

Filtrage exact dans les systèmes linéaires à sauts markoviens

Stéphane DERRODE¹, Wojciech PIECZYNSKI²

¹École Centrale Marseille & Aix-Marseille Université, Institut Fresnel (UMR 7249), 13013 Marseille, France.

²Télécom SudParis, Département CITI (UMR 5157), 91011 Evry, France.

stephane.derrode@centrale-marseille.fr, wojciech.pieczynski@telecom-sudparis.eu

Résumé – Nous considérons un système linéaire markovien gaussien $(\mathbf{X}, \mathbf{R}, \mathbf{Y})$, où \mathbf{X} un processus réel caché, \mathbf{R} une chaîne de Markov cachée, \mathbf{Y} un processus réel observé, et (\mathbf{X}, \mathbf{Y}) est gaussien conditionnellement à \mathbf{R} . Dans le modèle classique “Conditionally Gaussian Linear State-Space Model” (CGLSSM), le filtrage optimal n’est pas réalisable avec une complexité raisonnable. L’objet de ce papier est de proposer un modèle, proche du CGLSSM et appartenant à la famille de modèles dits “Conditionally Markov Switching Hidden Linear Models” (CMSHLMs), pour lesquels le filtrage exact optimal est réalisable avec un complexité linéaire du nombre d’observations. Le nouveau modèle et les algorithmes de filtrage proposés sont applicables dans toutes les situations où le filtrage approché pour le modèle CGLSSM est utilisé. Des résultats de simulation sont proposés, permettant de comparer, en termes d’EQM et de temps de calcul, les algorithmes avec des techniques de filtrage particulière.

Abstract – We consider triplet Markov Gaussian linear systems $(\mathbf{X}, \mathbf{R}, \mathbf{Y})$, where \mathbf{X} is hidden continuous random sequence, \mathbf{R} hidden discrete Markov chain, \mathbf{Y} observed continuous random sequence, and (\mathbf{X}, \mathbf{Y}) is Gaussian conditionally on \mathbf{R} . In the classical “Conditionally Gaussian Linear State-Space Model” (CGLSSM), optimal filter is not workable with a reasonable complexity. The aim of the paper is to propose a model, quite close to the CGLSSM and belonging to the general family of models “Conditionally Markov Switching Hidden Linear Models” (CMSHLMs), in which the computation of optimal filter with complexity linear in the number of observations is feasible. The new model and related filtering are immediately applicable in all situations where the classical CGLSSM is used via approximated filtering. Simulations results allow to compare, by means of MSE and computation time, the proposed algorithms with respect to particle filter ones.

1 Introduction

Considérons trois processus $\mathbf{X}_1^N = (\mathbf{X}_1, \dots, \mathbf{X}_N)$, $\mathbf{R}_1^N = (R_1, \dots, R_N)$ et $\mathbf{Y}_1^N = (\mathbf{Y}_1, \dots, \mathbf{Y}_N)$, où \mathbf{X}_1^N et \mathbf{Y}_1^N sont à valeurs respectivement dans \mathbb{R}^m et \mathbb{R}^q , alors que \mathbf{R}_1^N est à valeurs dans $\Omega = \{1, \dots, K\}$. Les processus \mathbf{X}_1^N et \mathbf{R}_1^N sont cachés, alors que \mathbf{Y}_1^N est observé. Le problème de « filtrage optimal » que nous abordons dans cet article consiste en l’estimation séquentielle de $(\mathbf{R}_1^N, \mathbf{X}_1^N)$ à partir de \mathbf{Y}_1^N . Plus précisément, en utilisant les notations usuelles pour les probabilités conditionnelles et les moyennes et variances conditionnelles, nous cherchons $p(r_{n+1} | \mathbf{y}_1^{n+1})$, $E[\mathbf{X}_{n+1} | r_{n+1}, \mathbf{y}_1^{n+1}]$ et $\text{Var}[\mathbf{X}_{n+1} | r_{n+1}, \mathbf{y}_1^{n+1}]$ à partir de $p(r_n | \mathbf{y}_1^n)$, $E[\mathbf{X}_n | r_n, \mathbf{y}_1^n]$, $\text{Var}[\mathbf{X}_n | r_n, \mathbf{y}_1^n]$ et \mathbf{y}_{n+1} ; le filtre est alors donné par

$$E[\mathbf{X}_{n+1} | \mathbf{y}_1^{n+1}] = \sum_{r_{n+1}} p(r_{n+1} | \mathbf{y}_1^{n+1}) E[\mathbf{X}_{n+1} | r_{n+1}, \mathbf{y}_1^{n+1}].$$

