INSTITUTE OF MATHEMATICS
of the
Polish Academy of Sciences

ul. Śniadeckich 8, P.O.B. 21, 00-956 Warszawa 10, Poland http://www.impan.pl

IM PAN Preprint 698 (2008)

Jacek Jakubowski and Mariusz Niewęgłowski

F-Doubly Stochastic Markov Chains – A New Class of Processes for Modeling Credit Rating Migration Processes and Valuation of Claims

Presented by Jerzy Zabczyk

Published as manuscript

Received 19 December 2008

$\mathbb{F}$-DOUBLY STOCHASTIC MARKOV CHAINS — A NEW CLASS OF PROCESSES FOR MODELING CREDIT RATING MIGRATION PROCESSES AND VALUATION OF CLAIMS

MARIUSZ NIEWĘGŁOWSKI AND JACEK JAKUBOWSKI

Abstract. We define a new class of processes very useful in applications, $\mathbb{F}$-doubly stochastic Markov chains, which contains Markov chains and are fully characterized by some martingale property. We use this class to model credit rating migrations in financial markets. In such a framework we present a solution of the valuation problem for defaultable rating-sensitive claims.

1. Introduction

Our goal is to find a class of processes with good properties which can be used for modeling a credit rating migration process and which contains processes usually taken for this purpose. This allows us to include rating migration in the process of valuation of defaultable claims and generalize the case where only two states are considered: default and non-default. In Section 2 we introduce a new class of processes, which we call $\mathbb{F}$-doubly stochastic Markov chains. The reason for the name is that there are two sources of uncertainty in their definition, so in analogy to Cox processes, called doubly stochastic Poisson processes, we chose the name "$\mathbb{F}$- doubly stochastic Markov chains". This class contains Markov chains, compound Poisson processes with jumps in $\mathbb{Z}$, Cox processes and the process of rating migration obtained by the canonical construction in Bielecki and Rutkowski [10]. An $\mathbb{F}$-doubly stochastic Markov chain differs from a doubly stochastic Markov chain which is a Markov chain with a doubly stochastic transition matrix. In the following we use the shorthand "$\mathbb{F}$–DS Markov chain" for the "$\mathbb{F}$-doubly stochastic Markov chain". We are interested in using the class of $\mathbb{F}$–DS Markov chains to model rating migrations on financial markets, so we restrict ourselves to processes with values in a finite set $\mathcal{K} = \{1, \ldots, K\}$. We give examples of $\mathbb{F}$-doubly stochastic Markov chains and investigate their properties. We prove that an $\mathbb{F}$–DS Markov chain C is a conditional Markov chain and that any $\mathbb{F}$-martingale is a $\mathbb{F} \vee \mathbb{F}^C$-martingale. This means that the so called hypothesis H holds. Then we introduce the notion of intensity of an $\mathbb{F}$–DS Markov chain and formulate conditions which ensure its existence. We prove that an $\mathbb{F}$–DS Markov chain C with intensity is completely characterized by the martingale property of the compensated process describing the position of C (Theorem 2.14) as well as by the martingale property of the compensated processes counting the number of jumps of C from one state to another (Theorem 2.18). An $\mathbb{F}$–DS Markov chain with a given intensity is constructed. At the end of that section, we investigate how replacing the probability measure by an equivalent one affects the properties of an $\mathbb{F}$–DS Markov chain.

Research supported in part by Polish KBN Grant P03A 034 29 "Stochastic evolution equations driven by Lévy noise".

In Section 3 we apply $\mathbb{F}$–DS Markov chains to model rating migration in financial markets; we are convinced that these processes can also be applied in other fields, e.g. in insurance.

The problem of modeling credit risk taking into consideration rating migration was proposed by Jarrow, Lando and Turnbull [21]. They took Markov chains to model time evolution of credit ratings. Jarrow et al. [21] considered both discrete and continuous time case, and within this framework they derived a valuation formula for defaultable bonds expressed through risk-neutral transition probabilities. Moreover, they proposed an algorithm for calculating risk neutral transition probabilities from the real-world one under an additional assumption on risk premia. Kijima and Komoribayashi [22] argued that the assumption on risk premia formulated by Jarrow et al. [21] may fail to hold in practice and introduced a different assumption which avoids difficulties arising in practical implementation of the model by Jarrow et al. [21]. Subsequently, Lando [23] has extended the framework of Jarrow et al. [21] by constructing a rating migration process which follows a conditional Markov chain (see also Bielecki and Rutkowski [11] for a precise definition of conditional Markov property). In Lando [23] and Bielecki and Rutkowski [11] the generator of the credit rating process follows a matrix-valued stochastic process, which is also the case for an $\mathbb{F}$-DS Markov chain. We also stress that in fact their construction gives $\mathbb{F}$-DS Markov chains. Lando [23] has considered the problem of providing explicit formulae for some credit derivatives connected with ratings, which is also of interest to us in Section 3. Lando has shown that, under the assumption that the generator matrix process has eigenvectors constant in time, it is possible to solve the conditional Kolmogorov equation and obtain explicit formulae for bond prices and rating-dependent payoffs. However, the structure of payoffs considered in [23] was very simple compared to ours: only a terminal payoff contingent on rating at a terminal date was considered. There are many papers that are concerned with pricing of credit derivatives with rating migrations, for example Acharya, Das and Sundaram [2], and Das and Tufano [12] who consider an interesting discrete time HJM model with its tree implementation for pricing purposes. For extension of HJM methodology to the case of defaultable bonds with rating migrations see Bielecki and Rutkowski [9], and for models with Lévy noise see Eberlein and Özkan [14]. For recent papers considering infinite-dimensional noise in HJM type models see Schmidt [31] and Jakubowski and Niewęgłowski [17]. We also mention recent work of Bielecki et al. [6] which deals with the problem of pricing basket derivatives with rating migrations in a very efficient markovian setting. Recently Hurd and Kuznetsov [15], [16] introduced so called affine Markov chains models for valuation of basket credit derivatives with rating migrations. They constructed rating processes as continuous Markov chains with time change via an independent affine process. They show how to price efficiently simple instruments such as defaultable bonds and more complicated ones like CDO's tranches. In Section 3 we consider the problem of valuating defaultable rating-sensitive claims. We assume that the rating migrations process is an $\mathbb{F}$–DS Markov chain. We give a general formula for the form of ex-dividend price process of defaultable rating-sensitive claim in terms of processes defining this claim and characteristics of the rating migration process (see Theorem 3.7). This generalizes the known results obtained for the case without rating migration (see e.g. Bielecki, Jeanblanc and Rutkowski [7]). As an example we give formulas for some known claims such as a defaultable bond with fractional recovery of par value, Credit Sensitive Note and Credit Default Swap.

2. $\mathbb{F}$–doubly stochastic Markov chains — definition and properties

In this section we introduce and investigate a new class of processes, which will be called $\mathbb{F}$-doubly stochastic Markov chains. This class contains Markov chains and Cox processes. Under natural assumptions, belonging of a process X to the class of $\mathbb{F}$-doubly stochastic Markov chains is fully characterized by the martingale property of some processes strictly connected with X. This martingale property allows us to use these processes in modeling rating migrations.

2.1. **Definition and examples.** We assume that all processes are defined on a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$. We also fix a filtration $\mathbb{F}$ satisfying usual conditions, which plays the role of a reference filtration.

Definition 2.1. A càdlàg process C is called an $\mathbb{F}$–doubly stochastic Markov chain with state space $\mathcal{K} \subset \mathbb{Z} = \{\ldots, -1, 0, 1, 2, \ldots\}$ if there exists a family of stochastic matrices $P(s,t) = (p_{i,j}(s,t))_{i,j\in\mathcal{K}}$ for $0 \leq s \leq t$ such that

(1) the matrix $P(s,t)$ is $\mathcal{F}_t$–measurable, and $P(s,\cdot)$ is $\mathbb{F}$ progressively measurable,

(2) for any $t \geq s \geq 0$ and every $i, j \in \mathcal{K}$ we have

$$\mathbf{P}(C_t = j \mid \mathcal{F}_\infty \vee \mathcal{F}_s^C)\mathbf{1}_{\{C_s=i\}} = \mathbf{1}_{\{C_s=i\}} p_{i,j}(s,t). \tag{2.1}$$

The process P will be called the conditional transition probability process of C.

The equality (2.1) implies that $P(t,t) = \mathbb{I}$ a.s. for all $t \geq 0$. Definition 2.1 extends the notion of Markov chain with continuous time (when $\mathcal{F}_\infty$ is trivial). A process satisfying (1) and (2) is called a doubly stochastic Markov chain by analogy with Cox processes (doubly stochastic Poisson processes). In both cases there are two sources of uncertainty. As mentioned in the Introduction, we use the shorthand "$\mathbb{F}$–DS Markov chain" for the "$\mathbb{F}$–doubly stochastic Markov chain". Now, we give a few examples of processes which are $\mathbb{F}$–DS Markov chains.

Example 1 (Compound Poisson process)**.** A compound Poisson process X with jumps in $\mathbb{Z}$ is an $\mathbb{F}$–DS Markov chain. We know that $X_t = \sum_{i=1}^{N_t} Y_i$, where N is a Poisson process with intensity λ, Y_i is a sequence of independent identically distributed random variables with values in $\mathbb{Z}$ and distribution ν. Moreover $(Y_i)_i$ and N are independent. Hence for $\mathcal{F}_\infty = \sigma(N)$, $j \geq i$, $s \leq t$,

$$\begin{aligned}
&\mathbf{P}(X_t = j \mid \mathcal{F}_\infty \vee \mathcal{F}_s^X)\mathbf{1}_{\{X_s=i\}} = \mathbf{P}(X_t = j, X_s = i \mid \mathcal{F}_\infty \vee \mathcal{F}_s^X)\mathbf{1}_{\{X_s=i\}} \\
&= \mathbf{P}(X_t - X_s = j - i \mid \mathcal{F}_\infty)\mathbf{1}_{\{X_s=i\}} = \mathbf{P}\Big(\sum_{m=N_s+1}^{N_t} Y_m = j - i \Big| \mathcal{F}_\infty\Big)\mathbf{1}_{\{X_s=i\}} \\
&= \nu^{\otimes(N_t-N_s)}(j-i)\mathbf{1}_{\{X_s=i\}}.
\end{aligned}$$

Thus

$$p_{i,j}(s,t) = \nu^{\otimes(N_t-N_s)}(j-i)$$

satisfy conditions (1) and (2) of Definition 2.1.

Example 2. Let X be a compound Poisson process as in Example 1. By standard calculations we see that X is an $\mathbb{F}$–DS Markov chain with respect to the trivial filtration with deterministic transition matrix given by the formula

$$p_{i,j}(s,t) = \sum_{k=0}^{\infty} \nu^{\otimes k}(j-i) \frac{[\lambda(t-s)]^k}{k!} e^{-\lambda(t-s)}.$$

From these examples we have seen that the conditional transition probability matrix depends on the choice of the reference filtration $\mathbb{F}$, and $P(s,t)$ can be either continuous with respect to s,t or discontinuous.

