Exact optimal filtering in an approximating switching system

W. Pieczynski[1], S. Derrode[2], N. Abassi[1], Y. Petetin[1], et F. Desbouvries[1]

[1] Telecom SudParis
CITI Department and CNRS UMR 5157, 9 Rue Charles Fourier, 91000 Evry, France.
wojciech.pieczynski@telecom-sudparis.eu
[2] Ecole Centrale Marseille
Institut Fresnel (CNRS UMR 6133), 38, rue F. Joliot-Curie, 13451 Marseille cedex 20, France.
stephane.derrode@centrale-marseille.fr

Abstract We consider a triplet Markov Gaussian linear systems $(\mathbf{X}, \mathbf{R}, \mathbf{Y})$, where $\mathbf{X}$ is a sequence of continuous hidden states, $\mathbf{R}$ is a hidden discrete sequence, $\mathbf{Y}$ is an observed continuous sequence, and $(\mathbf{X}, \mathbf{Y})$ is Gaussian conditionally on $\mathbf{R}$. In the classical "Conditionally Gaussian Linear State-Space Model" (CGLSSM), the exact computation of the first and the second moments of the filtering distribution is a NP-hard problem. By contrast, in a recent family of triplet models called "Conditionally Markov Switching Hidden Linear Models" (CMSHLM), the exact computation of these moments can be done with complexity linear in the number of observations. In this paper, we show that it is possible to modify a given CGLSSM to obtain a quite close CMSHLM in which exact optimal filtering is possible. So we provide an alternative to classical approximative filtering techniques in CGLSSM.

Keywords Triplet Markov Chains, Exact filtering, Conditionally Gaussian Linear State-Space Model, Conditionally Markov Switching Hidden Linear Model.

1 Introduction

Let us consider three random sequences $\mathbf{X}_1^N = (\mathbf{X}_1, \cdots, \mathbf{X}_N)$, $\mathbf{R}_1^N = (R_1, \cdots, R_N)$ and $\mathbf{Y}_1^N = (\mathbf{Y}_1, \cdots, \mathbf{Y}_N)$, where $\mathbf{X}_n \in \mathbb{R}^m$, $\mathbf{Y}_n \in \mathbb{R}^q$ and $R_n \in \Omega = \{1, \cdots, K\}$. The problem we address in this paper is the sequential estimation of hidden $(\mathbf{R}_1^N, \mathbf{X}_1^N)$ from observed $\mathbf{Y}_1^N$. More precisely, we look for computing $p\left(r_{n+1}|\mathbf{y}_1^{n+1}\right)$ and $\mathrm{E}\left[\mathbf{X}_{n+1}|r_{n+1}, \mathbf{y}_1^{n+1}\right]$ from $p\left(r_n|\mathbf{y}_1^n\right)$, $\mathrm{E}\left[\mathbf{X}_n|r_n, \mathbf{y}_1^n\right]$ and $\mathbf{y}_{n+1}$ at a linear computational cost. The general model used for the distribution of the Markov chain (MC) $\mathbf{T}_1^N = (\mathbf{X}_1^N, \mathbf{R}_1^N, \mathbf{Y}_1^N)$ includes the two classical models that are Hidden Markov Chains (HMCs) [3] and Linear Gaussian State-Space Models (LGSSMs) [8]. Roughly speaking, $((\mathbf{X}_1^N, \mathbf{R}_1^N), \mathbf{Y}_1^N)$ is an HMC and given $\mathbf{R}_1^N$, $(\mathbf{X}_1^N, \mathbf{Y}_1^N)$ is a LGSSM:

$$\mathbf{R}_1^N \text{ is a MC with } p\left(r_{n+1}|\mathbf{x}_1^n, \mathbf{r}_1^n, \mathbf{y}_1^n\right) = p\left(r_{n+1}|r_n\right), \quad (1)$$

$$\mathbf{X}_{n+1} = \mathbf{A}_{n+1}(R_{n+1})\mathbf{X}_n + \mathbf{C}_{n+1}(R_{n+1})\mathbf{U}_{n+1}, \quad (2)$$

