Markov Chain Models of Portfolio Credit Risk

Tomasz R. Bielecki
Department of Applied Mathematics
Illinois Institute of Technology
Chicago, IL 60616, USA

Stéphane Crépey*
Département de Mathématiques
Université d'Évry Val d'Essonne
91025 Évry Cedex, France

Alexander Herbertsson†
Department of Economics
University of Gothenburg
SE-405 30 Göteborg, Sweden

This Version: November 16, 2009

Contents

*The research of S. Crépey benefited from the support of the 'Chaire Risque de crédit', Fédération Bancaire Française, and of the Europlace Institute of Finance.
†Research of A. Herbertsson supported by the Jan Wallander and Tom Hedelius Foundation.

1 Outline

This is a survey paper in which we present a selection of methods and results regarding various applications of the theory of continuous time Markov chains to valuation of credit derivatives. We present both theoretical and numerical aspects of the Markovian methodology.

After a review of some basic notions and results from the theory of continuous-time Markov chains in section 2, sections 3 to 5 are devoted to the study of a few specific Markovian models of portfolio credit risk. This survey article is intended to illustrate the power and flexibility of the Markov chain approach to portfolio credit risk, yet it is by no means exhaustive and we refer the reader to, for instance, Frey and Backhaus [21, 20, 19], Cont and Minca [16], Albanese [1], Halperin and Tomecek [24] or Lopatin and Misirshapaev [36], for other models, and more specifically to Frey and Backhaus [21], Bielecki et al. [12, 8] and Laurent,

Cousin and Fermanian [34] regarding the issue of hedging, which we do not really discuss here due space limitations.

The full version of this paper can be found in *Oxford Handbook of Credit Derivatives*, A. Lipton and A. Rennie, eds. .

2 Continuous-Time Markov Chains

The material presented in this section is, for the most part, taken from Chapter 11 of [11], to which we refer in particular for the proofs of all results. For a more exhaustive treatment of Markov chains, we refer to any of a large variety of available monographs on the theory of stochastic processes, to mention a few: Bhattacharya and Waymire [7], Syski [47], Last and Brandt [33], and Rogers and Williams [42].

Since most Markov chain models used in portfolio credit risk are time-continuous, we shall focus on continuous-time Markov chains (except for the notion of discrete-time Markov Chain embedded to a continuous-time Markov chain which is dealt with in Section 2.3). However it should be noted that completely analogous developments are in fact valid relatively to discrete time Markov chains (see Chapter 11 of [11]), which is relevant to credit risk when it comes to model implementation, which often proceeds via time discretization (see, for instance, Section 3.1.4).

After an introduction to time-inhomogeneous as well as time-inhomogeneous Markov chains, we examine conditional expectations, which in the financial interpretation is of course relevant for pricing purposes. We then consider some 'fundamental' martingales with respect to certain relevant filtrations, and we derive related martingale representation theorems. In the context of finance, such martingale representation are key to the issue of hedging (see, e.g., Frey and Backhaus [19], Bielecki et al. [12, 8] or Laurent, Cousin and Fermanian [34]). We also deal with various examples of random times associated with a Markov chain, such as the jump times and the absorption time, which in the context of credit risk are related to the issue of deriving univariate (see section 2.5) or multivariate (see section 4.1.2) default, survival and conditional default and survival distributions. Finally, we study the behavior of a time-homogeneous Markov chain under an equivalent change of a probability measure. In applications such equivalent changes of measure can be interpreted and used as a bridge between the statistical measure and a risk-neutral measure (see, e.g., Section 3.1.1 and [12]), or as a change of numeraire (see, e.g., [9], cf. Section 3.1).

In what follows, we fix the underlying probability space $(\Omega, \mathcal{G}, \mathbb{Q})$, as well as a finite set $\mathcal{K} = \{1, \ldots, K\}$, which plays the role of the *state space* for a Markov chain of interest. Since the state space is finite, it is clear that any function $h : \mathcal{K} \to \mathbb{R}$ is bounded and measurable, provided that we endow the state space with the σ-field of all its subsets.

Let C_t, $t \in \mathbb{R}_+$, be a right-continuous stochastic process on $(\Omega, \mathcal{G}, \mathbb{Q})$ with values in the finite set $\mathcal{K}$, and let $\mathbb{F}^C$ be the filtration generated by this process. Also, let $\mathbb{G}$ be some filtration such that $\mathbb{F}^C \subseteq \mathbb{G}$.

Definition 2.1 A process C is a *continuous-time $\mathbb{G}$-Markov chain* if for an arbitrary function $h : \mathcal{K} \to \mathbb{R}$ and any $s, t \in \mathbb{N}$ we have

$$\mathbb{E}_{\mathbb{Q}}(h(C_{t+s}) \,|\, \mathcal{G}_t) = \mathbb{E}_{\mathbb{Q}}(h(C_{t+s}) \,|\, C_t).$$

A continuous-time $\mathbb{G}$-Markov chain C is said to be *time-homogeneous* if, in addition, for any $s, t, u \in \mathbb{N}$ we have

$$\mathbb{E}_{\mathbb{Q}}(h(C_{t+s}) \,|\, C_t) = \mathbb{E}_{\mathbb{Q}}(h(C_{u+s}) \,|\, C_u).$$

Definition 2.2 A two-parameter family $\mathcal{P}(t, s)$, $t, s \in \mathbb{R}_+$, $t \leq s$, of stochastic matrices is called the family of *transition probability matrices* for the $\mathbb{G}$-Markov chain C under $\mathbb{Q}$ if, for every $t, s \in \mathbb{R}_+$, $t \leq s$,

$$\mathbb{Q}\{C_s = j \,|\, C_t = i\} = p_{ij}(t, s), \quad \forall\, i, j \in \mathcal{K}.$$

In particular, the equality $\mathcal{P}(t, t) = \mathrm{Id}$ is satisfied for every $t \in \mathbb{R}_+$.

2.1 Time-homogeneous chains

In case of a time-homogeneous Markov chain C, we introduce the following definition.

Definition 2.3 The one-parameter family $\mathcal{P}(t),\ t \in \mathbb{R}_+$, of stochastic matrices is called the family of *transition probability matrices* for the time-homogeneous $\mathbb{G}$-Markov chain C under $\mathbb{Q}$ if, for every $t, s \in \mathbb{R}_+$,

$$\mathbb{Q}\{C_{s+t} = j \,|\, C_s = i\} = p_{ij}(t), \quad \forall\, i, j \in \mathcal{K}. \tag{1}$$

If $\mathcal{P}(t),\ t \in \mathbb{R}_+$ is the family of transition matrices for C then for any subset $A \subseteq \mathcal{K}$ we have

$$\mathbb{Q}\{C_{t+s} \in A \,|\, C_t\} = \sum_{j \in A} p_{C_t j}(s), \quad \forall\, s, t \in \mathbb{R}_+.$$

Moreover, the Chapman-Kolmogorov equation is satisfied, namely,

$$\mathcal{P}(t+s) = \mathcal{P}(t)\mathcal{P}(s) = \mathcal{P}(s)\mathcal{P}(t), \quad \forall\, s, t \in \mathbb{R}_+.$$

Equivalently, for every $s, t \in \mathbb{R}_+$ and $i, j \in \mathcal{K}$,

$$p_{ij}(t+s) = \sum_{k=1}^{K} p_{ik}(t) p_{kj}(s) = \sum_{k=1}^{K} p_{ik}(s) p_{kj}(t).$$

Let the K-dimensional (row) vector $\mu_0 = [\mu_0(i)]_{1 \le i \le K} = [\mathbb{Q}\{C_0 = i\}]_{1 \le i \le K}$ denote the initial probability distribution for the Markov chain C under $\mathbb{Q}$.

Likewise, let the (row) vector $\mu_t = [\mu_t(i)]_{1 \le i \le K} = [\mathbb{Q}\{C_t = i\}]_{1 \le i \le K}$ stand for the probability distribution of C at time $t \in \mathbb{R}_+$. It can be easily checked that

$$\mu_{t+s} = \mu_0 \mathcal{P}(t+s) = \mu_t \mathcal{P}(s) = \mu_s \mathcal{P}(t), \quad \forall\, s, t \in \mathbb{R}_+.$$

We now impose an important assumption on the family $\mathcal{P}(\cdot)$, specifically, that this family is right-continuous at time $t = 0$, that is, $\lim_{t \downarrow 0} \mathcal{P}(t) = \mathcal{P}(0)$. By virtue of the Chapman-Kolmogorov equation, this implies that

$$\lim_{s \to 0} \mathcal{P}(t+s) = \mathcal{P}(t), \quad \forall\, t > 0,$$

and thus

$$\lim_{s \to 0} \mathbb{Q}\{C_{t+s} = j \,|\, C_t = i\} = \delta_{ij}, \quad \forall\, i, j \in \mathcal{K},\ t > 0.$$

It is a well-known fact (see, for instance, Theorem 8.1.2 in Rolski et al. (1998)) that the right-hand side continuity at time $t = 0$ of the family $\mathcal{P}(\cdot)$ implies the right-hand side differentiability at $t = 0$ of this family. More specifically, the following finite limits exist, for every $i, j \in \mathcal{K}$,

$$\lambda_{ij} := \lim_{t \downarrow 0} \frac{p_{ij}(t) - p_{ij}(0)}{t} = \lim_{t \downarrow 0} \frac{p_{ij}(t) - \delta_{ij}}{t}\,. \tag{2}$$

Observe that for every $i \neq j$ we have $\lambda_{ij} \ge 0$, and $\lambda_{ii} = -\sum_{j=1,\, j \neq i}^{K} \lambda_{ij}$. The matrix $\Lambda := [\lambda_{ij}]_{1 \le i,j \le K}$ is called the *infinitesimal generator matrix* for a Markov chain associated with the family $\mathcal{P}(\cdot)$ via (1). Since each entry λ_{ij} of the matrix Λ can be shown to represent the intensity of transition from the state i to the state j, the infinitesimal generator matrix Λ is also commonly known as the *intensity matrix.*

Invoking the Chapman-Kolmogorov equation and equality (2), one may derive the *backward Kolmogorov equation*

$$\frac{d\mathcal{P}(t)}{dt} = \Lambda \mathcal{P}(t), \quad \mathcal{P}(0) = \mathrm{Id}, \tag{3}$$

and the *forward Kolmogorov equation*

$$\frac{d\mathcal{P}(t)}{dt} = \mathcal{P}(t) \Lambda, \quad \mathcal{P}(0) = \mathrm{Id}, \tag{4}$$

where, at time $t = 0$, we take the right-hand side derivatives. It is well known that both these equations have the same unique solution:

$$\mathcal{P}(t) = e^{t\Lambda} := \sum_{n=0}^{\infty} \frac{\Lambda^n t^n}{n!}, \quad \forall\, t \in \mathbb{R}_+. \tag{5}$$

We conclude that the generator matrix Λ uniquely determines all relevant probabilistic properties of a time-homogeneous Markov chain.

The following important result provides a martingale characterization of a time-homogeneous Markov chain C in terms of its infinitesimal generator. For the proof of Proposition 2.1, we refer to Last and Brandt [33] or Rogers and Williams [42]. In the quoted references, the corresponding result is stated for an $\mathbb{F}^C$-Markov chain, rather than for a $\mathbb{G}$-Markov chain. However, the proof of this more general version is analogous. For any state $i \in \mathcal{K}$ and any function $h : \mathcal{K} \to \mathbb{R}$, we denote

$$(\Lambda h)(i) = \sum_{j=1}^{K} \lambda_{ij} h(j).$$

Proposition 2.1 *A process C is a time-homogeneous $\mathbb{G}$-Markov chain under $\mathbb{Q}$, with the initial distribution μ_0 and with the infinitesimal generator matrix Λ, if and only if the following conditions are satisfied:*
(i) $\mathbb{Q}\{C_0 = i\} = \mu_0(i)$ *for every* $i \in \mathcal{K}$,
(ii) *for any function $h : \mathcal{K} \to \mathbb{R}$ the process M^h, defined by the formula*

$$M_t^h = h(C_t) - \int_0^t (\Lambda h)(C_u)\, du, \quad \forall\, t \in \mathbb{R}_+,$$

follows a $\mathbb{G}$-martingale under $\mathbb{Q}$.

Example 2.4 Let C be a time-homogeneous $\mathbb{G}$-Markov chain with the infinitesimal generator matrix Λ. Applying Proposition 2.1 to the function $h(\cdot) = \mathbb{1}_{\{i\}}(\cdot)$, we conclude that the process

$$M_t^i = H_t^i - \int_0^t \lambda_{C_u i}\, du, \quad \forall\, t \in \mathbb{R}_+, \tag{6}$$

follows a $\mathbb{G}$-martingale (and an $\mathbb{F}^C$-martingale) under $\mathbb{Q}$. Conversely, if for every $i \in \mathcal{K}$ the process M^i follows a $\mathbb{G}$-martingale, then for any function $h : \mathcal{K} \to \mathbb{R}$ the process M^h is a $\mathbb{G}$-martingale under $\mathbb{Q}$.

2.2 Time-inhomogeneous chains

If a Markov chain is time-inhomogeneous, the time-dependent transition intensities are introduced through the formula[1]

$$\lambda_{ij}(t) = \lim_{h \downarrow 0} \frac{p_{ij}(t, t+h) - \delta_{ij}}{h}\,.$$

It is obvious that $\lambda_{ij}(t) \geq 0$ for arbitrary $i \neq j$, and

$$\lambda_{ii}(t) = \lim_{h \downarrow 0} \frac{p_{ii}(t, t+h) - 1}{h} = -\lim_{h \downarrow 0} \frac{\sum_{j=1,\, j \neq i}^{K} p_{ij}(t, t+h)}{h} = -\sum_{j=1,\, j \neq i}^{K} \lambda_{ij}(t),$$

where

$$p_{ij}(t, t+h) = \mathbb{Q}\{C_{t+h} = j \,|\, C_t = i\}, \quad \forall\, i, j \in \mathcal{K}.$$

We shall write $\Lambda(t) = [\lambda_{ij}(t)]_{1 \leq i, j \leq K}$ to denote the infinitesimal generator matrix function associated with a time-inhomogeneous Markov chain C.

[1] Let us mention that mild regularity conditions need to be satisfied by the probabilities $p_{ij}(s, t)$ for the results of this subsection to be valid.

The two parameter family $\mathcal{P}(t,s) = [p_{ij}(t,s)]_{1\leq i,j\leq K}$, $0 \leq t \leq s$, of transition matrices for C satisfies the Chapman-Kolmogorov equation:

$$\mathcal{P}(t,s) = \mathcal{P}(t,u)\mathcal{P}(u,s), \quad \forall\, t \leq u \leq s,$$

the forward Kolmogorov equation:

$$\frac{d\mathcal{P}(t,s)}{ds} = \mathcal{P}(t,s)\Lambda(s), \quad \mathcal{P}(t,t) = \mathrm{Id}, \tag{7}$$

and the backward Kolmogorov equation:

$$\frac{d\mathcal{P}(t,s)}{dt} = -\Lambda(t)\mathcal{P}(t,s), \quad \mathcal{P}(s,s) = \mathrm{Id}. \tag{8}$$

The next result is an immediate consequence of Kolmogorov's equations.

Corollary 2.2 *The family $\mathcal{P}(t,s)$, $0 \leq t \leq s$, satisfies the integral equations*

$$\mathcal{P}(t,s) = \mathrm{Id} + \int_t^s \mathcal{P}(t,u)\Lambda(u)\,du$$

and

$$\mathcal{P}(t,s) = \mathrm{Id} + \int_t^s \Lambda(u)\mathcal{P}(u,s)\,du.$$

The above equations can be used in order to derive some remarkable representations, which are important from the computational point of view, and which are counterparts of (5). For the proof of Corollary 2.3, the interested reader is referred to Rolski et al. [43] (see Theorem 8.4.4 therein).

Corollary 2.3 *For every $0 \leq t \leq s$ we have*

$$\mathcal{P}(t,s) = \mathrm{Id} + \sum_{n=1}^{\infty} \int_t^s \int_{u_1}^s \cdots \int_{u_{n-1}}^s \Lambda(u_1)\ldots\Lambda(u_n)\,du_n\ldots du_1,$$

and

$$\mathcal{P}(t,s) = \mathrm{Id} + \sum_{n=1}^{\infty} \int_t^s \int_t^{u_1} \cdots \int_t^{u_{n-1}} \Lambda(u_1)\ldots\Lambda(u_n)\,du_n\ldots du_1.$$

Assume that the matrix function $\Lambda(t) = [\lambda_{ij}(t)]_{1\leq i,j\leq K}$ satisfies the conditions, which characterize the infinitesimal generator of an inhomogeneous Markov chain, namely,

$$\lambda_{ij}(t) \geq 0, \;\; i \neq j, \quad \lambda_{ii}(t) = -\sum_{j=1,\, j\neq i}^{K} \lambda_{ij}(t).$$

For any function $h : \mathcal{K} \to \mathbb{R}$, we introduce the mapping $\Lambda h : \mathcal{K} \times \mathbb{R}_+ \to \mathbb{R}$ by setting:

$$(\Lambda h)(i,t) = \sum_{j=1}^{K} \lambda_{ij}(t)h(j), \quad \forall\, i \in \mathcal{K},\, t \in \mathbb{R}_+.$$

The following result is a natural extension of Proposition 2.1.

Proposition 2.4 *A process C is a $\mathbb{G}$-Markov chain under $\mathbb{Q}$, with the initial distribution μ_0 and with the infinitesimal generator matrix function $\Lambda(\cdot)$, if and only if:*
(i) $\mathbb{Q}\{C_0 = i\} = \mu_0(i)$ *for every $i \in \mathcal{K}$,*
(ii) *for any function $h : \mathcal{K} \to \mathbb{R}$ the process M^h, defined by the formula*

$$M_t^h = h(C_t) - \int_0^t (\Lambda h)(C_u, u)\,du, \quad t \in \mathbb{R}_+, \tag{9}$$

follows a $\mathbb{G}$-martingale under $\mathbb{Q}$.

Example 2.5 Let C be a time-inhomogeneous $\mathbb{G}$-Markov chain with the infinitesimal generator matrix function $\Lambda(\cdot)$. By applying Proposition 2.4 to the function $h(\cdot) = \mathbb{1}_{\{i\}}(\cdot)$, we find that the process

$$M_t^i = H_t^i - \int_0^t \lambda_{C_u i}(u)\, du, \quad \forall\, t \in \mathbb{R}_+, \tag{10}$$

follows a $\mathbb{G}$-martingale (and an $\mathbb{F}^C$-martingale) under $\mathbb{Q}$. Similarly as in the time-homogeneous case, if for every $i \in \mathcal{K}$ the process M^i is a $\mathbb{G}$-martingale, then for any function $h : \mathcal{K} \to \mathbb{R}$ the process M^h, given by formula (9), also follows a $\mathbb{G}$-martingale.

2.3 Embedded Discrete-Time Markov Chain

Let C_t, $t \in \mathbb{R}_+$, stand for a continuous-time time-homogeneous $\mathbb{G}$-Markov chain (and thus a Markov chain w.r.t. its own filtration) under $\mathbb{Q}$ with the infinitesimal generator Λ. Let τ_n, $n \in \mathbb{N}$, denote the random sequence of successive jump times of C. More explicitly, for any $n \in \mathbb{N}$, the random variable τ_n defined as (by convention, $\tau_0 = 0$):

$$\tau_n = \inf\{t > \tau_{n-1} : C_t \neq C_{\tau_{n-1}}\}, \tag{11}$$

represents the time of the n^{th} jump (or transition) for C. Let us recall few classic results related to the behavior of a continuous-time Markov chain at its jump times.

First, it is well known that the following property holds, for any $n \in \mathbb{N}$ and every $t \in \mathbb{R}_+$,

$$\mathbb{Q}\{\tau_n - \tau_{n-1} > t \,|\, C_{\tau_{n-1}} = i\} = e^{\lambda_{ii} t}, \quad \forall\, i = 1, \ldots, K. \tag{12}$$

Equality (12) makes it clear that, conditionally on the position $C_{\tau_{n-1}} = i$ at the jump time τ_{n-1}, the random time that elapses until the next jump occurs has an exponential probability law with the parameter $-\lambda_{ii} > 0$.

Second, the conditional probabilities of transitions are known to satisfy:

$$\mathbb{Q}\{C_{\tau_n} = j \,|\, C_{\tau_{n-1}} = i\} = p_{ij} := -\frac{\lambda_{ij}}{\lambda_{ii}}, \quad \forall\, i, j \in \mathcal{K},\ i \neq j. \tag{13}$$

Formula (13) specifies the conditional probability law of a continuous-time Markov chain C after its n^{th} jump, given the position after the $(n-1)^{\text{th}}$ jump (it coincides, of course, with the position of C just before the n^{th} jump).

Let us emphasize that since C is assumed to be a time-homogeneous Markov chain, both probability laws introduced above do not depend on the number of transitions in the past (that is, on n). They only depend on the value taken by C after the previous jump.

Define a random sequence $\widehat{C}_n = C_{\tau_n}$ for every $n \in \mathbb{N}$. It is well known that the sequence $\widehat{C}$ is a time-homogeneous Markov chain under $\mathbb{Q}$ with the one-step transition probability matrix $P = [p_{ij}]_{1 \leq i,j \leq K}$. The discrete-time Markov chain $\widehat{C}_n$, $n \in \mathbb{N}$, is called the *embedded Markov chain* corresponding to the continuous-time Markov chain C.

