principle results in this decomposition.

Theorem 2.8. *Let $f \in \mathcal{C}_c^2(\mathbb{R}^m)$. Then $\varrho(X^t) = f(X_1(t), \ldots, X_m(t))$ admits the ISU decompostion $D(t) = (D_1(t), \ldots, D_m(t))$ with $D_i(t)$ given by* (2.9)*. In particular, the ISU decomposition does not depend on the update order or the choice of partitions.*

Proof. To avoid a cumbersome notation, we assume without loss of generality that the order of the risk basis to be $(X_1, \ldots, X_m)$. For $i \in \{1, \ldots, m\}$, we write

$$X^i(t_k, t_{k+1}) = (X_1(t_{k+1}), \ldots, X_i(t_{k+1}), X_{i+1}(t_k), \ldots X_m(t_k)),\ t_k, t_{k+1} \in \mathcal{T}_n,$$

$$X^i(s) = (X_1(s), \ldots, X_i(s), X_{i+1}(s-), \ldots X_m(s-)),\ s \in [0, t],$$

and

$$X^{i,t} = (X_1, \ldots, X_i, X_{i+1}^t, \ldots, X_m^t),$$

where X_j^t denotes the stopped process (at time t) of X_j. The stopped process is still a semimartingale, and thus $X^{i,t}$ is an m-tuple of semimartingales.

Let $(\mathcal{T}_n(t))_n$ be a sequence of partitions on $[0, t]$ with vanishing step lengths. For $t_k, t_{k+1} \in \mathcal{T}_n(t)$, Itô's formula (Protter, 2005, Chapter II, Theorem 33) yields

$$
\begin{aligned}
&f(X^i(t_k,t_{k+1})) - f(X(t_k)) \\
&= f(X^{i,t_k}(t_{k+1})) - f(X^{i,t_k}(t_k)) \\
&= \sum_{j=1}^{i} \int_{t_k}^{t_{k+1}} f_j'(X^{i,t_k}(s-))\mathrm{d}X_j(s) + \frac{1}{2}\sum_{j=1}^{i} \int_{t_k}^{t_{k+1}} f_{jj}''(X^{i,t_k}(s-))\mathrm{d}[X_j,X_j]^c(s) \\
&\quad + \sum_{t_k<s\leqslant t_{k+1}} \left(f(X^{i,t_k}(s)) - f(X^{i,t_k}(s-)) - \sum_{j=1}^{i} f_j'(X^{i,t_k}(s-))\Delta X_j(s) \right).
\end{aligned}
$$

In the following, let $X^{i,n}(s) := X^{i,t_k}(s)$, if $s \in (t_k, t_{k+1}]$. Then we have

$$
\begin{aligned}
&\sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} f(X^i(t_k,t_{k+1})) - f(X(t_k)) \\
&= \sum_{j=1}^{i} \int_0^t f_j'(X^{i,n}(s-))\mathrm{d}X_j(s) + \frac{1}{2}\sum_{j=1}^{i} \int_0^t f_{jj}''(X^{i,n}(s-))\mathrm{d}[X_j,X_j]^c(s) \\
&\quad + \sum_{0<s\leqslant t} \left(f(X^{i,n}(s)) - f(X^{i,n}(s-)) - \sum_{j=1}^{i} f_j'(X^{i,n}(s-))\Delta X_j(s) \right).
\end{aligned}
\tag{2.10}
$$

Next we want to investigate the limit of the SU decompositions for $n \to \infty$. Observe that

$$
\lim_{n\to\infty} X^{i,n}(s-) = \lim_{n\to\infty} (X_1(s-),\ldots,X_i(s-),X_{i+1}(t_k),\ldots,X_m(t_k)) = X(s-)
$$

and

$$
\lim_{n\to\infty} X^{i,n}(s) = (X_1(s),\ldots,X_i(s),X_{i+1}(s-),\ldots,X_m(s-)) = X^i(s)
$$

for every $s \in [0,t]$. Since $f \in \mathcal{C}_c^2(\mathbb{R}^m)$, we also have $\lim_{n\to\infty} f(X^{i,n}(s-)) = f(X(s-))$, $\lim_{n\to\infty} f(X^{i,n}(s)) = f(X^i(s))$ and $\lim_{n\to\infty} f_j'(X^{i,n}(s-)) = f_j'(X(s-))$, $j = 1,\ldots,m$, almost surely for every $s \in [0,t]$.

For the two integrals in (2.10), we apply the stochastic dominated convergence theorem (Protter, 2005, Chapter IV, Theorem 32). Exploiting that $f \in \mathcal{C}_c^2(\mathbb{R}^m)$, we can define random variables

$$
G_j = \sup_{s_1,\ldots,s_m\in[0,t]} |f_j'(X_1(s_1),\ldots,X_m(s_m))|,\ j = 1,\ldots,m
$$

and

$$
G_{jj} = \sup_{s_1,\ldots,s_m\in[0,t]} |f_{jj}''(X_1(s_1),\ldots,X_m(s_m))|,\ j = 1,\ldots,m
$$

with values in $[0,\infty)$ that dominate the integrands in (2.10). Thus, we get

$$
\operatorname*{plim}_{n\to\infty} \sum_{j=1}^{i} \int_0^t f_j'(X^{i,n}(s-))\mathrm{d}X_j(s) = \sum_{j=1}^{i} \int_0^t f_j'(X(s-))\mathrm{d}X_j(s),
$$

$$
\operatorname*{plim}_{n\to\infty} \frac{1}{2}\sum_{j=1}^{i} \int_0^t f_{jj}''(X^{i,n}(s-))\mathrm{d}[X_j,X_j]^c(s) = \frac{1}{2}\sum_{j=1}^{i} \int_0^t f_{jj}''(X(s-))\mathrm{d}[X_j,X_j]^c(s).
$$

For the summation in (2.10), the multidimensional Taylor Theorem (Forster, 2017, Section I.7, Theorem 2) together with the assumption $[X_i, X_j] = 0$, $i \neq j$ gives us the upper bound

$$\left| f(X^{i,n}(s)) - f(X^{i,n}(s-)) - \sum_{j=1}^{i} f_j'(X^{i,n}(s-))\Delta X_j(s) \right| \leqslant C \sum_{j=1}^{i} (\Delta X_j(s))^2,$$

where C can be chosen independently of s as $f \in \mathcal{C}_c^2(\mathbb{R}^m)$. Thus, the dominator is indeed summable with

$$\sum_{0<s\leqslant t} C \sum_{j=1}^{i} (\Delta X_j(s))^2 = C \sum_{j=1}^{i} \sum_{0<s\leqslant t} (\Delta X_j(s))^2 \leqslant C \sum_{j=1}^{i} [X_j, X_j](t)$$

(see Protter, 2005, Chapter II, proof of Theorem 32). Having found a summable dominator, we can apply Tannery's Theorem (e.g. Bromwich, 1926, p. 136) for interchanging limit and summation to get

$$\begin{aligned}
&\lim_{n\to\infty} \sum_{0<s\leqslant t} \left(f(X^{i,n}(s)) - f(X^{i,n}(s-)) - \sum_{j=1}^{i} f_j'(X^{i,n}(s-))\Delta X_j(s) \right) \\
&= \sum_{j=1}^{i} \sum_{0<s\leqslant t} \big(f(X^i(s)) - f(X(s-)) - f_j'(X(s-))\Delta X_j(s) \big) \\
&= \sum_{j=1}^{i} \sum_{\substack{0<s\leqslant t \\ \Delta X_j(s)\neq 0}} \big(f(X(s)) - f(X(s-)) - f_j'(X(s-))\Delta X_j(s) \big),
\end{aligned}$$

where we used $[X_i, X_j] = 0$, $i \neq j$, for the last equality. In total, it holds

$$\begin{aligned}
&\operatorname*{plim}_{n\to\infty} \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} (f(X^i(t_k, t_{k+1})) - f(X(t_k))) \\
&= \sum_{j=1}^{i} \int_0^t f_j'(X(s-)) \mathrm{d}X_j(s) + \frac{1}{2} \sum_{j=1}^{i} \int_0^t f_{jj}''(X(s-)) \mathrm{d}[X_j, X_j]^c(s) \\
&\quad + \sum_{j=1}^{i} \sum_{\substack{0<s\leqslant t \\ \Delta X_j(s)\neq 0}} \big(f(X(s)) - f(X(s-)) - f_j'(X(s-))\Delta X_j(s) \big).
\end{aligned} \tag{2.11}$$

The desired result is now deduced from

$$\begin{aligned}
&\sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} (f(X^i(t_k, t_{k+1})) - f(X^{i-1}(t_k, t_{k+1}))) \\
&= \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} (f(X^i(t_k, t_{k+1})) - f(X(t_k))) - \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} (f(X^{i-1}(t_k, t_{k+1})) - f(X(t_k)))
\end{aligned}$$

by taking on both sides the limit in probability and using (2.11). □

In the previous theorem, we have shown that the ISU decomposition does not depend on the update order. Therefore, we can also draw a conclusion about the IOAT and AISU decompositions.

Corollary 2.9. *Let $f \in \mathcal{C}_c^2(\mathbb{R}^m)$. For $R(t) = f(X_1(t), \dots, X_m(t))$ the IOAT decomposition and the averaged ISU decomposition are both equal to the ISU decomposition.*

Proof. The result follows immediately with Theorem 2.7 and Theorem 2.8. □

Focusing on continuous semimartingales with zero quadratic covariation, Frei (2020, Proposition 1) derives the IOAT decomposition in a similar framework. With Theorem 2.8 and Corollary 2.9, we generalise his result for functions $f \in C_c^2(\mathbb{R}^m)$ by allowing for jumps of the semimartingales, while keeping the assumption of the zero quadratic covariation. A further generalisation to semimartingales with non-zero quadratic covariation and twice differentiable functions f has been carried out by Junike et al. (2024, Theorem 4.7). In particular, the results by Junike et al. (2024, Theorem 4.7 and Remark 4.12) show, that Corollary 2.9 does not hold, if the risk factors incorporate a non-zero covariation $[X_i, X_j]$ for some $i \neq j$.

3 Embedding of traditional surplus decompositions into the ISU concept

This chapter is derived in part from an article published in the Scandinavian Actuarial Journal in 2022 (copyright Taylor & Francis), available online: http://www.tandfonline.com/10.1080/03461238.2022.2049636. More precisely, (parts of) Sections 2, 3, 5 and 6 of the article Jetses and Christiansen (2022) are included in this chapter. In order to improve readability and standardise the notation in this dissertation, minor changes have been made compared to the original article.

In traditional life insurance, the valuation assumptions include safety margins to protect against adverse economic and demographic developments. As a result of the conservative assumptions, the actual development of the risk factors is usually favourable to the life insurer, which leads to systematic surplus. By statute, a part of this surplus belongs to the policyholder and is therefore refunded (see Norberg, 1999). The refund terms often require an allocation of the surplus to its individual sources, see e.g. the German national directive *Mindestzuführungsverordnung* (BMF, 2016). It is therefore essential to decompose life insurance surplus into its individual surplus contributions.

The decomposititon of surplus is an old actuarial question that has been the focus of several publications. In a time-discrete life insurance setup, the so-called contribution formula (*Kontributionsformel*) is not only frequently used by German life insurers, but also appears in many standard references (see e.g. Milbrodt and Helbig, 1999, Section 11.B; Saxer, 1955, Section 9.5). This formula decomposes the yearly surplus into mortality surplus, interest surplus, lapse surplus and cost surplus. In time-continuous model frameworks, early attempts to decompose the surplus for traditional life insurance policies were made by Lidstone (1905) and Berger (1939) (see also Simonsen, 1970; Sverdrup, 1969 and references therein), who investigated the impact of changing valuation bases on prospective reserves. In later literature on surplus, the terms 'first-order basis' (prudent valuation basis) and 'second-order basis' (experience valuation basis) have become established. Based on the Markov chain setup introduced by Hoem (1969), Ramlau-Hansen (1988, 1991) investigated surplus in a multistate life insurance framework with deterministic valuation bases. Extending these contributions, Norberg (1999, 2001) not only allows for a second-order stochastic basis, but systematically defines surplus as the difference between a second-order retrospective reserve and a first-order prospective reserve, distinguishing between individual and portfolio surplus. While the above-mentioned literature focuses on the investment of surplus in risk-free assets, Møller and Steffensen (2007) as well as Asmussen and Steffensen (2020) derived surplus decompositions in a more sophisticated model, including a risky asset. Recent papers by Bruhn and Lollike (2020) as well as Falden and Nyegaard (2021) analyse the dynamics of surplus in terms of management actions and policyholder behaviour; however, these topics are beyond the scope of this chapter. Furthermore, it is worth to mention that focus is placed in this work on the decomposition of surplus ('bonus'), and not on the distribution of surplus ('bonus schemes', 'dividends').

The above listed references show that surplus decompositions have been derived in several frameworks, which differ in terms of time grid (time-discrete vs. time-continuous), randomness (deterministic vs. stochastic valuation bases), perspective (retrospective vs. prospective surplus) and information (individual vs. portfolio surplus) (see Steffensen, 2001, Section 3.7). However, all the decomposition formulas and their interpretations rely on heuristic reasoning, making it challenging to compare the different frameworks and their formulas. Furthermore, the existing literature does not pave the way for extending the frameworks by further risk drivers (see Chapter 1). It is shown that the proposed ISU decomposition principle addresses both shortcomings while providing further legitimacy to existing surplus decomposition formulas. In a general multistate life insurance setup, the known decomposition formulas are derived using appropriate choices of the revaluation surplus process, the risk basis and the link mapping between them. In particular, the time-discrete contribution formula is represented as an SU decomposition, which can therefore be interpreted as an approximation of the corresponding time-continuous ISU decomposition. The embedding of traditional surplus decomposition formulas into the ISU decomposition principle will allow for a comprehensive comparison of the existing decomposition formulas. Moreover, the clear idea of the ISU decomposition principle will open the door to the incorporation of further risk factors.

In Section 3.1, the decomposition of life insurance surplus is embedded into the framework of the previously presented ISU decomposition principle. The concepts of individual surplus and portfolio surplus, in line with Norberg (1999), are introduced in Section 3.2. Furthermore, different choices of the risk basis are discussed. In Section 3.3, general surplus decomposition formulas are derived by applying the ISU decomposition principle. As special cases of the main theorem, Section 3.4 contains the derivation of the traditional surplus decomposition formulas known from the literature.

3.1 The surplus process of an individual insurance contract

We consider an individual insurance policy on a finite contract period $[0, T]$. For each $t \geqslant 0$ let $B(t)$ be the aggregated insurance cash flow on $[0, t]$ between insurer and insured. We use the convention that premiums have a negative sign and benefits have a positive sign. Let κ be a semimartingale with $\kappa(0) = 1$ that describes the value process of the insurer's self-financing investment portfolio. Then the value $A(t)$ of the assets accrued at time t is given by

$$A(t) = -\int_{[0,t]} \frac{\kappa(t)}{\kappa(s)} \mathrm{d}B(s), \tag{3.1}$$

assuming that B is a finite variation semimartingale and that κ is strictly positive. In the hypothetical case that the insurer knew the future, the liabilities at time t would be likewise calculated as

$$L^h(t) = \int_{(t,T]} \frac{\kappa(t)}{\kappa(s)} \mathrm{d}B(s).$$

The difference between assets and liabilities is the surplus,

$$S^h(t) = A(t) - L^h(t) = \kappa(t)(A(0) - L^h(0)). \tag{3.2}$$

In this hypothetical setting, the actual surplus emerges at time zero and any dynamics after zero just comes from the compounding factor $\kappa(t)$. By defining

$$\mathrm{d}\Phi(t) = \frac{\mathrm{d}\kappa(t)}{\kappa(t-)}$$

as the return on investment of the insurer's investment portfolio, the process S^h satisfies

$$\mathrm{d}S^h(t) = S^h(t-)\mathrm{d}\Phi(t)$$

for $t > 0$, which shows again that the dynamics of S^h on $(0, \infty)$ stems solely from investment gains earned on the existing surplus. Since $A(0) - L^h(0)$ depends on the future and is nowhere adapted to the available information, in real life the insurer has to replace $A(0) - L^h(0)$ at each time t by an $\mathcal{F}_t$-measurable proxy $R(t)$. Since the process

$$R = (R(t))_{t \geqslant 0}$$

describes profits and losses that result from the continuous revaluation of $A(0) - L^h(0)$ as the information $\mathcal{F}_t$ increases with time t, we call R the *revaluation surplus process* (see Chapter 2). Now the *total surplus process* is given by

$$S(t) = \kappa(t)R(t), \quad t \geqslant 0, \tag{3.3}$$

and its dynamics is driven by both, the compounding factor κ and the revaluation surplus process R. As described in Chapter 2, we assume that the life insurance model rests on a risk basis

$$X = (X_1, \dots, X_m),$$

which is a multivariate adapted process composed of the risk factors $X_1, \dots, X_m$ such that R is adapted to the right-continuous and complete filtration generated by X. The risk basis is assumed to be fixed, but depending on R, different choices of X are conceivable.

The goal is to decompose R with help of the introduced ISU decomposition principle (see Chapter 2) into

$$R(t) = R(0) + D_1(t) + \cdots + D_m(t), \quad t \geqslant 0, \tag{3.4}$$

where $D_1, \ldots, D_m$ are adapted processes that start at zero and describe the contributions of each risk factor $X_1, \ldots, X_m$ to the dynamics of R. The first addend $R(0)$ represents initial surplus, which is not decomposed here. Equation (3.4) is equivalent to the additive decomposition

$$S(t) = \kappa(t)S(0) + \kappa(t)D_1(t) + \cdots + \kappa(t)D_m(t), \quad t \geqslant 0, \tag{3.5}$$

for the total surplus process. The first addend $\kappa(t)S(0)$ represents the time-t value of the initial surplus $S(0) = R(0)$, and the addends $\kappa(t)D_1(t), \ldots, \kappa(t)D_m(t)$ describe the time-t values of the contributions that the risk factors $X_1, \ldots, X_m$ make to the dynamics of S. The additivity of the decompositions (3.4) and (3.5) allows us to distribute the surplus among different parties.

The dynamics of the total surplus in (3.3) is driven by investment gains on the surplus itself and by revaluation gains. In (3.5) the investment gains are subdivided among the different surplus contribution addends according to their shares in the total investment earnings. It is not uncommon in the actuarial literature to collect all the investment gains in a separate term, see for example Norberg (1999, formula (5.3)). The idea is to apply Itô's product rule on $S(t) = \kappa(t)R(t)$ and then to identify each of the resulting addends either as investment gains or as revaluation gains. However, this approach mixes up the investment earnings of the carefully separated surplus contribution addends, so it is not helpful in our opinion and therefore it is not further considered in this chapter.

3.2 The revaluation surplus in multistate models

Let the random pattern Z of the insured be a right-continuous and adapted jump process on a finite state space $\mathcal{Z}$ with starting value $Z(0) = a \in \mathcal{Z}$. We define corresponding state processes $(I_j)_j$ and counting processes $(N_{jk})_{jk:j \neq k}$ by $I_i(t) := \mathbb{1}_{\{Z(t)=i\}}$ and

$$N_{jk}(t) = \sharp\{s \in (0,t] : Z(s-) = j, Z(s) = k\}, \; j,k \in \mathcal{Z}, \; j \neq k, \; t \geqslant 0.$$

Additionally, we define $N_{jj} := -\sum_{k:k \neq j} N_{jk}$, $j \in \mathcal{Z}$, and the vector-valued process $N = (N_{jk})_{jk:j \neq k}$.

We call a pair $(\overline{\Phi}, \overline{\Lambda})$ a valuation basis if the following properties hold:

- $\overline{\Phi}$ is semimartingale with $\overline{\Phi}(0) = 0$ and $\Delta\overline{\Phi}(t) > -1$ for all $t > 0$,
- $\overline{\Lambda} = (\overline{\Lambda}_{jk})_{jk:j\neq k}$ is a vector-valued, right-continuous finite variation process with $\overline{\Lambda}(0) = 0$,
- the processes $\overline{\Lambda}_{jk}$, $j \neq k$ are non-decreasing and $\sum_{k:k\neq j} \Delta\overline{\Lambda}_{jk}(t) \leqslant 1$ for every $t > 0$ and every j.

The process $\overline{\Phi}$ represents cumulative returns on investment, and the solution $\overline{\kappa} = (\overline{\kappa}(t))_{t\geqslant 0}$ of the stochastic differential equation

$$\mathrm{d}\overline{\kappa}(t) = \overline{\kappa}(t-)\mathrm{d}\overline{\Phi}(t), \quad \overline{\kappa}(0) = 1, \tag{3.6}$$

is the value process of a self-financing investment portfolio with respect to $\overline{\Phi}$.

Furthermore, given the valuation basis $(\overline{\Phi}, \overline{\Lambda})$, let $\overline{p} = (\overline{p}(t,s))_{0\leqslant t\leqslant s}$, $\overline{p}(t,s) = (\overline{p}_{jk}(t,s))_{jk}$ denote the solution of the stochastic differential equation system

$$\overline{p}_{jk}(t, \mathrm{d}s) = \sum_i \overline{p}_{ji}(t, s-)\mathrm{d}\overline{\Lambda}_{ik}(s),\ \overline{p}_{jk}(t,t) = \delta_{jk},\ s > t. \tag{3.7}$$

Observe that we may pick N itself for $\overline{\Lambda}$. In this case the solution of (3.7) satisfies $p_{aj}(0,s) = I_j(s)$, since $I_j(0) = \delta_{aj}$ and

$$\mathrm{d}I_j(s) = \sum_{k:k\neq j} (\mathrm{d}N_{kj}(s) - \mathrm{d}N_{jk}(s)) = \sum_k I_k(s-)\mathrm{d}N_{kj}(s). \tag{3.8}$$

Throughout this chapter, let the valuation basis (Φ, Λ) represent the so-called *second order valuation basis*. The process Φ describes the real return in investment in the insurer's investment portfolio. Let κ denote the solution of (3.6) with respect to Φ. For the second-order basis we additionally assume that

(S.1) Λ is a predictable process,

(S.2) conditional on $(\Phi, \Lambda) = (E, F)$, the process Z is a Markov process under $\mathbb{P}$ with cumulative transition intensity matrix F.

Thus, the process $I_j(t-)\mathrm{d}\Lambda_{jk}(t)$ is a $\mathbb{P}$-compensator of $\mathrm{d}N_{jk}$ with respect to the natural completed filtration of the random vector $(Z^t, \Phi, \Lambda)_{t\geqslant 0}$. Due to the conditional Markov property, the stochastic differential equation (3.7) with respect to Λ corresponds to the Kolmogorov forward equation of Z conditional on (Φ, Λ), and its solution $p(t,s) = (p_{jk}(t,s))_{jk}$ is the transition probability matrix of Z conditional on (Φ, Λ).

Furthermore, let the valuation basis (Φ^*, Λ^*) represent the so-called *first order valuation basis*. For this specific valuation basis we additionally assume that

(F.1) Φ^* and Λ^* are deterministic,

(F.2) Z is a Markov process under a prudent probability measure $\mathbb{P}^*$ with cumulative transition intensities Λ^*_{jk}, $j \neq k$,

(F.3) $(\mathbb{I} + \Delta\Lambda^*_M(s))^{-1}$ exists for every $s > 0$,

where Λ^*_M denotes the matrix-valued process $\Lambda^*_M = (\Lambda^*_{jk})_{jk}$ with $\Lambda^*_{jj} := -\sum_{k:k\neq j} \Lambda^*_{jk}$. Let κ^* and p^* be the solutions of (3.6) and (3.7) with respect to Φ^* and Λ^*, respectively. Under the first order valuation, (3.7) is the classical Kolmogorov forward equation and p^* is the classical transition probability matrix of Z under $\mathbb{P}^*$. The existence of $(\mathbb{I} + \Delta\Lambda^*_M(s))^{-1}$ for every $s > 0$ ensures that the matrix $p^*(t, s)$ has an inverse for each $s > t$, denoted as $q^*(t, s)$, see Lemma A.2.1 in the appendix. In particular, q^* satisfies the stochastic differential equation

$$q^*(t, \mathrm{d}s) = -(\mathrm{d}G(s))q^*(t, s-),\ q^*(t, t) = \mathbb{I},\ s > t,$$

where $G(s) = \Lambda^*_M(s) - \sum_{0<u\leqslant s}(\Delta\Lambda^*_M(u))^2(\mathbb{I} + \Delta\Lambda^*_M(u))^{-1}$ (see Lemma A.2.1).

Recall that the insurance policy shall have a finite contract horizon in $[0, T]$. We assume that the insurance cash flow B has the form

$$\mathrm{d}B(t) = \sum_j I_j(t-)\,\mathrm{d}B_j(t) + \sum_{jk:j\neq k} b_{jk}(t)\,\mathrm{d}N_{jk}(t), \tag{3.9}$$

where $(B_j)_j$ are right-continuous finite variation functions that satisfy $\mathrm{d}B_j(t) = 0$ for $t > T$, and $(b_{jk})_{jk:j\neq k}$ are bounded and measurable functions with $b_{jk}(t) = 0$ for $t > T$.

We generally assume that

(J) the processes Φ^*, Φ and $(N, \Lambda^*, \Lambda, (B_j)_j)$ have no simultaneous jumps.

The latter condition implies that the covariation between the investment risk and all other risk drivers is zero. This fact will help us to build additive decompositions by applying Itô's formula, see Lemma 3.7 below.

3.2.1 Individual revaluation surplus

In with-profit life insurance, the remaining future liabilities of the individual insurance contract at time t are commonly evaluated as

$$\sum_j I_j(t)V^*_j(t),$$

where $V_j^*(t)$ shall be the prospective reserve at time t in state j with respect to the first order valuation basis, see Norberg (1999). According to Milbrodt and Helbig (1999, Chapter 10.A), it holds that

$$\begin{aligned} V_j^*(t) =& \mathbb{E}^* \left[\int_t^T \frac{\kappa^*(t)}{\kappa^*(s)} \mathrm{d}B(s) \middle| Z(t) = j \right] \\ =& \sum_k \int_{(t,T]} \frac{\kappa^*(t)}{\kappa^*(s)} p_{jk}^*(t, s-) \mathrm{d}B_k(s) + \sum_{k,l:k \neq l} \int_{(t,T]} \frac{\kappa^*(t)}{\kappa^*(s)} p_{jk}^*(t, s-) b_{kl}(s) d\Lambda_{kl}^*(s), \end{aligned}$$

where $\mathbb{E}^*$ denotes the expectation with respect to $\mathbb{P}^*$ (see (F.2)). The accrued assets of the individual insurance contract at time t equal (3.1), so the total surplus of the individual policy at time t is

$$S(t) = -\int_{[0,t]} \frac{\kappa(t)}{\kappa(s)} \mathrm{d}B(s) - \sum_j I_j(t) V_j^*(t), \tag{3.10}$$

see Norberg (1999). The corresponding revaluation process R equals

$$R(t) = \frac{S(t)}{\kappa(t)} = -\int_{[0,t]} \frac{1}{\kappa(s)} \mathrm{d}B(s) - \sum_j \frac{1}{\kappa(t)} I_j(t) V_j^*(t). \tag{3.11}$$

Proposition 3.1. *For R defined by* (3.11) *and $t \in [0,T]$ it holds that*

$$R(t) = -H((\Phi^*, \Lambda^*) + (\Phi - \Phi^*, N - \Lambda^*)^t), \tag{3.12}$$

where $(\cdot)^t$ denotes the corresponding stopped process (see (2.1)*) and where for any valuation basis $(\overline{\Phi}, \overline{\Lambda})$ the mapping H is defined by*

$$H((\overline{\Phi}, \overline{\Lambda})) := \sum_j \int_{[0,T]} \frac{1}{\overline{\kappa}(s)} \overline{p}_{aj}(0, s-) \, dB_j(s) + \sum_{j,k:j \neq k} \int_{(0,T]} \frac{1}{\overline{\kappa}(s)} \overline{p}_{aj}(0, s-) b_{jk}(s) d\overline{\Lambda}_{jk}(s) \tag{3.13}$$

with $\overline{p}_{aj}(0, 0-) := \delta_{aj}$.

Proof. The solution of (3.7) with respect to the cumulative transition intensity vector $\Lambda^* + (N - \Lambda^*)^t$ is

$$\begin{cases} I_j(s), \;\; s \leqslant t, \\ \sum_l I_l(t) p_{lj}^*(t, s), \;\; s > t, \end{cases}$$

where $p_{lj}^*(t, s)$ is the solution of (3.7) with respect to the first order valuation basis. The solution of (3.6) with respect to $\Phi^* + (\Phi - \Phi^*)^t$ is

$$\begin{cases} \kappa(s), \;\; s \leqslant t, \\ \kappa(t) \frac{\kappa^*(s)}{\kappa^*(t)}, \;\; s > t, \end{cases}$$

where κ^* is the solution of (3.6) with respect to the first order valuation basis. By plugging these solutions into (3.13), we obtain the desired result. ☐

Proposition 3.1 allows us to represent R by

$$R(t) = \varrho(X^t), \quad t \geqslant 0,$$

for various choices of X and ϱ. For example, we may define the risk basis X and the mapping ϱ by means of the mapping H (see (3.13)) as follows:

Example 3.2. By setting

$$X = (X_\Phi, X_u, X_s) = (\Phi - \Phi^*, N - \Lambda, \Lambda - \Lambda^*),$$

we distinguish between financial risk, unsystematic biometric risk and systematic biometric risk, and we may define ϱ by

$$\varrho(X^t) = -H\big((\Phi^*, \Lambda^*) + (X^t_\Phi, X^t_u + X^t_s)\big).$$

Example 3.3. By setting

$$X = (X_\Phi, (X_{jk})_{jk:j\neq k}) = (\Phi - \Phi^*, (N_{jk} - \Lambda^*_{jk})_{jk:j\neq k}),$$

we distinguish between financial risk and transition-wise biometric risks, and we may define ϱ by

$$\varrho(X^t) = -H\big((\Phi^*, \Lambda^*) + (X^t_\Phi, (X^t_{jk})_{jk:j\neq k})\big).$$

Example 3.4. Let the processes $(\Phi_j)_j$ and $(\Phi^*_j)_j$ be defined by $\mathrm{d}\Phi_j(t) = I_j(t-)\mathrm{d}\Phi(t)$, $\Phi_j(0) = 0$, and $\mathrm{d}\Phi^*_j(t) = I_j(t-)\mathrm{d}\Phi^*(t)$, $\Phi^*_j(0) = 0$, respectively. Further, we denote $\Lambda_j = (\Lambda_{jk})_{k:k\neq j}$ and $\Lambda^*_j = (\Lambda^*_{jk})_{k:k\neq j}$. By setting

$$X = (X_u, (X_j)_j) = (X_u, (X_{j,1}, X_{j,2})_j) = (N - \Lambda, (\Phi_j - \Phi^*_j, \Lambda_j - \Lambda^*_j)_j),$$

we distinguish between unsystematic biometric risk and state-wise remaining risks, and we may define ϱ by

$$\varrho(X^t) = -H\left((\Phi^*, \Lambda^*_{jk}) + (0, X^t_u) + \left(\sum_j X^t_{j,1}, (X^t_{j,2})_j\right)\right).$$

3.2.2 Mean portfolio revaluation surplus

In actuarial practice it is not uncommon to focus on mean portfolio values only. We can replicate this perspective by applying the expectation $\mathbb{E}[\,\cdot\,|\Phi, \Lambda]$ on the individual values (3.10) and (3.11). The revaluation surplus takes then the form

$$R(t) = \mathbb{E}\bigg[- \int_{[0,t]} \frac{1}{\kappa(s)} \mathrm{d}B(s) - \sum_j \frac{1}{\kappa(t)} I_j(t) V^*_j(t) \bigg| \Phi, \Lambda \bigg], \tag{3.14}$$

and the corresponding total surplus still satisfies the equation

$$S(t) = \kappa(t) R(t). \tag{3.15}$$

Note that Norberg (1999) applies the expectation $\mathbb{E}[\,\cdot\,|\Phi^t, \Lambda^t]$ instead, but his definition is equivalent since

$$\begin{aligned} S(t) = &- \int_{[0,t]} \frac{\kappa(t)}{\kappa(s)} \sum_j \bigg(p_{aj}(0, s-) \mathrm{d}B_j(s) + \sum_{k:k\neq j} b_{jk}(s) p_{aj}(0, s-) \mathrm{d}\Lambda_{jk}(s) \bigg) \\ &- \sum_j p_{aj}(0,t) V_j^*(t) \end{aligned}$$

is $\sigma(\Phi^t, \Lambda^t)$-measurable. The following corollary is a direct consequence of Proposition 3.1.

Corollary 3.5. *For R defined by* (3.14) *and $t \in [0,T]$ it holds that*

$$R(t) = \mathbb{E}\big[- H((\Phi^*, \Lambda^*) + (\Phi - \Phi^*, N - \Lambda^*)^t) \big| \Phi, \Lambda \big], \tag{3.16}$$

where H is given by (3.13)*.*

Because of the latter corollary, in the Examples 3.2 to 3.4 we just need to add the conditional expectation $\mathbb{E}[\,\cdot\,|\Phi, \Lambda]$ to the definition of ϱ in order to get to the mean portfolio perspective.

The next example is in particular relevant in German life insurance.

Example 3.6. Consider a life insurance contract with the states active, surrendered and dead,

$$\mathcal{Z} = \{a, s, d\},$$

of an x-year old insured. We assume that Λ^* and Λ are absolutely continuous with densities λ^* and λ, respectively. Let

$$\begin{aligned} {}_{k-l}p^*_{x+l} &= p^*_{aa}(x+l, x+k), \\ q^*_{x+k-1} &= p^*_{ad}(x+k-1, x+k), \\ r^*_{x+k-1} &= p^*_{as}(x+k-1, x+k), \end{aligned}$$

and define ${}_{k-l}p_{x+l}$, q_{x+k-1} and r_{x+k-1} likewise for the second-order valuation basis. We assume that sojourn payments occur only in state active and only as lump sum payments b_k at integer times k. Furthermore, we assume that the death benefit function and the surrender benefit function have the form

$$b_{ad}(t) = \frac{\kappa(\lfloor t \rfloor)}{\kappa(t)} d_{\lceil t \rceil}, \qquad b_{as}(t) = \frac{\kappa(\lfloor t \rfloor)}{\kappa(t)} s_{\lceil t \rceil},$$

where $d_{\lceil t \rceil}$ and $s_{\lceil t \rceil}$ represent the death benefit and surrender benefit in year $\lceil t \rceil$. This definition of b_{ad} and b_{as} discounts death benefits and surrender benefits as if they are paid out at the end of the year, so that V_a^* has at integer times l the representation

$$V_a^*(l) = \sum_{k=l+1}^{T} \frac{\kappa^*(l)}{\kappa^*(k)} \; {}_{k-l}p^*_{x+l} \, b_k + \sum_{k=l+1}^{T} \frac{\kappa^*(l)}{\kappa^*(k)} \; {}_{k-l-1}p^*_{x+l} \big(d_k \, q^*_{x+k-1} + s_k \, r^*_{x+k-1} \big).$$

We define yearly interest rates of first order and second order by

$$i_k^* = e^{\int_k^{k+1} \phi^*(u)\,\mathrm{d}u} - 1, \qquad i_k = e^{\int_k^{k+1} \phi(u)\,\mathrm{d}u} - 1, \qquad k \in \mathbb{N}_0.$$

One can show that the yearly increments of the mean portfolio revaluation surplus process equal

$$\begin{aligned}&R(k+1) - R(k)\\ &= \mathrm{e}^{-\int_0^{k+1} \phi(u)\,\mathrm{d}u}\, {}_kp_x \Big(V_a^*(k)\,(1+i_k) - q_{x+k}\, d_{k+1} - r_{x+k}\, s_{k+1} - p_{x+k}\, \big(b_{k+1} + V_a^*(k+1)\big)\Big).\end{aligned}$$

This formula is widely used in German life insurance (see Milbrodt & Helbig, 1999, Section 11.B). It is common in Germany to decompose the increments $R(k+1) - R(k)$ into investment surplus, mortality surplus and lapse surplus. For that purpose, analogously to Example 3.3 we choose

$$X = (X_\Phi, X_{ad}, X_{as}) = (\Phi - \Phi^*, N_{ad} - \Lambda_{ad}^*, N_{as} - \Lambda_{as}^*)$$

as risk basis.

3.3 ISU decompositions in multistate life insurance

This section contains general technical results that will be needed for the examples in the next section. For any valuation basis $(\overline{\Phi}, \overline{\Lambda})$, we write

$$\widetilde{\overline{\Phi}}(t) = \overline{\Phi}(t) - [\overline{\Phi}, \overline{\Phi}]^c(t) - \sum_{0<s\leqslant t} (1 + \Delta\overline{\Phi}(s))^{-1} (\Delta\overline{\Phi}(s))^2,$$

where $[\overline{\Phi}, \overline{\Phi}]^c$ signifies the continuous part of $[\overline{\Phi}, \overline{\Phi}]$.

Moreover, let R_{jk}^*, $j \neq k$ denote the first order sum at risk, i.e.

$$R_{jk}^*(t) = b_{jk}(t) + V_k^*(t) - V_j^*(t).$$

Recalling that

$$H((\overline{\Phi}, \overline{\Lambda})) := \sum_j \int_{[0,T]} \frac{1}{\overline{\kappa}(s)} \overline{p}_{aj}(0, s-) \mathrm{d}B_j(s) + \sum_{j,k:j\neq k} \int_{(0,T]} \frac{1}{\overline{\kappa}(s)} \overline{p}_{aj}(0, s-) b_{jk}(s) d\overline{\Lambda}_{jk}(s),$$

for any valuation basis $(\overline{\Phi}, \overline{\Lambda})$ (see (3.13)), we can pose the following results.

Lemma 3.7. *Let $(\overline{\Phi}, \overline{\Lambda})$ be a valuation basis such that $(\Phi^*, \overline{\Phi})$ and $(\Lambda^*, \overline{\Lambda}, (B_j)_j)$ have no simultaneous jumps. Then it holds that*

$$\begin{aligned}H\big((\Phi^*, \Lambda^*) + (\overline{\Phi} - \Phi^*, \overline{\Lambda} - \Lambda^*)^t\big) &= \int_{(0,t]} \frac{1}{\overline{\kappa}(s-)} \sum_j \overline{p}_{aj}(0, s-) V_j^*(s-)\, d(\widetilde{\overline{\Phi}} - \Phi^* + [\widetilde{\overline{\Phi}}, \Phi^*])(s)\\ &\quad - \sum_{jk:j\neq k} \int_{(0,t]} \frac{1}{\overline{\kappa}(s)} \overline{p}_{aj}(0, s-) R_{jk}^*(s)\, d(\overline{\Lambda}_{jk} - \Lambda_{jk}^*)(s).\end{aligned}$$

Proof. As a shorthand notation, we define multivariate processes $C^* = (C_1^*, \dots, C_n^*)^\top$ and $\overline{C} = (\overline{C}_1, \dots, \overline{C}_n)^\top$ by

$$\mathrm{d}C_j^*(s) = \mathrm{d}B_j(s) + \sum_{k:k\neq j} b_{jk}(s)\mathrm{d}\Lambda_{jk}^*(s),\ C_j^*(0) = 0,$$
$$\mathrm{d}\overline{C}_j(s) = \mathrm{d}B_j(s) + \sum_{k:k\neq j} b_{jk}(s)\mathrm{d}\overline{\Lambda}_{jk}(s),\ \overline{C}_j(0) = 0.$$

Note that C^* and $\overline{C}$ are column vectors. The vectorial process $I = (I_j)_j$ shall combine all state processes as a row vector. We further define

$$W(s) := -H((\Phi^*, \Lambda^*) + (\overline{\Phi} - \Phi^*, \overline{\Lambda} - \Lambda^*)^s),$$

where H is given by (3.13). Due to the assumptions made on the first-order valuation basis in Section 3.2 (see (F.3) and the follow-up remarks), $p^*(0,s)$ is invertible with inverse $q^*(0,s)$.
Thus, for $s \in (0,t]$, we get

$$\begin{aligned}
W(s) &= -\int_{[0,s]} \frac{1}{\overline{\kappa}(u)} I(0)\overline{p}(0,u-)\mathrm{d}\overline{C}(u) - \frac{1}{\overline{\kappa}(s)} \int_{(s,T]} \frac{\kappa^*(s)}{\kappa^*(u)} I(0)\overline{p}(0,s)p^*(s,u-)\mathrm{d}C^*(u) \\
&= -\int_{[0,s]} \frac{1}{\overline{\kappa}(u)} I(0)\overline{p}(0,u-)\mathrm{d}\overline{C}(u) - \frac{\kappa^*(s)}{\overline{\kappa}(s)} I(0)\overline{p}(0,s)q^*(0,s)\,Y(s),
\end{aligned}$$

for $Y(s) = \int_{(s,T]} \frac{1}{\kappa^*(u)} p^*(0,u-)\mathrm{d}C^*(u)$. Analogously to Λ_M^*, let $\overline{\Lambda}_M$ denote the matrix-valued process $\overline{\Lambda}_M = (\overline{\Lambda}_{jk})_{jk}$ with $\overline{\Lambda}_{jj} := -\sum_{k:k\neq j} \overline{\Lambda}_{jk}$. By applying Itô's formula and using the assumption that $(\Phi^*, \overline{\Phi})$ and $(\Lambda^*, \overline{\Lambda}, (B_j)_j)$ have no common jumps, we can show that

$$\begin{aligned}
\mathrm{d}W(s) &= -\frac{1}{\overline{\kappa}(s)} I(0)\overline{p}(0,s-)\mathrm{d}(\overline{C} - C^*)(s) \\
&\quad - I(0)\frac{\kappa^*(s-)}{\overline{\kappa}(s-)}\overline{p}(0,s-)q^*(0,s-)Y(s-)\mathrm{d}(\Phi^* - \tilde{\overline{\Phi}} - [\Phi^*, \tilde{\overline{\Phi}}])(s) \\
&\quad - I(0)\frac{\kappa^*(s-)}{\overline{\kappa}(s-)}\overline{p}(0,s-)\mathrm{d}(\overline{\Lambda}_M - \Lambda_M^*)(s)\, q^*(0,s)Y(s) \\
&= -\frac{1}{\overline{\kappa}(s)} I(0)\overline{p}(0,s-)\mathrm{d}(\overline{C} - C^*)(s) \\
&\quad - \frac{1}{\overline{\kappa}(s-)} I(0)\overline{p}(0,s-)\left(\int_{[s,T]} \frac{\kappa^*(s-)}{\kappa^*(u)} p^*(s-,u-)\mathrm{d}C^*(u)\right)\mathrm{d}(\Phi^* - \tilde{\overline{\Phi}} - [\Phi^*, \tilde{\overline{\Phi}}])(s) \\
&\quad - \frac{1}{\overline{\kappa}(s)} I(0)\overline{p}(0,s-)\mathrm{d}(\overline{\Lambda}_M - \Lambda_M^*)(s)\left(\int_{(s,T]} \frac{\kappa^*(s)}{\kappa^*(u)} p^*(s,u-)\mathrm{d}C^*(u)\right),
\end{aligned}$$

where we used Lemma A.2.2 to get

$$\begin{aligned}
&\mathrm{d}\left(\frac{\kappa^*(s)}{\overline{\kappa}(s)}\overline{p}(0,s)q^*(0,s)\right) \\
&= \frac{\kappa^*(s-)}{\overline{\kappa}(s-)}\mathrm{d}\left(\overline{p}(0,s)q^*(0,s)\right) + \overline{p}(0,s-)q^*(0,s-)\mathrm{d}\left(\frac{\kappa^*(s)}{\overline{\kappa}(s)}\right) + \mathrm{d}\left[\frac{\kappa^*}{\overline{\kappa}}, \overline{p}(0,\cdot)q^*(0,\cdot)\right](s)
\end{aligned}$$

$$= \frac{\kappa^*(s-)}{\overline{\kappa}(s-)}\overline{p}(0,s-)\mathrm{d}(\overline{\Lambda}_M - \Lambda_M^*)(s)\, q^*(0,s) + \frac{\kappa^*(s-)}{\overline{\kappa}(s-)}\overline{p}(0,s-)q^*(0,s-)\mathrm{d}(\Phi^* - \tilde{\overline{\Phi}} - [\Phi^*, \tilde{\overline{\Phi}}])(s)$$

with $\tilde{\overline{\Phi}}(s) = \overline{\Phi}(s) - [\overline{\Phi}, \overline{\Phi}]^c(s) - \sum_{0<u\leqslant s}(1+\Delta\overline{\Phi}(u))^{-1}(\Delta\overline{\Phi}(u))^2$. Component-wise evaluation and integration on $(0,t]$ gives us the assertion. □

Theorem 3.8. *Let the processes $(\Phi_j)_j$ and $(\Phi_j^*)_j$ be defined by $d\Phi_j(t) = I_j(t-)d\Phi(t)$, $\Phi_j(0) = 0$, and $d\Phi_j^*(t) = I_j(t-)d\Phi^*(t)$, $\Phi_j^*(0) = 0$, respectively. For $j,k \in \mathcal{Z}$ let*

$$X_{\Phi,j} = \Phi_j - \Phi_j^*,$$
$$X_{u,jk} = N_{jk} - \Lambda_{jk},$$
$$X_{s,jk} = \Lambda_{jk} - \Lambda_{jk}^*,$$

and set $X = ((X_{\Phi,j})_j, (X_{u,jk})_{j,k:j\neq k}, (X_{s,jk})_{j,k:j\neq k})$. Then

$$\varrho(X^t) = -H\left((\Phi^*, \Lambda^*) + \left(\sum_j X_{\Phi,j}, (X_{u,jk} + X_{s,jk})_{jk:j\neq k}\right)^t\right)$$

has the ISU decomposition

$$D_{\Phi,j}(t) = \int_{(0,t]} \frac{1}{\kappa(s-)} I_j(s-) V_j^*(s-)\, d(\tilde{\Phi} - \Phi^*)(s),$$
$$D_{u,jk}(t) = -\int_{(0,t]} \frac{1}{\kappa(s)} I_j(s-) R_{jk}^*(s)\, d(N_{jk} - \Lambda_{jk})(s),$$
$$D_{s,jk}(t) = -\int_{(0,t]} \frac{1}{\kappa(s)} I_j(s-) R_{jk}^*(s)\, d(\Lambda_{jk} - \Lambda_{jk}^*)(s).$$

In particular, the ISU decomposition does not depend on the update order or the choice of partitions.

Proof. Let $J_\Phi \subseteq \mathcal{Z}$ and $J_u, J_s \subseteq \mathcal{J} := \{(j,k) \in \mathcal{Z}^2 : j \neq k\}$. For $r \leqslant s$, we define

$$X_{\Phi,J_\Phi,j}^{r,s} := \begin{cases} X_{\Phi,j}^r, & j \notin J_\Phi, \\ X_{\Phi,j}^s, & j \in J_\Phi, \end{cases}$$

as well as

$$X_{u,J_u,jk}^{r,s} := \begin{cases} X_{u,jk}^r, & (j,k) \notin J_u, \\ X_{u,jk}^s, & (j,k) \in J_u, \end{cases} \quad X_{s,J_s,jk}^{r,s} := \begin{cases} X_{s,jk}^r, & (j,k) \notin J_s, \\ X_{s,jk}^s, & (j,k) \in J_s. \end{cases}$$

We further set

$$X_{\Phi,J_\Phi} := \sum_j X_{\Phi,J_\Phi,j}^{0,T},\ X_{u,J_u} := \left(X_{u,J_u,jk}^{0,T}\right)_{jk},\ X_{s,J_s} := \left(X_{s,J_s,jk}^{0,T}\right)_{jk},$$

where $X_{u,J_u,jj}^{0,T} = -\sum_{k:k\neq j} X_{u,J_u,jk}^{0,T}$ and $X_{s,J_s,jj}^{0,T} = -\sum_{k:k\neq j} X_{s,J_s,jk}^{0,T}$. Let $\Phi^{J_\Phi} := X_{\Phi,J_\Phi} + \Phi^*$ and let κ^{J_Φ} denote the solution of $\mathrm{d}\kappa^{J_\Phi}(s) = \kappa^{J_\Phi}(s-)\mathrm{d}\Phi^{J_\Phi}(s)$ with $\kappa^{J_\Phi}(0) = 1$. Similarly, for $J = (J_u, J_s)$ let $\Lambda^J := X_{u,J_u} + X_{s,J_s} + \Lambda_M^*$ and let $p^J = (p_{jk})_{j,k}$ denote the solution of

$p^J(r,\mathrm{d}s) = p^J(r,s-)\mathrm{d}\Lambda^J(s)$ with $p^J(r,r)$ being the identity matrix.