Ce type de problème est très important en pratique et des centaines d’articles [1, 2, 3, 4] ou livres [5, 6, 7, 8, 9] traitent du sujet depuis plusieurs décennies. Dans les modèles linéaires conditionnellement gaussiens (dits CGLSSM pour «Conditionally Gaussian Linear State-Space Models»), la distribution de $\mathbf{T}_1^N = (\mathbf{X}_1^N, \mathbf{R}_1^N, \mathbf{Y}_1^N)$ est obtenue en spécifiant que le couple $(\mathbf{X}_1^N, \mathbf{R}_1^N)$ a la structure d’une chaîne de Markov cachée et, conditionnellement à \mathbf{R}_1^N , $(\mathbf{X}_1^N, \mathbf{Y}_1^N)$ est un système linéaire gaussien. Ainsi lorsque \mathbf{R}_1^N est connu, le problème est résolu

par le filtre de Kalman classique. Cependant, lorsque ce n’est pas le cas, le problème n’a pas de solution connue avec une complexité raisonnable et différentes approximations, dont celles utilisant des filtres particuliers [2, 5, 6] ont été proposées. L’objectif de ce papier est d’introduire un modèle alternatif, appartenant à la famille des modèles linéaires à saut conditionnellement markoviens (dits CMSHLM pour «Conditionally Markov Switching Hidden Linear Models» [10], et pour lequel il est possible de proposer des algorithmes de filtrage optimal qui soit rapide et exact. Il est important de noter que le modèle alternatif proposé est très proche du modèle classique CGLSSM et peut être utilisé dans des situations identiques.

Plus précisément, un CGLSSM $\mathbf{T}_1^N = (\mathbf{X}_1^N, \mathbf{R}_1^N, \mathbf{Y}_1^N)$ est donné par la distribution $p(\mathbf{x}_1, r_1, \mathbf{y}_1)$ de $(\mathbf{X}_1, R_1, \mathbf{Y}_1)$ et par les transitions $p(\mathbf{t}_{n+1} | \mathbf{t}_n)$ vérifiant

$$\begin{aligned} \mathbf{R}_1^N \text{ markovien avec } p(r_{n+1} | \mathbf{x}_1^n, \mathbf{r}_1^n, \mathbf{y}_1^n) &= p(r_{n+1} | r_n); \quad (1) \\ \mathbf{X}_{n+1} &= \mathbf{A}_{n+1}(R_{n+1}) \mathbf{X}_n + \mathbf{C}_{n+1}(R_{n+1}) \mathbf{U}_{n+1}; \quad (2) \\ \mathbf{Y}_{n+1} &= \mathbf{B}_{n+1}(R_{n+1}) \mathbf{X}_{n+1} + \mathbf{D}_{n+1}(R_{n+1}) \mathbf{V}_{n+1}, \quad (3) \end{aligned}$$

avec $\mathbf{A}_{n+1}(R_{n+1})$, $\mathbf{B}_{n+1}(R_{n+1})$, $\mathbf{C}_{n+1}(R_{n+1})$, $\mathbf{D}_{n+1}(R_{n+1})$ des matrices de tailles adéquates dépendant des sauts, et \mathbf{U}_{n+1} , \mathbf{V}_{n+1} des bruits blancs gaussiens indépendants entre eux et tels que pour tout $n = 1, \dots, N-1$, $(\mathbf{U}_{n+1}, \mathbf{V}_{n+1})$ est indépendant de $\mathbf{T}_1^n = (\mathbf{X}_1^n, \mathbf{R}_1^n, \mathbf{Y}_1^n)$. Dans un tel modèle, les distribu-

tions marginales $p(\mathbf{x}_n, r_n, \mathbf{y}_n)$ sont, dans le cas général, des mélanges de distributions gaussiennes avec un nombre de composantes croissant de manière exponentielle avec n .

Nous proposons deux contributions :

1. Nous modifions le modèle CGLSSM en remplaçant $\mathbf{C}_{n+1}(R_{n+1})$ par $\mathbf{C}_{n+1}(\mathbf{R}_n^{n+1})$ de manière à ce que, dans le modèle modifié, les lois marginales $p(\mathbf{x}_n, r_n, \mathbf{y}_n)$ soient telles que $p(\mathbf{x}_n, \mathbf{y}_n | r_n)$ sont gaussiennes. Alors la forme générale des lois marginales $p(\mathbf{x}_n, r_n, \mathbf{y}_n)$ ne varie pas avec n . Notons cependant que les paramètres peuvent varier avec n , il ne s'agit donc pas de la stationnarité. De tels modèles nous semblent mieux adaptés aux applications ; en effet, dans le modèle ci-dessus ces mélanges deviennent de plus en plus riches, et donc la forme des marginales dépend du début de l'observation ;
2. Nous associons avec le modèle modifié, appelé Modèle ①, un nouveau modèle, appelé Modèle ②, qui est très « proche » (dans un sens qui sera précisé) du Modèle ① et qui permet un filtrage optimal rapide (complexité linéaire en temps). Cette dernière propriété est due au fait que le Modèle ② fait partie de la famille des modèles dits « Conditionally Markov Switching Hidden Linear Models » (CMSHLMs) récemment proposée dans [10].