$$\mathbf{Y}_{n+1} = \mathbf{B}_{n+1}(R_{n+1})\mathbf{X}_{n+1} + \mathbf{D}_{n+1}(R_{n+1})\mathbf{V}_{n+1}, \quad (3)$$

with $\mathbf{A}_{n+1}(R_{n+1})$, $\mathbf{B}_{n+1}(R_{n+1})$, $\mathbf{C}_{n+1}(R_{n+1})$ and $\mathbf{D}_{n+1}(R_{n+1})$ appropriate matrices depending on switches, and $\mathbf{U}_{n+1}$, $\mathbf{V}_{n+1}$ appropriate Gaussian vectors. When the sequence $\mathbf{R}_1^N$ is known, the problem is solved by the Kalman Filter (KF); however, when $\mathbf{R}_1^N$ is not known the problem is NP-hard [9], so numerical or stochastic approximation methods need to be used [2,5].

Recently, another family of distributions for triplet $\mathbf{T}_1^N$ has been proposed[7] which makes possible the computation of $p\left(r_{n+1}|\mathbf{y}_1^{n+1}\right)$ and $\mathrm{E}\left[\mathbf{X}_{n+1}|r_{n+1},\mathbf{y}_1^{n+1}\right]$ with complexity linear in n:

$$p\left(r_{n+1},\mathbf{y}_{n+1}|\mathbf{x}_n,r_n,\mathbf{y}_n\right)=p\left(r_{n+1},\mathbf{y}_{n+1}|r_n,\mathbf{y}_n\right), \tag{4}$$

$$\mathbf{X}_{n+1}=\mathbf{F}_{n+1}(\mathbf{R}_n^{n+1},\mathbf{Y}_n^{n+1})\mathbf{X}_n+\mathbf{G}_{n+1}(\mathbf{R}_n^{n+1},\mathbf{Y}_n^{n+1})\mathbf{W}_{n+1}+\mathbf{H}_{n+1}(\mathbf{R}_n^{n+1},\mathbf{Y}_n^{n+1}), \tag{5}$$

where $\mathbf{F}_{n+1}(\mathbf{R}_n^{n+1},\mathbf{Y}_n^{n+1})$, $\mathbf{G}_{n+1}(\mathbf{R}_n^{n+1},\mathbf{Y}_n^{n+1})$ are appropriate matrices, $\mathbf{W}_{n+1}$ is an appropriate zero-mean independence sequence and $\mathbf{H}_{n+1}(\mathbf{R}_n^{n+1},\mathbf{Y}_n^{n+1})$ are appropriate vectors. The main difference between classical family (1)-(3) and recent family (4)-(5) lies in the fact that in the former, the couple $(\mathbf{X}_1^N,\mathbf{R}_1^N)$ is Markov while the couple $(\mathbf{R}_1^N,\mathbf{Y}_1^N)$ is not necessarily Markov, while in the latter the converse is true: the couple $(\mathbf{R}_1^N,\mathbf{Y}_1^N)$ is Markov while the couple $(\mathbf{X}_1^N,\mathbf{R}_1^N)$ is not necessarily Markov. Of course, in both classical and recent families, $\mathbf{T}_1^N$ is Markov. In this paper, we look for models(4)-(5) (so in which $p\left(r_n|\mathbf{y}_1^n\right)$, $\mathrm{E}\left[\mathbf{X}_n|r_n,\mathbf{y}_1^n\right]$ are computable at a linear computational cost) which are "close"to a given classical model (1)-(3). By close, we mean that we start from model (1)-(3) where jumps $\mathbf{R}_1^N$ are given and we look for a Pairwise MC (PMC) model [4,6] in which $\mathbf{Y}_1^N$ is a MC and the pdf of $(\mathbf{X}_n,\mathbf{Y}_n)$ coincides in both models. Finally, we introduce the discrete MC $\mathbf{R}_1^N$ in these particular PMC models and we apply the exact filtering technique[7].

