Affine processes and Markov chains: Interest Rate-Dependent Transition Rates in Life Insurance

Lars Frederik Brandt Henriksen
Department of Mathematical Sciences
Universitetsparken 5
DK-2100 Copenhagen Ø
Denmark
lars.brandt@math.ku.dk

This version: Tuesday 18th March, 2014

Abstract. We obtain results for calculations of life insurance reserves where one can model policyholder behavior governed by stochastic intensities within hierarchical Markov chain models. We model the interest rate and the transition rates as affine processes, where we allow for dependence between the interest rate and the transition rates. For many years affine processes have been popular choices when modelling interest rates and mortality rates due to mathematical and computational tractability. Our main example of dependence is interest rate-dependent surrender rates, which are widely acknowledged in the literature, but the presented framework also allows for many other types of dependencies. The paper extends some results of transforms of affine presented Duffie et al. (2000). We find two representations of transforms of affine processes needed for calculation of life insurance reserves and discuss their differences.

Key words: Affine processes, Solvency II, expected policy behavior.

1 Introduction

This paper considers valuation of life insurance contracts within a Markov chain model in the case where the interest rate and the transition rates are dependent stochastic processes. We are working in an affine setup such that we are able to obtain results without needing to solve partial differential equations but only systems of ordinary differential equations as long as we consider models without cycles. Affine processes have for some time been used to model the interest rate and transition rates in life insurance, see e.g. Dahl and Møller (2006) and Biffis (2005). The results in this paper can be seen as a continuation of the work in Duffie et al. (2000). The recent paper Buchardt (2012) addresses some of the same questions as this paper. However, the focus and methods for proving the results in the two papers are quite different. Moreover, this paper contains a generalization to an arbitrary number of transitions within hierarchical Markov chain models and two different representations of transforms of affine processes and discussions of subjects in continuation to these. By a transform of an affine process we mean the conditional expected value of the product of the exponential of the integral of an affine transformation of the process, the exponential of an affine transformation of the process and a polynomial transform of the process.

The surrender option is embedded in many life insurance contracts and the attention to the modelling of surrender rates and valuation of life insurance contracts where one takes into account the surrender option, has increased highly the last years. This is, in particular, motivated by the forthcoming Solvency II legislation, where the lapse risk plays an important role. One way to model the surrender rate is to make it dependent of the interest rate, which is a fairly natural approach as for example shown in Kuo et al. (2003). In De Giovanni (2010) a model for insurance liabilities is applied, where the surrender intensity is split into a rational part and an irrational part. The rational part is modeled as a function of the short rate squared, whereas the rate of the irrational part is assumed be constant. The interest rate-dependent surrender rates form our main example of our general results on dependent interest and transition rates.

The affine processes are widely used for interest rate modelling, credit risk modelling and modelling of mortality intensities because of the calculational tractability. That is, one only needs to solve ordinary differential equations (ODEs) instead of partial differential equations (PDEs) which in general is a much easier task. When we use affine processes to model transition intensities in a multistate Markov chain model the tractability can be a little reduced due to the fact that one needs to solve quite a lot of ODEs. The number of calculations necessary to obtain a result depends highly on the number of possible transitions (between different states) before reaching an absorbing state. Secondarily, it also depends on the dimension of the underlying processes.

The paper is organized as follows: In Section 2 and 3 we motivate the work and introduce the basic insurance Markov chain model, which we are going to consider throughout the rest of the paper. Moreover, we introduce the class of affine processes and state conditions to ensure positivity of these processes. In Section 4 we present the main results which allow us to calculate statewise reserves in general hierarchical Markov models. Section 5 contains a small selection of different subsections relating to the results in Section 4. These are a description of the workload relating to the results in Section 4, a presentation of a special case, where one can allow for cycles in the Markov chain and still be able to benefit from the same results as in the case with no cycles, and a description of the class of processes called linear-quadratic processes.

2 Motivating example

The forthcoming Solvency II legislation demands that the life insurance companies take into account future policyholder behaviour such as the likelihood of lapse during the remaining period of an insurance contract. In general, lapse refers to the exercise of policyholder options. In this paper we will focus at the two most import policyholder options. These are the surrender option and the free policy option, which are quite common features in many life insurance contracts. Moreover, lapse risk is a significant risk for many life insurance companies, and recognizing lapse can change the cash flow and reserves of a company significantly, see e.g. Henriksen et al. (2013) and Buchardt et al. (2013). In the following we will consider valuation of a general life insurance contract including both disability and lapse. The state space for such a contract is shown in Figure 1, where we have disregarded premium resumption, i.e. the possibility of a transition from state *fa* to state *a*.

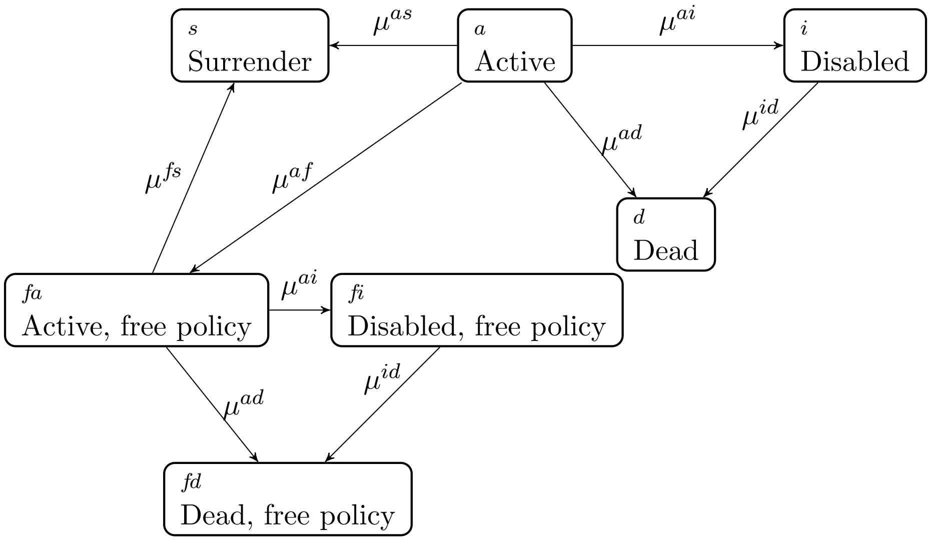

Figure 1: State space for a disability model with lapse.

In comparison with a standard disability model we have added four extra states to allow for lapse. It is not important for the results in this paper that e.g. the intensity from state a to state i is the same as from state *fa* to state *fi*, but for most modelling purposes this is the case. Note that even though the pure jump process, which determines the state of the insurance contract, is Markovian, i.e. the intensities only depend on the current state of the process, the payments do not only depend on the state of the pure jump process. The reason for this is that the payments, after the contract has entered the free policy state, depend on the time of this jump. That is, the introduction of a free policy state gives rise to duration depend payments in the model. The multistate model given in Figure 1 is a hierarchical Markov model, i.e. a model without cycles. Within this model, the state process of the Markov chain can jump at most 3 times before reaching an absorbing state. Since the model has no cycles, the results for the statewise reserves can be obtained by integration of solutions to ODEs in the case, where the interest rate and transition intensities are dependent affine processes. We show this in Section 4.

For the purpose of illustrating the general challenge when evaluating contracts in this type of models and justified by wanting to keep the terms as simple as possible, we consider the (not very realistic) case where the only payment of the insurance contract is a death sum $b_{id}(t)$ triggered by the transition from state *fi* to state *fd*. We allow for the discount factor, r, and the transition intensities to have an affine dependence of an underlying stochastic process X. Under the assumption that the insurance contract only includes one payment, the statewise reserve in state a at time t, denoted $V_a(t)$, is given by (assuming that we are allowed to interchange the expectation and the integrals):

$$\begin{aligned} V_a(t) =& \mathrm{E}\bigg[\int_t^T \int_s^T \int_u^T e^{-\int_t^v r(\tau,X(\tau))d\tau} e^{-\int_t^s \mu^{aa}(\tau,X(\tau))d\tau} e^{-\int_s^u \mu^{ff}(\tau,X(\tau))d\tau} e^{-\int_u^v \mu^{id}(\tau,X(\tau))d\tau} \\ & \times f(s)\mu^{af}(s,X(s))\mu^{ai}(u,X(u))\mu^{id}(v,X(v))b_{id}(v)dvduds\bigg|\mathcal{F}(t)\bigg] \\ =& \int_t^T \int_s^T \int_u^T \mathrm{E}\bigg[e^{-\int_t^v r(\tau,X(\tau))d\tau} e^{-\int_t^s \mu^{aa}(\tau,X(\tau))d\tau} e^{-\int_s^u \mu^{ff}(\tau,X(\tau))d\tau} e^{-\int_u^v \mu^{id}(\tau,X(\tau))d\tau} \\ & \times \mu^{af}(s,X(s))\mu^{ai}(u,X(u))\mu^{id}(v,X(v))\bigg|\mathcal{F}(t)\bigg] f(s)b_{id}(v)dvduds, \end{aligned} \tag{2.1}$$

where $\mu^{aa} = \mu^{ai} + \mu^{ad} + \mu^{af} + \mu^{as}$, $\mu^{ff} = \mu^{ai} + \mu^{ad} + \mu^{fs}$ and f is the free policy factor. The free policy factor is determined at time s and is equal to 1 in this example, since the example does not include any premium payments. More generally, the free policy factor is given by $f(\tau) = \frac{V^*(\tau)}{V^{*+}(\tau)}$, where V^* is the technical reserve, and V^{*+} is the technical reserve where we only take into account the benefits of the contract. This means, that the free policy factor is a time-dependent, deterministic function.

From equation (2.1) one sees that the important thing is to be able to calculate terms of the form

$$\begin{aligned} &\mathrm{E}\bigg[e^{-\int_t^v r(\tau,X(\tau))d\tau} e^{-\int_t^s \mu^{aa}(\tau,X(\tau))d\tau} e^{-\int_s^u \mu^{ff}(\tau,X(\tau))d\tau} e^{-\int_u^v \mu^{id}(\tau,X(\tau))d\tau} \\ &\quad \times \mu^{af}(s,X(s))\mu^{ai}(u,X(u))\mu^{id}(v,X(v))\bigg|\mathcal{F}(t)\bigg] \end{aligned} \tag{2.2}$$

for different assumptions regarding the interest rate and transition intensities. The rest of the paper is about how to calculate terms like this. Note that for the general case with continuous and lump sum payments, we are still able to calculate statewise reserves, if we are able to calculate terms of the form (2.2). That is, being able to calculate terms in the form (2.2) and integrating the results makes it possible to obtain the statewise reserves for all states in Figure 1. More generally, the results make us able to calculate transition probabilities and statewise reserves in arbitrary large hierarchical Markov chain models.

