European Actuarial Journal
https://doi.org/10.1007/s13385-023-00363-3

ORIGINAL RESEARCH PAPER

Two-dimensional forward and backward transition rates

Theis Bathke[1] **· Marcus C. Christiansen**[1]

Received: 22 April 2022 / Revised: 10 February 2023 / Accepted: 1 September 2023

Abstract
Forward transition rates were originally introduced with the aim to evaluate life insurance liabilities market-consistently. While this idea turned out to have its limitations, recent literature repurposes forward transition rates as a tool for avoiding Markov assumptions in the calculation of life insurance reserves. While life insurance reserves are some form of conditional first-order moments, the calculation of conditional second-order moments needs an extension of the forward transition rate concept from one dimension to two dimensions. Two-dimensional forward transition rates are also needed for the calculation of path-dependent life insurance cash-flows as they occur upon contract modifications. Forward transition rates are designed for doing prospective calculations, and by a time-symmetric definition of so-called backward transition rates one can do retrospective calculations.

Keywords Life & health insurance · Non-Markov modelling · Prospective & retrospective reserves · Second-order moments · Free-policy option

1 Introduction

Forward transition rates describe the expected number of future transitions conditional on the currently available information. If the current information is incomplete, then backward transition rates serve as a proxy for the number of past transitions. However, forward and backward transition rates do not describe the inter-temporal dependency structure of two successive jump events, so they are unsuitable for the calculation of second-order moments or for dealing with path-dependent insurance cash-flows as they occur upon contract modifications. For

✉ Theis Bathke
theis.bathke@uni-oldenburg.de

Marcus C. Christiansen
marcus.christiansen@uni-oldenburg.de

[1] Carl von Ossietzky Universität Oldenburg, Institut für Mathematik, Oldenburg, Germany

Published online: 07 October 2023

this reason, this paper introduces two-dimensional forward and backward transition rates and explains their use in life insurance calculations.

The research on forward transition rates originally emerged from the desire to calculate market values for life insurance liabilities. Miltersen and Persson [12] suggested to define mortality rates implicitly in such a way that the classical actuarial formulas reproduce market values instead of real-world expectations. They denoted these implicit mortality rates as forward mortality rates, inspired by the concept of forward interest rates from financial mathematics. Norberg [14] developed generalizations of the forward concept to multi-state life insurance but observed that the implicit definitions are lacking uniqueness. Christiansen and Niemeyer [7] observed that the implicit definitions often do not have any solution at all. So it turned out that the implicit concept has serious limitations. Buchardt, Furrer and Steffensen [3] suggested to overcome these limitations by introducing additional artificial states, but their concept is restricted to insurance contracts that have sojourn payments only and no transition payments.

Buchardt [2] discarded the implicit concept and presented an explicit definition that works for a doubly stochastic Markov framework. Buchardt actually drifts away from market-consistent valuation but without clearly mentioning.

Christiansen [5] discovered that Buchardt's approach can be extended to a general recipe for calculating real-world expectations in fully non-Markovian life insurance frameworks. Christiansen and Furrer [6] enriched this non-Markov concept with a change of measure technique that makes it possible to deal also with path-dependent insurance payments. To put it into a nutshell, we find two divergent concepts of forward transition rates in the actuarial literature. First, there is the implicit concept which aims to reproduce market values, first suggested by Miltersen and Persson [12] and taken on by numerous further authors. Second, there are the explicit definitions of Buchardt [2], Christiansen [5], and Christiansen and Furrer [6] that put the focus back on real-world expectations and use forward transition rates as a tool to cope with complex inter-temporal dependency structures in life insurance models. This paper follows the second line of thought.

The change of measure technique in Christiansen and Furrer [6] transfers the complexity of path-dependent insurance cash-flows to auxiliary probabilistic models. This way the forward transition rates may stay one-dimensional, but each insurance cash-flow needs another probabilistic model. Our aim is to have one probabilistic model only, and we achieve that by expanding the concept of forward transition rates to two dimensions. The price for having just one probabilistic model is an increased numerical effort that comes with the extra dimension. So the approach of Christiansen and Furrer [6] is beneficial when many scenarios are calculated for one and the same cash-flow, whereas the results of this paper are favourable when many different cash-flows ought to be calculated.

One-dimensional forward transition rates are designed as a tool for calculating expectations. Our two-dimensional concept can be used as a tool for calculating second-order moments. The calculation of second-order moments in classical Markov models was first outlined in Hoem [9]. Helwich [8] presents general formulas for the calculation of variances in semi-Markov models.

Calculation formulas for second-order moments in fully non-Markovian models do not exist yet in the literature. By introducing two-dimensional forward and backward transition rates we help to close that gap.

Markov modelling has a long tradition in life insurance. The classical concept to model the random pattern of states of the insured as a Markov process has been extended to semi-Markov modelling and further sophisticated Markov structures by numerous authors. The problem with any kind of Markov assumptions is that they come with model risk, while at the same time there is a lack of tools for quantifying the model error. This motivates the search for non-Markovian calculation techniques, and this paper is a major step into that direction.

The paper is structured as follows: Sect. 2 introduces the fundamental life insurance modelling framework. Section 3 introduces the definition of two-dimensional forward and backward transitions rates and develops a corresponding integral equation that generalizes Kolmogorov's forward equation. For the latter integral equation Sect. 4 verifies the uniqueness of solutions. Section 5 then turns to the main purpose of two-dimensional forward and backward transitions rates, namely the calculation of certain conditional moments. In Sect. 6 we explain the calculation of conditional variances. Section 7 illustrates the calculation of path-dependent cash-flows. Section 8 concludes and gives an outlook on open research questions.

2 Life insurance modelling framework

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$. We consider an individual life insurance contract and describe the status of the individual insured by an adapted càdlàg jump process

$$Z = (Z(t))_{t\geq 0}$$

on a finite state space $\mathcal{Z}$. Additionally, we set $Z_{0-} := Z_0$. Throughout this paper we assume that we are currently at time $s \geq 0$. So the time interval $[0, s]$ represents to the past and the present, and the time interval (s, ∞) represents the future. The real number $s \geq 0$ is arbitrary but fixed. Based on Z we define state indicator processes $(I_i)_{i\in\mathcal{Z}}$ by

$$I_i(t) := \mathbb{1}_{\{Z(t)=i\}}, \quad t \geq 0,$$

and transition counting processes $(N_{ij})_{i,j\in\mathcal{Z}:i\neq j}$ by

$$N_{ij}(t) := \#\{u \in (0, t] : Z(u-) = i, Z(u) = j\}, \quad t \geq 0.$$

We generally assume that

$$\mathbb{E}[N_{ij}(t)] < \infty, \quad t \geq 0\,, i,j \in \mathcal{Z},\ i \neq j, \tag{2.1}$$

which in particular implies that Z has almost surely no explosions. Let

$$
\begin{aligned}
N_{ii}(t) &:= -\sum_{\substack{j \in \mathcal{Z} \\ j \neq i}} (N_{ij}(t) - N_{ij}(s)), \quad t > s, \\
N_{ii}(t) &:= -\sum_{\substack{j \in \mathcal{Z} \\ j \neq i}} (N_{ji}(t) - N_{ji}(s)), \quad t \leq s.
\end{aligned} \tag{2.2}
$$

This construction is made so that $N_{ii}(t)$ satisfies

$$
\begin{aligned}
\mathrm{d}N_{ii}(t) &= -\sum_{\substack{j \in \mathcal{Z} \\ j \neq i}} \mathrm{d}N_{ij}(t), \quad t > s, \\
\mathrm{d}N_{ii}(t) &= -\sum_{\substack{j \in \mathcal{Z} \\ j \neq i}} \mathrm{d}N_{ji}(t), \quad t \leq s,
\end{aligned}
$$

and is càdlàg everywhere. Definition (2.2) and many further definitions following below involve a case differentiation between $t > s$ and $t \leq s$. As time s is fixed, we omit it in the notation, but one should keep in mind the dependence of many of our definitions on the parameter s. The following equations show a useful direct link between the processes $(N_{ij})_{i,j\in\mathcal{Z}}$ and $(I_i)_{i\in\mathcal{Z}}$:

$$
I_i(t) = I_i(s) + \sum_{j\in\mathcal{Z}} \int_{(s,t]} N_{ji}(\mathrm{d}u), \quad t \geq s, i \in \mathcal{Z}. \tag{2.3}
$$

and

$$
I_i(t) = I_i(s) + \sum_{j\in\mathcal{Z}} \int_{(t,s]} N_{ij}(\mathrm{d}u), \quad s \geq t, i \in \mathcal{Z}. \tag{2.4}
$$

The latter integrals and all following integrals in this paper are meant as path-wise Lebesgue-Stieltjes integrals.

The sigma-algebra $\mathcal{F}_s$ represents the available information at time s. In insurance practice, the insurer often uses a reduced information set $\mathcal{G}_s$ for actuarial evaluations. In this paper we generally assume that

$$
\sigma(Z(s)) \subseteq \mathcal{G}_s \subseteq \mathcal{F}_s. \tag{2.5}
$$

The special case $\mathcal{G}_s = \sigma(Z(s))$ is known as the as-if-Markov model, since it uses information of Markov-type only. The choice of $\mathcal{G}_s$ can be influenced by many factors. Some of these are listed here:

- By cutting down the information used, one can reduce the complexity of numerical calculations.

- A lack of data for statistical inference may make it necessary to simplify the information model.
- For some actuarial tasks it is sufficient to study mean portfolio values only, and then it is convenient to minimize the individual information.
- Anti-discrimination regulation may be a limiting factor for the use of information, such as unisex calculation.
- Data privacy regulation can restrict the information that the insurer actually gathers and stores. For example, the General Data Protection Regulation of the European Union gives individuals considerable control and rights over their personal data.

