Decomposing life insurance liabilities into risk factors

Katja Schilling*
Institut für Versicherungswissenschaften, Universität Ulm
Helmholtzstaße 20; 89081 Ulm; Germany
Email: katja.schilling@uni-ulm.de

Daniel Bauer
Department of Risk Management and Insurance, Georgia State University
35 Broad Street, 11th Floor; Atlanta, GA 30303; USA
Email: dbauer@gsu.edu

Marcus C. Christiansen
Department of Actuarial Mathematics and Statistics, Heriot-Watt University
Edinburgh Campus; Edinburgh EH14 4AS; United Kingdom
Email: m.c.christiansen@hw.ac.uk

Alexander Kling
Institut für Finanz- und Aktuarwissenschaften
Lise-Meitner-Straße 14; 89081 Ulm; Germany
Email: a.kling@ifa-ulm.de

November 21, 2015

Abstract

The decomposition of (life) insurance liabilities into risk factors associated with various sources of risk such as equity, interest, or mortality is of great relevance in view of risk management and product design. Nevertheless, although several decomposition approaches have been proposed, no systematic analysis is available. The present paper closes this gap in literature by introducing properties for *meaningful* risk decompositions and demonstrating that existing approaches violate at least one of these properties. As an alternative, we propose a novel *MRT decomposition* that relies on martingale representation and show that it satisfies all of the properties. We discuss its calculation using techniques from stochastic and Malliavin calculus, and present a detailed example illustrating its applicability.

Keywords: life insurance liabilities, risk decomposition, martingale representation, Malliavin calculus, mortality risk

*Corresponding author.

1 Introduction

Decomposing insurance liabilities into risk factors associated with different sources of risk is a problem of great practical significance, particularly in the life sector, in view of risk management, product design, and capital regulation. The primary contributions of this paper are twofold: On the one hand, we introduce properties for a *meaningful risk decomposition* and show that decomposition methods proposed in literature suffer shortcomings in view of these properties. On the other hand, we propose a novel decomposition approach based on martingale representation, labeled *MRT decomposition*, and show that it satisfies all the meaningful risk decomposition properties. We discuss the calculation of the MRT decomposition in a relatively general life insurance setting with an arbitrary (finite) insurance portfolio, where the homogeneous group of policyholders is modeled by a counting process and the (systematic) sources of risk are driven by a finite-dimensional Brownian motion. We derive explicit formulas in terms of Malliavin derivatives in the general case and in terms of derivatives of conditional expectations in the Markov case. Moreover, we provide detailed example calculations in the context of a Variable Annuity contract with a Guaranteed Minimum Death Benefit (GMDB).

Insurance liabilities are influenced by various sources of risk such as equity, interest, and insurance-specific risk. The interaction of these sources can be quite complex, so that the individual risk contributions are typically neither obvious nor readily available. This is particularly the case in life insurance, where the final payoffs – that commonly occur years or even decades after the origination of the contracts – depend on the interaction of financial factors and guarantees, aggregate demographic trends, and actual deaths observed in the portfolio of insured. Nonetheless, insurance companies need to assess the relative importance of each source of risk in order to be able to devise adequate risk management strategies. This may simply be a matter of identifying the most significant source of risk for focusing efforts in case resources for risk management are limited (Hoem, 1988; Kling et al., 2014). Alternatively, the decomposition may allow to gage the sufficiency of risk loadings to each source of risk, taking into account its contribution to the aggregate risk (Christiansen, 2013; Niemeyer, 2015). Evaluating the impact of different sources of risk is also important in view of product design, particularly when there are different risk penalties for different sources of risk (Kochanski and Karnarski, 2011), and in view of calculating solvency capital requirements. For instance, within Solvency II, individual risk contributions need to be quantified explicitly in partial internal models. Also, the decomposition may help to adequately calibrate standard formulas used in regulatory frameworks.

Given the relevance of risk decompositions, it is not surprising that there are a number of papers suggesting different methodologies for deriving risk factors, particularly in the life insurance context. Bühlmann (1995), Fischer (2004), Martin and Tasche (2007), and Christiansen and Helwich (2008) use a conditional expectation approach, which is the probabilistic foundation of the well-known *variance decomposition*. Another approach also based on conditional expectations – the so-called Hoeffding decomposition – is used, for example, by Rosen and Saunders (2010). The Taylor expansion method (Christiansen, 2007) uses derivatives for decomposing functionals of different sources of risk. A completely different method, applied by Gatzert and Wesker (2014), Artinger (2010), and also implicitly used in the Solvency II framework, "switches" off the randomness of all sources of risk which are momentarily not under consideration. Karabey et al. (2014) rely on several of these approaches (variance decomposition, Hoeffding, and Taylor) and show how the contributions of different sources of risk can be derived from the risk decompositions using the Euler allocation principle.[1]

In this paper, we commence by introducing a number of properties that define a *meaningful* risk

[1]Other papers also consider the decomposition of the total "risk" as defined based on a particular risk measure. However, it is important to note that those approaches implicitly also make use of a decomposition in corresponding random variables, and then apply risk allocation techniques (see also Section 5 and Karabey (2012)). Thus, we focus on the decomposition of random variables (see property P1).

decomposition for insurance liabilities.[2] In particular, we posit that a decomposition should consider the entire distribution of the company's risk (P1), that resulting decompositions should be unique (P3) and independent of the ordering of the risks (P4), that the different risk factors can be clearly attributed to the different sources of risk (P2), that the risk factors are invariant to changes in the scale of the sources of risk (P5), and, finally, that the decomposition should aggregate to the (normalized) entire risk (P6). However, it turns out that when benchmarking the decomposition approaches proposed in literature with this list of desirable properties, for each method at least one of the properties fails to hold.

This leads us to propose our alternative *MRT decomposition*. We show that this approach satisfies each property P1 to P6, and furthermore that the risk factor associated with *unsystematic mortality risk* vanishes as the portfolio size increases – whereas the *systematic* risk factors approach a non-zero limit. We provide explicit formulas for the decomposition, assuming a general definition of the payoff of the insurance contract entailing discrete as well as continuous survival and death benefits, by relying on the Clark-Ocone formula (in the general case) and Itô's lemma for diffusion processes (in the Markov case).

Our detailed numerical example relies on an affine specification of the interest and the mortality rates following Cox et al. (1985) and Dahl and Møller (2006), respectively, and a geometric Brownian motion for the underlying Variable Annuity account. We decompose the total liabilities associated with a return-of-premium GMDB – which presents a very common product in the U.S. market – into four sources of risk: equity risk, interest rate risk, systematic mortality risk, and unsystematic mortality risk. Our calculations show that for an unhedged exposure, equity risk is by far the most dominant risk, particularly when considering moderately sized insurance portfolios. More advanced examples for Guaranteed Annuity Options and Guaranteed Minimum Income Benefits within Variable Annuities that also consider the impact of hedging are considered in a companion paper (Schilling, 2015).

From a technical perspective, the derivation of our MRT decomposition is closely related to quadratic hedging approaches for life insurance liabilities under a martingale measure (Barbarin, 2008; Biagini et al., 2012, 2013; Biagini and Schreiber, 2013; Møller, 2001; Dahl and Møller, 2006; Dahl et al., 2008; Norberg, 2013), with the conceptual difference that we operate under the physical measure since we are interested in risk assessments. We rely on this analogy in our derivations, but we also present some new results in this direction such as the decomposition of arbitrary insurance payoffs within our general setting and the integration with the Clark-Ocone formula from Malliavin calculus.

The remainder of the paper is organized as follows. Section 2 presents the properties that define a meaningful risk decomposition and analyzes whether conventional approaches from literature satisfy these properties. Section 3 lays out the considered life insurance modeling framework and introduces our MRT decomposition within this framework. Properties and the calculation of the MRT decomposition are discussed in Section 4. Section 5 describes and analyzes our Variable Annuity example. Finally, Section 6 concludes.

2 Meaningful risk decompositions

2.1 Definition of meaningful risk decompositions

As outlined in the Introduction, the primary concern of this paper is decomposing a (life) *insurer's total risk* – which we suppose is given via the (normalized) *loss* random variable L, $\mathrm{E}[L] = 0$ – into different risk factors. More precisely, we assume there are k sources of risk, where $Z_i = (Z_i(t))_{0 \leq t \leq T^*}$ denotes the i-th source of risk and $Z = (Z_1, \ldots, Z_k)$. We assume that the loss variable L is $\sigma(Z)$-measurable,

[2] Fischer (2004) also provides a list of desirable properties for a reasonable decomposition method. However, he focuses on a decomposition of life insurance liabilities into financial risk and unsystematic mortality risk, where a number of these properties are trivial or irrelevant (e.g., because of independence of the sources of risk).

and we consider decomposition methodologies that assign each source of risk a corresponding risk factor.

While several papers in the actuarial literature propose a variety of decomposition methods, thus far there has been no systematic assessment and comparison among these different approaches. In what follows, we introduce a list of properties we argue a *meaningful risk decomposition* should satisfy (equalities between random variables are in the almost sure sense):

P1 *Randomness*

Individual risk factors are given by *random variables* $R_1,\ R_2,\dots,\ R_k$, where random variable R_i corresponds to risk factor $i \in \{1,2,\dots,k\}$. We introduce the relation $\leftrightarrow$ for a decomposition methodology and write $(L, Z_1,\dots,Z_k) \leftrightarrow (R_1, R_2,\dots,R_k)$ to indicate that the loss L depending on $(Z_1,\dots,Z_k)$ corresponds to the decomposition $(R_1,R_2,\dots,R_k)$.

P2 *Attribution*

R_i represents the risk factor related to risk i. Formally, we require that whenever the loss L is $\sigma(Z_i)$-measurable and Z_i is independent of $(Z_1,\dots,Z_{i-1},Z_{i+1},\dots,Z_k)$, then $R_j = 0$ for all $j \neq i$.

P3 *Uniqueness*

The decomposition methodology yields a unique decomposition. Formally, we require that $(L, Z_1,\dots,Z_k) \leftrightarrow (R_1,R_2,\dots,R_k)$ and $(L, Z_1,\dots,Z_k) \leftrightarrow (\tilde{R}_1,\tilde{R}_2,\dots,\tilde{R}_k)$ implies $R_i = \tilde{R}_i$, $i \in \{1,2,\dots,k\}$.

P4 *Order invariance*

The decomposition is invariant to the order of the risks $1,\ 2,\dots,\ k$. Formally, consider a permutation $\pi : \{1,2,\dots,k\} \to \{1,2,\dots,k\}$ and assume $(L, Z_1,\dots,Z_k) \leftrightarrow (R_1,R_2,\dots,R_k)$. Then we require:

$$(L, Z_{\pi(1)},\dots,Z_{\pi(k)}) \leftrightarrow (R_{\pi(1)},R_{\pi(2)},\dots,R_{\pi(k)}).$$

P5 *Scale invariance*

The decomposition is invariant to changes in the scale of the sources of risk. Formally, assume $(L, Z_1,\dots,Z_k) \leftrightarrow (R_1,R_2,\dots,R_k)$, and let $\tilde{Z}_i(t) := f_i(Z_i(t))$ for all $i = 1,\dots,k,\ 0 \le t \le T^*$, where, for each i, $f_i : \mathbb{R} \to \mathbb{R}$ is a smooth, invertible function. If $(L, \tilde{Z}_1,\dots,\tilde{Z}_k) \leftrightarrow (\tilde{R}_1,\tilde{R}_2,\dots,\tilde{R}_k)$, then we require that $R_i = \tilde{R}_i$ for all $i \in \{1,\dots,k\}$.

P6 *Aggregation*

The decomposition aggregates to the total risk faced by the company. Formally, we require that for each loss L and risks Z with $(L, Z_1,\dots,Z_k) \leftrightarrow (R_1,R_2,\dots,R_k)$, there exists a function $A_{(L,Z)} : \mathbb{R}^k \to \mathbb{R}$ such that

$$L = A_{(L,Z)}(R_1,R_2,\dots,R_k).$$

P6* *Additive aggregation*

A special case of P6 is an additive aggregation function, i.e. the case where L is given as the sum of the individual risk factors:

$$L = \sum_{i=1}^{k} R_i.$$

Note that the relation $\leftrightarrow$ will be a function if P3 is satisfied. Furthermore, if additionally P6 holds and the function $A_{(L,Z)}$ does not depend on L (as is e.g. the case under P6*), then $\leftrightarrow$ is injective in L for fixed Z since

$$(R_1, \dots, R_k) = (\tilde{R}_1, \dots, \tilde{R}_k) \Rightarrow L = A_Z(R_1, \dots, R_k) = A_Z(\tilde{R}_1, \dots, \tilde{R}_k) = \tilde{L}.$$

The scale invariance property (P5) is necessary that the risk factors are quantitatively comparable even if they are related to different loss variables. An additive decomposition (P6*) is desirable for multiple reasons. For instance, it allows for the natural interpretation that the risk factors *sum up* to the total risk. Moreover, for a decomposition into summands, it is straightforward to derive decompositions for homogeneous risk measures as within the well-known Euler allocation principle (Karabey et al., 2014).

2.2 Discussion: Are conventional approaches meaningful?

For discussing conventional decomposition approaches with regards to the meaningful risk decomposition properties, we consider the time-0 present value L_0 of an insurer's future losses and, for simplicity, assume that it is only influenced by two sources of risk $Z_1 = (Z_1(t))_{0\le t\le T^*}$ and $Z_2 = (Z_2(t))_{0\le t\le T^*}$. The insurer's risk is identified with $L := L_0 - \mathrm{E}^{\mathbb{P}}(L_0)$. To preview our results, we find that each considered decomposition approach fails to satisfy at least one property, which leads us to propose a new decomposition method in the next section.

Variance decomposition

A common approach for decomposing the insurer's risk $L_0 - \mathrm{E}^{\mathbb{P}}(L_0)$ into risk factors is a conditional expectation approach. Bühlmann (1995) and Fischer (2004) use this approach to decompose the profit/loss of a life insurer into a financial and a biometric part; Martin and Tasche (2007) determine the systematic and unsystematic risk in a credit portfolio by this method; and Christiansen and Helwich (2008) extend the approach to three sources of risk of a life insurance portfolio, namely unsystematic and systematic mortality risk as well as financial risk.

The basic idea is that the conditional expectation $R_1 := \mathrm{E}^{\mathbb{P}}(L|\, Z_1)$ captures the randomness of L caused by Z_1. Since the remaining risk $R_2 := L - R_1 = L - \mathrm{E}^{\mathbb{P}}(L|\, Z_1)$ must represent the randomness caused by Z_2, the decomposition for $L = L_0 - \mathrm{E}^{\mathbb{P}}(L_0)$ reads as

$$L_0 - \mathrm{E}^{\mathbb{P}}(L_0) = \left[\mathrm{E}^{\mathbb{P}}(L_0|\, Z_1) - \mathrm{E}^{\mathbb{P}}(L_0)\right] + \left[L_0 - \mathrm{E}^{\mathbb{P}}(L_0|\, Z_1)\right] = R_1 + R_2, \tag{2.1}$$

where R_1 and R_2 represent the two risk factors. As a result of the orthogonality property of conditional expectations, a straightforward consequence of (2.1) is

$$\mathrm{Var}(L) = \mathrm{Var}(R_1) + \mathrm{Var}(R_2).$$

Commonly, the latter equation is referred to as *variance decomposition* and frequently is the basis for applications (thus, we simply refer to the general decomposition (2.1) as "variance decomposition"). Note that for an arbitrary loss L the variance decomposition directly implies that $\mathrm{E}^{\mathbb{P}}(R_1) = \mathrm{E}^{\mathbb{P}}(L)$ and $\mathrm{E}^{\mathbb{P}}(R_2) = 0$. Of course, this asymmetry is irrelevant when considering the variance but potentially relevant when applying different risk measures. This emphasizes the necessity to first standardize the loss L_0 to mean zero, i.e. considering $L_0 - \mathrm{E}^{\mathbb{P}}(L_0)$, and then apply the decomposition approach, resulting in $\mathrm{E}^{\mathbb{P}}(R_1) = \mathrm{E}^{\mathbb{P}}(R_2) = 0$.

Obviously, the risk factors R_1 and R_2 are random variables (P1) and they add up to the total risk (P6*/P6). Since conditional expectations are unique almost surely, so is the variance decomposition (P3). To check the attribution property P2, for independent Z_1 and Z_2 and a $\sigma(Z_1)$-measurable loss

L, $R_2 = L - \mathrm{E}^{\mathbb{P}}(L|\, Z_1) = 0$. Conversely, if L is $\sigma(Z_2)$-measurable, $R_1 = \mathrm{E}^{\mathbb{P}}(L|\, Z_1) = \mathrm{E}^{\mathbb{P}}(L)$. Therefore, P2 is satisfied since L is standardized to mean zero. The variance decomposition is also scale invariant (P5), since for two smooth, invertible functions f_1 and f_2, with $\tilde{Z}_i(t) := f_i(Z_i(t)),\ i = 1, 2$, we have $\sigma(\tilde{Z}_i) = \sigma(Z_i)$, so that $\tilde{R}_1 = \mathrm{E}^{\mathbb{P}}(L|\, \tilde{Z}_1) = \mathrm{E}^{\mathbb{P}}(L|\, Z_1) = R_1$ and $\tilde{R}_2 = L - \mathrm{E}^{\mathbb{P}}(L|\, \tilde{Z}_1) = L - \mathrm{E}^{\mathbb{P}}(L|\, Z_1) = R_2$.

However, as the following example illustrates, the order invariance property P4 is not satisfied:

Example 2.1. Assume that $L_0 = Z_1(T)Z_2(T)$, where Z_1 and Z_2 are two independent processes. Then the variance decomposition with respect to $Z = (Z_1, Z_2)$ is

$$\begin{aligned} L_0 - \mathrm{E}^{\mathbb{P}}(L_0) &= \left[\mathrm{E}^{\mathbb{P}}(L_0|\, Z_1) - \mathrm{E}^{\mathbb{P}}(L_0)\right] + \left[L_0 - \mathrm{E}^{\mathbb{P}}(L_0|\, Z_1)\right] \\ &= \underbrace{\mathrm{E}^{\mathbb{P}}(Z_2(T))\left[Z_1(T) - \mathrm{E}^{\mathbb{P}}(Z_1(T))\right]}_{=:R_1} + \underbrace{Z_1(T)[Z_2(T) - \mathrm{E}^{\mathbb{P}}(Z_2(T))]}_{=:R_2}. \end{aligned}$$

In contrast, switching the order of Z_1 and Z_2, i.e. considering $\tilde{Z} = (\tilde{Z}_1, \tilde{Z}_2) := (Z_2, Z_1)$, the variance decomposition approach yields

$$L_0 - \mathrm{E}^{\mathbb{P}}(L_0) = \underbrace{\mathrm{E}^{\mathbb{P}}(Z_1(T))\left[Z_2(T) - \mathrm{E}^{\mathbb{P}}(Z_2(T))\right]}_{=:\tilde{R}_1} + \underbrace{Z_2(T)[Z_1(T) - \mathrm{E}^{\mathbb{P}}(Z_1(T))]}_{=:\tilde{R}_2}.$$

Clearly, in general $R_1 \neq \tilde{R}_2$ and $R_2 \neq \tilde{R}_1$. In particular, if $Z_1(T)$ and $Z_2(T)$ both have mean zero, the first decomposition will imply $R_1 = 0$ and $R_2 = Z_1(T)Z_2(T)$, whereas the second decomposition will yield $\tilde{R}_1 = 0$ and $\tilde{R}_2 = Z_1(T)Z_2(T)$, i.e. either no risk or the total risk will be attributed to Z_1 (vice versa for Z_2).

In addition, although Z_1 and Z_2 might be correlated, R_1 and R_2 will be uncorrelated. This means that correlated risk must be allocated in an independent way, which can further result in arbitrary, order-dependent decompositions. Just consider $L_0 = Z_1(T) + Z_2(T)$ with dependent risks Z_1 and Z_2.

Hoeffding decomposition

Another decomposition approach is based on the so-called *Hoeffding decomposition* and is, for example, used by Rosen and Saunders (2010) to determine the factor contributions to a portfolio's credit risk. For convenience, we again call this approach Hoeffding decomposition.

Similarly to the previous approach, it relies on conditional expectations. For the insurer's liability L_0 the Hoeffding decomposition reads:

$$\begin{aligned} L_0 - \mathrm{E}^{\mathbb{P}}(L_0) = &\underbrace{\mathrm{E}^{\mathbb{P}}(L_0|\, Z_1) - \mathrm{E}^{\mathbb{P}}(L_0)}_{=:R_1} + \underbrace{\mathrm{E}^{\mathbb{P}}(L_0|\, Z_2) - \mathrm{E}^{\mathbb{P}}(L_0)}_{=:R_2} \\ &+ \underbrace{L_0 - \mathrm{E}^{\mathbb{P}}(L_0|\, Z_1) - \mathrm{E}^{\mathbb{P}}(L_0|\, Z_2) + \mathrm{E}^{\mathbb{P}}(L_0)}_{=:R_{1,2}}, \end{aligned}$$

where R_1 and R_2 are the risk factors attributed to Z_1 and Z_2 in isolation, and $R_{1,2}$ represents the risk due to "joint effects". This illustrates the decomposition's primary drawback, namely that the total risk is not completely allocated to the individual sources of risk. For instance, if in Example 2.1 $Z_1(T)$ and $Z_2(T)$ both have mean zero, the Hoeffding approach will yield $R_1 = R_2 = 0$ and $R_{1,2} = L_0 - \mathrm{E}^{\mathbb{P}}(L_0)$, i.e. the total risk results from joint effects, which does not give any insights on the influence of the different sources of risk. In particular, this example shows that the aggregation property P6 is generally not satisfied since for every function $A_{(L,Z)} : \mathbb{R}^2 \to \mathbb{R}$ we have $A_{(L,Z)}(R_1, R_2) = A_{(L,Z)}(0, 0) \neq L$ whenever $L_0 = Z_1(T)Z_2(T)$ is not deterministic.

However, properties P1 to P5 are satisfied: Clearly, the risk factors R_1 and R_2 are random variables (P1), and the Hoeffding decomposition is unique in the almost sure sense as a result of the uniqueness of the conditional expectations (P3). Furthermore, it can be easily seen that, contrary to the variance decomposition, this approach is order invariant (P4). The scale invariance follows by the same argument as for the variance decomposition (P5). For the attribution property P2, let Z_1 and Z_2 be two independent processes; if L is $\sigma(Z_1)$-measurable, then L and thus also L_0 are independent of Z_2, so that $R_2 = \mathrm{E}^{\mathbb{P}}\left(L_0 \middle| Z_2\right) - \mathrm{E}^{\mathbb{P}}\left(L_0\right) = 0$ (analogously for $\sigma(Z_2)$-measurable L).

Since we only consider the individual risk factors of a decomposition approach and thus ignore the joint term $R_{1,2}$ of the Hoeffding decomposition, the preceding discussion effectively covers the so-called Hájek projection $L_0 - \mathrm{E}^{\mathbb{P}}\left(L_0\right) \approx R_1 + R_2$ as sum of first-order terms of the Hoeffding decomposition (Rosen and Saunders, 2010, p. 341).

Taylor expansion

Christiansen (2007, p. 80) proposes to approximate functionals of random variables by their first order Taylor expansion and to interpret the resulting summands as risk factors. Assume the insurer's loss L_0 is of the form $F(Z_1(T), Z_2(T))$. Then this approach yields:

$$L = L_0 - \mathrm{E}^{\mathbb{P}}(L_0) \approx \left[F(z_1, z_2) - \mathrm{E}^{\mathbb{P}}(L_0)\right] + \underbrace{\frac{\partial F}{\partial z_1}(z_1, z_2)(Z_1(T) - z_1)}_{=:R_1} + \underbrace{\frac{\partial F}{\partial z_2}(z_1, z_2)(Z_2(T) - z_2)}_{=:R_2},$$

where (z_1, z_2) denotes the (deterministic) expansion point. By using a generalized definition of the corresponding gradients, Christiansen (2007) extends this approach to an infinite-dimensional setting such that the loss L_0 may also depend on the entire path of the stochastic processes Z_1 and Z_2.