Let $t \in [0,T]$ and let $(\mathcal{T}_n(t))_n$ be a sequence of partitions of $[0,t]$. For a simpler notation, we only write t_k instead of t_k^n for the grid points in $\mathcal{T}_n$. Throughout the proof, let $\alpha_n(s)$ be the left point of s in $\mathcal{T}_n(t)$, i.e. $\alpha_n(s) := t_k$ if $s \in (t_k, t_{k+1}]$. For notational convenience, we write

$$\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s} = \varrho((X^{t_k,t_{k+1}}_{\Phi,J_\Phi,j})_j, (X^{t_k,t_{k+1}}_{u,J_u,jk})_{jk:j\neq k}, (X^{t_k,t_{k+1}}_{s,J_s,jk})_{jk:j\neq k}).$$

It is sufficient to show that

i) $\operatorname*{plim}_{n\to\infty} \sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} \left(\varrho^{t_k,t_{k+1}}_{J_\Phi\cup\{j_0\},J_u,J_s} - \varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s}\right) = D_{\Phi,j_0}(t),\ j_0 \in \mathcal{Z}\backslash J_\Phi,$

ii) $\operatorname*{plim}_{n\to\infty} \sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} \left(\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u\cup\{(j_0,k_0)\},J_s} - \varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s}\right) = D_{u,j_0k_0}(t),\ (j_0,k_0) \in \mathcal{J}\backslash J_u,$

iii) $\operatorname*{plim}_{n\to\infty} \sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} \left(\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s\cup\{(j_0,k_0)\}} - \varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s}\right) = D_{s,j_0k_0}(t),\ (j_0,k_0) \in \mathcal{J}\backslash J_s.$

We prove the convergences consecutively.

i) Let $\overline{J}_\Phi = J_\Phi \cup \{j_0\}$, $j_0 \in \mathcal{Z}\backslash J_\Phi$ and let

$$\Delta(r,s) = \frac{\kappa^{\overline{J}_\Phi}(r)}{\kappa^{\overline{J}_\Phi}(s)} - \frac{\kappa^{J_\Phi}(r)}{\kappa^{J_\Phi}(s)},\ r \leqslant s.$$

We define stochastic processes

$$\begin{aligned}
\xi_{\Phi,j_0,n}(s) &= \frac{1}{\kappa(\alpha_n(s))}\frac{\kappa^{\overline{J}_\Phi}(\alpha_n(s))}{\kappa^{\overline{J}_\Phi}(s-)}\sum_g I_g(t_k)\sum_j p^J_{gj}(\alpha_n(s),s-)V_j^*(s-)I_{j_0}(s-),\\
\xi_{\Phi,j,n}(s) &= \frac{\Delta(\alpha_n(s),s-)}{\kappa(\alpha_n(s))}\sum_g I_g(\alpha_n(s))\sum_j p^J_{gj}(\alpha_n(s),s-)V_j^*(s-)I_j(s-),\ j \in J_\Phi,\\
\xi_{us,jk,n}(s) &= -\sum_g I_g(\alpha_n(s))\frac{\Delta(\alpha_n(s),s)}{\kappa(\alpha_n(s))}p^J_{gj}(\alpha_n(s),s-)R^*_{jk}(s),\ (j,k) \in J_u \cup J_s,
\end{aligned}$$

where $s \in [0,t]$. Due to (J), we can apply Lemma 3.7, which gives us

$$\begin{aligned}
&\sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} (\varrho^{t_k,t_{k+1}}_{J_\Phi\cup\{j_0\},J_u,J_s} - \varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s})\\
&= \sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} (\varrho^{t_k,t_{k+1}}_{J_\Phi\cup\{j_0\},J_u,J_s} - \varrho^{t_k,t_{k+1}}_{\varnothing,\varnothing,\varnothing} - (\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s} - \varrho^{t_k,t_{k+1}}_{\varnothing,\varnothing,\varnothing}))\\
&= \sum_{j\in\overline{J}_\Phi}\int_{(0,t]}\xi_{\Phi,j,n}(s)\mathrm{d}(\widetilde{\Phi}-\Phi^*)(s) + \sum_{(j,k)\in J_u}\int_{(0,t]}\xi_{us,jk,n}(s)\mathrm{d}(N_{jk}-\Lambda_{jk})(s)\\
&\quad + \sum_{(j,k)\in J_s}\int_{(0,t]}\xi_{us,jk,n}(s)\mathrm{d}(\Lambda_{jk}-\Lambda^*_{jk})(s),
\end{aligned}$$

where $\widetilde{\Phi}(s) = \Phi(s) - [\Phi,\Phi]^c(s) - \sum_{0<u\leqslant s}(1+\Delta\Phi^1(u))^{-1}(\Delta\Phi(u))^2$. Here, we used that

$$\mathrm{d}(\widetilde{\Phi^{J_\Phi}} - \Phi^* + [\widetilde{\Phi^{J_\Phi}},\Phi^*])(s) = \sum_{j\in J_\Phi} I_j(s-)\mathrm{d}(\widetilde{\Phi}-\Phi^*)(s),$$

exploiting that Φ and Φ^* have no common jumps (see (J)). Since for every $s \in [0,t]$ we almost surely have

$$\begin{aligned}
\lim_{n\to\infty} \xi_{\Phi,j_0,n}(s) &= \frac{1}{\kappa(s-)} I_{j_0}(s-)V^*_{j_0}(s-),\\
\lim_{n\to\infty} \xi_{\Phi,j,n}(s) &= 0,\ j \in J_\Phi,\\
\lim_{n\to\infty} \xi_{us,jk,n}(s) &= I_j(s-)\frac{\Delta(s-,s)}{\kappa(s-)} R^*_{jk}(s),\ (j,k) \in J_u \cup J_s,
\end{aligned}$$

and since $\Delta(s-,s)\mathrm{d}(N_{jk}-\Lambda_{jk})(s) = \Delta(s-,s)\mathrm{d}(\Lambda_{jk}-\Lambda^*_{jk})(s) = 0$ almost surely, the dominated convergence theorem for stochastic integrals (Protter, 2005, Chapter IV, Theorem 32) yields

$$\operatorname*{plim}_{n\to\infty} \sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} \left(\varrho^{t_k,t_{k+1}}_{J_\Phi\cup\{j_0\},J_u,J_s} - \varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s}\right) = D_{\Phi,j_0}(t).$$

ii) Let $\overline{J} = (J_u \cup \{j_0,k_0\}, J_s)$, $(j_0,k_0) \in \mathcal{J}\backslash J_u$ and let

$$\Delta_{jk}(r,s) \coloneqq p^{\overline{J}}_{jk}(r,s) - p^{J}_{jk}(r,s),\ r \leqslant s.$$

We define stochastic processes

$$\begin{aligned}
\xi_{\Phi,j,n}(s) &= \frac{1}{\kappa(\alpha_n(s))}\frac{\kappa^{J_\Phi}(\alpha_n(s))}{\kappa^{J_\Phi}(s-)}\sum_g I_g(\alpha_n(s))\sum_j \Delta_{gj}(\alpha_n(s),s-)V^*_j(s-)I_j(s-),\ j \in J_\Phi,\\
\xi_{us,jk,n}(s) &= -\sum_g I_g(\alpha_n(s))\frac{1}{\kappa(\alpha_n(s))}\frac{\kappa^{J_\Phi}(\alpha_n(s))}{\kappa^{J_\Phi}(s)}\Delta_{gj}(\alpha_n(s),s-)R^*_{jk}(s),\ (j,k) \in J_u \cup J_s,\\
\xi_{u,j_0k_0,n}(s) &= -\sum_g I_g(\alpha_n(s))\frac{1}{\kappa(\alpha_n(s))}\frac{\kappa^{J_\Phi}(\alpha_n(s))}{\kappa^{J_\Phi}(s)} p^{\overline{J}}_{gj_0}(\alpha_n(s),s-)R^*_{j_0k_0}(s),
\end{aligned}$$

where $s \in [0,t]$. Again with Lemma 3.7, we have

$$\begin{aligned}
&\sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} (\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u\cup\{(j_0,k_0)\},J_s} - \varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s})\\
&= \sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} (\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u\cup\{(j_0,k_0)\},J_s} - \varrho^{t_k,t_{k+1}}_{\varnothing,\varnothing,\varnothing} - (\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s} - \varrho^{t_k,t_{k+1}}_{\varnothing,\varnothing,\varnothing}))\\
&= \sum_{j\in J_\Phi}\int_{(0,t]} \xi_{\Phi,j,n}(s)\mathrm{d}(\widetilde{\Phi}-\Phi^*)(s) + \sum_{(j,k)\in J_u}\int_{(0,t]} \xi_{us,jk,n}(s)\mathrm{d}(N_{jk}-\Lambda_{jk})(s)\\
&\quad + \int_{(0,t]} \xi_{u,j_0k_0,n}(s)\mathrm{d}(N_{j_0k_0}-\Lambda_{j_0k_0})(s) + \sum_{(j,k)\in J_s}\int_{(0,t]} \xi_{us,jk,n}(s)\mathrm{d}(\Lambda_{jk}-\Lambda^*_{jk})(s)
\end{aligned}$$

Since for every $s \in [0,t]$ we almost surely have

$$\begin{aligned}
\lim_{n\to\infty} \xi_{u,j_0k_0,n}(s) &= -\frac{1}{\kappa(s-)} I_{j_0}(s-)\frac{\kappa^{J_\Phi}(s-)}{\kappa^{J_\Phi}(s)} R^*_{j_0k_0}(s)\\
&= -\frac{1}{\kappa(s)} I_{j_0}(s-)\frac{1+\Delta\Phi(s)}{1+\Delta\Phi^{J_\Phi}(s)} R^*_{j_0k_0}(s),\\
\lim_{n\to\infty} \xi_{\Phi,j,n}(s) &= \lim_{n\to\infty} \xi_{us,jk,n}(s) = 0,\ j \in J_\Phi,\ (j,k) \in J_u \cup J_s,
\end{aligned}$$

and since $\frac{1+\Delta\Phi(s)}{1+\Delta\Phi^{J_\Phi}(s)}\mathrm{d}(N_{j_0k_0}-\Lambda_{j_0k_0})(s) = \mathrm{d}(N_{j_0k_0}-\Lambda_{j_0k_0})(s)$ almost surely, the dominated convergence theorem for stochastic integrals (Protter, 2005, Chapter IV, Theorem 32) yields

$$\operatorname*{plim}_{n\to\infty}\sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)}\left(\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u\cup\{(j_0,k_0)\},J_s}-\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s}\right)=D_{u,j_0k_0}(t).$$

iii) Let $\overline{J}=(J_u,J_s\cup\{j_0,k_0\})$, $(j_0,k_0)\in\mathcal{J}\backslash J_s$ and let

$$\Delta_{jk}(r,s):=p^{\overline{J}}_{jk}(r,s)-p^{J}_{jk}(r,s),\ r\leqslant s.$$

We define stochastic processes

$$\xi_{\Phi,j,n}(s)=\frac{1}{\kappa(\alpha_n(s))}\frac{\kappa^{J_\Phi}(\alpha_n(s))}{\kappa^{J_\Phi}(s-)}\sum_g I_g(\alpha_n(s))\sum_j\Delta_{gj}(\alpha_n(s),s-)V^*_j(s-)I_j(s-),\ j\in J_\Phi,$$

$$\xi_{us,jk,n}(s)=-\sum_g I_g(\alpha_n(s))\frac{1}{\kappa(\alpha_n(s))}\frac{\kappa^{J_\Phi}(\alpha_n(s))}{\kappa^{J_\Phi}(s)}\Delta_{gj}(\alpha_n(s),s-)R^*_{jk}(s),\ (j,k)\in J_u\cup J_s,$$

$$\xi_{s,j_0k_0,n}(s)=-\sum_g I_g(\alpha_n(s))\frac{1}{\kappa(\alpha_n(s))}\frac{\kappa^{J_\Phi}(\alpha_n(s))}{\kappa^{J_\Phi}(s)}p^{\overline{J}}_{gj_0}(\alpha_n(s),s-)R^*_{j_0k_0}(s),$$

where $s\in[0,t]$. Again with Lemma 3.7, we have

$$\begin{aligned}
&\sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)}(\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s\cup\{(j_0,k_0)\}}-\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s})\\
&=\sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)}(\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s\cup\{(j_0,k_0)\}}-\varrho^{t_k,t_{k+1}}_{\varnothing,\varnothing,\varnothing}-(\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s}-\varrho^{t_k,t_{k+1}}_{\varnothing,\varnothing,\varnothing}))\\
&=\sum_{j\in J_\Phi}\int_{(0,t]}\xi_{\Phi,j,n}(s)\mathrm{d}(\widetilde{\Phi}-\Phi^*)(s)+\sum_{(j,k)\in J_u}\int_{(0,t]}\xi_{us,jk,n}(s)\mathrm{d}(N_{jk}-\Lambda_{jk})(s)\\
&\quad+\sum_{(j,k)\in J_s}\int_{(0,t]}\xi_{us,jk,n}(s)\mathrm{d}(\Lambda_{jk}-\Lambda^*_{jk})(s)+\int_{(0,t]}\xi_{s,j_0k_0,n}(s)\mathrm{d}(\Lambda_{j_0k_0}-\Lambda^*_{j_0k_0})(s).
\end{aligned}$$

Since for every $s\in[0,t]$ we almost surely have

$$\begin{aligned}
\lim_{n\to\infty}\xi_{s,j_0k_0,n}(s)&=-\frac{1}{\kappa(s-)}I_{j_0}(s-)\frac{\kappa^{J_\Phi}(s-)}{\kappa^{J_\Phi}(s)}R^*_{j_0k_0}(s)\\
&=-\frac{1}{\kappa(s)}I_{j_0}(s-)\frac{1+\Delta\Phi(s)}{1+\Delta\Phi^{J_\Phi}(s)}R^*_{j_0k_0}(s),\\
\lim_{n\to\infty}\xi_{\Phi,j,n}(s)&=\lim_{n\to\infty}\xi_{us,jk,n}(s)=0,\ j\in J_\Phi,\ (j,k)\in J_u\cup J_s,
\end{aligned}$$

and since $\frac{1+\Delta\Phi(s)}{1+\Delta\Phi^{J_\Phi}(s)}\mathrm{d}(\Lambda_{j_0k_0}-\Lambda^*_{j_0k_0})(s) = \mathrm{d}(\Lambda_{j_0k_0}-\Lambda^*_{j_0k_0})(s)$ almost surely, the dominated convergence theorem for stochastic integrals (Protter, 2005, Chapter IV, Theorem 32) yields

$$\operatorname*{plim}_{n\to\infty}\sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)}\left(\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s\cup\{(j_0,k_0)\}}-\varrho^{t_k,t_{k+1}}_{J_\Phi,J_u,J_s}\right)=D_{s,j_0k_0}(t).$$

□

Lemma 3.9. *Let $X = (X_1, \dots, X_m)$ be a given risk basis with*

$$R(t) = \varrho((X_1, \dots, X_m)^t)$$

for a suitable mapping ϱ, generating the ISU decomposition $D(t) = (D_1(t), \dots, D_m(t))$ with respect to $(\mathcal{T}_n(t))_n$, and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{A}$. Suppose that the SU decomposition $D^n(t) = (D_1^n(t), \dots, D_m^n(t))$ of $R(t) - R(0)$ with respect to $\mathcal{T}_n(t)$ satisfies $|D_i^n(t)| \leqslant Y$, $i = 1, \dots, m$, $n \in \mathbb{N}$, for some integrable random variable Y. Then the ISU decomposition of

$$\widetilde{R}(t) = \widetilde{\varrho}((X_1, \dots, X_m)^t) := \mathbb{E}\left[\varrho((X_1, \dots, X_m)^t)|\mathcal{G}\right]$$

is given by

$$\widetilde{D}(t) = (\mathbb{E}[D_1(t)|\mathcal{G}], \dots, \mathbb{E}[D_m(t)|\mathcal{G}]).$$

Proof. Since the revaluation surplus surfaces U and $\widetilde{U}$ are linked via the equation

$$\widetilde{U}(t_1, \dots, t_m) = \mathbb{E}[U(t_1, \dots, t_m)|\mathcal{G}],$$

the SU decomposition of $\widetilde{R}(t) - \widetilde{R}(0)$ is given by $\widetilde{D}^n(t) = (\mathbb{E}[D_1^n(t)|\mathcal{G}], \dots, \mathbb{E}[D_m^n(t)|\mathcal{G}])$. Using that $|D_i^n(t)| \leqslant Y$, $i = 1, \dots, m$, for some integrable random variable Y and the fact that stochastically converging sequences have almost surely converging subsequences, the dominated convergence theorem for conditional expectations almost surely yields

$$\widetilde{D}_i(t) = \lim_{n \to \infty} \mathbb{E}[D_i^n(t)|\mathcal{G}] = \mathbb{E}[D_i(t)|\mathcal{G}], \; i = 1, \dots, m.$$

□

Theorem 3.10. *Let X be defined as in Theorem 3.8. Then*

$$\varrho(X^t) = \mathbb{E}\left[-H\left((\Phi^*, \Lambda^*) + \left(\sum_j X_{\Phi,j}, (X_{u,jk} + X_{s,jk})_{jk:j \neq k} \right)^t \right) \middle| \Phi, \Lambda \right]$$

has the ISU decomposition

$$\begin{aligned} D_{\Phi,j}(t) &= \int_{(0,t]} \frac{1}{\kappa(s-)} p_{aj}(0, s-) V_j^*(s-)\, d(\widetilde{\Phi} - \Phi^*)(s), \\ D_{u,jk}(t) &= 0, \\ D_{s,jk}(t) &= -\int_{(0,t]} \frac{1}{\kappa(s)} p_{aj}(0, s-) R_{jk}^*(s)\, d(\Lambda_{jk} - \Lambda_{jk}^*)(s). \end{aligned}$$

In particular, the ISU decomposition does not depend on the update order or the choice of partitions.

Proof. The model framework, introduced in Section 3.2, entails that the integrability assumption in Lemma 3.9 for the SU decomposition (see proof of Theorem 3.8) is satisfied. Thus, applying Lemma 3.9 with $\mathcal{G} = \sigma(\Phi, \Lambda)$ to the ISU decomposition in Theorem 3.8

and using the martingale property of $\mathrm{d}N_{jk}(t) - I_j(t-)\mathrm{d}\Lambda_{jk}(t)$ with respect to the natural completed filtration of the random vector $(Z^t, \Phi, \Lambda)_{t\geqslant 0}$ (see (S.2)) give the desired result. $\square$

Proposition 3.11. *Let $X = (X_1, \dots, X_m)$ be a given risk basis with*

$$R(t) = \varrho((X_1 + X_2, X_3, \dots, X_m)^t)$$

for a suitable mapping ϱ, generating the ISU decomposition $D(t) = (D_1(t), \dots, D_m(t))$. Then the partially aggregated risk basis

$$\widetilde{X} = (X_1 + X_2, (X_3, X_4), X_5 \dots, X_m)$$

generates the ISU decomposition

$$\widetilde{D}(t) = (D_1(t) + D_2(t), D_3(t) + D_4(t), D_5(t) \dots, D_m(t)).$$

Proof. Since the revaluation surplus surfaces U and $\widetilde{U}$ are linked via the equation

$$\widetilde{U}(t_1, t_3, t_5 \dots, t_m) = U(t_1, t_1, t_3, t_3, t_5 \dots, t_m),$$

the SU decompositions D^n and $\widetilde{D}^n$ with respect to $\mathcal{T}_n(t)$ satisfy

$$\widetilde{D}^n(t) = (D_1^n(t) + D_2^n(t), D_3^n(t) + D_4^n(t), D_5^n(t), \dots, D_m^n(t)).$$

The latter equation carries through the limit (2.5) to the ISU decompositions. $\square$

3.4 Examples

We continue with the examples for the risk basis X and the mapping ϱ from Section 3.2 and present the corresponding ISU decompositions.

3.4.1 Decomposition of the individual revaluation surplus

Let R be the individual revaluation surplus according to (3.11).

Example 3.12. Suppose that we are in the setting of Example 3.2, where we distinguish between financial risk, unsystematic biometric risk and systematic biometric risk. By applying Theorem 3.8 and Proposition 3.11 we obtain the ISU decomposition

$$\begin{aligned}
D_\Phi(t) &= \int_{(0,t]} \frac{1}{\kappa(s-)} \sum_j I_j(s-) V_j^*(s-) \mathrm{d}(\widetilde{\Phi} - \Phi^*)(s), \\
D_u(t) &= -\sum_{jk:j\neq k} \int_{(0,t]} \frac{1}{\kappa(s)} I_j(s-) R_{jk}^*(s) \mathrm{d}(N_{jk} - \Lambda_{jk})(s), \\
D_s(t) &= -\sum_{jk:j\neq k} \int_{(0,t]} \frac{1}{\kappa(s)} I_j(s-) R_{jk}^*(s) \mathrm{d}(\Lambda_{jk} - \Lambda_{jk}^*)(s).
\end{aligned}$$

Example 3.13. Suppose that we are in the setting of Example 3.3, where we distinguish between financial risk and transition-wise biometric risks. By applying Theorem 3.8 and Proposition 3.11 we obtain the ISU decomposition

$$\begin{aligned} D_\Phi(t) &= \int_{(0,t]} \frac{1}{\kappa(s-)} \sum_j I_j(s-) V_j^*(s-) \mathrm{d}(\widetilde{\Phi} - \Phi^*)(s), \\ D_{jk}(t) &= -\int_{(0,t]} \frac{1}{\kappa(s)} I_j(s-) R_{jk}^*(s) \mathrm{d}(N_{jk} - \Lambda_{jk}^*)(s), \quad j,k \in \mathcal{Z},\, j \neq k. \end{aligned}$$

As a special case this ISU decomposition includes the heuristic approach of Ramlau-Hansen (1988, formula (4.7)) for subdividing biometric surplus in a transition-wise way.

Example 3.14. Suppose that we are in the setting of Example 3.4, where we distinguish unsystematic biometric risk and state-wise remaining risks. By applying Theorem 3.8 and Proposition 3.11 we obtain the ISU decomposition

$$\begin{aligned} D_u(t) &= -\sum_{jk:j\neq k} \int_{(0,t]} \frac{1}{\kappa(s)} I_j(s-) R_{jk}^*(s) \mathrm{d}(N_{jk} - \Lambda_{jk})(s), \\ D_j(t) &= \int_{(0,t]} \frac{1}{\kappa(s-)} I_j(s-) \Big(V_j^*(s-) \mathrm{d}(\widetilde{\Phi} - \Phi^*)(s) - \sum_{k:k\neq j} R_{jk}^*(s) \mathrm{d}(\Lambda_{jk} - \Lambda_{jk}^*)(s) \Big), \quad j \in \mathcal{Z}. \end{aligned}$$

As a special case this ISU decomposition includes heuristic approaches of Ramlau-Hansen (1988, formula before (4.10)) and Norberg (1999, formula (5.4)) for splitting off unsystematic biometric surplus and then subdividing the remaining surplus in a state-wise way.

In Example 3.12 and Example 3.14 we split off the surplus contribution of the unsystematic biometric risk. Since this unsystematic biometric risk is diversifiable in the insurance portfolio, its contribution $\kappa(t)D_u(t)$ to the total surplus $S(t)$, see (3.5), is typically credited or debited to the insurer. Møller and Steffensen (2007, Chapter 6.3) denote the remaining surplus $S(t) - \kappa(t)D_u(t)$ as the 'systematic surplus'. This systematic surplus mainly belongs to the policyholder.

Asmussen and Steffensen (2020, Chapter VI.4) split also the financial risk into an unsystematic part and a systematic part and argue that the unsystematic financial risk surplus contribution should be fully credited or debited to the insurer, similarly to the unsystematic biometric risk surplus contribution. They distinguish unsystematic and systematic financial risk by splitting Φ into a martingale part and a remaining systematic part. If we likewise split $\Phi - \Phi^*$ in the risk basis X into a martingale part and a remaining systematic part, then the resulting ISU decomposition allows us to distinguish between systematic and unsystematic surplus contributions. If we collect the systematic biometrical and systematic financial surplus contributions, then we just end up with the systematic surplus formula of Asmussen and Steffensen (2020, Chapter VI.4). We do not show the detailed calculations here but leave them to the reader.

3.4.2 Decomposition of the mean portfolio revaluation surplus

Let R be the mean portfolio revaluation surplus according to (3.14).

Example 3.15. We choose the setting from Example 3.2 but adopt the mean portfolio perspective. By applying Theorem 3.10 and Proposition 3.11 we obtain the ISU decomposition

$$\begin{aligned}
D_\Phi(t) &= \int_{(0,t]} \frac{1}{\kappa(s-)} \sum_j p_{aj}(0,s-) V_j^*(s-) \mathrm{d}(\widetilde{\Phi} - \Phi^*)(s), \\
D_u(t) &= 0, \\
D_s(t) &= -\sum_{jk:j\neq k} \int_{(0,t]} \frac{1}{\kappa(s)} p_{aj}(0,s-) R_{jk}^*(s) \mathrm{d}(\Lambda_{jk} - \Lambda_{jk}^*)(s).
\end{aligned}$$

The conditional expectation in (3.14) and (3.15) completely eliminates the unsystematic biometric risk, which explains why we have $D_u(t) = 0$ here.

Example 3.16. We choose the setting from Example 3.3 but adopt the mean portfolio perspective. By applying Theorem 3.10 and Proposition 3.11 we obtain the ISU decomposition

$$\begin{aligned}
D_\Phi(t) &= \int_{(0,t]} \frac{1}{\kappa(s-)} \sum_j p_{aj}(0,s-) V_j^*(s-) \mathrm{d}(\widetilde{\Phi} - \Phi^*)(s), \\
D_{jk}(t) &= -\int_{(0,t]} \frac{1}{\kappa(s)} p_{aj}(0,s-) R_{jk}^*(s) \mathrm{d}(\Lambda_{jk} - \Lambda_{jk}^*)(s), \quad j,k \in \mathcal{Z},\ j \neq k.
\end{aligned}$$

The next example shows an application of this formula.

Example 3.17. We continue with the previous example but focus on the specific setting of Example 3.6. One can show that the SU decomposition of $R(k+1) - R(k)$ with respect to an integer partition equals

$$\begin{aligned}
U(k+1,k,k) - U(k,k,k) &= e^{-\int_0^{k+1} \phi(u)\,\mathrm{d}u}\ {}_kp_x\, V_a^*(k) \Delta(i), \\
U(k+1,k+1,k) - U(k+1,k,k) &= e^{-\int_0^{k+1} \phi(u)\,\mathrm{d}u}\ {}_kp_x \big(V_a^*(k+1-) - d_{k+1}\big) \Delta(q), \\
U(k+1,k+1,k+1) - U(k+1,k+1,k) &= e^{-\int_0^{k+1} \phi(u)\,\mathrm{d}u}\ {}_kp_x \big(V_a^*(k+1-) - s_{k+1}\big) \Delta(r),
\end{aligned} \tag{3.17}$$

where $\Delta(i) = i_k - i_k^*$, $\Delta(q) = q_{x+k} - q_{x+k}^*$ and $\Delta(r) = r_{x+k} - r_{x+k}^*$, see Section A.1 in the appendix. This decomposition is the standard surplus decomposition formula used in German life insurance (see Milbrodt & Helbig, 1999, Section 11.B). We can interpret the latter SU decomposition as an approximation of the ISU decomposition of $R(k+1) - R(k)$, which equals here

$$\begin{aligned}
D_\Phi(k+1) - D_\Phi(k) &= \int_{(k,k+1]} e^{-\int_0^s \phi(u)\,\mathrm{d}u} \; {}_s p_x \, V_a^*(s)\,\mathrm{d}(\Phi - \Phi^*)(s),\\
D_{ad}(k+1) - D_{ad}(k) &= \int_{(k,k+1]} e^{-\int_0^s \phi(u)\,\mathrm{d}u} \; {}_s p_x \big(V_a^*(s) - b_{ad}(s)\big)\,\mathrm{d}(\Lambda_{ad} - \Lambda_{ad}^*)(s),\\
D_{as}(k+1) - D_{as}(k) &= \int_{(k,k+1]} e^{-\int_0^s \phi(u)\,\mathrm{d}u} \; {}_s p_x \big(V_a^*(s) - b_{as}(s)\big)\,\mathrm{d}(\Lambda_{as} - \Lambda_{as}^*)(s).
\end{aligned} \tag{3.18}$$

The latter decomposition is invariant with respect to a reordering of the components of X, whereas the SU decomposition changes. Therefore, we recommend to replace the traditional SU decomposition (3.17) by the ISU decomposition (3.18).

Example 3.18. We choose the setting from Example 3.4 but adopt the mean portfolio perspective. By applying Theorem 3.10 and Proposition 3.11 we obtain the ISU decomposition

$$\begin{aligned}
D_u(t) &= 0,\\
D_j(t) &= \int_{(0,t]} \frac{1}{\kappa(s-)} p_{aj}(0,s-)\Big(V_j^*(s-)\mathrm{d}(\widetilde{\Phi}-\Phi^*)(s) - \sum_{k:k\neq j} R_{jk}^*(s)\mathrm{d}(\Lambda_{jk}-\Lambda_{jk}^*)(s)\Big), \; j \in \mathcal{Z}.
\end{aligned}$$

As a special case this ISU decomposition includes heuristic approaches of Ramlau-Hansen (1991, formula (3.2)) and Norberg (1999, formula (5.7)) for subdividing mean portfolio surplus in a state-wise manner.

As shown in Section 3.3, the ISU decompositions in our life insurance model do not depend on the update order. Thus, together with Theorem 2.7, we directly get the following result.

Corollary 3.19. *For all examples in Section 3.4 the IOAT decomposition and the averaged ISU decomposition are both equal to the ISU decomposition.*

The results in this chapter have shown that the ISU decomposition principle is a suitable tool for deriving invididual surplus contributions in traditional life insurance. Moreover, the clarity of the decomposition opens up new prospects for future research to take policyholder behaviour into account.

4 Relating the ISU concept to the martingale representation theorem

A company's surplus is often exposed to a variety of risks, such as financial, legal and economic risks. To master these risks, companies establish risk management processes that identify, quantify and control potential risks. In particular, risk quantification entails the assessment of how an individual risk affects the overall risk. However, individual risk contributions are usually not directly accessible (see Schilling et al., 2020). Especially if the different risks contribute to the company's surplus in a non-linear way, the decomposition of risk into its individual risk contributions is a very challenging task (see Frei, 2020). Despite its relevance, risk decompositions have only been addressed by a few authors in the literature (see Karabey et al., 2014 and references therein). With a focus on an insurer's surplus, an explaination is provided about how the ISU decomposition principle can also serve as a useful tool for decomposing risk. But first of all the term *risk* needs to be clarified.

According to the ISO 31000:2018 standard on risk management published by the International Organisation for Standardisation (ISO, 2018), risk is an 'effect of uncertainty on objectives'. The *uncertainty* is driven by a lack of knowledge about the future development of factors affecting the surplus (*objective*). An *effect* manifests itself as a 'deviation from the expected', which can be positive or negative (see Wuorikoski, 2018). Translating this definition into mathematical terms requires the taking into account of both, the different levels of information and the understanding of risk as a deviation from expected values. To reflect the different levels of information, a filtration $\mathbb{G} = (\mathcal{G}_t)_{t\geqslant 0}$ describes the information available over time. The expected surplus changes as information increases, therefore the revaluation surplus process (see Chapter 2) is defined by

$$R(t) = \mathbb{E}[\xi|\mathcal{G}_t], \ t \geqslant 0, \tag{4.1}$$

where ξ is a $\mathcal{G}_T$-measurable random variable that e.g. represents an insurer's surplus on a finite time horizon $[0, T]$. This time-dynamic approach allows one to interpret $R(t) - R(0)$, $t \geqslant 0$, as the risk on $[0, t]$ that the expectation for the total surplus needs to be adjusted. In particular, taking the entire time interval $[0, T]$ into account,

$$R(T) - R(0) = \xi - \mathbb{E}[\xi|\mathcal{G}_0] \tag{4.2}$$

describes the deviation of the actual surplus ξ from the initially expected surplus $\mathbb{E}[\xi|\mathcal{G}_0]$, and thus corresponds to the definition above.

Having derived a common understanding of risk, it is possible to return to the original question of how to decompose the total risk $R(t) - R(0)$ into individual risk contributions.

A systematic approach to this question has been carried out by Schilling et al. (2020), who introduce axioms of a meaningful risk decomposition for the time-static risk (4.2). On the basis of these axioms, Schilling et al. (2020) not only compare risk decompositions known from the literature, but also suggest another risk decomposition fulfilling all axioms. Based on the martingale representation theorem, Schilling et al. (2020) present the so-called MRT decomposition that decomposes the total risk into martingales, each of which characterises an individual risk contribution. In particular, this approach uses the martingale property of the revaluation surplus process (4.1), which is also conducive to the application of the ISU decomposition principle. More precisely, if $\mathcal{G}_t$ describes the information provided by a risk basis $X = (X_1, \cdots, X_m)$ until time t, the ISU decomposition principle is directly applicable and a natural link mapping ϱ between the risk basis and the revaluation surplus (4.1) satisfying $R(t) = \varrho(X^t)$ is given by

$$\varrho(X_1^{t_1}, \dots, X_m^{t_m}) = \mathbb{E}[\xi | \sigma(X_1^{t_1}, \dots, X_m^{t_m})], \ t_i \in [0, T], \ i = 1, \dots, m. \tag{4.3}$$

In this chapter, it is shown that the ISU decomposition principle is indeed a useful tool for the decomposition of risk. In particular, conditions are presented under which the ISU and MRT decompositions are equivalent.

So far, the motivation to consider ISU decompositions of martingales has been focused on risk management. Though risk assessment usually refers to a real-world measure (see Karabey et al., 2014; Schilling et al., 2020), other choices of the probability measure in (4.1) are conceivable. This raises the prospect of further applications. Under a risk-neutral probability measure, martingales are closely related to the pricing theory for finance products in arbitrage-free markets (see e.g. Harrison & Pliska, 1981 for option pricing). Not only financial mathematics, but also recent actuarial mathematics strive for a market-consistent valuation of insurance products (see e.g. Biagini, 2013 and references therein). The relevance of surplus decompositions in modern valuation setups is illustrated by Fischer (2004), who splits the gains associated with a life insurance contract into a biometric and a financial part.

Another application of martingale surplus processes relates to the decomposition of life insurance bonus in the context of Chapter 3. In the literature on traditional life insurance surplus (see Chapter 3), different valuation bases are used to value the earned surplus under best-estimate assumptions (‘second-order basis’) on the one hand, and to value the expected future surplus under conservative assumptions (‘first-order basis’) on the other hand. This valuation pattern can also be achieved by considering the revaluation surplus process (4.1) with respect to a conservative probability measure. Whereas in classical bonus theory, the conservative valuation basis is fixed at the beginning of the contract, the martingale approach takes full advantage of the available information. More precisely, the empirical observations are not only used to evaluate the earned surplus, but also to involve an adjustment to the expected future surplus. Though understanding surplus as a martingale is in line with a modern market-consistent valuation of insurance contracts, only a few authors take up this idea in bonus theory, references are Steffensen (2001, Section 3.5.1) and Dufresne (2001). Nevertheless, this application provides a further motivation to study the ISU decomposition of martingales in more detail.

After introducing the underlying model framework, Section 4.1 presents a property of risk bases that will be crucial from a technical perspective. In Section 4.2, it is shown that this property is satisfied by a number of stochastic processes commonly used in actuarial and financial modelling. The chapter concludes with Section 4.3, which contains the derivation of the ISU decomposition for martingales and a discussion of its relationship to the MRT decomposition.

4.1 Model framework

We generally assume that we have a complete probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with a right-continuous and complete filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geqslant 0}$. Let the system $\mathcal{N}$ contain the subsets of $\mathbb{P}$-null sets, i.e. $\mathcal{N} = \{N \subseteq \Omega | \exists A \in \mathcal{A} : N \subseteq A \wedge \mathbb{P}(A) = 0\}$. Moreover, for sub-$\sigma$-algebras $\mathcal{A}_i \subseteq \mathcal{A}$, $i = 1, \ldots, n$, we define the operator $\vee$ by

$$\bigvee_{i=1}^{n} \mathcal{A}_i := \mathcal{A}_1 \vee \ldots \vee \mathcal{A}_n := \sigma\left(\bigcup_{i=1}^{n} \mathcal{A}_i\right),$$

where the right-hand side denotes the smallest σ-algebra that contains all $\mathcal{A}_i$, $i = 1, \ldots, n$. We consider a finite time horizon $[0, T]$ and suppose that the risk basis $X = (X_1, \ldots, X_m)$ is given by a vector of $\mathbb{F}$-semimartingales. In the following, we write $\mathbb{G}^i = (\mathcal{G}_t^i)_{t \in [0,T]}$ for the completed natural filtration of X_i, $i = 1, \ldots, m$, and $\mathbb{G} = (\mathcal{G}_t)_{t \in [0,T]}$ with $\mathcal{G}_t = \mathcal{G}_t^1 \vee \ldots \vee \mathcal{G}_t^m$ for the joint filtration. For a subset $J = \{j_1, \ldots, j_r\} \subseteq \{1, \ldots, m\}$, we further introduce

- the subfiltration $\mathbb{G}^J = (\mathcal{G}_t^J)_{t \in [0,T]}$ with $\mathcal{G}_t^J = \mathcal{G}_t^{j_1} \vee \ldots \vee \mathcal{G}_t^{j_r}$, whereas $\mathcal{G}_t^{\varnothing} = \{\varnothing, \Omega\}$,
- the family of sub-σ-algebras $\mathcal{G}_{s,t}^J = \mathcal{G}_s \vee \mathcal{G}_t^J$ for $s \leqslant t$.

Moreover, we make the following assumption about the risk basis:

(S) Each X_i has a decomposition

$$X_i(t) = X_i(0) + M_i(t) + A_i(t), \tag{4.4}$$

where $X_i(0) \in \mathbb{R}$, M_i is a square integrable $\mathbb{G}$-martingale and A_i is a $\mathbb{G}$-predictable finite variation process with $M_i(0) = A_i(0) = 0$.

The decomposition in (4.4) is unique (see Protter, 2005, Chapter III, Theorem 34). In general, (S) holds true for a large class of semimartingales, see also the examples provided in the next section. In other words, (S) assumes that X_i is a so-called *special semimartingale* with the addition of stronger integrability conditions for the martingale part M_i. The class of special semimartingales covers all semimartingales with bounded jumps. Further characterisations of special semimartingales can be found in Protter (2005, Chapter III.7).

In this chapter, we focus on surplus revaluation processes given as martingales with respect to the information generated by X. More precisely, let

$$R(t) = \mathbb{E}[\xi|\mathcal{G}_t],\ t \geqslant 0, \tag{4.5}$$

where ξ is a square-integrable $\mathcal{G}_T$-measurable random variable. We interpret ξ as an discounted insurance claim and $R(t)$ as the $\mathcal{G}_t$-measurable proxy of ξ (see Section 2.1). It is worth noting that, unlike in Chapter 3, we do not assume any particular form of insurance claim.

Recall that the choice of link mapping ϱ satisfying $R(t) = \varrho(X^t)$, $t \geqslant 0$ is essential for the application of the ISU decomposition principle. As $\mathcal{G}_t$ describes the information provided by the risk basis X, a natural link mapping ϱ is given by

$$\varrho(X_1^{t_1}, \ldots, X_m^{t_m}) = \mathbb{E}[\xi|\sigma(X_1^{t_1}, \ldots, X_m^{t_m}) \vee \mathcal{N}],\ t_i \in [0,T],\ i = 1, \ldots, m. \tag{4.6}$$

Using this link mapping ϱ, we are able to apply the ISU decomposition principle (see Section 2). Let $(\mathcal{T}_n(t))_n$ be a sequence of partitions on $[0,t]$ with vanishing step lengths. For $R(t) = \mathbb{E}[\xi|\mathcal{G}_t]$, the ISU decomposition principle gives the additive decomposition

$$R(t) = D_1(t) + \ldots + D_m(t),$$

where

$$D_i(t) = \operatorname*{plim}_{n\to\infty} \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} (\mathbb{E}[\xi|\mathcal{G}^{\{1,\ldots,i\}}_{t_k,t_{k+1}}] - \mathbb{E}[\xi|\mathcal{G}^{\{1,\ldots,i-1\}}_{t_k,t_{k+1}}]),\ i = 1, \ldots, m.$$

To further analyse the ISU decomposition, we introduce the following property of a risk basis, which will be crucial for our main results.

Definition 4.1. Suppose the risk basis X satisfies (S). We say the risk basis X fulfils (M) if for every sequence of partitions $(\mathcal{T}_n(t))_n$ on $[0,t]$ with vanishing step lengths and every subset $J \subseteq \{1, \ldots, m\}$, it holds

$$\operatorname*{plim}_{n\to\infty} \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}] = 0, \quad \text{if } i \notin J, \tag{4.7}$$

$$\operatorname*{plim}_{n\to\infty} \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}] = M_i(t), \quad \text{if } i \in J, \tag{4.8}$$

for all $t \in [0,T]$.

The convergence (4.7) depicts an infinitesimally forward shifted martingale property of M_i, while the convergence (4.8) describes an infinitesimally backward shifted measurability of M_i $(i = 1, \ldots, m)$ (cf. Christiansen, 2021).

Remark 4.2. The convergence in (M) is stated for intervals $[0, t]$, $t \in [0, T]$. However, this already implies convergences (4.7) and (4.8) on intervals $[s, t]$, $s \leqslant t$: Let $(\mathcal{T}_n)_n$ be a sequence of partitions on $[s, t]$ with vanishing step lengths, let $J \subseteq \{1, \ldots, m\}$ and let property (M) hold. Suppose $(\widetilde{\mathcal{T}}_n)_n$ defines a sequence of partitions on $[0, s]$ with vanishing step lengths. Then $(\widetilde{\widetilde{\mathcal{T}}}_n)_n$ with $\widetilde{\widetilde{\mathcal{T}}}_n = \widetilde{\mathcal{T}}_n \cup \mathcal{T}_n$ is a sequence of partitions on $[0, t]$ with vanishing step lengths. Therefore, property (M) implies

$$
\begin{aligned}
&\operatorname*{plim}_{n\to\infty} \sum_{t_k, t_{k+1} \in \mathcal{T}_n} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] \\
&= \operatorname*{plim}_{n\to\infty} \sum_{t_k, t_{k+1} \in \widetilde{\widetilde{\mathcal{T}}}_n} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] - \operatorname*{plim}_{n\to\infty} \sum_{t_k, t_{k+1} \in \widetilde{\mathcal{T}}_n} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] \\
&= \begin{cases} 0, & i \notin J, \\ M_i(t) - M_i(s), & i \in J. \end{cases}
\end{aligned}
$$

In the following section, we provide evidence for the plausibility of property (M). Therefore, we show that property (M) is satisfied for several examples commonly used in actuarial and financial modelling.

4.2 Examples

This section contains of five subsections, each of which investigates a class of risk bases with regard to property (M). The notations introduced in Section 4.1 always refer to the respective risk basis under consideration. In addition, the following notation is needed. For a stochastic process Z, we denote by $\mathbb{F}^Z = (\mathcal{F}^Z_t)_{t \in [0,T]}$ its natural completed filtration. Furthermore, let $L_2(\mathbb{P})$ denote the usual L_2 space of square-integrable random variables with norm $\| \cdot \|_2$. Moreover, for sub-σ-algebras $\mathcal{A}_i \subseteq \mathcal{A}$, $i = 1, 2, 3$, we say $\mathcal{A}_1$ and $\mathcal{A}_2$ are conditionally independent given $\mathcal{A}_3$, if $\mathbb{P}(A_1 \cap A_2 | \mathcal{A}_3) = \mathbb{P}(A_1 | \mathcal{A}_3) \mathbb{P}(A_2 | \mathcal{A}_3)$ for all $A_i \in \mathcal{A}_i$, in symbols $\mathcal{A}_1 \perp\!\!\!\perp \mathcal{A}_2 | \mathcal{A}_3$. Furthermore, for a set $J \subseteq \{1, \ldots, m\}$, we write J^{c} for its complement in $\{1, \ldots, m\}$.

4.2.1 Independent sources of risk

We start with a simple setup of independent risk factors.

Proposition 4.3. *Suppose that the risk basis $X = (X_1, \ldots, X_m)$ consists of independent $\mathbb{F}$-semimartingales that satisfy (S). If $\mathbb{F}^{M_i} \subseteq \mathbb{G}^i$ for all $i = 1, \ldots, m$, then $X = (X_1, \ldots, X_m)$ fulfils (M).*

Proof. Let $\mathcal{T}_n(t)$ be a sequence of partitions on $[0,t]$ with vanishing step lengths, and let $J \subseteq \{1,\ldots,m\}$. If $i \notin J$, the σ-algebras $\mathcal{G}^J_{t_{k+1}}$ and $\sigma(\sigma(M_i(t_{k+1}) - M_i(t_k)), \mathcal{G}^{J^c}_{t_k})$ are independent. Thus, we can omit the information from X_j, $j \in J$ on $(t_k, t_{k+1}]$ (see Zitkovic, 2015, Proposition 10.5.9 (9)) and get

$$\sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}] = \sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k)|\mathcal{G}_{t_k}] = 0.$$

If $i \in J$, the assumption $\mathbb{F}^{M_i} \subseteq \mathbb{G}^i$ immediately implies

$$\sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}] = \sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} (M_i(t_{k+1}) - M_i(t_k)) = M_i(t).$$

Thus, (M) is fulfilled. $\square$

This leads us to the following examples.

Example 4.4. Let the risk basis $X = (X_1,\ldots,X_m)$ consist of independent $\mathbb{F}$-martingales. Then, by Proposition 4.3, (M) is fulfilled.

Example 4.5. Let $W = (W_1,\ldots,W_m)$ be a standard m-dimensional Brownian motion, and let the risk basis be given by $X = (X_1,\ldots,X_m)$, where X_i satisfies the stochastic differential equation

$$\mathrm{d}X_i(t) = \mu_i(t, X_i(t))\mathrm{d}t + \sigma_i(t, X_i(t))\mathrm{d}W_i(t),\ X_i(0) = x_i(0) \in \mathbb{R},$$

for continuous functions $\mu_i, \sigma_i\colon [0,T] \times \mathbb{R} \to \mathbb{R}$ $(i = 1,\ldots,m)$. The definition of X implies that X_i, $i = 1,\ldots,m$ are independent. Furthermore, it holds

$$\mathrm{d}M_i(t) = \sigma_i(t, X_i(t))\mathrm{d}W_i(t) = \mathrm{d}X_i(t) - \mu_i(t, X_i(t))\mathrm{d}t,$$

which implies $\mathbb{F}^{M_i} \subseteq \mathbb{G}^i$. Thus, by Proposition 4.3, (M) is fulfilled.

4.2.2 Grid-dependent sources of risk

Clearly, independence of risk factors is a strong assumption, which in reality is usually not satisfied for economic and demographic risk factors. Fortunately, the property (M) can still be verified, if we allow for dependency on a fixed grid. Let the corresponding grid π be given as $\pi = \{0 = u_0 < \ldots < u_d = T\}$, $d \in \mathbb{N}$.

Definition 4.6. Let Z be a stochastic process and let $\mathbb{H} = (\mathcal{H}_t)_{t\geqslant 0}$ be a subfiltration of $\mathbb{F}$. Then Z is called $\mathbb{H}$-π-*adapted* if $Z(t)$ is $\mathcal{H}_{u_l}$-measurable for $t \in (u_l, u_{l+1}]$.

We suppose that the risk basis X consists of $\mathbb{F}$-semimartingales X_i, $i = 1,\ldots,m$, which satisfy (S). Therefore, each source of risk X_i admits a canonical decomposition

$$X_i(t) = X_i(0) + M_i(t) + A_i(t),\ t \in [0,T]. \tag{4.9}$$

In this subsection, we focus on martingale parts having an integral representation. More precisely, we consider martingale parts M_i, $i = 1, \ldots, m$, satisfying

a) $M_i(t) = \int_0^t \varphi_i(s) \mathrm{d}\Gamma_i(s)$,

where

b) $\Gamma = (\Gamma_1, \ldots, \Gamma_m)$ consists of independent $\mathbb{F}$-martingales Γ_i with $\Gamma_i(0) = 0$, $i = 1, \ldots, m$,

c) φ_i is a càglàd $\mathbb{F}^\Gamma$-π-adapted process,

d) $\varphi_i \neq 0$ almost surely

For the compensator A_i, $i = 1, \ldots, m$, we assume

e) A_i is $\mathbb{F}^\Gamma$-π-adapted.

The measurability assumptions c) and e) allow A_i and the integrand of M_i (and thus X_i) to depend on the common past information (up to the grid point) of Γ. However, between the grid points, newly generated information for M_i and X_i stems only from the newly generated information of Γ_i. To be precise, let $t_k, t_{k+1} \in \mathcal{T}_n(t)$ such that $(t_k, t_{k+1}) \cap \pi = \varnothing$. For $s \in (t_k, t_{k+1}]$, we have

$$\sigma(M_i(u) - M_i(t_k) : t_k < u \leqslant s) \subseteq \mathcal{F}^\Gamma_{t_k} \vee \sigma(\Gamma_i(u) - \Gamma_i(t_k) : t_k < u \leqslant s). \tag{4.10}$$

With the canonical decomposition (4.9) and assumption e), this further implies

$$\sigma(X_i(u) - X_i(t_k) : t_k < u \leqslant s) \subseteq \mathcal{F}^\Gamma_{t_k} \vee \sigma(\Gamma_i(u) - \Gamma_i(t_k) : t_k < u \leqslant s). \tag{4.11}$$

The latter observation will help us to prove the following lemma.

Lemma 4.7. *Let the risk basis X consist of $\mathbb{F}$-semimartingales that fulfil (S), where the canonical decompositions (A_i, M_i), $i = 1, \ldots, m$, satisfy a) − e). Let $J \subseteq \{1, \ldots, m\}$ and $t_k < s$, such that $(t_k, s) \cap \pi = \varnothing$. Then it holds*

$$\mathcal{G}^J_{t_k,s} = \mathcal{F}^\Gamma_{t_k} \vee \sigma(\Gamma_j(u) - \Gamma_j(t_k) : t_k < u \leqslant s, j \in J). \tag{4.12}$$

Furthermore, we have $\mathbb{G} = \mathbb{F}^\Gamma$.

Proof. We prove the result in the reverse order to that stated, i.e. we first show that

$$\mathbb{G} = \mathbb{F}^\Gamma. \tag{4.13}$$

Due to our assumptions, we have $\mathcal{G}_0 = \mathcal{F}^\Gamma_0$. Furthermore, assumptions a), c) and e) directly imply

$$\mathcal{G}_s \subseteq \mathcal{F}^\Gamma_s$$

for $s \in [0,T]$. Vice versa, we use an induction argument. Let $s \in (0, u_1]$. Then the associativity for stochastic integrals (Protter, 2005, Chapter II, Theorem 19) and assumption e) yield

$$\mathrm{d}\Gamma_i(s) = \varphi_i^{-1}(s)\mathrm{d}X_i(s). \tag{4.14}$$

As $\varphi_i^{-1}(s)$ is $\mathcal{F}_0^\Gamma$-measurable (see d)) on $(0, u_1]$, we get $\mathcal{F}_s^{\Gamma_i} \subseteq \mathcal{G}_s^i$ for $s \in (0, u_1]$. Since the argument does not depend on i, we also have $\mathcal{F}_s^\Gamma \subseteq \mathcal{G}_s$ for $s \in (0, u_1]$. Let us now assume that $\mathcal{F}_s^\Gamma \subseteq \mathcal{G}_s$ holds for $s \in [0, u_l]$. For $s \in (u_l, u_{l+1}]$, we have that

$$\mathcal{F}_s^\Gamma = \mathcal{F}_{u_l}^\Gamma \vee \sigma(\Gamma_i(u) - \Gamma_i(u_l) : i = 1, \ldots, m, u_l < u \leqslant s).$$

Since $\mathcal{F}_{u_l}^\Gamma \subseteq \mathcal{G}_{u_l}$ and $\varphi_i(s)$ is $\mathcal{F}_{u_l}^\Gamma$-measurable, thus $\mathcal{G}_{u_l}$-measurable, the representation (4.14) implies $\mathcal{F}_s^\Gamma \subseteq \mathcal{G}_s$ for all $s \in (u_l, u_{l+1}]$. Hence, we have shown that $\mathbb{G} = \mathbb{F}^\Gamma$.

Together with (4.11), we get the inclusion

$$\mathcal{G}_{t_k,s}^J \subseteq \mathcal{F}_{t_k}^\Gamma \vee \sigma(\Gamma_j(u) - \Gamma_j(t_k) : t_k < u \leqslant s, j \in J).$$

For the other direction, the equation (4.13) together with the representation (4.14) gives the desired result. □

We are now in the position to prove property (M).

Proposition 4.8. *Let the risk basis $X = (X_1, \ldots, X_m)$ consist of $\mathbb{F}$-semimartingales that fulfil (S), where the canonical decompositions (A_i, M_i), $i = 1, \ldots, m$, satisfy a) − e). Furthermore, let $\langle M_i, M_i \rangle$, $i = 1, \ldots, m$ be continuous processes. Then X fulfils (M).*

Proof. Let $(\mathcal{T}_n(t))_n$ be a sequence of partitions on $[0,t]$ with vanishing step lengths and let $J \subseteq \{1, \ldots, m\}$. We have to prove that

$$\operatorname*{plim}_{n\to\infty} \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}\left[\int_{t_k}^{t_{k+1}} \varphi_i(s)\mathrm{d}\Gamma_i(s) \middle| \mathcal{G}_{t_k,t_{k+1}}^J\right] = 0, \;\; i \notin J, \tag{4.15}$$

and

$$\operatorname*{plim}_{n\to\infty} \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}\left[\int_{t_k}^{t_{k+1}} \varphi_i(s)\mathrm{d}\Gamma_i(s) \middle| \mathcal{G}_{t_k,t_{k+1}}^J\right] = \int_0^t \varphi_i(s)\mathrm{d}\Gamma_i(s), \;\; i \in J. \tag{4.16}$$

Firstly, for $t_k, t_{k+1} \in \mathcal{T}_n$ with $(t_k, t_{k+1}) \cap \pi = \varnothing$ and $i \notin J$, we verify

$$\sigma(M_i(t_{k+1}) - M_i(t_k)) \perp\!\!\!\perp \sigma(X_j(s) : t_k < s \leqslant t_{k+1}, j \in J) \,|\, \mathcal{G}_{t_k}. \tag{4.17}$$

Due to (4.10) and (4.11), it suffices to show

$$\mathcal{F}_{t_k}^\Gamma \vee \mathcal{H}^{\{i\}} \perp\!\!\!\perp \mathcal{F}_{t_k}^\Gamma \vee \mathcal{H}^J \,|\, \mathcal{G}_{t_k},$$

where $\mathcal{H}^{\widetilde{J}} = \sigma(\Gamma_i(s) - \Gamma_i(t_k) : t_k < s \leqslant t_{k+1}, j \in \widetilde{J})$ for any $\widetilde{J} \subseteq \{1, \ldots, m\}$. This, however, follows immediately from $\mathcal{G}_{t_k} = \mathcal{F}^{\Gamma}_{t_k}$ (see Lemma 4.7) and the independence between Γ_i and Γ_j, $j \in J$.

Without loss of generality we assume that $|\mathcal{T}_n(t)| < |\pi|$ for all n. If $u_l < t$ and $u_l \in (t_j, t_{j+1}]$ for an index j, we set $\underline{t}_l = t_j$ and $\bar{t}_l = t_{j+1}$ as the neighbouring points of u_l in $\mathcal{T}_n(t)$. If $u_l \geqslant t$, we set $\underline{t}_l = \bar{t}_l = t$. Moreover, we write $\mathcal{T}_n^0(t) = \mathcal{T}_n(t) \cap [u_0, u_1]$ and $\mathcal{T}_n^l(t) = \mathcal{T}_n(t) \cap (u_l, u_{l+1}]$, $l = 1, \ldots, d-1$.

For (4.15), i.e. $i \notin J$, the conditional independence (4.17) yields

$$
\begin{aligned}
&\sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] \\
&= \sum_{l=0}^{d-1} \sum_{t_k, t_{k+1} \in \mathcal{T}_n^l(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] + \sum_{l=1}^{d-1} \mathbb{E}[M_i(\bar{t}_l) - M_i(\underline{t}_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}] \\
&= \sum_{l=0}^{d-1} \sum_{t_k, t_{k+1} \in \mathcal{T}_n^l(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}_{t_k}] + \sum_{l=1}^{d-1} \mathbb{E}[M_i(\bar{t}_l) - M_i(\underline{t}_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}] \\
&= \sum_{l=1}^{d-1} \mathbb{E}[M_i(\bar{t}_l) - M_i(\underline{t}_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}].
\end{aligned}
$$

Let us recall that $\| \cdot \|_2$ denotes the norm on $L_2(\mathbb{P})$. With the Jensen's inequality for conditional expectations, the Itô isometry and the dominated convergence theorem, we observe

$$
\begin{aligned}
&\lim_{n \to \infty} \| \sum_{l=1}^{d-1} \mathbb{E}[M_i(\bar{t}_l) - M_i(\underline{t}_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}] \|_2^2 = \lim_{n \to \infty} \sum_{l=1}^{d-1} \| \mathbb{E}[M_i(\bar{t}_l) - M_i(\underline{t}_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}] \|_2^2 \\
&\leqslant \lim_{n \to \infty} \sum_{l=1}^{d-1} \mathbb{E}[(M_i(\bar{t}_l) - M_i(\underline{t}_l))^2] = \lim_{n \to \infty} \sum_{l=1}^{d-1} \mathbb{E}\left[\langle M_i, M_i \rangle(\bar{t}_l) - \langle M_i, M_i \rangle(\underline{t}_l)\right] = 0,
\end{aligned}
$$

where we used that $\langle M_i, M_i \rangle$ is assumed to be continuous. As L_2-convergence implies convergence in probability, we have shown (4.15).

Next, we prove (4.16), i.e. let $i \in J$. By (4.10) and Lemma 4.7, we get

$$
\begin{aligned}
&\sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] \\
&= M_i(t) + \sum_{l=1}^{d-1} \mathbb{E}[M_i(\bar{t}_l) - M_i(u_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}] - \sum_{l=1}^{d-1} (M_i(\bar{t}_l) - M_i(u_l)).
\end{aligned}
$$

With similar arguments as above, the last two terms tend again to 0 in L_2 (and therefore also in probability). □

This leads us to the following examples. We start with a risk basis consisting of Itô processes. With more general assumptions on measurability, this setup is used by Schilling et al. (2020) to model (systematic) financial and biometric risks.

Example 4.9. Let $W = (W_1, \ldots, W_m)$ be a standard m-dimensional Brownian motion, and let the risk basis be given by $X = (X_1, \ldots, X_m)$, where X_i satisfies the stochastic differential equation

$$\mathrm{d}X_i(t) = \mu_i(t)\mathrm{d}t + \sigma_i(t)\mathrm{d}W_i(t),\ X_i(0) = x_i(0) \in \mathbb{R},$$

with μ_i and σ_i being $\mathbb{F}^W$-π-adapted processes, such that $\sigma_i \neq 0$ almost surely for every i and

$$\mathbb{E}\bigg[\bigg(\int_0^t \sigma_i(s)\mathrm{d}W_i(s)\bigg)^2\bigg] < \infty,\ t \in [0, T].$$

Then, by Proposition 4.8, (M) is fulfilled.

Example 4.10. Let $W = (W_1, \ldots, W_m)$ be a standard m-dimensional Brownian motion, and let the risk basis be given by $X = (X_1, \ldots, X_m)$, where X_i is an $\mathbb{F}^W$-adapted process satisfying the stochastic differential equation

$$\mathrm{d}X_i(t) = \mu_i(t, X|_{[0,t)})\mathrm{d}t + \sigma_i(t, X|_{[0,t)})\mathrm{d}W_i(t),\ X_i(0) = x_i(0) \in \mathbb{R}$$

with

$$\mu_i(t, X|_{[0,t)}) = \sum_{l=1}^{d} \mu_i(u_l, X(u_l))\mathbb{1}_{(u_l,u_{l+1}]}(t), \quad \sigma_i(t, X|_{[0,t)}) = \sum_{l=1}^{d} \sigma_i(u_l, X(u_l))\mathbb{1}_{(u_l,u_{l+1}]}(t)$$

for a partition $\{0 = u_0 < \ldots < u_d = T\}$ of $[0, T]$ and functions $\mu_i, \sigma_i \colon \mathbb{R}^2 \to \mathbb{R}$, $\sigma_i \neq 0$, $i = 1, \ldots, m$. Then, by Proposition 4.8, (M) is fulfilled.

By applying similar techniques as in this section, we can extend the range of example risk bases that fulfil (M) to risk factors, that rely on Poisson random measures.

4.2.3 Sources of risk driven by Poisson random measures

In this paragraph we consider risk bases driven by Poisson random measures, which are widely used in actuarial applications (e.g. in claims modelling). Furthermore, Poisson random measures are crucial in the Lévy-Itô decomposition, describing the jumps of a Lévy process. An introduction to Poisson random measures and stochastic integrals with respect to Poisson random measures can be found e.g. in Ikeda and Watanabe (1989) or Jacod and Shiryaev (2003, Chapter II.1). The former reference also serves as the main reference for this paragraph. The stochastic integrals that appear are interpreted as pathwise Lebesgue-Stieltjes integrals.

Let $\mu = (\mu_1, \ldots, \mu_m)$ consist of independent Poisson random measures μ_i, $i = 1, \ldots, m$, on $E = \mathbb{R}_+ \times \mathbb{R}_+$. Here, the assumption of independence is to be understood as the independence of its natural filtrations $\mathbb{F}^{\mu_i} = (\mathcal{F}_t^{\mu_i})_{t \geqslant 0}$ defined by

$$\mathcal{F}_0^{\mu_i} = \{\Omega, \varnothing\},\ \mathcal{F}_t^{\mu_i} = \sigma(\mu_i((0, s] \times B) : s \in (0, t], B \in \mathcal{B}(\mathbb{R}_+)),\ t \geqslant 0.$$

We further denote the common filtration of μ by $\mathbb{F}^\mu = (\mathcal{F}_t^\mu)_{t\geqslant 0}$ with $\mathbb{F}^\mu = \bigvee_{i=1}^m \mathbb{F}^{\mu_i}$. The definition of a Poisson random measure entails the independence of increments, i.e. $\mu_i((u,s] \times B)$ is independent of $\mathbb{F}_u^{\mu_i}$ (see Jacod & Shiryaev, 2003, Definition 1.20). For each μ_i, let $\overline{\mu}_i$ denote its $\mathbb{F}$-compensator, given by the (deterministic) intensity measure $\overline{\mu}_i = \lambda \otimes n_i$, where λ is the Lebesgue measure and n_i is any σ-finite measure on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$, $i = 1, \ldots, m$. That is, the stochastic processes

$$t \mapsto m_i((0,t] \times B) := \mu_i((0,t] \times B) - \overline{\mu}_i((0,t] \times B),\ \ B \in \mathcal{B}(\mathbb{R}_+),$$

define $\mathbb{F}$-martingales.

