Detecting stressed periods from history of credit migrations : a point process filtering approach

Areski Cousin, Picard Tom
Nexialog Consulting, Paris, France

April 17, 2020

Abstract

Analyzing the effect of business cycle on rating transitions has been a subject of great interest these last fifteen years, particularly due to the increasing pressure coming from regulators for stress testing. In this paper, we consider that the dynamics of rating migrations is governed by an unobserved latent variable. Under a point process filtering framework, we explain how the current state of the hidden factor can be efficiently inferred from observations of rating histories. We then adapt the classical Baum-Welsh EM algorithm to our setting and show how to estimate the latent factor driving parameters. We demonstrate the efficiency of our approach on a credit rating database. Once calibrated, the filtering formula can be used in real-time to detect economic changes affecting the dynamics of rating migration.

nexialog CONSULTING

Contents

1 Introduction 3

2 Filtering of credit migration process 4

2.1 Filtering of counting process . . . 4

2.2 Extension to rating migration process . . . 5

2.2.1 Filtering multivariate point process . . . 5

2.2.2 Latent factor with finite number of states . . . 5

2.2.3 Application to rating migration processes . . . 6

2.3 Relation to Kalman filter . . . 7

3 Adjusting EM algorithm to our setting 8

3.1 Hypothesis . . . 8

3.2 The computations of the estimators . . . 9

3.3 Estimation of the parameters . . . 10

4 Implementation and results 11

4.1 Description of the data . . . 11

4.2 Algorithm's steps . . . 12

4.2.1 Initial conditions and Perturbations . . . 13

4.2.2 Discretization and implementation of the model . . . 14

4.3 The Experiment: stressed periods detection . . . 15

4.4 Summary . . . 18

Contacts 21

1 Introduction

Credit risk is the risk of loss associated with a borrower's default relative to the repayment of its debt. It may be related to the nominal or interest payments, to the uncertainty on the recovery rate in case of default but also to the deterioration of the credit quality. Credit risk research has been considerably intensified in the last 20 years. In particular, the challenges raised by the previous financial crisis prompted researchers to develop credit risk valuation models that capture evolution of the business cycle. The evolution of the banking supervisory regulation and accounting rules follows this trend (see official articles on IFRS9 as [1] and EBA guidelines on LDG downturn [10], PD and LGD [8], and stress testing [9]).
A credit rating evaluates the confidence in the capacity of the borrower to comply with credit's terms. Under the RWA capital rule, Basel regulations (see [2]), the amount of capital required for a credit depends on its rating. These ratings may be provided by rating agencies (like Moody's, Standard and Poor's, Fitch Ratings...) or may be the outcome of internal rating systems. An entity's rating may evolve through time according to its health and the economic cycle. Therefore predicting the evolution of rating migration is of primary importance for for every financial institution. The migrations of a group of credit entities can be described by migration matrices, defining the probabilities to move from one rating state to another in a given period of time. Transition matrices are required for the computation of important risk indicators such as Conditional Value-at-Risk (CVaR) or Incremental Risk Charge (IRC) on credit portfolios and can also be used for the pricing of some credit derivatives. Given recent evolution in banking supervisory and accounting rules, the challenge is to explain changes in transition probabilities due to change in the business cycle. For example, business-cycle dependent transition matrices - also known as *Point in Time* PIT matrices - may be used to achieve credit stress tests. In fact, the mechanism that binds transition probabilities to some macro-economic factors may serve to propagate stress test scenarios on credit portfolios.

Factor-based migration models provide a nice modeling framework for capturing migration sensitivities to macro-economic changes. In [7], different modelling and estimation approaches have been considered and estimation performance have been compared on real data. Factor migration models allow transition probabilities to depend on dynamic factors. Two types of models are considered: the "ordered Probit" (or structural approach) as described, e.g., in [13] and the "multi-state latent factor intensity model" (or intensity approach) studied, e.g., in [15]. In the latter factor intensity approach, the migration dynamics of each credit entity is described by a stochastic intensity matrix (or generator) whose components depend on a pool of common factors. The underlying factors aim at representing the evolution of the business cycle. The estimation procedure of such models differs given that the factors are considered to be observable or unobservable. In the first approach, explanatory observable factors (macroeconomic variables for example) have to be selected first. Then, sensitivities of transition intensities with respect to the selected factors can be estimated by a Cox regression, although the choice of explanatory variables may be arbitrary resulting in an important loss of information.

This article focuses on factor intensity models with unobservable driving factors. In this setting, the hidden factor may be inferred by maximizing the unconditional likelihood. The latter being very heavy to compute (see for example, [15]), an approach which consists of representing the dynamics of the transition probabilities in the form of a state-space linear Gaussian model has been developed in [7]. The non-observed factor is then filtered by the Kalman recursive algorithm. The model parameters are then estimated by maximizing the unconditional likelihood. However, the construction of the linear Gaussian model heavily relies on estimated quantities (transition intensities) which are reliable only when asymptotic normality applies. But this assumption does not hold in many applications where the sample size is too small. [12] use a point process filtering approaches for estimation of latent factor given default observation in a credit risk portfolio.

In this paper, we assume that the unobserved driving factor is given as finite state Markov chain. The method, presented here, consists in filtering the underlying driving factor through the observations of a multivariate point process counting migration events in the pool of entities. We explain how to adapt the classical univariate point process filtering formula ([5]) to observations given as sample path of multivariate (migration) counting processes. An EM algorithm is then used to estimate the hidden factor parameters in this setting. This methodology is illustrated on a data set of financial instruments ratings covering the period June 15, 2012 to December 14, 2016 and considering a two-state hidden factor. We observe that the dynamics of rating transition is switching among a steady migration regime and a volatile migration regime and that the steady state is clearly the more persistent one.

The paper is organised as follows. Section 2 presents the point process filtering method and application to credit migration processes.
Section 3 describes the Baum-Welsh algorithm adapted to the model in order to estimate the parameters at stake.
Finally the data, results of the crises detection framework are detailed along a fourth and last part.

2 Filtering of credit migration process

For a bond portfolio, the dynamics of rating migrations can be mathematically represented as a multidimensional counting process, each component representing the cumulative number of transitions from one rating category to another. Assuming that the counting process admits a latent intensity determined by the hidden factor's state, is it possible to estimate this intensity by only using observations of the point process ?

2.1 Filtering of counting process

We consider a probability space with a filtration: $(\Omega, F, F = (F_t)_{t \ [0,T]}, \mathrm{P})$. All the stochastic processes are supposed to be adapted to F and integrable on $[0, T]$. The problem is to estimate the states of an unobserved process Θ_t using only the information F_t^N resulting from the observation of a counting process N. By using the least squares criterion, this leads to determining

$$\hat{\Theta}_t = \mathrm{E}[\Theta_t | F_t^N].$$

With the same notations, all the processes O filtered by F_t^N are written :

$$\hat{O}_t = \mathrm{E}[O_t | F_t^N].$$

Unfortunately, the explicit computation of $\hat{\Theta}$ (by the Bayes formula) is only possible in very simple cases, for example, when N is a mixture of Poisson processes. Moreover, numerical approaches can be carried out but they are extremely heavy to implement as soon as non-trivial models are considered.

The filtering approach is different. It deals with $\hat{\Theta}$ as the solution of a forward EDS (differential stochastic equation) that allows, given $\hat{\Theta}_t$ to compute $\hat{\Theta}_{t+dt}$ per a simple recursive update. The main result on point process filtering can be stated in the following way (see [5], [14], [17], [23]). Let Θ be a process of the form

$$d\Theta_t = a_t \, \mathrm{d}t + \mathrm{d}M_t \,,$$

where M is a martingale with no jumps in common with N. Then, the $\hat{\Theta}$ process satisfies the following equation:

$$\mathrm{d}\hat{\Theta}_t = \hat{a}_t \, \mathrm{d}t + \eta_t \, (\mathrm{d}N_t - \hat{\nu}_t \, \mathrm{d}t) \,, \tag{1}$$

with

$$\eta_t = \frac{\widehat{(\Theta\nu)}_{t-}}{\hat{\nu}_{t-}} - \hat{\Theta}_{t-} \,, \tag{2}$$

and where ν is the F-intensity of N. The dynamic (1) suggests a recursive algorithm to update the $\hat{\Theta}$ process.

A consequence of (1) is that the dynamics of $\hat{\Theta}$ implies $(\widehat{\Theta\nu})$ whose dynamic implies the term $(\widehat{\Theta\nu^2})$, and so on ... Therefore this filtering formula induces an infinite nesting problem.

In order to study the dynamics of defaults in a credit portfolio or the dynamics of rating migrations, this filtering formula must be extended to a multivariate case.

2.2 Extension to rating migration process

In this section, we show that the previous filtering formula admits a natural multivariate extension and we solve the infinite nesting problem by considering a finite states latent factor. The details of this part are presented in article [6].

2.2.1 Filtering multivariate point process

In order to study the dynamics of defaults more generally the dynamics of rating migrations in a credit portfolio, the classical filtering formula also studied in [5] and [17], must first be generalized to a multivariate setting.

Let $N^j = (N_t^j)_{t \in [0,T]}$, $j = 1, \dots, p$, be a set of simple counting processes:
$N_t^j = \sum_{0<s\leq t} \Delta N_s^j < \infty$ and $\Delta N_s^j \in \{0,1\}$.
It is assumed that these processes admit a $\mathcal{F}$ - intensity $\nu^j = (\nu_t^j)_{t \in [0,T]}$, $j = 1, \dots, p$, and that they do not have any common jumps, i.e., $\Delta N_t^j \Delta N_t^k = \delta_{jk}$ (1 if $j = k$ and 0 otherwise).
In the following, we deal with the multivariate counting process $N = (N^1, \dots, N^p)$ and we write $\nu = (\nu^1, \dots, \nu^p)$ its intensity from filtration $\mathcal{F}$.