In view of its properties, the method's applicability is restricted since the derivatives do not necessarily exist. Also, in case of non-linear functionals the first-order Taylor expansion and its summands only approximate the risk L. Moreover, the approximation error at a certain point highly depends on the choice of the expansion point, i.e. the Taylor expansion is "local". The following example illustrates those aspects.

Example 2.2. Assume that $L_0 = Z_1(T)Z_2(T)$. Then the Taylor expansion with expansion point (z_1, z_2) yields

$$\begin{aligned} L = L_0 - \mathrm{E}^{\mathbb{P}}(L_0) &\approx \left[z_1 z_2 - \mathrm{E}^{\mathbb{P}}(L_0)\right] + \underbrace{z_2(Z_1(T) - z_1)}_{=:R_1} + \underbrace{z_1(Z_2(T) - z_2)}_{=:R_2} \\ &= L_0 - \mathrm{E}^{\mathbb{P}}(L_0) - (Z_1(T) - z_1)(Z_2(T) - z_2). \end{aligned}$$

Obviously, the approximation error amounts to $-(Z_1(T) - z_1)(Z_2(T) - z_2)$, i.e. the more $Z_1(T)$ and $Z_2(T)$ deviate from z_1 and z_2, respectively, the higher is the approximation error. In the special case of choosing $(z_1, z_2) = (0, 0)$ as expansion point, the decomposition results in $R_1 = R_2 = 0$, i.e. a risk is neither allocated to Z_1 nor to Z_2.

As a result of the example with $(z_1, z_2) = (0, 0)$ as expansion point, the aggregation property P6 generally is not satisfied since for every function $A_{(L,Z)} : \mathbb{R}^2 \to \mathbb{R}$ we have $A_{(L,Z)}(R_1, R_2) = A_{(L,Z)}(0, 0) \neq L$ (assuming that $Z_1(T)Z_2(T)$ is not deterministic). Furthermore, due to the dependence on the expansion point, the Taylor expansion approach is also not unique (P3). To show that scale invariance (P5) is violated, assume that $L_0 = e^{Z_1(T)}$. Then the Taylor expansion yields $L \approx e^{z_1} - \mathrm{E}^{\mathbb{P}}\left(e^{Z_1(T)}\right) + e^{z_1}(Z_1(T) - z_1)$ for some expansion point z_1. However, for $\tilde{Z}_1(T) := e^{Z_1(T)}$ and $\tilde{z}_1 := e^{z_1}$ we have $L \approx \tilde{z}_1 - \mathrm{E}^{\mathbb{P}}\big(\tilde{Z}_1(T)\big) + (\tilde{Z}_1(T) - \tilde{z}_1)$, and in general $R_1 = e^{z_1}(Z_1(T) - z_1) \neq e^{Z_1(T)} - e^{z_1} = \tilde{Z}_1(T) - \tilde{z}_1 = \tilde{R}_1$.

Still, the Taylor expansion satisfies properties P1, P2, and P4. The risk factors are obviously random variables, and the order invariance can be easily shown. For the attribution property, independence of the other source of risk will yield a zero derivative and thus a zero risk factor.

Solvency II approach

A different risk decomposition approach is "switching off" the randomness of all but one of the sources of risk, see e.g. Artinger (2010) or Gatzert and Wesker (2014). Since this method is in principle implied in the Solvency II framework for measuring the influence of different sources of risk (cf. CEIOPS, 2010), in what follows we refer to this decomposition method as the *Solvency II approach*.

To illustrate, assume the insurer's loss L_0 is of the form $F(Z_1, Z_2)$. Then the method suggests to model the risk factors corresponding to Z_1 and Z_2 via $F(Z_1, z_2) - \mathrm{E}^{\mathbb{P}}(L_0)$ and $F(z_1, Z_2) - \mathrm{E}^{\mathbb{P}}(L_0)$, respectively. In the context of Solvency II, z_1 and z_2 are typically chosen as best estimates of Z_1 and Z_2. However, in general there is no clear answer on how z_1 and z_2 should be chosen. In fact, the decomposition heavily depends on the choice of z_1 and z_2 and is thus not unique (P3). This is illustrated in the following example.

Example 2.3. Assume that $L_0 = F(Z_1(T), Z_2(T)) = Z_1(T)\max\{K - Z_2(T), 0\}$, where Z_1 and Z_2 are two arbitrary positive-valued processes and K is a constant, and that $\mathrm{E}^{\mathbb{P}}(Z_2(T)) > K$. Measuring the risk factor related to Z_1 by replacing $Z_2(T)$ with its expectation, the Solvency II approach yields

$$\begin{aligned} R_1 &= F(Z_1(T), \mathrm{E}^{\mathbb{P}}(Z_2(T))) - \mathrm{E}^{\mathbb{P}}(L_0) = Z_1(T)\max\{K - \mathrm{E}^{\mathbb{P}}(Z_2(T)), 0\} - \mathrm{E}^{\mathbb{P}}(L_0) \\ &= -\mathrm{E}^{\mathbb{P}}(L_0). \end{aligned}$$

Thus, although $L_0 > 0$ with positive probability (under certain conditions on Z_1 and Z_2) and L_0 is proportionally increasing in $Z_1(T)$ in that case, the risk attributed to Z_1 is constant and even negative. However, choosing any deterministic approximation $z_2(T) < K$ would yield $R_1 = Z_1(T)(K - z_2(T)) - \mathrm{E}^{\mathbb{P}}(L_0)$ with a different distribution for each choice of $z_2(T)$. Assuming $z_1(T) = \mathrm{E}^{\mathbb{P}}(Z_1(T))$, the second risk factor equals $R_2 = \mathrm{E}^{\mathbb{P}}(Z_1(T))\max\{K - Z_2(T), 0\} - \mathrm{E}^{\mathbb{P}}(L_0)$.

Besides uniqueness, the Solvency II approach also does not satisfy the aggregation property P6 (and thus also not P6*), as can be seen from Example 2.3: for any function $A_{(L,Z)} : \mathbb{R}^2 \to \mathbb{R}$, $A_{(L,Z)}(R_1, R_2)$ will be $\sigma(Z_2)$-measurable and thus $A_{(L,Z)}(R_1, R_2) \neq L$ (assuming that Z_1 is not $\sigma(Z_2)$-measurable). Furthermore, the attribution property is generally not satisfied (P2). To see this, consider two sources of risk Z_1 and Z_2 and assume that $L_0 = e^{Z_1(T)} = F(Z_1(T), Z_2(T))$. Then, L is $\sigma(Z_1)$-measurable, but for every $z_1(T) \neq \log(\mathrm{E}^{\mathbb{P}}(L_0))$ it holds that $R_2 = F(z_1, Z_2) = e^{z_1(T)} - \mathrm{E}^{\mathbb{P}}(L_0) \neq 0$.

In contrast, the Solvency II approach satisfies properties P1, P4, and P5. Again, the risk factors obviously are random variables (P1), and order invariance can be easily shown (P4). For scale invariance, let f_1 and f_2 be two smooth, invertible functions and define $\tilde{Z}_i(t) := f_i(Z_i(t))$ and $\tilde{z}_i(t) := f_i(z_i(t))$, $i = 1, 2$. It follows that $L_0 = F(Z_1, Z_2) = F((f_1^{-1}(\tilde{Z}_1(t)))_{0 \leq t \leq T^*}, (f_2^{-1}(\tilde{Z}_2(t)))_{0 \leq t \leq T^*}) =: \tilde{F}(\tilde{Z}_1, \tilde{Z}_2)$. Hence, $\tilde{R}_1 = \tilde{F}(\tilde{Z}_1, \tilde{z}_2) - \mathrm{E}^{\mathbb{P}}(L_0) = F(Z_1, z_2) - \mathrm{E}^{\mathbb{P}}(L_0) = R_1$ and analogously $\tilde{R}_2 = R_2$, which proves the scale invariance property P5 (at least if the change of scale is the same for $\tilde{z}_i$ as for $\tilde{Z}_i$, $i = 1, 2$).

Table 2.1 summarizes the results. In the next section, we introduce a novel decomposition approach labeled *MRT decomposition* that satisfies all *meaningful risk decomposition* properties. It is also added to the table.

	P1	P2	P3	P4	P5	P6	P6*
Variance decomposition	✓	✓	✓	×	✓	✓	✓
Hoeffding decomposition	✓	✓	✓	✓	✓	×	×
Taylor expansion	✓	✓	×	✓	×	×	×
Solvency II approach	✓	×	×	✓	✓	×	×
MRT decomposition	✓	✓	✓	✓	✓	✓	✓

Table 2.1: Summary of decomposition approaches with regards to whether (✓) or not (×) they satisfy the properties P1 to P6*.

3 MRT decomposition in life insurance

Due to the relevance of risk decompositions in life insurance and to keep the exposition comprehensible, we frame our approach in a life insurance seeting. Generalizations to other situations in insurance and beyond are certainly possible. The first subsection lays out the framework for the remainder of the paper, and Section 3.2 introduces the MRT decomposition.

3.1 Life insurance modeling framework

We fix a finite time horizon T^* and a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T^*}$ satisfying the usual conditions of right-continuity and $\mathbb{P}$-completeness.[3] Throughout, $\mathcal{F}_t$ describes the "full" information available at time t, where we assume $\mathcal{F}_0$ to be trivial and set $\mathcal{F} = \mathcal{F}_{T^*}$. We assume that the uncertainty of the life insurer's future profit/loss arises from the uncertain evolution of a number of financial and demographic factors as well as the actual occurrence of deaths in the insurance portfolio. For the former, we introduce an n-dimensional, locally bounded process $X = ((X_1(t), \dots, X_n(t))^\top)_{0\leq t\leq T^*}$, the so-called *state process*, and assume that all financial and demographic factors are functions of X. Specifically, we assume that the time-t prices of all risky assets on the financial market as well as the short rate $r(t) = r(t, X(t))$ and the mortality intensity $\mu(t) = \mu(t, X(t))$ (see below) can be expressed in terms of $X(t)$.[4] The state process itself is driven by a d-dimensional standard Brownian motion $W = ((W_1(t), \dots, W_d(t))^\top)_{0\leq t\leq T^*}$:[5]

Assumption 3.1. *The state process $X = ((X_1(t), \dots, X_n(t))^\top)_{0\leq t\leq T^*}$ is an n-dimensional Itô process satisfying*

$$dX(t) = \theta(t)dt + \sigma(t)dW(t), \tag{3.1}$$

with deterministic initial value $X(0) = x_0 \in \mathbb{R}^n$, where the n-dimensional drift vector $\theta = (\theta(t))_{0\leq t\leq T^*}$ and the $n \times d$-dimensional volatility matrix $\sigma = (\sigma(t))_{0\leq t\leq T^*}$ are $\mathbb{G}$-adapted with continuous paths. We assume that there exists a unique strong solution to (3.1).

Let $\mathbb{G}$ denote the augmented filtration generated by W, which is assumed to be a sub-filtration of $\mathbb{F}$. Furthermore, we impose the existence of a bank account $(B(t))_{0\leq t\leq T^*}$ defined as $B(t) = e^{\int_0^t r(s)ds}$.

For notional convenience and without much loss of generality, we consider m homogeneous policyholders aged x at time 0. The remaining lifetime τ_x^i of policyholder i as seen from time 0, $i = 1, \dots, m$,

[3] In principle, $\mathbb{P}$ could be any probability measure and the technical results in this paper are not fixed to a particular interpretation. However, since our focus is on a company's risk, we interpret $\mathbb{P}$ as the real-word measure and the interpretation is reflected in our language.

[4] Within particular models, the prices of risky assets or the short rate may be components of the state process X.

[5] This assumption is primarily for simplicity. We discuss possible extensions in the Conclusion (Section 6).

is defined as the first jump time of a doubly stochastic process with $\mathbb{G}$-predictable intensity $(\mu(t))_{0\le t\le T^*}$, i.e.

$$\tau_x^i = \inf\left\{t \in [0,T^*] : \int_0^t \mu(s)ds \ge E_i\right\}, \ \ i = 1,\dots,m,$$

where $E_i,\ i = 1,\dots,m$, are i.i.d. unit exponential random variables independent of $\mathcal{G}_{T^*}$. We use the convention $\inf\emptyset = \infty$. A motivation for this definition of the remaining lifetimes can be found in Biffis et al. (2010, p. 287). Defining the sub-filtration $\mathbb{I} = \bigvee_{i=1}^m \mathbb{I}^i$ of $\mathbb{F}$, where $\mathbb{I}^i = (\mathcal{I}_t^i)_{0\le t\le T^*}$ is the augmented filtration generated by the death indicator process $(\mathbb{1}_{\{\tau_x^i \le t\}})_{0\le t\le T^*}$, it is natural to assume that $\mathbb{F}$ is given by $\mathbb{G}\vee\mathbb{I}$.

Assuming that μ is non-negative and continuous, it follows that for any $t \in [0,T^*]$ (Lando, 1998, p. 102)

$$\mathbb{P}\left(\tau_x^i > T|\mathcal{G}_T\right) = e^{-\int_0^T \mu(s)ds},$$

in particular $\mathbb{P}\left(\tau_x^i > 0\right) = 1$, and that (Bielecki and Rutkowski, 2004, p. 268)

$$\mathbb{P}\left(\tau_x^i > t\middle|\,\mathcal{G}_T\right) = \mathbb{P}\left(\tau_x^i > t\middle|\,\mathcal{G}_s\right) \tag{3.2}$$

for all $0 \le t \le s \le T \le T^*,\ i = 1,\dots,m$.[6] Generalizations of the setting that preserve these results are possible (Jeanblanc and Rutkowski, 2000; Biffis et al., 2010). We write $\Gamma(t) := \int_0^t \mu(s)ds$ for the so-called cumulative mortality intensity, so that $\mathbb{P}\left(\tau_x^i > t\middle|\,\mathcal{G}_t\right) = e^{-\Gamma(t)}$. Moreover, we obtain for the conditional survival probability given $\mathcal{F}_t$ (Bielecki and Rutkowski, 2004, p. 145):

$$\mathbb{P}\left(\tau_x^i > T|\mathcal{F}_t\right) = \mathbb{P}\left(\tau_x^i > T|\mathcal{G}_t\vee\mathcal{I}_t^i\right) = \mathbb{1}_{\{\tau_x^i>t\}}\frac{\mathbb{P}\left(\tau_x^i > T|\mathcal{G}_t\right)}{\mathbb{P}\left(\tau_x^i > t|\mathcal{G}_t\right)} = \mathbb{1}_{\{\tau_x^i>t\}}\mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^T \mu(s)ds}\middle|\,\mathcal{G}_t\right).$$

Note that the residual lifetimes $\tau_x^i,\ i = 1,\dots,m$, of the homogeneous policyholders are by construction conditionally identically distributed and conditionally independent given the σ-algebra $\mathcal{G}_{T^*}$.

Each life insurance contract in the company's portfolio is assumed to entail the same cash flows, the only difference being the respective remaining lifetime. We denote by $N(t) = \sum_{i=1}^m \mathbb{1}_{\{\tau_x^i\le t\}}$ the number of policyholders that have died until time t and by $\mathbb{F}^{W,N} = \left(\mathcal{F}_t^{W,N}\right)_{0\le t\le T^*}$ the augmentation of the filtration generated by the processes W and N. Note that $\mathbb{F}^{W,N}$ is a sub-filtration of $\mathbb{F}$.

The sum of the (possibly discounted) future cash flows as from time 0, which represents the *insurer's total net liability at time* 0, is given by:

$$\begin{aligned} L_0 = &\ C_0 + \sum_{k=0}^{\ell}(m - N(t_k))C_{a,k} + \int_0^{T^*} C_a(s)ds \\ &+ \sum_{k=1}^{\ell}(N(t_k) - N(t_{k-1}))C_{ad,k} + \int_0^{T^*} C_{ad}(s)dN(s), \end{aligned} \tag{3.3}$$

where $0 = t_0 < t_1 < \ldots < t_\ell = T^*, \ell \in \mathbb{N}$, are discrete points in time. To keep the setting general, we allow for payments independent of the lifetimes (C_0) and discrete ($C_{a,k}; C_{ad,k}$) as well as continuous ($C_a(t); C_{ad}(t)$) insurance benefits. More precisely, the different components in (3.3) denote:

[6] According to Jeanblanc and Rutkowski (2000), the latter is equivalent to the so-called $\mathcal{H}$-hypothesis, which says that every $\mathbb{G}$-martingale remains a martingale with respect to the larger filtration $\mathbb{F}$.

C_0 the sum of all (possibly discounted) payments at or after time 0 that are independent of $\tau_x^i,\ i = 1, \ldots, m,\ t \in [0, T^*]$, such as hedging returns, benefits from a fixed-term insurance, etc.;

$C_{a,k}$ the sum of all (possibly discounted) payments at or after time t_k that are conditional on survival until time t_k, $k = 0, \ldots, l$, such as single premiums, discrete premium payments, discrete annuity payments, benefits from pure endowment insurances, benefits from a period certain of deferred annuities, etc.;

$C_a(t)$ time-t intensity of all continuous payments that are conditional on survival until time t or, in other words, $C_a(t)dt$ is the sum of all payments in the infinitesimal period $[t, t + dt]$ that are conditional on survival until time t, such as continuous premium payments, continuous annuity payments, etc.;

$C_{ad,k}$ the sum of all (possibly discounted) payments at or after time t_k that are conditional on death within $(t_{k-1}, t_k]$, $k = 1, \ldots, l$, such as death benefits paid at the end of a period;

$C_{ad}(t)$ the sum of all (possibly discounted) payments at time t that are conditional on death at time t, $t \in [0, T^*]$, such as death benefits paid immediately upon death.

We assume that $C_0, C_{a,k}$, and $C_{ad,k}$ are $\mathcal{G}_{T^*}$-measurable, i.e. the cash flows may only be known at time T^*, whereas $C_a(t)$ and $C_{ad}(t)$ are assumed to be $\mathcal{G}_t$-measurable. This assumption is one of the reasons why we explicitly distinguish between discrete and continuous cash flows. Note that due to its form, the net liability L_0 is $\mathcal{F}_{T^*}^{W,N}$-measurable. In general, each cash flow from above may include several payments from and to the insurance company. Positive payments are interpreted as payments made by the insurer and negative payments are interpreted as payments received by the insurance company. Thus, each cash flow corresponds to the insurer's net liability at a certain point, justifying the interpretation of L_0 as the insurer's total net liability.

The *insurer's risk* at time 0 is identified with $L_0 - \mathrm{E}^{\mathbb{P}}(L_0)$, i.e. the insurer's net liability as from time 0 less its expectation. The net liability L_0 is exactly the (stochastic) amount of money the insurance company needs at time 0 in order to be able to meet its future contract obligations (possibly conditional on certain investment assumptions introduced by the discount factors). Since the insurance company should at least prepare for the expected value $\mathrm{E}^{\mathbb{P}}(L_0)$, only liabilities exceeding the expectation are interpreted as risk.

Remark 3.2. In order to keep the presentation concise, we focus on the insurer's risk at time 0, but with appropriate modifications all related definitions and results can be transferred to the insurer's risk at any future time $t \in [0, T^*]$ (considered from time 0). For example, in analogy to L_0 and Equation (3.3), the insurer's net liability at time t, L_t, can be defined as the sum of all future cash flows as from time t, where possible discount factors of the cash flows need to be adjusted. The insurer's risk at time t then follows as $L_t - \mathrm{E}^{\mathbb{P}}(L_t | \mathcal{F}_t)$.

3.2 Definition of the MRT decomposition

Within the life insurance modeling framework introduced in the previous section, the objective is to find an approach that decomposes the insurer's risk $L_0 - \mathrm{E}^{\mathbb{P}}(L_0|\mathcal{F}_t)$ into risk factors attributed to the sources of risk the insurer faces in a *meaningful* way (cf. Section 2.1). Inspired by the martingale representation theorem, we propose a decomposition into stochastic integrals with respect to the compensated sources of risk and interpret each integral as the risk factor of the respective source of risk. We show in Section 4.1 that this approach satisfies all meaningful risk decomposition properties formulated in Section 2.1.

The sources of risk are identified with, on the one hand, the state processes $X_i = (X_i(t))_{0\leq t\leq T^*}$, $i = 1,\dots,n$, and, on the other hand, with the number of deaths in the portfolio $N = (N(t))_{0\leq t\leq T^*}$. The corresponding compensated processes, i.e. the processes less their $\mathbb{F}$-compensators, are denoted by $M_i^W = (M_i^W(t))_{0\leq t\leq T^*}$, $i = 1,\dots,n$, and $M^N = (M^N(t))_{0\leq t\leq T^*}$, respectively. We immediately obtain (a proof can be found in the Appendix):

Lemma 3.3. *i) The unique compensator of X_i is given by $A_i^W = (A_i^W(t))_{0\leq t\leq T^*}$, where $A_i^W(t) = \int_0^t \theta_i(s)ds$, $i = 1,\dots,n$. Thus,*

$$M_i^W(t) = \sum_{j=1}^d \int_0^t \sigma_{ij}(s)dW_j(s), \quad 0 \leq t \leq T^*,\ i = 1,\dots,n.$$

ii) The unique compensator of N is given by $A^N = (A^N(t))_{0\leq t\leq T^*}$, where $A^N(t) = \int_0^t (m - N(s-))\mu(s)ds$. Thus,

$$M^N(t) = N(t) - \int_0^t (m - N(s-))\mu(s)ds, \quad 0 \leq t \leq T^*,$$

where, for completeness, we define $N(0-) := 0$.

Accordingly, we look for a decomposition

$$L := L_0 - \mathrm{E}^{\mathbb{P}}(L_0) = \sum_{i=1}^n \int_0^{T^*} \psi_i^W(t)dM_i^W(t) + \int_0^{T^*} \psi^N(t)dM^N(t), \tag{3.4}$$

where $\psi_i^W(t)$, $i = 1\dots,n$, and $\psi^N(t)$ are $\mathbb{F}$-predictable processes. Each integral is interpreted as the portion of the total randomness of $L_0 - \mathrm{E}^{\mathbb{P}}(L_0)$ caused by the associated source of risk.[7] Thus, the risk factors are given by $R_i := \int_0^{T^*} \psi_i^W(t)dM_i^W(t)$, $i = 1,\dots,n$, and $R_{n+1} := \int_0^{T^*} \psi^N(t)dM^N(t)$, where the latter describes the randomness introduced by N, i.e. by the random occurrence of deaths in the portfolio, and thus corresponds to the inherent *unsystematic mortality risk*.[8] A decomposition of the form (3.4) is henceforth called *MRT decomposition*, since the idea and the decomposition's existence and uniqueness are implied by the martingale representation theorem, as shown by the following proposition.

Remark 3.4. To emphasize the points in Remark 3.2, assume that a decomposition of the form (3.4) can be analogously found for $L_t - \mathrm{E}^{\mathbb{P}}(L_t)$ as for $L_0 - \mathrm{E}^{\mathbb{P}}(L_0)$ (with possibly different integrands). Then, since all integrals in (3.4) are martingales, it follows for the insurer's risk at time t that (for simplicity, we use the same notation for the integrands as above)

$$L_t - \mathrm{E}^{\mathbb{P}}(L_t|\,\mathcal{F}_t) = \sum_{i=1}^n \int_t^{T^*} \psi_i^W(t)dM_i^W(t) + \int_t^{T^*} \psi^N(t)dM^N(t),$$

and the corresponding MRT risk factors can be defined as integrals starting from t. Thus, although we analyze the MRT decomposition only for the insurer's risk at time 0, with appropriate modifications all results generalize to the insurer's risk at any time t.

[7] Similar interpretations of stochastic integrals can be found e.g. in Christiansen (2013) for unsystematic risk and in Biagini et al. (2013) under a risk-neutral measure.

[8] Note that a decomposition consisting of stochastic integrals with respect to the different sources of risk X_i, $i = 1,\dots,n$, and N (instead of the compensated processes) does not necessarily exist, since the risk processes are not $\mathbb{P}$-martingales.