3 Insurance models and affine processes

Let T be a fixed finite time horizon and $(\Omega, \mathcal{F}, P)$ a probability space equipped with a filtration $I\!F = (\mathcal{F}(t))_{0\leq t\leq T}$ satisfying the usual conditions of right-continuity and completeness. In the following we consider a class of stochastic processes called affine processes. Affine processes are stochastic processes with the property that the conditional characteristic functions are exponentially affine. One normally divides affine processes into diffusion processes and jump diffusion processes, where affine diffusion processes are the continuous subgroup of affine jump diffusion processes. For detailed descriptions of affine processes, see e.g. Duffie et al. (2003) or Filipović (2009, Chapter 10).

The dynamics of a d-dimensional time-inhomogeneous affine jump diffusion are given by

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t) + dJ(t), \quad X(0) = x, \tag{3.3}$$

where $\mu(t,x) = K_0(t) + K_1(t)x$ and $\left(\sigma(t,x)\sigma^{\mathrm{tr}}(t,x)\right)_{ij} = H_{0,ij}(t) + \sum_{k=1}^{d} H_{1,ij}^{k}(t)x_k$. Moreover, W is a d-dimensional standard Brownian motion and J is a pure jump process with arrival intensity $\lambda(t,x) = \lambda_0(t) + \lambda_1^{\mathrm{tr}}(t)x$ and jump distribution ν. We assume that the interest rate has the form $r(t,x) = \rho_0(t) + \rho_1^{\mathrm{tr}}(t)x$. In the above equations we have that $\lambda_0, \rho_0 \in \mathbb{R}$, $K_0, \lambda_1, \rho_1 \in \mathbb{R}^d$, $K_1, H_0 \in \mathbb{R}^{d\times d}$ and $H_1 \in \mathbb{R}^{d\times d\times d}$.

The class of processes given by (3.3) is fairly general and living on the statespace $\mathbb{R}^d$. The purpose of this paper is to use affine processes to model the interest rate and the transition intensities of the Markov chain. That is, we want to restrict the class of processes in such a way, that the interest rate is non-negative and such that the transition intensities are *strictly* positive. To obtain non-negativity (in the case of time-homogeneous affine processes without jumps), Filipović (2009, Theorem 10.2) states the following conditions:

- $K_0 \in \mathbb{R}_+^d$.
- K_1 has non-negative off-diagonal elements.
- $H_0 = 0$.
- The only non-zero entries of H_1^k are $H_{1,kk}^k$, which are positive.

This result can be generalized to time-inhomogeneous affine processes (the conditions are the same), see Buchardt (2012). Positive jumps of X do not destroy the positivity, whereas negative jumps in general will. For the rest of the paper, we disregard jumps to keep the notation as simple as possible. Note, however, that the results presented here can easily be generalized from affine diffusion processes to affine jump diffusion processes.

To obtain the strict positivity for the intensities we need an extra condition to be fulfilled. The condition is known as the multivariate Feller condition and is given by . The condition is that $2K_{0,k}(t) \geq H_{1,kk}^k(t)$. That is, the structure of the non-negative affine processes is quite restricted compared to the general case and the structure leads to simplifications of the ODEs for the transforms of the affine processes. For example, the ODEs relating to the standard transform in Duffie et al. (2000) given by

$$\mathrm{E}\left[e^{-\int_t^T \rho_0(s) + \rho_1^{\mathrm{tr}}(s)X(s)ds + u^{\mathrm{tr}}X(T)} \middle| \mathcal{F}(t) \right]$$

simplify to:

$$\frac{\mathrm{d}}{\mathrm{d}t}\beta(t,T) = \rho_1(t) - K_1^{\mathrm{tr}}(t)\beta(t,T) - \frac{1}{2}\beta^{\mathrm{tr}}(t)\mathrm{diag}\left(H_{1,11}^1(t), \ldots, H_{1,dd}^d(t)\right)\beta(t,T), \quad \beta(T,T) = u$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\alpha(t,T) = \rho_0(t) - K_0^{\mathrm{tr}}(t)\beta(t,T), \quad \alpha(T,T) = 0.$$

However, for the rest of the paper, we consider, for sake of completeness, the case where for example H_0 can be different from 0. This also allows us to use e.g. a Vasiček process, which can be negative, to model the interest rate.

4 Main result

In this section we present results enabling us to calculate transforms like (2.2) and similar transforms for more general hierarchical Markov chain models. We give two different representations of the results; a *dense* and a *non-dense* representation. We make these two concepts clear in Section 4.1. In both cases we rely on conditioning by using the tower property.

The structure of this section is the following: First, we present the results for a one-dimensional underlying stochastic process. Afterwards, we state some of the results for the multidimensional setting without giving additional proofs.

4.1 The conditional approach

The literature has for many years investigated the dependence between the interest rate and policy behavior. This possible dependence is called the interest rate hypothesis, see e.g. Dar and Dodds (1989) and references therein. A natural way of modelling this dependence is to model the lapse rates as functions of a one-dimensional interest rate process like in De Giovanni (2010). Due to this reasoning and because we are not gaining much more insight by extending the approach to a multidimensional setting, we introduce what we call "the conditional approach" in a one-dimensional setting. This also makes the results and expressions much more readable. We consider three intermediate time points t_1, t_2 and t_3 fulfilling that $t < t_1 < t_2 < t_3 < T$. Assuming that all the intensities in Figure 1 are given as affine transformations of an underlying one-dimensional stochastic process, we need to calculate a quantity of the form[1]

$$\begin{aligned} \mathrm{E}\Big[& e^{-\int_t^{t_1} f_1(\tau,X(\tau))d\tau} e^{-\int_{t_1}^{t_2} f_2(\tau,X(\tau))d\tau} e^{-\int_{t_2}^{t_3} f_3(\tau,X(\tau))d\tau} e^{-\int_{t_3}^{T} f_4(\tau,X(\tau))d\tau} \\ & \times \mu_1(t_1, X(t_1))\mu_2(t_2, X(t_2))\mu_3(t_3, X(t_3)) \Big| \mathcal{F}(t)\Big], \end{aligned} \tag{4.4}$$

where μ_i and f_i, $i = 1, 2, 3$ are affine functions of X. Note that the term (4.4) is closely related to the calculation of transition probabilities in the Markov chain model, since the triple integral of (4.4) for $f_i = \mu_i$ is exactly a transition probability. By successively using the tower property we obtain that (4.4) equals

$$\begin{aligned} \mathrm{E}\Big[& e^{-\int_t^{t_1} f_1(\tau,X(\tau))d\tau}\mu_1(t_1, X(t_1))\mathrm{E}\Big[e^{-\int_{t_1}^{t_2} f_2(\tau,X(\tau))d\tau}\mu_2(t_2, X(t_2)) \\ & \times \mathrm{E}\left[e^{-\int_{t_2}^{t_3} f_3(\tau,X(\tau))d\tau}\mu_3(t_3, X(t_3))\mathrm{E}\left[e^{-\int_{t_3}^{T} f_4(\tau,X(\tau))d\tau}\Big| \mathcal{F}(t_3)\right]\Big| \mathcal{F}(t_2)\right] \Big| \mathcal{F}(t_1)\Big] \Big| \mathcal{F}(t)\Big]. \end{aligned} \tag{4.5}$$

Towards the end of this section we obtain closed form solutions for (4.5) by first calculating the innermost conditional expected value and then using this result for calculating the next conditional expected value and so on. To do so we need the results of the following two lemmas. In the following, we let ρ_i, u_i, $i = 1, 2$ and g_j, $j = 0, \ldots, n$ be deterministic functions. To shorten the notation, we define by Υ_X the term

$$\Upsilon_X(s, v) := e^{u_0(v)+u_1(v)X(v)-\int_s^v \rho_0(\tau)+\rho_1(\tau)X(\tau)d\tau}.$$

[1] We can e.g. think of the term with the integral from t_3 to T as giving us the possibility of modelling a contract where the death sum is due at time T.

Lemma 4.1 (Dense representation). *Let $s, v \in [t, T]$ with $s < v$. Assume that the system of ODEs stated in the lemma is uniquely solved by the functions $\beta, C_i, i = 0, \dots, n$, and that the following integrability conditions are fulfilled:*

$$E[|\Phi(v)|] < \infty$$

and

$$E\left[\left(\eta^2(t)\right)^{\frac{1}{2}}\right] < \infty \text{ for } \eta(t) = \left(\Phi(t)\beta(t,v) + \Upsilon_X(t,v)\sum_{i=1}^{n} C_i(t,v) i x^{i-1}\right)\sigma(X(t)),$$

where $\Phi(t) = e^{-\int_0^t \rho_0(\tau)+\rho_1(\tau)X(\tau)d\tau} e^{\beta(t,v)X(t)} \sum_{i=0}^n C_i(t,v) X^i(t)$.