Remark 2.6 In the case of a *time-inhomogeneous* $\mathbb{G}$-Markov chain C, one may define likewise the embedded *Markov process in discrete time* $(\tau_n, C_{\tau_n})\, n \in \mathbb{N}$, defined over the extended time-space state by the following analogs of (12), (13)

$$\mathbb{Q}\{\tau_n - \tau_{n-1} > t \,|\, C_{\tau_{n-1}} = i\} = e^{\int_{\tau_{n-1}}^t \lambda_{ii}(u) du}, \quad \forall\, i = 1, \ldots, K. \tag{14}$$

$$\mathbb{Q}\{C_{\tau_n} = j \,|\, C_{\tau_{n-1}} = i\,, \tau_n = t\} = p_{ij}(t) := -\frac{\lambda_{ij}(t)}{\lambda_{ii}(t)}, \quad \forall\, i, j \in \mathcal{K},\ i \neq j. \tag{15}$$

The results of this subsection are important in practice regarding the issue, for instance, of numerical simulation (See e.g., Section 3.1.4).

2.4 Conditional Expectations

We say that a state $k \in \mathcal{K}$ is *absorbing* for a time-homogeneous $\mathbb{G}$-Markov chain $C_t,\ t \in \mathbb{R}_+$, if the following holds:

$$\mathbb{Q}\{C_s = k \,|\, C_t = k\} = 1, \quad \forall\, t, s \in \mathbb{R}_+,\ t \leq s.$$

In view of (2), it is clear that if a state $k \in \mathcal{K}$ is absorbing, then we have $\lambda_{kj} = 0$ for every $j = 1, \ldots, K$.

>From now on, we shall postulate that the state K is absorbing. This implies that the infinitesimal generator of C under $\mathbb{Q}$ is given by the intensity matrix Λ of the following form:

$$\Lambda = \begin{pmatrix} \lambda_{1,1} & \ldots & \lambda_{1,K-1} & \lambda_{1,K} \\ \cdot & \ldots & \cdot & \cdot \\ \lambda_{K-1,1} & \ldots & \lambda_{K-1,K-1} & \lambda_{K-1,K} \\ 0 & \ldots & 0 & 0 \end{pmatrix}.$$

We assume that the initial state $C_0 = x \neq K$ is fixed, and we denote by τ the random time of absorption at K, i.e., $\tau = \inf\{t > 0 : C_t = K\}$. We assume that $\tau < \infty$, $\mathbb{Q}$-a.s.; this implies that the state K is the only recurrent state for C. As usual, we write $H_t^i = \mathbb{1}_{\{C_t = i\}}$ and $H_t = \mathbb{1}_{\{\tau \leq t\}} = \mathbb{1}_{\{C_t = K\}} = H_t^K$.

In the next few auxiliary results, we shall deal with the conditional expectations with respect to the filtrations $\mathbb{G}$ and $\mathbb{F}^C$. The absorption time τ is, of course, an $\mathbb{F}^C$-stopping time and a $\mathbb{G}$-stopping time. In what follows, Y will denote an integrable random variable, which is defined on the reference probability space $(\Omega, \mathcal{G}, \mathbb{Q})$.

Lemma 2.5 *We have*

$$\mathbb{1}_{\{\tau \leq t\}} \mathbb{E}_{\mathbb{Q}}(Y \,|\, \mathcal{G}_t) = \mathbb{E}_{\mathbb{Q}}(\mathbb{1}_{\{\tau \leq t\}} Y \,|\, \mathcal{G}_t \vee \sigma(\tau))$$

and

$$\mathbb{1}_{\{\tau \leq t\}} \mathbb{E}_{\mathbb{Q}}(Y \,|\, \mathcal{F}_t^C) = \mathbb{E}_{\mathbb{Q}}(\mathbb{1}_{\{\tau \leq t\}} Y \,|\, \mathcal{F}_t^C \vee \sigma(\tau)).$$

In the next lemma, we examine the case when the random variable Y has the form $h(C_{\tau-}, \tau)$ for some function $h : \mathcal{K} \times \mathbb{R}_+ \to \mathbb{R}$.

Lemma 2.6 *Let $Y = h(C_{\tau-}, \tau)$ for some function $h : \mathcal{K} \times \mathbb{R}_+ \to \mathbb{R}$. Then*

$$\mathbb{E}_{\mathbb{Q}}(\mathbb{1}_{\{\tau > t\}} Y \,|\, \mathcal{G}_t) = \sum_{i=1}^{K-1} H_t^i\, \mathbb{E}_{\mathbb{Q}}(\mathbb{1}_{\{\tau > t\}} Y \,|\, C_t = i).$$

The next two auxiliary results are simple corollaries to Lemma 2.6.

Corollary 2.7 *Let $Y = h(C_{\tau-}, \tau)$ for some function $h : \mathcal{K} \times \mathbb{R}_+ \to \mathbb{R}$. Then*

$$\mathbb{E}_{\mathbb{Q}}(Y \,|\, \mathcal{G}_t) = \mathbb{1}_{\{\tau \leq t\}} Y + \sum_{i=1}^{K-1} H_t^i\, \mathbb{E}_{\mathbb{Q}}(\mathbb{1}_{\{\tau > t\}} Y \,|\, C_t = i).$$

Corollary 2.8 *For any $s, t \in \mathbb{R}_+$, the following equalities are valid*

$$\mathbb{Q}\{\tau > s \,|\, \mathcal{G}_t\} = \mathbb{1}_{\{s \leq t\}} \mathbb{1}_{\{\tau > s\}} + \mathbb{1}_{\{s > t\}} \sum_{i=1}^{K-1} H_t^i\, \mathbb{Q}\{\tau > s \,|\, C_t = i\},$$

and

$$\mathbb{Q}\{\tau \geq s \,|\, \mathcal{G}_t\} = \mathbb{1}_{\{s \leq t\}} \mathbb{1}_{\{\tau \geq s\}} + \mathbb{1}_{\{s > t\}} \sum_{i=1}^{K-1} H_t^i\, \mathbb{Q}\{\tau \geq s \,|\, C_t = i\}.$$

2.5 Probability Distribution of the Absorption Time

We maintain the assumptions of section 2.4. More explicit formulae for the conditional expectations with respect to the σ-field $\mathcal{G}_t$ can be obtained, if the knowledge of conditional laws of C is used. Notice that for every $0 \leq t \leq s$ we have

$$\mathbb{Q}\{\tau > s \,|\, C_t = i\} = 1 - \mathbb{Q}\{C_s = K \,|\, C_t = i\} = 1 - p_{iK}(s-t),$$

hence, the first formula of Corollary 2.8 can be rewritten as follows:

$$\mathbb{Q}\{\tau > s \,|\, \mathcal{G}_t\} = \mathbb{1}_{\{s \leq t\}} \mathbb{1}_{\{\tau > s\}} + \mathbb{1}_{\{s > t\}} \sum_{i=1}^{K-1} H_t^i \big(1 - p_{iK}(s-t)\big). \tag{16}$$

To derive an alternative representation for the probability distribution of the absorption time, let us denote by $\widetilde{\Lambda}$ the matrix obtained from Λ by deleting the last row and the last column. Also, let $\widetilde{\mathcal{P}}(t) = [\widetilde{p}_{ij}(t)]_{i,j \in \widetilde{\mathcal{K}}}$ stand for the associated transition matrix, where $\widetilde{\mathcal{K}} = \{1, \ldots, K-1\}$.

Recall that the quantity $\mathbb{Q}\{\tau > s \,|\, C_t = i\}$ is of fundamental importance when studying the spread for a defaultable bond, see e.g. p.496-497 in [31].

It is not difficult to check that the so-called *taboo probabilities* $\widetilde{p}_{ij}(t)$, $i, j \in \widetilde{\mathcal{K}}$ can be found by solving the following differential equation:

$$\frac{d}{dt}\widetilde{\mathcal{P}}(t) = \widetilde{\Lambda}\widetilde{\mathcal{P}}(t), \quad t > 0, \tag{17}$$

with the initial condition $\widetilde{\mathcal{P}}(0) = \mathrm{Id}$.

It is also clearly seen that (recall that we have assumed that $C_0 = i \in \widetilde{\mathcal{K}}$)

$$F(t) = 1 - \sum_{j=1}^{K-1} \widetilde{p}_{ij}(t) = 1 - \sum_{j=1}^{K-1} p_{ij}(t). \tag{18}$$

Since $F(t) < 1$ for every $t \in \mathbb{R}_+$, we may introduce the hazard function Γ of τ by setting $\Gamma(t) = -\ln(1 - F(t))$. Denoting by $f(t)$ the density of $F(t)$ with respect to the Lebesgue measure, and setting $\gamma(t) = f(t)(1 - F(t))^{-1}$, we obtain $\Gamma(t) = \int_0^t \gamma(u)\, du$. In view of (18), we have

$$f(t) = -\sum_{j=1}^{K-1} \frac{d\widetilde{p}_{ij}(t)}{dt} = -\sum_{j=1}^{K-1} \frac{dp_{ij}(t)}{dt}\,.$$

Corollary 2.9 *For every $t \in \mathbb{R}_+$ and any $i = 1, \ldots, K-1$, the conditional law of the absorption time τ is given by the formula*

$$\mathbb{Q}\{\tau \leq t \,|\, C_0 = i\} = 1 - \sum_{j=1}^{K-1} p_{ij}(0,t).$$

Remark 2.7 (**Phase-type distributions**) Let $\boldsymbol{\alpha}$ be the initial distribution of C on $\boldsymbol{E} = \{1, \ldots, K-1\}$, i.e. $\boldsymbol{\alpha} = (\mathbb{Q}\{C_0 = j\})_{j=1}^{K-1}$ and assume that $\mathbb{Q}\{C_0 = K\} = 0$. Furthermore, note that we can rewrite the generator Λ as

$$\Lambda = \begin{pmatrix} \boldsymbol{T} & \boldsymbol{t} \\ \mathbf{0} & 0 \end{pmatrix}$$

where $\boldsymbol{t}$ is a column vector with $K-1$ rows. We assume that C_t is transient on $\boldsymbol{E} = \{1, \ldots, K-1\}$ so that $\boldsymbol{T}$ is invertible. Recall the definition of the time of absorbtion $\tau = \inf\{t > 0 : C_t = K\}$ and let $F(t) = \mathbb{Q}\{\tau \leq t\}$. We then say that $F(t)$, is a phase-type distribution (PH-distribution) with representation $(\boldsymbol{E}, \boldsymbol{\alpha}, \boldsymbol{T})$, or for short $(\boldsymbol{\alpha}, \boldsymbol{T})$. Sometime we also say that τ is phase-type distributed with representation $(\boldsymbol{\alpha}, \boldsymbol{T})$. The matrix $\boldsymbol{T}$ is called the *phase generator* and $\boldsymbol{t}$ the *exit vector*. The following proposition can be found in [3] and [43].

Proposition 2.10 *Let F be a PH-distribution with representation $(\boldsymbol{E}, \boldsymbol{\alpha}, \boldsymbol{T})$. Then*

$$\begin{aligned} F(t) &= 1 - \boldsymbol{\alpha} e^{\boldsymbol{T}t} \mathbf{1} \\ f(t) &= \boldsymbol{\alpha} e^{\boldsymbol{T}t} \boldsymbol{t} \\ \mathbb{E}\left[\tau^n\right] &= (-1)^n n! \boldsymbol{\alpha} \boldsymbol{T}^{-n} \mathbf{1} \\ \widehat{F}[s] &= \int_0^\infty e^{-st} f(t) dt = \boldsymbol{\alpha} \left(s\boldsymbol{I} - \boldsymbol{T^{-1}}\right) \boldsymbol{t} \end{aligned}$$

where $f(t)$ is the density of $F(t)$, $\widehat{F}[s]$ is the Laplace-transform of $f(t)$, $\boldsymbol{t} = -\boldsymbol{T}\mathbf{1}$ and $\mathbf{1}^T = (1, 1, \ldots, 1) \in \mathbb{R}^{K-1}$.

Note that the above quantities are computationally tractable given that we use mathematical software with a matrix-package. A short discussion of computing the matrix exponential is given in Appendix in the full version of the paper.

Phase-type distributions are dense (in a weak convergence sense) in the set of all probability distributions on $\mathbb{R}_+ = [0, \infty)$, see [3], [43]. Hence, any random variable on $[0, \infty)$ can be approximated by a properly chosen PH-distribution. Furthermore, PH-distributions are closed under convolution and mixing.

All of the above properties have made phase-type distributions to be an important tool in queuing and reliability theory, but also in insurance, see e.g. [2], [3], [39], [40] and [43]. In Sections 4.1 and 4.2 we will discuss the multivariate extensions of PH-distribution, so called multivariate phase type distributions (MPH) and their applications in portfolio credit risk.

2.6 Martingales Associated with Transitions

We shall now introduce some important examples of martingales associated with the absorption time τ and with the number of transitions. For any fixed $i \neq j$, let H_t^{ij} stand for the number of jumps of the process C from i to j in the interval $(0, t]$. Formally, for any $i \neq j$ we set

$$H_t^{ij} := \sum_{0<u\leq t} H_{u-}^i H_u^j, \quad \forall\, t \in \mathbb{R}_+.$$

The following result is classic (see Brémaud [14], Last and Brandt [33] or Rogers and Williams [42]).

Lemma 2.11 *For every $i, j \in \mathcal{K}$, $i \neq j$, the processes*

$$M_t^{ij} = H_t^{ij} - \int_0^t \lambda_{ij} H_u^i \, du = H_t^{ij} - \int_0^t \lambda_{C_u j} H_u^i \, du \tag{19}$$

and

$$M_t^K = H_t - \int_0^t \sum_{i=1}^{K-1} \lambda_{iK} H_u^i \, du = H_t - \int_0^t \lambda_{C_u K} (1 - H_u) \, du \tag{20}$$

follow $\mathbb{G}$-martingales (and $\mathbb{F}^C$-martingales).

2.7 Change of a Probability Measure

We shall now examine how the Markov property and the generator Λ of the time-homogeneous Markov chain C are affected by a change of the reference probability measure $\mathbb{Q}$ to an equivalent probability measure $\mathbb{Q}^*$ on $(\Omega, \mathcal{G}_{T^*})$ for some fixed $T^* > 0$. Let us emphasize that we do not need to assume here that the state K is absorbing.

Consider a family $\widetilde{\kappa}^{kl}$, $k,l \in \mathcal{K}$, $k \neq l$, of bounded, $\mathbb{F}^C$-predictable, real-valued processes, such that $\widetilde{\kappa}^{kl}_t > -1$. For the sake of notational convenience, we also introduce processes $\widetilde{\kappa}^{kk} \equiv 0$ for $k = 1, \ldots, K$. Let us define an auxiliary $\mathbb{G}$-martingale M (which is also an $\mathbb{F}^C$-martingale) by setting

$$M_t = \int_{]0,t]} \sum_{k,l=1}^K \widetilde{\kappa}^{kl}_u \, dM^{kl}_u = \int_{]0,t]} \sum_{k,l=1}^K \widetilde{\kappa}^{kl}_u \, dH^{kl}_u - M^c_t, \tag{21}$$

where M^c_t is the path-by-path continuous component of the process M, i.e.,

$$M^c_t = \int_0^t \sum_{k,l=1}^K \widetilde{\kappa}^{kl}_u \lambda_{kl} H^k_u \, du.$$

Remark 2.8 From Theorem 21.15 in Rogers and Williams [42], we know that an arbitrary $\mathbb{F}^C$-local martingale M under $\mathbb{Q}$ admits the following representation

$$M_t = \sum_{0<u\leq t} h_u(C_{u-}, C_u) - \int_0^t \sum_{j=1}^K \lambda_{C_{u-}j} \, h_u(C_{u-}, j) \, du,$$

where, for any states $i, j \in \mathcal{K}$, the process $h(i,j)$ is $\mathbb{F}^C$-predictable. In addition, we postulate that $h(j,j) \equiv 0$. Notice that the process M as in (21) can be obtained by setting:

$$h_t(i,j) = \sum_{k,l=1}^K \widetilde{\kappa}^{kl}_t \delta_{ik} \delta_{jl}.$$

Let us return to our problem. We fix a horizon date $T^* < \infty$, and we define an $\mathbb{G}$-martingale η_t, $t \in [0, T^*]$, by postulating that

$$\eta_t = 1 + \int_{]0,t]} \sum_{k,l=1}^K \eta_{u-} \widetilde{\kappa}^{kl}_u \, dM^{kl}_u. \tag{22}$$

It is known that the unique solution to the SDE (22) equals, for every $t \in [0, T^*]$,

$$\eta_t = e^{-M^c_t} \prod_{0<u\leq t} (1 + \Delta M_u).$$

$$\eta_t = e^{-M^c_t} \prod_{0<u\leq t} \Big(1 + \sum_{k,l=1}^K \widetilde{\kappa}^{kl}_u (M^{kl}_u - M^{kl}_{u-})\Big).$$

Observe that

$$1 + \sum_{k,l=1}^K \widetilde{\kappa}^{kl}_u (M^{kl}_u - M^{kl}_{u-}) = 1 + \sum_{k,l=1}^K \widetilde{\kappa}^{kl}_u (H^{kl}_u - H^{kl}_{u-}). \tag{23}$$

Since at most one of the differentials $H^{kl}_u - H^{kl}_{u-}$ is equal to one, and all those that are not equal to one are equal to zero, we see that the right-hand side of (23) is either equal to $1 + \widetilde{\kappa}^{ij}_u$ for some $i \neq j \in \mathcal{K}$, or it is equal to 1. Thus, in view of our assumption that $\widetilde{\kappa}^{kl}_u > -1$ for all $k \neq l$, we conclude that the product

$$\prod_{0<u\leq t} \Big(1 + \sum_{k,l=1}^K \widetilde{\kappa}^{kl}_u (M^{kl}_u - M^{kl}_{u-})\Big)$$

is strictly positive. Consequently, the process η is strictly positive. Since, in addition, $\mathbb{E}_{\mathbb{Q}}(\eta_{T^*}) = 1$, we may define a probability measure $\mathbb{Q}^*$, equivalent to $\mathbb{Q}$ on $(\Omega, \mathcal{G}_{T^*})$, by setting

$$\frac{d\mathbb{Q}^*}{d\mathbb{Q}}\Big|_{\mathcal{G}_{T^*}} = \eta_{T^*}, \quad \mathbb{Q}\text{-a.s.} \tag{24}$$

It is clear that for any date $t \in [0, T^*]$ we have

$$\frac{d\mathbb{Q}^*}{d\mathbb{Q}}\Big|_{\mathcal{G}_t} = \eta_t, \quad \mathbb{Q}\text{-a.s.}$$

Before proceeding further, we need to impose an additional measurability condition on processes $\widetilde{\kappa}^{kl}$, namely, we postulate that, for any fixed $k, l \in \mathcal{K}$ and $t \in \mathbb{R}_+$, the random variable $\widetilde{\kappa}^{kl}_t$ is measurable with respect to the σ-field $\sigma(C_t)$. This implies that, for any fixed $k, l \in \mathcal{K}$ and $t \in \mathbb{R}_+$, there exists a function $g^{kl}_t : \mathcal{K} \to \mathbb{R}$ such that $\widetilde{\kappa}^{kl}_t = g^{kl}_t(C_t)$.

We assume that for any $i \in \mathcal{K}$ there exists a version of $g^{kl}_t(i)$, $t \in \mathbb{R}_+$ that is Borel measurable as a function of t, and we introduce a family of functions $\kappa_{kl} : \mathbb{R}_+ \to (-1, \infty)$ by setting $\kappa_{kl}(t) := g^{kl}_t(k)$ for every $k, l \in \mathcal{K}$ and $t \in \mathbb{R}_+$.

To further simplify the exposition, we shall only consider processes $\widetilde{\kappa}^{kl}_t$ of the special form: $\widetilde{\kappa}^{kl}_t = \kappa_{kl}(t)$, where for every $k, l \in \mathcal{K}$, $k \neq l$, the function $\kappa_{kl} : \mathbb{R}_+ \to (-1, \infty)$ is Borel measurable and bounded. We thus may and do assume that $\kappa_{kk} \equiv 0$ for every $k = 1, \ldots, K$. Under this assumption, we have the following result that provides sufficient conditions for a $\mathbb{G}$-Markov chain C to remain a (time-inhomogeneous, in general) $\mathbb{G}$-Markov chain under $\mathbb{Q}^*$.

Proposition 2.12 *Let the probability measure $\mathbb{Q}^*$ by defined by* (24) *with the Radon-Nikodým density η_{T^*} given by* (22). *Then*
(i) *the process C_t, $t \in [0, T^*]$, is a $\mathbb{G}$-Markov chain under $\mathbb{Q}^*$,*
(ii) *the infinitesimal generator matrix function $\Lambda^*(t) = [\lambda^*_{ij}(t)]_{1 \leq i,j \leq K}$ for C under $\mathbb{Q}^*$ satisfies, for $i \neq j$,*

$$\lambda^*_{ij}(t) = (1 + \kappa_{ij}(t))\lambda_{ij}, \quad \forall\, t \in [0, T^*], \tag{25}$$

and

$$\lambda^*_{ii}(t) = -\sum_{j=1,\, j\neq i}^{K} \lambda^*_{ij}(t), \quad \forall\, t \in [0, T^*], \tag{26}$$

(iii) *the two parameter family $\mathcal{P}^*(t, s)$, $0 \leq s \leq t \leq T^*$, of transition matrices for C relative to $\mathbb{Q}^*$ satisfies the forward Kolmogorov equation*

$$\frac{d\mathcal{P}^*(t, s)}{ds} = \mathcal{P}^*(t, s)\Lambda^*(s), \quad \mathcal{P}^*(t, t) = \text{Id},$$

and the backward Kolmogorov equation

$$\frac{d\mathcal{P}^*(t, s)}{dt} = -\Lambda^*(t)\mathcal{P}^*(t, s), \quad \mathcal{P}^*(s, s) = \text{Id}.$$

It is clear that $\Lambda^*(t) = \Lambda$ for every $t \in [0, T^*]$ if and only if $\kappa_{kl} \equiv 0$ for all $k \neq l$. Letting $\phi_{ij}(t) = 1 + \kappa_{ij}(t)$, we obtain $\lambda^*_{ij}(t) = \phi_{ij}(t)\lambda_{ij}$.

