Article

Risk minimization for insurance products via $\mathbb{F}$-doubly stochastic Markov chains

Francesca Biagini [1,*], Andreas Groll [2] and Jan Widenmann [3]

[1] Department of Mathematics, University of Munich, Theresienstraße 39, 80333 Munich, Germany; francesca.biagini@math.lmu.de
[2] Department of Mathematics, University of Munich, Theresienstraße 39, 80333 Munich, Germany; groll@math.lmu.de
[3] BMW Financial Services, BMW Bank GmbH, 80787 Munich, Germany; jan-maximilian.widenmann@bmw.de
* Correspondence: francesca.biagini@math.lmu.de; Tel.: +49-89-2180-4492

Academic Editor: name
Version May 27, 2016 submitted to Risks; Typeset by LaTeX using class file mdpi.cls

Abstract: We study risk-minimization for a large class of insurance contracts. Given that the individual progress in time of visiting an insurance policy's states follows an $\mathbb{F}$-doubly stochastic Markov chain, we describe different state-dependent types of insurance benefits. These cover single payments at maturity, annuity-type payments and payments at the time of a transition. Based on the intensity of the $\mathbb{F}$-doubly stochastic Markov chain, we provide the Galtchouk-Kunita-Watanabe decomposition for a general insurance contract and specify risk-minimizing strategies in a Brownian financial market setting. The results are further illustrated explicitly within an affine structure for the intensity.

Keywords: insurance liabilities; doubly stochastic Markov chains; risk minimization.

MSC: 60J27, 62P05, 91G99

JEL: C02

---

1. Introduction

The management of an insurance portfolio's risk is one of the core challenges in actuarial science. While the classic form of risk mitigation is based on reinsurance contracts, in some cases it is also possible to hedge claim payments by appropriately trading in different assets. This particularly applies if the assets are correlated to the insurance contract's benefits or their (conditional) probability of occurrence. Practical examples in this direction are unit-linked life insurance products, where benefits depend on the performance of the assets, or unemployment insurance products, where the occurrence of a claim payment may depend to some extend on economic and financial conditions of the markets. Moreover there is an ongoing discussion about the introduction of so called longevity bonds which would establish the possibility for life insurance companies and pension funds to hedge parts of their longevity risk, see [1], [2] or [3]. Due to their unsystematic risk part, most insurance claims are not hedgeable completely through a self-financing trading strategy which particularly means that a hybrid market, consisting among others of financial and insurance markets, is incomplete. A reasonable method for optimally choosing an investment strategy is then important to cover at least parts of the risk.

In the present paper we choose the risk-minimization approach and determine hedging strategies in the sense of this criterion for insurance contracts in a very general setting. This quadratic hedging approach bases on the results in [4] for European type payments and to [5] for payment processes. In most cases, the risk-minimizing strategies can be derived from the well known Galtchouk-Kunita-Watanabe (GKW-) decomposition, see [6] or [7].

Similar to the works of [8], [5] or [9–11], we describe an insured person's progress of sojourning different states of an insurance policy as a right continuous stochastic process with finite state space $\mathcal{K} = \{1, ..., N\}$, 1 being a.s. the initial state. More specifically, we adopt the class of $\mathbb{F}$-doubly stochastic Markov chains as introduced in [12], see Appendix A and the comments therein. This family of processes has several properties which make them very suitable for applications in credit risk and insurance market modeling. Being a sub-class of $\mathbb{F}$-conditional Markov chains, they extend the classic notion of Markov chains by including a reference filtration $\mathbb{F}$ which in our case represents additional market information. In this way we are able to take in consideration the influence of external risk factors and economic and financial conditions on transition probabilities of an insured person's progress. In particular, $\mathbb{F}$-doubly stochastic Markov chains behave like time inhomogeneous Markov chains, if we know all the information concerning the underlying risk factors. This corresponds to the intuition that the transition probabilities would be completely specified, if we would dispose of full knowledge on the underlying economic and financial situation. Another important feature is that, if we specify the information as given by the filtration $\mathbb{G} := \mathbb{F}^X \vee \mathbb{F}$, where $\mathbb{F}^X$ is the natural filtration of the $\mathbb{F}$-doubly stochastic Markov chain X, then we have that predictable representation theorems and the so-called hypothesis (H)[1], or immersion property, hold. These properties play a fundamental role in order to compute the optimal strategy for insurance contracts according to the risk-minimization method.
Furthermore, $\mathbb{F}$-doubly stochastic Markov chains may admit matrix-valued stochastic intensity processes. This allows to investigate more flexible models compared to the results e.g. in Møller [5] where a (classical) Markov chain with deterministic intensity matrix function is considered. One further advantage is that $\mathbb{F}$-doubly stochastic Markov chains with intensity are fully characterized by some martingale properties, which can be used for the estimation of the underlying intensity processes, see Biagini *et al.* [13].
Well known examples of $\mathbb{F}$-doubly stochastic Markov chains are reduced form or intensity based models in the case that hypothesis (H) is satisfied. Here, the state space consists of two states with the second state being absorbing such that there can only occur one transition in time. There exist many works on quadratic hedging for these models particularly in the context of credit risk or life insurance theory, see e.g. [14], [15–17], [1,2,18,19] or [20]. In particular, the present paper extends these works to a multi-state framework where several subsequent transitions, driven by $\mathbb{F}$-adapted stochastic intensity processes, are considered. This general setting allows to investigate a larger class of insurance contracts, e.g. income protection insurance contracts with the states "healthy", "sick" and "deceased", and to include the influence of market conditions and external risk factors on the insured person's progress.

Given an $\mathbb{F}$-doubly stochastic Markov chain, we propose a general insurance contract, defined by three different types of insurance benefits: state-dependent payments at maturity, state-dependent annuity-type payments, and (transition-dependent) payments at the time of a transition from one state to another. This definition covers a large set of currently adopted insurance policies. In particular, we illustrate the definitions for pure endowment, term insurance, general life annuity and payment protection insurance contracts. Similar to the results in [21], who applied $\mathbb{F}$-doubly stochastic Markov chains in the context of hedging rating-sensitive financial claims, we obtain the GKW-decomposition for the payment process of general insurance contracts with respect to a particular $\mathbb{F}$-martingale. In this context, we generalize and complement the proofs in [21] in order to adapt the results for the risk-minimization approach which is not investigated there.

---

[1] For the definition and further comments on hypothesis (H), see Proposition A.4 and the text below.

Given that the reference information $\mathbb{F}$ is generated by an N-dimensional Brownian motion $\mathbf{W}$, we then introduce a financial market model, driven by $\mathbf{W}$. In this setting we infer risk-minimizing hedging strategies for insurance contracts with deterministic payment structure with respect to the assets on the financial market. Similarly to the work in [2] we then assume a general affine structure for the intensity of the underlying $\mathbb{F}$-doubly stochastic Markov chain and obtain explicit formulas for the strategies and their residual risk processes. We apply these results in the specific example of an income protection insurance, where we assume that the intensities follow a (multi-dimensional) Ornstein-Uhlenbeck process. We discuss the resulting expected cumulative payment, which may be considered as a fair premium in the interpretation of [22] and [23], as a function of the time horizon, the payment amounts and the underlying interest rates.

The paper is organized as follows. In Section 2 we introduce the notion of general insurance contracts and discuss several examples. In Section 3 we prove our main results for the risk minimization of this kind of contracts in full generality. The risk minimizing strategies are then further illustrated within a general affine specification for the intensities and in a numerical example in an Ornstein-Uhlenbeck framework. We conclude the paper with Appendix A and B, where we summarize important results and concepts of risk-minimization and $\mathbb{F}$-doubly stochastic Markov chains for the reader's convenience.

2. General insurance contracts

We now introduce the notion of general insurance contracts and provide some well known examples of actuarial practice.

In the same notation as in Appendix A, let $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ be a filtered probability space with $\mathbb{G} = \mathbb{F}^X \vee \mathbb{F}$ for some $\mathbb{F}$-doubly stochastic Markov chain X with state space $\mathcal{K} = \{1, ..., N\}$. We assume $\mathbb{P}(X_0 = 1) = 1$. The following definition of general insurance contracts is based on the definitions for payment processes on rating sensitive claims as given e.g. in [21] or [20]. The definition also covers the concepts of insurance contracts as given in [5] or [9].

Definition 1. A *general insurance contract* is given by the quadruple $(X; \mathbf{A}; \mathbf{Y}; \mathbf{Z})$, where $X = (X_t)_{t\in[0,T]}$ is an $\mathbb{F}$-doubly stochastic Markov chain, $\mathbf{A} = (A_t^1, ..., A_t^N)_{t\in[0,T]}$ is an $\mathbb{F}$-adapted, N-dimensional process of finite variation, $\mathbf{Y} = (Y^1, ..., Y^N)$ is an $\mathcal{F}_T$-measurable, N-dimensional random vector, and $\mathbf{Z} = (\mathbf{Z}_t)_{t\in[0,T]}$ with $\mathbf{Z}_t = \left[Z_t^{j,k}\right]_{j,k\in\mathcal{K}}$ is an $\mathbb{F}$-adapted, $N \times N$-dimensional process with zeros on the diagonal.

The different elements of a general insurance product's quadruple are interpreted as follows. The process X is the insured person's progress in time of sojourning in the states $j \in \mathcal{K}$, considered by the insurance policy. The N-dimensional process $\mathbf{A}$ characterizes the cumulative state-dependent payment streams which are continuously paid up to maturity. For example, one can take $\mathbf{A}_t = \mathbf{C}_t - \mathbf{P}_t$, $t \in [0, T]$ with $\mathbf{C}_t = (C_t^1, ...C_t^N)^\intercal$ representing the cumulative state-dependent claim payments (e.g. annuities) and $\mathbf{P}_t = (P_t^1, ..., P_t^N)$ the cumulative state-dependent insurance premiums up to maturity. Both processes, $\mathbf{P}$ and $\mathbf{C}$ are then taken to be $\mathbb{F}$-adapted, càdlàg and increasing. The vector $\mathbf{Y}$ characterizes state-dependent "extra" claim payments at maturity T and the process $\mathbf{Z}$ the "immediate" claim payments at the transition times from one state to another.

For every general insurance contract $(X;\mathbf{A};\mathbf{Y};\mathbf{Z})$ the cumulative payment process $D = (D_t)_{t\in[0,T]}$ is given by

$$\begin{aligned} D_t &:= \mathbf{Y}^\intercal \mathbf{H}_T \mathbb{1}_{\{t=T\}} + \int_{[0,t]} \mathbf{H}_s^\intercal d\mathbf{A}_s + \int_{]0,t]} (\mathbf{Z}_s^\intercal \mathbf{H}_{s-})^\intercal d\mathbf{H}_s \\ &= \sum_{j=1}^{N} \left(Y^j H_T^j \mathbb{1}_{\{t=T\}} + \int_{[0,t]} H_s^j \, dA_s^j + \sum_{\substack{k=1\\k\neq j}}^{N} \int_{]0,t]} Z_s^{j,k} \, dN_s^{jk} \right), \end{aligned} \tag{1}$$

with $H_t^j, j \in \mathcal{K}$, as defined in (48) and counting processes N_t^{jk} from (49). Note that D is of finite variation. We now provide some well known examples of insurance contracts.

Example 1. A *pure endowment* is an insurance contract which guarantees to the insured person some fixed payment if she is alive at maturity. For the sake of simplicity, we only consider the payment to be equal to 1.
We set $\mathcal{K} = \{1,2\}$ with 1 being the state "alive" and 2 the absorbing state "deceased". A pure endowment contract is then given as the quadruple $(X;0;(1,0)^\intercal;0)$ or $(X;-\mathbf{P};(0,1)^\intercal;0)$ if premium payments are considered, respectively.

Example 2. A *term insurance* is an insurance contract which guarantees the heirs of an insured person some fixed payment at the time of decease. For the sake of simplicity, we only consider the payment to be equal to 1.
Again, we set $\mathcal{K} = \{1,2\}$ with 1 being the state "alive" and 2 the absorbing state "deceased". Then a term insurance contract is given as the quadruple $(X;0;0;\mathbf{Z})$ or $(X;-\mathbf{P};0;\mathbf{Z})$ if premium payments are considered, respectively, with $\mathbf{Z} := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

Example 3. A *general annuity* as defined in [2] is an insurance contract which guarantees the insured person an $\mathbb{F}$-progressively measurable, non-negative continuous rate payment $(c_t)_{t\in[0,T]}$ as long as she is alive. The state space is again $\mathcal{K} = \{1,2\}$ with 1 being the state "alive" and 2 the absorbing state "deceased". Then a general annuity contract is given as the quadruple $(X;\left(\int_{]0,t]} c_s ds, 0\right)^\intercal_{t\in[0,T]};0;0)$ or $(X;\left(\int_{]0,t]} c_s ds, 0\right)^\intercal_{t\in[0,T]} - \mathbf{P};0;0)$ if premium payments are considered, respectively.

Example 4. A *payment protection insurance (PPI)* is an insurance contract which is usually offered as an add-on product to some payment obligations, e.g. a loan. In the case of an insured event, the insurance company takes over the respective instalments of the payment obligation for the insured person or her heirs. Generally, the insured events are "disability", "unemployment" and "decease". Hence, the state space for PPI products is given as $\mathcal{K} = \{1,2,3,4\}$ with "2" being the state "disabled", "3" the state "unemployed", "4" the absorbing state "deceased" and "1" the state where no insured event is present.

Then a PPI contract is given as the quadruple $(X;(0,C_t^2,C_t^3,C_t^4)^\intercal_{t\in[0,T]};(0,Y,Y,Y)^\intercal;0)$ or $(X;(0,C_t^2,C_t^3,C_t^4)^\intercal_{t\in[0,T]} - \mathbf{P};(0,Y,Y,Y)^\intercal;0)$ if premium payments are considered, respectively.

As the underlying payment obligation usually stipulates fixed instalments $c_1,...,c_K$ at some given payment dates $0 < T_1 < ... < T_K = T$, the processes C_t^i, $i = 2,3,4$, are generally given as $C_t^i = \sum_{j=1}^{K} c_j \mathbb{1}_{\{T_j \le t\}}$.

Moreover, some payment obligations also contain a so-called balloon rate B at the end of the contract, which has to be paid on top of the usual instalment. If there exists a balloon rate and it is insured, then we set $Y = B$, if there exists no balloon rate or it is not insured, then we set $Y = 0$.

Remark 1. 1) The extra claim payment could also be included in the continuous claim payments. For the reader's convenience, however, we explicitly separate continuous and extra claim payments.

2) The main concepts of premium payment are

- a *single premium* P, where the complete price for the insurance contract is paid at its beginning. In this case, the vector $\mathbf{P}$ would be given as $\mathbf{P} = \left(P\,\mathbb{1}_{\{t\geq 0\}}, P\,\mathbb{1}_{\{t\geq 0\}}, ..., P\,\mathbb{1}_{\{t\geq 0\}}\right)^{\intercal}_{t\in[0,T]}$.
- periodically paid premiums. Here, the insurance price is paid according to periodically paid premiums p_i at a priori specified dates $0 = T_0 < T_1 < ... < T_L \leq T$. Moreover, some insurance policies consider premium freedoms which allow the insured person to intermit premium payments while sojourning (some) insured states. In this case, we have for each vector entry P^i, $i \in \{1, ..., N\}$, of $\mathbf{P}$

$$P^i_t = \begin{cases} \sum_{j=1}^{L} p_j \mathbb{1}_{\{T_j \leq t\}} \\ 0 \end{cases},$$

depending on whether state i is guarantees premium freedom or not.

3. Risk-minimization for general insurance contracts

Aim of this paper is to provide the risk minimizing strategy for a general insurance contract by applying the approach presented in Appendix B. Risk minimization provides hedging strategies which perfectly replicate the claim. Since the market is incomplete, these strategies may not be self-financing and a readjustment (or cost) is needed to achieve perfect replication. According to this method we choose then the optimal strategy, i.e. the strategy with minimal cost.
For this sake, we first consider a general setting and then focus on a deterministic payment structure and an underlying market which is driven by some N-dimensional Brownian motion. These results are then further specified within a general affine setting for the different entries of the matrix-valued intensity.

