Conditional Markov Chains: Properties, Construction and Structured Dependence

Tomasz R. Bielecki

Illinois Institute of Technology, Chicago, IL 60616, USA

Jacek Jakubowski

Institute of Mathematics, University of Warsaw, Banacha 2, 02-097 Warszawa, Poland

Faculty of Mathematics and Information Science, Warsaw University of Technology, ul. Koszykowa 75, 00-662 Warszawa, Poland

Mariusz Niewęgłowski

Faculty of Mathematics and Information Science, Warsaw University of Technology, ul. Koszykowa 75, 00-662 Warszawa, Poland

---

Abstract

In this paper we contribute to the theory of conditional Markov chains (CMCs) that take finitely many values and that admit intensity. We provide a method for constructing a CMC with given intensity and with given conditional initial law, and which is also a doubly stochastic Markov chain. We provide a martingale characterization for such process, and we discuss other useful properties. We define and give sufficient and necessary conditions for strong Markovian consistency and weak Markovian consistency of a multivariate CMC. We use these results to model structured dependence between univariate CMCs, that is, to model a multivariate CMC whose components are univariate CMCs with given laws. An example of potential application of our theory is presented.

Keywords: conditional Markov chain, doubly stochastic Markov chain, compensator of a random measure, change of probability measure
2010 MSC: 60J27, 60G55

Email addresses: `bielecki@iit.edu` (Tomasz R. Bielecki), `jakub@mimuw.edu.pl` (Jacek Jakubowski), `m.nieweglowski@mini.pw.edu.pl` (Mariusz Niewęgłowski)

Contents

1. Introduction

In this paper we contribute to the study of conditional Markov chains (CMCs) with finite state spaces, which was initiated in Bielecki, Jakubowski and Niewęgłowski Bielecki et al. (2015) in an effort to enrich the theory of CMCs that was originated in Bielecki and Rutkowski Bielecki and Rutkowski (2004).

CMCs were conceptualized in the context of credit risk, where they have been found to provide a useful tool for modeling credit migrations. In many ways, a CMC is an important generalization of the concept of a default time with stochastic compensator, a key concept in the models of financial markets allowing for default of parties of a financial contract. Such a model of default time is really just a special example of a CMC: it is a CMC taking values in a state space consisting of only two states, say 0 and 1, where 0 is the transient state and 1 is the absorbing state.

In Bielecki et al. (2015) we proposed a modified definition of the conditional Markov property, which was less general than Definition 11.3.1 used in Chapter 11.3 in Bielecki and Rutkowski (2004). The reason for this was that the definition of conditional Markov property proposed in Bielecki et al. (2015) was aimed at providing a suitable framework for study of structured dependence between conditional Markov chains, a feature that can not be achieved within the CMC framework proposed in Bielecki and Rutkowski (2004). Still, the construction of a CMC presented in Bielecki et al. (2015) was not general enough, as it did not allow for study of conditional Markov families, because in that paper we only dealt with processes starting from a fixed, non-random initial state. Here, we generalize the construction of a CMC to allow for the initial state of the chain to have a non-degenerate conditional initial distribution, and, consequently, to allow for study of conditional Markov families, a study that will be conducted elsewhere.

Classical conditional Markov chains, that is, the ones defined originally in Bielecki and Rutkowski (2004), have already proven to play important role in applications in finance and in insurance (cf. Bielecki and Rutkowski Bielecki and Rutkowski (2000, 2003, 2004), Jakubowski and Niewęgłowski Jakubowski and Niewęgłowski (2010b, 2011), Eberlein and Özkan Eberlein and Özkan (2003), Eberlein and Grbac Eberlein and Grbac (2013), Biagini, Groll and Widenmann Biagini et al. (2013)). The main advantage of these processes is that, via appropriate conditioning, their primary Markov properties are mixed with dependence of their infinitesimal characteristics on relevant random factors, that do not have to be Markovian. The study of structured dependence between CMCs, which is the main theme of this paper, is crucial in applications to credit and counterparty risk, among other applications.

An important family of jump processes, so called doubly stochastic Markov chains, was introduced in Jakubowski and Niewęgłowski Jakubowski and Niewęgłowski (2010a). The conditional Markov chains constructed in the present paper turn out to be doubly stochastic Markov chains. In particular, the benefit from the construction provided here is three-fold:

• The constructed CMCs enjoy the conditional Markov property, which has unquestionable practical appeal, and

• The constructed CMCs enjoy the doubly stochastic Markov property, which has critical theoretical implications allowing for applying important tools from stochastic analysis to studying CMCs.

• The family of CMCs that can be constructed using our method is quite rich, as minimal postulates are imposed on the intensity processes used in the construction (cf. Section 3).

As said above, the main theme of the present paper is the study of structured dependence between CMCs. In a nutshell, the modeling of structured dependence between CMCs can be summarized as follows (we use a bivariate example for simplicity): given two conditional Markov chains, say Y^1 and Y^2, the modeling problem is to construct a non-trivial bivariate conditional Markov chain, say $X = (X^1, X^2)$, such that the coordinate processes X^1 and X^2 are conditionally Markovian (in some filtration), and such that the conditional law of X^i is the same as the conditional law of $Y^i, i = 1, 2$. The process X is a model for the structured dependence between Y^1 and Y^2, in the sense that the marginal processes X^1 and X^2, whose conditional laws coincide with the conditional laws of Y^1 and Y^2, are Markovian, and that the process X is conditionally Markovian. Two important comments are in order: (1) A trivial bivariate Markov process $X = (X^1, X^2)$ would be a process such that components X^1 and X^2 are conditionally independent - a situation, that is generally not of much interest. (2) Note that, in general, the bivariate process (Y^1, Y^2) may not be conditionally Markov in any filtration.

The study of structured dependence between CMCs is composed of the study of Markovian consistency and the study of Markovian copulae. The concepts of Markovian consistency and Markovian copulae were developed in the context of the problem of constructing a multivariate stochastic process in such a way that distributional laws of the univariate components of that process agree with given, predetermined laws (cf. Bielecki, Vidozzi and Vidozzi Bielecki et al. (2008c), Bielecki, Jakubowski, Vidozzi and Vidozzi Bielecki et al. (2008b), Bielecki, Jakubowski and Niewęgłowski Bielecki et al. (2010, 2013d)).

Markov copulae have been already successfully used in various applications (cf. Bielecki, Vidozzi and Vidozzi Bielecki et al. (2008c), Liang and Dong Liang and Dong (2014), Goutte and Ngoupeyou Goutte and Ngoupeyou (2013), Bielecki, Crepey, Jeanblanc and Zagrari Bielecki et al. (2012a), Bielecki, Cialenco and Iyigunler Bielecki et al. (2013a), and Bielecki, Cousin, Crepey and Herbertsson Bielecki et al. (2014d,c,a,b)).

In the present paper, we focus on the investigation of Markovian consistency and Markovian copulae with regard to finite CMCs, which are also doubly stochastic Markov chains. We follow up on the study of Markovian consistency and Markovian copulae for ordinary Markov chains that we presented in Bielecki, Jakubowski and Niewęgłowski Bielecki et al. (2013d). In fact, here we elevate the study done in Bielecki et al. (2013d) to the world of conditional Markov chains. We introduce and study the concepts of *strong Markovian consistency* and *weak Markovian consistency* for conditional Markov chains. Accordingly, we introduce and study the concepts of strong Markov copulae and weak Markov copulae for conditional Markov chains, which we call *strong CMC copulae* and *weak CMC copulae*, respectively. We refer to the discussion of practical relevance of the concepts of strong/weak Markov copulae that was done in Bielecki et al. (2013d) (see also Bielecki, Cousin, Crépey and Herbertsson Bielecki et al. (2014d) and Jakubowski and Pytel Jakubowski and Pytel (2016)). Much of what was said there applies in the context of strong/weak CMC copulae.

As already said, we confine our study to the case of finite CMCs. One might question the choice of finite CMCs as the object of interest in this paper, as one might think that this choice is very restrictive. In Bielecki, Jakubowski and Niewęgłowski Bielecki et al. (2012b) we studied strong Markovian dependence in the context of (nice) Feller processes, whereas in Bielecki et al. (2013d) we studied strong and weak Markovian dependence in the context of finite Markov chains. What we have learned from work on both these papers is that from the point of view of mathematical challenges involved in modeling of structured dependence between components of a multivariate Markov process, the finite state space set-up is actually not restrictive at all! Likewise, as it will be seen throughout this paper, studying the concepts of consistency and copulae in case of finite CMCs is quite challenging and by no means restrictive.

From the mathematical perspective, the problem of modeling of dependence between CMCs, generalizes the problem of modeling of dependence between random times. The latter problem is one of the key problems studied in the context of portfolio credit risk and counterparty risk, in case when one only considers two possible states of financial obligors: the pre-default state and the default state, with the additional caveat that the default state is absorbing, and the issue in question is the issue of modeling dependence between default times of various obligors (cf. e.g. Bielecki, Cousin, Crépey and Herbertsson Bielecki et al. (2014d, 2013b,c)). The study done in this paper allows for tackling more general problems, such as the problem of modeling of dependence between evolutions of credit ratings of financial obligors in cases where conditioning reference information is relevant; in particular, it opens a door for generalizing the set-up that was used in Biagini, Groll and Widenmann Biagini et al. (2013) to deal with an interesting problem of evaluation of premia for unemployment insurance products for a pool of individuals. This is done in response to the need for modeling dependence between dynamic systems in cases when some conditional properties of a system are important and should be accounted for. We refer to Section 8, where we discuss a relevant practical problem, which is motivated by the problem considered in Biagini et al. (2013).

The paper is organized as follows: In Section 2 we introduce the concept of CMC, which underlies the present study. In this section we also introduce and discuss the relevant concept of stochastic generator (or an intensity matrix) of a CMC. Section 3 is devoted to presentation of a specific method for constructing a CMC with a given intensity. In Section 4 we relate conditional Markov chains to doubly stochastic Markov chains. In particular, we show that any conditional Markov chain constructed using the change of measure technique used in Section 3 is also a doubly stochastic Markov chain. In Sections 5 and 6 we introduce and study the concepts of strong and weak consistency for CMCs, respectively. Section 7 is devoted to presentation of strong and weak CMC copulae, the concepts that are in the core of modeling of structured dependence between CMCs, and we also present related examples there. In Section 8 we propose a possible application of the theory developed in this paper. Finally, in the Appendix, we collect several technical results that are used throughout the text.

2. Conditional Markov Chain and Its Intensity

Let $T > 0$ be a fixed finite time horizon. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be an underlying complete probability space, which is endowed with two filtrations, $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ and $\mathbb{G} = (\mathcal{G}_t)_{t\in[0,T]}$, that are assumed to satisfy the usual conditions. The standing assumption throughout will be that all filtrations used in the paper are completed, with respect to relevant probability measures. For the future reference we also define

$$\widehat{\mathcal{G}}_t := \mathcal{F}_T \vee \mathcal{G}_t, \quad t \in [0, T], \tag{2.1}$$

as well as the corresponding filtration $\widehat{\mathbb{G}} := (\widehat{\mathcal{G}}_t)_{t\in[0,T]}$. In what what follows, we will not require that it is right-continuous.

Typically, processes considered in this paper are defined on $(\Omega, \mathcal{A}, \mathbb{P})$, and are restricted to the time interval $[0, T]$. Moreover, for any process U we denote by $\mathbb{F}^U$ the completed right-continuous filtration generated by this process. In addition, we fix a finite set S, and we denote by d the cardinality of S. Without loss of generality we take $S = \{1, 2, 3, \dots, d\}$.

Definition 2.1. *An S-valued, $\mathbb{G}$-adapted càdlàg process X is called an $(\mathbb{F}, \mathbb{G})$–conditional Markov chain if for every $x_1, \dots, x_k \in S$ and for every $0 \le t \le t_1 \le \dots \le t_k \le T$ it satisfies the following property*

$$\mathbb{P}(X_{t_1} = x_1, \dots, X_{t_k} = x_k | \mathcal{F}_t \vee \mathcal{G}_t) = \mathbb{P}(X_{t_1} = x_1, \dots, X_{t_k} = x_k | \mathcal{F}_t \vee \sigma(X_t)). \tag{2.2}$$

Remark 2.2. (i) We will call filtration $\mathbb{G}$ the *base* filtration, and we will call filtration $\mathbb{F}$ the *reference* filtration. Usually $\mathbb{G} = \mathbb{F}^X$.
(ii) It needs to be stressed that an $(\mathbb{F}, \mathbb{G})$–conditional Markov chain may not

be a classical Markov chain (in any filtration). However, if $\mathbb{G}$ is independent of $\mathbb{F}$, then the above definition reduces to the case of a classical Markov chain with respect to filtration $\mathbb{G}$, or $\mathbb{G}$–Markov chain. In other words, a classical $\mathbb{G}$-Markov chain is an $(\mathbb{F},\mathbb{G})$–conditional Markov chain for the reference filtration independent of the base filtration.

In what follows we shall write $(\mathbb{F},\mathbb{G})$-CMC, for short, in place of $(\mathbb{F},\mathbb{G})$-conditional Markov chain.

2.1. Intensity of an $(\mathbb{F},\mathbb{G})$-CMC

Let X be an $(\mathbb{F},\mathbb{G})$-CMC. For each $x \in S$ we define the corresponding state indicator process of X,

$$H^x_t := \mathbb{1}_{\{X_t=x\}}, \quad t \in [0,T]. \tag{2.3}$$

Accordingly, we define a column vector $H_t = (H^x_t, x \in S)^\top$, where $\top$ denotes transposition. Similarly, for $x, y \in S,\ x \neq y$, we define the process H^{xy} that counts the number of transitions from x to y,

$$H^{xy}_t := \#\{u \leq t : X_{u-} = x \text{ and } X_u = y\} = \int_{]0,t]} H^x_{u-} dH^y_u, \quad t \in [0,T]. \tag{2.4}$$

The following definition generalizes the concept of the generator matrix (or intensity matrix) of a Markov chain.

Definition 2.3. *We say that an $\mathbb{F}$-adapted (matrix valued) process $\Lambda_t = [\lambda^{xy}_t]_{x,y\in S}$ such that*

$$\lambda^{xy}_t \geq 0, \quad \forall x, y \in S, x \neq y, \quad \text{and} \quad \sum_{y\in S} \lambda^{xy}_t = 0, \quad \forall x \in S, \tag{2.5}$$

is an $\mathbb{F}$-stochastic generator or an $\mathbb{F}$-intensity matrix process for X, if the process $M := (M^x, x \in S)^\top$ defined as

$$M_t = H_t - \int_0^t \Lambda^\top_u H_u du, \quad t \in [0,T], \tag{2.6}$$

is an $\mathbb{F} \vee \mathbb{G}$ –local martingale (with values in $\mathbb{R}^d$).

Remark 2.4. We remark that even though the above definition is stated for an $(\mathbb{F},\mathbb{G})$-CMC process X, it applies to S-valued semimartingales.

We will now discuss the question of uniqueness of $\mathbb{F}$-intensity.

Definition 2.5. *We say that two processes Λ and $\widehat{\Lambda}$ are equivalent relative to X if*

$$\int_0^t (\Lambda_u - \widehat{\Lambda}_u)^\top H_u du = 0, \quad \forall t \in [0,T]. \tag{2.7}$$

Proposition 2.6. *Let X be an $(\mathbb{F},\mathbb{G})$-CMC.*

i) If Λ and $\widehat{\Lambda}$ are $\mathbb{F}$-intensities of X, then they are equivalent relative to X. In particular, $\mathbb{F}$-intensity of X is unique up to equivalence relative to X.

ii) Let Λ be an $\mathbb{F}$-intensity of X. If $\widehat{\Lambda}$ is an $\mathbb{F}$-adapted process equivalent to Λ relative to X, then $\widehat{\Lambda}$ is $\mathbb{F}$-intensity of X.

Proof. i) By assumption, M given by (2.6) and $\widehat{M}$ defined as

$$\widehat{M}_t = H_t - \int_0^t \widehat{\Lambda}_u^\top H_u du, \quad t \in [0,T],$$

are $\mathbb{F} \vee \mathbb{G}-$local martingales. We have that

$$\widehat{M}_t - M_t = \int_0^t (\Lambda_u - \widehat{\Lambda}_u)^\top H_u du.$$

Thus $\widehat{M} - M$ is a continuous finite variation $\mathbb{F} \vee \mathbb{G}$-martingale starting from 0, and hence it is a constant null process. Thus (2.7) holds.

ii) Note that (2.7) implies that for $\mathbb{F} \vee \mathbb{G}$ martingale M given by (2.6) it holds

$$M_t = H_t - \int_0^t \Lambda_u^\top H_u du + \int_0^t (\Lambda_u - \widehat{\Lambda}_u)^\top H_u du = H_t - \int_0^t \widehat{\Lambda}_u^\top H_u du, \quad t \in [0,T].$$

Thus $\widehat{\Lambda}$ is an $\mathbb{F}$-intensity of X. □

In Example 5.12 we exhibit an $(\mathbb{F},\mathbb{G})$-CMC X, which admits two different intensities that are equivalent relative to X.

In the case of classical Markov chains with finite state space, intensity matrix may not exist if the matrix of transition probabilities is not differentiable (e.g. when X is not quasi left continuous). In the case of $(\mathbb{F},\mathbb{G})$-CMC the situation is similar. That is, there exist $(\mathbb{F},\mathbb{G})$-CMCs that do not admit $\mathbb{F}$-intensities. We illustrate this possibility by means of the following example (see Bielecki et al. (2015) for details):

Example 2.7. Let W be a Brownian motion and let E be a random variable with unit exponential distribution and independent from W. It is shown in Bielecki et al. (2015) that the process X defined by $X_t := \mathbb{1}_{\{\tau \le t\}}$, $t \ge 0$, where $\tau := \inf\left\{t > 0 : \sup_{u\in[0,t]} W_u > E\right\}$ is an $(\mathbb{F}^W, \mathbb{F}^X)$-CMC which does not admit an $\mathbb{F}^W$-intensity matrix.

Theorem 2.8 below provides more insight into the issue of existence of $\mathbb{F}$-intensity for an $(\mathbb{F}, \mathbb{G})$-CMC.

The $\mathbb{F}$-intensity matrix of an $(\mathbb{F}, \mathbb{G})$-CMC X is related to the $\mathbb{F} \vee \mathbb{G}$-compensators of processes H^{xy}, $x, y \in S$, $x \neq y$. In fact, we have the following result, which is a special case of (Jakubowski and Niewęgłowski, 2010a, Lemma 4.3), and thus its proof is omitted.

Theorem 2.8. *Let X be an $(\mathbb{F}, \mathbb{G})$-CMC.*

1) *Suppose that X admits an $\mathbb{F}$-intensity matrix process Λ. Then for every $x, y \in S$, $x \neq y$, the process H^{xy} admits an absolutely continuous $\mathbb{F}\vee\mathbb{G}$–compensator given as $\int_0^\cdot H^x_u \lambda^{xy}_u du$, i.e. the process K^{xy} defined by*

$$K^{xy}_t = H^{xy}_t - \int_0^t H^x_u \lambda^{xy}_u du, \quad t \in [0, T], \tag{2.8}$$

is an $\mathbb{F} \vee \mathbb{G}$ –local martingale.

2) *Suppose that we are given a family of nonnegative $\mathbb{F}$-progressively measurable processes λ^{xy}, $x, y \in S$, $x \neq y$, such that for every $x, y \in S$, $x \neq y$, the process K^{xy} given in (2.8) is an $\mathbb{F} \vee \mathbb{G}$ –local martingale. Then, the matrix valued process $\Lambda_t = [\lambda^{xy}_t]_{x,y\in S}$, with diagonal elements defined as*

$$\lambda^{xx} = -\sum_{y\in S, y\neq x} \lambda^{xy}, \quad x \in S,$$

is an $\mathbb{F}$-intensity matrix of X.

We see that the $\mathbb{F}$-intensity may not exist since $\mathbb{F} \vee \mathbb{G}$-compensators of H^{xy} may not be absolutely continuous with respect to Lebesgue measure. On the other hand, absolute continuity of $\mathbb{F} \vee \mathbb{G}$-compensators of all processes H^{xy}, for $x, y \in S$, $x \neq y$, is not sufficient for existence of an $\mathbb{F}$-intensity. This is due to the fact that the density of $\mathbb{F} \vee \mathbb{G}$ compensator is, in general, $\mathbb{F} \vee \mathbb{G}$-adapted, whereas the $\mathbb{F}$-intensity is only $\mathbb{F}$-adapted.

In order to focus our study, we now introduce the following restriction:

In the rest of this paper we restrict ourselves to CMCs, which admit $\mathbb{F}$-intensity.

2.2. $(\mathbb{F},\mathbb{G})$-*CMC as a pure jump semimartingale*

It is important to note that an $(\mathbb{F},\mathbb{G})$-CMC X, admitting $\mathbb{F}$-intensity process Λ, can be viewed as a pure jump semimartingale,[1] with values in S, whose corresponding random jump measure μ defined by (cf. Jacod Jacod (1974/75))

$$\mu(\omega,dt,dz)=\sum_{s<T}\delta_{(s,\Delta X_s(\omega))}(dt,dz)\mathbb{1}_{\{\Delta X_s(\omega)\neq 0\}}=\sum_{n\geq 1}\delta_{(T_n(\omega),\Delta X_{T_n(\omega)}(\omega))}(dt,dz)\mathbb{1}_{\{T_n(\omega)<T\}},$$

where

$$T_n:=\inf\left\{t:T_{n-1}<t\leq T,\ X_t\neq X_{T_{n-1}}\right\}\wedge T,\quad T_0=0,$$

has the $\mathbb{F}\vee\mathbb{G}$ predictable projection under $\mathbb{P}$ (the $(\mathbb{F}\vee\mathbb{G},\mathbb{P})$–compensator) given as

$$\nu(\omega,dt,dz)=\sum_{x\in S}H^x_t\Big(\sum_{y\in S\setminus\{x\}}\delta_{y-x}(dz)\lambda^{xy}_t\Big)dt=\sum_{x\in S}\mathbb{1}_{\{X_t=x\}}\Big(\sum_{y\in S\setminus\{x\}}\delta_{y-x}(dz)\lambda^{xy}_t\Big)dt.$$

So the problem of construction of an $(\mathbb{F},\mathbb{G})$-CMC with an $\mathbb{F}$-intensity (matrix) process Λ is equivalent to the problem of construction of any $\mathbb{G}$-adapted, S-valued pure jump semimartingale with the $(\mathbb{F}\vee\mathbb{G},\mathbb{P})$–compensator ν given as above, and additionally satisfying condition (2.2).

Remark 2.9. With a slight abuse of terminology, we shall refer to a $\mathbb{G}$-adapted, S-valued pure jump semimartingale X with the $\mathbb{F}\vee\mathbb{G}$ compensator ν given above, as to a $\mathbb{G}$-adapted, S-valued pure jump semimartingale admitting the $\mathbb{F}$-intensity process Λ. In particular, this also means that the process M corresponding to X as in (2.6) (see Remark 2.4) is an $\mathbb{F}\vee\mathbb{G}$ – local martingale and, even though X is not necessarily $(\mathbb{F},\mathbb{G})$-CMC, the conclusions 1) and 2) of Theorem 2.8 hold.

Theorem 2.11 below shows that a $\mathbb{G}$-adapted, S-valued pure jump semimartingale admitting $\mathbb{F}$-intensity process Λ is, under some additional conditions, an $(\mathbb{F},\mathbb{G})$-CMC with the same $\mathbb{F}$-intensity process Λ. Before stating the theorem, we recall the notion of immersion between two filtrations.

[1] We adhere to the standard convention that semimartingale processes (taking values in finite dimensional spaces) are càdlàg.