Then for $n \in \mathbb{N}$ there exist functions β and C_i for $i = 0, \dots, n$ given as solutions to a system of ODEs such that,

$$E\left[\Upsilon_X(s,v)\sum_{i=0}^{n} g_i(v) X^i(v) \middle| \mathcal{F}(s)\right] = e^{\beta(s,v)X(s)} \sum_{i=0}^{n} C_i(s,v) X^i(s). \tag{4.6}$$

The functions β and $C_i, i = 0, \dots, n$ are given by the following systems of ODEs (suppressing the time arguments of the functions)

$$\begin{aligned}
\frac{\partial}{\partial s}\beta =& \rho_1 - \left(K_1 + \frac{1}{2}H_1\beta\right)\beta, \\
\frac{\partial}{\partial s}C_0 =& \rho_0 C_0 - \left(\beta C_0 + C_1\right)K_0 - \left(\beta^2 C_0 + 2\beta C_1 + 2C_2\right)\frac{1}{2}H_0, \\
\frac{\partial}{\partial s}C_1 =& C_0\left(\frac{\partial}{\partial s}\beta - \rho_1\right) + \rho_0 C_1 - 2C_2K_0 - 2C_2\beta H_0 - C_1K_1 - 3C_3H_0 - C_2H_1 \\
& - \left(K_0 + \frac{1}{2}H_0\beta + H_1\right)C_1\beta - \left(K_1 + \frac{1}{2}H_1\beta\right)C_0\beta.
\end{aligned}$$

For $n \geq 4$ and $i = 2, \dots, n-2$ we have the ODEs given by

$$\begin{aligned}
\frac{\partial}{\partial s}C_i =& - C_{i-1}\left(\frac{\partial}{\partial s}\beta - \rho_1\right) + \rho_0 C_i - (i+1)K_0C_{i+1} - (i+1)H_0C_{i+1}\beta \\
& - iK_1C_i - iH_1C_i\beta - \frac{1}{2}(i+2)(i+1)C_{i+2}H_0 - (i+1)iC_{i+1}\frac{1}{2}H_1 \\
& - \left(K_0 + \frac{1}{2}H_0\beta\right)\beta C_i - \left(K_1 + \frac{1}{2}H_1\beta\right)C_{i-1}\beta\Big).
\end{aligned}$$

For $n \geq 3$ we have the ODE given by

$$\begin{aligned}
\frac{\partial}{\partial s}C_{n-1} =& - C_{n-2}\left(\frac{\partial}{\partial s}\beta - \rho_1\right) + \rho_0 C_{n-1} - n(K_0 + \beta H_0)C_n - (n-1)(K_1 + \beta H_1)C_{n-1} \\
& - n(n-1)C_n\frac{1}{2}H_1 - \left(K_0 + \frac{1}{2}H_0\beta\right)C_{n-1}\beta - \left(K_1 + \frac{1}{2}H_1\beta\right)C_{n-2}\beta.
\end{aligned}$$

Finally, for $n \geq 2$ we have the ODE given by

$$\begin{aligned}
\frac{\partial}{\partial s}C_n =& - C_{n-1}\left(\frac{\partial}{\partial s}\beta - \rho_1\right) + \rho_0 C_n - n(K_1 + \beta H_1)C_n - \left(K_0 + \frac{1}{2}H_0\beta\right)C_n\beta \\
& - \left(K_1 + \frac{1}{2}H_1\beta\right)C_{n-1}\beta.
\end{aligned}$$

The corresponding boundary conditions are

$$\beta(v,v) = u_1(v), \quad C_i(v,v) = e^{u_0(v)}g_i(v), \quad i = 0, \dots, n.$$

Proof of dense representation. The proof of the dense representation follows the lines of proof of Duffie et al. (2000, Proposition 1). We are going to show that

$$\mathcal{H}(s,X(s);v) := \mathrm{E}\left[\Upsilon_X(s,v)\sum_{i=0}^{n} g_i(v)X^i(v)\middle|\,\mathcal{F}(s)\right] = e^{\beta(s,v)X(s)}\sum_{i=0}^{n} C_i(s,v)X^i(s),$$

where β and C_i are fulfilling the ODEs stated in the lemma.

We denote by $\tilde{\mathcal{H}}$ the martingale corresponding to $\mathcal{H}$. The martingale is given by

$$\tilde{\mathcal{H}}(s,X(s);v) = e^{-\int_0^s \rho_0(\tau)+\rho_1(\tau)X(\tau)d\tau}\mathcal{H}(s,X(s);v).$$

By utilizing that the drift of the martingale $\tilde{\mathcal{H}}(s,X(s);v)$ is 0 we get by Itô's formula that $\tilde{\mathcal{H}}$ is given by the following PDE:

$$0 = \frac{\partial}{\partial s}\tilde{\mathcal{H}}(s,x;v) + (K_0(s)+K_1(s)x)\frac{\partial}{\partial x}\tilde{\mathcal{H}}(s,x;v) + \frac{1}{2}(H_0(s)+H_1(s)x)\frac{\partial^2}{\partial x^2}\tilde{\mathcal{H}}(s,x;v). \tag{4.7}$$

We guess that the solution to this PDE is given by the function $\mathcal{G}$:

$$\begin{aligned}\mathcal{G}(s,x;v) &= e^{-\int_0^s \rho_0(\tau)+\rho_1(\tau)X(\tau)d\tau}e^{\beta(s,v)x}\sum_{i=0}^{n} C_i(s,v)x^i\\ &:= \mathcal{E}(s,x;v)\sum_{i=0}^{n} C_i(s,v)x^i,\end{aligned} \tag{4.8}$$

where we have introduced the shorthand notation $\mathcal{E}$ for the exponential terms. We want to find functions β and C_i, such that $\mathcal{G}$ is a martingale. If $\mathcal{G}$ is a martingale, we have that $\mathrm{E}\left[\mathcal{G}(v,x;v)|\,\mathcal{F}(s)\right] = \mathcal{G}(s,x;v)$ and we can get the result by multiplying both sides of (4.8) by $e^{\int_0^s \rho_0(\tau)+\rho_1(\tau)X(\tau)d\tau}$.

Inserting $\mathcal{G}$ in the place of $\tilde{\mathcal{H}}$ in the right-hand side of (4.7) gives us

$$\begin{aligned}&\frac{\partial}{\partial s}\mathcal{G}(s,x;v) + \frac{\partial}{\partial x}\mathcal{G}(s,x;v)(K_0(s)+K_1(s)x) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\mathcal{G}(s,x;v)(H_0(s)+H_1(s)x)\\ &= \mathcal{E}(s,x;v)\left(\left(-\rho_0(s)-\rho_1(s)x+\frac{\partial}{\partial s}\beta(s,v)x\right)\sum_{i=0}^{n} C_i(s,v)x^i + \sum_{i=0}^{n}\frac{\partial}{\partial s}C_i(s,v)x^i\right)\\ &\quad + \mathcal{E}(s,x;v)\left(\beta(s,v)\sum_{i=0}^{n} C_i(s,v)x^i + \sum_{i=1}^{n} C_i(s,v)ix^{i-1}\right)(K_0(s)+K_1(s)x)\\ &\quad + \frac{1}{2}\mathcal{E}(s,x;v)\left(\beta^2(s,v)\sum_{i=0}^{n} C_i(s,v)x^i + 2\beta(s,v)\sum_{i=1}^{n} iC_i(s,v)x^{i-1}\right.\\ &\qquad\qquad \left. + \sum_{i=2}^{n} C_i(s,v)i(i-1)x^{i-2}\right)(H_0(s)+H_1(s)x).\end{aligned}$$

We divide the right-hand side of the above term by $\mathcal{E}(s,x;v)$, suppress the arguments of all the functions, insert indicator functions to take care of small numbers of n, collect terms wrt. x, and obtain that the above term is equal to

$$
\begin{aligned}
&-\rho_0 C_0 + \frac{\partial}{\partial s} C_0 + (\beta C_0 + C_1) K_0 + \left(\beta^2 C_0 + 2\beta C_1 + 2C_2\right) \frac{1}{2} H_0 \\
&+ x\Bigg(C_0 \left(\frac{\partial}{\partial s}\beta - \rho_1 \right) - \rho_0 C_1 + \frac{\partial}{\partial s} C_1 + 2C_2 K_0 + 2C_2\beta H_0 + C_1 K_1 + 3C_3 H_0 + C_2 H_1 \\
&\qquad + \left(K_0 + \frac{1}{2} H_0 \beta + H_1 \right) C_1 \beta + \left(K_1 + \frac{1}{2} H_1 \beta \right) C_0 \beta \Bigg) \\
&+ \mathbf{1}_{\{n\geq 4\}} \sum_{i=2}^{n-2} x^i \Bigg(C_{i-1} \left(\frac{\partial}{\partial s}\beta - \rho_1 \right) - \rho_0 C_i + \frac{\partial}{\partial s} C_i + (i+1) K_0 C_{i+1} + (i+1) H_0 C_{i+1} \beta \\
&\qquad\qquad + i K_1 C_i + i H_1 C_i \beta + \frac{1}{2}(i+2)(i+1) C_{i+2} H_0 + (i+1) i C_{i+1} \frac{1}{2} H_1 \\
&\qquad\qquad + \left(K_0 + \frac{1}{2} H_0 \beta \right) \beta C_i + \left(K_1 + \frac{1}{2} H_1 \beta \right) C_{i-1} \beta \Bigg) \\
&+ \mathbf{1}_{\{n\geq 3\}} x^{n-1} \Bigg(C_{n-2} \left(\frac{\partial}{\partial s}\beta - \rho_1 \right) - \rho_0 C_{n-1} + \frac{\partial}{\partial s} C_{n-1} + n(K_0 + \beta H_0) C_n \\
&\qquad\qquad + (n-1)(K_1 + \beta H_1) C_{n-1} \\
&\qquad\qquad + n(n-1) C_n \frac{1}{2} H_1 + \left(K_0 + \frac{1}{2} H_0 \beta \right) C_{n-1} \beta + \left(K_1 + \frac{1}{2} H_1 \beta \right) C_{n-2} \beta \Bigg) \\
&+ \mathbf{1}_{\{n\geq 2\}} x^n \Bigg(C_{n-1} \left(\frac{\partial}{\partial s}\beta - \rho_1 \right) - \rho_0 C_n + \frac{\partial}{\partial s} C_n + n(K_1 + \beta H_1) C_n + \left(K_0 + \frac{1}{2} H_0 \beta \right) C_n \beta \\
&\qquad\quad + \left(K_1 + \frac{1}{2} H_1 \beta \right) C_{n-1} \beta \Bigg) \\
&+ x^{n+1} C_n \left(\frac{\partial}{\partial s}\beta - \rho_1 + \left(K_1 + \frac{1}{2} H_1 \beta \right) \beta \right).
\end{aligned}
\tag{4.9}
$$

From (4.9) we can see, that we can make the term become zero by setting each of the $n+2$ terms equal to zero. In total this gives us a system of $(n+2)$ ODEs that β and $C_0, \ldots, C_n$ need to fulfill for $\mathcal{G}$ to be a martingale. These ODEs are the ones stated in the lemma. The boundary values are the obvious ones: $\beta(v,v) = u_1(v)$ and $C_i(v,v) = e^{u_0(v)} g_i(v)$, $i = 0, \ldots, n$.