3.1. Martingale decomposition for payment processes of general insurance claims

We consider the payment process D in (1). Let S^0 denote the market's discounting factor, which will be further specified in Section 3.2. If $\int_{[0,T]} \frac{1}{S^0_u} d|D|_u < \infty$, we get by (59) that the *discounted* cumulative payment stream $\widehat{D} = (\widehat{D}_t)_{t\in[0,T]}$ is given as

$$\begin{aligned}
\widehat{D}_t &= \frac{\mathbf{Y}^{\intercal}\mathbf{H}_T}{S^0_T}\mathbb{1}_{\{t=T\}} + \int_{[0,t]} \frac{1}{S^0_s}\mathbf{H}^{\intercal}_s d\mathbf{A}_s + \int_{]0,t]} \frac{1}{S^0_s}(\mathbf{Z}^{\intercal}_s \mathbf{H}_{s-})^{\intercal} d\mathbf{H}_s \\
&= \sum_{j=1}^{N} \left(\frac{Y^j H^j_T}{S^0_T}\mathbb{1}_{\{t=T\}} + \int_{[0,t]} \frac{1}{S^0_s} H^j_s dA^j_s + \sum_{\substack{k=1\\k\neq j}}^{N} \int_{]0,t]} \frac{1}{S^0_s} Z^{j,k}_s \, dN^{jk}_s \right). \qquad (2)
\end{aligned}$$

We further assume the underlying $\mathbb{F}$-doubly stochastic Markov chain X to admit an intensity $\mathbf{\Psi} = \left([\psi^{j,k}_t]_{j,k\in\mathcal{K}}\right)_{t\in[0,T]}$, as introduced in A.5.

Assumption 2. For every general insurance contract $(X; \mathbf{A}; \mathbf{Y}; \mathbf{Z})$, let

$$\mathbb{E}\left[\left(\frac{Y^j}{S_T^0}\right)^2\right] < \infty\,, \quad j \in \mathcal{K} \tag{3}$$

$$\sup_{s\in[0,T]} \mathbb{E}\left[\left(\int_{[0,s]} \frac{1}{S_u^0} dA_u^j\right)^2\right] < \infty\,, \quad j \in \mathcal{K}, \tag{4}$$

$$\mathbb{E}\left[\int_{]0,T]} \left(\int_{[0,u]} \frac{1}{S_v^0} dA_v^j\right)^2 |\psi_{X_u,j}| du\right] < \infty\,, \quad j \in \mathcal{K}, \tag{5}$$

$$\mathbb{E}\left[\int_{]0,T]} \left(\frac{Z_u^{j,k}}{S_u^0}\right)^2 \psi_{j,k}(u) du\right] < \infty\,, \quad j,k \in \mathcal{K}, j \neq k\,, \tag{6}$$

$$\mathbb{E}\left[\left(\int_{]0,T]} \frac{Z_u^{j,k}}{S_u^0} \psi_{j,k}(u) du\right)^2\right] < \infty\,, \quad j,k \in \mathcal{K}, j \neq k\,, \tag{7}$$

where H_t^j, $t \in [0,T]$ is defined in (48).

Note that (3), (4), (6) and (7) ensure that the discounted payment stream $\widehat{D}$ generated by the general insurance contract $(X; \mathbf{A}; \mathbf{Y}; \mathbf{Z})$ is square integrable.

We remark that the following Lemma is given similarly in [21, Theorem 16.38] under the assumption the local martingale $\mathbf{M}$, defined in (50), is square integrable and the processes $\mathbf{A}$ and $\mathbf{Z}$ are bounded. Here, we generalize their proof to the case where $\mathbf{A}$ and $\mathbf{Z}$ satisfy the conditions of Assumption 2.

For notational convenience, we introduce the process $\mathbf{G} = (\mathbf{G}_t)_{t\in[0,T]} = \left(G_t^1, ..., G_t^N\right)^\intercal$ with

$$G_t^j := \left[\mathbf{Z}_t \mathbf{\Psi}_t^\intercal\right]_{j,j} = \sum_{\substack{k=1\\k\neq j}}^N Z_t^{j,k} \psi_{j,k}(t)\,, \quad j \in \mathcal{K}, t \in [0,T]\,. \tag{8}$$

Lemma 3. *Let $(X; \boldsymbol{A}; \boldsymbol{Y}; \boldsymbol{Z})$ be a general insurance contract, satisfying Assumption 2, then*

$$\begin{aligned}
\mathbb{E}\left[\widehat{D}_T - \widehat{D}_t \,\middle|\, \mathcal{G}_t\right] &= \sum_{j=1}^N \mathbb{E}\left[\frac{Y^j H_T^j}{S_T^0} + \int_{]t,T]} \frac{1}{S_u^0} H_u^j dA_u^j + \sum_{\substack{k=1\\k\neq j}}^N \int_{]t,T]} \frac{Z_u^{j,k}}{S_u^0} dN_u^{jk} \,\middle|\, \mathcal{G}_t\right] \\
&= \sum_{i=1}^N H_t^i \sum_{j=1}^N \mathbb{E}\left[\frac{Y^j p_{i,j}(t,T)}{S_T^0} + \int_{]t,T]} \frac{1}{S_u^0} p_{i,j}(t,u) dA_u^j + \sum_{\substack{k=1\\k\neq j}}^N \int_{]t,T]} \frac{Z_u^{j,k}}{S_u^0} p_{i,j}(t,u)\psi_{j,k}(u) du \,\middle|\, \mathcal{F}_t\right] \\
&= \mathbb{E}\left[\frac{\boldsymbol{P}(t,T)\boldsymbol{Y}}{S_T^0} + \int_{]t,T]} \frac{\boldsymbol{P}(t,u)}{S_u^0} d\boldsymbol{A}_u + \int_{]t,T]} \frac{\boldsymbol{P}(t,u)}{S_u^0} \boldsymbol{G}_u du \,\middle|\, \mathcal{F}_t\right]^\intercal \boldsymbol{H}_t\,,
\end{aligned} \tag{9}$$

where the conditional transition probability process $\boldsymbol{P} = \boldsymbol{P}(s,t) = \left[p_{i,j}(s,t)\right]_{i,j\in\mathcal{K}}$, $0 \leq s \leq t \leq T$ is defined in A.1.

Proof. We proof the theorem by investigating the different conditional expectations separately.

First note that because $\mathbf{Y}$ is taken to be $\mathcal{F}_T$-measurable and S^0 to be $\mathbb{F}$-adapted, by (44) we get for every $j \in \mathcal{K}$

$$\mathbb{E}\left[\frac{Y^j H_T^j}{S_T^0} \,\Big|\, \mathcal{G}_t\right] = \mathbb{E}\left[\frac{Y^j}{S_T^0}\sum_{i=1}^N H_t^i \mathbb{E}\left[H_T^j \,|\, \widetilde{\mathcal{G}}_t\right] \Big|\, \mathcal{G}_t\right] = \mathbb{E}\left[\frac{Y^j}{S_T^0}\sum_{i=1}^N H_t^i p_{i,j}(t,T) \,\Big|\, \mathcal{G}_t\right]$$
$$= \sum_{i=1}^N H_t^i \mathbb{E}\left[\frac{Y^j}{S_T^0} p_{i,j}(t,T) \,\Big|\, \mathcal{F}_t\right],$$

where $\widetilde{\mathcal{G}}_t := \mathcal{F}_T \vee \mathcal{F}_t^X$.

Next, by (51) of Theorem A.7, we get for $j,k \in \mathcal{K}$, $j \neq k$ that

$$\int_{]t,T]} \frac{Z_u^{j,k}}{S_u^0} dN_u^{j,k} = \int_{]t,T]} \frac{Z_u^{j,k}}{S_u^0} dM_u^{j,k} + \int_{]t,T]} \frac{Z_u^{j,k}}{S_u^0} H_u^j \psi_{j,k}(u) du\,.$$

Note that because of (56) and (6), the integral-process with respect to M^{jk} is a square integrable $\mathbb{G}$-martingale. Hence, for every $j,k \in \mathcal{K}$, $j \neq k$, we have

$$\begin{aligned}
\mathbb{E}\left[\int_{]t,T]} \frac{Z_u^{jk}}{S_u^0} dN_u^{jk} \,\Big|\, \mathcal{G}_t\right] &= \mathbb{E}\left[\int_{]t,T]} \frac{Z_u^{jk}}{S_u^0} H_u^j \psi_{j,k}(u) du \,\Big|\, \mathcal{G}_t\right] \\
&= \int_{]t,T]} \mathbb{E}\left[\frac{Z_u^{jk}}{S_u^0} H_u^j \psi_{j,k}(u) \,\Big|\, \mathcal{G}_t\right] du \\
&= \int_{]t,T]} \mathbb{E}\left[\mathbb{E}\left[\frac{Z_u^{jk}}{S_u^0} H_u^j \psi_{j,k}(u) \,\Big|\, \widetilde{\mathcal{G}}_t\right] \Big|\, \mathcal{G}_t\right] du \\
&= \int_{]t,T]} \mathbb{E}\left[\frac{Z_u^{jk}}{S_u^0} \psi_{j,k}(u) \left(\sum_{i=1}^N H_t^i p_{i,j}(t,u)\right) \Big|\, \mathcal{G}_t\right] du \\
&= \sum_{i=1}^N H_t^i \int_{]t,T]} \mathbb{E}\left[\frac{Z_u^{jk}}{S_u^0} p_{i,j}(t,u) \psi_{j,k}(u) \,\Big|\, \mathcal{F}_t\right] du \\
&= \sum_{i=1}^N H_t^i \mathbb{E}\left[\int_{]t,T]} \frac{Z_u^{jk}}{S_u^0} p_{i,j}(t,u) \psi_{j,k}(u) du \,\Big|\, \mathcal{F}_t\right],
\end{aligned}$$

by the conditional version of Fubini's theorem, the definition of $\mathbb{F}$-doubly stochastic Markov chains and hypothesis (H).

Finally, for every $j \in \mathcal{K}$ and for fixed $t \in [0,T]$ we define $\widetilde{A}_u^j := \int_{]t,u]} \frac{1}{S_v^0} dA_v^j$, $u \in [t,T]$. By Proposition A.9 we get

$$\begin{aligned}
\mathbb{E}\left[\int_{]t,T]} H_u^j \frac{1}{S_u^0} dA_u^j \,|\, \mathcal{G}_t\right] &= \mathbb{E}\left[\int_{]t,T]} H_u^j d\widetilde{A}_u^j \,|\, \mathcal{G}_t\right] \\
&= \mathbb{E}\left[\widetilde{A}_T^j H_T^j - \widetilde{A}_t^j H_t^j - \int_{]t,T]} \widetilde{A}_{u-}^j dH_u^j \,|\, \mathcal{G}_t\right] \\
&= \mathbb{E}\left[\widetilde{A}_T^j H_T^j - \int_{]t,T]} \widetilde{A}_{u-}^j dH_u^j \,|\, \mathcal{G}_t\right] = I_1 - I_2\,,
\end{aligned}$$

with

$$I_1 := \mathbb{E}[\widetilde{A}_T^j H_T^j \,|\, \mathcal{G}_t], \quad I_2 := \mathbb{E}\left[\int_{]t,T]} \widetilde{A}_{u-}^j dH_u^j \,|\, \mathcal{G}_t\right].$$

Since $\widetilde{A}_T$ is $\mathcal{F}_T$-measurable, it follows by the hypothesis (H) that

$$I_1 = \mathbb{E}\left[\widetilde{A}^j_T \mathbb{E}\left[H^j_T \mid \widetilde{\mathcal{G}}_t\right] \mid \mathcal{G}_t\right] = \sum_{i=1}^{K} H^i_t \mathbb{E}\left[\widetilde{A}^j_T p_{i,j}(t,v) \mid \mathcal{F}_t\right].$$

Again by the conditional version of Fubini's theorem, hypothesis (H) and with the Kolmogorov forward equation (47) it follows that

$$\begin{aligned}
I_2 &= \mathbb{E}\left[\int_{]t,T]} \widetilde{A}^j_{u-} dH^j_u \,\Big|\, \mathcal{G}_t\right] = \mathbb{E}\left[\int_{]t,T]} \widetilde{A}^j_{u-} dM^j_u + \int_{]t,T]} \widetilde{A}^j_{u-} \psi_{X_u,j}(u) du \,\Big|\, \mathcal{G}_t\right] \\
&= \mathbb{E}\left[\int_{]t,T]} \widetilde{A}^j_{u-} \sum_{k=1}^{K} H^k_u \psi_{k,j}(u) du \,\Big|\, \mathcal{G}_t\right] = \int_{]t,T]} \mathbb{E}\left[\widetilde{A}^j_{u-} \sum_{k=1}^{K} H^k_u \psi_{k,j}(u) \,\Big|\, \mathcal{G}_t\right] du \\
&= \int_{]t,T]} \mathbb{E}\left[\widetilde{A}^j_{u-} \sum_{k=1}^{K} \mathbb{E}\left[H^k_u \mid \widetilde{G}_t\right] \psi_{k,j}(u) \,\Big|\, \mathcal{G}_t\right] du \\
&= \sum_{i=1}^{K} H^i_t \int_{]t,T]} \mathbb{E}\left[\widetilde{A}^j_{u-} \left(\sum_{k=1}^{K} p_{i,k}(t,u)\psi_{k,j}(u)\right) \,\Big|\, \mathcal{F}_t\right] du \\
&= \sum_{i=1}^{K} H^i_t \mathbb{E}\left[\int_{]t,T]} \widetilde{A}^j_{u-} \left(\sum_{k=1}^{K} p_{i,k}(t,u)\psi_{k,j}(u)\right) du \,\Big|\, \mathcal{F}_t\right] \\
&= \sum_{i=1}^{K} H^i_t \mathbb{E}\left[\int_{]t,T]} \widetilde{A}^j_{u-} dp_{i,j}(t,u) \,\Big|\, \mathcal{F}_t\right].
\end{aligned}$$

Hence, by integration by parts and since $p(t,\cdot)$ is continuous, we get

$$\begin{aligned}
I_1 - I_2 &= \sum_{i=1}^{K} H^i_t \mathbb{E}\left[\widetilde{A}_T p_{i,j}(t,T) - \int_{]t,T]} \widetilde{A}^j_{u-} dp_{i,j}(t,u) \mid \mathcal{F}_t\right] \\
&= \sum_{i=1}^{K} H^i_t \mathbb{E}\left[\widetilde{A}_t p_{i,j}(t,t) + \int_{]t,T]} p_{i,j}(t,u) d\widetilde{A}^j_u \mid \mathcal{F}_t\right] \\
&= \sum_{i=1}^{K} H^i_t \mathbb{E}\left[\int_{]t,T]} p_{i,j}(t,u) dA^j_u \mid \mathcal{F}_t\right].
\end{aligned}$$

183 This completes the proof. □

184 Now we are ready to provide the Galtchouk-Kunita-Watanabe decomposition for payment
185 processes of general insurance contracts.