Proposition *Let Θ be a square integrable process of the form*

$$\Theta_t = \int_0^t a_s \,\mathrm{d}s + M_t\,, \tag{3}$$

where a is a $\mathcal{F}$-adapted process and M is a $\mathcal{F}$-square integrable martingale with no jumps in common with N. Therefore the process $\hat{\Theta}$ is solution of the EDS

$$\mathrm{d}\hat{\Theta}_t = \hat{a}_t \,\mathrm{d}t + \sum_j \eta_t^j \,(\mathrm{d}N_t^j - \hat{\nu}_t^j \,\mathrm{d}t)\,, \tag{4}$$

with

$$\eta_t^j = \frac{(\widehat{\Theta\nu^j})_{t-}}{\hat{\nu}_{t-}^j} - \hat{\Theta}_{t-}\,, \tag{5}$$

with initial conditions

$$\hat{\Theta}_0 = \mathbb{E}[\Theta_0]\,. \tag{6}$$

Proof: see [6]

2.2.2 Latent factor with finite number of states

To solve the issue of infinite imbrication mentioned above, we assume that the intensity $\nu = (\nu_j)_{j=1..p}$ is governed by a finite state Markov chain. Consequently the hidden factor driving process Θ is assumed to be a Markov chain with finite number of states in $\mathbb{T} = \{1, \dots, m\}$ and with constant transition intensities k^{hi}, $i \neq h$, such that $k^{hh} = -\sum_{i;i\neq h} k^{hi}$ and

$$\forall i \neq h \in [1...m] : \mathbb{P}(\Theta_{t+dt} = h \,|\, \Theta_t = i) = k^{ih}\,dt + o(dt). \tag{7}$$

The initial distribution of Θ is defined:

$$\forall i \in [1...m] : \Pi_i = \mathbb{P}(\Theta_0 = i). \tag{8}$$

Let us introduce the processes I^h, $h \in T$, defined by

$$I_t^h = 1_{[\Theta_t = h]}, \; h \in T \tag{9}$$

The process Θ can be represented as a finite sum of indicator processes

$$\Theta_t = \sum_{h \in T} h \, I_t^h. \tag{10}$$

2.2.3 Application to rating migration processes

The filtering theory with a counting process has been applied to credit risk for example in [11] and [17].
In order to remain realistic and fix the terminology, a bond market containing a finite number of individual bonds should be considered. All bonds are affected by variable and random market conditions represented by a hidden Θ Markov chain. A bond q of the sample is observed between the dates s^q and u^q, $0 \leq s^q \leq u^q \leq T$. At all times, the bond can only be in a state belonging to a finite set of states, $J = \{1, \dots, n\}$.
This space represents different credit risk scores in descending order, n being the default state. For example, Standard and Poor's long-term investment ratings can be translated to AAA = 1, AA = 2, A = 3, BBB =4, ..., D (Default) = 10.
Let $Z_t^q \in J$ be the state of the bond at time t and $Z^q = (Z_t^q)_{t \in [s^q, u^q]}$ the migration process that describes its evolution. The indicator process and the point process associated to Z^q are

$$J_t^{q,j} = 1_{[Z_t^q = j]}, \qquad N_t^{q,jk} = \#\{s \in [0,t]; Z_{s-}^q = j, Z_s^q = k\}. \tag{11}$$

$(Z^q)_q$ are assumed to be identically distributed. Then the processes $(Z^q)_q$ are governed as described before by their common intensity:

$$P[Z_{t+dt}^q = k \,|\, \Theta_t = h, Z_t^q = j] = \lambda^{h,jk} dt + o(dt). \tag{12}$$

We can illustrate the situation by the following figure. It shows the transitions between two states of Θ.

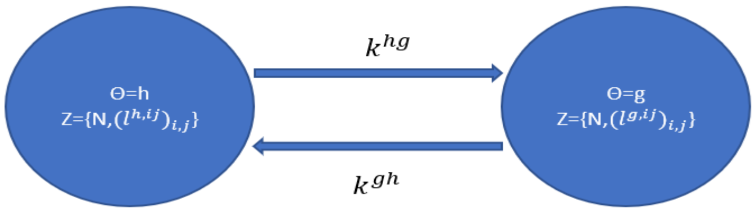

Figure 1 Θ's transitions

The processes of each individual Z^q are assumed to be conditionally independent, knowing F_T^Θ. However unconditionally, they are dependent through the effect of the factor Θ. Indeed the change of state of a bond will induce the change of state of other bond. The censorship mechanism governing (s^q, u^q) is assumed to be non-informative and can therefore be considered deterministic and belonging to F_0^N.

To infer the underlying process Θ, It is sufficient to observe the aggregated counting process N^{jk} defined by

$$N_t^{jk} = \sum_q N_t^{q,jk}, \tag{13}$$

governed by the F intensities

$$\nu_t^{jk} = Y_t^j \sum_h \lambda^{h,jk} I_{t-}^h, \tag{14}$$

with

$$Y_t^j = \sum_{q;\, s^q \le t < u^q} J_t^{q,j} . \tag{15}$$

The exposure process Y^j is left continuous. It increases by 1 when $k = j$: N^{kj} has jumped, or when a new bonds entered the pool with j. It decreases by 1 when a bound jumps outside equivalently $k = j$: $N^{jk} = 1$, or when a bond expires with the note j.
Recall that I_{t-}^l is the indicator function of the unobserved factor that is equal to 1 in state l.
As Y_j is F^N adapted and using the innovation theorem as in [6], the F^N intensity of N may be written:

$$\hat{\lambda}_t^{jk} = Y_t^j \sum_h l^{h,jk}\, \hat{I}_{t-}^h , \tag{16}$$

With such assumptions, we obtain the following form of the filter formula applied to the indicator of the unobserved process:

$$d\hat{I}_t^h = \sum_{i=1}^{m} k^{ih} \hat{I}_{t-}^i dt + \hat{I}_{t-}^h \sum_{j=k} \left(\frac{l^{h,jk}}{\sum_r l^{r,jk} \hat{I}_{t-}^r} - 1\right)\left(dN_t^{jk} - Y_t^j \sum_{i=1}^{m} l^{i,jk} \hat{I}_{t-}^i dt\right) \tag{17}$$

In this formula, some parameters are unknown: the initial distribution $(\Pi_i)_i$, the transition intensities of Θ, $(k^{ij})_{i,j}$ and the conditional transition intensities of the observable rating process $(l^{h,ij})_{h,i,j}$.
In order to make this formula usable, it is necessary to estimate these parameters. The major study of this note consists on seeking a way to make a such estimation.

2.3 Relation to Kalman filter

We try in this section to find similarities with the classical Kalman filter.

The set up of the Kalman filter can be defined in two phases. A first equation, of prediction, provides the filtered estimation of the hidden state knowing the observations of the previous times. A second one, that of correction which makes possible to correct the error made during the prediction phase, calculates the image of the hidden state knowing the observations of the previous and current time (the observation at the present time "corrects" the estimation of the previous step). The reader can refer to the studies made on the classical Kalman filter applied to probability laws in [19].
Thanks to the linearity of the equations and the Gaussian form of the handled processes, the Kalman filter compute recursively the like-hood of the model (conditioning a gaussian process by a gaussian is still normally distributed). Maximizing the like-hood could give the desired estimation of the parameters.
With the notations of the model, I_t the hidden process filtered and N_t the observation at time t, the Kalman filter would be as follows:

<u>Equation of prediction</u>

$$E[I_t|F_{t-}^N] = g(E[I_{t-}|F_{t-}^N])$$

<u>Equation of correction</u>

$$E[I_t|F_t^N] = f(E[I_t|F_{t-}^N], N_t)$$

Despite the no-linearity of the equations and forms processes without Gaussian assumptions, it is possible to see in the filter formula, (17), a similar decomposition:

$$E[I_t^h|F_{t-}^N] = \hat{I}_{t-}^h + \sum_i k^{ih} \hat{I}_{t-}^i dt \tag{18a}$$

$$E[I_t^h|F_t^N] = \underbrace{E[I_t^h|F_{t-}^N]}_{prediction} + \hat{I}_{t-}^h \sum_{j=k} \left(\frac{l^{h,jk}}{\sum_r l^{r,jk} \hat{I}_{t-}^r} - 1\right) \underbrace{\left(dN_t^{jk} - Y_t^j \sum_i l^{i,jk} \hat{I}_{t-}^i dt\right)}_{new\ Information} \tag{18b}$$

We can identify $Y_t^j \sum_i l^{i,jk}\hat{l}^i_{t-}dt$ as the information predicted from the previous step and $dN_t^{jk} - Y_t^j \sum_i l^{i,jk}\hat{l}^i_{t-}dt$ as the new information observed and brought by a jump at t. It would therefore be possible to estimate the parameters using extended Kalman filter. Derivations of such extensions can be found in [20]. However this solution turns out to be very heavy.

3 Adjusting EM algorithm to our setting

In this section, we propose an adaptation of the classical HMM Baum-Welsh algorithm in order to use the parameters estimations in the filtering formula (17).

If we consider the problem with a scale of discreet time, we must furnish new notations. For a space time $\Phi = (0, ..., t-1, t, t+1, ...T)$, we define the transition probabilities of Θ:

$$\forall (i,j) \in [1...m]^2, \quad \forall t \in \Phi, \ t \leq T-1 : K^{ij} = P(\Theta_{t+1} = j | \Theta_t = i).$$

We also define the conditional transition probabilities of $(Z^q)_q$:

$$\forall (i,j) \in [1...n]^2, \quad \forall h \in [1...m], \quad \forall t \in \Phi, \ t \leq T-1 : L^{h,ij} = P(Z^q_{t+1} = j | Z^q_t = i, \Theta_t = h).$$

We call $\forall t \in \Phi : Z_t = (Z_t^1, ..., Z_t^d)$ the total sample at time t.
We call $Z_{t_a|t_b} = (Z_{t_a}, ..., Z_{t_b})$ the sample of processes describing the evolution of the bonds ratings on Φ between the times t_a and t_b.
We write $Z \setminus Z_r = (Z_0...Z_{r-1}, Z_{r+1}...Z_T)$, the vector without the r^{th} component.