Example 2.9 Suppose that $\kappa_{kl}(t) = \upsilon_k(t)$, for every $k, l \in \mathcal{K}$, $l \neq k$, and $t \in [0, T^*]$. Then we obtain

$$\Lambda^*(t) = U(t)\Lambda, \quad \forall\, t \in [0, T^*],$$

where $U(t)$ is a diagonal K-dimensional matrix, specifically,

$$U(t) = \text{diag}\,[1 + \upsilon_1(t), \ldots, 1 + \upsilon_K(t)].$$

This type of relation between the pricing and real measure was to the best of our knowledge first used in Jarrow, Lando and Turnbull [31], see pp. 495 in [31].

3 Markovian Models of Portfolio Credit Risk

In the remainder of this article we shall review some Markovian models of portfolio credit risk. We shall be primarily concerned with models of dependent defaults, as well as models of the portfolio loss process. However, in Subsection 3.1, we shall describe briefly a model of credit migrations; for a much more comprehensive study of Markovian models of credit migrations we refer to Chapter 12 in [11] and to references therein. The models of dependent defaults discussed in the following sections may, in principle, be considered as special cases of the model of Subsection 3.1.

3.1 Market Model

We begin with a brief description of the Markovian market model that was studied in [9, 13] with view at valuation and hedging of basket credit instruments. This framework allows one to incorporate information relative to the dynamic evolution of credit ratings in the pricing of basket instruments. We begin with some notation.

As before we denote the underlying probability space by $(\Omega, \mathcal{G}, \mathbb{Q})$, where $\mathbb{Q}$ is here and henceforth a risk neutral measure inferred from the market. We endow this space with filtration $\mathbb{G} := \mathbb{H} \vee \mathbb{F}$, which contains all information available to market agents. Specifically, the filtration $\mathbb{H}$ carries information about the evolution of credit events, such as changes in credit ratings or defaults of respective credit names. The filtration $\mathbb{F}$ is a reference filtration containing information pertaining to the evolution of relevant macroeconomic variables.

We consider N obligors (or credit names) and we assume that the current credit quality of each reference entity can be classified into $\mathcal{K} := \{1, 2, \ldots, K\}$ rating categories. By convention, the category K corresponds to default. Let X^l, $l = 1, 2, \ldots, N$ be processes on $(\Omega, \mathcal{G}, \mathbb{Q})$ taking values in the finite state space $\mathcal{K}$. The process X^l represents the evolution of credit ratings of the l^{th} reference entity. We define the *default time* τ_l of the l^{th} reference entity by setting

$$\tau_l = \inf\{\, t > 0 \,:\, X^l_t = K\} \tag{27}$$

We assume that the default state K is absorbing, so that for each name the default event can only occur once.

We denote by $X = (X^1, X^2, \ldots, X^N)$ the joint credit rating process of the portfolio of N credit names. The state space of X is $\mathcal{X} := \mathcal{K}^N$ and the elements of $\mathcal{X}$ will be denoted by $x = \{x^1, \ldots, x^N\}$. We postulate that the filtration $\mathbb{H}$ is the natural filtration of the process X and that the filtration $\mathbb{F}$ is generated by a $\mathbb{R}^n$ valued factor process, Y, representing the evolution of relevant economic variables, like short rate or equity price processes.

We assume that the process $M = (X, Y)$ is jointly Markov under $\mathbb{Q}$, so that we have, for every $0 \leq t \leq s$, $x \in \mathcal{X}$, and any set $\mathcal{Y}$ from the state space of Y,

$$\mathbb{Q}(X_s = x, Y_s \in \mathcal{Y} \,|\, \mathcal{H}_t \vee \mathcal{F}^Y_t) = \mathbb{Q}(X_s = x, Y_s \in \mathcal{Y} \,|\, X_t, Y_t). \tag{28}$$

The process M is constructed as a Markov chain modulated by a Lévy process. We shall refer to X (Y) as the *Markov chain* (*Lévy*) component of M. Given $X_t = x$ and $Y_t = y$, the intensity matrix of the Markov chain component is given by $\Lambda_t = [\lambda(x, x'; y)]_{x' \in \mathcal{X}}$. The Lévy component satisfies the SDE:

$$dY_t = b(X_t, Y_t)\, dt + \sigma(X_t, Y_t)\, dW_t + \int_{\mathbb{R}^n} g(X_{t-}, Y_{t-}, y')\, N(dy', dt),$$

where, for a fixed $y \in \mathbb{R}^n$, $N(dy', dt)$ is a counting process with Lévy measure $\nu(x, y, dy')$ and $\sigma(x, y)$ satisfies $\sigma(x, y)\sigma(x, y)^{\mathsf{T}} = a(x, y)$. We provide the following structure to the generator of the process M.

$$
\begin{aligned}
Af(x,y) = (1/2)\sum_{i,j=1}^{n} a_{ij}(x,y)\partial_i\partial_j f(x,y) + \sum_{i=1}^{n} b_i(x,y)\partial_i f(x,y) \\
+ \int_{\mathbb{R}^n} \big(f(x, y+g(x,y,y')) - f(x,y)\big)\nu(x,y;dy') \\
+ \sum_{x'\in\mathcal{X}\setminus\{x\}} \lambda(x,x';y)f(x',y).
\end{aligned} \tag{29}
$$

We stress that, within the present set-up, the current credit rating of the credit name l directly impacts the intensity of transition of the rating of the credit name l', and vice versa. This property, known as *frailty*, may contribute to default contagion.

Remark 3.1 (Valuation of Basket Credit Derivatives) The model described above can be used to price various basket credit derivatives. In particular, computing the fair spreads of such products involves evaluating conditional expectations, under the risk neutral measure $\mathbb{Q}$, of some quantities related to the cash flows associated to each instrument. For example, the fair spread at time 0 of the CDO equity tranche is (we refer for details to [13], and to Section 3.1.2 for implementation issues):

$$
\kappa_0 = \frac{1}{C_0}\mathrm{E}_{\mathbb{Q}}^{x,y}\Big(\int_0^T \beta_t dM_t^0 - \sum_{j=0}^{J}\beta_{t_j}.05(C_0 - M_{t_j}^0)\Big), \tag{30}
$$

and the fair spread at time 0 of a CDS index is

$$
\eta_0 = \frac{\mathrm{E}_{\mathbb{Q}}^{x,y}\sum_{i=1}^{L}\beta_{\tau_i}(1-\delta)H_T^i}{\mathrm{E}_{\mathbb{Q}}^{x,y}\sum_{j=0}^{J}\beta_{t_j}\sum_{i=1}^{N}\Big(1 - H_{t_j}^i(1-\delta)\Big)}. \tag{31}
$$

3.1.1 Markovian Changes of Measure

For applications like performing changes of numeraire (which can be useful for simplifying the evaluation of quantities of the form (30) or (31, see [9]), or switching between the statistical and a risk-neutral measure (in a situation where the above model $M = (X, Y)$ would in fact be given under the statistical measure, which for pricing purposes would then need to be changed into a risk-neutral measure, see [13]), it is important to be able to apply changes of measure to the model M, whilst preserving Markovianity.

Towards this end we briefly state some facts concerning Markovian changes of measure. Let M_t (such as $M = (X, Y)$ above) be an E valued Markov process under $\mathbb{P}$ with extended generator A (see [41]). In addition define the process

$$
M_t^f := \frac{f(M_t)}{f(M_0)}\exp\left(-\int_0^t \frac{Af(M_s)}{f(M_s)}ds\right). \tag{32}
$$

Definition 3.2 We say that a strictly positive function $f \in \mathcal{D}(A)$ is a good function if M_t^f is a genuine martingale with $\mathbb{E}_{\mathbb{P}}(M_t^f) = 1$.

Let $f \in \mathcal{D}(A)$ and h be a good function in $C(E)$ or $\mathcal{M}_b(E)$ and define the operator

$$
A^h f = h^{-1}A(fh) - fA(h). \tag{33}
$$

In view of Definition 3.2, process M^h may play the rôle of the Radon-Nikodym density between measure $\mathbb{Q}$ and the resulting measure, say $\mathbb{Q}^h$. We have the following result (cf. [41]).

Theorem 3.1 *Let $\mathbb{Q}^h$ be the probability measure associated to the density process M_t^h. Then M_t is a Markov process under $\mathbb{Q}^h$ with extended generator $(A^h, \mathcal{D}(A))$.*

We now apply the above theorem to our model. The domain of $\mathcal{D}(A)$ contains all functions $f(x,y)$ with compact support that are twice continuously differentiable with respect to y. Let h be a good function. By application of Theorem 3.1, the generator of M under $\mathbb{P}^h$ is given as (see [9])

$$A^h f(x,y) = (1/2)\sum_{i,j=1}^{n} a_{ij}(x,y)\partial_i\partial_j f(x,y) + \sum_{i=1}^{n} b_i^h(x,y)\partial_i f(x,y)$$
$$+ \int_{\mathbb{R}^n} \left(f(x,y+g(x,y,y')) - f(x,y)\right)\nu^h(x,y;dy') + \sum_{x'\in\mathcal{X}} \lambda^h(x,x';y)f(x',y),$$

where

$$\begin{aligned} b_i^h(x,y) &= b_i(x,y) + \frac{1}{h(x,y)}\sum_{i,j=1}^{n} a_{ij}(x,y)\partial_j h(x,y), \\ \nu^h(x,y;dy') &= \frac{h(x,y+g(x,y,y'))}{h(x,y)}\,\nu(x,y;dy'), \\ \lambda^h(x,x';y) &= \lambda(x,x';y)\frac{h(x',y)}{h(x,y)},\ x\neq x', \quad \lambda^h(x,x;y) = -\sum_{x'\neq x}\lambda^h(x,x';y). \end{aligned} \tag{34}$$

3.1.2 Model Implementation

Expectations such as the ones appearing in the fair spread valuation formulae (30) or (31) above, can in principle be computed by numerical resolution of the related Kolmogorov valuation systems of reaction-diffusions equations, or, in the special case of a time-homogeneous Markov chain model, by numerical exponentiation of the model generator (see Sections 3.2-3.2.1). However all these analytical methods are limited in practice to low-dimensional models by the curse of dimensionality. In general simulation methods are then the only viable alternative. Implementation of the above Markovian model via simulation will now be described in the special case where the dynamics of the factor process Y do not depend on the credit migrations process X. The general case appears to be much harder.

3.1.3 Specification of Credit Ratings Transition Intensities

In order to alleviate the simulation of the model we specify the credit migrations intensity measure λ in (29) to be of the following form:

$$\sum_{x'\in\mathcal{X}\setminus\{x\}} \lambda(x,x';y)f(x',y) = \sum_{l=1}^{N}\sum_{\xi\in\mathcal{K}} \lambda^l(x,x_l^\xi;y)f(x_l^\xi,y), \tag{35}$$

where we write $x_l^\xi = (x^1, x^2, \ldots, x^{l-1}, \xi, x^{l+1}, \ldots, x^N)$.

Note that the model specified by (35) does not allow for simultaneous jumps of the components X^l and $X^{l'}$ for $l \neq l'$. In other words, the ratings of different credit names may not change simultaneously. The advantage is that, for the purpose of simulation of paths of process X, rather than dealing with $\mathcal{X}\times\mathcal{X}$ intensity matrix $[\lambda(x,x';y)]$, we shall deal with N intensity matrices $[\lambda^l(x,x_l^\xi;y)]$, each of dimension $\mathcal{K}\times\mathcal{K}$ (for any fixed y).

We now provide further structure to the generator of the Markov chain component of the joint process $M=(X,Y)$ and specify a general functional form for its transition intensities. We shall then briefly describe a recursive procedure for simulating the evolution of the process X.

Because we need to simulate the joint process (X,Y), it is important to specify its form in such a way to avoid unnecessary computational complexity. As noted earlier, the structure of the generator A that we postulate makes it so that simulation of the evolution of process X reduces to recursive simulation of the

evolution of processes X^l, whose state spaces are only of size K each. In order to facilitate simulations even further, we also postulate that each migration process X^l behaves like a birth-and-death process with absorption at default, and with possible jumps to default from every intermediate state. In addition, we shall assume that the factor process, Y, is independent of X. Conditional upon $(X_t, Y_t) = (x, y)$, the infinitesimal generator governing the evolution of the credit ratings of the l^{th} name is the sub-stochastic matrix:

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ \\ \vdots \\ K-1 \\ K \end{array} \begin{array}{c} \begin{array}{ccccccc} 1 & 2 & 3 & \cdots & K-1 & K \end{array} \\ \begin{pmatrix} \lambda^l_{1,1} & \lambda^l_{1,2} & 0 & \cdots & 0 & \lambda^l_{1,K} \\ \lambda^l_{2,1} & \lambda^l_{2,2} & \lambda^l_{2,3} & \cdots & 0 & \lambda^l_{2,K} \\ 0 & \lambda^l_{3,2} & \lambda^l_{3,3} & \cdots & 0 & \lambda^l_{3,K} \\ \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda^l_{K-1,K-1} & \lambda^l_{K-1,K} \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \end{array},$$

where, with a slight change of notation, $\lambda^l_{x^l,\xi} = \lambda^l_{x^l,\xi}(x, y) = \lambda^l(x, x^\xi_l; y)$. The functional form of the transition intensities should reflect the specific characteristics of the instruments we need to price and should be chosen to obtain the best possible fit in the calibration phase.

3.1.4 Simulation Algorithm

In general, a simulation of the evolution of the process X entails high computational costs, as the the cardinality of the state space of X is equal to K^N. Thus, for example, in case of $K = 18$ rating categories, as in Moody's ratings, and in case of a portfolio of $N = 100$ credit ratings, the state space has 18^{100} elements. However, the specific assumptions on the structure of the generator allow to simulate the process in a recursive fashion, which has a relatively low computational complexity. We consider here simulations of sample paths over a generic time interval, $[t_1, t_2]$, where $0 \le t_1 < t_2$, and assume that the time t_1 state of the process (X, Y) is (x, y). Generating one sample path will, in general, involve the following steps:

Step 1: in Step 1, a sample path of the process Y is simulated. Recall that the dynamics of the factor process are described by the SDE

$$\begin{aligned} dY_t &= b(Y_t)\,dt + \sigma(Y_t)\,dW_t + \int_{\mathbb{R}^n} g(Y_{t-}, y')\,N(dy', dt) \\ Y_{t_1} &= y \end{aligned}$$

Any standard procedure can be used to simulate a sample path of Y (the reader is referred, for example, to Cont and Tankov [17]). We denote by $\widehat{Y}$ the simulated sample path of Y.

Step 2: generate a sample path of X on the interval $[t_1, t_2]$.

Step 2.1: simulate the first jump time of the process X in the time interval $[t_1, t_2]$. Towards this end, draw from a unit exponential distribution. We denote by $\widehat{\eta}_1$ the value of the first draw. The simulated value of the first jump time, τ, is then given by:

$$\tau = \inf\Big\{ t > t_1 \,:\, \int_{t_1}^{t} \lambda(x, \widehat{Y}_u)\,du \ge \widehat{\eta}_1 \Big\},$$

where

$$\lambda(x, \widehat{Y}_t) := -\sum_{i=1}^{N} \lambda^i_{x^i,x^i}(x, \widehat{Y}_t)$$

and

$$\lambda^l_{x^l,x^l}(x,\widehat{Y}_t) = -\lambda^l_{x^l,x^l-1}(x,\widehat{Y}_t) - \lambda^l_{x^l,x^l+1}(x,\widehat{Y}_t) - \lambda^l_{x^l,K}(x,\widehat{Y}_t).$$

If $\tau > t_2$ return to step 1, otherwise go to Step 2.2.

Step 2.2: simulate which component of the vector process X jumps at time τ, by drawing from the conditional distribution:

$$(X^l_\tau \neq X^l_{\tau-}) = -\frac{\lambda^l_{x^l,x^l}(x,\widehat{Y}_\tau)}{\lambda(x,\widehat{Y}_\tau)} \tag{36}$$

Recall that $\lambda^l_{x^l,x^l}(x,\widehat{Y}_t) = 0$ if $x^l = K$, since K is an absorbing state.

Step 2.3: assume the i^{th} obligor jumps at τ. Simulate the direction of the jump by drawing from the conditional distribution

$$Q^i(X^i_\tau = \xi) = -\frac{\lambda^i_{x^i,\xi}(x,\widehat{Y}_\tau)}{\lambda^i_{x^i,x^i}(x,\widehat{Y}_\tau)} \tag{37}$$

where

$$\xi = \{x^i - 1; x^i + 1; K\}$$

Step 2.4: update the state of X and set $t_1 = \tau$. Repeat Steps 2.1-2.3 on $[t_1, t_2]$ until $\tau > t_2$

Step 3: calculate the simulated value of a relevant functional. For instance, assume that Y represents the short rate process, and is used as a discount factor, i.e $\int_0^t Y_t = -\ln B_t$. In order to compute the protection leg of a CDS index, one would evaluate

$$\sum_{i=1}^{L} \frac{B_{\tau_i}}{B_t}(1-\delta)(H^i_T - H^i_t)(\omega)$$

at each run ω, and obtain the Monte Carlo estimate by averaging over all sample paths.

Remark 3.3 An important issue in regard to simulation is *variance reduction*. Importance sampling is often regarded as the method of choice when it comes to variance reduction. Importance sampling and related particle methods for Markovian models of portfolio credit risk (in the Homogeneous Groups Model of Section 3.2 particularly) are dealt with in [15].

3.2 Homogeneous Groups Model

We now describe in some detail a more specific version of the previous model, considered for different purposes by various authors in [8, 15, 20, 25], among others. In this specification of the model, there is no factor process Y involved. We thus deal with a continuous-time Markov Chain denoted in this subsection by $\mathcal{N}$ (cf. X above), relative to the filtration and $\mathbb{F} = \mathbb{F}^{\mathcal{N}}$.

More precisely, a pool of n credit names is organized in d homogeneous groups of $(\nu - 1)$ obligors (so $n = (\nu-1)d$, assuming $\frac{n}{d}$ integer), and N^l represents the number of defaulted obligors in the l^{th} group (instead of X^l representing the credit rating of obligor l previously; so the interpretation of the Markov chain has changed, but the mathematical structure of the model is preserved). Moreover we assume that the N^l's can only jump one at a time *and by one*, so that we in fact deal with a d-variate Markov point process $\mathcal{N} = (N^1, \cdots, N^d)$. For each l, the ($\mathbb{F}$-)intensity of N^l is assumed to be of the form

$$\lambda^l(\mathcal{N}_t) = (\nu - 1 - N^l_t)\widetilde{\lambda}^l(\mathcal{N}_t)\ , \tag{38}$$

for an *aggregated intensity function* $\lambda^l = \lambda^l(\imath)$, and *pre-default individual* intensity function $\widetilde{\lambda}^l(\imath)$, where $\imath = (i_1, \cdots, i_d) \in \mathcal{I} = \{0, 1, \cdots, \nu - 1\}^d$. We thus deal with a d-variate Markov point process $\mathcal{N} =$

$(N^1, \cdots, N^d)$.

Since we assume that there are no common jumps between processes N^l, so the jump intensities λ^l are in one-to-one correspondence with the generator Λ of $\mathcal{N}$, which consists of a $\nu^d \otimes \nu^d$ matrix Λ (a very sparse matrix, since the components of $\mathcal{N}$ may only jump by one and only one at a time).

For $d = 1$, we recover the so called *Local Intensity Model* (birth-and-death process stopped at level n) used by Laurent, Cousin and Fermanian [34], Cont and Minca [16] or Herbertsson [25] for modeling a credit portfolio cumulative default process N. This model, which will be intensively used in Sections 4.1–4.2, can be considered as the analog for credit derivatives of the local volatility model for equity and equity index derivatives (analogous in the sense that at any given time in any loss derivatives model, there exists a local intensity model with the same marginals for the portfolio loss process, see Cont and Minca [16] and Gyöngy [23]).

At the other end of the spectrum, for $d = n$ (i.e. when each group has only a single element), we are in effect modeling the vector of default indicator processes $H = (H^i)_{1 \leq i \leq n}$ of the pool.

As d varies between 1 and n, we thus get a variety of models of credit risk, ranging from pure 'top-down' models for $d = 1$, to pure 'bottom-up' models for $d = n$ (see [8]). Introducing parsimonious parameterizations of the intensities allows one to account for inhomogeneity between groups, and/or for defaults contagion.

Other examples related to intensities as in (38) can be found in Section 2.3 and Section 2.5 in [6], where a portfolio with 125 obligors is split in first three and then six subportfolios. Both examples uses intensities similar to those in equation (38), with the aim to price CDO-tranches.

3.2.1 Pricing in the Homogeneous Groups Model

Since $\mathcal{N}$ is a Markov process and the portfolio cumulative default process N is a function of $\mathcal{N}$, the model price process of a (stylized) loss derivative (protection leg of a CDO tranche, say) with payoff $\phi(N_T)$ writes, for $t \in [0, T]$:

$$\Pi_t = \mathbb{E}(\phi(N_T)|\mathcal{F}_t) = u(t, \mathcal{N}_t) \ , \tag{39}$$

where $u(t, \imath)$ or $u_\imath(t)$ for $t \in [0, T]$ and $\imath \in \mathcal{I} = I^d$, is the *pricing function* (system of time-functionals $u_\imath$), solution to the following *pricing equation* (system of ODEs) with generator Λ:

$$(\partial_t + \Lambda)u = 0 \text{ on } [0, T) \ ,$$

with terminal condition $u_\imath(T) = \phi(\imath)$, for $\imath \in \mathcal{I}$.
Likewise, the groups losses distribution at time t, that is, $q_\imath(t) = \mathbb{Q}(\mathcal{N}_t = \imath)$ for $t \in [0, T]$ and $\imath \in \mathcal{I}$, can be characterized in terms of the associated forward Kolmogorov equations (see, e.g., [15]).
These pricing and transition probability backward and forward Kolmogorov equations can then be solved by various means, like numerical matrix exponentiation (since the model is time-homogeneous, see Appendix in the full version of the paper).