Definition 2.10. We say that a filtration $\mathbb{F}$ is $\mathbb{P}$-*immersed* in a filtration $\mathbb{H}$ if $\mathbb{F} \subset \mathbb{H}$ and if every $(\mathbb{P}, \mathbb{F})$ – local martingale is a $(\mathbb{P}, \mathbb{H})$ – local martingale.

We now have,

Theorem 2.11. *Assume that*

$$\mathbb{F} \textit{ is } \mathbb{P}\textit{–immersed in } \mathbb{F} \vee \mathbb{G}. \tag{2.9}$$

Let X be a $\mathbb{G}$-adapted, S-valued pure jump semimartingale admitting the $\mathbb{F}$-intensity process Λ. Moreover suppose that

$$\textit{all real valued } \mathbb{F} - \textit{local martingales are orthogonal to components } M^x,\ x \in S, \tag{2.10}$$

of process M given by (2.6).

Then X is an $(\mathbb{F}, \mathbb{G})$-CMC with the $\mathbb{F}$-intensity process Λ.[2]

Proof. Let us fix $0 = t_0 \leq t_1 \leq \ldots \leq t_k \leq T$, and $x_1, \ldots, x_k \in S$. It is enough to show that the martingale N, given as

$$N_t = \mathbb{P}(X_{t_1} = x_1, \ldots, X_{t_k} = x_k | \mathcal{F}_t \vee \mathcal{G}_t), \quad t \in [0, T],$$

is such that N_t is $\mathcal{F}_t \vee \sigma(X_t)$ measurable for any $t \in [0, t_1]$. Indeed, this implies that

$$\mathbb{P}(X_{t_1} = x_1, \ldots, X_{t_k} = x_k | \mathcal{F}_t \vee \mathcal{G}_t) = \mathbb{P}(X_{t_1} = x_1, \ldots, X_{t_k} = x_k | \mathcal{F}_t \vee \sigma(X_t)), \quad t \in [0, t_1],$$

which is the $(\mathbb{F}, \mathbb{G})$-CMC property. To this end, for each $n = 1, \ldots, k$, we define a process V_t^n by

$$V_t^n := \prod_{l=1}^{n-1} \mathbb{1}_{\{X_{t_l} = x_l\}} H_t^\top \mathbb{E}\left(Z_t Y_{t_n} e_{x_n} \prod_{m=n}^{k-1} e_{x_m}^\top Z_{t_m} Y_{t_{m+1}} e_{x_{m+1}} | \mathcal{F}_t \right), \quad t \in [0, T],$$

where e_x denotes a column vector in $\mathbb{R}^d$ with 1 at the coordinate corresponding to state x and with zeros otherwise, and Z, Y are solutions of the random

[2] We refer to He, Wang and Yan (He et al., 1992, Definition 7.33), for notion of orthogonality of local martingales.

ODE's[3]

$$
\begin{aligned}
dZ_t &= -\Lambda_t Z_t dt, \qquad && Z_0 = \mathrm{I}, \quad t \in [0,T], \\
dY_t &= Y_t \Lambda_t dt, \qquad && Y_0 = \mathrm{I}, \quad t \in [0,T].
\end{aligned}
$$

We will show, that

$$V_t^n = N_t, \text{ for } t \in [t_{n-1}, t_n], \; n = 1, 2, \ldots, k, \tag{2.11}$$

which, in particular, implies that for every $t \in [0, t_1]$ the random variable $N_t = V_t^1$ is measurable with respect to $\mathcal{F}_t \vee \sigma(X_t)$.

We first note that, in view of Lemma 9.5 in Appendix B, the process V^n is an $\mathbb{F} \vee \mathbb{G}$ martingale on $[t_{n-1}, t_n]$. Moreover, we have that

$$V_{t_n}^n = V_{t_n}^{n+1}. \tag{2.12}$$

Indeed,

$$
\begin{aligned}
V_{t_n}^{n+1} &= \prod_{l=1}^{n} \mathbb{1}_{\{X_{t_l} = x_l\}} H_{t_n}^\top \mathbb{E}\Big(Z_{t_n} Y_{t_{n+1}} e_{x_{n+1}} \prod_{m=n+1}^{k-1} e_{x_m}^\top Z_{t_m} Y_{t_{m+1}} e_{x_{m+1}} | \mathcal{F}_{t_n} \Big) \\
&= \prod_{l=1}^{n-1} \mathbb{1}_{\{X_{t_l} = x_l\}} H_{t_n}^\top e_{x_n} H_{t_n}^\top \mathbb{E}\Big(Z_{t_n} Y_{t_{n+1}} e_{x_{n+1}} \prod_{m=n+1}^{k-1} e_{x_m}^\top Z_{t_m} Y_{t_{m+1}} e_{x_{m+1}} | \mathcal{F}_{t_n} \Big) \\
&= \prod_{l=1}^{n-1} \mathbb{1}_{\{X_{t_l} = x_l\}} H_{t_n}^\top \mathbb{E}\Big(Z_{t_n} Y_{t_n} e_{x_n} e_{x_n}^\top Z_{t_n} Y_{t_{n+1}} e_{x_{n+1}} \prod_{m=n+1}^{k-1} e_{x_m}^\top Z_{t_m} Y_{t_{m+1}} e_{x_{m+1}} | \mathcal{F}_{t_n} \Big) \\
&= \prod_{l=1}^{n-1} \mathbb{1}_{\{X_{t_l} = x_l\}} H_{t_n}^\top \mathbb{E}\Big(Z_{t_n} Y_{t_n} e_{x_n} \prod_{m=n}^{k-1} e_{x_m}^\top Z_{t_m} Y_{t_{m+1}} e_{x_{m+1}} | \mathcal{F}_{t_n} \Big) = V_{t_n}^n,
\end{aligned}
$$

where the third equality follows from (9.7) in Appendix B, and from the fact that

$$H_{t_n}^\top e_{x_n} H_{t_n}^\top = H_{t_n}^\top e_{x_n} e_{x_n}^\top.$$

We will finish the proof by demonstrating (2.11) with use of backward induction. Towards this end, we start from the last interval, i.e. we take $n = k$.

[3] The symbol "I" used below is a generic symbol for the identity matrix, whose dimension may vary depending on the context.

Observing that

$$V_{t_k}^k = \prod_{l=1}^{k-1} \mathbb{1}_{\{X_{t_l}=x_l\}} H_{t_k}^\top \mathbb{E}\Big(Z_{t_k} Y_{t_k} e_{x_k} | \mathcal{F}_{t_k} \Big) = \prod_{l=1}^{k-1} \mathbb{1}_{\{X_{t_l}=x_l\}} H_{t_k}^\top e_{x_k} = \prod_{l=1}^{k} \mathbb{1}_{\{X_{t_l}=x_l\}},$$

and using the martingale property of V^k on $[t_{k-1}, t_k]$, we conclude that for $t \in [t_{k-1}, t_k]$

$$V_t^k = \mathbb{E}(V_{t_k}^k | \mathcal{F}_t \vee \mathcal{G}_t) = \mathbb{P}(X_{t_1} = x_1, \ldots, X_{t_k} = x_k | \mathcal{F}_t \vee \mathcal{G}_t) = N_t.$$

Now, suppose that for some $n = 2, \ldots, k-1$, the process V^n coincides with N on $[t_{n-1}, t_n]$. This, together with (2.12), yields that

$$N_{t_{n-1}} = V_{t_{n-1}}^n = V_{t_{n-1}}^{n-1}.$$

Thus, by the martingale property of V^{n-1} on the interval $[t_{n-2}, t_{n-1}]$, we obtain that

$$V_t^{n-1} = \mathbb{E}(N_{t_{n-1}} | \mathcal{F}_t \vee \mathcal{G}_t) = N_t, \quad t \in [t_{n-2}, t_{n-1}].$$

So the (backward) induction principle completes the proof. □

Remark 2.12. A sufficient condition for orthogonality of real valued $\mathbb{F}$ – local martingales and components of process M is that $\mathbb{F}$ – local martingales and the process M do not have common jumps or, equivalently, that $\mathbb{F}$ – local martingales and the process X do not have common jumps. Indeed, let Z be an $(\mathbb{F}, \mathbb{P})$ – local martingale. Since M^x is a local martingale of finite variation we have that

$$[Z, M^x]_t = \sum_{0<u\le t} \Delta Z_u \Delta M_u^x = \sum_{0<u\le t} \Delta Z_u \Delta H_u^x, \quad t \in [0, T].$$

Now, note that X jumps iff one of the processes $H^x, x \in S$, jumps. Thus if X and Z do not have common jumps then $[Z, M^x]$ is the null process, hence it is a local martingale. Consequently Z and M^x are orthogonal local martingales.

We refer to Proposition 2.13 in Bielecki et al. (2015a) for an interesting example of filtrations $\mathbb{F}$ and $\mathbb{G}$ that satisfy conditions (2.9) and (2.10) of Theorem 2.11.

3. Construction of $(\mathbb{F}, \mathbb{G})$–CMC via change of measure

The construction of CMC given in this section generalizes the construction done in Bielecki et al. (2015). In Bielecki et al. (2015) the authors constructed CMCs that are starting from a given state with probability one. Here, we construct a process $(X_t)_{t\in[0,T]}$ such that X is an $(\mathbb{F}, \mathbb{G})$–CMC with the $\mathbb{F}$-intensity matrix process Λ, and with X_0 satisfying

$$\mathbb{P}(X_0 = x|\mathcal{F}_T) = \mathbb{P}(X_0 = x|\mathcal{F}_0), \quad x \in S. \tag{3.1}$$

Even though in case of ordinary Markov chains a construction of a chain starting from a given state with probability one directly leads to construction of a chain with arbitrary initial distribution, this, in general, is not the case when one deals with CMCs. In fact, some non-trivial modifications of the construction argument used in Bielecki et al. (2015) will need to be introduced below.

3.1. Preliminaries

In our construction we start from some underlying probability space, say $(\Omega, \mathcal{A}, \mathbb{Q})$, on which we are given:

(I1) A (reference) filtration $\mathbb{F}$.

(I2) An S-valued random variable ξ, such that for any $x \in S$ we have that

$$\mathbb{Q}(\xi = x|\mathcal{F}_T) = \mu_x, \tag{3.2}$$

for some $\mathcal{F}_0$–measurable random variable μ_x taking values in $[0, 1]$.

(I3) A family $\mathcal{N} = (N^{xy})_{\substack{x,y\in S\\ y\neq x}}$ of mutually independent Poisson processes, that are independent of $\mathcal{F}_T \vee \sigma(\xi)$ and with non-negative intensities $(a^{xy})_{\substack{x,y\in S\\ y\neq x}}$ (of course for $a^{xy} = 0$ we put $N^{xy} = 0$).

In what follows we take

$$\mathcal{G}_t = \Big(\bigvee_{\substack{x,y\in S\\ y\neq x}} \mathcal{F}_t^{N^{xy}} \Big) \vee \sigma(\xi), \; t \in [0, T]; \tag{3.3}$$

the corresponding filtration $\mathbb{G}$ is known to be right-continuous for $t \in [0, T)$ (see (Amendinger, 2000, Proposition 3.3)).

Given $\mathbb{G}$ defined via (3.3), we will construct $\widehat{\mathbb{G}}$-Markov chain, say X, as a solution of an appropriate stochastic differential equation. This is an intermediate step in our goal of constructing an $(\mathbb{F},\mathbb{G})$–CMC with the $\mathbb{F}$-intensity matrix process Λ, and with X_0 satisfying

$$\mathbb{P}(X_0 = x|\mathcal{F}_T) = \mathbb{P}(X_0 = x|\mathcal{F}_0), \quad x \in S, \tag{3.4}$$

for a measure $\mathbb{P}$ to be constructed later.

Proposition 3.1. *Let $A = [a^{xy}]_{x,y\in S}$, where the diagonal elements of A are defined as $a^{xx} := -\sum_{\substack{y\in S\\ y\neq x}} a^{xy}$. Assume that ξ is an S-valued random variable and $\mathcal{N} = (N^{x,y})$ are Poisson processes satisfying (I3). Then the unique strong solution of the following SDE*

$$dX_t = \sum_{\substack{x,y\in S\\ x\neq y}} (y-x)\mathbb{1}_{\{x\}}(X_{t-})dN_t^{xy}, \quad t \in [0,T], \quad X_0 = \xi, \tag{3.5}$$

is a $\widehat{\mathbb{G}}$–Markov chain with the infinitesimal generator A. Moreover, A is an $\mathbb{F}$-intensity of X under $\mathbb{Q}$.

Proof. In view of (I3), the processes N^{xy} and $N^{xy'}$, $y \neq y'$, do not jump together. Thus, the process H^{xy} defined for $x, y \in S$, $x \neq y$ by

$$H_t^{xy} = \int_0^t H_{u-}^x dN_u^{xy}, \quad t \in [0,T],$$

(cf. (2.3) for definition of H^x) counts number of transitions of X from state x to state y. Independence of N^{xy} from $\mathcal{F}_T \vee \sigma(\xi)$ implies that N^{xy} is also a $\widehat{\mathbb{G}}$-Poisson processes with intensity a^{xy}. Thus, by boundedness and $\widehat{\mathbb{G}}$-predictability of $(H_{t-}^x)_{t\in[0,T]}$, the process L^{xy} given as

$$L_t^{xy} = \int_0^t H_{u-}^x(dN_u^{xy} - a^{xy}du) = H_t^{xy} - \int_0^t H_{u-}^x a^{xy}du = H_t^{xy} - \int_0^t H_u^x a^{xy}du, \quad t \in [0,T], \tag{3.6}$$

is a $\widehat{\mathbb{G}}$-martingale. Consequently, application of relevant characterization theorem (Jakubowski and Niewęgłowski, 2010a, Thm. 4.1) yields that X is a $\widehat{\mathbb{G}}$-Markov chain with the infinitesimal generator A.

To finish the proof we observe that since X given by (3.5) is a pure jump process with finite variation, it is a semimartingale. The $(\widehat{\mathbb{G}},\mathbb{Q})$–compensator of the jump measure of X, that is, the jump characteristic of X relative to $(\widehat{\mathbb{G}},\mathbb{Q})$, is given in terms of matrix A (cf. (3.6)). Moreover, since X is adapted to filtration $\mathbb{F}\vee\mathbb{G}\subseteq\widehat{\mathbb{G}}$, then we see that X is a semimartingale with the $(\mathbb{F}\vee\mathbb{G},\mathbb{Q})$–compensator of its jump measure given in terms of matrix A. Now, A is $\mathbb{F}$–adapted (since it is deterministic), so, in view of the terminology introduced earlier (cf. Definition 2.3), A is an $\mathbb{F}$-intensity of X under $\mathbb{Q}$. □

The fact that X is a Markov chain in filtration $\widehat{\mathbb{G}}$ will be critically important below.

3.2. Canonical conditions

Definition 3.2. We say that a matrix valued process $\Lambda=[\lambda^{xy}]_{x,y\in S}$ satisfies canonical conditions relative to the pair $(S,\mathbb{F})$ if:

(C1) Λ is an $\mathbb{F}$-progressively measurable and it fulfills (2.5).

(C2) The processes λ^{xy}, $x,y\in S$, $x\neq y$, have countably many jumps $\mathbb{Q}$-a.s, and their trajectories admit left limits.

Any $\mathbb{F}$-adapted càdlàg process $\Lambda_t=[\lambda^{xy}_t]_{x,y\in S}$, for which (2.5) holds, satisfies canonical conditions. We are now ready to proceed with construction of a CMC via change of measure.

3.3. Construction of a CMC

In this section we provide a construction of a probability measure $\mathbb{P}$, under which the process X following the dynamics (3.5) is an $(\mathbb{F},\mathbb{G})$-CMC with a given $\mathbb{F}$-intensity matrix Λ and with $\mathcal{F}_T$-conditional initial distribution satisfying (3.4).

Theorem 3.3. *Let Λ satisfy canonical conditions relative to the pair $(S,\mathbb{F})$ and assume that ξ satisfies (I2). Suppose that a^{xy}, introduced in (I3), is strictly positive for all $x,y\in S$, $x\neq y$. Moreover, let X be the unique solution of SDE* (3.5)*. For each pair $x,y\in S,x\neq y$, define the processes κ^{xy} as*

$$\kappa^{xy}_t=\frac{\lambda^{xy}_{t-}}{a^{xy}}-1,\quad t\in[0,T],$$

and assume that the random variable ϑ given as

$$\vartheta=\prod_{x,y\in S:x\neq y}\exp\left(-\int_0^T H^x_{u-}a^{xy}\kappa^{xy}_u du\right)\prod_{0<u\leq T}(1+\kappa^{xy}_u\Delta H^{xy}_u),$$

satisfies $\mathbb{E}_{\mathbb{Q}}\vartheta = 1$.[4] *Finally, define on* $(\Omega, \widehat{\mathcal{G}}_T)$ *the probability* $\mathbb{P}$ *by*

$$\frac{d\mathbb{P}}{d\mathbb{Q}}\Big|_{\widehat{\mathcal{G}}_T} = \vartheta. \tag{3.7}$$

Then

(i) $\mathbb{P}_{|\mathcal{F}_T} = \mathbb{Q}_{|\mathcal{F}_T}$.

(ii) X *is an* $(\mathbb{F}, \mathbb{G})$*–CMC under* $\mathbb{P}$ *with the* $\mathbb{F}$*-intensity matrix process* Λ*, and with the initial distribution satisfying*

$$\mathbb{P}(X_0 = x|\mathcal{F}_T) = \mathbb{P}(X_0 = x|\mathcal{F}_0) = \mathbb{Q}(X_0 = x|\mathcal{F}_T), \quad x \in S. \tag{3.8}$$

Proof. In view of Theorem 2.11, in order to prove (ii) it suffices to prove that:
(a) under measure $\mathbb{P}$ process X has an $\mathbb{F}$-intensity Λ,
(b) $\mathbb{F}$ is $\mathbb{P}$–immersed in $\mathbb{F} \vee \mathbb{G}$,
(c) all real valued $(\mathbb{F}, \mathbb{P})$-martingales are orthogonal (under $\mathbb{P}$) to martingales M^x, $x \in S$,
(d) (3.8) holds.
We will prove these claims in separate steps. In the process, we will also demonstrate (i) (see Step 2).
Step 1: Here we will show that Λ is an $\mathbb{F}$-intensity of X under $\mathbb{P}$. Towards this end, we consider a $\widehat{\mathbb{G}}$-adapted process η given as

$$\eta_t = \prod_{x,y\in S:x\neq y} \exp\left(-\int_0^t H^x_{u-} a^{xy} \kappa^{xy}_u du\right) \prod_{0<u\leq t} (1 + \kappa^{xy}_u \Delta H^{xy}_u), \quad t \in [0, T],$$

so that

$$d\eta_t = \eta_{t-}\left(\sum_{x,y\in S:x\neq y} \kappa^{xy}_t dL^{xy}_t\right), \quad \eta_0 = 1,$$

where L^{xy} is a $(\widehat{\mathbb{G}}, \mathbb{Q})$-martingale given by (3.6). Consequently, process η is a $(\widehat{\mathbb{G}}, \mathbb{Q})$-local martingale. Now, note that $\eta_T = \vartheta$, and thus $\mathbb{E}_{\mathbb{Q}}\eta_T = 1 = \eta_0$. Thus η is $(\widehat{\mathbb{G}}, \mathbb{Q})$–martingale (on $[0, T]$).

Since κ^{xy} is a left-continuous and $\mathbb{F}$-adapted process, and since $\mathbb{F} \subset \widehat{\mathbb{G}}$, we conclude that κ^{xy} is $\widehat{\mathbb{G}}$-predictable. Thus, by the Girsanov theorem (see

[4] There exist many different sufficient conditions ensuring that $\mathbb{E}_{\mathbb{Q}}\vartheta = 1$. For example uniform boundedness of Λ is such a condition.

Brémaud (Brémaud, 1981, Thm. VI.T3)), we conclude that the $(\widehat{\mathbb{G}}, \mathbb{P})$ compensator of H^{xy} has density with respect to the Lebesgue measure given as[5]

$$\mathbb{1}_{\{x\}}(X_{t-})a^{xy}(1+\kappa_t^{xy}) = \mathbb{1}_{\{x\}}(X_{t-})a^{xy}\left(1+\frac{\lambda_{t-}^{xy}}{a^{xy}}-1\right) = \mathbb{1}_{\{x\}}(X_{t-})\lambda_{t-}^{xy}, \quad t \in [0,T].$$

So, for any $x \neq y$, the process $\widehat{K}^{xy}$ defined as

$$\widehat{K}_t^{xy} := H_t^{xy} - \int_0^t \mathbb{1}_{\{x\}}(X_{u-})\lambda_{u-}^{xy} du,$$

is a $\widehat{\mathbb{G}}$ – local martingale under $\mathbb{P}$. Since X is a càdlàg process and since λ^{xy} satisfies condition (C2) we see that

$$\widehat{K}_t^{xy} = H_t^{xy} - \int_0^t H_u^x \lambda_u^{xy} du, \quad t \in [0,T], \tag{3.9}$$

is a $\widehat{\mathbb{G}}$ – local martingale under $\mathbb{P}$. Note that $\mathbb{F} \vee \mathbb{G} \subset \widehat{\mathbb{G}}$ and that the process $\widehat{K}^{xy}$ is $\mathbb{F} \vee \mathbb{G}$-adapted. Taking $\tau_n := \inf\left\{ t \geq 0 : H_t^{xy} \geq n \text{ or } \int_0^t \lambda_u^{xy} du \geq n \right\}$ gives us a reducing sequence of $\widehat{\mathbb{G}}$ stopping times for $\widehat{K}^{xy}$, which are also $\mathbb{F} \vee \mathbb{G}$ stopping times. So, in view of (Föllmer and Protter, 2011, Theorem 3.7), we have that $\widehat{K}^{xy}$ is also a $\mathbb{F} \vee \mathbb{G}$ – local martingale. Thus according to Remark 2.9 we can use Theorem 2.8 to conclude that Λ is an $\mathbb{F}$-intensity of X under $\mathbb{P}$.

Step 2: Here we will prove (i). In Step 1 we proved that η is $(\widehat{\mathbb{G}}, \mathbb{Q})$–martingale. By definition of $\mathbb{P}$ and by the tower property of conditional expectations we conclude that for an arbitrary $\psi \in L^\infty(\mathcal{F}_T)$

$$\mathbb{E}_\mathbb{P}(\psi) = \mathbb{E}_\mathbb{Q}(\psi\eta_T) = \mathbb{E}_\mathbb{Q}(\mathbb{E}_\mathbb{Q}(\psi\eta_T|\widehat{\mathcal{G}}_0)) = \mathbb{E}_\mathbb{Q}(\psi\mathbb{E}_\mathbb{Q}(\eta_T|\widehat{\mathcal{G}}_0)) = \mathbb{E}_\mathbb{Q}(\psi).$$

Step 3: Next, we show that $\mathbb{F}$ is $\mathbb{P}$-immersed in $\mathbb{F} \vee \mathbb{G}$. In view of Proposition 5.9.1.1 in Jeanablanc, Yor and Chesney Jeanblanc et al. (2009) it suffices to show that for any $\psi \in L^\infty(\mathcal{F}_T)$ and any $t \in [0,T]$ it holds

$$\mathbb{E}_\mathbb{P}(\psi|\mathcal{F}_t \vee \mathcal{G}_t) = \mathbb{E}_\mathbb{P}(\psi|\mathcal{F}_t), \quad \mathbb{P} - a.s. \tag{3.10}$$

[5] We use the usual convention that $U_{0-} := 0$ for any real valued process U.