Proposition 3.5. *Assume $n = d$, $\det \sigma(t) \neq 0$ for all $t \in [0,T^*]$ $\mathbb{P}$-almost surely, and that L_0 is square integrable. Then there exist $\mathbb{F}^{W,N}$-predictable processes $\psi_1^W, \dots, \psi_n^W, \psi^N : [0,T^*] \times \Omega \to \mathbb{R}$ such that the MRT decomposition (3.4) holds. The representation is unique in the sense that the integrands $\psi_1^W, \dots, \psi_n^W$ and the integrand ψ^N are a.s. unique on $[0,T^*] \times \Omega$ and $\{(t,\omega) \in [0,T^*] \times \Omega : N(t-) < m\}$, respectively, both with respect to $\lambda \otimes \mathbb{P}$, where λ denotes the Lebesgue measure on $[0,T^*]$. Moreover,*

$$\mathrm{E}^{\mathbb{P}}\left(\left[\int_0^{T^*} \psi^N(t)dM^N(t)\right]^2\right) < \infty. \tag{3.5}$$

Proof. As indicated, L_0 and thus L is $\mathcal{F}_{T^*}^{W,N}$-measurable. Applying the martingale representation theorem for point processes combined with Brownian motions (Björk, 2011, Theorem 4.1.2) to the martingale

$$M(t) := \mathrm{E}^{\mathbb{P}}\left(L_0 - \mathrm{E}^{\mathbb{P}}(L_0) \middle| \mathcal{F}_t^{W,N}\right),\ 0 \leq t \leq T^*,$$

together with the $\mathcal{F}_{T^*}^{W,N}$-measurability of L_0, it follows that there exist $\mathbb{F}^{W,N}$-predictable processes $\widetilde{\psi}_1^W, \dots, \widetilde{\psi}_d^W, \psi^N : [0,T^*] \times \Omega \to \mathbb{R}$ such that

$$L_0 - \mathrm{E}^{\mathbb{P}}(L_0) = \int_0^{T^*} \widetilde{\psi}^W(t)dW(t) + \int_0^{T^*} \psi^N(t)dM^N(t), \tag{3.6}$$

where $\widetilde{\psi}^W := (\widetilde{\psi}_1^W, \dots, \widetilde{\psi}_d^W)$. Since $n = d$ and $\det \sigma(t) \neq 0$ by assumption, the inverse of σ exists (and is unique). Thus, if $\psi_i^W(t) := \sum_{j=1}^d \tilde{\psi}_j^W(t)\sigma_{ji}^{-1}(t),\ i = 1, \dots, n$, denotes the i-th entry of the vector $\widetilde{\psi}^W(t)\sigma^{-1}(t)$, the first summand of (3.6) can be transformed into

$$\int_0^{T^*} \widetilde{\psi}^W(t)dW(t) = \int_0^{T^*} \widetilde{\psi}^W(t)\sigma^{-1}(t)\sigma(t)dW(t) = \sum_{i=1}^n \int_0^{T^*} \psi_i^W(t)dM_i^W(t),$$

which together with (3.6) proves the existence of the MRT decomposition (3.4).

Since $\langle W_i, W_j\rangle(t) = 0$ for all $i \neq j$, and $\langle W_i, M^N\rangle(t) = 0$ for $i = 1, \dots, d$, the Itô isometry yields

$$\mathrm{E}^{\mathbb{P}}\left(\left(L_0 - \mathrm{E}^{\mathbb{P}}(L_0)\right)^2\right) = \sum_{i=1}^d \mathrm{E}^{\mathbb{P}}\left(\left(\int_0^{T^*} \tilde{\psi}_i^W(t)dW_i(t)\right)^2\right) + \mathrm{E}^{\mathbb{P}}\left(\left(\int_0^{T^*} \psi^N(t)dM^N(t)\right)^2\right).$$

Thus, by the square-integrability of L_0, all integrals in (3.6) are square integrable, and in particular (3.5) holds.

To show uniqueness, suppose there exist $\mathbb{F}^{W,N}$-predictable processes $\widetilde{\xi}_1^W, \dots, \widetilde{\xi}_d^W, \xi^N : [0,T^*] \times \Omega \to \mathbb{R}$ such that

$$L_0 - \mathrm{E}^{\mathbb{P}}(L_0) = \int_0^{T^*} \widetilde{\xi}^W(t)dW(t) + \int_0^{T^*} \xi^N(t)dM^N(t).$$

Then we have $\int_0^{T^*} \left(\widetilde{\psi}^W(t) - \widetilde{\xi}^W(t)\right)dW(t) + \int_0^{T^*} \left(\psi^N(t) - \xi^N(t)\right) dM^N(t) = 0$. From Andersen et al. (1997, p. 78), we know that the predictable quadratic variation of $M^N(t)$ equals $\langle M^N, M^N\rangle(t) = \int_0^t (m - N(s-))\mu(s)ds$. Together with the Itô isometry, we thus obtain

$$0 = \sum_{i=1}^d \mathrm{E}^{\mathbb{P}}\left(\int_0^{T^*} \left(\widetilde{\psi}_i^W(t) - \widetilde{\xi}_i^W(t)\right)^2 dt\right) + \mathrm{E}^{\mathbb{P}}\left(\int_0^{T^*} \left(\psi^N(t) - \xi^N(t)\right)^2 (m - N(t-))\mu(t)dt\right).$$

This directly implies that $\tilde{\psi}_i^W = \tilde{\xi}_i^W$ $\lambda\otimes\mathbb{P}$-almost surely, $i = 1, \ldots, d$. Since μ is assumed to be positive, it also follows that $\psi^N = \xi^N$ on $\{(t, \omega) \in [0, T^*] \times \Omega : N(t-) < m\}$ with respect to $\lambda \otimes \mathbb{P}$. Finally, the uniqueness of $\psi_1^W, \ldots, \psi_n^W, \psi^N$ is a result of the uniqueness of $\widetilde{\psi}_1^W, \ldots, \widetilde{\psi}_n^W, \psi^N$ and the uniqueness of the inverse of σ.

□

Remark 3.6. The previous proposition is based on the assumption that each insurance contract in the considered portfolio entails the same cash flows. For relaxing this assumption, it is sufficient to split the considered portfolio into sub-portfolios with identical cash flows and to apply the result from above to each sub-portfolio separately. Moreover, if the payments depend on the sequence of deaths as within joint life policies, it is possible to extend the setting and consider the processes in the general filtration $\mathbb{F}$ implying $d + m$ driving martingales. We focus on $\mathbb{F}^{W,N}$ here since it is the most relevant setup and to keep the presentation concise.

If $n \neq d$, existence and uniqueness of the MRT decomposition (3.4) are not necessarily given. In fact, as follows from the proof, we need to look for $\psi^W(t)$ such that the equation $\widetilde{\psi}^W(t) = \psi^W(t)\sigma(t)$ holds true, where existence and uniqueness of $\widetilde{\psi}^W(t)$ result from the martingale representation theorem. If $n > d$, there are fewer equations than unknowns so that uniqueness is not guaranteed. On the other hand, if $n < d$, there are more equations than unknowns so that a solution may not exist. In what follows, we focus on the case $n = d$. If $n \neq d$, we assume that either redundant state processes (which can be represented via other state processes) are removed or additional state processes are artificially added, both along with an adjustment of the interpretation of the risk factors. In contrast to a hedging problem, where the number of state processes – or rather securities – is exogenously given, this procedure seems viable for a risk decomposition problem.

4 Analysis of the MRT decomposition

This section provides a detailed analysis of the MRT decomposition (3.4) introduced in the previous section. We first establish its advantages by showing that it satisfies the *meaningful* risk decomposition properties from Section 2.1. We then discuss its calculation, where we rely on analogies to hedging problems for insurance liabilities, and finally analyze diversification properties in the last subsection.

4.1 Meaningful risk decomposition properties

Proposition 4.1. *Assume that the state process $X = (X_1, \ldots, X_n)$ is defined as in Assumption 3.1 with $n = d$ and $\det \sigma(t) \neq 0$ for all $t \in [0, T^*]$ $\mathbb{P}$-almost surely, and let L_0 be square integrable. Then the MRT decomposition*

$$(L, X_1, \ldots, X_n, N) \overset{MRT}{\leftrightarrow} (R_1, \ldots, R_{n+1})$$

as defined in (3.4) *satisfies the properties P1, P2, P3, P4, P5, P6, and P6*.*

Proof. Obviously, the risk factors $R_1, \ldots, R_{n+1}$ are random variables, and $L = \sum_{i=1}^{n+1} R_i$, so that P1 and P6* (and thus also P6) are satisfied. The uniqueness property P3 directly follows from Proposition 3.5 and the fact that $\Delta M^N(t) = 0$ on $\{(t, \omega) \in [0, T^*] \times \Omega : N(t-) = m\}$.

To simplify the proof of the remaining properties, we let $\psi_i := \psi_i^W,\ M_i := M_i^W,\ i = 1, \ldots, n$, and $\psi_{n+1} := \psi^N,\ M_{n+1} := M^N$. Furthermore, we write $Z = (Z_1, \ldots, Z_{n+1}) := (X_1, \ldots, X_n, N)$. Assume that $(L, Z_1, \ldots, Z_{n+1}) \overset{\text{MRT}}{\leftrightarrow} (R_1, \ldots, R_{n+1})$.

P2: Let $i \in \{1, \dots, n+1\}$. Assume that L is $\sigma(Z_i)$-measurable and that Z_i is independent of $Z_{i-} := (Z_1, \dots, Z_{i-1}, Z_{i+1}, \dots, Z_{n+1})$. This directly implies that L is independent of Z_{i-}. Furthermore, since $\det \sigma(t) \neq 0$ for all $t \in [0, T^*]$ $\mathbb{P}$-almost surely, we have $\mathcal{F}_t^{W,N} = \mathcal{F}_t^Z = \mathcal{F}_t^{Z_i} \vee \mathcal{F}_t^{Z_{i-}}$, where $\mathbb{F}^Z = (\mathcal{F}_t^Z)_{0 \le t \le T^*}$, $\mathbb{F}^{Z_i} = (\mathcal{F}_t^{Z_i})_{0 \le t \le T^*}$, and $\mathbb{F}^{Z_{i-}} = (\mathcal{F}_t^{Z_{i-}})_{0 \le t \le T^*}$ denote the augmented filtrations generated by $Z,\ Z_i$, and Z_{i-}, respectively. Thus,

$$L(t) := \sum_{j=1}^{n+1} \int_0^t \psi_j(s) dM_j(s) = \mathrm{E}^{\mathbb{P}}\left(L | \mathcal{F}_t^{W,N} \right) = \mathrm{E}^{\mathbb{P}}\left(L | \mathcal{F}_t^{Z_i} \vee \mathcal{F}_t^{Z_{i-}} \right) = \mathrm{E}^{\mathbb{P}}\left(L | \mathcal{F}_t^{Z_i} \right).$$

This implies that the process $(L(t))_{0 \le t \le T^*}$ is independent of each process $Z_j,\ j \neq i$, so that the predictable covariation process satisfies $\langle L, Z_j \rangle (t) = 0$ for all $j \neq i,\ 0 \le t \le T^*$.

(a) Assume that $i = n+1$. Then $\langle M_i, Z_j \rangle (t) = \langle M^N, X_j \rangle (t) = 0$ for all $j \neq i$, so that

$$\begin{aligned} 0 = d\langle L, Z_j \rangle (t) &= \sum_{k=1}^{n+1} \psi_k(t) d \langle M_k, Z_j \rangle (t) = \sum_{k=1}^{n} \psi_k(t) d \langle M_k, Z_j \rangle (t) \\ &= \sum_{k=1}^{n} \psi_k(t) \sigma_{k,\cdot}(t) \sigma_{j,\cdot}^\top(t) dt, \quad j \neq i,\ 0 \le t \le T^*, \end{aligned} \tag{4.1}$$

where $\sigma_{k,\cdot}(t)$ denotes the k-th row of $\sigma(t)$. For any $0 \le t \le T^*$, this yields the linear system of equations $A_t^\top \psi_t = 0$, where $\psi_t = (\psi_1(t), \dots, \psi_n(t))^\top$ and $A_t = \sigma(t)\sigma(t)^\top$, so that $\det A_t^\top = (\det \sigma(t))^2 \neq 0$ for all $t \in [0, T^*]$ $\mathbb{P}$-almost surely, implying $\psi_t = 0$ for all $t \in [0, T^*]$ $\mathbb{P}$-almost surely. Thus, we have $R_j = \int_0^{T^*} \psi_j(t) dM_j(t) = 0$ almost surely for all $j \neq i$.

(b) Now assume that $i \neq n+1$ (w.l.o.g. $i = 1$). Then we know that

$$\begin{aligned} 0 = d\langle L, Z_{n+1} \rangle (t) &= \sum_{k=1}^{n+1} \psi_k(t) d \langle M_k, Z_{n+1} \rangle (t) = \psi_{n+1}(t) d \langle M_{n+1}, Z_{n+1} \rangle (t) \\ &= \psi_{n+1}(t) d \langle M_{n+1}, M_{n+1} \rangle (t), \end{aligned}$$

so by the Itô isometry it follows that

$$\mathrm{E}^{\mathbb{P}}\left(\left[\int_0^{T^*} \psi_{n+1}(t) dM_{n+1}(t) \right]^2 \right) = \mathrm{E}^{\mathbb{P}}\left(\int_0^{T^*} \psi_{n+1}^2(t) d \langle M_{n+1}, M_{n+1} \rangle (t) \right) = 0,$$

and thus $R_{n+1} = \int_0^{T^*} \psi_{n+1}(t) dM_{n+1}(t) = 0$ almost surely.
Since Z_1 is by assumption independent of Z_{1-} and thus independent of Z_j for all $j = 2, \dots, n+1$, it follows that $\sigma_{1,\cdot}(t)\sigma_{j,\cdot}(t)^\top dt = d\langle Z_1, Z_j \rangle (t) = 0$ for all $j \notin \{1, n+1\}$. Thus, for $A_t = \sigma(t)\sigma(t)^\top$, we obtain

$$A_t = \begin{pmatrix} \sigma_{1,\cdot}(t)\sigma_{1,\cdot}(t)^\top & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \tilde{A}_t & \\ 0 & & & \end{pmatrix}, \ \tilde{A}_t = \begin{pmatrix} \sigma_{2,\cdot}(t)\sigma_{2,\cdot}(t)^\top & \dots & \sigma_{2,\cdot}(t)\sigma_{n,\cdot}(t)^\top \\ \vdots & & \vdots \\ \sigma_{n,\cdot}(t)\sigma_{2,\cdot}(t)^\top & \dots & \sigma_{n,\cdot}(t)\sigma_{n,\cdot}(t)^\top \end{pmatrix},$$

and since $0 \neq \det A_t = \sigma_{1,\cdot}(t)\sigma_{1,\cdot}(t)^\top \det \tilde{A}_t$, $\det \tilde{A}_t \neq 0$ for all $t \in [0, T^*]$ $\mathbb{P}$-almost surely. Furthermore, following the same calculation steps as in (4.1) for $j \notin \{1, n+1\}$ and using $\langle M_1, Z_j \rangle (t) = d \langle Z_1, Z_j \rangle (t) = 0$ and $\langle M_{n+1}, Z_j \rangle (t) = 0,\ j \notin \{1, n+1\}$, we obtain the linear system $\tilde{A}_t^\top \tilde{\psi}_t = 0$, where $\tilde{\psi}_t = (\psi_2(t), \ldots, \psi_n(t))^\top$. Since $\det \tilde{A}_t \neq 0$ for all $t \in [0, T^*]$ $\mathbb{P}$-almost surely, it follows that $\tilde{\psi}_t = 0$ for all $t \in [0, T^*]$ $\mathbb{P}$-almost surely, and thus $R_j = \int_0^{T^*} \psi_j(t) dM_j(t) = 0$ almost surely for all $j \notin \{1, n+1\}$.

P4: Consider a permutation $\pi : \{1, \ldots, n+1\} \to \{1, \ldots, n+1\}$. Let $(L, Z_{\pi(1)}, \ldots, Z_{\pi(n+1)}) \overset{\text{MRT}}{\leftrightarrow} (\tilde{R}_1, \ldots, \tilde{R}_{n+1})$ with $\tilde{R}_i = \int_0^{T^*} \tilde{\psi}_i(t) dM_{\pi(i)}(t),\ i = 1, \ldots, n+1$, where $\tilde{\psi}_i$ are $\mathbb{F}$-predictable processes. Since

$$
\begin{aligned}
&\sum_{i=1}^{n+1} \int_0^{T^*} \tilde{\psi}_i(t) dM_{\pi(i)}(t) = \sum_{i=1}^{n+1} \tilde{R}_i \overset{\text{P6}^*}{=} L \overset{\text{P6}^*}{=} \sum_{i=1}^{n+1} R_i = \sum_{i=1}^{n+1} \int_0^{T^*} \psi_i(t) dM_i(t) \\
&= \sum_{i=1}^{n+1} \int_0^{T^*} \psi_{\pi(i)}(t) dM_{\pi(i)}(t),
\end{aligned}
$$

P4 follows by the uniqueness of the MRT decomposition.

P5: Let $\tilde{Z}_i(t) := f_i(Z_i(t)),\ i = 1, \ldots, n+1$, where the functions $f_i : \mathbb{R} \to \mathbb{R}$ are smooth and invertible, and consider $(L, \tilde{Z}_1, \ldots, \tilde{Z}_{n+1}) \overset{\text{MRT}}{\leftrightarrow} (\tilde{R}_1, \ldots, \tilde{R}_{n+1})$.

For each $i \neq n+1$, by Itô's lemma:

$$
d\tilde{Z}_i(t) = f_i'(X_i(t)) \sum_{j=1}^{d} \sigma_{ij}(t) dW_j(t) + \left(f_i'(X_i(t))\theta(t) + \frac{1}{2} f_i''(X_i(t)) \sum_{j=1}^{d} \sigma_{ij}^2(t) \right) dt.
$$

Thus, $(\tilde{Z}_1, \ldots, \tilde{Z}_n)$ is again an Itô process as in Assumption 3.1 and by Lemma 3.3 the corresponding compensated risk processes equal

$$
\tilde{M}_i(t) = f_i'(X_i(t)) dM_i(t),\ i = 1, \ldots, n.
$$

As a result, for $i = 1, \ldots, n$, the MRT risk factors equal

$$
\tilde{R}_i = \int_0^{T^*} \tilde{\psi}_i(t) d\tilde{M}_i(t) = \int_0^{T^*} \tilde{\psi}_i(t) f_i'(X_i(t)) dM_i(t). \tag{4.2}
$$

For $i = n+1$, we have

$$
\begin{aligned}
\tilde{Z}_{n+1}(t) &= f_{n+1}(N(0)) + \sum_{0<s\le t} \left(f_{n+1}(N(s)) - f_{n+1}(N(s-)) \right) \\
&= f_{n+1}(N(0)) + \sum_{0<s\le t} \underbrace{\left[\sum_{k=0}^{m} \mathbb{1}_{\{N(s-)=k\}} \left(f_{n+1}(k+1) - f_{n+1}(k) \right) \right]}_{=: a(s)} (N(s) - N(s-)) \\
&= f_{n+1}(N(0)) + \int_0^t a(s) dN(s) \\
&= f_{n+1}(N(0)) + \int_0^t a(s) dM_{n+1}(s) + \int_0^t a(s)(m - N(s-))\mu(s) ds,
\end{aligned}
$$

exploiting in the second equality that $\mathbb{P}(\tau_x^i = \tau_x^j) = 0$ for $i \neq j$ (Bielecki and Rutkowski, 2004, p. 269). Since $a(s) \neq 0$ (invertible) and predictable, $\tilde{A}_{n+1}(t) := \int_0^t a(s)(m - N(s-))\mu(s)ds$ is a predictable finite variation process and $\tilde{M}_{n+1}(t) := \int_0^t a(s)dM_{n+1}(s)$ is a local martingale. Thus:

$$\tilde{R}_{n+1} = \int_0^{T^*} \tilde{\psi}_{n+1}(t)d\tilde{M}_{n+1}(t) = \int_0^{T^*} \tilde{\psi}_{n+1}(t)a(t)dM_{n+1}(t). \tag{4.3}$$

The uniqueness of the MRT decomposition together with (4.2) and (4.3) implies that $R_i = \tilde{R}_i$ almost surely, $i = 1, \ldots, n+1$, and thus P5.

□

Remark 4.2. While the MRT decomposition in (3.4) is formally defined in terms of the Itô process X and the *counting process* N, in the proof of P5 we consider a generalized notion in terms of an Itô process and the *jump process* $\int_0^{\cdot} a(s)dN(s)$. However, since the generalization is straightforward and to keep the presentation in Section 3 concise, we accept this slight inconsistency.

Remark 4.3. While the notion of uniqueness of the MRT decomposition (P3) is based on the description of X in Assumption 3.1, it is important to note that it will not depend on the representation of X. In particular, we will clearly obtain the same MRT decomposition if we choose an equivalent representation of X (e.g., in terms of correlated Brownian motions) since the compensated risk processes $M_i^W,\ i = 1, \ldots, n,$ coincide for each representation. The restriction to Itô processes and to the life insurance setting, on the other hand, are limitations and we leave an analysis of the *meaningfulness* of the MRT decomposition in more general contexts for future research.

4.2 Calculation of the MRT decomposition

General case

As is evident from the proof of Proposition 3.5, the calculation of the MRT decomposition amounts to the determination of $\widetilde{\psi}_1^W, \ldots, \widetilde{\psi}_n^W, \psi^N$ in the martingale representation (3.6). A relatively general result for their calculation is the Clark-Ocone formula from Malliavin calculus (Di Nunno et al., 2009; Nualart, 2006). However, the Clark-Ocone formula is only applicable to independent driving processes, and for stochastic mortality intensities within our setting such an independence is usually not given between the number of deaths in the portfolio N and the standard Brownian motion W driving, among others, the mortality intensity. Thus, the following three lemmas reduce the problem to finding the martingale representation of a $\mathbb{G}$-martingale instead of an $\mathbb{F}$-martingale, i.e. the problem is reduced to a Brownian motion setting.

Remark 4.4. A very similar problem arises in the context of (quadratic) hedging of insurance liabilities, and a number of papers have taken a similar approach (Barbarin, 2008; Biagini et al., 2012, 2013; Biagini and Schreiber, 2013; Møller, 2001; Dahl and Møller, 2006; Dahl et al., 2008; Norberg, 2013). We heavily rely on this line of research but present several extensions and new adaptations.

The first lemma covers discrete survival cash flows.

Lemma 4.5. *Let Z be a random variable of the form $Z = (m - N(T))\,F,\ 0 \leq T \leq T^*$, where F is $\mathcal{G}_{T^*}$-measurable and integrable. Then there exist $\mathbb{G}$-predictable processes $\varphi_1, \ldots, \varphi_d$ such that*

$$\mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(T)}F\middle|\,\mathcal{G}_t\right) = \mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(T)}F\right) + \sum_{i=1}^{d}\int_0^t \varphi_i(u)dW_i(u),\ \ t \leq T^*, \tag{4.4}$$

and the martingale representation of Z is given by

$$Z = \mathrm{E}^{\mathbb{P}}(Z) + \sum_{i=1}^{d} \int_0^{T^*} \left[(m - N(t-))\, e^{\Gamma(t)} \mathbb{1}_{[0,T]}(t) + (m - N(T))\, e^{\Gamma(T)} \mathbb{1}_{(T,T^*]}(t)\right] \varphi_i(t) dW_i(t)$$
$$- \int_0^T \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(T)} F \middle| \mathcal{G}_t\right) dM^N(t). \tag{4.5}$$

For a single policyholder, the proof of Lemma 4.5 given in the Appendix mainly follows the ideas of the proof of Proposition 5.2.2 in Bielecki and Rutkowski (2004, pp. 159/160), albeit we modify their result so that it fits our later application and extend it to an entire (homogeneous) portfolio. For F $\mathcal{G}_T$-measurable instead of (more generally) $\mathcal{G}_{T^*}$-measurable, similar results (usually in a specific process setting) have been derived in the context of risk-minimizing hedging strategies, see e.g. Barbarin (2008, Prop. 4.10, Prop. 5.11), Biagini et al. (2012, Prop. 3.5), Biagini et al. (2013, Prop. 2, Prop. 9), and Biagini and Schreiber (2013, Lemma 4.2). In particular, most of them also consider entire portfolios.

