Lecture Notes of
Seminario Interdisciplinare di Matematica
Vol. 12(2015), pp. 273 – 288.

Modeling and filtering of credit health for a set of firms

Paola TARDELLI[1]

Abstract. Commercial banks have devoted many resources to develop models aiming to better quantify their financial risk and to assign economic capital. These efforts have been recognized and encouraged by bank regulators, and have been extended into the field of credit risk modeling. To differentiating credit quality, a heterogeneous population of identical firms is divided into a finite number of classes. In order to have a dynamical model, the partition can change during the time. A suitable exchangeability assumption is made to preserve the property of having identical items with different credit levels. In a partially observed setting, given the cardinality of the class of the firms already defaulted, the aim is to recover information about the time to default of the remaining firms. This topic is discussed by using stochastic filtering techniques.

1. INTRODUCTION

The field of credit risk modeling has developed rapidly over the past few years, and it is going to become a key tool in the risk management systems at financial institutions. More, agency rating has achieved wide investor acceptance as convenient tool for evaluating credit quality.
A credit rating is a risk indicator, telling an investor how likely or unlikely an entity's bankruptcy is. The investor, either an equity investor or a lender, will use this information to decide which return he wants to receive. A poor credit rating indicates that the risk of default for the company or the government is high.
Hence, the purpose of the rating classes is to provide investors with a simple system of gradation by which future relative creditworthiness of securities may be gauged. Gradations of creditworthiness are indicated by rating symbols, with each symbol representing a group in which the credit characteristics are broadly the same. There are nine symbols, from those used to designate least credit risk to those denoting greatest credit risk: Aaa Aa A Baa Ba B Caa Ca C. The symbol D denotes the class of the defaulted firms.
While their reputation has suffered in recent years, the three major rating agencies Moody's, Standard & Poor's and Fitch are still relied on the most. These agencies record past and present rating documents, as well as movements of firms from one rating category to another. They fulfill as direct indicators of risk dynamics. The

[1]P. Tardelli, Università degli Studi dell'Aquila, Dipartimento di Ingegneria Industriale e dell'Informazione e di Economia, Via Giovanni Gronchi 18, 67100 L'Aquila, Italy;
paola.tardelli@univaq.it

Keywords. Heterogeneous populations; exchangeability; default time.
AMS Subject Classification. 62N05, 60J75, 60G09, 60G55.

aim is to give information on the future expected changes in credit quality. This information is used for computing the risk on portfolios of corporate credits and for pricing credit derivatives.

The methods, used in the banking sector for modeling and for risk prediction, have an important effect on the performance of internal credit pricing models and the operating of risk monitoring systems. There exists a wide variety of credit risk models that differ in their fundamental assumptions, as well as their definition of credit losses.

The credit quality of most issuers and their obligations is not fixed and steady over a period of time, but tends to undergo change. For this reason, changes in rating reflect variations in the intrinsic relative position of issuers and their obligations. A change in rating may occur at any time in the case of an individual issue.

Hence, in order to develop a credit risk model, a population of identical firms, $\{U_1, \cdots, U_H\}$, is divided into different classes, $\{C_1, \cdots, C_d\}$, according to their level of credit quality and C_0 is the class of firms already defaulted. The model is dynamic, in the sense that the level of credit quality of any single firm can increase or can decrease, while the time goes on, according to several reasons. As first assumption, each firm has the chance to default, whatever class it belongs.

This model is inspired by the idea presented in [7, 14, 15, 13], which are generalizations of [5] and [6]. There, the unobservable partition cannot change during the time even if each firm could default. Instead, in this paper, the partition of the population can change, since the credit level of any single firm can increase or can decrease while the time goes on. In addition to this, in every class each firm can also default.

Here, to define the class of each firm at every time t, the components $Z_1(t), \cdots, Z_H(t)$ of the Markov chain $Z(t) = \{Z_i(t)\}_{1 \leq i \leq H}$ are assumed to be exchangeable. Furthermore, to let the model be dynamic and to preserve the property of the firms of being undistinguishable, the exchangeability property, has to be extended by considering the finite dimensional distribution of $Z(t)$, for any fixed t. To this end, an exchangeability assumption on the finite dimensional distribution of $Z(t)$ is made.

The problem of constructing a process $Z(t)$, satisfying such exchangeability assumption, will be dealt with in. Moreover, $Z(t)$ is assumed to be a Markov process and some general properties of the model are discussed, both in the discrete and in the continuous time case.

The exchangeability property is investigated and the relation between the exchangeability and the Markov property is clarified. Such properties are essential tools to prove that the lifetimes are an exchangeable sequence of random variables.

The partition of the population is supposed to be non-observable. The observation process is just the number of defaults up to t, namely the history of the process $Y(t)$. The aim is to find a prediction of the residual lifetimes given the history of the observations.

In order to find the law of the lifetimes given $\{Y(s), s \leq t\}$, the number of firms belonging to the class of the defaulted firms, up to time t, is computed first. Then, the law of the lifetimes given the variable $Z(t)$ is constructed as well as the filter of $Z(t)$ given $\{Y(s), s \leq t\}$.

The continuous time filtering problem is studied, with stochastic filtering techniques. Strong uniqueness for the filtering equation is obtained. A linearized equation is introduced and a useful representation of its solution is derived following a method which is a modification of the one proposed in [9, 3]. Hence, a construction of the approximating process is derived and its convergence is studied.
On the other hand, a different point of view is investigated by introducing the occupancy numbers $X(t) = \{X_j(t)\}_{0\leq j\leq d}$, which are defined as the cardinality of each class. These numbers are introduced and their role is highlighted, generalizing some concepts presented in [6]. More, the relations are investigated between the properties of $Z(t)$ and of the occupancy numbers taking into account all the results given in [8, 15, 14].
The particular case in which the dynamics of $Z(t)$ only depends on the crowding of each class is analyzed, and a self-contained construction of the process $X(t)$ is given. Again, to recover information about the time to default of each firm, a filtering problem is set up. The procedure to deal with follows the line of the classical innovation method, see [1] for the general theory. In general, the filter π, which is a valued process probability measure, is not finite dimensional. But, in this setting, [8, 14], π is purely atomic with a finite number of atoms and completely described by the value on its atoms. Hence, regarding the approximation, the classical approach, pioneered by [12], allows us to obtain a really efficient algorithm to compute the filter itself.

2. Preliminaries

Given $\mathcal{P} = \{U_j\}_{j\geq 1}$, a finite or countable population of firms and H, a positive integer value, $\mathcal{P}_H = \{U_j\}_{j=1,\cdots H}$ denotes a finite subpopulation of $\mathcal{P}$. It is heterogeneous in that its elements can be of d different types, labelled by the natural numbers $1, \cdots, d$, and representing the credit ratings. For $t \in \mathbb{R}^+$, $C_k(t)$ is the subset of all firms of type k, $k = 1, \cdots, d$, at time t, and $C_0(t)$ is the class of the firms defaulted up to time t. Therefore, $\mathcal{P}_H = \cup_{k=0,1,\cdots,d}\, C_k(t)$.

Let us introduce the random variables $Z_i(t)$, defining the class of rating of each firm U_i,

$$(1) \qquad \forall\, t \in \mathbb{R}^+ \quad Z_i(t) = k \iff U_i \in C_k(t) \quad , \quad i = 1, \cdots, H,\ k = 0, 1, \cdots, d\,.$$

For $i = 1, \cdots, H$, $k = 0, 1, \cdots, d$ and $t > 0$, events of the form $\{U_i \in C_k(t)\} = \{Z_i(t) = k\}$ are measurable when $Z(t)$ is a stochastic process. Fixed $t \in \mathbb{R}^+$, $Z_i(t)$ is a random variable taking value in $E = \{0, 1, \cdots, d\}$ and $Z(t) = (Z_1(t), \cdots, Z_H(t)) \in \mathcal{H} = E^H$.