The processes $(Z^1, ..., Z^d)$ are not independent. Indeed, only one trajectory's realisation of Θ, the hidden factor, governs the sample's elements one.
The proposed method is a maximisation expectation (EM) algorithm for hidden Markov chains (HMM), adapted to the framework. We can find studies on the classical model in [22], [3], [24] and [21].
However the classical algorithm is not totally adapted to the framework linked to the filter formula (17) that we want apply. We highlight the following problematical classical model's characteristics:

- Baum-Welsh algorithm is an estimation in discrete time (the idea of duration is not considered).
- Rating status depends only on the state of the hidden factor. In the considered case, it also depends on the previous rating.
- Elements of the observations sample (the different ratings of the different bonds) are independent whereas they are dependent throw the same effect of Θ in our framework.

The first step of the classical algorithm assigns initial values to the parameters we want to estimate. Then the algorithm replaces the missing data (the states of Θ) with Bayesian estimators using the observations and the current parameters values.
The second one consists on improving a conditional likelihood that leads to obtain better estimates of the parameters. Then these new estimates will be used to repeat the first step along the next iteration.
At each step, the value of the parameters increases a like-hood and converges up to a local equilibrium which describes as well as possible the data.

3.1 Hypothesis

In hidden Markov chain models, an unobservable chain governs another observable one.
We recall the fundamental assumption of a Markov chain: the transition from one state to another

depends only on the current state (no memory):

$$t \geq 1 : P(Z_t|Z_0, ...Z_{t-1}) = P(Z_t|Z_{t-1}). \tag{19}$$

We recall the classical HMM's strong hypotheses mentioned before:

$$t \geq 1 : P(Z_t|Z_0, .., Z_{t-1}, \Theta_0, ..., \Theta_t) = P(Z_t|\Theta_t). \tag{20}$$

Formally, the future observation depends only on the state in which the unobserved process is. The hidden factor brings all the observable factor's information.
In the reality, there should be a response time between a jump of Θ and the effects on the rating process Z. A jump of Θ at time t can not directly affect the rating process Z at time t. That's why, we change the conventions here by supposing that rating process law is governed by the hidden factor Θ at the previous time.
We rather have:

$$t \geq 1 : P(Z_t|Z_0, .., Z_{t-1}, \Theta_0, ..., \Theta_{t-1}) = P(Z_t|\Theta_{t-1}). \tag{21}$$

Along the definition of a self financing portfolio, one can find the same kind of considerations. Indeed the process strategy is assumed to be F_t predictable (see [4].
It is also not acceptable to suppose that the new rating only depends on the economic cycle (state of Θ). It should also depends on the previous rating. Therefore we mix the two previous hypotheses by assuming that:

$$t \geq 1 : P(Z_t|Z_0, ..., Z_{t-1}, \Theta_0, ..., \Theta_{t-1}) = P(Z_t|Z_{t-1}, \Theta_{t-1}). \tag{22}$$

An other equality emerges from the previous assumptions:

$$t \geq 1 : P(\Theta_t|Z_0, ..., Z_t, \Theta_0, ..., \Theta_{t-1}) = P(\Theta_t|\Theta_{t-1}). \tag{23}$$

Moreover, to be consistent with the classical HMM algorithm, the observations must emerge from independent realizations of Θ.
Since the observations come from one and an unique realization of Θ (we cannot have different and independent realizations of the cycle economy), we keep the dependence through the sample along the computations.
To tackle the difference of time dimension, the discreet time parameters are estimated in order to use them in the filtering formula, (17), which describes a continuous phenomenon. One time step of subdivision and an other of extrapolation will be necessary. This manipulation seems to be quite reasonable but may not provide a completely trustful description of the model.

3.2 The computations of the estimators

This part only presents the main results of the algorithm. Evidences and estimators derivations are not furnished here. We define the probability forward:

$$\alpha_t(j) = P(Z_{0|t} = z_{0|t}, \Theta_{t-1} = j) \tag{24}$$

and the probability backward:

$$\beta_t(j) = P(Z_{t+1|T} = z_{t+1|T}|Z_t = z_t, \Theta_{t-1} = j). \tag{25}$$

We find the following recursive formula in order to compute the two probabilities:

$$\alpha_t(j) = \sum_{k=1}^{m} \alpha_{t-1}(k) K^{kj} \prod_{d} L^{j,z^d_{t-1} z^d_t}, \tag{26}$$

$$\beta_t(j) = \sum_{l=1}^{m} \beta_{t+1}(l) K^{jl} \prod_{d} L^{l,z^d_t z^d_{t+1}}. \tag{27}$$

We define the two random variables useful to describe Θ:

$$u_t(j) = 1_{(\Theta_t=j)} = I_t^j, \tag{28}$$

$$v_t(k,j) = 1_{(\Theta_t=j,\Theta_{t-1}=k)}. \tag{29}$$

The forward and backward probabilities help to compute the following Bayesian estimators:

$$\hat{u}_t(j) = P(\Theta_t = j|Z_{0|T} = z_{0|T}) = \frac{\beta_{t+1}(j)\alpha_{t+1}(j)}{L_T}, \tag{30}$$

with

$$L_T = P(Z_{0|T} = z_{0|T}) = \sum_j \alpha_T(j), \tag{31}$$

that is the likelihood of the whole sample.

and:

$$\hat{v}_t(k,j) = P(\Theta_t = j, \Theta_{t-1} = k|Z_{0|T} = z_{0|T}) = \frac{\beta_{t+1}(j)K^{kj}\alpha_t(k)\prod_d L^{j,z_t^d z_{t+1}^d}}{L_T}. \tag{32}$$

3.3 Estimation of the parameters

The maximization phase consists on finding better parameters than those of the previous iteration. We call $\tilde{M} = (\Pi^{(\rho+1)}, L^{(\rho+1)}, K^{(\rho+1)})$, the new parameters at iteration ρ that we want to estimate and $M = (\Pi^{(\rho)}, L^{(\rho)}, K^{(\rho)})$, the old ones from the iteration ρ.
We define:

$$w_i = P(Z, \Theta = \theta_i|M), \tag{33}$$

$$r_i = P(Z, \Theta = \theta_i|\tilde{M}). \tag{34}$$

By maximizing $\ln \frac{P(Z|\tilde{M})}{P(Z|M)}$, we obtain better estimations of the parameters than in M.
By this way, we try to have a higher likelihood with the new parameters:

$$P(Z|\tilde{M}) \geq P(Z|M). \tag{35}$$

To maximize this last is equivalent to maximize

$$Q(M,\tilde{M}) = \sum_{i=1}^{m} P(\Theta = i, Z|M)\ln(P(\Theta = i, Z|\tilde{M}). \tag{36}$$

We obtain then the following new estimations:

$$\Pi_i = P(\Theta_0 = i|Z, M) = \hat{u}_0(i), \tag{37}$$

$$L^{i,kr} = \frac{\sum_d \sum_t \hat{u}_t(i) 1_{(z_t^d = r, z_{t-1}^d = k)}}{\sum_d \sum_t \hat{u}_t(i) 1_{(z_{t-1}^d = k)}}, \tag{38}$$

$$K^{ji} = \frac{\sum_t \hat{v}_t(j,i)}{\sum_t \hat{u}_{t-1}(j)}. \tag{39}$$

4 Implementation and results

In this section, the different steps and choices for the algorithm's implementation are presented.

4.1 Description of the data

The data chosen in this report come from OpenDataSoft. We choose the base "Credit Rating Agency Ratings History Data" and more precisely the ratings given by Moody's Investor's services.

Aaa	Aa1	Aa2	Aa3	A1	A2	A3	Baa1	Baa2	Baa3	Ba1
0	1	2	3	4	5	6	7	8	9	10
Ba2	Ba3	B1	B2	B3	Caa1	Caa2	Caa3	Ca	C	
11	12	13	14	15	16	17	18	19	20	

Figure 2 Ratings used

It remains difficult to know whether it is better to have a homogeneous sample or not (to compare the ratings of really different instruments might make no sense).
Therefore we choose the data which concerns the financial instruments and we limit the sample at 30000 ratings reportings for a sake of rapidity. For instance, we can find euro medium term notes from JPMorgan Chase and Co., japan bonds from Mizuho Bank...
The main attributes of a data value are the following:

- Issuer name
- Rating
- Maturity Date
- Rating action date

The study covers the period June 15, 2012 to December 14, 2016 with a daily granularity on a portfolio of 1769 bonds at the initial moment. At the initial date, the ratings proportions are described as on 3 :

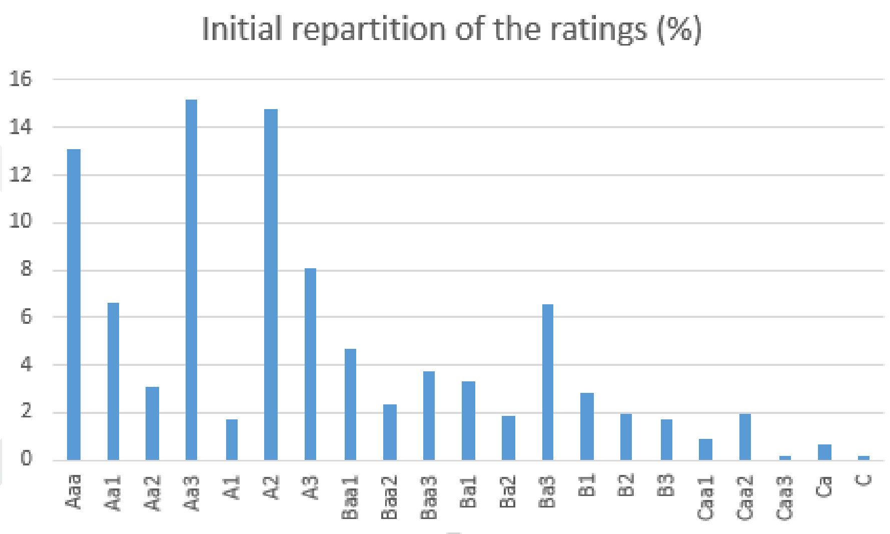

Figure 3 Ratings proportions

We count 2570 jumps during the considered period.