Dans la suite, nous explicitons le filtre et proposons des simulations illustrant la « proximité » des deux modèles.

2 CGLSSM et CMSHLM associé

Considérons un CGLSSM défini par (1)-(3). Soit $\Gamma_{\mathbf{X}_1}(r_1), \dots, \Gamma_{\mathbf{X}_N}(r_N)$ une suite de matrices de covariance donnée. Considérons le Modèle ① défini par

$$\mathbf{R}_1^N \text{ vérifie (1);} \quad (4)$$

$$\mathbf{X}_{n+1} = \mathbf{A}_{n+1}(R_{n+1}) \mathbf{X}_n + \mathbf{C}_{n+1}(\mathbf{R}_n^{n+1}) \mathbf{U}_{n+1} \quad (5)$$

$$\text{avec } \mathbf{C}_{n+1}(\mathbf{R}_n^{n+1}) \mathbf{C}_{n+1}^T(\mathbf{R}_n^{n+1}) = \Gamma_{\mathbf{X}_{n+1}}(R_{n+1}) - \mathbf{A}_{n+1}(R_{n+1}) \Gamma_{\mathbf{X}_n}(R_n) \mathbf{A}_{n+1}^T(R_{n+1}); \quad (6)$$

$$\mathbf{Y}_{n+1} = \mathbf{B}_{n+1}(R_{n+1}) \mathbf{X}_{n+1} + \mathbf{D}_{n+1}(R_{n+1}) \mathbf{V}_{n+1}. \quad (7)$$

Sachant que (5) implique

$$\Gamma_{\mathbf{X}_{n+1}}(R_{n+1}) = \mathbf{A}_{n+1}(R_{n+1}) \Gamma_{\mathbf{X}_n}(R_n) \mathbf{A}_{n+1}^T(R_{n+1}) + \mathbf{C}_{n+1}(\mathbf{R}_n^{n+1}) \mathbf{C}_{n+1}^T(\mathbf{R}_n^{n+1}),$$

nous notons, grâce à (6), que la suite $\Gamma_{\mathbf{X}_1}(r_1), \dots, \Gamma_{\mathbf{X}_N}(r_N)$ est la suite des matrices de covariance des lois $p(\mathbf{x}_1 | r_1), \dots, p(\mathbf{x}_N | r_N)$.

Posons $\mathbf{Z}_n = (\mathbf{X}_n, \mathbf{Y}_n)^T$ et $\mathbf{W}_n = (\mathbf{U}_n, \mathbf{V}_n)^T$. En reportant \mathbf{X}_{n+1} donné par (5) dans (7), (5)-(7) s'écrivent

$$\mathbf{Z}_{n+1} = \mathbf{A}_{n+1}^{\textcircled{1}}(R_{n+1}) \mathbf{Z}_n + \mathbf{B}_{n+1}^{\textcircled{1}}(\mathbf{R}_n^{n+1}) \mathbf{W}_{n+1}, \quad (8)$$

avec $\mathbf{A}_{n+1}^{\textcircled{1}}(R_{n+1})$ et $\mathbf{B}_{n+1}^{\textcircled{1}}(\mathbf{R}_n^{n+1})$ données par

$$\mathbf{A}_{n+1}^{\textcircled{1}}(R_{n+1}) = \begin{bmatrix} \mathbf{A}_{n+1}(R_{n+1}) & 0 \\ \mathbf{B}_{n+1}(R_{n+1})\mathbf{A}_{n+1}(R_{n+1}) & 0 \end{bmatrix} \quad (9)$$

$$\mathbf{B}_{n+1}^{\textcircled{1}}(\mathbf{R}_n^{n+1}) = \begin{bmatrix} \mathbf{C}_{n+1}(\mathbf{R}_n^{n+1}) & 0 \\ \mathbf{B}_{n+1}(R_{n+1})\mathbf{C}_{n+1}(\mathbf{R}_n^{n+1}) & \mathbf{D}_{n+1}(R_{n+1}) \end{bmatrix} \quad (10)$$