Assumption 1. *For all $t > 0$, $C_0(t)$ is an absorbing class, namely for $i = 1, \cdots, H$*

$$(2) \qquad Z_i(s) = 0 \implies Z_i(t) = 0 \quad , \quad \forall\, t \geq s,\ a.s.\,.$$

Fixed a t, $(Z_1(t), \cdots Z_H(t))$ is an exchangeable sequence, i.e. for each π, finite permutation of index, and, for each $a_1, \cdots, a_H \in E$,

$$(3) \qquad P(Z_1(t) = a_1, \cdots Z_H(t) = a_H) = P(Z_1(t) = a_{\pi(1)}, \cdots Z_H(t) = a_{\pi(H)})\,.$$

In this way, different firms can have different labels even if they are considered undistinguishable. But, since we are constructing a dynamical model, the exchangeability property has to be extended by considering the finite dimensional distribution of the process $Z(t) = (Z_1(t), \cdots, Z_H(t))$, for any fixed t. Thus,

Assumption 2. *For* $n \geq 1$, $t_1, \cdots, t_n \in \mathbb{R}^+$ *with* $t_1 \leq \cdots \leq t_n$, π *a permutation on* $\{1, \cdots, H\}$, $k^{(1)}, \cdots, k^{(n)} \in E^H$ *where* $k^{(i)} = \{k_1^{(i)}, \cdots, k_H^{(i)}\}$ *and* $\pi k^{(i)} = \{k_{\pi_1}^{(i)}, \cdots, k_{\pi_H}^{(i)}\}$, *then*

$$P\left(Z(t_1) = k^{(1)}, \cdots, Z(t_n) = k^{(n)}\right) = \tag{4}$$
$$= P\left(Z(t_1) = \pi k^{(1)}, \cdots, Z(t_n) = \pi k^{(n)}\right) .$$

Roughly speaking, this assumption is related to an "exchangeability property" of the trajectories of $Z(t)$. More, note that, if $t_1 = \cdots = t_n = t$, equation (4) implies equation (3).

2.1. **Lifetimes.** In order to discuss the most important consequence of (4), let $\{T_i\}_{i=1,\cdots,H}$ be such that

$$T_i = \inf\{t \in \mathbb{R}^+ \,:\, Z_i(t) = 0\} , \tag{5}$$

i.e. $\{T_i\}_{i=1,\cdots,H}$ is a sequence and each element T_i is the time to default of U_i, $i = 1, \cdots, H$.

While, from a mathematical point of view, (5) makes sense in any case, the condition given in (2) assures that T_i is just the time during which the firm U_i is in "some sense" alive, since it is not defaulted yet and it is still in the market.

Taking into account the structural properties of the model, the most important consequences of (4) is that $\{T_i\}_{i=1,\cdots,H}$ is a sequence of exchangeable random variables. Condition (4), which is very natural in our context, implies the result given in Proposition 1 below, while the exchangeability at fixed time does not. This result was already proved in [7].

Proposition 1. *If* (4) *holds, then* $\{T_i\}_{i\geq 1}$ *is a family of exchangeable random variables.*

To obtain the existence of a process $Z(t)$ verifying the conditions given in (2) and (4), a crucial tool is the following assumption.

Assumption 3. *The process* $Z(t)$ *is Markovian.*

2.2. **The discrete and the continuous time model.** For each $a, b \in \mathcal{H}$, let $\mu(a, b)$ be a transition probability and let ν_0 be a probability on $\mathcal{H}$.

Assumption 4. *The following conditions hold*

i) *If there exists an index* i *such that* $a_i = 0$ *and* $b_i \neq 0$ *then* $\mu(a, b) = 0$.
ii) *For* $a, b \in \mathcal{H}$ *and for each* β, *permutation of index, then* $\mu(a, b) = \mu(\beta a, \beta b)$.

Fixed a discrete time $t \in \mathbb{N}$, $Z(t)$ is the Markov chain with initial law ν_0 and with transitions defined as $P(Z(t) = b \mid Z(t-1) = a) := \mu(a, b)$, $a = (a_1, \cdots a_H)$ and $b = (b_1, \cdots b_H) \in \mathcal{H}$. More, the result given below can be proved as in [7].

Proposition 2. *If Assumption 4 holds and if* $Z(0) = (Z_1(0), \cdots, Z_H(0))$ *is an exchangeable sequence of random variables, then* $Z(t)$ *satisfies the conditions given in Assumptions 1 and 2.*

Let the continuous time model be defined by a continuous time Markov process $Z(t)$ with generator L given by

$$Lf(z) = l(z) \sum_{z' \in \mathcal{H}} [f(z') - f(z)]\, p(z, z') \tag{6}$$

where $l(z)$ is a positive function and $\{p(z,z')\}$ is a family of transition probabilities.

Assumption 5. *For any β, permutation of index on $\{1,\cdots,H\}$, let $l(z)=l(\beta z)$ and let $\{p(z,z')\}$ be a family of transition probabilities verifying Assumption 4.*

Since $\mathcal{H}$ is a finite set, there exist $\underline{l}$ and $\bar{l}$, constants, such that $0<\underline{l}\leq l(z)\leq \bar{l}$, $\forall\ z\in\mathcal{H}$. By construction, the generator L, given in (6), is a bounded operator. Then, there exists a unique solution of the $MgP(L,\nu_0)$, the martingale problem given the generator L and the initial condition ν_0, see [4]. This means that, for every probability measure ν_0 on $\mathcal{H}$, there exists $Z(t)$, a unique Markov process with sample paths in $D_{\mathcal{H}}[0,+\infty)$ (the space of right continuous $\mathcal{H}$-valued functions on $[0,\infty)$ having left limits), initial law ν_0 and generator L. Consequently, $Z(t)$ is the process defined as the unique solution of the $MgP(L,\nu_0)$.
The construction of a realization of $Z(t)$ consists in choosing a specific probability space and in defining a stochastic process having the same law of the unique solution of the $MgP(L,\nu_0)$. Any realization can be used to prove properties related to the law of a process, but of course, one of its realizations cannot be used to obtain properties of its paths.
Therefore, given a probability space $(\Omega,\mathcal{F},P)$, let $\{Z(n)\}_{n\geq 0}$ be a Markov chain with initial law ν_0 and transition probabilities $p(z,z')$. Let $\{V_i\}_{i\geq 1}$ be a sequence of random variables independent and esponentially distributed with parameter 1 and independent of $\{Z(n)\}_{n\geq 0}$. Finally, let

(7)
$$\tau_0=0\quad,\quad \tau_n=\sum_{i=1}^{n}\frac{V_i}{l(Z(i-1))}\quad\text{for } n>0\quad\text{and}$$
$$Z(t)=\sum_{n\geq 0}Z(n)\mathbb{I}_{\{\tau_n\leq t<\tau_{n+1}\}}$$

and $\mathcal{F}_t=\sigma\{Z(s),s\leq t\}$. Hence, on $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq 0},P)$, $Z(t)$ is a continuous time pure jump Markov process with generator L verifying Assumption 5, $\{\tau_i\}_{i\geq 1}$ is the sequence of its jump times and $N_t=\sum_{n\geq 0}\mathbb{I}_{\tau_n\leq t}$ is the counting process of all jumps up to time t.

Since all these properties are properties of the law of the process $Z(t)$, as in [15], without loss of generality, next results can be achieved by using the realization given in (7).