However, even if the matrix Λ is very sparse, its size is prohibitive in most cases as far as deterministic numerical methods are concerned. For instance, in the case of $d = 5$ groups of $\nu - 1 = 25$ names, one gets $\nu^{2d} = 26^{10}$. So for high values of d, Monte Carlo methods as of Section 3.1.4 appear to be the only viable computational alternative. Appropriate variance reduction techniques may help in this regard (cf. Remark 3.3).

3.3 Markov Copulae

Modeling of stochastic dependence between evolutions of credit migration processes (default processes, in particular) in a pool of credit names is of key importance of course. In the Markovian model presented in the

previous section, the dependence was modeled in a way that did not, automatically, guarantee that desired marginal distributions (of credit migration processes of individual obligors) are preserved. In this section we shall present a methodology that, in the Markovian framework, allows for construction of a multivariate process with given univariate marginals. For mathematical details pertaining to this section we refer to [10].

To simplify notation we shall only consider the case of bivariate Markov chains. The general multivariate case can be treated accordingly (it is even possible to treat the case of a general joint Markov process $M = (X, Y)$ as of Section 3.1, see [12]).

Given a bivariate processes $Z = (X, Y)$, which is a finite Markov chain with respect to to its natural filtration $\mathbb{F}^Z = \mathbb{F}^{X,Y}$, one is naturally confronted with the following two questions, which we shall address in this section:

(Q1): what are the sufficient and necessary conditions on the infinitesimal generator of Z so that the components X and Y are Markov chains with respect to their natural filtrations?

(Q2): how do we construct a bivariate Markov chain, whose components are themselves Markov chains w.r.t their natural filtration and have desired infinitesimal characteristics?

We denote by $\mathcal{S}$ and $\mathcal{O}$ two finite sets. Let $Z = (X, Y)$ denote a two dimensional Markov chain on $\mathcal{Z} = \mathcal{S} \times \mathcal{O}$, with generator function $A^Z(t) = [\lambda^{ih}_{jk}(t)]_{i,j\in\mathcal{S},k,h\in\mathcal{O}}$. Consider the following condition

Condition (M)

$$\sum_{k\in\mathcal{O}} \lambda^{ih}_{jk}(t) = \sum_{k\in\mathcal{O}} \lambda^{ih'}_{jk}(t), \ \ \forall h, h' \in \mathcal{O}, \ \forall i, j, \in \mathcal{S}, \ i \neq j,$$

and

$$\sum_{j\in\mathcal{S}} \lambda^{ih}_{jk}(t) = \sum_{j\in\mathcal{S}} \lambda^{i'h}_{jk}(t), \ \ \forall i, i' \in \mathcal{S}, \ \forall k, h \in \mathcal{O} \ \ h \neq k.$$

The following proposition addresses the sufficiency part in question (Q1),

Proposition 3.2 *Suppose that condition (M) holds, and define*

$$f^i_j(t) := \sum_{k\in\mathcal{O}} \lambda^{ih}_{jk}(t), \quad i, j \in \mathcal{S}, \ i \neq j, \quad f^i_i(t) = -\sum_{j\in\mathcal{S}, j\neq i} f^i_j(t), \quad \forall i \in \mathcal{S},$$

and

$$g^h_k(t) := \sum_{j\in\mathcal{S}} \lambda^{ih}_{jk}(t), \quad k, h \in \mathcal{O}, \ h \neq k, \quad g^h_h(t) = -\sum_{k\in\mathcal{O}, k\neq h} g^h_k(t), \quad \forall h \in \mathcal{O}.$$

Then the components X and Y of the Markov chain Z are Markov chains with respect to their natural filtrations with generator functions $A^X(t) = [f^i_j(t)]_{i,j\in\mathcal{S}}$ and $A^Y(t) = [g^h_k(t)]_{k,h\in\mathcal{O}}$, respectively.

For the necessity part of question (Q1) we have

Proposition 3.3 *For the components X and Y of the Markov chain Z to be Markov chains with respect to their natural filtrations, with generator functions $A^X(t) = [f^i_j(t)]_{i,j\in\mathcal{S}}$ and $A^Y(t) = [g^h_k(t)]_{k,h\in\mathcal{O}}$, respectively, it is necessary that the following conditions hold for almost all $t \geq 0$, $\mathbb{Q} - a.s.$:*

$$\left({}^{o_x}\Lambda^{ij}\right)^{p_x}_t = \int_0^t \mathbb{1}_{\{X_{u-}=i\}} f^i_j(u)du, \tag{40}$$

$$\left({}^{o_y}\Gamma^{hk}\right)^{p_y}_t = \int_0^t \mathbb{1}_{\{Y_{u-}=i\}} g^h_k(u)du, \tag{41}$$

where ${}^{o_x}(\cdot)$ $[(\cdot)^{p_x}]$ and ${}^{o_y}(\cdot)$ $[(\cdot)^{p_y}]$ denote the optional [predictable] projection on $\mathbb{F}^X$ and $\mathbb{F}^Y$ respectively, and where

$$\Lambda^{ij}_t = \int_0^t \sum_{k\in\mathcal{O}} \sum_{h\in\mathcal{O}} \mathbb{1}_{\{X_{u-}=i, Y_{u-}=h\}} \lambda^{ih}_{jk}(u)du$$

and

$$\Gamma_t^{hk} = \int_0^t \sum_{j\in\mathcal{S}} \sum_{i\in\mathcal{S}} \mathbb{1}_{\{X_{u-}=i, Y_{u-}=h\}} \lambda_{jk}^{ih}(u) du.$$

The following corollary addresses question (Q2),

Corollary 3.4 *Consider two Markov chains X and Y, with respect to their own filtrations, and with values in $\mathcal{S}$ and $\mathcal{O}$, respectively. Suppose that their respective generators are $A^X(t) = [\alpha_j^i(t)]_{i,j\in\mathcal{S}}$ and $A^Y(t) = [\beta_k^h(t)]_{h,k\in\mathcal{O}}$. Next, consider the system of equations in the unknowns $\lambda_{jk}^{ih}(t)$, where $i, j \in \mathcal{S}$, $h, k \in \mathcal{O}$ and $(i,h) \neq (j,k)$:*

$$\sum_{k\in\mathcal{O}} \lambda_{jk}^{ih}(t) = \alpha_j^i(t), \ \forall h \in \mathcal{O}, \ \forall i, j \in \mathcal{S}, \ i \neq j \tag{42}$$

$$\sum_{j\in\mathcal{S}} \lambda_{jk}^{ih}(t) = \beta_k^h(t), \ \forall i \in \mathcal{S}, \ \forall h, k \in \mathcal{O}, \ h \neq k. \tag{43}$$

Suppose that the above system admits solution such that the matrix function $A(t) = [\lambda_{jk}^{ih}(t)]_{i,j\in\mathcal{S},k,h\in\mathcal{O}}$, with

$$\lambda_{ih}^{ih}(t) = - \sum_{(j,k)\in\mathcal{S}\times\mathcal{O},(j,k)\neq(i,h)} \lambda_{jk}^{ih}(t), \tag{44}$$

properly defines an infinitesimal generator function of a Markov chain with values in $\mathcal{S}\times\mathcal{O}$. Consider, a bivariate Markov chain $Z := (Z_1, Z_2)$ on $\mathcal{S}\times\mathcal{O}$ with generator function $A^Z(t) = A(t)$. Then, the components Z_1 and Z_2 are Markov chains with respect to to their own filtrations, and their generators are $A^{Z_1}(t) = A^X(t)$ and $A^{Z_2}(t) = A^Y(t)$.

Note that, typically, system (42) –(43) contains many more unknowns than equations. In fact, given that cardinalities of $\mathcal{S}$ and $\mathcal{O}$ are $K_\mathcal{S}$ and $K_\mathcal{O}$, respectively, the system consists of $K_\mathcal{S}(K_\mathcal{S}-1) + K_\mathcal{O}(K_\mathcal{O}-1)$ equations in $K_\mathcal{S}K_\mathcal{O}(K_\mathcal{S}K_\mathcal{O}-1)$ unknowns.

Thus, in principle, one can create several bivariate Markov chains Z with the given margins X and Y. Thus, indeed, the system (42) –(43) essentially serves as a "copula" between the Markovian margins X, Y and the bivariate Markov chain Z. This observation leads to the following definition,

Definition 3.4 A *Markov copula* between the Markov chains X and Y is any solution to system (42) –(43) such that the matrix function $A(t) = [\lambda_{jk}^{ih}(t)]_{i,j\in\mathcal{S},k,h\in\mathcal{O}}$, with $\lambda_{ih}^{ih}(t)$ given in (44), properly defines an infinitesimal generator function of a Markov chain with values in $\mathcal{S}\times\mathcal{O}$.

Different Markov copulae will entail different dependence structure between the margins X and Y.

Markovian Changes of Measure For pricing purposes the probability $\mathbb{P}$ above typically denotes the statistical probability, which needs to be changed to the EMM. Typically, the Radon-Nikodym density is chosen in such a way that the resulting (risk-neutral) default probabilities are consistent with the term structure of CDS spreads. In addition, we require that the process Z, which is Markovian under the statistical measure, is also Markovian under the pricing measure. As a consequence, such change of measure must be chosen with some care.

In the case of a finite state Markov chain, Theorem 3.1 yields the following corollary (cf. [41])

Corollary 3.5 *Let Z_t be a finite state Markov chain on $\mathcal{K}$ with cardinality K and generator $A = a_{ij}$. In addition let $h = (h_1, \ldots, h_K)$ be a positive vector. Then Z_t is a Markov process under $\mathbb{Q}^h$with generator $A^h = [a_{ij}h_jh_i^{-1}]$.*

Remark 3.5 We note that in case of Markov chains the formula for the Markovian change of measure presented in Proposition 2.12 (cf. formula (25)) appears to be more general than the change of measure resulting from the theory of Palmowski and Rolski presented in [41], and leading to $A^h = [a_{ij}h_jh_i^{-1}]$. However, for the purpose of the following section about ratings triggered bonds, this less general result is sufficient.

3.3.1 Application to Ratings Triggered Step-Up Bonds

This application is taken from [12], to which we refer the reader for more examples of application of the above theory to valuation and hedging of credit derivatives. Even though the ratings triggered setp-up bonds are not credit derivatives per se, we nevertheless give a brief account of their valuation, as the techniques exploited in this case may as well be used for valuation of credit derivatives whose cash flows may depend on history of credit ratings assigned to an obligor by various rating agencies.

Ratings triggered step-up bonds were issued by some European telecom companies in the recent 6-7 years. As of now, to our knowledge, these products are not traded in baskets, however they are of interest because they offer protection against credit events other than defaults. In particular, ratings triggered corporate step-up bonds (step-up bonds for short) are corporate coupon issues for which the coupon payment depends on the issuer's credit quality: in principle, the coupon payment increases when the credit quality of the issuer declines. In practice, for such bonds, credit quality is reflected in credit ratings assigned to the issuer by at least one credit ratings agency (Moody's-KMV or Standard&Poor's). The provisions linking the cash flows of the step-up bonds to the credit rating of the issuer have different step amounts and different rating event triggers. In some cases, a step-up of the coupon requires a downgrade to the trigger level by both rating agencies. In other cases, there are step-up triggers for actions of each rating agency. Here, a downgrade by one agency will trigger an increase in the coupon regardless of the rating from the other agency. Provisions also vary with respect to step-down features which, as the name suggests, trigger a lowering of the coupon if the company regains its original rating after a downgrade. In general, there is no step-down below the initial coupon for ratings exceeding the initial rating.

Let R_t stand for some indicator of credit quality at time t (note that in this case, the process R may be composed of two, or more, distinct rating processes). Assume that t_i, $i = 1, 2, \ldots, n$ are coupon payment dates. In this paper we assume the convention that coupon paid at date t_n depends only on the rating history through date t_{n-1}, that is: $c_n = c(R_t,\, t \leq t_{n-1})$ are the coupon payments. In other words, we assume that no accrual convention is in force.

Assuming that the bond's notional amount is $1,$ the cumulative discounted cash flow of the step-up bond is (as usual we assume that the current time is 0):

$$(1 - H_T)\beta_T + \int_{(0,T]} (1 - H_u)\beta_u \, dC_u + \beta_\tau Z_\tau H_T, \tag{45}$$

where $C_t = \sum_{t_i \leq t} c_i$, τ is the bond's default time, $H_t = \mathbb{1}_{\tau \leq t},$ and where Z_t is a (predictable) recovery process.

Pricing Ratings Triggered Step-Up Bonds via Simulation Here, using our results on Markov copulae, we shall apply a simulation approach to pricing ratings triggered step-up bonds.

Let us consider a ratings triggered step-up bond issued by an obligor XYZ. Recall that, typically, cash-flows associated with a step-up bond depend on ratings assigned to XYZ by both Moody's Investors Service (Moody's in what follows) and Standard & Poor's (S&P in what follows). Thus, a straightforward way to model joint credit migrations would be to consider a credit migration process K such that $R_t = (M_t, SP_t)$, where M_t and SP_t denote the time t credit rating assigned to XYZ by Moody's and SP_t, respectively. We assume that process M is a time-homogeneous Markov chain w.r.t. its natural filtration, under the statistical probability $\mathbb{P},$ and that its state space is $\mathcal{K} = \{1, 2, \ldots, K\}$. Likewise, we assume that process SP is a time-homogeneous Markov chain w.r.t. its natural filtration, under the statistical probability $\mathbb{P},$ and that its state space is $\mathcal{K} = \{1, 2, \ldots, K\}$.

Typically, we are only provided with individual statistical characteristics of each of the processes M and SP. Thus, in a sense, we know the marginal distributions of the joint process R under the measure $\mathbb{P}$ (where M and SP are considered as the "univariate" margins). The crucial issue is thus the appropriate modeling of dependence between processes M and SP. In particular, we want to model dependence, under $\mathbb{P},$ between M and SP so that the joint process R is a time-homogeneous Markov chain, and so that the components M and SP are time-homogeneous Markov chains with given $\mathbb{P}$-generators, say A^M and A^{SP}, respectively.

Thus, essentially, we need to model a $\mathbb{P}$-generator matrix, say A^R, so that process R is a time-homogeneous Markov chain with $\mathbb{P}$-generator A^R and that processes M and SP are time-homogeneous Markov chains with $\mathbb{P}$-generators A^M and A^{SP}. We can of course deal with this problem using the theory of Markov copulae.

Towards this end, we fix an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$. On this space we consider two univariate Markov chains M and SP, with given infinitesimal $\mathbb{P}$-generators $A^M = [a^M_{ij}]$ and $A^{SP} = [a^{SP}_{hk}]$, respectively. Next, we consider the system equations in variables

$$\sum_{k\in\mathcal{K}} a^R_{ih,jk} = a^M_{ij}, \ \forall i, j \in \mathcal{K}, i \neq j, \ \forall h \in \mathcal{K}, \tag{46}$$

$$\sum_{j\in\mathcal{K}_1} a^R_{ih,jk} = a^{SP}_{hk}, \ \forall h, k, \in \mathcal{K}, h \neq k, \ \forall i \in \mathcal{K}. \tag{47}$$

Now, provided that the system (46) –(47) has a positive solution, then it follows from Corollary 3.4 that resulting matrix[2] $A^R = [a^R_{ih,jk}]_{i,j\in\mathcal{K}_1,\, h,k\in\mathcal{K}_2}$ satisfies conditions for a $\mathbb{P}-$generator matrix of a bivariate time-homogenous Markov chain, say $R = (R^1, R^2)$ whose components take values in finite state spaces $\mathcal{K}_1$ and $\mathcal{K}_2$ with cardinalities K_1 and K_2, respectively, and, more importantly, they are Markov chains with the same distributions as M and SP under under $\mathbb{P}$. Thus, indeed, the system (46)–(47) essentially serves as a Markov copula between the Markovian margins M, SP and the bivariate Markov chain R.

Note that, typically, the system (46)–(47) contains many more variables than equations. Thus, one can create several bivariate Markov chains R with the given margins M and SP. In financial applications this feature leaves a lot of room for various modeling options and for calibration of the model. For example, as observed by Lando and Mortensen [35] although the ratings assigned by S&P and Moody's to the same company do not necessarily coincide, split ratings are rare and are usually only observed in short time intervals. This feature can easily be modelled using the Markovian copula system (46) –(47) via imposing side constraints for the unknowns $a^R_{ih,jk}$'s. In order to model such observed behavior of the joint rating process, we thus impose additional constraints on the variables in the system (46) –(47). Specifically, we postulate that

$$a^R_{ih,jk} = \begin{cases} 0, & \text{if } i \neq j \text{ and } h \neq k \text{ and } j \neq k, \\ \alpha \min(a^M_{ij}, a^{SP}_{hk}), & \text{if } i \neq j \text{ and } h \neq k \text{ and } j = k, \end{cases} \tag{48}$$

where $\alpha \in [0, 1]$ is a modelling parameter. Using constraint (48) we can easily solve system (46) –(47) (in this case the system actually becomes fully decoupled) and we can obtain the generator of the joint process. The interpretation of constraint (48) is the following: The components M and SP of the process R migrate according to their marginal laws, but they tend to join, that is, they tend to both take the same values. The strength of such tendency is measured by the parameter α. When $\alpha = 0$ then, in fact, the two components are independent processes; when $\alpha = 1$ the intensity of both components migrating simultaneously to the same rating category is maximum (given the specified functional form for the intensities of common jumps).

For pricing purposes the statistical probability measure is changed to the EMM. Typically, the Radon-Nikodym density is chosen in such a way that the resulting (risk-neutral) default probabilities are consistent with the term structure of CDS spreads. In addition, we require that the process R, which is Markovian under the statistical measure, is also Markovian under the pricing measure.

We recall that $A^R = [a^R_{ih,jk}]$ is the generator of R under the statistical measure $\mathbb{P}$. In view of Corollary 3.5, or, more generally, in view of formula (25) in Proposition 2.12, given a vector $h = [h_{11}, \cdots, h_{KK}] \in \mathbb{R}^{K^2}$, we can change statistical measure $\mathbb{P}$ to an equivalent "risk-neutral" measure $\mathbb{Q}$ in such a way that process R is a time-homogeneous Markov chain under $\mathbb{Q}$, and its $\mathbb{Q}$-infinitesimal generator is given by

$$\widetilde{A}^R = [\widetilde{a}_{ih,jk}],$$

where $\widetilde{a}_{ih,jk} = a_{ih,jk}\frac{h_{jk}}{h_{ih}}$ for $ih \neq jk$ and $\widetilde{a}_{ih,jk} = -\sum_{jk\neq ih} a_{ih,jk}\frac{h_{jk}}{h_{ih}}$ for $ih = jk$.

[2] System (46) –(47) does not include diagonal elements of A^R. These elements are obtained as $a^R_{ih,ih} = -\sum_{(j,k)\in\mathcal{K}} a^R_{ih,jk}$.

Remark 3.6 Not that, although the change of measure preserves Markov property of the joint process R, its components may not be Markov (in their natural filtration) under the new probability measure. This however is not an issue for us, as all we need to conduct computations is the Markov property of the joint process R under the new measure.

An arbitrary choice of vector h may lead to a heavy parametrization of the pricing model. We suggest that the vector h_{ij} be chosen as follows:

$$h_{ij} = \exp(\alpha_1 i + \alpha_2 j), \quad \forall i, j \in \mathcal{K},$$

where α_1 and α_2 are parameters to be calibrated. It turns out, as the calibration results provided in [12] indicate, that this is a good choice.

Remark 3.7 Note that the formalism of Markovian copulae can be exploited in a wide range of applications. See for instance [18] for a recent use in the modeling of *counterparty credit risk*.

4 Multivariate phase-type distributions and matrix-analytical methods

In the remaining sections of this article we consider the intensity-based models for default contagion (with constant coefficients) which are studied in [25, 26, 27, 28, 29]. These intensity-based models are reinterpreted in terms of a time-homogeneous Markov jump process, a so called multivariate phase-type distribution, introduced in [4]. The translation makes it possible to use a matrix-analytic approach to derive practical formulas for all quantities that we want to study. To be more specific, we present convenient analytical formulas for multivariate default and survival distributions, conditional multivariate distributions, marginal default distributions, multivariate default densities, default correlations, and expected default times. Furthermore, computationally tractable closed-form expressions for credit derivatives such as synthetic CDO tranche spreads, index CDS spreads, k^{th}-to-default swap spreads and single-name CDS spreads can also be obtained. However, these formulas are omitted in this article and we instead refer to the details in the litterature.

Subsection 4.1 is devoted to inhomogeneous portfolios, whilst Subsection 4.2 deals with homogenous portfolios. Section 5 presents numerical results for some of the above quantities, in a homogeneous portfolio calibrated against CDO tranches from the iTraxx Europe series.

Note that multivariate phase-type distributions (MPH) can be viewed as a special case of the Markov models presented in Section 2 and Section 3. However, the practical formulas that can be derived in MPH settings, used particulary in reliability and queuing theory, has made MPH-distributions and matrix-analytical approaches, to grow into a subject of its own, see e.g. [2], [3], [4], [39], [40] and [43].

4.1 Inhomogeneous Portfolios

In this subsection we study inhomogeneous credit portfolios. First, Subsection 4.1.1 presents an intensity-based model with intuitive and explicit contagion effects. This model is then reformulated into a time-homogeneous Markov jump process which is used in Subsection 4.1.2 - 4.1.4 in order to find practical formulas for multivariate default and survival distributions, conditional multivariate distributions, marginal default distributions, multivariate default densities, default correlations, and expected default times.