The next lemma covers the case of continuous survival cash flows.

Lemma 4.6. *Let Z be a random variable of the form $Z = \int_0^T (m - N(v))F(v)dv,\ 0 \le T \le T^*$, where $F = (F(t))_{0\le t\le T}$ is a $\mathbb{G}$-predictable process with $\mathrm{E}^{\mathbb{P}}\left(\sup_{t\in[0,T]} |F(t)|\right) < \infty$. Then there exist $\mathbb{G}$-predictable processes $\varphi_1, \ldots, \varphi_d$ such that*

$$\mathrm{E}^{\mathbb{P}}\left(\int_0^T e^{-\Gamma(v)} F(v) dv \middle| \mathcal{G}_t\right) = \mathrm{E}^{\mathbb{P}}\left(\int_0^T e^{-\Gamma(v)} F(v) dv\right) + \sum_{i=1}^{d} \int_0^t \varphi_i(u) dW_i(u), \ \ t \le T, \tag{4.6}$$

and the martingale representation of Z is given by

$$\begin{aligned} Z = \mathrm{E}^{\mathbb{P}}(Z) &+ \sum_{i=1}^{d} \int_0^T (m - N(t-)) e^{\Gamma(t)} \varphi_i(t) dW_i(t) \\ &- \int_0^T \int_t^T \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(v)} F(v) \middle| \mathcal{G}_t\right) dv\, dM^N(t). \end{aligned} \tag{4.7}$$

In particular, if additionally $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left([F(t)]^2\right) < \infty$, then

$$\varphi_i(t) = \int_t^T \varphi_i^v(t) dv, \ \ t \le T, \tag{4.8}$$

where $\varphi_i^v,\ i = 1, \ldots, d,\ v \in [0, T]$, are the $\mathbb{G}$-predictable integrands of the martingale representation of $e^{-\Gamma(v)} F(v)$ (cf. Eq. (4.4)*).*

A proof of the lemma is given in the Appendix. Except for some details, the proof of the first part (4.7) mainly follows the proof of Proposition 4.12 in Barbarin (2008). The specification (4.8) may simplify the derivation of (4.6). For bounded F, it has already been shown in Biagini et al. (2013, Proposition 5).

The next lemma covers continuous cash flows contingent on death.

Lemma 4.7. *Let Z be a random variable of the form $Z = \int_0^T F(v) dN(v),\ 0 \le T \le T^*$, where $F = (F(t))_{0\le t\le T}$ is a continuous and $\mathbb{G}$-predictable process with $\mathrm{E}^{\mathbb{P}}\left(\sup_{t\in[0,T]} |F(t)|\right) < \infty$. Then there exist $\mathbb{G}$-predictable processes $\varphi_1, \ldots, \varphi_d$ such that for $t \le T$:*

$$\mathrm{E}^{\mathbb{P}}\left(\int_0^T e^{-\Gamma(v)} F(v) d\Gamma(v) \middle| \mathcal{G}_t\right) = \mathrm{E}^{\mathbb{P}}\left(\int_0^T e^{-\Gamma(v)} F(v) d\Gamma(v)\right) + \sum_{i=1}^{d} \int_0^t \varphi_i(u) dW_i(u), \tag{4.9}$$

and the martingale representation of Z is given by

$$\begin{aligned} Z = \mathrm{E}^{\mathbb{P}}(Z) + \sum_{i=1}^{d} \int_0^T (m - N(t-))\, e^{\Gamma(t)} \varphi_i(t) dW_i(t) \\ - \int_0^T \left[\int_t^T \mathrm{E}^{\mathbb{P}} \left(e^{\Gamma(t)-\Gamma(v)} F(v)\mu(v) \middle| \mathcal{G}_t \right) dv - F(t) \right] dM^N(t). \end{aligned} \tag{4.10}$$

In particular, if additionally $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}} \left([F(t)]^4 \right) < \infty$ *and* $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}} \left(\mu^4(t) \right) < \infty$, *then*

$$\varphi_i(t) = \int_t^T \varphi_i^v(t) dv, \;\; t \leq T, \tag{4.11}$$

where $\varphi_i^v, \; i = 1, \ldots, d, \; v \in [0, T]$, *are the* $\mathbb{G}$*-predictable integrands of the martingale representation of* $e^{-\Gamma(v)} F(v)\mu(v)$ *(cf. Eq.* (4.4)*).*

The proof of the first part (4.10) relies on a generalization of Proposition 4.11 in Barbarin (2008) and Proposition 4 in Biagini et al. (2013, p. 130, 138). Similar results were independently derived in Section 3.3 of Biagini et al. (2012) and in Section 4 of Biagini and Schreiber (2013). We added the specification (4.11) in analogy to (4.8). The whole proof is provided in the Appendix.

Combining the previous three lemmas with the Clark-Ocone formula from Malliavin calculus, we obtain the MRT decompositions for each summand of L_0 defined in Section 3.1 – and thus the MRT decomposition of L_0 itself by summing up the individual decompositions.

In what follows, let $\mathbb{D}_{1,2}$ denote the set of random variables that are Malliavin differentiable with respect to each one-dimensional Brownian motion W_i of $W = (W_1, \ldots, W_d)$, and let $D_{t,i}(\cdot)$ denote the respective time-t Malliavin derivative with respect to $W_i, \; i = 1, \ldots, d$. For a definition of Malliavin differentiability and the Malliavin derivative, we refer to Definition 3.1 in Di Nunno et al. (2009, p. 27). Note that all random variables in $\mathbb{D}_{1,2}$ are by definition in $L^2(\mathbb{P})$ and $\mathcal{G}_{T^*}$-measurable.

Proposition 4.8. *Assume that* $n = d$ *and that the inverse* $\sigma^{-1}(t) = \left(\sigma_{ij}^{-1}(t) \right)_{i,j=1,\ldots,n}$ *exists for all* $t \in [0, T^*]$ $\mathbb{P}$*-almost surely. Let* $0 \leq t_k \leq T^*, \; 0 \leq T \leq T^*$.

i) Let $L_0 = C_0$. *If* $C_0 \in \mathbb{D}_{1,2}$, *then the unique integrands of the MRT decomposition* (3.4) *of* L_0 *are given by*

$$\begin{aligned} \psi_i^W(t) &= \sum_{j=1}^{d} \mathrm{E}^{\mathbb{P}} \left(D_{t,j} \left(C_0 \right) \middle| \mathcal{G}_t \right) \sigma_{ji}^{-1}(t), \; i = 1, \ldots, n, \\ \psi^N(t) &= 0. \end{aligned}$$

ii) Let $L_0 = (m - N(t_k)) C_{a,k}$. *If* $e^{-\Gamma(t_k)} C_{a,k} \in \mathbb{D}_{1,2}$, *then the unique integrands of the MRT decomposition* (3.4) *of* L_0 *are given by*

$$\begin{aligned} \psi_i^W(t) &= \left[(m - N(t-)) e^{\Gamma(t)} \mathbb{1}_{[0,t_k]}(t) + (m - N(t_k)) e^{\Gamma(t_k)} \mathbb{1}_{(t_k,T^*]}(t) \right] \\ &\qquad \times \sum_{j=1}^{d} \mathrm{E}^{\mathbb{P}} \left(D_{t,j} \left(e^{-\Gamma(t_k)} C_{a,k} \right) \middle| \mathcal{G}_t \right) \sigma_{ji}^{-1}(t), \;\; i = 1, \ldots, n, \\ \psi^N(t) &= -\mathbb{1}_{[0,t_k]}(t)\, \mathrm{E}^{\mathbb{P}} \left(e^{\Gamma(t)-\Gamma(t_k)} C_{a,k} \middle| \mathcal{G}_t \right). \end{aligned}$$

iii) Let $L_0 = \int_0^T (m - N(t))C_a(t)dt$. *If* $C_a = (C_a(t))_{0\le t\le T}$ *is a* $\mathbb{G}$*-predictable process with* $\mathrm{E}^{\mathbb{P}}\left(\sup_{t\in[0,T]} |C_a(t)|\right) < \infty$, $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left([C_a(t)]^2\right) < \infty$, *and* $e^{-\Gamma(t)}C_a(t) \in \mathbb{D}_{1,2}$ *for all* $t \in [0,T]$, *then the unique integrands of the MRT decomposition* (3.4) *of* L_0 *are given by*

$$\psi_i^W(t) = \mathbb{1}_{[0,T]}(t)\,(m - N(t-))e^{\Gamma(t)} \times \sum_{j=1}^{d} \int_t^T \mathrm{E}^{\mathbb{P}}\left(D_{t,j}\left(e^{-\Gamma(v)}C_a(v)\right)\middle|\,\mathcal{G}_t\right) dv\, \sigma_{ji}^{-1}(t),\ \ i = 1,\dots,n,$$

$$\psi^N(t) = -\mathbb{1}_{[0,T]}(t) \int_t^T \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(v)}C_a(v)\middle|\mathcal{G}_t\right) dv.$$

iv) Let $L_0 = \int_0^T C_{ad}(t)dN(t)$. *If* $C_{ad} = (C_{ad}(t))_{0\le t\le T}$ *is a continuous and* $\mathbb{G}$*-predictable process with* $\mathrm{E}^{\mathbb{P}}\left(\sup_{t\in[0,T]} |C_{ad}(t)|\right) < \infty$ *and* $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left([C_{ad}(t)]^4\right) < \infty$, $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\mu^4(t)\right) < \infty$, *and* $e^{-\Gamma(t)}C_{ad}(t)\mu(t) \in \mathbb{D}_{1,2}$ *for all* $t \in [0,T]$, *then the unique integrands of the MRT decomposition* (3.4) *of* L_0 *are given by*

$$\psi_i^W(t) = \mathbb{1}_{[0,T]}(t)\,(m - N(t-))e^{\Gamma(t)} \times \sum_{j=1}^{d} \int_t^T \mathrm{E}^{\mathbb{P}}\left(D_{t,j}\left(e^{-\Gamma(v)}C_{ad}(v)\mu(v)\right)\middle|\,\mathcal{G}_t\right) dv\, \sigma_{ij}^{-1}(t),\ \ i = 1,\dots,n,$$

$$\psi^N(t) = -\mathbb{1}_{[0,T]}(t) \left[\int_t^T \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(v)}C_{ad}(v)\mu(v)\middle|\,\mathcal{G}_t\right) dv - C_{ad}(t)\right].$$

Proof. The integrands follow directly from Lemma 4.5, Lemma 4.6, and Lemma 4.7 together with the Clark-Ocone formula (Di Nunno et al., 2009, p. 196) and the proof of Proposition 3.5. Since each L_0 is square integrable as a result of the respective assumptions, the uniqueness follows from Proposition 4.1. □

Remark 4.9. Note that we omitted the MRT decomposition of discrete cash flows contingent on death since

$$(N(t_k) - N(t_{k-1}))\,C_{ad,k} = (m - N(t_{k-1}))C_{ad,k} - (m - N(t_k))C_{ad,k}$$

can be represented as a sum of two discrete survival cash flows. The MRT decompositions of these two summands can then be determined via Proposition 7 ii).

An application of the above proposition is given in the following example, where the MRT decomposition of a pure endowment portfolio is determined.

Example 4.10. Consider a portfolio of m pure endowment policies with survival benefit 1 at time T and single premium P_0 at time 0. For simplicity, assume a zero interest rate, so that the insurer's time-0 loss equals $L_0 = -mP_0 + (m - N(T))$. The mortality intensity is assumed to be a non-negative affine diffusion process

$$d\mu(t) = \theta(t, \mu(t))dt + \sigma(t, \mu(t))dW(t),\ \mu(0) = \mu_0,$$

where $(W(t))_{0\le t\le T}$ is a one-dimensional standard Brownian motion, so that (Biffis, 2005)

$$\mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^T \mu(s)ds}\middle|\,\mathcal{G}_t\right) = e^{\alpha(t)+\beta(t)\mu(t)},\ T \in (t, T^*], \tag{4.12}$$

where α and β satisfy certain Riccati ordinary differential equations. Clearly, since $-mP_0$ is deterministic, the integrands of its MRT decomposition are zero. Assume that $\sigma(t,\mu(t)) \neq 0$ for all $t \in [0,T]$ $\mathbb{P}$-almost surely and that $\mu(t),\ e^{\Gamma(t)},\ e^{-\Gamma(t)} \in \mathbb{D}_{1,2}$ with $D_t(\mu(t)) = \sigma(t,\mu(t))$ for all $t \in [0,T]$.[9] By applying part ii) of Proposition 4.8 to $(m - N(T))$ it follows that

$$L_0 - \mathrm{E}^{\mathbb{P}}(L_0) = \int_0^T (m - N(t-))e^{\Gamma(t)} \frac{\mathrm{E}^{\mathbb{P}}\left(D_t\left(e^{-\Gamma(T)}\right)\middle|\,\mathcal{G}_t\right)}{\sigma(t,\mu(t))} dM^W(t) - \int_0^T \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(T)}\middle|\,\mathcal{G}_t\right) dM^N(t).$$

Since $D_t(\mu(s)) = 0$ for all $t > s$, and thus $D_t(\Gamma(t)) = 0$, the chain rule (Proposition 1.2.3 in Nualart, 2006, p. 28) implies

$$D_t\left(e^{\Gamma(t)-\Gamma(T)}\right) = -e^{\Gamma(t)-\Gamma(T)} D_t(\Gamma(T) - \Gamma(t)) = -e^{\Gamma(t)-\Gamma(T)} D_t(\Gamma(T)) = e^{\Gamma(t)} D_t\left(e^{-\Gamma(T)}\right),$$

i.e. $e^{\Gamma(t)}\mathrm{E}^{\mathbb{P}}\left(D_t\left(e^{-\Gamma(T)}\right)\middle|\,\mathcal{G}_t\right) = \mathrm{E}^{\mathbb{P}}\left(D_t\left(e^{\Gamma(t)-\Gamma(T)}\right)\middle|\,\mathcal{G}_t\right)$. Furthermore, exchanging conditional expectation and Malliavin derivative operator (Di Nunno et al., 2009, Proposition 3.12, p. 33) together with (4.12) we have

$$\mathrm{E}^{\mathbb{P}}\left(D_t\left(e^{\Gamma(t)-\Gamma(T)}\right)\middle|\,\mathcal{G}_t\right) = D_t\left(\mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(T)}\middle|\,\mathcal{G}_t\right)\right) = D_t\left(e^{\alpha(t)+\beta(t)\mu(t)}\right).$$

The chain rule from Malliavin calculus (Proposition 1.2.3 in Nualart, 2006, p. 28) finally yields

$$D_t\left(e^{\alpha(t)+\beta(t)\mu(t)}\right) = e^{\alpha(t)+\beta(t)\mu(t)}\beta(t) D_t(\mu(t)) = e^{\alpha(t)+\beta(t)\mu(t)}\beta(t)\sigma(t,\mu(t)).$$

All-in-all, we obtain

$$L_0 - \mathrm{E}^{\mathbb{P}}(L_0) = \int_0^T (m - N(t-))\, e^{\alpha(t)+\beta(t)\mu(t)}\beta(t) dM^W(t) - \int_0^T e^{\alpha(t)+\beta(t)\mu(t)} dM^N(t),$$

where the first summand represents the systematic mortality risk and the second summand the unsystematic mortality risk.

Markov case

In what follows, we assume that the state process X is a Markovian diffusion process and that the insurance payments are functions of the state variables. In this case, given further conditions, we can directly evaluate the decompositions via Itô's formula rather than relying on Malliavin derivatives as in Proposition 4.8. We write $f \in C^{1,2}([0,T] \times \mathbb{R}^n)$ for a function $f : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ if the partial derivatives $\frac{\partial f}{\partial t},\ \frac{\partial f}{\partial x_i},\ \frac{\partial^2 f}{\partial x_i \partial x_j},\ 1 \le i,j \le n$, exist and are continuous on $(0,T) \times \mathbb{R}^n$, and if the indicated partial derivatives have continuous extensions to $[0,T] \times \mathbb{R}^n$.

Assumption 4.11. *The state process $X = ((X_1(t),\ldots,X_n(t))^\top)_{0\le t\le T^*}$ is an n-dimensional diffusion process satisfying*

$$dX(t) = \theta(t, X(t))dt + \sigma(t, X(t))dW(t), \tag{4.13}$$

with deterministic initial value $X(0) = x_0 \in \mathbb{R}^n$, where the drift vector $\theta : [0,T^*] \times \mathbb{R}^n \to \mathbb{R}^n$ and the volatility matrix $\sigma : [0,T^*] \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$ are continuous functions such that a unique strong solution to (4.13) *exists.*

[9] Note that for globally Lipschitz-continuous coefficients θ and σ with at most linear growth, diffusion processes are Malliavin differentiable (Nualart, 2006, Theorem 2.2.1, p. 119). However, as the discussion on the Malliavin differentiability of square-root processes shows (Alòs and Ewald, 2008), the general Malliavin differentiability of diffusion processes – and even affine processes – is not guaranteed.

Proposition 4.12. *Let X be an n-dimensional diffusion process as specified in Assumption 4.11 and assume that $n = d$ and that $\det \sigma(t, X(t)) \neq 0$ for all $t \in [0, T^*]$ $\mathbb{P}$-almost surely. Let $0 \leq t_k \leq T^*$, $0 \leq T \leq T^*$.*

i) Let $L_0 = C_0$. Assume that C_0 is square integrable and of the form

$$C_0 = e^{-\int_0^T g(s,X(s))ds} h(X(T))$$

for some measurable functions $g : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to [0, \infty)$. Define $f(t,x) := \mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^T g(s,X(s))ds} h(X(T)) \Big| X(t) = x\right)$. If $f \in C^{1,2}([0,T] \times \mathbb{R}^n)$, then the unique integrands of the MRT decomposition (3.4) *of L_0 are given by*

$$\psi_i^W(t) = \mathbb{1}_{[0,T]}(t)\, e^{-\int_0^t g(s,X(s))ds} \frac{\partial f}{\partial x_i}(t, X(t)),\ i = 1, \ldots, n,$$
$$\psi^N(t) = 0.$$

ii) Let $L_0 = (m - N(t_k))C_{a,k}$. Assume that $C_{a,k}$ is square integrable and of the form

$$C_{a,k} = e^{-\int_0^T g(s,X(s))ds} h(X(T))$$

for some measurable functions $g : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to [0, \infty)$.

(a) Assume that $T > t_k$, and define $f^A : [0, t_k] \times \mathbb{R}^n \to \mathbb{R}$ and $f^B : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ by

$$f^A(t,x) := \mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^{t_k} \mu(s,X(s))ds} e^{-\int_t^T g(s,X(s))ds} h(X(T)) \Big| X(t) = x\right),$$
$$f^B(t,x) := \mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^T g(s,X(s))ds} h(X(T)) \Big| X(t) = x\right).$$

If $f^A \in C^{1,2}([0,t_k] \times \mathbb{R}^n)$ and $f^B \in C^{1,2}([0,T] \times \mathbb{R}^n)$, then the unique integrands of the MRT decomposition (3.4) *of L_0 are given by*

$$\begin{aligned}\psi_i^W(t) = {} & \mathbb{1}_{[0,t_k]}(t)\,(m - N(t-))e^{-\int_0^t g(s,X(s))ds} \frac{\partial f^A}{\partial x_i}(t, X(t)) \\ & + \mathbb{1}_{(t_k,T]}(t)\,(m - N(t_k))e^{-\int_0^t g(s,X(s))ds} \frac{\partial f^B}{\partial x_i}(t, X(t)),\quad i = 1, \ldots, n,\end{aligned}$$
$$\psi^N(t) = -\mathbb{1}_{[0,t_k]}(t)\, e^{-\int_0^t g(s,X(s))ds} f^A(t, X(t)).$$

(b) Assume that $T \leq t_k$, and define $f^A : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ and $f^B : [0,t_k] \times \mathbb{R}^n \to \mathbb{R}$ by

$$f^A(t,x) := \mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^{t_k} \mu(s,X(s))ds} e^{-\int_t^T g(s,X(s))ds} h(X(T)) \Big| X(t) = x\right),$$
$$f^B(t,x) := \mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^{t_k} \mu(s,X(s))ds} \Big| X(t) = x\right).$$

If $f^A \in C^{1,2}([0,T] \times \mathbb{R}^n)$ and if in case $T < t_k$ additionally $f^B \in C^{1,2}([0,t_k] \times \mathbb{R}^n)$, then the unique integrands of the MRT decomposition (3.4) *of L_0 are given by*

$$\begin{aligned}\psi_i^W(t) = {} & \mathbb{1}_{[0,T]}(t)\,(m - N(t-))e^{-\int_0^t g(s,X(s))ds} \frac{\partial f^A}{\partial x_i}(t, X(t)) \\ & + \mathbb{1}_{(T,t_k]}(t)\,(m - N(t-))C_{a,k} \frac{\partial f^B}{\partial x_i}(t, X(t)),\quad i = 1, \ldots, n,\end{aligned}$$
$$\psi^N(t) = -\mathbb{1}_{[0,T]}(t)\, e^{-\int_0^t g(s,X(s))ds} f^A(t, X(t)) - \mathbb{1}_{(T,t_k]}\, C_{a,k} f^B(t, X(t)).$$

iii) Let $L_0 = \int_0^T (m - N(t))C_a(t)dt$. Assume that $C_a(t)$ is of the form

$$C_a(t) = e^{-\int_0^t g(s,X(s))ds} h(X(t))$$

for some measurable functions $g : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to [0,\infty)$, and assume that $\mathrm{E}^{\mathbb{P}}\left(\sup_{t\in[0,T]} |C_a(t)|\right) < \infty$ and $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left([C_a(t)]^2\right) < \infty$. Furthermore, define $f^v(t,x) := \mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^v [g(s,X(s))+\mu(s,X(s))]ds} h(X(v)) \middle| X(t) = x\right)$. If $f^v \in C^{1,2}([0,v] \times \mathbb{R}^n)$ for all $v \in [0,T]$, then the unique integrands of the MRT decomposition (3.4) *of L_0 are given by*

$$\psi_i^W(t) = \mathbb{1}_{[0,T]}(t)\,(m - N(t-))e^{-\int_0^t g(s,X(s))ds} \int_t^T \frac{\partial f^v}{\partial x_i}(t,X(t))dv, \quad i = 1,\dots,n,$$
$$\psi^N(t) = -\mathbb{1}_{[0,T]}(t)\, e^{-\int_0^t g(s,X(s))ds} \int_t^T f^v(t,X(t))dv.$$

iv) Let $L_0 = \int_0^T C_{ad}(t)dN(t)$. Assume that $C_{ad}(t)$ is of the form

$$C_{ad}(t) = e^{-\int_0^t g(s,X(s))ds} h(X(t))$$

for some measurable and continuous functions $g : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ and $h : \mathbb{R}^n \to [0,\infty)$, and that $\mathrm{E}^{\mathbb{P}}\left(\sup_{t\in[0,T]} |C_{ad}(t)|\right) < \infty$, $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left([C_{ad}(t)]^4\right) < \infty$, and $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\mu^4(t)\right) < \infty$. Define $f^v(t,x) := \mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^v [g(s,X(s))+\mu(s,X(s))]ds} h(X(v))\mu(v,X(v)) \middle| X(t) = x\right)$. If $f^v \in C^{1,2}([0,v] \times \mathbb{R}^n)$ for all $v \in [0,T]$, then the unique integrands of the MRT decomposition (3.4) *of L_0 are given by*

$$\psi_i^W(t) = \mathbb{1}_{[0,T]}(t)\,(m - N(t-))e^{-\int_0^t g(s,X(s))ds} \int_t^T \frac{\partial f^v}{\partial x_i}(t,X(t))dv, \quad i = 1,\dots,n,$$
$$\psi^N(t) = -\mathbb{1}_{[0,T]}(t)\,\left[e^{-\int_0^t g(s,X(s))ds} \int_t^T f^v(t,X(t))dv - C_{ad}(t)\right].$$

The proof of Proposition 4.12 uses Lemmas 4.5, 4.6, and 4.7 together with Itô's lemma. It is provided in the Appendix. Note that even if $n \neq d$ this proof implies the existence of the stated MRT decompositions. However, uniqueness is in general no longer guaranteed.