Theorem 1. *For any fixed t, the random variable $Z(t)=\{Z_i(t)\}_{1\leq i\leq H}$ is an exchangeable sequence of random variables and for $s\leq t$,*

$$P\left(Z(t)=k\mid Z(s)=h\right)=P\left(Z(t)=\beta(k)\mid Z(s)=\beta(h)\right)\ .$$

Furthermore, Assumption 2 holds, that is, for $t_0\leq t_1\leq\cdots\leq t_n$ and $a_0,a_1\cdots a_n\in\mathcal{H}$,

$$P\left(Z(t_1)=a_1,\cdots,Z(t_n)=a_n\right)=P\left(Z(t_1)=\beta(a_1),\cdots,Z(t_n)=\beta(a_n)\right)\ .$$

3. First approach: the indicator process Z

To get the conditional law of the lifetimes, given the history of $Y(t)$, which is the cardinality of the class $C_0(t)$, first the law of the lifetimes given $Z(t)$ has to be computed. Then the distribution of $Z(t)$, given the history of the process $Y(t)$, is obtained by solving a filtering problem. Setting, for each $z\in\mathcal{H}$, $\Phi(z)=$

$\sum_{i=1}^{H} \mathbb{I}_{\{z_i=0\}}$, then $Y(t) = \Phi(Z(t))$. Assuming that $t_1 \leq \cdots \leq t_H$, if t is such that $t_1 \leq t_2 \leq \cdots \leq t_m \leq t \leq t_{m+1} \leq \cdots \leq t_H$, for $P(Z(t) = k) \neq 0$ and $k \in \mathcal{H}$, the distribution of the lifetimes given the partition of the population at time t is

$$P(T_1 \leq t_1, T_2 \leq t_2, \cdots, T_H \leq t_H \mid Z(t) = k) = \frac{1}{P(Z(t) = k)} \times$$

$$\times P(Z_1(t_1) = 0, \cdots, Z_m(t_m) = 0, Z(t) = k, Z_{m+1}(t_{m+1}) = 0, \cdots, Z_H(t_H) = 0) .$$

This distribution is expressed as a function of $k \in \mathcal{H}$. Hence, to find the law of the lifetimes given the observed failures, only the law of $Z(t)$ given $\{Y(s),\ s \leq t\}$ is needed.

3.1. **The filtering problem.** Let $\pi_t(z) = P(Z(t) = z \mid \mathcal{F}_t^Y)$, where $\mathcal{F}_t^Y = \sigma\{Y(s), s \leq t\}$. At discrete time, since $Z(t)$ is a Markov chain and $Y = \Phi(Z)$, then the pair (Z, Y) is still Markovian. By using a Bayes formula and since the law of $Z(0)$ is known, the conditional law $\mathcal{L}(Z \mid Y)$, which is the filter, results to be

$$\pi_t(z) = \frac{\displaystyle\sum_{z' \in \mathcal{H}} P(Z(t) = z, Y(t) \mid Z(t-1) = z', Y(t-1))\, \pi_{t-1}(z')}{\displaystyle\sum_{z'', z''' \in \mathcal{H}} P(Z(t) = z''', Y(t) \mid Z(t-1) = z'', Y(t-1)\pi_{t-1}(z'')} .$$

Remark 1. Summing up,

$$P(T_1 \leq t_1, T_2 \leq t_2, \cdots, T_H \leq t_H \mid \mathcal{F}_t^Y) =$$
$$= \sum_{z \in \mathcal{H}} P(T_1 \leq t_1, T_2 \leq t_2, \cdots, T_H \leq t_H \mid Z(t) = z)\, \pi_t(z) .$$

At continuous time, for a function f smooth enough, the filter $\pi_t(f) = \mathbb{E}[f(Z(t)) \mid \mathcal{F}_t^Y]$ satisfies the Kushner-Stratonovich equation and, as in [7, 15], we get

Proposition 3. *For the continuous time model presented in this paper and for any real valued function $f(z), z \in \mathcal{H}$, the equation for the filter can be written as*

$$\pi_t(f) =$$

$$= \nu_0(f) + \int_0^t \{\pi_s(Gf) - \pi_s(mf) + \pi_s(m)\pi_s(f)\}\, ds + \tag{8}$$

$$+ \sum_{j=1}^{H} \int_0^t (\pi_{s-}(m_j))^+ \{\pi_{s-}(m_j f) - \pi_{s-}(m_j)\pi_{s-}(f) + \pi_{s-}(R_j f)\}\, dU_s^j ,$$

where, as usual in the filtering theory, $a^+ = (1/a)\mathbb{I}_{\{a>0\}}$ and where

$$Gf(z) = l(z) \sum_{z'} [f(z') - f(z)] \mathbb{I}_{\{\Phi(z')=\Phi(z)\}} \cdot p(z, z') ,$$
$$R_j f(z) = l(z) \sum_{z'} [f(z') - f(z)] \mathbb{I}_{\{\Phi(z') \neq \Phi(z)\}} \cdot \mathbb{I}_{\{\Phi(z')=j\}} \cdot p(z, z') .$$

In general, equation (8) does not have a unique solution. Thus, in order to deduce the properties of the filter, some kind of uniqueness for (8) is needed. Weak uniqueness could be obtained, by using the filtering martingale problem approach, as done in [9], [10] and [11].

On the other hand, since this model has a finite state space, a stronger kind of uniqueness can be reached. In particular, pathwise uniqueness for the solutions of the filtering equation is obtained with a procedure which has some kind of similarity with the one used in [7, 15]. Note that in this procedure we do not require $\pi_t(m_j) > 0$.

Theorem 2. *Let π'_t be a probability measure valued process, with cadlag trajectories, $\mathcal{F}^U_t$-adapted, satisfying equation* (8) *driven by the process $U(t)$. Then π'_t coincides pathwise with the filter.*

In order to assure that if none of defaultable firms have brought back to the market again, every transition must be possible, for the transition probabilities $p(z, z')$ verify conditions (4), a further condition is needed.

Assumption 6. *For all $z, z' \in \mathcal{H}$, such that if $z_i = 0$ then $z'_i = 0$, $i = 1, \cdots, H$, we get that $p(z, z') > 0$, and we are able to define a positive constant*

$$\underline{p} := \min_{z,z' \in \mathcal{H}} \{p(z, z') \; : \; \text{if } z_i = 0 \text{ then } z'_i = 0, \; i = 1, \cdots, H\} \; .$$

Remark 2. Consequently, if $j > \Phi(z)$ with $j = 1, \cdots, H$, then

$$m_j(z) \geq \underline{l} \sum_{z' \in \mathcal{H}} \mathbb{I}_{\{\Phi(z') > \Phi(z)\}} \mathbb{I}_{\{\Phi(z') = j\}} p(z, z') \geq \underline{l} \cdot \underline{p} \; .$$

An explicit expression for the solutions of equation (8) is not available, while in the discrete time model a simple recursive formula was given. This is the reason why, in the following an approximating discrete time model is constructed and then, the filter of the approximating discrete time model is proven to converge to the exact one. To this aim, we use a method which is a modification of the one proposed in [9]. Let the linearized equation be

$$\begin{aligned} \rho_t(f) = \nu_0(f) + \int_0^t \{\rho_s(Gf) - \rho_s(mf)\} \, ds + \\ + \sum_{j=0}^{H} \int_0^t \{\rho_{s-}(m_j f) - \rho_{s-}(f) + \rho_{s-}(R_j f)\} \, dU^j_s \; . \end{aligned} \tag{9}$$

This is a modification of equation (8). By classical arguments, not only equation (9) admits a unique solution in the weak sense but, by Lipschitz arguments, a unique strong solution exists which is necessarily $\mathcal{F}^U_t$-adapted.