Given the integrability conditions in the lemma (ensuring that the integral wrt. to the Brownian motion is a martingale) and functions $\beta, C_0, \ldots, C_n$ fulfilling the system of ODEs, we have

$$
\begin{aligned}
e^{\int_0^s \rho_0(\tau)+\rho_1(\tau)X(\tau)d\tau} \mathrm{E}\left[\mathcal{G}(v,x;v) \middle| \mathcal{F}(s) \right] &= e^{\int_0^s \rho_0(\tau)+\rho_1(\tau)X(\tau)d\tau} \mathcal{G}(s,x;v) \\
&= e^{\beta(s,v)X(s)} \sum_{i=0}^n C_i(s,v) X^i(s).
\end{aligned}
\tag{4.10}
$$

On the other hand we have that

$$\begin{aligned}
&e^{\int_0^s \rho_0(\tau)+\rho_1(\tau)X(\tau)d\tau}\mathrm{E}\left[\mathcal{G}(v,x;v)\middle|\,\mathcal{F}(s)\right]\\
&=\mathrm{E}\left[e^{-\int_s^v \rho_0(\tau)+\rho_1(\tau)X(\tau)d\tau}e^{\beta(v,v)X(v)}\sum_{i=0}^{n}C_i(v,v)X^i(v)\middle|\,\mathcal{F}(s)\right]\\
&=\mathrm{E}\left[\Upsilon_X(s,v)\sum_{i=0}^{n}g_i(v)X^i(v)\middle|\,\mathcal{F}(s)\right]\\
&=\mathcal{H}(s,X(s);v).
\end{aligned} \tag{4.11}$$

Since the left-hand sides of (4.10) and (4.11) are the same, we have that

$$\mathrm{E}\left[\Upsilon_X(s,v)\sum_{i=0}^{n}g_i(v)X^i(v)\middle|\,\mathcal{F}(s)\right]=e^{\beta(s,v)X(s)}\sum_{i=0}^{n}C_i(s,v)X^i(s),$$

and we have proved the lemma. □

Next, we give a non-dense representation of the transform. In the lemma below, f_1 and f_2 are deterministic functions.

Lemma 4.2 (Non-dense representation). *Let $n\in\mathbb{N}$ be fixed an let $s,v\in[t,T]$ with $s<v$. Assume that the system of ODEs stated in the lemma is uniquely solved by the functions $\alpha,\beta,A_j,B_j,j=1,\dots,n$ and that we have enough integrability to interchange differentiation and integration. That is, there exists a stochastic variable Y with finite expectation such that*

$$\left|\Upsilon_X(s,v)\left(f_0(v)+f_1(v)X(v)\right)^n\right|\le Y \text{ for all } \omega\in\Omega.$$

Then for $n\in\mathbb{N}$ there exist functions α,β, A_j and B_j for $j=1,\dots,n$ given as solutions to a system of ODEs such that,

$$\begin{aligned}
&E\left[\Upsilon_X(s,v)\left(f_0(v)+f_1(v)X(v)\right)^n\middle|\,\mathcal{F}(s)\right]\\
&=e^{\alpha(s,v)+\beta(s,v)X(s)}\sum_{(i_1,\dots,i_n)\in\mathcal{C}_n}\kappa^n(i_1,\dots,i_n)\prod_{j=1}^{n}\left(A_j(s,v)+B_j(s,v)X(s)\right)^{i_j},
\end{aligned} \tag{4.12}$$

where the set $\mathcal{C}_n$ is given by

$$\mathcal{C}_n=\left\{v\in\mathbb{N}_0^n\,\middle|\,\sum_{i=1}^{n}iv_i=n\right\},$$

and $\kappa^n\in\mathbb{N}_0$ is given recursively by

$$\begin{aligned}
\kappa^n(i_1,\dots,i_n)=&\mathbf{1}_{\{i_n=1\}}+\mathbf{1}_{\{i_n\neq1\}}\Bigg(\kappa^{n-1}(i_1-1,i_2,\dots,i_{n-1})\\
&+\sum_{k=1}^{n-2}(i_k+1)\kappa^{n-1}(i_1,\dots,i_k+1,i_{k+1}-1,\dots,i_{n-1})\Bigg),
\end{aligned}$$

where $\kappa^{n-1}(v_1, \ldots, v_{n-1}) = 0$ *for* $\min(v_1, \ldots, v_{n-1}) < 0$.

The functions α, β *and* $A_i, B_i, i = 1, \ldots, n$ *are given by the following systems of ODEs (suppressing the time arguments of the functions), where we use the notation* $B_0 := \beta$*:*

$$\begin{aligned}
\frac{\partial}{\partial s}\beta =& \rho_1 - \left(K_1 + \frac{1}{2}H_1\beta\right)\beta,\\
\frac{\partial}{\partial s}\alpha =& \rho_0 - \left(K_0 + \frac{1}{2}H_0\beta\right)\beta,\\
\frac{\partial}{\partial s}B_m =& -K_1B_m - \frac{1}{2}\sum_{i=0}^{m}\binom{m}{i}B_iH_1B_{m-i}, \quad m = 1, \ldots, n,\\
\frac{\partial}{\partial s}A_m =& -K_0B_m - \frac{1}{2}\sum_{i=0}^{m}\binom{m}{i}B_iH_0B_{m-i}, \quad m = 1, \ldots, n.
\end{aligned}$$

The boundary conditions are given by $\beta(v, v) = u_1(v)$, $\alpha(v, v) = u_0(v)$, $B_1(v, v) = f_1(v)$, $A_1(v, v) = f_0(v)$ *and* $A_i(v, v) = B_i(v, v) = 0$ *for* $i > 1$.

Remark 4.3. *We notice that the class of transforms we are able to calculate with Lemma 4.1 is much broader than the class of transforms we are able to calculate with Lemma 4.2. However, this is not the entire story about the non-dense representation. By following the lines of the proof of Lemma 4.2 one can see, that we can calculate the same type of transforms with the two lemmas. The price to achieve this, we pay in terms of less simplicity for the non-dense representation. This means, that the results are not as simple as the ones given by (4.12), nor are we be able to have so simple representations of the ODEs as the ones stated in Lemma 4.2. This is due to the fact, that in the general case, there are different "versions" of the functions* A_i *and* B_i*, since they have different boundary values. This also implies different ODEs for the different "versions" of* A_i *and* B_i*, since their dynamics are mutually dependent. Later in this section, we make these statements clear, and the results in the general case for* $n = 3$ *are outlined.*

Remark 4.4. *For a fixed* n *the values of the functions* κ^n *in Lemma 4.2 are given by the recursive formula. For* $n = 4$ *we for example get the coefficients in Table 1.*

$\kappa^4(0,0,0,1)$	$\kappa^4(1,0,1,0)$	$\kappa^4(0,2,0,0)$	$\kappa^4(2,1,0,0)$	$\kappa^4(4,0,0,0)$
1	4	3	6	1

Table 1: Example of the coefficients denoted by $\kappa^4()$.

By comparison, we see that Lemma 4.1 and Lemma 4.2 give us two representations for similar quantities. The reason to give two different representations is, that each of them have advantages in different situations. We refer to the representation (4.6) as the *dense representation*, whereas we refer to the representation (4.12) as the *non-dense representation*. The reason for this is that the dense representation contains at maximum $(n + 2)$ functions, which can be found by solving $n + 2$ ODEs, whereas the non-dense representation is overparameterized in the sense that the maximum number of functions evolve in an exponential manner. For some real numbers, see Table 2 on page 21.

The dense representation is interesting because it minimizes the number of ODEs we need to solve. On the other hand, the reason why the non-dense representation is interesting, even though it is overparameterized, is due to that we can use the structure of the problem to obtain a nice representation, where we can reuse the results of the former calculation steps by “gluing” solutions to differential equations together. This is explained in Section 4.2, where the resulting system of ODEs is stated in the case of 3 jumps of the pure jump process. This “gluing” method is not available for the dense representation, since we collect all terms no matter from where they come from to end up solving as few ODEs as possible. Moreover, when using this non-dense representation, results are easily extended to a multidimensional setting, whereas higher dimensions add significantly extra calculational workload to the calculation of the dense representation.

Now we give the proof of the non-dense representation.

Proof of non-dense representation. We are going to show that

$$\begin{aligned}&\mathrm{E}\left[\Upsilon_X(s,v)\left(f_0(v)+f_1(v)X(v)\right)^n\middle|\,\mathcal{F}(s)\right]\\&=e^{\alpha(s,v)+\beta(s,v)X(s)}\sum_{(i_1,\dots,i_n)\in\mathcal{C}_n}\kappa^n(i_1,\dots,i_n)\prod_{j=1}^{n}\left(A_j(s,v)+B_j(s,v)X(s)\right)^{i_j},\end{aligned}\tag{4.13}$$

where the functions incorporated in the right-hand side of the equation are fulfilling a system of ODEs. We prove the lemma by an induction proof.

The induction basis: $(n=1)$
The result follows in the same way as the “extended transform” in Duffie et al. (2000), where one includes time-dependence of the functions of the transform.

The induction step:
We assume that (4.13) holds for n and show that this implies, that it also holds for $n+1$. To make the notation simpler, we define $\mathcal{B}_n$ by

$$\mathcal{B}_n(s,v):=\Upsilon_X(s,v)\left(f_0(v)+f_1(v)X(v)\right)^n.$$

We take the derivative of the left-hand side of (4.13) wrt. v and obtain

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}v}\mathrm{E}\left[\mathcal{B}_n(s,v)\middle|\,\mathcal{F}(s)\right]\\&=\mathrm{E}\Bigg[\bigg(-\rho_0(v)-\rho_1(v)X(v)+\frac{\mathrm{d}}{\mathrm{d}v}u_0(v)+\frac{\mathrm{d}}{\mathrm{d}v}\left(u_1(v)\right)X(v)\\&\qquad+u_1(v)\left(K_0(v)+K_1(v)X(v)\right)\bigg)\mathcal{B}_n(s,v)\bigg|\mathcal{F}(s)\Bigg]\\&\quad+\mathrm{E}\Bigg[\Upsilon_X(s,v)(n-1)\left(f_0(v)+f_1(v)X(v)\right)^{n-1}\\&\qquad\times\left(\frac{\mathrm{d}}{\mathrm{d}v}f_0(v)+\frac{\mathrm{d}}{\mathrm{d}v}f_1(v)X(v)+f_1(v)\left(K_0(v)+K_1(v)X(v)\right)\right)\bigg|\mathcal{F}s\Bigg]\\&=\mathrm{E}\Bigg[\Upsilon_X(s,v)\bigg(\left(f_0(v)+f_1(v)X(v)\right)^{n-1}\left(L_0(v)+L_1(v)X(v)\right)\\&\qquad+\left(f_0(v)+f_1(v)X(v)\right)^{n}\left(L_2(v)+L_3(v)X(v)\right)\bigg)\bigg|\mathcal{F}s\Bigg]\end{aligned}\tag{4.14}$$

for some functions L_i.