Nous définissons comme Modèle ② associé au Modèle ①, le modèle défini par

$$\mathbf{Z}_{n+1} = \mathbf{A}_{n+1}^{\textcircled{2}}(\mathbf{R}_n^{n+1}) \mathbf{Z}_n + \mathbf{B}_{n+1}^{\textcircled{2}}(\mathbf{R}_n^{n+1}) \mathbf{W}_{n+1}, \quad (11)$$

avec

— Les matrices $\mathbf{A}_{n+1}^{\textcircled{2}}(\mathbf{R}_n^{n+1})$ données par

$$\mathbf{A}_{n+1}^{\textcircled{2}}(\mathbf{R}_n^{n+1}) = \begin{bmatrix} \mathbf{A}_{n+1}(R_{n+1}) & 0 \\ 0 & \mathbf{E}_{n+1}(\mathbf{R}_n^{n+1}) \end{bmatrix} \quad (12)$$

$$\text{et } \mathbf{E}_{n+1}(\mathbf{R}_n^{n+1}) = \mathbf{B}_{n+1}(R_{n+1}) \mathbf{A}_{n+1}(R_{n+1}) \Gamma_{\mathbf{X}_n \mathbf{Y}_n}(R_n) (\Gamma_{\mathbf{Y}_n}(R_n))^{-1};$$

— Les matrices $\mathbf{B}_{n+1}^{\textcircled{2}}(\mathbf{R}_n^{n+1})$ définies de manière à ce que les matrices de covariance $\Gamma_{\mathbf{Z}_{n+1}}(R_{n+1})$ soient identiques dans les deux modèles, ce qui s'écrit

$$\Gamma_{\mathbf{Z}_{n+1}}(R_{n+1}) = \mathbf{B}_{n+1}^{\textcircled{2}}(\mathbf{R}_n^{n+1}) (\mathbf{B}_{n+1}^{\textcircled{2}}(\mathbf{R}_n^{n+1}))^T + \mathbf{A}_{n+1}^{\textcircled{2}}(\mathbf{R}_n^{n+1}) \Gamma_{\mathbf{Z}_n}(R_n) (\mathbf{A}_{n+1}^{\textcircled{2}}(\mathbf{R}_n^{n+1}))^T \quad (13)$$

Ainsi $\mathbf{B}_{n+1}^{\textcircled{2}}(\mathbf{R}_n^{n+1})$ est donné récursivement par

$$\Gamma_{\mathbf{Z}_{n+1}}^{\textcircled{1}}(R_{n+1}) \text{ et } \mathbf{A}_{n+1}^{\textcircled{2}}(\mathbf{R}_n^{n+1}).$$

La proximité des deux modèles se traduit par le fait que les lois gaussiennes $p(\mathbf{x}_n | r_n), p(\mathbf{y}_{n+1} | r_{n+1}), p(\mathbf{x}_n^{n+1} | \mathbf{r}_n^{n+1}), p(\mathbf{x}_{n+1}, \mathbf{y}_n | \mathbf{r}_n^{n+1}), p(\mathbf{y}_n^{n+1} | \mathbf{r}_n^{n+1}), p(\mathbf{x}_n, \mathbf{y}_n | r_n)$ et $p(\mathbf{x}_{n+1}, \mathbf{y}_{n+1} | r_{n+1})$ sont identiques pour tout $n = 1, \dots, N$, et seules les lois $p(\mathbf{x}_n, \mathbf{y}_{n+1} | \mathbf{r}_n^{n+1})$ sont différentes. Cette propriété vient du fait que, en posant

$$\Gamma_{\mathbf{Z}_n}(R_n) = \begin{bmatrix} \Gamma_{\mathbf{X}_n}(R_n) & \Gamma_{\mathbf{X}_n \mathbf{Y}_n}(R_n) \\ \Gamma_{\mathbf{Y}_n \mathbf{X}_n}(R_n) & \Gamma_{\mathbf{Y}_n}(R_n) \end{bmatrix},$$

la matrice de covariances $\Gamma_{\mathbf{Z}_{n+1} \mathbf{Z}_n}^{\textcircled{1}}(\mathbf{R}_n^{n+1})$ vaut

$$\Gamma_{\mathbf{Z}_{n+1} \mathbf{Z}_n}^{\textcircled{1}}(\mathbf{R}_n^{n+1}) = \mathbf{A}_{n+1}^{\textcircled{1}}(R_{n+1}) \Gamma_{\mathbf{Z}_n}(R_n) = \begin{bmatrix} \alpha & \beta \\ \chi^1 & \delta \end{bmatrix}$$