Proposition 4. *Equation* (9) *admits at least one solution $\mathcal{F}^U_t$-adapted. In addition, denoted by $\rho_t(f)$ such solution, the following properties hold for any t*

i) *$\rho_t(f)$ is a finite positive measure,*

ii) $e^{-t \cdot \underline{l}}(1 \wedge \underline{l} \cdot \underline{p}) < \rho_t(1) \leq \bar{l} \vee 1$,

iii) $\pi_t(f) = \rho_t(f)/\rho_t(1)$.

3.2. **The approximating model.** The construction of the approximating process follows the same lines as in [2] and is related to the construction performed in the previous section. Fix $h > 0$ and set

$$\tau^h_0 = 0 \quad , \quad \theta^h_0 := 0 \quad \text{and} \quad \tau^h_n = h \sum_{i=1}^{n} \theta^h_i$$

where

$$\theta_i^h := \left[\frac{V_i}{h\ l(Z(i-1))}\right] + 1 \qquad i > 0,\ n > 0\ .$$

Then, over a finite time horizon $[0,T]$, with $T > 0$, for $t = kh$ and $k = 0, 1, \cdots$ such that $kh \leq T$, the approximating process is defined as $\zeta^h(t) = \sum_{n\geq 0} Z(n) \times \mathbb{I}_{\{\tau_n^h \leq t < \tau_{n+1}^h\}}$, where $Z(n)$ is the same Markov chain defined in (7). Hence, on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$, $\zeta^h(t)$ is a discrete time Markov chain, $\{\tau_i^h\}_{i\geq 1}$ is the sequence of its jump times and $N_t^h = \sum_{n\geq 0} \mathbb{I}_{\tau_n^h \leq t}$ denotes the counting process of all jumps of $\zeta^h(\cdot)$ up to time t.

Proposition 5. *The process $\zeta^h(t)$ is a discrete time Markov chain with transition probabilities*

$$\begin{aligned} \mu^h(z,z') &= P\left(\zeta^h((n+1)h) = z' \mid \zeta^h(nh) = z\right) = \\ &= \delta_{\{z,z'\}} e^{-hl(z)} + p(z,z')(1 - e^{-hl(z)})\ . \end{aligned} \tag{10}$$

Consequently, the discrete time observation processes are

$$Y^h(t) = \Phi(\zeta^h(t)) \quad , \quad U_t^{jh} = \sum_{k=1}^{H} \mathbb{I}_{\{T_k^h \leq t\}}\ \mathbb{I}_{Y^h(T_k^h) = j}$$

and

$$N_t^{Y^h} = \sum_{j=0}^{H} U_t^{jh} = \sum_{k=1}^{H} \mathbb{I}_{\{T_k^h \leq t\}}\ ,$$

where $\{T_k^h\}$ is the sequence of the jump times of $Y^h(t)$. Again, as in the continuous time case, $\mathcal{F}_t^{Y^h} = \mathcal{F}_t^{U^h}$, with $U^h = \{U^{jh}\}_{j=0,1,\cdots,H}$. Hence, we have to dealt with a filtering problem, in order to find the conditional law of $\zeta^h(t)$ given $\mathcal{F}_t^{U^h}$, that is $\pi_t^h(f) = \mathbb{E}[f(\zeta^h(t)) \mid \mathcal{F}_t^{U^h}]$ for a function f smooth enough. Note that $\pi_t^h(f)$ satisfies the equation

$$\begin{aligned} &\pi^h_{(n+1)h}(f) = \\ &= \pi^h_{nh}(f) + \left[\pi^h_{nh}(G^h f) + \pi^h_{nh}(m^h)\pi^h_{nh}(f) - \pi^h_{nh}(m^h f)\right] \cdot \\ &\cdot (1 - \pi^h_{nh}(m^h))^+ (1 - \Delta N^{Y^h}_{(n+1)h}) + \\ &+ \sum_{j=0}^{H} \pi^h_{nh}(m^h_j)^+ \left[\pi^h_{nh}(m^h_j f) - \pi^h_{nh}(m^h_j)\ \pi^h_{nh}(f) + \pi^h_{nh}(R^h_j f)\right] \Delta U^{jh}_{(n+1)h}\ , \end{aligned} \tag{11}$$

where

$$\begin{aligned} G^h f(z) &= \sum_{z'} [f(z') - f(z)]\, \mathbb{I}_{\Phi(z') = \Phi(z)}\ \mu^h(z,z')\ , \\ R^h f(z) &= \sum_{j=1}^{H} R^h_j f(z) = \sum_{j=1}^{H} \sum_{z'} [f(z') - f(z)] \cdot \mathbb{I}_{\Phi(z') \neq \Phi(z)} \cdot \mathbb{I}_{\Phi(z') = j} \cdot \mu^h(z,z') \end{aligned}$$

and also

$$m_j^h(z) = \sum_{z'} \mathbb{I}_{\Phi(z')\neq\Phi(z)} \cdot \mathbb{I}_{\Phi(z')=j} \cdot \mu^h(z,z') \ ,$$

$$m^h(z) = \sum_{j=1}^{H} m_j^h(z) = \sum_{z'} \mathbb{I}_{\Phi(z')\neq\Phi(z)} \cdot \mu^h(z,z') \ .$$

Equation (11) has a unique solution as a consequence of its recursive structure, taking into account (10) and the inequality

$$m_j^h(z) \leq m^h(z) \leq 1 - e^{-h\cdot\bar{l}} < 1 \quad , \quad \forall\, j = 1, \cdots, H \ . \tag{12}$$

A useful tool to prove the convergence of the discrete time model to the continuous time one is the linearized version of equation (11) given in Proposition 6 and already proved in [7].

Proposition 6. *The equation*

$$\rho_{nh}^h(f) = \nu_0(f) + \sum_{k=1}^{n}\{\rho_{(k-1)h}^h(G^h f) - \rho_{(k-1)h}^h(m^h f)\}(1 - \Delta N_{kh}^{Y^h}) + \tag{13}$$

$$+\sum_{j=1}^{H}\sum_{k=1}^{n}(1-e^{-h})^{+}\left\{\rho_{(k-1)h}^h(m_j^h f) - (1-e^{-h})\rho_{(k-1)h}^h(f) + \rho_{(k-1)h}^h(R_j^h f)\right\}\Delta U_{kh}^{jh}$$

admits a unique solution $\mathcal{F}_t^U$-adapted. Denoted by $\rho_t^h(f)$ this solution, the following properties hold for any $t = nh$

i) *$\rho_t^h(f)$ is a finite positive measure,*

ii) $0 < \rho_t^h(1) \leq (2\cdot\bar{l})^{N_t^{Y^h}} \vee 1$, $\forall\, h\cdot(\underline{l}\vee 1) < \log 2$,

iii) $\pi_t^h(f) = \rho_t^h(f)/\rho_t^h(1)$.

3.3. **Convergence.** Let $S^h := (\zeta^h, Y^h, U^h, N^h, \pi^h)$ denote the piecewise constant cadlag continuous time interpolation of the processes introduced previously. Set $S := (\zeta, Y, U, N, \pi)$, and $\mathcal{S} = \mathcal{H}\times\{0,1,\cdots,H\}\times\{0,1\}^H\times\mathbb{N}\times\Pi(\mathcal{H})$, where $\Pi(\mathcal{H})$ is the space of probability measure on $\mathcal{H}$.