Example 4.13. We again consider the setting from Example 4.10 but now determine the MRT decomposition of $L_0 - \mathrm{E}^{\mathbb{P}}(L_0)$ by applying Proposition 4.12. Obviously, the mortality intensity in this setting is a one-dimensional diffusion process, and we have that $C_{a,1} = e^{-\int_0^T g(s,X(s))ds} h(X(T))$ for $g \equiv 0$ and $h \equiv 1$, where $t_0 = 0$ and $t_1 = T$. The affine property of the mortality model yields

$$\mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^T [g(s,X(s))+\mu(s,X(s))]ds} h(X(T)) \middle| \mathcal{G}_t\right) = \mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^T \mu(v)dv} \middle| \mathcal{G}_t\right) = e^{\alpha(t)+\beta(t)\mu(t)} =: f^A(t,\mu(t)).$$

Since the function f^A obviously satisfies the smoothness requirements and since $T = t_1$, part b) of Proposition 4.12 ii) yields the MRT decomposition

$$L_0 - \mathrm{E}^{\mathbb{P}}(L_0) = \int_0^T (m - N(t-))\, e^{\alpha(t)+\beta(t)\mu(t)}\beta(t)dM^W(t) - \int_0^T e^{\alpha(t)+\beta(t)\mu(t)}dM^N(t),$$

where $-mP_0$ again does not contribute to the integrands since it is deterministic. Of course, this coincides with the result of Example 4.10.

To verify whether the functions f satisfy the required smoothness conditions imposed in Proposition 4.12, one can for instance rely on the (sufficient) conditions in Heath and Schweizer (2000). Of course, in case an analytic expression cannot be determined, the respective function f can be computed numerically.

The following proposition illustrates the relation between the integrands from Proposition 4.8 and Proposition 4.12, and thus generalizes the last part of Example 4.10.

Proposition 4.14. *Let X be an n-dimensional diffusion process as specified in Assumption 4.11. Assume that $n = d$ and that the inverse $\sigma^{-1}(t, X(t)) = \left(\sigma_{ij}^{-1}(t, X(t))\right)_{i,j=1,\dots,n}$ exists for all $t \in [0, T^*]$ $\mathbb{P}$-almost surely. If $X_i(t) \in \mathbb{D}_{1,2}$ with $D_{t,j}X_k(t) = \sigma_{kj}(t, X(t))$ for all $i = 1, \dots, n,\ t \in [0, T^*]$, and if $f : [0, T^*] \times \mathbb{R}^n \to \mathbb{R},\ (t, x) \mapsto f(t, x)$, is a continuously differentiable function with bounded partial derivatives, then for all $t \in [0, T^*]$ we have $f(t, X(t)) \in \mathbb{D}_{1,2}$ and*

$$\sum_{j=1}^{d} D_{t,j} f(t, X(t)) \sigma_{ji}^{-1}(t, X(t)) = \frac{\partial}{\partial x_i} f(t, X(t)) \quad \forall i = 1, \dots, n.$$

Proof. Using the chain rule (Proposition 1.2.3 in Nualart, 2006, p. 28) and $D_{t,j}t = 0$ (Theorem 2.2.1 in Nualart, 2006, p. 119), it follows that

$$\begin{aligned}
D_{t,j} f(t, X(t)) &= \frac{\partial}{\partial t} f(t, X(t)) D_{t,j} t + \sum_{k=1}^{n} \frac{\partial}{\partial x_k} f(t, X(t)) D_{t,j} X_k(t) \\
&= \sum_{k=1}^{n} \frac{\partial}{\partial x_k} f(t, X(t)) \sigma_{kj}(t, X(t)) \in \mathbb{D}_{1,2}.
\end{aligned}$$

This, together with $\sum_{j=1}^{d} \sigma_{kj}(t, X(t)) \sigma_{ji}^{-1}(t, X(t)) = \mathbb{1}_{\{k=i\}}$, implies that

$$\begin{aligned}
\sum_{j=1}^{d} D_{t,j} f(t, X(t)) \sigma_{ji}^{-1}(t, X(t)) &= \sum_{k=1}^{n} \frac{\partial}{\partial x_k} f(t, X(t)) \sum_{j=1}^{d} \sigma_{kj}(t, X(t)) \sigma_{ji}^{-1}(t, X(t)) \\
&= \frac{\partial}{\partial x_i} f(t, X(t)).
\end{aligned}$$

□

Thus, comparing parts i) – iv) of Propositions 4.8 and 4.12 – and possibly exchanging conditional expectations and Malliavin derivative operators (Di Nunno et al., 2009, Proposition 3.12, p. 33) and applying the product rule (Di Nunno et al., 2009, Theorem 3.4, p. 30) – it is easy to verify that the respective expressions coincide if all assumptions are satisfied.

4.3 Diversification properties

It is well known that unsystematic mortality risk arising from finite insurance portfolios vanishes as the number of policyholders goes to infinity, i.e. it is *diversifiable*. In the next proposition, we show that the risk factor associated with unsystematic mortality risk within the MRT decomposition also satisfies this property. On the one hand, this corroborates the adequacy of the MRT decomposition, and, on the other hand, it allows for a crisp definition of unsystematic (mortality) risk within an insurance payoff. The following lemma will simplify the proof.

Lemma 4.15. *Let $T \in [0, T^*]$ be fixed. If $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\mu^2(t)\right) < \infty$ and if $\left(\psi^N(t)\right)_{0\leq t\leq T}$ is a $\mathbb{G}$-predictable process with $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\left[\psi^N(t)\right]^4\right) < \infty$, then*

$$\frac{1}{m}\int_0^T \psi^N(t)dM^N(t) \xrightarrow[m\to\infty]{L^2} 0.$$

Proof. We need to show that

$$\mathrm{E}^{\mathbb{P}}\left(\left[\frac{1}{m}\int_0^T \psi^N(t)dM^N(t) - 0\right]^2\right) = \frac{1}{m^2}\mathrm{E}^{\mathbb{P}}\left(\left[\int_0^T \psi^N(t)dM^N(t)\right]^2\right) \xrightarrow[m\to\infty]{} 0.$$

From Andersen et al. (1997, p. 78), we know that the predictable quadratic variation of $M^N(t)$ equals $\left\langle M^N, M^N\right\rangle(t) = \int_0^t (m - N(s-))\mu(s)ds$. Since $M^N(t)$ is a martingale, since ψ^N is assumed to be predictable, and since by the calculations below $\mathrm{E}^{\mathbb{P}}\left(\int_0^T \left[\psi^N(t)\right]^2 d\left\langle M^N, M^N\right\rangle(t)\right) < \infty$, it follows that $\int_0^T \psi^N(t)dM^N(t)$ is a square integrable martingale and that the Itô isometry applies (for both, see Klebaner, 2005, p. 234) yielding

$$\begin{aligned}\frac{1}{m^2}\mathrm{E}^{\mathbb{P}}\left(\left[\int_0^T \psi^N(t)dM^N(t)\right]^2\right) &= \frac{1}{m^2}\mathrm{E}^{\mathbb{P}}\left(\int_0^T \left[\psi^N(t)\right]^2 \underbrace{(m - N(t-))}_{\leq m}\mu(t)dt\right) \\ &\leq \frac{1}{m}\mathrm{E}^{\mathbb{P}}\left(\int_0^T \left[\psi^N(t)\right]^2 \mu(t)dt\right).\end{aligned} \tag{4.14}$$

Since by assumption $C_1 := \sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\left[\psi^N(t)\right]^4\right) < \infty$ and $C_2 := \sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\mu^2(t)\right) < \infty$, the theorem of Fubini-Tonelli and the Cauchy-Schwarz inequality yield

$$\begin{aligned}\mathrm{E}^{\mathbb{P}}\left(\int_0^T \left[\psi^N(t)\right]^2 \mu(t)dt\right) &= \int_0^T \mathrm{E}^{\mathbb{P}}\left(\left[\psi^N(t)\right]^2 \mu(t)\right) dt \\ &\overset{\text{Cauchy-Schwarz}}{\leq} \int_0^T \sqrt{\mathrm{E}^{\mathbb{P}}\left(\left[\psi^N(t)\right]^4\right)\mathrm{E}^{\mathbb{P}}\left(\mu^2(t)\right)}dt \leq \int_0^T \sqrt{C_1C_2} = T\sqrt{C_1C_2} =: C < \infty.\end{aligned}$$

Together with (4.14), we obtain

$$0 \leq \frac{1}{m^2}\mathrm{E}^{\mathbb{P}}\left(\left[\int_0^T \psi^N(t)dM^N(t)\right]^2\right) \leq \frac{1}{m}\mathrm{E}^{\mathbb{P}}\left(\int_0^T \left[\psi^N(t)\right]^2 \mu(t)dt\right) \leq \frac{1}{m}C \xrightarrow[m\to\infty]{} 0.$$

$\square$

In order to show L^2-convergence, the following proposition is restricted to bounded C_a and C_{ad}.

Proposition 4.16. *Assume the setting and assumptions from Proposition 4.8 with resulting unsystematic mortality risks in part ii), iii), and iv) of, respectively,*

$$\begin{aligned}R^{(m)}_{n+1,ak} &= -\int_0^{t_k} \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(t_k)}C_{a,k}\middle|\,\mathcal{G}_t\right)dM^N(t), \\ R^{(m)}_{n+1,a} &= -\int_0^T\int_t^T \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(s)}C_a(s)\middle|\mathcal{G}_t\right)ds\,dM^N(t), \\ R^{(m)}_{n+1,ad} &= -\int_0^T\int_t^T \left[\mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(s)}C_{ad}(s)\mu(s)\middle|\,\mathcal{G}_t\right)ds - C_{ad}(t)\right]dM^N(t).\end{aligned}$$

i) If $C_{a,k} \in L^4(\mathbb{P})$ and $\sup_{t\in[0,t_k]} \mathrm{E}^{\mathbb{P}}\left(\mu^2(t)\right) < \infty$, then $\frac{1}{m} R^{(m)}_{n+1,ak} \xrightarrow[m\to\infty]{L^2} 0$.

ii) If $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\mu^2(t)\right) < \infty$ and C_a is bounded, then $\frac{1}{m} R^{(m)}_{n+1,a} \xrightarrow[m\to\infty]{L^2} 0$.

iii) If C_{ad} is bounded, then $\frac{1}{m} R^{(m)}_{n+1,ad} \xrightarrow[m\to\infty]{L^2} 0$.

The proof of Proposition 4.16 essentially checks the assumptions of Lemma 4.15 and can be found in the Appendix. While unsystematic mortality risk diversifies, Proposition 4.19 shows that the remaining risk factors also converge with the number of contracts, but in general not to zero, i.e. they are *not diversifiable*. This confirms their interpretation as systematic risks, particularly since the limits no longer depend on $N(t)$. In applications, if the portfolio is sufficiently large, the limits can be used as risk approximations.

Remark 4.17. Convergence in probability of the unsystematic risk factors (instead of L^2-convergence) can be shown under less restrictive assumptions, e.g. by applying the (stochastic) dominated convergence theorem similarly as below for the systematic risk factors.

Lemma 4.18. *If $\zeta = (\zeta(t))_{0\le t\le T}$ is $\mathbb{G}$-predictable and $\int_0^T \zeta(t)^2 dt < \infty$ almost surely, then for $0 \le t_k \le T \le T^*$*

$$\frac{1}{m}\int_0^T \left[(m - N(t-))e^{\Gamma(t)}\mathbb{1}_{[0,t_k]} + (m - N(t_k))e^{\Gamma(t_k)}\mathbb{1}_{(t_k,T]}\right]\zeta(t)dW(t) \xrightarrow[m\to\infty]{P} \int_0^T \zeta(t)dW(t),$$

where $(W(t))_{0\le t\le T^*}$ is a one-dimensional Brownian motion.

Proof. Define

$$\zeta_m(t) := \left[\frac{(m - N(t-))}{m}e^{\Gamma(t)}\mathbb{1}_{[0,t_k]}(t) + \frac{(m - N(t_k))}{m}e^{\Gamma(t_k)}\mathbb{1}_{(t_k,T]}(t)\right]\zeta(t).$$

If $\zeta_m = (\zeta_m(t))_{0\le t\le T}$, $m \in \mathbb{N}$, are predictable processes with $\zeta_m(t) \xrightarrow[m\to\infty]{a.s.} \zeta(t)$ for all $t \in [0,T]$, and if there exists a W-integrable process $\alpha = (\alpha(t))_{0\le t\le T}$ such that $|\zeta_m(t)| \le \alpha(t)$ for all $m \in \mathbb{N}$, $t \in [0,T]$, then the statement of the lemma follows by the dominated convergence theorem for stochastic integrals (Protter, 2005, p. 176) . Since ζ and μ are by assumption predictable, it follows that ζ_m is predictable for each $m \in \mathbb{N}$. Furthermore, since the remaining lifetimes τ_x^i, $i \in \mathbb{N}$, are assumed to be conditionally i.i.d., a conditional version of Kolmogorov's strong law of large numbers (Majerek et al., 2005, p. 154) together with the continuity of $\mu(t)$ yields that

$$\frac{m - N(t-)}{m} \xrightarrow[m\to\infty]{a.s.} e^{-\int_0^t \mu(s)ds}.$$

As a result, $\zeta_m(t) \xrightarrow[m\to\infty]{a.s.} \zeta(t)$ for all $t \in [0,T]$. Furthermore, since $\frac{m-N(t-)}{m} \le 1$ and $\mu(t)$ is positive for all $t \in [0,T]$, we have

$$|\zeta_m(t)| \le \left[e^{\Gamma(t)}\mathbb{1}_{[0,t_k]}(t) + e^{\Gamma(t_k)}\mathbb{1}_{(t_k,T]}(t)\right]|\zeta(t)| \le e^{\Gamma(T)}|\zeta(t)| =: \alpha(t).$$

Since $\mu(t)$ has continuous paths (particularly on [0,T]), it follows that $e^{2\Gamma(T)} < \infty$ a.s. Together with the assumption $\int_0^T \zeta(t)^2 dt < \infty$ a.s., we obtain that $\int_0^T \alpha(t)^2 dt < \infty$ with probability one. Since α is also $\mathbb{G}$-predictable, this implies that α is W-integrable (Klebaner, 2005, p. 96), and the statement follows. $\square$

Proposition 4.19. *Assume the setting and assumptions from Proposition 4.8 with resulting systematic risks in part ii), iii), and iv) of*

$$R^{(m)}_{i,\cdot} := \int_0^T \left[(m - N(t-))e^{\Gamma(t)}\mathbb{1}_{[0,t_k]}(t) + (m - N(t_k))e^{\Gamma(t_k)}\mathbb{1}_{(t_k,T]}(t)\right] \sum_{j=1}^{d} \varphi_{j,\cdot}(t)\sigma^{-1}_{ji}(t)\, dM^W_i(t),$$

where $0 \le T \le T^*$*, and for the different parts*

$$\begin{aligned}
\varphi_{j,ak}(t) &= \mathrm{E}^{\mathbb{P}}\left(D_{t,j}\left(e^{-\Gamma(t_k)}C_{a,k}\right)\middle|\, \mathcal{G}_t\right) \quad \textit{(part ii))},\\
\varphi_{j,a}(t) &= \int_t^T \mathrm{E}^{\mathbb{P}}\left(D_{t,j}\left(e^{-\Gamma(s)}C_a(s)\right)\middle|\, \mathcal{G}_t\right) ds \quad \textit{(part iii) where } t_k = T\textit{)},\\
\varphi_{j,ad}(t) &= \int_t^T \mathrm{E}^{\mathbb{P}}\left(D_{t,j}\left(e^{-\Gamma(s)}C_{ad}(s)\mu(s)\right)\middle|\, \mathcal{G}_t\right) ds \quad \textit{(part iv) where } t_k = T\textit{)}.
\end{aligned}$$

Then it follows for $i = 1, \dots, n$*:*

$$\frac{1}{m} R^{(m)}_{i,\cdot} \xrightarrow[m\to\infty]{P} \int_0^T \sum_{j=1}^{d} \varphi_{j,\cdot}(t)\sigma^{-1}_{ji}(t) dM^W_i(t).$$

The proof of Proposition 4.19 is based on Lemma 4.18 and is relegated to the Appendix. The following corollary emphasizes that the limits of the considered risk factors exactly equal the risk factors of the limit of the corresponding total risk, i.e. MRT decomposition and limit can be interchanged. The proof of this statement can again be found in the Appendix.

Corollary 4.20. *Assume the setting and assumptions from Proposition 4.8 with total risks in part ii), iii), and iv) of, respectively,*

$$L^{(m)}_{0,ak} := (m - N(t_k))C_{a,k}, \;\; L^{(m)}_{0,a} := \int_0^T (m - N(t))C_a(t)dt, \;\; L^{(m)}_{0,ad} := \int_0^T C_{ad}(t)dN(t).$$

Then the following holds:

i) $\frac{1}{m} L^{(m)}_{0,\cdot} \xrightarrow[m\to\infty]{a.s.} \mathrm{E}^{\mathbb{P}}\left(L^{(1)}_{0,\cdot}\middle|\, \mathcal{G}_{T^*}\right)$.

ii) Defining the MRT decompositions

- $(L^{(m)}_{0,\cdot} - \mathrm{E}^{\mathbb{P}}(L^{(m)}_{0,\cdot}), X_1, \dots, X_n, N) \overset{MRT}{\leftrightarrow} (R^{(m)}_{1,\cdot}, \dots, R^{(m)}_{n+1,\cdot}),\ m \in \mathbb{N}$, *and*
- $(\mathrm{E}^{\mathbb{P}}(L^{(1)}_{0,\cdot}|\mathcal{G}_{T^*}) - \mathrm{E}^{\mathbb{P}}(L^{(1)}_{0,\cdot}), X_1, \dots, X_n, N) \overset{MRT}{\leftrightarrow} (R^*_{1,\cdot}, \dots, R^*_{n+1,\cdot})$,

and additionally assuming the respective assumptions of Proposition 4.16, then it follows for $i = 1, \dots, n+1$*:*

$$\frac{1}{m} R^{(m)}_{i,\cdot} \xrightarrow[m\to\infty]{P} R^*_{i,\cdot}\,.$$

5 Numerical example

In order to demonstrate the applicability and usefulness of the MRT decomposition, we derive the equity, interest, systematic, and unsystematic mortality risk factor of a return-of-premium GMDB within a Variable Annuity (VA). VAs are deferred, fund-linked annuity contracts, and GMDBs are common embedded riders that guarantee a minimal amount paid upon the policyholder's death (see Bauer et al. (2008) for details on VAs with guaranteed minimum benefits).

We assume that the VA is offered against a single premium P_0 paid at time 0, which is fully invested in a fund $S = (S(t))_{0\le t\le T^*}$ modeled as a geometric Brownian motion with drift μ_S and volatility σ_S:

$$dS(t) = \mu_S S(t)dt + \sigma_S S(t)dW_S(t), \;\; S(0) > 0,$$

where $W_S = (W_S(t))_{0\le t\le T^*}$ denotes a $\mathbb{P}$-Brownian motion. In case the policyholder dies during the deferment period $[0,T]$, the GMDB guarantees that the death benefit paid at the end of the year of death equals at least the single premium P_0 (return of premium death benefit). We focus on the insurer's risk from the GMDB guarantee and assume that the company charges no fee for this embedded rider. Thus, the policyholder's account value equals $A(t) = \frac{P_0}{S(0)} S(t), \; t \in [0,T]$, and if identical contracts are issued to m homogeneous individuals, the total discounted loss of the insurance company amounts to

$$L_0 = \sum_{k=1}^{T} \left(N(t_k) - N(t_{k-1})\right) e^{-\int_0^{t_k} r(s)ds} \max\{P_0 - A(t_k), 0\}, \tag{5.1}$$

where $r(t)$ denotes the short rate, and $t_k = k, \; k = 0, 1, \ldots, T$. We can interpret L_0 as the (stochastic) amount of money the insurance company needs at time 0 for covering the GMDB liabilities given it invests in the bank account. Note that a single upfront fee on top of P_0 would not change $L = L_0 - \mathrm{E}^{\mathbb{P}}(L_0)$ thus leading to the same MRT decomposition.

The short rate $r = (r(t))_{0\le t\le T^*}$ is assumed to follow a positive Cox-Ingersoll-Ross (CIR) process

$$dr(t) = \kappa(\theta - r(t))dt + \sigma_r \sqrt{r(t)}dW_r(t), \;\; r(0) > 0,$$

where κ, θ, $\sigma_r \in \mathbb{R}$, $2\kappa\theta \ge \sigma_r^2$, and $W_r = (W_r(t))_{0\le t\le T^*}$ is a $\mathbb{P}$-Brownian motion. Since r is an affine process, it follows that

$$\mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^T r(s)ds} \middle| \mathcal{G}_t \right) = e^{\alpha_r(t,T) - \beta_r(t,T)r(t)}, \; T \in [t, T^*],$$

where (Brigo and Mercurio, 2006, p. 66):

$$\alpha_r(t,T) = \tfrac{2\kappa\theta}{\sigma_r^2} \log\left(\tfrac{2he^{(\kappa+h)\frac{T-t}{2}}}{2h + (\kappa+h)(e^{h(T-t)} - 1)} \right), \; \beta_r(t,T) = \tfrac{2(e^{h(T-t)} - 1)}{2h + (\kappa+h)(e^{h(T-t)} - 1)}, \; h = \sqrt{\kappa^2 + 2\sigma_r^2}.$$

Following Dahl and Møller (2006), we assume that under $\mathbb{P}$ the mortality intensity process $\mu = (\mu(t))_{0\le t\le T^*}$ follows a positive time-inhomogeneous CIR process

$$d\mu(t,x) = (\gamma(t,x) - \delta(t,x)\mu(t,x))dt + \sigma_\mu(t,x)\sqrt{\mu(t,x)}dW_\mu(t), \;\; \mu(0,x) = \mu^0(x),$$

where x denotes the policyholder's age at time 0, $W_\mu = (W_\mu(t))_{0\le t\le T^*}$ is a $\mathbb{P}$-Brownian motion, the initial mortality intensities $\mu^0(x+t) = a + bc^{x+t}$ are assumed to follow the Gompertz-Makeham mortality law, and

$$\gamma(t,x) = \frac{1}{2}\hat{\sigma}^2 \mu^0(x+t), \;\; \delta(t,x) = \hat{\delta} - \frac{\frac{d}{dt}\mu^0(x+t)}{\mu^0(x+t)}, \;\; \sigma_\mu(t,x) = \hat{\sigma}\sqrt{\mu^0(x+t)},$$

for some deterministic parameters $a, b, c, \hat{\delta}$, and $\hat{\sigma}$. The specified mortality intensity process is again affine so that $\mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^T \mu(s,x)ds}\Big|\, \mathcal{G}_t\right) = e^{\alpha_\mu(t,T,x)-\beta_\mu(t,T,x)\mu(t,x)},\ T \in [t, T^*]$, where α_μ and β_μ satisfy the ordinary differential equations specified in Proposition 3.1 of Dahl and Møller (2006, p. 197). In what follows, we only consider a single age cohort, i.e. we fix the initial age x, so that we no longer indicate the dependency on the age cohort but just write $\mu(t)$ and $\sigma_\mu(t)$.

Since we assume that $W_S,\ W_r$, and W_μ are independent, one-dimensional Brownian motions, the volatility function of the process $X := (S, r, \mu)^\top$ is $\sigma(t, x) = \mathrm{diag}\{\sigma_S x_1, \sigma_r\sqrt{x_2}, \sigma_\mu(t)\sqrt{x_3}\}$. Thus, it follows that $\det \sigma(t, x) \neq 0$ for all $t \in [0, T^*]$ and all values x the process $X(t),\ t \in [0, T^*]$, assumes.