$$\Gamma_{\mathbf{Z}_{n+1} \mathbf{Z}_n}^{\textcircled{2}}(\mathbf{R}_n^{n+1}) = \mathbf{A}_{n+1}^{\textcircled{2}}(\mathbf{R}_n^{n+1}) \Gamma_{\mathbf{Z}_n}(R_n) = \begin{bmatrix} \alpha & \beta \\ \chi^2 & \delta \end{bmatrix}$$

avec $\alpha = \mathbf{A}_{n+1}(R_{n+1}) \Gamma_{\mathbf{X}_n}(R_n), \delta = \mathbf{B}_{n+1}(R_{n+1}) \beta, \chi^1 = \mathbf{B}_{n+1}(R_{n+1}) \alpha, \beta = \mathbf{A}_{n+1}(R_{n+1}) \Gamma_{\mathbf{X}_n \mathbf{Y}_n}(R_n)$ et

$$\chi^2 = \delta \Gamma_{\mathbf{Y}_n}(R_n)^{-1} \Gamma_{\mathbf{Y}_n \mathbf{X}_n}(R_n)$$

Aussi, par construction, la loi markovienne de \mathbf{R}_1^N est la même dans les deux modèles.

Montrons maintenant que le Modèle ② fait partie de la famille de modèle CMSHLM proposée dans [10], ce qui rend le filtrage rapide, avec une complexité linéaire en temps, possible. Commençons par rappeler la définition de cette famille :

$\mathbf{T}_1^N = (\mathbf{X}_1^N, \mathbf{R}_1^N, \mathbf{Y}_1^N)$ est markovien avec

$$p(r_{n+1}, \mathbf{y}_{n+1} | \mathbf{x}_n, r_n, \mathbf{y}_n) = p(r_{n+1}, \mathbf{y}_{n+1} | r_n, \mathbf{y}_n) \quad ((\mathbf{R}_1^N, \mathbf{Y}_1^N) \text{ est alors markovien}); \quad (14)$$

$$\mathbf{X}_{n+1} = \mathbf{F}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1}) \mathbf{X}_n + \mathbf{G}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1}) \mathbf{W}_{n+1} + \mathbf{H}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1}). \quad (15)$$

avec $\mathbf{F}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1})$, $\mathbf{G}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1})$ et $\mathbf{H}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1})$ des matrices et vecteurs aux dimensions appropriées, et \mathbf{W}_{n+1} un bruit blanc. $p(r_{n+1} | \mathbf{y}_1^{n+1})$ et $E[\mathbf{X}_{n+1} | r_{n+1}, \mathbf{y}_1^{n+1}]$ peuvent être calculées à partir de $p(r_{n+1}, \mathbf{y}_{n+1} | r_n, \mathbf{y}_n)$, $\mathbf{F}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})$, $\mathbf{H}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})$, $p(r_n | \mathbf{y}_1^n)$ et $E[\mathbf{X}_n | r_n, \mathbf{y}_1^n]$ avec une complexité indépendante de n selon :

$$p(r_{n+1} | \mathbf{y}_1^{n+1}) = \frac{\sum_{r_n} p(r_{n+1}, \mathbf{y}_{n+1} | r_n, \mathbf{y}_n) p(r_n | \mathbf{y}_1^n)}{\sum_{r_n, r_n^*} p(r_{n+1}, \mathbf{y}_{n+1} | r_n, \mathbf{y}_n) p(r_n | \mathbf{y}_1^n)},$$

$$E[\mathbf{X}_{n+1} | r_{n+1}, \mathbf{y}_1^{n+1}] = \sum_{r_n} p(r_n | r_{n+1}, \mathbf{y}_1^{n+1}) (\mathbf{F}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}) E[\mathbf{X}_n | r_n, \mathbf{y}_1^n] + \mathbf{H}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})), \quad (16)$$

avec $p(r_{n+1}, \mathbf{y}_{n+1} | r_n, \mathbf{y}_n) = p(r_{n+1} | r_n) p(\mathbf{y}_{n+1} | \mathbf{r}_n^{n+1}, \mathbf{y}_n)$ et

$$p(r_n | r_{n+1}, \mathbf{y}_1^{n+1}) = \frac{p(r_{n+1}, \mathbf{y}_{n+1} | r_n, \mathbf{y}_n) p(r_n | \mathbf{y}_1^n)}{\sum_{r_n^*} p(r_{n+1}, \mathbf{y}_{n+1} | r_n^*, \mathbf{y}_n) p(r_n^* | \mathbf{y}_1^n)},$$