Subsection 4.1.6 shortly discuss the calibration of the parameters in our framework. Finally, Subsection 4.1.7 outlines alternative parameterizations of the model presented in Subsection 4.1.1

4.1.1 Intensity based models reinterpreted as Markov jump processes

For the default times $\tau_1, \tau_2 \ldots, \tau_m$, define the point process $N_{t,i} = 1_{\{\tau_i \leq t\}}$ and introduce the filtrations

$$\mathcal{F}_{t,i} = \sigma\left(N_{s,i}; s \leq t\right), \quad \mathcal{F}_t = \bigvee_{i=1}^{m} \mathcal{F}_{t,i}.$$

Let $\lambda_{t,i}$ be the $\mathcal{F}_t$-intensity of the point processes $N_{t,i}$. Below, we will for convenience often omit the filtration and just write intensity or "default intensity". With a further extension of language we will sometimes also write that the default times $\{\tau_i\}$ have intensities $\{\lambda_{t,i}\}$. The model studied in this section is specified by requiring that the default intensities have the following form,

$$\lambda_{t,i} = a_i + \sum_{j \neq i} b_{i,j} 1_{\{\tau_j \leq t\}}, \qquad t \leq \tau_i, \tag{49}$$

and $\lambda_{t,i} = 0$ for $t > \tau_i$. Further, $a_i \geq 0$ and $b_{i,j}$ are constants such that $\lambda_{t,i}$ is non-negative.

The financial interpretation of (49) is that the default intensities are constant, except at the times when defaults occur: then the default intensity for obligor i jumps by an amount $b_{i,j}$ if it is obligor j which has defaulted. Thus a positive $b_{i,j}$ means that obligor i is put at higher risk by the default of obligor j, while a negative $b_{i,j}$ means that obligor i in fact benefits from the default of j, and finally $b_{i,j} = 0$ if obligor i is unaffected by the default of j.

Equation (49) determines the default times through their intensities as well as their joint distribution. However, it is by no means obvious how to find these expressions. Here we will use the following observation, proved in [26].

Proposition 4.1 *There exists a Markov jump process $(Y_t)_{t\geq 0}$ on a finite state space $\boldsymbol{E}$ and a family of sets $\{\Delta_i\}_{i=1}^m$ such that the stopping times*

$$\tau_i = \inf\left\{t > 0 : Y_t \in \Delta_i\right\}, \qquad i = 1, 2, \ldots, m, \tag{50}$$

have intensities (49). Hence, any distribution derived from the multivariate stochastic vector $(\tau_1, \tau_2, \ldots, \tau_m)$ can be obtained from $\{Y_t\}_{t\geq 0}$.

The joint distribution of $(\tau_1, \tau_2, \ldots, \tau_m)$ is sometimes called a multivariate phase-type distribution (MPH), and was first introduced in [4]. Such constructions have largely been developed for queueing theory and reliability applications, see e.g. [2] and [4]). In this section, Proposition 4.1 is throughout used for computing distributions. However, we still use Equation (49) to describe the dependencies in a credit portfolio since it is more compact and intuitive.

Each state $\boldsymbol{j}$ in $\boldsymbol{E}$ is of the form $\boldsymbol{j} = \{j_1, \ldots j_k\}$ which is a subsequence of $\{1, \ldots m\}$ consisting of k integers, where $1 \leq k \leq m$. The interpretation is that on $\{j_1, \ldots j_k\}$ the obligors in the set have defaulted. Furthermore, every permutation of $\{j_1, \ldots j_k\}$ is treated as the same state, that is, the order in which the obligors $j_1, \ldots j_k$ default is of no concern to us, which also is clear from Equation (49). This implies that the cardinality of $\boldsymbol{E}$ will be 2^m, while keeping track of the ordering of $\{j_1, \ldots j_k\}$ implies that $|\boldsymbol{E}| = \sum_{n=0}^{m} n! \binom{m}{n}$ which increases the number of states in $\boldsymbol{E}$ violently. For a more detailed discussion about ordered and unordered default contagion, see [26].

Before we continue, further notation are needed. In the sequel, we let $\boldsymbol{Q}$ and $\boldsymbol{\alpha}$ denote the generator and initial distribution on $\boldsymbol{E}$ for the Markov jump process in Proposition 4.1. The generator $\boldsymbol{Q}$ is found by using the structure of $\boldsymbol{E}$, the definition of the states $\boldsymbol{j}$, and Equation (49). To be more specific, for a state $\boldsymbol{j} = \{j_1, j_2, \ldots, j_k\}$ a transition can only occur to a state $\boldsymbol{j}' = (\boldsymbol{j}, j_{k+1})$ where $j_{k+1} \neq j_i$ for $i = 1, 2, \ldots, k$. Further, the intensity for transitions from $\boldsymbol{j} = \{j_1, j_2, \ldots, j_k\}$ to such a $\boldsymbol{j}'$ is

$$\boldsymbol{Q}_{\boldsymbol{j},\boldsymbol{j}'} = a_{j_{k+1}} + \sum_{i=1}^{k} b_{j_{k+1},j_i} \tag{51}$$

where we remind the reader that every permutation of $\{j_1, \dots j_k\}$ is treated as the same state. The diagonal elements of $\boldsymbol{Q}$ are determined by the requirement that the row sums of an intensity matrix is zero. The set Δ_i is defined as

$$\Delta_i = \{\boldsymbol{j} \in \boldsymbol{E} : j_n = i \text{ for some } j_n \in \boldsymbol{j}\}$$

and since we define τ_i as $\tau_i = \inf\{t > 0 : Y_t \in \Delta_i\}$ for $i = 1, 2, \dots, m$, is clear from the construction that $\tau_1, \dots, \tau_m$ have the intensities (49), see e.g. [30], Chapter 4. The construction is illustrated in Figure 1 for the case $m = 3$.

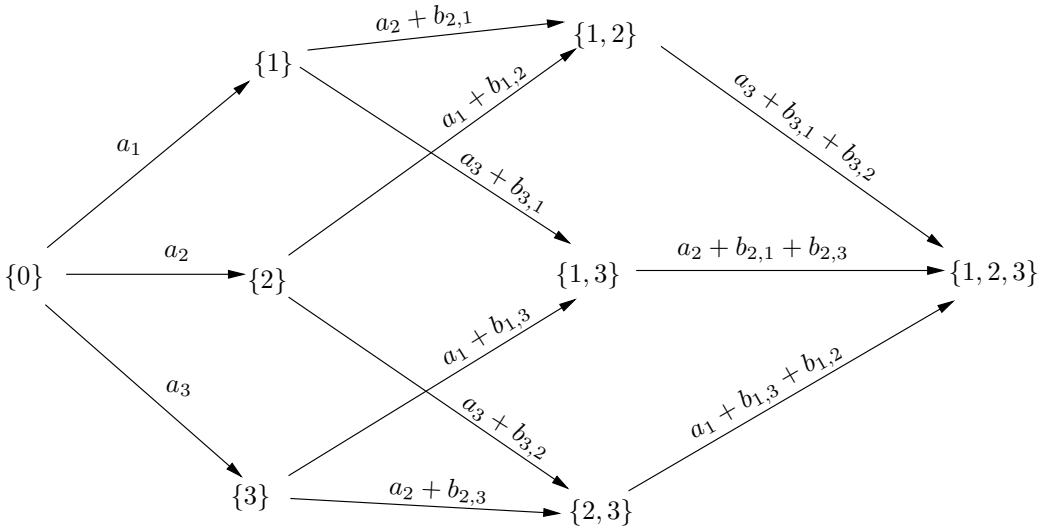


Figure 1: Illustration of the construction for $m = 3$. Arrows indicate possible transitions, and the transition intensities are given on top of the arrows.

The states in $\boldsymbol{E}$ are ordered so that $\boldsymbol{Q}$ is upper triangular, see [26]. In particular, the final state $\{1, \dots m\}$ is absorbing and $\{0\}$ is always the starting state. The latter implies that $\boldsymbol{\alpha} = (1, 0, \dots, 0)$. Furthermore, define the probability vector $\boldsymbol{p}(t) = (\mathbb{P}[Y_t = \boldsymbol{j}])_{\boldsymbol{j} \in \boldsymbol{E}}$. From Markov theory we know that

$$\boldsymbol{p}(t) = \boldsymbol{\alpha} e^{\boldsymbol{Q}t}, \quad \text{and} \quad \mathbb{P}[Y_t = \boldsymbol{j}] = \boldsymbol{\alpha} e^{\boldsymbol{Q}t} \boldsymbol{e}_{\boldsymbol{j}}, \tag{52}$$

where $\boldsymbol{e}_{\boldsymbol{j}} \in \mathbb{R}^{|\boldsymbol{E}|}$ is a column vector where the entry at position $\boldsymbol{j}$ is 1 and the other entries are zero. Recall that $e^{\boldsymbol{Q}t}$ is the matrix exponential which has a closed form expression in terms of the eigenvalue decomposition of $\boldsymbol{Q}$.

4.1.2 The multivariate default distributions

In this subsection we present some formulas for multivariate default and survival distributions, conditional multivariate default distributions, and multivariate default densities. Let $\boldsymbol{G}_i$ be $|\boldsymbol{E}| \times |\boldsymbol{E}|$ diagonal matrices, defined by

$$(\boldsymbol{G}_i)_{\boldsymbol{j},\boldsymbol{j}} = 1_{\{\boldsymbol{j} \in \Delta_i^C\}} \quad \text{and} \quad (\boldsymbol{G}_i)_{\boldsymbol{j},\boldsymbol{j}'} = 0 \quad \text{if} \quad \boldsymbol{j} \neq \boldsymbol{j}'. \tag{53}$$

Further, for a vector $(t_1, t_2, \dots, t_m)$ in $\mathbb{R}_+^m = [0, \infty)^m$, let the ordering of $(t_1, t_2, \dots, t_m)$ be $t_{i_1} < t_{i_2} < \dots < t_{i_m}$ where $(i_1, i_2, \dots, i_m)$ is a permutation of $(1, 2, \dots, m)$. The following proposition was stated in [4], but without a proof. A detailed proof is given in [29].

Proposition 4.2 *Consider m obligors with default intensities (49). Let $(t_1, t_2, \dots, t_m) \in \mathbb{R}_+^m$ and let $t_{i_1} < t_{i_2} < \dots < t_{i_m}$ be its ordering. Then,*

$$\mathbb{P}[\tau_1 > t_1, \dots, \tau_m > t_m] = \boldsymbol{\alpha} \left(\prod_{k=1}^{m} e^{\boldsymbol{Q}\left(t_{i_k} - t_{i_{k-1}}\right)} \boldsymbol{G}_{i_k} \right) \boldsymbol{1} \tag{54}$$

where $t_{i_0} = 0$.

Let $(t_{i_1}, t_{i_2}, \ldots, t_{i_m})$ be the ordering of $(t_1, t_2, \ldots, t_m) \in \mathbb{R}_+^m$ and fix a p, $1 \leq p \leq m-1$. We next consider conditional distributions of the types

$$\mathbb{P}\left[\tau_{i_{p+1}} > t_{i_{p+1}}, \ldots, \tau_{i_m} > t_{i_m} \,\middle|\, \tau_{i_1} \leq t_{i_1}, \ldots, \tau_{i_p} \leq t_{i_p}\right].$$

These probabilities may of course be computed from (54) without any further use of the structure of the problem. However, using this structure leads to compact formulas. For this, further notation is needed. Define Δ as the final absorbing state for Y_t, i.e.

$$\Delta = \bigcap_{i=1}^{m} \Delta_i, \tag{55}$$

and let $\boldsymbol{F}_i$ and $\boldsymbol{H}_i$ be $|\boldsymbol{E}| \times |\boldsymbol{E}|$ diagonal matrices, defined by

$$(\boldsymbol{F}_i)_{\boldsymbol{j},\boldsymbol{j}} = 1_{\{\boldsymbol{j} \in \Delta_i \setminus \Delta\}} \quad \text{and} \quad (\boldsymbol{F}_i)_{\boldsymbol{j},\boldsymbol{j}'} = 0 \quad \text{if} \quad \boldsymbol{j} \neq \boldsymbol{j}'. \tag{56}$$

$$(\boldsymbol{H}_i)_{\boldsymbol{j},\boldsymbol{j}} = 1_{\{\boldsymbol{j} \in \Delta_i\}} \quad \text{and} \quad (\boldsymbol{H}_i)_{\boldsymbol{j},\boldsymbol{j}'} = 0 \quad \text{if} \quad \boldsymbol{j} \neq \boldsymbol{j}'. \tag{57}$$

Then we can state the following proposition, proved in [29].

Proposition 4.3 *Consider* m *obligors with default intensities* (49). *Let* $(t_1, t_2, \ldots, t_m) \in \mathbb{R}_+^m$ *and let* $t_{i_1} < t_{i_2} < \ldots < t_{i_m}$ *be its ordering. If* $1 \leq p \leq m-1$ *then,*

$$\begin{aligned} &\mathbb{P}\left[\tau_{i_1} \leq t_{i_1}, \ldots, \tau_{i_p} \leq t_{i_p}, \tau_{i_{p+1}} > t_{i_{p+1}}, \ldots, \tau_{i_m} > t_{i_m}\right] \\ &= \boldsymbol{\alpha} \left(\prod_{k=1}^{p} e^{\boldsymbol{Q}\left(t_{i_k} - t_{i_{k-1}}\right)} \boldsymbol{F}_{i_k} \right) \left(\prod_{k=p+1}^{m} e^{\boldsymbol{Q}\left(t_{i_k} - t_{i_{k-1}}\right)} \boldsymbol{G}_{i_k} \right) \mathbf{1}. \end{aligned} \tag{58}$$

and

$$\begin{aligned} &\mathbb{P}\left[\tau_{i_{p+1}} > t_{i_{p+1}}, \ldots, \tau_{i_m} > t_{i_m} \,\middle|\, \tau_{i_1} \leq t_{i_1}, \ldots, \tau_{i_p} \leq t_{i_p}\right] \\ &= \frac{\boldsymbol{\alpha} \left(\prod_{k=1}^{p} e^{\boldsymbol{Q}\left(t_{i_k} - t_{i_{k-1}}\right)} \boldsymbol{F}_{i_k} \right) \left(\prod_{k=p+1}^{m} e^{\boldsymbol{Q}\left(t_{i_k} - t_{i_{k-1}}\right)} \boldsymbol{G}_{i_k} \right) \mathbf{1}}{\boldsymbol{\alpha} \left(\prod_{k=1}^{p} e^{\boldsymbol{Q}\left(t_{i_k} - t_{i_{k-1}}\right)} \boldsymbol{H}_{i_k} \right) \mathbf{1}}. \end{aligned} \tag{59}$$

where $t_{i_0} = 0$.

The following corollary is an immediate consequence of Equation (58) in Proposition 4.3.

Corollary 4.4 *Consider* m *obligors with default intensities* (49). *Let* $\{i_1, \ldots, i_p\}$ *and* $\{j_1, \ldots, j_q\}$ *be two disjoint subsequences in* $\{1, \ldots, m\}$. *If* $t < s$ *then*

$$\mathbb{P}\left[\tau_{i_1} > t, \ldots, \tau_{i_p} > t, \tau_{j_1} < s, \ldots, \tau_{j_q} < s\right] = \boldsymbol{\alpha} e^{\boldsymbol{Q}t} \left(\prod_{k=1}^{p} \boldsymbol{G}_{i_k} \right) e^{\boldsymbol{Q}(s-t)} \left(\prod_{k=1}^{q} \boldsymbol{H}_{j_k} \right) \mathbf{1}$$

A similar expression can be found when $s < t$.

We can of course generalize, the above proposition for three time points $t < s < u$, four time points $t < s < u <$ etc. Using the notation of Corollary 4.4 we conclude that if $t < s$ then

$$\mathbb{P}\left[\tau_{j_1} < s, \ldots, \tau_{j_q} < s \,\middle|\, \tau_{i_1} > t, \ldots, \tau_{i_p} > t\right] = \frac{\boldsymbol{\alpha} e^{\boldsymbol{Q}t} \left(\prod_{k=1}^{p} \boldsymbol{G}_{i_k} \right) e^{\boldsymbol{Q}(s-t)} \left(\prod_{k=1}^{q} \boldsymbol{H}_{j_k} \right) \mathbf{1}}{\boldsymbol{\alpha} e^{\boldsymbol{Q}t} \left(\prod_{k=1}^{p} \boldsymbol{G}_{i_k} \right) \mathbf{1}}$$

and a similar expression can be found for $s < t$.

Let $f(t_1,\dots,t_m)$ be the density of the multivariate random variable $(\tau_1,\dots,\tau_m)$. For $(t_1,t_2,\dots,t_m)$, let $(t_{i_1},t_{i_2},\dots,t_{i_m})$ be its ordering where $(i_1,i_2,\dots,i_m)$ is a permutation of $(1,2,\dots,m)$. We denote $(i_1,i_2,\dots,i_m)$ by $\boldsymbol{i}$, that is, $\boldsymbol{i}=(i_1,i_2,\dots,i_m)$. Furthermore, in view of the above notation, we let $f_{\boldsymbol{i}}(t_1,\dots,t_m)$ denote the restriction of $f(t_1,\dots,t_m)$ to the set $t_{i_1}<t_{i_2}<\ldots<t_{i_m}$. The following proposition was stated in [4], but without a proof. A detailed proof can be found in [29].

Proposition 4.5 *Consider m obligors with default intensities (49). Let $(t_1,t_2,\dots,t_m)\in\mathbb{R}_+^m$ and let $t_{i_1}<t_{i_2}<\ldots<t_{i_m}$ be its ordering. Then, with notation as above*

$$f_{\boldsymbol{i}}(t_1,\dots,t_m)=(-1)^m\boldsymbol{\alpha}\left(\prod_{k=1}^{m-1}e^{\boldsymbol{Q}\left(t_{i_k}-t_{i_{k-1}}\right)}\left(\boldsymbol{Q}\boldsymbol{G}_{i_k}-\boldsymbol{G}_{i_k}\boldsymbol{Q}\right)\right)e^{\boldsymbol{Q}\left(t_{i_m}-t_{i_{m-1}}\right)}\boldsymbol{Q}\boldsymbol{G}_{i_m}\mathbf{1} \tag{60}$$

where $t_{i_0}=0$.

4.1.3 The marginal distributions

In this subsection we state expressions for the marginal survival distributions $\mathbb{P}[\tau_i>t]$ and $\mathbb{P}[T_k>t]$, and for $\mathbb{P}[T_k>t,T_k=\tau_i]$ which is the probability that the k-th default is by obligor i and that it not occurs before t. The first ones are more or less standard, while the second one is less so. These marginal distributions are needed to compute single-name CDS spreads and k^{th}-to-default spreads, see e.g [26]. Note that CDS-s are used as calibration instruments when pricing portfolio credit derivatives. The following lemma is trivial, and borrowed from [26].

Lemma 4.6 *Consider m obligors with default intensities (49). Then,*

$$\mathbb{P}[\tau_i>t]=\boldsymbol{\alpha}e^{\boldsymbol{Q}t}\boldsymbol{g}^{(i)}\quad\text{and}\quad\mathbb{P}[T_k>t]=\boldsymbol{\alpha}e^{\boldsymbol{Q}t}\boldsymbol{m}^{(k)} \tag{61}$$

where the column vectors $\boldsymbol{g}^{(i)}$, $\boldsymbol{m}^{(k)}$ of length $|\boldsymbol{E}|$ are defined as

$$\boldsymbol{g}^{(i)}_{\boldsymbol{j}}=1_{\{\boldsymbol{j}\in(\Delta_i)^C\}}\quad\text{and}\quad\boldsymbol{m}^{(k)}_{\boldsymbol{j}}=1_{\{\boldsymbol{j}\in\cup_{n=0}^{k-1}\boldsymbol{E}_n\}}$$

and $\boldsymbol{E}_n$ is set of states consisting of precisely n elements of $\{1,\dots m\}$ where $\boldsymbol{E}_0=\{0\}$.

The lemma immediately follows from the definition of τ_i in Proposition 4.1. The same holds for the distribution for T_k, where we also use that $\boldsymbol{m}^{(k)}$ sums the probabilities of states where there has been less than k defaults. For more on this, see in [26].

We next restate the following result, proved in [26].

Proposition 4.7 *Consider m obligors with default intensities (49). Then,*

$$\mathbb{P}[T_k>t,T_k=\tau_i]=\boldsymbol{\alpha}e^{\boldsymbol{Q}t}\sum_{l=0}^{k-1}\left(\prod_{p=l}^{k-1}\boldsymbol{G}^{i,p}\boldsymbol{P}\right)\boldsymbol{h}^{i,k}, \tag{62}$$

for $k=1,\dots m$, where

$$\boldsymbol{P}_{\boldsymbol{j},\boldsymbol{j}'}=\frac{\boldsymbol{Q}_{\boldsymbol{j},\boldsymbol{j}'}}{\sum_{\boldsymbol{k}\neq\boldsymbol{j}}\boldsymbol{Q}_{\boldsymbol{j},\boldsymbol{k}}},\quad\boldsymbol{j},\boldsymbol{j}'\in\boldsymbol{E},$$

and $\boldsymbol{h}^{i,k}$ is column vectors of length $|\boldsymbol{E}|$ and $\boldsymbol{G}^{i,k}$ is $|\boldsymbol{E}|\times|\boldsymbol{E}|$ diagonal matrices, defined by

$$\boldsymbol{h}^{i,k}_{\boldsymbol{j}}=1_{\{\boldsymbol{j}\in\Delta_i\cap\boldsymbol{E}_k\}}\quad\text{and}\quad\boldsymbol{G}^{i,k}_{\boldsymbol{j},\boldsymbol{j}}=1_{\left\{\boldsymbol{j}\in(\Delta_i)^C\cap\boldsymbol{E}_k\right\}}\quad\text{and}\quad\boldsymbol{G}^{i,k}_{\boldsymbol{j},\boldsymbol{j}'}=0\quad\text{if}\quad\boldsymbol{j}\neq\boldsymbol{j}'.$$

Equipped with the above distributions, we can derive closed-form solutions for single-name CDS spreads and k^{th}-to-default swaps for a nonhomogeneous portfolio, see [26].