The paper is organized as follows. In Section 2 we derive PMC models which are close to a given HMC one and in which $\mathbf{Y}_1^N$ is a MC. In Section 3, we extend the previous model in order to derive particular TMC models in which exact filtering is possible and which are close to a given model (1)-(3). In Section 4 we perform some simulations and end the paper with a conclusion.

2 Models with known switches

In the whole paper, $\mathbf{R}_1^N$ will be assumed to be a MC. In this section, we assume that switches $\mathbf{R}_1^N$ are known and we consider distributions conditional on $\mathbf{R}_1^N$. Therefore, although all matrices depend on R_{n+1} or R_n, we will temporarily forget this dependence. Thus, when $\mathbf{R}_1^N$ is given, model (1)-(3) reads

$$\mathbf{X}_{n+1}=\mathbf{A}_{n+1}(R_{n+1})\mathbf{X}_n+\mathbf{C}_{n+1}(R_{n+1})\mathbf{U}_{n+1}, \tag{6}$$

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

where $\mathbf{U}_1^N$, $\mathbf{V}_1^N$ are zero mean-sequences which are independent, mutually independent and independent of $\mathbf{X}_0$, with $\mathrm{E}\left[\mathbf{X}_0\right] = \mathbf{0}$.

Assuming that $(\mathbf{B}_{n+1}^1(R_{n+1}))^{-1}$ exists, let us consider the following equivalent form which will be called "Model 1":

$$\mathbf{Z}_{n+1} = \mathbf{A}_{n+1}^1(R_{n+1})\mathbf{Z}_n + \mathbf{B}_{n+1}^1(R_{n+1})\mathbf{W}_{n+1}, \text{ with } \mathbf{Z}_n = \begin{bmatrix}\mathbf{X}_n \\ \mathbf{Y}_n\end{bmatrix}, \tag{8}$$

$$\mathbf{W}_{n+1} = \begin{bmatrix}\mathbf{U}_{n+1} \\ \mathbf{V}_{n+1}\end{bmatrix}, \mathbf{A}_{n+1}^1 = \begin{bmatrix}\mathbf{A}_{n+1} & \mathbf{0} \\ \mathbf{B}_{n+1}\mathbf{A}_{n+1} & \mathbf{0}\end{bmatrix}, \mathbf{B}_{n+1}^1 = \begin{bmatrix}\mathbf{C}_{n+1} & \mathbf{0} \\ \mathbf{B}_{n+1}\mathbf{C}_{n+1} & \mathbf{D}_{n+1}\end{bmatrix}. \tag{9}$$

Let $\boldsymbol{\Gamma}_{\mathbf{Z}_n}$ be the covariance matrix of $\mathbf{Z}_n$: $\boldsymbol{\Gamma}_{\mathbf{X}_n} = \mathrm{E}\left[\mathbf{X}_n\mathbf{X}_n^T\right]$, $\boldsymbol{\Gamma}_{\mathbf{Y}_n} = \mathrm{E}\left[\mathbf{Y}_n\mathbf{Y}_n^T\right]$, $\boldsymbol{\Gamma}_{\mathbf{X}_n\mathbf{Y}_n} = \mathrm{E}\left[\mathbf{X}_n\mathbf{Y}_n^T\right]$ and $\boldsymbol{\Gamma}_{\mathbf{Y}_n\mathbf{X}_n} = \mathrm{E}\left[\mathbf{Y}_n\mathbf{X}_n^T\right]$. From (8) that sequence $(\boldsymbol{\Gamma}_{\mathbf{Z}_n})$ satisfies the classic following recursion:

$$\boldsymbol{\Gamma}_{\mathbf{Z}_{n+1}} = \mathbf{A}_{n+1}^1\boldsymbol{\Gamma}_{\mathbf{Z}_n}(\mathbf{A}_{n+1}^1)^T + \mathbf{B}_{n+1}^1(\mathbf{B}_{n+1}^1)^T. \tag{10}$$

Our goal is to look for a "Model 2"with the following properties:

(i) Model 2 is "close" to the model (6)-(7) in the sense that pdf of couple $(\mathbf{X}_n, \mathbf{Y}_n)$ are the same in Models 1 and 2;

(ii) In Model 2, $\mathbf{Y}_1^N$ is a MC and its transitions are identical to pdf $p\left(\mathbf{y}_{n+1}|\mathbf{y}_n\right)$ of Model 1.