sachant que la distribution $p(\mathbf{y}_{n+1} | \mathbf{r}_n^{n+1}, \mathbf{y}_n)$ est gaussienne de moyenne $\mathbf{E}_{n+1}(\mathbf{r}_n^{n+1})\mathbf{y}_n$ et de matrice de covariance $\mathbf{Q}_{n+1}^4(\mathbf{r}_n^{n+1})$, avec

$$\mathbf{Q}_{n+1}(\mathbf{r}_n^{n+1}) = \mathbf{B}_{n+1}^{\otimes 2}(\mathbf{r}_n^{n+1})(\mathbf{B}_{n+1}^{\otimes 2}(\mathbf{r}_n^{n+1}))^T = \begin{bmatrix} \mathbf{Q}_{n+1}^1(\mathbf{r}_n^{n+1}) & \mathbf{Q}_{n+1}^2(\mathbf{r}_n^{n+1}) \\ \mathbf{Q}_{n+1}^3(\mathbf{r}_n^{n+1}) & \mathbf{Q}_{n+1}^4(\mathbf{r}_n^{n+1}) \end{bmatrix}$$

Le Modèle ② défini par (1) et (11)-(13) est un CMSHLM (14)-(15) avec $\mathbf{F}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1})$, $\mathbf{G}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1})$ et $\mathbf{H}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1})$ définis respectivement par (17)-(19).

Finalement, le filtre optimal dans le système à sauts (5)-(7) est, pour $\Gamma_{\mathbf{Z}_n}(r_n)$, $p(r_n | \mathbf{y}_1^n)$, $E[\mathbf{X}_n | r_n, \mathbf{y}_1^n]$ et \mathbf{y}_{n+1} donnés :

1. soit $\Gamma_{\mathbf{X}_{n+1}}(r_{n+1})$ et $\mathbf{C}_{n+1}(\mathbf{R}_n^{n+1})$ vérifiant (6), calculer $\Gamma_{\mathbf{Z}_n}(r_{n+1})$ avec (7) et $\mathbf{A}_{n+1}^{\otimes 2}(\mathbf{R}_n^{n+1})$ avec (12);
2. calculer $\mathbf{Q}_{n+1}(\mathbf{r}_n^{n+1}) = \mathbf{B}_{n+1}^{\otimes 2}(\mathbf{r}_n^{n+1})(\mathbf{B}_{n+1}^{\otimes 2}(\mathbf{r}_n^{n+1}))^T$ avec (13);
3. calculer $\mathbf{F}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})$, $\mathbf{H}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})$ et $\mathbf{G}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1})(\mathbf{G}_{n+1}(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}))^T$ avec (17)-(19);
4. calculer $p(r_{n+1}, \mathbf{y}_{n+1} | r_n, \mathbf{y}_n) = p(r_{n+1} | r_n) p(\mathbf{y}_{n+1} | \mathbf{r}_n^{n+1}, \mathbf{y}_n)$;
5. calculer $p(r_{n+1} | \mathbf{y}_1^{n+1})$, $E[\mathbf{X}_{n+1} | r_{n+1}, \mathbf{y}_1^{n+1}]$ et $E[\mathbf{X}_{n+1}\mathbf{X}_{n+1}^T | r_{n+1}, \mathbf{y}_1^{n+1}]$ avec (16).

3 Résultats expérimentaux

Dans cette section, nous illustrons la « proximité numérique » des deux modèles grâce à la restauration de données simulées

selon un scénario de poursuite de cible où $R_n = \{1, 2\}$ représente le changement des paramètres donnant les lois de probabilité qui décrivent le comportement de la cible. Nous considérons le cas où \mathbf{X}_1^N et \mathbf{Y}_1^N sont vectoriels ($m = q = 2$), et le cas stationnaire pour lequel les distributions de $(\mathbf{Z}_n, \mathbf{Z}_{n+1})$ sont indépendantes de $1, \dots, N-1$ dans les deux modèles. Ainsi $\mathbf{A}_{n+1} = \mathbf{A}$, $\mathbf{B}_{n+1} = \mathbf{B}$, $\mathbf{C}_{n+1} = \mathbf{C}$ et $\mathbf{D}_{n+1} = \mathbf{D}$, avec pour $R_{n+1} = 1$

$$\mathbf{A}(1) = \begin{bmatrix} 0.30 & 0.05 \\ 0.10 & 0.40 \end{bmatrix}, \mathbf{C}(1,1)\mathbf{C}^T(1,1) = \begin{bmatrix} 0.90 & 0.21 \\ 0.21 & 0.72 \end{bmatrix},$$