Let B be the insurance cash-flow of the individual life insurance contract, here assumed to be an adapted càdlàg process with paths of finite variation. We assume that the insurance contract has a maximum contract horizon of T, which means that

$$B(\mathrm{d}t) = 0, \quad t > T.$$

Throughout this paper we assume that

$$0 \leq s \leq T < \infty.$$

Let κ be a càdlàg function that describes the value development of a savings account. We assume that κ is strictly positive so that the corresponding discounting function $1/\kappa$ exists. The random variable

$$Y^{+} := \int_{(s,T]} \frac{\kappa(s)}{\kappa(u)} B(\mathrm{d}u) \tag{2.6}$$

describes the discounted accumulated future cash-flow seen from time s, and the random variable

$$Y^{-} := \int_{[0,s]} \frac{\kappa(s)}{\kappa(u)} B(\mathrm{d}u) \tag{2.7}$$

is the compounded accumulated past cash-flow. In order to ensure integrability, we generally assume that $1/\kappa(\cdot)$ is bounded on $[0, T]$.

Classical Markov modelling focusses on benefit payment functions that may depend on current states or current jumps of the insured but not on past random events. This means that B is supposed to have a so-called one-dimensional canonical representation.

Definition 2.1 A stochastic process A is said to have a *one-dimensional canonical representation* if there exist real-valued functions $(A_i)_i$ on $[0, \infty)$ that generate finite signed measures $A_i(\mathrm{d}u)$ and there exist measurable and bounded real functions $(a_{ij})_{ij:i \neq j}$ such that

$$A(t) = \sum_{i \in \mathcal{Z}} \int_{[0,t]} I_i(u^-) A_i(\mathrm{d}u) + \sum_{\substack{i,j \in \mathcal{Z} \\ i \neq j}} \int_{[0,t]} a_{ij}(u) N_{ij}(\mathrm{d}u), \quad t \geq 0. \tag{2.8}$$

Insurance cash-flows that involve contract modifications, such as a free-policy option, can not be brought into the form (2.8). For this reason we need to allow for more complex structures.

Definition 2.2 A stochastic process A is said to have a *two-dimensional canonical representation* if there exist real-valued functions $(A_i)_i$ on $[0,\infty)$ that generate finite signed measures $A_i(\mathrm{d}u_1)$, real-valued functions $(A_{ij})_{ij}$ on $[0,\infty)^2$ that generate finite signed measures $A_{ij}(\mathrm{d}u_1, \mathrm{d}u_2)$, and measurable and bounded real-valued functions $(a_{ikl})_{ikl:k\neq l}$, $(a_{ijkl})_{ijkl:i\neq j,k\neq l}$ on $[0,\infty)^2$ such that

$$\begin{aligned} A(t) = &\sum_{i,j \in \mathcal{Z}} \int_{[0,t]^2} I_i(u_1^-) I_j(u_2^-) A_{ij}(\mathrm{d}u_1, \mathrm{d}u_2) \\ &+ \sum_{\substack{i,k,l \in \mathcal{Z} \\ k \neq l}} \int_{[0,t]^2} I_i(u_1^-) a_{ikl}(u_1,u_2) A_i(\mathrm{d}u_1) N_{kl}(\mathrm{d}u_2) \\ &+ \sum_{\substack{i,j,k,l \in \mathcal{Z} \\ i \neq j, k \neq l}} \int_{[0,t]^2]} a_{ijkl}(u_1,u_2) N_{ij}(\mathrm{d}u_1) N_{kl}(\mathrm{d}u_2), \quad t_1, t_2 \geq 0. \end{aligned} \tag{2.9}$$

Example 2.3 Suppose that B is an insurance cash flow that has a one-dimensional canonical representation with respect to suitable functions $(B_i)_i$ and $(b_{ij})_{ij:i\neq j}$. Then the squared process B^2 has a two-dimensional canonical representation:

$$\begin{aligned} B^2(t) = &\sum_{i,j \in \mathcal{Z}} \int_{[0,t]^2} I_i(u_1^-) I_j(u_2^-) B_i(\mathrm{d}u_1) B_j(\mathrm{d}u_2) \\ &+ \sum_{\substack{i,k,l \in \mathcal{Z} \\ k \neq l}} \int_{[0,t]^2} 2 I_i(u_1^-) b_{kl}(u_2) B_i(\mathrm{d}u_1) N_{kl}(\mathrm{d}u_2) \\ &+ \sum_{\substack{i,j,k,l \in \mathcal{Z} \\ i \neq j, k \neq l}} \int_{[0,t]^2} b_{ij}(u_1) b_{kl}(u_2) N_{ij}(\mathrm{d}u_1) N_{kl}(\mathrm{d}u_2), \quad t \geq 0. \end{aligned} \tag{2.10}$$

In order to see this, use integration by parts according to Proposition A.2 and Fubini's theorem.

Example 2.4 Suppose that $\mathcal{Z} = \mathcal{S} \times \{0,1\}$, where the elements of $\mathcal{S}$ describe the health status of the individual insured and $\{0,1\}$ indicates whether the policyholder has exercised a free-policy option. We assume here that the free-policy option can be exercised only once and that the policy cannot move back to a premium-paying status. Let τ be the random time where the free-policy option is actually exercised, i.e. τ gives the time where Z moves from the set $\mathcal{S}_0 := \mathcal{S} \times \{0\}$ to the set $\mathcal{S}_1 := \mathcal{S} \times \{1\}$. At time τ the insurance payment scheme is rescaled by a factor $\rho(\tau, Z_{\tau^-}, Z_\tau)$ in order to maintain actuarial equivalence, cf. [6]. So, by writing C for the insurance payment scheme, the insurance cash-flow B equals

$$B(t) = \int_{[0,t]} \rho(\tau, Z_{\tau^-}, Z_\tau)^{\mathbb{1}_{\{\tau \le u\}}} C(\mathrm{d}u).$$

We assume that C has a finite horizon of T. For the sake of simplicity, let

$$\mathbb{1}_{\{\tau=u\}} C(\mathrm{d}u) = 0, \quad \forall i \in \mathcal{Z}, \tag{2.11}$$

almost surely, so that we almost surely have no lump sum payments at time τ. Suppose that the payment scheme C has a one-dimensional canonical representation with respect to suitable functions $(C_i)_i$ and $(c_{ij})_{ij:i\neq j}$. The cash-flow B can be decomposed to

$$B(t) = \int_{[0,t]} \mathbb{1}_{\{u<\tau\}} C(\mathrm{d}u) + \int_{[0,t]} \mathbb{1}_{\{u \ge \tau\}} \rho(\tau, Z_{\tau^-}, Z_\tau) C(\mathrm{d}u). \tag{2.12}$$

We will now analyse both integrals separately. Beginning with the first one, we have

$$\int_{[0,t]} \mathbb{1}_{\{u<\tau\}} C(\mathrm{d}u) = \sum_{i\in\mathcal{S}_0} \int_{[0,t]} \mathbb{1}_{\{u<\tau\}} I_i(u^-) C_i(\mathrm{d}u) + \sum_{\substack{i,j \in \mathcal{S}_0 \\ i \neq j}} \int_{[0,t]} \mathbb{1}_{\{u<\tau\}} c_{ij}(u) N_{ij}(\mathrm{d}u)$$

since $\mathbb{1}_{\{u<\tau\}} I_i(u^-) = 0$ and $\mathbb{1}_{\{u<\tau\}} N_{ij}(\mathrm{d}u) = 0$ for all $(i,j) \notin \mathcal{S}_0^2$. Because of (2.11) and the fact that $\mathbb{1}_{\{u\le\tau\}} I_i(u^-) = I_i(u^-)$ and $\mathbb{1}_{\{u<\tau\}} N_{ij}(\mathrm{d}u) = N_{ij}(\mathrm{d}u)$ for $i,j \in \mathcal{S}_0$, we furthermore get

$$\int_{[0,t]} \mathbb{1}_{\{u<\tau\}} C(\mathrm{d}u) = \sum_{i\in\mathcal{S}_0} \int_{[0,t]} I_i(u^-) C_i(\mathrm{d}u) + \sum_{\substack{i,j \in \mathcal{S}_0 \\ i \neq j}} \int_{[0,t]} c_{ij}(u) N_{ij}(\mathrm{d}u)$$

almost surely. So the insurance cash-flow prior to τ almost surely has a one-dimensional canonical representation. Since τ is the unique jump time of the counting process $\sum_{k\in\mathcal{S}_0} \sum_{l\in\mathcal{S}_1} N_{kl}$, by using assumption (2.11) the second integral in (2.12) can be almost surely transformed to

$$
\begin{aligned}
&\int_{[0,t]} \mathbb{1}_{\{u \geq \tau\}} \rho(\tau, Z_{\tau^-}, Z_\tau) C(\mathrm{d}u) \\
&= \sum_{k\in\mathcal{S}_0} \sum_{l\in\mathcal{S}_1} \int_{[0,t]^2} \mathbb{1}_{\{u_1 \geq u_2\}} \rho(u_2, k, l) C(\mathrm{d}u_1) N_{kl}(\mathrm{d}u_2) \\
&= \sum_{k\in\mathcal{S}_0} \sum_{i,l\in\mathcal{S}_1} \int_{[0,t]^2} I_i(u_1^-) \rho(u_2, k, l) C_i(\mathrm{d}u_1) N_{kl}(\mathrm{d}u_2) \\
&\quad + \sum_{k\in\mathcal{S}_0} \sum_{\substack{i,j,l \in \mathcal{S}_1 \\ i \neq j}} \int_{[0,t]^2} \rho(u_2, k, l) c_{ij}(u_1) N_{ij}(\mathrm{d}u_1) N_{kl}(\mathrm{d}u_2).
\end{aligned}
$$

So, at and after time τ the insurance cash-flow almost surely has a two-dimensional canonical representation. All in all, we obtain for the insurance cash-flow B the almost sure representation

$$
\begin{aligned}
B(t) = &\sum_{i\in\mathcal{S}_0} \int_{[0,t]} I_i(u^-) C_i(\mathrm{d}u) + \sum_{\substack{k,l \in \mathcal{S}_0 \\ k \neq l}} \int_{[0,t]} c_{kl}(u) N_{kl}(\mathrm{d}u) \\
&+ \sum_{k\in\mathcal{S}_0} \sum_{i,l\in\mathcal{S}_1} \int_{[0,t]^2} I_i(u_1^-) \rho(u_2, k, l) C_i(\mathrm{d}u_1) N_{kl}(\mathrm{d}u_2) \\
&+ \sum_{k\in\mathcal{S}_0} \sum_{\substack{i,j,l \in \mathcal{S}_1 \\ i \neq j}} \int_{[0,t]^2} \rho(u_2, k, l) c_{ij}(u_1) N_{ij}(\mathrm{d}u_1) N_{kl}(\mathrm{d}u_2), \quad t \geq 0.
\end{aligned}
\tag{2.13}
$$

3 Two-dimensional transition rates

This section generalizes the forward and backward transition rates of [5] from one dimension to two dimensions. We still suppose that we are currently at time s and have the information $\mathcal{G}_s$ available. As the parameter s is fixed we omit in the notation.