Theorem 3. *The process S^h converges to the process S a.s., as $h\to 0$, with respect to the Skorohod topology on the space $D_{\mathcal{S}}[0,T]$.*

The proof of Theorem 3 is a consequence of the next results already proved in [7]. Note that in our model $P(\tau_{N_T} = T) = 0$.

Proposition 7. *If*

$$h < \frac{T - \tau_{N_T}}{N_T} \ ,$$

then $\tau_{N_T}^h \leq T$ which in turn implies $N_T = N_T^h$.

On $(N_T = N_T^h)$, a function $\alpha_h(\cdot)$ of $[0,T]$ into itself can be defined such that:

(i) $\alpha_h(\cdot)$ is a piecewise linear map and transforms the intervals $[\tau_k^h, \tau_{k+1}^h)$ into $[\tau_k, \tau_{k+1})$, $\forall\, k < N_T$, and $[\tau_{N_T}^h, T)$ into $[\tau_{N_T}, T)$.

(ii) $\sup_{t\in[0,T]} |\alpha_h(t) - t| = \max_{k\leq N_T} |\tau_k - \tau_k^h| \leq \max_{k\leq N_T} kh = N_T h$.

(iii) $|\zeta(\alpha_h(t)) - \zeta^h(t)| = |Y(\alpha_h(t)) - Y^h(t)| = |\ U(\alpha_h(t)) - U^{jh}(t)| =$

$= |N_{\alpha_h(t)} - N_t^h| = 0.$

Then, to reach the result claimed in Theorem 3, the convergence of the filters is just enough.

Theorem 4. *Under the assumptions given in this paper (in particular Assumption 6), we have*

i) $\|\pi_{\alpha_h(t)} - \pi_t^h\| \leq \dfrac{2 \cdot e^{T\cdot \bar{l}}}{1 \vee \underline{l} \cdot \underline{p}} \cdot \left\|\rho_{\alpha_h(t)} - \rho_t^h\right\|$

ii) $\left\|\rho_{\alpha_h(t)} - \rho_t^h\right\| \leq (1 + 2\ \bar{l})^{N_t}\ e^{2t\cdot \bar{l}}\ C\,h$, *for a suitable quantity*

$C = C(T, N_T, \bar{l}) > 0$

and, finally, $\|\pi_{\alpha_h(t)} - \pi_t^h\| \leq C \cdot h$.

4. Second approach: the occupancy numbers X

Under the assumption of exchangeability on the trajectories of the indicator process, a different approach could be followed. Taking into account that the relevant informations are given by the number of firms belonging to a given credit class, at any time t, and the number of defaulted firms.

As a consequence, from now on, the state of the model is described by the number of firms in any class at time t, and by the number of firms defaulted up to time t.

Definition 1. For $t \in \mathbb{R}^+$ and for $i = 1, \cdots, d$, let $\Phi_i(z) = \sum_{j=1}^H \mathbb{I}_{z_j=i}$ and

$$X^i(t) = \#C_i(t) = \Phi_i(Z(t)) \quad , \quad Y(t) = \#C_0(t) = \Phi_0(Z(t))$$

be the occupancy numbers. In particular, $Y(t)$ is the number of firms defaulted up to time t.

Assumption 2 implies that $Y(t)$ is non-decreasing with respect to t a.s. The pair $(X(t), Y(t))$ takes values in $\mathcal{K} := \{(x_1, \cdots, x_d, y) : x_i, y \in \mathbb{N} \cup \{0\}, \forall\, i;\ x_1 + \cdots + x_d + y = H\}$ while the process $X(t)$ takes values in $\mathcal{X} := \{(x_1, \cdots, x_d) : x_i \in \mathbb{N} \cup \{0\},\ \forall\, i,\ x_1 + \cdots + x_d \leq H\}$.
Choosing the occupancy numbers as state variables, some of the results already obtained has to be expressed in terms of $X(t)$. To this end, a relation between the law of Z, $\mathcal{L}(Z)$, and the joint law of the occupancy numbers X and Y, $\mathcal{L}(X, Y)$ have to be established.
Since $X(t) = (X^1(t), \cdots, X^d(t))$, with $\Phi = (\Phi_1, \cdots, \Phi_d, \Phi_0)$ then $(X(t), Y(t)) = \Phi(Z(t))$ and for $s^{(1)}, \cdots, s^{(n)} \in \mathcal{K}$, $n > 0$,

$$P\Big((X(t_i), Y(t_i)) = s^{(i)},\ i = 1, \cdots, n\Big) = P\Big(Z(t_i) \in \Phi^{-1}(s^{(i)}),\ i = 1, \cdots, n\Big)\ .$$

Note that Φ is a deterministic function not necessarily one-to-one. Therefore, given $\mathcal{L}(Z)$, $\mathcal{L}(X, Y)$ is known. In general the converse is not true. For a fixed t, given $\mathcal{L}(X(t), Y(t))$, $\mathcal{L}(Z(t))$ is known, [5], as follows from Proposition 8.

Proposition 8. *For $k \in \{0,1,\cdots,d\}^H$, if $Z(t)$ is an exchangeable sequence for any fixed $t \in \mathbb{R}^+$, then*

$$P(Z(t)=k) = \frac{\Phi_0(k)!\Phi_1(k)!\cdots\Phi_d(k)!}{H!} \cdot P((X(t),Y(t)) = \Phi(k)) \; .$$

Since $\Phi(k) = \Phi(\beta k)$ for all k, this last result agrees with Assumption 2. However, a generalization of this result in a dynamic context is needed. This generalization can be achieved when the dynamics of the process $Z(t)$ only depend on the number of particles belonging to each class. For the following assumption, recall that $\Phi = (\Phi_1,\cdots,\Phi_d,\Phi_0)$ and $\Phi_i(z) = \sum_{j=1}^H \mathbb{I}_{z_j=i}$, for $i=0,1,\cdots,d$.

Assumption 7. *Let $\gamma(h,\widetilde{h}) := \prod_{i=1}^H (\mathbb{I}_{h_i\neq 0} + \mathbb{I}_{h_i=0,\widetilde{h}_i=0})$, for $h,\widetilde{h} \in \{0,1,\cdots,d\}^H$ and $t_1 \le \cdots \le t_n$ with $n \ge 1$. We assume that*

$$(14)\quad P\left(Z(t_1)=h^{(1)},\cdots,Z(t_n)=h^{(n)}\right) = P\left(Z(t_1)={h'}^{(1)},\cdots,Z(t_n)={h'}^{(n)}\right)$$

for all $(h^{(1)},\cdots,h^{(n)})$ and $\left({h'}^{(1)},\cdots,{h'}^{(n)}\right)$ with $h^{(i)},{h'}^{(i)} \in \{0,1,\cdots,d\}^H$ such that

i. $\Phi(h^{(i)}) = \Phi({h'}^{(i)})$ *, $i=1,\cdots,n$,*

ii. $\gamma(h^{(i)},h^{(i+1)}) = \gamma({h'}^{(i)},{h'}^{(i+1)})$*, $i=1,\cdots,n-1$.*

Note that, choosing ${h'}^{(i)} = \beta h^{(i)}$ for $i=1,\cdots,n$, Assumption 7 implies Assumption 2 while the converse is not true. To better understand the meaning of Assumption 7, notice that Equation (14) and Condition i. imply that the finite dimensional distributions of the process $Z(t)$ coincide if the vectors $h^{(i)}$ and ${h'}^{(i)}$, for $i=1,\cdots,n$, produce the same vector of occupancy numbers. On the other hand, Condition ii. guarantees Assumption 2. In Proposition 9 below, proved by combinatorial techniques, $\mathcal{L}(Z)$ is obtained in terms of $\mathcal{L}(X,Y)$. This is a generalization of Proposition 8 and an essential tool to reach this result is Assumption 7.