Similar to the previous subsection, we allow again for dependency of the risk factors on a grid. Therefore, let the grid π be given as $\pi = \{0 = u_0 < \ldots < u_d = T\}$, $d \in \mathbb{N}$ and let $B_1, \ldots, B_K \in \mathcal{B}(\mathbb{R}_+)$ be disjoint Borel sets with $n_i(B_r) < \infty$, $r = 1, \ldots, K$. We define a family of σ-algebras $\mathbb{I}^i = (\mathcal{I}_{s,t}^i)_{s\leqslant t}$ by $\mathcal{I}_{s,t}^i = \sigma(\mu_i((s,u] \times B_r) : s < u \leqslant t, r = 1, \ldots, K)$. Furthermore, let $\mathbb{I} = (\mathcal{I}_t)_{t\geqslant 0}$, $\mathcal{I}_t = \bigvee_{i=1}^m \mathcal{I}_{0,t}^i$ denote a subfiltration of $\mathbb{F}^\mu$.

Let $f_i \colon \mathbb{R}_+ \times \mathbb{R}_+ \times \Omega \to \mathbb{R}$ be simple predictable functions, i.e.

$$f_i(t,z) = \sum_{l=1}^d \sum_{r=1}^K c_r^i Y_l^i \mathbb{1}_{(u_l, u_{l+1}]}(t) \mathbb{1}_{B_r}(z),$$

where $c_r^i \in \mathbb{R}$, $r = 1, \ldots, K$, and Y_l^i are bounded $\mathcal{I}_{u_l}$-measurable random variables with $Y_l^i \neq 0$, $l = 1, \ldots, d$. In this paragraph, the risk basis $X = (X_1, \ldots, X_m)$ is defined via

$$X_i(t) = \int_{(0,t]\times\mathbb{R}_+} f_i(s,z)\mu_i(\mathrm{d}s, \mathrm{d}z) = \sum_{j=1}^d \sum_{r=1}^K c_r^i Y_j^i \mu_i(((u_j, u_{j+1}] \cap (0,t]) \times B_r), \tag{4.18}$$

are $\mathbb{F}$-martingales ($i = 1, \ldots, m$).

In contrast to the previous subsections, the risk driver μ has a further dimension. To reflect that, we define the available information slighty different compared to Section 4.1. More precisely, let $\mathbb{G}^i = (\mathcal{G}_t^i)_{t\geqslant 0}$ describe the information generated by the i-th source of risk with

$$\mathcal{G}_t^i = \sigma(X_i(s, B_r) : s \leqslant t, r = 1, \ldots, K),$$

where the random variables $X_i(t, B_r)$, $t \geqslant 0$, $r = 1, \ldots, K$, are given by

$$X_i(t, B_r) = \int_{(0,t]\times B_r} f_i(s,z)\mu_i(\mathrm{d}s, \mathrm{d}z) = \sum_{j=1}^d c_r^i Y_j^i \mu_i(((u_j, u_{j+1}] \cap (0,t]) \times B_r).$$

As a consequence, the joint filtration $\mathbb{G} = (\mathcal{G}_t)_{t\in[0,T]}$ with $\mathcal{G}_t = \mathcal{G}_t^1 \vee \ldots \vee \mathcal{G}_t^m$ also slightly differs from Section 4.1.

Following Ikeda and Watanebe (1989, Section II.3), X_i, $i = 1, \ldots, m$ are special semimartingale under $\mathbb{F}$ with compensators

$$A_i(t) = \int_{(0,t]\times\mathbb{R}_+} f_i(s,z)\overline{\mu}_i(\mathrm{d}s,\mathrm{d}z) = \sum_{j=1}^{d}\sum_{r=1}^{K} c_r^i Y_j^i \overline{\mu}_i(((u_j,u_{j+1}] \cap (0,t]) \times B_r) \qquad (4.19)$$

and square-integrable martingale parts

$$M_i(t) = \int_{(0,t]\times\mathbb{R}_+} f_i(s,z)m_i(\mathrm{d}s,\mathrm{d}z) = \sum_{j=1}^{d}\sum_{r=1}^{K} c_r^i Y_j^i m_i(((u_j,u_{j+1}] \cap (0,t]) \times B_r). \quad (4.20)$$

The measurability assumption on the functions f_i allows X_i (and thus M_i) to depend on the common past information (up to the grid point) of μ. However, between the grid points, newly generated information for M_i and X_i stems only from the newly generated information of μ_i. To be precise, let $t_k, t_{k+1} \in \mathcal{T}_n(t)$, such that $(t_k, t_{k+1}) \cap \pi = \varnothing$. For $s \in (t_k, t_{k+1}]$, we have

$$\sigma(M_i(s) - M_i(t_k) : t_k < s \leqslant t_{k+1}) \subseteq \mathcal{I}_{t_k} \vee \mathcal{I}^i_{t_k,s} \qquad (4.21)$$

and

$$\sigma(X_i(s) - X_i(t_k) : t_k < s \leqslant t_{k+1}) \subseteq \mathcal{I}_{t_k} \vee \mathcal{I}^i_{t_k,s}. \qquad (4.22)$$

The latter observation will help us to prove the following lemma.

Lemma 4.11. *Let $J \subseteq \{1,\ldots,m\}$ and let $t_k < s$, such that $(t_k, s) \cap \pi = \varnothing$. Then it holds*

$$\mathcal{G}^J_{t_k,s} = \mathcal{I}_{t_k} \vee \bigvee_{j\in J} \mathcal{I}^j_{t_k,s}. \qquad (4.23)$$

Furthermore, we have $\mathbb{G} = \mathbb{I}$.

Proof. The arguments are similar to those in the proof of Lemma 4.7, but adapted to the setup of Poisson random measures. Again, we prove the result in the reverse order to that stated, i.e. we first show that

$$\mathbb{G} = \mathbb{I}. \qquad (4.24)$$

Due to our assumptions, we have $\mathcal{G}_0 = \mathcal{I}_0 = \{\Omega, \varnothing\}$. Furthermore, (4.18) implies

$$\mathcal{G}_s \subseteq \mathcal{I}_s$$

for $s \in [0,T]$. Vice versa, we use an induction argument. Let $s \in (0, u_1]$ and $r \in \{1,\ldots,K\}$. Rearranging (4.18) yields

$$\mu_i((0,s] \times B_r) = (c_r^i Y_1^i)^{-1} X_i(s, B_r).$$

As $(c_r^i Y_1^i)^{-1}$ is $\mathcal{I}_0$-measurable on $(0,u_1]$, we get $\mathcal{I}_s \subseteq \mathcal{G}^i_s$ for $s \in (0,u_1]$. Since the argument does not depend on i, we also have $\mathcal{I}_s \subseteq \mathcal{G}_s$ for $s \in (0,u_1]$. Let us now assume that $\mathcal{I}_s \subseteq \mathcal{G}_s$

holds for $s \in [0, u_l]$. For $s \in (u_l, u_{l+1}]$ we have that

$$\mathcal{I}_s = \mathcal{I}_{u_l} \vee \sigma(\mu_i((u_l, u] \times B_r) : i = 1, \ldots, m, u_l < u \leqslant s, r = 1, \ldots, K).$$

Again with (4.18), we observe that

$$\mu_i((u_l, u] \times B_r) = (c_r^i Y_l^i)^{-1}(X_i(u, B_r) - X_i(u_l, B_r)). \tag{4.25}$$

Since $\mathcal{I}_{u_l} \subseteq \mathcal{G}_{u_l}$ and since $(c_r^i Y_l^i)^{-1}$ is $\mathcal{I}_{u_l}$-measurable, thus $\mathcal{G}_{u_l}$-measurable, the representation (4.25) implies $\mathcal{I}_s \subseteq \mathcal{G}_s$ also for all $s \in (u_l, u_{l+1}]$. So, via induction we have shown that $\mathbb{G} = \mathbb{I}$. Together with (4.22), we get the inclusion

$$\mathcal{G}^J_{t_k,s} \subseteq \mathcal{I}_{t_k} \vee \bigvee_{j \in J} \mathcal{I}^J_{t_k,s}.$$

For the other direction, the equation (4.24) together with the representation (4.25) gives the desired result. □

We can conclude from the second part in the previous lemma, that the risk basis X satisfies (S) with decompositions (A_i, M_i), $i = 1, \ldots, m$, given by (4.22) and (4.21).

Proposition 4.12. *Let the risk basis be defined by* (4.18)*. Then (M) is fulfilled.*

Proof. The arguments are similar to those in the proof of Proposition 4.8, but adapted to the setup of Poisson random measures. Let i be fixed. Furthermore, let $(\mathcal{T}_n(t))_n$ be a vanishing sequence of partitions on $[0, t]$ and let $J \subseteq \{1, \ldots, m\}$. We have to show that

$$\operatorname*{plim}_{n\to\infty} \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] = 0, \quad i \notin J, \tag{4.26}$$

and

$$\operatorname*{plim}_{n\to\infty} \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] = M_i(t), \quad i \in J, \tag{4.27}$$

where $M_i(t) = \int_{(0,t] \times \mathbb{R}_+} f_i(s, z) m_i(\mathrm{d}s, \mathrm{d}z)$. Therefore, we first show that

$$\sigma(M_i(t_{k+1}) - M_i(t_k)) \perp\!\!\!\perp \sigma(X_j(s) : t_k < s \leqslant t_{k+1}, j \in J) \,|\, \mathcal{G}_{t_k}, \tag{4.28}$$

where $u_l < t_k < t_{k+1} \leqslant u_{l+1}$ and $J \subseteq \{1, \ldots, m\} \backslash \{i\}$.

Due to (4.21), (4.22) and the fact that $\mathcal{G}_{t_k} = \mathcal{I}_{t_k}$ (see Lemma 4.11), it suffices to show

$$\mathcal{I}_{t_k} \vee \mathcal{I}^i_{t_k, t_{k+1}} \perp\!\!\!\perp \mathcal{I}_{t_k} \vee \bigvee_{j \in J} \mathcal{I}^j_{t_k, t_{k+1}} \,|\, \mathcal{G}_{t_k}, \ j \in J.$$

This, however, follows immediately from $\mathbb{G} = \mathbb{I}$ (see Lemma 4.11) and the independence between μ_i and μ_j, $j \in J$.

Without loss of generality we assume that $|\mathcal{T}_n(t)| < |\pi|$ for all n. If $u_l < t$ and $u_l \in (t_j, t_{j+1}]$ for some index j, we set $\underline{t}_l = t_j$ and $\bar{t}_l = t_{j+1}$ as the neighbouring points of

u_l in $\mathcal{T}_n(t)$. If $u_l \geqslant t$, we set $\underline{t}_l = \bar{t}_l = t$. Moreover, we write $\mathcal{T}_n^0(t) = \mathcal{T}_n(t) \cap [u_0, u_1]$ and $\mathcal{T}_n^l(t) = \mathcal{T}_n(t) \cap (u_l, u_{l+1}]$, $l = 0, \ldots, d-1$.

For (4.26), i.e. $i \notin J$, the conditional independence (4.28), yields

$$\begin{aligned}
&\sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] \\
&= \sum_{l=0}^{m-1} \sum_{\mathcal{T}_n^l} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] + \sum_{l=1}^{d-1} \mathbb{E}[M_i(\bar{t}_l) - M_i(\underline{t}_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}] \\
&= \sum_{l=0}^{m-1} \sum_{\mathcal{T}_n^l} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}_{t_k}] + \sum_{l=1}^{d-1} \mathbb{E}[M_i(\bar{t}_l) - M_i(\underline{t}_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}] \\
&= \sum_{l=1}^{d-1} \mathbb{E}[M_i(\bar{t}_l) - M_i(\underline{t}_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}]
\end{aligned}$$

Let us recall that $\| \cdot \|_2$ denotes the norm on $L_2(\mathbb{P})$. With the Jensen's inequality for conditional expectations, the Itô isometry (see Ikeda & Watanabe, 1989, Section II.3) and the dominated convergence theorem, we observe

$$\begin{aligned}
\lim_{n\to\infty} \left\| \sum_{l=1}^{d-1} \mathbb{E}[M_i(\bar{t}_l) - M_i(\underline{t}_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}] \right\|_2^2 &= \lim_{n\to\infty} \sum_{l=1}^{d-1} \|\mathbb{E}[M_i(\bar{t}_l) - M_i(\underline{t}_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}]\|_2^2 \\
&\leqslant \lim_{n\to\infty} \sum_{l=1}^{d-1} \mathbb{E}[(M_i(\bar{t}_l) - M_i(\underline{t}_l))^2] \\
&= \lim_{n\to\infty} \sum_{l=1}^{d-1} \mathbb{E}\left[\langle M_i, M_i\rangle(\bar{t}_l) - \langle M_i, M_i\rangle(\underline{t}_l)\right] = 0,
\end{aligned}$$

where we used the continuity of

$$\langle M_i, M_i \rangle(t) = \int_{(0,t]\times\mathbb{R}_+} (f_i(s,z))^2 \overline{\mu}_i(\mathrm{d}s, \mathrm{d}z)$$

in t, see Ikeda and Watanabe (1989). Thus, we have shown (4.26).

Next, we prove (4.27) (i.e. $i \in J$). With (4.21) and Lemma 4.11, we get

$$\begin{aligned}
&\sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] \\
&= M_i(t) + \sum_{l=1}^{d-1} E[M_i(\bar{t}_l) - M_i(u_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}] - \sum_{l=1}^{d-1} (M_i(\bar{t}_l) - M_i(u_l)).
\end{aligned}$$

With similar arguments as above, the last two terms tend again to 0 in L_2, thus in probability. □

With this proof, we close the third example and move on to the widely used multistate Markov models.

4.2.4 Competing risks in life insurance

In this section, we investigate property (M) in a life insurance context with competing risks, i.e. we model a policyholder whose life insurance contract can be terminated due to different, exclusive causes (e.g. lapse or death). Our setup follows Milbrodt and Helbig (1999, Section 3.C).

Let $I_W = \{1, \dots, m\}$ denote the set of withdrawal causes. The idea is to model each cause of withdrawal with its own clock, with the clock that stops first determining the time of withdrawal and the cause of withdrawal (see Milbrodt & Helbig, 1999, Remark after Definition 3.22). Therefore, let T_i, $i \in I_W$, be independent random variables with values in $[0, \infty)$. We assume that each T_i has a continuous intensity, i.e. its distribution function F_{T_i} fulfils

$$1 - F_{T_i}(t) = e^{-\int_0^t \lambda_i(u)\mathrm{d}u}, \ t \geqslant 0,$$

for a continuous function $\lambda_i : \mathbb{R} \to \mathbb{R}$. In the following, the *withdrawal time* or *contract lifetime* is set as

$$T = \bigwedge_{i \in I_W} T_i,$$

where $a \wedge b = \min(a, b)$, $a, b, \in \mathbb{R}$. In particular, T is well-defined, i.e. $\mathbb{P}(T_i = T_j \in [0, \infty)) = 0$ (see Milbrodt & Helbig, 1999, Theorem 3.23). We now define the counting processes $N_i = (N_i(t))_{t \geqslant 0}$ by

$$N_i(t) = \mathbb{1}_{\{T = T_i \leqslant t\}}, \ t \geqslant 0, \tag{4.29}$$

which describes if the policyholder terminates his contract due to cause i until time t.

Let the risk basis be given as $X = (N_1, \dots, N_m)$, and we suppose that $\mathbb{F} = \mathbb{G}$. Each N_i satisfies (S) with a compensator

$$A_i(t) = \int_0^t \mathbb{1}_{\{T \geqslant s\}} \lambda_i(s)\mathrm{d}s$$

and a square-integrable martingale part

$$M_i(t) = N_i(t) - \int_0^t \mathbb{1}_{\{T \geqslant s\}} \lambda_i(s)\mathrm{d}s,$$

see Milbrodt & Helbig (1999, Theorem 10.37).

For the following lemma, which helps us to prove property (M), we need some further notation. For a subset $J \subseteq I_W$, we write $T^J = \bigwedge_{i \in J} T_i$ and $N^J = \sum_{i \in J} N_i$, whereas $T^{\varnothing} = +\infty$ and $N^{\varnothing} = 0$. Then it holds

$$1 - F_{T^J}(t) = \prod_{i \in J} (1 - F_{T_i}(t)) = e^{-\int_0^t \sum_{i \in J} \lambda_i(s)\mathrm{d}s}, \tag{4.30}$$

where $F_{T^J}(t)$ denotes the distribution function of T^J.

Lemma 4.13. *Let $J \subseteq \{1, \ldots, M\}$ and $i \notin J$. For $t_k, t_{k+1} \in \mathcal{T}_n(t)$ and $s \in (t_k, t_{k+1}]$, we denote $U = \mathbb{E}[\mathbb{1}_{\{T \geqslant s\}} | \mathcal{G}^J_{t_k, t_{k+1}}]$ and $V = \mathbb{E}[N_i(t_{k+1}) - N_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}]$. Furthermore, let $A = \{T > t_k\}$, $B = \{N^J(t_{k+1}) - N^J(t_k) = 1\}$, $B^c = \{N^J(t_{k+1}) - N^J(t_k) = 0\}$. Then the following statements are true:*

a) For $\omega \in A \cap B$, it holds

$$U(\omega) = \mathbb{1}_{\{T^J \geqslant s\}}(\omega).$$

b) For $\omega \in A \cap B^c$, it holds

$$U(\omega) = \frac{\mathbb{P}(T^J > t_{k+1})\mathbb{P}(T^{J^c} \geqslant s) + \mathbb{P}(T^J \in [s, t_{k+1}], s \leqslant T^{J^c} \leqslant T^J)}{\mathbb{P}(T^J > t_{k+1})\mathbb{P}(T^{J^c} > t_k) + \mathbb{P}(T^J \in (t_k, t_{k+1}], t_k < T^{J^c} \leqslant T^J)}.$$

c) For $\omega \in A \cap B^c$, it holds

$$V(\omega) = \frac{\mathbb{P}(T = T_i \in (t_k, t_{k+1}])}{\mathbb{P}(T^J > t_{k+1})\mathbb{P}(T^{J^c} > t_k) + \mathbb{P}(T^J \in (t_k, t_{k+1}], t_k < T^{J^c} \leqslant T^J)}.$$

Proof. Let $J \subseteq \{1, \ldots, M\}$ and let the sets A and B defined as in the Lemma. Then equality a) follows from

$$\begin{aligned}
\mathbb{1}_{A \cap B}\mathbb{E}[\mathbb{1}_{\{T \geqslant s\}} | \mathcal{G}^J_{t_k, t_{k+1}}] &= \mathbb{1}_A \mathbb{E}[\mathbb{1}_B \mathbb{1}_{\{T^J \geqslant s\}} | \mathcal{G}^J_{t_k, t_{k+1}}] = \mathbb{1}_A \mathbb{E}[\mathbb{1}_B \mathbb{1}_{\{T = T^J \geqslant s\}} | \mathcal{G}^J_{t_k, t_{k+1}}] \\
&= \mathbb{1}_A \mathbb{E}[\mathbb{1}_B \mathbb{1}_{\{N^J(s-) = 0\}} | \mathcal{G}^J_{t_k, t_{k+1}}] = \mathbb{1}_{A \cap B} \mathbb{1}_{\{N^J(s-) = 0\}} \\
&= \mathbb{1}_{A \cap B} \mathbb{1}_{\{T^J \geqslant s\}}.
\end{aligned}$$

For the equality b), we first show that

$$\mathbb{1}_{A \cap B^c}\mathbb{E}[\mathbb{1}_{\{T \geqslant s\}} | \mathcal{G}^J_{t_k, t_{k+1}}] = \mathbb{1}_{A \cap B^c}\mathbb{E}[\mathbb{1}_{\{T \geqslant s\}} | \sigma(\mathbb{1}_{\{T > t_k\}}, N^J(t_{k+1}) - N^J(t_k))]. \tag{4.31}$$

Therefore, we observe that $\sigma(\mathbb{1}_{\{T > t_k\}}, N^J(t_{k+1}) - N^J(t_k)) \subseteq \mathcal{G}^J_{t_k, t_{k+1}}$ due to

$$\{T > t_k\} = \{N_i(t_k) = 0, i \in I_W\}.$$

Furthermore, for any $C \in \mathcal{G}^J_{t_k, t_{k+1}}$, we find $\widetilde{C} \in \sigma(\mathbb{1}_{\{T > t_k\}}, N^J(t_{k+1}) - N^J(t_k))$, such that $\mathbb{1}_{A \cap B^c \cap C} = \mathbb{1}_{A \cap B^c \cap \widetilde{C}}$, which implies

$$\begin{aligned}
\mathbb{E}[\mathbb{1}_{A \cap B^c \cap C}\mathbb{1}_{\{T \geqslant s\}}] &= \mathbb{E}[\mathbb{1}_{A \cap B^c \cap \widetilde{C}}\mathbb{1}_{\{T \geqslant s\}}] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_{A \cap B^c \cap \widetilde{C}}\mathbb{1}_{\{T \geqslant s\}} | \sigma(\mathbb{1}_{\{T > t_k\}}, N^J(t_{k+1}) - N^J(t_k))]] \\
&= \mathbb{E}[\mathbb{1}_C \mathbb{E}[\mathbb{1}_{A \cap B^c}\mathbb{1}_{\{T \geqslant s\}} | \sigma(\mathbb{1}_{\{T > t_k\}}, N^J(t_{k+1}) - N^J(t_k))]],
\end{aligned}$$

thus (4.31) holds. Since $\mathbb{1}_{\{T > t_k\}}$ and $N^J(t_{k+1}) - N^J(t_k)$ assumes only values in $\{0, 1\}$, we may interpret the conditional expectation on the right-hand side of (4.31) as a conditional expectation given an event (see Jakubowski & Niewęgłowski, 2008, Lemma 3). More precisely, we have

$$\mathbb{1}_{A \cap B^c}\mathbb{E}[\mathbb{1}_{\{T \geqslant s\}} | \sigma(\mathbb{1}_{\{T > t_k\}}, N^J(t_{k+1}) - N^J(t_k))] = \mathbb{1}_{A \cap B^c}\mathbb{E}[\mathbb{1}_{\{T \geqslant s\}} | A \cap B^c].$$

The conditional event can be written as a disjoint union of the following two sets

$$A \cap B^{\mathsf{c}} = \{T^J > t_{k+1}, T^{J^{\mathsf{c}}} > t_k\} \cup \{T^J \in (t_k, t_{k+1}], t_k < T^{J^{\mathsf{c}}} \leqslant T^J\}. \tag{4.32}$$

That gives us, with the independence of T_j, $j = 1, \dots, m$,

$$\begin{aligned}
&\mathbb{E}[\mathbb{1}_{\{T \geqslant s\}} | A \cap B^{\mathsf{c}}] \\
&= \frac{\mathbb{P}(\{T \geqslant s\} \cap (\{T^J > t_{k+1}, T^{J^{\mathsf{c}}} > t_k\} \cup \{T^J \in (t_k, t_{k+1}], t_k < T^{J^{\mathsf{c}}} \leqslant T^J\}))}{\mathbb{P}(\{T^J > t_{k+1}, T^{J^{\mathsf{c}}} > t_k\} \cup \{T^J \in (t_k, t_{k+1}], t_k < T^{J^{\mathsf{c}}} \leqslant T^J\})} \\
&= \frac{\mathbb{P}(T^J > t_{k+1})\mathbb{P}(T^{J^{\mathsf{c}}} \geqslant s) + \mathbb{P}(T^J \in (s, t_{k+1}], s \leqslant T^{J^{\mathsf{c}}} \leqslant T^J)}{\mathbb{P}(T^J > t_{k+1})\mathbb{P}(T^{J^{\mathsf{c}}} > t_k) + \mathbb{P}(T^J \in (t_k, t_{k+1}], t_k < T^{J^{\mathsf{c}}} \leqslant T^J)},
\end{aligned}$$

so we have shown b).

Finally, we prove c). For $i \notin J$, this follows from

$$\begin{aligned}
&\mathbb{1}_{A \cap B^{\mathsf{c}}} \mathbb{E}[N_i(t_{k+1}) - N_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] \\
&= \mathbb{1}_{A \cap B^{\mathsf{c}}} \mathbb{P}(N_i(t_{k+1}) - N_i(t_k) = 1 | T > t_k, N^J(t_{k+1}) - N^J(t_k) = 0) \\
&= \mathbb{1}_{A \cap B^{\mathsf{c}}} \mathbb{P}(T = T_i \in (t_k, t_{k+1}] | T > t_k, N^J(t_{k+1}) - N^J(t_k) = 0) \\
&= \mathbb{1}_{A \cap B^{\mathsf{c}}} \frac{\mathbb{P}(T = T_i \in (t_k, t_{k+1}], T > t_k, N^J(t_{k+1}) - N^J(t_k) = 0)}{\mathbb{P}(T > t_k, N^J(t_{k+1}) - N^J(t_k) = 0)} \\
&= \mathbb{1}_{A \cap B^{\mathsf{c}}} \frac{\mathbb{P}(T = T_i \in (t_k, t_{k+1}])}{\mathbb{P}(T^J > t_{k+1})\mathbb{P}(T^{J^{\mathsf{c}}} > t_k) + \mathbb{P}(T^J \in (t_k, t_{k+1}], t_k < T^{J^{\mathsf{c}}} \leqslant T^J)},
\end{aligned}$$

where we again used (4.32) and the independence of T_j, $j = 1, \dots, m$. □

With the Lemma 4.13, we are now able to prove the property (M) in the framework of competing risks.

Proposition 4.14. *Let the risk basis $X = (N_1, \dots, N_m)$ be defined via (4.29). Then the property (M) is fulfilled.*

Proof. Let $(\mathcal{T}_n(t))_n$ be a vanishing sequence of partitions on $[0, t]$ and let $J \subseteq \{1, \dots, m\}$. We have to show that

$$\operatorname*{plim}_{n \to \infty} \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] = 0, \quad i \notin J, \tag{4.33}$$

and

$$\operatorname*{plim}_{n \to \infty} \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] = M_i(t), \quad i \in J, \tag{4.34}$$

where $M_i(t) = N_i(t) - \int_0^t \mathbb{1}_{\{T \geqslant s\}} \lambda_i(s) \mathrm{d}s$.

Suppose $i \in J$. We denote $A = \{T > t_k\}$, $B = \{N^J(t_{k+1}) - N^J(t_k) = 1\}$ and $B^{\mathsf{c}} = \{N^J(t_{k+1}) - N^J(t_k) = 0\}$ as in Lemma 4.13. It holds $A = \{\sum_{i=1}^m N_i(t_k) = 0\}$.

Together with Lemma 4.13, this yields

$$
\begin{aligned}
&\mathbb{E}[M_i(t_{k+1}) - M_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}] \\
&= \mathbb{1}_A \mathbb{E}[M_i(t_{k+1}) - M_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}] \\
&= \mathbb{1}_{A\cap B} \mathbb{E}[(M_i(t_{k+1}) - M_i(t_k))|\mathcal{G}^J_{t_k,t_{k+1}}] + \mathbb{1}_{A\cap B^c} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}] \\
&= \mathbb{1}_{A\cap B} \left(N_i(t_{k+1}) - N_i(t_k) - \int_{t_k}^{t_{k+1}} \mathbb{1}_{\{T^J \geqslant s\}} \lambda_i(s) \mathrm{d}s \right) - \mathbb{1}_{A\cap B^c} \int_{t_k}^{t_{k+1}} p_n(s) \lambda_i(s) \mathrm{d}s,
\end{aligned}
$$

where

$$
p_n(s) = \frac{\mathbb{P}(T^J > t_{k+1})\mathbb{P}(T^{J^c} \geqslant s) + \mathbb{P}(T^J \in [s, t_{k+1}], s \leqslant T^{J^c} \leqslant T^J)}{\mathbb{P}(T^J > t_{k+1})\mathbb{P}(T^{J^c} > t_k) + \mathbb{P}(T^J \in (t_k, t_{k+1}], t_k < T^{J^c} \leqslant T^J)}, \; s \in (t_k, t_{k+1}]. \tag{4.35}
$$

Furthermore, we observe that

$$
\begin{aligned}
M_i(t) &= \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{1}_{\{T > t_k\}} (M_i(t_{k+1}) - M_i(t_k)) \\
&= \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{1}_{\{T > t_k\}} \mathbb{1}_{\{N^J(t_{k+1}) - N^J(t_k) = 1\}} \left(N_i(t_{k+1}) - N_i(t_k) - \int_{t_k}^{t_{k+1}} \mathbb{1}_{\{T^J \geqslant s\}} \lambda_i(s) \mathrm{d}s \right) \\
&\quad - \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{1}_{\{T > t_k\}} \mathbb{1}_{\{N^J(t_{k+1}) - N^J(t_k) = 0\}} \int_{t_k}^{t_{k+1}} \mathbb{1}_{\{T^{J^c} \geqslant s\}} \lambda_i(s) \mathrm{d}s.
\end{aligned}
$$

Thus, we get

$$
\begin{aligned}
&\left| \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}] - M_i(t) \right| \\
&= \left| \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{1}_{\{T > t_k\}} \mathbb{1}_{\{N^J(t_{k+1}) - N^J(t_k) = 0\}} \int_{t_k}^{t_{k+1}} (p_n(s) - \mathbb{1}_{\{T^{J^c} \geqslant s\}}) \lambda_i(s) \mathrm{d}s \right| \leqslant \int_0^t |f_n(s)| \mathrm{d}s
\end{aligned}
$$

with

$$
f_n(s) = \mathbb{1}_{\{T > t_k\}} \mathbb{1}_{\{N^J(t_{k+1}) - N^J(t_k) = 0\}} (p_n(s) - \mathbb{1}_{\{T^{J^c} \geqslant s\}}) \lambda_i(s)
$$

for $s \in (t_k, t_{k+1}]$. To apply the dominated convergence theorem, we need to investigate $\lim_{n\to\infty} f_n(s)$ for fixed s. Therefore, we first examine $p_n(s)$. Since

$$
\begin{aligned}
\mathbb{P}(T^J \in [s, t_{k+1}], s \leqslant T^{J^c} \leqslant T^J) &= \int_s^{t_{k+1}} \int_s^u f_{T^J}(u) f_{T^{J^c}}(v) \mathrm{d}v \mathrm{d}u \\
&= \int_s^{t_{k+1}} f_{T^J}(u) (F_{T^{J^c}}(u) - F_{T^{J^c}}(s)) \mathrm{d}u
\end{aligned}
$$

we can conclude $\lim_{n\to\infty} \mathbb{P}(T^J \in [s, t_{k+1}], s \leqslant T^{J^c} \leqslant T^J) = 0$. Having continuous densities, the same argumentation yields $\lim_{n\to\infty} \mathbb{P}(T^J \in (t_k, t_{k+1}], t_k < T^{J^c} \leqslant T^J) = 0$. Exploiting the continuity of the distribution function, we end up with

$$
\lim_{n\to\infty} p_n(s) = \frac{\mathbb{P}(T^J > s)\mathbb{P}(T^{J^c} \geqslant s)}{\mathbb{P}(T^J > s)\mathbb{P}(T^{J^c} \geqslant s)} = 1.
$$

That gives us

$$\lim_{n\to\infty} f_n(s) = \mathbb{1}_{\{T\geqslant s\}}\mathbb{1}_{\{N^J(s)-N^J(s-)=0\}}(1-\mathbb{1}_{\{T^{J^c}\geqslant s\}})\lambda_i(s) = 0,$$

almost surely, since $\mathbb{1}_{\{T^{J^c}\geqslant s\}} = 1$ on $\{T \geqslant s\}$. Finally, the dominated convergence theorem yields

$$\lim_{n\to\infty} \Big|\sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1})-M_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}] - M_i(t)\Big| = 0$$

almost surely, thus in probability.

Now, let $i \notin J$. With help of Lemma 4.13, we get

$$\begin{aligned}
&\mathbb{E}[M_i(t_{k+1})-M_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}]\\
&= \mathbb{1}_A\mathbb{E}[M_i(t_{k+1})-M_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}]\\
&= \mathbb{1}_{A\cap B}\mathbb{E}[M_i(t_{k+1})-M_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}] + \mathbb{1}_{A\cap B^c}\mathbb{E}[M_i(t_{k+1})-M_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}]\\
&= -\mathbb{1}_{A\cap B}\mathbb{E}\left[\int_{t_k}^{t_{k+1}} \mathbb{1}_{\{T\geqslant s\}}\lambda_i(s)\mathrm{d}s\,\middle|\,\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&\quad + \mathbb{1}_{A\cap B^c}\mathbb{E}\left[N_i(t_{k+1})-N_i(t_k)-\int_{t_k}^{t_{k+1}} \mathbb{1}_{\{T\geqslant s\}}\lambda_i(s)\mathrm{d}s\,\middle|\,\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= -\mathbb{1}_{A\cap B}\int_{t_k}^{t_{k+1}} \mathbb{1}_{\{T^J\geqslant s\}}\lambda_i(s)\mathrm{d}s + \mathbb{1}_{A\cap B^c}q_n - \mathbb{1}_{A\cap B^c}\int_{t_k}^{t_{k+1}} p_n(s)\lambda_i(s)\mathrm{d}s,
\end{aligned}$$

where $p_n(s)$ is given by (4.35), and $q_n \in [0,1]$ is given by

$$q_n = \frac{\mathbb{P}(T = T_i \in (t_k,t_{k+1}])}{\mathbb{P}(T^J > t_{k+1})\mathbb{P}(T^{J^c} > t_k) + \mathbb{P}(T^J \in (t_k,t_{k+1}], t_k < T^{J^c} \leqslant T^J)}.$$

For q_n, we derive an upper bound by

$$\begin{aligned}
q_n &\leqslant \frac{\mathbb{P}(T = T_i \in (t_k,t_{k+1}])}{\mathbb{P}(T^J > t_{k+1})\mathbb{P}(T^{J^c} > t_k)} = \frac{\mathbb{P}(T_i \in (t_k,t_{k+1}], T_i \leqslant T_j, j\neq i)}{\mathbb{P}(T^J > t_{k+1})\mathbb{P}(T^{J^c} > t_k)}\\
&= \frac{\int_{t_k}^{t_{k+1}}\int_u^\infty f_{T_i}(u) f_{T^{\{1,\dots,m\}\setminus\{i\}}}(v)\mathrm{d}v\mathrm{d}u}{\mathbb{P}(T^J > t_{k+1})\mathbb{P}(T^{J^c} > t_k)} = \int_{t_k}^{t_{k+1}} \frac{f_{T_i}(u)(1-F_{T^{\{1,\dots,m\}\setminus\{i\}}}(u))}{(1-F_{T^J}(t_{k+1}))(1-F_{T^{J^c}}(t_k))}\mathrm{d}u\\
&= \int_{t_k}^{t_{k+1}} \frac{\lambda_i(u)(1-F_{T_i}(u))(1-F_{T^{\{1,\dots,m\}\setminus\{i\}}}(u))}{(1-F_{T^J}(t_{k+1}))(1-F_{T^{J^c}}(t_k))}\mathrm{d}u.
\end{aligned}$$

Thus, we have

$$\sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1})-M_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}] \leqslant \int_0^t (f_1^n(s)+f_2^n(s)+f_3^n(s))\mathrm{d}s,$$

where

$$\begin{aligned}
f_1^n(s) &= -\mathbb{1}_{\{T>t_k\}}\mathbb{1}_{\{N^J(t_{k+1})-N^J(t_k))=1\}}\mathbb{1}_{\{T^J\geqslant s\}}\lambda_i(s),\\
f_2^n(s) &= -\mathbb{1}_{\{T>t_k\}}\mathbb{1}_{\{N^J(t_{k+1})-N^J(t_k)=0\}}\frac{\lambda_i(s)(1-F_{T_i}(s))(1-F_{T^{\{1,\dots,m\}\setminus\{i\}}}(s))}{(1-F_{T^J}(t_{k+1}))(1-F_{T^{J^c}}(t_k))},\\
f_3^n(s) &= \mathbb{1}_{\{T>t_k\}}\mathbb{1}_{\{N^J(t_{k+1})-N^J(t_k)=0\}}p_n(s)\lambda_i(s)
\end{aligned}$$

for $s \in (t_k,t_{k+1}]$. With the continuity of the distribution functions and (4.30), we get the limiting functions

$$\lim_{n\to\infty} f_1^n(s) = -\mathbb{1}_{\{T\geqslant s\}}\mathbb{1}_{\{N^J(s)-N^J(s-)=1\}}\mathbb{1}_{\{T^J\geqslant s\}}\lambda_i(s),$$
$$\lim_{n\to\infty} f_2^n(s) = -\mathbb{1}_{\{T\geqslant s\}}\mathbb{1}_{\{N^J(s)-N^J(s-)=0\}}\lambda_i(s),$$
$$\lim_{n\to\infty} f_3^n(s) = \mathbb{1}_{\{T\geqslant s\}}\mathbb{1}_{\{N^J(s)-N^J(s-)=0\}}\lambda_i(s).$$

This leads, with help of dominated convergence theorem, to

$$\lim_{n\to\infty}\int_0^t (f_1^n(s)+f_2^n(s)+f_3^n(s))\mathrm{d}s = -\int_0^t \mathbb{1}_{\{T\geqslant s\}}\mathbb{1}_{\{N^J(s)-N^J(s-)=1\}}\mathbb{1}_{\{T^J\geqslant s\}}\lambda_i(s)\mathrm{d}s = 0$$

almost surely, where the last equality follows from the fact, that each path of N^J has at most one jump, which is neglectible for the Lebesgue integral. □

This result will be the basis to derive ISU decompositions for life insurance contracts that include several withdrawal causes. For multistate Markov processes the situation between the jump times is quite similar to the setup of multiple withdrawal causes considered in this section. Therefore, extensions to general multi-state frameworks are conceivable. We leave this for future research, but focus on the addition of random systematic risks.

4.2.5 Doubly stochastic Markov processes in life insurance

In the previous paragraph, we have regarded unsystematic, competing risks. As a special case this included the classical life insurance setup, where the policyholder can assume only two states, *alive* and *dead*, so that dying is the only withdrawal cause. In extension to this setup, we also model systematic risks, including financial risks, with help of diffusion processes. In particular, the unsystematic risk is linked to the systematic risks via its mortality intensity. In the following, the risk basis is composed of diffusion processes Θ_i, $i = 1,\dots,m$, modelling the systematic sources of risk, and a jump process N, representing the unsystematic source of risk.

Let $W = (W_1,\dots,W_r)$ be a standard r-dimensional Brownian motion. Furthermore, let the systematic risk $\Theta = (\Theta_1,\dots,\Theta_r)$ consist of risk factors Θ_i, $i = 1,\dots,r$, each of which satisfies a stochastic differential equation

$$\mathrm{d}\Theta_i(t) = \mu_i(t,\Theta_i(t))\mathrm{d}t + \sigma_i(t,\Theta_i(t))\mathrm{d}W_i(t),\ \ \Theta_i(0) = \theta_i(0) \in \mathbb{R}, \tag{4.36}$$

for continuous functions $\mu_i, \sigma_i\colon [0,T]\times\mathbb{R}\to\mathbb{R}$ ($i = 1,\dots,r$), such that

$$\mathbb{E}\left[\int_0^t \sigma_i(u,\Theta_i(u))^2\mathrm{d}u\right] < \infty,\ \ t\in[0,T]. \tag{4.37}$$

The state process of the policyholder is modelled by an $\mathbb{F}^\Theta$-conditional Markov process $Z = (Z_t)_{t\geqslant 0}$ with state space $\mathcal{Z} = \{a,d\}$ and initial value $Z(0) = a$. That is, for every path

$\Theta|_{[0,T]} = \theta$ there exists a usual Markov process Z^θ with state space $\mathcal{Z} = \{a, d\}$, such that

$$\mathcal{L}(Z, \mathbb{P}(\cdot|\Theta|_{[0,T]} = \theta)) = \mathcal{L}(Z^\theta, \mathbb{P}),$$

and such that the family of stochastic matrices $(P(s,t))_{s\leqslant t} = ((p_{jk}(s,t,\Theta))_{j,k})_{s\leqslant t}$, where $((p_{jk}(s,t,\theta))_{j,k})_{s\leqslant t}$ is given by the transition probabilities of Z^θ, satisfies

a) $P(s,t)$ is $\mathcal{F}_t^\Theta$-measurable,

b) $P(s,\cdot)$ is $\mathbb{F}^\Theta$-progressively measurable.

Such a construction has already been used for the stochastic second-order basis in Chapter 3. A detailed introduction on $\mathbb{F}^\Theta$-conditional Markov processes and a proof of existence can be found in Jetses (2018), see also Jakubowski and Niewęgłowski (2010). The corresponding jump process $N = (N(t))_{t\geqslant 0}$ defined via

$$N(t) = \#\{s \in [0,t] : Z(s-) = a, Z(s) = d\}. \tag{4.38}$$

Moreover, let $I = (I(t))_{t\geqslant 0}$ denote the indicator function $I(t) = \mathbb{1}_{\{Z(t)=a\}}$, which keeps the value one as long as the policyholder is alive.

We assume that there exists a non-negative, piecewise-continuous jump intensity $\lambda = (\lambda(t))_{t\geqslant 0}$ such that $\lambda(t) = \lambda(t, \Theta(t \wedge u_l))$, $t \in (u_l, u_{l+1}]$ for a partition

$$\pi = \{0 = u_0 \leqslant \ldots \leqslant u_d = T\}$$

of $[0,T]$, and

$$\mathbb{E}\left[\int_0^t \lambda(s)\mathrm{d}s\right] < \infty,\ \ t \in [0,T]. \tag{4.39}$$

In particular, the definition of an intensity process requires λ to solve the Kolmogorov equations with respect to transition probabilities P (see Section 3.2 and Jetses, 2018), which in our setting yield the survival probability

$$p(s,t) := p_{aa}(s,t) = \exp\left(-\int_s^t \lambda(s)\mathrm{d}s\right).$$

Throughout this section, we consider the risk basis $X = (\Theta_1, \ldots, \Theta_r, N)$, i.e. the number of risk factors is $m = r + 1$. The following argumentation shows that X satisfies (S). With Jakubowski and Niewęgłowski (2010, Theorem 4.1), we conclude that the process $M_N = (M_N(t))_{t\geqslant 0}$ specified by

$$M_N(t) = N(t) - \Lambda(t) \tag{4.40}$$

with

$$\Lambda(t) = \int_0^t I(s)\lambda(s)\mathrm{d}s,$$

defines a $\overline{\mathbb{G}}$-martingale, where $\overline{\mathbb{G}} = (\overline{\mathcal{G}}_t)_{t\geqslant 0}$ is the filtration given by $\overline{\mathcal{G}}_t = \mathcal{F}_T^\Theta \vee \mathcal{F}_t^N$.

In particular, M_N also defines a $\mathbb{G}$-martingale. As Λ is a $\mathbb{G}$-predictable finite variation process, it remains to verify the square integrability of M_N. This, however, follows from the Assumption (4.39) and $\langle M_N, M_N\rangle(t) = \Lambda(t)$ (see Andersen et al., 1993, p. 74; Klebaner, 2005, Theorem 8.2). Thus, N satisfies (S) with decomposition (Λ, M_N).

For the systematic risks Θ_i, $i = 1, \dots, m$, the $\mathbb{G}$-martingale parts M_i, $i = 1, \dots, r$ are given by

$$M_i(t) = \Theta_i(t) - \int_0^t \mu_i(s, \Theta_i(s))\mathrm{d}s = \int_0^t \sigma_i(s, \Theta_i(s))\mathrm{d}W_i(s).$$

Clearly, M_i defines a $\mathbb{F}^{\Theta}$-martingale. With Jakubowski and Niewęgłowski (2010, Proposition 3.4) we find, that M_i is also a $\mathbb{G}$-martingale. The square integrability of M_i results from assumption (4.37) and $\langle M_i, M_i\rangle(t) = \int_0^t(\sigma_i(s, \Theta_i(s)))^2\mathrm{d}s$ (see Klebaner, 2005, Theorem 8.27). As $A_i(t) = \int_0^t \mu_i(s, \Theta_i(s))\mathrm{d}s$ is a $\mathbb{G}$-predictable finite variation process, Θ_i satisfies (S) with decomposition (M_i, A_i).

Recall from Section 4.1 that we have introduced filtrations $\mathbb{G}$, $\mathbb{G}^J$ and a family of sub-σ-algebras $(\mathcal{G}^J_{s,t})_{s,t}$ for subsets $J \subseteq \{1, \dots, m\}$, which refer to the entire risk basis X. In the following, we denote by $\mathbb{H}$, $\mathbb{H}^J$ the filtrations and by $(\mathcal{H}^J_{s,t})_{s,t}$ the family of sub-σ-algebras with $J \subseteq \{1, \dots, m-1\}$ that follow from exluding the unsystematic risk N in the definitions of Section 4.1. Thus $\mathbb{H}$ always refers to the information that is provided by the systematic risks. For showing evidence of property (M), the following lemma is needed.

Lemma 4.15. *Let $u \leqslant v \leqslant s \leqslant t$ and $J \subseteq \{1, \dots, m-1\}$, such that $(u, s) \cap \pi = \varnothing$. Then it holds $\mathcal{H}_t \perp\!\!\!\perp \mathcal{F}^N_v | \mathcal{H}^J_{u,s}$.*

Proof. Let $T := \inf\{s \geqslant 0 : N(s) = 1\}$ whereas $\inf \varnothing = +\infty$. With Proposition 13 of Rao and Swift (2006, Chapter 3), it is sufficient to prove $\mathbb{P}(A|\mathcal{H}_t) = \mathbb{P}(A|\mathcal{H}^J_{u,s})$ for $A \in \mathcal{F}^N_v$. Neglecting the null sets $\mathcal{N}$, the σ-algebra $\mathcal{F}^N_v$ consists of sets $\{T \leqslant z\}$, $\{T > z\}$, $z \in [0, v]$. With the grid measurability of λ and the martingale property of M_N, we have

$$\begin{aligned}\mathbb{P}(T \leqslant z|\mathcal{H}_t) &= \mathbb{P}(N(z) = 1|\mathcal{H}_t) = \mathbb{E}[N(z)|\mathcal{H}_t] = \mathbb{E}[\Lambda(z)|\mathcal{H}_t] = \int_0^z p(0, s)\lambda(s)\mathrm{d}s \\ &= \mathbb{E}[\Lambda(z)|\mathcal{H}^J_{u,s}] = \mathbb{E}[N(z)|\mathcal{H}^J_{u,s}] = \mathbb{P}(N(z) = 1|\mathcal{H}^J_{u,s}) = \mathbb{P}(T \leqslant z|\mathcal{H}^J_{u,s}).\end{aligned}$$

With $\{T > z\} = \{T \leqslant z\}^{\mathrm{c}}$, it also follows $\mathbb{P}(T > v|\mathcal{H}_t) = \mathbb{P}(T > z|\mathcal{H}^J_{u,s})$, which gives us the assertion. $\square$

Proposition 4.16. *Let the risk basis be given by $X = (\Theta_1, \dots, \Theta_r, N)$. Then (M) is fulfilled.*

Proof. Let $J \subseteq \{1, \ldots, m\}$ with $m \in J$. Without loss of generality we assume that $|\mathcal{T}_n(t)| < |\pi|$ for all n. If $u_l < t$ and $u_l \in (t_j, t_{j+1}]$ for some index j, we set $\underline{t}_l = t_j$ and $\bar{t}_l = t_{j+1}$ as the neighbouring points of u_l in $\mathcal{T}_n(t)$. If $u_l \geqslant t$, we set $\underline{t}_l = \bar{t}_l = t$. With the martingale property of M_N (see (4.40)) and the grid measurability of λ, it follows

$$\begin{aligned}
&\sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_N(t_{k+1}) - M_N(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] \\
&= M_N(t) + \sum_{l=0}^{d-1} \mathbb{E}[M_N(\bar{t}_l) - M_N(u_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}] - \sum_{l=0}^{d-1} \mathbb{E}[M_N(\bar{t}_l) - M_N(u_l) | \mathcal{G}^J_{u_l, \bar{t}_l}].
\end{aligned}$$

Thus, it remains to show that the last two addends tend to 0 as $n \to \infty$. It holds

$$\begin{aligned}
&\Big\| \sum_{l=0}^{d-1} \mathbb{E}[M_N(\bar{t}_l) - M_N(u_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}] \Big\|_2^2 = \sum_{l=0}^{d-1} \|\mathbb{E}[M_N(\bar{t}_l) - M_N(u_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}]\|_2^2 \\
&\leqslant \sum_{l=0}^{d-1} \mathbb{E}[(M_N(\bar{t}_l) - M_N(u_l))^2] = \sum_{l=0}^{d-1} \mathbb{E}[\langle M_N, M_N\rangle(\bar{t}_l) - \langle M_N, M_N\rangle(u_l)],
\end{aligned}$$

which tends to 0 as $n \to \infty$ with the dominated convergence theorem and the continuity of $\langle M_N, M_N\rangle(t) = \int_0^t I(s)\lambda(s)\mathrm{d}s$. Similarly, the third summand tends to 0 as $n \to \infty$.

Next, we suppose $m \notin J$. With the martingale property of M_N with respect to $\overline{\mathbb{G}}$, we directly get

$$\begin{aligned}
&\sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_N(t_{k+1}) - M_N(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] \\
&= \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[\mathbb{E}[M_N(t_{k+1}) - M_N(t_k) | \overline{\mathcal{G}}_{t_k}] | \mathcal{G}^J_{t_k, t_{k+1}}] = 0.
\end{aligned}$$

In the following, we fix $i \in \{1, \ldots, m-1\}$ and we suppose $i \in J$. Since M_i is $\mathbb{G}^i$-adapted, we immediately have

$$\sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] = \sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} (M_i(t_{k+1}) - M_i(t_k)) = M_i(t).$$

Last, we consider the case $i \in \{1, \ldots, m-1\}$, $i \notin J$ and write $\mathcal{T}^0_n(t) = \mathcal{T}_n(t) \cap [u_0, u_1]$, $\mathcal{T}^l_n(t) = \mathcal{T}_n(t) \cap (u_l, u_{l+1}]$, $l = 1, \ldots, d-1$. With Lemma 4.15, the independence of Θ_i, $i = 1, \ldots, r$, and the martingale property of M_i, it holds

$$\begin{aligned}
&\sum_{t_k, t_{k+1} \in \mathcal{T}_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}^J_{t_k, t_{k+1}}] \\
&= \sum_{l=0}^{d-1} \sum_{t_k, t_{k+1} \in \mathcal{T}^l_n(t)} \mathbb{E}[M_i(t_{k+1}) - M_i(t_k) | \mathcal{G}_{t_k}] + \sum_{l=0}^{d-1} \mathbb{E}[M_i(\bar{t}_l) - M_i(\underline{t}_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}] \\
&= \sum_{l=0}^{d-1} \mathbb{E}[M_i(\bar{t}_l) - M_i(\underline{t}_l) | \mathcal{G}^J_{\underline{t}_l, \bar{t}_l}].
\end{aligned}$$

The remaining term tends to 0 as $n \to \infty$: With the Itô isometry and the Jensen inequality we get

$$\begin{aligned}\Big\| \sum_{l=0}^{d-1} \mathbb{E}[M_i(\bar{t}_l) - M_i(\underline{t}_l)|\mathcal{G}^J_{\underline{t}_l,\bar{t}_l}]\Big\|_2^2 &= \sum_{l=0}^{d-1} \Big\|\mathbb{E}[M_i(\bar{t}_l) - M_i(\underline{t}_l)|\mathcal{G}^J_{\underline{t}_l,\bar{t}_l}]\Big\|_2^2 \\ &\leqslant \sum_{l=0}^{d-1} \mathbb{E}[(M_i(\bar{t}_l) - M_i(\underline{t}_l))^2] \\ &= \sum_{l=0}^{d-1} \mathbb{E}\bigg[\int_{\underline{t}_l}^{\bar{t}_l} \sigma_i(s, \Theta_i(s))^2 \mathrm{d}s\bigg]\end{aligned}$$

which tends to 0 as $n \to \infty$ by continuity of the Lebesgue integral. □

This result concludes the Section 4.2, which has supported the plausibility of the property (M) by several examples known from actuarial modelling. With that in mind, we now turn our attention back to the ISU decompositions of martingales.

4.3 ISU decompositions of martingales

Recalling that the considered revaluation surplus process R is assumed to be a martingale, we want to focus on the case that R admits a martingale representation, i.e.

$$R(t) - R(0) = \mathbb{E}[\xi|\mathcal{G}_t] - \mathbb{E}[\xi|\mathcal{G}_0] = \sum_{i=1}^{m} \int_0^t H_i(s)\mathrm{d}M_i(s),\ t \in [0, T], \tag{4.41}$$

for unique $\mathbb{G}$-predictable processes H_i. Classical results on the martingale representation are provided by Protter (2005, Chapter IV, Theorem 43) for Brownian motions, by Kunita (2004, Theorem 1.1) for Lévy processes and by Jacobsen (2006, Theorem 4.6.1) for marked point processes. Given that ξ is square-integrable and the representation (4.41) exists, a sufficient condition for the uniqueness of the integrands is a zero predictable quadratic covariation between the martingale parts, i.e. $\langle M_i, M_j\rangle = 0,\ i \neq j$. In that case, the integrands H_i are $\mathbb{P} \otimes \langle M_i, M_i\rangle$ - almost surely unique and it further holds

$$\mathbb{E}\bigg[\int_0^t H_i^2(u)\mathrm{d}\langle M_i, M_i\rangle(u)\bigg] < \infty,\ t \in [0, T], \tag{4.42}$$

see Lemma A.2.3 in the appendix.

Based on a representation (4.41), Schilling et al. (2020) have recently proposed the so-called *MRT decomposition*, which defines the i-th risk contribution as the corresponding martingale, i.e.

$$D_i(t) = \sum_{i=1}^{m} \int_0^t H_i(s)\mathrm{d}M_i(s),\ i = 1, \ldots, m. \tag{4.43}$$

The MRT decomposition fulfils a list of desirable properties of a risk decomposition (see Schilling et al., 2020). The goal of this section is to derive the ISU decomposition for the martingale R and to investigate its relation to the MRT decomposition.