For deriving the MRT decomposition of L_0 defined in (5.1), first note that L_0 can be rewritten as

$$\begin{aligned} L_0 = &\sum_{k=1}^{T} (m - N(t_{k-1}))\, e^{-\int_0^{t_k} r(s)ds} \max\{P_0 - A(t_k), 0\} \\ &- \sum_{k=1}^{T} (m - N(t_k))\, e^{-\int_0^{t_k} r(s)ds} \max\{P_0 - A(t_k), 0\}, \end{aligned} \tag{5.2}$$

i.e. it is a sum of survival benefits. Since X is a Markov process, we can define the functions

$$\begin{aligned} f_k^{A1}(t, x) &:= \mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^{t_{k-1}} \mu(s,X(s))ds} e^{-\int_t^{t_k} r(s)ds} \max\{P_0 - A(t_k), 0\}\Big|\, X(t) = x\right),\quad 0 \le t \le t_{k-1}, \\ f_k^{B1}(t, x) &:= \mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^{t_k} r(s)ds} \max\{P_0 - A(t_k), 0\}\Big|\, X(t) = x\right),\quad 0 \le t \le t_k, \\ f_k^{A2}(t, x) &:= \mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^{t_k} [r(s)+\mu(s,X(s))]ds} \max\{P_0 - A(t_k), 0\}\Big|\, X(t) = x\right),\quad 0 \le t \le t_k, \end{aligned}$$

which can be simplified by using the independence of $S,\ r$, and μ, as well as exploiting the log-normal distribution of S, and the affine property of r and μ. This immediately shows that all three functions are sufficiently smooth, so that we can apply Proposition 4.12 ii). We obtain the MRT decomposition

$$L = L_0 - \mathrm{E}^{\mathbb{P}}(L_0) = \sum_{i=1}^{n+1} R_i,$$

where the systematic risk factors implied by $X_i,\ i = 1, 2, 3$, are given by

$$\begin{aligned} R_i = &\sum_{k=1}^{T} \Bigg(\int_0^{t_{k-1}} (m - N(t-))e^{-\int_0^t r(s)ds} \frac{\partial}{\partial x_i} f_k^{A1}(t, X(t))dM_i^W(t) \\ &+ \int_{t_{k-1}}^{t_k} (m - N(t_{k-1}))e^{-\int_0^t r(s)ds} \frac{\partial}{\partial x_i} f_k^{B1}(t, X(t))dM_i^W(t) \Bigg) \\ &- \sum_{k=1}^{T} \int_0^{t_k} (m - N(t-))e^{-\int_0^t r(s)ds} \frac{\partial}{\partial x_i} f_k^{A2}(t, X(t))dM_i^W(t), \end{aligned}$$

respectively, and the unsystematic mortality risk factor is given by

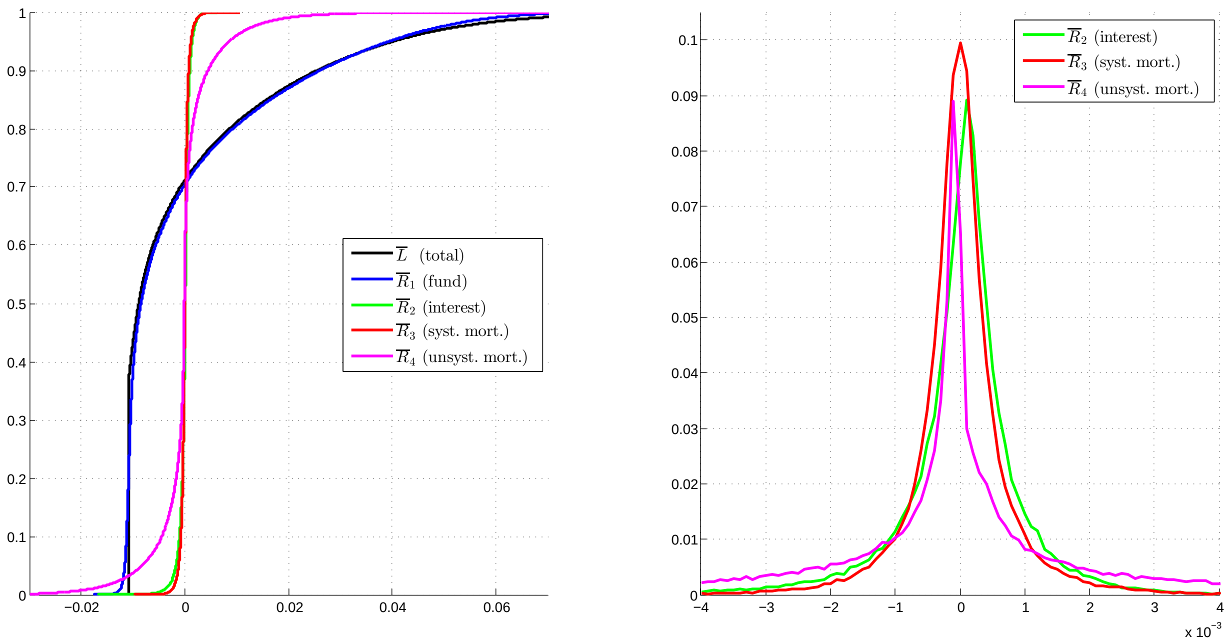
$$R_4 = -\sum_{k=1}^{T} \int_0^{t_{k-1}} e^{-\int_0^t r(s)ds} f_k^{A1}(t, X(t))dM^N(t) + \sum_{k=1}^{T} \int_0^{t_k} e^{-\int_0^t r(s)ds} f_k^{A2}(t, X(t))dM^N(t).$$

For the numerical calculations, we consider $m = 100$ GMDB contracts with maturity $T = 15$ and single premium $P_0 = 100{,}000$. All policyholders are assumed to be of age $x = 50$ at time 0. We perform

$N = 100{,}000$ simulations for determining the distributions of L, R_1, R_2, R_3, and R_4. For projecting the risk drivers r and μ as well as for approximating the stochastic integrals, we use an Euler scheme with $n = 100$ time steps per year. The number of survivors in the portfolio is projected by means of the binomial distribution conditioned on the mortality intensities. The ODEs implied by the mortality model are solved numerically. With respect to the mortality model, we adopt the parameter values for year 2003, case II, males, from Tables 1 to 3 in Dahl and Møller (2006, p. 211): $a = 0.000134$, $b = 0.0000353$, $c = 1.1020$, $\hat{\delta} = 0.008$, and $\hat{\sigma} = 0.02$. For the interest model, we assume $\kappa = 0.2$, $\theta = 0.025$, $\sigma_r = 0.075$, and $r(0) = 0.0029$. Thus, the Feller condition $2\kappa\theta > \sigma_r^2$ is satisfied. The parameters of the geometric Brownian motion are set to $\mu_S = 0.06$ and $\sigma_S = 0.22$.

We focus on the distributions scaled by the number of policyholders in the portfolio and the single premium, i.e. we consider $\overline{L} := \frac{1}{mP_0}L$, $\overline{R}_i := \frac{1}{mP_0}R_i$, $i = 1, \dots, 4$. The resulting empirical distribution functions of the total risk $\overline{L}$, the fund risk $\overline{R}_1$, the interest risk $\overline{R}_2$, the systematic mortality risk $\overline{R}_3$, and the unsystematic mortality risk $\overline{R}_4$ are shown in Figure 5.1(a). We find that the distribution function of the fund risk factor is right-skewed while the distribution functions of all other risk factors are approximately symmetric. Moreover, the plots indicate that the fund is the most relevant risk driver since the distribution of the risk factor closely resembles the distribution of the total risk. This seems intuitive since the fund value determines whether and to what extent the GMDB guarantee is in the money in case of death.

For $m = 100$ contracts, the randomness of the number of deaths within $[0, T]$, which trigger possible payoffs, also seems to be rather high: The range of likely outcomes of the unsystematic mortality risk factor is rather wide compared to the ranges of the interest risk factor and the systematic mortality risk factor. To further illustrate their relationship, we sort the respective outcomes into equally spaced bins of size $1e$-4 and plot the corresponding relative frequencies in Figure 5.1(b). We observe that the tails of the interest risk are heavier than the tails of the systematic mortality risk, but considerably lighter than the tails of the unsystematic mortality risk.



(a) Empirical distribution functions.

(b) Relative frequencies using bins of size $1e$-4.

Figure 5.1: GMDB portfolio with $m = 100$ contracts.

The resulting decomposition can now be used to allocate risk capital as cast by a homogeneous risk measure to the different sources of risk via the so-called Euler principle (see Bauer and Zanjani (2013) for an extensive discussion on allocation principles and Karabey et al. (2014) for a related application). More precisely, for a homogeneous risk measure ρ, Euler's homogeneous function theorem yields

$$\rho(\overline{L}) = \sum_{i=1}^{4} \left. \frac{\partial \rho\left(a_1\overline{R}_1 + a_2\overline{R}_2 + a_3\overline{R}_3 + a_4\overline{R}_4\right)}{\partial a_i} \right|_{a_1=a_2=a_3=a_4=1}, \tag{5.3}$$

and each summand can be interpreted as the risk contribution of the respective risk factor.

Table 5.1 provides results for three risk measures: standard deviation (Std. Dev.), Value-at-Risk at the 99% level ($\mathrm{VaR}_{0.99}$), and Tail-Value-at-Risk at the 99% level ($\mathrm{TVaR}_{0.99}$). For each risk measure, we report the total risk capital $\rho(\overline{L})$ (per unit per contract) as well as the risk factor contributions according to the Euler principle (absolute and as a percentage of the sum of the four risk contributions), where we use finite difference approximations for the derivatives in (5.3). As a result of the numerical approximation, the allocated values do not perfectly add up to the total risk capital $\rho(\overline{L})$, although the deviation is small (0.0002 for VaR).

		$\overline{L}$ (total)	$\overline{R}_1$ (fund)	$\overline{R}_2$ (interest)	$\overline{R}_3$ (syst. mort.)	$\overline{R}_4$ (unsyst. mort.)
Std. Dev.	abs.	0.0179	0.0160	0.0001	0.0000	0.0018
	perc.		89.3%	0.3%	0.2%	10.1%
$\mathrm{VaR}_{0.99}$	abs.	0.0675	0.0553	0.0005	0.0002	0.0117
	perc.		81.7%	0.8%	0.2%	17.3%
$\mathrm{TVaR}_{0.99}$	abs.	0.0813	0.0592	0.0008	0.0004	0.0208
	perc.		72.9%	1.0%	0.5%	25.6%

Table 5.1: The total risk capital and the Euler risk contributions for a GMDB portfolio with $m = 100$ contracts in absolute terms (abs.), and relative to the sum of the four Euler risk contributions (perc.).

The allocated risk contributions confirm our observations from the empirical distribution functions and the relative frequencies. The fund risk makes up between about 73% and 89% of the total risk capital, depending on the measure, whereas unsystematic mortality risk is the second-most significant factor accounting for between roughly 10% and 26%. It appears that unsystematic mortality risk becomes more relevant in the tail of the aggregate risk, which is intuitive since high losses can only occur if the policyholder actually dies. The risk contributions of interest and systematic mortality together amount to only about 1% of the total risk for all measures.

This order changes when increasing the number of policies within the insurer's portfolio since unsystematic mortality risk is *diversifiable* (cf. Section 4.3). For instance, Figure 5.2 illustrates the distribution functions of the four risk factors and the total risk for a portfolio of $10{,}000$ policies. While the fund risk still is the dominant risk factor, now unsystematic mortality risk exhibits the most concentrated distribution – and thus the smallest risk contribution. Interest rate risk and systematic mortality risk now present the second- and third-most important risk factors, although their influence seems marginal relative to the total risk. However, it is important to keep in mind that fund risk may be hedged, whereas at least for the systematic mortality risk hedging opportunities are scarce. We leave the further exploration of risk decompositions of hedged positions for future research (for a more detailed study in the context of annuitization options based on the methods presented in this paper, see Schilling (2015)).

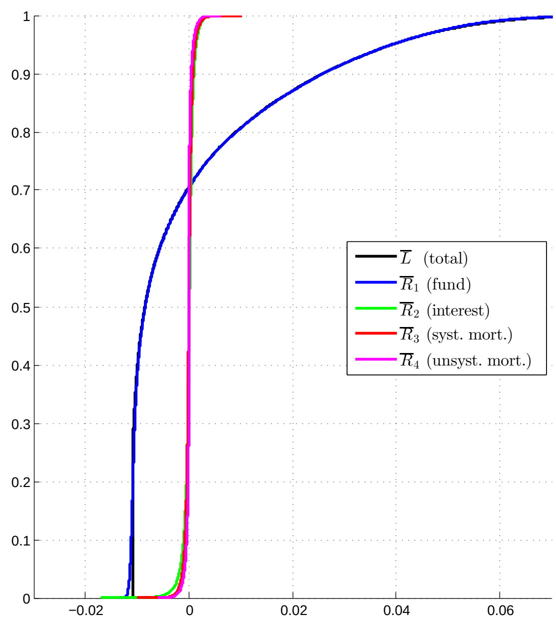


Figure 5.2: Empirical distribution functions of the total risk $\overline{L}$ and the risk factors $\overline{R}_1$, $\overline{R}_2$, $\overline{R}_3$, and $\overline{R}_4$ for a GMDB portfolio with $m = 10{,}000$ contracts.

6 Conclusion

The present paper provides a detailed analysis of risk decomposition methods which allocate the risk in (life) insurance liabilities to risk factors associated with different sources of uncertainty. For evaluating the usefulness of different approaches, we introduce a list of properties we posit a *meaningful* risk decomposition should satisfy. Then we propose a decomposition method, labeled *MRT decomposition*, that satisfies all of these desirable properties – as opposed to other approaches applied in the quantitative insurance literature.

The introduced MRT decomposition method is based on martingale representation. We discuss existence, uniqueness, and its calculation in a relatively general life insurance setting with an arbitrary (finite) insurance portfolio modeled by a counting process and an insurance payoff entailing discrete as well as continuous survival and death benefits. The (systematic) sources of risk are assumed to be driven by a finite-dimensional Brownian motion. We derive explicit formulas for the decomposition by means of the Clark-Ocone formula from Malliavin calculus in the general case and by Itô's lemma for diffusion processes in the Markov case. We also show that the unsystematic mortality risk as specified by the MRT decomposition is diversifiable, i.e. it vanishes as the portfolio increases, whereas the systematic risk factors approach a non-zero limit. As an example we provide the MRT decomposition of a Variable Annuity policy with return-of-premium Guaranteed Minimum Death Benefit.

Extensions include the generalization of the setting to a broader class of driving processes. This seems particularly relevant for the application in non-life insurance. Moreover, the considered decomposition approaches from literature are all static in nature, which translates to our list of desirable properties, whereas the MRT decomposition is already formulated time-dynamically. A closer look at this aspect in future research might be worthwhile. Finally, the application of the methods in this paper to study the risk drivers of advanced insurance guarantees such as Guaranteed Annuity Options or Guaranteed Minimum Income Benefits and Guaranteed Lifetime Withdrawal Benefits within Variable Annuities present interesting problems for applied research.

Acknowledgment

Katja Schilling acknowledges financial support from the DFG Research Training Group 1100 at the University of Ulm.

Appendix: Proofs

Proof of Lemma 3.3. i) Since the drift vector θ is $\mathbb{G}$-adapted with continuous paths, it follows that A_i^W is a predictable finite variation process. Since M_i^W is a local martingale and $X_i(t) = X_i(0) + M_i^W(t) + A_i^W(t)$ for all $t \in [0, T^*]$, A_i^W is a compensator of X_i. The uniqueness follows by Theorem 34 in Protter (2005, p. 130).

ii) From the assumptions, A^N is a predictable finite variation process and M^N is a martingale (for the latter, cf. Bielecki and Rutkowski, 2004, p. 153). Thus, A^N is a compensator of N and the uniqueness again follows by Theorem 34 in Protter (2005, p. 130).

$\square$

Proof of Lemma 4.5. Since $U = (U(t))_{0 \leq t \leq T^*}$ with $U(t) := \mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(T)}F \middle| \mathcal{G}_t\right)$ is a $\mathbb{G}$-martingale, it follows by the martingale representation theorem that there exist predictable processes $\varphi_1, \ldots, \varphi_d$ such that (4.4) holds.

We first show the lemma for a single policyholder with remaining lifetime τ_x^i, i.e. $m = 1$ and $\mathbb{F} = \mathbb{G} \vee \mathbb{I}^i$ for some arbitrary but fixed $i \in \{1, \ldots, m\}$. Define $L_i(t) := \mathbb{1}_{\{\tau_x^i > t\}} e^{\Gamma(t)}$ and $\tilde{L}_i(t) := \mathrm{E}^{\mathbb{P}}\left(L_i(T) \middle| \mathcal{F}_t\right)$. Since $L_i(t)$ is an $\mathbb{F}$-martingale (Bielecki and Rutkowski, 2004, p. 152), it follows that $\tilde{L}_i(t) = L_i(t)$ for $t \leq T$ and $\tilde{L}_i(t) = L_i(T)$ for $t \geq T$. Furthermore, $U(T^*) = e^{-\Gamma(T)}F$, which implies $Z_i := \mathbb{1}_{\{\tau_x^i > T\}} F = \tilde{L}_i(T^*) U(T^*)$. Thus, applying the Itô integration by parts formula (Protter, 2005, p. 68) to the product $\tilde{L}_i(t)U(t)$ and considering the continuity of $U(t)$ yields

$$\begin{aligned} Z_i &= \tilde{L}_i(0)U(0) + \int_0^{T^*} \tilde{L}_i(t-) dU(t) + \int_0^{T^*} U(t) d\tilde{L}_i(t) + [\tilde{L}_i, U](T^*) \\ &= L_i(0)U(0) + \int_0^{T^*} \left[L_i(t-)\mathbb{1}_{[0,T]}(t) + L_i(T)\mathbb{1}_{(T,T^*]}(t) \right] dU(t) + \int_0^T U(t) dL_i(t) + [L_i, U](T), \end{aligned} \tag{A.1}$$

where the second equality follows from the definition of $\tilde{L}_i$. Using $\mathbb{1}_{\{\tau_x^i > 0\}} = 1$ a.s. (which follows from the assumptions on μ), (3.2), and the $\mathcal{G}_{T^*}$-measurability of F, we have that

$$L_i(0)U(0) \overset{\text{a.s.}}{=} \mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(T)}F\right) = \mathrm{E}^{\mathbb{P}}\left(\mathrm{E}^{\mathbb{P}}\left(\mathbb{1}_{\{\tau_x^i > T\}} \middle| \mathcal{G}_{T^*}\right) F\right) = \mathrm{E}^{\mathbb{P}}\left(\mathbb{1}_{\{\tau_x^i > T\}} F\right) = \mathrm{E}^{\mathbb{P}}(Z_i).$$

Also note that

$$M_i^N(t) := \mathbb{1}_{\{\tau_x^i \leq t\}} - \int_0^t \mathbb{1}_{\{\tau_x^i > s-\}} \mu(s) ds = \mathbb{1}_{\{\tau_x^i \leq t\}} - \int_0^{t \wedge \tau_x^i} \mu(s) ds.$$

Thus, since the $\mathbb{G}$-adapted cumulative mortality intensity Γ of τ_x^i is continuous and increasing, Proposition 5.1.3 (i) from Bielecki and Rutkowski (2004, p. 153) implies that

$$dL_i(t) = -L_i(t-) dM_i^N(t).$$

Plugging in the definitions of L_i and M_i^N, this can be further written as

$$dL_i(t) = -e^{\Gamma(t)} \left(\mathbb{1}_{\{\tau_x^i > t-\}} d\mathbb{1}_{\{\tau_x^i \le t\}} - \mathbb{1}_{\{\tau_x^i > t-\}} \mathbb{1}_{\{\tau_x^i > t\}} \mu(t) dt \right) = -e^{\Gamma(t)} dM_i^N(t).$$

Moreover, $[L_i, U](t) = 0$ for every $t \in [0, T^*]$ (Bielecki and Rutkowski, 2004, p. 160). Thus, using the martingale representation of $U(t)$, Equation (A.1) becomes

$$Z_i = \mathrm{E}^{\mathbb{P}}(Z_i) + \sum_{j=1}^{d} \int_0^{T^*} \left[L_i(t-) \mathbb{1}_{[0,T]}(t) + L_i(T) \mathbb{1}_{(T,T^*]}(t) \right] \varphi_j(t) dW_j(t) - \int_0^T U(t) e^{\Gamma(t)} dM_i^N(t).$$

Together with the continuity and adaptedness of μ, this proves the statement of the proposition for any single policyholder.

In the portfolio case, where $\mathbb{F} = \mathbb{G} \vee \bigvee_{i=1}^m \mathbb{I}^i$, the conditional independence of the τ_x^i's implies that $\mathrm{E}^{\mathbb{P}}\left(Z_i \middle| \mathcal{F}_t \right) = \mathrm{E}^{\mathbb{P}}\left(Z_i \middle| \mathcal{G}_t \vee \mathcal{I}_t^i \right)$. Thus, by using the conditionally identical distribution of τ_x^i, $i = 1, \ldots, m$, the proposition follows for the entire portfolio from applying the previous part of the proof to each summand of $Z = \sum_{i=1}^m \mathbb{1}_{\{\tau_x^i > T\}} F$ separately and adding the respective decompositions. □

Proof of Lemma 4.6. Note that by the martingale representation theorem, there exist predictable processes $\varphi_1, \ldots, \varphi_d$ such that (4.6) holds. Again, we first show the statement for a single policyholder with remaining lifetime τ_x^i, i.e. $m = 1$ and $\mathbb{F} = \mathbb{G} \vee \mathbb{I}^i$ for an arbitrary but fixed $i \in \{1, \ldots, m\}$. Since F is assumed to be $\mathbb{G}$-predictable with $\mathrm{E}^{\mathbb{P}}\left(\sup_{t \in [0,T]} |F(t)| \right) < \infty$, it follows from Proposition 5.1.2 in Bielecki and Rutkowski (2004, p. 149) that

$$\begin{aligned}
&\mathrm{E}^{\mathbb{P}}\left(\int_0^T \mathbb{1}_{\{\tau_x^i > v\}} F(v) dv \middle| \mathcal{F}_t \right) \\
&= \int_0^t \mathbb{1}_{\{\tau_x^i > v\}} F(v) dv + L_i(t)\, \mathrm{E}^{\mathbb{P}}\left(\int_t^T e^{-\Gamma(v)} F(v) dv \middle| \mathcal{G}_t \right) \\
&= \int_0^t \mathbb{1}_{\{\tau_x^i > v\}} F(v) dv - L_i(t) \int_0^t e^{-\Gamma(v)} F(v) dv + L_i(t)\, \mathrm{E}^{\mathbb{P}}\left(\int_0^T e^{-\Gamma(v)} F(v) dv \middle| \mathcal{G}_t \right), \qquad \text{(A.2)}
\end{aligned}$$

where $L_i(t) := \mathbb{1}_{\{\tau_x^i > t\}} e^{\Gamma(t)}$. Note that Proposition 5.1.2 in Bielecki and Rutkowski (2004) actually requires $\int_0^T F(s) ds$ to be bounded. However, via dominated convergence it can be shown that the result still holds if F satisfies $\mathrm{E}^{\mathbb{P}}\left(\sup_{t \in [0,T]} |F(t)| \right) < \infty$ (Biagini et al. (2012, p. 22) already point out a possible relaxation to $\mathrm{E}^{\mathbb{P}}\left(\sup_{t \in [0,T]} |F(t)|^2 \right) < \infty$).
As in the proof of Lemma 4.5, it follows by applying integration by parts that

$$L_i(t) \int_0^t e^{-\Gamma(v)} F(v) dv = \int_0^t \mathbb{1}_{\{\tau_x^i > s-\}} F(s) ds - \int_0^t \left(\int_0^s e^{-\Gamma(v)} F(v) dv \right) e^{\Gamma(s)} dM_i^N(s),$$

and

$$\begin{aligned}
L_i(t)\, \mathrm{E}^{\mathbb{P}}\left(\int_0^T e^{-\Gamma(v)} F(v) dv \middle| \mathcal{G}_t \right) &= \mathrm{E}^{\mathbb{P}}\left(\int_0^T e^{-\Gamma(v)} F(v) dv \right) + \sum_{i=1}^{d} \int_0^t \mathbb{1}_{\{\tau_x^i > s-\}} e^{\Gamma(s)} \varphi_i(s) dW_i(s) \\
&\quad - \int_0^t \mathrm{E}^{\mathbb{P}}\left(\int_0^T e^{-\Gamma(v)} F(v) dv \middle| \mathcal{G}_s \right) e^{\Gamma(s)} dM_i^N(s)
\end{aligned}$$

where $M_i^N(t) := \mathbb{1}_{\{\tau_x^i \le t\}} - \int_0^t \mathbb{1}_{\{\tau_x^i > s-\}} \mu(s) ds$. Summing up the representations of all summands from (A.2) and using the $\mathcal{F}_T$-measurability of $\int_0^T \mathbb{1}_{\{\tau_x^i > v\}} F(v) dv$, we obtain

$$
\begin{aligned}
\int_0^T \mathbb{1}_{\{\tau_x^i > v\}} F(v) dv = \mathrm{E}^{\mathbb{P}} \left(\int_0^T e^{-\Gamma(v)} F(v) dv \right) + \sum_{i=1}^d \int_0^T \mathbb{1}_{\{\tau_x^i > t-\}} e^{\Gamma(t)} \varphi_i(t) dW_i(t) \\
- \int_0^T \mathrm{E}^{\mathbb{P}} \left(\int_t^T e^{\Gamma(t)-\Gamma(v)} F(v) dv \middle| \mathcal{G}_t \right) dM_i^N(t).
\end{aligned}
\tag{A.3}
$$

Since we assume that $\mathrm{E}^{\mathbb{P}} \left(\sup_{t \in [0,T]} |F(t)| \right) < \infty$, the theorem of Fubini-Tonelli together with the construction of τ_x^i implies that

$$
\mathrm{E}^{\mathbb{P}} \left(\int_0^T e^{-\Gamma(v)} F(v) dv \right) = \mathrm{E}^{\mathbb{P}} \left(\int_0^T \mathbb{1}_{\{\tau_x^i > v\}} F(v) dv \right),
$$

and the theorem of Fubini-Tonelli for conditional expectations yields

$$
\int_0^T \mathrm{E}^{\mathbb{P}} \left(\int_t^T e^{\Gamma(t)-\Gamma(v)} F(v) dv \middle| \mathcal{G}_t \right) dM_i^N(t) = \int_0^T \int_t^T \mathrm{E}^{\mathbb{P}} \left(e^{\Gamma(t)-\Gamma(v)} F(v) \middle| \mathcal{G}_t \right) dv \, dM_i^N(t),
$$

so that (4.7) follows from (A.3).