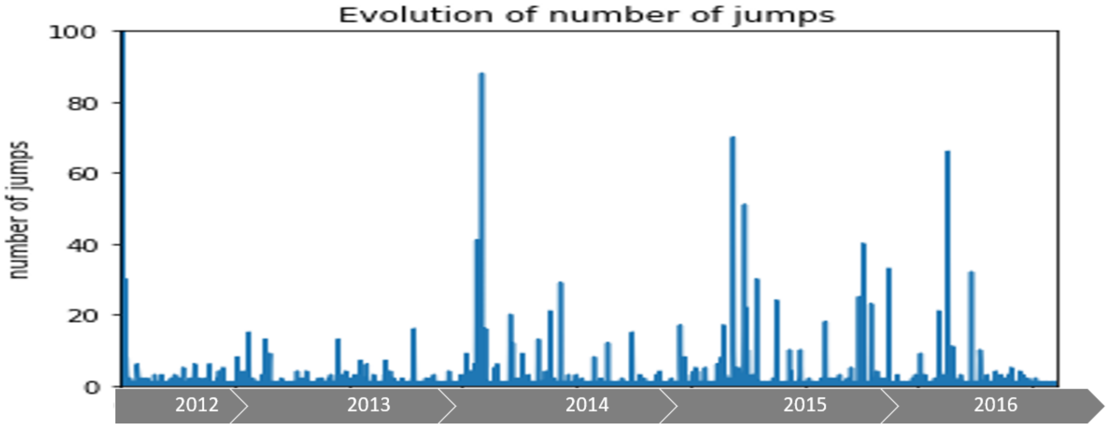

Figure 4 Evolution of number of jumps

4.2 Algorithm's steps

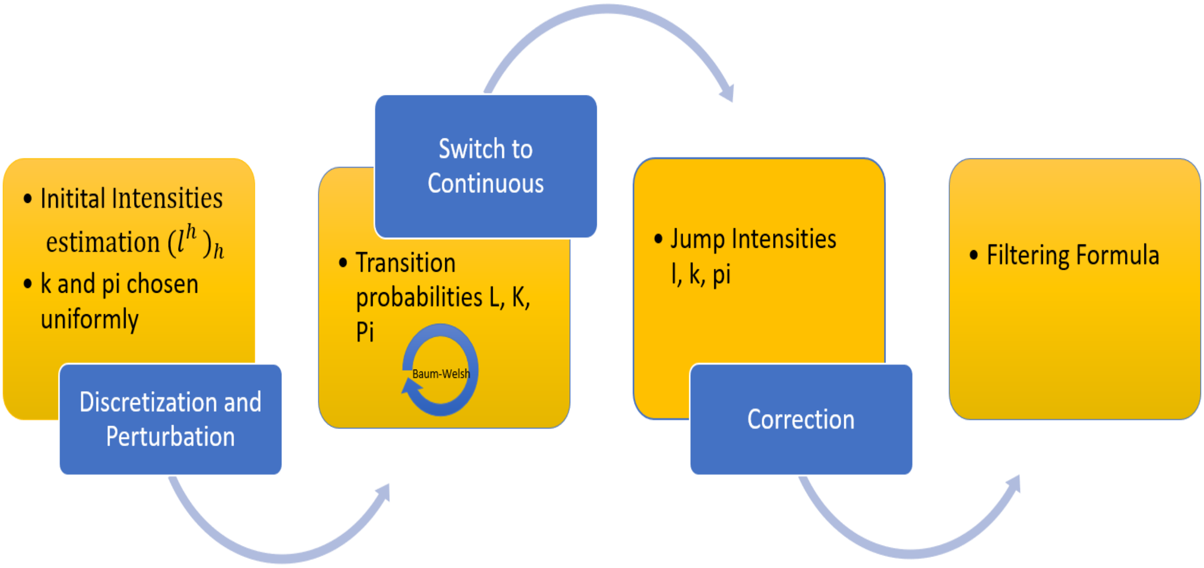

Figure 5 Steps

The framework is composed of three steps.
The first one consists on choosing initial values for the intensities $(k^{ij})_{i,j}$ and $(l^h)_h$ and the initial distribution Π.
Then the intensities are used to compute transition probabilities for a duration of 1 day.
To obtain transition probabilities, on a duration t, from an intensities matrix $\Lambda = (\lambda_{ij})_{i,j}$, the exponential matrix is used:

$$P(t) = \exp(\Lambda t) \tag{40}$$

By this way, we obtain the transition probabilities K and $(L^h)_h$.
If the initial values of $(l^h)_h$ have been chosen equal for all h $[0..m]$, that induces that Θ has no effect on the migrations. Therefore after the transformation in discreet time, we add a perturbation to the probabilities $(L^h)_h$ to differentiate the conditioning by Θ.
With these initial parameters values, the recursive Baum-Welsh algorithm sends us back his parameters estimations.

Along a last step, in order to compute $\hat{I}_t$, we estimate the final intensities values $(l^h)_h$ and $(k^{ij})_{i,j}$ from the Baum-Welsh's outputs to be used in the filtering formula.

4.2.1 Initial conditions and Perturbations

As described in the previous part, we have to furnish initial value of parameters to estimate along the Baum-Welsh step.
The first idea consists on running the Baum-Welsh algorithm with initial uniform distributions for Π and K, respectively, the initial distribution of the hidden factor and its transitions probabilities matrix.

θ 's states	
0	0,5
1	0,5

(a) initial value of Π

θ's states	0	1
0	0,5	0,5
1	0,5	0,5

(b) initial value of K

Figure 6 Initial Θ's parameters

It is more difficult to find consistent values for the initial values of $(L^h)_h$, the conditional transition probabilities matrices. Indeed, the choice of uniform probabilities does not fit with the form expected.
Furthermore it is impossible to simulate the conditioning by Θ since this last is unobservable. The first idea proposed is to consider that Θ has no effect on the transitions. By this way, we attribute the same value to transition intensities matrices for all the states of Θ. The conditioning by Θ is ignored and an estimators only based on the data, is used.
The classical intensity estimator of a Markov chain intensity used is:

$$i = j, \quad h : \hat{\lambda}_{ij} = \frac{N_{ij}(T)}{\int_0^T Y_i(s)ds}. \tag{41}$$

(see the studies in [16] and [7]).

With $N_{ij}(T)$ the number of jumps from i to j over the duration of the sampling and $Y_i(s)$ the number of individuals in state i at date s. Computing the denominator is similar than calculating the aggregate duration of all individuals spent in state i. Subsequently, we compute:

$$\hat{\lambda}_i = \sum_{i=j} \hat{\lambda}_{ij}.$$

By this way we consider from the beginning that Θ has no effect on the ratings.
Since we bring no information, without any perturbation, the Baum-Welsh algorithm is not able to differentiate the states of Θ and assimilate them to the events. Why would it promote some states compared to others if it has no profile, no way to differentiate them from the beginning ? Therefore it sends back exactly the same parameters that we have provided to it. $\hat{I}_\Theta$ remains constant at 0.5.

According to the previous result, it seems to be necessary to bring some information about the states of the hidden parameter to help the algorithm to differentiate them.
Then we can add different perturbations to each matrices of transition to differentiate the conditioning by the different states of Θ.

Along the framework, we choose a Θ with 2 states. With 2 states, it seems to be interesting to choose one state which tends to influence the bonds to jump to a better rating, the favourable state, and an other which tends to influence the bonds to jump to a worse rating, the unfavourable state.
These two states would describe the health of the global economy. This profiles choice may indicate

whether the economy in a favourable state which provides better confidence to credit or not.
For modeling that difference, we decide to add to the rating improvement transitions probabilities and to subtract it to the rating degradation transitions for the good state, the state 1. The opposite manipulation is done for 0, the unfavourable state for the credit.

	Aaa	Aa1	Aa2	...	Ca	C
Aaa		"+ε"	"+ε"	"+ε"	"+ε"	"+ε"
Aa1	"-ε"		"+ε"	"+ε"	"+ε"	"+ε"
Aa2	"-ε"	"-ε"		"+ε"	"+ε"	"+ε"
...	"-ε"	"-ε"	"-ε"		"+ε"	"+ε"
Ca	"-ε"	"-ε"	"-ε"	"-ε"		"+ε"
C	"-ε"	"-ε"	"-ε"	"-ε"	"-ε"	

Figure 7 Perturbation for the bad state

4.2.2 Discretization and implementation of the model

Two discretizations have been necessary for the implementation.
The first one concerns the Baum-Welsh algorithm. Indeed the Baum-Welsh algorithm's inputs are transition probabilities while the estimators $\hat{\lambda}$ at (41) represent intensities.
The time interval's length reaches almost 2 thousands days.
If the discretization is too fine, the algorithm may be very slow. If the interval of cutting in too large, we might miss too much information. In deed, the algorithm can only consider one jump of the same bond during a period. If several jumps occur during the same period, the algorithm will miss this piece of information.
The second risk of multiple jumps being much more significant, a length of 1 day has been chosen.