Let $P_i = (P_i(t))_{t\geq 0}$ and $Q_{ij} = (Q_{ij}(t))_{t\geq 0}$ for $i, j \in \mathcal{Z}$ be the almost surely unique càdlàg processes that satisfy

$$
\begin{aligned}
P_i(t) &= \mathbb{E}[I_i(t)|\mathcal{G}_s], \\
Q_{ij}(t) &= \mathbb{E}[N_{ij}(t)|\mathcal{G}_s].
\end{aligned}
$$

As we already mentioned after definition (2.2), the process N_{ii} is càdlàg.

Let $P_{ik} = (P_{ik}(t_1, t_2))_{t_1,t_2 \geq 0}$ and $Q_{ijkl} = (Q_{ijkl}(t_1, t_2))_{t_1,t_2 \geq 0}$ for $i, j, k, l \in \mathcal{Z}$ be the almost surely unique random surfaces that are càdlàg in each variable and satisfy

$$\begin{aligned} P_{ik}(t_1, t_2) &= \mathbb{E}[I_i(t_1) I_k(t_2) | \mathcal{G}_s], \\ Q_{ijkl}(t_1, t_2) &= \mathbb{E}[N_{ij}(t_1) N_{kl}(t_2) | \mathcal{G}_s]. \end{aligned}$$

In order to find suitable modifications of the conditional expectations so that the processes have the postulated càdlàg path properties, one can calculate the conditional expectations on the basis of a fixed regular conditional distribution $\mathbb{P}(\cdot | \mathcal{G}_s)$. Then the càdlàg path properties follow directly from the càdlàg path properties of $(I_i)_i$ and $(N_{ij})_{ij}$ and the dominated convergence theorem. The càdlàg path properties imply that the processes P_i, Q_{ij} and the surfaces P_{ik}, Q_{ijkl} are uniquely given by their values at rational time points only, which are countably many, so the whole processes and surfaces are almost surely unique.

In the following the superscript '$\pm$' is a short notation for

$$f(u^{\pm}) := \begin{cases} f(u^-) & : u > s, \\ f(u) & : u \leq s. \end{cases}$$

Moreover, for the sake of brief formulas, we introduce the processes

$$\tilde{P}_{ij}(u) := \begin{cases} P_i(u) & : u > s, \\ P_j(u) & : u \leq s, \end{cases}$$

and

$$\tilde{P}_{ijkl}(u_1, u_2) := \begin{cases} P_{ik}(u_1, u_2) & : u_1, u_2 > s, \\ P_{jk}(u_1, u_2) & : u_1 \leq s, u_2 > s, \\ P_{il}(u_1, u_2) & : u_1 > s, u_2 \leq s, \\ P_{jl}(u_1, u_2) & : u_1, u_2 \leq s. \end{cases}$$

Definition 3.1 For $i, j \in \mathcal{Z}$ let the stochastic process $(\Lambda_{ij}(t))_{t \geq 0}$ be defined by

$$\Lambda_{ij}(t) = \int_{(s \wedge t, s \vee t]} \frac{\mathbb{1}_{\{\tilde{P}_{ij}(u^{\pm}) > 0\}}}{\tilde{P}_{ij}(u^{\pm})} Q_{ij}(\mathrm{d}u), \tag{3.1}$$

and for $i, j, k, l \in \mathcal{Z}$ let the random surface $(\Lambda_{ijkl}(t_1, t_2))_{t_1,t_2 \geq 0}$ be defined by

$$\Lambda_{ijkl}(t_1, t_2) = \int_{(s \wedge t_1, s \vee t_1] \times (s \wedge t_2, s \vee t_2]} \frac{\mathbb{1}_{\{\tilde{P}_{ijkl}(u_1^{\pm}, u_2^{\pm}) > 0\}}}{\tilde{P}_{ijkl}(u_1^{\pm}, u_2^{\pm})} Q_{ijkl}(\mathrm{d}u_1, \mathrm{d}u_2). \tag{3.2}$$

In case of $t_1, t_2 > s$ we denote $\Lambda_{ij}(\mathrm{d}t_1)$ as (one-dimensional) forward transition rate and $\Lambda_{ijkl}(\mathrm{d}t_1, \mathrm{d}t_2)$ as two-dimensional forward transition rate. In case of $t_1, t_2 \leq s$ we call them backward transition rates. In order to ensure existence of (3.1) and (3.2) we generally assume that

$$\int_{[0,t]} \frac{\mathbb{1}_{\{\tilde{P}_{ij}(t^{\pm})>0\}}}{\tilde{P}_{ij}(t^{\pm})} Q_{ij}(\mathrm{d}t) < \infty, \quad t \geq 0,$$
$$\int_{[0,t]^2} \frac{\mathbb{1}_{\{\tilde{P}_{ijkl}(t_1^{\pm},t_2^{\pm})>0\}}}{\tilde{P}_{ijkl}(t_1^{\pm}, t_2^{\pm})} Q_{ijkl}(\mathrm{d}t_1, \mathrm{d}t_2) < \infty, \quad t \geq 0, \tag{3.3}$$

almost surely for all $i, j, k, l \in \mathcal{Z}$.

According to [5] it holds that

$$P_i(t) = I_i(s) + \sum_{j \in \mathcal{Z}} \int_{(s \wedge t, s \vee t]} P_j(u^{\pm}) \tilde{\Lambda}_{ji}(\mathrm{d}u), \quad t \geq 0, \tag{3.4}$$

for all $i \in \mathcal{Z}$, where

$$\tilde{\Lambda}_{ij}(u) := \begin{cases} \Lambda_{ij}(u) & : \ u > s, \\ \Lambda_{ji}(u) & : \ u \leq s. \end{cases}$$

This is a generalization of Kolmogorov's forward equation to non-Markov models.

Remark 3.2 We introduce $u^{\pm}$ as it allows us to use forward calculation for $t > s$ and backward calculation for $t \leq s$. Consider the case $t > s$, where $u^{\pm} = u^{-}$. Then Definition 3.1 implies

$$P_i(t) = \sum_{\substack{j \in \mathcal{Z} \\ j \neq i}} P_j(t^-) \Delta\Lambda_{ji}(t) - P_i(t^-) \Big(\sum_{\substack{j \in \mathcal{Z} \\ j \neq i}} \Delta\Lambda_{ij}(t) - 1 \Big),$$

when $\Delta\Lambda = \Lambda(t) - \Lambda(t^-)$ contains the jump-discontinuities of Λ. Which allows us to obtain $P_i(t), i \in \mathcal{Z}$, from $P_i(t^-), i \in \mathcal{Z}$. If we replaced Definition 3.1 by

$$\Lambda_{ij}(t) = \int_{(s,t]} \frac{\mathbb{1}_{\{\tilde{P}_j(u)>0\}}}{\tilde{P}_j(u)} Q_{ij}(\mathrm{d}u), \quad t > s, \tag{3.5}$$

then we would get

$$P_i(t) = \frac{P_i(t^-) - \sum_{\substack{j \in \mathcal{Z} \\ j \neq i}} P_j(t) \Delta\Lambda_{ij}(t)}{1 - \sum_{\substack{j \in \mathcal{Z} \\ j \neq i}} \Delta\Lambda_{ji}(t)}, \quad t > s.$$

which fails as a forward formula in case that

$$\sum_{\substack{j \in \mathcal{Z} \\ j \neq i}} \Delta\Lambda_{ji}(t) = 1.$$

This case can occur for example when semi-Markov effects are modelled by state space expansions as in [1]. If we replaced Definition 3.1 by

$$\Lambda_{ij}(t) = \int_{(s,t]} \frac{\mathbb{1}_{\{\tilde{P}_i(u)>0\}}}{\tilde{P}_i(u)} Q_{ij}(\mathrm{d}u), \quad t < s,$$

then Theorem 6.3 in [5] fails. In the case $t < s$ analogous problems occur. Our definition allows us to do backward calculation on $[0, s]$.

We now extend (3.4) from one to two dimensions.

Theorem 3.3 *The processes* $(P_{ik})_{i,k\in\mathcal{Z}}$, $(\Lambda_{ijkl})_{i,j,k,l\in\mathcal{Z}}$,$(P_i)_{i\in\mathcal{Z}}$ *and* $(\Lambda_{ij})_{i,j\in\mathcal{Z}}$ *almost surely satisfy the equation*

$$\begin{aligned} P_{ik}(t_1,t_2) = I_i(s)I_k(s) &+ I_k(s) \int_{(s\wedge t_1, s\vee t_1]} \sum_{j\in\mathcal{Z}} P_j(u^{\pm})\tilde{\Lambda}_{ji}(\mathrm{d}u) + I_i(s) \int_{(s\wedge t_2, s\vee t_2]} \sum_{l\in\mathcal{Z}} P_l(u^{\pm})\tilde{\Lambda}_{lk}(\mathrm{d}u) \\ &+ \int_{(s\wedge t_1, s\vee t_1]\times(s\wedge t_2, s\vee t_2]} \sum_{j,l \in \mathcal{Z}} P_{jl}(u_1^{\pm}, u_2^{\pm})\tilde{\Lambda}_{jilk}(\mathrm{d}u_1, \mathrm{d}u_2) \end{aligned} \tag{3.6}$$

for $t_1, t_2 \geq 0$ *and* $i, k \in \mathcal{Z}$, *where*

$$\tilde{\Lambda}_{ijkl}(u_1,u_2) := \begin{cases} \Lambda_{ijkl}(u_1,u_2) & u_1, u_2 > s \\ \Lambda_{jikl}(u_1,u_2) & u_1 \leq s, u_2 > s \\ \Lambda_{ijlk}(u_1,u_2) & u_1 > s, u_2 \leq s \\ \Lambda_{jilk}(u_1,u_2) & u_1, u_2 \leq s. \end{cases}$$