Theorem 4. *Let $(X, \boldsymbol{A}, \boldsymbol{Y}, \boldsymbol{Z})$ be a general insurance contract, satisfying Assumption 2, with discounted payment process $\widehat{D}$, defined in (2). Then the GKW-decomposition of the square-integrable discounted value process $\widehat{U}^D = (\widehat{U}^D_t)_{t\in[0,T]}$ with $\widehat{U}^D_t = \mathbb{E}\left[\widehat{D}_T \mid \mathcal{G}_t\right]$ is given as*

$$\widehat{U}^D_t = \widehat{U}^D_0 + \int_{]0,t]} \boldsymbol{\alpha}^\intercal_u d\boldsymbol{m}_u + \int_{]0,t]} \boldsymbol{\beta}^\intercal_u d\boldsymbol{M}_u\,, \tag{10}$$

where $\boldsymbol{M}$ is given by (50), $\boldsymbol{m} = (\boldsymbol{m}_t)_{t\in[0,T]}$ is a square-integrable $\mathbb{F}$-martingale, given by

$$\boldsymbol{m}_t := \mathbb{E}\left[\frac{\boldsymbol{P}(0,T)\boldsymbol{Y}}{S^0_T} + \int_{[0,T]} \frac{\boldsymbol{P}(0,u)}{S^0_u} d\boldsymbol{A}_u + \int_{]0,T]} \frac{\boldsymbol{P}(0,u)}{S^0_u} \boldsymbol{G}_u du \,\Big|\, \mathcal{F}_t\right], \tag{11}$$

and $\boldsymbol{\alpha}, \boldsymbol{\beta}$ are $\mathbb{G}$-predictable $\mathbb{R}^N$-valued processes defined by

$$\boldsymbol{\alpha}_t = \boldsymbol{L}_{t-} = \boldsymbol{Q}^\intercal(0,t)\boldsymbol{H}_{t-}, \qquad \boldsymbol{\beta}_t = \frac{\boldsymbol{F}(t-,T) + \boldsymbol{Z}^\intercal \boldsymbol{H}_{t-}}{S_t^0}, \tag{12}$$

with $\boldsymbol{H}_t = (H_t^1, ..., H_t^N)$, $t \in [0,T]$, defined by (48), $\boldsymbol{Q}(0,t)$, $t \in [0,T]$, defined in (52), $\boldsymbol{F}(t,T)$, $t \in [0,T]$, defined by

$$\boldsymbol{F}(t,T) := S_t^0 \mathbb{E}\left[\frac{\boldsymbol{P}(t,T)\boldsymbol{Y}}{S_T^0} + \int_{]t,T]} \frac{\boldsymbol{P}(t,u)}{S_u^0} d\boldsymbol{A}_u + \int_{]t,T]} \frac{\boldsymbol{P}(t,u)}{S_u^0} \boldsymbol{G}_u du \,\Big|\, \mathcal{F}_t \right] \tag{13}$$

and $\widehat{U}_0^D = \mathbb{E}[\widehat{D}_T] = \boldsymbol{m}_0^\intercal \boldsymbol{H}_0$.

Proof. The statement and the proof of this theorem can be found in [21, Theorem 16.62]. The authors there, however, prove Decomposition 10 only for $t \in [0,T)$. Because the integrals on the r.h.s. are not all continuous, it is a priori not clear if the decomposition also holds for $\widehat{U}_T^D$. Here we refer to [24, Theorem 4.2.3] for an extension of the proof of [21] to the case $t = T$. □

3.2. Risk minimization for general insurance contracts with deterministic payment structure

In this section we focus on a more specific setting, where we specify the underlying financial market and derive risk-minimizing hedging strategies for insurance contracts with deterministic payment structures.

We start by specifying the underlying market. First of all, we assume the reference filtration $\mathbb{F} = \mathbb{F}^W$ to be the augmented filtration, generated by some N-dimensional Brownian motion $\mathbf{W}$. For computational reasons, particularly in the affine setting of the next section, we set the dimension N of the Brownian motion equal to the number of states under consideration.

Consider then a financial market consisting of $(d+1)$ traded assets $\mathbf{S} = (S_t^0, ..., S_t^d)^\intercal_{t\in[0,T]}$, assumed to be $\mathbb{F}$-adapted, non-negative stochastic processes. Let $\widehat{\mathbf{S}} = (\widehat{S}_t^1, ..., \widehat{S}_t^d)^\intercal_{t\in[0,T]}$ denote the $\mathbb{R}^d$-valued stochastic process of the primary assets $S^1, ..., S^d$, discounted with the asset S^0, i.e. $\widehat{S}_t^i = S_t^i / S_t^0$, $i = 1, ..., d$. Here S^0 is taken to be continuous with $S_t^0 > 0$ for all $t \in [0,T]$ and shall generally represent the value of a self-financing portfolio on the primary assets. In the sequel, we assume $\widehat{\mathbf{S}}$ to be a local $(\mathbb{F}, \mathbb{P})$-martingale, which particularly implies that the market model is arbitrage-free.

Remark 2. The requirement that $\widehat{\mathbf{S}}$ is a local martingale may appear restrictive. However it is always satisfied if we choose S^0 to be the numéraire portfolio defined in [25], since we assume the underlying financial market to contain only continuous asset price processes.

We could also start with a general situation where the discounted asset price processes are given by semimartingales. In this case one has to assume some technical conditions to guarantee the existence of the optimal strategy, see [26] and [27].

Here we prefer to avoid technical complications since our aim is to compute explicitly the risk-minimizing strategy when it exists.

By the representation theorem with respect to Brownian motion it follows that there exists a measurable map $\widetilde{\boldsymbol{\sigma}} : [0,T] \times \mathbb{R}^N \to \mathbb{R}^{d\times N}$, such that

$$\widehat{\mathbf{S}}_t = \widehat{\mathbf{S}}_0 + \int_{]0,t]} \widetilde{\boldsymbol{\sigma}}(s, \mathbf{S}_s) d\mathbf{W}_s \,.$$

Assumption 5. We assume that $\widetilde{\boldsymbol{\sigma}}(t,\mathbf{S}_t)$ is a.s. left-invertible, i.e. that for almost every $(\omega,t)\in\Omega\times[0,T]$ there exists an $\mathbb{F}^W$-adapted $N\times d$-valued matrix $\boldsymbol{\Gamma}_t(\omega)$ such that $\boldsymbol{\Gamma}_t\widetilde{\boldsymbol{\sigma}}(t,\mathbf{S}_t)=\mathbb{I}_N$. This particularly implies $N\geq d$.

From now on we focus on discount factors and insurance contracts with deterministic payment structure.

Assumption 6. 1) $\mathbf{Y}$ is a deterministic vector in $\mathbb{R}^N$.
2) The payment $\mathbf{A}=(\mathbf{A}_t)_{t\in[0,T]}$ is of the form $\mathbf{A}_t=\int_0^t\boldsymbol{\nu}(s)ds$ for some bounded deterministic function $\boldsymbol{\nu}:[0,T]\to\mathbb{R}^N$.
3) $\mathbf{Z}:[0,T]\to\mathbb{R}^{N\times N}$ is a bounded deterministic matrix-valued function.
4) $S^0:[0,T]\to\mathbb{R}$ is a deterministic continuous function.
5) For every $j,k\in\mathcal{K}$, $j\neq k$, $C:=\sup_{u\in[0,T]}\mathbb{E}\left[\left(\psi_u^{j,k}\right)^2\right]<\infty$.

Assumption 6 particularly implies that the integrability conditions of Assumption 2 hold. Note also that the insurance contracts, given in Examples 1, 2 and 4 all satisfy 1), 2) and 3) of Assumption 6. The assumption on S^0 being deterministic is applied very frequently in the literature, e.g if $\mathbb{P}$ is assumed to be some risk-neutral probability measure and $S_t^0=e^{rt}$ for some constant $r>0$.

Remark 3. Here we assume constant interest rates for the sake of simplicity, since the focus of this paper is primarily to evaluate the role of a multi-state progression of the insured person on the risk-minimizing strategy. The following computations can be easily extended to the case of stochastic interest rates if S^0 is assumed to be independent of X. In more general models, the investigation of dependency structures will become inevitable. This goes beyond the scope of the paper and is left to further research.

Due to the representation theorem with respect to Brownian motion, for every $u\in[0,T]$ and every $i,j\in\mathcal{K}$, there exists some $\boldsymbol{\xi}^{i,j}(u,\cdot)\in L^2(\mathbf{W})$ such that

$$\mathbb{E}\left[p_{i,j}(0,u)\,|\,\mathcal{F}_t\right]=\mathbb{E}\left[p_{i,j}(0,u)\right]+\int_{]0,t]}\mathbb{1}_{]0,u]}(s)\boldsymbol{\xi}^{i,j}(u,s)d\mathbf{W}_s\,. \tag{14}$$

Similarly, because of Assumption 6 5), for every $u\in[0,T]$ and every $i,j,k\in\mathcal{K}$, $j\neq k$, there exists some $\boldsymbol{\theta}^{i,j,k}(u,\cdot)\in L^2(\mathbf{W})$ such that

$$\mathbb{E}\left[p_{i,j}(0,u)\psi_u^{j,k}\,|\,\mathcal{F}_t\right]=\mathbb{E}\left[p_{i,j}(0,u)\psi_u^{j,k}\right]+\int_{]0,t]}\mathbb{1}_{]0,u]}(s)\boldsymbol{\theta}^{i,j,k}(u,s)d\mathbf{W}_s\,. \tag{15}$$

Theorem 7. *Given Assumptions 5 and 6, the unique risk-minimizing hedging strategy $\boldsymbol{\xi}=(\boldsymbol{\xi}_t)_{t\in[0,T]}$, characterized in B.5 for a general insurance claim $(X;\boldsymbol{A};\boldsymbol{Y};\boldsymbol{Z})$, satisfying Assumption 6, is given as*

$$\boldsymbol{\xi}_t=\sum_{i=1}^N L_t^i\sum_{j=1}^N\int_{]0,t]}\Big(\frac{Y_T^j}{S_T^0}\boldsymbol{\xi}^{i,j}(T,t)+\int_{[t,T]}\frac{1}{S_u^0}\boldsymbol{\xi}^{i,j}(u,t)\nu_u^j du+\sum_{\substack{k=1\\k\neq j}}^N\int_{[t,T]}\frac{Z_u^{j,k}}{S_u^0}\boldsymbol{\theta}^{i,j,k}(u,t)du\Big)\boldsymbol{\Gamma}_t\,, \tag{16}$$

$$\xi_t^0=\widehat{U}_t^D-\widehat{D}_t-\boldsymbol{\xi}_t^{\intercal}\widehat{\mathbf{S}}_t \tag{17}$$

where Γ_t is the left-inverse of the volatility matrix $\widetilde{\sigma}(t,\mathbf{S}_t)$ and $\widehat{U}^D=(\widehat{U}_t^D)_{t\in[0,T]}$ is the discounted value process of the cumulative payment process D.

Proof. Because of Assumption 6, the i-th component m^i of the martingale $\mathbf{m}$ in Equation (11) is given as

$$
\begin{aligned}
m_t^i &= \sum_{j=1}^{N}\Bigg(\frac{Y_T^j}{S_T^0}\mathbb{E}\left[p_{i,j}(0,T)\mid\mathcal{F}_t\right] + \int_{[0,T]}\frac{1}{S_u^0}\mathbb{E}\left[p_{i,j}(0,u)\mid\mathcal{F}_t\right]\nu_u^j du \\
&\quad + \sum_{\substack{k=1\\k\neq j}}^{N}\int_{]0,T]}\frac{Z_u^{j,k}}{S_u^0}\mathbb{E}\left[p_{i,j}(0,u)\psi_u^{j,k}\mid\mathcal{F}_t\right]du\Bigg) \\
&= \sum_{j=1}^{N}\Bigg(\frac{Y_T^j}{S_T^0}\mathbb{E}\left[p_{i,j}(0,T)\right] + \int_{[0,T]}\frac{1}{S_u^0}\mathbb{E}\left[p_{i,j}(0,u)\right]\nu_u^j du + \sum_{\substack{k=1\\k\neq j}}^{N}\int_{]0,T]}\frac{Z_u^{j,k}}{S_u^0}\mathbb{E}\left[p_{i,j}(0,u)\psi_u^{j,k}\right]du\Bigg) \\
&\quad + \sum_{j=1}^{N}\Bigg(\frac{Y_T^j}{S_T^0}\int_{]0,t]}\boldsymbol{\xi}^{i,j}(T,s)d\mathbf{W}_s + \int_{]0,T]}\frac{1}{S_u^0}\int_{]0,t]}\mathbb{1}_{]0,u]}(s)\boldsymbol{\xi}^{i,j}(u,s)d\mathbf{W}_s\nu_u^j du \\
&\quad\quad + \sum_{\substack{k=1\\k\neq j}}^{N}\int_{]0,T]}\frac{Z_u^{j,k}}{S_u^0}\int_{]0,t]}\mathbb{1}_{]0,u]}(s)\boldsymbol{\theta}^{i,j,k}(u,s)d\mathbf{W}_s du\Bigg)\,.
\end{aligned}
$$

By Assumption 6, Fubini's theorem and the Itô isometry it then follows for every $i,j\in\mathcal{K}$ that

$$
\begin{aligned}
\mathbb{E}\left[\int_{]0,T]}\int_{]0,T]}\left(\frac{1}{S_u^0}\mathbb{1}_{]0,u]}(s)\|\boldsymbol{\xi}^{i,j}(u,s)\|\nu_u^j\right)^2 duds\right] &\leq K_1^2\int_{]0,T]}\mathbb{E}\left[\int_{]0,T]}\|\boldsymbol{\xi}^{i,j}(u,s)\|^2 ds\right]du \\
&= K_1^2\int_{]0,T]}\mathbb{E}\left[\left(\int_{]0,T]}\boldsymbol{\xi}^{i,j}(u,s)d\mathbf{W}_s\right)^2\right]du \\
&= K_1^2\int_{]0,T]}\mathbb{E}\left[\left(\mathbb{E}\left[p_{i,j}(0,u)\mid\mathcal{F}_T\right]-\mathbb{E}\left[p_{i,j}(0,u)\right]\right)^2\right]du \\
&\leq K_1^2 T < \infty
\end{aligned}
$$

for some constant $K_1 > 0$.

Due to Assumption 6 5), we similarly have for every $i,j,k\in\mathcal{K}$, $j\neq k$ that

$$
\begin{aligned}
\mathbb{E}\left[\int_{]0,T]}\int_{]0,T]}\left(\frac{Z_u^{j,k}}{S_u^0}\mathbb{1}_{]0,u]}(s)\|\boldsymbol{\theta}^{i,j,k}(u,s)\|\right)^2 duds\right] &\leq K_2^2\int_{]0,T]}\mathbb{E}\left[\int_{]0,T]}\|\boldsymbol{\theta}^{i,j,k}(u,s)\|^2 ds\right]du \\
&= K_2^2\int_{]0,T]}\mathbb{E}\left[\left(\int_{]0,T]}\boldsymbol{\theta}^{i,j,k}(u,s)d\mathbf{W}_s\right)^2\right]du \\
&= K_2^2\int_{]0,T]}\mathbb{E}\left[\left(\mathbb{E}\left[p_{i,j}(0,u)\psi_u^{j,k}\mid\mathcal{F}_T\right]-\mathbb{E}\left[p_{i,j}(0,u)\right]\right)^2\right]du \\
&\leq K_2^2\int_{]0,T]}\mathbb{E}\left[(\psi_u^{j,k})^2\right]du \\
&\leq K_2^2 CT < \infty
\end{aligned}
$$

for some constant $K_2 > 0$.

Therefore, we can apply the stochastic version of Fubini's theorem, see [28], and obtain

$$m_t^i = \sum_{j=1}^{N}\Big(\frac{Y_T^j}{S_T^0}\mathbb{E}\left[p_{i,j}(0,T)\right] + \int_{[0,T]}\frac{1}{S_u^0}\mathbb{E}\left[p_{i,j}(0,u)\right]\nu_u^j du + \sum_{\substack{k=1\\k\neq j}}^{N}\int_{]0,T]}\frac{Z_u^{j,k}}{S_u^0}\mathbb{E}\left[p_{i,j}(0,u)\psi_u^{j,k}\right]du\Big)$$
$$+ \sum_{j=1}^{N}\int_{]0,t]}\Big(\frac{Y_T^j}{S_T^0}\boldsymbol{\xi}^{i,j}(T,s) + \int_{[s,T]}\frac{1}{S_u^0}\boldsymbol{\xi}^{i,j}(u,s)\nu_u^j du + \sum_{\substack{k=1\\k\neq j}}^{N}\int_{[s,T]}\frac{Z_u^{j,k}}{S_u^0}\boldsymbol{\theta}^{i,j,k}(u,s)du\Big)d\mathbf{W}_s\,.$$

This finally implies

$$dm_t^i = \sum_{j=1}^{N}\Big(\frac{Y_T^j}{S_T^0}\boldsymbol{\xi}^{i,j}(T,t) + \int_{[t,T]}\frac{1}{S_u^0}\boldsymbol{\xi}^{i,j}(u,t)\nu_u^j du + \sum_{\substack{k=1\\k\neq j}}^{N}\int_{[t,T]}\frac{Z_u^{j,k}}{S_u^0}\boldsymbol{\theta}^{i,j,k}(u,t)du\Big)\boldsymbol{\Gamma}_t d\widehat{\mathbf{S}}_t\,.$$

The result then follows immediately with Theorem A.7 and the results of Theorem B.5. □

3.3. Risk minimization for general insurance contracts with deterministic payment structure under an affine specification for the intensities

In the same setting as in Section 3.2, we now specify the risk-minimizing hedging strategies, computed in Theorem 7 within a general affine setting for the intensities. In addition to Assumption 2 we also consider

Assumption 8. 1) For every $0 \le t \le u \le T$ and every $j,k \in \mathcal{K}$, $j \neq k$, we assume the entries $p_{j,k}(t,u)$ of the transition matrix $\mathbf{P}(t,u)$ are of the form

$$p_{j,k}(t,u) = 1 - e^{-\int_t^u \psi_v^{j,k}dv}\,, \tag{18}$$

where $\psi^{j,k}$ are the respective entries of the intensity matrix $\boldsymbol{\Psi}$.