Then a parameters transformation is crucial for the transition between the end of the Baum-Welsh algorithm and the use of the filtering formula. The algorithm's outputs being transition probabilities, we have to compute an antecedent (which respects intensities format) of the exponential matrix function in order to retrieve intensities to insert in the filtering formula (17).

Finally the filtering formula has do be discretized to be implemented (17).
In deed, this equation describes a continues phenomenon thus we have to choose a cutting as fine as possible in discrete time. With an intuitive length of 1 day, the system diverges.
That might be explained by the fact that 1 days was still too large.
The time interval has to be very small to be consistent with the intensities.
In addition, the number of jumps ΔN, should not exceed 1 since it is assumed that there is not simultaneous jumps. Although more than one identical jump may be observed during the same day.
Furthermore data format does not allow to know the precise moment of a jump during a day.
To solve this issue, we decide to cut each days in 500 and then to attribute randomly the jumps between those daily parts.
Those moves from continuous to discrete and from discrete to continuous time are not perfectly suitable but remain acceptable.

The reader's attention should be drawn to many other numerical issues encountered in this model like the manipulation of tiny numbers. This issue is a direct consequence of the adaptation of the model. Details are furnished in the appendices.

4.3 The Experiment: stressed periods detection

With a perturbation of $= 10^{-8}$, we can observe the following conditional transition matrices after applying the EM algorithm.

	Aaa	Aa1	Aa2	Aa3	A1	A2	A3	Baa1	Baa2	Baa3	Ba1	Ba2	Ba3	B1	B2	B3	Caa1	Caa2	Caa3	Ca	C
Aaa	0,994695	0,0045764	0,0001319	6,63E-05	2,06E-24	0,0001326	6,63E-05	6,63E-05	9,12E-26	0,0001326	7,48E-44	6,63E-05	2,03E-26	1,56E-43	2,63E-36	7,27E-42	2,40E-47	6,63E-05	6,86E-87	0	0
Aa1	4,70E-38	0,995275	0,0001524	0,0019815	0,0019815	0,0004573	4,83E-34	1,72E-43	1,24E-33	2,73E-30	0	0,0001524	9,49E-43	9,00E-38	6,87E-40	0	0	0	0	0	0
Aa2	6,06E-31	0,0001847	0,996712	0,0011081	0,000186	0,0001847	2,86E-31	0,0001847	0,0007387	0,0001847	8,79E-27	1,68E-36	0,0003694	0,0001473	0	2,97E-35	1,62E-27	0	0	0	9,92E-36
Aa3	3,57E-34	2,83E-50	0,0001426	0,989946	0,0019967	0,0065597	0,0007131	0,0002139	0,0001426	7,13E-05	7,13E-05	4,72E-26	0,0001426	3,91E-38	4,45E-39	0	0	1,31E-38	0	0	0
A1	1,91E-26	4,84E-20	1,95E-08	2,10E-32	0,992004	0,0060061	0,0004323	1,86E-26	0,0010209	0,0002145	0,0002145	0	1,40E-29	2,05E-43	0,0001072	1,86E-16	2,95E-42	0	0	0	0
A2	4,63E-05	7,81E-37	4,61E-05	0,0005552	0,0024535	0,989773	0,0051372	0,0016194	0,0003239	4,58E-05	1,14E-34	2,84E-36	4,88E-42	6,64E-41	0	0	1,48E-27	2,63E-33	5,73E-48	0	0
A3	8,67E-37	5,99E-05	2,65E-31	0,0001797	0,0002996	0,000659	0,991871	0,0046729	0,0014794	0,000659	0,0001198	2,31E-23	0	2,13E-36	0	3,86E-42	0	2,36E-36	0	0	0
Baa1	2,35E-25	5,90E-26	2,15E-36	1,57E-39	9,56E-05	3,32E-34	0,0070759	0,987514	0,0029239	0,0008606	0,0011474	0,0003825	0	1,78E-41	9,06E-39	5,73E-24	0	0	1,76E-37	0	0
Baa2	0,0001064	8,03E-36	0,0001064	1,60E-35	6,61E-35	5,78E-36	0,000745	0,0035121	0,989874	0,0027853	0,0018092	0,000745	0,00021	0,0001064	2,96E-36	4,21E-46	0	0	0	9,20E-52	0
Baa3	7,43E-05	1,21E-30	1,09E-16	2,46E-36	0,0001244	1,79E-28	3,75E-39	0,0003733	0,000871	0,993582	0,0034815	0,0007466	0,0004977	8,32E-12	4,95E-31	0,0001244	1,37E-42	0	0	0	0,0001244
Ba1	0,0001767	4,25E-41	8,40E-48	7,04E-40	3,65E-26	3,77E-37	6,40E-45	1,46E-34	5,15E-39	0,0003534	0,993816	0,0015902	0,0031804	0,0001767	0	0	0	0	0	0,0007068	0
Ba2	1,79E-34	0	1,38E-43	0,000226	0,0002259	2,89E-33	4,62E-43	2,83E-30	5,55E-47	1,64E-48	0,0024793	0,994809	0,001582	4,73E-45	0,000226	0,000452	0	0	0	0	0
Ba3	1,10E-35	0	1,94E-17	1,14E-40	0,0001303	1,69E-44	0,0001291	1,28E-41	5,83E-53	1,30E-33	0	0,0019551	0,991919	0,0020939	0,0002529	0,0026068	0,0009124	4,77E-28	0	0	0
B1	1,01E-34	0	1,39E-31	1,26E-37	1,13E-23	4,26E-35	4,30E-33	1,70E-27	3,70E-13	2,33E-43	0,0002289	0,0018314	0,0025183	0,993795	0,000995	0,0006315	0	3,94E-31	2,72E-158	1,07E-25	1,64E-37
B2	1,25E-34	3,32E-32	1,40E-34	0	0	7,64E-35	3,33E-27	2,56E-37	8,18E-47	2,62E-37	0	0	1,85E-34	0,0007005	0,992737	0,0004532	0,0051755	1,76E-40	0	0,0002335	0,0007005
B3	0	0	1,21E-33	1,31E-49	0	1,82E-41	2,11E-42	0	2,94E-42	2,55E-34	3,97E-35	0,0002773	2,47E-38	0,0005546	5,37E-19	0,9942	0,0038598	0,0008317	0,0002771	1,57E-123	0
Caa1	1,02E-32	0	0	1,56E-22	0	0	0	0	9,78E-33	0	6,69E-48	0	0,0003026	9,90E-21	0	9,04E-85	0,998487	0,0006052	0,0003026	0,0003026	1,39E-57
Caa2	2,76E-33	0	0	1,57E-40	2,34E-32	1,13E-36	0	0	0	0	0,0004697	0	0	0,0004697	0	0	2,48E-51	0,997182	0,0018789	2,28E-20	5,62E-43
Caa3	0	4,34E-36	0	0	0	0	3,86E-46	0	7,40E-43	0	0	0	0	2,55E-32	0	0	1,42E-98	2,24E-64	0,998037	0,0019632	1,04E-204
Ca	0	0	5,23E-49	0	0	0	0	0	0	0	0,0006201	0	0	0,0006201	0	0	0	1,01E-52	6,97E-24	0,99814	0,0006201
C	0	0	0	4,57E-35	0	0	0	0	0	0	0	0	0	9,60E-34	0	0	1,87E-64	1,69E-85	8,47E-86	1,29E-43	1

Figure 8 1 day transition matrix L[Θ=0]