Now, observe that

$$\mathbb{P}(\eta_t > 0) = \mathbb{E}_{\mathbb{Q}}(\mathbb{1}_{\{\eta_t>0\}}\eta_T) \geq \mathbb{E}_{\mathbb{Q}}(\mathbb{1}_{\{\eta_T>0\}}\eta_T) = \mathbb{E}_{\mathbb{Q}}(\eta_T) = 1,$$

so that $\mathbb{P}(\eta_t > 0) = 1$. Moreover, η_t is $\mathcal{F}_t \vee \mathcal{G}_t$ measurable by (I3), (3.6) and (C1). Thus we have

$$\mathbb{E}_{\mathbb{P}}(\psi|\mathcal{F}_t \vee \mathcal{G}_t) = \frac{\mathbb{E}_{\mathbb{Q}}(\psi\eta_T|\mathcal{F}_t \vee \mathcal{G}_t)}{\mathbb{E}_{\mathbb{Q}}(\eta_T|\mathcal{F}_t \vee \mathcal{G}_t)} = \frac{\mathbb{E}_{\mathbb{Q}}(\psi\mathbb{E}_{\mathbb{Q}}(\eta_T|\widehat{\mathcal{G}}_t)|\mathcal{F}_t \vee \mathcal{G}_t)}{\eta_t}$$

$$= \frac{\mathbb{E}_{\mathbb{Q}}(\psi\eta_t|\mathcal{F}_t \vee \mathcal{G}_t)}{\eta_t} = \mathbb{E}_{\mathbb{Q}}(\psi|\mathcal{F}_t \vee \mathcal{G}_t) = \mathbb{E}_{\mathbb{Q}}(\psi|\mathcal{F}_t), \quad \mathbb{P} - a.s.,$$

where the third equality holds in view of the fact that η is $(\widehat{\mathbb{G}}, \mathbb{Q})$-martingale, and where the last equality holds since $\mathbb{F}$ is $\mathbb{Q}$-immersed in $\mathbb{F}\vee\mathbb{G}$ (see Corollary 9.2 in the Appendix). Hence, using (i) we conclude that

$$\mathbb{E}_{\mathbb{P}}(\psi|\mathcal{F}_t \vee \mathcal{G}_t) = \mathbb{E}_{\mathbb{P}}(\psi|\mathcal{F}_t), \quad \mathbb{P} - a.s.$$

Consequently, (3.10) holds.

Step 4: Now we show the required orthogonality, that is we prove claim (c). Towards this end it suffices to prove that all real valued $(\mathbb{F}, \mathbb{P})$-martingales do not have common jumps with X under $\mathbb{P}$ (see Remark 2.12). Let us take Z to be an arbitrary real valued $(\mathbb{F}, \mathbb{P})$-martingale. Then, in view of (i), Z is an $(\mathbb{F}, \mathbb{Q})$-martingale. By (I3), we have that $(\mathbb{F}, \mathbb{Q})$-martingales and Poisson processes in $\mathcal{N}$ are independent under $\mathbb{Q}$. Thus, by Lemma 9.3 in the Appendix A, the $\mathbb{Q}$ probability that process Z has common jumps with any process from family $\mathcal{N}$ is zero. Consequently, in view of (3.5), the $(\mathbb{F}, \mathbb{Q})$-martingale Z does not jump together with X, $\mathbb{Q}$-a.s. Therefore, by absolute continuity of $\mathbb{P}$ with respect to $\mathbb{Q}$, $\mathbb{P}$ probability that Z jumps at the same time as X is zero.

Step 5: Finally, we will show that (3.8) holds. Towards this end, let us take an arbitrary real valued function h on S. The abstract Bayes rule yields

$$\mathbb{E}_{\mathbb{P}}(h(X_0)|\mathcal{F}_T) = \frac{\mathbb{E}_{\mathbb{Q}}(h(X_0)\eta_T|\mathcal{F}_T)}{\mathbb{E}_{\mathbb{Q}}(\eta_T|\mathcal{F}_T)} = \frac{\mathbb{E}_{\mathbb{Q}}(\mathbb{E}_{\mathbb{Q}}(h(X_0)\eta_T|\widehat{\mathcal{G}}_0)|\mathcal{F}_T)}{\mathbb{E}_{\mathbb{Q}}(\mathbb{E}_{\mathbb{Q}}(\eta_T|\widehat{\mathcal{G}}_0)|\mathcal{F}_T)}$$

$$= \mathbb{E}_{\mathbb{Q}}(h(X_0)\mathbb{E}_{\mathbb{Q}}(\eta_T|\widehat{\mathcal{G}}_0)|\mathcal{F}_T) = \mathbb{E}_{\mathbb{Q}}(h(X_0)|\mathcal{F}_T) = \mathbb{E}_{\mathbb{Q}}(h(X_0)|\mathcal{F}_0),$$

where the last equality follows from the fact that by assumption (3.2) the initial condition of the process X satisfies

$$\mathbb{Q}(X_0 = x|\mathcal{F}_T) = \mathbb{Q}(X_0 = x|\mathcal{F}_0), \quad x \in S. \tag{3.11}$$

Consequently,

$$\mathbb{E}_{\mathbb{P}}(h(X_0)|\mathcal{F}_0) = \mathbb{E}_{\mathbb{P}}(\mathbb{E}_{\mathbb{P}}(h(X_0)|\mathcal{F}_T)|\mathcal{F}_0) = \mathbb{E}_{\mathbb{P}}(\mathbb{E}_{\mathbb{Q}}(h(X_0)|\mathcal{F}_0)|\mathcal{F}_0) = \mathbb{E}_{\mathbb{Q}}(h(X_0)|\mathcal{F}_0) = \mathbb{E}_{\mathbb{P}}(h(X_0)|\mathcal{F}_T).$$

This completes the proof of (3.8), and the proof of the theorem. □

4. $(\mathbb{F}, \mathbb{G})$–CMC vs $(\mathbb{F}, \mathbb{G})$–DSMC

In this section we first re-visit the concept of the doubly stochastic Markov chain. Then, we study relationships between conditional Markov chains and doubly stochastic Markov chains. These relationships are crucial for the theory of consistency of CMCs and for the theory of CMC copulae.

4.1. $(\mathbb{F}, \mathbb{G})$*–DSMC*

The concept of $(\mathbb{F}, \mathbb{G})$-doubly stochastic Markov chain ($(\mathbb{F}, \mathbb{G})$–DSMC for brevity), generalizes the notion of $\mathbb{F}$-doubly stochastic Markov chain (cf. Jakubowski and Niewęgłowski (2010a)), as well as the notion of continuous time $\mathbb{G}$-Markov chain.

Definition 4.1. *A $\mathbb{G}$-adapted càdlàg process $X = (X_t)_{t\in[0,T]}$ is called an $(\mathbb{F}, \mathbb{G})$–doubly stochastic Markov chain with state space S if for any $0 \le s \le t \le T$ and for every $y \in S$ it holds that*

$$\mathbb{P}(X_t = y \mid \mathcal{F}_T \vee \mathcal{G}_s) = \mathbb{P}(X_t = y \mid \mathcal{F}_t \vee \sigma(X_s)). \tag{4.1}$$

We refer to Jakubowski and Niewęgłowski (2010a) for examples of processes, which are $(\mathbb{F}, \mathbb{F}^X)$-DSMCs. We remark that in Jakubowski and Niewęgłowski (2010a) it was assumed that the chain X starts from some point $x \in S$ with probability one, whereas here, we allow for the initial state X_0 to be a non-constant random variable.

With any X, which is an $(\mathbb{F}, \mathbb{G})$-DSMC, we associate a matrix valued random field $P = (P(s,t),\ 0 \le s \le t \le T)$, where $P(s,t) = (p_{xy}(s,t))_{x,y\in S}$ is defined by

$$p_{x,y}(s,t) = \frac{\mathbb{P}(X_t = y, X_s = x \mid \mathcal{F}_t)}{\mathbb{P}(X_s = x|\mathcal{F}_t)} \mathbb{1}_{\{\mathbb{P}(X_s=x|\mathcal{F}_t)>0\}} + \mathbb{1}_{\{x=y\}} \mathbb{1}_{\{\mathbb{P}(X_s=x|\mathcal{F}_t)=0\}}. \tag{4.2}$$

The following result provides a characterization of $(\mathbb{F}, \mathbb{G})$-DSMC.

Proposition 4.2. *A process X is an $(\mathbb{F},\mathbb{G})$-DSMC iff there exists a stochastic matrix valued random field $\widetilde{P}(s,t) = (\widetilde{p}_{xy}(s,t))_{x,y\in S},\ 0 \le s \le t \le T$, such that:*

1) for every $s \in [0,T]$, the process $\widetilde{P}(s,\cdot)$ is $\mathbb{F}$–adapted on $[s,T]$,

2) for any $0 \le s \le t \le T$ and for every $x, y \in S$ we have

$$\mathbb{1}_{\{X_s=x\}}\mathbb{P}(X_t = y \mid \mathcal{F}_T \vee \mathcal{G}_s) = \mathbb{1}_{\{X_s=x\}}\widetilde{p}_{xy}(s,t). \tag{4.3}$$

Proof. We first prove the sufficiency. Using (4.3) we have

$$\mathbb{P}(X_t = y \mid \mathcal{F}_T \vee \mathcal{G}_s) = \sum_{x\in S} \mathbb{1}_{\{X_s=x\}}\widetilde{p}_{xy}(s,t). \tag{4.4}$$

So, taking conditional expectations with respect to $\mathcal{F}_t \vee \sigma(X_s)$ on both sides of (4.4), observing that $\mathcal{F}_t \vee \sigma(X_s) \subset \mathcal{F}_T \vee \mathcal{G}_s$, and using the tower property of conditional expectations, we obtain

$$\mathbb{P}(X_t = y \mid \mathcal{F}_t\vee\sigma(X_s)) = \mathbb{E}\Bigg(\sum_{x\in S} \mathbb{1}_{\{X_s=x\}}\widetilde{p}_{xy}(s,t) \mid \mathcal{F}_t\vee\sigma(X_s)\Bigg) = \sum_{x\in S} \mathbb{1}_{\{X_s=x\}}\widetilde{p}_{xy}(s,t),$$

where the last equality follows from measurability of $\sum_{x\in S} \mathbb{1}_{\{X_s=x\}}\widetilde{p}_{xy}(s,t)$ with respect to $\mathcal{F}_t \vee \sigma(X_s)$. This and (4.4) imply $\mathbb{P}(X_t = y \mid \mathcal{F}_T \vee \mathcal{G}_s) = \mathbb{P}(X_t = y \mid \mathcal{F}_t \vee \sigma(X_s))$, which is (4.1).

Now we prove the necessity. First we observe that, using similar arguments as in Jakubowski and Niewęgłowski (Jakubowski and Niewęgłowski, 2008, Lemma 3) (see also Bielecki, Crépey, Jeanblanc and Rutkowski (Bielecki et al., 2008a, Lemma 2.1)), we have, for $t \ge s$,

$$\begin{aligned} &\mathbb{P}(X_t = y \mid \mathcal{F}_t \vee \sigma(X_s)) \\ &= \sum_{x\in S} \mathbb{1}_{\{X_s=x\}} \left(\frac{\mathbb{P}(X_t = y, X_s = x \mid \mathcal{F}_t)}{\mathbb{P}(X_s = x|\mathcal{F}_t)} \mathbb{1}_{\{\mathbb{P}(X_s=x|\mathcal{F}_t)>0\}} + \mathbb{1}_{\{y=x\}}\mathbb{1}_{\{\mathbb{P}(X_s=x|\mathcal{F}_t)=0\}} \right) \quad \mathbb{P}-a.s. \end{aligned} \tag{4.5}$$

Consequently, in view of (4.2) we have $\mathbb{P}(X_t = y \mid \mathcal{F}_t\vee\sigma(X_s)) = \sum_{x\in S} \mathbb{1}_{\{X_s=x\}}p_{x,y}(s,t)$. It is enough now to let $\widetilde{p}_{x,y}(s,t) = p_{x,y}(s,t)$, for $x, y \in S, 0 \le s \le t \le T$. □

As we saw in the proof of Proposition 4.2 we can take $\widetilde{P} = P$, where P is given by (4.2). In addition, we note that in view of the results in Rao Rao (1972), for every $s \in [0,T]$ and for almost every $\omega \in \Omega$ the function $P(s,\cdot)$ is measurable on $[s,T]$, and for every $t \in [0,T]$ and almost every $\omega \in \Omega$ the function $P(\cdot,t)$ is measurable on $[0,t]$. This, and (4.3) justify the following definition

Definition 4.3. *The matrix valued random field $P = (P(s,t),\ 0 \leq s \leq t \leq T)$, defined by* (4.2) *is called the conditional transition probability matrix field (c–transition field) of X.*

Remark 4.4. For the future reference, we note that (4.3) can be written in the following form (recall that we take $\widetilde{P} = P$):

$$\mathbb{E}(H_t^y \mid \mathcal{F}_T \vee \mathcal{G}_s) = \sum_{x \in S} H_s^x p_{xy}(s,t), \quad \text{for } y \in S,$$

which is equivalent to

$$\mathbb{E}\left(H_t \mid \mathcal{F}_T \vee \mathcal{G}_s\right) = P(s,t)^\top H_s. \tag{4.6}$$

We know that in the case of classical Markov chains the transition semigroup and the initial distribution of the chain characterize the finite dimensional distributions of the chain, and thus they characterize the law of the chain. The next proposition is borrowed from Bielecki et al. (2015a) (see Proposition 4.6 and its proof there). It shows that, in case of an $(\mathbb{F}, \mathbb{G})$–DSMC X, the c-transition field P of X and the conditional law of X_0 given $\mathcal{F}_T$ characterize conditional law of X given $\mathcal{F}_T$.

Proposition 4.5. *If X is an $(\mathbb{F}, \mathbb{G})$–DSMC with c-transition field P, then, for arbitrary $0 = t_0 \leq t_1 \leq \ldots \leq t_n \leq t \leq T$ and $(x_1, \ldots, x_n) \in S^n$,*

$$\mathbb{P}(X_{t_1} = x_1, \ldots X_{t_n} = x_n \mid \mathcal{F}_T) = \sum_{x_0 \in S} \mathbb{P}(X_0 = x_0 | \mathcal{F}_T) \prod_{k=0}^{n-1} p_{x_k, x_{k+1}}(t_k, t_{k+1}). \tag{4.7}$$

Moreover, if

$$\mathbb{P}(X_0 = x_0 | \mathcal{F}_T) = \mathbb{P}(X_0 = x_0 | \mathcal{F}_0) \quad \textit{for every } x_0 \in S, \tag{4.8}$$

then, for arbitrary $0 \leq t_1 \leq \ldots \leq t_n \leq t \leq T$ and $(x_1, \ldots, x_n) \in S^n$,

$$\mathbb{P}(X_{t_1} = x_1, \ldots X_{t_n} = x_n \mid \mathcal{F}_T) = \mathbb{P}(X_{t_1} = x_1, \ldots X_{t_n} = x_n \mid \mathcal{F}_t). \tag{4.9}$$

Corollary 4.6. *Let X be an $(\mathbb{F}, \mathbb{G})$-DSMC with X_0 satisfying* (4.8)*. Then $\mathbb{F}$ is $\mathbb{P}$-immersed in $\mathbb{F} \vee \mathbb{F}^X$.*

Proof. In view of Proposition 4.5 process X satisfies (4.9). This, by (Jakubowski and Niewęgłowski, 2008, Lemma 2), is equivalent to $\mathbb{P}$-immersion of $\mathbb{F}$ in $\mathbb{F} \vee \mathbb{F}^X$. □

In analogy to the concept of $\mathbb{F}$-intensity for $(\mathbb{F}, \mathbb{G})$-CMCs, one considers the concept of intensity with regard to $(\mathbb{F}, \mathbb{G})$-DSMCs. Definition 4.7 introduces a concept of such intensity. This definition is stated in the form, which is consistent with the way the original definition of intensity for DSMCs was introduced in Jakubowski and Niewęgłowski (2010a). Later on, we will show that this definition can be equivalently stated in the form similar to Definition 2.3.

Definition 4.7. *We say that an $\mathbb{F}$–adapted matrix-valued process $\Gamma = (\Gamma_s)_{s\geq 0} = ([\gamma_s^{xy}]_{x,y\in S})_{s\geq 0}$ is an intensity of $(\mathbb{F}, \mathbb{G})$-DSMC X if the following conditions are satisfied:*
1)

$$\int_{]0,T]} \sum_{x\in S} |\gamma_s^{xx}|\, ds < \infty. \tag{4.10}$$

2)

$$\gamma_s^{xy} \geq 0 \quad \forall x, y \in S, x \neq y, \qquad \gamma_s^{xx} = -\sum_{y\in S: y\neq x} \gamma_s^{xy} \qquad \forall x \in S. \tag{4.11}$$

3) The Kolmogorov backward equation holds: for all $v \leq t$,

$$P(v,t) - \mathrm{I} = \int_v^t \Gamma_u P(u,t) du. \tag{4.12}$$

4) The Kolmogorov forward equation holds: for all $v \leq t$,

$$P(v,t) - \mathrm{I} = \int_v^t P(v,u) \Gamma_u du. \tag{4.13}$$

Remark 4.8. The above Kolmogorov equations admit unique solution provided that Γ satisfies (4.10) (see formulae given by Peano-Baker series or by Magnus expansion see e.g. Blanes, Casas, Oteo and Ros Blanes et al. (2009)).

4.1.1. Martingale characterizations of $(\mathbb{F}, \mathbb{G})$-DSMC

It turns out that the $(\mathbb{F}, \mathbb{G})$-DSMC property of process X is fully characterized by the martingale property (with respect to the filtration $\widehat{\mathbb{G}}$ given by (2.1)) of some processes related to X. These characterizations are given in the next theorem.

Theorem 4.9. *Let $(X_t)_{t\in[0,T]}$ be an S–valued stochastic process and $(\Gamma_t)_{t\in[0,T]}$ be an $\mathbb{F}$-adapted matrix valued process satisfying* (4.10) *and* (4.11). *The following conditions are equivalent:*
i) The process X is an $(\mathbb{F},\mathbb{G})$-DSMC with the intensity process Γ.
ii) The processes $\widehat{M}^x$ defined by

$$\widehat{M}^x_t := H^x_t - \int_{]0,t]} \gamma^{X_u,x}_u du, \qquad x \in S, \tag{4.14}$$

are $\widehat{\mathbb{G}}$ –local martingales.
iii) Let $H^{xy}_t = \int_{]0,t]} H^x_{u-} dH^y_u$. The processes K^{xy} defined by

$$K^{xy}_t := H^{xy}_t - \int_{]0,t]} H^x_s \gamma^{xy}_s ds, \quad x,y \in S,\ x \neq y, \tag{4.15}$$

are $\widehat{\mathbb{G}}$ –local martingales.
iv) The process L defined by

$$L_t := Z^\top_t H_t, \tag{4.16}$$

where Z is a unique solution to the random integral equation

$$dZ_t = -\Gamma_t Z_t dt, \quad Z_0 = \mathrm{I}, \tag{4.17}$$

is a $\widehat{\mathbb{G}}$ –local martingale.
v) For any $t \in [0,T]$, the process N^t defined as

$$N^t_s := P(s,t)^\top H_s \qquad \textit{for } 0 \le s \le t. \tag{4.18}$$

is a $\widehat{\mathbb{G}}$ martingale, where $P(s,t) := Z_s Y_t$ with

$$dY_t = Y_t \Gamma_t dt, \qquad Y_0 = \mathrm{I}, \quad t \in [0,T].$$

Proof. The proof of equivalence of (i)–(iv) goes along the lines of the proof of (Jakubowski and Niewęgłowski, 2010a, Theorem 4.1); only minor and straightforward modifications are needed, and therefore the proof is omitted. Equivalence of (iv) and (v) follows from formula $N^t_s = Y^\top_t L_s$ for $0 \le s \le t$ and the fact that Y_t is uniformly bounded $\widehat{\mathcal{G}}_0$ measurable invertible matrix (Lemma 9.4). □

The following result is direct counterpart of Proposition 2.6 and therefore we omit its proof.

Proposition 4.10. *Let X be an $(\mathbb{F}, \mathbb{G})$-DSMC.*

i) If Γ and $\widehat{\Gamma}$ are intensities of X, then they are equivalent relative to X. In particular, intensity of X is unique up to equivalence relative to X.

ii) Let Γ be an intensity of X. If $\widehat{\Gamma}$ is an $\mathbb{F}$-adapted process equivalent to Γ relative to X, then $\widehat{\Gamma}$ is intensity of X.

Since an $(\mathbb{F}, \mathbb{G})$-DSMC X is a S-valued càdlàg process, then it is a pure jump semimartingale. This observation sheds a new light on the intensity of X as the following corollary shows.

Corollary 4.11. *A intensity of an $(\mathbb{F}, \mathbb{G})$-DSMC X is an $\mathbb{F}$-intensity of X in the sense of Definition 2.3.*

Proof. The process $\widehat{M}$ is a $\widehat{\mathbb{G}}$-local martingale by Theorem 4.9.ii). In fact, it is also an $\mathbb{F} \vee \mathbb{G}$-local martingale. To see this, we take a reducing sequence of $\widehat{\mathbb{G}}$-stopping times $\tau_n := \inf \left\{ t \geq 0 : \int_0^t \sum_{y \in S} |\gamma_s^{yy}| ds \geq n \right\}$. Since $\widehat{M}$ is also $\mathbb{F} \vee \mathbb{G}$-adapted and $(\tau_n)_{n \geq 1}$ are also $\mathbb{F} \vee \mathbb{G}$ stopping times we see that $\widehat{M}$ is an $\mathbb{F} \vee \mathbb{G}$-local martingale (see e.g. (Föllmer and Protter, 2011, Theorem 3.7)). This implies that the $\mathbb{F}$-adapted process Γ is an $\mathbb{F}$-intensity of X. □

4.2. Relation between CMC and DSMC

In this section we present some aspects of relationship between the classes of $(\mathbb{F}, \mathbb{G})$-CMCs and $(\mathbb{F}, \mathbb{G})$-DSMCs.

Proposition 4.12. *Assume that $\mathbb{F}$ and $\mathbb{G}$ satisfy the immersion property* (2.9)*, and that X is an $(\mathbb{F}, \mathbb{G})$-DSMC. Then X is an $(\mathbb{F}, \mathbb{G})$-CMC. In addition if X considered as an $(\mathbb{F}, \mathbb{G})$-DSMC admits intensity Γ, then X considered as an $(\mathbb{F}, \mathbb{G})$-CMC admits $\mathbb{F}$-intensity $\Lambda = \Gamma$.*

Proof. Let us fix arbitrary $x_1, \ldots, x_k \in S$ and $0 \leq t \leq t_1 \leq \ldots \leq t_k \leq T$, and let us define a set A by $A = \{X_{t_1} = x_1, \ldots, X_{t_k} = x_k\}$. We need to show that $\mathbb{P}(A|\mathcal{F}_t \vee \mathcal{G}_t) = \mathbb{P}(A|\mathcal{F}_t \vee \sigma(X_t))$.