On the other hand, we can take the derivative of the right-hand side of (4.13) wrt. v and obtain:

$$\frac{\mathrm{d}}{\mathrm{d}v}\left(e^{\alpha(s,v)+\beta(s,v)X(s)}\sum_{(i_1,\ldots,i_n)\in\mathcal{C}_n}\kappa^n(i_1,\ldots,i_n)\prod_{j=1}^{n}(A_j(s,v)+B_j(s,v)X(s))^{i_j}\right)$$
$$= e^{\alpha(s,v)+\beta(s,v)X(s)}\sum_{(i_1,\ldots,i_n)\in\mathcal{C}_n}\Bigg(\kappa^n(i_1,\ldots,i_n)(A_1(s,v)+B_1(s,v)X(s))^{i_1+1}$$
$$\times\prod_{j=2}^{n}(A_j(s,v)+B_j(s,v)X(s))^{i_j}$$
$$+\kappa^n(i_1,\ldots,i_n)\sum_{k=1}^{n}\prod_{j=1,j\neq k}^{n}\left((A_j(s,v)+B_j(s,v)X(s))^{i_j}\right)$$
$$\times\left(i_k(A_k(s,v)+B_k(s,v)X(s))^{i_k-1}\left(A_{k+1}(s,v)+B_{k+1}(s,v)X(s)\right)\right)\Bigg).$$

We collect the terms and obtain:

$$e^{\alpha(s,v)+\beta(s,v)X(s)}\sum_{(i_1,\ldots,i_n)\in\mathcal{C}_n}\Bigg(\kappa^n(i_1,\ldots,i_n)\prod_{j=1}^{n}(A_j(s,v)+B_j(s,v)X(s))^{i_j+\mathbf{1}_{\{j=1\}}}$$
$$+\sum_{k=1}^{n}i_k\kappa^n(i_1,\ldots,i_n)\prod_{j=1}^{n+1}\left((A_j(s,v)+B_j(s,v)X(s))^{i_{j,k}}\right)$$
$$= e^{\alpha(s,v)+\beta(s,v)X(s)}\sum_{(i_1,\ldots,i_{n+1})\in\mathcal{C}_{n+1}}\kappa^{n+1}(i_1,\ldots,i_{n+1})\prod_{j=1}^{n+1}(A_j(s,v)+B_j(s,v)X(s))^{i_j}. \tag{4.15}$$

Here, $i_{j,k}=i_j\mathbf{1}_{\{j\leq n\}}-\mathbf{1}_{\{j=k\}}+\mathbf{1}_{\{j=k+1\}}$ and

$$\kappa^{n+1}(i_1,\ldots,i_{n+1})=\mathbf{1}_{\{i_{n+1}=1\}}$$
$$+\mathbf{1}_{\{i_{n+1}\neq 1\}}\left(\kappa^n(i_1-1,i_2,\ldots,i_n)+\sum_{k=1}^{n-1}(i_k+1)\kappa^n(i_1,\ldots,i_k+1,i_{k+1}-1,\ldots,i_n)\right),$$

where $\kappa^n(v_1,\ldots,v_n)=0$ for $\min(v_1,\ldots,v_n)<0$. We see, that also the right-hand side has the form stated in the lemma for $n+1$.

We notice that neither the left-hand side nor the right-hand side have the exact form stated in the lemma for $n+1$. Here, we outline why the results also hold, if we change some boundary conditions. We want to replace the boundary conditions $L_2(v)$ and $L_3(v)$ for $f_0(v)$ and $f_1(v)$ and the boundary conditions $L_0(v)$ and $L_1(v)$ for 0 and 0. The argument relies on the well-behaved structure of the ODEs. We have that the ODEs for both B_i and A_i, $i>0$ are linear ODEs with closed form solutions. As we show later in the proof, the ODE for B_j is given by

$$\frac{\partial}{\partial s}B_j(s,v)=-\left(K_1(s)+\beta(s,v)\right)B_j(s,v)-\frac{1}{2}\sum_{i=1}^{j-1}\binom{j}{i}B_i(s,v)H_1(s)B_{j-i}(s,v). \tag{4.16}$$

Given that $B(v,v)=\varsigma$, the solution to (4.16) is given by

$$\begin{aligned}B_j(s,v) =& e^{\int_s^v K_1(\tau)+\beta(\tau,v)H_1(\tau)d\tau} \\ &\times\left(\int_s^v \frac{1}{2}\sum_{i=1}^{j-1}\binom{j}{i} B_i(u,v)H_1(u)B_{j-i}(u,v)e^{\int_s^u -K_1(\tau)-\beta(\tau,v)H_1(\tau)d\tau}du+\varsigma\right).\end{aligned}$$

Moreover, given $A_j(v,v)=\varsigma$, A_j is given by

$$A_j(s,v)=\int_s^v K_0(\tau)B_j(\tau,v)+\frac{1}{2}\sum_{i=0}^{j}\binom{j}{i} B_i(\tau,v)H_0(\tau)B_{j-i}(\tau,v)d\tau+\varsigma.$$

We claim that the solutions of A_j and B_j in conjunction with the form of the partial derivative of the right-hand side of (4.13) given by (4.15) yields, that we can change the boundary conditions and obtain the result given by (4.12).

Regarding the form of the set $\mathcal{C}_n$: Again, we give an induction proof, where the induction basis is obvious since it follows from Duffie et al. (2000), that the only element in $\mathcal{C}_1$ is the singleton $\{1\}$.

Now, the induction step. We consider (4.15) for a fixed element $(i_1,\ldots,i_n)\in\mathcal{C}_n$. We notice, that we in total have $n+1$ additive terms. By the induction hypothesis for n we have that $\sum_{j=1}^n ji_j=n$. This gives us, that the $1+\sum_{j=1}^n ji_j=n+1$ for the first term of the left-hand side of (4.15). For term k of the last n terms the sums are given by $n-k+(k+1)=n+1$. That is, for a vector v in $\mathcal{C}_{n+1}$, we have that $\sum_{i=1}^{n+1} iv_i=n+1$ and we have proved the form of $\mathcal{C}_i$, $i\in I\!N$.

Regarding the system of ODEs. We describe the procedure for the terms β and B_i. The calculations for terms α and A_i follow in exactly the same manner. We know from Duffie et al. (2000), that β fulfills the ODE:

$$\frac{\partial}{\partial s}\beta(s,v)=\rho_1(s)-K_1(s)\beta(s,v)-\frac{1}{2}\beta^2(s,v)H_1(s). \tag{4.17}$$

Taken the derivative of (4.17) wrt. v and setting $\frac{\partial}{\partial v}\beta(s,v)=B_1(s,v)$ we get

$$\frac{\partial}{\partial s}B_1(s,v)=-K_1(s)B_1(s,v)-\beta(s,v)H_1(s)B_1(s,v). \tag{4.18}$$

Taken the derivative of (4.18) wrt. v and setting $\frac{\partial}{\partial v}B_1(s,v)=B_2(s,v)$ we get

$$\frac{\partial}{\partial s}B_2(s,v)=-K_1(s)B_2(s,v)-\beta(s,v)H_1(s)B_2(s,v)-H_1(s)B_1^2(s,v). \tag{4.19}$$

The general representation of the ODEs can be proven by an induction proof:

The induction basis: $(n=1)$

We use (4.18), the notation $B_0:=\beta$ and obtain:

$$\begin{aligned}\frac{\partial}{\partial s}B_1(s,v)=&-K_1(s)B_1(s,v)-2\frac{1}{2}B_0(s,v)H_1(s)B_1(s,v)\\ =&-K_1(s)B_1(s,v)-\frac{1}{2}\sum_{i=0}^{1}\binom{1}{i} B_i(s,v)H_1(s)B_{1-i}(s,v).\end{aligned}$$

That is, the induction basis is correct.

The induction step:
We assume that equation (4.20) holds for a natural number n:

$$\frac{\partial}{\partial s}B_n(s,v) = -K_1(s)B_n(s,v) - \frac{1}{2}\sum_{i=0}^{n}\binom{n}{i}B_i(s,v)H_1(s)B_{n-i}(s,v). \tag{4.20}$$

We want to show, that (4.20) also holds for $n+1$. Differentiating the equation wrt. v yields:

$$\begin{aligned}\frac{\partial}{\partial v}\frac{\partial}{\partial s}B_n(s,v) =& -K_1(s)\frac{\partial}{\partial v}B_n(s,v) - \frac{1}{2}\sum_{i=0}^{n}\binom{n}{i}\bigg(\frac{\partial}{\partial v}\left(B_i(s,v)\right)H_1(s)B_{n-i}(s,v)\\ &+ B_i(s,v)H_1(s)\frac{\partial}{\partial v}\left(B_{n-i}(s,v)\right)\bigg)\\ \Rightarrow \frac{\partial}{\partial s}B_{n+1}(s,v) =& -K_1(s)B_{n+1}(s,v) - \frac{1}{2}\sum_{i=0}^{n}\binom{n}{i}\bigg(B_{i+1}(s,v)H_1(s)B_{n-i}(s,v)\\ &+ B_i(s,v)H_1(s)B_{n+1-i}(s,v)\bigg)\\ =& -K_1(s)B_{n+1}(s,v) - \frac{1}{2}\sum_{i=0}^{n+1}\binom{n+1}{i}B_i(s,v)H_1(s)B_{n+1-i}(s,v).\end{aligned}$$

This concludes the proof. □

We now wrap up the quantity given by (4.5) from inside out using the results of Lemma 4.1 and Lemma 4.2. Note, we cannot directly apply the results of the two lemmas, since there is some time-dependence in (4.5) which the lemma does not allow for. For the non-dense representation it means, that we in general are not able to represent the term (4.4) on this form:

$$e^{\alpha(s,v)+\beta(s,v)X(s)}\sum_{(i_1,\ldots,i_n)\in\mathcal{C}_n}\kappa^n(i_1,\ldots,i_n)\prod_{j=1}^{n}(A_j(s,v)+B_j(s,v)X(s))^{i_j}.$$

As we show in the following, we still have the same overall structure, but instead of a term like $(A_1(t,t_1)+B_n(t,t_1)x)^n$ we have a term $\prod_{k=1}^{n}(A_1(t,t_1;t_k)+B_1(t,t_1;t_k)x)$ reflecting that the jumps of the Markov chain happen at different points in time as opposed to the assumptions in Lemma 4.2. The dynamics of B_1 are the same as stated in Lemma 4.2, but the boundary values are different. For the functions $B_i, i \geq 2$ things get even messier because different "versions" both have different boundary conditions and different dynamics. The structure of the dynamics are the same, but the dynamics are different, since the dynamics of different versions of B_i depend on different versions of $B_j, j < i$ resulting in different versions of B_i. For three jumps, which include the model represented in Figure 1, this system is stated explicitly in Section 4.2. Otherwise, the approach is the same as in Lemma 4.1 and Lemma 4.2. In the terms below, the functions $\alpha_i, \beta_i, C_{ij}, A_{ij}$ and B_{ij} are solutions to systems of ODEs. The exact form of these ODEs are specified in Section 4.2. We let the affine functions μ_i, $i = 1,2,3$ be given by $\mu_i(t,x) = \mu_{i0}(t) + \mu_{i1}(t)x$.