The first point justifies the use of Model 2 in situations where Model 1 is used. The second point is of importance as it will allow us to use the recent results in [1,7] and to propose a fast exact optimal filtering in the presence of switches.

Proposition 1. Let us consider the Model 2 defined by

$$\mathbf{Z}_{n+1} = \mathbf{A}_{n+1}^2\mathbf{Z}_n + \mathbf{B}_{n+1}^2\mathbf{W}_{n+1}, \tag{11}$$

$$\mathbf{A}_{n+1}^2 = \begin{bmatrix}\mathbf{A}_{n+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{n+1}\mathbf{A}_{n+1}\boldsymbol{\Gamma}_{\mathbf{X}_n}(\mathbf{B}_n)^T(\boldsymbol{\Gamma}_{\mathbf{Y}_n})^{-1}\end{bmatrix} \tag{12}$$

and the sequence $\mathbf{B}_{n+1}^2$ verifying

$$\mathbf{B}_{n+1}^2(\mathbf{B}_{n+1}^2)^T = \mathbf{A}_{n+1}^1\boldsymbol{\Gamma}_{\mathbf{Z}_n}(\mathbf{A}_{n+1}^1)^T + \mathbf{B}_{n+1}^1(\mathbf{B}_{n+1}^1)^T - \mathbf{A}_{n+1}^2\boldsymbol{\Gamma}_{\mathbf{Z}_n}(\mathbf{A}_{n+1}^2)^T. \tag{13}$$

Then we can state:

(P1) $\mathbf{Y}_1^N$ is a MC and its transitions are given by pdf $p\left(\mathbf{y}_{n+1}|\mathbf{y}_n\right)$ of Model 1;

(P2) for any $n \geq 0$, covariance matrices $\mathrm{E}\left[\mathbf{Y}_{n+1}(\mathbf{Y}_n)^T\right]$ are the same in (8) and (11);

(P3) the sequence of variance-covariance matrices $\boldsymbol{\Gamma}_{\mathbf{Z}_n}$ is the same in (8) and (11);

(P4) in Models 1 and 2, $\mathbf{X}_1^N$ is a MC with the same distribution;

(P5) for any $n \geq 0$, the distributions of $(\mathbf{X}_n, \mathbf{Y}_n)$, $(\mathbf{X}_n, \mathbf{Y}_{n+1})$, $(\mathbf{Y}_n, \mathbf{Y}_{n+1})$ and $(\mathbf{X}_{n+1}, \mathbf{Y}_n)$ are the same in both models; so the distributions of $(\mathbf{X}_n, \mathbf{Y}_n, \mathbf{X}_{n+1})$ and $(\mathbf{X}_{n+1}, \mathbf{Y}_n, \mathbf{Y}_{n+1})$ are also identical in both models.

3 Gaussian switching linear systems

Let us now consider the equivalent formulation of CGLSSM (1)-(3):

$$\mathbf{R}_1^N \text{ is a MC with } p\left(r_{n+1}|\mathbf{x}_1^n, \mathbf{r}_1^n, \mathbf{y}_1^n\right) = p\left(r_{n+1}|r_n\right), \tag{14}$$

$$\mathbf{Z}_{n+1} = \mathbf{A}_{n+1}^1(R_{n+1})\mathbf{Z}_n + \mathbf{B}_{n+1}^1(R_{n+1})\mathbf{W}_{n+1}, \tag{15}$$

where $\mathbf{A}_{n+1}^1(R_{n+1})$ and $\mathbf{B}_{n+1}^1(R_{n+1})$ are defined as in (9) in which the dependence in R_{n+1} is introduced:

$$\mathbf{A}_{n+1}^1(R_{n+1}) = \begin{bmatrix} \mathbf{A}_{n+1}(R_{n+1}) & \mathbf{0} \\ \mathbf{B}_{n+1}(R_{n+1})\mathbf{A}_{n+1}(R_{n+1}) & \mathbf{0} \end{bmatrix}, \tag{16}$$

$$\mathbf{B}_{n+1}^1(R_{n+1}) = \begin{bmatrix} \mathbf{C}_{n+1}(R_{n+1}) & \mathbf{0} \\ \mathbf{B}_{n+1}(R_{n+1})\mathbf{C}_{n+1}(R_{n+1}) & \mathbf{D}_{n+1}(R_{n+1}) \end{bmatrix}. \tag{17}$$

Let us notice that the sequence of the covariance matrices satisfies

$$\begin{aligned}\boldsymbol{\Gamma}_{\mathbf{Z}_{n+1}}(r_{n+1}) = \mathbf{A}_{n+1}^1(r_{n+1}) &\left[\sum_{r_n} p\left(r_n|r_{n+1}\right) \boldsymbol{\Gamma}_{\mathbf{Z}_n}(r_n)\right] (\mathbf{A}_{n+1}^1(r_{n+1}))^T \\ &+ \mathbf{B}_{n+1}^1(r_{n+1})(\mathbf{B}_{n+1}^1(r_{n+1}))^T.\end{aligned} \tag{18}$$

As in Section 2 above, we now look for matching a CGLSSM to a CMSHLM. More precisely, let us consider the extension of Model 2 in section 2:

$$\mathbf{R}_1^N \text{ is a MC with } p\left(r_{n+1}|\mathbf{x}_1^n, \mathbf{r}_1^n, \mathbf{y}_1^n\right) = p\left(r_{n+1}|r_n\right), \tag{19}$$

$$\mathbf{Z}_{n+1} = \mathbf{A}_{n+1}^2(\mathbf{R}_n^{n+1})\mathbf{Z}_n + \mathbf{B}_{n+1}^2(\mathbf{R}_n^{n+1})\mathbf{W}_{n+1}, \tag{20}$$

where

$$\mathbf{A}_{n+1}^2(\mathbf{R}_n^{n+1}) = \begin{bmatrix} \mathbf{A}_{n+1}(R_{n+1}) & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_{n+1}(\mathbf{R}_n^{n+1}) \end{bmatrix}, \tag{21}$$

$$\mathbf{E}_{n+1}(\mathbf{R}_n^{n+1}) = \mathbf{B}_{n+1}(R_{n+1})\mathbf{A}_{n+1}(R_{n+1})\boldsymbol{\Gamma}_{\mathbf{X}_n}(R_n)(\mathbf{B}_{n+1}(R_{n+1}))^T(\boldsymbol{\Gamma}_{\mathbf{Y}_n}(R_n))^{-1}. \tag{22}$$

$\mathbf{B}_{n+1}^2(\mathbf{R}_n^{n+1})$ is defined such that $\boldsymbol{\Gamma}_{\mathbf{Z}_{n+1}}(r_{n+1})$ in the considered CMSHLM is equal to $\boldsymbol{\Gamma}_{\mathbf{Z}_{n+1}}(r_{n+1})$ defined in (18). It gives

$$\begin{aligned}\mathbf{B}_{n+1}^2(\mathbf{R}_n^{n+1})(\mathbf{B}_{n+1}^2(\mathbf{R}_n^{n+1}))^T = \boldsymbol{\Gamma}_{\mathbf{Z}_{n+1}}(r_{n+1}) - \\ \sum_{r_n} p\left(r_n|r_{n+1}\right) \mathbf{A}_{n+1}^2(\mathbf{R}_n^{n+1})\boldsymbol{\Gamma}_{\mathbf{Z}_n}(r_n)(\mathbf{A}_{n+1}^2(\mathbf{R}_n^{n+1}))^T.\end{aligned} \tag{23}$$

These particular CMSHLM models are of practical interest when we address the filtering problem in a CGLSSM one. Indeed, when the switches are known, they reduce to models (11)-(13) which are themselves close to the classical model. In addition, the mean and the variance of couple $(\mathbf{X}_n, \mathbf{Y}_n)$ are identical in both models.