Proposition 9. *Under Assumption 7,*

$$P(Z(t_1)=h^{(1)},\cdots,Z(t_n)=h^{(n)}) =$$
$$= A(h^{(1)},\cdots,h^{(n)}) \cdot P\left((X(t_1),Y(t_1)) = \Phi(h^{(1)}),\cdots,(X(t_n),Y(t_n)) = \Phi(h^{(n)})\right)$$

where $A(h^{(1)},\cdots,h^{(n)})$ are deterministic quantities given by

$$A(h^{(1)},\cdots,h^{(n)}) = \frac{\Phi_0(h^{(1)})!\Phi_1(h^{(1)})!\cdots\Phi_d(h^{(1)})!}{H!} \cdot$$
$$\cdot \prod_{j=1}^{n-1} \frac{(\Phi_0(h^{(j+1)})-\Phi_0(h^{(j)}))!\Phi_1(h^{(j+1)})!\cdots\Phi_d(h^{(j+1)})!}{(H-\Phi_0(h^{(j)}))!} \cdot \gamma(h^{(j)},h^{(j+1)}) \; .$$

The Markov property of $(X(t),Y(t))$ is the key for the existence of a process $Z(t)$ satisfying Assumption 7. Indeed, in such a case, $(X(t),Y(t))$ can be constructed as the solution of a suitable Martingale problem and the law of $Z(t)$ can be deduced from Proposition 9. Unfortunately, if $Z(t)$ is a Markov process this, in general, does not implies that such is $(X(t),Y(t))$, see [15], but

Proposition 10. *Under Assumption 7, $(X(t),Y(t))$ is a Markov process if and only if $Z(t)$ is a Markov process.*

As a consequence of Assumption 7, for $\Phi(z) = \Phi(\overline{z})$ and $\Phi(z') = \Phi(\overline{z'})$, with $z, \overline{z}, z', \overline{z}' \in \mathcal{H}$, the coefficients of the generator of $Z(t)$, given in (6), verify that $l(z) = l(z')$ and $p(z, z') = p(\overline{z}, \overline{z}')$. This implies that, there exists a function $\Lambda(x, y)$, for each $(x, y) \in \mathcal{K}$, and there exists a transition function $M((x, y); (x', y'))$, for each (x, y) and $(x', y') \in \mathcal{K}$, such that

$$\Lambda(\Phi(z)) = l(z) \quad \text{and} \quad M(\Phi(z), \Phi(z')) = p(z, z') \; , \tag{15}$$

for some $z, z' \in \mathcal{H}$, and $M((x, y); (x', y')) = 0$, for $y > y'$. This in turn implies that, for each real valued function f, the process $(X(t), Y(t))$ has the generator

$$L^{x,y} f(x, y) = \Lambda(x, y) \sum_{(x',y') \in \mathcal{K}} [f(x', y') - f(x, y)] M((x, y); (x', y')) \; . \tag{16}$$

Again, by construction, the generator L^z, given in (6), is a bounded operator and, consequently, the generator given in (16) is a bounded operator. Thus, [4], for all $\tilde{\nu}_0$, probability measure on $\mathcal{K}$, there exists $(X(t), Y(t))$, a unique Markov process with sample paths in $D_{\mathcal{K}}[0, +\infty)$, initial condition $\tilde{\nu}_0$ and generator $L^{x,y}$. Once $(X(t), Y(t))$ is uniquely determined in law, the law of $Z(t)$ can be deduced from (15) and verifies Assumption 7 by construction.

4.1. **The filtering problem.** The partition of the population is assumed to be non-observable. The observation is just the cardinality $Y(t)$ of $C_0(t)$, which is an absorbing class. Setting $\mathcal{F}_t^Y = \sigma\{Y(s), s \le t\}$, our aim is to find $\mathcal{L}(T_1, \cdots, T_H \mid \mathcal{F}_t^Y)$, namely the conditional law of the times to the defaults given the history of $Y(t)$. To this end, we are going to compute first $\mathcal{L}(T_1, \cdots, T_H \mid X(t))$ and then the filter, the conditional law of $X(t)$ given $\mathcal{F}_t^Y$, that is $\pi_t(x) = P\left(X(t) = x \mid \mathcal{F}_t^Y\right)$. Set $|x| = x_1 + \cdots + x_d$ for each $x \in \mathcal{X}$. Then $\mathcal{L}(T_1, \cdots, T_H \mid X(t))$ coincides with $\mathcal{L}(T_1, \cdots, T_H \mid X(t), Y(t))$ since $Y(t) = H - |X(t)|$.

Proposition 11. *Under Assumption 7, if $t_1 \le \cdots \le t_H$ and t is such that*

$$t_1 \le t_2 \le \cdots \le t_m \le t \le t_{m+1} \le \cdots \le t_H \; ,$$

and if $P(X(t) = x) \neq 0$, with $x \in \mathcal{X}$ then

$$P(T_1 \le t_1, T_2 \le t_2, \cdots, T_H \le t_H \mid X(t) = x) = \tag{17}$$

$$= \frac{\displaystyle\sum_{k \in \Phi^{-1}(x, H-|x|)} P(T_1 \le t_1, T_2 \le t_2, \cdots, T_H \le t_H, Z(t) = k)}{\displaystyle\sum_{h \in \Phi^{-1}(x, H-|x|)} P(Z(t) = h)}$$

where

$$P(Z(t) = k) = \sum_{h \in \mathcal{H}} p^{(t)}(h, k) \cdot P(Z(0) = h)$$

and

$$P(T_1 \le t_1, T_2 \le t_2, \cdots, T_H \le t_H, Z(t) = k) =$$
$$= \sum_{h, x_{i,j} \in \mathcal{H},\ 2 \le i \le H,\ 1 \le j \le i-1} p^{(t_1)}(h, (0, x_{2,1}, \cdots, x_{H,1})) \cdot$$
$$\cdots p^{(t - t_m)}((0, \cdots, 0, x_{m+1,m}, \cdots, x_{H,m}), k) \cdot$$
$$\cdot p^{(t_{m+1} - t)}(k, (0, \cdots, 0, x_{m+2,m+1}, \cdots, x_{H,m+1})) \cdot$$
$$\cdots p^{(t_H - t_{H-1})}((0, \cdots, 0, x_{H,H-1}), (0, \cdots, 0)) \cdot$$

$$\cdot P(Z(0) = h)\mathbb{I}_{k_1=\cdots=k_m=0} \ .$$

At discrete time, $X(t)$ is a Markov chain and $Y(t)$ is a deterministic function of $X(t)$. In this case, using the Bayes formula, the filter results to be

$$\pi_t(x) = \frac{\displaystyle\sum_{x'\in\mathcal{X}} P(X(t) = x, Y(t) \mid X(t-1) = x', Y(t-1))\ \pi_{t-1}(x')}{\displaystyle\sum_{x'',x'''\in\mathcal{X}} P\left(X(t) = x''', Y(t) \mid X(t-1) = x'', Y(t-1)\right)\pi_{t-1}(x'')} \tag{18}$$

where the law of $X(0)$ is known.