2) For every $u \in [0,T]$ and every $j,k \in \mathcal{K}$, $j \neq k$, $\psi_u^{j,k}$ is of the form

$$\psi_u^{j,k} = (\mathbf{b}^{j,k})^\intercal \boldsymbol{\mu}_u + c^{j,k}\,, \tag{19}$$

where $\mathbf{b}^{j,k} \in \mathbb{R}^N$, $c^{j,k} \in \mathbb{R}$ and $\boldsymbol{\mu} = (\boldsymbol{\mu}_t)_{t\in[0,T]}$ an $\mathbb{R}^N$-valued affine process as specified e.g. in [29, Section 3 and Appendix A] or [30]. Here, $\boldsymbol{\mu}$ is a Markov process with respect to $\mathbb{F}^W$, given as the strong solution to the SDE

$$d\boldsymbol{\mu}_t = \boldsymbol{\delta}(t,\boldsymbol{\mu}_t)dt + \boldsymbol{\sigma}(t,\boldsymbol{\mu}_t)d\mathbf{W}_t\,, \tag{20}$$

where for $t \in [0,T]$, $\mathbf{x} \in \mathbb{R}^N$ and $i,j \in \{1,...,N\}$

$$\boldsymbol{\delta}(t,\mathbf{x}) = \mathbf{d}^0(t) + (\mathbf{d}^1(t))^\intercal \mathbf{x} \tag{21}$$
$$[\boldsymbol{\sigma}(t,\mathbf{x})\boldsymbol{\sigma}(t,\mathbf{x})^\intercal]_{i,j} = [\mathbf{V}^0(t)]_{i,j} + (\mathbf{V}^1(t))_{i,j}^\intercal \mathbf{x}\,, \tag{22}$$

with coefficient functions $\mathbf{d}^0, \mathbf{d}^1, \mathbf{V}^0$ and $\mathbf{V}^1$, taking values in $\mathbb{R}^N, \mathbb{R}^{N\times N}, \mathbb{R}^{N\times N}$ and $\mathbb{R}^{N\times N\times N}$, respectively.

3) The process $\boldsymbol{\mu}$ is such that

$$C := \sup_{u\in[0,T]}\mathbb{E}[\|\boldsymbol{\mu}_u\|^2] < \infty\,. \tag{23}$$

This particularly implies that for every $j,k \in \mathcal{K}$, $j \neq k$, we have

$$\sup_{u\in[0,T]} \mathbb{E}[(\psi_u^{j,k})^2] < \infty \,. \tag{24}$$

With these assumptions and under some technical conditions, presented in [30], we obtain for every $0 \le t \le u \le T$ and every $i,j \in \mathcal{K}$ with $i \neq j$ that

$$\mathbb{E}\left[p_{i,j}(t,u)\,|\,\mathcal{F}_t\right] = \mathbb{E}[1 - e^{-\int_t^u \psi_v^{i,j}dv}\,|\,\mathcal{F}_t] = 1 - e^{\alpha_u^{i,j}(t)+(\boldsymbol{\beta}_u^{i,j}(t))^\intercal \boldsymbol{\mu}_t}\,, \tag{25}$$

$$\begin{aligned}\mathbb{E}\left[p_{i,i}(t,u)\,|\,\mathcal{F}_t\right] &= \mathbb{E}[1 - \sum_{\substack{j=1\\ j\neq i}}^N p_{i,j}(t,u)\,|\,\mathcal{F}_t] = 1 - \sum_{\substack{j=1\\ j\neq i}}^N \mathbb{E}[1 - e^{-\int_t^u \psi_v^{i,j}dv}\,|\,\mathcal{F}_t] \\ &= 2 - N + \sum_{\substack{j=1\\ j\neq i}}^N e^{\alpha_u^{i,j}(t)+(\boldsymbol{\beta}_u^{i,j}(t))^\intercal \boldsymbol{\mu}_t}\,. \end{aligned} \tag{26}$$

Similarly, we obtain for every $0 \le t \le u \le T$ and every $i,j,k \in \mathcal{K}$ with $i \neq j$, $j \neq k$

$$\begin{aligned}\mathbb{E}\left[p_{i,j}(t,u)\psi_u^{j,k}\,|\,\mathcal{F}_t\right] =& \mathbb{E}\left[(1 - e^{-\int_t^u \psi_u^{i,j}du})\psi_u^{j,k}\,|\,\mathcal{F}_t\right] = \mathbb{E}\left[\psi_u^{j,k}\,|\,\mathcal{F}_t\right] - \mathbb{E}\left[e^{-\int_t^u \psi_u^{i,j}du}\psi_u^{j,k}\,|\,\mathcal{F}_t\right] \\ =& e^{\widetilde{\alpha}_u(t)+(\widetilde{\boldsymbol{\beta}}_u(t))^\intercal \boldsymbol{\mu}_t}(\widetilde{\alpha}_u^{j,k}(t) + (\widetilde{\boldsymbol{\beta}}_u^{j,k}(t))^\intercal \boldsymbol{\mu}_t) \\ & - e^{\alpha_u^{i,j}(t)+(\boldsymbol{\beta}_u^{i,j}(t))^\intercal \boldsymbol{\mu}_t}(\widehat{\alpha}_u^{j,k}(t) + (\widehat{\boldsymbol{\beta}}_u^{j,k}(t))^\intercal \boldsymbol{\mu}_t)\,, \end{aligned} \tag{27}$$

$$\begin{aligned}\mathbb{E}\left[p_{j,j}(t,u)\psi_u^{j,k}\,|\,\mathcal{F}_t\right] =& (2-N)\mathbb{E}\left[\psi_u^{j,k}\,|\,\mathcal{F}_t\right] + \sum_{\substack{l=1\\ l\neq j}}^N \mathbb{E}\left[e^{-\int_t^u \psi_u^{j,l}du}\psi_u^{j,k}\,|\,\mathcal{F}_t\right] \\ =& (2-N)e^{\widetilde{\alpha}_u(t)+(\widetilde{\boldsymbol{\beta}}_u(t))^\intercal \boldsymbol{\mu}_t}(\widetilde{\alpha}_u^{j,k}(t) + (\widetilde{\boldsymbol{\beta}}_u^{j,k}(t))^\intercal \boldsymbol{\mu}_t) \\ & + \sum_{\substack{l=1\\ l\neq j}}^N e^{\alpha_u^{j,l}(t)+(\boldsymbol{\beta}_u^{j,l}(t))^\intercal \boldsymbol{\mu}_t}(\widehat{\alpha}_u^{j,k}(t) + (\widehat{\boldsymbol{\beta}}_u^{j,k}(t))^\intercal \boldsymbol{\mu}_t)\,. \end{aligned} \tag{28}$$

For every $0 \le t \le u \le T$ and every combination of $i,j,k,l \in \mathcal{K}$, considered in equations (25), (26), (27) and (28), the functions $\alpha_u^{i,j}$, $\boldsymbol{\beta}_u^{i,j}$ solve the ODEs

$$\frac{d\boldsymbol{\beta}_u^{i,j}}{dt}(t) = \mathbf{b}^{i,j} - \mathbf{d}^1(t)^\intercal \boldsymbol{\beta}_u^{i,j}(t) - \frac{1}{2}(\boldsymbol{\beta}_u^{i,j}(t))^\intercal \mathbf{V}^1(t)\boldsymbol{\beta}_u^{i,j}(t)\,, \tag{29}$$

$$\frac{d\alpha_u^{i,j}}{dt}(t) = c^{i,j} - \mathbf{d}^0(t)^\intercal \boldsymbol{\beta}_u^{i,j}(t) - \frac{1}{2}(\boldsymbol{\beta}_u^{i,j}(t))^\intercal \mathbf{V}^0(t)\boldsymbol{\beta}_u^{i,j}(t)\,, \tag{30}$$

with terminal conditions $\alpha_u^{i,j}(u) = 0$ and $\boldsymbol{\beta}_u^{i,j}(u) = 0$, whereas the functions $\widetilde{\alpha}_u$, $\widetilde{\boldsymbol{\beta}}_u$ solve the ODEs

$$\frac{d\widetilde{\boldsymbol{\beta}}_u}{dt}(t) = -\mathbf{d}^1(t)^\intercal \widetilde{\boldsymbol{\beta}}_u(t) - \frac{1}{2}(\widetilde{\boldsymbol{\beta}}_u(t))^\intercal \mathbf{V}^1(t)\widetilde{\boldsymbol{\beta}}_u(t)\,, \tag{31}$$

$$\frac{d\widetilde{\alpha}_u}{dt}(t) = -\mathbf{d}^0(t)^\intercal \widetilde{\boldsymbol{\beta}}_u(t) - \frac{1}{2}(\widetilde{\boldsymbol{\beta}}_u(t))^\intercal \mathbf{V}^0(t)\widetilde{\boldsymbol{\beta}}_u(t) \tag{32}$$

248 with terminal conditions $\widetilde{\alpha}_u(u) = 0$ and $\widetilde{\boldsymbol{\beta}}_u(u) = 0$.

The functions $\widetilde{\alpha}_u^{k,l}$, $\widetilde{\boldsymbol{\beta}}_u^{k,l}$, $\widehat{\alpha}_u^{k,l}$ and $\widehat{\boldsymbol{\beta}}_u^{k,l}$, $k,l \in \mathcal{K}$, corresponding to $\widetilde{\alpha}_u$ and $\widetilde{\boldsymbol{\beta}}_u$ or to $\alpha_u^{i,j}$ and $\boldsymbol{\beta}_u^{i,j}$ for $i,j \in \mathcal{K}$ as considered in equations (25), (26), (27) and (28), then solve the ODEs

$$\frac{d\widetilde{\boldsymbol{\beta}}_u^{k,l}}{dt}(t) = -\mathbf{d}^1(t)^\intercal \widetilde{\boldsymbol{\beta}}_u^{k,l}(t) - (\widetilde{\boldsymbol{\beta}}_u(t))^\intercal \mathbf{V}^1(t)\widetilde{\boldsymbol{\beta}}_u^{k,l}(t)\,, \tag{33}$$

$$\frac{d\widetilde{\alpha}_u^{k,l}}{dt}(t) = -\mathbf{d}^0(t)^\intercal \widetilde{\boldsymbol{\beta}}_u^{k,l}(t) - (\widetilde{\boldsymbol{\beta}}_u(t))^\intercal \mathbf{V}^0(t)\widetilde{\boldsymbol{\beta}}_u^{k,l}(t) \tag{34}$$

with terminal conditions $\widetilde{\alpha}_u^{k,l}(u) = c^{k,l}$, $\widetilde{\boldsymbol{\beta}}_u^{k,l}(u) = \mathbf{b}^{k,l}$ and

$$\frac{d\widehat{\boldsymbol{\beta}}_u^{k,l}}{dt}(t) = -\mathbf{d}^1(t)^\intercal \widehat{\boldsymbol{\beta}}_u^{k,l}(t) - (\boldsymbol{\beta}_u^{i,j}(t))^\intercal \mathbf{V}^1(t)\widehat{\boldsymbol{\beta}}_u^{k,l}(t)\,, \tag{35}$$

$$\frac{d\widehat{\alpha}_u^{k,l}}{dt}(t) = -\mathbf{d}^0(t)^\intercal \widehat{\boldsymbol{\beta}}_u^{k,l}(t) - (\boldsymbol{\beta}_u^{i,j}(t))^\intercal \mathbf{V}^0(t)\widehat{\boldsymbol{\beta}}_u^{k,l}(t) \tag{36}$$

249 with terminal conditions $\widehat{\alpha}_u^{k,l}(u) = c^{k,l}$, $\widehat{\boldsymbol{\beta}}_u^{k,l}(u) = \mathbf{b}^{k,l}$.

Note that with these specifications, we obtain by (20) that for every $0 \le t < u \le T$ and every $i,j \in \mathcal{K}$, $i \neq j$

$$\begin{aligned}
&\mathbb{E}\left[p_{i,j}(0,u)\,|\,\mathcal{F}_t\right] = 1 - e^{-\int_0^t \psi_v^{i,j}dv}\mathbb{E}\left[e^{-\int_t^u \psi_v^{i,j}dv}\,|\,\mathcal{F}_t\right] = 1 - e^{-\int_0^t \psi_v^{i,j}dv} e^{\alpha_u^{i,j}(t)+(\boldsymbol{\beta}_u^{i,j}(t))^\intercal \boldsymbol{\mu}_t} \\
=&1 - e^{\alpha_u^{i,j}(0)+(\boldsymbol{\beta}_u^{i,j}(0))^\intercal \boldsymbol{\mu}_0} - \int_0^t e^{-\int_0^s \psi_v^{i,j}dv} e^{\alpha_u^{i,j}(s)+(\boldsymbol{\beta}_u^{i,j}(s))^\intercal \boldsymbol{\mu}_s} (\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal \boldsymbol{\beta}_u^{i,j}(s)d\mathbf{W}_s
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\left[p_{i,i}(0,u)\,|\,\mathcal{F}_t\right] = 1 - \sum_{\substack{j=1\\ j\neq i}}^N \mathbb{E}[1 - e^{-\int_t^u \psi_v^{i,j}dv}\,|\,\mathcal{F}_t] = 2 - N + \sum_{\substack{j=1\\ j\neq i}}^N e^{-\int_0^t \psi_v^{i,j}dv} e^{\alpha_u^{i,j}(t)+(\boldsymbol{\beta}_u^{i,j}(t))^\intercal \boldsymbol{\mu}_t} \\
=&2 - N + \sum_{\substack{j=1\\ j\neq i}}^N \left(e^{\alpha_u^{i,j}(0)+(\boldsymbol{\beta}_u^{i,j}(0))^\intercal \boldsymbol{\mu}_0} + \int_0^t e^{-\int_0^s \psi_v^{i,j}dv} e^{\alpha_u^{i,j}(s)+(\boldsymbol{\beta}_u^{i,j}(s))^\intercal \boldsymbol{\mu}_s} (\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal \boldsymbol{\beta}_u^{i,j}(s)d\mathbf{W}_s \right).
\end{aligned}$$

Moreover, as every martingale with respect to the Brownian filtration $\mathbb{F}^W$ is continuous, we have for $0 \le u \le t \le T$ that

$$\begin{aligned}
&\mathbb{E}\left[p_{i,j}(0,u)\,|\,\mathcal{F}_t\right] = p_{i,j}(0,u) = \lim_{w \nearrow u} \mathbb{E}\left[p_{i,j}(0,u)\,|\,\mathcal{F}_w\right] \\
=&\lim_{w \nearrow u}\left(1 - e^{\alpha_u^{i,j}(0)+(\boldsymbol{\beta}_u^{i,j}(0))^\intercal \boldsymbol{\mu}_0} - \int_0^w e^{-\int_0^s \psi_v^{i,j}dv} e^{\alpha_u^{i,j}(s)+(\boldsymbol{\beta}_u^{i,j}(s))^\intercal \boldsymbol{\mu}_s} (\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal \boldsymbol{\beta}_u^{i,j}(s)d\mathbf{W}_s\right) \\
=&1 - e^{\alpha_u^{i,j}(0)+(\boldsymbol{\beta}_u^{i,j}(0))^\intercal \boldsymbol{\mu}_0} - \int_0^u e^{-\int_0^s \psi_v^{i,j}dv} e^{\alpha_u^{i,j}(s)+(\boldsymbol{\beta}_u^{i,j}(s))^\intercal \boldsymbol{\mu}_s} (\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal \boldsymbol{\beta}_u^{i,j}(s)d\mathbf{W}_s
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\left[p_{i,i}(0,u)\,|\,\mathcal{F}_t\right] = p_{i,i}(0,u) = \lim_{w \nearrow u} \mathbb{E}\left[p_{i,i}(0,u)\,|\,\mathcal{F}_w\right] \\
=&2 - N + \sum_{\substack{j=1\\ j\neq i}}^N \left(e^{\alpha_u^{i,j}(0)+(\boldsymbol{\beta}_u^{i,j}(0))^\intercal \boldsymbol{\mu}_0} + \int_0^u e^{-\int_0^s \psi_v^{i,j}dv} e^{\alpha_u^{i,j}(s)+(\boldsymbol{\beta}_u^{i,j}(s))^\intercal \boldsymbol{\mu}_s} (\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal \boldsymbol{\beta}_u^{i,j}(s)d\mathbf{W}_s \right).
\end{aligned}$$

