Continuous Affine Processes: Transformations, Markov Chains and Life Insurance

Kristian Buchardt*
University of Copenhagen and PFA Pension

September 20, 2012

Abstract

Within the setup of a finite state Markov chain, one can associate payments with transitions between states and sojourns in states. This setup is widely used for modelling life insurance and credit risk, and problems have increasingly been studied with affine stochastic interest and transition rates. The class of affine processes is convenient because of its mathematical tractability when valuating the expected present value of payment streams within the Markov chain setup. In this paper we present new results that generalises the classic results in a survival model and other simple Markov chain models. These generalisations extend the range of Markov chain models where one can exploit the structure of affine processes.

Keywords: affine processes; doubly stochastic process; multi-state life insurance models; credit risk; surrender modelling; stochastic mortality; stochastic interest

2010 Mathematics Subject Classification: Primary: 97M30 Secondary: 60J60;91G40;91G30

1 Introduction

In this paper, we show how affine models can be applied for valuation of life insurance liabilities in certain multistate Markov chain models. One of the generalisations is that the interest and transition rates are allowed to be dependent, and we show how a result from [5] can be used for valuation of life insurance liabilities in a simple model of dependent interest and surrender rate, where the payment when surrendering the contract is agreed upon in advance. A generalisation of the result from [5] is then shown, and this

*Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen O and PFA Pension, Sundkrogsgade 4, DK-2100 Copenhagen O, Denmark, http://math.ku.dk/~buchardt, e-mail: buchardt@math.ku.dk

allows for dependent interest and transition rates so that life insurance liabilities can be valuated in multiple decrement (or hierarchical) Markov chain models. An example is given for a disability model where the mortality, disability and interest rate can be dependent.

In traditional life insurance mathematics, a finite state Markov chain is often chosen to represent the state of the insured. Associating payments with sojourns in states and transitions between states, one can find the expected present value of life insurance liabilities when an interest rate is given. Traditionally these models have been studied with a deterministic interest rate and deterministic transition rates, however in modern life insurance mathematics focus has been on more realistic modelling of the interest and transition rates by allowing for stochastic rates. This is of particular interest either when the interest or transition rates are not fully known, and also if one wants to model dependence between the rates. For example, the short rate fluctuates and is unknown beyond today, and for the mortality rate, there is a great deal of uncertainty associated with its long term development. For dependence, one might believe that the surrender rate, for the case where surrender behaviour is modelled as a transition in a Markov chain, is dependent on the economic environment, i.e. the interest rate. basic treatment of stochastic interest rates in life insurance is given in [11], and for stochastic mortality rates in life insurance see e.g. [3]. For combined models for stochastic interest and mortality rates see [4] and [1], where the interest and mortality rates are independent. The case of dependent interest and mortality rates has also been addressed in the litterature, see [12] and references therein for a discussion. Common for the most of the references mentioned is that the interest and mortality rates are modelled as affine processes. This class of processes leads to mathematically tractable models, where one can solve a system of ordinary differential equations instead of partial differential equations in order to find expected present values. For the stochastic transition rates, the focus has mainly been on a stochastic mortality rate, but allowing for other stochastic transition rates leads to equally tractable models. In [7] a stochastic non-affine surrender rate is modelled, and dependence on the economic environment is assumed, and in particular the surrender rate is positively correlated with the interest rate.

Finite state Markov chains are also used when modelling credit risk, see e.g. [8]. A basic credit risk model is a two state Markov chain, where a jump from the initial state represents a default. A popular extension of this model is to let the default transition rate be modelled as a stochastic process itself such that it is possible for it to be dependent on the interest rate and other economic factors. This approach is studied in [10] where various Markov chain models are considered. A more general treatment of the Markov chain approach to credit risk modelling with stochastic transition rates is given in [9], and it is shown how prices generally satisfy a system of partial differential equations. In both papers, it is shown how one can benefit from affine stochastic processes as transition intensities and economic factors. If the model is particularly simple, the Riccati equations and a result from [5] can be used to reduce the problem of solving a system of partial differential equations to that of solving a system of ordinary differential equations. The results presented in this paper generalise these methods. This allows us to find prices in more general decrement Markov chain models solving only ordinary differential equations instead of partial differential equations.

The paper is structured as follows. We begin with a study of continuous affine processes, allowing us to state the main results precisely. We consider the case of multidimensional affine processes and, inspired by Chapter 10 in [6], we present some results on affine processes with time-inhomogeneous parameter functions. From this we present the main results: We give a new proof for a result from [5] on the expectation of a transformation of affine processes and use this to formulate and prove a more general result on affine processes. This result is of interest in life insurance and credit risk modelling and, in general, for Markov chains with stochastic transition rates. The results are discussed and further generalisations are considered. We briefly discuss how to find transition probabilities for certain Markov chains with stochastic transition rates, and in the last section, we apply the results in life insurance giving a couple of examples of applications.

2 Continuous affine processes

In this section we study continuous affine processes. We begin by giving a proper definition and then we consider a canonical state space $\mathbb{R}_+^m \times \mathbb{R}^n$. For this state space, we present precise conditions on the parameter functions that are satisfied if and only if the process is affine, and the Riccati equations for the characteristic and moment generating functions are presented. These initial considerations provide the basis for our treatment and application of affine processes in this paper. In Section 2.2 we present the usual theorem of affine processes within our setup, which gives e.g. bond prices and survival probabilities in the case where the short rate, respectively the mortality rate, is a stochastic affine process. This provides the basis for the main results which are presented in Section 2.3.

Let $(\Omega, \mathcal{F}, (\mathcal{F}(t))_{t\in\mathbb{R}_+}, P)$ be a filtered probability space, satisfying the usual conditions, and denote by $\mathbf{W} = (W(t))_{t\in\mathbb{R}_+}$ an adapted d-dimensional Wiener process. Let $\mathbf{X} = (X(t))_{t\in\mathbb{R}_+}$ be a d-dimensional stochastic process satisfying the stochastic differential equation,

$$\begin{aligned} \mathrm{d}X(t) &= \delta(t, X(t))\,\mathrm{d}t + \rho(t, X(t))\,\mathrm{d}W(t), \\ X(0) &= x \in \mathcal{X}, \end{aligned} \tag{2.1}$$

where $\mathcal{X} \subset \mathbb{R}^d$ is the state space. The functions $\delta : \mathbb{R}_+ \times \mathcal{X} \to \mathbb{R}^d$ and $\rho : \mathbb{R}_+ \times \mathcal{X} \to \mathbb{R}^{d\times d}$ are assumed to be measurable. We have not assumed any Lipschitz continuity of the parameter functions δ and ρ, and we will allow for some discontinuities. In this paper we make the crucial assumption that $\mathbf{X}$ exists for all start values $x \in \mathcal{X}$, and start time points.

We use the notation $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$ and $\mathbb{R}_- = \{x \in \mathbb{R} \mid x \leq 0\}$, and similarly $\mathbb{C}_{+-} = \{z \in \mathbb{C} \mid \Re z \in \mathbb{R}_{+-}\}$, where $\Re z$ is the real part of z.

2.1 Definition and characterisation

The continuous stochastic process $\mathbf{X}$ is affine if the $\mathcal{F}(t)$-conditional characteristic function of $X(T)$ has an exponential affine form for $0 \leq t \leq T$ and all start values $X(0) = x \in \mathcal{X}$, and this is made precise in the following definition. We think of T as a fixed, large time horizon.

Definition 2.1. *The process* $\mathbf{X}$*, with initial value* $X(0) = x$*, is* affine *if there exist functions* ϕ *and* ψ *such that for all* $x \in \mathcal{X}$*,* $0 \leq t \leq T$ *and* $z \in \mathrm{i}\mathbb{R}^d$*,*

$$\mathrm{E}\left[e^{z^\top X(T)} \middle| \mathcal{F}(t) \right] = e^{\phi(t,T,z)+\psi(t,T,z)^\top X(t)} \tag{2.2}$$

holds, where $\phi(t,T,z)$ *is* $\mathbb{C}$*-valued and* $\psi(t,T,z)$ *is* $\mathbb{C}^d$*-valued.*

It can be shown that for $\mathbf{X}$ to be affine, it is a necessary condition that the drift and diffusion parameter functions are affine of the form

$$\begin{aligned} \delta(t,x) &= b(t) + \sum_{i=1}^d \beta_i(t)x_i = b(t) + \mathcal{B}(t)x, \\ \rho(t,x)\rho(t,x)^\top &= a(t) + \sum_{i=1}^d \alpha_i(t)x_i, \end{aligned} \tag{2.3}$$

for some vector functions $b : \mathbb{R}_+ \to \mathbb{R}^d$ and $\beta_i : \mathbb{R}_+ \to \mathbb{R}^d$, $i = 1,\dots,d$, and matrix functions $a : \mathbb{R}_+ \to \mathbb{R}^{d\times d}$ and $\alpha_i : \mathbb{R}_+ \to \mathbb{R}^{d\times d}$, $i = 1,\dots,d$. We have $\mathcal{B}(t) = (\beta_1(t),\dots,\beta_d(t))$, i.e. column i equals $\beta_i(t)$.

From now on we assume that (2.3) holds, and that the parameter functions b, β_i, a, α_i, $i = 1,\dots,i$ are bounded and piecewise continuous.

The affine requirement (2.3) gives some information about the parameters of affine processes, but in fact, a more precise characterisation of the parameter functions can be given. We consider the canonical state space,

$$\mathcal{X} = \mathbb{R}_+^m \times \mathbb{R}^n,$$

for $n + m = d$, $n, m \geq 0$. We let n, m and d be fixed throughout this paper. It is convenient to introduce the index sets

$$I = \{1, \ldots, m\}, \quad J = \{m+1, \ldots, d\}.$$

We also introduce the following notation. For a vector γ, a matrix Γ and index sets M and N,

$$\gamma_N = (\gamma_i)_{i \in N}, \quad \Gamma_{NM} = (\Gamma_{ij})_{i \in N, j \in M}.$$

We present a theorem on the parameter functions that ensures that the process is affine and that it belongs to the state space $\mathbb{R}_+^m \times \mathbb{R}^n$. This is convenient, because the information about the positivity of the first m dimensions is particularly convenient, for example for certain regularity conditions, as will be seen in Theorem 2.3 below. Also, differential equations giving the characteristic and moment generating function (2.2) are given. The theorem is presented because of its close relation with the definition of the affine processes, and because it provides the basis for the study of affine processes in the canonical state space.

Theorem 2.2. *The stochastic process* $\mathbf{X}$ *is affine on the canonical state space* $\mathbb{R}_+^m \times \mathbb{R}^n$ *if and only if the parameter functions are admissible in the sense,*

$$\begin{aligned}
a(t), \alpha_i(t) &\ \text{are symmetric and positive semi-definite, } i \in I, \\
a_{II}(t) &= 0, \\
a_{IJ}(t) = a_{JI}(t)^\top &= 0, \\
\alpha_j(t) &= 0, \quad j \in J, \\
\alpha_{i,kl}(t) = \alpha_{i,lk}(t) &= 0, \quad k \in I \setminus \{i\}, i \in I, l \in \{1, \ldots, d\}, \\
b(t) &\in \mathbb{R}_+^m \times \mathbb{R}^n, \\
\mathcal{B}_{IJ}(t) &= 0, \\
\beta_{i,k}(t) &\geq 0, \quad k \in I \setminus \{i\}, i \in I,
\end{aligned}$$

for all $t \in \mathbb{R}_+$.