To exemplify this difference due to the time-dependence, we calculate the term (4.5) from inside and out by using Lemma 4.1 and Lemma 4.2. We do so by taking each step (each jump time) at a time. This both covers the dense and the non-dense representation. In the notation we suppress that e.g. C_{30} implicitly depend on time T.

First we consider the **time point $\mathbf{t_3}$**, i.e. the time point when no jumps have occurred. The result follows directly from Duffie et al. (2000, Proposition 1):

$$\begin{aligned}&\mathrm{E}\left[e^{-\int_{t_3}^{T} f_4(\tau,X(\tau))d\tau}\middle|\,\mathcal{F}(t_3)\right]=e^{\alpha_4(t_3,T)+\beta_4(t_3,T)X(t_3)},\\&\alpha_4(T,T)=\beta_4(T,T)=0.\end{aligned}\tag{4.21}$$

Using the result (4.21), Lemma 4.1 and Lemma 4.2, we get that the **time $\mathbf{t_2}$** value, where one jump has occurred, is given by:

Dense representation:

$$\begin{aligned}&\mathrm{E}\left[e^{-\int_{t_2}^{t_3} f_3(\tau,X(\tau))d\tau}e^{\alpha_4(t_3,T)+\beta_4(t_3,T)X(t_3)}\mu_3(t_3,X(t_3))\middle|\,\mathcal{F}(t_2)\right]\\&=e^{\beta_3(t_2,t_3)X(t_2)}\left(C_{30}(t_2,t_3)+C_{31}(t_2,t_3)X(t_2)\right),\\&\beta_3(t_3,t_3)=\beta_4(t_3,T),C_{30}(t_3,t_3)=e^{\alpha_4(t_3,T)}\mu_{30}(t_3),C_{31}(t_3,t_3)=e^{\alpha_4(t_3,T)}\mu_{31}(t_3).\end{aligned}\tag{4.22}$$

Non-dense representation:

$$\begin{aligned}&\mathrm{E}\left[e^{-\int_{t_2}^{t_3} f_3(\tau,X(\tau))d\tau}e^{\alpha_4(t_3,T)+\beta_4(t_3,T)X(t_3)}\mu_3(t_3,X(t_3))\middle|\,\mathcal{F}(t_2)\right]\\&=e^{\alpha_3(t_2,t_3)+\beta_3(t_2,t_3)X(t_2)}\left(A_{31}(t_2,t_3)+B_{31}(t_2,t_3)X(t_2)\right),\\&\alpha_3(t_3,t_3)=\alpha_4(t_3,T),\beta_3(t_3,t_3)=\beta_4(t_3,T),A_{31}(t_3,t_3)=\mu_{30}(t_3),B_{31}(t_3,t_3)=\mu_{31}(t_3).\end{aligned}\tag{4.23}$$

To get the value at time **time $\mathbf{t_1}$**, where two jumps occurred, we use (4.22) and (4.23).

Dense representation:

$$\begin{aligned}&\mathrm{E}\left[e^{-\int_{t_1}^{t_2} f_2(\tau,X(\tau))d\tau}e^{\beta_3(t_2,t_3)X(t_2)}\left(C_{30}(t_2,t_3)+C_{31}(t_2,t_3)X(t_2)\right)\mu_2(t_2,X(t_2))\middle|\,\mathcal{F}(t_1)\right]\\&=e^{\beta_2(t_1,t_2)X(t_1)}\left(\sum_{i=0}^{2}C_{2i}(t_1,t_2)X^i(t_1)\right),\\&\beta_2(t_2,t_2)=\beta_3(t_2,t_3),C_{20}(t_2,t_2)=C_{30}(t_2,t_3)\mu_{20}(t_2),\\&C_{21}(t_2,t_2)=C_{31}(t_2,t_3)\mu_{20}(t_2)+C_{30}(t_2,t_3)\mu_{21}(t_2),C_{22}(t_2,t_2)=C_{31}(t_2,t_3)\mu_{21}(t_2).\end{aligned}\tag{4.24}$$

Non-dense representation:

$$\begin{aligned}&\mathrm{E}\Bigg[e^{-\int_{t_1}^{t_2} f_2(\tau,X(\tau))d\tau}e^{\alpha_3(t_2,t_3)+\beta_3(t_2,t_3)X(t_2)}\\&\qquad\times\left(A_{31}(t_2,t_3)+B_{31}(t_2,t_3)X(t_2)\right)\mu_2(t_2,X(t_2))\Bigg|\,\mathcal{F}(t_1)\Bigg]\\&=e^{\alpha_2(t_1,t_2)+\beta_2(t_1,t_2)X(t_1)}\Big(\left(A_{21}(t_1,t_2)+B_{21}(t_1,t_2)X(t_1)\right)\left(A_{22}(t_1,t_2)+B_{22}(t_1,t_2)X(t_1)\right)\\&\qquad+A_{23}(t_1,t_2)+B_{23}(t_1,t_2)X(t_1)\Big),\\&\alpha_2(t_2,t_2)=\alpha_3(t_2,t_3),\beta_2(t_2,t_2)=\beta_3(t_2,t_3),A_{21}(t_2,t_2)=A_{31}(t_2,t_3),B_{21}(t_2,t_2)=B_{31}(t_2,t_3),\\&A_{22}(t_2,t_2)=\mu_{20}(t_2),B_{22}(t_2,t_2)=\mu_{21}(t_2),A_{23}(t_2,t_2)=B_{23}(t_2,t_2)=0.\end{aligned}\tag{4.25}$$

In the same manner, we can also obtain results at **time t** of the quantities

$$\begin{aligned} &\mathrm{E}\left[e^{-\int_t^{t_1} f_1(\tau,X(\tau))d\tau}e^{\beta_2(t_1,t_2)X(t_1)}\sum_{i=0}^{2}\left(C_{2i}(t_1,t_2)X^i(t_1,t_2)\right)\mu_1(t_1,X(t_1))\middle|\mathcal{F}(t_1)\right] \\ &= e^{\beta(t,t_1)X(s)}\left(\sum_{i=0}^{3}C_i(t,t_1)X^i(t)\right). \end{aligned} \tag{4.26}$$

We can also calculate the non-dense representation

$$\begin{aligned} \mathrm{E}\Bigg[&e^{-\int_t^{t_1} f_1(\tau,X(\tau))d\tau}e^{\alpha_2(t_1,t_2)+\beta_2(t_1,t_2)X(t_1)}\Bigg(\left(A_{21}(t_1,t_2)+B_{21}(t_1,t_2)X(t_1)\right) \\ &\times\left(A_{22}(t_1,t_2)+B_{22}(t_1,t_2)X(t_1)\right)+\left(A_{23}(t_1,t_2)+B_{23}(t_1,t_2)X(t_1)\right)\Bigg)\mu_1(t_1,X(t_1))\Bigg|\mathcal{F}(t)\Bigg]. \end{aligned} \tag{4.27}$$

The result is quite lengthy and since it is stated in details in Section 4.2, we skip it here.

4.2 ODEs for the non-dense transforms

As mentioned in the former subsection we cannot get the ODEs for the terms (4.21)-(4.27) directly from Lemma 4.2. This is why we state the ODEs in this subsection. When one wants to calculate a statewise reserve of an insurance contract or the transition probabilities in a Markov chain model, the approach is to do the calculations backwards starting from the last time point and obtain the first systems of ODEs. Then we use these values as input for the next system of ODEs and so on. After discretizing the entire time interval, we need to calculate the transforms for all different time points and at last integrate.

Because of the dependence of the different jump times, it is not doable for the non-dense representation to give a representation as neat as the one in Lemma 4.2. By neat we mean a representation which holds for fixed time points s and v and does not depend on anything after time v. Therefore, the ODEs for a model with up to 3 jumps are stated here. That is, one is able to calculate all statewise reserves in the model given by Figure 1. In practice, it is not too interesting to do calculations for a very high n due to the computational workload one encounters, see Section 5.1. The workload does not come from solving the equations for a single transform, but rather from the massive number of transforms needed in order to obtain the value for a reserve.

Before stating the precise results of (4.23), (4.25) and (4.27), one should note, that due to the structure of the transforms, it is possible to "glue" together the solutions to some of the ODEs. Among others, it holds for the pairs $\{\alpha_4,\alpha_3\}$, $\{\beta_4,\beta_3\}$, $\{A_{31},A_{21}\}$ and $\{B_{31},B_{21}\}$. That is, we can specify some of the ODEs which have boundary conditions at time s wrt. boundary conditions for time points after time s. This works because for a given function f it is equivalent to specify a function ε either by

$$\frac{\mathrm{d}}{\mathrm{d}t}\varepsilon(t)=f(t),\quad \varepsilon(v)=K, \tag{4.28}$$

or by

$$\frac{\mathrm{d}}{\mathrm{d}t}\varepsilon(t)=f(t),\quad \varepsilon(s)=\tilde{\varepsilon}(s),\quad \frac{\mathrm{d}}{\mathrm{d}t}\tilde{\varepsilon}(t)=f(t),\quad \tilde{\varepsilon}(v)=K, \tag{4.29}$$

where $t < s < v$ and K is a constant. So by "gluing" we mean, that we instead of the representation (4.28) which arises in a natural way from the conditional approach, use the equivalent representation (4.29). The latter representation highlights that the function ε has its natural boundary condition at time v.

If we "glue" the various differential equations, we obtain the following results for the different points in time. To shorten notation, we write $(A_1 + B_1X)(t, t_3, T)$ rather than $A_1(t, t_3, T) + B_1(t, t_3, T)X(t)$. We mark by red text the time point for which the boundary condition is specified, and by blue text we mark the time points in which the functions (excluding the two functions for which we are specifying the dynamics), that the ODEs are depending on, have boundary conditions. For example the arguments of $A_1(t, t_3, T)$ follow in this way: The T is caused by that the dynamics depend on $\beta(t, T)$ with boundary condition in T. Moreover, the dynamics also depend on $B_1(t, t_3, T)$ but as stated previously, we suppress this. The argument t_3 arise due to that $A_1(t, t_3, T)$ has a boundary condition at time t_3.