Lemma 4.17. *Suppose the risk basis $X = (X_1, \ldots, X_m)$ fulfils (S) and (M), such that $\langle M_i, M_i\rangle$, $i = 1, \ldots, m$ are continuous processes. Let $(Z_i(t))_{t\in[0,T]}$ be defined by*

$$Z_i(t) = \int_0^t H_i(u)\,dM_i(u)$$

for $\mathbb{G}$-predictable processes H_i, $i = 1, \ldots, m$, that satisfy (4.42)*. Then it holds*

$$\operatorname*{plim}_{n\to\infty} \sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)} \mathbb{E}[Z_i(t_{k+1}) - Z_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}] = \begin{cases} 0, & i \notin J, \\ Z_i(t), & i \in J. \end{cases} \tag{4.44}$$

Proof. Let $(\mathcal{T}_n(t))_n$ be a sequence of partitions of $[0, t]$ with vanishing step lengths and $i \in \{1, \ldots, m\}$ fixed. For a better readability, this proof uses the short-hand notation $\sum_{\mathcal{T}_n(t)}$ instead of $\sum_{t_k,t_{k+1}\in\mathcal{T}_n(t)}$. Firstly, we prove the case where H_i is a bounded simple process, defined on a partition $\pi = \{0 = u_0 < \ldots < u_d = T\}$, $d \in \mathbb{N}$, i.e.

$$H_i = \sum_{l=0}^{d-1} H_{i,l}\mathbb{1}_{(u_l,u_{l+1}]}, \quad i = 1, \ldots, m,$$

and

$$C_i := \sup_{\omega\in\Omega,t\in[0,T]} |H_i(t)| < \infty.$$

Without loss of generality we assume that $|\mathcal{T}_n(t)| < |\pi|$ for all n. If $u_l < t$ and $u_l \in (t_j, t_{j+1}]$ for some index j, we set $\underline{t}_l = t_j$ and $\bar{t}_l = t_{j+1}$ as the neighbouring points of u_l in $\mathcal{T}_n(t)$. If $u_l \geqslant t$, we set $\underline{t}_l = \bar{t}_l = t$. Furthermore, we define partitions $\widetilde{\mathcal{T}}^l_n$ on $[u_l, u_{l+1} \wedge t]$, $n \in \mathbb{N}$, $l = 0, \ldots, d-1$, by

$$\widetilde{\mathcal{T}}^l_n = \begin{cases} (\mathcal{T}_n(t) \cap [u_l, u_{l+1}]) \cup \{u_l, u_{l+1} \wedge t\}, \text{ if } u_l < t, \\ \varnothing, \text{ if } u_l \geqslant t, \end{cases}$$

We start by showing (4.7), i.e. let $i \notin J$. With the boundedness of H_i, the Jensen inequality for conditional expectations and the Itô isometry, it holds

$$\begin{aligned}
&\Big|\sum_{\mathcal{T}_n(t)} \mathbb{E}[Z_i(t_{k+1}) - Z_i(t_k)|\mathcal{G}^J_{t_k,t_{k+1}}]\Big| \\
&= \Bigg|\sum_{l=0}^{d-1}\sum_{\widetilde{\mathcal{T}}^l_n} \mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} H_i(u)\mathrm{d}M_i(u)\bigg|\mathcal{G}^J_{t_k,t_{k+1}}\bigg] + \sum_{l=1}^{d-1} \mathbb{E}\bigg[\int_{\underline{t}_l}^{\bar{t}_l} H_i(u)\mathrm{d}M_i(u)\bigg|\mathcal{G}^J_{\underline{t}_l,\bar{t}_l}\bigg] \\
&\quad - \sum_{l=1}^{d-1} \mathbb{E}\bigg[\int_{\underline{t}_l}^{u_l} H_i(u)\mathrm{d}M_i(u)\bigg|\mathcal{G}^J_{\underline{t}_l,u_l}\bigg] - \sum_{l=1}^{d-1} \mathbb{E}\bigg[\int_{u_l}^{\bar{t}_l} H_i(u)\mathrm{d}M_i(u)\bigg|\mathcal{G}^J_{u_l,\bar{t}_l}\bigg]\Bigg| \\
&\leqslant \sum_{l=0}^{d-1} C_i\Bigg|\sum_{\widetilde{\mathcal{T}}^l_n}\mathbb{E}\bigg[M_i(t_{k+1}) - M_i(t_k)\bigg|\mathcal{G}^J_{t_k,t_{k+1}}\bigg]\Bigg| + \Bigg|\sum_{l=1}^{d-1} \mathbb{E}\bigg[\int_{\underline{t}_l}^{\bar{t}_l} H_i(u)\mathrm{d}M_i(u)\bigg|\mathcal{G}^J_{\underline{t}_l,\bar{t}_l}\bigg]\Bigg| \\
&\quad + \Bigg|\sum_{l=1}^{d-1} \mathbb{E}\bigg[\int_{\underline{t}_l}^{u_l} H_i(u)\mathrm{d}M_i(u)\bigg|\mathcal{G}^J_{\underline{t}_l,u_l}\bigg]\Bigg| + \Bigg|\sum_{l=1}^{d-1} \mathbb{E}\bigg[\int_{u_l}^{\bar{t}_l} H_i(u)\mathrm{d}M_i(u)\bigg|\mathcal{G}^J_{u_l,\bar{t}_l}\bigg]\Bigg|.
\end{aligned}$$

The first addend tends to 0 in probability by (M) and Remark 4.2. For the second addend, Jensen's inequality, Itô isometry and the dominated convergence theorem yield

$$\begin{aligned} &\lim_{n\to\infty} \Big\| \sum_{l=1}^{d-1} \mathbb{E}\Big[\int_{\underline{t}_l}^{\bar{t}_l} H_i(u)\mathrm{d}M_i(u) \Big| \mathcal{G}^J_{\underline{t}_l,\bar{t}_l} \Big] \Big\|_2^2 = \lim_{n\to\infty} \sum_{l=1}^{d-1} \Big\| \mathbb{E}\Big[\int_{\underline{t}_l}^{\bar{t}_l} H_i(u)\mathrm{d}M_i(u) \Big| \mathcal{G}^J_{\underline{t}_l,\bar{t}_l} \Big] \Big\|_2^2 \\ &\leqslant \lim_{n\to\infty} \sum_{l=1}^{d-1} \mathbb{E}\Big[\Big(\int_{\underline{t}_l}^{\bar{t}_l} H_i(u)\mathrm{d}M_i(u) \Big)^2 \Big] = \lim_{n\to\infty} \sum_{l=1}^{d-1} \mathbb{E}\Big[\int_{\underline{t}_l}^{\bar{t}_l} H_i(u)^2 \mathrm{d}\langle M_i, M_i\rangle(u) \Big] = 0. \end{aligned} \tag{4.45}$$

For the last equation, we have used the continuity of $\int_0^{\cdot} H_i(u)^2 \mathrm{d}\langle M_i, M_i\rangle(u)$ that follows from the continuity of $\langle M_i, M_i\rangle(\cdot)$ (see Protter, 2005, Chapter IV, Theorem 8). With similar arguments, the last two addends tend to 0 in L_2 and therefore also in probability. For $i \in J$, we have

$$\begin{aligned} &\Big| \sum_{\mathcal{T}_n(t)} \mathbb{E}\Big[Z_i(t_{k+1}) - Z_i(t_k) \Big| \mathcal{G}^J_{t_k,t_{k+1}} \Big] - Z_i(t) \Big| \\ &\leqslant \Big| \sum_{l=0}^{d-1} \Big(\sum_{\tilde{\mathcal{T}}^l_n(l)} \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} H_i(u)\mathrm{d}M_i(u) \Big| \mathcal{G}^J_{t_k,t_{k+1}} \Big] - \int_{u_l\wedge t}^{u_{l+1}\wedge t} H_i(u)\mathrm{d}M_i(u) \Big) \Big| \\ &\quad + \Big| \sum_{l=1}^{d-1} \mathbb{E}\Big[\int_{\underline{t}_l}^{\bar{t}_l} H_i(u)\mathrm{d}M_i(u) \Big| \mathcal{G}^J_{\underline{t}_l,\bar{t}_l} \Big] \Big| + \Big| \sum_{l=1}^{d-1} \mathbb{E}\Big[\int_{\underline{t}_l}^{u_l} H_i(u)\mathrm{d}M_i(u) \Big| \mathcal{G}^J_{\underline{t}_l,u_l} \Big] \Big| \\ &\quad + \Big| \sum_{l=1}^{d-1} \mathbb{E}\Big[\int_{u_l}^{\bar{t}_l} H_i(u)\mathrm{d}M_i(u) \Big| \mathcal{G}^J_{u_l,\bar{t}_l} \Big] \Big| \\ &\leqslant \sum_{l=0}^{d-1} C_i \Big| \sum_{\tilde{\mathcal{T}}^l_n(l)} \mathbb{E}\Big[M_i(t_{k+1}) - M_i(t_k) \Big| \mathcal{G}^J_{t_k,t_{k+1}} \Big] - (M_i(u_{l+1}\wedge t) - M_i(u_l \wedge t)) \Big| \\ &\quad + \Big| \sum_{l=1}^{d-1} \mathbb{E}\Big[\int_{\underline{t}_l}^{\bar{t}_l} H_i(u)\mathrm{d}M_i(u) \Big| \mathcal{G}^J_{\underline{t}_l,\bar{t}_l} \Big] \Big| + \Big| \sum_{l=1}^{d-1} \mathbb{E}\Big[\int_{\underline{t}_l}^{u_l} H_i(u)\mathrm{d}M_i(u) \Big| \mathcal{G}^J_{\underline{t}_l,u_l} \Big] \Big| \\ &\quad + \Big| \sum_{l=1}^{d-1} \mathbb{E}\Big[\int_{u_l}^{\bar{t}_l} H_i(u)\mathrm{d}M_i(u) \Big| \mathcal{G}^J_{u_l,\bar{t}_l} \Big] \Big| \end{aligned}$$

The first addend tends to 0 in probability by (M) and Remark 4.2. For the other addends, we can use similar arguments as above (Jensen's inequality, Itô isometry and the dominated convergence theorem) to show that they tend to 0 in L_2 and therefore in probability.

Next, we consider the general case, where

$$Z_i(t) = \int_0^t H_i(u)\mathrm{d}M_i(u)$$

for a $\mathbb{G}$-predictable process H_i, that fulfils the assumptions. By Kuo (2006, Chapter 6), there exist simple processes $H_{i,c}$ with $H_{i,c} \cdot M \in \mathcal{M}_2$, such that

$$\lim_{c\to\infty} \big\| ((H_{i,c} - H_i) \cdot M_i)(t) \big\|_2 = \lim_{c\to\infty} \Big(\mathbb{E}\Big[\int_0^t (H_{i,c} - H_i)^2(u) \mathrm{d}\langle M_i, M_i\rangle(u) \Big] \Big)^{\frac{1}{2}} = 0. \tag{4.46}$$

Without loss of generality, we assume that $H_{i,c}$, $c \in \mathbb{N}$ are bounded simple processes. Otherwise, consider $H_{i,c,N}$ defined through $H_{i,c,N}(s) := H_{i,c}(s)\mathbb{1}_{\{H_{i,c}(s)\leqslant N\}}$ (see Protter, 2005, Chapter IV, Theorem 14). Then $H_{i,c,N}(s)$ tends to $H_{i,c}(s)$ almost surely as $N \to \infty$ for every $s \in (0,t]$. Thus, the dominated convergence theorem gives us

$$\lim_{N\to\infty} \left(\mathbb{E}\left[\int_0^t (H_{i,c,N} - H_{i,c})^2(u) \mathrm{d}\langle M_i, M_i\rangle(u) \right] \right)^{\frac{1}{2}} = 0$$

for every c. In particular, we find a subsequence $(N_c)_c$ such that

$$\lim_{c\to\infty} \left(\mathbb{E}\left[\int_0^t (H_{i,c,N_c} - H_i)^2(u) \mathrm{d}\langle M_i, M_i\rangle(u) \right] \right)^{\frac{1}{2}} = 0$$

using triangle inequality for the L_2-norm.

For $J \subseteq \{1,\ldots,m\}$ with $i \notin J$, it holds

$$\begin{aligned}
&\sum_{\mathcal{T}_n(t)} \mathbb{E}[Z_i(t_{k+1}) - Z_i(t_k) | \mathcal{G}^J_{t_k,t_{k+1}}] \\
&= \sum_{\mathcal{T}_n(t)} \mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} (H_i - H_{i,c})(u) \mathrm{d}M_i(u) \bigg| \mathcal{G}^J_{t_k,t_{k+1}} \bigg] + \sum_{\mathcal{T}_n(t)} \mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} H_{i,c}(u) \mathrm{d}M_i(u) \bigg| \mathcal{G}^J_{t_k,t_{k+1}} \bigg]
\end{aligned}$$

for every c. We consider the two summands separately. For the first summand, similar arguments like in (4.45) yield

$$\bigg\| \sum_{\mathcal{T}_n(t)} \mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} (H_i - H_{i,c})(u) \mathrm{d}M_i(u) \bigg| \mathcal{G}^J_{t_k,t_{k+1}} \bigg] \bigg\|_2^2 \leqslant \mathbb{E}\left[\int_0^t (H_i - H_{i,c})^2(u) \mathrm{d}\langle M_i, M_i\rangle(u) \right]$$

for every c. Thus, we get

$$\lim_{n\to\infty} \bigg\| \sum_{\mathcal{T}_n(t)} \mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} (H_i - H_{i,c})(u) \mathrm{d}M_i(u) \bigg| \mathcal{G}^J_{t_k,t_{k+1}} \bigg] \bigg\|_2^2 \leqslant \mathbb{E}\left[\int_0^t (H_i - H_{i,c})^2(u) \mathrm{d}\langle M_i, M_i\rangle(u) \right].$$

Since c was arbitrary, (4.46) yields

$$\lim_{n\to\infty} \bigg\| \sum_{\mathcal{T}_n(t)} \mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} (H_i - H_{i,c})(u) \mathrm{d}M_i(u) \bigg| \mathcal{G}^J_{t_k,t_{k+1}} \bigg] \bigg\|_2 = 0.$$

For the second summand, the first part of the proof yields

$$\operatorname*{plim}_{n\to\infty} \sum_{\in\mathcal{T}_n(t)} \mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} H_{i,c}(u) \mathrm{d}M_i(u) \bigg| \mathcal{G}^J_{t_k,t_{k+1}} \bigg] = 0.$$

In total, we get

$$\operatorname*{plim}_{n\to\infty} \sum_{\mathcal{T}_n(t)} \mathbb{E}[Z_i(t_{k+1}) - Z_i(t_k) | \mathcal{G}^J_{t_k,t_{k+1}}] = 0.$$

Now, let $J \subseteq \{1, \dots, m\}$ and $i \in J$. Then we have

$$\begin{aligned}
&\sum_{\mathcal{T}_n(t)} \mathbb{E}[Z_i(t_{k+1}) - Z_i(t_k) | \mathcal{G}^J_{t_k,t_{k+1}}] - Z_i(t) \\
&= \sum_{\mathcal{T}_n(t)} \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} (H_i - H_{i,c})(u) \mathrm{d}M_i(u) \Big| \mathcal{G}^J_{t_k,t_{k+1}} \Big] + \int_0^t (H_{i,c} - H_i)(u) \mathrm{d}M_i(u). \\
&\quad + \Big(\sum_{\mathcal{T}_n(t)} \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} H_{i,c}(u) \mathrm{d}M_i(u) \Big| \mathcal{G}^J_{t_k,t_{k+1}} \Big] - \int_0^t H_{i,c}(u) \mathrm{d}M_i(u) \Big).
\end{aligned}$$

The first part of the proof yields for the last term

$$\operatorname*{plim}_{n\to\infty} \Big(\sum_{\mathcal{T}_n(t)} \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} H_{i,c}(u) \mathrm{d}M_i(u) \Big| \mathcal{G}^J_{t_k,t_{k+1}} \Big] - \int_0^t H_{i,c}(u) \mathrm{d}M_i(u) \Big) = 0, \tag{4.47}$$

so we focus on the other two terms. With similar arguments as used in the case $i \in J$, it holds

$$\Big\| \sum_{\mathcal{T}_n(t)} \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} (H_i - H_{i,c})(u) \mathrm{d}M_i(u) \Big| \mathcal{G}^J_{t_k,t_{k+1}} \Big] \Big\|_2 \leqslant \Big\| \int_0^t (H_{i,c} - H_i)(u) \mathrm{d}M_i(u) \Big\|_2$$

for every $c \in \mathbb{N}$. Thus, we get

$$\begin{aligned}
&\lim_{n\to\infty} \Big\| \sum_{\mathcal{T}_n(t)} \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} (H_i - H_{i,c})(u) \mathrm{d}M_i(u) \Big| \mathcal{G}^J_{t_k,t_{k+1}} \Big] + \int_0^t (H_{i,c} - H_i)(u) \mathrm{d}M_i(u) \Big\|_2 \\
&\leqslant 2 \Big\| \int_0^t (H_{i,c} - H_i)(u) \mathrm{d}M_i(u) \Big\|_2.
\end{aligned}$$

Since c was arbitrary, we can conclude with (4.46) that

$$\lim_{n\to\infty} \Big\| \sum_{\mathcal{T}_n(t)} \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} (H_i - H_{i,c})(u) \mathrm{d}M_i(u) \Big| \mathcal{G}^J_{t_k,t_{k+1}} \Big] + \int_0^t (H_{i,c} - H_i)(u) \mathrm{d}M_i(u) \Big\|_2 = 0.$$

In total, we get the desired result

$$\operatorname*{plim}_{n\to\infty} \Big(\sum_{\mathcal{T}_n(t)} \mathbb{E}[Z_i(t_{k+1}) - Z_i(t_k) | \mathcal{G}^J_{t_k,t_{k+1}}] - Z_i(t) \Big) = 0.$$

□

This lemma directly leads us to the main result of this chapter.

Theorem 4.18. *Let the risk basis $X = (X_1, \dots, X_m)$ fulfil (S) and (M), such that $\langle M_i, M_i \rangle$, $i = 1, \dots, m$, are continuous processes. Suppose R has a martingale representation* (4.41) *for unique $\mathbb{G}$-predictable processes H_i, $i = 1, \dots, m$, that fulfil* (4.42). *Then $\varrho(X^t) = \mathbb{E}[\xi | \sigma(X^t)]$ admits the ISU decomposition*

$$D_i(t) = \int_0^t H_i(u) dM_i(u), \;\; t \in [0, T], \;\; i = 1, \dots, m.$$

In particular, the ISU decomposition does not depend on the update order or the choice of partitions.

Proof. Let $t \in [0,t]$ and let $(\mathcal{T}_n(t))_n$ be a sequence of partitions of $[0,t]$ with $|\mathcal{T}_n(t)| \to 0$, $n \to \infty$. For a better readability, this proof uses the short-hand notation $\sum_{\mathcal{T}_n(t)}$ instead of $\sum_{t_k,t_{k+1} \in \mathcal{T}_n(t)}$. Observe that it is sufficient to verify

$$\operatorname*{plim}_{n\to\infty} \sum_{\mathcal{T}_n(t)} (\mathbb{E}[\xi|\mathcal{G}^J_{t_k,t_{k+1}}] - \mathbb{E}[\xi|\mathcal{G}_{t_k}]) = \sum_{i\in J} \int_0^t H_i(u)\mathrm{d}M_i(u) \tag{4.48}$$

for all $J \subseteq \{1,\dots,m\}$. It holds

$$\begin{aligned}
&\sum_{\mathcal{T}_n(t)} (\mathbb{E}[\xi|\mathcal{G}^J_{t_k,t_{k+1}}] - \mathbb{E}[\xi|\mathcal{G}_{t_k}]) \\
&= \sum_{\mathcal{T}_n(t)} \sum_{i\in J} \mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} H_i(u)\mathrm{d}M_i(u) \bigg| \mathcal{G}^J_{t_k,t_{k+1}} \bigg] + \sum_{\mathcal{T}_n(t)} \sum_{i\notin J} \mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} H_i(u)\mathrm{d}M_i(u) \bigg| \mathcal{G}^J_{t_k,t_{k+1}} \bigg].
\end{aligned}$$

Applying Lemma 4.17, we can conclude that the first summand tends to $\sum_{i\in J} \int_0^t H_i(u)\mathrm{d}M_i(u)$ in probability and that the second summand tends to 0 in probability ($n \to \infty$), which gives the assertion. □

The theorem not only states the ISU decomposition of a martingale revaluation process R, but also immediately reveals its relation to the MRT decomposition, introduced by Schilling et al. (2020). Furthermore, the ISU decomposition does not depend on the update order, so the Theorem 2.7 helps us to derive the IOAT and AISU decomposition. The results are summarized in the following corollary.

Corollary 4.19. *Under the assumptions of Theorem 4.18, the ISU decomposition, the IOAT decomposition, the AISU decomposition and the MRT decomposition are equal.*

The fact that the ISU decomposition coincides with the MRT decomposition underlines the meaningfulness of the decomposition principle presented in this thesis. However, both decompositions require knowledge of the martingale representation, in particular its integrands. In its canonical form, the martingale representation theorem is a pure statement of existence, without an explicit specification of the integrands, see the references stated at the beginning of this section. Only in certain frameworks, (semi-)explicit forms of the integrands can be derived (see Cont & Lu, 2016) using the theory on Markov processes, the Malliavin calculus (see Malliavin, 1978; Nualart, 2006) or the Functional Itô calculus (see Cont & Fournié, 2013). Examples in the setup of diffusion processes are supplied by Cont and Lu (2016). Regarding marked point processes, we refer to the constructive proofs of Davis (1976) and Elliott (1976). In most of the cases, evaluating the integrals of the martingale represenation is cumbersome and often only possible numerically, even if there are supposedly explicit forms. With the ISU decomposition principle, a new door is open to approximate martingale representations via SU decompositions. In the following, we will investigate its numerical feasibility.

5 On the numerical feasibility of the ISU concept

In the previous chapters, the practical relevance of the ISU decomposition principle has been highlighted through its numerous applications inside and outside the life insurance context. In particular, it has been shown that established decomposition formulas can be obtained by applying the ISU decomposition principle in suitable models (see Chapter 3 and 4). However, from a practical point of view, the numerical feasibility of the ISU decomposition principle has yet to be verified. The present chapter thus examines the numerical feasibility of the ISU decomposition principle using a pension insurance example.

The numerical complexity of the ISU decomposition principle depends strongly on the model framework, the choice of the link mapping ϱ and the resulting form of the contributions. In Chapter 4, an analysis of ISU decompositions in a general martingale framework has been presented, that allows for applications beyond the life insurance context. More precisely, we have assumed that the revaluation surplus process R has the form $R(t) = \mathbb{E}[\xi|\sigma(X^t) \vee \mathcal{N}]$. This representation intuitively reflects the available information on the risk basis X until time t, and leads to a natural link mapping ϱ given by

$$\varrho(X_1^{t_1}, \ldots, X_m^{t_m}) = \mathbb{E}[\xi|\sigma(X_1^{t_1}, \ldots, X_m^{t_m}) \vee \mathcal{N}], \tag{5.1}$$

with $t_i \in [0, T]$, $i = 1, \ldots, m$. In contrast to the setup of traditional surplus decompositions considered in Chapter 3, this approach takes full advantage of the observed information, especially for estimating future values (see Chapter 4). Therefore, the numerical analyses in this chapter are embedded in the model framework of Chapter 4.

The present study considers a single fund-linked pension insurance that is driven by the systematic risks interest, fund and mortality, modelled as solutions of stochastic differential equations. The unknown time of death of the policyholder depicts the unsystematic risk, which is represented by a doubly stochastic Markov process with two states *alive* and *dead*. Within this setup, the ISU decomposition coincides with the MRT decomposition under certain conditions (see Section 4.2 and Theorem 4.18). A key assumption is the existence of the martingale representation, for which reference is made to the publication of Schilling et al. (2020). In that paper, a proof of the martingale representation theorem in a general doubly stochastic Markov model is provided. Moreover, (semi-)explicit representations for the integrands are derived using the Malliavin calculus (Malliavin, 1978) and the theory of Markov processes. Based on these representations, Schilling et al. (2020) simulated the various risk components in the example of a Guaranteed Minimum Death Benefit (GMDB) insurance, assuming affine processes for the systematic interest risk and the systematic mortality risk. Another example of numerically approximating the martingale representation in an actuarial framework can be found in Biagini et al. (2016).

In a doubly stochastic multistate Markov framework with affine intensity processes, Biagini et al. (2016) computed the martingale factors for an income protection insurance.

The ISU decomposition principle, presented in this thesis, opens up a new possibility to approach the martingale representation numerically. By definition, the ISU decomposition is a limit of SU decompositions, which, in the situation of Chapter 4, involve the sum of conditional expectations like (5.1). To simulate the latter, a multilevel Monte Carlo (MLMC) approach based on a contribution by Giles (2008) is used. In that paper, Giles (2008) investigated the computational complexity of the MLMC method for expected values of random variables that are driven by solutions of stochastic differential equations. In this thesis, the MLMC approach of Giles (2008) is generalised to conditional expectations and a systematic notion of MLMC convergence is introduced. A key component to obtain convergence results is the approximate solution of the stochastic differential equations via numerical schemes. Relying on the Euler approximation scheme (see Kloeden & Platen, 1992, Section 10.2), convergent MLMC estimators are derived for the SU decompositions of a single fund-linked pension insurance claim.

The MLMC estimators are implemented using the statistical software R 4.4.2 (R Core Team, 2024). The implementation is designed to allow for an analysis with a threefold focus. Firstly, the numerical feasibility and the question of whether the numerical approximation of the ISU decomposition is possible in an acceptable amount of time is investigated. Secondly, the impact of the chosen SU grid width on the approximation is reviewed and thirdly, the effect of the update order on the surplus decomposition is scrutinised. Taken together, these aspects will demonstrate the expediency of the ISU principle from a numerical point of view.

In Section 5.1, the model framework of the numerical example is introduced. The derived convergent MLMC estimators obtained by extending the MLMC theory of Giles (2008) to conditional expectations are presented in Section 5.2. The last part of this chapter, namely Section 5.3, is focused on the numerical implementation. First, the chosen parameters and the numerical methodology are described, which is followed by the presentation of the numerical results.

5.1 Model framework

We consider a single fund-linked pension insurance of an x-year old policyholder and assume a retirement age of $x + \gamma$, while the a maximum age is supposed to be $x + T$ years, where $T > \gamma$ and $\gamma, T \in \mathbb{N}$. The policy includes a one-off premium p at the start as well as regular premiums paid until the policyholder retires (or dies). A large part of the premiums is invested in a capital market fund. In return, the insurance company pays a death benefit for the case of the policyholder dying before retirement, otherwise a regular pension payment is disbursed until the end of the policyholder's life. If a pension is paid, it will be paid to the policyholder or the heirs for at least 10 years, no matter if the policyholder survives this period ('guaranteed pension period'). The pension level is linked to the price of the capital market fund.

In the following we introduce the underlying model, which is a special case of the doubly stochastic Markov setup that we have investigated in Subsection 4.2.5. We consider the time horizon $[0, T]$ covering the range from the initial age x to the maximum age $x + T$. The systematic risk factors are represented by a mortality intensity $\lambda = (\lambda(t))_{t \geqslant 0}$, a interest intensity $r = (r(t))_{t \geqslant 0}$ and a market fund $Y = (Y(t))_{t \geqslant 0}$. More precisely, we assume that the systematic risk drivers, which are composed to $\Theta = (\lambda, r, Y)$, solve the following stochastic differential equations

$$\begin{aligned} \mathrm{d}\lambda(s) &= \lambda(s)\mu_\lambda \mathrm{d}s + \sigma_\lambda \mathrm{d}W_\lambda(s), \ \lambda(0) = \lambda_0, \\ \mathrm{d}r(s) &= (\beta - r(s)\mu_r)\mathrm{d}s + \sigma_r \mathrm{d}W_r(s), \ r(0) = r_0, \\ \mathrm{d}Y(s) &= Y(s)\mu_Y \mathrm{d}s + Y(s)\sigma_Y \mathrm{d}W_Y(s), \ Y(0) = y_0, \end{aligned} \tag{5.2}$$

where μ_λ, μ_r, μ_Y, β, σ_λ, σ_r, σ_Y, r_0, λ_0, y_0 are positive constants and $W = (W_\lambda, W_r, W_Y)$ is a three-dimensional standard Brownian motion. That is, the process λ, which will later drive the underlying mortality intensity, is modelled by an Ornstein-Uhlenbeck process without mean reversion. Moreover, the interest intensity r is described by an Ornstein-Uhlenbeck process with mean reversion (see Ahmad et al., 2022), while the market fund Y follows a geometric Brownian motion with positive drift.

Rather technically motivated, we introduce the functions g_λ, g_r and g_Y from $\mathbb{R}$ to $\mathbb{R}$, which will provide the necessary regularity of the risk drivers λ, r and Y to conclude the desired convergence. We assume that

- g_λ, g_r and g_Y are Lipschitz continuous,
- g_λ has values in $[0, B_\lambda]$ for some $B_\lambda > 0$,
- g_r has values in $[-B_r, B_r]$ for some $B_r > 0$,
- g_Y has values in $[B_Y, +\infty)$ for some $B_Y > 0$.

The constraints for g_λ ensure the non-negativity of the mortality intensity. Apart from that, g_λ, g_r and g_Y can be chosen to be close to the identity, with large parameters B_λ and B_r and a small parameter B_Y, so that these transformations have no practical effect in applications.

We model the state process of the insured by an $\mathbb{F}^\Theta$-conditional Markov chain $Z = (Z(t))_{t \in [0,T]}$ with state space $\mathcal{Z} = \{a, d\}$ and a corresponding right-continuous jump process N that jumps from 0 to 1 at time of death (see Subsection 4.2.5). Given a grid

$$\pi^\lambda = \{0 = u_0 < \ldots < u_d = T\},$$

the jump intensity process $\overline{\lambda} = (\overline{\lambda}(t))_{t\in[0,T]}$ is defined by $\overline{\lambda}(0) = g_\lambda(0)$ and $\overline{\lambda}(t) = g_\lambda(\lambda(u_l))$, $t \in (u_l, u_{l+1}]$. In particular, the process $M_N = (M_N(t))_{t\geqslant 0}$, specified by

$$M_N(t) = N(t) - \int_0^t I(s)\overline{\lambda}(s)\mathrm{d}s,$$

is assumed to be a $\overline{\mathbb{G}}$-martingale, where $\overline{\mathbb{G}} = (\overline{\mathcal{G}}_t)_{t\geqslant 0}$ is the filtration given by $\overline{\mathcal{G}}_t = \mathcal{F}_T^\Theta \vee \mathcal{F}_t^N$ (see Subsection 4.2.5). We define the indicator process $I = (I(t))_{t\geqslant 0}$ by $I(t) = \mathbb{1}_{\{Z(t)=a\}}$. Moreover, the survival probability $p_{aa}(s,t)$, $s \leqslant t$, has the representation

$$p_{aa}(s,t) = \exp\left(-\int_s^t \overline{\lambda}(s)\mathrm{d}s\right),$$

see Subsection 4.2.5.

Let the risk-free bank account $K = (K(t))_{t\geqslant 0}$ satisfy

$$\mathrm{d}K(t) = K(t)g_r(r(t))\mathrm{d}t, \ K(0) = 1.$$

For the discounting of cashflows, we introduce the discount factors $v(s,t)$, $0 \leqslant s \leqslant t$, by

$$v(s,t) = \frac{K(s)}{K(t)} = \exp\left(-\int_s^t g_r(r(u))\mathrm{d}u\right).$$

It is assumed that a deterministic savings rate, given by a bounded Lipschitz continuous function $a = (a(t))_{t\in[0,T]}$, is invested into Y. The shares held are described by a stochastic process $Q = (Q(t))_{t\in[0,T]}$, defined through

$$\mathrm{d}Q(s) = \frac{a(s)}{g_Y(Y(s))}\mathrm{d}s, \quad Q(0) = a(0). \tag{5.3}$$

Here, $a(0) \in [0,p]$ is the amount of shares that is bought with a one-off payment at the beginning of the contract. In the following, the process $V = (V(t))_{t\in[0,T]}$ with $V(t) = Q(t)g_Y(Y(t))$ represents the deposit value of the policyholder.

The insurance claim is given by

$$\xi = p + \xi_1 + \xi_2 + \xi_3 + \xi_4,$$

where

$$\begin{aligned}
\xi_1 &= I(\gamma)v(0,\gamma)V(\gamma),\\
\xi_2 &= \int_0^\gamma v(0,s)(1-f_d)V(s)\mathrm{d}N(s),\\
\xi_3 &= -I(\gamma)\int_\gamma^{\gamma+10} v(0,s)f_pV(\gamma)\mathrm{d}s,\\
\xi_4 &= -\int_{\gamma+10}^T I(s)v(0,s)f_pV(\gamma)\mathrm{d}s.
\end{aligned}$$

The insurance claim includes a one-off premium p, which covers the initial investment $a(0)$. The addend ξ_1 comprises the discounted deposit value at retirement resulting from the savings process a. If the policyholder dies before retirement, a death benefit $f_d \cdot V(s)$ with rate $f_d \in (0,1)$ is paid out, while the residual deposit value remains at the insurer. The latter cashflow is depicted by ξ_2. With $f_p \in (0,1)$, we denote the pension factor, i.e. the part of the depository that is yearly disbursed. The cashflow ξ_3 models the guaranteed pension period of ten years, if the policyholder reaches the retirement age. The fourth part ξ_4 represents the pension payment after the guaranteed pension period. The assumptions made in this paragraph imply $\mathbb{E}[\xi^2] < \infty$.

Throughout this chapter, the risk basis is given by $X = (N, \lambda, r, Y)$ and, as usual, its information flow is denoted by $\mathbb{G} = (\mathcal{G}_t)_{t \geqslant 0}$. The goal is to numerically approximate the ISU decomposition of the revaluation process $R = (R(t))_{t \geqslant 0}$ given by

$$R(t) = \mathbb{E}[\xi | \mathcal{G}_t], \; t \geqslant 0,$$

with the link mapping ϱ defined through

$$\varrho(X_1^{t_1}, \ldots, X_m^{t_m}) = \mathbb{E}[\xi | \sigma(X_1^{t_1}, \ldots, X_m^{t_m}) \vee \mathcal{N}], \tag{5.4}$$

for $t_i \in [0, T]$, $i = 1, \ldots, m$ (see Chapter 4). Since ξ is square-integrable, and since Θ and W generate the same information, Proposition 1 in Schilling et al. (2020) ensures the existence of a martingale representation. Furthermore, by Lemma 4.16, the risk basis X fulfils property (M), which has been crucial in Chapter 4 for the derivation of ISU decompositions. Thus, we know from Theorem 4.18 that the ISU decomposition exists. In particular, it coincides with the MRT decomposition.

The idea of this chapter is to approximate the ISU decomposition by means of SU decompositions. Therefore, let $\mathcal{T} = \{0 = t_0 < \ldots < t_n = t\}$ be a partition of $[0, t]$. We assume that the grid π^λ of the mortality intensity is included in $\mathcal{T}$, i.e. let $\pi^\lambda \subseteq \mathcal{T}$. Recall from Chapter 4, that $\mathbb{G}^J$ describes the complete natural filtration generated by the subset $J \subseteq \{N, \lambda, r, Y\}$ of risk sources, and $\mathcal{G}_{s,t}^J = \mathcal{G}_s \vee \mathcal{G}_t^J$. For better readability, we omit the curly brackets in J. The SU decomposition with respect to $\mathcal{T}$ is given by

$$\begin{aligned}
D_N(t) &= \sum_{k=0}^{n-1} (\mathbb{E}[\xi | \mathcal{G}_{t_k, t_{k+1}}^N] - \mathbb{E}[\xi | \mathcal{G}_{t_k, t_{k+1}}^\varnothing]), \\
D_\lambda(t) &= \sum_{k=0}^{n-1} (\mathbb{E}[\xi | \mathcal{G}_{t_k, t_{k+1}}^{N,\lambda}] - \mathbb{E}[\xi | \mathcal{G}_{t_k, t_{k+1}}^N]), \\
D_r(t) &= \sum_{k=0}^{n-1} (\mathbb{E}[\xi | \mathcal{G}_{t_k, t_{k+1}}^{N,\lambda,r}] - \mathbb{E}[\xi | \mathcal{G}_{t_k, t_{k+1}}^{N,\lambda}]), \\
D_Y(t) &= \sum_{k=0}^{n-1} (\mathbb{E}[\xi | \mathcal{G}_{t_k, t_{k+1}}^{N,\lambda,r,Y}] - \mathbb{E}[\xi | \mathcal{G}_{t_k, t_{k+1}}^{N,\lambda,r}]),
\end{aligned}$$

see Chapter 2. In general, conditional expectations cannot necessarily be computed analytically. In that case, a method to numerically approximate conditional expectations is needed. Inspired by Giles (2008), this chapter presents a general MLMC approach to approximate the conditional expectations of the SU decompostion. Since the SU decomposition itself is already an approximation of the ISU decomposition, the multilevel Monte Carlo approach represents a further level of approximation. The convergence of the SU decompositions towards the ISU decomposition has been established in Chapter 4. For the approach of the conditional expectations with the MLMC methods, the theoretical basis still has to be laid out. This is therefore the focus of the next section.

5.2 Numerical method

The goal of this section is to derive convergent estimators for the contributions of the SU decomposition. Before introducing the multilevel Monte Carlo methods, we study the SU decomposition in more detail.

5.2.1 Analysis of the SU decomposition

Let $\mathbb{H} = (\mathcal{H}_t)_{t \geqslant 0}$ denote the completed natural filtration generated by the systematic risks Θ. For $t \in (u_l, u_{l+1}]$, we define $u(t) = u_{l+1}$. Furthermore, for $s > u(t_k)$, $k = 0, \ldots, n$, we set

$$p_u^{\mathbb{E}}(t_k, s) = \mathbb{E}[p_{aa}(u(t_k), s)|\mathcal{H}_{t_k}], \tag{5.5}$$

$$p_{u\lambda}^{\mathbb{E}}(t_k, s) = \mathbb{E}[p_{aa}(u(t_k), s)\overline{\lambda}(s)|\mathcal{H}_{t_k}], \tag{5.6}$$

as well as

$$v^{\mathbb{E}}(t_k, s) = \mathbb{E}[v(t_k, s)|\mathcal{H}_{t_k}], \tag{5.7}$$

$$Y^{\mathbb{E}}(t_k, s) = \mathbb{E}[g_Y(Y(s))|\mathcal{H}_{t_k}], \tag{5.8}$$

$$Y_Q^{\mathbb{E}}(t_k, s) = \mathbb{E}[(Q(s) - Q(t_k))g_Y(Y(s))|\mathcal{H}_{t_k}] \tag{5.9}$$

for $s > t_k$. This leads us to the following short-hand notations

$$\begin{aligned} p^{\mathbb{E}}(t_k, s) &= p(t_k, u(t_k))p_u^{\mathbb{E}}(t_k, s), \\ p_\lambda^{\mathbb{E}}(t_k, s) &= p(t_k, u(t_k))p_{u\lambda}^{\mathbb{E}}(t_k, s), \\ V^{\mathbb{E}}(t_k, s) &= Q(t_k)Y^{\mathbb{E}}(t_k, s) + Y_Q^{\mathbb{E}}(t_k, s), \end{aligned}$$

for $s > t_k$, $k = 0, \ldots, n$, and

$$\begin{aligned} p_+^{\mathbb{E}}(t_k, s) &= p(t_{k+1}, u(t_k))p_u^{\mathbb{E}}(t_k, s), \\ p_{\lambda,+}^{\mathbb{E}}(t_k, s) &= p(t_{k+1}, u(t_k))p_{u\lambda}^{\mathbb{E}}(t_k, s) \end{aligned}$$

for $s > t_{k+1}$, $k = 0, \dots, n-1$. In the following, we provide compact formulas for the SU contributions. For the different claim components i, the SU contributions assume the following form

$$\begin{aligned}
D_N^i(t) &= \sum_{k=0}^{n-1} (\mathbb{E}[\xi_i|\mathcal{G}_{t_k,t_{k+1}}^N] - \mathbb{E}[\xi_i|\mathcal{G}_{t_k,t_{k+1}}^\varnothing]),\\
D_\lambda^i(t) &= \sum_{k=0}^{n-1} (\mathbb{E}[\xi_i|\mathcal{G}_{t_k,t_{k+1}}^{N,\lambda}] - \mathbb{E}[\xi_i|\mathcal{G}_{t_k,t_{k+1}}^N]),\\
D_r^i(t) &= \sum_{k=0}^{n-1} (\mathbb{E}[\xi_i|\mathcal{G}_{t_k,t_{k+1}}^{N,\lambda,r}] - \mathbb{E}[\xi_i|\mathcal{G}_{t_k,t_{k+1}}^{N,\lambda}]),\\
D_Y^i(t) &= \sum_{k=0}^{n-1} (\mathbb{E}[\xi_i|\mathcal{G}_{t_k,t_{k+1}}^{N,\lambda,r,Y}] - \mathbb{E}[\xi_i|\mathcal{G}_{t_k,t_{k+1}}^{N,\lambda,r}]).
\end{aligned}$$

Clearly, we have

$$D_j(t) = \sum_{i=1}^{4} D_j^i(t), \ \ j \in \{N, \lambda, r, Y\}.$$

After some tedious calculations (see Appendix A.3), one derives the following representations for the SU contributions.

Unsystematic biometric risk

The SU contributions with respect to N are given by

$$\begin{aligned}
D_N^1(t) &= \sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k<\gamma}} \mathbb{1}_{\{t_k<\gamma\}} v(0,t_k)\Delta I^p(t_k,t_{k+1},\gamma) v^{\mathbb{E}}(t_k,\gamma) V^{\mathbb{E}}(t_k,\gamma),\\
D_N^2(t) &= \sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k<\gamma}} v(0,t_k) \int_{t_k}^{t_{k+1}} v^{\mathbb{E}}(t_k,s)(1-f_d)V^{\mathbb{E}}(t_k,s)(\mathrm{d}N(s) - I(t_k)p_\lambda^{\mathbb{E}}(t_k,s)\mathrm{d}s)\\
&\quad + \sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k<\gamma}} v(0,t_k) \int_{t_{k+1}}^{\gamma} v^{\mathbb{E}}(t_k,s)(1-f_d)V^{\mathbb{E}}(t_k,s)\Delta I_\lambda^p(t_k,t_{k+1},s)\mathrm{d}s,\\
D_N^3(t) &= -\sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k<\gamma}} v(0,t_k)\Delta I^p(t_k,t_{k+1},\gamma) \int_{\gamma}^{\gamma+10} v^{\mathbb{E}}(t_k,s) f_p V^{\mathbb{E}}(t_k,\gamma)\mathrm{d}s,\\
D_N^4(t) &= -\sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k<\gamma+10}} v(0,t_k) \int_{\gamma+10}^{T} \Delta I^p(t_k,t_{k+1},s) v^{\mathbb{E}}(t_k,s) f_p V^{\mathbb{E}}(t_k,\gamma)\mathrm{d}s\\
&\quad - \sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k\geqslant\gamma+10}} v(0,t_k) \int_{t_k}^{t_{k+1}} (I(s) - I(t_k)p^{\mathbb{E}}(t_k,s)) v^{\mathbb{E}}(t_k,s) f_p V(\gamma)\mathrm{d}s\\
&\quad - \sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k\geqslant\gamma+10}} v(0,t_k) \int_{t_{k+1}}^{T} \Delta I^p(t_k,t_{k+1},s) v^{\mathbb{E}}(t_k,s) f_p V(\gamma)\mathrm{d}s,
\end{aligned}$$

where

$$\Delta I^p(t_k, t_{k+1}, s) = I(t_{k+1}) p_+^{\mathbb{E}}(t_k, s) - I(t_k) p^{\mathbb{E}}(t_k, s),$$

$$\Delta I_\lambda^p(t_k, t_{k+1}, s) = I(t_{k+1}) p_{\lambda,+}^{\mathbb{E}}(t_k, s) - I(t_k) p_\lambda^{\mathbb{E}}(t_k, s).$$

Systematic biometric risk

The SU contributions with respect to λ are given by

$$\begin{aligned}
D_\lambda^1(t) &= \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_{k+1}) v(0, t_k) \Delta p(t_k, t_{k+1}, \gamma) v^{\mathbb{E}}(t_k, \gamma) V^{\mathbb{E}}(t_k, \gamma), \\
D_\lambda^2(t) &= \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_{k+1}) v(0, t_k) \int_{t_{k+1}}^{\gamma} v^{\mathbb{E}}(t_k, s)(1 - f_d) V^{\mathbb{E}}(t_k, s) \Delta p_\lambda(t_k, t_{k+1}, s) \mathrm{d}s, \\
D_\lambda^3(t) &= - \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_{k+1}) \Delta p(t_k, t_{k+1}, \gamma) v(0, t_k) \int_{\gamma}^{\gamma+10} v^{\mathbb{E}}(t_k, s) f_p V^{\mathbb{E}}(t_k, \gamma) \mathrm{d}s, \\
D_\lambda^4(t) &= - \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma+10}} I(t_{k+1}) v(0, t_k) \int_{\gamma+10}^{T} \Delta p(t_k, t_{k+1}, s) v^{\mathbb{E}}(t_k, s)) f_p V^{\mathbb{E}}(t_k, \gamma) \mathrm{d}s \\
&\quad - \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k \geqslant \gamma+10}} I(t_{k+1}) v(0, t_k) \int_{t_{k+1}}^{T} \Delta p(t_k, t_{k+1}, s) v^{\mathbb{E}}(t_k, s) f_p V(\gamma) \mathrm{d}s,
\end{aligned}$$

where

$$\Delta p(t_k, t_{k+1}, s) = p^{\mathbb{E}}(t_{k+1}, s) - p_+^{\mathbb{E}}(t_k, s),$$

$$\Delta p_\lambda(t_k, t_{k+1}, s) = p_\lambda^{\mathbb{E}}(t_{k+1}, s) - p_{\lambda,+}^{\mathbb{E}}(t_k, s).$$

Systematic interest risk

The SU contributions with respect to r are given by

$$\begin{aligned}
D_r^1(t) &= \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_{k+1}) p^{\mathbb{E}}(t_{k+1}, \gamma) \Delta v(t_k, t_{k+1}, \gamma) V^{\mathbb{E}}(t_k, \gamma), \\
D_r^2(t) &= \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} v(0, t_k) \int_{t_k}^{t_{k+1}} (v(t_k, s) - v^{\mathbb{E}}(t_k, s))(1 - f_d) V^{\mathbb{E}}(t_k, s) \mathrm{d}N(s) \\
&\quad + \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_{k+1}) \int_{t_{k+1}}^{\gamma} \Delta v(t_k, t_{k+1}, s)(1 - f_d) V^{\mathbb{E}}(t_k, s) p_\lambda^{\mathbb{E}}(t_{k+1}, s) \mathrm{d}s,
\end{aligned}$$

$$
\begin{aligned}
D_r^3(t) = &- \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_{k+1}) p^{\mathbb{E}}(t_{k+1}, \gamma) \int_{\gamma}^{\gamma+10} \Delta v(t_k, t_{k+1}, s) f_p V^{\mathbb{E}}(t_k, \gamma) \mathrm{d}s \\
&- \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k \geqslant \gamma}} I(\gamma) f_p V(\gamma) \int_{t_k}^{t_{k+1}} (v(0,s) - v(0,t_k) v^{\mathbb{E}}(t_k, s)) \mathrm{d}s \\
&- \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k \geqslant \gamma}} I(\gamma) f_p V(\gamma) \int_{t_{k+1}}^{\gamma+10} \Delta v(t_k, t_{k+1}, s) \mathrm{d}s, \\
D_r^4(t) = &- \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma+10}} I(t_{k+1}) \int_{\gamma+10}^{T} p^{\mathbb{E}}(t_{k+1}, s) \Delta v(t_k, t_{k+1}, s) f_p V^{\mathbb{E}}(t_k, \gamma) \mathrm{d}s \\
&- \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k \geqslant \gamma+10}} \int_{t_k}^{t_{k+1}} I(s)(v(0,s) - v(0,t_k) v^{\mathbb{E}}(t_k, s)) f_p V(\gamma) \mathrm{d}s \\
&- \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k \geqslant \gamma+10}} I(t_{k+1}) \int_{t_{k+1}}^{T} p^{\mathbb{E}}(t_{k+1}, s) \Delta v(t_k, t_{k+1}, s) f_p V(\gamma) \mathrm{d}s,
\end{aligned}
$$

where

$$
\Delta v(t_k, t_{k+1}, s) = v(0, t_{k+1}) v^{\mathbb{E}}(t_{k+1}, s) - v(0, t_k) v^{\mathbb{E}}(t_k, s).
$$

Systematic fund risk

The SU contributions with respect to Y are given by

$$
\begin{aligned}
D_Y^1(t) = &\sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_{k+1}) p^{\mathbb{E}}(t_{k+1}, \gamma) v^{\mathbb{E}}(t_{k+1}, \gamma) \Delta V(t_k, t_{k+1}, \gamma), \\
D_Y^2(t) = &\sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} \int_{t_k}^{t_{k+1}} v(0,s)(1 - f_d)(V(s) - V^{\mathbb{E}}(t_k, s)) \mathrm{d}N(s) \\
&+ \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_{k+1}) v(0, t_{k+1}) \int_{t_{k+1}}^{\gamma} v^{\mathbb{E}}(t_{k+1}, s)(1 - f_d) \Delta V(t_k, t_{k+1}, s) p_{\lambda}^{\mathbb{E}}(t_{k+1}, s) \mathrm{d}s, \\
D_Y^3(t) = &- \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_{k+1}) v(0, t_{k+1}) p^{\mathbb{E}}(t_{k+1}, \gamma) \int_{\gamma}^{\gamma+10} v^{\mathbb{E}}(t_{k+1}, s) f_p \Delta V(t_k, t_{k+1}, \gamma) \mathrm{d}s, \\
D_Y^4(t) = &- \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_{k+1}) v(0, t_{k+1}) \int_{\gamma+10}^{T} p^{\mathbb{E}}(t_{k+1}, s) v^{\mathbb{E}}(t_{k+1}, s) f_p \Delta V(t_k, t_{k+1}, \gamma) \mathrm{d}s,
\end{aligned}
$$

where

$$
\Delta V(t_k, t_{k+1}, s) = V^{\mathbb{E}}(t_{k+1}, s) - V^{\mathbb{E}}(t_k, s).
$$

The formulas above reveal that each SU addend can be represented by pathwise Lebesgue integrals, where the randomness of the integrands stems from the risk factors and conditional expectations thereof like (5.5) - (5.9). Thus, our idea is as follows. We consider a fixed path for the risk factors and approximate the integrals pathwise by finite sums. The appearing conditional expectations, given the observed paths of the risk factors, are then approximated with MLMC methods. In the following, we develop a theoretical fundament to this approach.

5.2.2 Multilevel Monte Carlo estimators

In this section, we take a step back and introduce MLMC estimators in a general framework. The MLMC approach can be traced back to various papers by Heinrich (1998, 2000, 2001) and Heinrich and Sindambiwe (1999) on parametric integration and the solution of integral equations. A good overview about the MLMC approach and its applications can be found in Giles (2015). The definition of MLMC estimators and the basic notation follows the theory of Giles (2008), who laid the cornerstone for an application of the MLMC approach to estimate the expectation of random variables driven by stochastic differential equations. In extension to Giles (2008), we study MLMC estimators for *conditional* expectations and establish a systematic notion of (integral) MLMC convergence. For this thesis, focus is placed on convergence results, but not on the computational complexity of the approximation. Nevertheless, the basic notation and the convergence result is oriented towards Giles (2008).

Let $(\Omega, \mathcal{A}, \mathbb{F}, \mathbb{P})$ be a filtered probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geqslant 0}$ that satisfies the usual conditions. Furthermore, let W denote a standard m-dimensional Brownian motion that is independent from $\mathcal{F}_0$, and let S be the solution to the m-dimensional stochastic differential equation

$$\mathrm{d}S(t) = \mu(t, S(t))\mathrm{d}t + \sigma(t, S(t))\mathrm{d}W(t), \ S(0) = S_0,$$

where $\mu\colon [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$, $\sigma\colon [0,T] \times \mathbb{R}^m \to \mathbb{R}^{m \times m}$ are functions and S_0 is $\mathcal{F}_0$-measurable. The goal of this section is to find an appropriate estimator for

$$\mathfrak{E} = \mathbb{E}[f(S(t_1), .., S(t_r))|\mathcal{F}_0], \ 0 \leqslant t_1 < \ldots < t_r \leqslant T,$$

where f is any appropriate function. In the following, the abbreviation $P = f(S(t_1), .., S(t_r))$ is used.

Let $L \in \mathbb{N}$, $K \in \mathbb{R}_+$ and let S_l denote the Euler approximation of S with timestep[1] $h_l = 2^{-l}T/K$, $l = 0, \ldots, L$, i.e.

$$\begin{aligned}
&S_l(0) = S_0, \\
&S_l(u_{j+1}) = S_l(u_j) + \mu(u_j, S_l(u_j))(u_{j+1} - u_j) + \sigma(u_j, S_l(u_j))(W(u_{j+1}) - W(u_j)),
\end{aligned}$$

where $u_j = jh_l$, $j = 0, \ldots, 2^l K$. For a better readability, the index l in u_j is omitted.

[1]The definition of the step width slightly differs from Giles (2008), who uses $h_l = M^{-l}T$ for $M \in \mathbb{N}$, $M \geqslant 2$.

In line with Giles (2008), we set $P_l = f(S_l(t_1), \dots, S_l(t_r))$ and

$$\widehat{\mathfrak{E}}_l = N_l^{-1} \sum_{i=1}^{N_l} (\widehat{P}_l^i - \widehat{P}_{l-1}^i), \; l = 1, \dots, L, \tag{5.10}$$

where N_l is the number of independent samples and $\widehat{P}_l^i = f(\widehat{S}_l^i(t_1), .., \widehat{S}_l^i(t_r))$ with $\widehat{S}_l^i$ denoting a sample Euler approximation path. Following Giles (2008), the difference $\widehat{P}_l^i - \widehat{P}_{l-1}^i$ is based on two Euler approximations with different step lengths but with the same Brownian motion path. Furthermore, let $\widehat{\mathfrak{E}}_0$ be the usual Monte Carlo estimator for $\mathbb{E}[f(S(t_1), .., S(t_r))|\mathcal{F}_0]$ using N_0 samples for the Euler approximation with step width h_0, i.e.

$$\widehat{\mathfrak{E}}_0 = N_0^{-1} \sum_{i=1}^{N_0} \widehat{P}_0^i.$$

The combined estimator $\widehat{\mathfrak{E}}$ is given by

$$\widehat{\mathfrak{E}} = \sum_{l=0}^{L} \widehat{\mathfrak{E}}_l, \tag{5.11}$$

see Giles (2008). Let us now introduce the notion of MLMC convergence.

Definition 5.1. A family of random variables $\widehat{H} = \widehat{H}(L, (N_l)_l)$, $l = 0, \dots, L$, $L \in \mathbb{N}_0$ is called a *MLMC estimator* for a random variable H, if for any $\varepsilon > 0$, there exist L and N_l, $l = 0, \dots, L$ such that,

$$MSE = \mathbb{E}[(\widehat{H} - H)^2] < \varepsilon^2. \tag{5.12}$$

We write briefly $\widehat{H} \overset{MLMC}{\Rightarrow} H$. Furthermore, if the following two conditions are fulfilled, namely

(i) for every $\widetilde{L} > L$, there exist $(\widetilde{N}_l)_{l=0,\dots,\widetilde{L}}$ such that it still holds (5.12),

(ii) for every tupel $(\widetilde{N}_l)_{l=0,\dots,L}$ with $\widetilde{N}_l \geqslant N_l$, $l = 0, \dots, L$, it still holds (5.12),

we call the convergence *monotonic*.

By extending the ideas of Giles (2008) to our framework, we can state assumptions, under which $\widehat{\mathfrak{E}}$ converges to $\mathfrak{E}$ in terms of Definition 5.1.

Theorem 5.2. *If there exist constants $\alpha \geqslant \frac{1}{2}$, $\beta > 0$, and positive functions c_1, c_2, such that $\mathbb{E}[c_1(S_0)] < \infty$, $\mathbb{E}[c_2(S_0)] < \infty$ and*

(i) $\mathbb{E}[\widehat{P}_l - P|\mathcal{F}_0] \leqslant c_1(S_0) h_l^\alpha$,

(ii) $\mathrm{Var}[\widehat{\mathfrak{E}}_l|\mathcal{F}_0] \leqslant c_2(S_0) N_l^{-1} h_l^\beta$,

then it holds $\widehat{\mathfrak{E}} \overset{MLMC}{\Rightarrow} \mathbb{E}[f(S(t_1), .., S(t_r))|\mathcal{F}_0]$ monotonically.