In this case, the functions ϕ *and* ψ *in* (2.2) *solve the Riccati equations,*

$$\begin{aligned}
\frac{\partial}{\partial t}\phi(t,T,z) &= -\frac{1}{2}\psi_J(t,T,z)^\top a_{JJ}(t)\psi_J(t,T,z) - b(t)^\top \psi(t,T,z), \\
\phi(T,T,z) &= 0, \\
\frac{\partial}{\partial t}\psi_i(t,T,z) &= -\frac{1}{2}\psi(t,T,z)^\top \alpha_i(t)\psi(t,T,z) - \beta_i(t)^\top \psi(t,T,z), \quad i \in I, \\
\frac{\partial}{\partial t}\psi_J(t,T,z) &= -\mathcal{B}_{JJ}(t)^\top \psi_J(t,T,z), \\
\psi(T,T,z) &= z,
\end{aligned} \tag{2.4}$$

and there exists a unique solution $t \mapsto (\phi(t,T,z), \psi(t,T,z)) : \mathbb{R}_+ \to \mathbb{C}_- \times \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$ *for all* $z \in \mathbb{C}_-^m \times \mathrm{i}\mathbb{R}^n$ *and* $T > 0$.

The theorem is proven in [6] for time-homogeneous continuous affine processes, i.e. time-independent parameter functions. The time-inhomogeneous version is proven in [2] for continuous parameter functions, and this can be generalised to the piecewise continous case presented here.

The theorem gives a characterisation of the admissible parameters with respect to our state space $\mathbb{R}_+^m \times \mathbb{R}^n$. We briefly give an example for the case $d = 3$, $m = 2$, $n = 1$. The drift vectors satisfy

$$b(t) = \begin{bmatrix} + \\ + \\ * \end{bmatrix}, \quad \mathcal{B}(t) = \begin{bmatrix} \mathcal{B}_{II}(t) & \mathcal{B}_{IJ}(t) \\ \mathcal{B}_{JI}(t) & \mathcal{B}_{JJ}(t) \end{bmatrix} = \begin{bmatrix} * & + & 0 \\ + & * & 0 \\ * & * & * \end{bmatrix},$$

and the diffusion matrices satisfy

$$a(t) = \begin{bmatrix} 0 & 0 & 0 \\ & 0 & 0 \\ & & + \end{bmatrix}, \quad \alpha_1(t) = \begin{bmatrix} + & 0 & * \\ & 0 & 0 \\ & & + \end{bmatrix}, \quad \alpha_2(t) = \begin{bmatrix} 0 & 0 & 0 \\ & + & * \\ & & + \end{bmatrix},$$

and $\alpha_3(t) = 0$. Here, $+$ denotes non-negative real functions, and $*$ denotes real functions, such that $a(t)$, $\alpha_i(t)$, $i = 1, 2$ are symmetric and positive semi-definite for all t.

To give a more intuitive understanding of the conditions, we consider the conditions on the drift vectors. The stochastic differential equation for $\mathbf{X}_i$, $i \in I$ is

$$\mathrm{d}X_i(t) = (b_i(t) + \mathcal{B}_{iI}(t)X_I(t) + \mathcal{B}_{iJ}(t)X_J(t))\, \mathrm{d}t + e_i^\top \rho(t, X(t))\, \mathrm{d}W(t).$$

We inspect the stochastic differential equation with the requirement that $X_i(t) \geq 0$ for all initial values $x \in \mathcal{X}$ in mind. Consider then the case $x \in \mathcal{X}$ where $x_i = 0$. Here our requirement implies that the drift must be non-negative, which gives rise to the conditions

- $b_i(t) \geq 0$,
- if $k \in I \setminus \{i\}$ then $\mathcal{B}_{ik}(t) \geq 0$,
- if $j \in J$ then $\mathcal{B}_{ij}(t) = 0$.

These are the conditions on the drift parameters presented in the theorem. For $\mathbf{X}_i$, $i \in J$, there are no requirements on the drift parameters.

2.2 The basic pricing theorem

In order to allow for modelling flexibility, we consider affine transformations of $\mathbf{X}$. Let $p \geq 1$, and let c and Γ be a vector and matrix function respectively, $c : \mathbb{R}_+ \to \mathbb{R}^p$ and $\Gamma : \mathbb{R}_+ \to \mathbb{R}^{p\times d}$. Define the p-dimensional process $\mathbf{Y}$ by

$$Y(t) = c(t) + \Gamma(t)X(t). \tag{2.5}$$

We assume that c_j is integrable and that c_j and Γ_{ji}, $j = 1, \ldots, p$, $i = 1, \ldots, d$ are discontinuous in only a finite set of time points belonging to the interval $(0, T)$. We also assume that c_j and Γ_{ji} are bounded and with limits everywhere. It is noted, that if $\Gamma(t)$ has a left inverse for all t, or, equivalently if it is injective for all t, then the affine transformation $\mathbf{Y}$ is still an affine process, as can be seen by the definition (2.2).

We use the column sums of Γ, so define the d-dimensional function $\gamma(t) = \mathbb{1}^\top\Gamma(t)$, where $\mathbb{1} = (1, \ldots, 1)^\top$ is a column vector with 1 in all entries. Then $\gamma_i(t) = \mathbb{1}^\top\Gamma(t)e_i$ is the sum of column i in Γ. Using this notation, we have $\mathbb{1}^\top Y(t) = \mathbb{1}^\top c(t) + \sum_{i=1}^d \gamma_i(t)X_i(t)$.

The following theorem states the essential feature of the affine processes, and the result is the reason for the great interest in affine processes. For example, if $\mathbf{X}$ is a one-dimensional affine process modelling the short rate, the theorem enables us to price bonds by simply solving a system of Riccati differential equations.

Theorem 2.3. *Let $\mathbf{X}$ be an affine process in the canonical state space $\mathbb{R}_+^m \times \mathbb{R}^n$, $m+n = d$, $m, n \geq 0$. If either of the following hold*

- $n = 0$ *and* $\gamma_i \geq 0$ *for* $i = 1, \ldots, d$,
- (ϕ, ψ) *exists as a solution to* (2.8) *and the stochastic process*
$$t \mapsto e^{-\int_0^t \mathbb{1}^\top Y(s)\,\mathrm{d}s + \phi(t,T) + \psi(t,T)^\top X(t)} \tag{2.6}$$
is a martingale,

then

$$\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} \middle| \mathcal{F}(t) \right] = e^{\phi(t,T) + \psi(t,T)^\top X(t)}, \tag{2.7}$$

where the functions ϕ and ψ solve the Riccati differential equations,

$$\begin{aligned}
\frac{\partial}{\partial t}\phi(t,T) &= -\frac{1}{2}\psi_J(t,T)^\top a_{JJ}(t)\psi_J(t,T) - b(t)^\top\psi(t,T) + \mathbb{1}^\top c(t), \\
\phi(T,T) &= 0, \\
\frac{\partial}{\partial t}\psi_i(t,T) &= -\frac{1}{2}\psi(t,T)^\top\alpha_i(t)\psi(t,T) - \beta_i(t)^\top\psi(t,T) + \gamma_i(t), \quad i \in I, \\
\frac{\partial}{\partial t}\psi_J(t,T) &= -\mathcal{B}_{JJ}(t)^\top\psi_J(t,T) + \gamma_J(t), \\
\psi(T,T) &= 0.
\end{aligned} \tag{2.8}$$

For the case of time-homogeneous parameters, the theorem is proven in [6], and for the case of time-inhomogeneous continuous parameter functions, the theorem is proven in [2]. This can be generalised to the case of piecewise continuous parameter functions presented here.

We briefly comment on the two regularity conditions. If $n > 0$, the sign of $X_j(t)$, $j \in J$ is not controlled, and then it is difficult to find conditions that ensure the existence of a solution to the Riccati equations (2.8), and that (2.6) is a martingale, which is the second condition. Instead, if $n = 0$, it is simpler because $\mathbf{X}$ is positive, and the theorem ensures that the solution exists for suitable Γ. In that case we have $J = \emptyset$, and the differential equation for $\frac{\partial}{\partial t}\psi_J$ in (2.8) no longer appears. Also, the first term of the right hand side of $\frac{\partial}{\partial t}\phi(t,T)$ disappears.

Whereas Theorem 2.2 presented in the previous section gives the Riccati equations for the characteristic and moment generating function, which is not directly applicable for pricing, Theorem 2.3 presented here is a more important result, and can be applied directly. Consider the case where $\mathbf{X}$ is two-dimensional and jointly models the short rate and mortality rate, i.e. $X(t) = (X_1(t), X_2(t)) = (r(t), \mu(t))$. We assume that we are modelling the processes under some risk-neutral or pricing measure. Choosing

$$\Gamma(t) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

and $c(t) = 0$, we obtain by (2.5) that $Y^r(t) = r(t)$, i.e. $\mathbf{Y}^r$ is the short rate. Then Theorem 2.3 can be applied to obtain bond prices,

$$P(t,T) = \mathrm{E}\left[e^{-\int_t^T Y^r(s)\,\mathrm{d}s} \,\middle|\, \mathcal{F}(t) \right] = e^{\Phi^r(t,T) + \Psi^r(t,T)^\top X(t)},$$

where $P(t,T)$ is the market value at time t of a payoff of 1 at time T. Similarly, by choosing another Γ, we can let $\mathbf{Y}^\mu$ be the mortality rate, i.e. $Y^\mu(t) = \mu(t)$, and then survival probabilities are obtained,

$$S(t,T) = \mathrm{E}\left[e^{-\int_t^T Y^\mu(s)\,\mathrm{d}s} \,\middle|\, \mathcal{F}(t) \right] = e^{\Phi^\mu(t,T) + \Psi^\mu(t,T)^\top X(t)},$$

where $S(t,T)$ is the probability of surviving from time t till time T. Finally, choosing Γ as the identity matrix, $\mathbf{Y}$ is two-dimensional and equal to $\mathbf{X}$, hence jointly modelling the short rate and the mortality rate. In this case we obtain the price of a pure endowment,

$$V(t) = \mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} \,\middle|\, \mathcal{F}(t) \right] = e^{\Phi(t,T) + \Psi(t,T)^\top X(t)},$$

where $V(t)$ is the expected present value of a payoff of 1 at time T, conditional on survival to time T given that the individual is alive at time t. Note that here the interest and mortality rate can be dependent. If we consider the special case of independence between the interest and mortality rate, we obtain $V(t) = P(t,T)S(t,T)$ and in particular that

$$\begin{aligned} \Phi(t,T) &= \Phi^r(t,T) + \Phi^\mu(t,T), \\ \Psi(t,T) &= \Psi^r(t,T) + \Psi^\mu(t,T). \end{aligned}$$

For an endowment insurance, where there is a payoff upon death as well, Theorem 2.3 is not sufficient for pricing, and we have to show another result.

2.3 Main results

In this section we present the main contributions of the paper, which is basically a new proof of Theorem 2.4 below. The advantage of the proof is that it is constructive and the idea of the proof can be reapplied which allows us to state and prove Theorem 2.8. To the author's knowledge, this is a new result. Together with Theorem 2.3 the two theorems presented in this section have applications for Markov chain modelling with stochastic transition rates, e.g. for life insurance valuation or credit risk modelling.

Consider the transformation

$$\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} Y_k(u) \middle| \mathcal{F}(t) \right],$$

for $k \in \{1, \dots, p\}$ and $u \in [t,T]$, and recall that $\mathbf{Y}$ is defined by (2.5). In [5] differential equations are derived for the expectation under slightly different conditions for time-homogenenous affine jump diffusions, though only for the case $u = T$. The result presented here is for general $u \in [t,T]$ and the case of continuous affine processes with time-inhomogeneous coefficients. The system of differential equations found in [5] is essentially the same as the one presented in Theorem 2.4.