Towards this end we first note that, by Lemma 3.1 in Jakubowski and Niewęgłowski (2010a), we have

$$\mathbb{P}(A|\mathcal{F}_T \vee \mathcal{G}_t)\mathbb{1}_{\{X_t = x\}} = \mathbb{1}_{\{X_t = x\}} p_{x,x_1}(t, t_1) \prod_{n=1}^{k-1} p_{x_n, x_{n+1}}(t_n, t_{n+1}). \tag{4.19}$$

The tower property of conditional expectation and (4.19) imply

$$\begin{aligned}\mathbb{P}(A|\mathcal{F}_t \vee \mathcal{G}_t) &= \mathbb{E}\left(\sum_{x\in S} \mathbb{E}\left(\mathbb{1}_A|\mathcal{F}_T \vee \mathcal{G}_t\right) \mathbb{1}_{\{X_t=x\}}|\mathcal{F}_t \vee \mathcal{G}_t\right) \\ &= \mathbb{E}\left(\sum_{x\in S} \mathbb{1}_{\{X_t=x\}} p_{x,x_1}(t,t_1) \prod_{n=1}^{k-1} p_{x_n,x_{n+1}}(t_n,t_{n+1})|\mathcal{F}_t \vee \mathcal{G}_t\right) \\ &= \sum_{x\in S} \mathbb{1}_{\{X_t=x\}} \mathbb{E}\left(p_{x,x_1}(t,t_1) \prod_{n=1}^{k-1} p_{x_n,x_{n+1}}(t_n,t_{n+1})|\mathcal{F}_t \vee \mathcal{G}_t\right).\end{aligned}$$

Thus using the assumed immersion property of $\mathbb{F}$ in $\mathbb{F} \vee \mathbb{G}$ we obtain

$$\mathbb{P}(A|\mathcal{F}_t \vee \mathcal{G}_t) = \sum_{x\in S} \mathbb{1}_{\{X_t=x\}} \mathbb{E}\left(p_{x,x_1}(t,t_1) \prod_{n=1}^{k-1} p_{x_n,x_{n+1}}(t_n,t_{n+1})|\mathcal{F}_t\right),$$

which implies the CMC property.

The second claim of the theorem follows immediately from Corollary 4.11. □

Theorem 4.13. *Suppose that X is an $(\mathbb{F},\mathbb{G})$-CMC admitting an $\mathbb{F}$-intensity Λ. In addition, suppose that X is also an $(\mathbb{F},\mathbb{G})$-DSMC with an intensity Γ. Then Γ is an $\mathbb{F}$-intensity of X and Λ is an intensity of X.*

Proof. It follows from Corollary 4.11 that Γ is an $\mathbb{F}$-intensity. Thus by Proposition 2.6 Λ and Γ are equivalent relative to X. Consequentlyȷ, by Proposition 4.10 process Λ is an intensity of X. □

Let us note that, in view of Theorem 4.13, the intensity of X considered as an $(\mathbb{F},\mathbb{G})$–DSMC coincides, in the sense of Definition 2.5, with the $\mathbb{F}$-intensity Λ of X considered as an $(\mathbb{F},\mathbb{G})$–CMC. Consequently, we introduce the following definition

Definition 4.14. *We say that X is an $(\mathbb{F},\mathbb{G})$–CDMC with an $\mathbb{F}$-intensity, if it is both an $(\mathbb{F},\mathbb{G})$–CMC with an $\mathbb{F}$-intensity and an $(\mathbb{F},\mathbb{G})$–DSMC admitting an intensity.*

Remark 4.15. It is worth noting that Theorem 4.13 and Proposition 4.5 imply that if X is an $(\mathbb{F},\mathbb{G})$-CDMC with an $\mathbb{F}$-intensity, then this $\mathbb{F}$-intensity and the $\mathcal{F}_T$-conditional distribution of X_0 determine the $\mathcal{F}_T$-conditional distribution of X.

In case of process X constructed in Theorem 3.3 the result of Theorem 4.13 can be strengthen as follows.

Proposition 4.16. *Let X be a process constructed in Theorem 3.3, so that X is an $(\mathbb{F},\mathbb{G})$-CMC process with an $\mathbb{F}$-intensity process Λ. Then X is an $(\mathbb{F},\mathbb{G})$-DSMC with an intensity process $\Gamma = \Lambda$. Thus, X is an $(\mathbb{F},\mathbb{G})$-CDMC with $\mathbb{F}$-intensity.*

Proof. In Step 1 of the proof of Theorem 3.3 we showed that the processes $\widehat{K}^{xy}$, $x, y \in S$, $x \neq y$, given by (3.9), are $\widehat{\mathbb{G}}$ – local martingales. Thus, by Theorem 4.9, X is an $(\mathbb{F},\mathbb{G})$-DSMC with intensity Λ. □

The following example illustrates an alternative way for constructing an $(\mathbb{F},\mathbb{G})$-CDMC. In particular, it illustrates use of Proposition 4.12.

Example 4.17. (Time changed discrete Markov chain) Consider process $\bar{C}$, which is a discrete time Markov chain with values in $S = \{1, \ldots, K\}$ and with transition probability matrix P. In addition consider process N, which is a Cox process with càdlàg $\mathbb{F}$-intensity process $\tilde{\lambda}$. From (Jakubowski and Niewęgłowski, 2008, Theorem 7 and 9) we know that under assumption that the processes $(\bar{C}_k)_{k\geq 0}$ and $(N_t)_{t\in[0,T]}$ are independent and conditionally independent given $\mathcal{F}_T$, the process

$$C_t := \bar{C}_{N_t}$$

is an $(\mathbb{F},\mathbb{F}^C)$-DSMC. Moreover C admits intensity process $\Gamma = [\gamma^{xy}]$ given as $\gamma_t^{xy} = (P - I)_{x,y}\tilde{\lambda}_t$. Thus, by Corollary 4.6 and Proposition 4.12, the process C is an $(\mathbb{F},\mathbb{F}^C)$-CMC with $\mathbb{F}$-intensity $\Lambda = \Gamma$, and hence, it is also an $(\mathbb{F},\mathbb{F}^C)$-CDMC.

We end this section with providing sufficient conditions for a pure jump semimartingale to be a CDMC. The proof of the following theorem is omitted here, and can be found in Bielecki et al. (2015a).

Theorem 4.18. *Let $\mathbb{F}$, $\mathbb{G}$ satisfy the immersion property* (2.9). *Assume that S–valued $\mathbb{G}$–adapted pure jump semimartingale X admits an $\mathbb{F}$-intensity Λ. Moreover suppose that the orthogonality property* (2.10) *is fulfilled. Then X is an $(\mathbb{F},\mathbb{G})$-CDMC with an $\mathbb{F}$-intensity Λ.*

5. Strong Markovian Consistency of Conditional Markov Chains

Let $X = (X^1, \ldots, X^N)$ be a multivariate $(\mathbb{F}, \mathbb{F}^X)$-CMC [6] with values in $S := \times_{k=1}^N S_k$, where S_k is a finite set, $k = 1, \ldots, N$.

Definition 5.1. *(i) Let us fix $k \in \{1, \ldots, N\}$. We say that process X satisfies the strong Markovian consistency property with respect to $(X^k, \mathbb{F})$ if for every $x_1^k, \ldots, x_m^k \in S_k$ and for all $0 \le t \le t_1 \le \ldots \le t_m \le T$, it holds that*

$$\mathbb{P}\left(X_{t_m}^k = x_m^k, \ldots, X_{t_1}^k = x_1^k | \mathcal{F}_t \vee \mathcal{F}_t^X\right) = \mathbb{P}\left(X_{t_m}^k = x_m^k, \ldots, X_{t_1}^k = x_1^k | \mathcal{F}_t \vee \sigma(X_t^k)\right), \tag{5.1}$$

or, equivalently, if X^k is an $(\mathbb{F}, \mathbb{F}^X)$-CMC.[7]
(ii) If X satisfies the strong Markovian consistency property with respect to $(X^k, \mathbb{F})$ for all $k \in \{1, \ldots, N\}$, then we say that X satisfies the strong Markovian consistency property with respect to $\mathbb{F}$.

Remark 5.2. There is a relation between strong Markovian consistency of X with respect to $(X^k, \mathbb{F})$ and the concept of Granger's causality (cf. Granger Granger (1969)): Suppose that process X satisfies the strong Markovian consistency property with respect to $(X^k, \mathbb{F})$. If the reference filtration $\mathbb{F}$ is trivial, then the collection $\{X^i,\ i \neq k\}$ does not Granger cause X^k. By extenso, we may say that, in the case when reference filtration $\mathbb{F}$ is not trivial, then, "conditionally on $\mathbb{F}$", the collection $\{X^i,\ i \neq k\}$ does not Granger cause X^k.

The next definition extends the previous one by requiring that the laws of the marginal processes $X^k,\ k = 1, \ldots, N$, are predetermined. This definition will be a gateway to the concept of strong CMC copulae in Section 7.1.

Definition 5.3. *Let $\mathcal{Y} = \left\{Y^1, \ldots, Y^N\right\}$ be a family of processes such that each Y^k is an $(\mathbb{F}, \mathbb{F}^{Y^k})$-CMC with values in S_k.*
(i) Let us fix $k \in \{1, 2, \ldots, N\}$ and let process X satisfy the strong Markovian consistency property with respect to $(X^k, \mathbb{F})$. If the conditional law of X^k

[6] Definitions of strong and weak Markov consistency can be naturally extended to the case of process $X = (X^1, \ldots, X^N)$, which is a multivariate $(\mathbb{F}, \mathbb{G})$-CMC, with $\mathbb{F}^X \subseteq \mathbb{G}$. In the present paper we shall only work with $X = (X^1, \ldots, X^N)$ being a multivariate $(\mathbb{F}, \mathbb{F}^X)$-CMC.

[7] In more generality, one might define strong Markovian consistency with respect to a collection $X^I := \{X^k, k \in I \subset \{1, 2, \ldots\}\}$ of components of X. This will not be done in this paper though.

given $\mathcal{F}_T$ coincides with the conditional law of Y^k given $\mathcal{F}_T$, then we say that process X satisfies the strong Markovian consistency property with respect to $(X^k, \mathbb{F}, Y^k)$.
(ii) If X satisfies the strong Markovian consistency property with respect to $(X^k, \mathbb{F}, Y^k)$ for every $k \in \{1, 2, \dots, N\}$, then we say that X satisfies the strong Markovian consistency property with respect to $(\mathbb{F}, \mathcal{Y})$.

5.1. Sufficient and Necessary Conditions for Strong Markovian Consistency

In what follows we use the following assumption

Assumption (A):	(i) X is an $(\mathbb{F}, \mathbb{F}^X)$–CDMC admitting an intensity. (ii) $\mathbb{P}(X_0 = x_0 \vert \mathcal{F}_T) = \mathbb{P}(X_0 = x_0 \vert \mathcal{F}_0)$ for every $x_0 \in S$.

In view of Proposition 4.16, we see that Assumption (A) is satisfied for the large class of processes.

Our next goal is to provide condition characterizing strong Markovian consistency of process X with intensity Λ. Towards this end we first introduce the following condition[8]

Condition (SM-k): There exist $\mathbb{F}$-adapted processes $\lambda^{k;x^k y^k}$, $x^k, y^k \in S_k$, $x^k \neq y^k$, such that

$$\begin{aligned} &\mathbb{1}_{\{X^k_t = x^k\}} \sum_{\substack{y^n \in S_n, \\ n=1,2,\dots,N, n\neq k}} \lambda_t^{(X^1_t,\dots,X^{k-1}_t,x^k,X^{k+1}_t,\dots,X^N_t)(y^1,\dots,y^k,\dots,y^N)} \\ &= \mathbb{1}_{\{X^k_t = x^k\}} \lambda_t^{k;x^k y^k}, \quad dt \otimes d\mathbb{P}\text{-}a.e. \quad \forall x^k, y^k \in S_k, x^k \neq y^k. \end{aligned} \tag{5.2}$$

We have the following proposition, which is a direct consequence of Proposition 2.6, and thus we omit its proof.

Proposition 5.4. *Let X satisfy Assumption (A) and let Λ, $\widehat{\Lambda}$ be $\mathbb{F}$-intensities of X. Then (SM-k) holds for Λ if and only if it holds for $\widehat{\Lambda}$.*

The next theorem provides sufficient and necessary conditions for strong Markovian consistency property of process X with respect to $\mathbb{F}$. This theorem elevates the results of Theorem 1.8 from Bielecki et al. (2013d) to the universe

[8]The acronym SM comes from Strong Markov.

of conditional Markov chains studied here. We stress that Theorem 5.5 is not just a mere extension of Theorem 1.8 from Bielecki et al. (2013d), but that it opens a new horizons for studying structured dependence between jump semimartingales. Similar remarks apply to other results presented here, that are rooted in the study that we did in Bielecki et al. (2013d).

Theorem 5.5. *Let X satisfy Assumption (A), and let us fix $k \in \{1, 2, \ldots, N\}$. Then, X is strongly Markovian consistent with respect to $(X^k, \mathbb{F})$ if and only if Condition (SM-k) is satisfied. Moreover, in this case, process X^k admits the $\mathbb{F}$-intensity $\Lambda^k = [\lambda^{k;x^k y^k}]_{x^k, y^k \in S_k}$, with $\lambda^{k;x^k y^k}$ as in Condition (SM-k), and with $\lambda^{k;x^k x^k}$ given by*

$$\lambda_t^{k;x^k x^k} = - \sum_{y^k \in S_k, y^k \neq x^k} \lambda_t^{k;x^k y^k}.$$

Proof. Let Λ be an $\mathbb{F}$ intensity of X. For simplicity of notation, but without loss of generality, we give the proof for $k = 1$, and for $N = 2$, so that $S = S_1 \times S_2$, $X = (X^1, X^2)$. In this case, (5.2) takes the form

$$\mathbb{1}_{\{X_t^1 = x^1\}} \sum_{y^2 \in S_2} \lambda_t^{(x^1, X_t^2)(y^1, y^2)} = \mathbb{1}_{\{X_t^1 = x^1\}} \lambda_t^{1;x^1 y^1} \ dt \otimes d\mathbb{P}\text{-}a.e., \quad \forall x^1, y^1 \in S_1, x^1 \neq y^1. \tag{5.3}$$

Step 1: For $x^1, y^1 \in S_1$, $x^1 \neq y^1$ and for $x^2 \in S_2$ we define processes $H^{1;x^1 y^1}, H^{1;x^1}, H^{2;x^2}$ by

$$H_t^{1;x^1 y^1} := \sum_{0 < u \leq t} \mathbb{1}_{\{X_{u-}^1 = x^1, X_u^1 = y^1\}} = \sum_{x^2, y^2 \in S_2} H_t^{(x^1, x^2)(y^1, y^2)}$$

and $H_t^{1;x^1} := \mathbb{1}_{\{X_t^1 = x^1\}}, \quad H_t^{2;x^2} := \mathbb{1}_{\{X_t^2 = x^2\}}$, for $t \in [0, T]$.

Next, we consider process $K^{(x^1, x^2)(y^1, y^2)}$, given as

$$\begin{aligned} K_t^{(x^1, x^2)(y^1, y^2)} &= H_t^{(x^1, x^2)(y^1, y^2)} - \int_0^t H_u^{(x^1, x^2)} \lambda_u^{(x^1, x^2)(y^1, y^2)} \, du \\ &= H_t^{(x^1, x^2)(y^1, y^2)} - \int_0^t H_u^{1;x^1} H_u^{2;x^2} \lambda_u^{(x^1, x^2)(y^1, y^2)} \, du, \quad t \in [0, T]. \end{aligned}$$

In view of Theorem 2.8 process $K^{(x^1, x^2)(y^1, y^2)}$ is an $\mathbb{F} \vee \mathbb{F}^X$ – local martingale. Since, in view of Assumption (A), X is also an $(\mathbb{F}, \mathbb{F}^X)$-DSMC, then, Theorem

4.9 implies that $K^{(x^1,x^2)(y^1,y^2)}$ is also $\widehat{\mathbb{F}}^X$ – local martingale, where $\widehat{\mathbb{F}}^X := (\mathcal{F}_T \vee \mathcal{F}_t^X)_{t\in[0,T]}$. Consequently, process $K^{x^1y^1}$ given as

$$K_t^{x^1y^1} = \sum_{x^2,y^2\in S_2} K_t^{(x^1,x^2)(y^1,y^2)} = H_t^{1;x^1y^1} - \int_0^t \sum_{x^2,y^2\in S_2} H_u^{1;x^1} H_u^{2;x^2} \lambda_u^{(x^1,x^2)(y^1,y^2)} du \tag{5.4}$$

is an $\mathbb{F} \vee \mathbb{F}^X$ – local martingale as well as an $\widehat{\mathbb{F}}^X$ – local martingale.

Step 2: Now, assume that (5.3) holds. Then, using (5.4) we obtain that

$$\begin{aligned} K_t^{x^1y^1} &= H_t^{1;x^1y^1} - \int_0^t H_u^{1;x^1} \sum_{x^2\in S_2} \left[H_u^{2;x^2} \left(\sum_{y^2\in S_2} \lambda_u^{(x^1,x^2)(y^1,y^2)} \right) \right] du \\ &= H_t^{1;x^1y^1} - \int_0^t H_u^{1;x^1} \left[\sum_{y^2\in S_2} \lambda_u^{(x^1,X_u^2)(y^1,y^2)} \right] du = H_t^{1;x^1y^1} - \int_0^t H_u^{1;x^1} \lambda_u^{1;x^1y^1} du. \end{aligned}$$

Since $K^{x^1y^1}$ is a $\widehat{\mathbb{F}}^X$ – local martingale, then, by Theorem 4.9, the process X^1 is $(\mathbb{F},\mathbb{F}^X)$-DSMC with intensity process Λ^1. X is an $(\mathbb{F},\mathbb{F}^X)$-DSMC, so the filtration $\mathbb{F}$ is immersed in $\mathbb{F}\vee\mathbb{F}^X$ (see Corollary 4.6). Consequently, applying Proposition 4.12 we conclude that X^1 is $(\mathbb{F},\mathbb{F}^X)$-CMC.

Step 3: Conversely, assume that X^1 is an $(\mathbb{F},\mathbb{F}^X)$-CMC with $\mathbb{F}$-intensity Λ^1. So, process $\widehat{K}^{x^1y^1}$ given as

$$\widehat{K}_t^{x^1y^1} = H_t^{1;x^1y^1} - \int_0^t H_u^{1;x^1} \lambda_u^{1;x^1y^1} du, \quad t\in[0,T],$$

is an $\mathbb{F} \vee \mathbb{F}^X$ – local martingale. Recall that process $K^{x^1y^1}$ defined in (5.4) is an $\mathbb{F} \vee \mathbb{F}^X$–local martingale. Consequently, the difference $\widehat{K}^{x^1y^1} - K^{x^1y^1}$, which equals

$$\begin{aligned} \widehat{K}_t^{x^1y^1} - K_t^{x^1y^1} &= \int_0^t H_u^{1;x^1} \left(\sum_{x^2,y^2\in S_2} H_u^{2;x^2} \lambda_u^{(x^1,x^2)(y^1,y^2)} - \lambda_u^{1;x^1y^1} \right) du \\ &= \int_0^t H_u^{1;x^1} \left[\sum_{y^2\in S_2} \lambda_u^{(x^1X_u^2)(y^1y^2)} - \lambda_u^{1;x^1y^1} \right] du, \quad t\in[0,T], \end{aligned}$$

is a continuous $\mathbb{F} \vee \mathbb{F}^X$ – local martingale of finite variation, and therefore it is equal to the null process. This implies (5.3). The proof of the theorem is complete. □

The next theorem gives sufficient and necessary conditions for strong Markovian consistency property of X with respect to $(\mathbb{F}, \mathcal{Y})$. This theorem will be used to prove Proposition 5.9, which will be critically important in the study of strong CMC copulae in Section 7.1.

Theorem 5.6. *Let $\mathcal{Y} = \{Y^1, \ldots, Y^N\}$ be a family of processes such that each Y^k is an $(\mathbb{F}, \mathbb{F}^{Y^k})$-CDMC, with values in S_k, and with $\mathbb{F}$-intensity $\Psi_t^k = [\psi_t^{k;x^k y^k}]_{x^k, y^k \in S_k}$. Let process X satisfy Assumption (A). Then, X satisfies the strong Markovian consistency property with respect to $(\mathbb{F}, \mathcal{Y})$ if and only if for all $k = 1, 2, \ldots, N$, the following hold:*

(i) Condition (SM-k) is satisfied with

$$\lambda^{k;x^k y^k} = \psi^{k;x^k y^k}, \quad x^k, y^k \in S_k, x^k \neq y^k.$$

(ii) The law of X_0^k given $\mathcal{F}_T$ coincides with the law of Y_0^k given $\mathcal{F}_T$.

Proof. First we prove sufficiency. In view of (i) we conclude from Theorem 5.5 that process X is strongly Markovian consistent with respect to $(X^k, \mathbb{F})$, and that X^k admits the $\mathbb{F}$-intensity Ψ^k, for each $k = 1, 2, \ldots, N$. This, combined with (ii) implies, in view of Lemma 9.6, that X satisfies the strong Markovian consistency property with respect to $(\mathbb{F}, \mathcal{Y})$.
Now we prove necessity. Since X satisfies the strong Markovian consistency property with respect to $(\mathbb{F}, \mathcal{Y})$, then, clearly, the law of X_0^k given $\mathcal{F}_T$ coincides with the law of Y_0^k given $\mathcal{F}_T$ for all $k = 1, 2, \ldots, N$. In addition, in view of Theorem 5.5 and Lemma 9.6, we conclude that (5.2) is satisfied with $\Lambda^k = \Psi^k$, for all $k = 1, 2, \ldots, N$. □

5.2. Algebraic Conditions for Strong Markov Consistency

The necessary and sufficient condition for strong Markov consistency stated in Theorem 5.5 may not be easily verified. Here, we provide an algebraic sufficient condition for strong Markov consistency, which typically is easily verified. Towards this end let us fix $k \in \{1, 2, \ldots, N\}$, and let us consider the following condition[9]

[9]The acronym ASM comes from Algebraic Strong Markov.

Condition (ASM-k): *The $\mathbb{F}$-intensity process Λ of X satisfies, for every $t \in [0,T]$, and for all $x^k,\ y^k \in S_k$, $x^k \neq y^k$, and $\bar{x}^n, x^n \in S_n$, $n \neq k$,*

$$\sum_{\substack{y^n \in S_n,\\ n=1,2,\ldots,N, n\neq k}} \lambda_t^{(x^1,\ldots,x^k,\ldots,x^N)(y^1,\ldots,y^k,\ldots,y^N)} = \sum_{\substack{y^n \in S_n,\\ n=1,2,\ldots,N, n\neq k}} \lambda_t^{(\bar{x}^1,\ldots,\bar{x}^{k-1},x^k,\bar{x}^{k+1},\ldots,\bar{x}^N)(y^1,\ldots,y^k,\ldots,y^N)}.$$

Remark 5.7. Contrary to Condition (SM-k), whether condition (ASM-k) holds or not depends on the choice of version of $\mathbb{F}$ intensity (see Example 5.12).

We note that Condition (ASM-k) generalizes the analogous condition introduced in Bielecki, Jakubowski, Vidozzi and Vidozzi Bielecki et al. (2008b) for Markov chains, and called there Condition (M). The importance of Condition (ASM-k) stems from the fact that it is easily verifiable and from the next result.

Lemma 5.8. *Let process X satisfy Assumption (A), and let us fix $k \in \{1,2,\ldots,N\}$. Then, Condition (ASM-k) is sufficient for strong Markovian consistency of X relative to $(X^k,\mathbb{F})$ and for $\Lambda^k = [\lambda^{k;x^ky^k}]_{x^k,y^k\in S_k}$ to be an $\mathbb{F}$-intensity process of X^k, where $\lambda^{k;x^ky^k}$ is given as*

$$\lambda^{k;x^ky^k} = \sum_{\substack{y^n \in S_n,\\ n=1,2,\ldots,N, n\neq k}} \lambda^{(x^1,\ldots,x^k,\ldots,x^N)(y^1,\ldots,y^k,\ldots,y^N)} \tag{5.5}$$

for $x^k \neq y^k$, and

$$\lambda^{k;x^kx^k} = -\sum_{y^k\in S_k, y^k\neq x^k} \lambda^{k;x^ky^k}.$$

Proof. Condition (ASM-k) implies that for any $x^k, y^k \in S_k$, $x^k \neq y^k$, the following sum

$$\sum_{\substack{y^n \in S_n,\\ n=1,2,\ldots,N, n\neq k}} \lambda_t^{(x^1,\ldots,x^k,\ldots,x^N)(y^1,\ldots,y^k,\ldots,y^N)}, \tag{5.6}$$

does not depend on $x^1,\ldots,x^{k-1},x^{k+1},\ldots,x^N$. Thus, condition (5.2) holds with $\lambda^{k;x^ky^k}$ given by (5.5). Consequently, the result follows by application of Theorem 5.5. □

Proposition 5.9 below will play the key role in Section 7.1.