$$\mathbf{C}(2,1)\mathbf{C}^T(2,1) = \begin{bmatrix} 0.80 & 0.15 \\ 0.15 & 0.60 \end{bmatrix}, \mathbf{B}(1) = \begin{bmatrix} 0.30 & 0.40 \\ 0.40 & 0.30 \end{bmatrix},$$

$$\mathbf{D}(1)\mathbf{D}^T(1) = \begin{bmatrix} 0.69 & -0.10 \\ -0.10 & 0.69 \end{bmatrix},$$

et pour $R_{n+1} = 2$

$$\mathbf{A}(2) = \begin{bmatrix} 0.60 & 0.20 \\ 0.01 & 0.50 \end{bmatrix}, \mathbf{C}(1,2)\mathbf{C}^T(1,2) = \begin{bmatrix} 1.10 & 0.19 \\ 0.19 & 1.12 \end{bmatrix},$$

$$\mathbf{C}(2,2)\mathbf{C}^T(2,2) = \begin{bmatrix} 1.53 & 0.31 \\ 0.31 & 1.27 \end{bmatrix}, \mathbf{B}(2) = \begin{bmatrix} 0.20 & 0.40 \\ 0.40 & 0.30 \end{bmatrix},$$

$$\mathbf{D}(2)\mathbf{D}^T(2) = \begin{bmatrix} 1.60 & 0.05 \\ 0.05 & 0.93 \end{bmatrix}.$$

Finalement, la matrice de transition de la chaîne de Markov \mathbf{R}_1^N est donnée par $\begin{bmatrix} 0.98 & 0.02 \\ 0.02 & 0.98 \end{bmatrix}$.

Chaque expérience consiste en la simulation de $N = 1000$ données observées (\mathbf{y}_1^N) et de N couples de données cachées ($\mathbf{x}_1^N, \mathbf{r}_1^N$), selon le Modèle ① d'une part (série 1), et selon le Modèle ② d'autre part (série 2). Les données \mathbf{y}_1^N sont alors restaurées grâce

1. au filtre optimal selon le Modèle ①, en considérant les sauts \mathbf{r}_1^N connus (le filtre est appelé F1-SC¹);
2. au filtre optimal selon le Modèle ②, en considérant \mathbf{r}_1^N connus (le filtre est appelé F2-SC);
3. au filtre optimal selon le Modèle ②, en considérant \mathbf{r}_1^N inconnus (le filtre est appelé F2-SI);
4. au filtre particulière selon le Modèle ①, en considérant \mathbf{r}_1^N inconnus (le filtre est appelé FP1-SI);
5. au filtre particulière selon le Modèle ②, en considérant \mathbf{r}_1^N inconnus (le filtre est appelé FP2-SI).

Conditionnellement à \mathbf{r}_1^N , F1-SC est obtenu par un filtre de Kalman classique alors que F2-SC est obtenu par un filtre plus général (modèle de Markov couple), mais pour lequel des équations explicites existent [11, 12]. Enfin, les filtres particuliers FP1-SI et FP2-SI [2] ont été mis en oeuvre avec $P = 300$ particules et un ré-échantillonnage SIR de ces particules lorsque $N_{eff}/P < 0.3$ avec N_{eff} la somme des carrés des poids des particules. Notons que le filtre FP2-SI est une extension de la technique de filtrage particulières au cas de modèles couple [13]. Pour les filtres SI, les sauts sont estimés « en ligne », en maximisant $p(r_n | \mathbf{y}_1^n)$ à chaque instant n .

La proximité des modèles et l'efficacité des filtres sont évaluées par la moyenne des Erreurs Quadratiques Moyennes (EQM) de 300 expériences (cf. Table 1).

1. SC signifiant « Sauts Connus » et SI « Sauts Inconnus ».

$$\mathbf{F}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1}) = \mathbf{A}_{n+1}(R_{n+1}), \quad (17)$$

$$\mathbf{H}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1}) = \mathbf{Q}_{n+1}^2(\mathbf{R}_n^{n+1}) (\mathbf{Q}_{n+1}^4(\mathbf{R}_n^{n+1}))^{-1} (\mathbf{Y}_{n+1} - \mathbf{E}_{n+1}(\mathbf{R}_n^{n+1}) \mathbf{Y}_n), \quad (18)$$

$$\mathbf{G}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1}) (\mathbf{G}_{n+1}(\mathbf{R}_n^{n+1}, \mathbf{Y}_n^{n+1}))^T = \mathbf{Q}_{n+1}^1(\mathbf{R}_n^{n+1}) - \mathbf{Q}_{n+1}^2(\mathbf{R}_n^{n+1}) (\mathbf{Q}_{n+1}^4(\mathbf{R}_n^{n+1}))^{-1} \mathbf{Q}_{n+1}^3(\mathbf{R}_n^{n+1}). \quad (19)$$

TABLE 1 – EQM, taux d'erreur moyens pour les filtres à sauts inconnus et temps de restauration pour les différents filtres considérés.