Theorem 2.4. *Let $k \in \{1, \dots, p\}$ and $u \in [t,T]$. Then, if either Assumption 2.5 or Assumption 2.6 holds, and under the conditions of Theorem 2.3, it holds that*

$$\begin{aligned} &\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} Y_k(u) \middle| \mathcal{F}(t) \right] \\ &= e^{\phi(t,T) + \psi(t,T)^\top X(t)} \Big(A^k(t,T,u) + B^k(t,T,u)^\top X(t) \Big), \end{aligned} \tag{2.9}$$

where (ϕ, ψ) is a solution to (2.8) and (A^k, B^k) solves the linear differential equation *system,*

$$
\begin{aligned}
\frac{\partial}{\partial t} A^k(t,T,u) &= -\psi_J(t,T)^\top a_{JJ}(t) B_J^k(t,T,u) - b(t)^\top B^k(t,T,u), \\
A^k(u,T,u) &= e_k^\top c(u), \\
\frac{\partial}{\partial t} B_i^k(t,T,u) &= -\psi(t,T)^\top \alpha_i(t) B^k(t,T,u) - \beta_i(t)^\top B^k(t,T,u), \quad i \in I \qquad (2.10) \\
\frac{\partial}{\partial t} B_J^k(t,T,u) &= -\mathcal{B}_{JJ}(t)^\top B_J^k(t,T,u) \\
B^k(u,T,u) &= e_k^\top \Gamma(u).
\end{aligned}
$$

The original proof of the result, from [5], is the classic one. The result holds if, when multiplied with $e^{-\int_0^t \mathbb{1}^\top Y(s)\,\mathrm{d}s}$, the right hand side of (2.9) is a martingale. Applying Itô's lemma, and setting the drift equal to zero, one obtains the system of differential equations. Carrying out this proof, note that $A^k(t,T,u)$ and $B^k(t,T,u)$ are constant for $t > u$. For this proof, the integrability condition is the following.

Assumption 2.5. *(Classic integrability condition) The process*

$$
t \mapsto e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s + \phi(t,T) + \psi(t,T)^\top X(t)} \left(A^k(t,T,u) + B^k(t,T,u)^\top X(t) \right)
$$

is a real martingale (i.e. not only a local martingale).

Instead of presenting the original proof, we give a new proof. Below is presented an outline of the proof, where certain details are omitted. The result holds under a slightly different integrability condition than in the original proof. It holds whenever we can interchange differentiation and expectation, and this is allowed if we can dominate an integrand, which is then a sufficient integrability condition. Then, the integrability condition is the following.

Assumption 2.6. *(New integrability condition) Let $t < u < T$ and let J be an open interval containing u. The assumption is that there exists a stochastic bound Z with finite expectation, such that*

$$
\sup_{u' \in J} \left\{ e^{-\int_t^T \left(\mathbb{1} - e_k + \left(1_{(s \le u')} + 1_{(s > u)}\right) e_k\right)^\top \Gamma(s) X(s)\,\mathrm{d}s} e_k^\top \Gamma(u') X(u') \right\} \le Z. \qquad (2.11)
$$

For the case $n = 0$, i.e. where the state space is $\mathbb{R}_+^d$, and $\Gamma(t)$ only has positive entries for all $t \in \mathbb{R}_+$, then the exponential part in (2.11) is bounded by 1. In that case, a sufficient condition is that $\sup_{u' \in J} e_k^\top \Gamma(u') X(u') \le Z$ holds.

Proof of Theorem 2.4. (Outline of proof) The theorem clearly holds for $u = t$. We prove the case $c = 0$ and $u < T$. So, assume that $t < u \le r \le T$. Define now the matrix function $\tilde{\Gamma}(t,u,r)$ by

$$\tilde{\Gamma}_{ji}(t,u,r) = \begin{cases} \Gamma_{ji}(t), & j \neq k, \\ (1_{(-\infty,u]}(t) + 1_{(r,\infty)}(t))\Gamma_{ji}(t), & j = k, \end{cases}$$

for $j = 1,\dots,p$ and $i = 1,\dots,d$. Notice that $\Gamma(t) = \tilde{\Gamma}(t,u,r)$ when $u = r$. We consider k fixed throughout the proof and suppress the functions' dependence on k. First, using the definition of $\tilde{\Gamma}$, and then applying Theorem 2.3,

$$\begin{aligned} &\mathrm{E}\left[e^{-\int_t^T \left(\mathbb{1} - e_k + \left(1_{(s\le u)} + 1_{(s>r)}\right)e_k\right)\Gamma(s)X(s)\,\mathrm{d}s} \,\middle|\, \mathcal{F}(t)\right] \\ &= \mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top \tilde{\Gamma}(s,u,r)X(s)\,\mathrm{d}s} \,\middle|\, \mathcal{F}(t)\right] \\ &= e^{\tilde{\phi}(t,T,u) + \tilde{\psi}(t,T,u)^\top X(t)}. \end{aligned} \tag{2.12}$$

Here, $\tilde{\phi}$ and $\tilde{\psi}$ solve the differential equations (2.8) with $\tilde{\gamma} = \mathbb{1}^\top \tilde{\Gamma}$ in the place of γ, and we have added u as an argument to make clear the dependence of u in the solution. The solution also depends on r, but that is suppressed in the notation.

We apply $-\frac{\partial}{\partial u}$ on both sides of (2.12). On the left hand side we obtain,

$$\begin{aligned} &-\frac{\partial}{\partial u}\mathrm{E}\left[e^{-\int_t^T \left(\mathbb{1} - e_k + \left(1_{(s\le u)} + 1_{(s>r)}\right)e_k\right)^\top \Gamma(s)X(s)\,\mathrm{d}s} \,\middle|\, \mathcal{F}(t)\right] \\ &= -\frac{\partial}{\partial u}\mathrm{E}\left[e^{-\int_t^T \left(\mathbb{1} - e_k + 1_{(s>r)}e_k\right)^\top \Gamma(s)X(s)\,\mathrm{d}s - \int_t^u e_k^\top \Gamma(s)X(s)\,\mathrm{d}s} \,\middle|\, \mathcal{F}(t)\right] \\ &= \mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top \tilde{\Gamma}(s,u,r)X(s)\,\mathrm{d}s} e_k^\top \Gamma(u)X(u) \,\middle|\, \mathcal{F}(t)\right], \end{aligned}$$

where differentiation and expectation can be interchanged by the integrability condition (2.11). When $u = r$, this is the left hand side of (2.9) with $c = 0$. For differentiation, we have implicitly assumed that $e_k^\top \Gamma(t)$ is continuous in a neighbourhood around u, however this is not a necessary assumption for the theorem.

Applying $-\frac{\partial}{\partial u}$ on the right hand side of (2.12), we obtain

$$\begin{aligned} &-\frac{\partial}{\partial u} e^{\tilde{\phi}(t,T,u) + \tilde{\psi}(t,T,u)^\top X(t)} \\ &= e^{\tilde{\phi}(t,T,u) + \tilde{\psi}(t,T,u)^\top X(t)} \left(-\frac{\partial}{\partial u}\tilde{\phi}(t,T,u) + X(t)^\top \left(-\frac{\partial}{\partial u}\tilde{\psi}(t,T,u)\right)\right). \end{aligned}$$

Let $\tilde{A}(t,T,u) = -\frac{\partial}{\partial u}\tilde{\phi}(t,T,u)$ and $\tilde{B}(t,T,u) = -\frac{\partial}{\partial u}\tilde{\psi}(t,T,u)$. We find, using (2.8),

$$\begin{aligned} &\tilde{A}(t,T,u) \\ &= -\frac{\partial}{\partial u}\int_T^t \left(-\frac{1}{2}\tilde{\psi}_J(s,T,u)^\top a_{JJ}(s)\tilde{\psi}_J(s,T,u) - b(s)^\top \tilde{\psi}(s,T,u)\right) \mathrm{d}s \\ &= \int_T^t \left(-\tilde{\psi}_J(s,T,u)^\top a_{JJ}(s)\tilde{B}_J(s,T,u) - b(s)^\top \tilde{B}(s,T,u)\right) \mathrm{d}s. \end{aligned}$$

For $\tilde{B}$, note first that the differential equation for ψ_j in (2.8) is a special case of the one for ψ_i, obtained by setting $\alpha_j = 0$. Then, for $i = 1, \ldots, d$,

$$\begin{aligned}
&\tilde{B}_i(t,T,u) \\
&= -\frac{\partial}{\partial u}\int_T^t \left(-\frac{1}{2}\tilde{\psi}(s,T,u)^\top \alpha_i(s)\tilde{\psi}(s,T,u) - \beta_i(s)^\top \tilde{\psi}(s,T,u) + \tilde{\gamma}_i(s,u,r)\right) \mathrm{d}s \\
&= \int_T^t \left(-\tilde{\psi}(s,T,u)^\top \alpha_i(s)\tilde{B}(s,T,u) - \beta_i(s)^\top \tilde{B}(s,T,u)\right) \mathrm{d}s - \frac{\partial}{\partial u}\int_T^t \sum_{j=1}^p \tilde{\Gamma}_{ji}(s)\,\mathrm{d}s \\
&= \int_T^t \left(-\tilde{\psi}(s,T,u)^\top \alpha_i(s)\tilde{B}(s,T,u) - \beta_i(s)^\top \tilde{B}(s,T,u)\right) \mathrm{d}s - 1_{(t<u)}\frac{\partial}{\partial u}\int_u^t \tilde{\Gamma}_{ki}(s)\,\mathrm{d}s \\
&= \int_T^t \left(-\tilde{\psi}(s,T,u)^\top \alpha_i(s)\tilde{B}(s,T,u) - \beta_i(s)^\top \tilde{B}(s,T,u)\right) \mathrm{d}s + 1_{(t<u)}\Gamma_{ki}(u).
\end{aligned}$$

In particular, $\tilde{A}(T,T,u) = 0$ and $\tilde{B}(T,T,u) = 0$, and since the integrands are linear in $\tilde{B}$ for $t > u$, we have $\tilde{A}(t,T,u) = 0$ and $\tilde{B}(t,T,u) = 0$ for $t > u$.

Now let $r = u$. Then $\Gamma = \tilde{\Gamma}$ and thus $\psi = \tilde{\psi}$ and $\phi = \tilde{\phi}$, where ϕ and ψ are solutions to (2.8). Letting $A^k(t,T,u) = \tilde{A}(t,T,u)$ and $B^k(t,T,u) = \tilde{B}(t,T,u)$, the system of differential equations (2.10) is obtained, when $c = 0$.

For $u = T$, the result holds as well, which can be shown by taking limits as $u \nearrow T$. The left and right hand side of (2.9) are continuous in u when $u < T$, and it can be shown that the continuity also holds in the limit as $u \nearrow T$.