Proof We show the proof for the case $t_1, t_2 > s$ only. For the other cases the proof is similar. Let $\mathbb{P}_\omega(\cdot)$ be a regular version of the conditional distribution $\mathbb{P}(\cdot|\mathcal{G}_s)$ and let $\mathbb{E}_\omega[\cdot]$ be the Lebesgue-Stieltjes integral of the argument with respect to the integrator $\mathbb{P}_\omega$. By applying Campbell's theorem, see Section A.1, and Fubini's theorem, we get for $(a,b], (c,d] \subset [s,\infty)$ and almost all $\omega \in \Omega$

$$
\begin{aligned}
&\int_{(a,b]\times(c,d]} \mathbb{1}_{\{P_{ik}(u_1^-,u_2^-)(\omega)>0\}} Q_{ijkl}(\mathrm{d}(u_1,u_2))(\omega) \\
&\quad = \mathbb{E}_\omega\left[\int_{(a,b]\times(c,d]} \mathbb{1}_{\{P_{ik}(u_1^-,u_2^-)(\omega)>0\}} N_{ij}(\mathrm{d}u_1) N_{kl}(\mathrm{d}u_2)\right] \\
&\quad = \mathbb{E}_\omega\left[\int_{(a,b]\times(c,d]} I_i(u_1^-) I_k(u_2^-) N_{ij}(\mathrm{d}u_1) N_{kl}(\mathrm{d}u_2)\right] \\
&\quad = \int_{(a,b]\times(c,d]} Q_{ijkl}(\mathrm{d}(u_1,u_2))(\omega)
\end{aligned}
\tag{3.7}
$$

since $P_{ik}(u_1^-, u_2^-)(\omega) = 0$ implies $I_i(u_1^-)I_k(u_2^-) = 0$ $\mathbb{P}_\omega(\cdot)$-almost surely and $I_i(u_1^-)N_{ij}(\mathrm{d}u_1) = N_{ij}(\mathrm{d}u_1)$. By applying (2.3), we get for $t_1, t_2 \geq s$ and $i, k \in \mathcal{Z}$

$$
\begin{aligned}
P_{ik}(t_1,t_2) &= \mathbb{E}[I_i(t_1)I_k(t_2)|\mathcal{G}_s] \\
&= \mathbb{E}\Big[I_i(s)I_k(s)\Big|\mathcal{G}_s\Big] + \mathbb{E}\Big[I_k(s)\int_{(s,t_1]}\sum_{j\in\mathcal{Z}} N_{ji}(\mathrm{d}t)\Big|\mathcal{G}_s\Big] + \mathbb{E}\Big[I_i(s)\int_{(s,t_2]}\sum_{l\in\mathcal{Z}} N_{lk}(\mathrm{d}t)\Big|\mathcal{G}_s\Big] \\
&\quad + \mathbb{E}\Big[\sum_{j,l\in\mathcal{Z}}\int_{(s,t_1]\times(s,t_2]} N_{ji}(\mathrm{d}t_1)N_{lk}(\mathrm{d}t_2)\Big|\mathcal{G}_s\Big].
\end{aligned}
$$

By using the assumption (2.5), the definition of $(Q_{ijkl})_{ijkl}$, the definition of $(Q_{ij})_{i,j}$, Fubini's theorem, and Campbell's theorem, we can conclude that

$$
\begin{aligned}
P_{ik}(t_1,t_2) &= I_i(s)I_k(s) + I_k(s)\int_{(s,t_1]}\sum_{j\in\mathcal{Z}} Q_{ji}(\mathrm{d}u) + I_i(s)\int_{(s,t_2]}\sum_{l\in\mathcal{Z}} Q_{lk}(\mathrm{d}u) \\
&\quad + \sum_{j,l\in\mathcal{Z}}\int_{(s,t_1]\times(s,t_2]} Q_{jilk}(\mathrm{d}u_1,\mathrm{d}u_2)
\end{aligned}
$$

almost surely. The assertion follows now from (3.2) and (3.7). □

Remark 3.4 The integral equation (3.6) can be solved in different ways. For example one could use an algorithm similar to Finite Element methods. The idea is to use a simple box integration method to approximate the integral on small sub-rectangles of $(s, T] \times (s, T]$. Let $t_1, t_2, ...$ be a partition of $(s, T]$ we use the following approximation procedure:

$$
\int_{(t_i,t_{i+1}]\times(t_j,t_{j+1}]} f(x,y)G(\mathrm{d}x,\mathrm{d}y) \approx f(t_i,t_j)\int_{(t_i,t_{i+1}]\times(t_j,t_{j+1}]} G(\mathrm{d}x,\mathrm{d}y).
$$

This leaves us with the following formula:

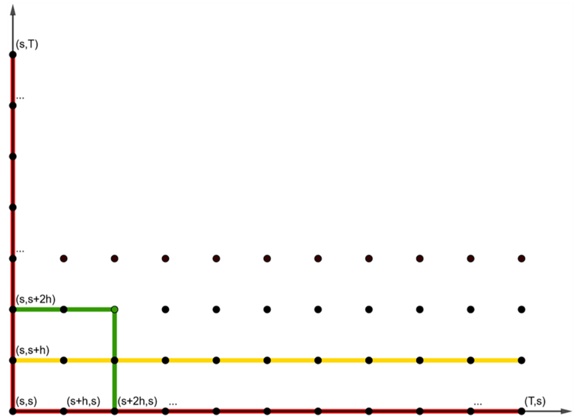

Fig. 1 Sketch of an algorithm to calculate $P_{ik}(t_1, t_2)$

$$
\begin{aligned}
P_{ik}(t_q, t_p) \approx\ & I_i(s)I_k(s) + I_k(s) \sum_{j\in\mathcal{Z}} \sum_{n=0}^{q-1} P_j(t_n)\Lambda_{ji}((t_n, t_{n+1}]) \\
& + I_i(s) \sum_{l\in\mathcal{Z}} \sum_{\tilde{m}=0}^{p-1} P_l(t_m)\Lambda_{lk}((t_m, t_{m+1}]) \\
& + \sum_{j,l\in\mathcal{Z}} \sum_{m=0}^{p-1} \sum_{n=0}^{q-1} P_{jl}(t_n, t_m)\Lambda_{jilk}((t_n, t_{n+1}] \times (t_m, t_{m+1}]),
\end{aligned}
$$

under the conventions $\Lambda_{ij}((a, b]) := \int_{(a,b]} \Lambda_{ij}(\mathrm{d}x)$, and $\Lambda_{ijkl}((a, b] \times (c, d]) := \int_{(a,b]\times(c,d]} \Lambda_{ijkl}(\mathrm{d}x, \mathrm{d}y)$. Combining this with an analogous procedure for the one-dimensional case, we obtain a recursion scheme.

We first calculate the points on the red lines in Fig. 1 from the centre to the edge. Then we calculate every point on the yellow line from left to right and proceed with the same notion for the points above. The reason this works is that the calculation of every point only depends on Points in the bottom left rectangle of the point, see the green rectangle for reference. This allows us to approximate P_{ik} on all the grid points, and then we approximate P_{ik} on the whole rectangle.

For an empirical Λ as a pure jump process, the approximation of the integral is exact, and our algorithm calculates $P_{ij}, i, j \in \mathcal{Z}$ accurately on the whole square.

4 Uniqueness of solutions of the integral equations

Equation (3.4) is commonly used for the calculation of the state occupation probabilities $(P_i)_i$ from given one-dimensional transition rates. Likewise, equation (3.6) may be used in order to calculate the two-dimensional state occupation probabilities $(P_{ik})_{ik}$ from the one-dimensional and two-dimensional transition rates, but it is crucial then that $(P_{ik})_{ik}$ is the only solution.

Theorem 4.1 *There exists an almost surely unique solution* $(P_i)_i$, $(P_{ik})_{ik}$ *to the stochastic integral equation system formed by equations* (3.4) *and* (3.6).

Proof The existence of a solution follows from Theorem 3.3, so it remains to show that the solution is almost surely unique. As the equations (3.4) and (3.6) are almost surely pathwise integral equations, in the remaining proof we identify without loss of generality all stochastic processes and random surfaces with just one of their paths. These paths have to be chosen such that (3.3) holds.

We show the proof for the case $t_1, t_2 > s$ only. For the other cases the proof is similar. We are going to use a fixed-point argument. For any choice of $T \in (s, \infty)$, the set

$$BV_2^{|\mathcal{Z}|} := \Big\{ f : [s,T] \times [s,T] \to \mathbb{R}^{|\mathcal{Z}|^2} \Big| \text{ there exist finite signed measures } \mu_{ik} \\ \text{with } f_{ik}(x,y) = \mu_{ik}([s,x] \times [s,y]), x, y \in [s,T] \times [s,T] \Big\}$$

is a linear space. The Hahn-Jordan decomposition offers for any finite signed measure a unique decomposition into a difference of two finite measures, and this decomposition can be furthermore used to decompose also any $f \in BV_2^{|\mathcal{Z}|}$ into a difference $f = f^+ - f^-$ of nonnegative mappings $f^+, f^- \in BV_2^{|\mathcal{Z}|}$. Based on this unique construction of f^+ and f^-, we then define $|f| := f^+ + f^-$. By equipping $BV_2^{|\mathcal{Z}|}$ with the norm

$$\|(f_{ik})_{i,k\in\mathcal{Z}}\| := \sum_{i,k\in\mathcal{Z}} \int_{[s,T]\times[s,T]} |f_{ik}|(\mathrm{d}t_1, \mathrm{d}t_2) + \sum_{i,k\in\mathcal{Z}} |f_{ik}(s,s)|$$

we obtain a metric space. On this metric space, we define an operator $O : BV_2^{|\mathcal{Z}|} \to BV_2^{|\mathcal{Z}|}$ as follows:

$$\big(O((f_{jl})_{j,l\in\mathcal{Z}})\big)_{ik}(t_1,t_2) := \sum_{j,l\in\mathcal{Z}} \int_{(s,t_1]\times(s,t_2]} f_{jl}(u_1^-, u_2^-) \Lambda_{jilk}(\mathrm{d}u_1, \mathrm{d}u_2)$$

for $t_1, t_2 \in [s,T]$. We want this operator to be a contraction, but unfortunately this is not true, so we need to replace our norm for $BV_2^{|\mathcal{Z}|}$ by another equivalent norm. Let

$$\nu(t_1,t_2) := 4 \sum_{\substack{i,j,k,l\in\mathcal{Z}\\ i\neq j\\ k\neq l}} \Lambda_{ijkl}(t_1,t_2), \quad t_1, t_2 \in [s,T].$$