Example 3 (Cox process)**.** A Cox process C is an $\mathbb{F}$–DS Markov chain with $\mathcal{K} = \mathbb{N}$. Indeed, the definition of a Cox process implies that

$$\mathbf{P}(C_t - C_s = k \mid \mathcal{F}_\infty \vee \mathcal{F}_s^C) = e^{-\int_s^t \lambda_u du} \frac{\left(\int_s^t \lambda_u du\right)^k}{k!} \tag{2.2}$$

for some $\mathbb{F}$–adapted process λ such that $\lambda \geq 0$, $\int_0^t \lambda_s ds < \infty$ for all $t \geq 0$ and $\int_0^\infty \lambda_s ds = \infty$ a.s. Hence

$$\mathbf{P}(C_t - C_s = k \mid \mathcal{F}_\infty \vee \mathcal{F}_s^C) = \mathbf{P}(C_t - C_s = k \mid \mathcal{F}_\infty),$$

so the increments and the past (i.e. $\mathcal{F}_s^C$) are conditionally independent given $\mathcal{F}_\infty$. Therefore for $j \geq i$,

$$\begin{aligned} &\mathbf{P}(C_t = j \mid \mathcal{F}_\infty \vee \mathcal{F}_s^C)\mathbf{1}_{\{C_s=i\}} = \mathbf{P}(C_t = j, C_s = i \mid \mathcal{F}_\infty \vee \mathcal{F}_s^C)\mathbf{1}_{\{C_s=i\}} \\ &= \mathbf{1}_{\{C_s=i\}}\mathbf{P}(C_t - C_s = j - i \mid \mathcal{F}_\infty \vee \mathcal{F}_s^C) = \mathbf{1}_{\{C_s=i\}} e^{-\int_s^t \lambda_u du} \frac{\left(\int_s^t \lambda_u du\right)^{j-i}}{(j-i)!}. \end{aligned}$$

Thus

$$p_{i,j}(s,t) = \begin{cases} \frac{\left(\int_s^t \lambda_u du\right)^{j-i}}{(j-i)!} e^{-\int_s^t \lambda_u du} & \text{for } j \geq i, \\ 0 & \text{for } j < i, \end{cases}$$

satisfy conditions (1) and (2) of Definition 2.1.

Example 4 (Time changed discrete Markov chain)**.** Assume that $\bar{C}$ is a discrete time Markov chain with values in $\mathcal{K} = \{1, \ldots, K\}$, N is a Cox process and the processes $(\bar{C}_k)_{k\geq 0}$ and $(N_t)_{t\geq 0}$ are independent and conditionally independent given $\mathcal{F}_\infty$. Then the process $C_t := \bar{C}_{N_t}$ is an $\mathbb{F}$–DS Markov chain (see [18, Theorem 7 and 9]).

Simple calculations give us another elementary example:

Example 5 (Truncated Cox process)**.** The process $C_t := \min\{N_t, K\}$, where N is a Cox process and $K \in \mathbb{N}$, is an $\mathbb{F}$–DS Markov chain with state space $\mathcal{K} = \{0, \ldots, K\}$.

2.2. **Properties of $\mathbb{F}$–DS Markov chains in the case of a finite state space.** Since we are interested in using the class of $\mathbb{F}$–DS Markov chains to model financial markets with rating migrations, we restrict ourselves to a finite set $\mathcal{K}$, i.e. $\mathcal{K} = \{1, \ldots, K\}$, with $K < \infty$. Moreover we assume that $C_0 = i_0$ for some $i_0 \in \mathcal{K}$.

We start the investigation of $\mathbb{F}$–DS Markov chains from the very useful lemma describing conditional finite–dimensional distributions of C.

Lemma 2.2. *If C is an $\mathbb{F}$–DS Markov chain, then*

$$\mathbf{P}(C_{u_1} = i_1, \ldots C_{u_n} = i_n \mid \mathcal{F}_\infty \vee \mathcal{F}^C_{u_0})\mathbf{1}_{\{C_{u_0}=i_0\}} = \mathbf{1}_{\{C_{u_0}=i_0\}} p_{i_0,i_1}(u_0,u_1) \prod_{k=1}^{n-1} p_{i_k,i_{k+1}}(u_k,u_{k+1}) \tag{2.3}$$

for arbitrary $0 \le u_0 \le \ldots \le u_n$ *and* $(i_0,\ldots,i_n) \in \mathcal{K}^{n+1}$.

Proof. The proof is by induction on n. For $n = 1$ the above formula obviously holds. Assume that it holds for n, arbitrary $0 \le u_0 \le \ldots \le u_n$ and $(i_0,\ldots,i_n) \in \mathcal{K}^{n+1}$. We will prove it for $n+1$ and arbitrary $0 \le u_0 \le \ldots \le u_n \le u_{n+1}$, $(i_0,\ldots,i_n,i_{n+1}) \in \mathcal{K}^{n+2}$. Because

$$\mathbf{E}(\mathbf{1}_{\{C_{u_1}=i_1,\ldots C_{u_{n+1}}=i_{n+1}\}} \mid \mathcal{F}_\infty \vee \mathcal{F}^C_{u_0})\mathbf{1}_{\{C_{u_0}=i_0\}}$$
$$= \mathbf{E}\left(\mathbf{E}\left(\mathbf{1}_{\{C_{u_2}=i_2,\ldots,C_{u_{n+1}}=i_{n+1}\}} \mid \mathcal{F}_\infty \vee \mathcal{F}^C_{u_1}\right)\mathbf{1}_{\{C_{u_1}=i_1\}} \mid \mathcal{F}_\infty \vee \mathcal{F}^C_{u_0}\right)\mathbf{1}_{\{C_{u_0}=i_0,\}}$$

by the induction assumption applied to $u_1 \le \ldots \le u_{n+1}$ and $(i_1,\ldots,i_{n+1}) \in \mathcal{K}^{n+1}$ we know that the left hand side of (2.3) is equal to

$$\mathbf{E}\left(\mathbf{1}_{\{C_{u_1}=i_1\}} p_{i_1,i_2}(u_1,u_2) \prod_{k=2}^{n} p_{i_k,i_{k+1}}(u_k,u_{k+1}) \mid \mathcal{F}_\infty \vee \mathcal{F}^C_{u_0}\right)\mathbf{1}_{\{C_{u_0}=i_0\}} = I.$$

Using $\mathcal{F}_\infty$–measurability of family of transition probabilities $(P(s,t))_{0\le s\le t<\infty}$, and the definition of $\mathbb{F}$–DS Markov chain, we obtain

$$I = \mathbf{E}\left(\mathbf{1}_{\{C_{u_1}=i_1\}} \mid \mathcal{F}_\infty \vee \mathcal{F}^C_{u_0}\right)\mathbf{1}_{\{C_{u_0}=i_0\}}\left(p_{i_1,i_2}(u_1,u_2)\prod_{k=2}^{n} p_{i_k,i_{k+1}}(u_k,u_{k+1})\right)$$
$$= \mathbf{1}_{\{C_{u_0}=i_0\}} p_{i_0,i_1}(u_0,u_1)\left(p_{i_1,i_2}(u_1,u_2)\prod_{k=2}^{n} p_{i_k,i_{k+1}}(u_k,u_{k+1})\right)$$
$$= \mathbf{1}_{\{C_{u_0}=i_0\}} p_{i_0,i_1}(u_0,u_1)\prod_{k=1}^{n} p_{i_k,i_{k+1}}(u_k,u_{k+1})$$

and this completes the proof. □

Remark 2.3. Of course, if (2.3) holds, then condition (2) of Definition 2.1 of $\mathbb{F}$–DS Markov chain is satisfied. Therefore (2.3) can be viewed as an alternative to equality (2.1).

As a consequence of our assumption that $C_0 = i_0$ we have

Corollary 2.4. *If C is an $\mathbb{F}$–DS Markov chain, then*

$$\mathbf{P}(C_{u_1} = i_1, \ldots C_{u_n} = i_n \mid \mathcal{F}_\infty) = p_{i_0,i_1}(0,u_1)\prod_{k=0}^{n-1} p_{i_k,i_{k+1}}(u_k,u_{k+1}) \tag{2.4}$$

for arbitrary $0 \le u_1 \le \ldots \le u_n$ *and* $(i_1,\ldots,i_n) \in \mathcal{K}^n$.

The following hypothesis is standard in credit risk theory
HYPOTHESIS H : For every bounded $\mathcal{F}_\infty$–measurable random variable Y and for each $t \ge 0$ we have

$$\mathbf{E}(Y \mid \mathcal{F}_t \vee \mathcal{F}^C_t) = \mathbf{E}(Y \mid \mathcal{F}_t).$$

It is well known that hypothesis H is equivalent to the martingale invariance property of the filtration $\mathbb{F}$ with respect to $\mathbb{F}\vee\mathbb{F}^C$ (see [10, Lemma 6.1.1, page 167]) i.e. any $\mathbb{F}$ martingale is an $\mathbb{F} \vee \mathbb{F}^C$ martingale. We will show that this hypothesis is satisfied for $\mathbb{F}$–DS Markov chains.

Proposition 2.5. *If C is an $\mathbb{F}$–DS Markov chain then hypothesis H holds.*

Proof. According to Lemma 2 from [18] we know that hypothesis H is equivalent to the following condition: for any n and arbitrary $0 \le u_1 \le \ldots \le u_n$, $t \ge u_n$, $(i_1, \ldots, i_n) \in \mathcal{K}^n$,

$$\mathbf{P}(C_{u_1} = i_1, \ldots C_{u_n} = i_n \mid \mathcal{F}_\infty) = \mathbf{P}(C_{u_1} = i_1, \ldots C_{u_n} = i_n \mid \mathcal{F}_t). \tag{2.5}$$

We prove that an $\mathbb{F}$–DS Markov chain satisfies (2.5). The left-hand side of (2.5) is, by Corollary 2.4 and Definition 2.1, $\mathcal{F}_{u_n}$ measurable as a product of $\mathcal{F}_{u_k}$–measurable random variables, $k = 1, \ldots, n$, and therefore equality (2.5) holds for $t \ge u_n$. □

Now, we will show that each $\mathbb{F}$–DS Markov chain is a conditional Markov chain (see Bielecki, Rutkowski [10, page 340] for a precise definition). For an example of a process which is an $\mathbb{F}$ conditional $\mathbb{F} \vee \mathbb{F}^C$ Markov chain and is not a $\mathbb{F}$–DS Markov chain we refer to Section 3 of Becherer and Schweizer [5].

Proposition 2.6. *Assume that C is an $\mathbb{F}$–DS Markov chain. Then C is an $\mathbb{F}$ conditional $\mathbb{F} \vee \mathbb{F}^C$ Markov chain.*

Proof. We have to check that for $s \le t$,

$$\mathbf{P}(C_t = i \mid \mathcal{F}_s \vee \mathcal{F}_s^C) = \mathbf{P}(C_t = i \mid \mathcal{F}_s \vee \sigma(C_s)).$$

By the definition of an $\mathbb{F}$–DS Markov chain,

$$\mathbf{P}(C_t = i \mid \mathcal{F}_s \vee \mathcal{F}_s^C) = \mathbf{E}(\mathbf{E}(\mathbf{1}_{\{C_t = i\}} \mid \mathcal{F}_\infty \vee \mathcal{F}_s^C) \mid \mathcal{F}_s \vee \mathcal{F}_s^C)$$

$$= \mathbf{E}\Big(\sum_{j=1}^K \mathbf{1}_{\{C_s = j\}} p_{j,i}(s,t) \Big| \mathcal{F}_s \vee \mathcal{F}_s^C \Big) = \sum_{j=1}^K \mathbf{1}_{\{C_s = j\}} \mathbf{E}\left(p_{j,i}(s,t) \mid \mathcal{F}_s \vee \mathcal{F}_s^C \right) = I.$$

But

$$I = \sum_{j=1}^K \mathbf{1}_{\{C_s = j\}} \mathbf{E}\left(p_{j,i}(s,t) \mid \mathcal{F}_s \right)$$

since hypothesis H holds (Proposition 2.5), and this ends the proof. □

Now we define processes H^i, which play a crucial role in our characterization of the class of $\mathbb{F}$–DS Markov chains:

$$H_t^i := \mathbf{1}_{\{C_t = i\}} \tag{2.6}$$

for $i \in \mathcal{K}$. The process H_t^i tells us whether at time t the process C is in state i or not. Let $H_t := (H_t^1, \ldots, H_t^K)^\top$, where $\top$ denotes transposition.