Proposition 5.9. *Let* $\mathcal{Y} = \{Y^1, \dots, Y^N\}$ *be a family of processes such that each* Y^k *is an* $(\mathbb{F}, \mathbb{F}^{Y^k})$*-CDMC with values in* S_k*, and with* $\mathbb{F}$*-intensity* $\Psi^k_t = [\psi^{k;x^k y^k}_t]_{x^k, y^k \in S_k}$*. Let process* X *satisfy Assumption (A). Assume that*

(i) There exists a version of $\mathbb{F}$*–intensity* Λ *which satisfies the following condition:*

for each $k = 1, 2, \dots, N,\ x^k, y^k \in S_k,\ x^k \neq y^k$*,*

$$\psi^{k;x^k y^k}_t = \sum_{\substack{y^n \in S_n, \\ n=1,2,\dots,N, n\neq k}} \lambda_t^{(x^1,\dots,x^k,\dots,x^N)(y^1,\dots,y^k,\dots,y^N)}. \tag{5.7}$$

(ii) The law of X^k_0 *given* $\mathcal{F}_T$ *coincides with the law of* Y^k_0 *given* $\mathcal{F}_T$ *for all* $k = 1, 2, \dots, N$*.*

Then, X *satisfies the strong Markovian consistency property with respect to* $(\mathbb{F}, \mathcal{Y})$*.*

Proof. We observe, that for $\mathbb{F}$–intensity Λ satisfying (i), Condition (ASM-k) holds for every $k = 1, 2, \dots, N$. Thus, by Lemma 5.8 it follows that (5.2) holds with $\lambda^{k;x^k y^k} = \psi^{k;x^k y^k}$, $\forall x^k, y^k \in S_k,\ x^k \neq y^k$. From (ii) and Theorem 5.6, we conclude that X is strongly Markovian consistent with respect to $(\mathbb{F}, \mathcal{Y})$. □

5.2.1. Condition (ASM-k) is not necessary for strong Markovian consistency

Example 5.10 below shows that, in general, Condition (ASM-k) is not necessary for strong Markovian consistency of X relative to $(X^k, \mathbb{F})$. Thus, Condition (SM-k) is (essentially) weaker than Condition (ASM-k). In fact, Condition (ASM-k) is so powerful that it implies strong Markovian consistency of X relative to $(X^k, \mathbb{F})$ regardless of the initial distribution of process X. However, whether or not Condition (SM-k) holds depends also on the initial distribution of X.[10]

[10] This observation suggests that the relation between Condition (ASM-k) and Condition (SM-k) is analogous to the relationship between strong lumpability property and weak lumpability property (cf. Ball and Yeo Ball and Yeo (1993), Burke and Rosenblatt Burke and Rosenblatt (1958)).

Example 5.10. Consider a bivariate process $X = (X^1, X^2)$ taking values in a finite state space $S = \{(0,0),(0,1),(1,0),(1,1)\}$, and such that it is an $(\mathbb{F},\mathbb{F}^X)$–CDMC. Assume that X admits the $\mathbb{F}$-intensity Λ of the form[11]

$$\Lambda_t = [\lambda_t^{xy}]_{x,y\in S} = \begin{matrix} & (0,0) & (0,1) & (1,0) & (1,1) \\ (0,0) & -a_t & 0 & 0 & a_t \\ (0,1) & 0 & 0 & 0 & 0 \\ (1,0) & 0 & 0 & 0 & 0 \\ (1,1) & b_t & 0 & 0 & -b_t \end{matrix}. \tag{5.8}$$

Let us suppose that $\mathcal{F}_T$–conditional distribution of X_0 is given as

$$\begin{aligned} \mathbb{P}(X_0 = (0,1)|\mathcal{F}_T) = \mathbb{P}(X_0 = (1,0)|\mathcal{F}_T) = 0, \\ \mathbb{P}(X_0 = (0,0)|\mathcal{F}_T) = m_0, \quad \mathbb{P}(X_0 = (1,1)|\mathcal{F}_T) = m_1, \end{aligned} \tag{5.9}$$

where m_0, m_1 are $\mathcal{F}_0$ measurable random variables.

Now let us investigate Condition (SM-1) relative to this X. One can verify that c–transition field of X (see Definition 4.3) has the following structure

$$\begin{matrix} & (0,0) & (0,1) & (1,0) & (1,1) \\ (0,0) & P^1_{00}(s,t) & 0 & 0 & P^1_{01}(s,t) \\ (0,1) & 0 & 1 & 0 & 0 \\ (1,0) & 0 & 0 & 1 & 0 \\ (1,1) & P^1_{10}(s,t) & 0 & 0 & P^1_{11}(s,t) \end{matrix}.$$

Thus, in view of Proposition 4.5, we conclude that for any $t \in [0,T]$

$$\mathbb{P}(X_t = (0,1)|\mathcal{F}_T) = \mathbb{P}(X_t = (1,0)|\mathcal{F}_T) = 0, \tag{5.10}$$

$$\mathbb{P}(X_t = (0,0)|\mathcal{F}_T) = m_0 P^1_{00}(0,t) + m_1 P^1_{10}(0,t).$$

Consequently, as we will show now Condition (SM-1) (i.e. (5.2) for $k = 1$) is satisfied here. In fact, taking $x^1 = 0, y^1 = 1$ and invoking (5.9), we obtain that

$$\begin{aligned} &\mathbb{1}_{\{X^1_t=0\}}(\lambda_t^{(0,X^2_t)(1,0)} + \lambda_t^{(0,X^2_t)(1,1)}) \\ &= \mathbb{1}_{\{X^1_t=0,X^2_t=0\}}(\lambda_t^{(0,0)(1,0)} + \lambda_t^{(0,0)(1,1)}) + \mathbb{1}_{\{X^1_t=0,X^2_t=1\}}(\lambda_t^{(0,1)(1,0)} + \lambda_t^{(0,1)(1,1)}) \\ &= \mathbb{1}_{\{X^1_t=0,X^2_t=0\}}a_t + \mathbb{1}_{\{X^1_t=0,X^2_t=1\}}0 = (\mathbb{1}_{\{X^1_t=0,X^2_t=0\}} + \mathbb{1}_{\{X^1_t=0,X^2_t=1\}})a_t = \mathbb{1}_{\{X^1_t=0\}}a_t, \end{aligned}$$

[11]It was shown in Proposition 4.16 that one can always construct an $(\mathbb{F},\mathbb{F}^X)$–CDMC with a given $\mathbb{F}$-intensity Λ.

where the third equality follows from the fact that

$$\mathbb{1}_{\{X_t^1=0,X_t^2=1\}} = 0, \quad dt \otimes d\mathbb{P}, \tag{5.11}$$

which is a consequence of (5.10). Analogously, for $x^1 = 1, y^1 = 0$ it holds

$$\begin{aligned}
&\mathbb{1}_{\{X_t^1=1\}}(\lambda_t^{(1,X_t^2)(0,0)} + \lambda_t^{(1,X_t^2)(0,1)}) \\
&= \mathbb{1}_{\{X_t^1=1,X_t^2=0\}}(\lambda_t^{(1,0)(0,0)} + \lambda_t^{(1,0)(0,1)}) + \mathbb{1}_{\{X_t^1=1,X_t^2=1\}}(\lambda_t^{(1,1)(0,0)} + \lambda_t^{(1,1)(0,1)}) \\
&= \mathbb{1}_{\{X_t^1=1,X_t^2=0\}}0 + \mathbb{1}_{\{X_t^1=1,X_t^2=1\}}b_t = (\mathbb{1}_{\{X_t^1=1,X_t^2=0\}} + \mathbb{1}_{\{X_t^1=1,X_t^2=1\}})b_t = \mathbb{1}_{\{X_t^1=1\}}b_t,
\end{aligned}$$

where we used the fact that

$$\mathbb{1}_{\{X_t^1=1,X_t^2=0\}} = 0, \quad dt \otimes d\mathbb{P}. \tag{5.12}$$

Thus, Condition (SM-1) holds here for $\lambda_t^{1;01} = a_t$, $\lambda_t^{1;11} = b_t$. Similarly, one can show that Condition (SM-2) is fulfilled for $\lambda_t^{2;01} = a_t$, $\lambda_t^{2;11} = b_t$. Thus, X is strongly Markovian consistent with respect to $\mathbb{F}$. However, Condition (ASM-1) is not satisfied here (regardless of the initial distribution of X) since for every $t \in [0,T]$ we have

$$\lambda_t^{(0,0)(1,0)} + \lambda_t^{(0,0)(1,1)} = a_t \neq 0 = \lambda_t^{(0,1)(1,0)} + \lambda_t^{(0,1)(1,1)}. \qquad \square$$

Remark 5.11. It is worth noting that strong Markovian consistency depends on the initial distribution of X. Consequently, we may have two processes: X which is $(\mathbb{F},\mathbb{F}^X)$-CMC and Y which is $(\mathbb{F},\mathbb{F}^Y)$-CMCs with the same $\mathbb{F}$-intensity, such that one of them is strongly Markovian consistent and the other one is not. In fact, let Y be an $(\mathbb{F},\mathbb{F}^Y)$-CMC taking values in a finite state space $S = \{(0,0),(0,1),(1,0),(1,1)\}$, endowed with the same $\mathbb{F}$-intensity as in Example 5.10, and with the conditional initial distribution such that either $\mathbb{P}(\mathbb{P}(Y_0 = (0,1)|\mathcal{F}_T) > 0) > 0$, or $\mathbb{P}(\mathbb{P}(Y_0 = (1,0)|\mathcal{F}_T) > 0) > 0$. For process Y equality (5.11) is not satisfied, and thus Condition (SM-1) does not hold. Consequently, process Y is not strongly Markovian consistent with respect to $\mathbb{F}$.

In the next example we will show that an $(\mathbb{F},\mathbb{F}^X)$-CMC X may have intensity for which Condition (ASM-k) does not hold, and it may admit another version of intensity, in the sense of Definition 2.5, for which Condition (ASM-k) is fulfilled.

Example 5.12. Let us take X as in Example 5.10. In that example we proved that Conditions (ASM-1) and (ASM-2) are not satisfied by the $\mathbb{F}$-intensity Λ given by (5.8). However there exists another version of $\mathbb{F}$-intensity of X, say Γ, for which Conditions (ASM-1) and (ASM-2) are satisfied. Indeed, let us consider that process Γ defined by

$$\Gamma_t = [\gamma_t^{xy}]_{x,y\in S} = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -a_t & 0 & 0 & a_t \\ b_t & -a_t - b_t & 0 & a_t \\ b_t & 0 & -a_t - b_t & a_t \\ b_t & 0 & 0 & -b_t \end{pmatrix} \end{array}.$$

Γ is an $\mathbb{F}$-intensity of X, by Proposition 2.6(ii), since in view of (5.11) and (5.12) it holds that

$$\int_0^t (\Gamma_u - \Lambda_u)^\top H_u du = 0, \quad t \in [0,T].$$

Finally, we see that Conditions (ASM-1) and (ASM-2) are satisfied for Γ, because

$$\begin{aligned} \gamma_t^{(0,0)(1,0)} + \gamma_t^{(0,0)(1,1)} &= a_t = \gamma_t^{(0,1)(1,0)} + \gamma_t^{(0,1)(1,1)}, \\ \gamma_t^{(1,1)(0,0)} + \gamma_t^{(1,1)(0,1)} &= b_t = \gamma_t^{(1,0)(0,0)} + \gamma_t^{(1,0)(0,1)}, \\ \gamma_t^{(0,0)(0,1)} + \gamma_t^{(0,0)(1,1)} &= a_t = \gamma_t^{(1,0)(0,1)} + \gamma_t^{(1,0)(1,1)}, \\ \gamma_t^{(1,1)(0,0)} + \gamma_t^{(1,1)(1,0)} &= b_t = \gamma_t^{(0,1)(0,0)} + \gamma_t^{(0,1)(1,0)}. \qquad \square \end{aligned}$$

6. Weak Markovian Consistency of Conditional Markov Chains

We will study here the concept of weak Markovian consistency. As in Section 5 let us consider $X = (X^1, \ldots, X^N)$ – a multivariate $(\mathbb{F}, \mathbb{F}^X)$-CMC with values in $S := \times_{k=1}^N S_k$ (recall that S_k is a finite set, $k = 1, \ldots, N$), and admitting an $\mathbb{F}$-intensity Λ.

Remark 6.1. In many respects, the concept of weak consistency is more important in practical applications than the concept of strong Markovian consistency. For example, in the context of credit risk strong Markovian consistency prohibits so called default contagion, except for the extreme case of joint defaults. On the contrary, weak Markovian consistency allows to model not only default contagion but, more generally, contagion of credit migrations. We refer to Example 2.4 in Bielecki et al. (2014d) for some insight into these issues.

As it will be seen below, the definitions and results regarding the weak Markovian consistency to some extent are parallel those regarding the strong Markovian consistency. But, as always, "the devil is in the details", so the reader is kindly asked to be patient with presentation that follows.

Definition 6.2. *(i) Let us fix $k \in \{1, \dots, N\}$. We say that the process X satisfies the weak Markovian consistency property relative to $(X^k, \mathbb{F})$ if for every $x_1^k, \dots, x_m^k \in S_k$ and for all $0 \le t \le t_1 \le \dots \le t_m \le T$, it holds*

$$\mathbb{P}\left(X_{t_m}^k = x_m^k, \dots, X_{t_1}^k = x_1^k | \mathcal{F}_t \vee \mathcal{F}_t^{X^k}\right) = \mathbb{P}\left(X_{t_m}^k = x_m^k, \dots, X_{t_1}^k = x_1^k | \mathcal{F}_t \vee \sigma(X_t^k)\right), \tag{6.1}$$

or, equivalently, if X^k is a $(\mathbb{F}, \mathbb{F}^{X^k})$-CMC.
(ii) If X satisfies the weak Markovian consistency property with respect to $(X^k, \mathbb{F})$ for all $k \in \{1, \dots, N\}$, then we say that X satisfies the weak Markovian consistency property with respect to $\mathbb{F}$.

Definition 6.3. *Let $\mathcal{Y} = \{Y^1, \dots, Y^N\}$ be a family of processes such that each Y^k is an $(\mathbb{F}, \mathbb{F}^{Y^k})$-CMC with values in S_k.*
(i) Let us fix $k \in \{1, 2, \dots, N\}$ and let the process X satisfy the weak Markovian consistency property with respect to $(X^k, \mathbb{F})$. If the conditional law of X^k given $\mathcal{F}_T$ coincides with the conditional law of Y^k given $\mathcal{F}_T$, then we say that X satisfies the weak Markovian consistency property with respect to $(X^k, \mathbb{F}, Y^k)$.
(ii) If X satisfies the weak Markovian consistency property with respect to $(X^k, \mathbb{F}, Y^k)$ for every $k \in \{1, 2, \dots, N\}$, then we say that X satisfies the weak Markovian consistency property with respect to $(\mathbb{F}, \mathcal{Y})$.

6.1. Sufficient and necessary conditions for weak Markovian consistency

We postulate in this subsection that the process X satisfies Assumption (A) (see Section 5.1), and we aim here at providing a condition characterizing weak Markovian consistency of the process X.

Let us start from introducing[12]

[12]The acronym WM comes from Weak Markov.

Condition (WM-k): There exist $\mathbb{F}$-adapted processes $\lambda^{k;x^k y^k}$, $x^k, y^k \in S_k$, $x^k \neq y^k$, such that

$$\mathbb{1}_{\{X_t^k = x^k\}} \sum_{\substack{x^n, y^n \in S_n \\ n=1,2,\dots,N, n\neq k}} \lambda_t^{(x^1,\dots,x^N)(y^1,\dots,y^N)} \mathbb{E}_{\mathbb{P}} \left(\mathbb{1}_{\{X_t^1 = x^1, \dots, X_t^{k-1} = x^{k-1}, X_t^{k+1} = x^{k+1}, \dots, X_t^N = x^N\}} \middle| \mathcal{F}_t \vee \mathcal{F}_t^{X^k} \right)$$
$$= \mathbb{1}_{\{X_t^k = x^k\}} \lambda_t^{k;x^k y^k}, \quad dt \otimes d\mathbb{P}\text{-a.e. } \forall x^k, y^k \in S_k,\ x^k \neq y^k. \tag{6.2}$$

Similarly as in the case of Condition (SM-k) we have the following proposition, which is a direct consequence of Proposition 2.6:

Proposition 6.4. *Let X satisfy Assumption (A) and Λ, $\widehat{\Lambda}$ be $\mathbb{F}$-intensities of X. Then (WM-k) holds for Λ if and only if it holds for $\widehat{\Lambda}$.*

The next theorem characterizes weak Markovian consistency in the present set-up.

Theorem 6.5. *The process X with an $\mathbb{F}$-intensity Λ is weakly Markovian consistent relative to $(X^k, \mathbb{F})$ if and only if Condition (WM-k) is satisfied. Moreover, X^k admits an $\mathbb{F}$-intensity process*

$$\Lambda^k := [\lambda^{k;x^k y^k}]_{x^k, y^k \in S_k}, \tag{6.3}$$

with $\lambda^{k;x^k x^k}$ given by

$$\lambda_t^{k;x^k x^k} = - \sum_{y^k \in S_k, y^k \neq x^k} \lambda_t^{k;x^k y^k}, \qquad \forall x^k \in S_k.$$

Proof. For simplicity of notation we give proof for $k = 1$ and $N = 2$. In this case, (6.2) takes the following form (recall our notation: $H_t^{k;x^k} = \mathbb{1}_{\{X_t^k = x^k\}}$)

$$H_t^{1;x^1} \sum_{x^2, y^2 \in S_2} \lambda_t^{(x^1 x^2)(y^1 y^2)} \mathbb{E}_{\mathbb{P}} \left(H_t^{2;x^2} | \mathcal{F}_t \vee \mathcal{F}_t^{X^1} \right) = H_t^{1;x^1} \lambda_t^{1;x^1 y^1}, \quad dt \otimes d\mathbb{P}\text{-a.e. } \forall x^1, y^1 \in \mathcal{X}^1,\ x^1 \neq y^1. \tag{6.4}$$

<u>Step 1:</u> In Step 1 of the proof of Theorem 5.5 we have shown that the process $K^{x^1 y^1}$ given in (5.4) is an $\mathbb{F} \vee \mathbb{F}^X$ – local martingale. Now let us denote by $\widetilde{K}^{x^1 y^1}$ the optional projection of $K^{x^1 y^1}$ on the filtration $\mathbb{F} \vee \mathbb{F}^{X^1}$.[13] Observe

[13] We note that for existence of optional projections we do not need right continuity of the filtration (see Ethier, Kurtz (Ethier and Kurtz, 1986, Theorem 2.4.2)).

that the sequence

$$\tau_n := \inf\left\{t \geq 0 : H_t^{1;x^1y^1} \geq n \text{ or } \int_0^t \Big(\sum_{x^2,y^2\in S_2} \lambda_u^{(x^1,x^2)(y^1,y^2)}\Big) du \geq n\right\},\ n=1,2,\dots,$$

is sequence of $\mathbb{F}\vee\mathbb{F}^X$–stopping times, as well as $\mathbb{F}\vee\mathbb{F}^{X^1}$–stopping times, and that is a reducing sequence for $K^{x^1y^1}$. So, by (Föllmer and Protter, 2011, Theorem 3.7), the process $\widetilde{K}^{x^1y^1}$ is an $\mathbb{F}\vee\mathbb{F}^{X^1}$– local martingale.

Following the reasoning in Theorem 5.25 in He et al. (1992), we obtain that

$$\mathbb{E}\bigg(\int_0^t H_u^{1;x^1} \sum_{x^2,y^2\in S_2} H_u^{2;x^2} \lambda_u^{(x^1,x^2)(y^1,y^2)} du \Big| \mathcal{F}_t\vee\mathcal{F}_t^{X^1}\bigg) = \int_0^t H_u^{1;x^1} \mathbb{E}\bigg(\sum_{x^2,y^2\in S_2} H_u^{2;x^2} \lambda_u^{(x^1,x^2)(y^1,y^2)} \Big| \mathcal{F}_u\vee\mathcal{F}_u^{X^1}\bigg) du$$

and hence the process $\widetilde{K}^{x^1y^1}$ given as

$$\widetilde{K}_t^{x^1y^1} = H_t^{1;x^1y^1} - \int_0^t H_u^{1;x^1} \mathbb{E}\bigg(\sum_{x^2,y^2\in S_2} H_u^{2;x^2} \lambda_u^{(x^1,x^2)(y^1,y^2)} \Big| \mathcal{F}_u\vee\mathcal{F}_u^{X^1}\bigg) du, \quad t\in[0,T], \tag{6.5}$$

is an $\mathbb{F}\vee\mathbb{F}^{X^1}$ –local martingale.
Step 2: Now, suppose that (6.4) holds. Then we have

$$\widetilde{K}_t^{x^1y^1} = H_t^{1;x^1y^1} - \int_0^t H_u^{1;x^1} \lambda_u^{1;x^1y^1} du, \quad t\in[0,T].$$

Thus according to Remark 2.9 we can apply Theorem 2.8 to process X^1 in order to conclude that Λ^1 is an $\mathbb{F}$-intensity of X^1, so that $\mathbb{R}^d$-valued process $\widetilde{M}^1 = (\widetilde{M}^{1;x^1}; x^1\in S_1)^\top$, given as

$$\widetilde{M}_t^1 = H_t^1 - \int_0^t (\Lambda_u^1)^\top H_u^1 du, \quad t\in[0,T],$$

is an $\mathbb{F}\vee\mathbb{F}^{X^1}$ –local martingale.

Next, using Theorem 2.11 we will show that X^1 is an $(\mathbb{F},\mathbb{F}^{X^1})$-CMC. Towards this end, we first observe that Assumption (A) implies, by Corollary 4.6, that $\mathbb{F}$ is immersed in $\mathbb{F}\vee\mathbb{F}^X$, and thus $\mathbb{F}$ is immersed $\mathbb{F}\vee\mathbb{F}^{X^1}$. Moreover, as we will show now, all real valued $\mathbb{F}$-local martingales are orthogonal to processes $M^x, x\in S$, that are components of process M defined in (2.6).

Indeed, let us take an arbitrary real valued $\mathbb{F}$-local martingale N. Then, by definition of M and the fact that M is a pure-jump local martingale we have, for any $(x^1, x^2) \in S$,

$$[N, M^{(x^1,x^2)}]_t = \sum_{0<u\le t} \Delta N_u \Delta M_u^{(x^1,x^2)} = \sum_{0<u\le t} \Delta N_u \Delta H_u^{(x^1,x^2)}, \quad t \in [0,T]. \tag{6.6}$$

Now, since the jump times of N are $\mathbb{F}$-stopping times, then by Proposition 6.1 in Jakubowski and Niewęgłowski (2010a) we conclude that N and X do not have common jump times, equivalently, N and M do not have common jump times. Therefore, $[N, M^{x^1,x^2}] = 0$, so that N is orthogonal to all processes M^{x^1,x^2}.