At continuous time, a dynamic version of Bayes' formula is obtained observing that the filter is a solution of a stochastic differential equation known as Kushner-Stratonovich equation, see [1]. Since $Y(t) = H - |X(t)|$ and $(X(t), Y(t))$ is Markovian, then $X(t)$ is Markovian with generator given by

$$Lf(x) = \lambda(x) \sum_{x'\in\mathcal{X}} [f(x') - f(x)]\mu(x, x')$$

where $\lambda(x) = \Lambda(x, H - |x|)$, $\mu(x, x') = M((x, H - |x|); (x', H - |x'|))$, and the functions Λ and M defined in (15). For the sequel, a suitable representation of the generator L is given by

$$Lf(x) = L_0 f(x) + \sum_{j=1}^{H} L_1^j f(x) \ ,$$

where

$$L_0 f(x) = \lambda(x) \sum_{x'\in\mathcal{X}} [f(x') - f(x)]\mu_0(x, x') \ ,$$

$$L_1^j f(x) = \lambda(x) \sum_{x'\in\mathcal{X}} [f(x') - f(x)]\ \mathbb{I}_{|x'|=j} \cdot \mu_1(x, x') \ .$$

The transitions are such that $\mu_0(x, x') = \mu(x, x')\mathbb{I}_{|x'|=|x|}$ and $\mu_1(x, x') = \mu(x, x')\cdot$ $\cdot\mathbb{I}_{0\le|x'|<|x|}$, which guarantee that $(X(t), Y(t)) \in \mathcal{K}$, whatever the initial condition does. Then, $X(t)$ is a jump process taking values in $\mathcal{X}$ and the observation $Y(t)$ counts all the jumps of $|X(t)|$.

This dynamics admits the following interpretation. For $X(t) = x$, $X(t)$ jumps following a transition function $\mu_0(x, x')$ and $Y(t)$ does not jump. Otherwise, $X(t)$ jumps following a transition function $\mu_1(x, x')$ and in this case, $Y(t) := H - |X(t)|$ increases. In this context, only the second kind of jumps are registered as observations.

By construction, there exists $\widetilde{\tau}$, the first time such that $Y(t) = H$. Thus, $\widetilde{\tau}$ is an $\mathcal{F}_t^Y$-stopping time, i.e. $\widetilde{\tau}$ is an observable random variable and for $\widetilde{\tau} < +\infty$ and for $t \ge \widetilde{\tau}$, $|X(t)| = 0$ and $Y(t) = H$.

Since, in general, the component of $Z(t)$ are not independent, $Y(t)$ cannot be considered a counting process. As in [3], to overcome this problem, let $U(t) = (U^1(t), \cdots, U^H(t))$ be the multivariate point processes, defined as

$$U^j(t) := \sum_{i\ge1} \mathbb{I}_{\{\tau_i\le t\}}\ \mathbb{I}_{\{Y(\tau_i)=j\}} \qquad j = 1, \cdots, H \ , \tag{19}$$

where $\{\tau_i\}_{i\ge1}$ is the sequence of the jump times of $Y(t)$ and a subset of $T_{(1)}, \cdots,$ $T_{(H)}$, an order statistic of the lifetimes. Then, $N_t = \sum_{j=1}^{H} U^j(t) = \sum_{i\ge1} \mathbb{I}_{\{\tau_i\le t\}}$ counts all the jumps of $Y(t)$, up to time t, and $U^j(t)$ counts the number of jumps bringing $Y(t)$ on j. Being $Y(t)$ non-decreasing, by definition, $U^j(t)$ is $\{0,1\}$-valued, for all j. Moreover, the relation

$$Y(t) = Y(0) + \int_0^t \sum_{j=1}^H [j - Y(s-)] dU^j(s)$$

implies that $\mathcal{F}_t^Y = \mathcal{F}_t^U$, where $\mathcal{F}_t^U = \sigma\{U^1(s), \cdots, U^H(s), s \le t\}$. Then, our problem reduces to find the filter that is the conditional law of the process $X(t)$ given $\mathcal{F}_t^U$, $\pi_t(f) = \mathbb{E}[f(X(t)) \mid \mathcal{F}_t^U]$, for any real valued function f. Now, the intensity function of $U(t)$ is given by

$$\lambda^j(x) := \lambda(x) \sum_{x' \in \mathcal{X}} \mathbb{I}_{|x'|=j} \cdot \mu_1(x,x') =$$

$$= \lambda(x) \sum_{x' \in \mathcal{X}} \mathbb{I}_{|x|>|x'|=j} \cdot \mu(x,x') \quad , \quad \forall\, j = 1, \cdots, H \ .$$

Theorem 5. *The process $U^j(t)$ has minimal intensity given by $\pi_t(\lambda^j)$, for all $j = 1, \cdots, H$ and, for any real function $f(x)$ and $x \in \mathcal{X}$, the Kushner-Stratonovich equation can be written as*

$$\pi_t(f) = \nu_0(f) + \int_0^t [\pi_s(L_0 f) + \pi_{s-}(f)\pi_{s-}(\lambda) - \pi_{s-}(\lambda f)] \cdot ds + \tag{20}$$

$$+ \sum_{j=1}^H \int_0^t (\pi_{s-}(\lambda^j))^+ \cdot [\pi_{s-}(L_1^j f) - \pi_{s-}(f)\pi_{s-}(\lambda^j) + \pi_{s-}(\lambda^j f)] \cdot dU^j(s) \ .$$

More, equation (20) *has a unique pathwise solution, necessarily, $\mathcal{F}_t^Y$-adapted.*

The procedure to deduce equation (20) follows the line of the classical innovation method (for instance see [1]). Its uniqueness can be obtained as in [14].

4.2. **Approximation.** In general, the filter, π, which is a probability measure valued process, is not finite dimensional. But, in our setting, π is purely atomic with a finite number of atoms and it is completely described by the value on its atoms. In the following, for all $x \in \mathcal{X}$ and $t \ge 0$, set $\Pi_t(x) := \int_{\mathcal{X}} \mathbf{I}_{\{\xi = x\}} \pi_t(d\xi) = \mathbb{E}[\mathbf{I}_{\{X_t = x\}} \mid \mathcal{F}_t^Y]$. Then $\Pi_t := \{\Pi_t(x)\}_{x \in \mathcal{X}}$ is completely determined by the finite family $\{\Pi_t(x)\}_{x \in \mathcal{X}}$, that is the vector indexed by $x \in \mathcal{X}$, taking values in $\mathcal{S}' := \big\{\{\Pi(x)\}_{x \in \mathcal{X}} \,:\, \Pi(x) \ge 0, \forall\, x \in \mathcal{X}; \sum_{x \in \mathcal{X}} \Pi(x) = 1\big\}$, where $\mathcal{S}' \subset [0,1]^{\mathcal{X}}$. Therefore, the problem turns into a finite dimensional controlled totally observed problem, where the state variables are Y_t and Π_t. Defining $\Lambda^j := \{\lambda^j(x)\}_{x \in \mathcal{X}}$ for $j = 1, \cdots, H$, the minimal intensity of U^j is $\pi_t(\lambda^j(\cdot)) := (\Lambda^j, \Pi_t) = \sum_{x \in \mathcal{X}} \lambda^j(x) \cdot \Pi_t(x)$. Then equation (20) becomes

$$\Pi_t = \Pi_0 + \int_0^t \mathbf{b}(\Pi_s)\, ds + \sum_{j=1}^H \int_0^t \mathbf{a}^j(\Pi_{s-})\, dU^j(s) \tag{21}$$

where $\Pi_0 = \{\Pi_0(x),\ x \in \mathcal{X}\} = \{\nu_0(\mathbb{I}_{\cdot = x}),\ x \in \mathcal{X}\}$ and for $j = 1, \cdots, H$,

$$\begin{cases} \mathbf{b}(\Pi_s)(x) = \displaystyle\sum_{x' \in \mathcal{X}} \lambda(x')\mu_0(x',x)\Pi_s(x') + (\Lambda, \Pi_s)\Pi_s(x) - 2\lambda(x)\Pi_s(x) \\ \mathbf{a}^j(\Pi_s)(x) = \left[\displaystyle\sum_{x' \in \mathcal{X}} \lambda^j(x')\mu_1(x',x)\Pi_s(x') - (\Lambda^j, \Pi_s)\Pi_s(x)\right] (\Lambda^j, \Pi_s)^+ \ . \end{cases} \tag{22}$$