>From Equation (61) and (62) we see that all model parameters, including the jump parameters $\{b_{i,j}\}$ creating the default dependence, influence the marginal distributions $\{\mathbb{P}[\tau_i > t]\}$, $\{\mathbb{P}[T_k > t]\}$ and $\{\mathbb{P}[T_k > t, T_k = \tau_i]\}$. This have to be compared with copula models used in portfolio credit risk, where $\{\mathbb{P}[\tau_i > t]\}$ are modelled by idiosyncratic parameters unique for each obligor. Further, in a copula model, the joint dependence is introduced by the copula and its parameters, which are separated from the parameters describing each individual default distribution.

4.1.4 The default correlations and expected default times

In this subsection we derive expressions for pairwise default correlations, i.e. $\rho_{i,j}(t) = \text{Corr}(1_{\{\tau_i \le t\}}, 1_{\{\tau_j \le t\}})$ between the obligors $i \neq j$ belonging to a portfolio of m obligors satisfying (49).

Lemma 4.8 *Consider m obligors with default intensities (49). Then, for any pair of obligors $i \neq j$,*

$$\rho_{i,j}(t) = \frac{\boldsymbol{\alpha} e^{\boldsymbol{Q}t}\boldsymbol{c}^{(i,j)} - \boldsymbol{\alpha} e^{\boldsymbol{Q}t}\boldsymbol{h}^{(i)}\boldsymbol{\alpha} e^{\boldsymbol{Q}t}\boldsymbol{h}^{(j)}}{\sqrt{\boldsymbol{\alpha} e^{\boldsymbol{Q}t}\boldsymbol{h}^{(i)}\boldsymbol{\alpha} e^{\boldsymbol{Q}t}\boldsymbol{h}^{(j)}\left(1 - \boldsymbol{\alpha} e^{\boldsymbol{Q}t}\boldsymbol{h}^{(i)}\right)\left(1 - \boldsymbol{\alpha} e^{\boldsymbol{Q}t}\boldsymbol{h}^{(j)}\right)}} \tag{63}$$

where the column vectors $\boldsymbol{h}^{(i)}$, $\boldsymbol{c}^{(i,j)}$ of length $|\boldsymbol{E}|$ are defined as

$$\boldsymbol{h}^{(i)}_{\boldsymbol{j}} = 1_{\{\boldsymbol{j} \in \Delta_i\}} \quad \textit{and} \quad \boldsymbol{c}^{(i,j)}_{\boldsymbol{j}} = 1_{\{\boldsymbol{j} \in \Delta_i \cap \Delta_j\}} = \boldsymbol{h}^{(i)}_{\boldsymbol{j}}\boldsymbol{h}^{(j)}_{\boldsymbol{j}}. \tag{64}$$

The default correlations $\{\rho_{i,j}(T)\}$ can be used to calibrate the parameters in (49), as will be shortly discussed in Subsection 4.1.6 (see also in [29]). In the standard copula model, $\rho_{i,j}(t)$ is assumed to be constant, and given by the correlation of some latent factor variables driving the individual defaults.

Next, let us consider the expected moments of $\{\tau_i\}$ and $\{T_k\}$. By construction (see Proposition 4.1), the intensity matrix $\boldsymbol{Q}$ for the Markov jump process Y_t on $\boldsymbol{E}$ has the form

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{T} & \boldsymbol{t} \\ \boldsymbol{0} & 0 \end{pmatrix}$$

where $\boldsymbol{t}$ is a column vector with $|\boldsymbol{E}| - 1$ rows. The $\boldsymbol{j}$-th element $\boldsymbol{t_j}$ is the intensity for Y_t to jump from the state $\boldsymbol{j}$ to the absorbing state $\Delta = \cap_{i=1}^m \Delta_i$. Furthermore, $\boldsymbol{T}$ is invertible since it is upper diagonal with strictly negative diagonal elements. Thus, we have the following lemma, proved in [29].

Lemma 4.9 *Consider m obligors with default intensities (49). Then, with notation as above*

$$\mathbb{E}[\tau_i^n] = (-1)^n n! \widetilde{\boldsymbol{\alpha}}\boldsymbol{T}^{-n}\widetilde{\boldsymbol{g}}^{(i)} \quad \textit{and} \quad \mathbb{E}[T_k^n] = (-1)^n n! \widetilde{\boldsymbol{\alpha}}\boldsymbol{T}^{-n}\widetilde{\boldsymbol{m}}^{(k)}$$

for $n \in \mathbb{N}$ where $\widetilde{\boldsymbol{\alpha}}, \widetilde{\boldsymbol{g}}^{(i)}, \widetilde{\boldsymbol{m}}^{(k)}$ are the restrictions of $\boldsymbol{\alpha}, \boldsymbol{g}^{(i)}, \boldsymbol{m}^{(k)}$ from $\boldsymbol{E}$ to $\boldsymbol{E} \setminus \Delta$.

In the paper [29], the implied quantities $\mathbb{E}[\tau_i]$ are computed for two different calibrated portfolios.

4.1.5 Pricing single-name credit default swaps and k^{th}-to-default swap spreads

Given the marginal distributions $\{\mathbb{P}[\tau_i > t]\}$, $\{\mathbb{P}[T_k > t]\}$ and $\{\mathbb{P}[T_k > t, T_k = \tau_i]\}$ presented in Lemma 4.6 and Proposition 4.7 we can find compact, computationally tractable closed-form expression for single-name credit default swaps (CDS-s) and k^{th}-to-default swap spreads. A detailed discussion derivation of these spreads can be found in [26] and [29].

The CDS spreads are used as our main calibration tools when finding the parameters in the model (49). A short discussion of this topic is given in the next section.

4.1.6 Calibration the model via CDS-spreads and correlation matrices

The parameters in (49) can be obtained by calibrating the model against market CDS spreads and market CDS correlations. In [26] and [29] the authors reparameterize the basic description (49) of the default intensities to the form

$$\lambda_{t,i} = a_i \left(1 + \sum_{j=1, j\neq i}^{m} \theta_{i,j} 1_{\{\tau_j \leq t\}} \right), \tag{65}$$

where the a_i-s are the base default intensities and the $\theta_{i,j}$ measure the "relative dependence structure". In [26] the authors assumed that the matrix $\{\theta_{i,j}\}$ is exogenously given and then calibrated the a_i-s against the m market CDS spreads. In [29] the author determine the $\{\theta_{i,j}\}$ from market data on CDS correltaions. To be more specific, if $\rho_{i,j}(T) = \text{Corr}(1_{\{\tau_i \leq T\}}, 1_{\{\tau_j \leq T\}})$ denotes the default correlation matrix computed under the risk neutral measure then [29] used $\beta\{\rho_{i,j}^{(\text{CDS})}(T)\}$ as a proxy for $\{\rho_{i,j}(T)\}$. Here $\{\rho_{i,j}^{(\text{CDS})}(T)\}$ is the observed correlation matrix for the T-years market CDS spreads, and β is a exogenously given parameter. The matrix $\{\rho_{i,j}(T)\}$ is a function of the parameters $\{\theta_{i,j}\}$, and [29] use this fact in the calibration by matching $\rho_{i,j}(T)$ against $\beta\{\rho_{i,j}^{(\text{CDS})}(T)\}$ together with the corresponding market spreads, in order to determine $\{\theta_{i,j}\}$, and the base default intensities a_i-s. However, in the calibration some restrictions have to be imposed on $\{\theta_{i,j}\}$ and we refer to [29] for more on this issue.

Furthermore, in the calibration as well as computation of the quantities presented in Subsection 4.1.2-Subsection 4.1.5, we need efficent methods to compute the matrix exponential $e^{\boldsymbol{Q}t}$. For such discussions, we refer to the Appendix in the full version of the paper.

4.1.7 Alternative parameterizations of the default-intensities

Finally we remark that MPH framework presented in this section also works for other parameterizations of the intensities than given by (49). To be more specific, in the inhomogeneous case, (49) can be replaced by

$$\lambda_{t,i} = f_i(N_{t,1}, \ldots, N_{t,i-1}, N_{t,i+1}, \ldots, N_{t,m}), \qquad t \leq \tau_i, \tag{66}$$

and $\lambda_{t,i} = 0$ for $t > \tau_i$ where the function f_i can be arbitrary chosen, as long as $\lambda_{t,i}$ is non-negative. Recall that $N_{t,j} = 1_{\{\tau_j \leq t\}}$. One can for example choose a multiplicative parametrization of the default intensities, that is

$$\lambda_{t,i} = a_i \prod_{j=1, j\neq i}^{m} (1 + b_{i,j})^{N_{t,j}}, \qquad t \leq \tau_i, \tag{67}$$

and $\lambda_{t,i} = 0$ for $t > \tau_i$ where $a_i > 0$. This parametrization has the intuitive feature that $\Delta\lambda_{\tau_j,i} = b_{i,j}\lambda_{\tau_j-,i}$ for $\tau_i > \tau_j$ which implies that the jump in the default intensity is given by the pre-jump intensity times the constant $b_{i,j}$. Furthermore, the only constraints on $b_{i,j}$ is that $b_{i,j} > -1$ where the case $0 > b_{i,j} > -1$ implies negative default contagion (the intensities jump down at a default). Note that in calibrations with negative jumps, the multiplicative form (67) is much more practical (from an implementation point of view) than the additive framework (49), where we have to put constraints on the parameters to make sure we have non-negative intensities.

Similar multiplicative parameterizations can also be done for the homogeneous model to be presented in Subsection 4.2.1.

4.2 Homogeneous Portfolios

In the nonhomogeneous portfolio presented in Subsection 4.1.1, we have $|\boldsymbol{E}| = 2^m$ which in practice will force us to work with portfolios of size m less or equal to 25, say ([26] used $m = 15$). Standard synthetic CDO portfolios typically contains 125 obligors so we will therefore, in this subsection, consider a special case of (49) which leads to a symmetric portfolio where the state space $\boldsymbol{E}$ can be simplified to make $|\boldsymbol{E}| = m + 1$.

This allows us to practically work with the Markov setup in Proposition 4.1 for large m, where $m \geq 125$ with no further complications.

First, Subsection 4.2.1 gives a short introduction the model. Then, Subsection 4.2.3-Subsection 4.2.5 discuss mariginal and multivariate distributions and related quantites. Finally, Subsection 4.2.6 gives a practical description how to calibrate the model against portfolio credit derivatives (which is performed in Section 5).

4.2.1 The intensity specification for a homogeneous portfolio

We consider a special case of (49) where all obligors have the same default intensities $\lambda_{t,i} = \lambda_t$ specified by parameters a and $b_1, \ldots, b_m$, as

$$\lambda_t = a + \sum_{k=1}^{m-1} b_k 1_{\{T_k \leq t\}} \tag{68}$$

where $\{T_k\}$ is the ordering of the default times $\{\tau_i\}$ and $\phi_1 = \ldots = \phi_m = \phi$ where ϕ is constant. In this model the obligors are exchangeable. The parameter a is the base intensity for each obligor i, and given that $\tau_i > T_k$, then b_k is how much the default intensity for each remaining obligor jumps at default number k in the portfolio. We start with the simpler version of Proposition 4.1. A detailed proof can be found in [27].

Corollary 4.10 *There exists a Markov jump process $(Y_t)_{t\geq 0}$ on a finite state space $\boldsymbol{E} = \{0, 1, 2, \ldots, m\}$, such that the stopping times*

$$T_k = \inf\{t > 0 : Y_t = k\}, \quad k = 1, \ldots, m$$

are the ordering of m exchangeable stopping times $\tau_1, \ldots, \tau_m$ with intensities (68). The generator $\boldsymbol{Q}$ to Y_t is given by

$$\boldsymbol{Q}_{k,k+1} = (m-k)\left(a + \sum_{j=1}^{k} b_j\right) \quad \textit{and} \quad \boldsymbol{Q}_{k,k} = -\boldsymbol{Q}_{k,k+1} \quad \textit{for} \quad k = 0, 1, \ldots, m-1$$

where the other entries in $\boldsymbol{Q}$ are zero. The Markov process always starts in $\{0\}$.

By Corollary 4.10, the states in $\boldsymbol{E}$ can be interpreted as the number of defaulted obligors in the portfolio. In the sequel, we let $\boldsymbol{\alpha} = (1, 0, \ldots, 0)$ denote the initial distribution on $\boldsymbol{E}$. Further, if k belongs to $\boldsymbol{E}$ then $\boldsymbol{e}_k$ denotes a column vector in $\mathbb{R}^{m+1}$ where the entry at position k is 1 and the other entries are zero. From Markov theory we know that $\mathbb{P}[Y_t = k] = \boldsymbol{\alpha} e^{\boldsymbol{Q}t} \boldsymbol{e}_k$ where $e^{\boldsymbol{Q}t}$ is the matrix exponential which has a closed form expression in terms of the eigenvalue decomposition of $\boldsymbol{Q}$.

We remark that the framework (68) is equivalent to the local intensity model which was the starting point in the papers [5], [36], [44] and [45].

4.2.2 Pricing CDO-s and index CDS-s in a homogeneous portfolio

By using Corollary 4.10 we can derive practical formulas for CDO tranche spreads and index CDS spreads. The derivations and other issues regarding the computations, can be found in [27] and [28].

The formulas for CDO tranches are used to calibrate the parameters in a homogeneous model specified by (68), under the risk neutral measure. We will discuss this in Subsection 4.2.6 and Section 5.

4.2.3 The multivariate distributions

In this subsection we present formulas for multivariate default and survival distributions both for ordered as well as unordered default times. We start with the latter. Let $\boldsymbol{M}_k$ be $(m+1) \times (m+1)$ diagonal matrices, defined by $(\boldsymbol{M}_k)_{j,j} = 1_{\{j<k\}}$ and $(\boldsymbol{M}_k)_{j,j'} = 0$ if $j \neq j'$. The following proposition is similar to Proposition 4.2.

Proposition 4.11 *Consider m obligors with default intensities (68) and let $k_1 < \ldots < k_q$ be an increasing subsequence in $\{1, \ldots, m\}$ where $1 \leq q \leq m$. Furthermore, let $t_1 < t_2 < \ldots < t_q$. Then,*

$$\mathbb{P}\left[T_{k_1} > t_1, \ldots, T_{k_q} > t_q\right] = \boldsymbol{\alpha}\left(\prod_{i=1}^{q} e^{\boldsymbol{Q}(t_i - t_{i-1})}\boldsymbol{M}_{k_i}\right)\mathbf{1} \tag{69}$$

where $t_{i_0} = 0$.

A similar expression can also be found for $\mathbb{P}\left[T_{k_1} \leq t_1, \ldots, T_{k_q} \leq t_q\right]$, see in [28]. An explicit proof of Proposition 4.11 is given in [28].

Finding joint distributions for $\{\tau_i\}$ in a homogeneous model with default intensities (68) is a more complicated task than in an inhomogeneous model. For $1 \leq q \leq m$, fix a vector $t_1, \ldots, t_q \in \mathbb{R}_+^q$. For a set of q distinct obligors $i_1, i_2, \ldots, i_q$, the probability $\mathbb{P}\left[\tau_{i_1} \leq t_1, \ldots, \tau_{i_q} \leq t_q\right]$ is by exchangeability the same for any such distinct sequence of q obligors. Therefore we will in this subsection without loss of generality only consider $\mathbb{P}\left[\tau_1 \leq t_1, \ldots, \tau_q \leq t_q\right]$ where $t_1 \leq \ldots \leq t_q$ and similarly for $\mathbb{P}\left[\tau_{i_1} > t_1, \ldots, \tau_{i_q} > t_q\right]$. To exemplify, we state the following proposition proved in [28], where we let $q = 2$ and $t_1 < t_2$.

Proposition 4.12 *Consider m obligors with default intensities (68) and let $t_1 < t_2$. Then,*

$$\mathbb{P}\left[\tau_1 \leq t_1, \tau_2 \leq t_2\right] = \frac{(m-2)!}{m!}\boldsymbol{\alpha}e^{\boldsymbol{Q}t_1}\boldsymbol{n} + \frac{(m-2)!}{m!}\sum_{k_1=1}^{m}\sum_{k_2=k_1+1}^{m}\boldsymbol{\alpha}e^{\boldsymbol{Q}t_1}\boldsymbol{N}_{k_1}e^{\boldsymbol{Q}(t_2-t_1)}\boldsymbol{N}_{k_2}\mathbf{1}. \tag{70}$$

where $\boldsymbol{n}$ is a column vector in $\mathbb{R}^{m+1}$ such that $\boldsymbol{n}_j = \frac{j(j-1)}{2}$.

A similar expression can also be found for $\mathbb{P}\left[\tau_1 > t_1, \tau_2 > t_2\right]$, see in [28].

It is possible to generalize Proposition 4.12 to more that two default times. These expressions do not seem to be easily simplified. However, if $t_1 = \ldots = t_q = t$ we can find compact formulas.

Proposition 4.13 *Consider m obligors with default intensities (2.1) and let q be a integer where $1 \leq q \leq m$. Then,*

$$\mathbb{P}\left[\tau_1 \leq t, \ldots, \tau_q \leq t\right] = \boldsymbol{\alpha}e^{\boldsymbol{Q}t}\boldsymbol{d}^{(q)} \quad \text{and} \quad \mathbb{P}\left[\tau_1 > t, \ldots, \tau_q > t\right] = \boldsymbol{\alpha}e^{\boldsymbol{Q}t}\boldsymbol{s}^{(q)} \tag{71}$$

where $\boldsymbol{d}^{(q)}$ and $\boldsymbol{s}^{(q)}$ are column vectors in $\mathbb{R}^{m+1}$ defined by

$$\boldsymbol{d}_j^{(q)} = \frac{\binom{j}{q}}{\binom{m}{q}}1_{\{j \geq q\}} \quad \text{and} \quad \boldsymbol{s}_j^{(q)} = \frac{\binom{m-j}{q}}{\binom{m}{q}}1_{\{j \leq m-q\}}. \tag{72}$$

A proof of Proposition 4.13 can be found in [28].

4.2.4 The marginal distributions

By Proposition 4.13 with $q = 1$ we get $\mathbb{P}\left[\tau_i > t\right] = \boldsymbol{\alpha}e^{\boldsymbol{Q}t}\boldsymbol{s}^{(1)}$ where $\boldsymbol{s}_j^{(1)} = (m-j)/m = 1 - j/m$. Furthermore, letting $\boldsymbol{m}^{(k)}$ denote $\boldsymbol{m}^{(k)} = \boldsymbol{M}_k\mathbf{1}$, then Proposition 4.11 with $q = 1$ for any $1 \leq k \leq m$, renders that $\mathbb{P}\left[T_k > t\right] = \boldsymbol{\alpha}e^{\boldsymbol{Q}t}\boldsymbol{m}^{(k)}$ where $\boldsymbol{m}_j^{(k)} = 1_{\{j<k\}}$.

Recall that $\mathbb{P}\left[\tau_i > t\right]$ is used to find formulas for the CDS spread in the model specified by (68).

4.2.5 The default correlations and expected default times

In this subsection we use Proposition 4.13 to state expressions for pairwise default correlations between two different obligors belonging to a homogeneous portfolio of m obligors satisfying (68). By exchangeability, $\text{Corr}(1_{\{\tau_i \leq t\}}, 1_{\{\tau_j \leq t\}})$ is the same for all pairs $i \neq j$ so we let $\rho(t)$ denote $\text{Corr}(1_{\{\tau_i \leq t\}}, 1_{\{\tau_j \leq t\}})$.

Lemma 4.14 *Consider m obligors with default intensities (68). Then, with notation as in Subsection 4.2.3*

$$\rho(t) = \frac{\boldsymbol{\alpha} e^{\boldsymbol{Q}t}\boldsymbol{d}^{(2)} - \left(\boldsymbol{\alpha} e^{\boldsymbol{Q}t}\boldsymbol{d}^{(1)}\right)^2}{\boldsymbol{\alpha} e^{\boldsymbol{Q}t}\boldsymbol{d}^{(1)}\left(1 - \boldsymbol{\alpha} e^{\boldsymbol{Q}t}\boldsymbol{d}^{(1)}\right)}. \tag{73}$$

In Section 5 we shall calibrate CDO portfolio for against market data on CDO's and then use Proposition 4.14 to plot the implied default correlation $\rho(t)$ as function of time t.

Next, let us consider the expected moments of $\{T_k\}$ in a homogeneous portfolio. By construction, the intensity matrix $\boldsymbol{Q}$ for the Markov jump process (see Proposition 4.10) has the form

$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{T} & \boldsymbol{t} \\ \boldsymbol{0} & 0 \end{pmatrix}$$

where $\boldsymbol{t}$ is a column vector such that $\boldsymbol{t}_{m-1}$ is nonzero and $\boldsymbol{t}_k = 0$ for $k = 0, 1, \ldots, m-2$, because the k-th element $\boldsymbol{t}_k$, $k \leq m-1$ is the intensity for the Markov jump process Y_t to jump from the state k to the absorbing state $\{m\}$. Furthermore, $\boldsymbol{T}$ is invertible since it is upper diagonal with strictly negative diagonal elements. The following lemma is proved as in Lemma 4.8.

Lemma 4.15 *Consider m obligors with default intensities (68). Then,*

$$\mathbb{E}\left[\tau_i^n\right] = (-1)^n n! \widetilde{\boldsymbol{\alpha}} \boldsymbol{T}^{-n} \widetilde{\boldsymbol{s}}^{(1)} \quad \text{and} \quad \mathbb{E}\left[T_k^n\right] = (-1)^n n! \widetilde{\boldsymbol{\alpha}} \boldsymbol{T}^{-n} \widetilde{\boldsymbol{m}}^{(k)}$$

for $n \in \mathbb{N}$ where $\widetilde{\boldsymbol{\alpha}}, \widetilde{\boldsymbol{s}}^{(1)}, \widetilde{\boldsymbol{m}}^{(k)}$ are the restrictions of $\boldsymbol{\alpha}, \boldsymbol{s}^{(1)}, \boldsymbol{m}^{(k)}$ from $\boldsymbol{E}$ to $\boldsymbol{E} \setminus \{m\}$.