Time $\mathbf{t_2}$:

$$(4.23) = e^{\alpha(t_2,T)+\beta(t_2,T)X(t_2)} (A_1 + B_1X)(t_2, t_3, T).$$

Time $\mathbf{t_1}$:

$$\begin{aligned}(4.25) =& e^{\alpha(t_1,T)+\beta(t_1,T)X(t_1)}((A_1 + B_1X)(t_1, t_3, T)(A_1 + B_1X)(t_1, t_2, T)\\ &+ (A_2 + B_2X)(t_1, t_2, t_2, t_3, T)).\end{aligned}$$

Time t :

$$\begin{aligned}(4.27) =& e^{\alpha(t,T)+\beta(t,T)X(t)}((A_1 + B_1X)(t, t_3, T)(A_1 + B_1X)(t, t_2, T)\\ &\times (A_1 + B_1X)(t, t_1, T) + (A_2 + B_2X)(t, t_2, t_2, t_3, T)(A_1 + B_1X)(t, t_1, T)\\ &+ (A_2 + B_2X)(t, t_1, t_3, T)(A_1 + B_1X)(t, t_2, T)\\ &+ (A_1 + B_1X)(t, t_3, T)(A_2 + B_2X)(t, t_1, t_2, T)\\ &+ (A_3 + B_3X)(t, t_1, t_1, t_2, t_3, T)).\end{aligned}$$

The above functions are solutions to the following systems of ODEs (we only state the ODEs for the β and the B_i functions, since the ODEs for α and the A_i functions are the same except that K_0 is substituted for K_1, H_0 is substituted for H_1, and μ_0 is substituted for μ_1 in the boundary conditions):

$$\frac{\partial}{\partial t}B_1(t,t_3,T) = -K_1(t)B_1(t,t_3,T) - \beta(t,T)H_1(t)B_1(t,t_3,T), \quad B_1(t_3,t_3,T) = \mu_1(t_3).$$

$$\frac{\partial}{\partial t}B_1(t,t_2,T) = -K_1(t)B_1(t,t_2,T) - \beta(t,T)H_1(t)B_1(t,t_2,T), \quad B_1(t_2,t_2,T) = \mu_1(t_2).$$

$$\frac{\partial}{\partial t}B_1(t,t_1,T) = -K_1(t)B_1(t,t_1,T) - \beta(t,T)H_1(t)B_1(t,t_1,T), \quad B_1(t_1,t_1,T) = \mu_1(t_1).$$

$$\begin{aligned}\frac{\partial}{\partial t}B_2(t,t_2,t_2,t_3,T) = &- K_1(t)B_2(t,t_2,t_2,t_3,T) - \beta(t,T)H_1(t)B_2(t,t_2,t_2,t_3,T)\\ &- H_1(t)B_1(t,t_2,T)B_1(t,t_3,T), \quad B_2(t_2,t_2,t_2,t_3,T) = 0.\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial t}B_2(t,t_1,t_2,T) = &- K_1(t)B_2(t,t_1,t_2,T) - B_1(t,t_1,T)H_1(t)B_1(t,t_2,T)\\ &- \beta(t,T)H_1(t)B_2(t,t_1,t_2,T), \quad B_2(t_1,t_1,t_2,T) = 0.\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial t}B_2(t,t_1,t_3,T) = &- K_1(t)B_2(t,t_1,t_3,T) - B_1(t,t_1,T)H_1(t)B_1(t,t_3,T)\\ &- \beta(t,T)H_1(t)B_2(t,t_1,t_3,T), \quad B_2(t_1,t_1,t_3,T) = 0.\end{aligned}$$

$$\begin{aligned}\frac{\partial}{\partial t}B_3(t,t_1,t_1,t_2,t_3,T) = &- K_1(t)B_3(t,t_1,t_1,t_2,t_3,T) - \beta(t,T)H_1(t)B_3(t,t_1,t_1,t_2,t_3,T)\\ &- B_1(t,t_1,T)H_1(t)B_2(t,t_2,t_2,t_3,T)\\ &- H_1(t)B_2(t,t_1,t_3,T)B_1(t,t_2,T) - H_1(t)B_2(t,t_1,t_2,T)B_1(t,t_3,T),\\ &B_3(t_1,t_1,t_1,t_2,t_3,T) = 0.\end{aligned}$$

To illustrate how one obtains the above ODEs, we show how to obtain the ODE for $B_2(t,t_1,t_3,T)$ by tanking the derivative of the ODE for $B_1(t,t_3,T)$ wrt. t_1.

$$\begin{aligned}\frac{\partial}{\partial t_1}\frac{\partial}{\partial t}B_1(t,t_3,T) =& \frac{\partial}{\partial t_1}\left(-K_1(t)B_1(t,t_3,T) - \beta(t,T)H_1(t)B_1(t,t_3,T)\right)\\ \Rightarrow \frac{\partial}{\partial t}\frac{\partial}{\partial t_1}B_1(t,t_3,T) =& - K_1(t)\frac{\partial}{\partial t_1}B_1(t,t_3,T) - \left(\frac{\partial}{\partial t_1}\beta(t,T)\right)H_1(t)B_1(t,t_3,T)\\ &- \beta(t,T)H_1(t)\frac{\partial}{\partial t_1}B_1(t,t_3,T)\\ \Rightarrow \frac{\partial}{\partial t}B_2(t,t_1,t_3,T) =& - K_1(t)B_2(t,t_1,t_3,T) - B_1(t,t_1,T)H_1(t)B_1(t,t_3,T)\\ &- \beta(t,T)H_1(t)B_2(t,t_1,t_3,T).\end{aligned}$$

4.3 Generalizing to higher dimensions

The results of the previous subsection can easily be generalized to the case where the underlying process is a multidimensional affine process. That is, we assume that X is given by stochastic differential equation (3.3) but without jumps. Unfortunately, it is rather hard to state the dense representation for a general dimension (d) of the affine process and for a general number of jumps (n), even though it is doable in practice for fixed numbers n and d by following the approach in the proof of Lemma 4.1. It is especially hard to state a general result for a dense representation, since one gets all different kinds of mixed terms. For the non-dense approach, the result is exactly the same except that the dimension of the ODEs is d and not 1. Here, we state a multidimensional version of Lemma 4.2 without giving any proof.

Theorem 4.5 (Non-dense representation). *Let $s, v \in [t,T]$ with $s < v$. Assuming the same integrability conditions as in Lemma 4.2 (in a multidimensional setting), we have* *that*

$$E\left[e^{-\int_s^v \rho_0(\tau)+\rho_1^{tr}(\tau)X(\tau)d\tau}e^{u_0(v)+u_1^{tr}(v)X(v)}\left(f_0(v)+f_1^{tr}(v)X(v)\right)^n \middle| \mathcal{F}(s)\right]$$
$$= e^{\alpha(s,v)+\beta^{tr}(s,v)x} \sum_{(i_1,\ldots,i_n)\in\mathcal{C}_n} \kappa^n(i_1,\ldots,i_n)\prod_{j=1}^{n}\left(A_j(s,v)+B_j^{tr}(s,v)x\right)^{i_j},$$

where $\kappa^n(i_1,\ldots,i_n)\in I\!N_0$ *and the set* $\mathcal{C}_n$ *is given by*

$$\mathcal{C}_n = \left\{ v\in I\!N_0^n \middle| \sum_{i=1}^n iv_i = n\right\}.$$

Moreover, $\alpha,\beta,A_i,B_i, i=1,\ldots,n$ *fulfill the ODEs*

$$\frac{\partial}{\partial s}\beta(s,v) = \rho_1(s) - K_1^{tr}(s)\beta(s,v) - \frac{1}{2}\beta^{tr}(s,v)H_1(s)\beta(s,v), \quad \beta(v,v)=u_1(v),$$
$$\frac{\partial}{\partial s}\alpha(s,v) = \rho_0(s) - K_0^{tr}(s)\beta(s,v) - \frac{1}{2}\beta^{tr}(s,v)H_1(s)\beta(s,v), \quad \alpha(v,v)=u_0(v),$$
$$\frac{\partial}{\partial s}B_n(s,v) = -K_1^{tr}(s)B_n(s,v) + \frac{1}{2}\sum_{i=0}^n \binom{n}{i} B_i^{tr}(s,v)H_1(s)B_{n-i}(s,v),$$
$$\frac{\partial}{\partial s}A_n(s,v) = -K_0^{tr}(s)B_n(s,v) + \frac{1}{2}\sum_{i=0}^n \binom{n}{i} B_i^{tr}(s,v)H_0(s)B_{n-i}(s,v).$$

$$B_1(v,v)=f_1(v), \quad A_1(v,v)=f_0(v), \quad B_j(v,v)=0, \quad A_j(v,v)=0, \quad j>1.$$

By similar arguments as in the one-dimensional case, we can "glue" the ODEs and obtain exactly the same results as in Section 4.2. We just need to replace the one-dimensional notation with the multidimensional vector notation.

5 Different subtopics relating to affine processes and reserves in Markov chain models

In this section we present different topics relating to the results in Section 4. These topics are a comparison of the dense and the non-dense representation and a description of a special case, where one can obtain simple results even in the case of cycles in the Markov chain model. Lastly, we comment on that one, under some additional assumptions, can include quadratic terms in the dynamics of the underlying stochastic process and quadratic terms in the transforms of the underlying stochastic processes and still be able to obtain results of the same type as the ones in Section 4.

5.1 Comparison of representations/computational efficiency

Affine processes have gained their popularity in areas as economics, life insurance and credit risk modelling because of their computational tractability. This tractability for some modelling purposes is caused by the opportunity to solve ODEs instead of PDEs. In general, it is much simpler and computationally faster to solve ODEs instead of PDEs. Since computational efficiency is one of the main advantages of the class of affine processes, it is relevant to consider the computational workload. Our main example in this paper is life to calculate life insurance reserves, why we in this section consider the computational workload of calculating these reserves. Moreover, this section includes some comments about the differences between the dense and the non-dense representation. In short, the main difference is that the dense representation leads to a minimum number of ODEs to solve, whereas the non-dense representation is overparameterized. On the other hand, the ODEs for the dense representation are in general more complicated and harder to obtain than the ODEs for the non-dense representation.