We can express condition (2.1) in the definition of an $\mathbb{F}$–DS Markov chain in the form

$$H_t^i \mathbf{E}(H_u^j \mid \mathcal{F}_\infty \vee \mathcal{F}_t^C) = H_t^i p_{i,j}(t,u),$$

or equivalently

$$\mathbf{E}(H_u^j \mid \mathcal{F}_\infty \vee \mathcal{F}_t^C) = \sum_{i \in \mathcal{K}} H_t^i p_{i,j}(t,u)$$

and so (2.1) is equivalent to

$$\mathbf{E}\left(H_u \mid \mathcal{F}_\infty \vee \mathcal{F}_t^C \right) = P(t,u)^\top H_t. \tag{2.7}$$

The next theorem states that the family of matrices $P(s,t) = [p_{i,j}(s,t)]_{i,j=1}^{K}$ satisfies the Chapman-Kolmogorov equations.

Theorem 2.7. *Let C be an $\mathbb{F}$–DS Markov chain with transition matrices $P(s,t)$. Then for any $u \geq t \geq s$ we have*

$$P(s,u) = P(s,t)P(t,u) \quad a.s., \tag{2.8}$$

so on the set $\{C_s = i\}$ we have

$$p_{i,j}(s,u) = \sum_{k=1}^{K} p_{i,k}(s,t)p_{k,j}(t,u).$$

Proof. It is enough to prove that (2.8) holds on each set $\{C_s = i\}$, $i \in \mathcal{K}$. So we have to prove that

$$H_s^\top P(s,u) = H_s^\top P(s,t)P(t,u).$$

By the chain rule for conditional expectation, equality (2.7) and the fact that $P(t,u)$ is $\mathcal{F}_\infty$–measurable it follows that for $s \leq t \leq u$,

$$\begin{aligned} P(s,u)^\top H_s &= \mathbf{E}\left(H_u \mid \mathcal{F}_\infty \vee \mathcal{F}_s^C\right) = \mathbf{E}\left(\mathbf{E}\left(H_u \mid \mathcal{F}_\infty \vee \mathcal{F}_t^C\right) \mid \mathcal{F}_\infty \vee \mathcal{F}_s^C\right) \\ &= \mathbf{E}\left(P(t,u)^\top H_t \mid \mathcal{F}_\infty \vee \mathcal{F}_s^C\right) = P(t,u)^\top \mathbf{E}\left(H_t \mid \mathcal{F}_\infty \vee \mathcal{F}_s^C\right) = P(t,u)^\top P(s,t)^\top H_s \\ &= (P(s,t)P(t,u))^\top H_s, \end{aligned}$$

and this completes the proof. □

Now, for $\mathbb{F}$–DS Markov chains we introduce the concept of intensity, analogous to that for continuous time Markov chains.

Definition 2.8. We say that an $\mathbb{F}$–DS Markov chain C has an intensity if there exists an $\mathbb{F}$–adapted matrix-valued process $\Lambda = (\Lambda(s))_{s\geq 0} = (\lambda_{i,j}(s))_{s\geq 0}$ such that:
1) Λ is locally integrable, i.e. for any $T > 0$

$$\int_{]0,T]} \sum_{i\in\mathcal{K}} |\lambda_{ii}(s)|\, ds < \infty. \tag{2.9}$$

2) Λ satisfies the conditions:

$$\lambda_{i,j}(s) \geq 0 \quad \forall i,j \in \mathcal{K}, i \neq j, \qquad \lambda_{i,i}(s) = -\sum_{j\neq i} \lambda_{i,j}(s) \qquad \forall i \in \mathcal{K}, \tag{2.10}$$

the Kolmogorov backward equation: for all $v \leq t$,

$$P(v,t) - \mathbb{I} = \int_v^t \Lambda(u)P(u,t)du, \tag{2.11}$$

the Kolmogorov forward equation: for all $v \leq t$,

$$P(v,t) - \mathbb{I} = \int_v^t P(v,u)\Lambda(u)du. \tag{2.12}$$

A process Λ satisfying the above conditions is called an intensity of the $\mathbb{F}$–DS Markov chain C.

It is not obvious that if we have a solution to the Kolmogorov backward equation then it also solves the Kolmogorov forward equation. This fact follows from the theory of differential equations, namely we have

Theorem 2.9. *Assume that Λ is locally integrable. Then the random ODE's*

$$dX(t) = -\Lambda(t)X(t)dt, \qquad X(0) = \mathbb{I}, \tag{2.13}$$

$$dY(t) = Y(t)\Lambda(t)dt, \qquad Y(0) = \mathbb{I}, \tag{2.14}$$

have unique solutions, and in addition $X(t) = Y^{-1}(t)$. Moreover, $Z(s,t) := X(s)Y(t)$ is a unique solution to the Kolmogorov forward equation (2.12) *and to the Kolmogorov backward equation* (2.11)*.*

Proof. The existence and uniqueness of solutions of the ODE's (2.13) and (2.14) follows by standard arguments. To deduce that $X(t) = Y^{-1}(t)$ we apply integration by parts to the product $X(t)Y(t)$ of finite variation continuous processes and get

$$d(Y(t)X(t)) = Y(t)dX(t) + (dY(t))X(t) = Y(t)\,(-\Lambda(t)X(t)dt) + Y(t)\Lambda(t)X(t)dt = 0.$$

From $Y(0) = X(0) = \mathbb{I}$ we have $Y(t)X(t) = \mathbb{I}$, which means that $X(t)$ is a right inverse matrix of $Y(t)$. It is also the left inverse, since we are dealing with square matrices.

Now we check that $Z(s,t)$ are solutions to the Kolmogorov backward equation and also the Kolmogorov forward equation. Indeed,

$$d_sZ(s,t) = (dX(s))Y(t) = -\Lambda(s)X(s)Y(t)ds = -\Lambda(s)Z(s,t)ds,$$

and

$$d_tZ(s,t) = X(s)dY(t) = X(s)Y(t)\Lambda(t)dt = Z(s,t)\Lambda(t)dt.$$

This ends the proof since $X(t) = Y^{-1}(t)$ implies that $Z(t,t) = \mathbb{I}$ for every $t \geq 0$. □

Corollary 2.10. *If an $\mathbb{F}$–DS–Markov chain C has intensity, then the conditional transition probability process $P(s,t)$ is jointly continuous at (s,t) for $s \leq t$.*

Proof. This follows immediately from Theorem 2.9, since

$$P(s,t) = X(s)Y(t)$$

and both factors are continuous in s and t respectively. □

Remark 2.11 (Construction of transition probabilities with a given intensity matrix). From Theorem 2.9 it follows that if we are given a matrix process $(\Lambda(s))_{s\geq 0}$ which satisfies conditions (2.10) and moreover is locally integrable, then the integral equation (2.11) has a unique solution. This solution is given by the formula

$$P(v,t) = \mathbb{I} + \sum_{n=1}^{\infty} \int_v^t \int_{v_1}^t \ldots \int_{v_{n-1}}^t \Lambda(v_1)\ldots\Lambda(v_n)dv_n \ldots dv_1, \tag{2.15}$$

and also (2.12) has a unique solution given by

$$P(v,t) = \mathbb{I} + \sum_{n=1}^{\infty} \int_v^t \int_v^{v_1} \ldots \int_v^{v_{n-1}} \Lambda(v_n)\ldots\Lambda(v_1)dv_n \ldots dv_1$$

(cf with Rolski et al. [30, § 8.4.1, page 348]).

Proposition 2.12. *Let $P = (P(s,t))$, $0 \leq s \leq t$, be a family of stochastic matrices such that the matrix $P(s,t)$ is $\mathcal{F}_t$–measurable, and $P(s,\cdot)$ is $\mathbb{F}$–progressively measurable. Let $\Lambda = (\Lambda(s))_{s\geq 0}$ be an $\mathbb{F}$–adapted matrix-valued locally integrable process such that the Kolmogorov backward equation* (2.11) *and Kolmogorov forward equation* (2.12) *hold. Then*
i) For each $s \in [0,t]$ there exists an inverse matrix of $P(s,t)$ denoted by $Q(s,t)$.

ii) There exists a version of $Q(\cdot,t)$ such that the process $Q(\cdot,t)$ is a unique solution to the integral (backward) equation

$$dQ(s,t) = Q(s,t)\Lambda(s)ds, \qquad Q(t,t) = \mathbb{I}. \tag{2.16}$$

This unique solution is given by the following series:

$$Q(s,t) = \mathbb{I} + \sum_{k=1}^{\infty}(-1)^k \int_s^t \int_{u_1}^t \dots \int_{u_{k-1}}^t \Lambda(u_k)\dots\Lambda(u_1)du_k\dots du_1. \tag{2.17}$$

iii) There exists a version of $Q(s,\cdot)$ such that the process $Q(s,\cdot)$ is a unique solution to the integral (forward) equation

$$dQ(s,t) = -\Lambda(t)Q(s,t)dt, \qquad Q(s,s) = \mathbb{I}. \tag{2.18}$$

This unique solution is given by the following series:

$$Q(s,t) = \mathbb{I} + \sum_{k=1}^{\infty}(-1)^k \int_s^t \int_s^{u_1} \dots \int_s^{u_{k-1}} \Lambda(u_1)\dots\Lambda(u_k)du_k\dots du_1. \tag{2.19}$$

Proof. i) From Theorem 2.9 it follows that $P(s,t) = X(s)Y(t)$, where X, Y are solutions to the random ODE's (2.13), (2.14) and moreover $Y = X^{-1}$. Therefore the matrix $P(s,t)$ is invertible and its inverse $Q(s,t)$ is given by $Q(s,t) = X(t)Y(s)$.
ii) We differentiate $Q(s,t)$ with respect to the first argument and obtain

$$d_sQ(s,t) = X(t)dY(s) = X(t)Y(s)\Lambda(s)ds = Q(s,t)\Lambda(s)ds.$$

Moreover $Q(t,t) = X(t)Y(t) = \mathbb{I}$. So $Q(\cdot,t)$ solves (2.16). Uniqueness of solutions to (2.16) follows by standard arguments based on Gronwall's lemma. Formula (2.17) is derived analogously to a similar formula for $P(s,t)$ in § 8.4.1, page 348 of Rolski et al. [30].
iii) The proof of iii) is analogous to that of ii). □

In the next theorem we prove that under some conditions imposed on the conditional transition probability process P, an $\mathbb{F}$–DS Markov chain C has intensity.

Theorem 2.13 (Existence of Intensity)**.** *Let C be an $\mathbb{F}$–DS–Markov chain with conditional transition probability process P. Assume that*

1. *P as a matrix–valued mapping is measurable, i.e.*
$$P : (\mathbb{R}_+ \times \mathbb{R}_+ \times \Omega, \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}_+) \otimes \mathcal{F}) \to (\mathbb{R}^{K\times K}, \mathcal{B}(\mathbb{R}^{K\times K})).$$
2. *There exists a version of P which is continuous in s and in t.*
3. *For every $t \geq 0$ the following limit exists almost surely*
$$\Lambda(t) := \lim_{h\downarrow 0}\frac{P(t,t+h)-\mathbb{I}}{h}, \tag{2.20}$$
and is locally integrable.

Then Λ is the intensity of C.