Proof. The proof follows the lines from Giles (2008, Theorem 3.1), adapted to our case. Let $\varepsilon > 0$. We decompose the MSE into bias and variance, i.e.

$$\mathbb{E}[(\widehat{\mathfrak{C}} - \mathfrak{C})^2|\mathcal{F}_0] = \mathbb{E}[\widehat{\mathfrak{C}} - \mathfrak{C}|\mathcal{F}_0]^2 + \mathrm{Var}[\widehat{\mathfrak{C}}|\mathcal{F}_0],$$

and investigate both addends separately. We set

$$L = \left\lceil \frac{\log(\sqrt{2\mathbb{E}[c_1(S_0)]}(T/K)^\alpha \varepsilon^{-1})}{\alpha \log(2)} \right\rceil,$$

where $\lceil \cdot \rceil$ is the ceiling function. Using that $x \leqslant \lceil x \rceil < x + 1$, $x \in \mathbb{R}$, we immediately get

$$h_L^\alpha \leqslant \frac{1}{\sqrt{2\mathbb{E}[c_1(S_0)]}}\varepsilon.$$

That gives us, together with the definition of $\widehat{\mathfrak{C}}$ and assumption (i),

$$\begin{aligned}
\mathbb{E}[\widehat{\mathfrak{C}} - \mathfrak{C}|\mathcal{F}_0]^2 &= \mathbb{E}\left[\frac{1}{N_0}\sum_{i=1}^{N_0} \widehat{P}_0^i + \sum_{l=1}^{L} \frac{1}{N_l}\sum_{i=1}^{N_l}(\widehat{P}_l^i - \widehat{P}_{l-1}^i) - P \middle| \mathcal{F}_0 \right]^2 \\
&= \mathbb{E}\left[P_L - P|\mathcal{F}_0\right]^2 \leqslant (c_1(S_0)h_L^\alpha)^2 \leqslant \frac{c_1(S_0)}{2\mathbb{E}[c_1(S_0)]}\varepsilon^2.
\end{aligned}$$

For the variance, we consider different cases for β (see assumption (ii)). If assumption (ii) holds for $\beta = 1$, let $N_l = \lceil 2\varepsilon^{-2}(L+1)\mathbb{E}[c_2(S_0)]h_l \rceil$ yielding

$$\mathrm{Var}[\widehat{\mathfrak{C}}|\mathcal{F}_0] = \sum_{l=0}^{L} \mathrm{Var}[\widehat{\mathfrak{C}}_l|\mathcal{F}_0] \leqslant \sum_{l=0}^{L} c_2(S_0)N_l^{-1}h_l \leqslant \frac{c_2(S_0)}{2\mathbb{E}[c_2(S_0)]}\varepsilon^2.$$

If $\beta > 1$, let $N_l = \lceil 2\varepsilon^{-2}\mathbb{E}[c_2(S_0)](T/K)^{(\beta-1)/2}(1 - 2^{-(\beta-1)/2})^{-1}h_l^{(\beta+1)/2} \rceil$. That gives

$$\begin{aligned}
\mathrm{Var}[\widehat{\mathfrak{C}}|\mathcal{F}_0] &= \sum_{l=0}^{L} \mathrm{Var}[\widehat{\mathfrak{C}}_l|\mathcal{F}_0] \leqslant \sum_{l=0}^{L} c_2(S_0)N_l^{-1}h_l^\beta \\
&\leqslant \frac{c_2(S_0)}{2\mathbb{E}[c_2(S_0)]}\varepsilon^2 \left(\frac{T}{K}\right)^{-(\beta-1)/2} (1 - 2^{-(\beta-1)/2}) \sum_{l=0}^{L} h_l^{(\beta-1)/2}.
\end{aligned}$$

Using the formula for the geometric series, we can estimate the last term by

$$\begin{aligned}
\sum_{l=0}^{L} h_l^{(\beta-1)/2} &= \left(\frac{T}{K}\right)^{(\beta-1)/2} \sum_{l=0}^{L} (2^{-(\beta-1)/2})^l \\
&= \left(\frac{T}{K}\right)^{(\beta-1)/2} \frac{1 - (2^{-(\beta-1)/2})^{L+1}}{1 - 2^{-(\beta-1)/2}} \\
&< \left(\frac{T}{K}\right)^{(\beta-1)/2} \frac{1}{1 - 2^{-(\beta-1)/2}},
\end{aligned}$$

which in total leads to

$$\mathrm{Var}[\widehat{\mathfrak{C}}|\mathcal{F}_0] \leqslant \frac{c_2(S_0)}{2\mathbb{E}[c_2(S_0)]}\varepsilon^2.$$

If $\beta < 1$, let $N_l = \lceil 2\varepsilon^{-2}\mathbb{E}[c_2(S_0)]h_L^{-(1-\beta)/2}(1-2^{-(1-\beta)/2})^{-1}h_l^{(\beta+1)/2}\rceil$. With that, we have

$$\begin{aligned}\operatorname{Var}[\widehat{\mathfrak{E}}|\mathcal{F}_0] &= \sum_{l=0}^{L}\operatorname{Var}[\widehat{\mathfrak{E}}_l|\mathcal{F}_0] \leqslant \sum_{l=0}^{L} c_2(S_0)N_l^{-1}h_l^{\beta} \\ &\leqslant \frac{c_2(S_0)}{2\mathbb{E}[c_2(S_0)]}\varepsilon^2 h_L^{(1-\beta)/2}(1-2^{-(\beta-1)/2})\sum_{l=0}^{L} h_l^{-(1-\beta)/2}.\end{aligned}$$

Using the formula for a geometric series, we derive the following upper bound for the last term

$$\begin{aligned}\sum_{l=0}^{L} h_l^{-(1-\beta)/2} &= h_L^{-(1-\beta)/2}\sum_{l=0}^{L}(2^{-(1-\beta)/2})^l \\ &= h_L^{-(1-\beta)/2}\frac{1-(2^{-(1-\beta)/2})^{L+1}}{(1-2^{-(1-\beta)/2})} \\ &< h_L^{-(1-\beta)/2}\frac{1}{(1-2^{-(1-\beta)/2})},\end{aligned}$$

which in total gives us again

$$\operatorname{Var}[\widehat{\mathfrak{E}}|\mathcal{F}_0] \leqslant \frac{c_2(S_0)}{2\mathbb{E}[c_2(S_0)]}\varepsilon^2.$$

Thus, we found L and N_l, $l = 0, \dots, L$, such that

$$\mathbb{E}[(\widehat{\mathfrak{E}} - \mathfrak{E})^2|\mathcal{F}_0] \leqslant \frac{c_1(S_0)}{2\mathbb{E}[c_1(S_0)]}\varepsilon^2 + \frac{c_2(S_0)}{2\mathbb{E}[c_2(S_0)]}\varepsilon^2.$$

Taking the expectation on both sides and using the monotonicity of expected values yields

$$\mathbb{E}[(\widehat{\mathfrak{E}} - \mathfrak{E})^2] \leqslant \frac{1}{2}\varepsilon^2 + \frac{1}{2}\varepsilon^2 = \varepsilon^2.$$

It remains to show the monotonicity of the MLMC convergence. Firstly, for $\widetilde{L} > L$, we clearly have

$$(c_1(S_0)h_{\widetilde{L}}^{\alpha})^2 \leqslant (c_1(S_0)h_L^{\alpha})^2 \leqslant \frac{c_1(S_0)}{2\mathbb{E}[c_1(S_0)]}\varepsilon^2.$$

Setting $N_l = 0$ for $l = L+1, \dots, \widetilde{L}$, we still have

$$\sum_{l=0}^{\widetilde{L}} c_2(S_0)N_l^{-1}h_l \leqslant \frac{c_2(S_0)}{2\mathbb{E}[c_2(S_0)]}\varepsilon^2,$$

so that the MSE $\mathbb{E}[(\widehat{\mathfrak{E}} - \mathfrak{E})^2]$ stays below ε^2. Secondly, the estimates are still true for $\widetilde{N}_l \geqslant N_l$, $l = 0, \dots, L$, since we have $\sum_{l=0}^{L} c_2(S_0)\widetilde{N}_l^{-1}h_l \leqslant \sum_{l=0}^{L} c_2(S_0)N_l^{-1}h_l$. Thus, the convergence is indeed monotonic. □

In this thesis, we focus on the convergence of the approximation. However, it is worth noting that the complexity theorem in Giles (2008) includes results on the numerical complexity of the MLMC approach. Moreover, the application of the theorem above requires the verification of the two assumptions (i) and (ii). While assumption (i) closely relates to the well-studied weak convergence of numerical schemes (see Kloeden & Platen, 1992), assumption (ii) might be difficult to prove (see Giles, 2008). Nevertheless, by applying the arguments of Giles (2008) to our setup, we derive sufficient conditions that imply (i) and (ii). In the following, $\|\cdot\|$ denotes the Euclidean norm.

Lemma 5.3. *Suppose that f is Lipschitz continuous and that for every $t \in [0, T]$, there exists a positive function c with $\mathbb{E}[c(S_0)] < \infty$, such that*

$$\mathbb{E}[\|S_l(t) - S(t)\|^2 | \mathcal{F}_0] \leqslant c(S_0) h_l.$$

Then both conditions of Theorem 5.2, (i) and (ii), are satisfied.

Proof. The proof follows the lines from Giles (2008, Section 2), but adapted to our case. The second part of the assumptions describes the strong convergence of the Euler scheme with order 1/2 for every t. The equality

$$\|(x_1 - x'_1, \ldots, x_r - x'_r)\|^2 = \sum_{i=1}^{r} (x_i - x'_i)^2$$

also implies the strong convergence of the Euler scheme for the vector $(S_l(t_1), \ldots, S_l(t_r))$ to $(S(t_1), \ldots, S(t_r))$. In particular, we find a positive function $\widetilde{c}$ such that

$$\mathbb{E}[\|(S_l(t_1) - S(t_1), \ldots, S_l(t_r) - S(t_r))\|^2 | \mathcal{F}_0] \leqslant \widetilde{c}(S_0) h_l$$

with $\mathbb{E}[\widetilde{c}(S_0)] < \infty$. The Lipschitz continuity of f with Lipschitz constant $\tau \geqslant 0$ and the Jensen inequality for conditional expectations yield

$$\begin{aligned}\mathbb{E}[|P_l - P||\mathcal{F}_0] &\leqslant \tau \mathbb{E}[\|(S_l(t_1) - S(t_1), \ldots, S_l(t_r) - S(t_r))\| | \mathcal{F}_0] \\ &\leqslant \tau \mathbb{E}[\|(S_l(t_1) - S(t_1), \ldots, S_l(t_r) - S(t_r))\|^2 | \mathcal{F}_0]^{1/2} \leqslant c_1(S_0) h_l^{1/2}\end{aligned}$$

for a positive function c_1 satisfying $\mathbb{E}[c_1(S_0)] < \infty$. This gives us condition (i) with $\alpha = 1/2$. Furthermore, we can write (see Giles, 2008, p. 608)

$$P_l - P_{l-1} = (P_l - P) - (P_{l-1} - P),$$

which implies together with the Minkowski inequality for conditional expectation (Doob, 1994, Section XI.3)

$$\operatorname{Var}[P_l - P_{l-1} | \mathcal{F}_0] \leqslant (\operatorname{Var}[P_l - P | \mathcal{F}_0]^{\frac{1}{2}} + \operatorname{Var}[P_{l-1} - P | \mathcal{F}_0]^{\frac{1}{2}})^2.$$

Exploiting again the Lipschitz continuity we get

$$\operatorname{Var}[P_l - P | \mathcal{F}_0] \leqslant \mathbb{E}[(P_l - P)^2 | \mathcal{F}_0] \leqslant \tau^2 \mathbb{E}[\|(S_l(t_1) - S(t_1), \ldots, S_l(t_r) - S(t_r))\|^2 | \mathcal{F}_0] \leqslant \widetilde{\widetilde{c}}(S_0) h_l,$$

which in total gives

$$\operatorname{Var}[E_l | \mathcal{F}_0] = \operatorname{Var}\left[N_l^{-1} \sum_{i=1}^{N_l} (\widehat{P}_l^i - \widehat{P}_{l-1}^i) \middle| \mathcal{F}_0 \right] = N_l^{-1} \operatorname{Var}[\widehat{P}_l - \widehat{P}_{l-1} | \mathcal{F}_0] \leqslant c_2(S_0) N_l^{-1} h_l,$$

for some positive function c_2 with $\mathbb{E}[c_2(s_0)] < \infty$, i.e. condition (*ii*) is fulfilled with $\beta = 1$. $\square$

The following lemma provides helpful results on the multiplicity and additivity of monotonically convergent MLMC estimators.

Lemma 5.4. *Let $\mathfrak{E}_1$, $\mathfrak{E}_2 \in L_2(\mathbb{P})$ and let B be a bounded random variable. Furthermore, let $\widehat{\mathfrak{E}}_1$, $\widehat{\mathfrak{E}}_2$ denote the corresponding estimators such that $\widehat{\mathfrak{E}}_1 \overset{MLMC}{\Rightarrow} \mathfrak{E}_1$ monotonically and $\widehat{\mathfrak{E}}_2 \overset{MLMC}{\Rightarrow} \mathfrak{E}_2$ monotonically. Then it holds*

(i) $\widehat{\mathfrak{E}}_1 + \widehat{\mathfrak{E}}_2 \overset{MLMC}{\Rightarrow} \mathfrak{E}_1 + \mathfrak{E}_2$ *monotonically,*

(ii) $B\widehat{\mathfrak{E}}_1 \overset{MLMC}{\Rightarrow} B\mathfrak{E}_1$ *monotonically.*

If $(\mathfrak{E}_1, \widehat{\mathfrak{E}}_1)$ and $(\mathfrak{E}_2, \widehat{\mathfrak{E}}_2)$ are independent, it further holds

(iii) $\widehat{\mathfrak{E}}_1\widehat{\mathfrak{E}}_2 \overset{MLMC}{\Rightarrow} \mathfrak{E}_1\mathfrak{E}_2$ *monotonically.*

Proof. (i) By Definition 5.1, we find L^j, N_l^j, $j = 1, 2$, $l = 0, \dots, L^j$, such that

$$\mathbb{E}[(\widehat{\mathfrak{E}}_j - \mathfrak{E}_j)^2] < \frac{\varepsilon^2}{4}, \ j = 1, 2.$$

Setting $L = \max\{L^1, L^2\}$, the monotonicity ensures that there exist $\widetilde{N}_l^j$, $l = 0, \dots, L$, such that we still have

$$\mathbb{E}[(\widehat{\mathfrak{E}}_j - \mathfrak{E}_j)^2] < \frac{\varepsilon^2}{4}, \ j = 1, 2.$$

Now, again monotonicity reveals that this inequality also holds in both cases ($j = 1, 2$) for L and $N_l = \max\{\widetilde{N}_l^1, \widetilde{N}_l^2\}$, $l = 0, \dots, L$. Since $2ab \leqslant a^2 + b^2$, $a, b \in \mathbb{R}$, we also have

$$\begin{aligned}\mathbb{E}[(\widehat{\mathfrak{E}}_1 + \widehat{\mathfrak{E}}_2 - \mathfrak{E}_1 - \mathfrak{E}_2)^2] &\leqslant 2\mathbb{E}[(\widehat{\mathfrak{E}}_1 - \mathfrak{E}_1)^2] + 2\mathbb{E}[(\widehat{\mathfrak{E}}_2 - \mathfrak{E}_2)^2] \\ &= 2\frac{\varepsilon^2}{4} + 2\frac{\varepsilon^2}{4} = \varepsilon^2.\end{aligned}$$

In particular, the monotonicity follows from individual monotonicities.

(ii) Suppose $|B| \leqslant M$ for some $M > 0$. Since $\widehat{\mathfrak{E}}_1 \overset{MLMC}{\Rightarrow} \mathfrak{E}_1$, we find L, N_l, $l = 0, \dots, L$, such that

$$\mathbb{E}[(\widehat{\mathfrak{E}}_1 - \mathfrak{E}_1)^2] < \frac{\varepsilon^2}{M^2}.$$

Thus, we directly get

$$\mathbb{E}[(B\widehat{\mathfrak{E}}_1 - B\mathfrak{E}_1)^2] = \mathbb{E}[B^2(\widehat{\mathfrak{E}}_1 - \mathfrak{E}_1)^2] \leqslant B^2\mathbb{E}[(\widehat{\mathfrak{E}}_1 - \mathfrak{E}_1)^2] < M^2\frac{\varepsilon^2}{M^2} = \varepsilon^2.$$

The convergence is monotonic as $\widehat{\mathfrak{E}}_1$ converges monotonically to $\mathfrak{E}_1$.

(iii) Let $\varepsilon > 0$. Following the argumentation on monotonicity in (i), we find L, N_l, $l = 0, \dots, L$ such that

$$\mathbb{E}[(\widehat{\mathfrak{E}}_j - \mathfrak{E}_j)^2] < \eta^2,$$

where $\eta = \min\{\varepsilon/(\|\mathfrak{E}_1\|_2 + \|\mathfrak{E}_2\|_2 + 1), 1\}$. This choice of η originates from standard proofs on convergent product sequences (see e.g. Hortmann, 2006). Since

$$\widehat{\mathfrak{E}}_1\widehat{\mathfrak{E}}_2 - \mathfrak{E}_1\mathfrak{E}_2 = (\widehat{\mathfrak{E}}_2 - \mathfrak{E}_2)(\widehat{\mathfrak{E}}_1 - \mathfrak{E}_1) + \mathfrak{E}_2(\widehat{\mathfrak{E}}_1 - \mathfrak{E}_1) + \mathfrak{E}_1(\widehat{\mathfrak{E}}_2 - \mathfrak{E}_2),$$

we can apply triangle inequality and exploit the assumed independence to get

$$\begin{aligned} &\|\widehat{\mathfrak{E}}_1\widehat{\mathfrak{E}}_2 - \mathfrak{E}_1\mathfrak{E}_2\|_2 \\ &\leqslant \|\widehat{\mathfrak{E}}_2 - \mathfrak{E}_2\|_2\|\widehat{\mathfrak{E}}_1 - \mathfrak{E}_1\|_2 + \|\mathfrak{E}_2\|_2\|\widehat{\mathfrak{E}}_1 - \mathfrak{E}_1\|_2 + \|\mathfrak{E}_1\|_2\|\widehat{\mathfrak{E}}_2 - \mathfrak{E}_2\|_2 \\ &\leqslant \eta^2 + \eta\|\mathfrak{E}_1\|_2 + \eta\|\mathfrak{E}_2\|_2 \leqslant \eta(1 + \|\mathfrak{E}_1\|_2 + \|\mathfrak{E}_2\|_2) \leqslant \varepsilon. \end{aligned}$$

In particular, the monotonicity follows from individual monotonicities.

□

As our insurance cash flow includes continuous payments, we are interested in objects like

$$\Psi = \int_a^b \mathfrak{E}(s)\mathrm{d}g(s),\ 0 \leqslant a < b \leqslant T,$$

where g is a finite variation process like $g(s) = s$ or $g(s) = N(s)$, that is bounded on $[a, b]$, and where for every $s \in [a, b]$, $\mathfrak{E}(s)$ can be approximated by a MLMC simulation. Thus, a natural idea is to approximate this integral via finite sums. Therefore, let $(\pi_d^\Psi)_d$ be a sequence of partitions on $[a, b]$ with $|\pi_d^\Psi| \to 0$, $d \to \infty$. We define

$$\Psi_d = \sum_{s_j, s_{j+1} \in \pi_d^\Psi} \mathfrak{E}(s_j)(g(s_{j+1}) - g(s_j)).$$

Assuming that there exists MLMC estimator $\widehat{\mathfrak{E}}(s)$ for $\mathfrak{E}(s)$, we can define the integral estimator

$$\widehat{\Psi}_d = \sum_{s_j, s_{j+1} \in \pi_d^\Psi} \widehat{\mathfrak{E}}(s_j)(g(s_{j+1}) - g(s_j)). \tag{5.13}$$

Definition 5.5. Let $(\pi_d^\Psi)_d$ be a sequence of partitions on $[a, b]$ with $|\pi_d^\Psi| \to 0$, $d \to \infty$. We call $(\widehat{\Psi}_d)_d$ an *integral MLMC estimator* for $\Psi = \int_a^b \mathfrak{E}(s)\mathrm{d}g(s)$, $0 \leqslant a < b \leqslant T$, if for any $\varepsilon > 0$, there exists $d_0 \in \mathbb{N}$ such that for every $d \geqslant d_0$, there are values L, N_l, $l = 0, \dots, L$, such that

$$MSE = \mathbb{E}\left[(\widehat{\Psi}_d - \Psi)^2\right] < \varepsilon^2. \tag{5.14}$$

We briefly write $(\widehat{\Psi}_d)_d \overset{I-MLMC}{\Rightarrow} \Psi$ (w.r.t. $(\pi_d^\Psi)_d$). We call the convergence *monotonic* if $(\widehat{\Psi}_d)_d \overset{MLMC}{\Rightarrow} \Psi_d$ monotonically for every d.

Theorem 5.6. *Let $(\pi_d^\Psi)_d$ be a sequence of partitions on $[a,b]$ with $|\pi_d^\Psi| \to 0$. If for every $s \in [a,b]$, there exists a MLMC estimator $\widehat{\mathfrak{E}}(s)$ such that $\widehat{\mathfrak{E}}(s) \overset{MLMC}{\Rightarrow} \mathfrak{E}(s)$ monotonically and if $\|\Psi - \Psi_d\|_2 \to 0$, $d \to \infty$, then it holds $(\widehat{\Psi}_d)_d \overset{I-MLMC}{\Rightarrow} \Psi$ monotonically (w.r.t. $(\pi_d^\Psi)_d$).*

Proof. Following the construction of the estimator (see the remarks before Definition 5.5) we have to show that for each $\varepsilon > 0$, there exist $d_0 \in \mathbb{N}$, such that for every $d \geqslant d_0$, there are L, N_l, $l = 0, \dots, L$ such that

$$\Big\| \int_a^b \mathfrak{E}(s) \mathrm{d}g(s) - \sum_{s_j, s_{j+1} \in \pi_d^\Psi} \widehat{\mathfrak{E}}(s_j)(g(s_{j+1}) - g(s_j)) \Big\|_2 < \varepsilon$$

for all $d \geqslant d_0$. For this purpose, we proceed in two steps. First, we apply the triangle inequality to get

$$\begin{aligned} &\Big\| \int_a^b \mathfrak{E}(s) \mathrm{d}g(s) - \sum_{s_j, s_{j+1} \in \pi_d^\Psi} \widehat{\mathfrak{E}}(s_j)(g(s_{j+1}) - g(s_j))) \Big\|_2 \\ &\leqslant \Big\| \int_a^b \mathfrak{E}(s) \mathrm{d}g(s) - \sum_{s_j, s_{j+1} \in \pi_d^\Psi} \mathfrak{E}(s_j)(g(s_{j+1}) - g(s_j)) \Big\|_2 \\ &\quad + \Big\| \sum_{s_j, s_{j+1} \in \pi_d^\Psi} (\mathfrak{E}(s_j) - \widehat{\mathfrak{E}}(s_j))(g(s_{j+1}) - g(s_j)) \Big\|_2. \end{aligned}$$

for every d. Since we assumed $\|\Psi - \Psi_d\|_2 \to 0$, $d \to \infty$, we find d_0 such that

$$\Big\| \int_a^b \mathfrak{E}(s) \mathrm{d}g(s) - \sum_{s_j, s_{j+1} \in \pi_d^\Psi} \mathfrak{E}(s_j)(g(s_{j+1}) - g(s_j)) \Big\|_2 < \frac{\varepsilon}{2}$$

for all $d \geqslant d_0$.

For the second part, the combination of (i) and (ii) in Lemma 5.4 tells us (g is bounded in both cases), that for every $d \geqslant d_0$ we find L and N_l, $l = 0, \dots, L$, such that

$$\Big\| \sum_{s_j, s_{j+1} \in \pi_d^\Psi} (\mathfrak{E}(s_j) - \widehat{\mathfrak{E}}(s_j))(g(s_{j+1}) - g(s_j)) \Big\|_2 < \frac{\varepsilon}{2}. \tag{5.15}$$

In particular, since $\widehat{\mathfrak{E}}(s) \overset{MLMC}{\Rightarrow} \mathfrak{E}(s)$ monotonically, the convergence in (5.15) is monotonic (see Lemma 5.4). In total, we have $(\widehat{\Psi}_d)_d \overset{I-MLMC}{\Rightarrow} \Psi$ monotonically. □

Lemma 5.7. *Let $0 \leqslant a_i < b_i \leqslant T$, and let $(\pi_d^{\Psi,i})_d$ be a sequence of partitions on $[a_i, b_i]$ with $|\pi_d^{\Psi,i}| \to 0$, $d \to \infty$, $i = 1, 2$. We consider integrals $\Psi_i = \int_{a_i}^{b_i} \mathfrak{E}_i(s) dg(s)$ with corresponding estimators $\widehat{\Psi}_{i,d}$ defined as in* (5.13). *Suppose that $\widehat{\Psi}_{i,d} \overset{I-MLMC}{\Rightarrow} \Psi_i$ monotonically, $i = 1, 2$. Then it holds*

$$(\widehat{\Psi}_{1,d} + \widehat{\Psi}_{2,d})_d \overset{I-MLMC}{\Rightarrow} \Psi_1 + \Psi_2$$

monotonically.

Proof. Let $\varepsilon > 0$. Since $(\widehat{\Psi}_{i,d})_d \overset{I-MLMC}{\Rightarrow} \Psi_i$, we find $d_i \in \mathbb{N}$, such that for every $d \geqslant d_i$ there are values L^i, N_l^i, $l = 0, \dots, L^i$, such that

$$\|\widehat{\Psi}_{i,d} - \Psi_i\|_2 < \frac{\varepsilon}{2}. \tag{5.16}$$

In particular, for $d \geqslant d_0 := \max\{d_1, d_2\}$, there are values L^i, N_l^i, $l = 0, \dots, L^i$ such that (5.16) holds. Setting $L = \max\{L^1, L^2\}$, the monotonicity ensures that we find N_l, $l = 0, \dots, L$, such that (5.16) holds, which implies

$$\|\Psi_1 + \Psi_2 - (\widehat{\Psi}_{1,d} + \widehat{\Psi}_{2,d})\|_2 \leqslant \|\Psi_1 - \widehat{\Psi}_{1,d}\|_2 + \|\Psi_2 - \widehat{\Psi}_{2,d}\|_2 < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

With Lemma 5.4, we conclude that the convergence is monotonic. □

Now that we have extended MLMC theory to conditional expectations, we can return to our original model and apply the results there.

5.2.3 Convergence results

Building on the previously discussed MLMC methods, this section derives convergent MLMC estimators for the SU contributions introduced in Section 5.1. At the beginning, we focus on estimating conditional expectations like (5.5) - (5.9), which are part of the SU integral representations (see Section 5.2.1).

First we need a few preparations. In the following, we fix a SU grid point $t_k \in \mathcal{T}$. Let $(\widetilde{\lambda}, \widetilde{r}, \Phi, \widetilde{Y}, \widetilde{Q})$ denote the solution to the shifted system of stochastic differential equations

$$\mathrm{d}\widetilde{\lambda}(t) = \widetilde{\lambda}(t)\mu_\lambda \mathrm{d}t + \sigma_\lambda \mathrm{d}W_\lambda(t),\ \widetilde{\lambda}(t_k) = \lambda(t_k) \tag{5.17}$$

$$\mathrm{d}\widetilde{r}(t) = (\beta - \widetilde{r}(t)\mu_r)\mathrm{d}t + \sigma_r \mathrm{d}W_r(t),\ \widetilde{r}(t_k) = r(t_k), \tag{5.18}$$

$$\mathrm{d}\Phi(t) = g_r(\widetilde{r}(t))\mathrm{d}t,\ \Phi(t_k) = 0, \tag{5.19}$$

$$\mathrm{d}\widetilde{Y}(t) = \widetilde{Y}(t)\mu_Y \mathrm{d}t + \widetilde{Y}(t)\sigma_Y \mathrm{d}W_Y(t),\ \widetilde{Y}(t_k) = Y(t_k), \tag{5.20}$$

$$\mathrm{d}\widetilde{Q}(t) = \frac{a(t)}{g_Y(Y(t))}\mathrm{d}t,\ \widetilde{Q}(t_k) = Q(t_k), \tag{5.21}$$

for $t \geqslant t_k$. Let $s \in (t_k, T]$ and suppose that $s \in (t_{k_1}, t_{k_1+1}]$ for some index k_1. Recall from Subsection 5.2.1 that $u(t) = u_{l+1}$, if $t \in (u_l, u_{l+1}]$. Assume $u(t_k) = t_{k_0} \geqslant t_k$ for some k_0. We define for $s > u(t_k)$,

$$f_p(\widetilde{\lambda}(t_{k_0}), \dots, \widetilde{\lambda}(t_{k_1})) = \exp\left(-\sum_{j=k_0}^{k_1} g_\lambda(\widetilde{\lambda}(t_j))(t_{j+1} - t_j)\right), \tag{5.22}$$

$$f_{p\lambda}(\widetilde{\lambda}(t_{k_0}), \dots, \widetilde{\lambda}(t_{k_1})) = \exp\left(-\sum_{j=k_0}^{k_1} g_\lambda(\widetilde{\lambda}(t_j))(t_{j+1} - t_j)\right) g_\lambda(\widetilde{\lambda}(t_{k_1})) \tag{5.23}$$

as well as for $s > t_k$,

$$f_v(\Phi(s)) = \exp(-\Phi(s)), \tag{5.24}$$

$$f_Y(\widetilde{Y}(s)) = g_Y(\widetilde{Y}(s)), \tag{5.25}$$

$$f_Q(\widetilde{Q}(t_k), \widetilde{Q}(s), \widetilde{Y}(s)) = (\widetilde{Q}(s) - \widetilde{Q}(t_k))g_Y(\widetilde{Y}(s)). \tag{5.26}$$

Using this notation, we can write

$$p_u^{\mathbb{E}}(t_k, s) = \mathbb{E}[f_p(\widetilde{\lambda}(t_{k_0}), \ldots, \widetilde{\lambda}(t_{k_1}))|\mathcal{H}_{t_k}], \tag{5.27}$$

$$p_{u\lambda}^{\mathbb{E}}(t_k, s) = \mathbb{E}[f_{p\lambda}(\widetilde{\lambda}(t_{k_0}), \ldots, \widetilde{\lambda}(t_{k_1}))|\mathcal{H}_{t_k}], \tag{5.28}$$

$$v^{\mathbb{E}}(t_k, s) = \mathbb{E}[f_v(\Phi(s))|\mathcal{H}_{t_k}], \tag{5.29}$$

$$Y^{\mathbb{E}}(t_k, s) = \mathbb{E}[f_Y(\widetilde{Y}(s))|\mathcal{H}_{t_k}], \tag{5.30}$$

$$Y_Q^{\mathbb{E}}(t_k, s) = \mathbb{E}[f_Q(\widetilde{Q}(t_k), \widetilde{Q}(s), \widetilde{Y}(s))|\mathcal{H}_{t_k}]. \tag{5.31}$$

These representations allow us to apply the results from the previous section. Let $(\widetilde{\lambda}_l^{t_k}, \widetilde{r}_l^{t_k}, \Phi_l^{t_k}, \widetilde{Y}_l^{t_k}, \widetilde{Q}_l^{t_k})$ denote the Euler approximation of $(\widetilde{\lambda}, \widetilde{r}, \Phi, \widetilde{Y}, \widetilde{Q})$ with respect to a step width h_l. With help of Theorem 10.2.2 in Kloeden and Platen (1992), we can conclude the strong convergence of the Euler scheme for the shifted system of stochastic differential equations.

Lemma 5.8. *The Euler approximation $X_l = (\lambda_l^{t_k}, r_l^{t_k}, \Phi_l^{t_k}, Y_l^{t_k}, Q_l^{t_k})$ with step width h_l converges strongly to $\widetilde{X} = (\widetilde{\lambda}, \widetilde{r}, \Phi, \widetilde{Y}, \widetilde{Q})$ with order $1/2$, i.e. there exists a positive function c with $\mathbb{E}[c(X(t_k))] < \infty$, such that*

$$\mathbb{E}[\|X_l(s) - \widetilde{X}(s)\|^2|\mathcal{H}_{t_k}] \leqslant c(X(t_k))h_l.$$

Proof. Firstly, we rewrite the system of stochastic differential equations (5.17) - (5.21) as

$$\mathrm{d}\widetilde{X}(t) = \mu(s, \widetilde{X}(t))\mathrm{d}t + \sigma(s, \widetilde{X}(t))\mathrm{d}\widetilde{W}(t), \;\; \widetilde{X}(t_k) = \widetilde{X}_0(t_k),$$

with the (extended) standard Brownian motion $\widetilde{W} = (W_\lambda, W_r, W_\Phi, W_Y, W_Q)$, and

$$\widetilde{X}_0(t_k) = (\lambda(t_k), r(t_k), 0, Y(t_k), Q(t_k)),$$

$$\mu(t, x) = (x_1\mu_\lambda, (\beta - x_2)\mu_r, -g_r(x_2), x_4\mu_Y, a(t)/g_Y(x_4))^\top$$

and

$$\sigma(t, x) = (\sigma_Y, \sigma_r, 0, x_4\sigma_Y, 0)^\top,$$

where $x = (x_1, \ldots, x_5) \in \mathbb{R}^5$, $t \in [t_k, s]$. Recall that g_r is assumed to be Lipschitz continuous. Furthermore, since g_Y is Lipschitz continuous with values in $[c, +\infty)$, also $1/g_Y$ defines a Lipschitz continuous function. Having this in mind, one easily shows that

$$\|\mu(t, x) - \mu(t, y)\| + \|\sigma(t, x) - \sigma(t, y)\| \leqslant K_1\|x - y\|, \;\; t \in [t_k, s], \;\; x, y \in \mathbb{R}^5,$$

for some constant $K_1 > 0$. Moreover, with the boundedness of $g_r(x_2)$ in x_2 and $a(t)/g_Y(x_4)$ in (t, x_4), it holds

$$\|\mu(t,x)\| + \|\sigma(t,x)\| \leqslant K_3(1 + \|x\|),\ t \in [t_k, s],\ x \in \mathbb{R}^5,$$

for some constant $K_2 > 0$. With the Lipschitz continuity of $a(t)$, one immediately verifies

$$\|\mu(t,x) - \mu(u,x)\| + \|\sigma(t,x) - \sigma(u,x)\| \leqslant K_3(1 + \|x\|)|t-u|^{1/2},\ u,t \in [t_k, s],\ x \in \mathbb{R}^5,$$

for some $K_3 > 0$. Furthermore, our assumptions made in Section 4.1 imply

$$\mathbb{E}[\|\widetilde{X}(t_k)\|^2] < \infty. \tag{5.32}$$

Thus, all assumptions of Theorem 10.2.2 in Kloeden and Platen (1992) are fulfilled. A closer look into their proof reveals that there exists a constant $C > 0$ such that

$$\mathbb{E}[\sup_{t_k \leqslant u \leqslant s} \|\widetilde{X}(u) - X_l(u)\|^2 | \mathcal{F}_{t_k}^{\widetilde{W}}] \leqslant C(1 + \|X(t_k)\|^2)h_l =: c(X(t_k))h_l.$$

With $\mathcal{H}_{t_k} \subseteq \mathcal{F}_{t_k}^{\widetilde{W}}$, the monotonicity of conditional expectations and (5.32), we get

$$\mathbb{E}[\|\widetilde{X}(s) - X_l(s)\|^2 | \mathcal{H}_{t_k}] \leqslant c(X(t_k))h_l,$$

where $\mathbb{E}[c(X(t_k))] < \infty$, which concludes the proof. □

Following the construction (5.10) - (5.11), we derive MLMC estimators $\widehat{p}_u(t_k, s)$, $\widehat{p}_{u\lambda}(t_k, s)$ for $s > u(t_k)$, and $\widehat{v}(t_k, s)$, $\widehat{Y}(s)$ and $\widehat{Y}_Q(s)$ for $s > t_k$. Applying the previous lemma and Theorem 5.2, we get the following convergence results.

Lemma 5.9. *It holds*

a) $\widehat{p}_u(t_k, s) \overset{MLMC}{\Rightarrow} p_u^{\mathbb{E}}(t_k, s)$, *b)* $\widehat{p}_{u\lambda}(t_k, s) \overset{MLMC}{\Rightarrow} p_{u\lambda}^{\mathbb{E}}(t_k, s)$,

c) $\widehat{v}(t_k, s) \overset{MLMC}{\Rightarrow} v^{\mathbb{E}}(t_k, s)$, *d)* $\widehat{Y}(s) \overset{MLMC}{\Rightarrow} Y^{\mathbb{E}}(s)$,

e) $\widehat{Y}_Q(s) \overset{MLMC}{\Rightarrow} Y_Q^{\mathbb{E}}(s)$.

Proof. The idea is to apply Theorem 5.2. To verify the assumptions of (i) and (ii) of Theorem 5.2, we use the Lemma 5.3. The latter requires both, the strong convergence of the Euler scheme and the Lipschitz continuity of the functions (5.22) - (5.26). The strong convergence of the Euler scheme follows from Lemma 5.8. Thus, it remains to prove the Lipschitz continuity.

a) First, we consider the functions

$$f_{p,i}(x) = \exp(-g_\lambda(x)(t_{i+1} - t_i)),\ i = k_0, \ldots, k_1. \tag{5.33}$$

The function $x \mapsto \exp(-x(t_{i+1} - t_i))$, $x \geqslant 0$ is a bounded Lipschitz continuous function. Since g_λ is assumed to be Lipschitz continuous with values in $[0, B_\lambda]$ (see Section 5.1), the functions $f_{p,i}$, $i = k_0, \dots, k_1$ are Lipschitz continuous as well. Thus, the product function $f_p = \prod_{i=k_0}^{k_1} f_{i,p}$, $f_p(x_{k_0}, \dots, x_{k_1}) = f_{k_0}(x_{k_0}) \cdots f_{k_1}(x_{k_1})$ is Lipschitz continuous.

b) Observe that $f_{p\lambda} = (\prod_{i=k_0}^{k_1} f_{p,i})g_\lambda$, where $f_{p,i}$ are defined by (5.33). With a) and the Lipschitz continuity of g_λ, we conclude the Lipschitz continuity of $f_{p\lambda}$.

c) Since Φ is bounded (since g_r is bounded), we can consider $f_v(x) = \exp(-x)$ on a compact interval, which is a Lipschitz continuous function.

d) The Lipschitz continuity of $f_Y = g_Y$ follows from our assumptions on g_Y in Section 5.1.

e) Since Q is bounded, it is sufficient to consider f_Q on $D \times \mathbb{R}_+$ for a compact set $D \subseteq \mathbb{R}^2$. Together with the Lipschitz continuity of g_Y, we have

$$\begin{aligned}
&|f_Q(x_1, x_2, x_3) - f_Q(y_1, y_2, y_3)| \\
&= |(x_2 - x_1)g_Y(x_3) - (y_2 - y_1)g_Y(y_3)| \\
&= |(x_2 - y_2 + y_1 - x_1)g_Y(x_3) + (y_2 - y_1)(g_Y(x_3) - g_Y(y_3))| \\
&\leqslant (|x_2 - y_2| + |y_1 - x_1|)|g_Y(x_3)| + |y_2 - y_1||g_Y(x_3) - g_Y(y_3)| \\
&\leqslant \widetilde{C}\|x - y\|_1(|g_Y(x_3)| + |y_2 - y_1|) \\
&\leqslant C\|x - y\|_2,
\end{aligned}$$

for $x, y \in D \times \mathbb{R}_+$ and some $C > 0$. Thus, f_Q is Lipschitz continuous on $D \times \mathbb{R}_+$.

With Lemma 5.3, Lemma 5.8 and Theorem 5.2, we can therefore conclude the asserted convergences a) to e).

□

Based on this lemma, we derive estimators for the SU risk factors. As presented in Section 5.2.1, each SU addend is a sum of integrals $\Psi = \int_a^b \mathfrak{E}(s)\mathrm{d}g(s)$ with $g \in \{id, N\}$. We approximate the integrals by finite sums $\Psi_d = \sum_{s_j, s_{j+1} \in \pi_d^\Psi} \mathfrak{E}(s_j)(g(s_{j+1}) - g(s_j))$, where π_d^Ψ is a sequence of partitions on $[a, b]$ with $|\pi_d^\Psi| \to 0$, $n \to \infty$. As $\mathfrak{E}(s_j)$ still includes conditional expectations, we substitute $\mathfrak{E}(s_j)$ by $\widehat{\mathfrak{E}}(s_j)$, which results from replacing the conditional expectations (5.5) - (5.9) in $\mathfrak{E}(s_j)$ by the corresponding MLMC estimators $\widehat{p}_u$, $\widehat{p}_{u\lambda}$, $\widehat{v}$, $\widehat{Y}$ and $\widehat{Y}_Q$. The corresponding MLMC estimators for the different contract components ξ_i are denoted by $\widehat{D}_j^i(t)$, $j \in \{N, \lambda, r, Y\}$, $i = 1, \dots, 4$. Consequently, the MLMC estimators for the total insurance cashflow ξ are given by

$$\widehat{D}_j(t) = \sum_{i=1}^{4} \widehat{D}_j^i(t), \; j \in \{N, \lambda, r, Y\}.$$

The next theorem verifies the desired convergence of the constructed estimators.

Theorem 5.10. *It holds* $\widehat{D}_j(t) \overset{I-MLMC}{\Rightarrow} D_j(t)$*,* $j \in \{N, \lambda, r, Y\}$*,* $t \in [0, T]$*.*

Proof. With help of Lemma 5.7, it is sufficient to show

$$\widehat{D}^i_j(t) \overset{I-MLMC}{\Rightarrow} D^i_j(t) \tag{5.34}$$

monotonically, for each $i = 1, \dots, 4$ and $j \in \{N, \lambda, r, Y\}$.
As described in the previous paragraph, each $\widehat{D}^i_j(t)$ is the sum of estimators

$$\widehat{\Psi}_d = \sum_{s_j, s_{j+1} \in \pi^\Psi_d} \widehat{\mathfrak{E}}(s_j)(g(s_{j+1}) - g(s_j)), \; g \in \{id, N\}.$$

As usual, we write

$$\Psi = \int_a^b \mathfrak{E}(s) \mathrm{d}g(s)$$

for the original integral and

$$\Psi_d = \sum_{s_j, s_{j+1} \in \pi^\Psi_d} \mathfrak{E}(s_j)(g(s_{j+1}) - g(s_j))$$

for its finite sum approximation. According to Theorem 5.6, we have to check that $\widehat{\mathfrak{E}}(s) \overset{MLMC}{\Rightarrow} \mathfrak{E}(s)$ monotonically, and $\|\Psi - \Psi_d\|_2 \to 0$, $d \to \infty$.

Clearly, each $\mathfrak{E}(s)$ is a product of bounded factors like $I(s)$, $v(0, s)$, $V(s)$ and conditional expectations like (5.5) - (5.9). Though the conditional expectations always refer to the information $\mathcal{H}_{t_k}$, the independence of λ, r and Y and the construction of the MLMC estimators imply the independence of $(p^{\mathbb{E}}_u, p^{\mathbb{E}}_{u\lambda}, \widehat{p}_u, \widehat{p}_{u\lambda})$, $(v^{\mathbb{E}}, \widehat{v})$ and $(Y^{\mathbb{E}}, Y^{\mathbb{E}}_Q, \widehat{Y}, \widehat{Y}_Q)$. Therefore, we can apply Lemma 5.4 and Lemma 5.9, which proves $\widehat{\mathfrak{E}}(s) \overset{MLMC}{\Rightarrow} \mathfrak{E}(s)$ monotonically.

Regarding the integral approximation, we start with $g = id$. As the appearing conditional expectations (see (5.5) - (5.9)) can be bounded by a constant and are continuous in s, and as I has at most one jump, the dominated convergence theorem implies $\|\Psi - \Psi_d\|_2 \to 0$, $d \to \infty$.

In the second case, i.e. $g = N$, the Itô isometry and $\mathrm{d}\langle N, N\rangle(s) = \overline{\lambda}(s)\mathrm{d}s$ yield

$$\|\Psi - \Psi_d\|_2^2 = \mathbb{E}\left[\int_a^b (\mathfrak{E}(s) - \mathfrak{E}_d(s))^2 \overline{\lambda}(s) \mathrm{d}s\right], \tag{5.35}$$

where $\mathfrak{E}_d(s) = \mathfrak{E}(s_j)$ for $s \in (s_j, s_{j+1}]$. With similar arguments as in the first case and the piecewise continuity and boundedness of $\overline{\lambda}$, this tends to 0 as $d \to \infty$ by the dominated convergence theorem. Thus, with Theorem 5.6, we can conclude the convergence (5.34), which proves the theorem. □

The derived MLMC estimators will be the basis of the numerical implementation, which is the subject of the next section.

5.3 Numerical example

In the previous section, we have laid the theoretical basis for the numerical approximation of the SU decomposition. This section presents the numerical implementation and starts with stating the utilized parameters and explaining the methodology. This section and thus this chapter concludes with analysing the numerical results.

5.3.1 Parameters

We consider a policyholder of age $x = 45$ with retirement age 67 (i.e. $\gamma = 22$), where a maximum age of 100 (i.e. $T = 55$) is assumed. The systematic risks are modelled by the stochastic differential equations (5.2) with parameters

$$\begin{aligned}
&\lambda_0 = 0.00329542,\ \mu_\lambda = 0.07731571,\ \sigma_\lambda = 0.00012212,\\
&r_0 = 0.025,\ \beta = 0.000199,\ \mu_r = 0.01,\ \sigma_r = 0.0035,\\
&Y_0 = 100,\ \mu_Y = 0.06,\ \sigma_Y = 0.2.
\end{aligned}$$

The parameters are based on calibrations under the real-world measure $\mathbb{P}$ that have been carried out in the actuarial literature. For the mortality intensity λ, we refer to Chen and Vigna (2017), who calibrated the parameters in an unisex mortality model. The parameters for the interest intensity are provided by Spangler (2018, Section 5.2), who obtained the Vasicek dynamics by means of six-month Euribor observations. For the market index Y, the parameters have been suggested by Bernard and Kwak (2016).

It is assumed that the one-off premium p covers the purchase of $Q(0) = 100$ shares of Y in $t = 0$. As the initial surplus is not decomposed (see Chapter 2), we do not need to further specify the one-off premium p. In addition to the one-off premium, the policyholder pays a savings rate $a = a(s) = 3600$ until retirement, which is invested in Y. For the death benefit factor and the pension factor we suppose $f_d = 0.9$ and $f_p = 0.032$.

The convergence proofs have required a certain regularity in the risk drivers, which has been ensured by the auxiliary functions g_r, g_λ and g_Y. In our example, we use

$$\begin{aligned}
g_\lambda(s) &= s \cdot \mathbb{1}_{[0,1000]}(s) + 1000 \cdot \mathbb{1}_{(1000,+\infty)}(s),\\
g_r(s) &= -1000 \cdot \mathbb{1}_{(-\infty,-1000)}(s) + s \cdot \mathbb{1}_{[-1000,1000]}(s) + 1000 \cdot \mathbb{1}_{(1000,+\infty)}(s),\\
g_Y(s) &= 0.0001 \cdot \mathbb{1}_{(-\infty,0.0001)}(s) + s \cdot \mathbb{1}_{[0.0001,+\infty)}(s).
\end{aligned}$$

As pointed out in Section 5.1, the function g_λ ensures the non-negativity of the mortality intensity. Apart from that, the practical effect of these transformations is neglectable due to the chosen boundaries. Now that the parameters have been established, we can present the implementation approach.

5.3.2 Methodology

The numerical implementation is carried out with the help of the statistical software R 4.4.2 (R Core Team, 2024). The goal is to approximate the ISU decomposition by means of SU decompositions. To simulate the latter, we focus on a fixed path and apply the presented MLMC methods. As described in the introduction of this chapter, the undertaken studies are focused on the following aspects:

- Numerical feasibility of calculating SU decompositions with a MLMC approach
- Impact of the grid width on the SU decomposition
- Impact of the update order on the SU decomposition

The first aspect addresses the question of whether the SU decomposition can be computed numerically in a reasonable time using the presented MLMC approach. In addition to feasibility, the stability of the decomposition is also of great importance, which is reflected in aspects two and three. In the previous chapters, we have promoted the ISU decomposition, because of its independence from the chosen grid and the chosen order of the risk basis. In order to investigate the stability with regard to the grid, we calculate the SU decompositions on a yearly, a quarterly and a 4-weekly grid. Afterwards, the resulting decompositions are compared with each other. The stability with respect to the order of risk factors can be verified by calculating SU decompositions for different orders. Therefore, we do not only compute the SU decomposition with respect to the presented risk basis (N, λ, r, Y), but also with respect to the vice-versa risk basis (Y, r, λ, N). The integral representations of the SU contributions for both orders are derived in the appendix (see Section A.3).

Before getting into the details of the numerical simulation, we would like to raise the awareness for the involved approximation steps:

1) We approximate the ISU decomposition with the help of SU decompositions. The convergence has been shown in Chapter 4 (see Section 4.2.5 and Theorem 4.18).

2) The contributions of the SU decomposition are approximated by integral MLMC estimators. The convergence has been shown in this chapter (see Theorem 5.10). In particular, this approximation encompasses two steps.

    2.1) The integrals, which amount to the SU contributions, are approximated by finite sums (Riemann sums).

    2.2) For each summand in the integral approximation 2.1), the appearing conditional expectations are approximated with the help of MLMC estimators.

Recall from the previous subsection that the total projection horizon amounts to 55 years. To approximate the ISU decomposition via SU decompositions (approximation step 1), we consider the following SU grids

$$\begin{aligned}
\mathcal{T}^{yearly} &= \{k : k = 0, \dots, 55\}, \\
\mathcal{T}^{quarterly} &= \left\{\frac{k}{4} : k = 0, \dots, 55 \cdot 4\right\}, \\
\mathcal{T}^{4-weekly} &= \left\{\frac{k}{13} : k = 0, \dots, 55 \cdot 13\right\}.
\end{aligned}$$

We calculate the corresponding SU estimators $\widehat{D}_j^{yearly}(t)$, $\widehat{D}_j^{quarterly}(t)$, $\widehat{D}_j^{4-weekly}(t)$, $j \in \{N, \lambda, r, Y\}$, for all t in the respective SU grid. This enables us to analyse not only the surplus contributions at the contract end, but also its development over time.

For the approximation of the appearing integrals via finite sums (approximation step 2.1), we use a weekly grid, namely

$$\pi_{INT} = \{k/52 : k = 0 \dots, 55 \cdot 52\},$$

which contains the various SU grids by definition. Each summand in the integral approximation includes conditional expectations (see (5.5) - (5.9)), which are approached with the help of MLMC methods (approximation step 2.2). The latter step entails the computation of the MLMC estimators $\widehat{p}_u(t_k, s)$, $\widehat{p}_{u\lambda}(t_k, s)$, $\widehat{v}(t_k, s)$, $\widehat{Y}(t_k, s)$ and $\widehat{Y}_Q(t_k, s)$ for all combinations of SU grid points t_k and integral grid points s. Most of the numerical effort associated with approximating the ISU decomposition is required in this step.

The procedure for the computation of the MLMC estimators is as follows. At the beginning, we calculate a fixed path for the risk factors, which represents the actual observed trend. For every SU grid point t_k, the observed trends up to t_k determine the starting points of the MLMC simulations. For every level $l = 0, \dots, L$, we simulate N_l paths of the risk factors using the Euler scheme with the given starting points and the step width h_l. For the latter, we use $h_l = 2^{-l}/52$, i.e. for $l = 0$ we start with the integration grid π_{INT}, which is then refined by the factor two with every level l. For the MLMC simulation, we use $L = 4$ levels with $N = (100000, 50000, 20000, 10000, 5000)$ describing the simulations per level. After evaluating the functions f_p, $f_{p\lambda}$, f_v and f_Q with respect to the simulated paths, we follow the construction in (5.10) and (5.11) to achieve the MLMC estimators $\widehat{p}_u(t_k, s)$, $\widehat{p}_{u\lambda}(t_k, s)$, $\widehat{v}(t_k, s)$, $\widehat{Y}(t_k, s)$ and $\widehat{Y}^Q(t_k, s)$. The computed MLMC estimators then feed, together with the fixed path of the risk factors and the integral approximation step 2.1, into the calculation of the SU contributions (see the integral representations in Subsection 5.2.1) for different the SU grids and the different update orders.

5.3.3 Results

In this section, we present the numerical results for decomposing the surplus related to a fund-linked pension insurance (see Section 5.1) by means of the ISU decomposition principle. The underlying R script can be found in the appendix (see Section A.4).

In our numerical example, we focus on a fixed path of the SU decompositons. Therefore, we first sampled a path for each systematic risk factor (mortality, interest, fund). With help of the sampled mortality intensity, we then simulated a path for the policyholder's state process. The example paths depict an observed trend on $[0, T]$. In particular, the realisations form the starting values for the MLMC simulations. Figure 1 shows the sampled paths.

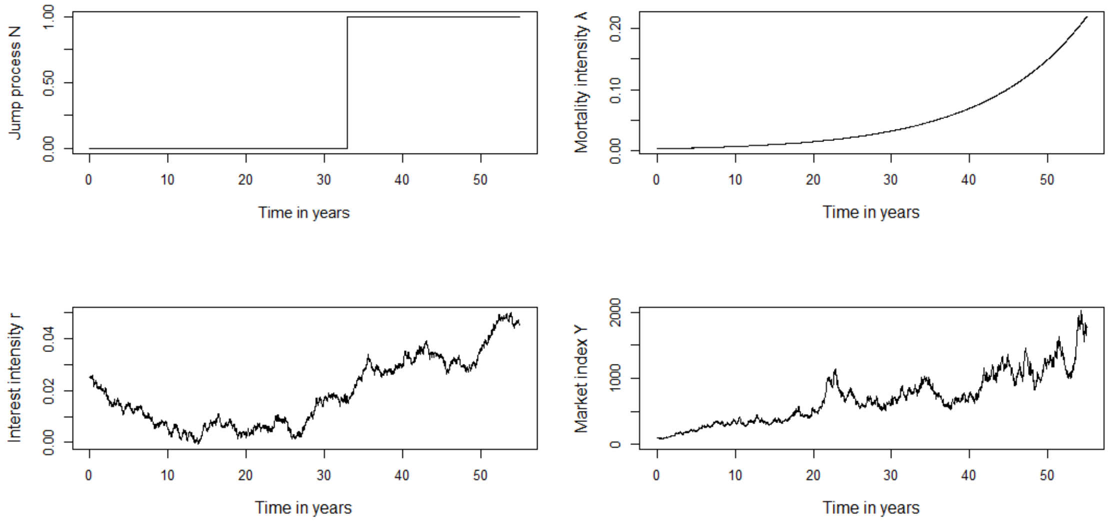

Figure 1: The sampled trend of the risk factors

The policyholder reaches the age of 77 years and has therefore been able to draw pension benefits from the insurance for 10 years. The mortality intensity λ shows an exponential increase with a maximum value of approx. 0.22 at the age 100. After a decreasing period of ten years, the interest intensity r shows a fluctuating trend for the next two decades. In the second half of the time horizon, the interest intensity increases significantly to up to 5%. The market index Y displays a steady growth with ordinary fluctuations in the first two decades. Afterwards volatility increases, although the market index continues to grow strongly. Fortunately for the policyholder, the index is growing particularly strongly in the 22nd year, which gives an average yield of 10.55% p.a. in the accumulation phase (for comparison: the average yield in the first 20 years is 7.2% p.a.).

In the following, our analysis focuses on the aspects stated in Subsection 5.3.2.

Numerical feasibility of the MLMC approach

With help of the introduced MLMC approach, we were able to calculate the SU decompositions for a single path in a resonable time. Using a personal computer with average random access memory (16 GB), the runtime was approximately 4 hours for the yearly, 12 hours for the quarterly and 48 hours for the 4-weekly SU grid. The results presented below clearly demonstrate the numerical feasibility of the ISU decomposition principle. However, to calculate the distribution of the SU contributions, further computational power would be needed.