Because of assumption (3.3), each Λ_{ijkl}, $i \neq j$, $k \neq l$, has an associated measure $\mu_{\Lambda_{ijkl}}$ that is almost surely finite on $[s,T]^2$. Hence, there exists a finite measure μ that satisfies

$$\nu(t_1,t_2) = \mu([s,t_1] \times [s,t_2]), \quad t_1, t_2 \in [s,T]. \tag{4.1}$$

We moreover define

$$
\begin{aligned}
\nu_1(u) &:= \mu([s,u] \times [s,T]),\\
\nu_2(u) &:= \mu([s,T] \times [s,u]), \quad u \in [s,T],
\end{aligned}
$$

which are càdlàg by construction. For each $K \in (0,\infty)$ the mapping $\|\cdot\|_K$ defined by

$$
\|(f_{ij})_{i,j\in\mathcal{Z}}\|_K := \sum_{i,j\in\mathcal{Z}} \int_{[s,T]\times[s,T]} e^{-K(\nu_1(u_1)+\nu_2(u_2))} |f_{ij}|(\mathrm{d}u_1,\mathrm{d}u_2) + \sum_{i,k\in\mathcal{Z}} |f_{ik}(s,s)|
$$

is a norm on $BV_2^{|\mathcal{Z}|}$ that is equivalent to the norm $\|\cdot\|$. This construction of $\|\cdot\|_K$ is inspired by [4] but is here extended to the two-dimensional case. For any $f \in BV_2^{|\mathcal{Z}|}$, $(a,b]\times(c,d] \subset [s,T]\times[s,T]$ and $k,i \in \mathcal{Z}$, the definitions of operator O and the two-dimensional transition rates $(\Lambda_{ijkl})_{i,j,k,l\in\mathcal{Z}}$ in conjunction with the triangle inequality yield that

$$
\begin{aligned}
\int_{(a,b]\times(c,d]} |O(f)_{ik}|(\mathrm{d}u_1,\mathrm{d}u_2) \leq \sum_{\substack{l\,:\,l\neq k\\ j\,:\,j\neq i}} \Bigg(&\int_{(a,b]\times(c,d]} |f_{jl}(u_1^-,u_2^-)|\Lambda_{jilk}(\mathrm{d}u_1,\mathrm{d}u_2)\\
&+ \int_{(a,b]\times(c,d]} |f_{jk}(u_1^-,u_2^-)|\Lambda_{jikl}(\mathrm{d}u_1,\mathrm{d}u_2)\\
&+ \int_{(a,b]\times(c,d]} |f_{il}(u_1^-,u_2^-)|\Lambda_{ijlk}(\mathrm{d}u_1,\mathrm{d}u_2)\\
&+ \int_{(a,b]\times(c,d]} |f_{ik}(u_1^-,u_2^-)|\Lambda_{ijkl}(\mathrm{d}u_1,\mathrm{d}u_2)\Bigg).
\end{aligned}
$$

Summation over i, k and a reordering of some of the resulting sums lead to the inequality

$$
\begin{aligned}
\sum_{i,k} \int_{(a,b]\times(c,d]} |O(f)_{ik}|(\mathrm{d}u_1,\mathrm{d}u_2) &\leq 4 \sum_{\substack{i,j,k,l\\ i\neq j,\,k\neq l}} \int_{(a,b]\times(c,d]} |f_{ik}(u_1^-,u_2^-)|\Lambda_{ijkl}(\mathrm{d}u_1,\mathrm{d}u_2)\\
&\leq 4 \sum_{i,k} \int_{(a,b]\times(c,d]} |f_{ik}(u_1^-,u_2^-)|\,\nu(\mathrm{d}u_1,\mathrm{d}u_2).
\end{aligned}
$$

Suppose now that $f(s,s)$ is zero. Then it holds that

$$
\sum_{i,k} \int_{(a,b]\times(c,d]} |O(f)_{ik}|(\mathrm{d}u_1,\mathrm{d}u_2) \leq 4 \sum_{k,i} \int_{(a,b]\times(c,d]} \int_{[s,u_1)\times[s,u_2)} |f_{ik}|(\mathrm{d}r_1,\mathrm{d}r_2)\,\nu(\mathrm{d}t_1,\mathrm{d}t_2).
$$

As a consequence, the norm $\|\cdot\|_K$ of $O(f)$ has an upper bound of

$$
\begin{aligned}
\|O(f)\|_K &= \sum_{i,k} \int_{[s,T]\times[s,T]} e^{-K(v_1(u_1)+v_2(u_2))} |O(f)_{ik}|(\mathrm{d}u_1, \mathrm{d}u_2) \\
&\leq \sum_{i,k} \int_{[s,T]\times[s,T]} e^{-K(v_1(u_1)+v_2(u_2))} \int_{[s,u_1)\times[s,u_2)} |f_{ik}|(\mathrm{d}r_1, \mathrm{d}r_2)\, v(\mathrm{d}u_1, \mathrm{d}u_2) \\
&= \sum_{i,k} \int_{[s,T]\times[s,T]} \int_{(r_1,T]\times(r_2,T]} e^{-K(v_1(u_1)+v_2(u_2))} v(\mathrm{d}u_1, \mathrm{d}u_2)) |f_{ik}|(\mathrm{d}r_1, \mathrm{d}r_2),
\end{aligned}
\tag{4.2}
$$

where the last equality uses Fubini's theorem. Moreover, for arbitrary but fixed $u_1, u_2 \in [s, T]$ let

$$
\tilde{v}(r_1, r_2) := \frac{\mu((u_1, r_1] \times (u_2, r_2])}{\mu((u_1, T] \times (u_2, T])}, \quad r_1 \in (u_1, T],\ r_2 \in (u_2, T],
$$

for the same μ is in equation (4.1). Without loss of generality we assume that $\mu((u_1, T] \times (u_2, T]) > 0$. Otherwise the conclusion that we want to draw is trivial. Then

$$
\begin{aligned}
\tilde{v}_1(r_1) &:= \tilde{v}(r_1, T), \\
\tilde{v}_2(r_2) &:= \tilde{v}(T, r_2), \quad (r_1, r_2) \in (u_1, T] \times (u_2, T],
\end{aligned}
$$

correspond to cumulative distribution functions. Let $C := \tilde{v}(u_1, u_2) > 0$ and let (A, B) be a random vector that has $\tilde{v}$ as its two-dimensional cumulative distribution function. Then, by applying Sklar's theorem, see e.g. Theorem 2.3.3 in [13], we can show that

$$
\begin{aligned}
&\int_{(u_1,T]\times(u_2,T]} e^{-K(v_1(r_1)+v_2(r_2))} v(\mathrm{d}r_1, \mathrm{d}r_2) \\
&\quad \leq C e^{-K(v_1(u_1)+v_2(u_2))} \int_{(u_1,T]\times(u_2,T]} e^{-CK(\tilde{v}_1(r_1)+\tilde{v}_2(r_2))} \tilde{v}(\mathrm{d}r_1, \mathrm{d}r_2) \qquad (4.3) \\
&\quad = C e^{-K(v_1(u_1)+v_2(u_2))}\, \mathbb{E}\left[e^{-CK(\tilde{v}_1(\tilde{v}_1^{-1}(U))+\tilde{v}_2(\tilde{v}_2^{-1}(V)))} \right]
\end{aligned}
$$

for a suitable random vector (U, V) whose components are uniformly distributed on $(0, 1)$ and such that (A, B) and $(\tilde{v}_1^{-1}(U), \tilde{v}_2^{-1}(V))$ have the same distribution. Note here that the copula of (U, V) may be non-trivial. The inverse functions $\tilde{v}_1^{-1}$ and $\tilde{v}_2^{-1}$ are here defined as $\tilde{v}_n^{-1}(t) := \inf\{x : \tilde{v}_n(x) \geq t\}$, $n = 1, 2$ for $t \in (0, 1)$. Since $\tilde{v}_1(\tilde{v}_1^{-1}(t)) \geq t$ for $t \in (0, 1)$, see e.g. Theorem 3.1 in [17], the inequality (4.3) has an upper bound of

$$
\int_{(u_1,T]\times(u_2,T]} e^{-K(v_1(r_1)+v_2(r_2))} v(\mathrm{d}r_1, \mathrm{d}r_2) \leq C e^{-K(v_1(u_1)+v_2(u_2))}\, \mathbb{E}\left[e^{-CK(U+V)} \right].
$$

By applying Theorem 10.6.4 in [10], we can show that the latter expectation has an upper bound of

$$\mathbb{E}\left[e^{-CK(U+V)}\right] \leq \mathbb{E}\left[e^{-CK(2U)}\right]$$

since $x \mapsto \exp\{-CKx\}$ is a convex function. As U is uniformly distributed on (0, 1), we moreover have

$$\begin{aligned}\mathbb{E}\left[e^{-CK(2U)}\right] &= \int_{(0,1)} e^{-CK(2u)}\mathrm{d}u \\ &= \frac{1}{2KC}\left(1 - e^{-2KC}\right) \\ &\leq \frac{1}{2KC}.\end{aligned}$$

We set $K = 1$. Then, all in all we can conclude that the inequality (4.2) has an upper bound of

$$\begin{aligned}\|O(f)\|_K &\leq \frac{1}{2K}\sum_{i,k}\int_{[s,T]\times[s,T]} e^{-K(v_1(u_1)+v_2(u_2))}|f_{ik}|(\mathrm{d}u_1, \mathrm{d}u_2) \\ &= \frac{1}{2}\|f\|_K\end{aligned}$$

whenever $f(s, s)$ equals zero. Suppose now that we have two solutions $R = (P_{ik})_{ik}$ and $R' = (P'_{ik})_{ik}$ of (3.6) for given $(P_i)_i$. Then $R - R' \in BV_2^{|\mathcal{Z}|}$ is a fixed point of the operator O and $(R - R')(s, s)$ is zero. So we have

$$\|R - R'\|_K \leq \frac{1}{2}\|R - R'\|_K,$$

which necessarily implies that $R - R'$ is zero on $[s, T] \times [s, T]$. In an analogous way it is possible to proof $R = R'$ also on the three rectangles $[0, s]^2$, $[0, s] \times (s, T]$ and $(s, T] \times [0, s]$. Since $T \in (s, \infty)$ was arbitrary, we can expand the uniqueness property to infinity.