Proof. By assumption (3) the process Λ is well defined and by (1) it is $(\mathbb{R}_+ \times \Omega, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F})$ measurable. By assumption (3), $\Lambda(t)$ is $\mathcal{F}_{t+}$–measurable, but $\mathbb{F}$ satisfies the usual conditions, so $\Lambda(t)$ is $\mathcal{F}_t$–measurable. It is easy to see that (2.10) holds.
It remains to prove that equations (2.11) and (2.12) are satisfied. Fix t. From

the assumptions and the Chapman–Kolmogorov equations it follows that for $v \leq v+h \leq t$,

$$P(v+h,t)-P(v,t) = P(v+h,t)-P(v,v+h)P(v+h,t) = -(P(v,v+h)-\mathbb{I})P(v+h,t),$$

so

$$\frac{P(v+h,t)-P(v,t)}{h} = -\frac{(P(v,v+h)-\mathbb{I})}{h}P(v+h,t).$$

Therefore $\frac{\partial^+}{\partial v}P(v,t)$ exists for a.e. v and is $(\mathbb{R}_+ \times \mathbb{R}_+ \times \Omega, \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}_+) \otimes \mathcal{F})$ measurable. Using assumption (2) and (3) we finally have

$$\frac{\partial^+}{\partial v}P(v,t) = -\Lambda(v)P(v,t), \qquad P(t,t) = \mathbb{I}. \tag{2.21}$$

Since elements of $P(u,t)$ are bounded by 1, and Λ is integrable over $[v,t]$ (by assumption (3)), we see that $\frac{\partial^+}{\partial u}P(u,t)$ is Lebesgue integrable on $[v,t]$, so (see Walker [32])

$$\mathbb{I} - P(v,t) = \int_v^t \frac{\partial^+}{\partial u}P(u,t)du.$$

Hence, by (2.21), we have

$$P(v,t) - \mathbb{I} = \int_v^t \Lambda(u)P(u,t)du,$$

and this is exactly the Kolmogorov backward equation (2.11).

Similar arguments apply to the case of right derivatives of $P(v,t)$ with respect to the second variable. Since for $v \leq t \leq t+h$,

$$P(v,t+h) - P(v,t) = P(v,t)(P(t,t+h) - \mathbb{I}),$$

we obtain

$$\frac{\partial^+}{\partial t}P(v,t) = P(v,t)\Lambda(t), \qquad P(v,v) = \mathbb{I},$$

which gives (2.12),

$$P(v,t) - \mathbb{I} = \int_v^t P(v,u)\Lambda(u)du.$$

□

Now, we find the intensity for the processes described in Examples 4 and 5.

Example 6. If $C_t = \min\{N_t, K\}$, where N is a Cox process with intensity process $\tilde{\lambda}$, then C has the intensity process of the form

$$\lambda_{i,j}(t) = \begin{cases} -\tilde{\lambda}(t) & \text{for } i = j \in \{0, \dots K-1\}; \\ \tilde{\lambda}(t) & \text{for } j = i+1 \text{ with } i \in \{0, \dots K-1\}; \\ 0 & \text{otherwise.} \end{cases}$$

Example 7. If an $\mathbb{F}$–DS Markov chain C is defined as in Example 4 with a discrete time Markov chain $\bar{C}$ with a transition matrix P, then

$$P(s,t) = e^{(P-I)\int_s^t \lambda(u)du}$$

(see Theorem 9 in [18]), so the intensity of C is given by

$$\lambda_{i,j}(t) = (P-I)_{i,j}\tilde{\lambda}(t).$$

Now we are ready to give a martingale characterization of $\mathbb{F}$–DS Markov chains. To do this, we introduce a filtration $\hat{\mathbb{G}} = (\hat{\mathcal{G}}_t)_{t\geq 0}$, where

$$\hat{\mathcal{G}}_t := \mathcal{F}_\infty \vee \mathcal{F}_t^C. \tag{2.22}$$

Theorem 2.14. *Let $(C_t)_{t\geq 0}$ be a $\mathcal{K}$–valued stochastic process and $(\Lambda(t))_{t\geq 0}$ be a matrix valued process satisfying* (2.9) *and* (2.10)*. The process C is an $\mathbb{F}$–DS Markov chain with intensity process Λ if and only if the processes*

$$M_t^i := H_t^i - \int_{]0,t]} \lambda_{C_u,i}(u)du, \qquad i \in \mathcal{K}, \tag{2.23}$$

are $\hat{\mathbb{G}}$ local martingales.

Proof. Denoting by M the vector valued process with coordinates M^i, we can write M as follows

$$M_t := H_t - \int_{]0,t]} \Lambda^\top(u) H_u du.$$

"$\Rightarrow$" Assume that C is an $\mathbb{F}$–DS Markov chain with intensity process $(\Lambda(t))_{t\geq 0}$. Fix $t \geq 0$ and set

$$N_s := P(s,t)^\top H_s \qquad \text{for } 0 \leq s \leq t. \tag{2.24}$$

The process C satisfies (2.7), which is equivalent to N being a $\hat{\mathbb{G}}$ martingale for $0 \leq s \leq t$. Using integration by parts and the Kolmogorov backward equation (2.11) we find that

$$\begin{aligned} dN_s &= (dP(s,t))^\top H_s + P^\top(s,t)dH_s = -P^\top(s,t)\Lambda^\top(s)H_s ds + P^\top(s,t)dH_s \\ &= P^\top(s,t)dM_s. \end{aligned} \tag{2.25}$$

Hence, using $Q(s,t)$ (the inverse of $P(s,t)$; we know that it exists, see Proposition 2.12), we conclude that

$$M_s - M_0 = \int_{]0,s]} Q^\top(u,t)P^\top(u,t)dM_u = \int_{]0,s]} Q^\top(u,t)dN_u.$$

Therefore, by the $\hat{\mathbb{G}}$ martingale property of N, we conclude that M is a $\hat{\mathbb{G}}$ local martingale.

"$\Leftarrow$" Assume that the process M associated with C and Λ is a $\hat{\mathbb{G}}$ martingale. Fix $t \geq 0$. To prove that C is an $\mathbb{F}$–DS Markov chain it is enough to show that for some process $(P(s,t))_{0\leq s\leq t}$ the process N defined by (2.24) is a $\hat{\mathbb{G}}$ martingale on $[0,t]$. Let $P(s,t)$ be defined by $P(s,t) := X(s)Y(t)$ with X, Y being solutions to the random ODE's (2.13) and (2.14). We know that $P(\cdot,t)$ satisfies the following integral equation (see Theorem 2.9):

$$dP(s,t) = -\Lambda(s)P(s,t)ds, \qquad P(t,t) = \mathbb{I}. \tag{2.26}$$

We also know that $P(s,t)$ is $\mathcal{F}_t$–measurable (Remark 2.11) and continuous in t, hence $\mathbb{F}$ progressively measurable. Using the same arguments as before, we find that (2.25) holds. So, using the martingale property of M we see that N is a local martingale. The definition of N implies that N is bounded (since H and P are bounded, see Last-Brandt [25, §7.4]). Therefore N has an integrable supremum, so it is a $\hat{\mathbb{G}}$ martingale, so C is an $\mathbb{F}$–DS Markov chain with transition matrix P. From Theorem 2.9 it follows that Λ is the intensity matrix process of C. The proof is complete. $\square$

Corollary 2.15. *If C is an $\mathbb{F}$–DS Markov chain, then M^i are $\mathbb{G}$ local martingales with $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{F}_t^C$.*

Proof. This follows from the fact that the M^i are adapted to $\mathbb{G}$, and $\mathbb{G}$ is a subfiltration of $\hat{\mathbb{G}}$. □

Remark 2.16. The process C obtained by the canonical construction in [10] is an $\mathbb{F}$–DS Markov chain. This is a consequence of Theorem 2.14, because Λ in the canonical construction is bounded and calculations analogous to those in [10, Lemma 11.3.2 page 347] show that M^i are $\hat{\mathbb{G}}$ martingales.

We can also exhibit another set of local martingales which characterize an $\mathbb{F}$–DS Markov chain with intensity process $(\Lambda(s))_{s\geq 0}$. This will be an immediate consequence of the fact which we formulate as a lemma in a slightly more general form, namely for an arbitrary filtration.

Lemma 2.17. *Let $\mathbb{A}$ be some filtration. The processes M^i, $i \in \mathcal{K}$, are $\mathbb{A}$ local martingales if and only if for all $i, j \in \mathcal{K}$, $i \neq j$, the processes $M^{i,j}$ defined by*

$$M_t^{i,j} := H_t^{i,j} - \int_{]0,t]} H_s^i \lambda_{i,j}(s) ds, \tag{2.27}$$

where

$$H_t^{i,j} := \int_{]0,t]} H_{u-}^i dH_u^j, \tag{2.28}$$

are $\mathbb{A}$ local martingales.

Proof. $\Rightarrow$ Fix $i \neq j$, $i, j \in \mathcal{K}$. Using the definition of $M_t^{i,j}$ and M^i we have

$$\begin{aligned} M_t^{i,j} &= \int_{]0,t]} H_{u-}^i dH_u^j - \int_{]0,t]} H_u^i \lambda_{i,j}(u) du = \int_{]0,t]} H_{u-}^i dH_u^j - \int_{]0,t]} H_u^i \lambda_{C_u,j}(u) du \\ &= \int_{]0,t]} H_{u-}^i dH_u^j - \int_{]0,t]} H_{u-}^i \lambda_{C_u,j}(u) du = \int_{]0,t]} H_{u-}^i dM_u^j. \end{aligned}$$

Hence $M_t^{i,j}$ is an $\mathbb{A}$ local martingale, since M^j is one and H_{u-}^i is bounded.
$\Leftarrow$ Assume that the $M_t^{i,j}$ are $\mathbb{A}$ martingales for all $i \neq j$, $i, j \in \mathcal{K}$. First notice that H^i can be obtained from $H^{j,i}$ by the formula

$$H_t^i = H_0^i + \sum_{j \neq i} \left(H_t^{j,i} - H_t^{i,j} \right).$$

Indeed, from (2.28) it follows that

$$\begin{aligned} \sum_{j \neq i} (H_t^{j,i} - H_t^{i,j}) &= \int_{]0,t]} \left(\sum_{j \neq i} H_{u-}^j \right) dH_u^i + \int_{]0,t]} H_{u-}^i d \left(- \sum_{j \neq i} H_u^j \right) \\ &= \int_{]0,t]} (1 - H_{u-}^i) dH_u^i + \int_{]0,t]} H_{u-}^i dH_u^i = H_t^i - H_0^i. \end{aligned}$$

Next by (2.27)

$$\begin{aligned} H_t^i &= H_0^i + \sum_{j\neq i}(M_t^{j,i} - M_t^{i,j}) + \sum_{j\neq i}\left(\int_{]0,t]} H_s^j \lambda_{j,i}(s) - H_s^i \lambda_{i,j}(s)ds\right) \\ &= H_0^i + \sum_{j\neq i}\left(M_t^{j,i} - M_t^{i,j}\right) + \int_{]0,t]} \sum_{j=1}^{K} H_s^j \lambda_{j,i}(s)ds \\ &= H_0^i + \sum_{j\neq i}\left(M_t^{j,i} - M_t^{i,j}\right) + \int_{]0,t]} \lambda_{C_s,i}(s)ds, \end{aligned}$$

which implies that

$$M_t^i = H_t^i - \int_{]0,t]} \lambda_{C_s,i}(s)ds = H_0^i + \sum_{j\neq i}\left(M_t^{j,i} - M_t^{i,j}\right)$$

and therefore M^i is an $\mathbb{A}$ local martingale for each $i \in \mathcal{K}$ as a finite sum of $\mathbb{A}$ local martingales. □

The process $H^{i,j}$ defined by (2.28) counts the number of jumps from state i to j over the time interval $(0,t]$. One can show that

$$H^{i,j} = \sum_{0<u\leq t} H_{u-}^i H_u^j.$$

Using Lemma 2.17 with $\mathbb{A} = \hat{\mathbb{G}}$ given by (2.22), we obtain

Theorem 2.18. *Let $(C_t)_{t\geq 0}$ be a $\mathcal{K}$-valued stochastic process and $(\Lambda(t))_{t\geq 0}$ be a matrix-valued process satisfying (2.9) and (2.10). The process C is an $\mathbb{F}$–DS Markov chain with intensity process Λ if and only if the processes*

$$M_t^{i,j} := H_t^{i,j} - \int_{]0,t]} H_s^i \lambda_{i,j}(s)ds \tag{2.29}$$

$i \neq j,\ i,j \in \mathcal{K}$, are $\hat{\mathbb{G}}$ local martingales.