In Section 5 we shall study the implied expected default times $\mathbb{E}\left[T_k\right]$ as function of the number of defaults k. This is done for three different calibrated CDO portfolios.

4.2.6 Calibrating the homogeneous portfolio using CDO-trances and index CDS-s

In this subsection we discuss how to calibrate the model (68) against portfolio credit derivatives.

Let $\boldsymbol{a} = (a, b_1, b_2, \ldots, b_{m-1})$ denote the m parameters in (68). Furthermore, let $\{C_j(T; \boldsymbol{a})\}$ be the $\kappa + 2$ model spreads which are: the CDS spread, the index CDS spread and the κ different CDO tranche spreads. We let $\{C_{j,M}(T)\}$ denote the corresponding market spreads. In $C_j(T; \boldsymbol{a})$ we have emphasized that the model spreads are functions of $\boldsymbol{a} = (a, b_1, b_2, \ldots, b_{m-1})$ but suppressed the dependence of interest rate, payment frequency, etc. The vector $\boldsymbol{a}$ is then obtained as

$$\boldsymbol{a} = \underset{\widehat{\boldsymbol{a}}}{\operatorname{argmin}} \sum_{j=1}^{\eta} \left(C_j(T; \widehat{\boldsymbol{a}}) - C_{j,M}(T)\right)^2 \tag{74}$$

with the constraint that all elements in $\boldsymbol{a}$ are nonnegative. For a fixed maturity T, we use $\kappa = 5$ tranche spreads. This gives us 7 market observations, while the model can contain up to $m = 125$ parameters. In order to reduce the number of unknown parameters to as many as the market observations, we make following assumption on the parameters b_k for $1 \leq k \leq m-1$

$$b_k = \begin{cases} b^{(1)} & \text{if} \quad 1 \leq k < \mu_1 \\ b^{(2)} & \text{if} \quad \mu_1 \leq k < \mu_2 \\ \vdots & \\ b^{(c)} & \text{if} \quad \mu_5 \leq k < \mu_6 = m \end{cases} \tag{75}$$

where $1, \mu_1, \mu_2, \ldots, \mu_6$ is an partition of $\{1, 2, \ldots, m\}$. This means that all jumps in the intensity at the defaults $1, 2, \ldots, \mu_1 - 1$ are same and given by $b^{(1)}$, all jumps in the intensity at the defaults $\mu_1, \ldots, \mu_2 - 1$ are same and given by $b^{(2)}$ and so on. Hence, in (74) we now minimize over the unknown vector $\boldsymbol{a} = (a, b^{(1)}, \ldots, b^{(6)})$. Furthermore, if we for example want to calibrate our model against CDO-tranches from the iTraxx-series, we can use a recovery of $\phi = 40\%$ with $m = 125$ and let $\mu_1, \mu_2, \ldots, \mu_6$ be given as in Table 1. In this way we assign one parameter for each trance $[0, 3], [3, 6], [6, 9], [6, 12]$ $[12, 22]$ and also one parameter for the loss interval $[22, 60]$.

Table 1: The integers $1, \mu_1, \mu_2, \ldots, \mu_c$ define a partition of $\{1, 2, \ldots, m\}$ used in the models that generates the spreads in Table 2.

partition	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
	7	13	19	25	46	125

5 Numerical studies

This section presents few illustrative numerical results on CDO tranches and related quantities. We consider the simple case of a time-homogeneous and exchangeable model of portfolio credit risk. We will therefore use the results of Subsection 4.2. First, Subsection 5.1 discuss the data sets used in the calibration, as well as the obtained parameters and related quantities. In Subsection 5.2 we present and study the implied loss distribution. The topic of Subsection 5.3 is the implied expected ordered default times. Finally, Subsection 5.4 studies the implied default correlations which posses some interesting features.

5.1 Model Calibration

We first calibrate the model (68) against credit derivatives on the iTraxx Europe series with maturity of five years. We do this for three different data sets, sampled on 2004-08-04, 2006-10-15 and 2008-03-07. The spreads in these series differ substantially, where the last set was collected during the subprime crises. Each data set contain five different CDO tranche spreads with tranches $[0, 3], [3, 6], [6, 9], [9, 12]$ and $[12, 22]$, the index CDS spreads and the average CDS spread. We use the same parametrization of the jumps $\{b_k\}_{k=1}^m$ as described in Subsection 4.2.6. We choose the partition $\mu_1, \mu_2, \ldots, \mu_6$ so that it roughly coincides with the number of defaults needed to reach the upper attachment point for each tranche, see Table 1. In all three calibrations the interest rate was set to 3%, the payment frequency was quarterly and the recovery rate was 40%.

In all three data sets we obtain perfect fits, although in the 2008 portfolio the accumulated calibration error is around nine times higher than it is in the 2006 portfolio. The relative calibration errors is, however, very good. Furthermore, due to the subprime crises some of the corresponding spreads in the 2008 data has increased by a factor 50 compared with the 2006 portfolio, see Table 2.

The calibrated parameters, displayed in Table 3 are obtained by using an stiff-ODE solver, as discussed in the Appendix in the full version of the paper. For more details on different numerical methods used in the calibration, see in [27]. In Figure 2 we have displayed the next-to-default intensities, that is $\boldsymbol{Q}_{k,k+1}$, for the three calibrated portfolios with parameters given by Table 3. The next-to-default intensities have similar shapes for the three portfolios, but differ in magnitude, especially when k is bigger than 25 defaults.

Having calibrated the portfolio, we can compute implied quantities that are relevant to credit portfolio management, for example the implied loss distribution (and the loss surface), the implied default correlations and the implied expected ordered default times. To do this we use the calibrated parameters in Table 3 and the closed formulas presented in Subsection 4.2.3 - 4.2.5. All computations of the matrix-exponential are performed with either a stiff ODE-solver, or the Padé-method, both discussed in the Appendix in the full version of the paper.

Besides the above quantities, we can also compute more exotic credit derivatives that are not liquidly quoted on the market, such as tranchlets and basket default swaps on subportfolios in the CDO portfolio. The latter quantities are not treated in this chapter, and we refer to [27] for such studies.

5.2 Loss distributions

The implied loss distribution on the interval $0 \leq x \leq 22\%$ for $t = 5$ is displayed in Figure 3 and the distribution for the whole loss-interval $0 \leq x \leq 60\%$ is shown in the first subfigure of Figure 4. Furthermore,

Table 2: iTraxx Europe Series 3, 6 and 8 collected at August 4^{th} 2004, November 28^{th}, 2006 and March 7^{th}, 2008. The market and model spreads and the corresponding absolute errors, both in bp and in percent of the market spread. The $[0, 3]$ spread is quoted in %. All maturities are for five years.

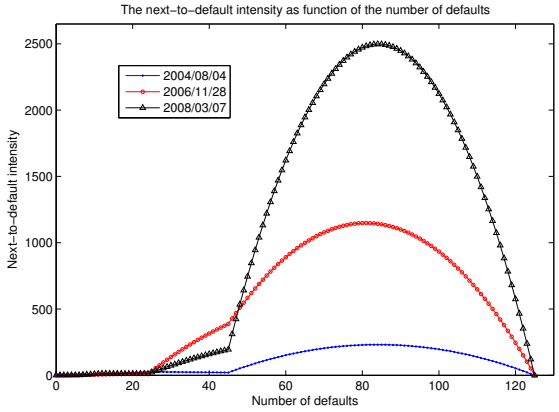
2004-08-04	Market	Model	error (bp)	error (%)
$[0, 3]$	27.6	27.6	3.851e-005	1.4e-006
$[3, 6]$	168	168	0.000316	0.0001881
$[6, 9]$	70	70	0.000498	0.0007115
$[9, 12]$	43	43	0.0005563	0.001294
$[12, 22]$	20	20	0.0004006	0.002003
index	42	42.02	0.01853	0.04413
avg CDS	42	41.98	0.01884	0.04486
Σ abs.cal.err			0.03918 bp	

2006-11-28	Market	Model	error (bp)	error (%)
$[0, 3]$	14.5	14.5	0.008273	0.0005705
$[3, 6]$	62.5	62.48	0.02224	0.03558
$[6, 9]$	18	18.07	0.07275	0.4042
$[9, 12]$	7	6.872	0.1282	1.831
$[12, 22]$	3	3.417	0.4169	13.9
index	26	26.15	0.1464	0.5632
avg CDS	26.87	26.13	0.7396	2.752
Σ abs.cal.err			1.534 bp	

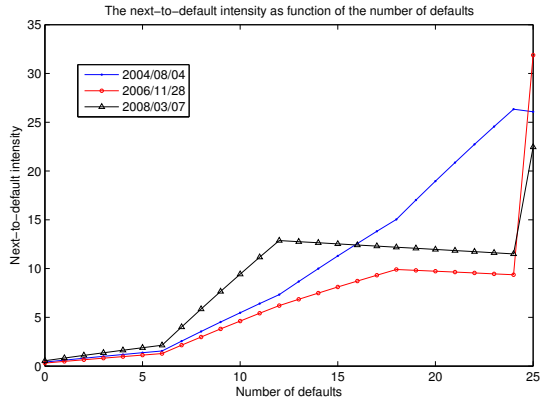
2008-03-07	Market	Model	error (bp)	error (%)
$[0, 3]$	46.5	46.5	0.0505	0.001086
$[3, 6]$	567.5	568	0.4742	0.08356
$[6, 9]$	370	370	0.04852	0.01311
$[9, 12]$	235	234	1.035	0.4404
$[12, 22]$	145	149.9	4.911	3.387
index	150.3	144.3	5.977	3.978
avg CDS	145.1	143.8	1.296	0.8933
Σ abs.cal.err			13.79 bp	

Table 3: The calibrated parameters that gives the model spreads in Table 2

	a	$b^{(1)}$	$b^{(2)}$	$b^{(3)}$	$b^{(4)}$	$b^{(5)}$	$b^{(6)}$	
2004/08/04	33.07	16.3	86.24	126.2	200.3	0	1379	$\times 10^{-4}$
2006/11/28	24.9	13.93	73.36	62.9	0.2604	2261	5904	$\times 10^{-4}$
2008/03/07	44.2	22.66	159.8	0	6e-008	1107	779700	$\times 10^{-4}$



(a) The next-to-default intensities for $0 \leq k \leq 125$



(b) The next-to-default intensities for $0 \leq k \leq 26$

Figure 2: The next-to-default intensities, i.e. $\boldsymbol{Q}_{k,k+1}$, in the three calibrated portfolios with parameters given by Table 3. The upper plot is for $0 \leq k \leq 125$, while the lower displays $\boldsymbol{Q}_{k,k+1}$ when $0 \leq k \leq 26$.

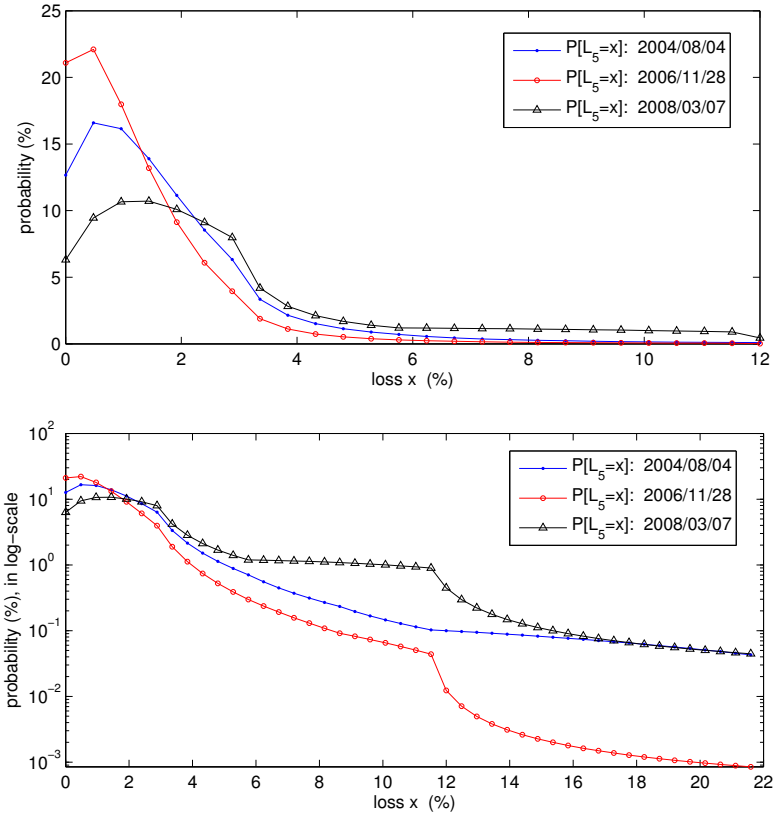


Figure 3: The five year implied loss distributions $\mathbb{P}[L_5 = x\%]$ (in %) for the 2004-08-04, 2006-11-28 and 2008-03-07 portfolios, where $0 \leq x \leq 12$ (upper) and $0 \leq x < 22$ (lower). The lower graph is in log-scale.

Table 4 displays the probabilities $\mathbb{P}[L_5 \geq x\%]$ for $x = 3, 6, 9, 12, 22$ and $x = 60$. With 40% recovery, $\mathbb{P}[L_5 \geq 60\%] = \mathbb{P}[L_5 = 60\%] = \mathbb{P}[Y_5 = 125]$ is the so called five-year "Armageddon probability", i.e. the probability that all obligors in the portfolio have defaulted within 5 years from the date the portfolio was calibrated. The five year "Armageddon probabilities" are negligible for the 2004 and 2006 portfolios(0.08 % and 0.127 respectively), but very big for the 2008 data-set, where $\mathbb{P}[L_5 = 60\%] = 7.11\%$. Thus, there is 7% probability (under the risk-neutral measure) that all 125 obligors in the portfolio have defaulted within 5 years from March 2008. The huge differences in the "Armageddon probabilities" between the 2006 and 2008 portfolios are due to the subprime-crises, that emerged 2007 and continued into 2008.

Table 4: The probabilities $\mathbb{P}[L_5 \geq x\%]$ (in %) where $x = 3, 6, 9, 12, 22$ and $x = 60$, for the 2004-08-04, 2006-11-28 and 2008-03-07 portfolios.

$\mathbb{P}[L_5 \geq x\%]$	$x = 3$	$x = 6$	$x = 9$	$x = 12$	$x = 22$	$x = 60$
2004/08/04	14.7	4.976	2.793	1.938	0.4485	0.07997
2006/11/28	6.466	1.509	0.5935	0.2212	0.1674	0.1265
2008/03/07	35.67	22.26	15.44	9.552	7.122	7.108

We also study the dynamics of the implied loss model over time, Figure 5 displays the loss-distribution at the time points $t = 1, 5, 10$ and $t = 15$ (time measured in years) and where the loss x ranges between 0% to 24%. Furthermore, Figure 6 shows the whole implied loss-surface, i.e. the loss probabilities as function of time and loss, for the calibrated 2006 portfolio. The four subpictures in Figure 6 clearly depicts the shift of probability mass due to contagion, as time progresses. This shift of probability mass can also be seen in the first ($t = 5$) and second ($t = 15$) subplot in Figure 4 where, for example in the 2006 portfolio it holds that $\mathbb{P}[L_5 = 60\%] = \mathbb{P}[Y_5 = 125] = 0.127\%$ while $\mathbb{P}[L_{15} = 60\%] = 64.5\%$.

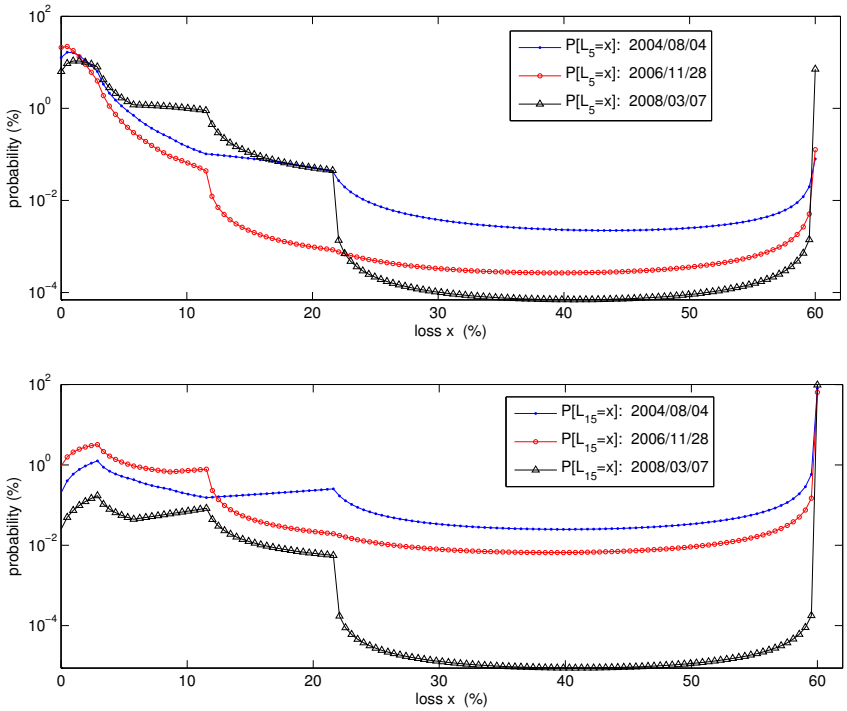


Figure 4: The five year (upper) and fifteen year (lower) implied loss distributions (in %) for the 2004-08-04, 2006-11-28 and 2008-03-07 portfolios, where $0 \leq x \leq 60$. Both graphs are in log-scale.

Further, for the 2006-11-28 and 2008-03-07 portfolios, we clearly see the effect of default contagion on

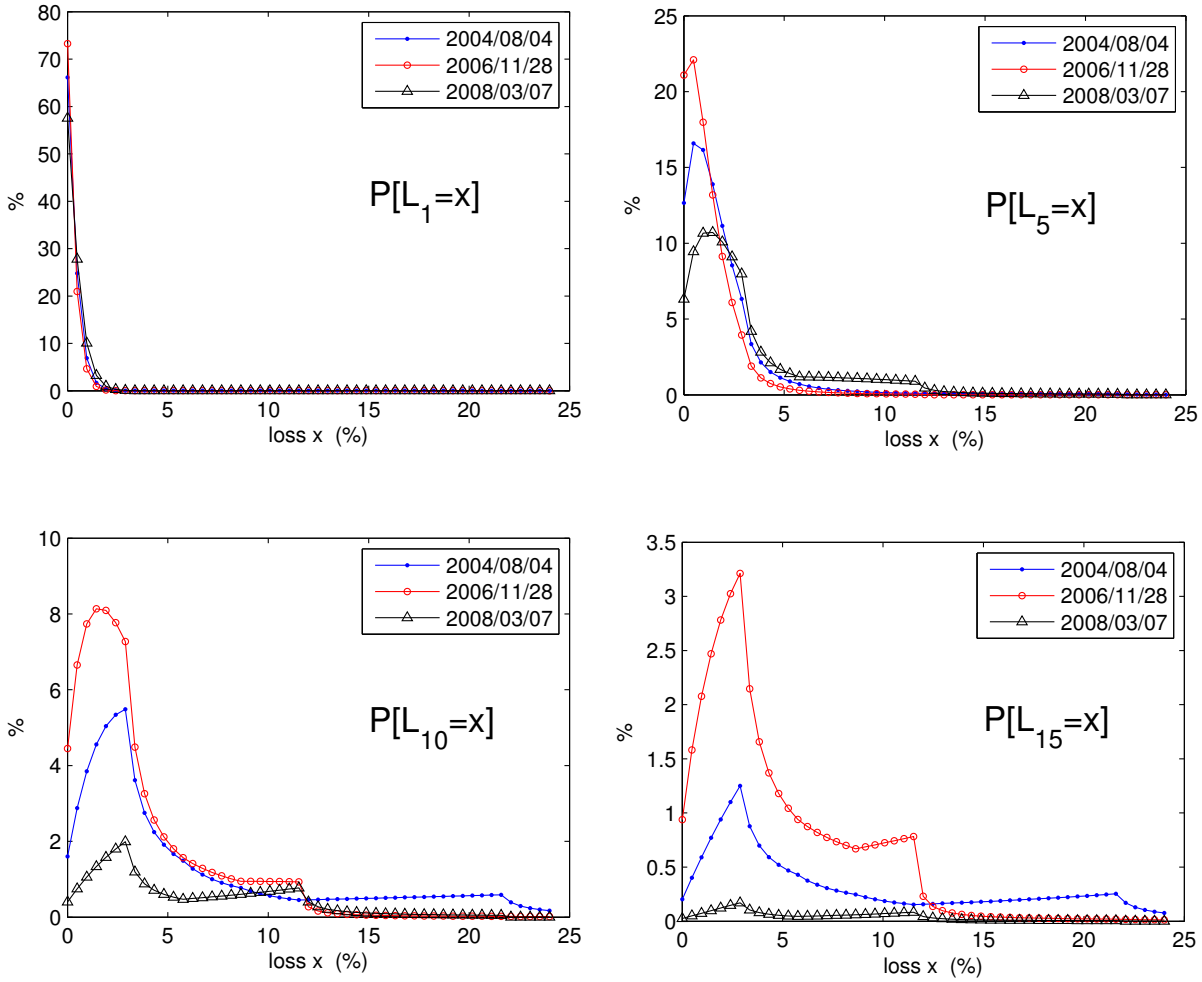
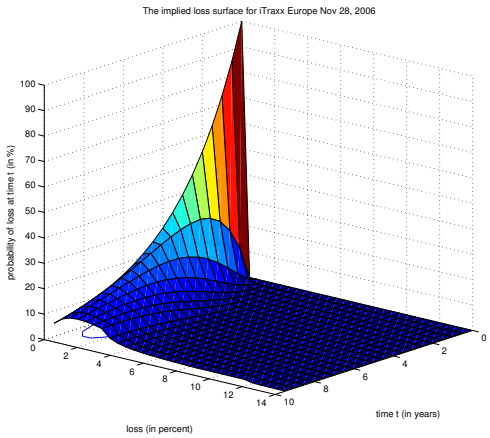
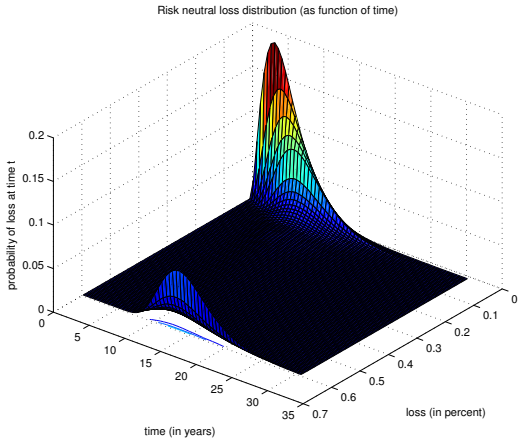


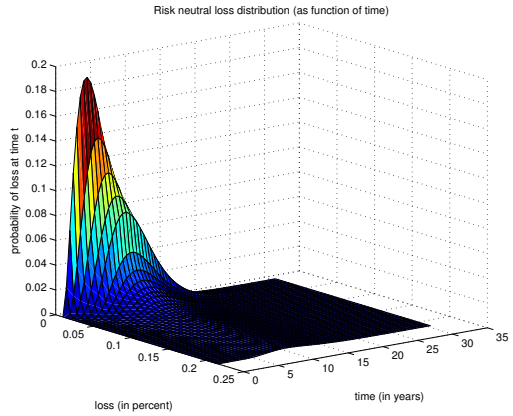
Figure 5: The implied loss distributions $\mathbb{P}[L_t = x\%]$ (in %) for the 2004-08-04, 2006-11-28 and 2008-03-07 portfolios at the time points $t = 1, 5, 10, 15$ and where the loss x ranges from 0% to 24%.