With the same ideas that we used in this proof, one can also show the uniqueness of a solution $(P_i)_{i\in\mathcal{Z}}$ of (3.4) with respect to $(\Lambda_{ij})_{i,j\in\mathcal{Z}}$. The one-dimensional case is actually even simpler. The equation system formed by equations (3.4) and (3.6) can then have only one solution $(P_i)_i$, $(P_{ik})_{ik}$. This completes the proof. □

5 Conditional expectations of canonical representations

According to [5], for any stochastic process A that has a one-dimensional canonical representation of the form (2.8), it holds that

$$\mathbb{E}\Big[\int_{(s,T]} A(\mathrm{d}t) \Big| \mathcal{G}_s \Big] = \sum_{i\in\mathcal{Z}} \int_{(s,T]} P_i(t^-) A_i(\mathrm{d}t) + \sum_{\substack{i,j\in\mathcal{Z}\\ i\neq j}} \int_{(s,T]} a_{ij}(t) P_i(t^-) \Lambda_{ij}(\mathrm{d}t) \tag{5.1}$$

and

$$\mathbb{E}\Big[\int_{[0,s]} A(\mathrm{d}t) \Big| \mathcal{G}_s \Big] = \sum_{i\in\mathcal{Z}} \int_{[0,s]} P_i(t^-) A_i(\mathrm{d}t) + \sum_{\substack{i,j\in\mathcal{Z}\\ i\neq j}} \int_{[0,s]} a_{ij}(t) P_j(t) \Lambda_{ij}(\mathrm{d}t) \tag{5.2}$$

almost surely. These formulas are classical in the life insurance literature in case that Z is a Markov process, and the recent contribution of [5] was to show that Markov assumptions are actually not needed here. The formulas (5.1) and (5.2) are used in life insurance for the calculation of prospective and retrospective reserves at time s. They are typically applied as follows: For given one-dimensional transition rates, calculate the corresponding one-dimensional state occupation probabilities by solving (3.4), and then solve the integrals in (5.1) and (5.2) in order to obtain the desired conditional expectations. The solving usually happens numerically.

The following three propositions will allow us to generalize (5.1) and (5.2) to stochastic processes A that have a two-dimensional canonical representation according to (2.9).

Proposition 5.1 *Let A_{ij}, $i,j \in \mathcal{Z}$, be a real-valued function on $[0,\infty)^2$ that generates a finite signed measure. Then we almost surely have*

$$\mathbb{E}\Big[\int_{[0,T]^2} I_i(u_1^-) I_j(u_2^-) A_{ij}(\mathrm{d}u_1, \mathrm{d}u_2) \Big| \mathcal{G}_s \Big] = \int_{[0,T]^2} P_{ij}(u_1^-, u_2^-) A_{ij}(\mathrm{d}u_1, \mathrm{d}u_2).$$

Proof The assertion follows from Fubini's theorem. ◻

Proposition 5.2 *Let A_i, $i \in \mathcal{Z}$, be a real-valued function on $[0,\infty)$ that generates a finite signed measure, and let a_{ikl}, $i,k,l \in \mathcal{Z}$, $k \neq l$, be a measurable and bounded real-valued function on $[0,\infty)^2$. Then we almost surely have*

$$\begin{aligned}
&\mathbb{E}\Big[\int_{[0,s]\times[0,T]} I_i(u_1^-) a_{ikl}(u_1,u_2) A_i(\mathrm{d}u_1) N_{kl}(\mathrm{d}u_2) \Big| \mathcal{G}_s \Big] \\
&= I_i(s) \int_{[0,s]\times[0,T]} a_{ikl}(u_1,u_2) A_i(\mathrm{d}u_1) \tilde{P}_{kl}(u_2^{\pm}) \Lambda_{kl}(\mathrm{d}u_2) \\
&\quad + \sum_{j\in\mathcal{Z}} \int_{[0,s]} \int_{[0,T]\times[u_1,s]} a_{ikl}(u_1,u_2) \tilde{P}_{klij}(u_2^{\pm}, u_3) \Lambda_{klij}(\mathrm{d}u_2, \mathrm{d}u_3) A_i(\mathrm{d}u_1)
\end{aligned}$$

and

$$
\begin{aligned}
&\mathbb{E}\Bigg[\int_{(s,T]\times[0,T]} I_i(u_1^-)a_{ikl}(u_1,u_2)A_i(\mathrm{d}u_1)N_{kl}(\mathrm{d}u_2)\Bigg|\mathcal{G}_s\Bigg]\\
&= I_i(s)\int_{(s,T]\times[0,T]} a_{ikl}(u_1,u_2)A_i(\mathrm{d}u_1)\tilde{P}_{kl}(u_2^{\pm})\Lambda_{kl}(\mathrm{d}u_2)\\
&\quad+\sum_{j\in\mathcal{Z}}\int_{(s,T]}\int_{[0,T]\times(s,u_1)} a_{ikl}(u_1,u_2)\tilde{P}_{klji}(u_2^{\pm},u_3^-)\Lambda_{klji}(\mathrm{d}u_2,\mathrm{d}u_3)A_i(\mathrm{d}u_1).
\end{aligned}
$$

Proof From (2.3), (2.4) we get

$$
I_i(t^-) = I_i(s) + \int_{W(t)} \sum_{j\in\mathcal{Z}} N_{ji}^{W(t)}(du)
$$

for $W(t) := (s,t)$ in case of $t > s$ and $W(t) := [t,s]$ in case of $t \le s$ and $N_{ij}^{[t,s]}(t) := N_{ji}(t)$ and $N_{ij}^{(s,t)}(t) := N_{ij}(t)$. This equation and assumption (2.5) imply for $D \in \{[0,s],(s,T]\}$ the equation

$$
\begin{aligned}
&\mathbb{E}\Bigg[\int_{D\times[0,T]} I_i(u_1^-)a_{ikl}(u_1,u_2)A_i(\mathrm{d}u_1)N_{kl}(\mathrm{d}u_2)\Bigg|\mathcal{G}_s\Bigg]\\
&= \mathbb{E}\Bigg[\int_{D\times[0,T]} I_i(s)a_{ikl}(u_1,u_2)A_i(\mathrm{d}u_1)N_{kl}(\mathrm{d}u_2)\Bigg|\mathcal{G}_s\Bigg]\\
&\quad+\sum_{j\in\mathcal{Z}}\mathbb{E}\Bigg[\int_D\int_{[0,T]\times W(u_1)} a_{ikl}(u_1,u_2)N_{kl}(\mathrm{d}u_2)N_{ji}^{W(u_1)}(\mathrm{d}u_3)A_i(\mathrm{d}u_1)\Bigg|\mathcal{G}_s\Bigg].
\end{aligned}
$$

Now apply Fubini's theorem and Campbell's theorem in order to obtain

$$
\begin{aligned}
&\mathbb{E}\Bigg[\int_{D\times[0,T]} I_i(u_1^-)a_{ikl}(u_1,u_2)A_i(\mathrm{d}u_1)N_{kl}(\mathrm{d}u_2)\Bigg|\mathcal{G}_s\Bigg]\\
&= \int_{D\times[0,T]} I_i(s)a_{ikl}(u_1,u_2)A_i(\mathrm{d}u_1)Q_{kl}(\mathrm{d}u_2)\\
&\quad+\int_D\int_{[0,T]\times W(u_1)} a_{ikl}(u_1,u_2)Q_{klji}^{W(u_1)}(\mathrm{d}u_2,\mathrm{d}u_3)A_i(\mathrm{d}u_1),
\end{aligned}
$$

for $Q_{klji}^{[0,s]}(t_1,t_2) = Q_{klij}(t_1,t_2)$ and $Q_{klji}^{(s,T]}(t_1,t_2) = Q_{klji}(t_1,t_2)$. Finally, use the equations

$$
\begin{aligned}
Q_{kl}(\mathrm{d}u_2) &= \tilde{P}_{kl}(u_2^{\pm})\Lambda_{kl}(\mathrm{d}u_2),\\
Q_{klij}(\mathrm{d}u_2,\mathrm{d}u_3) &= \tilde{P}_{klij}(u_2^{\pm},u_3^{\pm})\Lambda_{klij}(\mathrm{d}u_2,\mathrm{d}u_3),
\end{aligned}
$$

which follow from the definitions (3.1) and (3.2). □

Proposition 5.3 *Let a_{ijkl}, $i,j,k,l \in \mathcal{Z}$, $i \neq j$, $k \neq l$, be a measurable and bounded real-valued function on* $[0,\infty)^2$. *Then we almost surely have*

$$\mathbb{E}\Bigg[\int_{[0,T]^2} a_{ijkl}(u_1,u_2) N_{ij}(\mathrm{d}u_1) N_{kl}(\mathrm{d}u_2) \Bigg| \mathcal{G}_s \Bigg] = \int_{[0,T]^2} a_{ijkl}(u_1,u_2) \tilde{P}_{ijkl}(u_1^{\pm}, u_2^{\pm}) \Lambda_{ijkl}(\mathrm{d}u_1, \mathrm{d}u_2).$$