To end this subsection, we construct an $\mathbb{F}$–DS Markov chain with intensity given by an arbitrary $\mathbb{F}$ adapted matrix-valued locally bounded stochastic process which satisfies condition (2.10).

Theorem 2.19. *Let $(\Lambda(t))_{t\geq 0}$ be an arbitrary $\mathbb{F}$ adapted matrix-valued stochastic process which satisfies conditions (2.9) and (2.10). Then there exists an $\mathbb{F}$–DS Markov chain with intensity $(\Lambda(t))_{t\geq 0}$.*

Proof. We assume that on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$ with a filtration $\mathbb{F}$ we have a family of Cox processes $N^{i,j}$ for $i,j \in \mathcal{K}$ with intensities $(\lambda_{i,j}(t))$ such that the $N^{i,j}$ are conditionally independent given $\mathcal{F}_\infty$ (otherwise we enlarge the probability space). We construct on $(\Omega, \mathcal{F}, \mathcal{P})$ an $\mathbb{F}$–DS Markov chain C with intensity $(\Lambda(t))_{t\geq 0}$ and given initial state i_0. It is a pathwise construction. First, we define a sequence $(\tau_n)_n$ of jump times of C and a sequence $(\bar{C}_n)_n$ which describes the states of rating after change. We define these sequences by induction. We put

$$\bar{C}_0 = i_0, \qquad \tau_1 := \min_{j\in\mathcal{K}\setminus\bar{C}_0} \inf\left\{t > 0 : \Delta N_t^{\bar{C}_0,j} > 0\right\}$$

and $\bar{C}_1 := j$, where j is the element of $\mathcal{K} \setminus \bar{C}_0$ for which the above minimum is attained. By conditional independence of $N^{i,j}$ given $\mathcal{F}_\infty$, the processes $N^{i,j}$ have no common jumps, so $\bar{C}_1$ is uniquely determined. We now assume that $\tau_1, \ldots, \tau_k$, $\bar{C}_1, \ldots, \bar{C}_k$ are defined and we construct τ_{k+1} as the first jump time of the Cox processes after τ_k, i.e.

$$\tau_{k+1} := \min_{j \in \mathcal{K} \setminus \bar{C}_k} \inf \left\{ t > \tau_k : \Delta N_t^{\bar{C}_k, j} > 0 \right\},$$

and we put $\bar{C}_{k+1} := j$, where j is the element of $\mathcal{K} \setminus \bar{C}_k$ for which the above minimum is attained. Arguing as before, we see that τ_{k+1} and $\bar{C}_{k+1}$ are well defined.

Having the sequences $(\tau_n)_n$ and $(\bar{C}_n)_n$ we define a process C by the formula

$$C_t := \sum_{k=0}^{\infty} \bar{C}_k 1_{[\tau_k, \tau_{k+1})}(t). \tag{2.30}$$

This process C is càdlàg and adapted to the filtration $\mathcal{A} = (\mathcal{A}_t)_t$, where $\mathcal{A}_t := \mathcal{F}_t \vee \left(\bigvee_{i \neq j} \mathcal{F}_t^{N^{i,j}} \right)$, and hence it is also adapted to the larger filtration $\tilde{\mathbb{A}} = (\tilde{\mathcal{A}}_t)_t$, $\tilde{\mathcal{A}}_t := \mathcal{F}_\infty \vee \left(\bigvee_{i \neq j} \mathcal{F}_t^{N^{i,j}} \right)$. Notice that

$$H_t^{i,j} = \int_{]0,t]} \mathbf{1}_{\{i\}}(C_{s-}) dN_s^{i,j} \qquad a.s.$$

The processes $N_t^{i,j} - \int_{]0,t]} \lambda_{i,j}(s) ds$ are $\tilde{\mathbb{A}}$ martingales (since they are compensated Cox's processes, see e.g. [10]). Likewise, each $M^{i,j}$ is an $\tilde{\mathbb{A}}$ martingale, since

$$\begin{aligned} M_t^{i,j} &= H_t^{i,j} - \int_{]0,t]} H_s^i \lambda_{i,j}(s) ds = H_t^{i,j} - \int_{]0,t]} H_{s-}^i \lambda_{i,j}(s) ds \\ &= \int_{]0,t]} \mathbf{1}_{\{i\}}(C_{s-}) d(N_s^{i,j} - \lambda_{i,j}(s) ds). \end{aligned}$$

Hence, by Lemma 2.17 with $\tilde{\mathbb{A}} = (\tilde{\mathcal{A}}_t)_{t \geq 0}$ and $\lambda_{i,i}(t) := -\sum_{j \neq i} \lambda_{i,j}(t)$, we see that

$$M_t^i := H_t^i - \int_{]0,t]} \lambda_{C_s, i}(s) ds$$

is an $\tilde{\mathbb{A}}$ martingale. Recall that $\hat{\mathcal{G}}_t = \mathcal{F}_\infty \vee \mathcal{F}_t^C$, so $\hat{\mathbb{G}} \subseteq \tilde{\mathbb{A}}$. Therefore each M^i is also a $\hat{\mathbb{G}}$ martingale, since M^i is $\hat{\mathbb{G}}$ adapted. Hence, using Theorem 2.14, we see that C is an $\mathbb{F}$–DS Markov chain. $\square$

2.3. **Change of probability and doubly stochastic property.** Now, we investigate how changing the probability measure to an equivalent one affects the properties of an $\mathbb{F}$–DS Markov chain. We start from a lemma

Lemma 2.20. *Let* $\mathbf{Q}$, $\mathbf{P}$ *be equivalent probability measures with density factorizing as*

$$\frac{d\mathbf{Q}}{d\mathbf{P}}\Big|_{\mathcal{F}_\infty \vee \mathcal{F}_{T^*}^C} := \eta_1 \eta_2, \tag{2.31}$$

where η_1 *is an* $\mathcal{F}_\infty$*–measurable strictly positive random variable and* η_2 *is an* $\mathcal{F}_\infty \vee \mathcal{F}_{T^*}^C$*–measurable strictly positive random variable integrable under* $\mathbf{P}$*. Let* $(\eta_2(t))_{t \in [0,T^*]}$ *be defined by the formula*

$$\eta_2(t) := \mathbf{E}_{\mathbf{P}}(\eta_2 \mid \mathcal{F}_\infty \vee \mathcal{F}_t^C), \quad \eta_2(0) = 1. \tag{2.32}$$

Then $(N(t))_{t \in [0,T^*]}$ *is a* $\hat{\mathbb{G}}$ *martingale (resp. local martingale) under* $\mathbf{Q}$ *if and only if* $(N(t)\eta_2(t))_{t \in [0,T^*]}$ *is a* $\hat{\mathbb{G}}$ *martingale (resp. local martingale) under* $\mathbf{P}$*.*

Proof. $\Rightarrow$ By the abstract Bayes rule and the fact that η_1 is $\mathcal{F}_\infty$ measurable and hence also $\mathcal{G}_u$ measurable for all $u \geq 0$, we obtain, for $s < t$,

$$N(s) = \mathbf{E_Q}\left(N(t) \mid \hat{\mathcal{G}}_s\right) = \frac{\mathbf{E_P}\left(N(t)\eta_1\eta_2 \mid \hat{\mathcal{G}}_s\right)}{\mathbf{E_P}\left(\eta_1\eta_2 \mid \hat{\mathcal{G}}_s\right)} = \mathbf{E_P}\left(N(t)\frac{\mathbf{E_P}\left(\eta_1\eta_2 \mid \hat{\mathcal{G}}_t\right)}{\mathbf{E_P}\left(\eta_1\eta_2 \mid \hat{\mathcal{G}}_s\right)} \mid \hat{\mathcal{G}}_s\right)$$

$$= \mathbf{E_P}\left(N(t)\frac{\mathbf{E_P}\left(\eta_2 \mid \hat{\mathcal{G}}_t\right)}{\mathbf{E_P}\left(\eta_2 \mid \hat{\mathcal{G}}_s\right)} \mid \hat{\mathcal{G}}_s\right) = \mathbf{E_P}\left(N(t)\frac{\eta_2(t)}{\eta_2(s)} \mid \hat{\mathcal{G}}_s\right) = \frac{\mathbf{E_P}\left(N(t)\eta_2(t) \mid \hat{\mathcal{G}}_s\right)}{\eta_2(s)}.$$

$\Leftarrow$ The proof is similar. □

Lemma 2.21. *Let C be an $\mathbb{F}$–DS Markov chain under $\mathbf{P}$ with intensity $(\lambda_{i,j})$ and suppose that η_2 defined by* (2.32) *satisfies*

$$d\eta_2(t) = \eta_2(t-)\left(\sum_{k,l\in\mathcal{K}:k\neq l}\kappa_{k,l}(u)dM_u^{k,l}\right) \tag{2.33}$$

with some $\mathbb{G}$ predictable stochastic processes $\kappa_{i,j}$, $i,j \in \mathcal{K}$, such that $\kappa_{i,j} > -1$. Then

$$\tilde{M}_t^{i,j} = H_t^{i,j} - \int_{]0,t]} H_u^i(1+\kappa_{i,j}(u))\lambda_{i,j}(u)du, \tag{2.34}$$

$i,j \in \mathcal{K}$, is a $\hat{\mathbb{G}}$ local martingale under $\mathbf{Q}$ defined by (2.31).

Proof. By Lemma 2.20 it is enough to prove that $\tilde{M}^{i,j}\eta_2$ is a $\hat{\mathbb{G}}$ local martingale under $\mathbf{P}$. Integration by parts yields

$$d(\tilde{M}_t^{i,j}\eta_2(t)) = \tilde{M}_{t-}^{i,j}d\eta_2(t) + \eta_2(t-)d\tilde{M}_t^{i,j} + \Delta\tilde{M}_t^{i,j}\Delta\eta_2(t) =: I.$$

Since

$$\tilde{M}_t^{i,j} = M_t^{i,j} - \int_{]0,t]} H_u^i\kappa_{i,j}(u)\lambda_{i,j}(u)du,$$

we have

$$d\tilde{M}_t^{i,j} = dM_t^{i,j} - H_t^i\kappa_{i,j}(t)\lambda_{i,j}(t)dt$$

and

$$\Delta\tilde{M}_t^{i,j}\Delta\eta_2(t) = \Delta M_t^{i,j}\eta_2(t-)\left(\sum_{k,l\in\mathcal{K}:k\neq l}\kappa_{k,l}(t)\Delta M_t^{k,l}\right) = \eta_2(t-)\kappa_{i,j}(t)\left(\Delta M_t^{i,j}\right)^2$$

$$= \eta_2(t-)\kappa_{i,j}(t)\left(\Delta H_t^{i,j}\right)^2 = \eta_2(t-)\kappa_{i,j}(t)\Delta H_t^{i,j}.$$

Hence

$$\begin{aligned} I &= \tilde{M}_{t-}^{i,j}\eta_2(t-)\left(\sum_{k,l\in\mathcal{K}:k\neq l}\kappa_{k,l}(t)dM_t^{k,l}\right) + \eta_2(t-)dM_t^{i,j} \\ &\quad + \eta_2(t-)\kappa_{i,j}(t)(\Delta H_t^{i,j} - H_t^i\lambda_{i,j}(t)dt) \\ &= \tilde{M}_{t-}^{i,j}\eta_2(t-)\left(\sum_{k,l\in\mathcal{K}:k\neq l}\kappa_{k,l}(t)dM_t^{k,l}\right) + \eta_2(t-)(1+\kappa_{i,j}(t))dM_t^{i,j}, \end{aligned}$$

which completes the proof. □

Hence and from Theorem 2.18 we deduce that the doubly stochastic property is preserved by a wide class of equivalent changes of probability measures.