From the above we will deduce that all real valued $\mathbb{F}$-local martingales are orthogonal to processes $\widetilde{M}^{1;x^1}, x^1 \in S_1$, that are components of process $\widetilde{M}^1$ defined above. In fact taking N as above we see that orthogonality of N and $\widetilde{M}^{1;x^1}$ follows from the following equalities

$$\begin{aligned}
[N, \widetilde{M}^{1;x^1}]_t &= \sum_{0<u\le t} \Delta N_u \Delta \widetilde{M}_u^{1;x^1} = \sum_{0<u\le t} \Delta N_u \Delta H_u^{1;x^1} = \sum_{0<u\le t} \sum_{x^2\in S_2} \Delta N_u \Delta H_u^{(x^1,x^2)} \\
&= \sum_{x^2\in S_2} \sum_{0<u\le t} \Delta N_u \Delta H_u^{(x^1,x^2)} = \sum_{x^2\in S_2} [N, M^{(x^1,x^2)}]_t = 0, \quad t \in [0,T],
\end{aligned}$$

where the penultimate equality follows from (6.6).

Consequently, we see that assumptions of Theorem 2.11 are fulfilled (taking there $X = X^1$ and $\mathbb{G} = \mathbb{F}^{X^1}$), and thus we may conclude that X^1 is $(\mathbb{F}, \mathbb{F}^{X^1})$-CMC with $\mathbb{F}$-intensity $\Lambda^1(t) = [\lambda_t^{1;x^1y^1}]_{x^1,y^1\in S_1}$.

Step 3: Conversely, assume that X^1 is an $(\mathbb{F}, \mathbb{F}^{X^1})$-CMC with $\mathbb{F}$-intensity $\Lambda^1(t) = [\lambda_t^{1;x^1y^1}]_{x^1,y^1\in S_1}$. Fix $x^1, y^1 \in S_1$, $x^1 \neq y^1$. In an analogous way as in Step 3 of the proof of Theorem 5.5 we see that the difference $\widetilde{K}^{x^1y^1} - \widehat{K}^{x^1y^1}$, where $\widetilde{K}^{x^1y^1}$ given in (6.5) and

$$\widehat{K}_t^{x^1y^1} = H_t^{1;x^1y^1} - \int_0^t H_u^{1;x^1} \lambda_u^{1;x^1y^1} du, \quad t \in [0,T],$$

is a continuous $\mathbb{F} \vee \mathbb{F}^{X^1}$ – local martingale of finite variation. Therefore it is equal to 0, which implies (6.4). The proof of the theorem is complete. $\square$

The next theorem gives sufficient and necessary conditions for weak Markovian consistency property of X with respect to $(\mathbb{F}, \mathcal{Y})$. We omit the proof of this theorem, as its proof can be derived from the proof of Theorem 5.6 by using Theorem 6.5 instead of Theorem 5.5.

Theorem 6.6. *Let $\mathcal{Y} = \left\{Y^1, \ldots, Y^N\right\}$ be a family of processes such that each Y^k is an $(\mathbb{F}, \mathbb{F}^{Y^k})$-CDMC, with values in S_k, and with $\mathbb{F}$-intensity $\Psi_t^k = [\psi_t^{k;x^k y^k}]_{x^k, y^k \in S_k}$. Let process X satisfy Assumption (A) and let Λ be a version of its $\mathbb{F}$-intensity. Then, X satisfies the weak Markovian consistency property with respect to $(\mathbb{F}, \mathcal{Y})$ if and only if for all $k = 1, 2, \ldots, N$, the following hold:*

(i) Condition (WM-k) is satisfied with Ψ^k in place of Λ^k.

(ii) The law of X_0^k given $\mathcal{F}_T$ coincides with the law of Y_0^k given $\mathcal{F}_T$.

6.2. Necessary condition for weak Markovian consistency

Conditions (WM-k) are mathematically interesting, but they are difficult to verify since they entail computations of projections on the filtration $\mathbb{F} \vee \mathbb{F}^{X^k}$. Here we will formulate an "algebraic like" necessary condition for weak Markovian consistency, which is easier to verify.

We start with imposing the following simplifying assumption on process X:

> Assumption (B): For each $k \in \{1, 2, \ldots, N\}$ it holds that
>
> $$\mathbb{P}\left(X_t^k = x^k \middle| \mathcal{F}_t\right) > 0, \quad dt \otimes d\mathbb{P}\text{-a.e.}, \quad \forall x^k \in S_k.$$

Clearly, this assumption imposes constraints on the initial distribution of the chain, as well as constraints on the structure of the intensity process of X. However, it allows to simplify and to streamline the discussion below. The general case can be dealt with in a similar way, with special attention paid to sets of ω-s for which $\mathbb{P}\left(X_t^k = x^k | \mathcal{F}_t\right)(\omega) = 0$.

Before we proceed we observe that Assumption (B) implies that

$$\mathbb{P}\left(X_t^k = x^k\right) > 0, \quad dt\text{-a.e.}, \quad \forall x^k \in S.$$

We will also need a simple technical result regarding events $B(t, k, x^k)$ and $C(t, k, x^k)$ defined, for every $t \in [0, T]$, $x^k \in S_k$ and $k \in \{1, 2, \ldots, N\}$ as

$$B(t, k, x^k) = \left\{\omega : X_t^k(\omega) = x^k\right\}, \quad C(t, k, x^k) = \left\{\omega : \mathbb{P}\left(X_t^k = x^k | \mathcal{F}_t\right)(\omega) > 0\right\}.$$

Observe that (we write B and C in place of $B(t,k,x)$ and $C(t,k,x)$ to shorten the formulae)

$$\mathbb{P}(B\cap C) = \mathbb{E}\left(\mathbb{1}_B \mathbb{1}_C\right) = \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_B|\mathcal{F}_t\right)\mathbb{1}_C\right) = \mathbb{E}\left(\mathbb{P}\left(B|\mathcal{F}_t\right)\mathbb{1}_{\{\mathbb{P}(B|\mathcal{F}_t)>0\}}\right) = \mathbb{P}\left(B\right) > 0.$$

We are now ready to state the main result in this section.

Proposition 6.7. *Assume that X satisfies Assumptions (A) and (B). Fix $k \in \{1,\dots,N\}$. Suppose that X is weakly Markovian consistent relative to $(X^k,\mathbb{F})$. Then, the $\mathbb{F}$-intensity Λ^k of X^k defined by* (6.3) *satisfies*

$$\lambda_t^{k;x^k y^k}(\omega) = \sum_{\substack{x^n,y^n\in S_n\\ n=1,2,\dots,N, n\neq k}} \lambda_t^{(x^1,\dots,x^N)(y^1,\dots,y^N)}(\omega)\frac{\mathbb{P}\left(X_t^1 = x^1,\dots,X_t^N = x^N|\mathcal{F}_t\right)(\omega)}{\mathbb{P}\left(X_t^k = x^k|\mathcal{F}_t\right)(\omega)},$$
$$\forall x^k, y^k \in S_k,\ y^k \neq x^k \text{ and } \forall \omega \in B(t,k,x^k)\cap C(t,k,x^k), \tag{6.7}$$

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

Proof. Since weak Markovian consistency relative to $(X^k,\mathbb{F})$ holds, then Λ^k satisfies (6.2). Taking conditional expectations in (6.2), with respect to $\mathcal{F}_t \vee \sigma(X_t^k)$, yields

$$\begin{aligned}
\mathbb{1}_{\{X_t^k=x^k\}}\lambda_t^{k;x^k y^k} &= \mathbb{E}\left(\mathbb{1}_{\{X_t^k=x^k\}}\lambda_t^{k;x^k y^k}|\mathcal{F}_t\vee\sigma(X_t^k)\right)\\
&= \mathbb{E}\Bigg(\mathbb{1}_{\{X_t^k=x^k\}}\sum_{\substack{x^n,y^n\in S_n\\ n=1,2,\dots,N,n\neq k}}\lambda_t^{(x^1,\dots,x^N)(y^1,\dots,y^N)}\\
&\qquad\times\mathbb{E}\left(\mathbb{1}_{\{X_t^1=x^1,\dots,X_t^{k-1}=x^{k-1},X_t^{k+1}=x^{k+1},\dots,X_t^N=x^N\}}|\mathcal{F}_t\vee\mathcal{F}_t^{X^k}\right)\Bigg|\mathcal{F}_t\vee\sigma(X_t^k)\Bigg)\\
&= \mathbb{1}_{\{X_t^k=x^k\}}\sum_{\substack{x^n,y^n\in S_n\\ n=1,2,\dots,N,n\neq k}}\lambda_t^{(x^1,\dots,x^N)(y^1,\dots,y^N)}\mathbb{E}\left(\mathbb{1}_{\{X_t^1=x^1,\dots,X_t^{k-1}=x^{k-1},X_t^{k+1}=x^{k+1},\dots,X_t^N=x^N\}}|\mathcal{F}_t\vee\sigma(X_t^k)\right).
\end{aligned}$$

Now, let us take an arbitrary $\omega \in B(t,k,x^k)\cap C(t,k,x^k)$. By Assumption (B), using Jakubowski and Niewęgłowski (Jakubowski and Niewęgłowski, 2008, Lemma 3), we have

$$
\begin{aligned}
&\lambda_t^{k;x^k y^k}(\omega) \\
&= \sum_{\substack{x^n, y^n \in S_n \\ n=1,2,\ldots,N, n\neq k}} \lambda_t^{(x^1,\ldots,x^N)(y^1,\ldots,y^N)}(\omega) \mathbb{E}\left(\mathbb{1}_{\{X_t^1=x^1,\ldots,X_t^{k-1}=x^{k-1},X_t^{k+1}=x^{k+1},\ldots,X_t^N=x^N\}} | \mathcal{F}_t \vee \sigma(X_t^k) \right)(\omega) \\
&= \sum_{\substack{x^n, y^n \in S_n \\ n=1,2,\ldots,N, n\neq k}} \lambda_t^{(x^1,\ldots,x^N)(y^1,\ldots,y^N)}(\omega) \frac{\mathbb{P}\left(X_t^1 = x^1, \ldots, X_t^N = x^N | \mathcal{F}_t\right)(\omega)}{\mathbb{P}\left(X_t^k = x^k | \mathcal{F}_t\right)(\omega)},
\end{aligned}
$$

which shows that condition (6.7) is necessary for the weak Markovian consistency of X relative to $(X^k, \mathbb{F})$. □

The next proposition can be used in construction of weak CMC copulae.

Proposition 6.8. *Let $\mathcal{Y} = \{Y^1, \ldots, Y^N\}$ be a family of processes such that each Y^k is an $(\mathbb{F}, \mathbb{F}^{Y^k})$-CDMC, with values in S_k, and with an $\mathbb{F}$-intensity $\Psi_t^k = [\psi_t^{k;x^k y^k}]_{x^k, y^k \in S_k}$. Assume that X satisfies Assumptions (A) and (B), and let Λ be a version of its $\mathbb{F}$-intensity. In addition, suppose that X is weakly Markovian consistent relative to $(\mathbb{F}, \mathcal{Y})$. Then,*

(i) For almost every $t \in [0, T]$ we have

$$
\begin{aligned}
\psi_t^{k;x^k y^k}(\omega) = &\sum_{\substack{x^n, y^n \in S_n \\ n=1,2,\ldots,N, n\neq k}} \lambda_t^{(x^1,\ldots,x^N)(y^1,\ldots,y^N)}(\omega) \frac{\mathbb{P}\left(X_t^1 = x^1, \ldots, X_t^N = x^N | \mathcal{F}_t\right)(\omega)}{\mathbb{P}\left(X_t^k = x^k | \mathcal{F}_t\right)(\omega)}, \\
&\forall x^k, y^k \in S_k,\ y^k \neq x^k \text{ and } \forall \omega \in B(t,k,x^k) \cap C(t,k,x^k).
\end{aligned}
\tag{6.8}
$$

(ii) The law of X_0^k given $\mathcal{F}_T$ coincides with the law of Y_0^k given $\mathcal{F}_T$.

Proof. Since X is weakly Markovian consistent relative to $(\mathbb{F}, \mathcal{Y})$, then, X is weakly Markovian consistent relative to $(X^k, \mathbb{F})$ for each k. Thus, in view of (6.7) and Lemma 9.6 we conclude that (6.8) holds. This proves (i). The conclusion (ii) is clear from the weak Markovian consistency of X relative to $(\mathbb{F}, \mathcal{Y})$. □

Remark 6.9. Even though the above proposition gives a necessary, rather than a sufficient, condition for the weak Markovian consistency of X relative to $(\mathbb{F}, \mathcal{Y})$, it will be skillfully used in construction of weak CMC copulae, in

Section 7.3. In the present time we do not have a workable sufficient condition for the weak Markovian consistency of X relative to $(\mathbb{F}, \mathcal{Y})$ to hold. Thus, for the time being, our strategy for constructing CMC copulae will be to use the necessary condition (6.8) to construct process X which is a candidate for a CMC copula, and then to verify that this process indeed furnishes a weak CMC copula. We refer to Section 7.3 for details.

6.3. When does weak Markov consistency imply strong Markov consistency?

It is clear that the strong Markovian consistency for X implies the weak Markovian consistency for X. As it will be seen in Section 7.3.1, process X may be weakly Markovian consistent relative to $(X^k, \mathbb{F})$, but may fail to satisfy the strong Markovian consistency condition relative to $(X^k, \mathbb{F})$. The following result provides sufficient conditions under which the weak Markovian consistency of X relative to $(X^k, \mathbb{F})$ implies the strong Markovian consistency relative to $(X^k, \mathbb{F})$ for process X.

Theorem 6.10. *Assume that X satisfies the weak Markovian consistency condition relative to $(X^k, \mathbb{F})$. If $\mathbb{F} \vee \mathbb{F}^{X^k}$ is $\mathbb{P}$-immersed in $\mathbb{F} \vee \mathbb{F}^X$, then X satisfies the strong Markovian consistency condition relative to $(X^k, \mathbb{F})$.*

Proof. Suppose that $\mathbb{F} \vee \mathbb{F}^{X^k}$ is immersed in $\mathbb{F} \vee \mathbb{F}^X$. Fix arbitrary $x_1^k, \ldots, x_m^k \in S_k$ and $0 \leq t_1 \leq \ldots \leq t_m \leq T$. Let $A = \left\{X_{t_m}^k = x_m^k, \ldots, X_{t_1}^k = x_1^k\right\}$.

X^k is an $(\mathbb{F}, \mathbb{F}^{X^k})$-CMC, so we have, for $s \leq t_1$,

$$\mathbb{P}(A|\mathcal{F}_s \vee \sigma(X_s^k)) = \mathbb{P}(A|\mathcal{F}_s \vee \mathcal{F}_s^{X^k}) = \mathbb{P}(A|\mathcal{F}_s \vee \mathcal{F}_s^X),$$

where in the second equality we have used immersion of $\mathbb{F} \vee \mathbb{F}^{X^k}$ in $\mathbb{F} \vee \mathbb{F}^X$ (cf. Section 6.1.1 in Bielecki and Rutkowski Bielecki and Rutkowski (2002)). Thus X^k is an $(\mathbb{F}, \mathbb{F}^X)$-CMC. □

Remark 6.11. We note that the above theorem states only a sufficient condition for the weak Markovian consistency of X to imply the strong Markovian consistency of X (relative to $(X^k, \mathbb{F})$). As it is shown in (Bielecki et al., 2013d, Theorem 1.17), in case of trivial filtration $\mathbb{F}$, the condition that $\mathbb{F}^{X^k}$ is immersed in $\mathbb{F}^X$ is both sufficient and necessary for weak Markovian consistency of X to imply the strong Markovian consistency of X (relative to X^k).

7. CMC copulae

As mentioned in the Introduction, the objective of the theory and practice of Markov copulae for classical Markov chains was to construct a non-trivial family of multivariate Markov chains such that components of each chain in the family are Markov chains (in some relevant filtrations) with given laws. Here, our goal is to extend the theory of Markov copulae from the universe of classical (finite) Markov chains to the universe of (finite) conditional Markov chains. Accordingly, we now use the term CMC copulae. As it turns out such extension is not a trivial one. But, it is quite important both from the mathematical point of view and from the practical point of view.

We will first discuss the so called strong CMC copulae, and then we will study the concept of the weak CMC copulae. It needs to be stressed, that an important role in applications is played by the so called weak only CMC copulae, that is weak CMC copulae that are not strong CMC copulae (see discussion in (Bielecki et al., 2013d, Remark 2.3)). An example of such CMC copula will be given in Section 7.3.1.

We recall that in this paper the state space S of process $X = (X^1, \ldots, X^N)$ is given as the Cartesian product $S_1 \times S_2 \times \ldots \times S_N$.

7.1. Strong CMC copulae

Definition 7.1. Let $\mathcal{Y} = \{Y^1, \ldots, Y^N\}$ be a family of processes, defined on some underlying probability space $(\Omega, \mathcal{A}, \mathbb{Q})$, such that each Y^k is an $(\mathbb{F}, \mathbb{F}^{Y^k})$-CMC with values in S_k. A *strong CMC copula* between processes $Y^1, \ldots, Y^N$ is any multivariate process $X = (X^1, \ldots, X^N)$, given on $(\Omega, \mathcal{A})$ endowed with some probability measure $\mathbb{P}$, such that X is an $(\mathbb{F}, \mathbb{F}^X)$–CMC, and such that it satisfies the strong Markovian consistency property with respect to $(\mathbb{F}, \mathcal{Y})$.

The methodology developed in Section 3 allows us to construct strong CMC copulae between processes $Y^1, \ldots, Y^N$, that are defined on some underlying probability space $(\Omega, \mathcal{A}, \mathbb{Q})$ endowed with a reference filtration $\mathbb{F}$, and are such that each Y^k is $(\mathbb{F}, \mathbb{F}^{Y^k})$-CDMC with $\mathbb{F}$–intensity, say, $\Psi^k = [\psi^{k;x^k y^k}]_{x^k, y^k \in S_k}$. The additional feature of our construction is that, typically, the constructed CMC copulae X are also $(\mathbb{F}, \mathbb{F}^X)$-DSMC.

In view of Theorem 3.3, Proposition 5.9 and Lemma 9.6 a natural starting point for constructing a strong copula between $Y^1, \ldots, Y^N$ is to determine a system of stochastic processes $[\lambda^{xy}]_{x,y \in S}$ and an S-valued random variable $\xi = (\xi^1, \ldots, \xi^N)$ on $(\Omega, \mathcal{A})$, such that they satisfy the following conditions:

(CMC-1)

$$\psi_t^{k;x^k y^k} = \sum_{\substack{y^n \in S_n,\\ n=1,2,\dots,N, n\neq k}} \lambda_t^{(x^1,\dots,x^k,\dots,x^N)(y^1,\dots,y^k,\dots,y^N)}, \qquad \begin{array}{l} x^n \in S_n, n = 1,\dots,N, \\ y^k \in S_k, y^k \neq x^k, \\ k = 1,\dots,N, t \in [0,T]. \end{array}$$

(CMC-2) The matrix process $\Lambda_t = [\lambda_t^{xy}]_{x,y\in S}$ satisfies canonical conditions relative to the pair $(S,\mathbb{F})$ (cf. Definition 3.2).

(CMC-3)

$$\mathbb{Q}(\xi = y|\mathcal{F}_T) = \mathbb{Q}(\xi = y|\mathcal{F}_0), \quad \forall y \in S.$$

(CMC-4)

$$\mathbb{Q}(\xi^k = y^k|\mathcal{F}_T) = \mathbb{Q}(Y_0^k = y^k|\mathcal{F}_T), \quad \forall y^k \in S_k, k = 1,\dots,N.$$

We will call any pair (Λ,ξ) satisfying conditions (CMC-1)–(CMC-4) *strong CMC pre-copula* between processes $Y^1,\dots,Y^N$. Given a strong CMC pre-copula between processes $Y^1,\dots,Y^N$ we can construct on $(\Omega,\mathcal{A})$ a probability measure $\mathbb{P}$ and a process X, using Theorem 3.3 and starting from measure $\mathbb{Q}$ as above[14], such that X is an $(\mathbb{F},\mathbb{F}^X)$–CMC under $\mathbb{P}$, and which satisfies the strong Markovian consistency property with respect to $(\mathbb{F},\mathcal{Y})$, in view of Proposition 5.9 and Lemma 9.6.

Thus, it is a strong CMC copula between processes $Y^1,\dots,Y^N$.

Remark 7.2. It follows from (i) and (3.8) in Theorem 3.3 that for $\mathbb{P}$ constructed as above we have

$$\mathbb{P}(\xi = y|\mathcal{F}_T) = \mathbb{P}(\xi = y|\mathcal{F}_0), \quad \forall y \in S.$$
$$\mathbb{P}(\xi^k = y^k|\mathcal{F}_T) = \mathbb{Q}(Y_0^k = y^k|\mathcal{F}_T), \quad \forall y^k \in S_k, k = 1,\dots,N.$$

Remark 7.3. (i) Note that in the definition of strong CMC copula it is required that $\mathcal{F}_T$-conditional distribution of X_0^k coincides with $\mathcal{F}_T$-conditional distribution of Y_0^k, for $k \in 1,\dots,N$, but the $\mathcal{F}_T$-conditional distribution of the multivariate random variable $X_0 = (X_0^1,\dots,X_0^N)$ can be arbitrary. Thus, in principle, a strong CMC copula between processes $Y^1,\dots,Y^N$ can

[14]It is always tacitly assumed that the probability space $(\Omega,\mathcal{A},\mathbb{Q})$ is sufficiently rich so to support all stochastic processes and random variables that are considered throughout.

be constructed with help of a strong CMC pre-copula between processes $Y^1, \ldots, Y^N$, as well as a copula between the $\mathcal{F}_T$-conditional distributions of X_0^ks, for $k \in 1, \ldots, N$. For instance, in Example 7.2.1 below, we take the components $X^1, \ldots, X^N$ are conditionally independent given $\mathcal{F}_T$.
(ii) In general, there exist numerous systems of stochastic processes that satisfy conditions (CMC-1) and (CMC-2), so that there exist numerous strong pre-copulae between conditional Markov chains $Y^1, \ldots, Y^N$, and, consequently, there exists numerous strong CMC copulae between conditional Markov chains $Y^1, \ldots, Y^N$. This is an important feature in financial applications, as it allows to calibrate a CMC model to both marginal data and to the basket data.

Below we provide examples of strong CMC copulae. The first example, dealing with conditionally independent univariate CMCs, does not really address the issue of modeling dependence between components of a multivariate CMC. Nevertheless, on one hand, this example may have non-trivial practical applications in insurance, and, on the other hand, it is a non-trivial example from the mathematical point of view. Moreover, this example provides a sort of a reality check for the theory of strong CMC copulae: it would be not good for the theory if a multivariate conditional Markov chain $X = (X^1, \ldots, X^N)$ with conditionally independent components would not be a strong CMC copula.