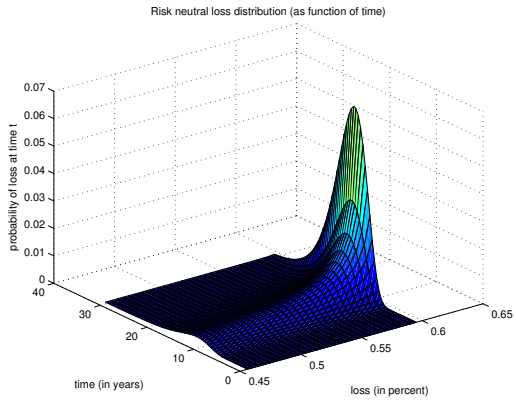
(a) Loss x for $0\% \leq x \leq 14\%$



(b) Loss x for $0\% < x < 60\%$



(c) Loss x for $0\% < x < 22\%$



(d) Loss x for $45\% < x \leq 59\%$

Figure 6: The implied loss surface for iTraxx Europe Nov 28, 2006, from different angels and for different loss regions.

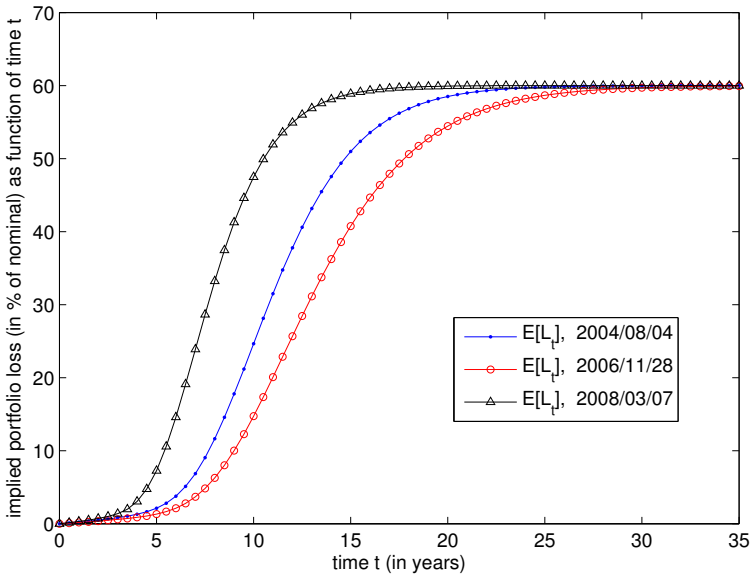


Figure 7: The implied portfolio losses in % of nominal, for the 2004-08-04, 2006-11-28 and 2008-03-07 portfolios.

the upper tranche-losses, making them lie close to each other, see Figure 8. >From Figure 7 we conclude that our model, with a constant recovery rate of 40%, calibrated to market spreads on the five year iTraxx Europe Series in Table 2, implies that the whole portfolio has defaulted within approximately 30 years, under the risk-neutral measure, (for all three data sets). In reality, this will likely not happen, since risk-neutral (implied) default probabilities are substantially larger than the "real", so called actuarial, default probabilities.

5.3 Expected ordered default times

Next, let us study the implied expected ordered default times $\mathbb{E}[T_k]$. In Figure 9, left, we note that the implied expected ordered default times take values roughly between 3.5 years and 14 years. A striking feature in the 2006-11-28 portfolio is that after the 25-th default, the $\mathbb{E}[T_k]$ cluster around 14 years. This is a consequence of the explosion in the jump intensities for $k \geq 25$, see Table 3. Under the risk-neutral measure, implied by the market data in Table 2, this clustering of $\mathbb{E}[T_k]$ means that we expect extreme losses in year 13 and 14 for the 2006-11-28 portfolio. The clustering effect is also present in the 2008 data-set, which indicates that the whole portfolio is expected to be wiped out within 9 years. Again, recall that all computations are under the risk-neutral measure, and should not be confused with real default probabilities and their expectations. These are likely to be substantially smaller for the loss probability and much bigger for the expected ordered default times.

5.4 Default correlations

Finally, we study the implied pairwise default correlation $\rho(t) = \text{Corr}(1_{\{\tau_i \leq t\}}, 1_{\{\tau_j \leq t\}})$ for two distinct obligors i, j, as function of time t, see Figure 10. In e.g. the 2006-11-28 portfolio, we see that $\rho(t)$ is less than 2% when $t \leq 4$, but then starts to increase rapidly, first to 4% for $t = 4.5$, then to 77% for $t = 10$ and reaches 88% at $t = 15$. After this drastic development, the implied default correlation flattens out and converges to 91% as time increases against 30 years. The explosive increase of $\rho(t)$ from 2% to 88% in the time interval $[4.5, 15]$ is due to the default contagion and is also consistent with the clustering of $\{T_k\}$ around

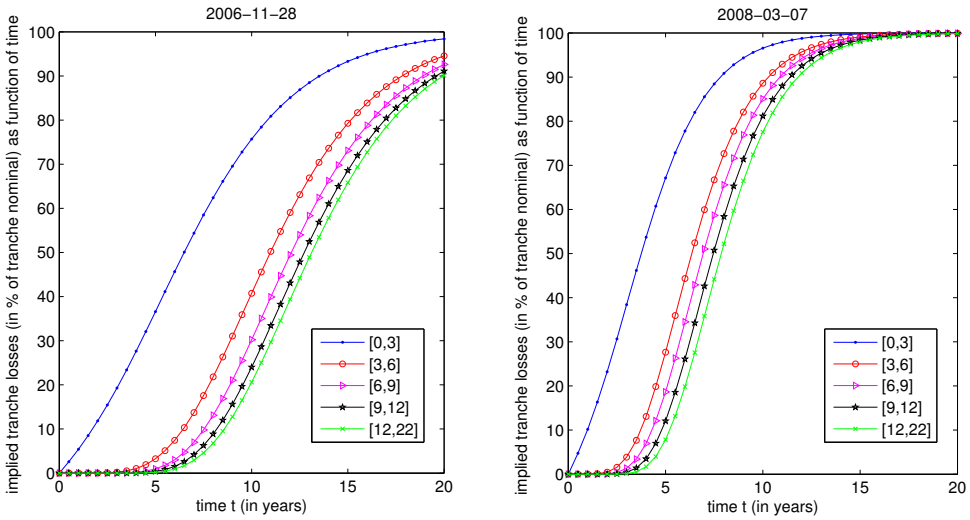


Figure 8: The implied tranche losses in % of tranche nominal for the 2006-11-28 (left) and 2008-03-07 (right) portfolios.

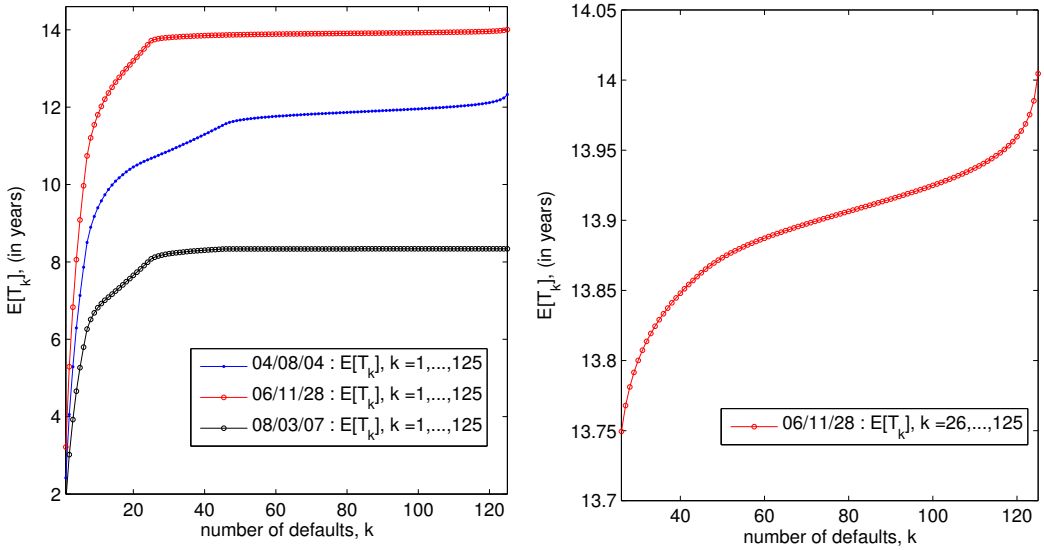


Figure 9: The implied expected ordered default times $\mathbb{E}[T_k]$ for the 2004-08-04, 2006-11-28 and 2008-03-07 portfolios where $k = 1, \ldots, 125$ (left) and $k = 26, \ldots, 125$ (right, for 2006-11-28).

$t = 14$. We also note that the implied default correlation for the 2004-08-04 portfolio follows an almost identical trend up to 8 years. This is consistent with the jump-to-default parameters for the first 13 defaults, which are in the same order as in 2006-11-28 case, see also Figure 9. Even though there is a big difference between the corresponding contagious parameters for $k > 13$ in the 2004 and 2006 portfolio, the implied default correlations never differ more than $10\% - 12\%$ during the first 30 years between these portfolios. The default correlation for 2008 portfolio has similar shape as the 2004 and 2006 portfolios, but the steep increase in the correlation curve (due to contagion) starts earlier in the 2008 portfolio. Given the huge spreads in the 2008 set compared with the 2004 and 2006 portfolios, this does not come as a surprise.

For more numerical studies of the model (68) we refer to the papers [27] and [28].

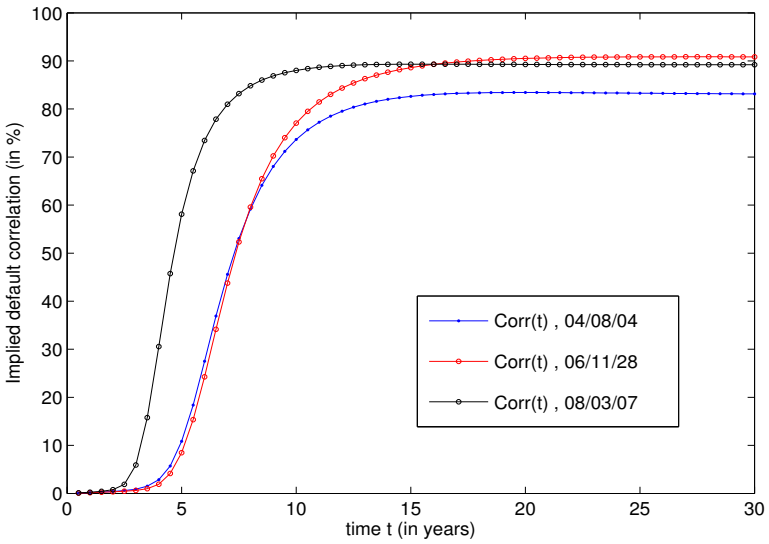


Figure 10: The implied default correlation $\rho(t) = \text{Corr}(1_{\{\tau_i \le t\}}, 1_{\{\tau_j \le t\}})$, $i \neq j$ as function of time for the 2004-08-04, 2006-11-28 and 2008-03-07 portfolios.

References

[1] ALBANESE, C. AND CHEN, O.: Discrete Credit Barrier Models. *Quantitative Finance*, vol. 5, pp. 247–256, 2005.

[2] ASMUSSEN, S.: Matrix-analytic models and their analysis. *Scandinavian Journal of Statistics*, vol. 27, pp. 193–226, 2000.

[3] ASMUSSEN, S.: *Applied Probability and Queues. Second Edition.* Springer, London, 2003.

[4] ASSAF, D., LANGBERT, N. A., SAVIS, T.H. AND SHAKED, M.: Multivariate phase-type distributions. *Operations Research*, vol. 32, pp. 688–701, 1984.

[5] ARNSDORF, M. AND HALPERIN, I.: BSLP: Markovian Bivariate Spread-Loss Model for Portfolio Credit. *The Journal of Computational Finance* vol. 12, no. 2, 2008.

[6] BACKHAUS, J.: Pricing and Hedging of Credit Derivatives in a Model with Interacting Default Intensities: A Markovian Approach. *PhD Thesis*, University of Leipzig, 2008.

[7] BHATTACHARYA, R.N., WAYMIRE, E.C.: *Stochastic Processes with Applications.* J. Wiley, Chichester, 1990.

[8] BIELECKI, T.R., CRÉPEY, S. AND JEANBLANC, M.: Up and Down Credit Risk. Forthcoming in *Quantitative Finance*.

[9] BIELECKI, T.R., CRÉPEY, S., JEANBLANC, M. AND RUTKOWSKI, M.: Valuation of basket credit derivatives in the credit migrations environment. *Handbook of Financial Engineering*, J. Birge and V. Linetsky eds., Elsevier, 2007.

[10] T.R. BIELECKI, J. JAKUBOWSKI, A. VIDOZZI AND L. VIDOZZI,: Study of Dependence for Some Classes of Stochastic Processes. *Stoch. Anal. and Appl.*, vol. 26, no. 4, pp. 903–924, 2008.

[11] BIELECKI, T.R. AND RUTKOWSKI, M.: *Credit Risk: Modeling, Valuation and Hedging.* Springer-Verlag, Berlin, 2002.

[12] BIELECKI, T.R., VIDOZZI, A. AND VIDOZZI, L.: A Markov Copulae Approach to Pricing and Hedging of Credit Index Derivatives and Ratings Triggered Step–Up Bonds. *J. of Credit Risk*, vol. 4, no.1, 2008.

[13] BIELECKI, T.R., VIDOZZI, A. AND VIDOZZI, L.: An efficient approach to valuation of credit basket products and ratings triggered step-up bonds. *Working Paper*, 2006.

[14] BRÉMAUD, P.: *Point processes and queues, martingale dynamics.* Springer-Verlag, New York, 1981.

[15] CARMONA, R. AND CRÉPEY, S.: Importance Sampling and Interacting Particle Systems for the Estimation of Markovian Credit Portfolios Loss Distribution, Forthcoming in *International Journal of Theoretical and Applied Finance*.

[16] CONT, R., MINCA, A.: Recovering Portfolio Default Intensities Implied by CDO Quotes. *Working Paper*, 2008.

[17] R. CONT AND P. TANKOV.: *Financial Modeling With Jump Processes.* Chapman & Hall/CRC, Boca Raton, 2003.

[18] CRÉPEY, S., JEANBLANC, M. AND ZARGARI, B.: CDS with Counterparty Risk in a Markov Chain Copula Model with Joint Defaults. *Working Paper*, 2009.

[19] FREY, R., BACKHAUS, J.: Dynamic hedging of synthetic CDO-tranches with spread- and contagion risk, Forthcoming in in *Journal of Economic Dynamics and Control*.

[20] FREY, R., BACKHAUS, J.: Pricing and Hedging of Portfolio Credit Derivatives with Interacting Default Intensities. *International Journal of Theoretical and Applied Finance*, vol. 11, no. 6, pp. 611–634, 2008

[21] FREY, R., BACKHAUS, J.: Portfolio credit risk models with interacting default intensities: a Markovian approach. *Working Paper*, 2004.

[22] GROSS, D AND MILLER, D. R.: The randomization technique as a modelling tool and solution procedure for transient markov processes. *Operations Research*, vol. 32, no. 2, pp. 343–361, 1984.

[23] GYÖNGY, I.: Mimicking the one dimensional distributions of processes having an Ito differential. *Probability theory and related fields*, vol. 71, no. 4, pp. 501–516, 1986.

[24] HALPERIN, I. AND TOMECEK, P.: Climbing Down from the Top: Single Name Dynamics in Credit Top Down Models. *Working paper*, 2008.

[25] HERBERTSSON, A.: Pricing portfolio credit derivatives. *PhD Thesis*, University of Gothenburg, 2007.

[26] HERBERTSSON, A. AND ROOTZÉN, H: Pricing k^{th}-to-default swaps under default contagion:the matrix-analytic approach. *The Journal of Computational Finance*, vol. 12, no. 1, pp. 49–78, 2008.

[27] HERBERTSSON, A.: Pricing synthetic CDO tranches in a model with Default Contagion using the Matrix-Analytic approach. *The Journal of Credit Risk*, vol. 4, no. 1, pp.3–35, 2008.

[28] HERBERTSSON, A.: Default contagion in large homogeneous portfolios. *The Credit Derivatives Handbook Global Perspectives, Innovations, and Market Drivers*, Gregoriou, G.N and Ali, P.U, (eds), McGraw-Hill, New York, 2008.

[29] HERBERTSSON, A.: Modelling default contagion using Multivariate Phase-Type distributions. *Working paper*, 2007.

[30] JACOBSEN, M. *Point process theory and applications, marked point and piecewise deterministic processes.* Birkhäuser, Boston, 2006.

[31] JARROW, R., LANDO, D. AND TURNBULL, S: A Markov Model for the Term Structure of Credit Risk Spreads. *The Review of Financial Studies*, vol. 10, no. 2, pp. 481–523, 1997.

[32] LANDO, D *Credit Risk Modeling: Theory and applications.* Princeton University Press, Princeton, 2004.

[33] LAST, G., BRANDT, A.: *Marked Point Processes on the Real Line: The Dynamical Approach*. Springer, 1995.

[34] LAURENT, J.-P., COUSIN, A. , FERMANIAN, J-D.: Hedging default risks of CDOs in Markovian contagion models. *Working Paper*, 2007.

[35] D. LANDO AND A. MORTENSEN.: On the pricing of step-up bonds in the European telecom sector, *J. Credit Risk*, vol. 1, no. 1, pp. 71–110, 2005

[36] LOPATIN, A., MISIRPASHAEV, T.: Two-Dimensional Markovian Model for Dynamics of Adequate Credit Loss. *Working Paper*, 2007.

[37] MOELER, C AND VAN LOAN, C.: Nineteen dubioius ways to compute the exponential of a matrix. *SIAM Review*, vol. 20, vol. 4, pp. 801–836, 1978.

[38] MOELER, C AND VAN LOAN, C.: Nineteen dubioius ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review*, vol. 45, no. 1, pp. 3–49, 2003.

[39] NEUTS, M. F.: *Matrix-geometric solutions in stochastic models. An algorithmic approach*. The Johns Hopkins University Press, Baltimore, 1981.

[40] NEUTS, M. F.: *Structured Stochastic Matrices of M/g/1 Type and Their Applications*. Dekker, New York, 1989.

[41] PALMOWSKI, Z. AND ROLSKI, T.:, A technique for exponential change of measure for Markov processes. *Bernoulli*, vol. 8, no. 6, pp. 767–785, 2002.

[42] ROGERS, L.C.G., WILLIAMS, D.: *Diffusions, Markov Processes and Martingales.* 2^{nd} edition. Cambridge University Press, 2000.

[43] ROLSKI, T., SCHMIDLI, H., SCHMIDT, V., TEUGGELS, J.: *Stochastic Processes for Insurance and Finance*, J. Wiley, Chichester, 1998.

[44] SCHÖNBUCHER, P. J.: Portfolio losses and the term structure of loss transition rates: A new methodology for the pricing of portfolio credit derivatives. *Working paper*, 2005.

[45] SIDENIUS, J., PITERBARG, V. AND ANDERSEN, L.: A new framework for dynamic credit portfolio loss modelling. *International Journal of Theoretical and Applied Finance*, vol. 11, no. 2, pp. 163–197, 2008.

[46] SIDJE, R. B. AND STEWART, W. J.: A numerical study of large sparse matrix exponentials arising in Markov chains. *Computational statistics and data analysis*, vol. 29, no. 3, pp. 345–368, 1999.

[47] SYSKI, R.: *Passage Times for Markov Chains.* IOS Press, Amsterdam, 1992.