Theorem 2.22. *Let C be an $\mathbb{F}$–DS Markov chain under $\mathbf{P}$ with intensity $(\lambda_{i,j})$, and $\mathbf{Q}$ be an equivalent probability measure with density given by* (2.31) *and η_2 satisfying* (2.33) *with an $\mathbb{F}$ predictable matrix-valued process κ. Then C is an $\mathbb{F}$–DS Markov chain under $\mathbf{Q}$ with intensity $((1+\kappa_{i,j})\lambda_{i,j})$.*

Now, we exhibit a broad class of equivalent probability measures such that the factorization (2.31) in Lemma 2.20 holds.

Example 8. Let $\mathbb{F} = \mathbb{F}^W$ be the filtration generated by some Brownian motion W under $\mathbf{P}$, and let C be an $\mathbb{F}$–DS Markov chain with intensity matrix process Λ. Let $\mathbf{Q}$ be a probability measure equivalent to $\mathbf{P}$ with Radon-Nikodym density process given as a solution to the SDE

$$d\eta_t = \eta_{t-}\left(\gamma_t dW_t + \sum_{k,l\in\mathcal{K}:k\neq l} \kappa_{k,l}(u) dM_u^{k,l}\right), \qquad \eta_0 = 1,$$

with $\mathbb{F}$ predictable stochastic processes γ and κ. It is easy to see that this density can be written as a product of the following two Doleans-Dade exponentials:

$$d\eta_1(t) = \eta_1(t-)\gamma_t dW_t, \qquad \eta_1(0) = 1;$$

and

$$d\eta_2(t) = \eta_2(t-)\left(\sum_{k,l\in\mathcal{K}:k\neq l} \kappa_{k,l}(u) dM_u^{k,l}\right), \qquad \eta_2(0) = 1.$$

Therefore a factorization

$$\eta(t) = \eta_1(t)\eta_2(t)$$

as in Lemma 2.20 holds, since η_1 is $\mathcal{F}_\infty$ measurable. As an immediate consequence we find that C is an $\mathbb{F}$–DS Markov chain under $\mathbf{Q}$ with intensity $[\lambda^{\mathbf{Q}}]_{i,j} = ((1+\kappa_{i,j})\lambda_{i,j})$ and moreover the process defined by $W_t^* := W_t - \int_0^t \gamma_u du$ is a Brownian motion under $\mathbf{Q}$.

3. Valuation of defaultable rating-sensitive claims with ratings given by a doubly stochastic Markov chain

3.1. **Description of claims.** We consider an arbitrage-free market with finite horizon on which defaultable instruments are also traded. We denote by $\mathbb{F}$ the reference filtration corresponding to observation of the market without credit rating, i.e. a filtration corresponding to the interest rate risk and other market factors that drive the credit risk. C is a credit rating process which takes values in the set of rating classes $\mathcal{K} = \{1,\dots,K\}$. If $K = 2$, then it is understood that there are only two states: default and non-default. We assume that the process C is càdlàg. Let $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{F}_t^C$. By a defaultable rating-sensitive claim we mean a classical one broadened by a migration process.

Definition 3.1. By a defaultable rating-sensitive claim maturing at T we mean a quintuple (X, A, Z, C, τ), where X is a $K-1$ dimensional vector of $\mathcal{F}_T$ measurable random variables, A is a $K-1$ dimensional vector valued $\mathbb{F}$-progressively measurable stochastic process of finite variation, Z is an $\mathbb{F}$-predictable $K \times K$ dimensional matrix-valued process with zero on the diagonal, C is a càdlàg process with values in $\mathcal{K}$, and τ is a positive random variable, defined by

$$\tau := \inf\{t \geq 0 : C_t = K\}.$$

In this definition X describes the promised payoff which is contingent on rating at maturity T, i.e. the payoff is equal to X_i provided that $\{C_T = i\}$; A models the process of promised dividends which can depend on current credit rating; the processes $Z^{i,j}$ describe the payments at times when the rating changes, in particular, $Z^{j,K}$ specifies the recovery payment at the default time τ provided that before the default time we are in state j; and C is the credit rating process. This definition of claim is very general and covers many different type of claims.

Remark 3.2. If we put $X_i = X$ for each i, then the promised payment depends only on the default time:

$$\sum_{i=1}^{K-1} X_i \mathbf{1}_{\{C_T=i\}} = X \sum_{i=1}^{K-1} \mathbf{1}_{\{C_T=i\}} = X\mathbf{1}_{\{C_T \neq K\}} = X\mathbf{1}_{\{\tau>T\}}.$$

Remark 3.3. Since

$$\sum_{i=1}^{K-1} \int_{]0,t\wedge T]} Z_u^{i,K} dH_u^{i,K} = \sum_{i=1}^{K-1} Z_\tau^{i,K} \mathbf{1}_{\{0<\tau\leq t\wedge T, C_{\tau-}=i\}} = Z_\tau^{C_{\tau-},K} \mathbf{1}_{\{0<\tau\leq t\wedge T\}}$$

the recovery process allows recovery depending on the rating of the bond before the default time τ.

Now, we define the dividend process which describes the cash flows from the claim in the interval $[0, T]$.

Definition 3.4. The dividend process $D = (D_t)_{t\geq 0}$ of the claim (X, A, Z, C, τ) maturing at T equals for $t \geq 0$

$$D_t = \sum_{i=1}^{K-1} \left(X_i H_T^i \mathbf{1}_{[T,+\infty[}(t) + \int_{]0,t\wedge T]} H_u^i dA_u^i + \sum_{j\neq i\in\mathcal{K}} \int_{]0,t\wedge T]} Z_u^{i,j} dH_u^{i,j} \right). \tag{3.1}$$

Remark 3.5. For fixed i, if at time t the rating process changes from state i to state j then the promised dividend $A_t^i - A_{t-}^i$ is not passed over to the holder of the claim, and if the rating process changes from some j to i then the promised dividend $A_t^i - A_{t-}^i$ is passed over to the holder of the claim.

Example 9. Consider a defaultable bond with fractional recovery of par value. In this case the bond's holder receives at maturity time T its face value (say 1 unit of cash) provided that default didn't occur before or at T. If the default occurred before or at time T the recovery $\delta_{C_{\tau-}}$ is paid at the default time τ to the bond holder. So, the recovery payment depends on the pre-default rating $C_{\tau-}$, and it is assumed that the recovery $\delta_i \in [0, 1)$ is a fixed number for each $i \in \mathcal{K} \setminus K$. By taking

$$X_i = 1, \quad A^i = 0, \quad Z^{i,K} = \delta_i \text{ for } i = 1, \ldots, K-1, \quad Z^{i,j} = 0 \text{ for } j \neq K,$$

we see that a defaultable bond is a claim in the sense of our definition. The dividend process for such a claim equals

$$D_t = \mathbf{1}_{\{\tau>T\}} \mathbf{1}_{[T,+\infty[}(t) + \delta_{C_{\tau-}} \mathbf{1}_{\{0<\tau\leq t\wedge T\}}.$$

Example 10. Another example is a defaultable credit-sensitive note. It is a corporate bond with coupons linked to the rating of corporation. The coupons of this note are paid at pre-specified coupon dates $0 < T_1 < T_2 < \ldots < T_n$ if default does not arise, and the value of the coupon is contingent on rating corporate at the coupon date. If a default occurred before or at time T, the recovery $\delta_{C_{\tau-}}$ is paid at the default time τ to the bond holders. It is assumed that $\delta_i \in [0,1)$ is fixed for each $i \in \mathcal{K} \setminus K$. So

$$X_i = 1, \quad Z^{i,K} = \delta_i \text{ for } i \leq K-1, \quad Z^{i,j} = 0 \text{ for } j \neq K, \quad A_t^i = \sum_{j=1}^{n} \mathbf{1}_{\{t \geq T_j\}} d_{i,j},$$

where $d_{i,j}$ are fixed constants chosen in advance, and the dividend process of this note is given by

$$D_t = \mathbf{1}_{\{\tau > T\}} \mathbf{1}_{[T,+\infty[}(t) + \sum_{i=1}^{K-1} \int_{]0,t\wedge T]} \delta_i dH_u^{i,K} + \sum_{i=1}^{K-1} \int_{]0,t\wedge T]} H_u^i dA_u^i.$$

Example 11. We can consider modification of the example above with $A_t^i = d_i t$, which corresponds to continuous payments at the rate d_i provided that at time t the rating is equal to i. The dividend process of this note is given by

$$D_t = \mathbf{1}_{\{\tau > T\}} \mathbf{1}_{[T,+\infty[}(t) + \sum_{i=1}^{K-1} \int_{]0,t\wedge T]} \delta_i dH_u^{i,K} + \sum_{i=1}^{K-1} \int_{]0,t\wedge T]} H_u^i d_i du.$$

3.2. **Pricing of rating-sensitive claims.** Now, we consider the problem of pricing of rating-sensitive claims. We put ourselves in an arbitrage free framework, which means that we assume existence of a spot martingale measure $\mathbf{P}$ for the underlying market. As usual, the spot martingale measure $\mathbf{P}$ is a measure related to the choice of the saving account B as a numéraire. Then the price process discounted by B of any tradable security (non-dividend paying) is a martingale under $\mathbf{P}$. We assume that the saving account B is given by

$$dB_t = r_t B_t dt, \quad B_0 = 1,$$

where r is an $\mathbb{F}$-progressively measurable stochastic process. So B^{-1} is a discount factor. In what follows, we assume that the process C of rating migration is an $\mathbb{F}$–DS Markov chain under $\mathbf{P}$ with intensity process Λ which satisfies the following integrability condition:

$$\mathbf{E}\left(\int_{]0,T]} \sum_{i \in \mathcal{K}} |\lambda_{i,i}(s)|\, ds \right) < \infty. \tag{3.2}$$

Rating-sensitive claims we are going to consider are dividend paying securities, and for such securities it is very common to define their price at time $t \in [0,T]$ as the conditional expectation of integral over the time interval $]t,T]$ of the discount factor process with respect to the dividend flow process D, see e.g. Duffie [13] or Bielecki et al. [7]. The natural idea that calculating the value at time t we take only discounted future cashflows (from the time interval $]t,T]$) goes back to Lucas [26] (see also a recent paper of Aase [1]) and leads us to the following definition of ex-dividend price of the claim.

Definition 3.6. The ex-dividend price process S of a defaultable rating-sensitive claim (X, A, F, Z, C, τ) is given by

$$S_t = B_t \mathbf{E}\Big(\int_{]t,T]} B_u^{-1} dD_u \Big| \mathcal{G}_t \Big)$$

for every $t \in [0, T]$.

The main theorem of this subsection gives a convenient form of the ex-dividend price process S of a defaultable rating-sensitive claim. It generalizes the results of Bielecki et al. [7] obtained for $K = 2$.