	Série 1		Série 2		Temps
F1-SC	0.8097		0.8038		0.0012
F2-SC	0.8138		0.8006		0.0098
F2-SI	0.8192	(21.1%)	0.8072	(20.7%)	0.0282
FP1-SI	0.8153	(26.3%)	0.8097	(26.1%)	2.50
FP2-SI	0.8192	(26.5%)	0.8072	(26.0%)	2.45

4 Conclusion

Dans ce travail, nous avons proposé un modèle alternatif au modèle CGLSSM (1)-(3) permettant un filtrage statistique optimal, exacte et rapide car faisant partie de la famille CM-SHLM [10].

Le point important à noter est que le modèle CGLSSM est généralement bâti à partir de la connaissance des lois $p(\mathbf{x}_{n+1} | \mathbf{x}_n, r_n, r_{n+1})$ et des lois d'« attache aux données » $p(\mathbf{y}_n | \mathbf{x}_n, r_n)$. Sachant qu'elles sont identiques dans les deux modèles la justification de l'utilisation d'un Modèle ① est immédiatement transposable à l'utilisation du Modèle ② associé.

Comme perspectives, on peut noter l'utilisation des nouveaux modèles à des fins de lissage [14] ou prédiction [15]. L'introduction des copules dans le nouveau modèle, très utiles dans les modèles cachés discrets [16], constitue une autre perspective intéressante.

Références

- [1] C. Andrieu, C. M. Davy, and A. Doucet, “Efficient particle filtering for jump Markov systems. Application to time-varying autoregressions,” *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1762–1770, 2003.
- [2] A. Doucet, N. J. Gordon, and V. Krishnamurthy, “Particle filter for states estimation of jump Markov linear systems,” *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 613–624, 2001.
- [3] E. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky, “Bayesian nonparametric inference of switching dynamic linear models,” *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1569–1585, 2011.
- [4] R. H. Shumway and D. S. Stoffer, “Dynamic linear models with switching,” *J. of American Statistical Association*, vol. 86, no. 415, pp. 763–769, 1991.
- [5] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman filter : particle filters for tracking applications*. Boston, MA : Artech House, 2004.
- [6] O. Cappé, E. Moulines, and T. Ryden, *Inference in hidden Markov models*. Springer, 2005.
- [7] S. Frühwirth-Schnatter, *Finite mixture and Markov switching models*. New York : Springer-Verlag, 2006.
- [8] C.-J. Kim and C. R. Nelson, *State-space models with regime switching*. Cambridge, MA : MIT Press, 1999.
- [9] O. L. V. Costa, M. D. Fragoso, and R. P. Marques, *Discrete time Markov jump linear systems*. New York : Springer-Verlag, 2005.
- [10] W. Pieczynski, “Exact filtering in conditionally Markov switching hidden linear models,” *Comptes Rendus Mathématique*, vol. 349, no. 9-10, pp. 587–590, 2011.
- [11] W. Pieczynski and F. Desbouvries, “Kalman filtering using pairwise Gaussian models,” in *Proc. IEEE Int. Conf. ICASSP'03*, Hong-Kong, Apr. 2003.
- [12] V. Némésin and S. Derrode, “Robust blind pairwise Kalman algorithms using QR decompositions,” *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 5–9, 2013.
- [13] F. Desbouvries and W. Pieczynski, “Particle filtering in pairwise and triplet Markov chains,” in *Proc. IEEE EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP 2003)*, Grado-Gorizia, Jun. 2003.
- [14] W. Pieczynski, “Exact smoothing in hidden conditionally Markov switching linear models,” *Communications in Statistics - Theory and Methods*, vol. 40, no. 16, pp. 2823–2829, 2011.
- [15] N. Bardel and F. Desbouvries, “Exact Bayesian prediction in a class of Markov-switching models,” *Methodology and Computing in Applied probability*, vol. 14, no. 1, pp. 125–134, 2012.
- [16] S. Derrode and W. Pieczynski, “Unsupervised data classification using pairwise Markov chains with automatic copulas selection,” *Computational Statistics & Data Analysis*, vol. 63, pp. 81–98, 2013.