	Aaa	Aa1	Aa2	Aa3	A1	A2	A3	Baa1	Baa2	Baa3	Ba1	Ba2	Ba3	B1	B2	B3	Caa1	Caa2	Caa3	Ca	C
Aaa	0,999731	1,09E-05	5,47E-05	3,94E-05	1,31E-05	3,28E-05	1,97E-05	1,31E-05	2,40E-05	2,62E-05	8,75E-06	2,19E-06	4,37E-06	4,37E-06	8,75E-06	2,19E-06	2,19E-06	1,31E-09	2,19E-06	0	0
Aa1	2,85E-05	0,999777	5,23E-05	9,51E-06	1,90E-05	1,43E-05	3,33E-05	1,43E-05	9,51E-06	2,38E-05	0	1,13E-281	4,75E-06	4,75E-06	4,75E-06	4,75E-06	0	0	0	0	0
Aa2	9,36E-05	2,34E-05	0,999303	5,85E-05	8,77E-05	6,43E-05	8,77E-05	4,09E-05	5,85E-05	7,60E-05	5,85E-06	1,75E-05	2,34E-05	3,63E-05	0	5,85E-06	5,85E-06	0	0	0	1,17E-05
Aa3	3,27E-05	4,67E-06	4,44E-05	0,999785	9,34E-06	1,17E-05	2,10E-05	9,34E-06	2,57E-05	2,57E-05	9,34E-06	4,67E-06	4,67E-06	2,34E-06	4,67E-06	0	0	4,67E-06	0	0	0
A1	3,96E-05	1,32E-05	1,98E-05	2,64E-05	0,999724	2,64E-05	2,63E-05	3,30E-05	2,80E-05	2,64E-05	9,91E-06	0	1,32E-05	3,30E-06	8,86E-12	6,61E-06	3,30E-06	0	0	0	0
A2	1,36E-05	7,55E-06	1,81E-05	1,36E-05	2,11E-05	0,999843	1,66E-05	6,04E-06	1,81E-05	1,51E-05	1,06E-05	1,51E-06	1,51E-06	6,04E-06	1,51E-06	0	3,02E-06	1,51E-06	1,51E-06	0	0
A3	1,72E-05	7,66E-06	2,30E-05	1,34E-05	1,72E-05	4,21E-05	0,999766	2,68E-05	3,51E-05	2,11E-05	3,83E-06	1,92E-06	1,53E-05	5,75E-06	0	1,92E-06	0	1,92E-06	0	0	0
Baa1	2,47E-05	1,54E-05	3,08E-05	1,85E-05	1,54E-05	2,47E-05	1,54E-05	0,999752	4,45E-05	3,08E-05	8,28E-12	6,17E-06	0	6,17E-06	6,17E-06	6,17E-06	0	0	3,08E-06	0	0
Baa2	4,44E-05	1,71E-05	5,47E-05	4,10E-05	2,05E-05	3,42E-05	2,39E-05	2,73E-05	0,999611	6,77E-05	6,83E-06	3,42E-06	6,92E-06	6,83E-06	1,37E-05	6,83E-06	6,83E-06	0	0	3,42E-06	3,42E-06
Baa3	6,95E-05	1,60E-05	4,79E-05	2,80E-05	3,99E-05	3,19E-05	1,60E-05	5,19E-05	6,79E-05	0,999519	3,60E-05	1,20E-05	3,99E-06	2,40E-05	2,00E-05	1,29E-59	7,99E-06	3,99E-06	0	3,99E-06	0
Ba1	1,11E-05	5,57E-06	5,57E-06	2,23E-05	1,11E-05	5,57E-06	1,11E-05	1,67E-05	1,67E-05	7,80E-05	0,999733	5,01E-05	1,11E-05	5,57E-06	0	0	1,11E-05	5,57E-06	0	0	0
Ba2	1,44E-05	0	7,20E-06	1,53E-13	3,90E-09	1,44E-05	7,20E-06	7,20E-06	7,20E-06	1,44E-05	2,18E-05	0,999841	3,60E-05	1,44E-05	7,20E-06	7,20E-06	0	0	0	0	0
Ba3	8,63E-06	0	8,63E-06	1,29E-05	3,00E-17	1,73E-05	1,30E-05	4,32E-06	4,32E-06	8,63E-06	0	1,29E-05	0,999745	5,58E-05	7,79E-05	4,32E-06	1,73E-05	8,63E-06	0	0	0
B1	1,46E-05	0	3,66E-05	1,46E-05	2,93E-05	1,46E-05	4,39E-05	1,46E-05	5,12E-05	1,46E-05	1,46E-05	1,46E-05	3,66E-05	0,999532	7,07E-05	5,30E-05	0	1,46E-05	1,46E-05	7,32E-06	7,32E-06
B2	7,20E-06	7,20E-06	1,44E-05	0	0	2,16E-05	2,88E-05	2,16E-05	7,20E-06	1,44E-05	0	0	2,88E-05	6,48E-05	0,999669	6,53E-05	2,76E-05	7,20E-06	0	7,20E-06	7,20E-06
B3	0	0	8,85E-06	8,85E-06	0	1,77E-05	8,85E-06	0	1,77E-05	3,54E-05	8,85E-06	1,35E-156	8,85E-06	8,85E-06	1,68E-04	0,999601	8,04E-05	5,31E-09	5,31E-09	1,77E-05	8,85E-06
Caa1	2,86E-05	0	0	1,91E-05	0	0	0	0	9,53E-06	0	9,53E-06	0	0	2,86E-05	0	3,81E-05	0,999771	1,91E-05	3,68E-32	6,67E-05	9,53E-06
Caa2	3,09E-05	0	0	1,54E-05	1,54E-05	3,09E-05	0	0	0	0	3,02E-202	0	0	1,92E-151	0	0	6,18E-05	0,999645	4,63E-05	7,72E-05	7,72E-05
Caa3	0	6,14E-05	0	0	0	0	6,14E-05	0	6,14E-05	0	0	0	0	6,14E-05	0	0	6,14E-05	6,14E-05	0,999202	0,0001842	2,46E-04
Ca	0	0	1,99E-05	0	0	0	0	0	0	0	3,89E-202	0	0	2,47E-151	0	0	0	1,99E-05	5,97E-05	0,999622	0,0002787
C	0	0	0	5,02E-05	0	0	0	0	0	0	0	0	0	2,51E-05	0	0	5,02E-05	5,02E-05	2,51E-05	2,51E-05	0,999774

Figure 9 1 day transition matrix L[Θ=1]

We can no longer observe a good or a bad state for the credit anymore. Although we can notice that the state 1 of Θ is more stable than 0 since the values on its diagonal are mainly higher than in 0.
Moreover its probabilities of jumps (which improve or deteriorate the ratings) are mainly lower than in 0.
The algorithm divides the information between the two states. 0 is the volatile state where a

bond tends to have more chance to change its rating. 0 represents a state of stress for the credit whereas 1 a state of stability.
It is interesting to remark that the probabilities to fall in the worst rate (C which is almost the state of default) are not the same in the two states. We compute 2.27084414030e-05 for the state 0 and 4.86572724696e-06 for the state 1.
Indeed, even if a bond seems to have more chance to improve its rating in 0 it also has more chance to fall in default. 0 remains the worst and the riskiest state.

To deeper understand the behaviours of the two states, it is more relevant to observe the transitions on 1 year: $L_{1year}[\Theta = 0] = L[\Theta = 0]^{365}$ and $L_{1year}[\Theta = 1] = L[\Theta = 1]^{365}$ (the time transition for L was 1 day). 0 is clearly the volatile state (poor diagonal) and not desirable for the market credit: ratings deterioration would be more often.

	Aaa	Aa1	Aa2	Aa3	A1	A2	A3	Baa1	Baa2	Baa3	Ba1	Ba2	Ba3	B1	B2	B3	Caa1	Caa2	Caa3	Ca	C
Aaa	0,144926	0,270009	0,022173	0,067277	0,110598	0,110464	0,083571	0,042953	0,031361	0,025014	0,027232	0,023	0,013277	0,004059	0,002147	0,004441	0,004397	0,006244	0,002823	0,00306	0,000976
Aa1	0,00245	0,179946	0,017672	0,069221	0,143338	0,157924	0,147109	0,076617	0,054567	0,036219	0,041305	0,027933	0,018325	0,005149	0,003263	0,005974	0,006843	0,000826	0,00047	0,00362	0,001231
Aa2	0,002759	0,017716	0,304628	0,054167	0,051297	0,075475	0,089069	0,059739	0,063619	0,049128	0,05383	0,041272	0,049349	0,025368	0,005287	0,01992	0,023266	0,003456	0,002087	0,006237	0,002331
Aa3	0,004274	0,004409	0,009702	0,041343	0,099811	0,147979	0,230424	0,128672	0,085324	0,064972	0,071585	0,03683	0,030243	0,008871	0,003471	0,009943	0,010696	0,001451	0,000841	0,00691	0,002248
A1	0,004856	0,004398	0,004058	0,014688	0,120727	0,13857	0,213561	0,121542	0,08812	0,070467	0,082598	0,04292	0,034694	0,010204	0,005429	0,011246	0,01558	0,001879	0,001131	0,009565	0,003769
A2	0,005378	0,005949	0,005052	0,015776	0,073305	0,102338	0,24277	0,14354	0,097206	0,080831	0,092358	0,048116	0,037542	0,010183	0,003459	0,011502	0,01036	0,001345	0,000723	0,009282	0,002986
A3	0,00561	0,006568	0,003224	0,010065	0,032402	0,043053	0,227037	0,144419	0,102124	0,09902	0,12324	0,069603	0,055712	0,01594	0,00404	0,018665	0,016286	0,00228	0,001251	0,014714	0,004746
Baa1	0,006427	0,005925	0,003097	0,008066	0,024349	0,030729	0,190438	0,12826	0,094424	0,097195	0,139943	0,087988	0,071207	0,022163	0,005567	0,026386	0,025304	0,003699	0,002102	0,020765	0,005968
Baa2	0,009465	0,008027	0,005967	0,006828	0,0171	0,018333	0,108525	0,083081	0,080523	0,104157	0,16778	0,11782	0,098223	0,034031	0,008502	0,040425	0,042497	0,006498	0,003803	0,029195	0,009221
Baa3	0,010206	0,007359	0,001378	0,004883	0,014186	0,011413	0,033515	0,028263	0,032374	0,126985	0,199938	0,140263	0,128095	0,044728	0,010937	0,057754	0,065244	0,010605	0,006468	0,041447	0,023958
Ba1	0,009732	0,008393	0,000675	0,005361	0,011452	0,008637	0,00977	0,005217	0,004762	0,018241	0,17841	0,163863	0,159216	0,073135	0,017764	0,084473	0,113504	0,018481	0,011612	0,08089	0,016413
Ba2	0,00468	0,00274	0,000647	0,009011	0,01972	0,018225	0,018371	0,009274	0,007006	0,012037	0,150209	0,24829	0,141928	0,054917	0,020583	0,084799	0,120066	0,020384	0,013242	0,035939	0,007934
Ba3	0,001303	0,000647	0,000214	0,003645	0,010247	0,00826	0,011904	0,005954	0,004293	0,005056	0,058526	0,134244	0,132517	0,093878	0,02508	0,122243	0,259951	0,04931	0,035433	0,029558	0,007737
B1	0,001574	0,000749	0,000167	0,003921	0,008686	0,006351	0,006874	0,003069	0,002318	0,004043	0,069681	0,160985	0,13837	0,162956	0,046456	0,101087	0,19361	0,030137	0,020197	0,023956	0,014811
B2	0,00023	8,41E-05	1,34E-05	0,00042	0,001037	0,000595	0,000776	0,000288	0,000218	0,000527	0,014143	0,023231	0,035109	0,03979	0,076398	0,030536	0,495799	0,064288	0,051672	0,066253	0,098595
B3	0,000413	0,000165	2,97E-05	0,000798	0,001675	0,001121	0,001152	0,000474	0,000371	0,000929	0,022604	0,037621	0,037751	0,042129	0,00691	0,136155	0,429914	0,116704	0,097129	0,059729	0,006229
Caa1	0,000257	9,71E-05	1,19E-05	0,000275	0,000884	0,000499	0,00082	0,00032	0,000237	0,000595	0,015493	0,015615	0,034392	0,021218	0,00305	0,013317	0,59124	0,103645	0,095547	0,091257	0,011231
Caa2	0,001271	0,000647	4,59E-05	0,000647	0,001346	0,00081	0,000916	0,000419	0,000401	0,002413	0,051763	0,032547	0,035341	0,051443	0,007233	0,0155	0,018014	0,359379	0,289523	0,118395	0,011947
Caa3	0,000497	0,000187	1,13E-05	0,000186	0,000379	0,000192	0,000231	9,81E-05	0,000103	0,000942	0,028736	0,011999	0,01424	0,028717	0,002826	0,005068	0,004666	0,000553	0,488374	0,359949	0,052046
Ca	0,001785	0,000904	6,39E-05	0,000904	0,00188	0,001127	0,001275	0,000583	0,000558	0,003388	0,072432	0,045679	0,04967	0,071998	0,010162	0,021714	0,025102	0,003361	0,001964	0,518642	0,166808
C	2,60E-34	1,09E-34	1,63E-34	5,39E-33	3,40E-33	4,45E-33	3,65E-33	1,67E-33	1,07E-33	1,05E-33	1,52E-32	4,72E-32	4,69E-32	1,51E-31	1,76E-32	2,57E-32	3,11E-32	3,99E-33	2,34E-33	2,98E-33	1