Hence, for arbitrary $u, t \in [0, T]$, we obtain

$$\mathbb{E}\left[p_{i,j}(0,u) \mid \mathcal{F}_t\right] = c_1^{i,j}(u) + \int_0^t \vartheta_1^{i,j}(s,u)\mathbb{1}_{[0,u]}(s)d\mathbf{W}_s \ , \tag{37}$$

$$\mathbb{E}\left[p_{i,i}(0,u) \mid \mathcal{F}_t\right] = c_2^{i}(u) + \int_0^t \vartheta_2^{i}(s,u)\mathbb{1}_{[0,u]}(s)d\mathbf{W}_s \ , \tag{38}$$

where for $u, s \in [0, T]$ and every $i, j \in \mathcal{K}$, $i \neq j$

$$\begin{aligned}
c_1^{i,j}(u) &:= 1 - e^{\alpha_u^{i,j}(0)+(\boldsymbol{\beta}_u^{i,j}(0))^\intercal \boldsymbol{\mu}_0} \ , \\
\vartheta_1^{i,j}(s,u) &:= -e^{-\int_0^s \psi_v^{i,j} dv} e^{\alpha_u^{i,j}(s)+(\boldsymbol{\beta}_u^{i,j}(s))^\intercal \boldsymbol{\mu}_s} (\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal \boldsymbol{\beta}_u^{i,j}(s) \ , \\
c_2^{i}(u) &:= 2 - N + \sum_{\substack{j=1 \\ j\neq i}}^{N} e^{\alpha_u^{i,j}(0)+(\boldsymbol{\beta}_u^{i,j}(0))^\intercal \boldsymbol{\mu}_0} \ , \\
\vartheta_2^{i}(s,u) &:= \sum_{\substack{j=1 \\ j\neq i}}^{N} e^{-\int_0^s \psi_v^{i,j} dv} e^{\alpha_u^{i,j}(s)+(\boldsymbol{\beta}_u^{i,j}(s))^\intercal \boldsymbol{\mu}_s} (\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal \boldsymbol{\beta}_u^{i,j}(s) \ .
\end{aligned}$$

Similarly we get for $0 \leq t < u \leq T$ and every $i, j, k \in \mathcal{K}$ with $i \neq j$, $j \neq k$ that

$$\begin{aligned}
&\mathbb{E}\left[p_{i,j}(0,u)\psi_u^{j,k} \mid \mathcal{F}_t\right] = \mathbb{E}\left[\psi_u^{j,k} \mid \mathcal{F}_t\right] - e^{-\int_0^t \psi_v^{i,j} dv}\mathbb{E}\left[e^{-\int_t^u \psi_v^{i,j} dv}\psi_u^{j,k} \mid \mathcal{F}_t\right] \\
&= e^{\widetilde{\alpha}_u(t)+(\widetilde{\boldsymbol{\beta}}_u(t))^\intercal \boldsymbol{\mu}_t}(\widetilde{\alpha}_u^{j,k}(t) + (\widetilde{\boldsymbol{\beta}}_u^{j,k}(t))^\intercal \boldsymbol{\mu}_t) - e^{-\int_0^t \psi_v^{i,j} dv} e^{\alpha_u^{i,j}(t)+(\boldsymbol{\beta}_u^{i,j}(t))^\intercal \boldsymbol{\mu}_t}(\widehat{\alpha}_u^{j,k}(t) + (\widehat{\boldsymbol{\beta}}_u^{j,k}(t))^\intercal \boldsymbol{\mu}_t) \\
&= e^{\widetilde{\alpha}_u(0)+(\widetilde{\boldsymbol{\beta}}_u(0))^\intercal \boldsymbol{\mu}_0}(\widetilde{\alpha}_u^{j,k}(0) + (\widetilde{\boldsymbol{\beta}}_u^{j,k}(0))^\intercal \boldsymbol{\mu}_0) - e^{\alpha_u^{i,j}(0)+(\boldsymbol{\beta}_u^{i,j}(0))^\intercal \boldsymbol{\mu}_0}(\widehat{\alpha}_u^{j,k}(0) + (\widehat{\boldsymbol{\beta}}_u^{j,k}(0))^\intercal \boldsymbol{\mu}_0) \\
&\quad + \int_0^t \Big\{ e^{\widetilde{\alpha}_u(s)+(\widetilde{\boldsymbol{\beta}}_u(s))^\intercal \boldsymbol{\mu}_s}(\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal \big((\widetilde{\alpha}_u^{j,k}(s) + (\widetilde{\boldsymbol{\beta}}_u^{j,k}(s))^\intercal \boldsymbol{\mu}_s)\widetilde{\boldsymbol{\beta}}_u(s) + \widetilde{\boldsymbol{\beta}}_u^{j,k}(s)\big) \\
&\qquad + e^{-\int_0^s \psi_v^{j,k} dv} e^{\alpha_u^{i,j}(s)+(\boldsymbol{\beta}_u^{i,j}(s))^\intercal \boldsymbol{\mu}_s}(\widehat{\alpha}_u^{j,k}(s) + (\widehat{\boldsymbol{\beta}}_u^{j,k}(s))^\intercal \boldsymbol{\mu}_s)(\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal \boldsymbol{\beta}_u^{i,j}(s) \\
&\qquad + e^{-\int_0^s \psi_v^{j,k} dv} e^{\alpha_u^{i,j}(s)+(\boldsymbol{\beta}_u^{i,j}(s))^\intercal \boldsymbol{\mu}_s}(\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal \widehat{\boldsymbol{\beta}}_u^{j,k}(s) \Big\} d\mathbf{W}_s
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\left[p_{j,j}(0,u)\psi_u^{j,k} \mid \mathcal{F}_t\right] = (2-N)\mathbb{E}\left[\psi_u^{j,k} \mid \mathcal{F}_t\right] + \sum_{\substack{l=1 \\ l\neq j}}^{N} e^{-\int_0^t \psi_v^{j,l} dv}\mathbb{E}\left[e^{-\int_t^u \psi_v^{j,l} dv}\psi_u^{j,k} \mid \mathcal{F}_t\right] \\
&= (2-N)e^{\widetilde{\alpha}_u(0)+(\widetilde{\boldsymbol{\beta}}_u(0))^\intercal \boldsymbol{\mu}_0}(\widetilde{\alpha}_u^{j,k}(0) + (\widetilde{\boldsymbol{\beta}}_u^{j,k}(0))^\intercal \boldsymbol{\mu}_0) + \sum_{\substack{l=1 \\ l\neq j}}^{N} e^{\alpha_u^{j,l}(0)+(\boldsymbol{\beta}_u^{j,l}(0))^\intercal \boldsymbol{\mu}_0}(\widehat{\alpha}_u^{j,k}(0) + (\widehat{\boldsymbol{\beta}}_u^{j,k}(0))^\intercal \boldsymbol{\mu}_0) \\
&\quad + \int_0^t (2-N)\Big\{ e^{\widetilde{\alpha}_u(s)+(\widetilde{\boldsymbol{\beta}}_u(s))^\intercal \boldsymbol{\mu}_s}(\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal \big((\widetilde{\alpha}_u^{j,k}(s) + (\widetilde{\boldsymbol{\beta}}_u^{j,k}(s))^\intercal \boldsymbol{\mu}_s)\widetilde{\boldsymbol{\beta}}_u(s) + \widetilde{\boldsymbol{\beta}}_u^{j,k}(s)\big) \\
&\qquad + \sum_{\substack{l=1 \\ l\neq j}}^{N} e^{-\int_0^s \psi_v^{j,l} dv} e^{\alpha_u^{j,l}(s)+(\boldsymbol{\beta}_u^{j,l}(s))^\intercal \boldsymbol{\mu}_s}(\widehat{\alpha}_u^{j,k}(s) + (\widehat{\boldsymbol{\beta}}_u^{j,k}(s))^\intercal \boldsymbol{\mu}_s)(\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal \boldsymbol{\beta}_u^{j,l}(s) \\
&\qquad + e^{-\int_0^s \psi_v^{j,l} dv} e^{\alpha_u^{j,l}(s)+(\boldsymbol{\beta}_u^{j,l}(s))^\intercal \boldsymbol{\mu}_s}(\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal \widehat{\boldsymbol{\beta}}_u^{j,k}(s)\Big) \Big\} d\mathbf{W}_s.
\end{aligned}$$

Note that by Jensen's inequality and Assumption 8 4), we get for every $0 \leq t < u \leq T$ and every $i, j, k \in \mathcal{K}$ with $j \neq k$ that

$$\mathbb{E}[\mathbb{E}[p_{i,j}(0,u)\psi_u^{j,k} \mid \mathcal{F}_t]^2] \leq \mathbb{E}[\mathbb{E}[p_{i,j}(0,u)^2(\psi_u^{j,k})^2 \mid \mathcal{F}_t]] \leq \mathbb{E}[(\psi_u^{j,k})^2] \leq C \ .$$

Finally, with the same limit-arguments as above, we obtain for arbitrary $u, t \in [0, T]$ and every $i, j, k \in \mathcal{K}$ with $i \neq j$, $j \neq k$ that

$$\mathbb{E}\left[p_{i,j}(0,u)\psi_u^{j,k} \mid \mathcal{F}_t\right] = c_3(u) + \int_0^t \vartheta_3^{i,j,k}(s,u)\mathbb{1}_{[0,u]}(s)d\mathbf{W}_s \,, \tag{39}$$

$$\mathbb{E}\left[p_{j,j}(0,u)\psi_u^{j,k} \mid \mathcal{F}_t\right] = c_4(u) + \int_0^t \vartheta_4^{j,k}(s,u)\mathbb{1}_{[0,u]}(s)d\mathbf{W}_s \,, \tag{40}$$

where for $u, s \in [0, T]$, $i, j, k \in \mathcal{K}$ with $i \neq j$, $j \neq k$

$$c_3(u) := e^{\widetilde{\alpha}_u(0)+(\widetilde{\boldsymbol{\beta}}_u(0))^\intercal \boldsymbol{\mu}_0}(\widetilde{\alpha}_u^{j,k}(0) + (\widetilde{\boldsymbol{\beta}}_u^{j,k}(0))^\intercal \boldsymbol{\mu}_0) - e^{\alpha_u^{i,j}(0)+(\boldsymbol{\beta}_u^{i,j}(0))^\intercal \boldsymbol{\mu}_0}(\widehat{\alpha}_u^{j,k}(0) + (\widehat{\boldsymbol{\beta}}_u^{j,k}(0))^\intercal \boldsymbol{\mu}_0)\,,$$

$$\vartheta_3^{i,j,k}(u,s) := \Big\{ e^{\widetilde{\alpha}_u(s)+(\widetilde{\boldsymbol{\beta}}_u(s))^\intercal \boldsymbol{\mu}_s}(\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal\big((\widetilde{\alpha}_u^{j,k}(s) + (\widetilde{\boldsymbol{\beta}}_u^{j,k}(s))^\intercal \boldsymbol{\mu}_s)\widetilde{\boldsymbol{\beta}}_u(s) + \widetilde{\boldsymbol{\beta}}_u^{j,k}(s)\big)$$
$$+ e^{-\int_0^s \psi_v^{j,k} dv} e^{\alpha_u^{i,j}(s)+(\boldsymbol{\beta}_u^{i,j}(s))^\intercal \boldsymbol{\mu}_s}(\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal\Big((\widehat{\alpha}_u^{j,k}(s) + (\widehat{\boldsymbol{\beta}}_u^{j,k}(s))^\intercal \boldsymbol{\mu}_s)\boldsymbol{\beta}_u^{i,j}(s) + \widehat{\boldsymbol{\beta}}_u^{j,k}(s)\Big)\Big\}\,,$$

$$c_4(u) := (2-N)e^{\widetilde{\alpha}_u(0)+(\widetilde{\boldsymbol{\beta}}_u(0))^\intercal \boldsymbol{\mu}_0}(\widetilde{\alpha}_u^{j,k}(0) + (\widetilde{\boldsymbol{\beta}}_u^{j,k}(0))^\intercal \boldsymbol{\mu}_0)$$
$$+ \sum_{\substack{l=1\\ l\neq j}}^{N} e^{\alpha_u^{j,l}(0)+(\boldsymbol{\beta}_u^{j,l}(0))^\intercal \boldsymbol{\mu}_0}(\widehat{\alpha}_u^{j,k}(0) + (\widehat{\boldsymbol{\beta}}_u^{j,k}(0))^\intercal \boldsymbol{\mu}_0)\,,$$

$$\vartheta_4^{j,k}(u,s) := \Big\{(2-N)e^{\widetilde{\alpha}_u(s)+(\widetilde{\boldsymbol{\beta}}_u(s))^\intercal \boldsymbol{\mu}_s}(\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal\Big((\widetilde{\alpha}_u^{j,k}(s) + (\widetilde{\boldsymbol{\beta}}_u^{j,k}(s))^\intercal \boldsymbol{\mu}_s)\widetilde{\boldsymbol{\beta}}_u(s) + \widetilde{\boldsymbol{\beta}}_u^{j,k}(s)\Big)$$
$$+ \sum_{\substack{l=1\\ l\neq j}}^{N} e^{-\int_0^s \psi_v^{j,l} dv} e^{\alpha_u^{j,l}(s)+(\boldsymbol{\beta}_u^{j,l}(s))^\intercal \boldsymbol{\mu}_s}(\boldsymbol{\sigma}(s,\boldsymbol{\mu}_s))^\intercal\Big((\widehat{\alpha}_u^{j,k}(s) + (\widehat{\boldsymbol{\beta}}_u^{j,k}(s))^\intercal \boldsymbol{\mu}_s)\boldsymbol{\beta}_u^{j,l}(s) + \widehat{\boldsymbol{\beta}}_u^{j,k}(s)\Big)\Big\}\,.$$

We can now apply these results to compute explicitly the risk minimizing strategy as given in Theorem 4 and more specifically by Theorem 7 in the Brownian setting in consideration. From Equations (37), (38), (39) and (40) it follows immediately that the processes $\xi^{i,j}(u,\cdot)$ and $\theta^{i,j,k}(u,\cdot)$, $i, j, k \in \mathcal{K}$, $j \neq k$, of (14) and (15) are given as

$$\xi^{i,j}(u,t) = \vartheta_1^{i,j}(u,t)\,, \quad \xi^{i,i}(u,t) = \vartheta_2^{i}(u,t)\,,$$
$$\theta^{i,j,k}(u,t) = \vartheta_3^{i,j,k}(u,t)\,, \quad \theta^{j,j,k}(u,t) = \vartheta_4^{j,k}(u,t)\,.$$