Theorem 3.7. *Let (X, A, Z, τ, C) be a defaultable rating-sensitive claim. Under the condition* (3.2)*, the ex-dividend price process is given by the formula:*

$$S_t \mathbf{1}_{\{C_t = i\}} = \mathbf{1}_{\{C_t = i\}} \sum_{j=1}^{K-1} B_t \mathbf{E}\Big(\frac{X_j p_{i,j}(t,T)}{B_T} + \int_{]t,T]} B_u^{-1} p_{i,j}(t,u) dA_u^j + \int_{]t,T]} \sum_{k=1}^{K} \frac{Z_u^{j,k}}{B_u} p_{i,j}(t,u) \lambda_{j,k}(u) du \Big| \mathcal{F}_t \Big). \tag{3.3}$$

Proof. The theorem follows immediately from (3.1) and the following three lemmas. □

Lemma 3.8. *Let X be a bounded $\mathcal{F}_T$–measurable random variable, and $j \in \mathcal{K} \setminus K$. Then*

$$\mathbf{E}\left(X \mathbf{1}_{\{C_T = j\}} \mid \mathcal{G}_t \right) = \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t = i\}} \mathbf{E}\left(X p_{i,j}(t,T) \mid \mathcal{F}_t \right).$$

Proof. This is an easy consequence of the definition of $\mathbb{F}$–DS Markov chain and Proposition 2.5. Indeed,

$$\begin{aligned} &\mathbf{E}\left(X \mathbf{1}_{\{C_T = j\}} \mid \mathcal{G}_t \right) = \mathbf{E}\left(X \mathbf{E}\left(\mathbf{1}_{\{C_T = j\}} \mid \mathcal{F}_\infty \vee \mathcal{F}_t^C \right) \mid \mathcal{G}_t \right) \\ &= \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t = i\}} \mathbf{E}\left(X p_{i,j}(t,T) \mid \mathcal{G}_t \right) = \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t = i\}} \mathbf{E}\left(X p_{i,j}(t,T) \mid \mathcal{F}_t \right). \end{aligned}$$

□

Lemma 3.9. *Let Z be a bounded $\mathbb{F}$ predictable stochastic process, and $j \in \mathcal{K} \setminus K$. Under the condition* (3.2)*, for $k \neq j$ we have*

$$\mathbf{E}\left(\int_{]t,T]} Z_u dH_u^{j,k} \mid \mathcal{G}_t \right) = \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t = 1\}} \mathbf{E}\left(\int_{]t,T]} Z_u p_{i,j}(t,u) \lambda_{j,k}(u) du \Big| \mathcal{F}_t \right). \tag{3.4}$$

Proof. Fix $k \neq j \in \mathcal{K}$. Because

$$\int_{]t,T]} Z_u dH_u^{j,k} = \int_{]t,T]} Z_u dM_u^{j,k} + \int_{]t,T]} Z_u H_u^j \lambda_{j,k}(u) du,$$

and $M^{j,k}$ is a martingale and Z a bounded process, we have

$$\mathbf{E}\left(\int_{]t,T]} Z_u dH_u^{j,k} \mid \mathcal{G}_t \right) = \mathbf{E}\left(\int_{]t,T]} Z_u H_u^j \lambda_{j,k}(u) du \mid \mathcal{G}_t \right) = I.$$

Using the conditional version of Fubini's theorem (see e.g. Applebaum [3] p. 12), the definition of an $\mathbb{F}$–DS Markov chain and hypothesis H we have

$$
\begin{aligned}
I &= \int_{]t,T]} \mathbf{E}\left(Z_u H^j_u \lambda_{j,k}(u) \mid \mathcal{G}_t\right) du = \int_{]t,T]} \mathbf{E}\left(\mathbf{E}\left(Z_u H^j_u \lambda_{j,k}(u) \mid \mathcal{F}_\infty \vee \mathcal{F}^C_t\right) \mid \mathcal{G}_t\right) du \\
&= \int_{]t,T]} \mathbf{E}\left(Z_u \lambda_{j,k}(u) \left(\sum_{i=1}^{K-1} \mathbf{1}_{\{C_t=i\}} p_{i,j}(t,u)\right) \mid \mathcal{G}_t\right) du \\
&= \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t=i\}} \int_{]t,T]} \mathbf{E}\left(Z_u p_{i,j}(t,u)\lambda_{j,k}(u) \mid \mathcal{F}_t\right) du \\
&= \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t=i\}} \mathbf{E}\left(\int_{]t,T]} Z_u p_{i,j}(t,u)\lambda_{j,k}(u) du \mid \mathcal{F}_t\right),
\end{aligned}
$$

and this completes the proof. □

Remark 3.10. Notice that for any pair j, K we have

$$
\int_{]t,T]} Z_u dH^{j,K}_u = Z_\tau \mathbf{1}_{\{t<\tau\leq T, C_{\tau-}=j\}},
$$

which is the recovery at the default time depending on a pre-default rating state, and (3.2) implies

(3.5)

$$
\mathbf{E}\left(Z_\tau \mathbf{1}_{\{t<\tau\leq T, C_{\tau-}=j\}} \mid \mathcal{G}_t\right) = \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t=1\}} \mathbf{E}\Big(\int_{]t,T]} Z_u p_{i,j}(t,u)\lambda_{j,K}(u) du \Big| \ \mathcal{F}_t\Big).
$$

Lemma 3.11. *Let A be an $\mathbb{F}$–adapted bounded stochastic process of finite variation. Then for any $j \in \mathcal{K} \setminus K$ we have*

$$
\mathbf{E}\Big(\int_{]t,v]} H^j_u dA_u \Big| \ \mathcal{G}_t\Big) = \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t=i\}} \mathbf{E}\Big(\int_{]t,v]} p_{i,j}(t,u) dA_u \Big| \ \mathcal{F}_t\Big).
$$

Proof. We follow the idea from Bielecki and Rutkowski's book [10] in which the case with two states (default and no default) is considered. Fix t. Define $\tilde{A}_u := A_u - A_t$ for $u \in [t,v]$. Obviously, this is an $\mathbb{F}$–predictable bounded process of finite variation and $\tilde{A}_t = 0$. Integrals with respect to A and $\tilde{A}$ are equal and therefore

$$
\begin{aligned}
\mathbf{E}\Big(\int_{]t,v]} H^j_u dA_u \Big| \ \mathcal{G}_t\Big) &= \mathbf{E}\Big(\int_{]t,v]} H^j_u d\tilde{A}_u \Big| \ \mathcal{G}_t\Big) = \mathbf{E}\Big(\tilde{A}_v H^j_v - \tilde{A}_t H^j_t - \int_{]t,v]} \tilde{A}_{u-} dH^j_u \Big| \ \mathcal{G}_t\Big) \\
&= \mathbf{E}\Big(\tilde{A}_v H^j_v - \int_{]t,v]} \tilde{A}_{u-} dH^j_u \Big| \ \mathcal{G}_t\Big) = I_1 + I_2,
\end{aligned}
$$

where

$$
I_1 = \mathbf{E}\Big(\tilde{A}_v H^j_v \mid \mathcal{G}_t\Big), \qquad I_2 = \mathbf{E}\Big(\int_{]t,v]} \tilde{A}_{u-} dH^j_u \Big| \ \mathcal{G}_t\Big).
$$

Since $\tilde{A}_v$ is $\mathcal{F}_\infty$ measurable, it follows that

$$
I_1 = \mathbf{E}\Big(\tilde{A}_v \mathbf{E}\Big(H^j_v \mid \mathcal{F}_\infty \vee \mathcal{F}^C_t\Big) \mid \mathcal{G}_t\Big) = \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t=i\}} \mathbf{E}\Big(\tilde{A}_v p_{i,j}(t,v) \mid \mathcal{F}_t\Big).
$$

Now we calculate I_2. By boundedness of $\tilde{A}$ and λ, using martingale property of M^j, the conditional Fubini theorem, hypothesis H and the Kolmogorov forward equation (2.12) we have

$$\begin{aligned}
I_2 &= \mathbf{E}\Big(\int_{]t,v]} \tilde{A}_{u-} dM_u^j + \int_{]t,v]} \tilde{A}_{u-}\lambda_{C_u,j}(u)du\Big|\ \mathcal{G}_t\Big) \\
&= \mathbf{E}\Big(\int_{]t,v]} \tilde{A}_{u-} \sum_{k=1}^{K-1} H_u^k \lambda_{k,j}(u)du\Big|\ \mathcal{G}_t\Big) = \int_{]t,v]} \mathbf{E}\Big(\tilde{A}_{u-} \sum_{k=1}^{K-1} H_u^k \lambda_{k,j}(u)\Big|\ \mathcal{G}_t\Big)du \\
&= \int_{]t,v]} \mathbf{E}\Big(\tilde{A}_{u-} \sum_{k=1}^{K-1} \mathbf{E}\Big(H_u^k \mid \mathcal{F}_\infty \vee \mathcal{F}_t^C\Big)\lambda_{k,j}(u)\Big|\ \mathcal{G}_t\Big)du \\
&= \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t=i\}} \int_{]t,v]} \mathbf{E}\Big(\tilde{A}_{u-}\Big(\sum_{k=1}^{K-1} p_{i,k}(t,u)\lambda_{k,j}(u)\Big)\Big|\ \mathcal{F}_t\Big)du \\
&= \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t=i\}} \mathbf{E}\Big(\int_{]t,v]} \tilde{A}_{u-}\Big(\sum_{k=1}^{K-1} p_{i,k}(t,u)\lambda_{k,j}(u)\Big)du\Big|\ \mathcal{F}_t\Big) \\
&= \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t=i\}} \mathbf{E}\Big(\int_{]t,v]} \tilde{A}_{u-} dp_{i,j}(t,u)\Big|\ \mathcal{F}_t\Big).
\end{aligned}$$

Hence

$$I_1 + I_2 = \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t=i\}} \mathbf{E}\Big(\tilde{A}_v p_{i,j}(t,v) - \int_{]t,v]} \tilde{A}_{u-} dp_{i,j}(t,u)\Big|\ \mathcal{F}_t\Big)$$

and by integration by parts

$$\begin{aligned}
I_1 + I_2 &= \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t=i\}} \mathbf{E}\Big(\tilde{A}_t p_{i,j}(t,t) + \int_{]t,v]} p_{i,j}(t,u) d\tilde{A}_u\Big|\ \mathcal{F}_t\Big) \\
&= \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t=i\}} \mathbf{E}\Big(\int_{]t,v]} p_{i,j}(t,u) dA_u\Big|\ \mathcal{F}_t\Big).
\end{aligned}$$

□

3.3. **Examples of pricing of selected instruments.** In a series of propositions we now give examples of application of the general Theorem 3.7. All these results are stated under the assumption that the migration process C is an $\mathbb{F}$–DS Markov chain with intensity process $(\Lambda(t))_{t\geq 0}$. Whenever case we apply results based on Lemma 3.9 we assume that $(\Lambda(t))_{t\geq 0}$ satisfies condition (3.2).

3.3.1. *Defaultable bond with fractional recovery of par value.* This simplest example of rating-sensitive claim is described in Example 9. We only stress that the recovery payment is contingent on the pre-default rating, i.e. on $C_{\tau-}$.

Proposition 3.12. *The ex-dividend price D^δ of a defaultable bond with fractional recovery of par value is equal to*

$$D^\delta(t,T)\mathbf{1}_{\{C_t=i\}} = \mathbf{1}_{\{C_t=i\}} \sum_{j=1}^{K-1} B_t \mathbf{E}\Big(\frac{p_{i,j}(t,T)}{B_T} + \int_{]t,T]} \frac{\delta_j}{B_u} p_{i,j}(t,u)\lambda_{j,K}(u)du\Big|\ \mathcal{F}_t\Big)$$

for $t < T$.