Remark 1. When we apply the exact filtering technique [7], we need distributions $p\left(\mathbf{y}_{n+1}|\mathbf{r}_n^{n+1}, \mathbf{y}_n\right)$ and $p\left(\mathbf{x}_{n+1}|\mathbf{x}_n, \mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right)$. In models (19)-(23), these distributions are given by applying classical results on Gaussian distributions defined by $\mathbf{A}_{n+1}^2(\boldsymbol{r}_n^{n+1})$ and by $\mathbf{B}_{n+1}^2(\boldsymbol{r}_n^{n+1})(\mathbf{B}_{n+1}^2(\boldsymbol{r}_n^{n+1}))^T$.

So we propose a new filtering technique in CGLSSM models; starting from (1)-(3):

1. we derive a CMSHLM (19)-(20) where matrices $\mathbf{A}_{n+1}^2(\mathbf{R}_n^{n+1})$ and $\mathbf{B}_{n+1}^2(\mathbf{R}_n^{n+1})$ are respectively defined in (21) and (23);
2. we apply the filtering technique [7] where $p\left(\mathbf{y}_{n+1}|\mathbf{r}_n^{n+1}, \mathbf{y}_n\right)$ and $p\left(\mathbf{x}_{n+1}|\mathbf{x}_n, \mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right)$ are computed from $\mathbf{A}_{n+1}^2(\mathbf{R}_n^{n+1})$ and from $\mathbf{B}_{n+1}^2(\mathbf{R}_n^{n+1})(\mathbf{B}_{n+1}^2(\mathbf{R}_n^{n+1}))^T$.

4 Simulations

In experiment below, $\mathbf{X}_1^N$ and $\mathbf{Y}_1^N$ are assumed to be real valued processes, $\mathbf{Z}_n$ is assumed homogeneous in both models and $\Omega = \{1, 2\}$. We set $\boldsymbol{A}_n(1) = 0.3$, $\boldsymbol{A}_n(2) = 0.6$, $\boldsymbol{B}_n(1) = b_1$, $\boldsymbol{B}_n(2) = 0.2$, $\boldsymbol{\Gamma}_{\mathbf{X}_n}(1) = 1$, $\boldsymbol{\Gamma}_{\mathbf{Y}_n}(1) = 2$, $\boldsymbol{\Gamma}_{\mathbf{X}_n}(2) = 2$, $\boldsymbol{\Gamma}_{\mathbf{Y}_n}(2) = 4$, and for $r_n \in \Omega$, $\boldsymbol{C}_n^2(r_n) = \boldsymbol{\Gamma}_{\mathbf{X}_n}(r_n)(1 - \boldsymbol{A}_n^2(r_n))$ and $\boldsymbol{D}_n^2(r_n) = \boldsymbol{\Gamma}_{\mathbf{Y}_n}(r_n)(1 - \boldsymbol{B}_n^2(r_n))$. The jump transition matrix is set symetric with $p\left(R_1 = 1|R_2 = 1\right) = p\left(R_1 = 2|R_2 = 2\right) = 0.95$.