Figure 10 1 year transition matrix $L_{1year}[\Theta = 0]$

	Aaa	Aa1	Aa2	Aa3	A1	A2	A3	Baa1	Baa2	Baa3	Ba1	Ba2	Ba3	B1	B2	B3	Caa1	Caa2	Caa3	Ca	C
Aaa	0,907191	0,003851	0,017177	0,013566	0,004839	0,011558	0,007089	0,004749	0,008331	0,008936	0,003051	0,000878	0,00163	0,00166	0,003005	0,000809	0,000826	4,79E-05	0,000669	5,27E-05	8,43E-05
Aa1	0,010083	0,921871	0,016511	0,003624	0,006785	0,005359	0,011622	0,005139	0,003623	0,008102	0,000134	0,000102	0,001756	0,001752	0,001726	0,00163	8,30E-05	2,40E-05	1,23E-05	1,87E-05	4,64E-05
Aa2	0,029783	0,007619	0,776556	0,019019	0,027505	0,02115	0,027823	0,013432	0,018692	0,02337	0,002226	0,005686	0,007557	0,011094	0,000544	0,001989	0,002001	0,000135	6,97E-05	9,38E-05	0,003657
Aa3	0,011462	0,00175	0,014103	0,924988	0,003539	0,00446	0,007481	0,003481	0,008837	0,008753	0,003253	0,001712	0,001716	0,000964	0,001659	7,30E-05	8,65E-05	0,001558	2,21E-05	3,79E-05	6,35E-05
A1	0,013606	0,004583	0,006612	0,009232	0,904666	0,009362	0,00913	0,011272	0,009555	0,008922	0,003423	0,000124	0,004495	0,001254	0,000248	0,002201	0,001197	3,77E-05	1,68E-05	3,57E-05	2,95E-05
A2	0,004954	0,002686	0,005958	0,004895	0,007346	0,944617	0,005894	0,00229	0,006291	0,005252	0,003662	0,000606	0,00062	0,002085	0,000591	6,53E-05	0,00108	0,000532	0,000477	4,71E-05	5,27E-05
A3	0,00624	0,002741	0,007517	0,004887	0,006079	0,014688	0,918417	0,009246	0,011828	0,007214	0,001411	0,00075	0,005222	0,002036	0,000204	0,000693	8,46E-05	0,000657	2,12E-05	2,75E-05	3,82E-05
Baa1	0,00878	0,005356	0,009988	0,006625	0,005521	0,008819	0,005599	0,913792	0,014904	0,010371	0,000169	0,002177	0,00016	0,002182	0,002222	0,002096	0,00011	3,74E-05	0,000944	6,22E-05	8,60E-05
Baa2	0,015256	0,00584	0,017052	0,013989	0,007262	0,011955	0,008437	0,009431	0,868475	0,021721	0,002484	0,001303	0,002468	0,002501	0,004646	0,002308	0,002378	7,52E-05	3,69E-05	0,001152	0,00123
Baa3	0,02297	0,005445	0,014954	0,00975	0,0133	0,011156	0,005891	0,017122	0,021872	0,84006	0,011668	0,004134	0,001588	0,007684	0,006546	0,000268	0,002712	0,001324	6,44E-05	0,001332	0,000159
Ba1	0,004304	0,00201	0,002194	0,007764	0,004004	0,002289	0,00399	0,005941	0,005947	0,02513	0,907309	0,017031	0,003889	0,002033	0,00023	9,59E-05	0,003802	0,001867	2,56E-05	9,63E-05	4,79E-05
Ba2	0,005075	7,19E-05	0,002473	0,000211	0,000173	0,00518	0,002655	0,002606	0,002629	0,004965	0,007448	0,943849	0,012292	0,004912	0,00275	0,002475	0,000121	4,90E-05	1,72E-05	2,45E-05	2,27E-05
Ba3	0,003171	9,09E-05	0,002992	0,004535	0,000238	0,006169	0,004745	0,001698	0,001788	0,003031	0,000132	0,004462	0,911566	0,018265	0,025896	0,001931	0,005976	0,002933	7,40E-05	0,000179	0,000128
B1	0,005357	0,000284	0,011359	0,005265	0,009748	0,005506	0,01475	0,005192	0,0166	0,00531	0,004846	0,004924	0,012058	0,843692	0,023086	0,016873	0,000598	0,004754	0,004329	0,002621	0,00285
B2	0,002755	0,002473	0,004772	0,000365	0,000335	0,007629	0,009894	0,007333	0,002864	0,004962	0,000174	0,000124	0,009696	0,02075	0,887394	0,021161	0,009529	0,002467	0,000115	0,002573	0,002634
B3	0,000522	0,000158	0,003049	0,003206	0,000219	0,006263	0,003352	0,000426	0,005992	0,011372	0,003025	8,92E-05	0,003229	0,003661	0,053937	0,865258	0,026543	0,000245	8,81E-05	0,006037	0,003328
Caa1	0,009673	4,34E-05	0,000333	0,006629	0,000126	0,000206	0,000185	9,13E-05	0,003325	0,000267	0,003256	6,80E-05	0,00011	0,009261	0,000531	0,012514	0,920219	0,006408	0,000319	0,021974	0,00446
Caa2	0,010279	0,000217	0,000284	0,005534	0,005109	0,010384	0,000284	8,16E-05	0,000331	0,000133	9,25E-05	1,41E-05	3,20E-05	0,00041	3,15E-05	0,000155	0,020711	0,878917	0,014114	0,025524	0,027363
Caa3	0,0006	0,01874	0,000764	0,00114	0,000369	0,000501	0,018979	0,000303	0,01846	0,000451	0,000128	7,48E-05	0,000228	0,018336	0,000318	0,000366	0,019605	0,019235	0,748349	0,055241	0,077811
Ca	0,000159	0,000238	0,005994	0,000942	0,000125	0,000122	0,000317	5,24E-05	0,000282	9,27E-05	1,15E-05	2,23E-05	3,13E-05	0,00066	7,84E-06	1,66E-05	0,001144	0,007404	0,018073	0,872178	0,092127
C	0,000312	0,000107	0,000219	0,017057	0,000125	0,000161	0,000228	5,62E-05	0,000277	0,000107	8,22E-05	3,85E-05	7,24E-05	0,008268	0,000126	0,000196	0,017161	0,016711	0,007865	0,008926	0,921904

Figure 11 1 year transition matrix $L_{1year}[\Theta = 1]$

We compute probabilities to fall in C: 0.008309693346555597 for 0 and 0.0017274814489362618 for 1 at 1 year. A bond is almost 8 times more likely to fall in the worst rate in 0 than in 1.

The following figures show the outputs of the Θ 's parameters.

θ 's states	
0	0,393582
1	0,606418

(a) Π

θ's states	0	1
0	0,067941	0,93206
1	0,030798	0,969202

(b) K

Figure 12 Θ's parameters

By analyzing these values, we can predict that, Θ would be used to stays in 1, the normal and favourable state for the credit with fast and regular jumps to 0, the riskier one.
Consequently the global economy would be used to be in a steady state but would know sometimes short moments of stress for the credit.
the following plot shows the outputs of the filtering formula, $\hat{I}_\Theta$, the hidden parameter filtered.