3.3.2. *Credit Sensitive Note (CNS) — resetting at coupon payment date.* Recall that CSN are generally speaking corporate coupon bonds that pay coupons which are sensitive to credit rating of a firm assigned by some rating agency (see Example 10). Coupons are usually greater if the rating is worse.

Proposition 3.13. *The ex-dividend price of a Credit Sensitive Note with coupons with resetting at coupon payment date is for $t < T$ equal to*

$$B_t \mathbf{E}\Big(\sum_{k:t<T_k} d_{C_{T_k}} \frac{\mathbf{1}_{\{\tau > T_k\}}}{B_{T_k}} \Big| \mathcal{F}_t \Big) \mathbf{1}_{\{C_t = i\}}$$
$$= \mathbf{1}_{\{C_t=i\}} \mathbf{E}\Big(\sum_{k:t<T_k} e^{-\int_t^{T_k} r_u du} \Big(\sum_{j=1}^{K-1} d_j p_{i,j}(t, T_k) \Big) \Big| \mathcal{F}_t \Big).$$

Remark 3.14. Specifying d by

$$d_j = s_U (j - i_U)_+$$

where s_U is constant, one can include a rating-triggering step-up feature to coupon payments. If the rating crosses level i_U (step-up), then the coupon will increase proportionally. The so called Rating-Triggered Step-Up Bonds were issued by some European telecom companies, e.g. Deutsche Telecom, France Telecom; for details see Lando and Mortensen [24].

3.3.3. *Credit Sensitive Note — continuous coupon payments.* One can consider CSN with coupons that are paid continuously in time at rate d_{C_t} depending on rating state at t (see Example 11). This is a mathematical idealization of previous case rather than real-life example but it might be seen as approximation of discrete payments considered in the previous subsection.

Proposition 3.15. *The ex-dividend price of the Credit Sensitive Note continuous coupon payments is for $t \le T$ equal to*

$$B_t \mathbf{E}\Big(\sum_{j=1}^{K-1} \int_{]t,T]} H_u^j \frac{d_j}{B_u} du \Big| \mathcal{F}_t \Big) \mathbf{1}_{\{C_t=i\}}$$
$$= \mathbf{1}_{\{C_t=i\}} \mathbf{E}\Big(\sum_{j=1}^{K-1} \int_{]t,T]} d_j e^{-\int_t^u r_v dv} p_{i,j}(t,u) du \Big| \mathcal{F}_t \Big).$$

3.3.4. *Credit Default Swap.* Credit Default Swap is an agreement between two parties: protection buyer and protection seller. This agreement has two legs:

Premium Leg: The protection buyer agrees to pay a fixed amount κ (CDS spread) at fixed times $\mathcal{T} = \{T_1 < T_2 < \ldots < T_n\}$. He pays $\kappa\Delta_k$ at time T_k (where $\Delta_k := T_k - T_{k-1}$) provided that no default has occurred before or at T_k. Then for $t < T_n$ the value of the premium leg is equal to

$$V_P(t) = B_t \mathbf{E}\Big(\frac{\kappa}{B_\tau}(\tau - T_{\beta(\tau)-1}) \mathbf{1}_{\{t<\tau\le T_n\}} + \sum_{k=\beta(t)}^{n} \frac{\kappa\Delta_k}{B_{T_k}} \mathbf{1}_{\{\tau > T_k\}} \Big| \mathcal{G}_t \Big),$$

where $\beta(t) = \inf\{j : T_j \ge t\}$.

Default leg: The protection seller agrees to cover all losses on the reference bond provided that the loss occurs before the protection horizon T_n. For $t < T_n$, the value of this default leg is equal to

$$V_D(t) = B_t \mathbf{E}\Big(\frac{1-\delta_{C_{\tau-}}}{B_\tau} \mathbf{1}_{\{t<\tau\le T_n\}} \Big| \ \mathcal{G}_t \Big).$$

If we know the value of the spread, i.e. κ, then the CDS value at time t is the difference between the premium leg and the default leg:

$$\mathrm{CDS}(t, \mathcal{T}, \kappa) = V_P(t) - V_D(t).$$

A market CDS spread (fair spread) $\kappa = \kappa(t, \mathcal{T})$ is agreed at contract's inception (at some time $t < T_1$) in such a way that the value of the contract is 0, i.e. $\mathrm{CDS}(t, \mathcal{T}, \kappa) = 0$. The next theorem, which is an easy consequence of Theorem 3.7, provides formulae for the value of both legs expressed through the conditional transition probability process P and intensity process Λ, so we can calculate $\mathrm{CDS}(t, \mathcal{T}, \kappa)$.

Theorem 3.16. *Assume that C is an $\mathbb{F}$–DS Markov chain with intensity matrix process $(\Lambda(u))_{u\ge 0}$ and conditional transition probability process $P(s,t)$. The value of the default leg of CDS for $t < T_n$ is equal to*

$$V_D(t) = \sum_{i=1}^{K-1} \mathbf{1}_{\{C_t=i\}} \Big(\sum_{j=1}^{K-1} (1-\delta_j) \int_t^{T_n} \mathbf{E}\Big(e^{-\int_t^u r_v dv} p_{i,j}(t,u) \lambda_{j,K}(u) \Big| \mathcal{F}_t \Big) du \Big).$$

The value of the premium leg is given by

$$\begin{aligned} V_P(t) = &\sum_{i=1}^{K-1} \mathbf{1}_{\{C_t=i\}} \Big(\sum_{k=\beta(t)}^{n} \mathbf{E}\Big(e^{-\int_t^{T_k} r_u du} (1 - p_{i,K}(t,T_k)) \Big) \Delta_k \Big| \ \mathcal{F}_t \Big) \\ &+ \sum_{j=1}^{K-1} \int_t^{T_n} \Big(u - T_{\beta(u)-1} \Big) \mathbf{E}\Big(e^{-\int_t^u r_u du} p_{i,j}(t,u) \lambda_{j,K}(u) \Big| \ \mathcal{F}_t \Big) du \Big). \end{aligned}$$

In the case of the model given in Example 4 one can obtain a more explicit formula for values of the default leg and the premium leg of the CDS (see section 4.2 in Jakubowski and Niewęgłowski [18]).

References

[1] K. K. Aase, *On the consistency of the Lucas pricing formula*, Mathematical Finance , 18: 293–303, 2008.

[2] V. V. Acharya, S. R. Das, and R. K. Sundaram. *Pricing credit derivatives with rating transitions.* Financial Analysts Journal, 58: 28-44, 2002.

[3] D. Applebaum *Lévy Processes and Stochastic Calculus*, Cambridge University Press, 2004.

[4] A. Arvanitis, J. P. Laurent, and J. Gregory, *Building model for credit spreads.* Journal of Derivatives, 27–42, Spring 1999.

[5] D. Becherer, M. Schweizer. *Classical solutions to reaction-diffusion systems for hedging problems with interacting Ito and point processes* , Annals of Applied Probability 15: 1111–1144, 2005.

[6] T.R. Bielecki, S. Crepey, M. Jeanblanc and M. Rutkowski *Valuation of basket credit derivatives in the credit migrations environment*, Handbooks in Operations Research and Management Science: Financial Engineering, J. Birge and V. Linetsky eds, Elsevier, Volume 15 , pp. 471–510, 2008.

[7] T. Bielecki, M. Jeanblanc and M. Rutkowski. *Pricing and trading Credit default Swaps in a hazard process model.* Working Paper 2008.

[8] C. Blanchet-Scalliet and M. Jeanblanc. *Hazard rate for credit risk and hedging defaultable contingent claims.* Finance and Stochastics 8:145-159, 2004.

[9] T. Bielecki and M. Rutkowski. *Multiple ratings model of defaultable term structure*. Mathematical Finance, 10:125–139, 2000.
[10] T. Bielecki and M. Rutkowski. *Credit Risk: Modeling, Valuation and Hedging*. Springer Finance. Springer-Verlag Berlin Heidelberg, New York, 2001.
[11] T. Bielecki and M. Rutkowski. *Modeling of the defaultable term structure: Conditionally Markov approach*. IEEE Transactions on Automatic Control, 49:361–373, 2004.
[12] S. R. Das and P. Tufano. *Pricing credit-sensitive debt when interest rates, credit ratings and credit spreads are stochastic*. Journal Of Financial Engineering, 5:161–198, 1996.
[13] D. Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, New Jersey 1992.
[14] E. Eberlein and F. Özkan. *The defaultable Lévy term structure: Ratings and restructuring*. Mathematical Finance, 13:277–300, 2003.
[15] T. R. Hurd and A. Kuznetsov. *Fast cdo computations in the affine Markov chain model*. Working Paper, 2006.
[16] T. R. Hurd and A. Kuznetsov. *Affine Markov chain model of multifirm credit migration*. Journal of Credit Risk, 3:3–29, 2007.
[17] J. Jakubowski and M. Niewęgłowski. *Defaultable bonds with an infinite number of Lévy factors*. Preprint, 2008.
[18] J. Jakubowski and M. Niewęgłowski. *Pricing bonds and CDS in the model with rating migration induced by Cox process* Advances in Mathematics of Finance, ed. L. Stettner, Banach Centre Publication, vol 83, p. 159–182, 2008.
[19] J. Jakubowski and M. Niewęgłowski. *Hedging strategies of payment streams and application to defaultable rating-sensitive claims, in preparation.*
[20] J. Jakubowski and M. Niewęgłowski. *Local risk minimalization for defaultable rating-sensitive claims on market driven by Lévy noise, in preparation.*
[21] R. A. Jarrow, D. Lando, and T. S. Turnbull *A markov model for the term structure of credit risk spreads*. The Review of Financial Studies, 10:481–523, 1997.
[22] M. Kijima and K. Komoribayashi, *A Markov chain model for valuing credit risk derivatives* Journal of Derivatives 6:97-108, 1998.
[23] D. Lando. *On Cox processes and credit risky securities,* Review of Derivatives Research 2:99-120, 1998.
[24] D. Lando and A. Mortensen *On the pricing of step-up bonds in the European telecom sector* Journal of Credit Risk 1: 71-110, 2004.
[25] G. Last and A. Brandt. *Marked Point Processes on the Real Line: The Dynamic Approach*. Springer-Verlag Berlin Heidelberg, New York, 1995.
[26] R. Lucas. *Asset prices in an exchange economy*. Econometrica 16:1429-1445, 1978.
[27] M. Niewęgłowski. *Wybrane aspekty ryzyka obligacji modelowanych procesami Levy'ego*. PHD thesis, Warsaw University of Technology, 2007.
[28] P. Protter *Stochastic Integration and Differential Equations*. Springer-Verlag Berlin Heidelberg, New York, 2004.
[29] L.C.GT. Rogers, D. Williams *Diffusions, Markov Processes and Martingales*, Cambridge University Press, 2000.
[30] T. Rolski, H. Schmidli, V. Schmidt and J. Teugels , *Stochastic Processes for Insurance and Finance*, Wiley Series in Probability and Statistics, 1999.
[31] T. Schmidt. *An infinite factor model for credit risk*. International Journal of Theoretical and Applied Finance, 9:43–68, 2006.
[32] P. L. Walker, *On Lebesgue Integrable Derivatives*, The American Mathematical Monthly 84:287-288, 1977.

Institute of Mathematics, University of Warsaw, Banacha 2, 02-097 Warszawa, Poland, and Faculty of Mathematics and Information Science, Warsaw University of Technology, E-mail: jakub@mimuw.edu.pl

Faculty of Mathematics and Information Science, Warsaw University of Technology, Plac Politechniki 1, 00-661 Warszawa, Poland, E-mail: M.Nieweglowski@mini.pw.edu.pl