For a general integrable function c, we can rewrite in terms of the functions ϕ^0, ψ, $A^{k,0}$ and B^k corresponding to the case $c = 0$. (The functions ψ and B^k does not depend on c.) We get,

$$\begin{aligned}
&\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top (c(s)+\Gamma(s)X(s))\,\mathrm{d}s} e_k^\top \left(c(u) + \Gamma(u)X(u)\right)\middle| \mathcal{F}(t)\right] \\
&= e^{-\int_t^T \mathbb{1}^\top c(s)\,\mathrm{d}s}\Big(\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top \Gamma(s)X(s)\,\mathrm{d}s}\middle| \mathcal{F}(t)\right] e_k^\top c(u) \\
&\qquad + \mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top \Gamma(s)X(s)\,\mathrm{d}s} e_k^\top \Gamma(u)X(u)\middle| \mathcal{F}(t)\right]\Big) \\
&= e^{-\int_t^T \mathbb{1}^\top c(s)\,\mathrm{d}s + \phi^{c=0}(t,T) + \psi(t,T)^\top X(t)}\left(e_k^\top c(u) + A^{k,c=0}(t,T,u) + B^k(t,T,u)^\top X(t)\right) \\
&= e^{\phi(t,T) + \psi(t,T)^\top X(t)}\left(A^k(t,T,u) + B^k(t,T,u)^\top X(t)\right),
\end{aligned}$$

where the last equality sign holds whenever ϕ and A^k solve the differential equations (2.8) and (2.10), respectively. This completes the proof. □

The interesting result is the identity (2.9), and this can immediately be extended, which is done in the following corollary. Since the expectation operator is linear, we have

$$\begin{aligned}
&\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s}\left(Y_k(u)+Y_l(v)\right)\Big|\,\mathcal{F}(t)\right]\\
&=\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} Y_k(u)\Big|\,\mathcal{F}(t)\right]+\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} Y_l(v)\Big|\,\mathcal{F}(t)\right],
\end{aligned}$$

for $l \in \{1,\dots,p\}$ and $v \in [t,T]$. By Theorem 2.4, the two expectations on the right hand side can be calculated, thus enabling us to find the left hand side. However, using that the system of differential equations for (A,B) is linear, we can actually calculate the left hand side directly. This is stated in a general way in the following corollary to Theorem 2.4. For the corollary, note that for a finite linear combination of elements of the type $Y_k(u)$, there exist $q \geq 1$, vectors $\kappa^1,\dots,\kappa^q \in \mathbb{R}^p$ and time points $u_1,\dots,u_q \in [t,T]$ such that the linear combination can be written as $\sum_{l=1}^q {\kappa^l}^\top Y(u_l)$.

Corollary 2.7. *For $q \geq 1$ let $\kappa^1,\dots,\kappa^q \in \mathbb{R}^p$ be vectors and let $u_1,\dots,u_q \in [t,T]$ be time points. If the conditions of Theorem 2.4 are satisfied for all combinations of time points $u_1,\dots,u_q$ and dimensions $k = 1,\dots,p$, then*

$$\begin{aligned}
&\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s}\sum_{l=1}^q {\kappa^l}^\top Y(u_l)\Bigg|\,\mathcal{F}(t)\right]\\
&=e^{\phi(t,T)+\psi(t,T)^\top X(t)}\Big(A(t,T)+B(t,T)^\top X(t)\Big),
\end{aligned}$$

where (ϕ,ψ) is a solutions to (2.8) and (A,B) solves the linear system of differential equations with jumps,

$$\begin{aligned}
\frac{\partial}{\partial t}A(t,T) &= -\psi_J(t,T)^\top a_{JJ}(t)B_J(t,T) - b(t)^\top B(t,T),\\
A(u_l-,T) &= A(u_l,T) + {\kappa^l}^\top c(u_l), \quad l = 1,\dots,q\\
A(T,T) &= 0,\\
\frac{\partial}{\partial t}B_i(t,u) &= -\psi(t,T)^\top \alpha_i(t)B(t,u) - \beta_i(t)^\top B(t,u), \quad i \in I\\
\frac{\partial}{\partial t}B_J(t,u) &= -\mathcal{B}_{JJ}(t)^\top B_J(t,u)\\
B(u_l-,T) &= B(u_l,T) + {\kappa^l}^\top \Gamma(u_l), \quad l = 1,\dots,q\\
B(T,T) &= 0.
\end{aligned} \tag{2.13}$$

Proof. By linearity of the expectation operator and Theorem 2.4

$$
\begin{aligned}
&\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} \sum_{l=1}^{q} {\kappa^l}^\top Y(u_l) \middle| \mathcal{F}(t) \right] \\
&= \sum_{l=1}^{q} \sum_{k=1}^{p} \kappa_k^l \,\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} Y_k(u) \middle| \mathcal{F}(t) \right] \\
&= e^{\phi(t,T)+\psi(t,T)^\top X(t)} \sum_{l=1}^{q} \sum_{k=1}^{p} \kappa_k^l \left(A^k(t,T,u_l) + B^k(t,T,u_l)^\top X(t) \right),
\end{aligned}
$$

where A^k and B^k solve (2.10), where the superscript k refers to e_k in the boundary conditions of (2.10). Now, let

$$
\begin{aligned}
A(t,T) &= \sum_{l=1}^{q} \sum_{k=1}^{p} \kappa_k^l A^k(t,T,u_l), \\
B(t,T) &= \sum_{l=1}^{q} \sum_{k=1}^{p} \kappa_k^l B^k(t,T,u_l),
\end{aligned}
$$

and recall that $A^k(t,T,u) = 0$ and $B^k(t,T,u) = 0$ for $t > u$ and any $k = 1, \ldots, d$. Then for $t \leq T$,

$$
\begin{aligned}
&A(t,T) \\
&= \sum_{l=1}^{q} 1_{(t \leq u_l)} \sum_{k=1}^{p} \Big\{ \kappa_k^l e_k^\top c(u_l) \\
&\quad + \int_{u_l}^{t} \left(-\psi_J(s,T)^\top a_{JJ}(s) \kappa_k^l B^k(s,T,u_l) - b(s)^\top \kappa_k^l B^k(s,T,u_l) \right) \mathrm{d}s \Big\} \\
&= \sum_{l=1}^{q} 1_{(t \leq u_l)} {\kappa^l}^\top c(u_l) + \int_{T}^{t} \left(-\psi_J(s,T)^\top a_{JJ}(s) B(s,T) - b(s)^\top B(s,T) \right) \mathrm{d}s.
\end{aligned}
$$

The calculation for $B(t,T)$ is analogous, and the result is obtained. □

The functions A and B can be compared to A^k and B^k from Theorem 2.4. As is seen from the proof, they are sums of functions $A^{k_l}(t,T,u_l)$ and $B^{k_l}(t,T,u_l)$ for $l = 1, \ldots, q$, and thus the linear differential equation system for A and B is the same as the one for A^k and B^k in (2.10), except for the boundary conditions. In the system of differential equations for A and B extra jumps occur, which we can consider as gluing boundary conditions. These are exactly the boundary conditions of each of the system of differential equations for the functions A^k and B^k that add up to A and B.

The presentation in the corollary of the transformation yields another insight. We consider an affine transformation $\mathbf{Y}$ of $\mathbf{X}$, but the results also hold for the more general

$$\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} \kappa^\top X(u) \middle| \mathcal{F}(t) \right].$$

In this case, the boundary conditions in (2.13) are changed, such that $c(t) = 0$ and $\Gamma(u) = 1$. To give an intuition for the validity of this, consider for simplicity the case where $c(t) = 0$, and where $\Gamma(u)$ has a left inverse, $\Gamma^{-1}(u)$. Then apply the corollary for $q = 1$ and $\kappa = \tilde{\kappa}\Gamma^{-1}(u)$ for any $\tilde{\kappa}$. We get,

$$\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} \kappa^\top Y(u) \middle| \mathcal{F}(t) \right] = \mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} \tilde{\kappa}^\top X(u) \middle| \mathcal{F}(t) \right].$$

In other words, the affine transformation of $\mathbf{X}$ in the exponentiated integral, does not need to be the same as the affine transformation outside the exponentiation.

We now present a result which, to the author's knowledge, is new. Theorem 2.3 is the usual result about discount factors when affine processes are used in a multidimensional framework, and the essential result about affine processes. This result was used to obtain Theorem 2.4, by differentiating in a specific way. This approach can be applied again, and the following theorem is then obtained.

Theorem 2.8. *Let $k, l \in \{1, \ldots, p\}$ and $u, v \in [t, T]$. Assuming sufficient integrability, then*

$$\begin{aligned}
&\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} Y_k(u) Y_l(v) \middle| \mathcal{F}(t) \right] \\
&= e^{\phi(t,T) + \psi(t,T)^\top X(t)} \\
&\quad \times \Big\{ \Big(A^k(t,T,u) + B^k(t,T,u)^\top X(t) \Big) \Big(A^l(t,T,v) + B^l(t,T,v)^\top X(t) \Big) \\
&\qquad + C^{kl}(t,T,u,v) + D^{kl}(t,T,u,v)^\top X(t) \Big\},
\end{aligned} \tag{2.14}$$

where, for $(i, \eta) \in \{(k, u), (l, v)\}$, the functions $\big(A^i(t,T,\eta), B^i(t,T,\eta)\big)$ solve the system of differential equations, (2.10)*, with boundary conditions*

$$\begin{aligned}
A^i(\eta, T, \eta) &= e_i^\top c(\eta), \\
B^i(\eta, T, \eta) &= e_i^\top \Gamma(\eta).
\end{aligned}$$

The functions $C^{kl}(t,T,u,v)$ and $D^{kl}(t,T,u,v)$ are solutions to the following system of *differential equations*[1],

$$
\begin{aligned}
\frac{\partial}{\partial t} C^{kl}(t,T,u,v) &= -B^k_J(t,T,u)^\top a_{JJ}(t) B^l_J(t,T,v) \\
&\quad - \psi_J(t,T)^\top a_{JJ}(t) D^{kl}_J(t,T,u,v) - b(t)^\top D^{kl}(t,T,u,v), \\
C^{kl}(u \wedge v, T, u, v) &= 0, \\
\frac{\partial}{\partial t} D^{kl}_i(t,T,u,v) &= -B^k(t,T,u)^\top \alpha_i(t) B^l(t,T,v) \\
&\quad - \psi(t,T)^\top \alpha_i(t) D^{kl}(t,T,u,v) - \beta_i(t)^\top D^{kl}(t,T,u,v), \quad i \in I \\
\frac{\partial}{\partial t} D^{kl}_J(t,T,u,v) &= -\beta_J(t)^\top D^{kl}_J(t,T,u,v), \\
D^{kl}(u \wedge v, T, u, v) &= 0.
\end{aligned}
$$

Proof. (Outline of proof) The proof is analogous to the proof of Theorem 2.4. If either $u = t$ or $v = t$, the result follows from Theorem 2.4. We first prove the result for the case $c(t) = 0$ and $u, v < T$, so assume that $u \in (t,T)$ and $v \in (t,T)$. As in the proof of Theorem 2.4, define now the matrix function $\tilde{\Gamma}(t,u,r)$ by

$$
\tilde{\Gamma}_{ji}(t,u,r) = \begin{cases} \Gamma_{ji}(t) & j \neq k \\ \left(1_{(-\infty,u]}(t) + 1_{(r,\infty)}(t)\right) \Gamma_{ji}(t) & j = k \end{cases},
$$

for $j = 1,\ldots,p$ and $i = 1,\ldots,d$. Notice that $\Gamma(t) = \tilde{\Gamma}(t,u,r)$ when $u = r$. An application of Theorem 2.4 for $Y_l(v)$ yields

$$
\begin{aligned}
&\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top \tilde{\Gamma}(s,u,r) X(s)\,\mathrm{d}s} e_l^\top \tilde{\Gamma}(v,u,r) X(v) \,\middle|\, \mathcal{F}(t) \right] \\
&= e^{\tilde{\phi}(t,T) + \tilde{\psi}(t,T)^\top X(t)} \left(\tilde{A}^l(t,T,v) + \tilde{B}^l(t,T,v)^\top X(t) \right).
\end{aligned}
$$