Proof The result follows directly from Campbell's theorem and (3.2). □

Corollary 5.4 *Suppose that A is a stochastic process that has a two-dimensional canonical representation according to* (2.9). *Then we almost surely have*

$$\begin{aligned}
&\mathbb{E}\Bigg[\int_{[0,T]^2} A(\mathrm{d}u_1, \mathrm{d}u_2) \Bigg| \mathcal{G}_s \Bigg] \\
&= \sum_{i,j \in \mathcal{Z}} \int_{[0,T]^2} P_{ij}(u_1^-, u_2^-) A_{ij}(\mathrm{d}u_1, \mathrm{d}u_2) \\
&\quad + \sum_{\substack{i,k,l \\ k \neq l}} I_i(s) \int_{[0,T]^2} a_{ikl}(u_1,u_2) A_i(\mathrm{d}u_1) \tilde{P}_{kl}(u_2^{\pm}) \Lambda_{kl}(\mathrm{d}u_2) \\
&\quad + \sum_{\substack{i,j,k,l \\ k \neq l}} \int_{[0,s]} \int_{[0,T]\times[u_1,s]} a_{ikl}(u_1,u_2) \tilde{P}_{klij}(u_2^{\pm}, u_3) \Lambda_{klij}(\mathrm{d}u_2, \mathrm{d}u_3) A_i(\mathrm{d}u_1) \\
&\quad + \sum_{\substack{i,j,k,l \\ k \neq l}} \int_{(s,T]} \int_{[0,T]\times(s,u_1)} a_{ikl}(u_1,u_2) \tilde{P}_{klji}(u_2^{\pm}, u_3^-) \Lambda_{klji}(\mathrm{d}u_2, \mathrm{d}u_3) A_i(\mathrm{d}u_1) \\
&\quad + \sum_{\substack{i,j,k,l \\ i \neq j, k \neq l}} \int_{[0,T]^2} a_{ijkl}(u_1,u_2) \tilde{P}_{ijkl}(u_1^{\pm}, u_2^{\pm}) \Lambda_{ijkl}(\mathrm{d}u_1, \mathrm{d}u_2).
\end{aligned} \tag{5.3}$$

From formula 5.3 we can immediately obtain general expressions for the expected values of discounted future liabilities as well as the pure expected values. The discounting needs to be included in the functions A_{ij}, A_i, a_{ikl}, and a_{ijkl}.

In insurance practice, formula 5.3 may be used as follows: For given one-dimensional and two-dimensional transition rates, calculate the corresponding one-dimensional and two-dimensional state occupation probabilities by solving

(3.4) and (3.6). Then solve the integrals in (5.3) in order to obtain the desired conditional expectation.

6 Conditional variance of the future liabilities

We still think of s as an arbitrary but fixed parameter and omit this parameter in the notation. The discounted accumulated future payments Y^+ according to definition (2.6) are usually not measurable with respect to $\mathcal{G}_s$, so actuaries use the so-called *prospective reserve* at time s

$$V^+ := \mathbb{E}[Y^+|\mathcal{G}_s] \tag{6.1}$$

as a proxy for Y^+. In order to quantify the dispersion of the approximation error $Y^+ - V^+$, the actuary is also interested in the conditional variance

$$\mathrm{Var}[Y^+|\mathcal{G}_s] := \mathbb{E}[(Y^+)^2|\mathcal{G}_s] - (\mathbb{E}[Y^+|\mathcal{G}_s])^2. \tag{6.2}$$

This section discusses the calculation of V^+ and

$$S^+ := \mathbb{E}[(Y^+)^2|\mathcal{G}_s], \tag{6.3}$$

from which we can then get the conditional variance as

$$\mathrm{Var}[Y^+|\mathcal{G}_s] = S^+ - (V^+)^2.$$

Our results here are limited to insurance cash-flows B that have a one-dimensional canonical representation,

$$B(t) = \sum_i \int_{[0,t]} I_i(u^-)B_i(\mathrm{d}u) + \sum_{i,j:j\neq i} \int_{[0,t]} b_{ij}(u)N_{ij}(\mathrm{d}u), \quad t \geq 0. \tag{6.4}$$

From (2.6) and (5.1) we almost surely obtain

$$V^+ = \sum_{i\in\mathcal{Z}} \int_{(s,T]} \frac{\kappa(s)}{\kappa(t)} P_i(t^-)B_i(\mathrm{d}t) + \sum_{\substack{i,j\in\mathcal{Z}\\ i\neq j}} \int_{(s,T]} \frac{\kappa(s)}{\kappa(t)} b_{ij}(t)P_i(t^-)\Lambda_{ij}(\mathrm{d}t).$$

This formula and equation (3.4) allow us to calculate V^+ from given one-dimensional transition rates $(\Lambda_{ij}(t))_{t>s}$, $i,j \in \mathcal{Z}$.

We now turn to the calculation of S^+. Analogously to (2.10), one can show that

$$
\begin{aligned}
(Y^+)^2 = &\sum_{i,j\in\mathcal{Z}} \int_{(s,T]^2} \frac{2\kappa(s)^2}{\kappa(u_1)\kappa(u_2)} I_i(u_1^-)I_j(u_2^-)B_i(\mathrm{d}u_1)B_j(\mathrm{d}u_2)\\
&+ \sum_{\substack{i,j,k\in\mathcal{Z}\\ j\neq k}} \int_{(s,T]^2} \frac{2\kappa(s)^2}{\kappa(u_1)\kappa(u_2)} I_i(u_1^-)b_{jk}(u_2)B_i(\mathrm{d}u_1)N_{jk}(\mathrm{d}u_2)\\
&+ \sum_{\substack{i,j,k,l\in\mathcal{Z}\\ i\neq j, k\neq l}} \int_{(s,T]^2} \frac{2\kappa(s)^2}{\kappa(u_1)\kappa(u_2)} b_{ij}(u_1)b_{kl}(u_2)N_{ij}(\mathrm{d}u_1)N_{kl}(\mathrm{d}u_2).
\end{aligned}
$$

By using definition (6.3), interchanging $\int_{(s,T]^2}$ and $\int_{[0,T]^2} \mathbb{1}_{(s,T]^2}$, and applying Corollary 5.4, we obtain that

$$
\begin{aligned}
S^+ = &\sum_{i,j\in\mathcal{Z}} \int_{(s,T]^2} \frac{2\kappa(s)^2}{\kappa(u_1)\kappa(u_2)} P_{ij}(u_1^-,u_2^-)B_i(\mathrm{d}u_1)B_j(\mathrm{d}u_2)\\
&+ \sum_{\substack{i,k,l\in\mathcal{Z}\\ k\neq l}} I_i(s) \int_{(s,T]^2} \frac{2\kappa(s)^2}{\kappa(u_1)\kappa(u_2)} b_{kl}(u_2)P_k(u_2^-)B_i(\mathrm{d}u_1)\Lambda_{kl}(\mathrm{d}u_2)\\
&+ \sum_{\substack{i,j,k,l\in\mathcal{Z}\\ k\neq l}} \int_{(s,T]} \int_{(s,T]\times(s,u_1)} \frac{2\kappa(s)^2}{\kappa(u_1)\kappa(u_2)} b_{kl}(u_2)P_{kj}(u_2^-,u_3^-)\Lambda_{klji}(\mathrm{d}u_2,\mathrm{d}u_3)B_i(\mathrm{d}u_1)\\
&+ \sum_{\substack{i,j,k,l\in\mathcal{Z}\\ i\neq j, k\neq l}} \int_{(s,T]^2} \frac{2\kappa(s)^2}{\kappa(u_1)\kappa(u_2)} b_{ij}(u_1)b_{kl}(u_2)P_{ik}(u_1^-,u_2^-)\Lambda_{ijkl}(\mathrm{d}u_1,\mathrm{d}u_2)
\end{aligned}
$$

almost surely. This formula and the equations (3.4) and (3.6) allow us to calculate S^+ from given transition rates $(\Lambda_{ij}(t))_{t\in(s,T]}$ and $(\Lambda_{ijkl}(t))_{t\in(s,T]^2}$, $i,j,k,l\in\mathcal{Z}$.

For the conditional expectation and the conditional variance of Y^- we can obtain a similar result. We leave these analogous calculations to the reader.

7 Prospective and retrospective reserve for a path-dependent cash-flow

This section continues with Example 2.4. Recall that (6.1) is the so-called prospective reserve at time s. The *retrospective reserve* at time s is defined as

$$V^- := \mathbb{E}[Y^-|\mathcal{G}_s] \tag{7.1}$$

for Y^- defined by (2.7). Analogously to (2.13) one can show that

$$
\begin{aligned}
Y^- = &\sum_{i\in\mathcal{S}_0}\int_{[0,s]}\frac{\kappa(s)}{\kappa(u)}I_i(u^-)C_i(\mathrm{d}u) + \sum_{\substack{k,l\in\mathcal{S}_0\\k\neq l}}\int_{[0,s]}\frac{\kappa(s)}{\kappa(u)}c_{kl}(u)N_{kl}(\mathrm{d}u)\\
&+\sum_{k\in\mathcal{S}_0}\sum_{i,l\in\mathcal{S}_1}\int_{[0,s]^2}\frac{\kappa(s)}{\kappa(u_1)}I_i(u_1^-)\rho(u_2,k,l)C_i(\mathrm{d}u_1)N_{kl}(\mathrm{d}u_2)\\
&+\sum_{k\in\mathcal{S}_0}\sum_{\substack{i,j,l\in\mathcal{S}_1\\i\neq j}}\int_{[0,s]^2}\frac{\kappa(s)}{\kappa(u_1)}\rho(u_2,k,l)c_{ij}(u_1)N_{ij}(\mathrm{d}u_1)N_{kl}(\mathrm{d}u_2).
\end{aligned}
$$