Moreover, with equations (25), (26), (27) and (28) the i-th component $F^i(t,T)$, $i \in \mathcal{K}$, of $\mathbf{F}(t,T)$ in (13) is given as

$$
\begin{aligned}
F^i(t,T) =& \frac{S^0(t)}{S^0(T)} Y^i \mathbb{E}\left[p_{i,i}(t,T) \mid \mathcal{F}_t\right] + \sum_{\substack{j=1\\ j\neq i}}^{N} \frac{S^0(t)}{S^0(T)} Y^j \mathbb{E}\left[p_{i,j}(t,T) \mid \mathcal{F}_t\right] \\
&+ \int_t^T \frac{S^0(t)}{S^0(u)} \mathbb{E}\left[p_{i,i}(t,u) \mid \mathcal{F}_t\right] \nu^i(u) du + \sum_{\substack{j=1\\ j\neq i}}^{N} \int_t^T \frac{S^0(t)}{S^0(u)} \mathbb{E}\left[p_{i,j}(t,u) \mid \mathcal{F}_t\right] \nu^j(u) du \\
&+ \sum_{\substack{k=1\\ k\neq i}}^{N} \int_t^T \frac{S^0(t)}{S^0(u)} Z^{i,k}(u) \mathbb{E}\left[p_{i,i}(t,u) \psi^{i,k}(u) \mid \mathcal{F}_t\right] du \\
&+ \sum_{\substack{j=1\\ j\neq i}}^{N} \sum_{\substack{k=1\\ k\neq j}}^{N} \int_t^T \frac{S^0(t)}{S^0(u)} Z^{j,k}(u) \mathbb{E}\left[p_{i,j}(t,u) \psi^{j,k}(u) \mid \mathcal{F}_t\right] du \\
=& \frac{S^0(t)}{S^0(T)} Y^i \Big(2 - N + \sum_{\substack{j=1\\ j\neq i}}^{N} e^{\alpha^{i,j}(t,T) + \boldsymbol{\beta}^{i,j}(t,T)\cdot \boldsymbol{\mu}_t}\Big) + \sum_{\substack{j=1\\ j\neq i}}^{N} \frac{S^0(t)}{S^0(T)} Y^j (1 - e^{\alpha^{i,j}(t,T) + \boldsymbol{\beta}^{i,j}(t,T)\cdot \boldsymbol{\mu}_t}) \\
&+ \int_t^T \frac{S^0(t)}{S^0(u)} \Big(2 - N + \sum_{\substack{j=1\\ j\neq i}}^{N} e^{\alpha^{i,j}(t,u) + \boldsymbol{\beta}^{i,j}(t,u)\cdot \boldsymbol{\mu}_t}\Big) \nu^i(u) du \\
&+ \sum_{\substack{j=1\\ j\neq i}}^{N} \int_t^T \frac{S^0(t)}{S^0(u)} \Big(1 - e^{\alpha_t^{i,j}(u) + \boldsymbol{\beta}_t^{i,j}(u)\cdot \boldsymbol{\mu}_u}\Big) \nu^j(u) du \\
&+ \sum_{\substack{k=1\\ k\neq i}}^{N} \int_t^T \frac{S^0(t)}{S^0(u)} Z^{i,k}(u) \Big\{ (2-N) e^{\widetilde{\alpha}(t,u) + \widetilde{\boldsymbol{\beta}}(t,u)\cdot \boldsymbol{\mu}_t} (\widetilde{\alpha}^{i,k}(t,u) + \widetilde{\boldsymbol{\beta}}^{i,k}(t,u)\cdot \boldsymbol{\mu}_t) \\
&\qquad + \sum_{\substack{l=1\\ l\neq i}}^{N} e^{\alpha^{i,l}(t,u) + \boldsymbol{\beta}^{i,l}(t,u)\cdot \boldsymbol{\mu}_t} (\widehat{\alpha}^{i,k}(t,u) + \widehat{\boldsymbol{\beta}}^{i,k}(t,u)\cdot \boldsymbol{\mu}_t) \Big\} du \\
&+ \sum_{\substack{j=1\\ j\neq i}}^{N} \sum_{\substack{k=1\\ k\neq j}}^{N} \int_t^T \frac{S^0(t)}{S^0(u)} Z^{j,k}(u) \Big\{ e^{\widetilde{\alpha}(t,u) + \widetilde{\boldsymbol{\beta}}(t,u)\cdot \boldsymbol{\mu}_t} (\widetilde{\alpha}^{j,k}(t,u) + \widetilde{\boldsymbol{\beta}}^{j,k}(t,u)\cdot \boldsymbol{\mu}_t) \\
&\qquad - e^{\alpha^{i,j}(t,u) + \boldsymbol{\beta}^{i,j}(t,u)\cdot \boldsymbol{\mu}_t} (\widehat{\alpha}^{j,k}(t,u) + \widehat{\boldsymbol{\beta}}^{j,k}(t,u)\cdot \boldsymbol{\mu}_t) \Big\} du\,, \qquad (41)
\end{aligned}
$$

and can hence be expressed explicitly in terms of $\boldsymbol{\mu}$.

3.4. Application: The Expected Cumulative Payment in an Ornstein-Uhlenbeck Framework

In this section, we illustrate in a specific example how the expected (discounted) cumulative payment $\mathbb{E}[\widehat{D}_T]$ from Lemma 3 can be calculated by using the explicit expression of the components $F^i(t,T)$ from the previous section and the connection established through equation (9). In the following, we regard a specific insurance product, namely an income protection insurance, and for simplicity we assume that the corresponding process X of the insured person can only take the three states $\mathcal{K} = \{1 = \textit{healthy}, 2 = \textit{sick/unfit for work}, 3 = \textit{death}\}$, where state 3 is absorbing. The corresponding transitions are illustrated in Figure 1. Hence, for $\mathbf{b}^{j,k} \in \mathbb{R}^3$, $c^{j,k} \in \mathbb{R}$ from (19) we set $\mathbf{b}^{j,k} = \mathbf{0}$ and $c^{j,k} = 0$ for $(j,k) \notin \mathcal{I} := \{(1,2),(2,1),(1,3),(2,3)\}$.

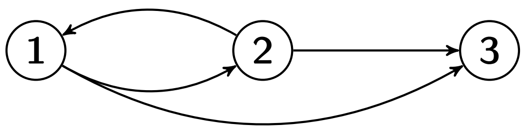

Figure 1. Possible transitions for an insured person's process in an income protection insurance with the three states $\mathcal{K} = \{1 = \text{healthy}, 2 = \text{sick/unfit for work}, 3 = \text{death}\}$ with absorbing state 3.

In the following, we assume a specific form for the Markov process $\boldsymbol{\mu}$ from (19), namely a simple, 3-dimensional Ornstein-Uhlenbeck (OU) process with corresponding SDE

$$d\boldsymbol{\mu}_t = \mathbf{a}\,\boldsymbol{\mu}_t dt + \boldsymbol{\sigma} d\mathbf{W}_t\,, \qquad \boldsymbol{\mu}_0 = \boldsymbol{\xi} \in \mathbb{R}^3\,.$$

Note that for the OU process it is well-known that condition (23) for the expectation is fulfilled. A major drawback, though, of choosing this process for the intensity is the undesirable feature that it can become negative with positive probability. However, [31] provide calibrated parameters for which the probability of negative values for $\boldsymbol{\mu}$ turns out to be negligible. For this reason, we choose similar parameter values and set

$$\mathbf{a} = (0.07, 0.11, 0.09)^\intercal\,, \boldsymbol{\sigma} = \begin{pmatrix} 0.0003 & 0 & 0 \\ 0 & 0.0007 & 0 \\ 0 & 0 & 0.0005 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\xi} = (0.1, 0.1, 0.1)^\intercal\,.$$

Hence, equations (21) and (22) simplify and yield to

$$\begin{aligned} \boldsymbol{\delta}(t, \mathbf{x}) &= (\mathbf{d}^1)^\intercal \mathbf{x} \\ [\boldsymbol{\sigma}(t, \mathbf{x})\boldsymbol{\sigma}(t, \mathbf{x})^\intercal]_{i,j} &= [\mathbf{V}^0]_{i,j}\,, \end{aligned}$$

with

$$\mathbf{d}^1 = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \quad \text{and} \quad \mathbf{V}^0 = \begin{pmatrix} \boldsymbol{\sigma}_{11}^2 & 0 & 0 \\ 0 & \boldsymbol{\sigma}_{22}^2 & 0 \\ 0 & 0 & \boldsymbol{\sigma}_{33}^2 \end{pmatrix}.$$

As a consequence of these assumptions, the ODEs (29)-(36) substantially simplify and can be explicitly solved, see e.g. [32]. For example, for every $0 \le t \le u \le T$ and every choice of $(i, j) \in \mathcal{I}$, the ODEs (29) and (30) now are given by

$$\begin{aligned} \frac{d\boldsymbol{\beta}_u^{i,j}}{dt}(t) &= \mathbf{b}^{i,j} - \mathbf{d}^1 \boldsymbol{\beta}_u^{i,j}(t)\,, \\ \frac{d\alpha_u^{i,j}}{dt}(t) &= c^{i,j} - \frac{1}{2}(\boldsymbol{\beta}_u^{i,j}(t))^\intercal \mathbf{V}^0 \boldsymbol{\beta}_u^{i,j}(t)\,, \end{aligned}$$

with terminal conditions $\alpha_u^{i,j}(u) = 0$ and $\boldsymbol{\beta}_u^{i,j}(u) = 0$. For the k-th components of $\boldsymbol{\beta}_u^{i,j}(t)$, i.e. ${}^{(k)}\beta_u^{i,j}(t)$, one obtains the explicit solutions

$${}^{(k)}\beta_u^{i,j}(t) = \frac{b_k^{i,j}}{d_{kk}^1}\left(1 - \exp(d_{kk}^1(u - t))\right)\,, \qquad k \in \{1, 2, 3\}\,,$$

and

$$\alpha_u^{i,j}(t) = -\, c^{i,j}(u-t) + \frac{1}{2}\sum_{k=1}^{3} \frac{\sigma_k^2 (b_k^{i,j})^2}{(d_{kk}^1)^2} \Bigg[(u-t) - \frac{2}{d_{kk}^1}\Big(1 - \exp(d_{kk}^1(u-t))\Big) + \frac{1}{2d_{kk}^1}\Big(1 - \exp(2d_{kk}^1(u-t))\Big) \Bigg].$$

Similarly, the ODEs of the remaining equations (31)-(36) can be derived.

Next, we need to specify all remaining parameters, i.e. $\mathbf{b}^{j,k}, c^{j,k}$ for $(j,k) \in \mathcal{I}$, the components $\mathbf{A}, \mathbf{Y}, \mathbf{Z}$ of the quadruple $(X; \mathbf{A}; \mathbf{Y}; \mathbf{Z})$ and the discounting factor S_0. For simplicity, we set $S_0 = e^{-rt}$ with a constant interest rate r and choose

$$\mathbf{b}^{1,2} = \mathbf{b}^{2,1} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}, \mathbf{b}^{1,3} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{b}^{2,3} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}, c^{1,2} = c^{2,1} = 1\,, c^{1,3} = 0.1\,, c^{2,3} = 0.2\,.$$

We have chosen the components of $\mathbf{b}^{j,k}$ to be equal in order to emphasize the general dependence of the process $\psi_t^{j,k}$ in (19) on $\boldsymbol{\mu}_t$ and not on the specific linear combination.

Furthermore, similar to Example 2, we set

$$\mathbf{Z} := \begin{pmatrix} 0 & 0 & z \\ 0 & 0 & z \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{Y} \equiv 0\,, \quad \text{and} \quad \mathbf{A}_t = \mathbf{C}_t - \mathbf{P}_t\,, \quad t \in [0,T]\,.$$

with $\mathbf{C}_t$ and $\mathbf{P}_t$ representing the cumulative state-dependent claim payments (e.g. annuities) and insurance premiums up to maturity, respectively, and z the "immediate" claim payment if the insured person dies. More precisely, we assume monthly equal insurance premiums and claim payments equal to 1, if the insured person is in the states 1 (*healthy*) or 2 (*sick/unfit for work*), respectively, and which are paid at the end of each month in a proportional way. Let the payment dates $0 < T_1 < T_2 < \ldots$ denote the final day of each month, i.e. $T_i = \frac{i}{12}$, then we set

$$\boldsymbol{\nu}(t) = \begin{pmatrix} 1/\Delta t \\ -1/\Delta t \\ 0 \end{pmatrix}$$

with $\Delta t = 1/12$ and $\mathbf{A}_t = \int_0^t \boldsymbol{\nu}(s)ds$.

Clearly, with these specifications Assumption 6 is fulfilled. Finally, assuming that the insured persons process X starts in the state 1 (*healthy*) at $t = 0$, one obtains $\mathbf{H}_0^\intercal = (1,0,0)$. Based on the explicit result for $F^i(0,T)$ from equation (41), now we are able to calculate the expected (discounted) cumulative payment $\mathbb{E}[\widehat{D}_T]$ from Lemma 3 in $t = 0$, depending on the claim payment amount z, which is payed when the insured person dies, and on the constant interest rate r. The corresponding integrals involved in equation (41) are approximated using the `integrate` function in the statistical software program R ([33]).

Figure 2 illustrates the expected cumulative payment $\mathbb{E}[\widehat{D}_T]$ as a function of the claim payment amount z, a time horizon of one year ($T = 1$) and three different values of the constant interest rate r. In Figure 3, the expected cumulative payment $\mathbb{E}[\widehat{D}_T]$ are displayed against the time horizon T, for three different values of the claim payment amount z and for a constant interest rate $r = 0.1$.

Acknowledgments: The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no [228087].

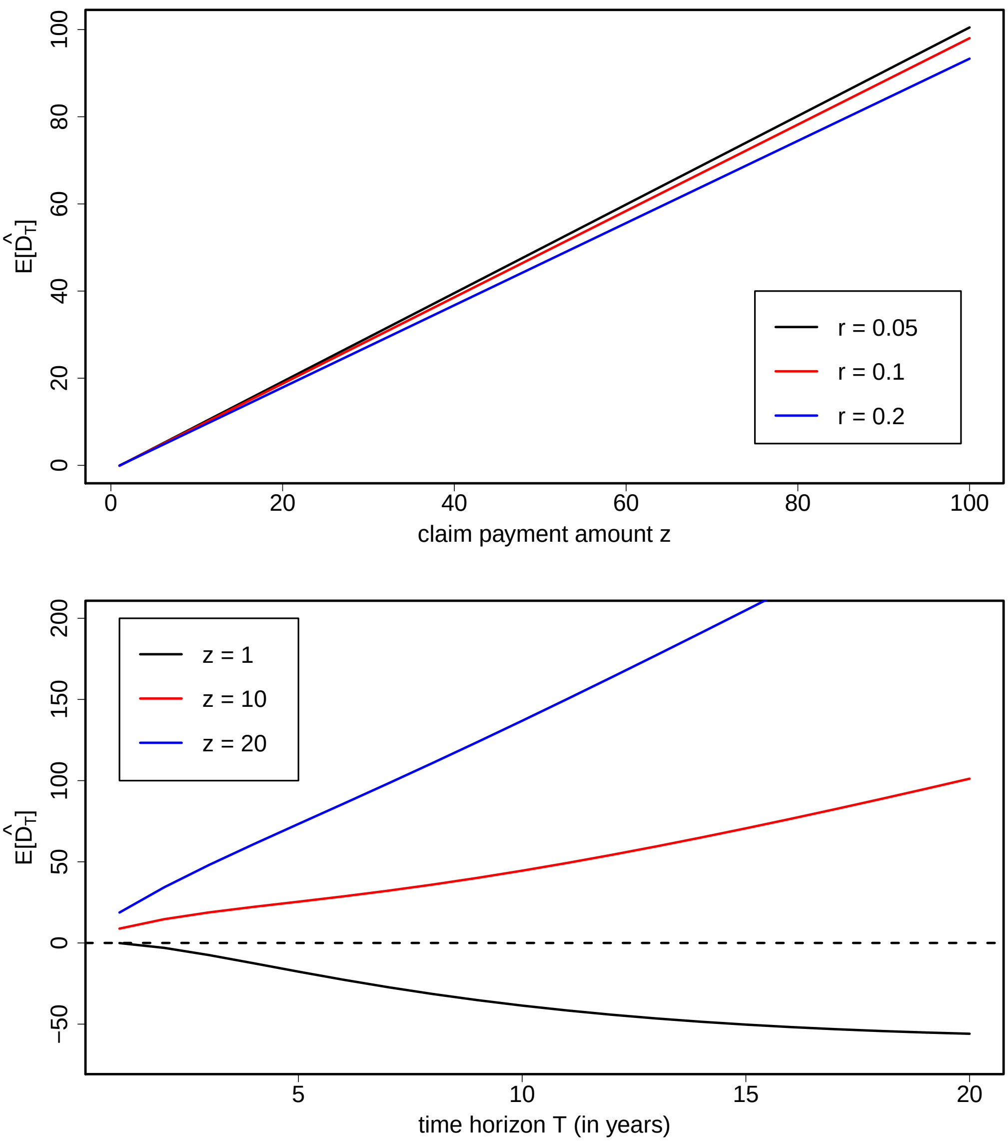

Figure 3. Expected (discounted) cumulative payments $\mathbb{E}[\widehat{D}_T]$ as a function of the time horizon T (in years) and three different claim payment amounts z and for a constant interest rate $r = 0.1$.