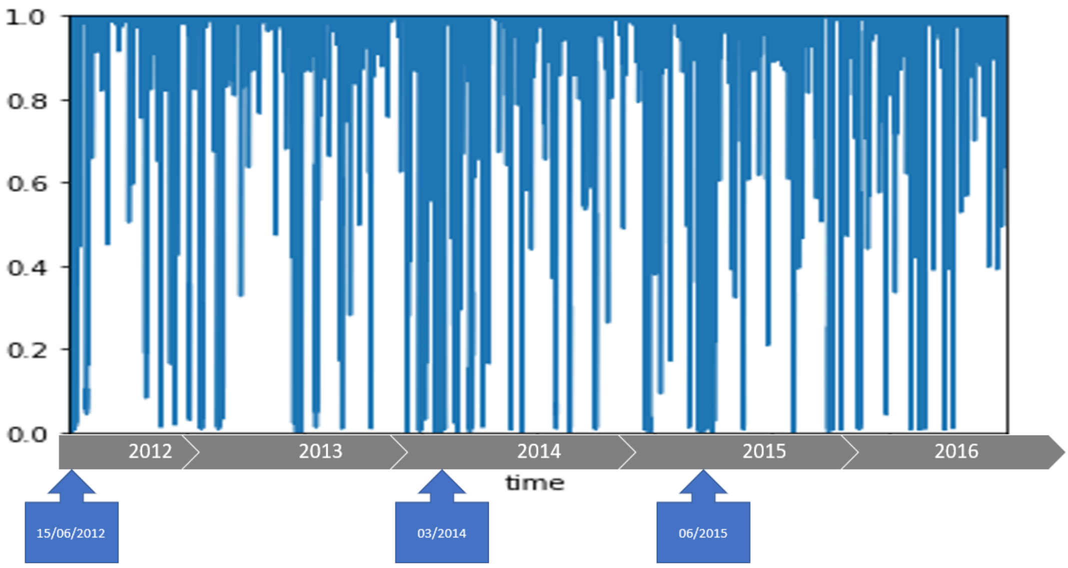

Figure 13 Out of $\hat{I}_\Theta$

The behaviour of $\hat{I}_\Theta$ confirms the previous predictions: the hidden factor filtered remains the major part of the time in the state 1, the steady state but is used to jump quickly to the volatile state 0.
We may notice many periods where Θ jumps a lot to the systemic state 0. According to the model, these moments may be interpreted as periods of crisis for the credit.
The most notable, during march 2014, is not specially linked to a famous economic event. Indeed, many other underlying economical or political reasons might explain a stressed period. It would be interesting to study and compare the evolution of highly correlated factors like borrow rates during those periods.
Nevertheless the model may able to detect real crisis periods. The unfavourable period for the credit identified by the model at the beginning of the summer 2015, corresponds to a stock market crashed in China (a third of the value of Shanghai Stock Exchange was lost in a month).

The following plots show the features of two filtered intensities of jumps according to (16).

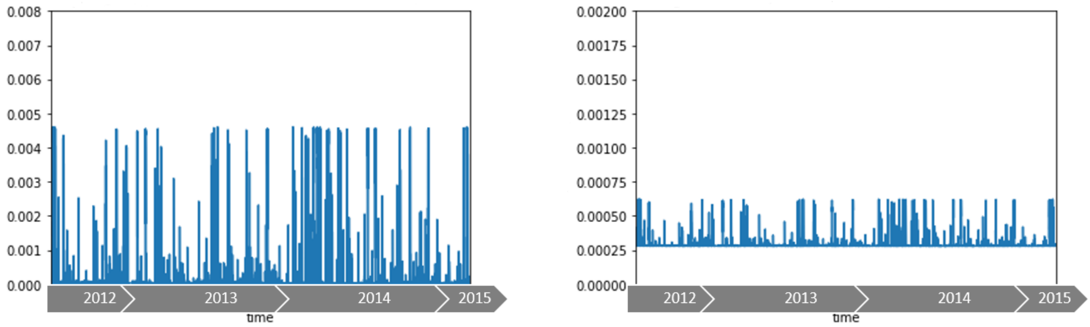

Figure 14 Outs of $\hat{\nu}^{AAA,Aa1}$ and $\hat{\nu}^{Ca,C}$

4.4 Summary

The filtered hidden factor values carry lot of information about the credit economy cycle. This last is used to remain in a stable and favourable state for the credit with regularly short jumps to a worse and riskier state for the credit.
A raise of the frequency of these jumps may be witness of a crisis.
This study may be useful to detect systemic periods in the credit market. These periods may relate scenarios useful to simulate historical stress tests.

These periods could be used in downturn period detection according to the Basel modelling parameters as well.
The filtered value of I_Θ could also identify the current states of the economy and detect cycle's disruptions.
The idea is to know in real time the state we are. According to the current observations, the model furnishes the filtered value of the hidden state. The filtered value has the advantage to be a continuous process.
Let's Imagine that we could observe Θ with 2 states, one stable and one of crisis. We could not prevent a period of stress since the break is brutal (in one day Θ jumps). We would have a discontinuity in our PD estimations since we change suddenly from one state to another.
Using the filtered outputs furnishes a continuous PD and rating transitions probabilities estimations according to the economic cycle.

On the other hand, the model could be improved by different ways:

- The type of data may be changed since financial instruments are stable elements and are not a lot sensible to the evolution of the economy's cycle.
- We can extend the Baum-Welsh algorithm to continuous time developed for instance in biology according to the studies in [18].
- The most promising improvement would consists on developing a discrete version of the filter to avoid the steps of discretization which lead to approximations. Furthermore the discrete version would be consistent with the daily data format.
- The initial conditions and perturbations may be the purpose of further studies.

References

[1] Amending regulation (ec) no 1126/2008 adopting certain international accounting standards in accordance with regulation (ec) no 1606/2002 of the european parliament and of the council as regards international financial reporting standard 9. *O cial Journal of the European Union*, 2016.

[2] Basel committee on banking supervision basel iii: Finalising post-crisis reforms. *O cial Journal of the European Union*, 2017.

[3] C. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, 2006.

[4] Tomas Björk. *Arbitrage theory in continuous time*. Oxford university press, 2009.

[5] P. Brémaud. *Point Processes and Queues: Martingale Dynamics*. New-York: Springer-Verlag, 1981.

[6] A. Cousin, N. Ragnar, and T. Picard. Working paper, filtering with multivariate counting processes and an application to credit risk. 2020.

[7] A. Cousin and M. Reda Kheliouen. A comparative study on the estimation of factor migration models, 2015.

[8] European Banking Authorities EBA. Guidelines on pd estimation, lgd estimation and the treatment of defaulted exposures. 2017.

[9] European Banking Authorities EBA. Guidelines on institutions' stress testing. 2018.

[10] European Banking Authorities EBA. Guidelines for the estimation of lgd appropriate for an economic downturn ('downturn lgd estimation'). 2019.

[11] Rüdiger Frey and Wolfgang Runggaldier. Pricing credit derivatives under incomplete information: a nonlinear-filtering approach. *Finance and Stochastics*, 14(4):495–526, 2010.

[12] Rüdiger Frey and Thorsten Schmidt. Pricing and hedging of credit derivatives via the innovations approach to nonlinear filtering. *Finance and Stochastics*, 16(1):105–133, 2012.

[13] P. Gagliardini and C. Gourieroux. Stochastic migration models with application to corporate risk. *Journal of Financial Econometrics*, 3(2):188–226, 2005.

[14] A. Karr. *Point Processes and their Statistical Inference*. Marcel Dekker Inc., New York, Basel, Hong Kong, 1991.

[15] S. J. Koopman, A. Lucas, and A. Monteiro. The multi-state latent factor intensity model for credit rating transitions. *Journal of Econometrics*, 142(1):399–424, 07 2007.

[16] D. Lando and M. Skodeberg. Statistical analysis of rating transisitions-a continuous- time approach. *Journal of Banking Finance*, 2002.

[17] V. Leijdekker and P. Spreij. Explicit computations for a filtering problem with point process observations with applications to credit risk, 2011.

[18] Li S. Li F. Le Song Liu Y., Moreno A. and M. Rehg J. *Learning Continuous-Time Hidden Markov Models for Event Data, Mobile Health*. Springer International Publishing, 2017.

[19] T. Proietti and A. Luati. Maximum likelihood estimation of time series models: the kalman filter and beyond, 2012.

[20] M.I. Ribeiro. Kalman and extended kalman filters: Concept, derivation and properties. 2004.

[21] L. R.Rabinet. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 1989.

[22] A. Tenyakov. Estimation of hidden markov models and their applications in finance. *Electronic Thesis and Dissertation Repository. 2348*, 2014.

[23] J.H. van Schuppen. Filtering, prediction, and smoothing for counting process observations – a martingale approach. *SIAM J. Appl. Math.*, 1997.

[24] Lindsten F. Fritsche C. Özkan, E. and Fredrik Gustafsson. Recursive maximum likelihood identification of jump markov nonlinear systems. *IEEE Transactions on Signal Processing*, 2013.

Nexialog Consulting est un cabinet de conseil spécialisé en Banque et en Assurance. Organisés autour de 3 domaines d'activité - Risques Bancaires, Financiers & Assurantiels - nous intervenons au sein des équipes métiers afin de les accompagner depuis le cadrage jusqu'à la mise en œuvre de leurs projets. Associant innovation et expertise, le savoir-faire de notre cabinet a permis de consolider notre positionnement sur ce segment et de bénéficier d'une croissance forte et régulière.

Les besoins de nos clients étant en constante évolution, nous nous adaptons continuellement pour proposer le meilleur accompagnement. Le département R&D de Nexialog Consulting se donne pour objectif de proposer des solutions innovantes à des problématiques métier ou d'actualité. Pour cela, nous nous appuyons sur des librairies internes et sur le travail de nos consultants. Le pôle R&D Nexialog a également pour mission de former les collaborateurs sur l'évolution des techniques et la réglementation en lien avec leur activité.

Site web du cabinet : https://www.nexialog.com
Publications : https://www.nexialog.com/publications-nexialog/

Contacts

Ali BEHBAHANI
Associé, Fondateur
Tél : + 33 (0) 1 44 73 86 78
Email : abehbahani@nexialog.com

Christelle BONDOUX
Associée, Directrice commerciale
Tél : + 33 (0) 1 44 73 75 67
Email : cbondoux@nexialog.com

Adrien MISKO
Manager R&D
Email : amisko@nexialog.com