In Table 2 we show a comparison between the dense and the non-dense representation in terms of the number of ODEs to solve to calculate the "standard transforms" given in Lemma 4.1 for the one-dimensional case. This comparison covers different numbers of jumps in the case underlying stochastic process is of dimension 1 and 2, respectively. When we state the number of differential equations that needs to be solved, we count two-dimensional ODE as 2 equations.

Dimension: 1			**Dimension: 2**		
Jumps	Dense	Non-dense	Jumps	Dense	Non-dense
0	2	2	0	3	3
1	3	4	1	5	6
2	4	8	2	8	12
3	5	16	3	12	24

Table 2: Comparison of the number of ODEs in case of 1- and two-dimensional affine processes.

Regardless of whether we consider the dense or the non-dense representation of the transform, the main problem with either approach is that one needs to calculate lots of ODEs to obtain a reserve. The computational challenge is not so much the dimension of the underlying stochastic process as it is the number of jumps. It does for example hold for the non-dense representation, that the number of ODEs is linear in the dimension in the sense, that the number of ODEs is given as a multiple of $(d+1)$, where d is the dimension of the affine stochastic process. This is not too bad and the greater problems arise when calculating the value of a reserve. In that case, one needs to integrate the solutions of the ODEs. Moreover, the different ODEs depend on each other through the fact that the solutions to some ODEs are needed as boundary values for other ODEs. To exemplify this, we consider the case with an underlying one-dimensional stochastic process and a calculation of the quantity given by (2.1). If the length of the contract $(T-t)$ is 30 years, and we assume that we discretize each year into 10 intervals, we in total need to calculate ODEs in the magnitude of 1 billion time points.

5.2 Affine processes and models with cycles

Usually when dealing with Markov chains, affine processes are almost exclusively used for models *without* cycles, which are also called hierarchical Markov chain models. The reason for this is that transition probabilities in such models can be expressed as integrals of solutions to ODEs as shown in the previous sections. Nonetheless, various products in the life insurance industry would more naturally be valuated in Markov chain models with cycles. The most prominent of these examples is the case of disability insurance, where it is naturally to assume, that the insurance policy can "jump" between the states "active/paying premiums" and "disabled/receiving disability benefits".

In models with cycles the affine structure is in the general case not be of any help. However, there is a class of interesting examples, where the affine structure is still useful, which we outline here. The affine structure is useful in the case, where on can diagonalize an intensity matrix Q (also known as a transition rate matrix) of a continuous time Markov chain such that

$$Q(t, X(t)) = VD(t, X(t))V^{-1},$$

where V is time-homogeneous and D is a diagonal matrix whose non-zero entries are affine in the underlying affine process.

The most common case where one is able to diagonalize the intensity matrix is in the case of proportional intensities. An example of such a matrix could be

$$Q(t, x) = \begin{pmatrix} -(\kappa + \rho)x & \kappa x & \rho x \\ \varrho x & -(\varrho + \phi)x & \phi x \\ 0 & 0 & 0 \end{pmatrix},$$

which corresponds to Figure 2.

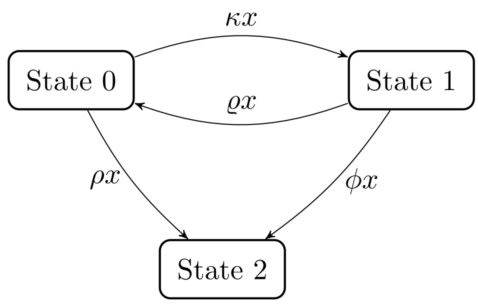

Figure 2: Diagonalizable three state model.

In such a model we can easily obtain results for transition probabilities. We denote by Z the state process of the Markov chain and by τ_T the integral $\int_t^T X(s)ds$. By the tower property and by making a time change in the matrix exponential, we get that for example the probability of going from "State z" to "State j" from time t to time T is given by

$$\begin{aligned} P_{Z(t)=z,X(t)}(Z(T) = j) &= \mathrm{E}\left[\mathbf{1}_{\{Z(T)=j\}} \middle| Z(t) = z, X(t)\right] \\ &= \mathrm{E}\left[\mathrm{E}\left[\mathbf{1}_{\{Z(T)=j\}} \middle| Z(t) = z, X(s), s \leq T,\right] \middle| Z(t) = z, X(t)\right] \\ &= \mathrm{E}\left[\left(e^{Q\tau_T}\right)_{zj} \middle| Z(t) = z, X(t)\right] \\ &= \mathrm{E}\left[\left(Ve^{D\tau_T}V^{-1}\right)_{zj} \middle| Z(t) = z, X(t)\right] \\ &= \sum_{i=0}^{2} \left(v_{zi}\tilde{v}_{ij}\mathrm{E}\left[e^{d_i \int_t^T X(s)ds} \middle| \mathcal{F}(t)\right]\right), \end{aligned} \tag{5.30}$$

where v_{ij} are the entries of V, $\tilde{v}_{ij}$ are the entries of V^{-1}, $D = \mathrm{diag}(d_1X, d_2X, d_3X)$ and $\mathrm{E}\left[e^{d_i \int_t^T X(s)ds} \middle| \mathcal{F}(t)\right]$ are a standard transform considered in this paper. That is, one can obtain transition probabilities by diagonalizing a matrix, calculating simple transforms of affine processes and summing the terms.

5.3 The linear quadratic class

In the literature one sometimes sees examples of transition rates modeled as linear quadratic functions. One example is De Giovanni (2010), where the surrender rate γ is given by

$$\gamma(t) = L_0 r^2(t) + L_1.$$

Here, r is the short rate and L_0 and L_1 are constants. That is, in general it could be interesting to be able to consider other transforms than the previously described affine ones. Moreover, it would also make the model more flexible, if we are able to have other types of dynamics than the affine ones. This section is about including these cases and still be able to obtain results by solving ODEs instead of PDEs.

If we add some more structure, we can extend the class of stochastic processes from the affine to the linear quadratic case. This requires some extra conditions to be fulfilled (for the coordinates, where one has quadratic terms). In one dimension, we can calculate transforms of the form

$$\mathrm{E}\left[e^{-\int_t^T \rho_0(s)+\rho_1(s)X(s)+\frac{1}{2}\rho_2(s)X^2(s)ds} \middle| \mathcal{F}(t) \right] = e^{c(t)+b(t)X(t)+\frac{1}{2}\Gamma(t)X^2(t)}, \tag{5.31}$$

where the dynamics of the underlying stochastic process are allowed to contain the following terms:

$$dX(t) = (K_0(t) + K_1(t)X(t))dt + \sqrt{H_0(t)}dW(t). \tag{5.32}$$

For a two-dimensional stochastic process, where there is a quadratic term in one of the dimensions but not in the other, we can calculate transforms of the form

$$\mathrm{E}\left[e^{-\int_t^T \rho_0(s)+\rho_1(s)X_1(s)+\rho_2(s)X_2(s)+\rho_3(s)X_1^2(s)ds} \middle| \mathcal{F}(t) \right] = e^{a(t)+b(t)X_1(t)+b_2(t)X_2(t)+c(t)X_1^2(t)},$$

where the dynamics are allowed to contain the following terms:

$$d\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \begin{pmatrix} K_{10}(t) + K_{11}(t)X_1(t) \\ K_{20}(t) + K_{21}^{\mathrm{tr}}(t)X(t) + K_{22}(t)X_1^2(t) \end{pmatrix} dt + \sigma(t)dW(t), \tag{5.33}$$

where $\sigma\sigma^{\mathrm{tr}}$ is given by

$$\sigma(t)\sigma^{\mathrm{tr}}(t) = \begin{pmatrix} H_{10}(t) & 0 \\ 0 & H_{20}(t)X_2(t) \end{pmatrix}.$$

This setup still covers many of the well known interest models e.g. the Vasiček and the Hull-White model. It can be shown in general, that the linear quadratic class is equivalent to the affine class, what can be proved by an expansion of the statespace for the affine processes. For a proof and more details about this equivalence, see Cheng and Scaillet (2007).

References

Biffis, E. (2005). Affine Processes for Dynamic Mortality and Actuarial Valuations, *Insurance: Mathematics and Economics*, 443–468.

Buchardt, K. (2012). Life insurance valuation with depent affine rates. Working paper.

Buchardt, K., Møller, T. and Schmidt, K. B. (2013). Cash flows and policyholder behaviour in the semi-markov life insurance setup. Preprint (submitted), available for download on http://www.math.ku.dk/~buchardt/.

Cheng, P. and Scaillet, O. (2007). Linear-quadratic jump-diffusion modeling with application to stochastic volatility, *Mathematical Finance* **17**(4), 575–598.
URL: *http://dx.doi.org/10.1111/j.1467-9965.2007.00316.x*

Dahl, M. and Møller, T. (2006). Valuation and Hedging of Life Insurance Liabilities with Systematic Mortality Risk, *Insurance: Mathematics and Economics* **39**, 193–217.

Dar, A. and Dodds, C. (1989). Interest rates, the emergency fund hypothesis and saving through endowment policies: Some empirical evidence for the u.k., *The Journal of Risk and Insurance* **56**(3), pp. 415–433.
URL: *http://www.jstor.org/stable/253166*

De Giovanni, D. (2010). Lapse rate modeling: a rational expectation approach, *Scandinavian Actuarial Journal* **2010**(1), 56–67.
URL: *http://www.tandfonline.com/doi/abs/10.1080/03461230802550649*

Duffie, D., Filipović, D. and Schachermayer, W. (2003). Affine processes and applications in finance, *Annals of Applied Probability* **13**(3), 984–1053.

Duffie, D., Pan, J. and Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* **68**(6), 1343–1376.
URL: *http://ideas.repec.org/a/ecm/emetrp/v68y2000i6p1343-1376.html*

Filipović, D. (2009). *Term-Structure Models: A Graduate Course*, Springer Finance, Springer.
URL: *http://books.google.at/books?id=KqcSh6CavaAC*

Henriksen, L. F. B., Nielsen, J. W., Steffensen, M. and Svensson, C. (2013). Markov chain modeling of policy holder behavior in life insurance and pension (september 27, 2013). Available at SSRN: http://ssrn.com/abstract=2289926 or http://dx.doi.org/10.2139/ssrn.2289926.

Kuo, W., Tsai, C. and Chen, W.-K. (2003). An empirical study on the lapse rate: The cointegration approach, *Journal of Risk and Insurance* **70**(3), 489–508.
URL: *http://dx.doi.org/10.1111/1539-6975.t01-1-00061*