Author Contributions: All authors have contributed equally to all aspects of this paper

Conflicts of Interest: The authors declare no conflict of interest.

4. Conclusion

In this paper we consider pricing and hedging of general insurance contracts by means of risk minimization. We model the individual progress in time of visiting an insurance policy's states by using $\mathbb{F}$-doubly stochastic Markov chains. In this way we are able to consider a multi-state setting to describe different types of insurance benefits and to include the influence of market conditions and external risk factors on the evolution of the insured person among the policy's states as well as on the insurance benefits, when they are linked to some financial performance. We explicitly provide the risk-minimizing strategy for an insurance contract in a Brownian financial market setting and specify it within an affine structure for the intensity. The results are illustrated by a numerical example, which shows how this technical setting can actually be easily implemented.

Appendix A. $\mathbb{F}$-doubly stochastic Markov chains

In this section we introduce briefly to some basic properties of $\mathbb{F}$-doubly stochastic Markov chains, which we are going to use in the sequel. Main references are [12] and [21].

On a probability space $(\Omega, \mathcal{G}, \mathbb{P})$, let $X = (X_t)_{t\in[0,T]}$ be a right-continuous stochastic process with state space $\mathcal{K} := \{1, ..., N\}$. We denote by $\mathbb{F}^X$ the filtration generated by X, i.e. $\mathcal{F}_t^X = \sigma(X(u) : u \leq t)$ for all $t \in [0, T]$, and consider the filtration $\mathbb{G}$ to be the enlargement of $\mathbb{F}^X$ through some reference filtration $\mathbb{F}$, i.e. we assume $\mathcal{G}_t = \mathcal{F}_t^X \vee \mathcal{F}_t$ for all $t \in [0, T]$. Further, we set $\widetilde{\mathcal{G}}_t = \mathcal{F}_t^X \vee \mathcal{F}_T$, $t \in [0, T]$ and assume that all filtrations satisfy the usual conditions of completeness and right-continuity, see [28].

Definition A.1. A process $X = (X_t)_{t\in[0,T]}$ is called an $\mathbb{F}$-doubly stochastic Markov chain with state space $\mathcal{K}$, if there exists a family of stochastic matrices

$$\mathbf{P}(s,t) = [p_{j,k}(s,t)]_{j,k\in\mathcal{K}}, \quad 0 \leq s \leq t \leq T,$$

such that

(1) the matrix $\mathbf{P}(s,t)$ is $\mathcal{F}_t$-measurable, and $\mathbf{P}(s,.)$ is progressively measurable,
(2) for every $j,k \in \mathcal{K}$ we have

$$\mathbb{1}_{\{X_s=j\}}\mathbb{P}(X_t = k \mid \widetilde{\mathcal{G}}_s) = \mathbb{1}_{\{X_s=j\}}p_{j,k}(s,t)\ . \tag{42}$$

The process $\mathbf{P}$ is called the *conditional transition probability process* of X.

By definition A.1 we can see that the class of $\mathbb{F}$-doubly stochastic Markov chains contains Markov chains, compound Poisson processes with integer-valued jumps, Cox processes as in [34] and processes of rating migration as in [35]. The adjective "double" refers to the fact that there are two sources of uncertainty in their definition. We remark that an $\mathbb{F}$-doubly stochastic Markov chain is a different object than a doubly stochastic Markov chain which is a Markov chain with a doubly stochastic transition matrix. Furthermore, in [12] it is shown that $\mathbb{F}$-doubly stochastic Markov chains are a subclass of $\mathbb{F}$-conditional $\mathbb{G} = \mathbb{F}^X \vee \mathbb{F}$ Markov chains. In particular, $\mathbb{F}$-doubly stochastic Markov chains behave like time inhomogeneous Markov chains conditioned on $\mathcal{F}_T$, i.e. if we know all the information concerning the underlying risk factors.

Definition A.2. We say that a state $N \in \mathcal{K}$ is an absorbing state, if $p_{N,j}(s,t) = 0$ for all $0 \leq s < t \leq T$ and all $j \in \mathcal{K}$ with $j \neq N$.

Proposition A.3. *Let X be an $\mathbb{F}$-doubly stochastic Markov chain with transition matrices $\boldsymbol{P}(s,t)$, then for every $0 \leq s < t < u \leq T$ we have*

$$\boldsymbol{P}(s,u) = \boldsymbol{P}(s,t)\boldsymbol{P}(t,u) \quad a.s. \tag{43}$$

Proposition A.4. *If X is an $\mathbb{F}$-doubly stochastic Markov chain, then for every bounded, $\mathcal{F}_T$-measurable random variable Y and for each $t \in [0,T]$, we have*

$$\mathbb{E}[Y \mid \mathcal{G}_t] = \mathbb{E}[Y \mid \mathcal{F}_t]. \tag{44}$$

Property (44) is well-known in the context of survival analysis and credit risk as *hypothesis (H)* or *immersion property*. According to Proposition A.4, $\mathbb{F}$-martingales remain martingales with respect to the enlarged filtration $\mathbb{G}$. If we think of a martingale as a process describing a fair game, this property means that the additional information contained in $\mathbb{G}$ does not change the valuation of processes which are considered fair by taking in account only the information $\mathbb{F}$.

Another property, which makes the class of $\mathbb{F}$-doubly stochastic Markov chains interesting for applications is that they may admit matrix-valued stochastic intensity processes in the following sense.

Definition A.5. An $\mathbb{F}$-doubly stochastic Markov chain X with state space $\mathcal{K}$ is said to have an *intensity*, if there exists an $\mathbb{F}$-adapted matrix-valued stochastic process $\mathbf{\Psi} = (\mathbf{\Psi}_t)_{t\in[0,T]}$ with $\mathbf{\Psi}_t = \left[\psi_t^{j,k}\right]_{j,k\in\mathcal{K}}$ such that

1) $\mathbf{\Psi}$ is integrable, i.e.

$$\int_{]0,T]} \sum_{j\in\mathcal{K}} |\psi_s^{j,j}| ds < \infty \,. \tag{45}$$

2) $\mathbf{\Psi}$ satisfies the following conditions:

$$\psi_t^{j,k} \geq 0 \quad \forall j,k \in \mathcal{K}, j \neq k, \quad \psi_t^{j,j} = -\sum_{k\neq j} \psi_t^{j,k} \quad \forall j \in \mathcal{K},\ t \in [0,T], \tag{46}$$

$$\mathbf{P}(v,t) - \mathbb{I} = \int_{]v,t]} \mathbf{\Psi}(u)\mathbf{P}(u,t)du \quad \forall v \leq t \quad \text{(Kolmogorov backward equation)},$$

$$\mathbf{P}(v,t) - \mathbb{I} = \int_{]v,t]} \mathbf{P}(v,u)\mathbf{\Psi}(u)du \quad \forall v \leq t \quad \text{(Kolmogorov forward equation)}. \tag{47}$$

A process $\mathbf{\Psi}$, satisfying the above conditions, is called an *intensity* of the $\mathbb{F}$-doubly stochastic Markov chain X.

Theorem A.6. *Let $(\widetilde{\mathbf{\Psi}}_t)_{t\in[0,T]}$ be an $\mathbb{F}$-adapted $N \times N$ matrix-valued stochastic process, satisfying the conditions (45) and (46) of Definition A.5. Then there exists an $\mathbb{F}$-doubly stochastic Markov chain X with intensity $(\widetilde{\mathbf{\Psi}}_t)_{t\in[0,T]}$.*

For $j \in \mathcal{K}$, let

$$H_t^j := \mathbb{1}_{\{X_t=j\}}\,, \quad t \in [0,T]\,, \tag{48}$$

be the indicator function for X, being in state j at time t and denote by $\mathbf{H}_t = (H_t^1, ..., H_t^N)^\intercal$ the corresponding N-variate vector. Moreover, for $j,k \in \mathcal{K}$, $j \neq k$, let $N^{jk} = (N_t^{jk})_{t\in[0,T]}$ with

$$N_t^{jk} := \int_{]0,t]} H_{u-}^j dH_u^k = \sum_{0<u\leq t} H_{u-}^j \triangle H_u^k\,, \tag{49}$$

define the counting processes of the jumps of X from state j to k up to time t, $t \in [0,T]$.

The following theorem provides a martingale characterization of $\mathbb{F}$-doubly stochastic Markov chains and is the core connection of the theory of $\mathbb{F}$-doubly stochastic Markov chains and the counting process theory, underlying for example several estimation schemes for intensity processes, see [13].

Theorem A.7. *Let $X = (X_t)_{t\in[0,T]}$ be a stochastic process with state space $\mathcal{K}$ and $\mathbf{\Psi} = (\mathbf{\Psi}_t)_{t\in[0,T]}$ be a matrix-valued process, satisfying (45) and (46) of Definition A.5. The following conditions are equivalent:*

i) X is an $\mathbb{F}$-doubly stochastic Markov chain.
ii) The process $\mathbf{M} = (\mathbf{M}_t)_{t\in[0,T]}$ with

$$\mathbf{M}_t := \mathbf{H}_t - \int_{]0,t]} \mathbf{\Psi}_u^\intercal \mathbf{H}_u du\,, \tag{50}$$

is a $\widetilde{\mathbb{G}}$-local martingale.

iii) *For $j,k \in \mathcal{K}$, $j \neq k$, the processes $M^{jk} = (M_t^{jk})_{t\in[0,T]}$ with*

$$M_t^{jk} := N_t^{jk} - \int_{]0,t]} H_u^j \psi_u^{j,k} du \tag{51}$$

337 *are $\widetilde{\mathbb{G}}$-local martingales.*

iv) *The process $\boldsymbol{L} = (\boldsymbol{L}_t)_{t\in[0,T]}$ with*

$$\boldsymbol{L}_t := \boldsymbol{Q}(0,t)^\intercal \boldsymbol{H}_t \ ,$$

is a $\widetilde{\mathbb{G}}$-local martingale. Here $\boldsymbol{Q}(0,t)$ is a unique solution to the random integral equation

$$d\boldsymbol{Q}(0,t) = -\boldsymbol{\Psi}_t \boldsymbol{Q}(0,t)dt, \ \boldsymbol{Q}(0,0) = \mathbb{I}, \tag{52}$$

Note that then

$$\boldsymbol{L}_t = \boldsymbol{H}_0 + \int_{]0,t]} \boldsymbol{Q}^\intercal(0,u) d\boldsymbol{M}_u, \quad t \in [0,T] \,. \tag{53}$$

338 **Remark A.1.** 1) For every $t \in [0,T]$, the matrix $\mathbf{Q}(0,t)$ is the unique inverse matrix of $\mathbf{P}(0,t)$.
339 More generally, for $0 \leq s \leq t \leq T$, we denote by $\mathbf{Q}(s,t)$ the unique inverse matrix of $\mathbf{P}(s,t)$. The
340 existence and further properties of the family $\mathbf{Q}(s,t)$ is given in [12].

It follows immediately from (43) that for every $0 \leq s < t < u \leq T$, we have

$$\mathbf{P}(t,u) = \mathbf{Q}(s,t)\mathbf{P}(s,u) \,. \tag{54}$$

341 2) As the processes $\mathbf{M}$, $\mathbf{L}$ and M^{jk}, $j,k \in \mathcal{K}$, $j \neq k$, are $\mathbb{G}$-adapted, they are also $\mathbb{G}$-local martingales.

Corollary A.8. *For every $j,k \in \mathcal{K}$, $j \neq k$, and for every $t \in [0,T]$ we have*

$$[M^{jk}]_t = N_t^{jk} \,, \tag{55}$$

$$\langle M^{jk} \rangle_t = \int_{]0,t]} H_u^j \psi_{j,k}(u) du \,. \tag{56}$$

Moreover, with $M_t^j = H_t^j - \int_{]0,t]} \sum_{k=1}^N \psi_u^{k,j} H_u^k du$, $j \in \mathcal{K}, t \in [0,T]$, we have

$$[M^j]_t = \sum_{0<s\leq t} (\Delta H_s^j)^2 = \sum_{\substack{k=1\\k\neq j}}^N (N_t^{kj} + N_t^{jk}) \,,$$

$$\langle M^j \rangle_t = \sum_{\substack{k=1\\k\neq j}}^N \int_{]0,t]} H_u^k \psi_{k,j}(u) du - \int_{]0,t]} H_u^j \psi_{j,j}(u) du \,. \tag{57}$$

342 **Proof.** Equalities (55) and (56) follow directly by the definition of M^{jk} in (51).

Moreover, we observe that $\left(\int_{]0,t]} \sum_{k\in\mathcal{K}} H_u^k \psi_u^{k,j} du\right)_{t\in[0,T]}$ is a continuous finite variation process. It follows that

$$[M^j]_t = \sum_{0<s\leq t} (\Delta H_s^j)^2 = \sum_{\substack{k=1\\k\neq j}}^N \left(N_t^{kj} + N_t^{jk}\right) \,,$$

as $\sum_{0<s\leq t}(\Delta H^j_s)^2$ counts the jumps of X into and out of the state j up to time t. As $\left(\int_{]0,t]} H^j_u \psi^{j,k}_u du\right)_{t\in[0,T]}$ is the compensator of N^{jk} it follows that

$$\langle M^j \rangle_t = \sum_{\substack{k=1\\k\neq j}}^N \left(\int_{]0,t]} H^k_u \psi^{k,j}_u du + \int_{]0,t]} H^j_u \psi^{j,k}_u du \right)$$
$$= \sum_{\substack{k=1\\k\neq j}}^N \int_{]0,t]} H^k_u \psi^{k,j}_u du - \int_{]0,t]} H^j_u \psi^{j,j}_u du \,,$$

where the last equality follows from (46). This ends the proof. □

Proposition A.9. *Let X be an $\mathbb{F}$-doubly stochastic Markov chain with intensity and jump times $\tau_0 := 0$ and*

$$\tau_k := \inf\{\tau_{k-1} < t \leq T : X_t \neq X_{\tau_{k-1}}\}. \tag{58}$$

Then every jump time τ_k, $k \geq 1$, avoids $\mathbb{F}$-stopping times, i.e. $\mathbb{P}(\tau_k = \varrho) = 0$ for every $\mathbb{F}$-stopping time ϱ, provided that $\tau_k < \infty$ a.s..

The following proposition is the crucial result in order to compute the risk-minimizing strategies for general insurance claims which we provide in Section 3.

Proposition A.10. *Let X be an $\mathbb{F}$-doubly stochastic Markov chain. Then the local martingale $\boldsymbol{M}$, given in (50), is orthogonal to every $\mathbb{F}$-local martingale N, in the sense that for each $i \in \mathcal{K}$, the product $M^i N$ is a $\mathbb{G}$-local martingale.*

Proof. First note that M^i is a finite variation local martingale. Its sequence $(\widetilde{\tau}^i_k)_{k\geq 0}$ of jump times with $\widetilde{\tau}^i_0 := 0$ and

$$\widetilde{\tau}^i_k := \inf\{t > \widetilde{\tau}^i_{k-1} | M^i_{t-} \neq M^i_t\}\,, \quad k \geq 1\,,$$

is a subsequence of the jump times $(\tau_j)_{j\geq 0}$ of X, as given by (58). As the jump times of the càdlàg local martingale N are $\mathbb{F}$-stopping times, the processes M^i and N have almost surely no common jumps due to Proposition A.9.