Different SU grid widths

One motivation to introduce the ISU decomposition in this thesis was the dependence of the SU decomposition on the considered time grid. The ISU decomposition overcomes this drawback by refining the SU grids. However, if the ISU decomposition is not readily available, one needs to use SU decompositions as an approximation. Therefore, a natural question is how fine the grid needs to be in order to observe a stable decomposition with respect to the SU grid width.

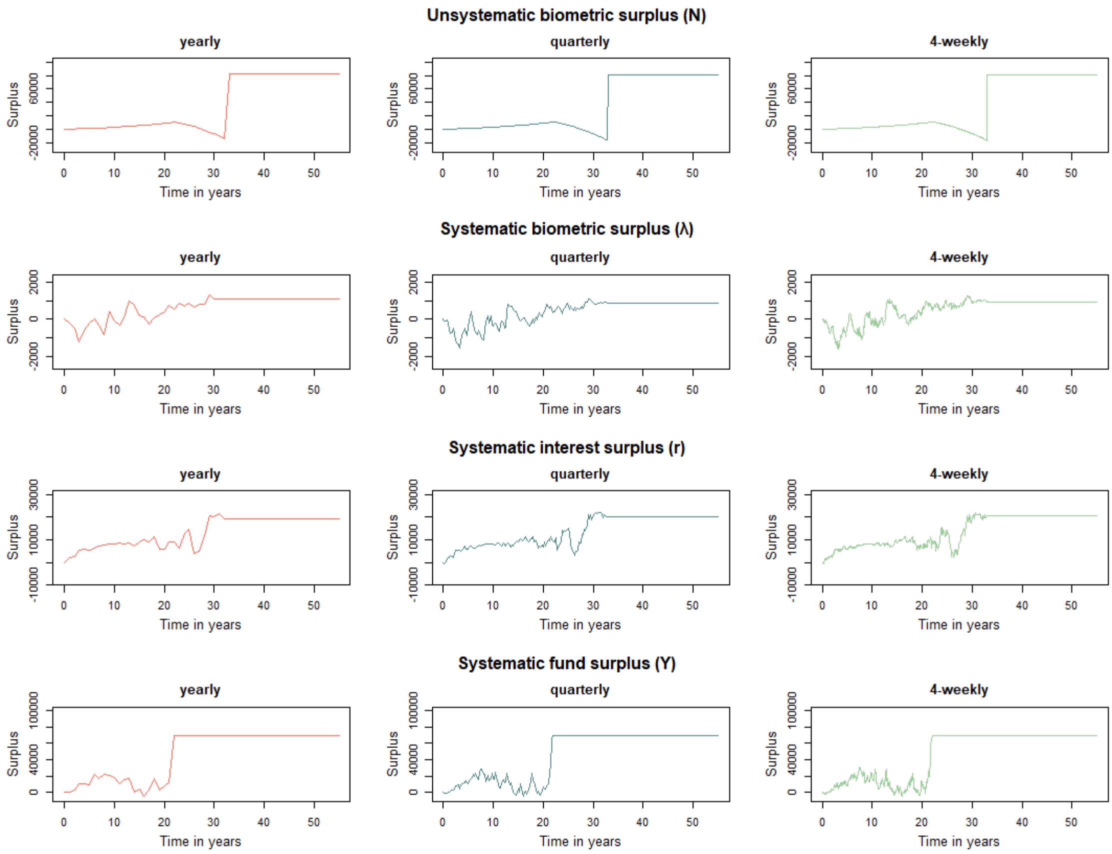

Figure 2: The SU decomposition for different time grids

To answer this question, we compare the SU decompositions for three different SU grids, namely the yearly, the quarterly and the 4-weekly grid, and analyse their impact on the surplus decomposition. The resulting risk factors for the risk basis (N, λ, r, Y) are shown in the Figure 2.

In particular, for the systematic risk factors, we observe that the finer the grid, the more sensitive the SU factors are to the ups and downs of the risk factors. However, it is clearly visible that the underlying time grid does not change the overall shape of the SU contributions. As the total surplus is given as the sum of the individual surpluses, these observations stay true when we look at the insurer's total surplus (see Figure 3). We note that the total surplus does not depend on the order of the risk basis.

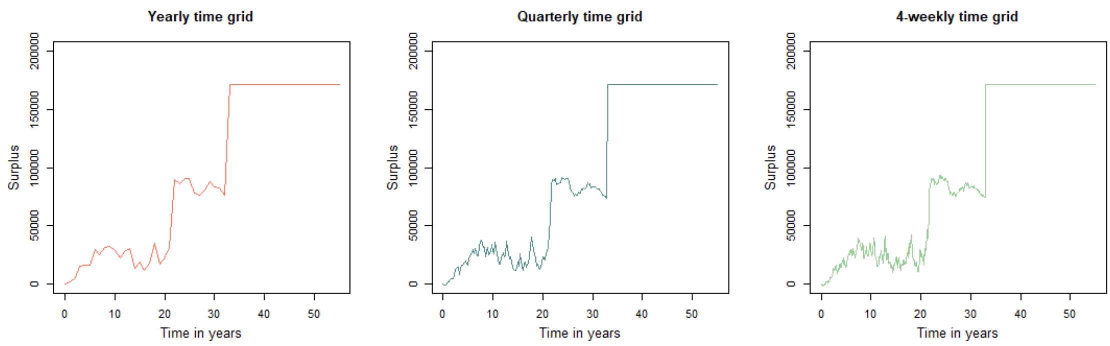

Figure 3: The total surplus for different time grids

To gain further insight into the effect of the SU grid on surplus decomposition, we focus on the surplus left when the contract ends. Figure 4 shows the contributed surplus of the different risk factors at the contract end.

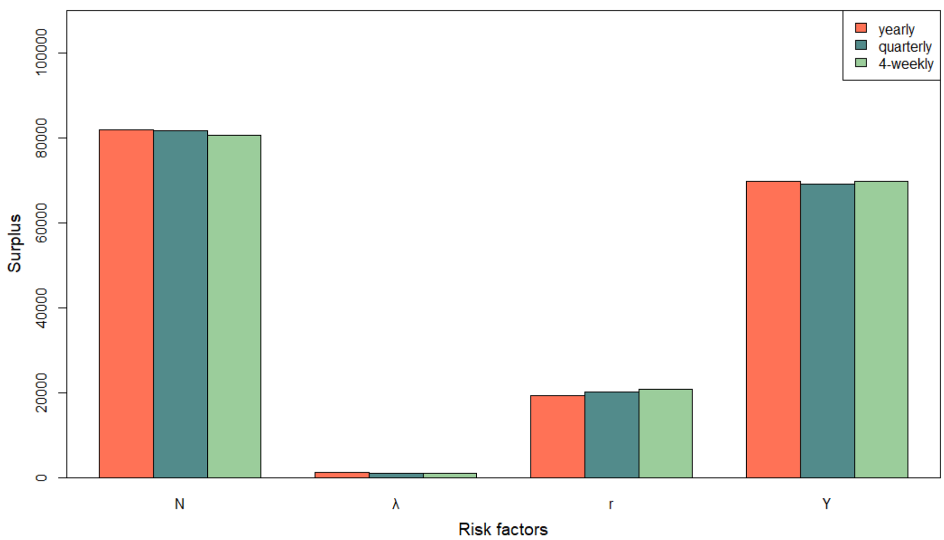

Figure 4: Impact of the time grid on the SU decomposition

The figure lets us surmise a certain stability of the SU decomposition with respect to the SU grid. A closer look at the numbers shows that the unsystematic biometric surplus decreases by 1.6%, while the systematic interest surplus increases by 7.9% and the systematic fund surplus decreases by 0.01% when moving from an annual to a 4-weekly updating frequency. The biggest change was in the systematic biometric surplus, which has been reduced by 15.7% due to the refinement of the SU grid.

Different orders of the risk basis

Not only the dependence of SU decompositions on the chosen time grid, but also the dependence of SU decompositions on the chosen update order of risk factors was a pivotal motivation for the introduction of the ISU decomposition principle. The applications in this thesis demonstrated that the ISU decomposition principle can help to overcome this drawback. However, if SU decompositions are used for approximating the ISU decomposition, the order dependence has to be taken into account. Therefore, one should aim for a SU grid that leads to neglectable order effects. In that regard, we have compared the order impact for the two update orders (N, λ, r, Y) and (Y, r, λ, N) with respect to the different SU grids (yearly, quarterly, 4-weekly). The relative deviations between the different orders are presented in Figure 5.

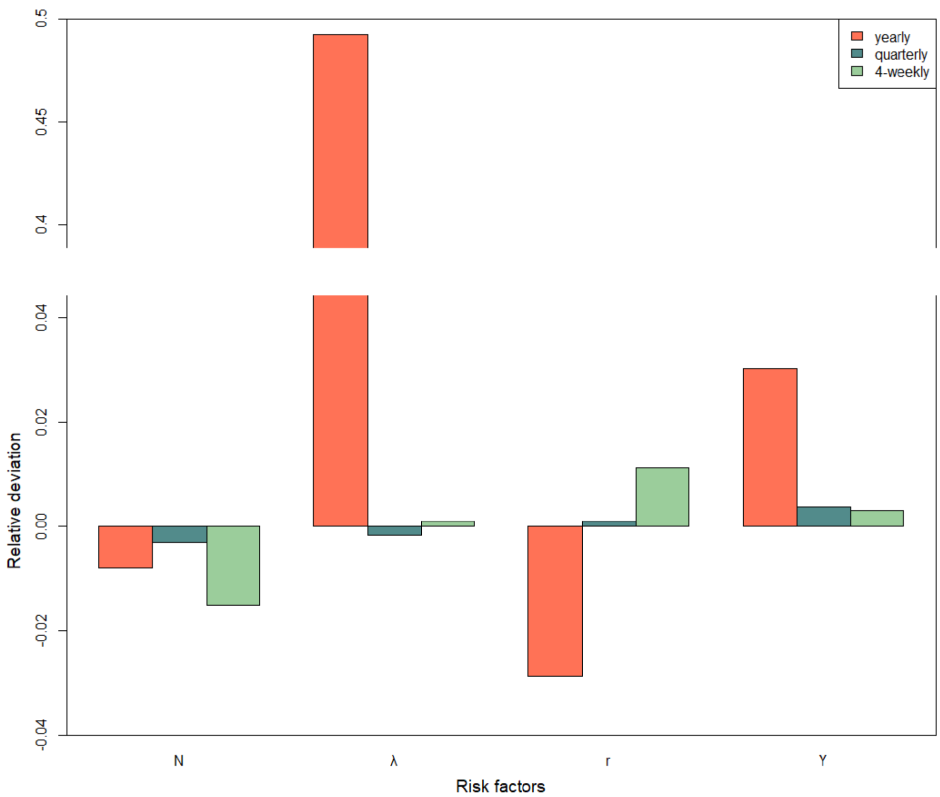

Figure 5: Impact of the update order on the SU decomposition

For the systematic biometric surplus and the fund surplus, we observe a huge harmonisation of the SU contributions for the different orders when the time grid is refined. In contrast, the unsystematic biometric surplus and the systematic risk surplus still show volatility during the refinement process. Nevertheless, it can be stated that the refinement of the grid already shows positive effects. For the practical implementation of the ISU decomposition, it is therefore recommended to first test different grids for the approximation in order to get a better insight into the volatility of the decomposition.

6 Conclusion and outlook

At the beginning of the thesis, the relevance of surplus decomposition in traditional life insurance has been highlighted. Especially, it has been pointed out that although several surplus decompositions exist in the literature, they are based on heuristic arguments. The lack of a general decomposition principle has made it difficult to compare and to extend the existing surplus decompositions. The thesis has remedied this deficiency by the introduction of the ISU decomposition principle (see Chapter 2). In particular, it was shown that all existing surplus decomposition formulas can be recovered from the ISU decomposition principle (see Chapter 3). This not only provides further evidence for the existing decomposition formulas, but also allows for a systematic comparison of decomposition formulas derived in separate model frameworks.

Moreover, the clarity and generality of the ISU decomposition principle paves the way to the inclusion of further risk factors. In an increasingly digital world, the available data, e.g. collected by wearables, allows for a more accurate risk profile of the policyholder. To benefit from this, modern life insurance products offer rewards to policyholders who provide their data and demonstrate a healthy lifestyle. The challenge for the life insurer is to find a fair reward, that is, to quantify the contribution of a particular policyholder behaviour to the total surplus. The ISU decomposition principle could also help here, so the study of ISU decompositions in life modelling frameworks that incorporate behavioural risks is an appealing task for future research.

In the pursuit of a suitable additive decomposition principle to help with the traditional surplus decompositions, the idea of SU stood out for its simplicity. Though the SU decomposition principle is highly regarded in economics literature, it strongly depends on the chosen update order and the chosen time grid. By pushing the update frequency to the limit, the ISU decomposition principle presented in this thesis helps to redress these shortcomings. Furthermore, the infinitesimal approach has also proven to be useful for the alternative OAT and ASU decomposition principles. For this reason, the thesis has not only proposed the ISU, IOAT and IASU decomposition principles, but also derived fundamental results on their relationship. While the underlying decomposition principles have their roots in economic science, the developed examples have been focused on actuarial modelling. Therefore, the application of the infinitesimal decomposition principles to the use cases of the SU, ASU and OAT decompositions principles (see Chapter 2) may be the subject of future studies (see Jetses & Christiansen, 2022).

Throughout the thesis, it has been stressed that the scope of the ISU decomposition principle goes far beyond the decomposition of traditional life insurance surplus. This claim has been supported by the application of the ISU decomposition principle to martingales, which has undergirded the great potential of the introduced decomposition approach in risk management. In particular, the main result in Chapter 4 has shown that the ISU and MRT decompositions match if the martingale representation exists and the risk basis satisfies a certain property (M). The verification of the property (M) for a number of examples that are common in actuarial modelling has provided evidence of its plausibility. However, the examples only allow for a limited dependency between risk factors. Generalising and extending the examples provided could be the focus of future research. Especially, the analysis of competing risks in life insurance has laid the foundation for exploring general multistate life insurance models.

As already briefly touched in Chapter 4, the underlying probability measure can be freely selected, which opens up new possibilities. In particular, the proposed switch to a conservative probability measure raises the question of whether and how martingale decompositions and traditional surplus decompositions are related to each other. This question could be likewise the subject of future work. Furthermore, instead of varying the probability measure, one might consider replacing the expectation operator with another appropriate risk measure. Follow-up research on the resulting ISU decompositions might give a valuable contribution in view of risk management.

Despite the many possible applications of the ISU decomposition principle, the actual use is also a question of numerical feasibility. Using the example of a fund-linked life insurance, this question has therefore been investigated in Chapter 5 of the thesis. There, not only integral representations for the SU contributions have been obtained but also a theoretical framework for approximating conditional expectations and their integrals has been developed, leveraging from MLMC methods. For both, the approximation of the ISU decomposition with SU decompositions as well as the approximation of SU contributions with MLMC methods, convergence results have been derived. However, future research is needed to analyse the numerical complexity of the approximation and to find error bounds.

The implementation with the statistical software R 4.4.2 (R Core Team, 2024) and the presentation of the numerical results have demonstrated the numerical feasibility of the ISU decomposition principle. While the expounded numerical analysis has fixated a single path, future studies might take the distribution of the SU contributions into account. In particular, this will shift the research focus more towards the efficiency of the approximation. Furthermore, the numerical example could be extended to more sophisticated multistate models, but this first requires a proof of the property (M) to obtain the convergence of the ISU decomposition (see above).

This thesis has introduced the ISU decomposition principle in such a general manner that its application is conceivable whenever an output needs to be decomposed additively and time-dynamically with respect to its input factors. The derived ISU decompositions of life insurance surplus processes have already proven its added value, but the wide range of applications remains to be explored.

Bibliography

Ahmad, J., Buchardt, K., & Furrer, C. (2022). Computation of bonus in multi-state life insurance. *ASTIN Bulletin: The Journal of the IAA, 52*(1), 291–331. https://doi.org/10.1017/asb.2021.32

Andersen, P. K., Borgan, Ø., Gill, R. D., & Keiding, N. (1993). *Statistical models based on counting processes*. Springer. https://doi.org/10.1007/978-1-4612-4348-9

Asmussen, S., & Steffensen, M. (2020). *Risk and insurance: A graduate text* (Vol. 96). Springer Nature. https://doi.org/10.1007/978-3-030-35176-2

Berger, A. (1939). *Mathematik der Lebensversicherung*. Verlag von Julius Springer.

Bernard, C., & Kwak, M. (2016). Semi-static hedging of variable annuities. *Insurance: Mathematics and Economics, 67*, 173–186. https://doi.org/10.1016/j.insmatheco.2016.01.004

Biagini, F. (2013). Evaluating hybrid products: The interplay between financial and insurance markets. In: R. Dalang, M. Dozzi & F. Russo (Eds.), *Seminar on stochastic analysis, random fields and applications VII. Progress in Probability* (Vol. 67, pp. 285–304). Birkhäuser. https://doi.org/10.1007/978-3-0348-0545-2_15

Biagini, F., Groll, A., & Widenmann, J. (2016). Risk minimization for insurance products via $\mathbb{F}$-doubly stochastic Markov chains. *Risks, 4*(3), 23. https://doi.org/10.3390/risks4030023

Biewen, M. (2014). A general decomposition formula with interaction effects. *Applied Economics Letters, 21*(9), 636–642. https://doi.org/10.1080/13504851.2013.879280

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources, 8*(4), 436–455. https://doi.org/10.2307/144855

BMF (2016). *Verordnung über die Mindestbeitragsrückerstattung in der Lebensversicherung (MindZV)*. https://www.gesetze-im-internet.de/mindzv_2016/BJNR083100016.html

Bromwich, T. J. (1926). *An introduction to the theory of infinite series* (2nd ed.). Macmillan.

Bruhn, K., & Lollike, A. S. (2020). Retrospective reserves and bonus. *Scandinavian Actuarial Journal, 2021*(6), 457–475. https://doi.org/10.1080/03461238.2020.1809509

Candland, A., & Lotz, C. (2014). Profit and loss attribution. In: P. Cadoni (Ed.), *Internal models and Solvency II: From regulation to implementation* (pp. 225–247). Risk Books.

CFO Forum (2016). *Presentation of analysis of earnings. Market Consistent Embedded Value Principles, Appendix B.* https://cfoforum.eu/mediaitem/3b8d66a9-7752-497e-8f71-2d0804ae9e73/CFO-Forum_MCEV_Principles_and_Guidance_April_2016.pdf

Chen, A., & Vigna, E. (2017). A unisex stochastic mortality model to comply with EU Gender Directive. *Insurance: Mathematics and Economics, 73*, 124–136. https://doi.org/10.1016/j.insmatheco.2017.01.007

Christiansen, M. C. (2021). Time-dynamic evaluations under non-monotone information generated by marked point processes. *Finance and Stochastics, 25*, 563–596. https://doi.org/10.1007/s00780-021-00456-5

Christiansen, M. C. (2022). On the decomposition of an insurer's profits and losses. *Scandinavian Actuarial Journal, 2023*(1), 51–70. https://doi.org/10.1080/03461238.2022.2079996

Cont, R., & Fournié, D.-A. (2013). Functional Itô calculus and stochastic integral representation of martingales. *The Annals of Probability, 41*(1), 109–133. https://doi.org/10.1214/11-AOP721

Cont, R., & Lu, Y. (2016). Weak approximation of martingale representations. *Stochastic Processes and their Applications, 126*(3), 857–882. https://doi.org/10.1016/j.spa.2015.10.002

Davis, M. H. A. (1976). The representation of martingales of jump processes. *SIAM Journal on Control and Optimization, 14*(4), 623–638. https://doi.org/10.1137/0314041

DiNardo, J., Fortin, N., & Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica, 64*(5), 1001–1044.

Doob, J. L. (1994). *Measure theory.* Springer. https://doi.org/10.1007/978-1-4612-0877-8

Dufresne, D. (2001). *A decomposition of actuarial surplus and applications* [working paper]. https://ozdaniel.com/A/DufresneActSurplus2001.pdf

Elliott, R. J. (1976). Stochastic integrals for martingales of a jump process with partially accessible jump times. *Zeitschrift für Wahrscheinlichkeitstheorie verwandte Gebiete, 36*, 213–226. https://doi.org/10.1007/BF00532546

European Parliament and the Council (2009). *Directive of the European Parliament and of the Council on the taking-up and pursuit of the business of insurance and reinsurance (Solvency II).* https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32009L0138

Falden, D. K., & Nyegaard, A. K. (2021). Retrospective reserves and bonus with policyholder behavior. *Risks, 9*(1), 15. https://www.mdpi.com/2227-9091/9/1/15

Fischer, T. (2004). *On the decomposition of risk in life insurance* [working paper]. Technische Universität Darmstadt. https://www.ressources-actuarielles.net/EXT/ISFA/1226.nsf/769998e0a65ea348c1257052003eb94f/77bfd4acb962f3eac1256f62006261a5/$FILE/On%20the%20decomposition%20of%20risk%20in%20life%20insurance.pdf

Flaig, S., & Junike, G. (2024). Profit and loss attribution: An empirical study. *European Actuarial Journal, 14*, 1013–1019. https://doi.org/10.1007/s13385-024-00380-w

Forster, O. (2017). *Analysis 2. Differentialrechnung im* $\mathbb{R}^n$*, gewöhnliche Differentialgleichungen* (11th ed.). Springer. https://doi.org/10.1007/978-3-658-19411-6

Fortin, N., Lemieux, T., & Firpo, S. (2011). Decomposition methods in economics. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 4, pp. 1–102). Elsevier. https://doi.org/10.1016/S0169-7218(11)00407-2

Frei, C. (2020). A new approach to risk attribution and its application in credit risk analysis. *Risks, 8*(2), 65. https://doi.org/10.3390/risks8020065

Friedman, E., & Moulin, H. (1999). Three methods to share joint costs or surplus. *Journal of Economic Theory, 87*(2), 275–312. https://doi.org/10.1006/jeth.1999.2534

Giles, M. B. (2008). Multilevel Monte Carlo path simulation. *Operations Research, 56*(3), 607–617. http://dx.doi.org/10.1287/opre.1070.04960

Giles, M. B. (2015). Multilevel Monte Carlo methods. *Acta Numerica, 24*, 259–328. https://doi.org/10.1017/S096249291500001X

Godin F., Hamel E., Gaillardetz P., & Edwin, H.-M. N. (2023). Risk allocation through shapley decompositions, with applications to variable annuities. *ASTIN Bulletin, 53*(2), 311–331. https://doi.org/10.1017/asb.2023.7

Harrison, J. M., & Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications, 11*(3), 215–260. https://doi.org/10.1016/0304-4149(81)90026-0

Heinrich, S. (1998). Monte Carlo complexity of global solution of integral equations. *Journal of Complexity, 14*(2), 151–175. https://doi.org/10.1006/jcom.1998.0471

Heinrich, S. (2000). The multilevel method of dependent tests. In: N. Balakrishnan, V. B. Melas & S. Ermakov, S. (Eds.), *Advances in stochastic simulation methods. Statistics for industry and technology* (pp. 47–61). Birkhäuser. https://doi.org/10.1007/978-1-4612-1318-5_4

Heinrich, S. (2001). Multilevel Monte Carlo methods. In: S. Margenov, J. Wasniewski & P. Yalamov (Eds.), *Large-scale scientific computing* (pp. 58–67). Springer.

Heinrich, S., & Sindambiwe, E. (1999). Monte Carlo complexity of parametric integration. *Journal of Complexity, 15*(3), 317–341.

Hoem, J. M. (1969). Markov chain models in life insurance. *Blätter DGVFM 9*, 91–107. https://doi.org/10.1007/BF02810082

Hortmann, M. (2006). *Produktfolgen* [lecture notes, Analysis 1]. Universität Bremen. https://user.informatik.uni-bremen.de/~michaelh/Lehrveranstaltungen/Ana1_WS06/Material/Produktfolgen.pdf

IASB (2017). *IFRS 17 Insurance Contracts.* https://www.ifrs.org/content/dam/ifrs/publications/pdf-standards/english/2022/issued/part-a/ifrs-17-insurance-contracts.pdf?bypass=on

Ikeda, N., & Watanabe, S. (1989). *Stochastic differential equations and diffusion processes* (2nd ed.). North-Holland/Kodansha.

ISO (2018). *ISO 31000:2018 - Risk Management - Guidelines.*https://www.iso.org/obp/ui/en/#iso:std:iso:31000:ed-2:v1:en

Jacobsen, M. (2006). *Point process theory and applications. Marked point and piecewise deterministic processes.* Birkhäuser. https://doi.org/10.1007/0-8176-4463-6

Jacod, J., & Shiryaev, A. N. (2003). *Limit theorems for stochastic processes* (2nd ed., Vol. 288). Springer. https://doi.org/10.1007/978-3-662-05265-5

Jakubowski, J., & Niewęgłowski, M. (2008). Pricing bonds and CDS in the model with rating migration induced by a Cox process. *Advances in Mathematics of Finance. Banach Center Publications, 83*, 159–182. https://doi.org/10.4064/bc83-0-10

Jakubowski, J., & Niewęgłowski, M. (2010). A class of $\mathbb{F}$-doubly stochastic Markov chains. *Electronic Journal of Probability, 15*, 1743–1771. https://doi.org/10.1214/EJP.v15-815

Jetses, J. (2018). *Martingalzerlegung von Verbindlichkeiten in der Personenversicherung* [master's thesis]. Carl von Ossietzky Universität Oldenburg. https://uol.de/f/5/inst/mathe/personen/Jetses_2018_Martingalzerlegung_von_Verbindlichkeiten_in_der_Personenversicherung.pdf

Jetses, J., & Christiansen, M. C. (2022). A general surplus decomposition principle in life insurance. *Scandinavian Actuarial Journal, 2022*(10), 901–925. https://doi.org/10.1080/03461238.2022.2049636

Junike, G., Stier, H., & Christiansen, M. C. (2024). Profit and loss decomposition in continuous time and approximations. arXiv preprint. https://doi.org/10.48550/arXiv.2212.06733

Karabey, U., Kleinow, T., & Cairns, J. G. C. (2014). Factor risk quantification in annuity models. *Insurance: Mathematics and Economics, 58*, 34–45. https://doi.org/10.1016/j.insmatheco.2014.06.00

Klebaner, F. C. (2005). *Introduction to stochastic calculus with applications* (2nd ed.). Imperial College Press.

Klenke, A. (2020). *Probability theory: A comprehensive course* (3rd ed.). Springer. https://doi.org/10.1007/978-3-030-56402-5

Kloeden, P. E., & Platen, E. (1992). *Numerical solution of stochastic differential equations*. Springer. https://doi.org/10.1007/978-3-662-12616-5

Kunita, H. (2004). Representation of martingales with jumps and applications to mathematical finance. In H. Kunita, S. Watanabe & Y. Takahashi (Eds.), *Stochastic analysis and related topics in Kyoto. In honour of Kiyoshi Itô* (Vol. 41, pp. 209–232). Mathematical Society of Japan. https://doi.org/10.2969/aspm/04110209

Kuo, H. (2006). *Introduction to stochastic integration*. Springer. https://doi.org/10.1007/0-387-31057-6

Lidstone, G. J. (1905). Changes in pure premium values consequent upon variations in the rate of interest or rate of mortality. *Journal of the Institute of Actuaries, 39*, 209–252.

Malliavin, P. (1978). Stochastic calculus of variations and hypoelliptic operators. In K. Itô (Ed.), *Proceedings of the International Symposium on Stochastic Differential Equations, Kyoto, 1976* (pp. 195–263). Wiley.

Milbrodt, H., & Helbig, M. (1999). *Mathematische Methoden der Personenversicherung*. Walter de Gruyter.

Møller, T., & Steffensen, M. (2007). *Market-valuation methods in life and pension insurance*. Cambridge University Press. https://doi.org/10.1017/CBO9780511543289

Norberg, R. (1999). A theory of bonus in life insurance. *Finance and Stochastics, 3*, 373–390.

Norberg, R. (2001). On bonus and bonus prognoses in life insurance. *Scandinavian Actuarial Journal, 2001*(2), 126–147. https://doi.org/10.1080/03461230152592773

Nualart, D. (2006). *The Malliavin calculus and related topics* (2nd ed.). Springer. https://doi.org/10.1007/3-540-28329-3.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review, 14*(3), 693–709. https://doi.org/10.2307/2525981

Protter, P. E. (2005). *Stochastic integration and differential equations* (2nd ed.). Springer. https://doi.org/10.1007/978-3-662-10061-5

R Core Team (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ramlau-Hansen, H. (1988). The emergence of profit in life insurance. *Insurance: Mathematics and Economics, 7*, 225–236.

Ramlau-Hansen, H. (1991). Distribution of surplus in life insurance. *ASTIN Bulletin: The Journal of the IAA, 21*(1), 57–71. https://doi.org/10.2143/AST.21.1.2005401

Rao, M. M., & Swift, R. J. (2006). *Probability theory with applications* (2nd ed.). Springer. https://doi.org/10.1007/0-387-27731-5

Saxer, W. (1955). *Versicherungsmathematik. Erster Teil.* Springer. https://doi.org/10.1007/978-3-642-88629-4

Schilling, K., Bauer, D., Christiansen, M. C., & Kling, A. (2020). Decomposing dynamic risks into risk components. *Management Science, 66*(12), 5485–6064. https://doi.org/10.1287/mnsc.2019.3522

Shapley, L. S. (1953). A value for n-person games. In H. W. Kuhn, A. W. Tucker (Eds.), *Contributions to the Theory of Games. Annals of Mathematical Studies* (Vol. 28, pp. 307-317). Princeton University Press. https://doi.org/10.1515/9781400881970-018

Shorrocks, A. F. (2013). Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. *The Journal of Economic Inequality, 11*, 99–126. https://doi.org/10.1007/s10888-011-9214-z

Shubik, M. (1962). Incentives, decentralized control, the assignment of joint costs and internal pricing. *Management Science, 8*(3), 325–343.

Simonsen, W. (1970). *Forsikringsmatematik*, hefte III. Københavns Universitets Fond til tilvejebringelse af læremidler.

Spangler, M. (2018). *Modelling german covered bonds.* Springer. https://doi.org/10.1007/978-3-658-23915-2

Sprumont, Y. (1998). Ordinal cost sharing. *Journal of Economic Theory, 81*(1), 126–162. https://doi.org/10.1006/jeth.1998.2408

Steffensen, M. (2001). *On valuation and control in life and pension insurance* [doctoral dissertation], University of Copenhagen. https://noter.math.ku.dk/phd01ms.pdf

Sundararajan, M. & Najmi, A. (2020). The many Shapley values for model explanation. arXiv preprint. https://doi.org/10.48550/arXiv.1908.08474

Sverdrup. E. (1969). *Noen forsikringsmatematiske emner. Statisticat memoirs No. 1. Insitute of Mathematics, University of Oslo. https://www.duo.uio.no/handle/10852/55912*

Ter Horst, H. J. (1984). Riemann-Stieltjes and Lebesgue-Stieltjes integrability. *The American Mathematical Monthly, 91*(3), 551–559. https://doi.org/10.1080/00029890.1984.11971492

Werner, D. (2018). *Funktionalanalysis* (8th ed.). Springer. https://doi.org/10.1007/978-3-662-55407-4

Wuorikoski, T. (2018). *Relationship of risk and quality management.* In I. Kīsnica (Ed.), *Organization and individual security* (pp. 169–176). Nordplus.

Zitkovic, G. (2015). *Conditional expectation* [lecture notes, Theory of Probability I]. University of Texas at Austin. https://web.ma.utexas.edu/users/gordanz/notes/conditional_expectation.pdf

Appendix

A.1 Proof of the SU decomposition in the time-discrete case

In the setting of Example 3.6, the functional H in (3.13) takes the form

$$H(\overline{\Phi}, \overline{\Lambda}_{ad}, \overline{\Lambda}_{as}) = \sum_{l=0}^{T} e^{-\int_0^l \overline{\phi}(u)\mathrm{d}u} \, {}_l\overline{p}_x b_l + \sum_{l=1}^{T} e^{-\int_0^l \overline{\phi}(u)\mathrm{d}u} \, {}_{l-1}\overline{p}_x (\overline{q}_{x+l} d_l + \overline{r}_{x+l} s_l)$$

Furthermore, for the risk basis $X = (\Phi - \Phi^*, \Lambda_{ad} - \Lambda_{ad}^*, \Lambda_{as} - \Lambda_{as}^*)$, the mapping ϱ is given by $\varrho(X^t) = -H((\Phi^*, \Lambda_{ad}^*, \Lambda_{as}^*) + X^t)$. We prove the three equations consecutively.

i) We have that

$$\begin{aligned}
&U(k+1,k,k) - U(k,k,k) \\
&= \varrho(\Phi^{k+1}, \Lambda_{ad}^k, \Lambda_{as}^k) - \varrho(\Phi^k, \Lambda_{ad}^k, \Lambda_{as}^k) \\
&= e^{-\int_0^{k+1} \phi(u)\,\mathrm{d}u} \, {}_kp_x((1+i_k)V_a^*(k) - p_{x+k}^* \, b_{k+1} - q_{x+k}^* \, d_{k+1} - r_{x+k}^* \, s_{k+1}) \\
&\quad - e^{-\int_0^{k+1} \phi(u)\,\mathrm{d}u} \, {}_kp_x p_{x+k}^* \, V_a^*(k+1)
\end{aligned}$$

Since

$$- p_{x+k}^* \, b_{k+1} - q_{x+k}^* \, d_{k+1} - r_{x+k}^* \, s_{k+1} - p_{x+k}^* \, V_a^*(k+1) = -(1+i_k^*)V_a^*(k),$$

we get the first equation.

ii) With similar calculations as in i) we get

$$\begin{aligned}
&U(k+1,k+1,k) - U(k,k,k) \\
&= \varrho(\Phi^{k+1}, \Lambda_{ad}^{k+1}, \Lambda_{as}^k) - \varrho(\Phi^k, \Lambda_{ad}^k, \Lambda_{as}^k) \\
&= -e^{-\int_0^{k+1} \phi(u)\,\mathrm{d}u} \, {}_kp_x((1 - q_{x+k} - r_{x+k}^*)b_{k+1} + q_{x+k} \, d_{k+1} + r_{x+k}^* \, s_{k+1}) \\
&\quad - e^{-\int_0^{k+1} \phi(u)\,\mathrm{d}u} \, {}_kp_x(1 - q_{x+k} - r_{x+k}^*)V_a^*(k+1) + e^{-\int_0^{k} \phi(u)\,\mathrm{d}u} \, {}_kp_x \, V_a^*(k) \\
&= e^{-\int_0^{k+1} \phi(u)\,\mathrm{d}u} \, {}_kp_x(V_a(k+1-) - d_{k+1})(q_{x+k} - q_{x+k}^*) \\
&\quad + e^{-\int_0^{k+1} \phi(u)\,\mathrm{d}u} \, {}_kp_x \, V_a^*(k)\big(i_k - i_k^*\big)
\end{aligned}$$

The second equality follows then by substracting $U(k+1,k,k) - U(k,k,k)$ (see i)) from $U(k+1,k+1,k) - U(k,k,k)$.

iii) For the third equality, we can use the results from i) and ii) to obtain

$$\begin{aligned}
&U(k+1,k+1,k+1) - U(k+1,k+1,k) \\
&= R(k+1) - R(k) \\
&\quad - (U(k+1,k+1,k) - U(k+1,k,k)) - (U(k+1,k,k) - U(k,k,k)) \\
&= e^{-\int_0^{k+1} \phi(u)\,\mathrm{d}u} \, {}_kp_x\big(V_a^*(k+1-) - s_{k+1}\big)\,\big(r_{x+k} - r_{x+k}^*\big).
\end{aligned}$$

A.2 Technical results

Analogously to Λ_M^*, let $\overline{\Lambda}_M$ denote the matrix-valued process $\overline{\Lambda}_M = (\overline{\Lambda}_{jk})_{jk}$ with $\overline{\Lambda}_{jj} := -\sum_{k:k\neq j}\overline{\Lambda}_{jk}$, and define Λ'_M likewise.

Lemma A.2.1. *Let $(\overline{\Phi},\overline{\Lambda})$ be a valuation basis.*

a) Let $\overline{\kappa}$ be the solution of the stochastic differential $d\overline{\kappa}(t) = \overline{\kappa}(t-)\,d\overline{\Phi}(t)$ with $\overline{\kappa}(0) = 1$. Then it holds that

$$d\left(\frac{1}{\overline{\kappa}(t)}\right) = -\frac{1}{\overline{\kappa}(t-)}\,d\tilde{\overline{\Phi}}(t),$$

where $\tilde{\overline{\Phi}}(t) = \overline{\Phi}(t) - [\overline{\Phi},\overline{\Phi}]^c(t) - \sum_{0<s\leqslant t}(1+\Delta\overline{\Phi}(s))^{-1}(\Delta\overline{\Phi}(s))^2$.

b) Let $\overline{p}(s,t)$ be the solution of the matrix-valued stochastic differential equation $\overline{p}(s,dt) = \overline{p}(s,t-)\,d\overline{\Lambda}_M(t)$ with $\overline{p}(s,s) = \mathbb{I}$. Assume that $(\mathbb{I}+\Delta\overline{\Lambda}_M(t))^{-1}$ exists for all $t>0$. Then $\overline{p}(s,t)$ is invertible, and the inverse $\overline{q}(s,t)$ solves the SDE

$$\overline{q}(s,dt) = -(dG(t))\overline{q}(s,t-) = -(d\overline{\Lambda}_M(t))\overline{q}(s,t),$$

where $G(t) = \overline{\Lambda}_M(t) - \sum_{0<s\leqslant t}(\Delta\overline{\Lambda}_M(s))^2(I+\Delta\overline{\Lambda}_M(s))^{-1}$.

Proof. a) Due to the properties of a valuation basis, $\tilde{\overline{\Phi}}$ is a well-defined semimartingale. Thus, with Theorem V.10.63 of Protter (2005), the assertion follows.

b) For applying Theorem V.10.63 of Protter (2005) later again, we firstly have to show that G is a well-defined semimartingale. Since $\overline{\Lambda}$ is a càdlàg finite variation process, it suffices to show $(\sum_{0<s\leqslant t}(\Delta\overline{\Lambda}_M(s))^2(I+\Delta\overline{\Lambda}_M(s))^{-1})_{jk} < \infty$ for all $t>0$ and j,k. Let $\|\cdot\|$, defined by $\|A\| = n\cdot\max_{j,k}|a_{jk}|$ for a matrix $A = (a_{jk})_{jk}\in\mathbb{R}^{n\times n}$, denote the maximum norm on $\mathbb{R}^{n\times n}$. If $\|\Delta\overline{\Lambda}_M(t)\|\leqslant 1/2$, then it holds

$$\|(I+\Delta\overline{\Lambda}_M(t))^{-1}\| \leqslant \frac{1}{1-\|\Delta\overline{\Lambda}_M(t)\|} \leqslant 2,$$

see for example Werner (2018, Theorem II.1.12). Using this upper bound, the subadditity and the submultiplicity of the norm, we get

$$\begin{aligned}
&\Bigg\|\sum_{0<s\leqslant t}(\Delta\overline{\Lambda}_M(s))^2(I+\Delta\overline{\Lambda}_M(s))^{-1}\Bigg\| \\
&\leqslant \sum_{\substack{0<s\leqslant t\\ \|\Lambda_M(s)\|>1/2}}\Big\|(\Delta\overline{\Lambda}_M(s))^2(I+\Delta\overline{\Lambda}_M(s))^{-1}\Big\| + \sum_{\substack{0<s\leqslant t\\ \|\overline{\Lambda}_M(s)\|\leqslant 1/2}}\Big\|(\Delta\overline{\Lambda}_M(s))^2(I+\Delta\overline{\Lambda}_M(s))^{-1}\Big\| \\
&\leqslant \sum_{\substack{0<s\leqslant t\\ \|\overline{\Lambda}_M(s)\|>1/2}}\Big\|(\Delta\overline{\Lambda}_M(s))^2(I+\Delta\overline{\Lambda}_M(s))^{-1}\Big\| + \sum_{\substack{0<s\leqslant t\\ \|\overline{\Lambda}_M(s)\|\leqslant 1/2}}\Big\|\Delta\overline{\Lambda}_M(s)\Big\|.
\end{aligned}$$

The first sum in the latter expression is finite, since $\|\Delta\overline{\Lambda}_M(s)\| > 1/2$ occurs only for finitely many $s \in [0,t]$. For the second term, observe that

$$\sum_{0<s\leqslant t} \|\overline{\Lambda}_M(s)\| \leqslant \sum_{j,k} \sum_{0<s\leqslant t} |\Delta\overline{\Lambda}_{jk}(s)| < \infty,$$

on account of the fact that $\overline{\Lambda}$ is a finite variation process. Thus, G is a well-defined semimartingale.

For a matrix-valued semimartingale Z, let $\mathcal{E}(Z)$ denote the (matrix-valued) exponential of Z and let $\mathbb{E}^R(Z)$ denote the (matrix-valued) right-stochastic exponential of Z (see Protter, 2005, Chapter V). By applying Theorem V.10.63 of Protter (2005), we get

$$\mathcal{E}(F)(t)\mathcal{E}^R(\overline{\Lambda}_M^\top)(t) = \mathbb{I}$$

for $F(t) = -\overline{\Lambda}_M^\top(t) + \sum_{0<s\leqslant t}(I + \Delta\overline{\Lambda}_M^\top(s))^{-1}(\Delta\overline{\Lambda}_M^\top(s))^2$. Because of $\mathcal{E}^R(Z) = \mathcal{E}(Z^\top)^\top$ and $F^\top = -G$, the latter equation is equivalent to

$$\mathcal{E}(\overline{\Lambda}_M)(t)\mathcal{E}^R(-G)(t) = \mathbb{I},$$

which proves the first equation of the assertion. In particular, we verified that $\overline{q}(s,t) - \overline{q}(s,t-) = -(\Delta G(t))\overline{q}(s,t-)$, which implies that

$$\begin{aligned} -(\Delta\overline{\Lambda}_M(t))\overline{q}(s,t) &= -(\Delta\overline{\Lambda}_M(t))\overline{q}(s,t-) + (\Delta\overline{\Lambda}_M(t))(\Delta G(t))\overline{q}(s,t-) \\ &= -\Delta\overline{\Lambda}_M(t))(\mathbb{I} - \Delta G(t))\overline{q}(s,t-) \\ &= -(\Delta G(t))\overline{q}(s,t-). \end{aligned}$$

Thus, the second equation of the assertion is also true.

□

Lemma A.2.2. *Let $(\Phi', (\Lambda'_{jk})_{jk:j\neq k})$, $(\overline{\Phi}, (\overline{\Lambda}_{jk})_{jk:j\neq k})$ be valuation bases.*

a) Let $d\kappa'(t) = \kappa'(t-)d\Phi'(t)$ with $\kappa'(0) = 1$ and $d\overline{\kappa}(t) = \overline{\kappa}(t-)d\overline{\Phi}(t)$ with $\kappa'(0) = 1$. Then it holds that

$$d\left(\frac{\kappa'(t)}{\overline{\kappa}(t)}\right) = \frac{\kappa'(t-)}{\overline{\kappa}(t-)}\Big(d\Phi' - d\tilde{\overline{\Phi}}(t) - d[\Phi', \tilde{\overline{\Phi}}](t)\Big),$$

where $\tilde{\overline{\Phi}}(t) = \overline{\Phi}(t) - [\overline{\Phi}, \overline{\Phi}]^c(t) - \sum_{0<s\leqslant t}(1 + \Delta\overline{\Phi}(s))^{-1}(\Delta\overline{\Phi}(s))^2$.

b) Let $p'(s,dt) = p'(s,t-)d\Lambda'_M(t)$ with $p'(s,s) = \mathbb{I}$ and $\overline{p}(s,dt) = \overline{p}(s,t-)d\overline{\Lambda}_M(t)$ with $\overline{p}(s,s) = \mathbb{I}$. Suppose that $\overline{p}(s,t)$ is invertible with inverse $\overline{q}(s,t)$. Then it holds that

$$d_t\left(p'(s,t)\overline{q}(s,t)\right) = p'(s,t-)d(\Lambda'_M - \overline{\Lambda}_M)(t)\overline{q}(s,t).$$

Proof. a) Integration by parts (Protter, 2005, Corollary II.6.2) and Lemma A.2.1a) yield

$$\begin{aligned}
\mathrm{d}\left(\frac{\kappa'(t)}{\overline{\kappa}(t)}\right) &= \kappa'(t-)\mathrm{d}\left(\frac{1}{\overline{\kappa}(t)}\right) + \frac{1}{\overline{\kappa}(t-)}\mathrm{d}\kappa'(t) + \mathrm{d}\left[\frac{1}{\overline{\kappa}}, \kappa'\right](t) \\
&= -\frac{\kappa'(t-)}{\overline{\kappa}(t-)}\mathrm{d}\widetilde{\overline{\Phi}}(t) + \frac{\kappa'(t-)}{\overline{\kappa}(t-)}\mathrm{d}\Phi'(t) - \frac{\kappa'(t-)}{\overline{\kappa}(t-)}\mathrm{d}[\widetilde{\overline{\Phi}}, \Phi'](t).
\end{aligned}$$

b) Integration by parts (Protter, 2005, Corollary II.6.2) and Lemma A.2.1b) yield

$$\begin{aligned}
\mathrm{d}_t\left(p'(s,t)\overline{q}(s,t)\right) &= p'(s,t-)\overline{q}(s,\mathrm{d}t) + p'(s,\mathrm{d}t)\overline{q}(s,t-) + \mathrm{d}\left[p'(s,\cdot), \overline{q}(s,\cdot)\right](t) \\
&= -p'(s,t-)(\mathrm{d}\overline{\Lambda}_M(t))\overline{q}(s,t) + p'(s,t-)(\mathrm{d}\Lambda'_M(t))\overline{q}(s,t) \\
&= p'(s,t-)\mathrm{d}(\Lambda'_M - \overline{\Lambda}_M)(t)\overline{q}(s,t).
\end{aligned}$$

□

Lemma A.2.3. *Let ξ be a $\mathcal{G}_T$-measurable, square-integrable random variable and let $\langle M_i, M_j\rangle = 0$, $i \neq j$. Suppose there exists a martingale representation*

$$\xi - \mathbb{E}[\xi|\mathcal{G}_0] = \sum_{i=1}^{m}\int_0^t H_i(s)\,dM_i(s)$$

for $\mathbb{G}$-predictable integrands H_i, $i = 1, \ldots, m$. Then the integrands H_i are almost surely unique with respect to $\mathbb{P}\otimes\langle M_i, M_i\rangle$. Furthermore, it holds

$$\mathbb{E}\left[\int_0^t H_i^2(u)\,d\langle M_i, M_i\rangle(u)\right] < \infty,\ \ t \in [0,T].$$

Proof. The proof follows the ideas from Schilling et al. (2020, Proposition 3.5). Suppose there exist $\mathbb{G}$-predictable integrands H_i, $\widetilde{H}_i$, $i = 1, \ldots, m$, such that

$$\xi - \mathbb{E}[\xi|\mathcal{G}_0] = \sum_{i=1}^{m}\int_0^t H_i(s)\mathrm{d}M_i(s) = \sum_{i=1}^{m}\int_0^t \widetilde{H}_i(s)\mathrm{d}M_i(s). \tag{A.1}$$

With the Itô isometry, we get

$$0 = \mathbb{E}\left[\left(\sum_{i=1}^{m}\int_0^t (H_i - \widetilde{H}_i)(s)\mathrm{d}M_i(s)\right)^2\right] = \sum_{i=1}^{m}\mathbb{E}\left[\int_0^t ((H_i - \widetilde{H}_i)(s))^2\mathrm{d}\langle M_i, M_i\rangle(s)\right].$$

This implies $H_i = \widetilde{H}_i$ almost surely with respect to $\mathbb{P}\otimes\langle M_i, M_i\rangle$. Again, with Itô's isometry and the square-integrability of ξ, we also have

$$\mathbb{E}[(\xi - \mathbb{E}[\xi|\mathcal{G}_0])^2] = \sum_{i=1}^{m}\mathbb{E}\left[\int_0^t H_i(s)^2\mathrm{d}\langle M_i, M_i\rangle(s)\right] < \infty.$$

□

A.3 Numerical example - analysis of SU contributions

In the following, we calculate the conditional expectations

$$\mathbb{E}[\xi_i|\mathcal{G}^J_{t_k,t_{k+1}}],\ i = 1, \dots, 4, \tag{A.2}$$

for grid points $t_k, t_{k+1} \in \mathcal{T}$ and subsets $J \subseteq \{1, \dots, m\}$. The increments then lead us to the contributions of the SU decomposition. Recall from Section 4.1 the definition of the filtrations $\mathbb{G}$, $\mathbb{G}^J$ and the family of sub-σ-algebras $(\mathcal{G}^J_{s,t})_{s,t}$, which refer to the entire risk basis $X = (N, \lambda, r, Y)$. Furthermore, for a subset $J \subseteq \{\lambda, r, Y\}$, the filtrations $\mathbb{H}$, $\mathbb{H}^J$ and the family of sub-σ-algebras $(\mathcal{H}^J_{s,t})_{s,t}$ refer to the systematic risks $\Theta = (\lambda, r, Y)$. To derive the integral representations for the SU decompositions, we need to separate the different risk factors from each other. Therefore, we need the following two lemmas.

Lemma A.3.1. *Let $J \subseteq \{N, \lambda, r, Y\}$ and $t_k, t_{k+1} \in \mathcal{T}$. For an integrable random variable $Z \in \mathcal{H}_T$, it holds*

$$\mathbb{E}[Z|\mathcal{G}^J_{t_k,t_{k+1}}] = \mathbb{E}[Z|\mathcal{H}^{\widetilde{J}}_{t_k,t_{k+1}}],$$

where $\widetilde{J} = J\backslash\{N\}$.

Proof. The result follows from Lemma 4.15, the fact that $\pi^\lambda \subseteq \mathcal{T}$ and Proposition 13 (*iv*) of Rao and Swift (2006, Chapter 3). □

The previous lemma enables us to separate the unsystematic risk from the systematic risk drivers. The next next lemma will allow us to divide the systematic risks.

Lemma A.3.2. *Let $J \subseteq \{\lambda, r, Y\}$ and $t_k, t_{k+1} \in \mathcal{T}$, then it holds*

$$(\mathcal{F}^\lambda_T \perp\!\!\!\perp \mathcal{F}^r_T \perp\!\!\!\perp \mathcal{F}^Y_T)|\mathcal{H}^J_{t_k,t_{k+1}}.$$

In particular, if $Z_1 \in \mathcal{F}^\lambda_T$, $Z_2 \in \mathcal{F}^r_T$ and $Z_3 \in \mathcal{F}^Y_T$ and their product are integrable random variables, then the equality

$$\mathbb{E}[Z_1Z_2Z_3|\mathcal{H}^J_{t_k,t_{k+1}}] = \mathbb{E}[Z_1|\mathcal{H}^J_{t_k,t_{k+1}}]\mathbb{E}[Z_2|\mathcal{H}^J_{t_k,t_{k+1}}]\mathbb{E}[Z_3|\mathcal{H}^J_{t_k,t_{k+1}}]$$

is true.

Proof. First, we show that $\mathcal{H}^\lambda_T \perp\!\!\!\perp \mathcal{H}^{r,Y}_T|\mathcal{H}^J_{t_k,t_{k+1}}$. Without loss of generality, we assume that $\lambda \in J$, otherwise the arguments are the same with $\mathcal{F}^\lambda_{t_{k+1}}$ substituted by $\mathcal{F}^\lambda_{t_k}$. With Rao and Swift (2006, Chapter 3, Proposition 13), it is sufficient to demonstrate

$$\mathbb{P}(B_\lambda|\mathcal{H}^J_{t_k,t_{k+1}} \vee \mathcal{H}^{r,Y}_T) = \mathbb{P}(B|\mathcal{H}^J_{t_k,t_{k+1}}),\ B_\lambda \in \mathcal{F}^\lambda_T. \tag{A.3}$$

Now, together with the independence of λ, r and Y, the Proposition 10.5.9 (9) of Zitkovic

(2015) helps us to derive

$$\mathbb{P}(B|\mathcal{H}^J_{t_k,t_{k+1}} \vee \mathcal{H}^{r,Y}_T) = \mathbb{P}(B|\mathcal{F}^\lambda_{t_{k+1}}) = \mathbb{P}(B|\mathcal{H}^J_{t_k,t_{k+1}}),$$

which proves (A.3). Following the same steps, one can also show $\mathcal{H}^r_T \perp\!\!\!\perp \mathcal{H}^Y_T|\mathcal{H}^J_{t_k,t_{k+1}}$.

In total, we get

$$\begin{aligned}\mathbb{P}(B_\lambda \cap B_r \cap B_Y|\mathcal{H}^J_{t_k,t_{k+1}}) &= \mathbb{P}(B_\lambda|\mathcal{H}^J_{t_k,t_{k+1}})\mathbb{P}(B_r \cap B_Y|\mathcal{H}^J_{t_k,t_{k+1}}) \\ &= \mathbb{P}(B_\lambda|\mathcal{H}^J_{t_k,t_{k+1}})\mathbb{P}(B_r|\mathcal{H}^J_{t_k,t_{k+1}})\mathbb{P}(B_Y|\mathcal{H}^J_{t_k,t_{k+1}}),\end{aligned}$$

which yields the assertion. The second part follows from standard arguments like the approximation of non-negative random variables with increasing simple random variables, the application of the monotone convergence theorem and the representation of general random variables as the difference of non-negative random variables (see e.g. Klenke, 2020, Proof of Theorem 5.4). □

Lemma A.3.3. *Let $\overline{J} \subseteq \{\lambda, r, Y\}$, $J = \overline{J} \cup \{N\}$ and $t_k, t_{k+1} \in \mathcal{T}$. Suppose Z is a continuous $\mathbb{F}$-adapted process with*

$$\sup_{\omega\in\Omega, s\in[t_k,t_{k+1}]} |Z(s)| < \infty.$$

Then it holds

$$\mathbb{E}\left[\int_{t_k}^{t_{k+1}} Z(s)dN(s)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] = \int_{t_k}^{t_{k+1}} \mathbb{E}[Z(s)|\mathcal{G}^J_{t_k,t_{k+1}}]dN(s).$$

Proof. Since N is a finite variation process and Z has continuous paths, we can interpret the integral in the Riemann-Stieltjes sense (see Ter Horst, 1984, Theorem C). Thus, let $(P_n)_n$ denote a vanishing sequence of partitions on $[t_k, t_{k+1}]$, such that it almost surely holds

$$\lim_{n\to\infty} \sum_{s_j,s_{j+1}\in P_n} Z(s_j)(N(s_{j+1}) - N(s_j)) = \int_{t_k}^{t_{k+1}} Z(s)\mathrm{d}N(s)$$

Let $m := \max_{s\in[t_k,t_{k+1}]} |Z(s)| < \infty$, then we almost surely have

$$\left|\sum_{s_j,s_{j+1}\in P_n} Z(s_j)(N(s_{j+1}) - N(s_j))\right| \leqslant m \sum_{s_j,s_{j+1}\in P_n} |N(s_{j+1}) - N(s_j)| \leqslant m.$$

Consequently, we can apply the dominated convergence theorem to get

$$\begin{aligned}\mathbb{E}\left[\int_{t_k}^{t_{k+1}} Z(s)\mathrm{d}N(s)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] &= \mathbb{E}\left[\lim_{n\to\infty} \sum_{s_j,s_{j+1}\in P_n} Z(s_j)(N(s_{j+1}) - N(s_j))\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\ &= \lim_{n\to\infty} \mathbb{E}\left[\sum_{s_j,s_{j+1}\in P_n} Z(s_j)(N(s_{j+1}) - N(s_j))\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\end{aligned}$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \sum_{s_j, s_{j+1}\in P_n} \mathbb{E}[Z(s_j)|\mathcal{G}^J_{t_k,t_{k+1}}](N(s_{j+1}) - N(s_j)) \\
&= \int_{t_k}^{t_{k+1}} \mathbb{E}[Z(s)|\mathcal{G}^J_{t_k,t_{k+1}}]\mathrm{d}N(s),
\end{aligned}
$$

where the latter equality uses $\sup_{\omega\in\Omega, s\in[t_k,t_{k+1}]} |Z(s)| < \infty$ to derive the continuity of $s \mapsto \mathbb{E}[Z(s)|\mathcal{G}^J_{t_k,t_{k+1}}]$. □

For both orders, (N, λ, r, Y) and (Y, r, λ, N), we simplify the conditional expectations (A.2) by investigating the different claim components one after the other. The main tools will be Lemma A.3.1, Lemma A.3.2 and the martingale property of M_N with respect to $\overline{\mathbb{G}}$, where $\overline{\mathbb{G}} = (\overline{\mathcal{G}}_t)_{t\geqslant 0}$ is the filtration given by $\overline{\mathcal{G}}_t = \mathcal{F}^{\Theta}_T \vee \mathcal{F}^N_t$ (see Subsection 4.2.5). The death cover claim will further require the application of Lemma A.3.3. For a better readability we avoid repeating the arguments in every line.