7.2. Examples

7.2.1. Conditionally independent strong CMC copula

This example is a counterpart of the independent Markov copula example presented in Section 2.1 in Bielecki et al. (2013d). Let $Y^1, \ldots, Y^N$ be processes such that each Y^k is an $(\mathbb{F}, \mathbb{F}^{Y^k})$-CDMC with values in S_k, and with $\mathbb{F}$–intensity $\Psi_t^k = [\psi_t^{k;x^k y^k}]_{x^k, y^k \in S_k}$. Assume that for each k the process Ψ^k satisfies canonical conditions relative to the pair $(S_k, \mathbb{F})$. Additionally assume that

$$\mathbb{Q}(Y_0^k = x^k | \mathcal{F}_T) = \mathbb{Q}(Y_0^k = x^k | \mathcal{F}_0), \quad \forall x^k \in S_k, \ k = 1, \ldots, N. \tag{7.1}$$

Consider a matrix valued random process Λ given as the following Kronecker sum

$$\Lambda_t = \sum_{k=1}^N I_1 \otimes \ldots \otimes I_{k-1} \otimes \Psi_t^k \otimes I_{k+1} \otimes \ldots \otimes I_N, \quad t \in [0, T], \tag{7.2}$$

where $\otimes$ is the Kronecker product (see e.g. Horn and Johnson Horn and Johnson (1994)), and where I_k denotes the identity matrix of dimensions $|S_k| \times |S_k|$. Moreover, let us take an S-valued random variable $\xi = (\xi^1, \dots, \xi^N)$, which has $\mathcal{F}_T$-conditionally independent coordinates, that is

$$\mathbb{Q}(\xi^1 = x^1, \dots, \xi^N = x^N | \mathcal{F}_T) = \prod_{i=1}^{N} \mathbb{Q}(\xi^i = x^i | \mathcal{F}_T), \quad \forall x = (x^1, \dots, x^N) \in S. \tag{7.3}$$

Additionally assume that $\mathcal{F}_T$-conditional distributions of coordinates of ξ and Y_0 coincide:

$$\mathbb{Q}(\xi^k = x^k | \mathcal{F}_T) = \mathbb{Q}(Y_0^k = x^k | \mathcal{F}_T), \quad \forall x^k \in S_k, \ k = 1, \dots, N. \tag{7.4}$$

As shown in (Bielecki et al., 2015b, Proposition 6.2), Λ satisfies conditions (CMC-1) and (CMC-2). Furthermore, by (7.3) and (7.1), ξ satisfies (CMC-3) and, by (7.4), also (CMC-4). Thus, (Λ, ξ) is a strong CMC pre-copula between conditional Markov chains Y^1, ..., Y^N. Now, we can construct, with the help of Theorem 3.3 , a multivariate $(\mathbb{F}, \mathbb{F}^X)$-CDMC (see also Proposition 4.16), say $X = (X^1, \dots, X^N)$, with values in S, which in view of Proposition 5.9 satisfies the strong Markovian consistency property with respect to $(\mathbb{F}, \mathcal{Y})$. Therefore, the process X furnishes a strong CMC copula between processes $Y^1, \dots, Y^N$. Finally, (Bielecki et al., 2015b, Proposition 6.4) demonstrates that components of X are conditionally independent given $\mathcal{F}_T$. It is quite clear from (7.2) that components X^i of X do not jump simultaneously; this, indeed, is the inherent feature of the conditional independent CMC copula.

Next, we will present an example of a strong CMC copula such that its components have common jumps.

7.2.2. Common jump strong CMC copula

Let us consider two processes, Y^1 and Y^2, such that each Y^i is an $(\mathbb{F}, \mathbb{F}^{Y^i})$-CDMC taking values in the state space $\{0, 1\}$. Suppose that their $\mathbb{F}$-intensities are

$$\Psi^1(t) = \begin{matrix} & \begin{matrix} 0 & 1 \end{matrix} \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{pmatrix} -a_t & a_t \\ 0 & 0 \end{pmatrix} \end{matrix}, \qquad \Psi^2(t) = \begin{matrix} & \begin{matrix} 0 & 1 \end{matrix} \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{pmatrix} -b_t & b_t \\ 0 & 0 \end{pmatrix} \end{matrix},$$

where a, b are nonnegative $\mathbb{F}$-progressively measurable stochastic processes, which have left limits and countably many jumps. Moreover assume that

$\mathbb{Q}(Y_0^1 = 0) = \mathbb{Q}(Y_0^2 = 0) = 1$. Next, let Λ be a matrix valued process given by

$$\Lambda_t = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -(a_t + b_t - c_t) & b_t - c_t & a_t - c_t & c_t \\ 0 & -a_t & 0 & a_t \\ 0 & 0 & -b_t & b_t \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array},$$

where c is an $\mathbb{F}$-progressively measurable stochastic processes, which has left limits and countably many jumps, and such that

$$0 \leq c_t \leq a_t \wedge b_t, \quad t \in [0, T].$$

Moreover, let ξ be an S-valued random variable satisfying $\mathbb{Q}(\xi = (0,0)) = 1$. It can be easily checked that (Λ, ξ) satisfies conditions (CMC-1)-(CMC-4), so that it is a strong CMC pre-copula between conditional Markov chains Y^1, Y^2. Now, in view of Theorem 3.3 and Proposition 4.16, one can construct a stochastic process $X = (X^1, X^2)$, which is a two-variate $(\mathbb{F}, \mathbb{F}^X)$-CDMC with an $\mathbb{F}$-intensity Λ and such that $X_0 = \xi$. Moreover, by Proposition 5.9, the process X is strongly Markovian consistent with respect to $(\mathbb{F}, \mathcal{Y})$ and hence X is a strong CMC copula between Y^1 and Y^2. Note also that, in view of interpretation of intensity, the coordinates of the process X have common jumps, provided that $c > 0$.

Remark 7.4. We have chosen this very simple example just to illustrate an idea of construction of strong copulae for CMC. One can, in a similar spirit as in Bielecki et al. (2014d), generalize it to arbitrary dimension N preserving that each marginal process is two-states absorbing CMC. Then the $\mathbb{F}$-intensity matrix has a similar structure as in the above example, i.e. its entries are marginal intensities minus intensities of "common jumps" to absorbing states. Generalization to a higher number of non-absorbing states is tricky and requires clever parametrization, since number of free parameters in strong CMC copula becomes enormously large (see, e.g. Bielecki et al. (2008c)).

7.2.3. Perfect dependence strong CMC copula

Let $Y^1, \ldots, Y^N$ be processes such that each Y^k is an $(\mathbb{F}, \mathbb{F}^{Y^k})$-CMC, and such that they have the same $\mathcal{F}_T$ conditional laws. Consider process $X = (X^1, \ldots, X^N)$, where $X^k = Y^1, \quad k = 1, 2, \ldots, N$. It is clear that X

furnishes a strong CMC copula between conditional Markov chains $Y^1, \ldots, Y^N$.

Obviously other CMC copulae between $Y^1, \ldots, Y^N$, such as conditionally independent copulae, can be constructed.

7.3. Weak CMC Copulae

Definition 7.5. Let $\mathcal{Y} = \{Y^1, \ldots, Y^N\}$ be a family of processes, defined on some underlying probability space $(\Omega, \mathcal{A}, \mathbb{Q})$, such that each Y^k is an $(\mathbb{F}, \mathbb{F}^{Y^k})$-CMC with values in S_k. A *weak CMC copula* between processes $Y^1, \ldots, Y^N$ is any multivariate process $X = (X^1, \ldots, X^N)$, defined on $(\Omega, \mathcal{A})$ endowed with some probability measure $\mathbb{P}$, such that X is an $(\mathbb{F}, \mathbb{F}^X)$–CMC, and such that it satisfies the weak Markovian consistency property with respect to $(\mathbb{F}, \mathcal{Y})$.

Similarly as in the case of the strong CMC copulae, the methodology developed in Section 3 allows us to construct weak CMC copulae between processes $Y^1, \ldots, Y^N$, that are defined on some underlying probability space $(\Omega, \mathcal{A}, \mathbb{Q})$ endowed with a reference filtration $\mathbb{F}$, and are such that each Y^k is $(\mathbb{F}, \mathbb{F}^{Y^k})$-CDMC with $\mathbb{F}$–intensity, say, $\Psi^k = [\psi^{k;x^k y^k}]_{x^k, y^k \in S_k}$.

In view of Theorem 3.3, Proposition 6.8, Lemma 9.6, as well as of Remark 6.9, a natural starting point for constructing a weak CMC copula X between $Y^1, \ldots, Y^N$ is to determine any system of stochastic processes $(\lambda^{xy})_{x,y \in S}$ and any S-valued random variable $\xi = (\xi^1, \ldots, \xi^N)$ on $(\Omega, \mathcal{A})$ and to find a probability measure $\mathbb{P}$, such that the following conditions are satisfied:

(WCMC-1) The matrix process $\Lambda_t = [\lambda_t^{xy}]_{x,y \in S}$ satisfies canonical conditions relative to the pair $(S, \mathbb{F})$.

(WCMC-2)

$$\mathbb{P}(\xi = y | \mathcal{F}_T) = \mathbb{P}(\xi = y | \mathcal{F}_0), \quad \forall y \in S.$$

(WCMC-3)

$$\mathbb{P}(\xi^k = y^k | \mathcal{F}_T) = \mathbb{Q}(Y_0^k = y^k | \mathcal{F}_T), \quad \forall y^k \in S_k, k = 1, \ldots, N.$$

(WCMC-4)

$$\psi_t^{k;x^ky^k}(\omega) = \sum_{\substack{x^n,y^n\in S_n\\ n=1,2,\dots,N,n\neq k}} \lambda_t^{(x^1,\dots,x^N)(y^1,\dots,y^N)}(\omega)\frac{\mathbb{P}\left(X_t^1=x^1,\dots,X_t^N=x^N|\mathcal{F}_t\right)(\omega)}{\mathbb{P}\left(X_t^k=x^k|\mathcal{F}_t\right)(\omega)},$$

$$\forall x^k, y^k \in S_k,\ y^k\neq x^k \text{ and } \forall\omega\in B(t,k,x^k)\cap C(t,k,x^k), \tag{7.5}$$

for almost every $t\in[0,T]$, where process $X=(X^1,\dots,X^N)$ is a $(\mathbb{F},\mathbb{F}^X)$–CMC (under probability measure $\mathbb{P}$) with intensity Λ and initial distribution given by ξ.

We will call any triple (Λ,ξ,X) satisfying conditions (WCMC-1)–(WCMC4) *a base for weak CMC copula* between processes $Y^1,\dots,Y^N$, and we will call the process X in (Λ,ξ,X) *a candidate for weak CMC copula* between processes $Y^1,\dots,Y^N$. So, a possible method for constructing a weak CMC copula between processes $Y^1,\dots,Y^N$ is to first construct a base (Λ,ξ,X) for weak CMC copula between processes $Y^1,\dots,Y^N$, and then to skillfully verify that the candidate process X satisfies the weak Markovian consistency property with respect to $(\mathbb{F},\mathcal{Y})$, and thus, that it is a weak CMC copula between processes $Y^1,\dots,Y^N$. We will illustrate application of this method in Section 7.3.1. Before we proceed to the next subsection, we observe that remark analogous to Remark 7.3 applies in the case of the weak CMC copulae.

The next section provides an example of a weak only CMC copula.

7.3.1. Example of a weak CMC copula that is not strong CMC copula

In Section 7.2 we gave three examples of strong CMC copulae. Consequently, they are also examples of weak CMC copulae. Here, we will give an example of a weak only CMC copula. In particular, this property implies that in the present example the immersion property formulated in Theorem 6.10 is not satisfied.

Let us consider processes Y^1 and Y^2, defined on some probability space $(\Omega,\mathcal{A},\mathbb{Q})$, such that each Y^i is an $(\mathbb{F},\mathbb{F}^{Y^i})$-CDMC taking values in the state space $S_i=\{0,1\}$. We assume that $\mathbb{F}$-intensities of Y^1 and Y^2 are, respectively,

$$\Psi_t^1 = \begin{pmatrix} -(a_t+c_t)+c_t\frac{\alpha_t}{\delta_t+\alpha_t} & (a_t+c_t)-c_t\frac{\alpha_t}{\delta_t+\alpha_t} \\ 0 & 0 \end{pmatrix},$$

$$\Psi_t^2 = \begin{pmatrix} -(b_t + c_t) + c_t\frac{\beta_t}{\delta_t+\beta_t} & (b_t + c_t) - c_t\frac{\beta_t}{\delta_t+\beta_t} \\ 0 & 0 \end{pmatrix},$$

where

$$\alpha_t = e^{-\int_0^t a_u du} \int_0^t b_u e^{-\int_0^u (b_v+c_v)dv} du, \qquad \beta_t = e^{-\int_0^t b_u du} \int_0^t a_u e^{-\int_0^u (a_v+c_v)dv} du,$$

$$\delta_t = e^{-\int_0^t (a_u+b_u+c_u)du},$$

for a, b, c being positive $\mathbb{F}$-progressive stochastic processes, which have left limits and countably many jumps. Moreover, suppose that $\mathbb{Q}(Y_0^i = 0) = 1,\ i = 1, 2$, which implies $\mathbb{Q}(Y_0^i = 0|\mathcal{F}_T) = 1$.

Our goal is to find a weak CMC copula between Y^1 and Y^2. Towards this end we will look for an $(\mathbb{F}, \mathbb{F}^X)$-CMC process X defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, satisfying condition (7.5) adapted to the present setup. In particular, the state space of process X needs to be equal to $S = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$.

However, since condition (7.5) is a necessary condition for weak Markovian consistency with respect to $(\mathbb{F}, \mathcal{Y})$, but not a sufficient one in general, then a process satisfying (7.5) may not be weakly Markovian consistent with respect to $(\mathbb{F}, \mathcal{Y})$. Nevertheless, we will construct an $(\mathbb{F}, \mathbb{F}^X)$-CMC process X that satisfies condition (7.5) and is weakly Markovian consistent with respect to $(\mathbb{F}, \mathcal{Y})$, so that it is a weak CMC copula between Y^1 and Y^2.

Let us consider stochastic process X with state space S, which is an $(\mathbb{F}, \mathbb{F}^X)$-CDMC with an $\mathbb{F}$-intensity matrix Λ given by

$$\Lambda_t = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -(a_t + b_t + c_t) & b_t & a_t & c_t \\ 0 & -a_t & 0 & a_t \\ 0 & 0 & -b_t & b_t \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}, \tag{7.6}$$

and with the initial distribution $\mathbb{P}(X_0 = (0, 0)) = 1$. The components X^1 and X^2 are processes with state space $S = \{0, 1\}$, and such that the state 1 is an absorbing state for both X^1 and X^2. Thus, by similar arguments as in (Bielecki et al., 2015, Example 2.4), X^1 (resp. X^2) is an $(\mathbb{F}, \mathbb{F}^{X^1})$-CDMC (resp. $(\mathbb{F}, \mathbb{F}^{X^2})$-CDMC). Consequently, X is a weakly Markovian consistent process relative to $(X^1, \mathbb{F})$ $((X^2, \mathbb{F})$ resp.).

It is shown in Bielecki et al. (2015b) that a version of $\mathbb{F}$-intensity of X^1 is given by

$$\Lambda_t^1 = \begin{pmatrix} -(a_t + c_t) + c_t\frac{\alpha_t}{\delta_t+\alpha_t} & (a_t + c_t) - c_t\frac{\alpha_t}{\delta_t+\alpha_t} \\ 0 & 0 \end{pmatrix} = \Psi_t^1,$$

and that a version of $\mathbb{F}$-intensity of X^2 is given by

$$\Lambda_t^2 = \begin{pmatrix} -(b_t + c_t) + c_t\frac{\beta_t}{\delta_t+\beta_t} & (b_t + c_t) - c_t\frac{\beta_t}{\delta_t+\beta_t} \\ 0 & 0 \end{pmatrix} = \Psi_t^2.$$

Consequently, X is a weak CMC copula for Y^1 and Y^2.

Finally, we will demonstrate that X is in fact weak only CMC copula for Y^1 and Y^2. We have

$$\begin{aligned}
&\mathbb{P}(X_t^1 = 0|\mathcal{F}_T \vee \sigma(X_s))\mathbb{1}_{\{X_s^1=0,X_s^2=0\}} = \mathbb{1}_{\{X_s^1=0,X_s^2=0\}}(p_{(0,0)(0,0)}(s,t) + p_{(0,0)(0,1)}(s,t)) \\
&= \mathbb{1}_{\{X_s^1=0,X_s^2=0\}}\left(e^{-\int_s^t(a_u+b_u+c_u)du} + e^{-\int_s^t a_u du}\int_s^t b_u e^{-\int_s^u(b_v+c_v)dv}du\right),
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{P}(X_t^1 = 0|\mathcal{F}_T \vee \sigma(X_s))\mathbb{1}_{\{X_s^1=0,X_s^2=1\}} &= \mathbb{1}_{\{X_s^1=0,X_s^2=1\}}(p_{(0,1)(0,0)}(s,t) + p_{(0,1)(0,1)}(s,t)) \\
&= \mathbb{1}_{\{X_s^1=0,X_s^2=1\}}e^{-\int_s^t a_u du}.
\end{aligned}$$

Clearly

$$\left(e^{-\int_s^t(a_u+b_u+c_u)du} + e^{-\int_s^t a_u du}\int_s^t b_u e^{-\int_s^u(b_v+c_v)dv}du\right) \neq e^{-\int_s^t a_u du}, \tag{7.7}$$

unless $c \equiv 0$ on $[s,t]$. In this case (7.7) implies that

$$\mathbb{P}\big(\mathbb{P}(X_t^1 = 0|\mathcal{F}_s \vee \mathcal{F}_s^X) \neq \mathbb{P}(X_t^1 = 0|\mathcal{F}_t \vee \sigma(X_s^1))\big) > 0.$$

Thus process X is not strongly Markovian consistent, so X is a weak only CMC copula between Y^1 and Y^2 unless $c \equiv 0$. For $c \equiv 0$, it follows from Section 7.2.2, that process X is a strong CMC copula between Y^1 and Y^2 .

Remark 7.6. Note that Λ_t admits the following representation

$$\Lambda_t = \Psi_t^1 \otimes I_2 + I_1 \otimes \Psi^2(t) + B_t^{12} - B_t^1 - B_t^2, \tag{7.8}$$

where the term $\Psi_t^1 \otimes I_2 + I_1 \otimes \Psi_t^2$ gives the conditionally independent copula between Y^1 and Y^2 (cf. Section 7.2.1), and the remaining terms

$$B_t^{12} = \begin{pmatrix} -c_t & 0 & 0 & c_t \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B_t^1 = \begin{pmatrix} -c_t\frac{\delta_t}{\delta_t+\beta_t} & c_t\frac{\delta_t}{\delta_t+\beta_t} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -c_t\frac{\delta_t}{\delta_t+\beta_t} & c_t\frac{\delta_t}{\delta_t+\beta_t} \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$B_t^2 = \begin{pmatrix} -c_t\frac{\delta_t}{\delta_t+\alpha_t} & 0 & c_t\frac{\delta_t}{\delta_t+\alpha_t} & 0 \\ 0 & -c_t\frac{\delta_t}{\delta_t+\alpha_t} & 0 & c_t\frac{\delta_t}{\delta_t+\alpha_t} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

introduce the dependence structure between Y^1 and Y^2.

Representations of the form (7.8) are important for construction of CMC copulae and will be studied in detail in Bielecki, Jakubowski and Niewęgłowski Bielecki et al. (????).

8. Applications to the premium evaluation for unemployment insurance products

In the recent paper by Biagini, Groll and Widenmann Biagini et al. (2013) a very interesting problem of evaluation of premia for unemployment insurance products, for a pool of individuals, was considered. We would like to suggest here a possible generalization of the model studied in Biagini et al. (2013); this generalization, we believe, may provide a more adequate way to deal with computation of the premia.

Biagini et al. Biagini et al. (2013) used the DSMC framework to model the dynamics of employment status of an individual. The dynamics are modeled in Biagini et al. (2013) under the probability measure, say $\mathbb{P}$, called a real-world measure. Then, using these dynamics they aim at computing for $t \in [0, T]$ the insurance premium, which is denoted as P_t. In Biagini et al. (2013), the evolution of the employment status of an individual k is given in terms of a Markov chain, say X^k, which takes values in the state space $S_k = \{1, 2\}$, where the state "1" indicates that the individual is employed, and the "2" indicates that the individual is unemployed. It is assumed that process

X^k is an $(\mathbb{F}^Z, \mathbb{F}^{X^k})$-DSMC, where $\mathbb{F}^Z$ is a reference filtration generated by some factor process Z.

As stated earlier, the quantity to be computed for the individual k is the value of the premium of insurance against unemployment. Roughly speaking, the premium P_t^k at time t is given as

$$P_t^k = \mathbb{E}_{\mathbb{P}}(\Phi_k(X^k)|\mathcal{G}_t^k),$$

where Φ_k is some random functional of process X^k, and where $\mathcal{G}_t^k = \mathcal{F}_t^Z \vee \mathcal{F}_t^{X^k}$. In particular, the premium at time $t = 0$ needs to be computed, that is

$$P_0^k = \mathbb{E}_{\mathbb{P}}(\Phi_k(X^k)|\mathcal{G}_0^k).$$

Note, that we have written P_0^k as a conditional expectation, given $\mathcal{G}_0^k$, rather than the unconditional expectation, as it is done in formula (2) in Biagini et al. (2013).

Proposed CMC copula approach. We think that, for the purpose of evaluation of premia for unemployment insurance products for a pool of individuals labeled as $k = 1, 2, \ldots, N$, it is important to account for possible dependence between processes $X^k, \ k = 1, 2, \ldots, N$.

Thus, we think that it may be advantageous to enrich the model studied in Biagini et al. (2013) by considering a process $Y = (Y^1, \ldots, Y^N)$, which is a CMC copula between processes $X^k, \ k = 1, 2, \ldots, N$.

Thanks to copula property, the characteristics of dependence between processes $X^k, \ k = 1, 2, \ldots, N$ can be estimated separately from estimation of the distributional characteristics of each process X^k. The latter task can be efficiently executed using the methodology outlined in Biagini et al. (2013).

The premium P_t^k at time t is given in the CMC copula model as

$$P_t^k = \mathbb{E}(\Phi_k(Y^k)|\widehat{\mathcal{G}}_t^k),$$

where $\widehat{\mathcal{G}}_t^k = \mathcal{F}_t^Z \vee \mathcal{F}_t^{Y^k}$. If process Y is constructed as a weak only CMC copula between processes $X^k, \ k = 1, 2, \ldots, N$, then we have that, with $\widehat{\mathcal{G}}_t = \mathcal{F}_t^Z \vee \mathcal{F}_t^Y$,

$$\mathbb{E}(\Phi_k(Y^k)|\widehat{\mathcal{G}}_t^k) \neq \mathbb{E}(\Phi_k(Y^k)|\widehat{\mathcal{G}}_t).$$

This, of course, means that the employment status of the entire pool influences the calculation of the individual premium, a feature, which we think is important.

The theory of strong and weak CMC copulae can be extended in straightforward manner to modeling structured dependence between subgroups of processes $X^k,\ k=1,2,\ldots,N,$. This will allow for study of insurance premia modeling for relevant subgroups of employees.

9. Appendix

In this appendix we provide technical results needed for derivations done in Section 3.

Lemma 9.1. *Let ξ be an S–valued random variable defined on a filtered probability space $(\Omega,\mathcal{A},\mathbb{H},\widetilde{\mathbb{P}})$ with $\mathbb{H}=\{\mathcal{H}_t\}_{t\in[0,T]}$. Suppose that*

$$\mathbb{E}_{\widetilde{\mathbb{P}}}(h(\xi)|\mathcal{H}_T)=\mathbb{E}_{\widetilde{\mathbb{P}}}(h(\xi)|\mathcal{H}_0) \tag{9.1}$$

for every real valued function h on S. Then $\mathbb{H}$ is $\widetilde{\mathbb{P}}$–immersed in $\mathbb{H}\vee\sigma(\xi)$.