Equation (21) is very useful to get discrete time approximation for the filter in many ways. In the previous section a method is described, providing an almost sure convergence in the Skorohod topology. But here, we follow the classical approach, as suggested by the results given in [12]. Even if this approach allows us to obtain only the weak convergence of the discrete time approximation to the true filter, a really efficient algorithm is obtained for computing the filter itself. The generator of the process Π_t is

$$\mathcal{L}f(\Pi) = (\nabla f, \mathbf{b}(\Pi)) + \sum_{j=1}^{H} (\Lambda^j, \Pi) \left[f(\Pi + \mathbf{a}^j(\Pi)) - f(\Pi) \right] =$$

$$= \sum_{x \in \mathcal{X}} \frac{\partial f}{\partial \Pi(x)} \, \mathbf{b}(\Pi)(x) + \sum_{j=1}^{H} (\Lambda^j, \Pi) \left[f(\Pi + \mathbf{a}^j(\Pi)) - f(\Pi) \right] .$$

Let us define, for $h > 0$,

$$\mathcal{L}_h f(\Pi) = \frac{f(\Pi + h \cdot \mathbf{b}(\Pi)) - f(\Pi)}{h} + \sum_{j=1}^{H} (\Lambda^j, \Pi) \left[f(\Pi + \mathbf{a}^j(\Pi)) - f(\Pi) \right] .$$

Remark 3. From equation (22), since $\sum_{x \in \mathcal{X}} \mathbf{b}(\Pi_s)(x) = 0$ and $\sum_{x \in \mathcal{X}} \mathbf{a}^j(\Pi_s)(x) = 0$, for $j = 1, \cdots, H$, if $\Pi \in \mathcal{S}'$ then $\Pi + h \cdot \mathbf{b}(\Pi)$ and $\Pi + \mathbf{a}^j(\Pi) \in \mathcal{S}'$, for any positive h and $j = 1, \cdots, H$.

Let the time interpolation interval be defined as

$$\Delta t^h(\Pi) := \frac{h}{1 + (\Lambda, \Pi) \cdot h}$$

and let Π_n^h be the Markov chain with transition probabilities given by

$$p^h(\Pi, \Pi + h \cdot \mathbf{b}(\Pi)) = \frac{1}{1 + (\Lambda, \Pi) \cdot h} \quad \text{and}$$

(23)

$$p^h(\Pi, \Pi + \mathbf{a}^j(\Pi)) = \frac{(\Lambda^j, \Pi) \cdot h}{1 + (\Lambda, \Pi) \cdot h} ,$$

for $j = 1, \cdots, H$, where $p^h(\Pi, \widetilde{\Pi})$ are zero for all nonlisted values of $\widetilde{\Pi}$.

Definition 2. Let the interpolated continuous time process $\Pi^h(\cdot)$ be defined as

$$\Pi^h(t) := \Pi_n^h \quad \text{on} \quad t \in [t_n^h, t_{n+1}^h) \quad , \quad t_n^h := \sum_{i=0}^{n-1} \Delta t^h(\Pi_i^h) . \tag{24}$$

Let $\mathbb{E}_n^h$ and $\mathbb{E}_t^h$ be, respectively, the conditional expectation w.r.t. the σ-algebras generated by $\{\Pi_i^h \; i \leq n\}$ and $\{\Pi^h(s) \; s \leq t\}$. As in [14], a direct computation provides that

Lemma 1. *There exists a constant $K \in \mathbb{R}^+$ such that*

$$\mathbb{E}_n^h \left[\sum_{x \in \mathcal{X}} \left| \Pi_{n+1}^h(x) - \Pi_n^h(x) \right| \right] \leq \Delta t^h(\Pi_n^h) \cdot K .$$

Finally, the convergence of $\{\Pi^h(\cdot)\}$ to $\{\Pi(\cdot)\}$ can be achieved, first, by proving the tightness of the sequence and then by the identification of the limit, as done in [14].

5. Conclusions

Real data on individual credit rating histories are not always available. This is the reason why, prediction of future credit ratings for some given set of companies is the aim not only for this paper but also for any credit risk analysis.
However, taking a look to a real financial market, many are the generalizations that can be made on this model. First of all, the cardinality of the population of the firms can be supposed increasing, in order to model the situation in which new firms enter in the market, for instance when the economy is going well.
Then, we can assume that there are two or more partitions of the populations: some given by the rating agencies and the actual one. The observations could be the partitions coming from the agencies and the time of the default of each firm. The aim is to achieve informations about the actual credit rating.
Furthermore, some numerical experiments can be performed by using both simulating and real data.

Acknowledgements
The author is grateful to Prof. Silvia Centanni for many useful comments and observations which help to developed this model in a real financial environment.

References

[1] P. Brémaud, *Point processes and queues*, Springer-Verlag, 1980.
[2] A. Calzolari & G. Nappo, *A filtering problem with counting observations: error bounds due to the uncertainty on the infinitesimal parameters*, Stoch. and Stoch.Rep., 61(1997), 1–19.
[3] A. Calzolari & G. Nappo, *The filtering problem in a model with grouped data and counting observation times*, Dipartimento di Matematica G. Castelnuovo, Universitá di Roma La Sapienza, Preprint, 2001.
[4] S. N. Ethier & T. G. Kurtz, *Markov processes: characterization and convergence*, J. Wiley, 1986.
[5] A. Gerardi, F. Spizzichino & B. Torti, *Exchangeable mixture models for lifetimes: the role of "occupation number"*, Statist. Probab. Lett., 49(2000), 365–375.
[6] A. Gerardi, F. Spizzichino & B. Torti, *Filtering equations for the conditional law of residual lifetimes from a hetereogenous population*, J. Appl. Probab., 37(2000), 823–834.
[7] A. Gerardi & P. Tardelli, *Heterogeneous population dynamical model: a filtering problem*, J. Appl. Probab., 42(2005), 346–361.
[8] A. Gerardi & P. Tardelli, *Controlled hetereogeneous collection: the role of occupation numbers*, J. Appl. Probab., 45(2008), 595–609.
[9] W. Kliemann, G. Koch & F. Marchetti, *On the unnormalized solution of the filtering problem with counting process observations*, IEEE Trans. Inf. Theory, 36(1990), 1415–1425.
[10] T.G. Kurtz, *Martingale problems for conditional distribution of Markov processes*, Electronic J. Probab., 3(1998), 29 pp. .
[11] T.G. Kurtz & D. Ocone, *Unique characterization of conditional distributions in nonlinear filtering*, Ann. Probab., 16(1988), 80–107.
[12] H.J. Kushner & P. G. Dupuis, *Numerical methods for stochastic control problems in continuous time*, Springer-Verlag, 1992.
[13] P. Tardelli, *Modelling a population under maintenance*, Atti del 5^o Congresso Nazionale della Società Italiana di Biometria, Sezione Italiana dell'International Biometric Society, (2003), 137–140.
[14] P. Tardelli, *Conditional law of lifetimes for a heterogeneous population of living particles via filtering techniques*, Sci. Math. Jpn., 64(2006), 915–932.
[15] P. Tardelli, *Occupancy numbers for dynamic heterogeneous populations: estimate of particles lifetimes*, J. Statist. Plann. Inference, 137(2007), 3127–3139.