Here, $\tilde{\phi}$ and $\tilde{\psi}$ solve the differential equations (2.8) with $\tilde{\gamma} = \mathbb{1}^\top \tilde{\Gamma}$ in place of γ. The functions $\tilde{A}^l$ and $\tilde{B}^l$ solve the differential equations (2.10) with boundary conditions $\tilde{A}^l(v,T,v) = 0$ and $\tilde{B}^l(v,T,v) = e_l^\top \tilde{\Gamma}(v,u,r)$. Note that the functions $\tilde{\phi}$, $\tilde{\psi}$, $\tilde{A}^l$ and $\tilde{B}^l$ all depend on u. Applying $-\frac{\partial}{\partial u}$ on both sides, we obtain, after a few calculations similar to those in the proof of Theorem 2.4,

$$
\begin{aligned}
&\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top \tilde{\Gamma}(s,u,r) X(s)\,\mathrm{d}s} e_k^\top \Gamma(u) X(u) e_l^\top \tilde{\Gamma}(v,u,r) X(v) \,\middle|\, \mathcal{F}(t) \right] \\
&= e^{\tilde{\phi}(t,T) + \tilde{\psi}(t,T)^\top X(t)} \\
&\quad \times \Bigg\{ \left(-\frac{\partial}{\partial u} \tilde{\phi}(t,T) - X(t)^\top \frac{\partial}{\partial u} \tilde{\psi}(t,T) \right) \left(\tilde{A}^l(t,T,v) + \tilde{B}^l(t,T,v)^\top X(t) \right) \\
&\qquad - \frac{\partial}{\partial u} \tilde{A}^l(t,T,v) - X(t)^\top \frac{\partial}{\partial u} \tilde{B}^l(t,T,v) \Bigg\}.
\end{aligned}
$$

[1] The notation $x \wedge y = \min\{x, y\}$ is used.

From the proof of Theorem 2.4 it follows that $\tilde{A}^k(t,T,u) = -\frac{\partial}{\partial u}\tilde{\phi}(t,T)$ and $\tilde{B}^k(t,T,u) = -\frac{\partial}{\partial u}\tilde{\psi}(t,T)$, where $\tilde{A}^k$ and $\tilde{B}^k$ are solutions to the differential equations (2.10), with boundary conditions $\tilde{A}^k(u,T,u) = 0$ and $\tilde{B}^k(u,T,u) = e_k^\top \Gamma(u,u)$, and $\tilde{\phi}$ and $\tilde{\psi}$ in the place of ϕ and ψ.

Now, let $\tilde{C}(t,T,u,v) = -\frac{\partial}{\partial u}\tilde{A}^l(t,T,v)$ and $\tilde{D}(t,T,u,v) = -\frac{\partial}{\partial u}\tilde{B}^l(t,T,v)$. By straight forward differentiation, we find

$$\begin{aligned}
\tilde{C}(t,T,u,v) &= -\frac{\partial}{\partial u}\int_v^t \Big(-\tilde{\psi}_J(s,T)^\top a_{JJ}(s)\tilde{B}^l_J(s,T,v) - b(s)^\top \tilde{B}^l(s,T,v)\Big)\,\mathrm{d}s \\
&= \int_v^t \Big\{ -\tilde{B}^k_J(s,T,u)^\top a_{JJ}(s)\tilde{B}^l_J(s,T,v) \\
&\qquad - \tilde{\psi}_J(s,T)^\top a_{JJ}(s)\tilde{D}_J(t,T,u,v) - b(s)^\top \tilde{D}(s,T,u,v)\Big\}\,\mathrm{d}s.
\end{aligned}$$

For $\tilde{D}$ we find $\tilde{D}_i$ for $i = 1,\dots,d$. Note that the differential equation for $\tilde{B}^l_j$, $j \in J$ in (2.10) is a special case of the one for $\tilde{B}^l_i$, $i \in I$, obtained by setting $\alpha_j = 0$. We find,

$$\begin{aligned}
\tilde{D}_i(t,T,u,v) &= -\frac{\partial}{\partial u}\int_v^t \Big(-\tilde{\psi}(s,T)^\top \alpha_i(s)\tilde{B}^l(s,T,v) - \beta_i(s)^\top \tilde{B}^l(s,T,v)\Big)\,\mathrm{d}s \\
&= \int_v^t \Big\{ -\tilde{B}^k(s,T,u)^\top \alpha_i(s)\tilde{B}^l(s,T,v) \\
&\qquad - \tilde{\psi}(s,T)^\top \alpha_i(s)\tilde{D}(s,T,u,v) - \beta_i(s)^\top \tilde{D}(s,T,u,v)\Big\}\,\mathrm{d}s.
\end{aligned}$$

We have $\tilde{D}(t,T,u,v) = 0$ for $t > v$. Since $\tilde{B}^k(t,T,u) = 0$ for $t > u$, we also have $\tilde{D}(t,T,u,v) = 0$ for $t > u$. Similarly $\tilde{C}(t,T,u,v) = 0$ for $t > u \wedge v$.

Let $r = u$. Then $\Gamma = \tilde{\Gamma}$ and thus $\phi = \tilde{\phi}$ and $\psi = \tilde{\psi}$, where (ϕ,ψ) solves (2.8). In this case, $A^i(t,T,\eta) = \tilde{A}^i(t,T,\eta)$ and $B^i(t,T,\eta) = \tilde{B}^i(t,T,\eta)$ for $(i,\eta) \in \{(k,u),(l,v)\}$, where (A^i, B^i) solves (2.10). Also, in this case, let $C^{kl} = \tilde{C}$ and $D^{kl} = \tilde{D}$. The result is now obtained for $c = 0$ and $u, v < T$.

If $u = T$ or $v = T$, then take limits for $u \nearrow T$ or $v \nearrow T$ respectively, which yields the result, since both the left hand and right hand side of (2.14) can be shown to be continuous in the arguments u and v on $[t,T]$.

The extension to $c \neq 0$ can be done analogously to the proof of Theorem 2.4. First use the linearity of the expectation operator, second apply Theorems 2.3 and 2.4, and last verify the differential equations. This completes the proof. □

2.4 Generalisations of the transformations

Theorem 2.3 provides the relation (2.7), repeated here,

$$\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} \middle| \mathcal{F}(t) \right] = e^{\phi(t,T)+\psi(t,T)^\top X(t)},$$

where the functions ϕ and ψ can be found by solving a set of Riccati differential equations, (2.8). By a differentiation argument, as carried out in the proof, Theorem 2.4 gives us the relation (2.9), repeated here,

$$\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} Y_k(u) \middle| \mathcal{F}(t) \right] = e^{\phi(t,T)+\psi(t,T)^\top X(t)} \left(A^k(t,T,u) + B^k(t,T,u)^\top X(t) \right).$$

The functions A^k and B^k are given by a set of differential equations, (2.10). Since they arise through a differentiation argument, we essentially think of them as ϕ and ψ differentiated, respectively. By an application of the exact same differentiation technique, but to relation (2.9) instead of (2.7), we then obtained Theorem 2.8, which is the relation

$$\begin{aligned} \mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} Y_k(u) Y_l(v) \middle| \mathcal{F}(t) \right] &= e^{\phi(t,T)+\psi(t,T)^\top X(t)} \\ \times \Big\{ \left(A^k(t,T,u) + B^k(t,T,u)^\top X(t) \right) &\left(A^l(t,T,v) + B^l(t,T,v)^\top X(t) \right) \\ + C^{kl}(t,T,u,v) + D^{kl}(t,T,u,v)^\top X(t) \Big\}, \end{aligned}$$

i.e. (2.14). The functions C^{kl} and D^{kl} are again given by a set of differential equations. Again, they arise because of a differentiation, and we essentially think of them as A^k and B^k (or A^l and B^l) differentiated, respectively. This can also be seen from the proofs given.

There is no particular reason to stop here. We can apply the differentiation technique to (2.14), and obtain an expression for

$$\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} Y_k(u) Y_l(v) Y_r(w) \middle| \mathcal{F}(t) \right],$$

for some $r \in \{1, \dots, p\}$ and $w \in [t,T]$. To find the expression, one will have to apply the differentiation technique to the right hand side of (2.14). The result is obtainable, but the notation and number of differential equations grow with every differentiation, and thus becomes even more cumbersome. In principle, one can reapply the technique, and obtain expressions for any expectation of the form,

$$\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} Y_{k_1}(u_1) \cdots Y_{k_q}(u_q) \middle| \mathcal{F}(t) \right], \tag{2.15}$$

for $k_1, \dots, k_q \in \{1, \dots, p\}$, $u_1, \dots, u_q \in [t,T]$ and $q \geq 0$. We can count the number of differential equations that need to be solved. For the expression (2.7), corresponding to the case $q = 0$, the functions ϕ and ψ must be found, which is a system of differential equations of dimension $d+1$. For the second expression, (2.9) (corresponding to $q = 1$), the functions A^k and B^k must also be found, which is $d+1$ extra dimensions, in total $2(d+1)$. For the third expression, (2.14) (corresponding to $q = 2$), A^l and B^l as well as C^{kl} and D^{kl} must be found, which is $2(d+1)$ extra equations, in total $4(d+1)$. It seems that the dimension of the system of differential equations that needs to be solved is increasing exponentially in dimension with q increasing.

In the relation (2.14), one can choose $k = l$ and $u = v$, to obtain

$$\begin{aligned} &\mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top Y(s)\,\mathrm{d}s} Y_k(u)^2 \middle| \mathcal{F}(t) \right] = e^{\phi(t,T)+\psi(t,T)^\top X(t)} \\ &\quad \times \left\{ \left(A^k(t,T,u) + B^k(t,T,u)^\top X(t) \right)^2 + C^{kk}(t,T,u,u) + D^{kk}(t,T,u,u)^\top X(t) \right\}. \end{aligned}$$

Similarly, for particular choices of $k_1, \ldots, k_q$ and $u_1, \ldots, u_q$ in (2.15), one can obtain all moments and combinations of moments of different $Y_k(u)$. Using linearity of the expectation operator, one can use this to construct any polynomial in Y as well.

A special case of transformations are moments of affine processes. For a process $\mathbf{Y}$ consider the modified process,

$$\hat{Y}(t) = 1_{\{u,v\}}(t) Y(t) = \hat{c}(t) + \hat{\Gamma}(t) X(t),$$

where $\hat{c}(t) = 1_{\{u,v\}}(t) c(t)$ and $\hat{\Gamma}(t) = 1_{\{u,v\}}(t)\Gamma(t)$. Then it holds that

$$\mathrm{E}\left[Y_k(u) \middle| \mathcal{F}(t) \right] = \mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top \hat{Y}(s)\,\mathrm{d}s} \hat{Y}_k(u) \middle| \mathcal{F}(t) \right] = \left(A^k(t,T,u) + B^k(t,T,u)^\top X(t) \right),$$

where $\phi(t,T) = \psi(t,T) = 0$, because $\hat{c}$ and $\hat{\Gamma}$ equal zero almost surely with respect to the Lebesgue measure. The system of differential equations for A^k and B^k simplifies to

$$\begin{aligned} \frac{\partial}{\partial t} A^k(t,T,u) &= -b(t)^\top B^k(t,T,u), \\ A^k(u,T,u) &= e_k^\top c(u), \\ \frac{\partial}{\partial t} B^k(t,T,u) &= -\mathcal{B}(t)^\top B^k(t,T,u), \\ B^k(u,T,u) &= e_k^\top \Gamma(u). \end{aligned}$$

Considering the second order moment, we can obtain an interpretation of the functions C^{kl} and D^{kl} as a covariance. See that

$$\begin{aligned} &\mathrm{E}\left[Y_k(u) Y_l(v) \middle| \mathcal{F}(t) \right] \\ &= \mathrm{E}\left[e^{-\int_t^T \mathbb{1}^\top \hat{Y}(s)\,\mathrm{d}s} \hat{Y}_k(u) \hat{Y}_l(v) \middle| \mathcal{F}(t) \right] \\ &= \left(A^k(t,T,u) + B^k(t,T,u)^\top X(t) \right) \left(A^l(t,T,v) + B^l(t,T,v)^\top X(t) \right) \\ &\quad + C^{kl}(t,T,u,v) + D^{kl}(t,T,u,v)^\top X(t). \end{aligned}$$

Combining with the result above, we conclude that

$$\operatorname{Cov}\left[Y_k(u), Y_l(v) \middle| \mathcal{F}(t)\right] = C^{kl}(t,T,u,v) + D^{kl}(t,T,u,v)^\top X(t).$$

We note here that the functions C^{kl} and D^{kl} of course depend on the functions $\hat{c}(t)$ and $\hat{\Gamma}(t)$ through $\psi(t,T)$, which is very special in this case because it is equal to zero. In general the functions C^{kl} and D^{kl} do not give the covariance of the stochastic variables $Y_k(u)$ and $Y_l(v)$.