A.3.1 Analysis of conditional expectations for the order $\{N, \lambda, r, Y\}$

Savings account ($i = 1$):

For $J = \varnothing$ and $t_k < \gamma$, it holds

$$
\begin{aligned}
\mathbb{E}\left[I(\gamma)v(0,\gamma)V(\gamma)\big|\mathcal{G}^J_{t_k,t_{k+1}}\right] &= I(t_k)v(0,t_k)\mathbb{E}\left[p_{aa}(t_k,\gamma)v(t_k,\gamma)V(\gamma)\big|\mathcal{G}_{t_k}\right] \\
&= I(t_k)v(0,t_k)\mathbb{E}\left[p_{aa}(t_k,\gamma)v(t_k,\gamma)V(\gamma)\big|\mathcal{H}_{t_k}\right] \\
&= I(t_k)v(0,t_k)p^{\mathbb{E}}(t_k,\gamma)v^{\mathbb{E}}(t_k,\gamma)V^{\mathbb{E}}(t_k,\gamma).
\end{aligned}
$$

For $J = \{N\}$ and $t_k < \gamma$, it holds

$$
\begin{aligned}
\mathbb{E}\left[I(\gamma)v(0,\gamma)V(\gamma)\big|\mathcal{G}^J_{t_k,t_{k+1}}\right] &= I(t_{k+1})v(0,t_k)\mathbb{E}\left[p_{aa}(t_{k+1},\gamma)v(t_k,\gamma)V(\gamma)\big|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= I(t_{k+1})v(0,t_k)\mathbb{E}\left[p_{aa}(t_{k+1},\gamma)v(t_k,\gamma)V(\gamma)\big|\mathcal{H}^J_{t_k,t_{k+1}}\right] \\
&= I(t_{k+1})v(0,t_k)p^{\mathbb{E}}_{+}(t_k,\gamma)v^{\mathbb{E}}(t_k,\gamma)V^{\mathbb{E}}(t_k,\gamma).
\end{aligned}
$$

For $J = \{N, \lambda\}$ and $t_k < \gamma$, it holds

$$
\begin{aligned}
\mathbb{E}\left[I(\gamma)v(0,\gamma)V(\gamma)\big|\mathcal{G}^J_{t_k,t_{k+1}}\right] &= I(t_{k+1})v(0,t_k)\mathbb{E}\left[p_{aa}(t_{k+1},\gamma)v(t_k,\gamma)V(t_k,\gamma)\big|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= I(t_{k+1})v(0,t_k)\mathbb{E}\left[p_{aa}(t_{k+1},\gamma)v(t_k,\gamma)V(\gamma)\big|\mathcal{H}^J_{t_k,t_{k+1}}\right] \\
&= I(t_{k+1})v(0,t_k)p^{\mathbb{E}}(t_{k+1},\gamma)v^{\mathbb{E}}(t_k,\gamma)V^{\mathbb{E}}(t_k,\gamma).
\end{aligned}
$$

For $J = \{N, \lambda, r\}$ and $t_k < \gamma$, it holds

$$
\begin{aligned}
\mathbb{E}\left[I(\gamma)v(0,\gamma)V(\gamma)\big|\mathcal{G}^J_{t_k,t_{k+1}}\right] &= I(t_{k+1})v(0,t_k)\mathbb{E}\left[p_{aa}(t_{k+1},\gamma)v(t_k,\gamma)V(\gamma)\big|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= I(t_{k+1})v(0,t_k)\mathbb{E}\left[p_{aa}(t_{k+1},\gamma)v(t_{k+1},\gamma)V(\gamma)\big|\mathcal{H}^J_{t_k,t_{k+1}}\right] \\
&= I(t_{k+1})v(0,t_{k+1})p^{\mathbb{E}}(t_{k+1},\gamma)v^{\mathbb{E}}(t_{k+1},\gamma)V^{\mathbb{E}}(t_k,\gamma).
\end{aligned}
$$

Death cover (i=2):

For $J = \varnothing$ and $t_k < \gamma$, it holds

$$\begin{aligned}
&\mathbb{E}\left[\int_0^\gamma v(0,s)(1-f_d)V(s)\mathrm{d}N(s)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= \int_0^{t_k} v(0,s)(1-f_d)V(s)\mathrm{d}N(s) + v(0,t_k)\mathbb{E}\left[\int_{t_k}^\gamma v(t_k,s)(1-f_d)V(s)I(s)\overline{\lambda}(s)\mathrm{d}s\middle|\mathcal{G}_{t_k}\right].
\end{aligned}$$

The second term can be further simplied via

$$\begin{aligned}
&v(0,t_k)\mathbb{E}\left[\int_{t_k}^\gamma v(t_k,s)(1-f_d)V(s)p_{aa}(t_k,s)\overline{\lambda}(s)\mathrm{d}s\middle|\mathcal{G}_{t_k}\right]\\
&= v(0,t_k)\mathbb{E}\left[\int_{t_k}^\gamma v(t_k,s)(1-f_d)V(s)p_{aa}(t_k,s)\overline{\lambda}(s)\mathrm{d}s\middle|\mathcal{G}_{t_k}\right]\\
&= v(0,t_k)\int_{t_k}^\gamma v^{\mathbb{E}}(t_k,s)(1-f_d)V^{\mathbb{E}}(t_k,s)p^{\mathbb{E}}_\lambda(t_k,s)\mathrm{d}s.
\end{aligned}$$

Thus, we get

$$\begin{aligned}
&\mathbb{E}\left[\int_0^\gamma v(0,s)(1-f_d)V(s)\mathrm{d}N(s)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= \int_0^{t_k} v(0,s)(1-f_d)V(s)\mathrm{d}N(s) + v(0,t_k)\int_{t_k}^\gamma v^{\mathbb{E}}(t_k,s)(1-f_d)V^{\mathbb{E}}(t_k,s)p^{\mathbb{E}}_\lambda(t_k,s)\mathrm{d}s.
\end{aligned}$$

For $J = \{N\}$ and $t_k < \gamma$, it holds

$$\begin{aligned}
&\mathbb{E}\left[\int_0^\gamma v(0,s)(1-f_d)V(s)\mathrm{d}N(s)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= \int_0^{t_k} v(0,s)(1-f_d)V(s)\mathrm{d}N(s) + v(0,t_k)\int_{t_k}^{t_{k+1}} \mathbb{E}\left[v(t_k,s)(1-f_d)V(s)\middle|\mathcal{H}^J_{t_k,t_{k+1}}\right]\mathrm{d}N(s)\\
&\quad + v(0,t_k)\mathbb{E}\left[\int_{t_{k+1}}^\gamma v(t_k,s)(1-f_d)V(s)p_{aa}(t_{k+1},s)\overline{\lambda}(s)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= \int_0^{t_k} v(0,s)(1-f_d)V(s)\mathrm{d}N(s) + v(0,t_k)\int_{t_k}^{t_{k+1}} v^{\mathbb{E}}(t_k,s)(1-f_d)V^{\mathbb{E}}(t_k,s)\mathrm{d}N(s)\\
&\quad + v(0,t_k)I(t_k)\int_{t_{k+1}}^\gamma v^{\mathbb{E}}(t_k,s)(1-f_d)V^{\mathbb{E}}(s)p^{\mathbb{E}}_{\lambda,+}(t_k,s)\mathrm{d}s
\end{aligned}$$

For $J = \{N, \lambda\}$ and $t_k < \gamma$, one shows analogously

$$\begin{aligned}
&\mathbb{E}\left[\int_0^\gamma v(0,s)(1-f_d)V(s)\mathrm{d}N(s)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= \int_0^{t_k} v(0,s)(1-f_d)V(s)\mathrm{d}N(s) + v(0,t_k)\int_{t_k}^{t_{k+1}} v^{\mathbb{E}}(t_k,s)(1-f_d)V^{\mathbb{E}}(t_k,s)\mathrm{d}N(s)\\
&\quad + v(0,t_k)I(t_k)\int_{t_{k+1}}^\gamma v^{\mathbb{E}}(t_k,s)(1-f_d)V^{\mathbb{E}}(s)p^{\mathbb{E}}_\lambda(t_{k+1},s)\mathrm{d}s
\end{aligned}$$

For $J = \{N, \lambda, r\}$ and $t_k < \gamma$, it follows

$$\begin{aligned}
&\mathbb{E}\left[\int_0^\gamma v(0,s)(1-f_d)V(s)\mathrm{d}N(s)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= \int_0^{t_k} v(0,s)(1-f_d)V(s)\mathrm{d}N(s) + \int_{t_k}^{t_{k+1}} v(0,s)(1-f_d)V^{\mathbb{E}}(t_k,s)\mathrm{d}N(s)\\
&\quad + v(0,t_{k+1})I(t_k)\int_{t_{k+1}}^{\gamma} v^{\mathbb{E}}(t_{k+1},s)(1-f_d)V^{\mathbb{E}}(s)p^{\mathbb{E}}_{\lambda}(t_{k+1},s)\mathrm{d}s
\end{aligned}$$

Guaranteed pension period (i=3):

We need to distinguish between $t_k < \gamma$ and $t_k \geqslant \gamma$.

For $J = \varnothing$ and $t_k < \gamma$, it holds

$$\begin{aligned}
&\mathbb{E}\left[-I(\gamma)\int_\gamma^{\gamma+10} v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= v(0,t_k)I(t_k)\mathbb{E}\left[-p_{aa}(t_k,\gamma)\int_\gamma^{\gamma+10} v(t_k,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}_{t_k}\right]\\
&= -v(0,t_k)I(t_k)p^{\mathbb{E}}(t_k,\gamma)\int_\gamma^{\gamma+10} v^{\mathbb{E}}(t_k,s)f_pV^{\mathbb{E}}(t_k,\gamma)\mathrm{d}s.
\end{aligned}$$

For $J = \varnothing$ and $t_k \geqslant \gamma$, it holds

$$\begin{aligned}
&\mathbb{E}\left[-I(\gamma)\int_\gamma^{\gamma+10} v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= -v(0,t_k)I(\gamma)\int_\gamma^{t_k} v(0,s)f_pV(\gamma)\mathrm{d}s - I(\gamma)v(0,t_k)\int_{t_k}^{\gamma+10} v^{\mathbb{E}}(t_k,s)f_pV(\gamma)\mathrm{d}s.
\end{aligned}$$

For $J = \{N\}$ and $t_k < \gamma$, it holds

$$\begin{aligned}
&\mathbb{E}\left[-I(\gamma)\int_\gamma^{\gamma+10} v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= v(0,t_k)I(t_{k+1})\mathbb{E}\left[-p_{aa}(t_{k+1},\gamma)\int_\gamma^{\gamma+10} v(t_k,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= -v(0,t_k)I(t_{k+1})p^{\mathbb{E}}_{+}(t_k,\gamma)\int_\gamma^{\gamma+10} v^{\mathbb{E}}(t_k,s)f_pV^{\mathbb{E}}(t_k,\gamma)\mathrm{d}s.
\end{aligned}$$

For $J = \{N\}$ and $t_k \geqslant \gamma$, it holds

$$\begin{aligned}
&\mathbb{E}\left[-I(\gamma)\int_\gamma^{\gamma+10} v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= -I(\gamma)\int_\gamma^{t_k} v(0,s)f_pV(\gamma)\mathrm{d}s - I(\gamma)v(0,t_k)\int_{t_k}^{\gamma+10} v^{\mathbb{E}}(t_k,s)f_pV(\gamma)\mathrm{d}s
\end{aligned}$$

For $J = \{N, \lambda\}$ and $t_k < \gamma$, one shows analogously

$$\begin{aligned}
&\mathbb{E}\left[-I(\gamma)\int_\gamma^{\gamma+10} v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= -v(0,t_k)I(t_{k+1})p^{\mathbb{E}}(t_{k+1},\gamma)\int_\gamma^{\gamma+10} v^{\mathbb{E}}(t_k,s)f_pV^{\mathbb{E}}(t_k,\gamma)\mathrm{d}s.
\end{aligned}$$

For $J = \{N, \lambda\}$ and $t_k \geqslant \gamma$, one shows analogously

$$\mathbb{E}\left[-I(\gamma)\int_{\gamma}^{\gamma+10} v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= -I(\gamma)\int_{\gamma}^{t_k} v(0,s)f_pV(\gamma)\mathrm{d}s - I(\gamma)v(0,t_k)\int_{t_k}^{\gamma+10} v^{\mathbb{E}}(t_k,s)f_pV(\gamma)\mathrm{d}s$$

For $J = \{N, \lambda, r\}$ and $t_k < \gamma$, it follows

$$\mathbb{E}\left[-I(\gamma)\int_{\gamma}^{\gamma+10} v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= -v(0,t_{k+1})I(t_{k+1})p^{\mathbb{E}}(t_{k+1},\gamma)\int_{\gamma}^{\gamma+10} v^{\mathbb{E}}(t_{k+1},s)f_pV^{\mathbb{E}}(t_k,\gamma)\mathrm{d}s.$$

For $J = \{N, \lambda, r\}$ and $t_k \geqslant \gamma$, it follows

$$\mathbb{E}\left[-I(\gamma)\int_{\gamma}^{\gamma+10} v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= -I(\gamma)\int_{\gamma}^{t_{k+1}} v(0,s)f_pV(\gamma)\mathrm{d}s - I(\gamma)v(0,t_{k+1})\int_{t_{k+1}}^{\gamma+10} v^{\mathbb{E}}(t_{k+1},s)f_pV(\gamma)\mathrm{d}s.$$

Pension cover (i=4):

We need to distinguish between $t_k < \gamma + 10$ and $t_k \geqslant \gamma + 10$.

For $J = \varnothing$ and $t_k < \gamma + 10$, it holds

$$\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= I(t_k)v(0,t_k)\mathbb{E}\left[-\int_{\gamma+10}^{T} p_{aa}(t_k,s)v(t_k,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}_{t_k}\right]$$
$$= -I(t_k)v(0,t_k)\int_{\gamma+10}^{T} p^{\mathbb{E}}(t_k,s)v^{\mathbb{E}}(t_k,s)f_pV^{\mathbb{E}}(t_k,\gamma)\mathrm{d}s.$$

For $J = \varnothing$ and $t_k \geqslant \gamma + 10$, it holds

$$\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= -\int_{\gamma+10}^{t_k} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s - I(t_k)v(0,t_k)\mathbb{E}\left[-\int_{t_k}^{T} p_{aa}(t_k,s)v(t_k,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}_{t_k}\right]$$
$$= -\int_{\gamma+10}^{t_k} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s - I(t_k)v(0,t_k)\int_{t_k}^{T} p^{\mathbb{E}}(t_k,s)v^{\mathbb{E}}(t_k,s)f_pV(\gamma)\mathrm{d}s.$$

For $J = \{N\}$ and $t_k < \gamma + 10$, it holds

$$\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= I(t_{k+1})v(0,t_k)\mathbb{E}\left[-\int_{\gamma+10}^{T} p_{aa}(t_{k+1},s)v(t_k,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= -I(t_{k+1})v(0,t_k)\int_{\gamma+10}^{T} p^{\mathbb{E}}_{+}(t_k,s)v^{\mathbb{E}}(t_k,s)f_pV^{\mathbb{E}}(t_k,\gamma)\mathrm{d}s.$$

For $J = \{N\}$ and $t_k \geqslant \gamma + 10$, it holds

$$
\begin{aligned}
&\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= -\int_{\gamma+10}^{t_k} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s - v(0,t_k)\mathbb{E}\left[\int_{t_k}^{t_{k+1}} I(s)v(t_k,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&\quad - I(t_{k+1})v(0,t_k)\mathbb{E}\left[\int_{t_{k+1}}^{T} p_{aa}(t_{k+1},s)v(t_k,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= -\int_{\gamma+10}^{t_k} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s - v(0,t_k)\int_{t_k}^{t_{k+1}} I(s)v^{\mathbb{E}}(t_k,s)f_pV(\gamma)\mathrm{d}s \\
&\quad - I(t_{k+1})v(0,t_k)\int_{t_{k+1}}^{T} p_+^{\mathbb{E}}(t_k,s)v^{\mathbb{E}}(t_k,s)f_pV(\gamma)\mathrm{d}s.
\end{aligned}
$$

For $J = \{N, \lambda\}$ and $t_k < \gamma + 10$, it holds

$$
\begin{aligned}
&\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= -I(t_{k+1})v(0,t_k)\int_{\gamma+10}^{T} p^{\mathbb{E}}(t_{k+1},s)v^{\mathbb{E}}(t_k,s)f_pV^{\mathbb{E}}(t_k,\gamma)\mathrm{d}s.
\end{aligned}
$$

For $J = \{N, \lambda\}$ and $t_k \geqslant \gamma + 10$, it holds

$$
\begin{aligned}
&\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= -\int_{\gamma+10}^{t_k} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s - v(0,t_k)\int_{t_k}^{t_{k+1}} I(s)v^{\mathbb{E}}(t_k,s)f_pV(\gamma)\mathrm{d}s \\
&\quad - I(t_{k+1})v(0,t_k)\int_{t_{k+1}}^{T} p^{\mathbb{E}}(t_{k+1},s)v^{\mathbb{E}}(t_k,s)f_pV(\gamma)\mathrm{d}s.
\end{aligned}
$$

For $J = \{N, \lambda, r\}$ and $t_k < \gamma + 10$, it holds

$$
\begin{aligned}
&\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= -I(t_{k+1})v(0,t_{k+1})\int_{\gamma+10}^{T} p^{\mathbb{E}}(t_{k+1},s)v^{\mathbb{E}}(t_{k+1},s)f_pV^{\mathbb{E}}(t_k,\gamma)\mathrm{d}s.
\end{aligned}
$$

For $J = \{N, \lambda, r\}$ and $t_k \geqslant \gamma + 10$, it holds

$$
\begin{aligned}
&\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= -\int_{\gamma+10}^{t_{k+1}} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s - I(t_{k+1})v(0,t_{k+1})\int_{t_{k+1}}^{T} p^{\mathbb{E}}(t_{k+1},s)v^{\mathbb{E}}(t_{k+1},s)f_pV(\gamma)\mathrm{d}s.
\end{aligned}
$$

A.3.2 Analysis of conditional expectations for the order $\{\boldsymbol{Y}, \boldsymbol{r}, \boldsymbol{\lambda}, \boldsymbol{N}\}$

Savings account ($i = 1$):

For $J = \varnothing$ and $t_k < \gamma$, it holds (see Section A.3.1)

$$\mathbb{E}\left[I(\gamma)v(0,\gamma)V(\gamma)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] = I(t_k)v(0,t_k)p^{\mathbb{E}}(t_k,\gamma)v^{\mathbb{E}}(t_k,\gamma)V^{\mathbb{E}}(t_k,\gamma).$$

For $J = \{Y\}$ and $t_k < \gamma$, it holds

$$\begin{aligned}\mathbb{E}\left[I(\gamma)v(0,\gamma)V(\gamma)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] &= I(t_k)v(0,t_k)\mathbb{E}\left[p_{aa}(t_k,\gamma)v(t_k,\gamma)V(\gamma)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\ &= I(t_{k+1})v(0,t_k)\mathbb{E}\left[p_{aa}(t_k,\gamma)v(t_k,\gamma)V(\gamma)\middle|\mathcal{H}^J_{t_k,t_{k+1}}\right]\\ &= I(t_k)v(0,t_k)p^{\mathbb{E}}(t_k,\gamma)v^{\mathbb{E}}(t_k,\gamma)V^{\mathbb{E}}(t_{k+1},\gamma).\end{aligned}$$

For $J = \{Y, r\}$ and $t_k < \gamma$, it holds

$$\begin{aligned}\mathbb{E}\left[I(\gamma)v(0,\gamma)V(\gamma)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] &= I(t_k)v(0,t_{k+1})\mathbb{E}\left[p_{aa}(t_k,\gamma)v(t_{k+1},\gamma)V(t_k,\gamma)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\ &= I(t_k)v(0,t_{k+1})\mathbb{E}\left[p_{aa}(t_k,\gamma)v(t_{k+1},\gamma)V(\gamma)\middle|\mathcal{H}^J_{t_k,t_{k+1}}\right]\\ &= I(t_k)v(0,t_{k+1})p^{\mathbb{E}}(t_k,\gamma)v^{\mathbb{E}}(t_{k+1},\gamma)V^{\mathbb{E}}(t_{k+1},\gamma).\end{aligned}$$

For $J = \{Y, r, \lambda\}$ and $t_k < \gamma$, it holds

$$\begin{aligned}\mathbb{E}\left[I(\gamma)v(0,\gamma)V(\gamma)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] &= I(t_k)v(0,t_k)\mathbb{E}\left[p_{aa}(t_k,\gamma)v(t_k,\gamma)V(\gamma)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\ &= I(t_k)v(0,t_{k+1})\mathbb{E}\left[p_{aa}(t_k,\gamma)v(t_{k+1},\gamma)V(\gamma)\middle|\mathcal{H}^J_{t_k,t_{k+1}}\right]\\ &= I(t_k)v(0,t_{k+1})p_{aa}(t_k,t_{k+1})p^{\mathbb{E}}(t_{k+1},\gamma)v^{\mathbb{E}}(t_{k+1},\gamma)V^{\mathbb{E}}(t_{k+1},\gamma).\end{aligned}$$

Death cover (i=2):

For $J = \varnothing$ and $t_k < \gamma$, it holds (see Section A.3.1)

$$\begin{aligned}&\mathbb{E}\left[\int_0^\gamma v(0,s)(1-f_d)V(s)\mathrm{d}N(s)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\ &= \int_0^{t_k} v(0,s)(1-f_d)V(s)\mathrm{d}N(s) + v(0,t_k)I(t_k)\int_{t_k}^\gamma v^{\mathbb{E}}(t_k,s)(1-f_d)V^{\mathbb{E}}(t_k,s)p^{\mathbb{E}}_\lambda(t_k,s)\mathrm{d}s.\end{aligned}$$

For $J = \{Y\}$ and $t_k < \gamma$, it holds

$$\begin{aligned}&\mathbb{E}\left[\int_0^\gamma v(0,s)(1-f_d)V(s)\mathrm{d}N(s)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\ &= \int_0^{t_k} v(0,s)(1-f_d)V(s)\mathrm{d}N(s)\\ &\quad + v(0,t_k)I(t_k)\mathbb{E}\left[\int_{t_k}^{t_{k+1}} v(t_k,s)(1-f_d)V(s)p_{aa}(t_k,s)\overline{\lambda}(s)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\ &\quad + v(0,t_k)I(t_k)\mathbb{E}\left[\int_{t_{k+1}}^{\gamma} v(t_k,s)(1-f_d)V(s)p_{aa}(t_k,s)\overline{\lambda}(s)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\end{aligned}$$

$$
\begin{aligned}
&= \int_0^{t_k} v(0,s)(1-f_d)V(s)\mathrm{d}N(s) + v(0,t_k)I(t_k)\int_{t_k}^{t_{k+1}} v^{\mathbb{E}}(t_k,s)(1-f_d)V(s)p_\lambda^{\mathbb{E}}(t_k,s)\mathrm{d}s \\
&\quad + v(0,t_k)I(t_k)\int_{t_{k+1}}^{\gamma} v^{\mathbb{E}}(t_k,s)(1-f_d)V^{\mathbb{E}}(t_{k+1},s)p_\lambda^{\mathbb{E}}(t_k,s)\mathrm{d}s
\end{aligned}
$$

For $J=\{Y,r\}$ and $t_k<\gamma$, one shows analogously

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^{\gamma} v(0,s)(1-f_d)V(s)\mathrm{d}N(s)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= \int_0^{t_k} v(0,s)(1-f_d)V(s)\mathrm{d}N(s) + I(t_k)\int_{t_k}^{t_{k+1}} v(0,s)(1-f_d)V(s)p_\lambda^{\mathbb{E}}(t_k,s)\mathrm{d}s \\
&\quad + v(0,t_{k+1})I(t_k)\int_{t_{k+1}}^{\gamma} v^{\mathbb{E}}(t_{k+1},s)(1-f_d)V^{\mathbb{E}}(t_{k+1},s)p_\lambda^{\mathbb{E}}(t_k,s)\mathrm{d}s.
\end{aligned}
$$

For $J=\{Y,r,\lambda\}$ and $t_k<\gamma$, it follows

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^{\gamma} v(0,s)(1-f_d)V(s)\mathrm{d}N(s)\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= \int_0^{t_k} v(0,s)(1-f_d)V(s)\mathrm{d}N(s) + I(t_k)\int_{t_k}^{t_{k+1}} v(0,s)(1-f_d)V(s)p_{aa}(t_k,s)\overline{\lambda}(s)\mathrm{d}s \\
&\quad + v(0,t_{k+1})I(t_k)p_{aa}(t_k,t_{k+1})\int_{t_{k+1}}^{\gamma} v^{\mathbb{E}}(t_{k+1},s)(1-f_d)V^{\mathbb{E}}(t_{k+1},s)p_\lambda^{\mathbb{E}}(t_{k+1},s)\mathrm{d}s.
\end{aligned}
$$

Guaranteed pension period (i=3):

We need to distinguish between $t_k<\gamma$ and $t_k\geqslant\gamma$.

For $J=\varnothing$ and $t_k<\gamma$, it holds (see Section A.3.1)

$$
\begin{aligned}
&\mathbb{E}\left[-I(\gamma)\int_{\gamma}^{\gamma+10} v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= -v(0,t_k)I(t_k)p^{\mathbb{E}}(t_k,\gamma)\int_{\gamma}^{\gamma+10} v^{\mathbb{E}}(t_k,s)f_pV^{\mathbb{E}}(t_k,\gamma)\mathrm{d}s.
\end{aligned}
$$

For $J=\varnothing$ and $t_k\geqslant\gamma$, it holds (see Section A.3.1)

$$
\begin{aligned}
&\mathbb{E}\left[-I(\gamma)\int_{\gamma}^{\gamma+10} v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= -v(0,t_k)I(t_{k+1})p_+^{\mathbb{E}}(t_k,\gamma)\int_{\gamma}^{\gamma+10} v^{\mathbb{E}}(t_k,s)f_pV^{\mathbb{E}}(t_k,\gamma)\mathrm{d}s.
\end{aligned}
$$

For $J=\{Y\}$ and $t_k<\gamma$, it holds

$$
\begin{aligned}
&\mathbb{E}\left[-I(\gamma)\int_{\gamma}^{\gamma+10} v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= -v(0,t_k)I(t_k)p^{\mathbb{E}}(t_k,\gamma)\int_{\gamma}^{\gamma+10} v^{\mathbb{E}}(t_k,s)f_pV^{\mathbb{E}}(t_{k+1},\gamma)\mathrm{d}s.
\end{aligned}
$$

For $J=\{Y\}$ and $t_k\geqslant\gamma$, it holds

$$
\begin{aligned}
&\mathbb{E}\left[-I(\gamma)\int_{\gamma}^{\gamma+10} v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right] \\
&= -I(\gamma)\int_{\gamma}^{t_k} v(0,s)f_pV(\gamma)\mathrm{d}s - I(\gamma)v(0,t_k)\int_{t_k}^{\gamma+10} v^{\mathbb{E}}(t_k,s)f_pV(\gamma)\mathrm{d}s
\end{aligned}
$$

For $J = \{Y, r\}$ and $t_k < \gamma$, one shows analogously

$$\mathbb{E}\left[-I(\gamma)\int_{\gamma}^{\gamma+10} v(0,s) f_p V(\gamma)\mathrm{d}s \middle| \mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= -v(0,t_{k+1}) I(t_k) p^{\mathbb{E}}(t_k,\gamma)\int_{\gamma}^{\gamma+10} v^{\mathbb{E}}(t_{k+1},s) f_p V^{\mathbb{E}}(t_{k+1},\gamma)\mathrm{d}s.$$

For $J = \{Y, r\}$ and $t_k \geqslant \gamma$, one shows analogously

$$\mathbb{E}\left[-I(\gamma)\int_{\gamma}^{\gamma+10} v(0,s) f_p V(\gamma)\mathrm{d}s \middle| \mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= -I(\gamma)\int_{\gamma}^{t_{k+1}} v(0,s) f_p V(\gamma)\mathrm{d}s - I(\gamma) v(0,t_{k+1})\int_{t_{k+1}}^{\gamma+10} v^{\mathbb{E}}(t_{k+1},s) f_p V(\gamma)\mathrm{d}s$$

For $J = \{Y, r, \lambda\}$ and $t_k < \gamma$, it follows

$$\mathbb{E}\left[-I(\gamma)\int_{\gamma}^{\gamma+10} v(0,s) f_p V(\gamma)\mathrm{d}s \middle| \mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= -v(0,t_{k+1}) I(t_k) p_{aa}(t_k,t_{k+1}) p^{\mathbb{E}}(t_{k+1},\gamma)\int_{\gamma}^{\gamma+10} v^{\mathbb{E}}(t_{k+1},s) f_p V^{\mathbb{E}}(t_{k+1},\gamma)\mathrm{d}s.$$

For $J = \{Y, r, \lambda\}$ and $t_k \geqslant \gamma$, it follows

$$\mathbb{E}\left[-I(\gamma)\int_{\gamma}^{\gamma+10} v(0,s) f_p V(\gamma)\mathrm{d}s \middle| \mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= -I(\gamma)\int_{\gamma}^{t_{k+1}} v(0,s) f_p V(\gamma)\mathrm{d}s - I(\gamma) v(0,t_{k+1})\int_{t_{k+1}}^{\gamma+10} v^{\mathbb{E}}(t_{k+1},s) f_p V(\gamma)\mathrm{d}s.$$

Pension cover (i=4):

We need to distinguish between $t_k < \gamma + 10$ and $t_k \geqslant \gamma + 10$.

For $J = \varnothing$ and $t_k < \gamma + 10$, it holds (see Section A.3.1)

$$\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s) v(0,s) f_p V(\gamma)\mathrm{d}s \middle| \mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= -I(t_k) v(0,t_k)\int_{\gamma+10}^{T} p^{\mathbb{E}}(t_k,s) v^{\mathbb{E}}(t_k,s) f_p V^{\mathbb{E}}(t_k,\gamma)\mathrm{d}s.$$

For $J = \varnothing$ and $t_k \geqslant \gamma + 10$, it holds (see Section A.3.1)

$$\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s) v(0,s) f_p V(\gamma)\mathrm{d}s \middle| \mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= -\int_{\gamma+10}^{t_k} I(s) v(0,s) f_p V(\gamma)\mathrm{d}s - I(t_k) v(0,t_k)\int_{t_k}^{T} p^{\mathbb{E}}(t_k,s) v^{\mathbb{E}}(t_k,s) f_p V(\gamma)\mathrm{d}s.$$

For $J = \{Y\}$ and $t_k < \gamma + 10$, it holds

$$\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s) v(0,s) f_p V(\gamma)\mathrm{d}s \middle| \mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= I(t_k) v(0,t_k)\mathbb{E}\left[-\int_{\gamma+10}^{T} p_{aa}(t_k,s) v(t_k,s) f_p V(\gamma)\mathrm{d}s \middle| \mathcal{G}^J_{t_k,t_{k+1}}\right]$$
$$= -I(t_k) v(0,t_k)\int_{\gamma+10}^{T} p^{\mathbb{E}}(t_k,s) v^{\mathbb{E}}(t_k,s) f_p V^{\mathbb{E}}(t_{k+1},\gamma)\mathrm{d}s.$$

For $J = \{Y\}$ and $t_k \geqslant \gamma + 10$, it holds

$$
\begin{aligned}
&\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= -\int_{\gamma+10}^{t_k} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s - I(t_k)v(0,t_k)\mathbb{E}\left[\int_{t_k}^{T} p_{aa}(t_k,s)v(t_k,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= -\int_{\gamma+10}^{t_k} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s - I(t_k)v(0,t_k)\int_{t_k}^{T} p^{\mathbb{E}}(t_k,s)v^{\mathbb{E}}(t_k,s)f_pV(\gamma)\mathrm{d}s.
\end{aligned}
$$

For $J = \{Y, r\}$ and $t_k < \gamma + 10$, it holds

$$
\begin{aligned}
&\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= -I(t_k)v(0,t_{k+1})\int_{\gamma+10}^{T} p^{\mathbb{E}}(t_k,s)v^{\mathbb{E}}(t_{k+1},s)f_pV^{\mathbb{E}}(t_{k+1},\gamma)\mathrm{d}s.
\end{aligned}
$$

For $J = \{Y, r\}$ and $t_k \geqslant \gamma + 10$, it holds

$$
\begin{aligned}
&\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= -\int_{\gamma+10}^{t_k} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s - I(t_k)\int_{t_k}^{t_{k+1}} p^{\mathbb{E}}(t_k,s)v(0,s)f_pV(\gamma)\mathrm{d}s\\
&\quad - I(t_k)v(0,t_{k+1})\int_{t_{k+1}}^{T} p^{\mathbb{E}}(t_k,s)v^{\mathbb{E}}(t_{k+1},s)f_pV(\gamma)\mathrm{d}s.
\end{aligned}
$$

For $J = \{Y, r, \lambda\}$ and $t_k < \gamma + 10$, it holds

$$
\begin{aligned}
&\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= -I(t_k)v(0,t_{k+1})p_{aa}(t_k,t_{k+1})\int_{\gamma+10}^{T} p^{\mathbb{E}}(t_{k+1},s)v^{\mathbb{E}}(t_{k+1},s)f_pV^{\mathbb{E}}(t_{k+1},\gamma)\mathrm{d}s.
\end{aligned}
$$

For $J = \{Y, r, \lambda\}$ and $t_k \geqslant \gamma + 10$, it holds

$$
\begin{aligned}
&\mathbb{E}\left[-\int_{\gamma+10}^{T} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s\middle|\mathcal{G}^J_{t_k,t_{k+1}}\right]\\
&= -\int_{\gamma+10}^{t_k} I(s)v(0,s)f_pV(\gamma)\mathrm{d}s - I(t_k)\int_{t_k}^{t_{k+1}} p_{aa}(t_k,s)v(0,s)f_pV(\gamma)\mathrm{d}s\\
&\quad - v(0,t_{k+1})I(t_k)p_{aa}(t_k,t_{k+1})\int_{t_{k+1}}^{T} p^{\mathbb{E}}(t_{k+1},s)v^{\mathbb{E}}(t_{k+1},s)f_pV(\gamma)\mathrm{d}s.
\end{aligned}
$$

A.3.3 SU decomposition with respect to the order $\{\boldsymbol{Y}, \boldsymbol{r}, \boldsymbol{\lambda}, \boldsymbol{N}\}$

This leads us to the following SU contributions for the order $\{Y, r, \lambda, N\}$, categorised according to the different risks.

Unsystematic biometric risk

The SU contributions with respect to N are given by

$$\begin{aligned}
D_N^1(t) &= \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} \mathbb{1}_{\{t_k < \gamma\}} v(0, t_{k+1}) \widetilde{\Delta I^p}(t_k, t_{k+1}, \gamma) v^{\mathbb{E}}(t_{k+1}, \gamma) V^{\mathbb{E}}(t_{k+1}, \gamma), \\
D_N^2(t) &= \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} \int_{t_k}^{t_{k+1}} v(0, s)(1 - f_d) V(s)(\mathrm{d}N(s) - I(t_k) p_\lambda^{\mathbb{E}}(t_k, s)\mathrm{d}s) \\
&\quad + \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} v(0, t_{k+1}) \int_{t_{k+1}}^{\gamma} v^{\mathbb{E}}(t_{k+1}, s)(1 - f_d) V^{\mathbb{E}}(t_{k+1}, s) \widetilde{\Delta I_\lambda^p}(t_k, t_{k+1}, s)\mathrm{d}s, \\
D_N^3(t) &= -\sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} v(0, t_{k+1}) \widetilde{\Delta I^p}(t_k, t_{k+1}, \gamma) \int_{\gamma}^{\gamma+10} v^{\mathbb{E}}(t_{k+1}, s) f_p V^{\mathbb{E}}(t_{k+1}, \gamma)\mathrm{d}s, \\
D_N^4(t) &= -\sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma+10}} v(0, t_{k+1}) \int_{\gamma+10}^{T} \widetilde{\Delta I^p}(t_k, t_{k+1}, s) v^{\mathbb{E}}(t_{k+1}, s) f_p V^{\mathbb{E}}(t_{k+1}, \gamma)\mathrm{d}s \\
&\quad - \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k \geqslant \gamma+10}} v(0, t_{k+1}) \int_{t_k}^{t_{k+1}} (I(s) - I(t_k) p(t_k, s)) v^{\mathbb{E}}(t_{k+1}, s) f_p V(\gamma)\mathrm{d}s \\
&\quad - \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k \geqslant \gamma+10}} v(0, t_{k+1}) \int_{t_{k+1}}^{T} \widetilde{\Delta I^p}(t_k, t_{k+1}, s) v^{\mathbb{E}}(t_{k+1}, s) f_p V(\gamma)\mathrm{d}s,
\end{aligned}$$

where

$$\begin{aligned}
\widetilde{\Delta I^p}(t_k, t_{k+1}, s) &= I(t_{k+1}) p^{\mathbb{E}}(t_{k+1}, s) - I(t_k) p_{aa}(t_k, t_{k+1}) p^{\mathbb{E}}(t_{k+1}, s), \\
\widetilde{\Delta I_\lambda^p}(t_k, t_{k+1}, s) &= I(t_{k+1}) p_\lambda^{\mathbb{E}}(t_{k+1}, s) - I(t_k) p_{aa}(t_k, t_{k+1}) p_\lambda^{\mathbb{E}}(t_{k+1}, s).
\end{aligned}$$

Systematic biometric risk

The SU contributions with respect to λ are given by

$$\begin{aligned}
D_\lambda^1(t) &= \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_k) v(0, t_{k+1}) \widetilde{\Delta p}(t_k, t_{k+1}, \gamma) v^{\mathbb{E}}(t_{k+1}, \gamma) V^{\mathbb{E}}(t_{k+1}, \gamma), \\
D_\lambda^2(t) &= \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_k) v(0, t_{k+1}) \int_{t_{k+1}}^{\gamma} v^{\mathbb{E}}(t_{k+1}, s)(1 - f_d) V^{\mathbb{E}}(t_{k+1}, s) \widetilde{\Delta p_\lambda}(t_k, t_{k+1}, s)\mathrm{d}s, \\
D_\lambda^3(t) &= -\sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_k) \widetilde{\Delta p}(t_k, t_{k+1}, \gamma) v(0, t_{k+1}) \int_{\gamma}^{\gamma+10} v^{\mathbb{E}}(t_{k+1}, s) f_p V^{\mathbb{E}}(t_{k+1}, \gamma)\mathrm{d}s,
\end{aligned}$$

$$D_\lambda^4(t) = -\sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k<\gamma+10}} I(t_k)v(0,t_{k+1})\int_{\gamma+10}^{T}\widetilde{\Delta p}(t_k,t_{k+1},s)v^{\mathbb{E}}(t_{k+1},s))f_pV^{\mathbb{E}}(t_{k+1},\gamma)\mathrm{d}s$$
$$-\sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k\geqslant\gamma+10}} I(t_k)v(0,t_{k+1})\int_{t_{k+1}}^{T}\widetilde{\Delta p}(t_k,t_{k+1},s)v^{\mathbb{E}}(t_{k+1},s)f_pV(\gamma)\mathrm{d}s,$$

where

$$\widetilde{\Delta p}(t_k,t_{k+1},s) = p_{aa}(t_k,t_{k+1})p^{\mathbb{E}}(t_{k+1},s) - p^{\mathbb{E}}(t_k,s),$$
$$\widetilde{\Delta p_\lambda}(t_k,t_{k+1},s) = p_{aa}(t_k,t_{k+1})p_\lambda^{\mathbb{E}}(t_k,s) - p_\lambda^{\mathbb{E}}(t_k,s).$$

Systematic interest risk

The SU contributions with respect to r are given by

$$D_r^1(t) = \sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k<\gamma}} I(t_k)p^{\mathbb{E}}(t_k,\gamma)\Delta v(t_k,t_{k+1},\gamma)V^{\mathbb{E}}(t_{k+1},\gamma),$$
$$D_r^2(t) = \sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k<\gamma}} I(t_k)v(0,t_k)\int_{t_k}^{t_{k+1}}(v(t_k,s)-v^{\mathbb{E}}(t_k,s))(1-f_d)V^{\mathbb{E}}(t_{k+1},s)p_\lambda^{\mathbb{E}}(t_k,s)\mathrm{d}s$$
$$+\sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k<\gamma}} I(t_k)\int_{t_{k+1}}^{\gamma}\Delta v(t_k,t_{k+1},s)(1-f_d)V^{\mathbb{E}}(t_{k+1},s)p_\lambda^{\mathbb{E}}(t_k,s)\mathrm{d}s,$$
$$D_r^3(t) = -\sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k<\gamma}} I(t_k)p^{\mathbb{E}}(t_k,\gamma)\int_{\gamma}^{\gamma+10}\Delta v(t_k,t_{k+1},s)f_pV^{\mathbb{E}}(t_{k+1},\gamma)\mathrm{d}s$$
$$-\sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k\geqslant\gamma}} I(\gamma)f_pV(\gamma)\int_{t_k}^{t_{k+1}}(v(0,s)-v(0,t_k)v^{\mathbb{E}}(t_k,s))\mathrm{d}s$$
$$-\sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k\geqslant\gamma}} I(\gamma)f_pV(\gamma)\int_{t_{k+1}}^{\gamma+10}\Delta v(t_k,t_{k+1},s)\mathrm{d}s,$$
$$D_r^4(t) = -\sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k<\gamma+10}} I(t_k)\int_{\gamma+10}^{T}p^{\mathbb{E}}(t_k,s)\Delta v(t_k,t_{k+1},s)f_pV^{\mathbb{E}}(t_{k+1},\gamma)\mathrm{d}s$$
$$-\sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k\geqslant\gamma+10}} \int_{t_k}^{t_{k+1}} I(s)(v(0,s)-v(0,t_k)v^{\mathbb{E}}(t_k,s))f_pV(\gamma)\mathrm{d}s$$
$$-\sum_{\substack{t_k,t_{k+1}\in\mathcal{T}\\ t_k\geqslant\gamma+10}} I(t_k)\int_{t_{k+1}}^{T}p^{\mathbb{E}}(t_k,s)\Delta v(t_k,t_{k+1},s)f_pV(\gamma)\mathrm{d}s,$$

where

$$\Delta v(t_k,t_{k+1},s) = v(0,t_{k+1})v^{\mathbb{E}}(t_{k+1},s) - v(0,t_k)v^{\mathbb{E}}(t_k,s).$$

Systematic fund risk

The SU contributions with respect to Y are given by

$$
\begin{aligned}
D_Y^1(t) &= \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_k) p^{\mathbb{E}}(t_k, \gamma) v^{\mathbb{E}}(t_k, \gamma) \Delta V(t_k, t_{k+1}, \gamma), \\
D_Y^2(t) &= \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_k) v(0, t_k) \int_{t_k}^{t_{k+1}} v^{\mathbb{E}}(t_k, s)(1 - f_d)(V(s) - V^{\mathbb{E}}(t_k, s)) p_\lambda^{\mathbb{E}}(t_k, s) \mathrm{d}s \\
&\quad + \sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_k) v(0, t_k) \int_{t_{k+1}}^{\gamma} v^{\mathbb{E}}(t_k, s)(1 - f_d) \Delta V(t_k, t_{k+1}, s) p_\lambda^{\mathbb{E}}(t_k, s) \mathrm{d}s, \\
D_Y^3(t) &= -\sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma}} I(t_k) v(0, t_k) p^{\mathbb{E}}(t_k, \gamma) \int_{\gamma}^{\gamma+10} v^{\mathbb{E}}(t_k, s) f_p \Delta V(t_k, t_{k+1}, \gamma) \mathrm{d}s, \\
D_Y^4(t) &= -\sum_{\substack{t_k, t_{k+1} \in \mathcal{T} \\ t_k < \gamma+10}} I(t_k) v(0, t_k) \int_{\gamma+10}^{T} p^{\mathbb{E}}(t_k, s) v^{\mathbb{E}}(t_k, s) f_p \Delta V(t_k, t_{k+1}, \gamma) \mathrm{d}s,
\end{aligned}
$$

where

$$\Delta V(t_k, t_{k+1}, s) = V^{\mathbb{E}}(t_{k+1}, s) - V^{\mathbb{E}}(t_k, s).$$