This implies that for all $t \in [0, T]$ we have

$$[M^i, N]_t = M_0 N_0 + \sum_{0<s\leq t} \Delta M^i_s \Delta N_s = 0$$

and ends the proof. □

Remark A.2. It is easily seen that hazard-rate models, as applied frequently in the context of credit risk or life insurance, are particular examples of $\mathbb{F}$-doubly stochastic Markov chains, provided they satisfy hypothesis (H). A thorough description of this relation is given in [24].

Appendix B. Risk-minimization for payment processes

The following survey of risk-minimization for payment processes is borrowed to some extend from [1], as well as [22]. As in the foregoing sections, we provide the results with respect to a general numéraire process S^0 such that one could also consider e.g. the $\mathbb{P}$-numéraire portfolio as discount factor, see [25]. The results base on the proofs, given in [4] for the case of European type contingent claims and in [5], [36] and [37, Chapter 4] for the case of payment processes.

In the market model, defined in Section 3.2, we would like to find a hedging strategy for a $\mathbb{G}$-adapted, square integrable payment process $\widehat{D} = (\widehat{D}_t)_{t\in[0,T]}$, representing the cumulative discounted payments up to time t, $t \in [0,T]$.

Note that if an *undiscounted* cumulative payment stream $D = (D_t)_{t\in[0,T]}$ is a stochastic process of finite variation and we have $\int_{[0,T]} \frac{1}{S_u^0} d|D|_u < \infty$ with $|D|$ denoting the absolute variation process of D, then $\widehat{D}$ is given as

$$\widehat{D}_t = \int_{[0,t]} \frac{1}{S_u^0} dD_u \,. \tag{59}$$

Definition B.1. If $\int_{[0,T]} \frac{1}{S_u^0} d|D|_u < \infty$, then the value process $U^D = (U_t^D)_{t\in[0,T]}$ of a payment process D is defined as

$$U_t^D := S_t^0 \mathbb{E}\left[\widehat{D}_T \middle| \mathcal{G}_t\right] = S_t^0 \mathbb{E}\left[\int_{[0,T]} \frac{1}{S_u^0} dD_u \middle| \mathcal{G}_t\right] . \tag{60}$$

Since the market is not necessarily complete, it is in general not possible to find a self-financing hedging strategy that perfectly replicates the discounted cumulative payment process $\widehat{D}$. In this context, the idea of risk-minimization is to relax the self-financing assumption, allowing for a wider class of admissible strategies, and to find an optimal hedging strategy with "minimal risk" within this class of strategies that perfectly replicate $\widehat{D}$.

For the local martingale $\mathbf{S}$, we denote

$$L^2(\mathbf{S}) := \left\{ \boldsymbol{\xi} = (\xi_t^1, ..., \xi^d)_{t\in[0,T]}^\intercal \,\middle|\, \boldsymbol{\xi} \text{ is } \mathbb{G}\text{-predictable}, \left(\mathbb{E}\left[\int_{[0,T]} \boldsymbol{\xi}_s^\intercal d[\mathbf{S}]_s \boldsymbol{\xi}_s\right]\right)^{\frac{1}{2}} < \infty \right\} . \tag{61}$$

It is well known that for every $\boldsymbol{\xi} \in L^2(\mathbf{S})$, the process $\left(\int_{[0,t]} \boldsymbol{\xi}_s^\intercal \, d\widehat{\mathbf{S}}\right)_{t\in[0,T]}$ is a square integrable martingale.

In the following we now explain how to find the risk-minimizing strategy and explain in what sense this strategy is optimal. We begin with some definitions.

Definition B.2. An L^2-*strategy* is a pair $\boldsymbol{\varphi} = (\boldsymbol{\xi}, \xi^0)$, such that $\boldsymbol{\xi} \in L^2(\widehat{S})$ and ξ^0 is a real-valued $\mathbb{G}$-adapted process, such that the discounted portfolio value process

$$\widehat{V}_t^{\varphi} = \boldsymbol{\xi}_t^\intercal \widehat{S}_t + \xi_t^0, \quad t \in [0,T],$$

is right-continuous and square integrable.

For an L^2-strategy $\boldsymbol{\varphi}$ the *discounted (cumulative) cost process* $\widehat{C}^{\varphi}$ is defined as

$$\widehat{C}_t^{\varphi} := \widehat{V}_t^{\varphi} - \int_{]0,t]} \boldsymbol{\xi}_s^\intercal \, d\, \widehat{S}_s + \widehat{D}_t, \quad t \in [0,T],$$

describing the accumulated costs of the trading strategy $\boldsymbol{\varphi}$ during $[0,t]$, including the payments $\widehat{D}_t$. Note that $\widehat{V}_t^{\varphi}$ should therefore be interpreted as the discounted value of the portfolio $\boldsymbol{\varphi}_t$ held at time t *after* the payments $\widehat{D}_t$ have been made. In particular, $\widehat{V}_t^{\varphi}$ is the discounted value of the portfolio upon settlement of all liabilities, and a natural condition is then to restrict to 0-*admissible strategies*, satisfying

$$\widehat{V}_T^{\varphi} = 0 \quad \mathbb{P}\text{-a.s.}$$

The *risk process* of $\boldsymbol{\varphi}$ is given by the conditional expected value of the squared future costs

$$R_t^{\varphi} = \mathbb{E}[(\widehat{C}_T^{\varphi} - \widehat{C}_t^{\varphi})^2 \,|\, \mathcal{G}_t], \quad t \in [0,T], \tag{62}$$

and is taken as a measure of the hedger's remaining risk. We would like to find a trading strategy that minimizes the risk in the following sense.

Definition B.3. An L^2-strategy $\boldsymbol{\varphi} = (\boldsymbol{\xi}, \xi^0)$ is called *risk-minimizing* for the discounted payment process $\widehat{D}$, if for any L^2-strategy $\widetilde{\boldsymbol{\varphi}} = (\widetilde{\boldsymbol{\xi}}, \widetilde{\xi}^0)$ such that $\widehat{V}_T^{\widetilde{\varphi}} = \widehat{V}_T^{\varphi} = 0$ $\mathbb{P}$-a.s., we have

$$R_t^{\varphi} \leq R_t^{\widetilde{\varphi}} \quad \mathbb{P}\text{-a.s.}, \; t \in [0,T],$$

i.e., φ minimizes pointwise the risk process introduced in (62).

The key to finding the strategy with minimal risk in our setting is the so-called Galtchouk-Kunita-Watana decomposition.

Definition B.4. Given a square integrable martingale $\widehat{U} \in \mathcal{M}^2$ and the local martingale $\widehat{\mathbf{S}}$, the *Galtchouk-Kunita-Watanabe decomposition* for $\widehat{U}$ with respect to $\widehat{\mathbf{S}}$ is given as

$$\widehat{U}_t = \widehat{U}_0 + \int_{]0,t]} (\boldsymbol{\vartheta}_s^U)^\intercal \, d\widehat{\mathbf{S}}_s + L_t^U, \quad t \in [0,T], \tag{63}$$

where $\boldsymbol{\vartheta}^U \in L^2(\widehat{\mathbf{S}})$ and $L^{\widehat{D}}$ is a square integrable martingale null at 0 which is strongly orthogonal to the space $\mathcal{I}^2(\widehat{\mathbf{S}})$ of all integral processes $\left(\int_{[0,t]} \boldsymbol{\psi}_s^\intercal \, d\widehat{\mathbf{S}}\right)_{t\in[0,T]}$ with $\boldsymbol{\psi} \in L^2(\widehat{\mathbf{S}})$.

It is well known that the set $\mathcal{I}^2(\widehat{\mathbf{S}})$ is a closed stable subset of $\mathcal{M}_0^2$, the set of all square integrable martingales, zero at 0.

Due to Jensen's inequality and the fact that $\widehat{D}$ is square-integrable, the discounted value process $\widehat{U}^D = \frac{U^D}{S^0}$ is a square-integrable martingale and may be decomposed according to (63).

Theorem B.5. *For every (discounted) square integrable payment stream $\widehat{D}$, there exists a unique 0-admissible risk-minimizing L^2-strategy $\boldsymbol{\varphi} = (\boldsymbol{\xi}, \xi^0)$, given by*

$$\begin{aligned} \boldsymbol{\xi}_t &:= \boldsymbol{\xi}_t^{\widehat{D}}, \\ \xi_t^0 &:= \widehat{U}_t^D - \widehat{D}_t - (\boldsymbol{\xi}_t^{\widehat{D}})^\intercal \widehat{\mathbf{S}}_t, \end{aligned}$$

with discounted portfolio value process

$$\widehat{V}_t^{\varphi} = \mathbb{E}[\widehat{D}_T \,|\, \mathcal{G}_t] - \widehat{D}_t = \mathbb{E}[\widehat{D}_T \,|\, \mathcal{G}_0] + \int_{]0,t]} \boldsymbol{\xi}_s^\intercal \, d\widehat{\mathbf{S}}_s + L_t^{\widehat{D}} - \widehat{D}_t,$$

discounted optimal cost process

$$\widehat{C}_t^{\varphi} = \mathbb{E}[\widehat{D}_T \,|\, \mathcal{G}_0] + L_t^{\widehat{D}} = C_0^{\varphi} + L_t^{\widehat{D}},$$

and minimal risk process

$$R_t^{\varphi} = \mathbb{E}[(L_T^{\widehat{D}} - L_t^{\widehat{D}})^2 \,|\, \mathcal{G}_t],$$

$t \in [0,T]$, where $\boldsymbol{\xi}^{\widehat{D}}$ and $L^{\widehat{D}}$ are given by (63) for the square integrable martingale $\widehat{U}^D$.

Proof. See [26] for the single payoff case or [5] and [36] for the extension to the case of payment streams. □

Note that the approach, described above, relies heavily on the fact that the discounted asset prices are local martingales under the measure $\mathbb{P}$. In a more general setting, when the vector of discounted asset is a semimartingale under $\mathbb{P}$, one has to apply the *local* risk-minimization technique, see [36] or [37, Chapter 4]. For more information on (local) risk-minimization and other quadratic hedging approaches we refer to the survey paper of [26].

Bibliography

1. Biagini, F.; Schreiber, I. Risk minimization for life insurance liabilities. *SIAM Journal on Financial Mathematics* **2013**, *4*, 243 – 264.
2. Biagini, F.; Rheinländer, T.; Widenmann, J. Hedging mortality claims with longevity bonds. *ASTIN Bulletin* **2013**, *43*, 123–157.
3. Blake, D.; Cairns, A.J.G.; Dowd, K. The birth of the life market. *Asia-Pacific Journal of Risk and Insurance* **2008**, *3*, 6–36.
4. Föllmer, H.; Sondermann, D. Hedging of non-redundant contingent claims. In *Contributions to Mathematical Economics*; Mas-Colell, A.; Hildenbrand, W., Eds.; North Holland, 1986; pp. 205–223.
5. Møller, T. Risk-Minimizing Hedging Strategies for Insurance Payment Processes. *Finance and Stochastics* **2001**, *5*, 419–446.
6. Ansel, J.P.; Stricker, C. Décomposition de Kunita-Watanabe. In *Séminaire de Probabilités XXVII*; Lecture Notes in Mathematics Volume 1557, Springer, 1993.
7. Kunita, H.; Watanabe, S. On square integrable martingales. *Nagoya Mathematical Journal* **1967**, *30*, 209–245.
8. Hoem, J. Markov Chain Models in Life Insurance. *Blätter der Deutschen Gesellschaft für Versicherungs- und Finanzmathematik* **1969**, *9*, 91–107.
9. Norberg, R. Hattendorff's theorem and Thiele's differential equation generalized. *Scandinavian Actuarial Journal* **1992**, *1*, 2–14.
10. Norberg, R. A Theory of Bonus in Life Insurance. *Finance and Stochastics* **1999**, *3*, 373–390.
11. Norberg, R. Optimal hedging of demographic risk in life insurance. *Finance and Stochastics* **2013**, *17*, 197 – 222.
12. Jakubowski, J.; Niewęgłowski, M. A class of $\mathbb{F}$-doubly stochastic Markov chains. *Electronic journal of probability* **2010**, *15*, 1743–1771.
13. Biagini, F.; Groll, A.; Widenmann, J. Intensity-based premium evaluation for unemployment insurance products. *Insurance: Mathematics and Economics* **2013**, *53*, 302–316.
14. Barbarin, J. Heath-Jarrow-Morton modelling of longevity bonds and the risk minimization of life insurance portfolios. *Insurance: Mathematics and Economics* **2008**, *43*, 41–55.
15. Biagini, F.; Cretarola, A. Quadratic Hedging Methods for Defaultable Claims. *Applied Mathematics and Optimization* **2007**, *56*, 425–443.
16. Biagini, F.; Cretarola, A. Local Risk-Minimization for Defaultable Markets. *Mathematical Finance* **2009**, *19*, 669–689.
17. Biagini, F.; Cretarola, A. Local Risk-Minimization for defaultable claims with recovery process. *Applied Mathematics & Optimization* **2012**, *65*, 293–314.
18. Biagini, F.; Rheinländer, T.; Schreiber, I. Risk-minimization for life insurance liabilities with basis risk. *Mathematics and Financial Economics* **2016**, *10*, 151–178.
19. Biagini, F.; Botero, C.; Schreiber, I. Risk-minimization for life insurance liabilities with dependent mortality risk. *accepted on Mathematical Finance* **2015**.
20. Bielecki, T.R.; Jeanblanc, M.; Rutkowski, M. Pricing and trading credit default swaps in a hazard process model. *Annals of Applied Probability* **2008**, *18*, 2495–2529.
21. Jakubowski, J.; Niewęgłowski, M. Pricing and Hedging of Rating-Sensitive Claims Modeled by $\mathbb{F}$-doubly stochastic Markov chains. Advanced Mathematical Methods for Finance; diNunno, G.; ksendal, B.O., Eds. Springer, 2011, pp. 417 – 454.
22. Biagini, F.; Cretarola, A.; Platen, E. Local risk-minimization under the benchmark approach. *Mathematics and Financial Economics* **2014**, *8*, 109–134.

23. Biagini, F. Evaluating hybrid products: the interplay between financial and insurance markets. Seminar on Stochastic Analysis, Random Fields and Applications VII, Centro Stefano Franscini, Ascona, May 2011; Dalang, R.; Dozzi, M.; Russo, F., Eds.; Springer & Birkhäuser: Basel, 2013; pp. 285–304.
24. Widenmann, J. Pricing and Hedging Insurance Contracts in Hybrid Markets. PhD thesis, Ludwig Maximilians University Munich, 2013.
25. Platen, E.; Heath, D. *A Benchmark Approach to Quantitative Finance*; Springer, 2007.
26. Schweizer, M. A guided tour through quadratic hedging approaches. In *Option Pricing, Interest Rates and Risk Management*; Jouini, E.; Cvitanic, J.; Musiela, M., Eds.; Cambridge University Press: Cambridge, UK, 2001; pp. 538–574.
27. Henriksen, L.F.B.; Møller, T. Local risk-minimization with survivor bonds. *Applied Stochastic Models in Business and Industry* **2015**, *31*, 241–263.
28. Protter, P.E. *Stochastic Integration and Differential Equations*, second ed.; Springer, New York, 2005.
29. Biffis, E. Affine processes for dynamic mortality and actuarial valuation. *Insurance: Mathematics and Economics* **2005**, *37*, 443–468.
30. Duffie, D.; Pan, J.; Singleton, K. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* **2000**, *68*, 1343–1376.
31. Luciano, E.; Vigna, E. Mortality Risk via affine stochastic intensities: calibration and empirical relevance. *Belgian Actuarial Bulletin* **2008**, *8*, 5–16.
32. Filipović, D. *Term-Structure Models*; Springer, 2009.
33. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015.
34. Lando, D. On Cox processes and credit risk securities. *Rev. Deriv. Res.* **1998**, *2*, 99–120.
35. Bielecki, T.; Rutkowski, M. *Credit Risk: Modelling, Valuation and Hedging*, second ed.; Springer-Finance, Springer, 2004.
36. Schweizer, M. Local Risk-Minimization for Multidimensional Assets and Payment Streams. *Banach Center Publications* **2008**, *83*, 213–229.
37. Barbarin, J. Valuation, Hedging and the Risk Management of Insurance Contracts. PhD thesis, Université Catholique de Louvain, 2008.