3 Decrement markov chains with stochastic transition rates

In this section, we consider so-called decrement Markov chains in finite state spaces with affine and dependent transition rates. The theorems presented above, and their generalisations, allow us to calculate transition probabilities for decrement Markov chains. We use the notion decrement (or multiple decrement) Markov chain for the case where, for each state i, the Markov chain cannot return to state i after leaving it. This is also sometimes refered to in the literature as a hierarchical Markov chain model. Here, we show examples in small state spaces.

Let a finite state space J be given. We associate a set of non-negative transition rates (μ_{ij}), $i,j \in \mathcal{J}$, with some of the transition rates identical to zero, such that it is a multiple decrement (hierarchical) Markov chain. Let the set of non-zero transition rates be modelled as an affine transformation of a d-dimensional affine process $\mathbf{X}$. That is, assuming that there are p non-zero transition rates, let functions $c : \mathbb{R}_+ \to \mathbb{R}^p$ and $\Gamma : \mathbb{R}_+ \to \mathbb{R}^{p\times d}$ be given, and define

$$Y(t) = c(t) + \Gamma(t)X(t).$$

Then each of the stochastic transition rates is modelled as an element in $\mathbf{Y}$, i.e. for each non-zero transition rate μ_{ij}, there is a dimension in $\mathbf{Y}$, k say, such that $\mu_{ij}(t) = Y_k(t)$.

Define now the stochastic process $\mathbf{Z} = (Z(t))_{t\in\mathbb{R}_+}$ with $Z(0) = 0$, and let $\mathbf{Z}$ be a Markov chain in $\mathcal{J}$ with distribution given by the transition rates (μ_{ij}), conditional on $\mathbf{X}$. That is, we have defined $\mathbf{Z}$ through the conditional distribution, given the stochastic transition rates. With $(N_{ij}(t))_{t\in\mathbb{R}_+}$, $i,j \in \mathcal{J}$, being the process that counts the number of jumps for $\mathbf{Z}$ from state i to j, the compensated process

$$N_{ij}(t) - \int_0^t 1_{(Z(s-)=i)}\mu_{ij}(s)\,\mathrm{d}s \tag{3.1}$$

is a martingale, conditional on $\mathbf{X}$.

Let the filtrations $\mathbb{F}^{\mathbf{Z}} = (\mathcal{F}^{\mathbf{Z}}(t))_{t\in\mathbb{R}_+}$ and $\mathbb{F}^{\mathbf{X}} = (\mathcal{F}^{\mathbf{X}}(t))_{t\in\mathbb{R}_+}$ be the ones generated by the processes $\mathbf{Z}$ and $\mathbf{X}$, respectively, satisfying the usual hypotheses. Let the general filtration be given by $\mathbb{F} = (\mathcal{F}(t))_{t\in\mathbb{R}_+} = \left(\mathcal{F}^{\mathbf{Z}}(t) \vee \mathcal{F}^{\mathbf{X}}(t)\right)_{t\in\mathbb{R}_+}$.

The transition probability from state i to j can then be written as

$$
\begin{aligned}
&P\left(Z(t) = j \middle| Z(s) = i\right) \\
&= p_{ij}(s,t) \\
&= \mathrm{E}\left[1_{(Z(t)=j)} \middle| \mathcal{F}^{\mathbf{X}}(s), Z(s) = i\right] \\
&= \mathrm{E}\left[\mathrm{E}\left[1_{(Z(t)=j)} \middle| \mathcal{F}^{\mathbf{X}}(\infty), Z(s) = i\right] \middle| \mathcal{F}^{\mathbf{X}}(s)\right] \\
&= \mathrm{E}\left[p_{ij}^{\mathbf{X}}(s,t) \middle| \mathcal{F}^{\mathbf{X}}(s)\right] \\
&= \mathrm{E}\left[P\left(Z(t) = j \middle| Z(s) = i, \mathcal{F}^{\mathbf{X}}(\infty)\right) \middle| \mathcal{F}^{\mathbf{X}}(s)\right],
\end{aligned}
$$

where $p_{ij}^{\mathbf{X}}(s,t)$ is the transition probability in the conditional distribution of $\mathbf{Z}$ given $\mathbf{X}$. From this calculation it is seen, that if we know the conditional transition probabilities, corresponding to the case of known transition rates, we can find the unconditional ones by applying the expectation operator.

With the Theorems 2.3, 2.4 and 2.8, and the generalisation (2.15), we can, in principle, find transition probabilities in the unconditional decrement Markov chain. In that case, the conditional transition probabilities $p_{ij}^{\mathbf{X}}(s,t)$ are known explicitly[2] and can be written in sums and integrals of expressions of the type (2.15). Here, we only show how it can be done for a simple state space, where we can apply Theorems 2.3, 2.4 and 2.8.

Example 3.1. Let $\mathcal{J} = \{0,1,2\}$, and assume that the non-zero transition rates are μ_{01}, μ_{02} and μ_{12}. This could be a life insurance disability model with state 0, 1 and 2 corresponding to *active*, *disabled* and *dead*. The transition rates are modelled by $Y(t) = c(t) + \Gamma(t)X(t)$ such that $(\mu_{01}(t), \mu_{02}(t), \mu_{12}(t))^\top = Y(t)$, for some affine process $\mathbf{X}$ and functions c and Γ.

Conditional on $\mathbf{X}$, the transition probabilities are known explicitly, and for state 0 we have

$$
\begin{aligned}
p_{00}^{\mathbf{X}}(t,s) &= e^{-\int_t^s (\mu_{01}(\tau)+\mu_{02}(\tau))\,\mathrm{d}\tau}, \\
p_{01}^{\mathbf{X}}(t,s) &= \int_t^s e^{-\int_t^u (\mu_{01}(\tau)+\mu_{02}(\tau))\,\mathrm{d}\tau} \mu_{01}(u) e^{-\int_u^s \mu_{12}(\tau)\,\mathrm{d}\tau}\,\mathrm{d}u, \\
p_{02}^{\mathbf{X}}(t,s) &= \int_t^s e^{-\int_t^u (\mu_{01}(\tau)+\mu_{02}(\tau))\,\mathrm{d}\tau} \left(\mu_{02}(u) + \mu_{01}(u) \int_u^s e^{-\int_u^v \mu_{12}(\tau)\,\mathrm{d}\tau} \mu_{12}(v)\,\mathrm{d}v \right) \mathrm{d}u \\
&= \int_t^s p_{00}^{\mathbf{X}}(t,u)\mu_{02}(u)\,\mathrm{d}u + \int_t^s p_{01}^{\mathbf{X}}(t,u)\mu_{12}(u)\,\mathrm{d}u,
\end{aligned}
$$

[2]For a Markov chain, the transition probabilities can be found using e.g. Kolmogorov's backward differential equations.

and they can be verified by e.g. Kolmogorov's differential equations. Define

$$I^1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad I^{2,u}(t) = \begin{bmatrix} 1_{(t\leq u)} & 0 & 0 \\ 0 & 1_{(t\leq u)} & 0 \\ 0 & 0 & 1_{(t>u)} \end{bmatrix},$$

and let the processes $\mathbf{Y}^1$ and $\mathbf{Y}^{2,u}$ be given by

$$\begin{aligned} Y^1(t) &= I^1 Y(t) = c^1(t) + \Gamma^1(t)X(t), \\ Y^{2,u}(t) &= I^{2,u}(t)Y(t) = c^{2,u}(t) + \Gamma^{2,u}(t)X(t), \end{aligned}$$

where $c^1 = I^1 c$, $\Gamma^1 = I^1\Gamma$, $c^{2,u} = I^{2,u}c$ and $\Gamma^{2,u} = I^{2,u}\Gamma$. With these definitions, we have

$$\begin{aligned} \mathbb{1}^\top Y^1(t) &= \mu_{01}(t) + \mu_{02}(t), \\ \mathbb{1}^\top Y^{2,u}(t) &= 1_{(t\leq u)}\left(\mu_{01}(t) + \mu_{02}(t)\right) + 1_{(t>u)}\mu_{12}(t) \end{aligned}$$

and the above conditional probabilities can be rewritten,

$$\begin{aligned} p_{00}^{\mathbf{X}}(t,s) &= e^{-\int_t^s \mathbb{1}^\top Y^1(\tau)\,\mathrm{d}\tau}, \\ p_{01}^{\mathbf{X}}(t,s) &= \int_t^s e^{-\int_t^s \mathbb{1}^\top Y^{2,u}(\tau)\,\mathrm{d}\tau} Y_1^{2,u}(u)\,\mathrm{d}u, \\ p_{02}^{\mathbf{X}}(t,s) &= \int_t^s e^{-\int_t^u \mathbb{1}^\top Y^1(\tau)\,\mathrm{d}\tau} Y_2^1(u)\,\mathrm{d}u + \int_t^s \int_t^v e^{-\int_t^v \mathbb{1}^\top Y^{2,u}(\tau)\,\mathrm{d}\tau} Y_1^{2,u}(u) Y_3^{2,u}(v)\,\mathrm{d}u\,\mathrm{d}v \end{aligned}$$

The real transition probabilities can then be found as the conditional expectations,

$$p_{ij}(t,s) = \mathrm{E}\left[p_{ij}^{\mathbf{X}}(t,s) \middle| \mathcal{F}(t)\right],$$

and using linearity and interchanging expectation and integration, we see that we need to find the following quantities,

$$\begin{aligned} &\mathrm{E}\left[e^{-\int_t^s \mathbb{1}^\top Y^1(\tau)\,\mathrm{d}\tau} \middle| \mathcal{F}(t)\right], \\ &\mathrm{E}\left[e^{-\int_t^s \mathbb{1}^\top Y^{2,u}(\tau)\,\mathrm{d}\tau} Y_1^{2,u}(u) \middle| \mathcal{F}(t)\right], \quad u \in [t,s], \\ &\mathrm{E}\left[e^{-\int_t^u \mathbb{1}^\top Y^1(\tau)\,\mathrm{d}\tau} Y_2^1(u) \middle| \mathcal{F}(t)\right], \quad u \in [t,s], \\ &\mathrm{E}\left[e^{-\int_t^v \mathbb{1}^\top Y^{2,u}(t)\,\mathrm{d}\tau} Y_1^{2,u}(u) Y_3^{2,u}(v) \middle| \mathcal{F}(t)\right], \quad u,v \in [t,s], u \leq v. \end{aligned}$$

For fixed parameters u and v, this can be done using Theorem 2.3 for the first line, Theorem 2.4 for the second and third line, and Theorem 2.8 for the fourth line. To find the solution for e.g. u in an interval, as is required here, one must in practice discretise the interval and solve the differential equations for each u in the discretisation. ○

4 Applications in life insurance

We adopt the setup from Section 3 above, and consider life insurance contracts defined within the setup of such a Markov chain. Let a conditional decrement Markov chain $\mathbf{Z}$ be given. We recall that conditional on the intensitites (μ_{ij}), the process $\mathbf{Z}$ is a Markov chain, in a state space $\mathcal{J}$. The intensitites (μ_{ij}) are modelled by a process $Y(t) = c(t) + \Gamma(t)X(t)$, where $\mathbf{X}$ is a continuous affine process. Also, we let $(N_{ij}(t))_{t\in\mathbb{R}_+}$, $i, j \in \mathcal{J}$, be the process that counts the number of jumps for $\mathbf{Z}$ from state i to j.