Next we prove expression (4.8). By the martingale representation theorem, there exist for every $v \in [0,T]$ $\mathbb{G}$-predictable processes $\varphi_1^v, \ldots, \varphi_d^v$ such that

$$
\mathrm{E}^{\mathbb{P}} \left(e^{-\Gamma(v)} F(v) \middle| \mathcal{G}_t \right) = \mathrm{E}^{\mathbb{P}} \left(e^{-\Gamma(v)} F(v) \right) + \sum_{i=1}^d \int_0^t \varphi_i^v(u) \mathbb{1}_{[0,v]}(u) dW_i(u), \;\; t \in [0,T].
$$

Thus, using the theorem of Fubini-Tonelli and the stochastic Fubini theorem (Protter, 2005, Theorem 65), it follows that

$$
\begin{aligned}
\mathrm{E}^{\mathbb{P}} \left(\int_0^T e^{-\Gamma(v)} F(v) dv \middle| \mathcal{G}_t \right) &= \int_0^T \mathrm{E}^{\mathbb{P}} \left(e^{-\Gamma(v)} F(v) \middle| \mathcal{G}_t \right) dv \\
&= \int_0^T \mathrm{E}^{\mathbb{P}} \left(e^{-\Gamma(v)} F(v) \right) dv + \sum_{i=1}^d \int_0^T \int_0^t \varphi_i^v(u) \mathbb{1}_{[0,v]}(u) dW_i(u) \, dv \\
&= \mathrm{E}^{\mathbb{P}} \left(\int_0^T e^{-\Gamma(v)} F(v) dv \right) + \sum_{i=1}^d \int_0^t \int_0^T \varphi_i^v(u) \mathbb{1}_{[0,v]}(u) dv \, dW_i(u),
\end{aligned}
$$

where for applying the stochastic Fubini theorem it is sufficient to note that

$$
\begin{aligned}
&\mathrm{E}^{\mathbb{P}} \left(\int_0^t \int_0^T [\varphi_i^v(u)]^2 \, \mathbb{1}_{[0,v]}(u) dv \, du \right) \le \int_0^T \mathrm{E}^{\mathbb{P}} \left(\int_0^T [\varphi_i^v(u)]^2 \, \mathbb{1}_{[0,v]}(u) du \right) dv \\
&\le T \sup_{v \in [0,T]} \mathrm{E}^{\mathbb{P}} \left(\int_0^T [\varphi_i^v(u)]^2 \, \mathbb{1}_{[0,v]}(u) du \right) \le T \sup_{v \in [0,T]} \mathrm{E}^{\mathbb{P}} \left(\left[e^{-\Gamma(v)} F(v) \right]^2 \right) \\
&\le T \sup_{v \in [0,T]} \mathrm{E}^{\mathbb{P}} \left([F(v)]^2 \right) < \infty.
\end{aligned}
$$

The uniqueness of the martingale representation finally implies (4.8).

By the conditional independence assumption on $\tau_x^i, \; i = 1, \ldots, m$, we have in the portfolio case with $\mathbb{F} = \mathbb{G} \vee \bigvee_{i=1}^m \mathbb{I}^i$ that $\mathrm{E}^{\mathbb{P}} \left(\int_0^T \mathbb{1}_{\{\tau_x^i > v\}} F(v) dv \middle| \mathcal{F}_t \right) = \mathrm{E}^{\mathbb{P}} \left(\int_0^T \mathbb{1}_{\{\tau_x^i > v\}} F(v) dv \middle| \mathcal{G}_t \vee \mathcal{I}_t^i \right)$. Thus, the statement for the portfolio directly follows by applying the obtained equation to each summand $\int_0^T \mathbb{1}_{\{\tau_x^i > v\}} F(v) dv, \; i = 1, \ldots, m$, separately and adding the respective decompositions. $\square$

Proof of Lemma 4.7. Note that by the martingale representation theorem, there exist predictable processes $\varphi_1, \dots, \varphi_d$ such that (4.9) holds. Since F is continuous, it follows from the definition of Lebesgue integrals that

$$Z = \int_0^T F(v)dN(v) = \sum_{i=1}^m \mathbb{1}_{\{\tau_x^i \leq T\}} F(\tau_x^i). \tag{A.4}$$

Again, we first show the statement for a single policyholder with remaining lifetime τ_x^i, i.e. $m = 1$ and $\mathbb{F} = \mathbb{G} \vee \mathbb{I}^i$ for an arbitrary but fixed $i \in \{1, \dots, m\}$. Note that $\mathbb{1}_{\{\tau_x^i \leq t\}} F(\tau_x^i)$ is $\mathcal{F}_t$-measurable, so that

$$\mathrm{E}^{\mathbb{P}}\left(\mathbb{1}_{\{\tau_x^i \leq T\}} F(\tau_x^i) \middle| \mathcal{F}_t\right) = \mathrm{E}^{\mathbb{P}}\left(\mathbb{1}_{\{t < \tau_x^i \leq T\}} F(\tau_x^i) \middle| \mathcal{F}_t\right) + \mathbb{1}_{\{\tau_x^i \leq t\}} F(\tau_x^i). \tag{A.5}$$

Since F is assumed to be $\mathbb{G}$-predictable with $\mathrm{E}^{\mathbb{P}}\left(\sup_{t\in[0,T]} |F(t)|\right) < \infty$, it follows from Corollary 5.1.3 in Bielecki and Rutkowski (2004, p. 148) that

$$\begin{aligned}
\mathrm{E}^{\mathbb{P}}\left(\mathbb{1}_{\{t < \tau_x^i \leq T\}} F(\tau_x^i) \middle| \mathcal{F}_t\right) &= \mathbb{1}_{\{\tau_x^i > t\}} \mathrm{E}^{\mathbb{P}}\left(\int_t^T e^{\Gamma(t)-\Gamma(v)} F(v) d\Gamma(v) \middle| \mathcal{G}_t\right) \\
&= L_i(t) \mathrm{E}^{\mathbb{P}}\left(\int_0^T e^{-\Gamma(v)} F(v) d\Gamma(v) \middle| \mathcal{G}_t\right) - L_i(t) \int_0^t e^{-\Gamma(v)} F(v) d\Gamma(v),
\end{aligned}$$

where $L_i(t) := \mathbb{1}_{\{\tau_x^i > t\}} e^{\Gamma(t)}$. Again, Proposition 5.1.1 and thus Corollary 5.1.3 in Bielecki and Rutkowski (2004, p. 148) actually require F to be bounded, but a generalization to non-bounded F satisfying $\mathrm{E}^{\mathbb{P}}\left(\sup_{t\in[0,T]} |F(t)|\right) < \infty$ can be shown via dominated convergence (Biagini et al. (2012, p. 19) already point out a possible relaxation to $\mathrm{E}^{\mathbb{P}}\left(\sup_{t\in[0,T]} |F(t)|^2\right) < \infty$). As in the proof of Lemma 4.5, it then follows by applying integration by parts to both addends that

$$\begin{aligned}
&\mathrm{E}^{\mathbb{P}}\left(\mathbb{1}_{\{t < \tau_x^i \leq T\}} F(\tau_x^i) \middle| \mathcal{F}_t\right) \\
&= \mathrm{E}^{\mathbb{P}}\left(\int_0^T e^{-\Gamma(v)} F(v) d\Gamma(v)\right) + \sum_{i=1}^d \int_0^t \mathbb{1}_{\{\tau_x^i > s-\}} e^{\Gamma(s)} \varphi_i(s) dW_i(s) \\
&\quad - \int_0^t \mathrm{E}^{\mathbb{P}}\left(\int_s^T e^{\Gamma(s)-\Gamma(v)} F(v) d\Gamma(v) \middle| \mathcal{G}_s\right) dM_i^N(s) - \int_0^t \mathbb{1}_{\{\tau_x^i > s\}} F(s) d\Gamma(s),
\end{aligned}$$

where $M_i^N(t) := \mathbb{1}_{\{\tau_x^i \leq t\}} - \int_0^t \mathbb{1}_{\{\tau_x^i > s-\}} \mu(s) ds$. On the other hand, we obtain by (A.4) that

$$\mathbb{1}_{\{\tau_x^i \leq t\}} F(\tau_x^i) = \int_0^t F(s) d\mathbb{1}_{\{\tau_x^i \leq s\}} = \int_0^t F(s) dM_i^N(s) + \int_0^t F(s) \mathbb{1}_{\{\tau_x^i > s\}} d\Gamma(s).$$

Summing up the representations of the two summands from Equation (A.5) and using the $\mathcal{F}_T$-measurability of $\mathbb{1}_{\{\tau_x^i \leq T\}} F(\tau_x^i)$, we obtain

$$\begin{aligned}
\mathbb{1}_{\{\tau_x^i \leq T\}} F(\tau_x^i) = \mathrm{E}^{\mathbb{P}}\left(\int_0^T e^{-\Gamma(v)} F(v) d\Gamma(v)\right) + \sum_{i=1}^d \int_0^T \mathbb{1}_{\{\tau_x^i > t-\}} e^{\Gamma(t)} \varphi_i(t) dW_i(t) \\
- \int_0^T \left[\mathrm{E}^{\mathbb{P}}\left(\int_t^T e^{\Gamma(t)-\Gamma(v)} F(v) d\Gamma(v) \middle| \mathcal{G}_t\right) - F(t)\right] dM_i^N(t).
\end{aligned} \tag{A.6}$$

Corollary 5.1.3 in Bielecki and Rutkowski (2004) implies that

$$\mathrm{E}^{\mathbb{P}}\left(\int_0^T e^{-\Gamma(v)} F(v) d\Gamma(v)\right) = \mathrm{E}^{\mathbb{P}}\left(\mathbb{1}_{\{\tau_x^i \leq T\}} F(\tau_x^i)\right),$$

and since $\mathrm{E}^{\mathbb{P}}\left(\int_t^T |F(v)|\, e^{\Gamma(t)-\Gamma(v)}\mu(v)dv\right) \leq \mathrm{E}^{\mathbb{P}}\left(\sup_{v\in[0,T]}|F(v)|\right) < \infty$, the theorem of Fubini-Tonelli for conditional expectations yields

$$\mathrm{E}^{\mathbb{P}}\left(\int_t^T F(v)e^{\Gamma(t)-\Gamma(v)}d\Gamma(v)\middle|\,\mathcal{G}_t\right) = \int_t^T \mathrm{E}^{\mathbb{P}}\left(F(v)e^{\Gamma(t)-\Gamma(v)}\mu(v)\middle|\,\mathcal{G}_t\right)dv,$$

so that (4.10) follows from (A.6). The proof of expression (4.11) works analogously to the proof of (4.8), additionally using the Cauchy-Schwarz inequality.

By the conditional independence assumption on $\tau_x^i,\ i = 1,\ldots,m$, we have in the portfolio case with $\mathbb{F} = \mathbb{G} \vee \bigvee_{i=1}^m \mathbb{I}^i$ that $\mathrm{E}^{\mathbb{P}}\left(\mathbb{1}_{\{\tau_x^i\leq T\}}F(\tau_x^i)\middle|\,\mathcal{F}_t\right) = \mathrm{E}^{\mathbb{P}}\left(\mathbb{1}_{\{\tau_x^i\leq T\}}F(\tau_x^i)\middle|\,\mathcal{G}_t \vee \mathcal{I}_t^i\right)$. Thus, the statement for the portfolio directly follows by applying the obtained equation to each summand $\mathbb{1}_{\{\tau_x^i\leq t\}}F(\tau_x^i),\ i = 1,\ldots,m$, separately and adding the respective decompositions. □

Proof of Proposition 4.12. Since $n = d$, $\det\sigma(t,X(t)) \neq 0$ for all $t \in [0,T^*]$ $\mathbb{P}$-almost surely, and each L_0 is square integrable as a result of the respective assumptions, the uniqueness of the decompositions follows by Proposition 4.1. Furthermore, Assumption 4.11 implies that X is a Markov process, which together with the factorization lemma yields for all cases i) to iv) below that

$$\mathrm{E}^{\mathbb{P}}\left(\cdot|\,\mathcal{G}_t\right) = \mathrm{E}^{\mathbb{P}}\left(\cdot|\,X(t)\right) \tag{A.7}$$

is a function of $X(t)$. Define $G(t) := \int_0^t g(s,X(s))ds,\ 0 \leq t \leq T$, and note that (Shreve, 2004, p. 480)

$$d[G,G](t) = d[G,\Gamma](t) = d[\Gamma,\Gamma](t) = d[G,X_i](t) = d[\Gamma,X_i](t) = 0. \tag{A.8}$$

i) The assumption on the form of C_0 together with (A.7) yields that

$$\begin{aligned}\mathrm{E}^{\mathbb{P}}\left(C_0|\,\mathcal{G}_t\right) &= e^{-G(t)}\mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^T g(s,X(s))ds}h(X(T))\middle|\,\mathcal{G}_t\right) = e^{-G(t)}f(t,X(t))\\ &=: \tilde{f}(t,G(t),X(t)).\end{aligned}$$

Since f is assumed to be smooth, this holds for $\tilde{f}$ as well. Thus, Itô's formula yields for $0 \leq t \leq T$ (Theorem 33 in Protter, 2005, p. 81)

$$\mathrm{E}^{\mathbb{P}}\left(C_0|\,\mathcal{G}_t\right) - \mathrm{E}^{\mathbb{P}}\left(C_0\right) = \sum_{i=1}^n \int_0^t e^{-G(s)}\frac{\partial f}{\partial x_i}(s,X(s))dM_i^W(s) + \int_0^t a(s)ds,$$

where $a = (a(t))_{0\leq t\leq T^*}$ is short-hand for all ds-quantities. We have used (A.8) and that $(t,G(t),X(t))$ has continuous paths. The right-hand side $\mathrm{E}^{\mathbb{P}}\left(C_0|\,\mathcal{G}_t\right) - \mathrm{E}^{\mathbb{P}}\left(C_0\right)$ is a martingale. On the other hand, the stochastic integrals with respect to $M_i^W,\ i = 1,\ldots,n$, are martingales as well. Thus, it follows by the uniqueness of the Doob-Meyer decomposition (Theorem 16 in Protter, 2005, p. 116) that the ds-term vanishes. Since C_0 is $\mathcal{G}_T$-measurable, the statement follows.

ii) In both cases, $T > t_k$ and $T \leq t_k$, we derive the MRT decomposition with the help of Lemma 4.5. Thus, we mainly need to determine the MRT decomposition of $e^{-\Gamma(t_k)}C_{a,k}$ less its expectation.

(a) If $T > t_k$, we consider the decomposition

$$\begin{aligned}&e^{-\Gamma(t_k)}C_{a,k} - \mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}C_{a,k}\right)\\ &= \left[e^{-\Gamma(t_k)}\mathrm{E}^{\mathbb{P}}\left(C_{a,k}|\,\mathcal{G}_{t_k}\right) - \mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}C_{a,k}\right)\right] + e^{-\Gamma(t_k)}\left[C_{a,k} - \mathrm{E}^{\mathbb{P}}\left(C_{a,k}|\,\mathcal{G}_{t_k}\right)\right],\end{aligned} \tag{A.9}$$

and separately derive the MRT decompositions of the two parts.

The assumption on the form of $C_{a,k}$ together with (A.7) yield for $0 \le t \le t_k$ that

$$\begin{aligned}\mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}\mathrm{E}^{\mathbb{P}}\left(C_{a,k}\middle|\,\mathcal{G}_{t_k}\right)\middle|\,\mathcal{G}_t\right) &= e^{-\Gamma(t)}e^{-G(t)}\mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(t_k)}e^{G(t)-G(T)}h(X(T))\middle|\,\mathcal{G}_t\right)\\ &= e^{-\Gamma(t)}e^{-G(t)}f^A(t,X(t))\\ &=: \tilde{f}^A(t,\Gamma(t),G(t),X(t)).\end{aligned}$$

Since f^A is assumed to be smooth, this holds for $\tilde{f}^A$ as well. Thus, Itô's formula yields for $0 \le t \le t_k$ (Theorem 33 in Protter, 2005, p. 81)

$$\begin{aligned}&\mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}\mathrm{E}^{\mathbb{P}}\left(C_{a,k}\middle|\,\mathcal{G}_{t_k}\right)\middle|\,\mathcal{G}_t\right) - \mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}\mathrm{E}^{\mathbb{P}}\left(C_{a,k}\middle|\,\mathcal{G}_{t_k}\right)\right)\\ &= \sum_{i=1}^n \int_0^t e^{-\Gamma(s)}e^{-G(s)}\frac{\partial f^A}{\partial x_i}(s,X(s))dM_i^W(s) + \int_0^t a(s)ds,\end{aligned}$$

where $a = (a(t))_{0\le t\le T^*}$ is short-hand for all ds-quantities. We have used (A.8) and that $(t,\Gamma(t),G(t),X(t))$ has continuous paths. By the same arguments as in i) the ds-term vanishes, and since $e^{-\Gamma(t_k)}\mathrm{E}^{\mathbb{P}}\left(C_{a,k}\middle|\,\mathcal{G}_{t_k}\right)$ is $\mathcal{G}_{t_k}$-measurable, it follows that

$$e^{-\Gamma(t_k)}\mathrm{E}^{\mathbb{P}}\left(C_{a,k}\middle|\,\mathcal{G}_{t_k}\right) - \mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}C_{a,k}\right) = \sum_{i=1}^n \int_0^{t_k} e^{-\Gamma(s)}e^{-G(s)}\frac{\partial f^A}{\partial x_i}(s,X(s))dM_i^W(s).$$

Furthermore, applying part i) to $C_{a,k}$ it holds that

$$e^{-\Gamma(t_k)}\left[C_{a,k} - \mathrm{E}^{\mathbb{P}}\left(C_{a,k}\middle|\,\mathcal{G}_{t_k}\right)\right] = \sum_{i=1}^n \int_{t_k}^T e^{-\Gamma(t_k)}e^{-G(s)}\frac{\partial f^B}{\partial x_i}(s,X(s))dM_i^W(s).$$

In total, by (A.9) we have

$$\begin{aligned}&e^{-\Gamma(t_k)}C_{a,k} - \mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}C_{a,k}\right)\\ &= \sum_{i=1}^n \int_0^T \Big[e^{-\Gamma(s)}e^{-G(s)}\frac{\partial f^A}{\partial x_i}(s,X(s))\mathbb{1}_{[0,t_k]}(s)\\ &\qquad\qquad + e^{-\Gamma(t_k)}e^{-G(s)}\frac{\partial f^B}{\partial x_i}(s,X(s))\mathbb{1}_{(t_k,T]}(s)\Big]dM_i^W(s).\end{aligned}$$

The statement then follows by Lemma 4.5 using the equality

$$\begin{aligned}&\sum_{i=1}^n \int_0^t \tilde{\varphi}_i(u)dM_i^W(u) = \int_0^t \tilde{\varphi}(u)dM^W(u) = \int_0^t \tilde{\varphi}(u)\sigma(u)dW(u)\\ &= \sum_{j=1}^d \int_0^t (\tilde{\varphi}(u)\sigma(u))_j dW_j(u),\end{aligned} \tag{A.10}$$

where $\tilde{\varphi} = (\tilde{\varphi}_1,\dots,\tilde{\varphi}_n)$ is any vector, $M^W = (M_1^W,\dots,M_n^W)$, and $(\cdot)_j$ denotes the j-th component of a vector.