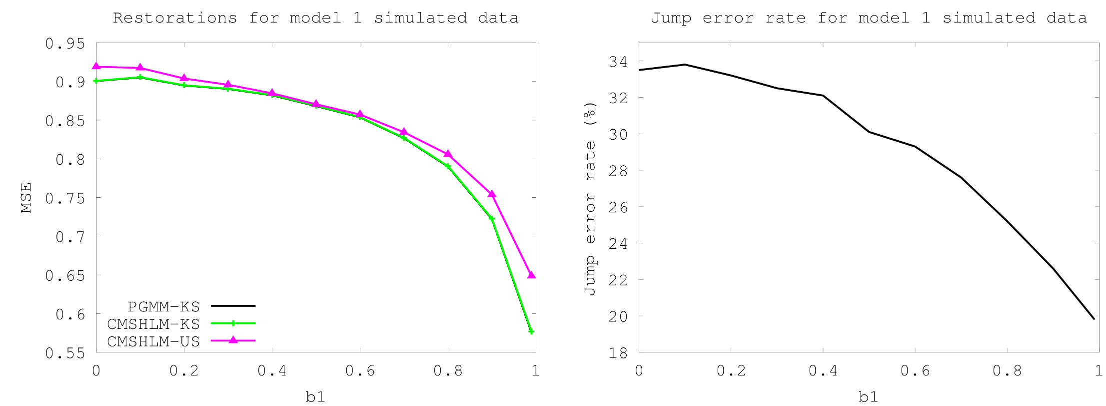
For this experiment, data were sampled according to model 1 given by (14)-(15) and restored by (1) the model 1 based optimal filter with known switches (denoted by PGMM-KS), (2) the model 2 based optimal filter with known switches (denoted by CMSHLM-KS), and (3) the model 2 based optimal filter with unknown switches (CMSHLM-US).

The results in Figure 1 are means of 300 independent experiments, each of them with $N = 1000$ data. Figure 1(a) reports the influence of b_1 on the Mean Square Error (MSE) of the estimated states by the three filters (when compared to the true states), while Figure 1(b) reports the influence of b_1 on the jump error rate when jumps are estimated by the third filter.

Particularly interesting, the performances between the two models are very close, whatever the value of $\boldsymbol{B}_n(1)$. The second interesting point is that the filter CMSHLM-US provides MSE which close to the MSE obtained from the first two filters ; the MSEs becoming almost equal when $\boldsymbol{B}_n(1) = \boldsymbol{B}_n(2)$. In the setting of this experiment, we conclude that model (11)-(13) is a good approximation of model (6)-(7).

5 Conclusion

In this paper we have proposed a new approximation filtering technique for CGLSSMs. Starting from a given CGLSSM, we have derived a close CMSHLM



(a) (b)

Figure 1. (a) MSEs for the three filters; (b) Jump error rate estimation for the CMSHLM-US filter.

in which the computation of the first and the second moments of the filtering distribution is possible. The main conclusion is that the two models are so close that it is difficult to see any difference at the results level, at least in the case of real-valued sequences considered. In addition, the results obtained with the new model with known switches are very close to those obtained when the switches are not known.

References

1. N. Abbassi, D. Benboudjema, and W. Pieczynski. Kalman filtering approximations in triplet Markov Gaussian switching models. In *IEEE Workshop on Statistical Signal Processing*, Nice, France, June 28-30 2011.
2. H. A. P. Blom and Y. Bar-Shalom. The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Transactions on Automatic Control*, 33(8):780–783, 1988.
3. O. Cappé, E. Moulines, and T. Rydén. *Inference in Hidden Markov Models*. Springer-Verlag, 2005.
4. S. Derrode and W. Pieczynski. Signal and image segmentation using pairwise Markov chains. *IEEE Transactions on Signal Processing*, 52(9):2477–89, 2004.
5. A. Doucet, N. J. Gordon, and V. Krishnamurthy. Particle filters for state estimation of jump Markov linear systems. *IEEE Transactions on Signal Processing*, 49(3):613–24, March 2001.
6. V. Némesin and S. Derrode. Robust pairwise Kalman filter using QR decompositions. *IEEE Transactions on Signal Processing*, 61(1):5–9, 2013.
7. W. Pieczynski. Exact filtering in conditionally Markov switching hidden linear models. *Comptes Rendus de l'Académie des Sciences Mathématiques*, 349(9-10):587–590, 2011.
8. B. Ristic, S. Arulampalam, and N. Gordon. *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, 2004.
9. J. K. Tugnait. Adaptive estimation and identification for discrete systems with Markov jump parameters. *IEEE Transactions on Automatic Control*, 27(5):1054–65, October 1982.