The process $\mathbf{Z}$ indicates the state of the insured and the states in $\mathcal{J}$ can be e.g. *alive*, *dead* etc.

The interest rate process $(r(t))_{t\in\mathbb{R}_+}$ is also allowed to be stochastic. This will be modelled in the same way, so we let r be an element in $\mathbf{Y}$. By design of Γ and $\mathbf{X}$, the interest and transition rates can be dependent, independent or deterministic.

We consider a life insurance policy, with payments specified by the process $\mathbf{B} = (B(t))_{t\in\mathbb{R}_+}$, such that $B(t)$ is the total payment until time t. Then we can think of $\mathrm{d}B(t)$ as the payment at time t, and we can specify $\mathbf{B}$ as

$$\mathrm{d}B(t) = \sum_{i\in\mathcal{J}} 1_{(Z(t)=i)} b_i(t)\,\mathrm{d}t + \sum_{\substack{i,j\in\mathcal{J}\\ i\neq j}} b_{ij}(t)\,\mathrm{d}N_{ij}(t),$$

for deterministic payment functions b_i and b_{ij}, $i, j \in \mathcal{J}$. Then $b_i(t)$ is the payment while in state i at time t, and $b_{ij}(t)$ is the payment if jumping from state i to j at time t.

The present value at time t of the life insurance policy is then given by

$$PV(t) = \int_t^\infty e^{-\int_t^s r(\tau)\,\mathrm{d}\tau}\,\mathrm{d}B(s).$$

For pricing and reserving, one considers the expected present value

$$V(t) = \mathrm{E}\left[\int_t^\infty e^{-\int_t^s r(\tau)\,\mathrm{d}\tau}\,\mathrm{d}B(s)\middle|\,\mathcal{F}(t)\right],$$

where the expectation is taken using a pricing or risk neutral measure. In valuation of $V(t)$, we use the tower property, that is, we condition on $\mathcal{F}^{\mathbf{X}}(\infty)$ to get

$$V^{\mathbf{X}}(t) = \mathrm{E}\left[\int_t^\infty e^{-\int_t^s r(\tau)\,\mathrm{d}\tau}\,\mathrm{d}B(s)\middle|\,\mathcal{F}^{\mathbf{Z}}(t)\vee\mathcal{F}^{\mathbf{X}}(\infty)\right],$$

so that $V(t) = \mathrm{E}\left[V^{\mathbf{X}}(t)\middle|\,\mathcal{F}(t)\right]$. Here, $V^{\mathbf{X}}(t)$ is the reserve conditional on the interest and transition rates, thus corresponding to the case of deterministic rates. When valuating $V^{\mathbf{X}}$ we need the conditional distribution of $\mathbf{Z}$, and thus $\mathbf{B}$, given the transition rates. By construction this is known, and the usual results hold.

Example 4.1. Consider a surrender model with 3 states $\mathcal{J} = \{0, 1, 2\}$, corresponding to *alive*, *dead* and *surrendered* respectively. The Markov model is shown in Figure 1. This example illustrates a simple way of modelling policyholder behaviour dependent on the economic environment. Here, the economic environment is simply the short interest rate process.

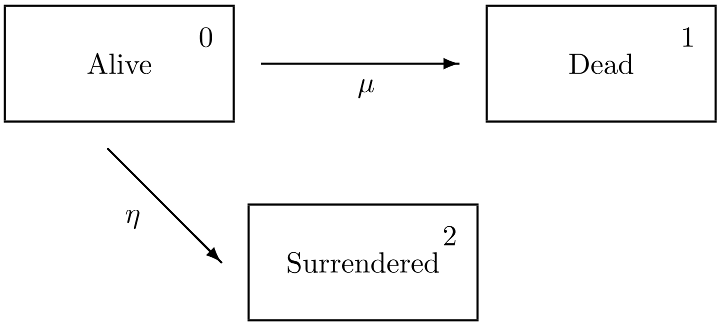

Figure 1: *State space for the survival-surrender model.*

Let the transition rate from state *alive* to state *dead*, i.e. the mortality rate, be deterministic. Further, let the interest rate r and the surrender rate η be modelled as dependent affine processes of the form

$$(r(t), \eta(t))^\top = Y(t) = c(t) + \begin{bmatrix} \gamma_1(t) & 0 \\ 0 & \gamma_2(t) \end{bmatrix} X(t),$$

for a 2-dimensional affine process $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$. This is analogous to (2.5). By the design of $\mathbf{X}$, the dimensions $\mathbf{X}_1$ and $\mathbf{X}_2$ can be dependent processes, such that the interest rate r and the surrender rate η are dependent processes. For notational ease, we have chosen $\Gamma(t)$ to be a diagonal matrix. As noted in Section 2, we use the notation $\gamma(t) = (\gamma_1(t), \gamma_2(t))^\top$ for the column sums, which is equal to the diagonal elements in this case.

Let the payments be defined by

$$\mathrm{d}B(t) = b(t)1_{(Z(t)=0)}\,\mathrm{d}t + b^d(t)\,\mathrm{d}N_{01}(t) + G(t)\,\mathrm{d}N_{02}(t),$$

where $b(t)$ is the continuous payment at time t while alive, $b^d(t)$ is the single payment if death occurs at time t, and $G(t)$ is the payment upon surrender at time t. We consider the three payment functions to be deterministic. We assume that B is constant a.s. after some finite time T, which is a natural assumption for life insurance contracts.

Conditioning on the intensitites, the expected present value $V^{\mathbf{X}}(t)$ is the classic result,

$$\begin{aligned}V^{\mathbf{X}}(t) &= \mathrm{E}\left[PV(t) \middle| \mathcal{F}^{\mathbf{X}}(t), Z(t)=0\right] \\ &= \int_t^T e^{-\int_t^s (r(\tau)+\mu(\tau)+\eta(\tau))\,\mathrm{d}\tau} \left(b(s)+\mu(s)b_d(s)+\eta(s)G(s)\right)\,\mathrm{d}s.\end{aligned}$$

Removing the condition, we find, using Theorems 2.3 and 2.4,

$$\begin{aligned}V(t) &= \mathrm{E}\left[V^{\mathbf{X}}(t) \middle| \mathcal{F}(t)\right] \\ &= \int_t^T e^{-\int_t^s \mu(\tau)\,\mathrm{d}\tau} \Big\{ \mathrm{E}\left[e^{-\int_t^s (r(\tau)+\eta(\tau))\,\mathrm{d}\tau} \middle| \mathcal{F}(t)\right] \left(b(s)+\mu(s)b_d(s)\right) \\ &\qquad + \mathrm{E}\left[e^{-\int_t^s (r(\tau)+\eta(\tau))\,\mathrm{d}\tau}\eta(s) \middle| \mathcal{F}(t)\right] G(s) \Big\}\,\mathrm{d}s \\ &= \int_t^T e^{-\int_t^s \mu(\tau)\,\mathrm{d}\tau + \phi(t,s)+\psi(t,s)^\top X(t)} \Big(b(s)+\mu(s)b_d(s) \\ &\qquad + \left(A^\eta(t,T,s)+B^\eta(t,T,s)^\top X(t)\right) G(s)\Big)\,\mathrm{d}s.\end{aligned}$$

Depending on the choice of $\mathbf{X}$, the system of differential equations (2.8) and (2.10) may simplify a bit. Recall that the state space is $\mathcal{X} = \mathbb{R}_+^m \times \mathbb{R}^n$, with the indicator sets $I = \{1,\dots,m\}$ and $J = \{m+1,\dots,m+n\}$. In this example, we have $m+n=2$. If either $I=\emptyset$ or $J=\emptyset$ the system simplifies, and we put up the system of differential equations for the functions ϕ, ψ, A^η and B^η, for the case $J=\emptyset$,

$$\begin{aligned}\frac{\partial}{\partial t}\phi(t,s) &= -b(t)^\top \psi(t,s) + \mathbb{1}^\top c(s), \\ \frac{\partial}{\partial t}\psi_i(t,s) &= -\frac{1}{2}\psi(t,s)^\top \alpha_i(t)\psi(t,s) - \beta_i(t)^\top\psi(t,s) + \gamma_i(t), \\ \frac{\partial}{\partial t}A^\eta(t,T,s) &= -b(t)^\top B^\eta(t,T,s), \\ \frac{\partial}{\partial t}B_i^\eta(t,T,s) &= -\psi(t,T)^\top \alpha_i(t) B^\eta(t,T,s) - \beta_i(t)^\top B^\eta(t,T,s),\end{aligned}$$

with boundary conditions $\phi(s,s) = \psi_i(s,s) = 0$, $A^\eta(s,T,s) = c_2(s)$ and $B^\eta(s,T,s) = (0,\gamma_2(s))^\top$, for $i=1,2$.

We see that in order to evaluate $V(t)$ one needs to integrate over s, so the functions ϕ,ψ (and A^η and B^η) need to be evaluated in the points (t,s) (and (t,T,s)) for t fixed and $s\in(t,T)$. However, solving the differential equations once yields a solution for fixed s, and $t\le s$, which as desired. Therefore, in practice one needs to solve the differential equation system completely for each desired value of $s\in(t,T)$. ○

Example 4.2. Consider a disability model with 3 states $\mathcal{J}=\{0,1,2\}$, corresponding to states *active*, *disabled* and *dead*, respectively. The Markov model is shown in Figure 2.

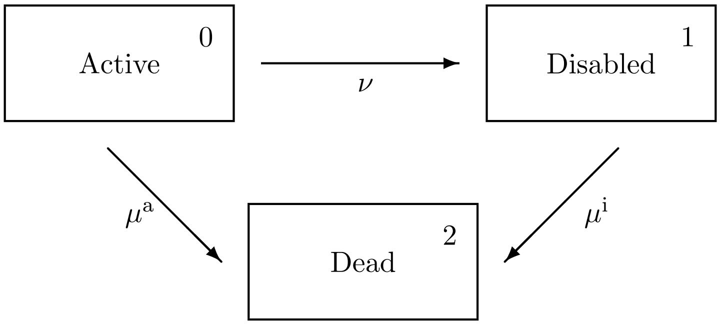

Figure 2: *State space for the disability model.*

The mortality rate while active is denoted by μ^{a} and the mortality rate while disabled is denoted by μ^{i}. The disability rate is denoted by ν. The interest rate is denoted r.