(b) If $T \le t_k$, we consider the decomposition

$$\begin{aligned}&e^{-\Gamma(t_k)}C_{a,k} - \mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}C_{a,k}\right)\\ &= \left[\mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}\middle|\,\mathcal{G}_T\right)C_{a,k} - \mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}C_{a,k}\right)\right] + \left[e^{-\Gamma(t_k)} - \mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}\middle|\,\mathcal{G}_T\right)\right]C_{a,k}\end{aligned}$$

and again separately derive the MRT decompositions of the two parts. Analogously to above, we obtain

$$\mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}\middle|\,\mathcal{G}_T\right)C_{a,k}-\mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}C_{a,k}\right)=\sum_{i=1}^{n}\int_0^T e^{-\Gamma(s)}e^{-G(s)}\frac{\partial f^A}{\partial x_i}(s,X(s))dM_i^W(s),$$

and

$$\left[e^{-\Gamma(t_k)}-\mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}\middle|\,\mathcal{G}_T\right)\right]C_{a,k}=\sum_{i=1}^{n}\int_T^{t_k} e^{-\Gamma(s)}C_{a,k}\frac{\partial f^B}{\partial x_i}(s,X(s))dM_i^W(s),$$

so that

$$\begin{aligned}
&e^{-\Gamma(t_k)}C_{a,k}-\mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(t_k)}C_{a,k}\right)\\
&=\sum_{i=1}^{n}\int_0^{t_k}\Big[e^{-\Gamma(s)}e^{-G(s)}\frac{\partial f^A}{\partial x_i}(s,X(s))\mathbb{1}_{[0,T]}(s)\\
&\qquad\qquad+e^{-\Gamma(s)}C_{a,k}\frac{\partial f^B}{\partial x_i}(s,X(s))\mathbb{1}_{(T,t_k]}(s)\Big]dM_i^W(s).
\end{aligned}$$

The statement then follows by Lemma 4.5 using (A.10).

iii) The assumption on the form of $C_a(v)$ together with (A.7) yield that, for each $v\in[0,T]$,

$$\begin{aligned}
\mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(v)}C_a(v)\middle|\,\mathcal{G}_t\right)&=e^{-\Gamma(t)}e^{-G(t)}\mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^v[g(s,X(s))+\mu(s,X(s))]ds}h(X(v))\middle|\,\mathcal{G}_t\right)\\
&=e^{-\Gamma(t)}e^{-G(t)}f^v(t,X(t))\\
&=:\tilde{f}^v(t,\Gamma(t),G(t),X(t)),\;\; t\le v,
\end{aligned}$$

where $f^v:[0,v]\times\mathbb{R}^n\to\mathbb{R}$. Since f^v is assumed to be smooth, this holds for $\tilde{f}^v$ as well. Thus, Itô's formula yields for $t\le v$ (Protter, 2005, p. 81)

$$\begin{aligned}
&\mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(v)}C_a(v)\middle|\,\mathcal{G}_t\right)-\mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(v)}C_a(v)\right)\\
&=\sum_{i=1}^{n}\int_0^t e^{-\Gamma(s)}e^{-G(s)}\frac{\partial f^v}{\partial x_i}(s,X(s))dM_i^W(s)+\int_0^t a(s)ds,
\end{aligned}$$

where $a=(a(t))_{0\le t\le T^*}$ is short-hand for all ds-quantities. We have used (A.8) and that $(t,\Gamma(t),G(t),X(t))$ has continuous paths. By the same arguments as in i), the ds-term has to vanish. Thus, we obtain by Lemma 4.6 for $t\in[0,T]$ (for $t>T$ all integrands are zero) that

$$\begin{aligned}
\psi_i^W(t)&=(m-N(t-))e^{\Gamma(t)}\int_t^T\varphi_i^v(t)dv\\
&=(m-N(t-))e^{-G(t)}\int_t^T\frac{\partial f^v}{\partial x_i}(t,X(t))dv,\;\; i=1,\dots,n,
\end{aligned}$$

and

$$\begin{aligned}
\psi^N(t)&=-\int_t^T\mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(v)}C_a(v)\middle|\,\mathcal{G}_t\right)dv\\
&=-\int_t^T e^{-G(t)}\mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^v[g(s,X(s))+\mu(s,X(s))]ds}h(X(v))\middle|\,\mathcal{G}_t\right)dv\\
&=-e^{-\int_0^t g(s,X(s))ds}\int_t^T f^v(t,X(t))dv.
\end{aligned}$$

iv) As in part iii), the assumption on the form of $C_{ad}(t)$ and $\mu(t)$ together with (A.7) yield that, for each $v \in [0,T]$,

$$
\begin{aligned}
&\mathrm{E}^{\mathbb{P}}\left(e^{-\Gamma(v)}C_{ad}(v)\mu(v)\middle|\,\mathcal{G}_t\right)\\
&= e^{-\Gamma(t)}e^{-G(t)}\mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^v[g(s,X(s))+\mu(s,X(s))]ds}h(X(v))\mu(v,X(v))\middle|\,\mathcal{G}_t\right)\\
&= e^{-\Gamma(t)}e^{-G(t)}f^v(t,X(t))\\
&=: \tilde{f}^v(t,\Gamma(t),G(t),X(t)),\;\; t\le v,
\end{aligned}
$$

where $f^v : [0,v]\times\mathbb{R}^n \to \mathbb{R}$. Thus, the integrands $\psi_i^W(t)$, $i = 1,\dots,n$, of part iv) follow analogously to part iii) using Lemma 4.7 instead of Lemma 4.6. Lemma 4.7 also yields for $t \le T$ (otherwise it is equal to zero) that

$$
\begin{aligned}
\psi^N(t) &= -\left[\int_t^T \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(v)}C_{ad}(v)\mu(v)\middle|\,\mathcal{G}_t\right)dv - C_{ad}(t)\right]\\
&= -\left[\int_t^T e^{-G(t)}\mathrm{E}^{\mathbb{P}}\left(e^{-\int_t^v[g(s,X(s))+\mu(s,X(s))]ds}h(X(v))\mu(v)\middle|\,\mathcal{G}_t\right)dv - C_{ad}(t)\right]\\
&= -\left[e^{-\int_0^t g(s,X(s))ds}\int_t^T f^v(t,X(t))dv - C_{ad}(t)\right].
\end{aligned}
$$

□

Proof of Proposition 4.16. Note that any conditional expectation $\mathrm{E}^{\mathbb{P}}(\cdot|\mathcal{G}_t)$ is predictable, since it is by definition $\mathcal{G}_t$-measurable and $\mathcal{G}_t$ is left-continuous as a result of the continuity of Brownian motions.

i) Define $\psi_{ak}^N(t) := \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(t_k)}C_{a,k}\middle|\,\mathcal{G}_t\right)$ for all $t\in[0,t_k]$, and the process $(\psi_{ak}^N(t))_{0\le t\le t_k}$ is predictable. Furthermore, applying Jensen's inequality for conditional expectations (Protter, 2005, p. 11), and using that $\Gamma(t)$ is non-decreasing in t, it follows that

$$
\begin{aligned}
&\sup_{t\in[0,t_k]}\mathrm{E}^{\mathbb{P}}\left(\left[\psi_{ak}^N(t)\right]^4\right) = \sup_{t\in[0,t_k]}\mathrm{E}^{\mathbb{P}}\left(\left[\mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(t_k)}C_{a,k}\middle|\,\mathcal{G}_t\right)\right]^4\right)\\
&\le \sup_{t\in[0,t_k]}\mathrm{E}^{\mathbb{P}}\left(\mathrm{E}^{\mathbb{P}}\left(\left[e^{\Gamma(t)-\Gamma(t_k)}C_{a,k}\right]^4\middle|\,\mathcal{G}_t\right)\right) \le \sup_{t\in[0,t_k]}\mathrm{E}^{\mathbb{P}}\left(\mathrm{E}^{\mathbb{P}}\left([C_{a,k}]^4\middle|\,\mathcal{G}_t\right)\right)\\
&= \sup_{t\in[0,t_k]}\mathrm{E}^{\mathbb{P}}\left([C_{a,k}]^4\right) = \mathrm{E}^{\mathbb{P}}\left([C_{a,k}]^4\right) < \infty \quad \text{(by assumption).}
\end{aligned}
$$

Since we also assume that $\sup_{t\in[0,t_k]}\mathrm{E}^{\mathbb{P}}\left(\mu^2(t)\right) < \infty$, the statement follows by Lemma 4.15.

ii) Define $\psi_a^N(t) := \int_t^T \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(s)}C_a(s)\middle|\mathcal{G}_t\right)ds$ for all $t\in[0,T]$, and the process $(\psi_a^N(t))_{0\le t\le T}$ is predictable. Furthermore, since $0\le e^{\Gamma(t)-\Gamma(s)}\le 1$ for $s\ge t$ and since $C := \sup_{t\in[0,T]}\mathrm{E}^{\mathbb{P}}\left(|C_a(t)|\right) < \infty$ as a result of the boundedness of $C_a(t)$, it follows by applying Jensen's inequality for integrals and for conditional expectations (for the latter, cf. Protter, 2005, p. 11) that for any $t\in[0,T]$

$$
\begin{aligned}
\left|\psi_a^N(t)\right| &= \left|\int_t^T \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(s)}C_a(s)\middle|\mathcal{G}_t\right)ds\right| \le \int_t^T\left|\mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(s)}C_a(s)\middle|\mathcal{G}_t\right)\right|ds\\
&\le \int_t^T \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(s)}\left|C_a(s)\right|\middle|\mathcal{G}_t\right)ds \le C\,T.
\end{aligned}
$$

Thus, we have

$$
\sup_{t\in[0,T]}\mathrm{E}^{\mathbb{P}}\left(\left[\psi_a^N(t)\right]^4\right) \le \sup_{t\in[0,T]}\mathrm{E}^{\mathbb{P}}\left([CT]^4\right) = C^4T^4 < \infty.
$$

Since we also assume that $\sup_{t\in[0,T]}\mathrm{E}^{\mathbb{P}}\left(\mu^2(t)\right) < \infty$, the statement follows by Lemma 4.15.

iii) Since $X_m, Y_m, X, Y \in L^2(\mathbb{P})$ and $X_m \xrightarrow{L^2} X,\ Y_m \xrightarrow{L^2} Y$ implies that $X_m + Y_m \xrightarrow{L^2} X + Y$, it is sufficient to show that

a) $$\frac{1}{m}\int_0^T \left[-\int_t^T \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(s)} C_{ad}(s)\mu(s)\middle|\, \mathcal{G}_t\right) ds\right] dM^N(t) \xrightarrow[m\to\infty]{L^2} 0, \quad \text{and}$$

b) $$\frac{1}{m}\int_0^T C_{ad}(t) dM^N(t) \xrightarrow[m\to\infty]{L^2} 0.$$

Define $\psi^N_{ad,1}(t) := -\int_t^T \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(s)} C_{ad}(s)\mu(s) ds\middle|\, \mathcal{G}_t\right)$ and $\psi^N_{ad,2}(t) := C_{ad}(t)$ for all $t \in [0,T]$. Note that since by assumption $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\mu^4(t)\right) < \infty$, it also follows by Jensen's inequality that

$$\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\mu^2(t)\right) \le \sup_{t\in[0,T]} \sqrt{\mathrm{E}^{\mathbb{P}}\left(\mu^4(t)\right)} < \infty.$$

ad a): Since the process $(\psi^N_{ad,1}(t))_{0\le t\le T}$ is predictable, since $0 \le e^{\Gamma(t)-\Gamma(s)} \le 1$ for $s \ge t$, and since $C_1 := \sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(|C_{ad}(t)|\right) < \infty$ as a result of the boundedness of C_{ad}, it follows by applying Jensen's inequality for integrals and for conditional expectations (for the latter, cf. Protter, 2005, p. 11) that

$$\begin{aligned}
\left|\psi^N_{ad,1}(t)\right| &= \left|\int_t^T \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(s)} C_{ad}(s)\mu(s)\middle|\, \mathcal{G}_t\right) ds\right| \\
&\le \int_t^T \mathrm{E}^{\mathbb{P}}\left(e^{\Gamma(t)-\Gamma(s)} \left|C_{ad}(s)\right| \mu(s)\middle|\, \mathcal{G}_t\right) ds \\
&\le C_1 \int_t^T \mathrm{E}^{\mathbb{P}}\left(\mu(s)\middle|\, \mathcal{G}_t\right) ds \le C_1 \int_0^T \mathrm{E}^{\mathbb{P}}\left(\mu(s)\middle|\, \mathcal{G}_t\right) ds.
\end{aligned}$$

Since by assumption $C_2 := \sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\mu^4(t)\right) < \infty$, this implies

$$\begin{aligned}
\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\left[\psi^N_{ad,1}(t)\right]^4\right) &\le \sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\left[C_1 \int_0^T \mathrm{E}^{\mathbb{P}}\left(\mu(s)\middle|\, \mathcal{G}_t\right) ds\right]^4\right) \\
&\overset{(*)}{\le} \sup_{t\in[0,T]} C_1^4\, \mathrm{E}^{\mathbb{P}}\left(\int_0^T \mathrm{E}^{\mathbb{P}}\left(\mu^4(s)\middle|\, \mathcal{G}_t\right) ds\right) \\
&\overset{(**)}{=} \sup_{t\in[0,T]} C_1^4 \int_0^T \mathrm{E}^{\mathbb{P}}\left(\mu^4(s)\right) ds \le C_1^4\, C_2\, T < \infty,
\end{aligned}$$

where $(*)$ again follows by Jensen's inequality for integrals and conditional expectations and $(**)$ from the theorem of Fubini-Tonelli. Since $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\mu^2(t)\right) < \infty$ as shown above, the statement follows by Lemma 4.15.

ad b): The process $(\psi^N_{ad,2}(t))_{0\le t\le T}$ is predictable. As a result of the boundedness of $C_{ad}(t)$, it also holds $C_1 := \sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(|C_{ad}(t)|\right) < \infty$, so that

$$\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\left[\psi^N_{ad,2}(t)\right]^4\right) = \sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\left[C_{ad}(t)\right]^4\right) \le C_1^4 < \infty.$$

Since $\sup_{t\in[0,T]} \mathrm{E}^{\mathbb{P}}\left(\mu^2(t)\right) < \infty$ as shown above, the statement follows by Lemma 4.15.

□

Proof of Proposition 4.19. Since $M_i^W(t) = \sum_{k=1}^d \int_0^t \sigma_{ik}(s)dW_k(s),\ 0 \le t \le T^*$, it follows that

$$R_{i,\cdot}^{(m)} = \sum_{k=1}^{d}\sum_{j=1}^{d}\int_0^T \left[(m - N(t-))e^{\Gamma(t)}\mathbb{1}_{[0,t_k]}(t) + (m - N(t_k))e^{\Gamma(t_k)}\mathbb{1}_{(t_k,T]}(t)\right] \times \varphi_{j,\cdot}(t)\sigma_{ji}^{-1}(t)\sigma_{ik}(t)dW_k(t). \tag{A.11}$$

Because of the additivity of integration and the continuous mapping theorem, it is sufficient to prove the convergence of each summand in (A.11), $i = 1, \dots, n,\ j, k = 1, \dots, d$, separately. For this, by Lemma 4.18, we only need to show that each $\varphi_{j,\cdot}(t)\sigma_{ji}^{-1}(t)\sigma_{ik}(t)$ is $\mathbb{G}$-predictable with $\int_0^T \left(\varphi_{j,\cdot}(t)\sigma_{ji}^{-1}(t)\sigma_{ik}(t)\right)^2 dt < \infty$ almost surely. We have:

- By assumption, $\sigma(t)$ is $\mathbb{G}$-adapted with continuous paths.
- When determining the inverse of $\sigma(t)$ with Cramer's rule and the necessary determinants with Laplace's formula, it can be seen that $\sigma_{ij}^{-1}(t)$ is a continuous function of the matrix components $\sigma_{ij}(t),\ i = 1, \dots n, j = 1 \dots, d$. So $\sigma_{ij}^{-1}(t)$ has itself continuous paths and is $\mathbb{G}$-adapted.
- In all parts ii), iii), and iv), $\varphi_{j,\cdot}(t)$ is a conditional expectation of the form $\mathrm{E}^{\mathbb{P}}\left(\cdot \,|\, \mathcal{G}_t\right)$ or can be transformed into such an expectation using the theorem of Fubini-Tonelli for conditional expecations. As a result, $\varphi_{j,\cdot}(t)$ is by definition $\mathbb{G}$-adapted.
- The $\mathbb{D}_{1,2}$-assumptions in Proposition 4.8 and particularly the implicit square integrability of the respective quantities yield that

$$\mathrm{E}\left(\int_0^T \varphi_{j,\cdot}(t)^2 dt\right) = \mathrm{E}\Bigg(\left(\int_0^T \varphi_{j,\cdot}(t)dW_j(t)\right)^2\Bigg) < \infty,$$

  implying that $\int_0^T \varphi_{j,\cdot}(t)^2 dt < \infty$ almost surely.

Since $\mathbb{G} = (\mathcal{G}_t)_{0\le t\le T}$ is left-continuous as a result of the continuity of Brownian motions, every $\mathbb{G}$-adapted process is also $\mathbb{G}$-predictable. Thus, the product $\varphi_{j,\cdot}(t)\sigma_{ji}^{-1}(t)\sigma_{ik}(t)$ is not only $\mathbb{G}$-adapted, but also $\mathbb{G}$-predictable. Furthermore, since $\sigma_{ji}^{-1}(t)\sigma_{ik}(t)$ has continuous paths and $\int_0^T \varphi_{j,\cdot}(t)^2 dt < \infty$ almost surely, it follows similarly as in the proof of Lemma 4.18 that $\int_0^T \left(\varphi_{j,\cdot}(t)\sigma_{ji}^{-1}(t)\sigma_{ik}(t)\right)^2 dt < \infty$ almost surely. The statement then directly follows by Lemma 4.18. □

Proof of Corollary 4.20. i) Since each $L_{0,\cdot}^{(m)}$ is the sum of m random variables which are conditionally identically distributed and conditionally independent given the σ-algebra $\mathcal{G}_{T^*}$, the statement follows by a conditional version of Kolmogorov's strong law of large numbers (Majerek et al., 2005, p. 154).

ii) Since $\mathrm{E}^{\mathbb{P}}(L_{0,\cdot}^{(m)}) = m\,\mathrm{E}^{\mathbb{P}}(L_{0,\cdot}^{(1)})$ as a result of the conditionally identical distribution of $\tau_x^k,\ k = 1, \dots, m$, it follows by i) that

$$\frac{1}{m}\left(L_{0,\cdot}^{(m)} - \mathrm{E}^{\mathbb{P}}(L_{0,\cdot}^{(m)})\right) \xrightarrow[m\to\infty]{P} \mathrm{E}^{\mathbb{P}}(L_{0,\cdot}^{(1)}|\mathcal{G}_{T^*}) - \mathrm{E}^{\mathbb{P}}(L_{0,\cdot}^{(1)}).$$

Furthermore, by Proposition 4.16 (note that L^2-convergence implies convergence in probability) and Proposition 4.19 we have

$$\frac{1}{m}\left(L_{0,\cdot}^{(m)} - \mathrm{E}^{\mathbb{P}}(L_{0,\cdot}^{(m)})\right) = \sum_{i=1}^{n+1}\frac{1}{m}R_{i,\cdot}^{(m)} \xrightarrow[m\to\infty]{P} \sum_{i=1}^{n}\int_0^T \sum_{j=1}^{d}\varphi_{j,\cdot}(t)\sigma_{ji}^{-1}(t)dM_i^W(t).$$

Since the limit in probability is almost surely unique, it follows that

$$\mathrm{E}^{\mathbb{P}}(L_{0,\cdot}^{(1)}|\mathcal{G}_{T^*}) - \mathrm{E}^{\mathbb{P}}(L_{0,\cdot}^{(1)}) = \sum_{i=1}^{n} \int_0^T \sum_{j=1}^{d} \varphi_{j,\cdot}(t)\sigma_{ji}^{-1}(t) dM_i^W(t),$$

which is an MRT decomposition. By the uniqueness of the MRT decomposition (Proposition 4.1), each risk factor R_i^* thus equals almost surely the limit in probability of $\frac{1}{m}R_{i,\cdot}^{(m)},\ i = 1, \ldots, n+1$, so that the statement follows.

□

References

Alòs, E. and Ewald, C.-O. (2008). Malliavin differentiability of the Heston volatility and applications to option pricing. *Advances in Applied Probability*, 40(1):144–162.

Andersen, P. K., Borgan, Ø., Gill, R. D., and Keiding, N. (1997). *Statistical models based on counting processes*. Springer Verlag, New York.

Artinger, H. (2010). Longevity Risk in the Pension Context. Diploma thesis, Technische Universität München.

Barbarin, J. (2008). *Valuation, hedging and the risk management of insurance contracts*. Doctoral thesis, Université catholique de Louvain.

Bauer, D., Kling, A., and Ruß, J. (2008). A universal pricing framework for guaranteed minimum benefits in variable annuities. *Astin Bulletin*, 38(2):621–651.

Bauer, D. and Zanjani, G. H. (2013). Capital Allocation and Its Discontents. In *Handbook of Insurance*, pages 863–880. Springer Verlag, New York.

Biagini, F., Rheinländer, T., and Schreiber, I. (2012). Risk-minimization for life insurance liabilities with basis risk. *Preprint, University of Munich and London School of Economics and Political Science*.

Biagini, F., Rheinländer, T., and Widenmann, J. (2013). Hedging mortality claims with longevity bonds. *Astin Bulletin*, 43(2):123–157.

Biagini, F. and Schreiber, I. (2013). Risk-minimization for life insurance liabilities. *SIAM Journal on Financial Mathematics*, 4(1):243–264.

Bielecki, T. and Rutkowski, M. (2004). *Credit Risk: Modeling, Valuation and Hedging*. Springer Verlag, Berlin.

Biffis, E. (2005). Affine processes for dynamic mortality and actuarial valuations. *Insurance: Mathematics and Economics*, 37(3):443–468.

Biffis, E., Denuit, M., and Devolder, P. (2010). Stochastic mortality under measure changes. *Scandinavian Actuarial Journal*, 2010(4):284–311.

Björk, T. (2011). An Introduction to Point Processes from a Martingale Point of View. Lecture notes, KTH.

Brigo, D. and Mercurio, F. (2006). *Interest Rate Models: Theory and Practice: With Smile, Inflation, and Credit*. Springer Verlag, Berlin.

Bühlmann, H. (1995). Life Insurance with Stochastic Interest Rates. In *Financial Risk in Insurance*, pages 1–24. Springer Verlag, Berlin.

CEIOPS (2010). QIS 5 Technical Specifications. Available online at `https://eiopa.europa.eu/fileadmin/tx_dam/files/consultations/QIS/QIS5/QIS5-technical_specifications_20100706.pdf`; accessed April 28th, 2014.

Christiansen, M. C. (2007). *A joint analysis of financial and biometrical risks in life insurance*. Doctoral thesis, Universität Rostock.

Christiansen, M. C. (2013). Safety margins for unsystematic biometric risk in life and health insurance. *Scandinavian Actuarial Journal*, 2013(4):286–323.

Christiansen, M. C. and Helwich, M. (2008). Some further ideas concerning the interaction between insurance and investment risks. *Blätter der DGVFM*, 29(2):253–266.

Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407.

Dahl, M., Melchior, M., and Møller, T. (2008). On systematic mortality risk and risk-minimization with survivor swaps. *Scandinavian Actuarial Journal*, 2008(2-3):114–146.

Dahl, M. and Møller, T. (2006). Valuation and hedging of life insurance liabilities with systematic mortality risk. *Insurance: Mathematics and Economics*, 39(2):193–217.

Di Nunno, G., Øksendal, B., and Proske, F. (2009). *Malliavin calculus for Lévy processes with applications to finance*. Springer Verlag, Berlin.

Fischer, T. (2004). On the decomposition of risk in life insurance. Working paper, Technische Universität Darmstadt.

Gatzert, N. and Wesker, H. (2014). Mortality Risk and Its Effect on Shortfall and Risk Management in Life Insurance. *Journal of Risk and Insurance*, 81(1):57–90.

Heath, D. and Schweizer, M. (2000). Martingales versus PDEs in finance: an equivalence result with examples. *Journal of Applied Probability*, 37(4):947–957.

Hoem, J. M. (1988). The versatility of the Markov chain as a tool in the mathematics of life insurance. In *Transactions of the 23rd International Congress of Actuaries*, pages 171–202.

Jeanblanc, M. and Rutkowski, M. (2000). Modelling of default risk: an overview. In *Mathematical Finance: Theory and Practice*, pages 171–269. Higher Education Press, Beijing.

Karabey, U. (2012). *Risk Capital Allocation and Risk Quantification in Insurance Companies*. Doctoral thesis, Heriot-Watt University.

Karabey, U., Kleinow, T., and Cairns, A. J. (2014). Factor risk quantification in annuity models. *Insurance: Mathematics and Economics*, 58:34–45.

Klebaner, F. C. (2005). *Introduction to Stochastic Calculus with Applications*. Imperial College Press, London.

Kling, A., Ruß, J., and Schilling, K. (2014). Risk analysis of annuity conversion options in a stochastic mortality environment. *Astin Bulletin*, 44(2):197–236.

Kochanski, M. and Karnarski, B. (2011). Solvency capital requirement for hybrid products. *European Actuarial Journal*, 1(2):173–198.

Lando, D. (1998). On cox processes and credit risky securities. *Review of Derivatives Research*, 2(2-3):99–120.

Majerek, D., Nowak, W., and Zięba, W. (2005). Conditional strong law of large number. *International Journal of Pure and Applied Mathematics*, 20(2):143–157.

Martin, R. and Tasche, D. (2007). Shortfall: a tail of two parts. *Risk*, 2007(2):84–89.

Møller, T. (2001). Risk-minimizing hedging strategies for insurance payment processes. *Finance and Stochastics*, 5(4):419–446.

Niemeyer, A. (2015). Safety margins for systematic biometric and financial risk in a semi-Markov life insurance framework. *Risks*, 3(1):35–60.

Norberg, R. (2013). Optimal hedging of demographic risk in life insurance. *Finance and Stochastics*, 17(1):197–222.

Nualart, D. (2006). *The Malliavin calculus and related topics*. Springer Verlag, Berlin.

Protter, P. E. (2005). *Stochastic Integration and Differential Equations*. Springer Verlag, Berlin.

Rosen, D. and Saunders, D. (2010). Risk factor contributions in portfolio credit risk models. *Journal of Banking & Finance*, 34(2):336–349.

Schilling, K. (2015). Comparing financial and biometric risks in annuity conversion options via the MRT decomposition. Working paper, Universität Ulm.

Shreve, S. E. (2004). *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer Verlag, New York.