By using (5.1), interchanging $\int_{[0,s]}$ and $\int_{[0,T]}\mathbb{1}_{[0,s]}$, interchanging $\int_{[0,s]^2}$ and $\int_{[0,T]^2}\mathbb{1}_{[0,s]^2}$, and applying Corollary 5.4, we can conclude that

$$
\begin{aligned}
V^- = &\sum_{i\in\mathcal{S}_0}\int_{[0,s]}\frac{\kappa(s)}{\kappa(u)}P_i(u^-)C_i(\mathrm{d}u) + \sum_{\substack{k,l\in\mathcal{S}_0\\k\neq l}}\int_{[0,s]}\frac{\kappa(s)}{\kappa(u)}c_{kl}(u)P_l(u)\Lambda_{kl}(\mathrm{d}u)\\
&+\sum_{k\in\mathcal{S}_0}\sum_{l,i\in\mathcal{S}_1}I_i(s)\int_{[0,s]^2}\frac{\kappa(s)}{\kappa(u_1)}\rho(u_2,k,l)P_l(u_2)C_i(\mathrm{d}u_1)\Lambda_{kl}(\mathrm{d}u_2)\\
&+\sum_{k\in\mathcal{S}_0}\sum_{l,i\in\mathcal{S}_1}\sum_{j\in\mathcal{Z}}\int_{[0,s]}\int_{[0,s]\times[u_1,s]}\frac{\kappa(s)}{\kappa(u_1)}\rho(u_2,k,l)P_{lj}(u_2,u_3)\Lambda_{klij}(\mathrm{d}u_2,\mathrm{d}u_3)C_i(\mathrm{d}u_1)\\
&+\sum_{k\in\mathcal{S}_0}\sum_{\substack{l,i,j\in\mathcal{S}_1\\i\neq j}}\int_{[0,s]^2}\frac{\kappa(s)}{\kappa(u_1)}\rho(u_2,k,l)c_{ij}(u_1)P_{jl}(u_1,u_2)\Lambda_{ijkl}(\mathrm{d}u_1,\mathrm{d}u_2)
\end{aligned}
$$

almost surely. This formula and the equations (3.4) and (3.6) allow us to calculate V^- from given transition rates $(\Lambda_{ij}(t))_{t\leq s}$ and $(\Lambda_{ijkl}(t))_{t\leq s}$, $i,j,k,l\in\mathcal{Z}$.

By arguing analogously to (2.13), one can moreover show that the discounted future liabilities Y^+ have the representation

$$
\begin{aligned}
Y^+ = &\sum_{i\in\mathcal{S}_0}\int_{(s,T]}\frac{\kappa(s)}{\kappa(u)}I_i(u^-)C_i(\mathrm{d}u) + \sum_{\substack{k,l\in\mathcal{S}_0\\k\neq l}}\int_{(s,T]}\frac{\kappa(s)}{\kappa(u)}c_{kl}(u)N_{kl}(\mathrm{d}u)\\
&+\sum_{k\in\mathcal{S}_0}\sum_{i,l\in\mathcal{S}_1}\int_{(s,T]\times[0,T]}\frac{\kappa(s)}{\kappa(u_1)}I_i(u_1^-)\rho(u_2,k,l)C_i(\mathrm{d}u_1)N_{kl}(\mathrm{d}u_2)\\
&+\sum_{k\in\mathcal{S}_0}\sum_{\substack{i,j,l\in\mathcal{S}_1\\i\neq j}}\int_{(s,T]\times[0,T]}\frac{\kappa(s)}{\kappa(u_1)}\rho(u_2,k,l)c_{ij}(u_1)N_{ij}(\mathrm{d}u_1)N_{kl}(\mathrm{d}u_2).
\end{aligned}
$$

By applying (5.1), interchanging $\int_{(s,T]}$ and $\int_{[0,T]} \mathbb{1}_{(s,T]}$, interchanging $\int_{(s,T]\times[0,T]}$ and $\int_{[0,T]^2} \mathbb{1}_{(s,T]\times[0,T]}$, and applying Corollary 5.4, we almost surely get

$$
\begin{aligned}
V^+ = & \sum_{i\in\mathcal{S}_0} \int_{(s,T]} \frac{\kappa(s)}{\kappa(u)} P_i(u^-) C_i(\mathrm{d}u) + \sum_{\substack{i,j\in\mathcal{S}_0 \\ i\neq j}} \int_{(s,T]} \frac{\kappa(s)}{\kappa(u)} c_{ij}(u) P_i(u^-) \Lambda_{ij}(\mathrm{d}u) \\
& + \sum_{k\in\mathcal{S}_0} \sum_{l,i\in\mathcal{S}_1} I_i(s) \int_{(s,T]\times[0,T]} \frac{\kappa(s)}{\kappa(u_1)} \rho(u_2,k,l) \tilde{P}_{kl}(u_2^{\pm}) C_i(\mathrm{d}u_1) \Lambda_{kl}(\mathrm{d}u_2) \\
& + \sum_{k\in\mathcal{S}_0} \sum_{l,i\in\mathcal{S}_1} \sum_{j\in\mathcal{Z}} \int_{(s,T]} \int_{[0,T]\times(s,u_1)} \frac{\kappa(s)}{\kappa(u_1)} \rho(u_2,l,k) \tilde{P}_{klji}(u_2^{\pm},u_3^{-}) \Lambda_{klji}(\mathrm{d}u_2,\mathrm{d}u_3) C_i(\mathrm{d}u_1) \\
& + \sum_{k\in\mathcal{S}_0} \sum_{\substack{l,i,j\in\mathcal{S}_1 \\ i\neq j}} \int_{(s,T]\times[0,T]} \frac{\kappa(s)}{\kappa(u_1)} \rho(u_2,k,l) c_{ij}(u_1) \tilde{P}_{ijkl}(u_1^{-},u_2^{\pm}) \Lambda_{ijkl}(\mathrm{d}u_1,\mathrm{d}u_2).
\end{aligned}
$$

This formula and the equations (3.4) and (3.6) allow us to calculate V^- from given transition rates $(\Lambda_{ij}(t))_{t\leq T}$ and $(\Lambda_{ijkl}(t))_{t\leq T}$, $i,j,k,l \in \mathcal{Z}$.

8 Conclusion and outlook

So far, the non-Markov calculation technique of Christiansen [5] has been limited to the calculation of first-order moments. By introducing two-dimensional forward and backward transition rates, we make second-order moments accessible as well. The two-dimensional rates capture intertemporal dependency structures that the one-dimensional rates miss. In order to calculate also third-order or even higher-order moments, one may envision further extensions of the forward and backward transition rate concept to three dimensions or even higher dimensions. Such extensions are beyond the scope of this paper and left for future research.

Intertemporal dependency structures play a major role when life insurance cash-flows are path-dependent. Such path dependencies occur for example upon contract modifications. We illustrated how two-dimensional forward and backward transition rates are of help here for actuarial calculations by studying insurance policies with free-policy option.

The non-Markov approach exchanges systematic model risk for unsystematic estimation risk. The latter risk has much nicer asymptotic properties. For the one-dimensional case, the landmark Nelson-Aalen estimator of Putter and Spitoni [16] can be suitably adapted, see [5]. The idea is to use subsampling, so in order to achieve a small estimation error the information model $\mathcal{G}_s$ should be rather small. A positive example is the as-if-Markov model, which records only the current state of the insured. For the two-dimensional case, efficient generalizations of the landmark Nelson-Aalen estimator have yet to be explored.

A Appendix

Theorem A.1 (*Campbell theorem*) *Let η be a point process on $(\mathbb{R}^d, \mathbb{B}(\mathbb{R}^d))$ with intensity measure λ and let $u : \mathbb{R}^d \to \mathbb{R}$ be a measurable function. Then*

$$\int u(x)\eta(\mathrm{d}x)$$

is a random variable and

$$\mathbb{E}\left[\int u(x)\eta(\mathrm{d}x)\right] = \int u(x)\lambda(\mathrm{d}x),$$

whenever $u \geq 0$ or $\int |u(x)|\lambda(\mathrm{d}x) < \infty$.

For the proof see Sect. 2.2 in [11].

Proposition A.2 (integration by parts) *Suppose that $(F(t))_{t\geq 0}$ and $(G(t))_{t\geq 0}$ are real-valued càdlàg processes with paths of finite variation. Then*

$$\int_{(a,b]} F(x)\mathrm{d}G(x) = F(b)G(b) - F(a)G(a) - \int_{(a,b]} G(x^-)\mathrm{d}F(x)$$

for $(a, b] \subset (0, \infty)$.

Proof For the proof see Corollary 2 following Theorem 22 in [15]. □

Acknowledgements We would like to thank the reviewers for insightful remarks that helped to improve the paper.

Funding Open Access funding enabled and organized by Projekt DEAL.

References

1. Amsler MH (1968) Les chaines de markov des assurances vie, invalidité et maladie. In: Transactions of the 18th international congress of actuaries, vol 5, pp 731–746
2. Buchardt K (2017) Kolmogorov's forward pide and forward transition rates in life insurance. Scand Actuar J 2017(5):377–394

3. Buchardt K, Furrer C, Steffensen M (2019) Forward transition rates. Finance Stoch 23(4):975–999
4. Christiansen MC (2010) Biometric worst-case scenarios for multi-state life insurance policies. Insur Math Econ 47(2):190–197
5. Christiansen MC (2021) On the calculation of prospective and retrospective reserves in non-markov models. Eur Actuar J, pp 1–22
6. Christiansen MC, Furrer C (2022) Extension of as-if-markov modeling to scaled payments. Insur Math Econ 107:288–306
7. Christiansen MC, Niemeyer A (2015) On the forward rate concept in multi-state life insurance. Finance Stochas 19(2):295–327
8. Helwich M (2008) Durational effects and non-smooth semi-markov models in life insurance. Doctoral dissertation. University of Rostock
9. Hoem JM (1969) Markov chain models in life insurance. Blätter der DGVFM 9(2):91–107
10. Kaas R, Goovaerts M, Dhaene J, Denuit M (2002) Modern actuarial risk theory: using R. Springer
11. Last G, Penrose M (2017) Lectures on the Poisson process, vol 7. Cambridge University Press
12. Miltersen KR, Persson S-A (2005) Is mortality dead? stochastic forward force of mortality rate determined by no arbitrage. Technical Report
13. Nelsen RB (2007) An introduction to copulas. Springer
14. Norberg R (2010) Forward mortality and other vital rates-are they the way forward? Insur Math Econ 47(2):105–112
15. Protter PE (2005) Stochastic integration and differential equations. Stoch Model Appl Probab 21
16. Putter H, Spitoni C (2018) Non-parametric estimation of transition probabilities in non-markov multi-state models: the landmark Aalen-Johansen estimator. Stat Methods Med Res 27(7):2081–2092
17. Shorack GR, Shorack G (200) Probability for statisticians, vol 951. Springer