We model the interest rate and transition rates as (dependent) affine processes. Thus, let $\mathbf{X}$ be a d-dimensional affine process, let $c(t)$ be a 4-dimensional deterministic integrable function, and let $\Gamma(t)$ be a $4 \times d$ dimensional matrix function. Then $c(t) + \Gamma(t)X(t)$ is a 4-dimensional process, and the interest and transition rates are then given as

$$(r(t), \mu^{\mathrm{a}}(t), \mu^{\mathrm{i}}(t), \nu(t))^\top = c(t) + \Gamma(t)X(t).$$

We have thus specified a model where the interest and transition rates can have any dependent or independent affine structure. For applications, one could argue that the interest rate might not be dependent on the transition rates, however, because it does not increase the complexity of the mathematics in any significant way, the general case is considered here.

Consider a policy in this model, with total payments given by the process $\mathbf{B}$, satisfying

$$\mathrm{d}B(t) = b(t)1_{(Z(t)=0)}\,\mathrm{d}t + b^{\mathrm{i}}(t)1_{(Z(t)=1)}\,\mathrm{d}t + b^{\mathrm{ad}}(t)\,\mathrm{d}N_{01}(t) + b^{\mathrm{id}}(t)\,\mathrm{d}N_{12}(t).$$

Here, $b(t)$ are continuous payments while active which, when negative, corresponds to premium payments. The function $b^{\mathrm{i}}(t)$ represents continuous payments while disabled, $b^{\mathrm{ad}}(t)$ is a payment upon death at time t from the state active, and $b^{\mathrm{id}}(t)$ is a payment upon death at time t from the state disabled. The payment functions are deterministic, and we assume that there are no payments after time T.

Conditional on the transition rates, we find the statewise reserves. For state 2, the reserve is trivially 0. For state 1, disabled, we find

$$\begin{aligned} V_1^{\mathbf{X}}(t) &= \mathrm{E}\left[\int_t^T e^{-\int_t^s r(\tau)\,\mathrm{d}\tau}\,\mathrm{d}B(s) \middle| \mathcal{F}^{\mathbf{X}}(\infty), Z(t) = 1 \right] \\ &= \int_t^T e^{-\int_t^s \left(r(\tau)+\mu^{\mathrm{i}}(\tau)\right)\mathrm{d}\tau} \left(b^{\mathrm{i}}(s) + \mu^{\mathrm{i}}(s) b^{\mathrm{id}}(s) \right) \mathrm{d}s \end{aligned}$$

For state 0, active, we know from the case of deterministic transition rates, that we can express it in terms of the statewise reserve for disabled, thus

$$
\begin{aligned}
&V_0^{\mathbf{X}}(t) \\
&= \int_t^T e^{-\int_t^s (r(\tau)+\mu^a(\tau)+\nu(\tau))\,\mathrm{d}\tau} \Big(b(s) + \mu^a(s) b^{\mathrm{ad}}(s) + \nu(s) V_1^{\mathbf{X}}(s) \Big)\,\mathrm{d}s \\
&= \int_t^T \Big\{ e^{-\int_t^s (r(\tau)+\mu^a(\tau)+\nu(\tau))\,\mathrm{d}\tau} \Big(b(s) + \mu^a(s) b^{\mathrm{ad}}(s) \Big) \\
&\qquad + \int_s^T e^{-\int_t^u \big(r(\tau)+1_{(\tau\le s)}(\mu^a(\tau)+\nu(\tau))+1_{(\tau\ge s)}\mu^{\mathrm{i}}(\tau)\big)\,\mathrm{d}\tau} \nu(s) \Big(b^{\mathrm{i}}(u) + \mu^{\mathrm{i}}(u) b^{\mathrm{id}}(u) \Big)\,\mathrm{d}u \Big\}\,\mathrm{d}s
\end{aligned}
$$

Now, the liabilities are, assuming the policyholder is active,

$$
\begin{aligned}
&V(t) \\
&= \mathrm{E}\left[V_0^{\mathbf{X}}(t) \middle| \mathcal{F}(t) \right] \\
&= \int_t^T \Big\{ \mathrm{E}\left[e^{-\int_t^s (r(\tau)+\mu^a(\tau)+\nu(\tau))\,\mathrm{d}\tau} \middle| \mathcal{F}(t) \right] b(s) \\
&\quad + \mathrm{E}\left[e^{-\int_t^s (r(\tau)+\mu^a(\tau)+\nu(\tau))\,\mathrm{d}\tau} \mu^a(s) \middle| \mathcal{F}(t) \right] b^{\mathrm{ad}}(s) \\
&\quad + \int_s^T \Big(\mathrm{E}\left[e^{-\int_t^u \big(r(\tau)+1_{(\tau\le s)}(\mu^a(\tau)+\nu(\tau))+1_{(\tau\ge s)}\mu^{\mathrm{i}}(\tau)\big)\,\mathrm{d}\tau} \nu(s) \middle| \mathcal{F}(t) \right] b^{\mathrm{i}}(u) \\
&\quad + \mathrm{E}\left[e^{-\int_t^u \big(r(\tau)+1_{(\tau\le s)}(\mu^a(\tau)+\nu(\tau))+1_{(\tau\ge s)}\mu^{\mathrm{i}}(\tau)\big)\,\mathrm{d}\tau} \nu(s)\mu^{\mathrm{i}}(u) \middle| \mathcal{F}(t) \right] b^{\mathrm{id}}(u) \Big)\,\mathrm{d}u \Big\}\,\mathrm{d}s.
\end{aligned}
$$

To find the four expectations, we use Theorem 2.3, Theorem 2.4 twice and Theorem 2.8, respectively. First, let

$$
I^1(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad I^{2,s}(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1_{(t\le s)} & 0 & 0 \\ 0 & 0 & 1_{(t>s)} & 0 \\ 0 & 0 & 0 & 1_{(t\le s)} \end{bmatrix}.
$$

With this notation, we have

$$
r(t) + \mu^{\mathrm{a}}(t) + \nu(t) = \mathbb{1}^\top I^1(t) \left(c(t) + \Gamma(t) X(t) \right).
$$

Therefore, define $c^1(t) = I^1(t)c(t)$ and $\Gamma^1(t) = I^1(t)\Gamma(t)$, and let $Y^1(t) = c^1(t) + \Gamma^1(t)X(t)$. Then the two first expectations can be rewritten as

$$
\begin{aligned}
\mathrm{E}\left[e^{-\int_t^s (r(\tau)+\mu^a(\tau)+\nu(\tau))\,\mathrm{d}\tau} \middle| \mathcal{F}(t) \right] &= \mathrm{E}\left[e^{-\int_t^s \mathbb{1}^\top Y^1(\tau)\,\mathrm{d}\tau} \middle| \mathcal{F}(t) \right], \\
\mathrm{E}\left[e^{-\int_t^s (r(\tau)+\mu^a(\tau)+\nu(\tau))\,\mathrm{d}\tau} \mu^a(s) \middle| \mathcal{F}(t) \right] &= \mathrm{E}\left[e^{-\int_t^s \mathbb{1}^\top Y^1(\tau)\,\mathrm{d}\tau} Y_2^1(s) \middle| \mathcal{F}(t) \right].
\end{aligned}
$$

By applying Theorem 2.3 to the first expectation and Theorem 2.4 to the second expectation, solutions are obtained.

For the third and fourth expectation, we can proceed analogously. First, see that

$$r(t) + 1_{(t\leq s)}\left(\mu^a(t) + \nu(t)\right) + 1_{(t>s)}\mu^{\mathrm{i}}(t) = \mathbb{1}^\top\left(c^{2,s}(t) + \Gamma^{2,s}(t)X(t)\right).$$

Thus, let $c^{2,s}(t) = I^{2,s}(t)c(t)$ and $\Gamma^{2,s}(t) = I^{2,s}(t)\Gamma(t)$, and let $Y^{2,s}(t) = c^{2,s}(t) + \Gamma^{2,s}(t)X(t)$. Then the third and fourth expectation can be rewritten as

$$\begin{aligned}
&\mathrm{E}\left[e^{-\int_t^u \left(r(\tau)+1_{(\tau\leq s)}(\mu^a(\tau)+\nu(\tau))+1_{(\tau>s)}\mu^{\mathrm{i}}(\tau)\right)\mathrm{d}\tau}\nu(s)\,\middle|\,\mathcal{F}(t)\right]\\
&= \mathrm{E}\left[e^{-\int_t^u Y^{2,s}(\tau)\,\mathrm{d}\tau}Y_4^{2,s}(s)\,\middle|\,\mathcal{F}(t)\right]\\
&\mathrm{E}\left[e^{-\int_t^u \left(r(\tau)+1_{(\tau\leq s)}(\mu^a(\tau)+\nu(\tau))+1_{(\tau>s)}\mu^{\mathrm{i}}(\tau)\right)\mathrm{d}\tau}\nu(s)\mu^{\mathrm{i}}(u)\,\middle|\,\mathcal{F}(t)\right]\\
&= \mathrm{E}\left[e^{-\int_t^u Y^{2,s}(\tau)\,\mathrm{d}\tau}Y_4^{2,s}(s)Y_3^{2,s}(u)\,\middle|\,\mathcal{F}(t)\right].
\end{aligned}$$

Then, applying Theorem 2.4 and Theorem 2.8 respectively, solutions are obtained.

We note, as earlier, that the expectations are found for each value of s (and u). Therefore, to integrate the solution, one needs to discretise the integrals and solve the differential equations for the solution of the expectations for each s (and u). ○

Acknowledgements

I would like to thank my ph.d.-supervisors Thomas Møller and Mogens Steffensen for their guidance and support during this work.

References

[1] Enrico Biffis. Affine processes for dynamic mortality and actuarial valuations. *Insurance: Mathematics & Economics*, 37:443–468, 2005.

[2] Kristian Buchardt. Affine Processes in Life Insurance Mathematics - With a View Towards Solvency II. `http://math.ku.dk/~buchardt`, 2011. Thesis for the Master degree in Insurance Mathematics, Department of Mathematical Sciences, University of Copenhagen.

[3] Mikkel Dahl. Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts. *Insurance: Mathematics & Economics*, 35 (1):113–136, 2004.

[4] Mikkel Dahl and Thomas Møller. Valuation and Hedging of Life Insurance Liabilities with Systematic Mortality Risk. *Insurance: Mathematics & Economics*, 39 (2):193–217, 2006.

[5] Darrell Duffie, Jun Pan, and Kenneth Singleton. Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica*, 68, No. 6:1343–1376, 2000.

[6] Damir Filipovic. *Term Structure Models.* Spring Finance. Springer, 2009.

[7] Domenico De Giovanni. Lapse Rate Modeling: A Rational Expectation Approach. Finance Research Group Working Papers, University of Aarhus, Aarhus School of Business, Department of Business Studies, 2007.

[8] Jarrow, R. A., Lando D., and Turnbull, S. M. A Markov model for the term structure of credit risk spreads. *Journal of Finance*, 50:53–85, 1995.

[9] Holger Kraft and Mogens Steffensen. Bankruptcy, Counterparty Risk, and Contagion. *Review of Finance*, pages 209–252, 2007.

[10] David Lando. On Cox Processes and Credit Risky Securities. *Review of Derivatives Research*, 2:99–120, 1998.

[11] Thomas Møller and Mogens Steffensen. *Market-Valuation Methods in Life and Pension Insurance.* International Series on Actuarial Science. Cambridge University Press, 2007.

[12] Ragnar Norberg. Forward mortality and other vital rates - Are they the way forward? *Insurance: Mathematics & Economics*, 47 (2):105–112, 2010.