Proof. It is sufficient to prove (c.f. (Bielecki and Rutkowski, 2004, Lemma 6.1.1)) that for every $\psi\in L^\infty(\mathcal{H}_T)$ it holds that

$$\mathbb{E}_{\widetilde{\mathbb{P}}}(\psi|\mathcal{H}_t\vee\sigma(\xi))=\mathbb{E}_{\widetilde{\mathbb{P}}}(\psi|\mathcal{H}_t),\quad \forall t\in[0,T]. \tag{9.2}$$

Let us fix $t\in[0,T]$ and $\psi\in L^\infty(\mathcal{H}_T)$. By the standard $\pi-\lambda$ system arguments it is enough to show that

$$\mathbb{E}_{\widetilde{\mathbb{P}}}(\psi\mathbb{1}_A\mathbb{1}_B(\xi))=\mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{E}_{\widetilde{\mathbb{P}}}(\psi|\mathcal{H}_t)\mathbb{1}_A\mathbb{1}_B(\xi)),\quad \forall A\in\mathcal{H}_t, B\subseteq S, \tag{9.3}$$

where

$$\mathbb{1}_B(\xi)=\begin{cases}1, & \xi\in B,\\ 0, & \xi\notin B.\end{cases}$$

Towards this end we first derive another representation of the right hand side in (9.3),

$$\begin{aligned}\mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{E}_{\widetilde{\mathbb{P}}}(\psi|\mathcal{H}_t)\mathbb{1}_A\mathbb{1}_B(\xi)) &= \mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{E}_{\widetilde{\mathbb{P}}}(\psi\mathbb{1}_A|\mathcal{H}_t)\mathbb{1}_B(\xi)) = \mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{E}_{\widetilde{\mathbb{P}}}(\psi\mathbb{1}_A|\mathcal{H}_t)\mathbb{1}_B(\xi)|\mathcal{H}_T))\\ &= \mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{E}_{\widetilde{\mathbb{P}}}(\psi\mathbb{1}_A|\mathcal{H}_t)\mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{1}_B(\xi)|\mathcal{H}_T)) = \mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{E}_{\widetilde{\mathbb{P}}}(\psi\mathbb{1}_A|\mathcal{H}_t)\mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{1}_B(\xi)|\mathcal{H}_0))\\ &= \mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{E}_{\widetilde{\mathbb{P}}}(\psi\mathbb{1}_A\mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{1}_B(\xi)|\mathcal{H}_0)|\mathcal{H}_t)) = \mathbb{E}_{\widetilde{\mathbb{P}}}(\psi\mathbb{1}_A\mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{1}_B(\xi)|\mathcal{H}_0)),\end{aligned}$$

where the fourth equality follows from (9.1). The left hand side of (9.3) can be rewritten as

$$\mathbb{E}_{\widetilde{\mathbb{P}}}(\psi\mathbb{1}_A\mathbb{1}_B(\xi))=\mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{E}_{\widetilde{\mathbb{P}}}(\psi\mathbb{1}_A\mathbb{1}_B(\xi)|\mathcal{H}_T))=\mathbb{E}_{\widetilde{\mathbb{P}}}(\psi\mathbb{1}_A\mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{1}_B(\xi)|\mathcal{H}_T))=\mathbb{E}_{\widetilde{\mathbb{P}}}(\psi\mathbb{1}_A\mathbb{E}_{\widetilde{\mathbb{P}}}(\mathbb{1}_B(\xi)|\mathcal{H}_0)),$$

where the last equality follows from (9.1). This proves (9.3) and thus concludes the proof of the lemma. □

Corollary 9.2. *Let* $\mathbb{K}$ *be a filtration on* $(\Omega, \mathcal{A}, \widetilde{\mathbb{P}})$*, such that it is independent of* $\mathbb{H} \vee \sigma(\xi)$*. Suppose that* ξ *satisfies* (9.1)*. Then* $\mathbb{H}$ *is* $\widetilde{\mathbb{P}}$*–immersed in* $\mathbb{H} \vee \mathbb{K} \vee \sigma(\xi)$*.*

Proof. The result follows from Lemma 9.1 and from the fact that if $\mathbb{H}^1$ and $\mathbb{H}^2$ are two independent filtrations on $(\Omega, \mathcal{A}, \widetilde{\mathbb{P}})$, then $\mathbb{H}^1$ is $\widetilde{\mathbb{P}}$-immersed in $\mathbb{H}^1 \vee \mathbb{H}^2$. □

In the next lemma we use the same probabilistic setup as in Section 2.

Lemma 9.3. *Let* X *be an* $\mathbb{F}$ *adapted càdlàg process, and let* N *be a Poisson process. Suppose that* N *and* $\mathbb{F}$ *are independent. Then*

$$\mathbb{P}\left(\{\omega \in \Omega : \exists t \in [0,T] \ s.t. \ \Delta X_t(\omega)\Delta N_t(\omega) \neq 0\}\right) = 0.$$

Proof. First note that both X and N have countable number of jumps on $[0,T]$, and let denote their jump times as $(T_n)_{n\geq 1}$ and $(S_n)_{n\geq 1}$, respectively. Independence of N and $\mathbb{F}$ implies that $(T_n)_{n\geq 1}$ and $(S_n)_{n\geq 1}$ are independent. Since each random variable S_n is Gamma distributed and thus has density, then for any $n, k \geq 1$ it holds that $\mathbb{P}(T_n = S_k) = 0$. Since

$$A := \{\omega : \exists t \in [0,T] \ s.t. \ \Delta X_t(\omega)\Delta N_t(\omega) \neq 0\} = \bigcup_{n,k\geq 1} \{\omega : T_n(\omega) = S_k(\omega)\}$$

we have

$$\mathbb{P}(A) \leq \sum_{n,k\geq 1} \mathbb{P}(T_n = S_k) = 0. \qquad \square$$

□

Lemma 9.4. *Let* Z *and* Y *be solutions of the random ODE's*

$$dZ_t = -\Psi_t Z_t dt, \qquad Z_0 = \mathrm{I}, \quad t \in [0,T], \tag{9.4}$$

$$dY_t = Y_t \Psi_t dt, \qquad Y_0 = \mathrm{I}, \quad t \in [0,T], \tag{9.5}$$

where Ψ *is an appropriately measurable matrix valued process satisfying* (2.5)[15] *and*

$$\sum_{x\in S} \int_0^T |\psi_u^{xx}| du < \infty. \tag{9.6}$$

[15]For any ω for which Ψ does not satisfy (2.5), we set $\Psi_t(\omega) = 0$ for all $t \in [0,T]$.

Then, the matrix valued random processes $(Y_t)_{0\le t\le T}$ *and* $(Z_tY_v)_{0\le t\le v}$, $v \in [0,T]$, *have elements that are nonnegative and bounded by* 1. *Moreover*

$$Z_tY_t = \mathrm{I} \quad \textit{for } t \in [0,T]. \tag{9.7}$$

Proof. Using Remark 4.8 one can verify that for each t the functions $Y_t(\cdot)$ and $Z_t(\cdot)$ are measurable, so that Y and Z are matrix valued random processes.

Since Ψ satisfies (2.5), then for every ω, $Y_{\cdot}(\omega)$ is a solution of matrix forward Kolmogorov equation, and so its elements belong to the interval $[0,1]$ (since they give conditional probabilities, see e.g. Gill and Johansen (Gill and Johansen, 1990, Thm. 12 and Thm. 13)).

Next, observe that, letting $Z(t,v) = Z_tY_v$ we have

$$d_tZ(t,v) = (dZ_t)Y_v = -\Psi_tZ_tY_vdt = -\Psi_tZ(t,v)dt, \quad 0 \le t \le v.$$

Moreover, it is easy to verify that $Z(v,v) = Z_vY_v = Z_0Y_0 = \mathrm{I}$. We thus see that for every ω, $Z(\cdot,v)(\omega)$ satisfies the Kolmogorov backward equation,

$$d_tZ(t,v) = -\Psi_tZ(t,v)dt, \quad 0 \le t \le v, \quad Z(v,v) = \mathrm{I},$$

and so, it has non-negative elements bounded by 1. □

The following lemma is used in the proof of Theorem 2.11.

Lemma 9.5. *Suppose that assumptions of Theorem 2.11 are satisfied. Let* U *be an* $\mathbb{R}^d$*-valued bounded random variable, and let* Z *and* Y *be solutions of the random ODE's* (9.4) *and* (9.5), *respectively. Fix* u *and* v *satisfying* $0 \le u < v \le T$, *and fix set* $A \in \mathcal{F}_u \vee \mathcal{G}_u$. *Then, the process* V *given by*

$$V_t = \mathbb{1}_A H_t^\top Z_t \mathbb{E}(Y_vU|\mathcal{F}_t), \quad t \in [0,T],$$

is an $\mathbb{F} \vee \mathbb{G}$ *martingale on the interval* $[u,v]$.

Proof. It suffices to prove that the process $\widehat{V}$ given as

$$\widehat{V}_t = H_t^\top Z_t \mathbb{E}(Y_vU|\mathcal{F}_t), \quad t \in [0,T],$$

is an $\mathbb{F} \vee \mathbb{G}$ martingale on $[0,v]$. Furthermore, since all components of H_t and Z_tY_v are non-negative and bounded by 1 (for the latter see Lemma 9.4), and since U is bounded, then it suffices to show that $\widehat{V}$ is an $\mathbb{F} \vee \mathbb{G}$ local martingale.

Towards this end we first verify that vector valued process $L = (L^x,\, x \in S)^\top$ defined by $L_t := H_t^\top Z_t,\ t \in [0,T]$, is an $\mathbb{F} \vee \mathbb{G}$ – local martingale with the following representation

$$L_t = H_0^\top + \int_0^t dM_u^\top \cdot Z_u, \quad t \in [0,T]. \tag{9.8}$$

Indeed, since Λ is an $\mathbb{F}$-intensity, integration by parts yields that

$$dL_t = d(H_t^\top Z_t) = H_{t-}^\top dZ_t + dH_t^\top \cdot Z_t = -H_{t-}^\top \Lambda_t Z_t dt + dH_t^\top \cdot Z_t = dM_t^\top \cdot Z_t.$$

Next, we observe that the vector valued process $U(\cdot, v) = (U^x(\cdot, v),\, x \in S)^\top$ defined by

$$U^x(t,v) = \sum_{y \in S} \mathbb{E}\left(Y_v^{xy} U^y | \mathcal{F}_t\right), \quad t \in [0,T],$$

is an $\mathbb{F}$-martingale. Since we assume that $\mathbb{F}$ is right-continuous we can take right-continuous modification of $U(\cdot, v)$.

Thus, by assumptions (2.9) and (2.10) in Theorem 2.11, its components are orthogonal to components of M. Hence the square bracket processes $[M^y, U^x(\cdot, v)]$, $x, y \in S$, are $\mathbb{F} \vee \mathbb{G}$-local martingales. By properties of square brackets (cf. Protter (Protter, 2005, Thm. II.6.29)) we obtain

$$[L^x, U^x(\cdot, v)]_t = \sum_{y \in S} \int_0^t Z_u^{y,x} d[M^y, U^x(\cdot, v)]_u.$$

Thus, by predictability and local boundedness of Z, and by (Protter, 2005, Thm. IV.2.29), we conclude that the process $[L^x, U^x(\cdot, v)]$ is a local martingale, and consequently that the local martingales L^x and $U^x(\cdot, v)$ are orthogonal. Since,

$$\widehat{V}_t = L_t U(t,v) = \sum_{x \in S} L_t^x U^x(t,v), \quad t \in [0,T],$$

we conclude that $\widehat{V}$ is an $\mathbb{F} \vee \mathbb{G}$ – local martingale as a sum of local martingales. □

Lemma 9.6. *Let Z be an $(\mathbb{F}, \mathbb{F}^Z)$-CDMC and let U be an $(\mathbb{F}, \mathbb{F}^U)$-CDMC, with values in some (finite) state space $\widehat{S}$, and with intensities Γ^Z and Γ^U,* *respectively. Then, the conditional law of Z given $\mathcal{F}_T$ coincides with the conditional law of U given $\mathcal{F}_T$ if and only if*

$$\Gamma^Z = \Gamma^U \quad du \otimes d\mathbb{P} - a.e., \tag{9.9}$$

$$\mathbb{P}(Z_0 = x|\mathcal{F}_T) = \mathbb{P}(U_0 = x|\mathcal{F}_T) \quad \forall x \in \widehat{S}. \tag{9.10}$$

Proof. First we prove sufficiency. Suppose that (9.9) and (9.10) hold. Recall that the c-transition fields P^Z (P^U respectively) satisfy Kolmogorov equations (4.12) and (4.13). Since (9.9) holds we see, by uniqueness of solutions of Kolmogorov equations, that $P^Z = P^U$. This and (9.10), by Proposition 4.5 (see eq. (4.7)) imply that conditional law of Z and U given $\mathcal{F}_T$ coincide. Now we prove necessity. Suppose that conditional laws of Z and U given $\mathcal{F}_T$ coincide, we want to show that (9.9) and (9.10) hold. First, note that the equality of conditional laws of Z and U given $\mathcal{F}_T$ implies (9.10). To show that (9.9) holds it suffices to show that their c-transition fields are equal. Indeed, this equality implies that, for any $0 \le v \le t \le T$,

$$0 = P^Z(v,t) - P^U(v,t) = \int_v^t \left(P^Z(v,u)\Gamma^Z - P^U(v,u)\Gamma^U_u \right) du = \int_v^t P^Z(v,u)(\Gamma^Z_u - \Gamma^U_u) du.$$

Consequently

$$P^Z(v,u)(\Gamma^Z_u - \Gamma^U_u) = 0, \quad du \otimes d\mathbb{P} - a.e. \ (on\ [v,T]),$$

since $P^Z(v,u)$ is invertible (cf. (Jakubowski and Niewęgłowski, 2010a, Proposition 3.11)). This in turn implies (9.9). This ends the proof. □

Acknowledgments

Research of T.R. Bielecki was partially supported by NSF grant DMS-1211256.

The authors extend sincere gratitude to two anonymous referees for their very careful reading of the first version of the paper, and for very insightful and helpful comments and advice.

Amendinger, J., 2000. Martingale representation theorems for initially enlarged filtrations. Stochastic Process. Appl. 89 (1), 101–116.
URL `http://dx.doi.org/10.1016/S0304-4149(00)00015-6`

Ball, F., Yeo, G. F., 1993. Lumpability and marginalisability for continuous-time Markov chains. J. Appl. Probab. 30 (3), 518–528.

Biagini, F., Groll, A., Widenmann, J., 2013. Intensity-based premium evaluation for unemployment insurance products. Insurance Math. Econom. 53 (1), 302–316.
URL http://dx.doi.org/10.1016/j.insmatheco.2013.06.001

Bielecki, T. R., Cialenco, I., Iyigunler, I., 2013a. Collateralized CVA valuation with rating triggers and credit migrations. International Journal of Theoretical and Applied Finance 16 (2), 1350009.

Bielecki, T. R., Cousin, A., Crépey, S., Herbertsson, A., 2013b. Forthcoming in Recent Advances in Financial Engineering 2012. World Scientific, Ch. A bottom-up dynamic model of portfolio credit risk - Part I: Markov copula perspective.

Bielecki, T. R., Cousin, A., Crépey, S., Herbertsson, A., 2013c. Forthcoming in Recent Advances in Financial Engineering 2012. World Scientific, Ch. A bottom-up dynamic model of portfolio credit risk - Part II: Common-shock interpretation, calibration and hedging issues.

Bielecki, T. R., Cousin, A., Crepey, S., Herbertsson, A., 2014a. A bottom-up dynamic model of portfolio credit risk - part I: Markov copula perspective. In: Recent Advances in Financial Engineering 2012. World Scientific.

Bielecki, T. R., Cousin, A., Crepey, S., Herbertsson, A., 2014b. A bottom-up dynamic model of portfolio credit risk - part II: Common-shock interpretation, calibration and hedging issues. In: Recent Advances in Financial Engineering 2012. World Scientific.

Bielecki, T. R., Cousin, A., Crepey, S., Herbertsson, A., 2014c. A bottom-up dynamic model of portfolio credit risk with stochastic intensities and random recoveries. Communications in Statistics âĂŞ Theory and Methods 43, 1362–1389.

Bielecki, T. R., Cousin, A., Crepey, S., Herbertsson, A., 2014d. Dynamic hedging of portfolio credit risk in a Markov copula model. Journal of Optimization Theory and Applications 161 (1).

Bielecki, T. R., Crépey, S., Jeanblanc, M., Rutkowski, M., 2008a. Valuation of basket credit derivatives in the credit migrations environment. Birge, J. and Linetsky, V. (ed.) Handbooks in Operations Research and Management Science: Financial Engineering, vol.15. 471-510.

Bielecki, T. R., Crepey, S., Jeanblanc, M., Zargari, B., 2012a. Valuation and hedging of cds counterparty exposure in a Markov copula model. Int. J. Theor. Appl. Finance 15 (1), 1250004, 39.

Bielecki, T. R., Jakubowski, J., Niewęgłowski, M., ???? Fundamentals of Theory of Structured Dependence between Stochastic Processes, book in preparation.

Bielecki, T. R., Jakubowski, J., Niewęgłowski, M., 2010. Dynamic modeling of dependence in finance via copulae between stochastic processes. In: Copula theory and its applications. Vol. 198 of Lect. Notes Stat. Proc. Springer, Heidelberg, pp. 33–76.
URL http://dx.doi.org/10.1007/978-3-642-12465-5_2

Bielecki, T. R., Jakubowski, J., Niewęgłowski, M., 2012b. Study of dependence for some stochastic processes: symbolic Markov copulae. Stochastic Process. Appl. 122 (3), 930–951.
URL http://dx.doi.org/10.1016/j.spa.2011.11.001

Bielecki, T. R., Jakubowski, J., Niewęgłowski, M., 2013d. Intricacies of dependence between components of multivariate Markov chains: weak Markov consistency and weak Markov copulae. Electron. J. Probab. 18, no. 45, 21.
URL http://dx.doi.org/10.1214/EJP.v18-2238

Bielecki, T. R., Jakubowski, J., Niewęgłowski, M., 2015. Conditional Markov chains – construction and properties. In: Banach Center Publications 105(2015) Stochastic analysis. Special volume in honour of Jerzy Zabczyk. Warsaw: Polish Academy of Sciences, Institute of Mathematics, pp. 33–42.

Bielecki, T. R., Jakubowski, J., Niewęgłowski, M., 2015a. Conditional Markov chains, part I: construction and properties.
URL http://arxiv.org/abs/1501.05531

Bielecki, T. R., Jakubowski, J., Niewęgłowski, M., 2015b. Conditional Markov chains, part II: consistency and copulae.
URL http://arxiv.org/abs/1501.05535

Bielecki, T. R., Jakubowski, J., Vidozzi, A., Vidozzi, L., 2008b. Study of dependence for some stochastic processes. Stoch. Anal. Appl. 26 (4), 903–

924.
URL http://dx.doi.org/10.1080/07362990802128958

Bielecki, T. R., Rutkowski, M., 2000. Multiple ratings model of defaultable term structure. Math. Finance 10 (2), 125–139, iNFORMS Applied Probability Conference (Ulm, 1999).
URL http://dx.doi.org/10.1111/1467-9965.00085

Bielecki, T. R., Rutkowski, M., 2002. Credit risk: Modelling, valuation and hedging. Springer Finance. Berlin: Springer. xviii, 500 p.

Bielecki, T. R., Rutkowski, M., 2003. Dependent defaults and credit migrations. Appl. Math. (Warsaw) 30 (2), 121–145.
URL http://dx.doi.org/10.4064/am30-2-1

Bielecki, T. R., Rutkowski, M., 2004. Modeling of the defaultable term structure: Conditionally Markov approach. IEEE Transactions on Automatic Control 49, 361–373.

Bielecki, T. R., Vidozzi, A., Vidozzi, L., 2008c. Markov copulae approach to pricing and hedging of credit index derivatives and ratings triggered step - up bonds. Journal of Credit Risk 4 (1), 47–76.

Blanes, S., Casas, F., Oteo, J. A., Ros, J., 2009. The Magnus expansion and some of its applications. Phys. Rep. 470 (5-6), 151–238.
URL http://dx.doi.org/10.1016/j.physrep.2008.11.001

Brémaud, P., 1981. Point Processes and Queues. Springer-Verlag, New York, martingale dynamics, Springer Series in Statistics.

Burke, C. J., Rosenblatt, M., 1958. A Markovian function of a Markov chain. Ann. Math. Statist. 29, 1112–1122.

Eberlein, E., Grbac, Z., 2013. Rating based Lévy Libor model. Math. Finance 23 (4), 591–626.
URL http://dx.doi.org/10.1111/j.1467-9965.2011.00514.x

Eberlein, E., Özkan, F., 2003. The defaultable Lévy term structure: ratings and restructuring. Math. Finance 13 (2), 277–300.
URL http://dx.doi.org/10.1111/1467-9965.00017

Ethier, S. N., Kurtz, T. G., 1986. Markov processes. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, characterization and convergence.

Föllmer, H., Protter, P., 2011. Local martingales and filtration shrinkage. ESAIM Probab. Stat. 15 (In honor of Marc Yor, suppl.), S25–S38.
URL http://dx.doi.org/10.1051/ps/2010023

Gill, R. D., Johansen, S., 1990. A survey of product-integration with a view toward application in survival analysis. Ann. Statist. 18 (4), 1501–1555.
URL http://dx.doi.org/10.1214/aos/1176347865

Goutte, S., Ngoupeyou, A., 2013. Defaultable bond pricing using regime switching intensity model. J. Appl. Math. Inform. 31 (5-6), 711–732.
URL http://dx.doi.org/10.14317/jami.2013.711

Granger, C. W. J., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica.

He, S. W., Wang, J. G., Yan, J. A., 1992. Semimartingale Theory and Stochastic Calculus. Kexue Chubanshe (Science Press), Beijing.

Horn, R. A., Johnson, C. R., 1994. Topics in matrix analysis. Cambridge University Press, Cambridge, corrected reprint of the 1991 original.

Jacod, J., 1974/75. Multivariate point processes: predictable projection, Radon-Nikodým derivatives, representation of martingales. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 31, 235–253.

Jakubowski, J., Niewęgłowski, M., 2008. Pricing bonds and CDS in the model with rating migration induced by a Cox process. Stettner, Łukasz (ed.), Advances in mathematics of finance. Banach Center Publications 83, 159-182.

Jakubowski, J., Niewęgłowski, M., 2010a. A class of $\mathbb{F}$-doubly stochastic Markov chains. Electron. J. Probab. 15, no. 56, 1743–1771.
URL http://dx.doi.org/10.1214/EJP.v15-815

Jakubowski, J., Niewęgłowski, M., 2010b. Defaultable bonds with an infinite number of Lévy factors. Appl. Math. (Warsaw) 37 (3), 275–307.
URL http://dx.doi.org/10.4064/am37-3-2

Jakubowski, J., Niewęgłowski, M., 2011. Pricing and hedging of rating-sensitive claims modeled by $\mathbb{F}$-doubly stochastic Markov chains. In: Advanced mathematical methods for finance. Springer, Heidelberg, pp. 417–453.
URL http://dx.doi.org/10.1007/978-3-642-18412-3_16

Jakubowski, J., Pytel, A., June 2016. The Markov consistency of Archimedean survival processes. Journal of Applied Probability 53 (02), 392–409.
URL http://dx.doi.org/10.1017/jpr.2016.8

Jeanblanc, M., Yor, M., Chesney, M., 2009. Mathematical methods for financial markets. Springer Finance. Springer-Verlag London, Ltd., London.
URL http://dx.doi.org/10.1007/978-1-84628-737-4

Liang, X., Dong, Y., 2014. A Markov chain copula model for credit default swaps with bilateral counterparty risk. Comm. Statist. Theory Methods 43 (3), 498–514.
URL http://dx.doi.org/10.1080/03610926.2012.665555

Protter, P. E., 2005. Stochastic Integration and Differential Equations. Vol. 21 of Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin, second edition. Version 2.1, Corrected third printing.

Rao, M., 1972. On modification theorems. Trans. Amer. Math. Soc. 167, 443–450.