A.4 Numerical example - R code

Within the R code, A represents Q and μ depicts $\overline{\lambda}$.

```
set.seed(50)

#INITIALIZATION PARAMETERS
T<-55
gamma<-22;
f_p=0.032;
f_d=0.9;
lambda_0<-0.00329542;
mu_lambda<-0.07731571;
sigma_lambda<-0.00012212;
r_0<-0.025;
beta<-0.000199;
mu_r<-0.01;
sigma_r<-0.0035;
Y_0<-100;
mu_Y<-0.06;
sigma_Y<-0.2;
A_0<-100;
a<-3600;

#INITIALIZATION FUNCTIONS
g_lambda <- function(x){
 if(x<0){
        u<-0
        return(u)
 } else if(0<= x & x<=1000){
        u<- x
        return(u)
 } else if(1000<x){
        u<- 1000
        return(u)
 }
}

g_r <- function(x){
 if(x< -1000){
        u<- -1000
        return(u)
 } else if(-1000<=x & x<=1000){
```

```r
        u<-x
        return(u)
 } else if(1000<x){
        u<- 1000
        return(u)
 }
}

g_Y <- function(x){
 if(x<0.0001){
        u<-0.0001
        return(u)
 } else if(0.0001<=x){
        u<-x
        return(u)
 }
}

#AUXILIARY FUNCTION
#v is a vector, x is a real number
#returns the position of an entry in an increasing vector, that is
   the closest smaller or equal to x
#helps us to relate the integration grid to the grid of mu
nextsmaller <- function(x,v){
 o=max(which(v-x<=0));
 return(o);
}

#CALCULATION MAIN PATH lambda, r, Y
#GRID MU
h_mu=1;
M_mu=T/h_mu;
I_mu=0:M_mu;
partition_mu=I_mu*h_mu;

#GRID INTEGRAL APPROXIMATION
h_INT=1/52;
M<-T/h_INT;
INT=0:M;
partitionINT=INT*T/M;
```

```
gamma_INT=which(partitionINT==gamma); #GRID index_MLMC FOR
   RETIREMENT
gamma_INT10=which(partitionINT==gamma+10);

#MAIN PATH - MORTALITY INTENSITY DRIVER lambda AND MORTALITY
   INTENSITY mu
increments_lambda=rnorm(M,mean=0,sd=sqrt(T/M));
lambda<-rep(0,M+1);
lambda[1]=lambda_0;
for (i in 2:(M+1)){
 lambda[i]=lambda[i-1]+mu_lambda*lambda[i-1]*T/M+sigma_lambda*
    increments_lambda[i-1];
}

mu<-rep(0,M+1);
mu[1]=lambda[1];
x<-rep(0,M+1)
for (i in 2:M){
 k_mu=min(which(partitionINT[i]-partition_mu<0))-1;
 l_mu=min(which(partitionINT==partition_mu[k_mu]))
 mu[i]=g_lambda(lambda[l_mu]);
}
mu[M+1]=lambda[M+1];
mu_help=mu[INT %% (h_mu/h_INT)==0]; #MU ON GRID MU

#MAIN PATH - INTEREST INTENSITY r AND DISCOUNT FACTOR v(0,-)
increments_r=rnorm(M,mean=0,sd=sqrt(T/M));
r<-rep(0,M+1);
r[1]=r_0;
for (i in 2:(M+1)){
 r[i]=r[i-1]+(beta-mu_r*r[i-1])*T/M+sigma_r*increments_r[i-1];
}
v<-rep(0,M+1);
v[1]=1;
g_r_val<- unlist(lapply(r,g_r), use.names=FALSE)
v[2:(M+1)]<-exp(-T/M*cumsum(g_r_val[1:M]));

#MAIN PATH - MARKET INDEX Y, SHARES HELD A AND DEPOSIT VALUE V
increments_Y=rnorm(M,mean=0,sd=sqrt(T/M));
Y<- Y_0*exp((mu_Y-sigma_Y^2/2)*partitionINT+sigma_Y*c(0,cumsum(
   increments_Y)));
```

```
g_Y_val<- unlist(lapply(Y,g_Y), use.names=FALSE)
A<- c(A_0, A_0+cumsum(a/g_Y_val[1:M]*T/M));
V<-A*g_Y_val;

#MAIN PATH - SURVIVAL PROBABILITIES
p<-matrix(0,M+1,M+1);
for(k in 1:M){
 p[k,k]=1;
 p[k,(k+1):(M+1)]=exp(-cumsum(mu[k:M]*T/M));
}
p[M+1,M+1]=1;

#MAIN PATH - INDICATOR FUNCTION
Ind<-rep(0,M+1);
Ind[1]=1;
rn<-runif(M);
run<-1;
while(run<(M+1)& 0 <= rn[run] & rn[run]<= exp(-mu[run]*T/M) & Ind[
   run]==1){
 run=run+1;
 Ind[run]=1;
}

#INTRODUCTION MLMC ESTIMATORS
#MLMC estimator p and p_mu
mlmc_p<-function(n,j,start_val){
 I=0:n;
 partitionSU=I*T/n;
 start_INT=which(partitionINT==partitionSU[j]);
 exp_p=0
 exp_pm=0
 for (l in 1:(L+1)){
   exp_p_lvl=0
   exp_pm_lvl=0
   hl0=2^(-l+2)*h_MLMC; #step width MLMC previous level
   hl1=2^(-l+1)*h_MLMC; #step width MLMC this level
   int0=round((T-partitionSU[j])/hl0,1);
   I_MLMC0=1:(int0+1);
   int1=round((T-partitionSU[j])/hl1,1);
   I_MLMC1=1:(int1+1);
```

```
partitionMLMC0=partitionSU[j]+hl0*(I_MLMC0-1); #MLMC grid
    previous level
partitionMLMC1=partitionSU[j]+hl1*(I_MLMC1-1); #MLMC grid next
    level
I_MLMC_mu0=I_MLMC0[round(partitionMLMC0,6) %% h_mu == 0]; #
    relate MLMC grid to mu grid
I_MLMC_mu1=I_MLMC1[round(partitionMLMC1,6) %% h_mu == 0];#relate
     MLMC grid to mu grid
if(partitionSU[j]%%h_mu==0){ #if starting point is on mu grid,
    we omit the first index_MLMC
    index_MLMC=I_MLMC_mu1[2]
    I_MLMC_mu0=I_MLMC_mu0[2:length(I_MLMC_mu0)]
    I_MLMC_mu1=I_MLMC_mu1[2:length(I_MLMC_mu1)]
}else{
    index_MLMC=I_MLMC_mu1[1]
    I_MLMC_mu0=I_MLMC_mu0[1:length(I_MLMC_mu0)]
    I_MLMC_mu1=I_MLMC_mu1[1:length(I_MLMC_mu1)]
}
mu_grid_INT=unlist(lapply(partitionINT,nextsmaller,v=partition_
    mu)); #helps to duplicate values for mu on integration grid
index_INT=which(partitionINT==partitionMLMC1[index_MLMC])
numpck=N[l] #number of packages in level l
for (i in (1:numpck)){
  lambda_est0<-matrix(0,pck,int0+1)
  lambda_est0[,1]<-rep(start_val,pck)
  lambda_est1<-matrix(0,pck,int1+1)
  lambda_est1[,1]<-rep(start_val,pck)
  mu_shift0=matrix(0,pck,M_mu+1);
  mu_shift1=matrix(0,pck,M_mu+1);
  if (l==1) {
        MLMC_incr_lambda=matrix(rnorm(int1*pck,mean=0,sd=sqrt(hl1
            )),nrow=pck, byrow=TRUE)
        for (k in 2:(int1+1)){
        lambda_est1[,k]=lambda_est1[,k-1]+mu_lambda*lambda_est1[,
            k-1]*hl1+sigma_lambda*MLMC_incr_lambda[,k-1]
        }
        lambda_est_help1=lambda_est1[,I_MLMC_mu1];#estimated
            lambda on grid mu
        mu_est1<- matrix(sapply(lambda_est_help1,g_lambda), nrow=
            pck)
        mu_shift1[,(M_mu-length(I_MLMC_mu1)+2):(M_mu+1)]=mu_est1;
```

```
    if(index_INT<(M+1)){
     p_aa_est1<-cbind(matrix(rep(p[start_INT,start_INT:index_
        INT],pck),nrow=pck,byrow=TRUE),p[start_INT,index_INT
        ]*exp(-t(apply(h_INT*mu_shift1[,mu_grid_INT[index_INT
        :M]],1,cumsum))));
     p_aa_est1p<-cbind(matrix(rep(p[start_INT,start_INT:index
        _INT]*mu[start_INT:index_INT],pck),nrow=pck, byrow=
        TRUE),p[start_INT,index_INT]*exp(-t(apply(h_INT*mu_
        shift1[,mu_grid_INT[index_INT:M]],1,cumsum)))*mu_
        shift1[,mu_grid_INT[(index_INT+1):(M+1)]]);
    }else{
     p_aa_est1<-matrix(rep(p[start_INT,start_INT:index_INT],
        pck),nrow=pck,byrow=TRUE);
     p_aa_est1p<-matrix(rep(p[start_INT,start_INT:index_INT]*
        mu[start_INT:index_INT],pck),nrow=pck,byrow=TRUE);
    }
    exp_p_lvl=exp_p_lvl+(1/(numpck*pck))*colSums(p_aa_est1);
    exp_pm_lvl=exp_pm_lvl+(1/(numpck*pck))*colSums(p_aa_est1p
       );
} else {
    MLMC_incr_lambda=matrix(rnorm(int1*pck,mean=0,sd=sqrt(hl1
       )),nrow=pck, byrow=TRUE)
    for (k in 2:(int1+1)){
     lambda_est1[,k]=lambda_est1[,k-1]+mu_lambda*lambda_est1
        [,k-1]*hl1+sigma_lambda*MLMC_incr_lambda[,k-1]
    }
    lambda_est_help1=lambda_est1[,I_MLMC_mu1];
    mu_est1<- matrix(sapply(lambda_est_help1,g_lambda), nrow=
       pck)
    mu_shift1[,(M_mu-length(I_MLMC_mu1)+2):(M_mu+1)]=mu_est1;
    if(index_INT<(M+1)){
     p_aa_est1<-cbind(matrix(rep(p[start_INT,start_INT:index_
        INT],pck),nrow=pck,byrow=TRUE),p[start_INT,index_INT
        ]*exp(-t(apply(h_INT*mu_shift1[,mu_grid_INT[index_INT
        :M]],1,cumsum))));
     p_aa_est1p<-cbind(matrix(rep(p[start_INT,start_INT:index
        _INT]*mu[start_INT:index_INT],pck),nrow=pck, byrow=
        TRUE),p[start_INT,index_INT]*exp(-t(apply(h_INT*mu_
        shift1[,mu_grid_INT[index_INT:M]],1,cumsum)))*mu_
        shift1[,mu_grid_INT[(index_INT+1):(M+1)]]);
    }else{
```

```
          p_aa_est1<-matrix(rep(p[start_INT,start_INT:index_INT],pck),nrow=pck,byrow=TRUE);
          p_aa_est1p<-matrix(rep(p[start_INT,start_INT:index_INT]*mu[start_INT:index_INT],pck),nrow=pck,byrow=TRUE);
         }
         for (k in 2:(int0+1)){
          lambda_est0[,k]=lambda_est0[,k-1]+mu_lambda*lambda_est0[,k-1]*hl0+sigma_lambda*(rowSums(MLMC_incr_lambda[,(2*(k-2)+1):(2*(k-1))]));
         }
         lambda_est_help0=lambda_est0[,I_MLMC_mu0];
         mu_est0<- matrix(sapply(lambda_est_help0,g_lambda), nrow=pck)
         mu_shift0[,(M_mu-length(I_MLMC_mu0)+2):(M_mu+1)]=mu_est0;
         if(index_INT<(M+1)){
          p_aa_est0<-cbind(matrix(rep(p[start_INT,start_INT:index_INT],pck),nrow=pck,byrow=TRUE),p[start_INT,index_INT]*exp(-t(apply(h_INT*mu_shift0[,mu_grid_INT[index_INT:M]],1,cumsum))));
          p_aa_est0p<-cbind(matrix(rep(p[start_INT,start_INT:index_INT]*mu[start_INT:index_INT],pck),nrow=pck,byrow=TRUE),p[start_INT,index_INT]*exp(-t(apply(h_INT*mu_shift0[,mu_grid_INT[index_INT:M]],1,cumsum)))*mu_shift0[,mu_grid_INT[(index_INT+1):(M+1)]]);
         }else{
          p_aa_est0<-matrix(rep(p[start_INT,start_INT:index_INT],pck),nrow=pck,byrow=TRUE);
          p_aa_est0p<-matrix(rep(p[start_INT,start_INT:index_INT]*mu[start_INT:index_INT],pck),nrow=pck,byrow=TRUE);
         }
         exp_p_lvl=exp_p_lvl+1/(numpck*pck)*colSums(p_aa_est1-p_aa_est0)
         exp_pm_lvl=exp_pm_lvl+1/(numpck*pck)*colSums(p_aa_est1p-p_aa_est0p)
         }
       }
       exp_p=exp_p+exp_p_lvl
       exp_pm=exp_pm+exp_pm_lvl
}
List <- list("exp_p" = exp_p, "exp_pm"=exp_pm);
return(List);
```

```
}

#MLMC estimator v
mlmc_v<-function(n,j,start_val){
 exp_v=0
 I=0:n;
 partitionSU=I*T/n;
 start_INT=which(partitionINT==partitionSU[j]);
 for (l in 1:(L+1)){
  exp_v_lvl=0
  hl0=2^(-l+2)*h_MLMC; #step width MLMC previous level
  hl1=2^(-l+1)*h_MLMC; #step width MLMC this level
  int0=round((T-partitionSU[j])/hl0,1);
  I_help0=1:(int0+1);
  int1=round((T-partitionSU[j])/hl1,1);
  I_help1=1:(int1+1);
  partitionMLMC0=partitionSU[j]+hl0*(I_help0-1);
  partitionMLMC1=partitionSU[j]+hl1*(I_help1-1);
  z=N[l];
  I_MLMC_INT0= (I_help0-1) %% (h_INT/hl0)==0
  I_MLMC_INT1= (I_help1-1) %% (h_INT/hl1)==0
  for(i in (1:z)){
   r_est0<-matrix(0,pck,int0+1)
   r_est0[,1]<-rep(start_val,pck)
   r_est1<-matrix(0,pck,int1+1)
   r_est1[,1]<-rep(start_val,pck)
   if (l==1) {
    MC_incr_r=matrix(rnorm(int1*pck,mean=0,sd=sqrt(hl1)),nrow=pck,
        byrow=TRUE)
        for (k in 2:(int1+1)){
         r_est1[,k]=r_est1[,k-1]+(beta-mu_r*r_est1[,k-1])*hl1+sigma
             _r*MC_incr_r[,k-1]
        }
        r_est1_help<-r_est1[,I_MLMC_INT1];
        bool1<- (r_est1_help>-995 & r_est1_help<995)
        if(sum(bool1)<length(r_est1_help)){
         g_r_est1<- matrix(sapply(r_est1_help,g_r), nrow=pck)
        } else {
         g_r_est1<-r_est1_help
        }
```

```
        v_est1<-exp(-t(apply(h_INT*g_r_est1[,1:(M-start_INT+1)],1,
           cumsum)));
        exp_v_lvl=exp_v_lvl+(1/(z*pck))*colSums(v_est1);
        } else {
         MC_incr_r=matrix(rnorm(int1*pck,mean=0,sd=sqrt(hl1)),nrow=
            pck, byrow=TRUE)
         for (k in 2:(int1+1)){
         r_est1[,k]=r_est1[,k-1]+(beta-mu_r*r_est1[,k-1])*hl1+sigma
            _r*MC_incr_r[,k-1]
         }
         r_est1_help<-r_est1[,I_MLMC_INT1];
         bool1<- (r_est1_help>-995 & r_est1_help<995)
         if(sum(bool1)<length(r_est1_help)){
          g_r_est1<- matrix(sapply(r_est1_help,g_r), nrow=pck)
         } else {
          g_r_est1<-r_est1_help
         }
         v_est1<-exp(-t(apply(h_INT*g_r_est1[,1:(M-start_INT+1)],1,
            cumsum)));
         for (k in 2:(int0+1)){
          r_est0[,k]=r_est0[,k-1]+(beta-mu_r*r_est0[,k-1])*hl0+
             sigma_r*(rowSums(MC_incr_r[,(2*(k-2)+1):(2*(k-1))]));
         }
         r_est0_help<-r_est0[,I_MLMC_INT0];
         bool0<- (r_est0_help>-995 & r_est0_help<995)
         if(sum(bool0)<length(r_est0_help)){
          g_r_est0<- matrix(sapply(r_est0_help,g_r), nrow=pck)
         } else {
          g_r_est0<-r_est0_help
         }
         v_est0<-exp(-t(apply(h_INT*g_r_est0[,1:(M-start_INT+1)],1,
            cumsum)));
         exp_v_lvl=exp_v_lvl+1/(z*pck)*colSums(v_est1-v_est0)
        }
  }
  exp_v=exp_v+exp_v_lvl
 }
 List <- list("exp_v" = exp_v);
 return(List);
}
```

```
#MLMC estimator Y
mlmc_Y<-function(n,j,start_val){
 exp_Y=0
 exp_AY=0
 I=0:n;
 partitionSU=I*T/n;
 start_INT=which(partitionINT==partitionSU[j]);
 for (l in 1:(L+1)){
  exp_Y_lvl=0
  exp_AY_lvl=0
  hl0=2^(-l+2)*h_MLMC;
  hl1=2^(-l+1)*h_MLMC;
  int0=round((gamma-partitionSU[j])/hl0,1);
  I_help0=1:(int0+1);
  int1=round((gamma-partitionSU[j])/hl1,1);
  I_help1=1:(int1+1);
  partitionMLMC0=partitionSU[j]+hl0*(I_help0-1);
  partitionMLMC1=partitionSU[j]+hl1*(I_help1-1);
  z=N[l];
  for (i in (1:z)){
   Y_est0<-matrix(0,pck,int0+1)
   Y_est0[,1]<-rep(start_val,pck)
   Y_est1<-matrix(0,pck,int1+1)
   Y_est1[,1]<-rep(start_val,pck)
   if (l==1) {
    MC_incr_Y=matrix(rnorm(int1*pck,mean=0,sd=sqrt(hl1)),nrow=pck,
        byrow=TRUE);
        for (k in 2:(int1+1)){
         Y_est1[,k]=Y_est1[,k-1]+mu_Y*Y_est1[,k-1]*hl1+sigma_Y*Y_
             est1[,k-1]*MC_incr_Y[,k-1]
        }
        Y_est1_INT=Y_est1[,(I_help1-1) %% (h_INT/hl1)==0];
        bool1<- (Y_est1_INT>0.01)
        if(sum(bool1)<length(Y_est1_INT)){
         g_Y_est1_INT<- matrix(sapply(Y_est1_INT,g_Y), nrow=pck);
        } else {
         g_Y_est1_INT<- Y_est1_INT
        }
        A_est1_INT=a*t(apply(1/g_Y_est1_INT[,1:(gamma_INT-start_INT
           )]*T/M,1,cumsum));
```

```
        exp_Y_lvl=exp_Y_lvl+(1/(z*pck))*colSums(Y_est1_INT[,2:(
           gamma_INT-start_INT+1)]);
        exp_AY_lvl=exp_AY_lvl+(1/(z*pck))*colSums(A_est1_INT*Y_est1
           _INT[,2:(gamma_INT-start_INT+1)]);
} else {
        MC_incr_Y=matrix(rnorm(int1*pck,mean=0,sd=sqrt(hl1)),nrow=
           pck, byrow=TRUE);
        for (k in 2:(int1+1)){
         Y_est1[,k]=Y_est1[,k-1]+mu_Y*Y_est1[,k-1]*hl1+sigma_Y*Y_
            est1[,k-1]*MC_incr_Y[,k-1]
        }
        Y_est1_INT=Y_est1[,(I_help1-1) %% (h_INT/hl1)==0];
        bool1<- (Y_est1_INT>0.01)
        if(sum(bool1)<length(Y_est1_INT)){
         g_Y_est1_INT<- matrix(sapply(Y_est1_INT,g_Y), nrow=pck);
        } else {
         g_Y_est1_INT<- Y_est1_INT
        }
        A_est1_INT=a*t(apply(1/g_Y_est1_INT[,1:(gamma_INT-start_INT
           )]*T/M,1,cumsum))
        for (k in 2:(int0+1)){
         Y_est0[,k]=Y_est0[,k-1]+mu_Y*Y_est0[,k-1]*hl0+sigma_Y*Y_
            est0[,k-1]*(rowSums(MC_incr_Y[,(2*(k-2)+1):(2*(k-1))]))
            ;
        }
        Y_est0_INT=Y_est0[,(I_help0-1) %% (h_INT/hl0)==0];
        bool0<- (Y_est0_INT>0.01)
        if(sum(bool0)<length(Y_est0_INT)){
         g_Y_est0_INT<- matrix(sapply(Y_est0_INT,g_Y), nrow=pck);
        } else {
         g_Y_est0_INT<- Y_est0_INT
        }
        A_est0_INT=a*t(apply(1/g_Y_est0_INT[,1:(gamma_INT-start_INT
           )]*T/M,1,cumsum))
        exp_Y_lvl=exp_Y_lvl+1/(z*pck)*colSums(g_Y_est1_INT[,2:(
           gamma_INT-start_INT+1)]-g_Y_est0_INT[,2:(gamma_INT-start
           _INT+1)]);
        exp_AY_lvl=exp_AY_lvl+1/(z*pck)*colSums(A_est1_INT*g_Y_est1
           _INT[,2:(gamma_INT-start_INT+1)]-A_est0_INT*g_Y_est0_INT
           [,2:(gamma_INT-start_INT+1)]);
}
```

```
  }
  exp_Y=exp_Y+exp_Y_lvl
  exp_AY=exp_AY+exp_AY_lvl
 }
 List <- list("exp_Y" = exp_Y, "exp_AY" = exp_AY);
 return(List);
}

#INTRODUCTION MLMC FUNCTION
#calculates MLMC estimators for a given SU step width
incr_N=Ind[1:(M)]-Ind[2:(M+1)]; #increments indicator function
Njump=1-Ind; #counting process

MLMC <- function(h_SU){
 n=T/h_SU;
 I=0:n; #partition initialization SU Decomposition
 partitionSU=I*T/n; #initialization partition SU Decomposition
 Ind_SU=Ind[INT %% (h_SU/h_INT)==0] #indicator function on SU grid
 gamma_SU=which(partitionSU==gamma);
 gamma_SU10=which(partitionSU==gamma+10);
 time_of_death=min(which(Ind_SU==0));
 max=max(time_of_death+1,gamma+10)
 exp_p=matrix(0,n,M+1);
 exp_pm=matrix(0,n,M+1);
 exp_forward_p=matrix(0,n,M+1);
 exp_forward_pm=matrix(0,n,M+1);
 exp_v=matrix(0,n,M+1);
 exp_Y=matrix(0,n,gamma_INT);
 exp_AY=matrix(0,n,gamma_INT);
 for(j in 1:max){
  print(paste("SU grid point", j))
  print(Sys.time())
  start_INT=which(partitionINT==partitionSU[j]);
  start_INT_forward=which(partitionINT==partitionSU[j+1]);
  exp_p_help=mlmc_p(n,j, lambda[start_INT]);
  exp_p[j,start_INT:(M+1)]=exp_p_help$exp_p;
  exp_pm[j,start_INT:(M+1)]=exp_p_help$exp_pm;
  if(start_INT_forward<(M+1)){
   exp_forward_p[j,start_INT_forward:(M+1)]=1/p[start_INT,start_INT
      _forward]*exp_p_help$exp_p[(start_INT_forward-start_INT+1):(M
      +2-start_INT)];
```

```
    exp_forward_pm[j,start_INT_forward:(M+1)]=1/p[start_INT,start_
       INT_forward]*exp_p_help$exp_pm[(start_INT_forward-start_INT+1
       ):(M+2-start_INT)];
   }else{
    exp_forward_p[j,start_INT_forward]=1;
    exp_forward_pm[j,start_INT_forward]=mu[start_INT_forward];
   }
   exp_v[j,start_INT]=1;
   exp_v_help=mlmc_v(n,j, r[start_INT]);
   exp_v[j,(start_INT+1):(M+1)]=exp_v_help$exp_v;
   if(j<gamma_SU){
    exp_Y[j,start_INT]=Y[start_INT];
    exp_AY[j,start_INT]=0;
    exp_Y_help=mlmc_Y(n,j, Y[start_INT]);
    exp_Y[j,(start_INT+1):gamma_INT]=exp_Y_help$exp_Y;
    exp_AY[j,(start_INT+1):gamma_INT]=exp_Y_help$exp_AY;
   } else if(j==gamma_SU){
    exp_Y[j,start_INT]=Y[gamma_INT];
   }
  }
  List <- list("exp_p" = exp_p, "exp_forward_p"=exp_forward_p, "exp_
     pm"=exp_pm, "exp_forward_pm"=exp_forward_pm, "exp_v"=exp_v, "
     exp_AY"=exp_AY, "exp_Y"=exp_Y);
  return(List);
}

#INTRODUCTION SU FUNCTIONS
#calculate SU decompositions for different orders in the risk basis
SU1<-function(h_SU,Exp){
  n=T/h_SU;
  I=0:n; #partition initialization ISU Decomposition
  partitionSU=I*T/n;
  D=matrix(0,16,n+1);
  gamma_SU=which(partitionSU==gamma);
  gamma_SU10=which(partitionSU==gamma+10);
  exp_p=unlist(Exp$exp_p)
  exp_pm=unlist(Exp$exp_pm)
  exp_forward_p=unlist(Exp$exp_forward_p)
  exp_forward_pm=unlist(Exp$exp_forward_pm)
  exp_v=unlist(Exp$exp_v)
  exp_AY=unlist(Exp$exp_AY)
```

```
exp_Y=unlist(Exp$exp_Y)
for (j in 2:gamma_SU){
 print(j)
 b=which(partitionINT==partitionSU[j-1]);
 b_f=which(partitionINT==partitionSU[j]);
 D[1,j]=D[1,j-1]+v[b]*(Ind[b_f]*exp_forward_p[j-1,gamma_INT]-Ind[b
    ]*exp_p[j-1,gamma_INT])*exp_v[j-1,gamma_INT]*(exp_AY[j-1,gamma
    _INT]+A[b]*exp_Y[j-1,gamma_INT]);
 D[2,j]=D[2,j-1]+v[b]*Ind[b_f]*(exp_p[j,gamma_INT]-exp_forward_p[j
    -1,gamma_INT])*exp_v[j-1,gamma_INT]*(exp_AY[j-1,gamma_INT]+A[b
    ]*exp_Y[j-1,gamma_INT]);
 D[3,j]=D[3,j-1]+Ind[b_f]*exp_p[j,gamma_INT]*(v[b_f]*exp_v[j,gamma
    _INT]-v[b]*exp_v[j-1,gamma_INT])*(exp_AY[j-1,gamma_INT]+A[b]*
    exp_Y[j-1,gamma_INT]);
 D[4,j]=D[4,j-1]+Ind[b_f]*exp_p[j,gamma_INT]*v[b_f]*exp_v[j,gamma_
    INT]*((exp_AY[j,gamma_INT]+A[b_f]*exp_Y[j,gamma_INT])-(exp_AY[
    j-1,gamma_INT]+A[b]*exp_Y[j-1,gamma_INT]));
 D[5,j]=D[5,j-1]+sum(v[b]*exp_v[j-1,b:(b_f-1)]*(1-f_d)*(exp_AY[j-1
    ,b:(b_f-1)]+A[b]*exp_Y[j-1,b:(b_f-1)])*(incr_N[b:(b_f-1)]-Ind[
    b]*p[b,b:(b_f-1)]*mu[b:(b_f-1)]*T/M))+(j<gamma_SU)*sum(v[b]*(
    Ind[b_f]*exp_forward_pm[j-1,b_f:(gamma_INT-1)]-Ind[b]*exp_pm[j
    -1,b_f:(gamma_INT-1)])*exp_v[j-1,b_f:(gamma_INT-1)]*(1-f_d)*(
    exp_AY[j-1,b_f:(gamma_INT-1)]+A[b]*exp_Y[j-1,b_f:(gamma_INT-1)
    ])*T/M);
 D[6,j]=D[6,j-1]+(j<gamma_SU)*sum(v[b]*Ind[b_f]*(exp_pm[j,b_f:(
    gamma_INT-1)]-exp_forward_pm[j-1,b_f:(gamma_INT-1)])*exp_v[j-1
    ,b_f:(gamma_INT-1)]*(1-f_d)*(exp_AY[j-1,b_f:(gamma_INT-1)]+A[b
    ]*exp_Y[j-1,b_f:(gamma_INT-1)])*T/M);
 D[7,j]=D[7,j-1]+sum((v[b:(b_f-1)]-v[b]*exp_v[j-1,b:(b_f-1)])*(1-f
    _d)*(exp_AY[j-1,b:(b_f-1)]+A[b]*exp_Y[j-1,b:(b_f-1)])*incr_N[b
    :(b_f-1)])+(j<gamma_SU)*sum(Ind[b_f]*exp_pm[j,b_f:(gamma_INT-1
    )]*(v[b_f]*exp_v[j,b_f:(gamma_INT-1)]-v[b]*exp_v[j-1,b_f:(
    gamma_INT-1)])*(1-f_d)*(exp_AY[j-1,b_f:(gamma_INT-1)]+A[b]*exp
    _Y[j-1,b_f:(gamma_INT-1)])*T/M);
 D[8,j]=D[8,j-1]+sum(v[b:(b_f-1)]*(1-f_d)*(V[b:(b_f-1)]-(exp_AY[j-
    1,b:(b_f-1)]+A[b]*exp_Y[j-1,b:(b_f-1)]))*incr_N[b:(b_f-1)])+(j
    <gamma_SU)*sum(Ind[b_f]*exp_pm[j,b_f:(gamma_INT-1)]*v[b_f]*exp
    _v[j,b_f:(gamma_INT-1)]*(1-f_d)*((exp_AY[j,b_f:(gamma_INT-1)]+
    A[b_f]*exp_Y[j,b_f:(gamma_INT-1)])-(exp_AY[j-1,b_f:(gamma_INT-
    1)]+A[b]*exp_Y[j-1,b_f:(gamma_INT-1)]))*T/M);
```

```
  D[9,j]=D[9,j-1]-sum(v[b]*(Ind[b_f]*exp_forward_p[j-1,gamma_INT]-
      Ind[b]*exp_p[j-1,gamma_INT])*exp_v[j-1,gamma_INT:(gamma_INT10-
      1)]*f_p*(exp_AY[j-1,gamma_INT]+A[b]*exp_Y[j-1,gamma_INT])*T/M)
      ;
  D[10,j]=D[10,j-1]-sum(v[b]*Ind[b_f]*(exp_p[j,gamma_INT]-exp_
      forward_p[j-1,gamma_INT])*exp_v[j-1,gamma_INT:(gamma_INT10-1)
      ]*f_p*(exp_AY[j-1,gamma_INT]+A[b]*exp_Y[j-1,gamma_INT])*T/M);
  D[11,j]=D[11,j-1]-sum(Ind[b_f]*exp_p[j,gamma_INT]*(v[b_f]*exp_v[j
      ,gamma_INT:(gamma_INT10-1)]-v[b]*exp_v[j-1,gamma_INT:(gamma_
      INT10-1)])*f_p*(exp_AY[j-1,gamma_INT]+A[b]*exp_Y[j-1,gamma_INT
      ])*T/M);
  D[12,j]=D[12,j-1]-sum(Ind[b_f]*exp_p[j,gamma_INT]*v[b_f]*exp_v[j,
      gamma_INT:(gamma_INT10-1)]*f_p*((exp_AY[j,gamma_INT]+A[b_f]*
      exp_Y[j,gamma_INT])-(exp_AY[j-1,gamma_INT]+A[b]*exp_Y[j-1,
      gamma_INT]))*T/M);
  D[13,j]=D[13,j-1]-sum(v[b]*(Ind[b_f]*exp_forward_p[j-1,gamma_INT1
      0:M]-Ind[b]*exp_p[j-1,gamma_INT10:M])*exp_v[j-1,gamma_INT10:M
      ]*f_p*(exp_AY[j-1,gamma_INT]+A[b]*exp_Y[j-1,gamma_INT])*T/M);
  D[14,j]=D[14,j-1]-sum(v[b]*Ind[b_f]*(exp_p[j,gamma_INT10:M]-exp_
      forward_p[j-1,gamma_INT10:M])*exp_v[j-1,gamma_INT10:M]*f_p*(
      exp_AY[j-1,gamma_INT]+A[b]*exp_Y[j-1,gamma_INT])*T/M);
  D[15,j]=D[15,j-1]-sum(Ind[b_f]*exp_p[j,gamma_INT10:M]*(v[b_f]*exp
      _v[j,gamma_INT10:M]-v[b]*exp_v[j-1,gamma_INT10:M])*f_p*(exp_AY
      [j-1,gamma_INT]+A[b]*exp_Y[j-1,gamma_INT])*T/M);
  D[16,j]=D[16,j-1]-sum(Ind[b_f]*exp_p[j,gamma_INT10:M]*v[b_f]*exp_
      v[j,gamma_INT10:M]*f_p*((exp_AY[j,gamma_INT]+A[b_f]*exp_Y[j,
      gamma_INT])-(exp_AY[j-1,gamma_INT]+A[b]*exp_Y[j-1,gamma_INT]))
      *T/M);
}
D[1:8,(gamma_SU+1):(n+1)]=rep(D[1:8,gamma_SU],n-gamma_SU+1);
for (j in (gamma_SU+1):gamma_SU10){
  print(j)
  b=which(partitionINT==partitionSU[j-1]);
  b_f=which(partitionINT==partitionSU[j]);
  D[9,j]=D[9,j-1];
  D[10,j]=D[10,j-1];
  D[11,j]=D[11,j-1]-sum(Ind[gamma_INT]*f_p*V[gamma_INT]*(v[b:(b_f-1
      )]-v[b]*exp_v[j-1,b:(b_f-1)])*T/M)-(j<gamma_SU10)*sum(Ind[
      gamma_INT]*f_p*V[gamma_INT]*(v[b_f]*exp_v[j,b_f:(gamma_INT10-1
      )]-v[b]*exp_v[j-1,b_f:(gamma_INT10-1)])*T/M);
  D[12,j]=D[12,j-1];
```

```
  D[13,j]=D[13,j-1]-sum(v[b]*(Ind[b_f]*exp_forward_p[j-1,gamma_INT1
     0:M]-Ind[b]*exp_p[j-1,gamma_INT10:M])*exp_v[j-1,gamma_INT10:M
     ]*f_p*V[gamma_INT]*T/M);
  D[14,j]=D[14,j-1]-sum(v[b]*Ind[b_f]*(exp_p[j,gamma_INT10:M]-exp_
     forward_p[j-1,gamma_INT10:M])*exp_v[j-1,gamma_INT10:M]*f_p*V[
     gamma_INT]*T/M);
  D[15,j]=D[15,j-1]-sum(Ind[b_f]*exp_p[j,gamma_INT10:M]*(v[b_f]*exp
     _v[j,gamma_INT10:M]-v[b]*exp_v[j-1,gamma_INT10:M])*f_p*V[gamma
     _INT]*T/M);
  D[16,j]=D[16,j-1];
 }
 D[9:12,(gamma_SU10+1):(n+1)]=rep(D[9:12,gamma_SU10],n-gamma_SU10+1
    );
 for (j in (gamma_SU10+1):n){
  print(j)
  b=which(partitionINT==partitionSU[j-1]);
  b_f=which(partitionINT==partitionSU[j]);
  D[13,j]=D[13,j-1]-sum(v[b]*(Ind[b:(b_f-1)]-Ind[b]*p[b,b:(b_f-1)])
     *exp_v[j-1,b:(b_f-1)]*f_p*V[gamma_INT]*T/M)-sum(v[b]*(Ind[b_f
     ]*exp_forward_p[j-1,b_f:M]-Ind[b]*exp_p[j-1,b_f:M])*exp_v[j-1,
     b_f:M]*f_p*V[gamma_INT]*T/M);
  D[14,j]=D[14,j-1]-sum(v[b]*Ind[b_f]*(exp_p[j,b_f:M]-exp_forward_p
     [j-1,b_f:M])*exp_v[j-1,b_f:M]*f_p*V[gamma_INT]*T/M);
  D[15,j]=D[15,j-1]-sum((v[b:(b_f-1)]-v[b]*exp_v[j-1,b:(b_f-1)])*
     Ind[b:(b_f-1)]*f_p*V[gamma_INT]*T/M)-sum(Ind[b_f]*exp_p[j,b_f:
     M]*(v[b_f]*exp_v[j,b_f:M]-v[b]*exp_v[j-1,b_f:M])*f_p*V[gamma_
     INT]*T/M);
  D[16,j]=D[16,j-1];
 }
 b=which(partitionINT==partitionSU[n]);
 b_f=which(partitionINT==partitionSU[n+1]);
 D[13,n+1]=D[13,n]-sum(v[b]*(Ind[b:(b_f-1)]-Ind[b]*p[b,b:(b_f-1)])*
    exp_v[n,b:(b_f-1)]*f_p*V[gamma_INT]*T/M);
 D[14,n+1]=D[14,n];
 D[15,n+1]=D[15,n]-sum((v[b:(b_f-1)]-v[b]*exp_v[n,b:(b_f-1)])*Ind[b
    :(b_f-1)]*f_p*V[gamma_INT]*T/M);
 D[16,n+1]=D[16,n];
 return(D);
}

SU2<-function(h_SU,Exp){
```

```
n=T/h_SU;
I=0:n; #partition initialization ISU Decomposition
partitionSU=I*T/n;
D=matrix(0,16,n+1);
gamma_SU=which(partitionSU==gamma);
gamma_SU10=which(partitionSU==gamma+10);
exp_p=unlist(Exp$exp_p)
exp_pm=unlist(Exp$exp_pm)
exp_forward_p=unlist(Exp$exp_forward_p)
exp_forward_pm=unlist(Exp$exp_forward_pm)
exp_v=unlist(Exp$exp_v)
exp_AY=unlist(Exp$exp_AY)
exp_Y=unlist(Exp$exp_Y)
for (j in 2:gamma_SU){
 print(j)
 b=which(partitionINT==partitionSU[j-1]);
 b_f=which(partitionINT==partitionSU[j]);
 D[1,j]=D[1,j-1]+Ind[b]*exp_p[j-1,gamma_INT]*v[b]*exp_v[j-1,gamma_
    INT]*((exp_AY[j,gamma_INT]+A[b_f]*exp_Y[j,gamma_INT])-(exp_AY[
    j-1,gamma_INT]+A[b]*exp_Y[j-1,gamma_INT]));
 D[2,j]=D[2,j-1]+Ind[b]*exp_p[j-1,gamma_INT]*(v[b_f]*exp_v[j,gamma
    _INT]-v[b]*exp_v[j-1,gamma_INT])*(exp_AY[j,gamma_INT]+A[b_f]*
    exp_Y[j,gamma_INT]);
 D[3,j]=D[3,j-1]+Ind[b]*v[b_f]*(p[b,b_f]*exp_p[j,gamma_INT]-exp_p[
    j-1,gamma_INT])*exp_v[j,gamma_INT]*(exp_AY[j,gamma_INT]+A[b_f
    ]*exp_Y[j,gamma_INT]);
 D[4,j]=D[4,j-1]+(Ind[b_f]-Ind[b]*p[b,b_f])*v[b_f]*exp_p[j,gamma_
    INT]*exp_v[j,gamma_INT]*(exp_AY[j,gamma_INT]+A[b_f]*exp_Y[j,
    gamma_INT]);
 D[5,j]=D[5,j-1]+sum(Ind[b]*v[b]*p[b,b:(b_f-1)]*mu[b:(b_f-1)]*exp_
    v[j-1,b:(b_f-1)]*(1-f_d)*(V[b:(b_f-1)]-(exp_AY[j-1,b:(b_f-1)]+
    A[b]*exp_Y[j-1,b:(b_f-1)]))*T/M)+(j<gamma_SU)*sum(v[b]*Ind[b]*
    exp_pm[j-1,b_f:(gamma_INT-1)]*exp_v[j-1,b_f:(gamma_INT-1)]*(1-
    f_d)*((exp_AY[j,b_f:(gamma_INT-1)]+A[b_f]*exp_Y[j,b_f:(gamma_
    INT-1)])-(exp_AY[j-1,b_f:(gamma_INT-1)]+A[b]*exp_Y[j-1,b_f:(
    gamma_INT-1)]))*T/M);
 D[6,j]=D[6,j-1]+sum(Ind[b]*p[b,b:(b_f-1)]*mu[b:(b_f-1)]*(v[b:(b_f
    -1)]-v[b]*exp_v[j-1,b:(b_f-1)])*(1-f_d)*V[b:(b_f-1)]*T/M)+(j<
    gamma_SU)*sum(Ind[b]*exp_pm[j-1,b_f:(gamma_INT-1)]*(v[b_f]*exp
    _v[j,b_f:(gamma_INT-1)]-v[b]*exp_v[j-1,b_f:(gamma_INT-1)])*(1-
    f_d)*(exp_AY[j,b_f:(gamma_INT-1)]+A[b_f]*exp_Y[j,b_f:(gamma_
```

```
    INT-1)])*T/M);
 D[7,j]=D[7,j-1]+(j<gamma_SU)*sum(Ind[b]*(p[b,b_f]*exp_pm[j,b_f:(
    gamma_INT-1)]-exp_pm[j-1,b_f:(gamma_INT-1)])*v[b_f]*exp_v[j,b_
    f:(gamma_INT-1)]*(1-f_d)*(exp_AY[j,b_f:(gamma_INT-1)]+A[b_f]*
    exp_Y[j,b_f:(gamma_INT-1)])*T/M);
 D[8,j]=D[8,j-1]+sum(v[b:(b_f-1)]*(1-f_d)*V[b:(b_f-1)]*(incr_N[b:(
    b_f-1)]-Ind[b]*p[b,b:(b_f-1)]*mu[b:(b_f-1)]*T/M))+(j<gamma_SU)
    *sum((Ind[b_f]-Ind[b]*p[b,b_f])*exp_pm[j,b_f:(gamma_INT-1)]*v[
    b_f]*exp_v[j,b_f:(gamma_INT-1)]*(1-f_d)*(exp_AY[j,b_f:(gamma_
    INT-1)]+A[b_f]*exp_Y[j,b_f:(gamma_INT-1)])*T/M);
 D[9,j]=D[9,j-1]-sum(v[b]*Ind[b]*exp_p[j-1,gamma_INT]*exp_v[j-1,
    gamma_INT:(gamma_INT10-1)]*f_p*((exp_AY[j,gamma_INT]+A[b_f]*
    exp_Y[j,gamma_INT])-(exp_AY[j-1,gamma_INT]+A[b]*exp_Y[j-1,
    gamma_INT]))*T/M);
 D[10,j]=D[10,j-1]-sum(Ind[b]*exp_p[j-1,gamma_INT]*(v[b_f]*exp_v[j
    ,gamma_INT:(gamma_INT10-1)]-v[b]*exp_v[j-1,gamma_INT:(gamma_
    INT10-1)])*f_p*(exp_AY[j,gamma_INT]+A[b_f]*exp_Y[j,gamma_INT])
    *T/M);
 D[11,j]=D[11,j-1]-sum(Ind[b]*(p[b,b_f]*exp_p[j,gamma_INT]-exp_p[j
    -1,gamma_INT])*v[b_f]*exp_v[j,gamma_INT:(gamma_INT10-1)]*f_p*(
    exp_AY[j,gamma_INT]+A[b_f]*exp_Y[j,gamma_INT])*T/M);
 D[12,j]=D[12,j-1]-sum((Ind[b_f]-Ind[b]*p[b,b_f])*exp_p[j,gamma_
    INT]*v[b_f]*exp_v[j,gamma_INT:(gamma_INT10-1)]*f_p*(exp_AY[j,
    gamma_INT]+A[b_f]*exp_Y[j,gamma_INT])*T/M);
 D[13,j]=D[13,j-1]-sum(v[b]*Ind[b]*exp_p[j-1,gamma_INT10:M]*exp_v[
    j-1,gamma_INT10:M]*f_p*((exp_AY[j,gamma_INT]+A[b_f]*exp_Y[j,
    gamma_INT])-(exp_AY[j-1,gamma_INT]+A[b]*exp_Y[j-1,gamma_INT]))
    *T/M);
 D[14,j]=D[14,j-1]-sum((v[b_f]*exp_v[j,gamma_INT10:M]-v[b]*exp_v[j
    -1,gamma_INT10:M])*Ind[b]*exp_p[j-1,gamma_INT10:M]*f_p*(exp_AY
    [j,gamma_INT]+A[b_f]*exp_Y[j,gamma_INT])*T/M);
 D[15,j]=D[15,j-1]-sum(Ind[b]*(p[b,b_f]*exp_p[j,gamma_INT10:M]-exp
    _p[j-1,gamma_INT10:M])*v[b_f]*exp_v[j,gamma_INT10:M]*f_p*(exp_
    AY[j,gamma_INT]+A[b_f]*exp_Y[j,gamma_INT])*T/M);
 D[16,j]=D[16,j-1]-sum((Ind[b_f]-Ind[b]*p[b,b_f])*exp_p[j,gamma_
    INT10:M]*v[b_f]*exp_v[j,gamma_INT10:M]*f_p*(exp_AY[j,gamma_INT
    ]+A[b_f]*exp_Y[j,gamma_INT])*T/M);
}
D[1:8,(gamma_SU+1):(n+1)]=rep(D[1:8,gamma_SU],n-gamma_SU+1);
for (j in (gamma_SU+1):gamma_SU10){
 print(j)
```

```
 b=which(partitionINT==partitionSU[j-1]);
 b_f=which(partitionINT==partitionSU[j]);
 D[9,j]=D[9,j-1];
 D[10,j]=D[10,j-1]-sum(Ind[gamma_INT]*f_p*V[gamma_INT]*(v[b:(b_f-1
    )]-v[b]*exp_v[j-1,b:(b_f-1)])*T/M)-(j<gamma_SU10)*sum(Ind[
    gamma_INT]*f_p*V[gamma_INT]*(v[b_f]*exp_v[j,b_f:(gamma_INT10-1
    )]-v[b]*exp_v[j-1,b_f:(gamma_INT10-1)])*T/M);
 D[11,j]=D[11,j-1];
 D[12,j]=D[12,j-1];
 D[13,j]=D[13,j-1];
 D[14,j]=D[14,j-1]-sum((v[b_f]*exp_v[j,gamma_INT10:M]-v[b]*exp_v[j
    -1,gamma_INT10:M])*Ind[b]*exp_p[j-1,gamma_INT10:M]*f_p*V[gamma
    _INT]*T/M);
 D[15,j]=D[15,j-1]-sum(Ind[b]*(p[b,b_f]*exp_p[j,gamma_INT10:M]-exp
    _p[j-1,gamma_INT10:M])*v[b_f]*exp_v[j,gamma_INT10:M]*f_p*V[
    gamma_INT]*T/M);
 D[16,j]=D[16,j-1]-sum((Ind[b_f]-Ind[b]*p[b,b_f])*exp_p[j,gamma_
    INT10:M]*v[b_f]*exp_v[j,gamma_INT10:M]*f_p*V[gamma_INT]*T/M);
}
D[9:12,(gamma_SU10+1):(n+1)]=rep(D[9:12,gamma_SU10],n-gamma_SU10+1
   );
for (j in (gamma_SU10+1):n){
 print(j)
 b=which(partitionINT==partitionSU[j-1]);
 b_f=which(partitionINT==partitionSU[j]);
 D[13,j]=D[13,j-1];
 D[14,j]=D[14,j-1]-sum(Ind[b]*(v[b:(b_f-1)]-v[b]*exp_v[j-1,b:(b_f-
    1)])*p[b,b:(b_f-1)]*f_p*V[gamma_INT]*T/M)-sum(Ind[b]*exp_p[j-1
    ,b_f:M]*(v[b_f]*exp_v[j,b_f:M]-v[b]*exp_v[j-1,b_f:M])*f_p*V[
    gamma_INT]*T/M);
 D[15,j]=D[15,j-1]-sum(Ind[b]*v[b_f]*(p[b,b_f]*exp_p[j,b_f:M]-exp_
    p[j-1,b_f:M])*exp_v[j,b_f:M]*f_p*V[gamma_INT]*T/M);
 D[16,j]=D[16,j-1]-sum(v[b:(b_f-1)]*(Ind[b:(b_f-1)]-Ind[b]*p[b,b:(
    b_f-1)])*f_p*V[gamma_INT]*T/M)-sum(v[b_f]*(Ind[b_f]-Ind[b]*p[b
    ,b_f])*exp_p[j,b_f:M]*exp_v[j,b_f:M]*f_p*V[gamma_INT]*T/M);
}
b=which(partitionINT==partitionSU[n]);
b_f=which(partitionINT==partitionSU[n+1]);
D[13,n+1]=D[13,n];
D[14,n+1]=D[14,n]-sum(Ind[b]*(v[b:(b_f-1)]-v[b]*exp_v[j-1,b:(b_f-1
   )])*p[b,b:(b_f-1)]*f_p*V[gamma_INT]*T/M);
```

```
 D[15,n+1]=D[15,n];
 D[16,n+1]=D[16,n]-sum(v[b:(b_f-1)]*(Ind[b:(b_f-1)]-Ind[b]*p[b,b:(b
    _f-1)])*f_p*V[gamma_INT]*T/M);
 return(D);
}

### MLMC estimation ###
# MLMC parameters
L=4; #number of MLMC levels
k_MLMC<-1;
h_MLMC=(1/2)^(k_MLMC-1)*h_INT; #MLMC starting step width
pck=1000 #simulation package
N=c(100,50,20,10,5); #number of packages per level
#MLMC calculation
Exp_yearly=MLMC(1);
Exp_quarterly=MLMC(1/4);
Exp_4weekly=MLMC(1/13);

#SU decompositions for both orders
D_yearly1=SU1(1,Exp_yearly);
D_yearly2=SU2(1,Exp_yearly);
D_quarterly1=SU1(1/4,Exp_quarterly);
D_quarterly2=SU2(1/4,Exp_quarterly);
D_4weekly1=SU1(1/13,Exp_4weekly);
D_4weekly2=SU2(1/13,Exp_4weekly);

#Add up SU addends for the different claim components and form the
   differences between the two orders
D_yearly1_total<- matrix(0,4,ncol(D_yearly1))
D_yearly2_total<- matrix(0,4,ncol(D_yearly2))
D_quarterly1_total<- matrix(0,4,ncol(D_quarterly1))
D_quarterly2_total<- matrix(0,4,ncol(D_quarterly2))
D_4weekly1_total<- matrix(0,4,ncol(D_4weekly1))
D_4weekly2_total<- matrix(0,4,ncol(D_4weekly2))
D_yearly_diff<- matrix(0,4,ncol(D_yearly1))
D_quarterly_diff<- matrix(0,4,ncol(D_quarterly1))
D_4weekly_diff<- matrix(0,4,ncol(D_4weekly1))

D_yearly1_total[1,]=D_yearly1[1,]+D_yearly1[5,]+D_yearly1[9,]+D_
   yearly1[13,]
```

```
D_yearly1_total[2,]=D_yearly1[2,]+D_yearly1[6,]+D_yearly1[10,]+D_
    yearly1[14,]
D_yearly1_total[3,]=D_yearly1[3,]+D_yearly1[7,]+D_yearly1[11,]+D_
    yearly1[15,]
D_yearly1_total[4,]=D_yearly1[4,]+D_yearly1[8,]+D_yearly1[12,]+D_
    yearly1[16,]
D_quarterly1_total[1,]=D_quarterly1[1,]+D_quarterly1[5,]+D_
    quarterly1[9,]+D_quarterly1[13,]
D_quarterly1_total[2,]=D_quarterly1[2,]+D_quarterly1[6,]+D_
    quarterly1[10,]+D_quarterly1[14,]
D_quarterly1_total[3,]=D_quarterly1[3,]+D_quarterly1[7,]+D_
    quarterly1[11,]+D_quarterly1[15,]
D_quarterly1_total[4,]=D_quarterly1[4,]+D_quarterly1[8,]+D_
    quarterly1[12,]+D_quarterly1[16,]
D_4weekly1_total[1,]=D_4weekly1[1,]+D_4weekly1[5,]+D_4weekly1[9,]+D
    _4weekly1[13,]
D_4weekly1_total[2,]=D_4weekly1[2,]+D_4weekly1[6,]+D_4weekly1[10,]+
    D_4weekly1[14,]
D_4weekly1_total[3,]=D_4weekly1[3,]+D_4weekly1[7,]+D_4weekly1[11,]+
    D_4weekly1[15,]
D_4weekly1_total[4,]=D_4weekly1[4,]+D_4weekly1[8,]+D_4weekly1[12,]+
    D_4weekly1[16,]
D_yearly2_total[4,]=D_yearly2[1,]+D_yearly2[5,]+D_yearly2[9,]+D_
    yearly2[13,]
D_yearly2_total[3,]=D_yearly2[2,]+D_yearly2[6,]+D_yearly2[10,]+D_
    yearly2[14,]
D_yearly2_total[2,]=D_yearly2[3,]+D_yearly2[7,]+D_yearly2[11,]+D_
    yearly2[15,]
D_yearly2_total[1,]=D_yearly2[4,]+D_yearly2[8,]+D_yearly2[12,]+D_
    yearly2[16,]
D_quarterly2_total[4,]=D_quarterly2[1,]+D_quarterly2[5,]+D_
    quarterly2[9,]+D_quarterly2[13,]
D_quarterly2_total[3,]=D_quarterly2[2,]+D_quarterly2[6,]+D_
    quarterly2[10,]+D_quarterly2[14,]
D_quarterly2_total[2,]=D_quarterly2[3,]+D_quarterly2[7,]+D_
    quarterly2[11,]+D_quarterly2[15,]
D_quarterly2_total[1,]=D_quarterly2[4,]+D_quarterly2[8,]+D_
    quarterly2[12,]+D_quarterly2[16,]
D_4weekly2_total[4,]=D_4weekly2[1,]+D_4weekly2[5,]+D_4weekly2[9,]+D
    _4weekly2[13,]
```

```
D_4weekly2_total[3,]=D_4weekly2[2,]+D_4weekly2[6,]+D_4weekly2[10,]+D_4weekly2[14,]
D_4weekly2_total[2,]=D_4weekly2[3,]+D_4weekly2[7,]+D_4weekly2[11,]+D_4weekly2[15,]
D_4weekly2_total[1,]=D_4weekly2[4,]+D_4weekly2[8,]+D_4weekly2[12,]+D_4weekly2[16,]
D_yearly_diff=D_yearly1_total-D_yearly2_total
D_quarterly_diff=D_quarterly1_total-D_quarterly2_total
D_4weekly_diff=D_4weekly1_total-D_4weekly2_total

###PLOTS###
x_INT=seq(0,T,h_INT);
x_yearly=seq(0,T,1);
x_quarterly=seq(0,T,1/4);
x_4weekly=seq(0,T,1/13);

#PLOT RISK DRIVERS
par(mfrow=c(2, 2), oma=c(0,0,3,0))
plot(x_INT, Njump, type="l", xlab = "Time␣in␣years", ylab= "Jump␣process␣N", cex.lab=1.1, ylim=c(0,1), yaxt='n')
axis(side = 2, at = seq(0,1,0.25))
plot(x_INT, lambda, type="l", xlab = "Time␣in␣years", ylab= "Mortality␣intensity␣\u03BB", cex.lab=1.2)
plot(x_INT, r, type="l", xlab = "Time␣in␣years", ylab= "Interest␣intensity␣r", cex.lab=1.2)
plot(x_INT, Y, type="l", xlab = "Time␣in␣years", ylab= "Market␣index␣Y", yaxt='n',cex.lab=1.2, ylim=c(0,2000))
axis(side = 2, at = seq(0,2000,1000))
axis(side = 2, at = seq(0,2000,500),labels=NA)
options(scipen=999)

#PLOT SU ADDENDS FOR DIFFERENT SU GRID WIDTHS
par(mfrow=c(4, 3), oma=c(0,0,3,0))
plot(x_yearly, D_yearly1_total[1,], type="l", xlab = "", ylab= "", main="\nyearly", ylim=c(-30000,100000),cex.lab=1.1, col="coral1", font.main = 2)
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
plot(x_quarterly, D_quarterly1_total[1,], type="l",xlab="", ylab="", main= "\nquarterly", ylim=c(-30000,100000),cex.lab=1.1, col="darkslategray4", font.main = 2)
```

```
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
plot(x_4weekly, D_4weekly1_total[1,], type="l", xlab = "", ylab= ""
    , main= "\n4-weekly", ylim=c(-30000,100000),cex.lab=1.1, col="
    darkseagreen3", font.main = 2)
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
plot(x_yearly,D_yearly1_total[2,], type="l", xlab = "",ylab= "",
    main="\nyearly",cex.lab=1.1, ylim=c(-2500,2200),col="coral1",
    font.main = 2)
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
plot(x_quarterly, D_quarterly1_total[2,], type="l", xlab = "", ylab
    = "", main= "\nquarterly",  ylim=c(-2500,2200),cex.lab=1.1, col=
    "darkslategray4", font.main = 2)
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
plot(x_4weekly, D_4weekly1_total[2,], type="l", xlab = "", ylab= ""
    , main= "\n4-weekly",  ylim=c(-2500,2200),cex.lab=1.1, col="
    darkseagreen3", font.main = 2)
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
plot(x_yearly, D_yearly1_total[3,], type="l", xlab = "", ylab= "",
    main="\nyearly", ylim=c(-8500,30000),cex.lab=1.1,col="coral1",
    font.main = 2)
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
plot(x_quarterly, D_quarterly1_total[3,], type="l", xlab = "", ylab
    = "",main="\nquarterly", ylim=c(-8500,30000),cex.lab=1.1, col="
    darkslategray4", font.main = 2)
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
plot(x_4weekly, D_4weekly1_total[3,], type="l", xlab = "", ylab= ""
    ,main="\n4-weekly", ylim=c(-8500,30000),cex.lab=1.1, col="
    darkseagreen3", font.main = 2)
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
plot(x_yearly, D_yearly1_total[4,], type="l", xlab = "", ylab= "",
    main="\nyearly", ylim=c(-7000,100000),cex.lab=1.1,col="coral1",
    font.main = 2)
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
```

```
title(ylab="Surplus", line=2.5, cex.lab=1.2)
plot(x_quarterly, D_quarterly1_total[4,], type="l", xlab = "", ylab
   = "",main="\nquarterly", ylim=c(-7000,100000),cex.lab=1.1, col="
   darkslategray4", font.main = 2)
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
plot(x_4weekly, D_4weekly1_total[4,], type="l", xlab = "", ylab= ""
   , main="\n4-weekly", ylim=c(-7000,100000),cex.lab=1.1, col="
   darkseagreen3", font.main = 2)
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
mtext("Unsystematic␣biometric␣surplus␣(N)", side = 3, line = -1.4,
   outer = TRUE, font = 2)
mtext("Systematic␣biometric␣surplus␣(\u03BB)", side = 3, line = -17
   .5, outer = TRUE, font = 2)
mtext("Systematic␣interest␣surplus␣(r)", side = 3, line = -34,
   outer = TRUE, font = 2)
mtext("Systematic␣fund␣surplus␣(Y)", side = 3, line = -50.3, outer
   = TRUE, font = 2)
options(scipen=999)

#PLOT TOTAL SURPLUS
par(mfrow=c(1, 3), oma=c(0,0,3,0))
plot(x_yearly, D_yearly1_total[1,]+D_yearly1_total[2,]+D_yearly1_
   total[3,]+D_yearly1_total[4,], type="l", xlab = "", ylab= "",
   main="Yearly␣time␣grid", ylim=c(0,200000),col="coral1")
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
plot(x_quarterly, D_quarterly1_total[1,]+D_quarterly1_total[2,]+D_
   quarterly1_total[3,]+D_quarterly1_total[4,], type="l", xlab = ""
   , ylab= "", main="Quarterly␣time␣grid", ylim=c(0,200000), col="
   darkslategray4")
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
plot(x_4weekly, D_4weekly1_total[1,]+D_4weekly1_total[2,]+D_4weekly
   1_total[3,]+D_4weekly1_total[4,], type="l", xlab = "", ylab= "",
    main="4-weekly␣time␣grid", ylim=c(0,200000), col="darkseagreen3
   ")
title(xlab="Time␣in␣years", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
```

```
#Barplot - Impact of the time grid on the SU decomposition
par(mfrow=c(1, 1), oma=c(0,0,3,0))
m1<- matrix(c(D_yearly1_total[1,ncol(D_yearly1)],D_yearly1_total[2,
   ncol(D_yearly1)],D_yearly1_total[3,ncol(D_yearly1)],D_yearly1_
   total[4,ncol(D_yearly1)],D_quarterly1_total[1,ncol(D_quarterly1)
   ],D_quarterly1_total[2,ncol(D_quarterly1)],D_quarterly1_total[3,
   ncol(D_quarterly1)],D_quarterly1_total[4,ncol(D_quarterly1)],D_4
   weekly1_total[1,ncol(D_4weekly1)],D_4weekly1_total[2,ncol(D_4
   weekly1)],D_4weekly1_total[3,ncol(D_4weekly1)],D_4weekly1_total[
   4,ncol(D_4weekly1)]), byrow=TRUE, nrow=3)
sources<- c("N","\u03BB","r","Y")
time <- c("yearly", "quarterly", "4-weekly")
colnames(m1) <- sources
rownames(m1) <- time
colours=c("coral1", "darkslategray4", "darkseagreen3")
#colours=c("coral1", "cyan4", "darkgoldenrod3")
barplot(m1,ylab='', xlab='',beside = TRUE, col=colours, ylim=c(0,11
   0000))
title(xlab="Risk␣factors", line=2.5, cex.lab=1.2)
title(ylab="Surplus", line=2.5, cex.lab=1.2)
options(scipen=999)
box()
legend('topright',fill=colours,legend=time)

#Barplot - Impact of the update order on the SU decomposition Part
   I
m2<- matrix(c(D_yearly_diff[1,ncol(D_yearly1)]/D_yearly1_total[1,
   ncol(D_yearly1)],D_yearly_diff[2,ncol(D_yearly1)]/D_yearly1_
   total[2,ncol(D_yearly1)],D_yearly_diff[3,ncol(D_yearly1)]/D_
   yearly1_total[1,ncol(D_yearly1)],D_yearly_diff[4,ncol(D_yearly1)
   ]/D_yearly1_total[1,ncol(D_yearly1)],D_quarterly_diff[1,ncol(D_
   quarterly1)]/D_quarterly1_total[1,ncol(D_quarterly1)],D_
   quarterly_diff[2,ncol(D_quarterly1)]/D_quarterly1_total[1,ncol(D
   _quarterly1)],D_quarterly_diff[3,ncol(D_quarterly1)]/D_quarterly
   1_total[1,ncol(D_quarterly1)],D_quarterly_diff[4,ncol(D_
   quarterly1)]/D_quarterly1_total[1,ncol(D_quarterly1)],D_4weekly_
   diff[1,ncol(D_4weekly1)]/D_4weekly1_total[1,ncol(D_4weekly1)],D_
   4weekly_diff[2,ncol(D_4weekly1)]/D_4weekly1_total[1,ncol(D_4
   weekly1)],D_4weekly_diff[3,ncol(D_4weekly1)]/D_4weekly1_total[1,
   ncol(D_4weekly1)],D_4weekly_diff[4,ncol(D_4weekly1)]/D_4weekly1_
   total[1,ncol(D_4weekly1)]), byrow=TRUE, nrow=3)
```

```
sources <- c("N","\u03BB","r","Y")
time <- c("yearly", "quarterly", "4-weekly")
colnames(m2) <- sources
rownames(m2) <- time
colours=c("coral1", "darkslategray4", "darkseagreen3")
barplot(m2,ylab='', xlab='',beside = TRUE, col=colours, ylim=c(-0.0
   4,0.05))
title(xlab="Risk␣factors", line=2.5, cex.lab=1.2)
title(ylab="Relative␣deviation", line=2.5, cex.lab=1.2)
options(scipen=999)
box()

#Barplot - Impact of the update order on the SU decomposition Part
   II
barplot(m2,beside = TRUE, col=colours, ylim=c(0.38,0.5),xaxt='n',
   yaxt='n')
options(scipen=999)
box()
legend('topright',fill=colours,legend=time)
axis(side = 2, at = seq(0.4,0.5,0.05), labels=c(0.4,0.45,0.5))

#Further information in the results paragraph
#Time of death
time_death=min(which(Ind==0));
partitionINT[time_death]
#Maximum mortality intensity
max(lambda)
#Maximum interest intensity
max(r)
#Market fund average yield
((Y[1145]-Y[1])/Y[1])^(1/22)
((Y[1041]-Y[1])/Y[1])^(1/20)
#Changes between yearly and weekly grid for the contract-end
   surplus
(D_4weekly1_total[1,716]-D_yearly1_total[1,56])/D_yearly1_total[1,5
   6]
(D_4weekly1_total[3,716]-D_yearly1_total[3,56])/D_yearly1_total[3,5
   6]
(D_4weekly1_total[2,716]-D_yearly1_total[2,56])/D_yearly1_total[2,5
   6]
```

```
(D_4weekly1_total[4,716]-D_yearly1_total[4,56])/D_yearly1_total[4,5
  6]
```

Affidavit

Hereby, I declare, that the presented dissertation with the title

Decomposition of Stochastic Surplus Processes in Life Insurance

is my own work and, that I have not used other sources than the sources stated in the text or in the bibliography. The dissertation has, neither as a whole, nor in part, been submitted for assessment in a doctoral procedure at another university. I confirm that I am aware of the guidelines of good scientific practice of the Carl von Ossietzky University Oldenburg and that I observed them. Furthermore, I declare that I have not availed myself of any commercial placement or consulting services in connection with my doctoral procedure.

Oldenburg, 17th January 2025

(Julian Jetses)