Fast exact filtering in generalized conditionally observed Markov switching models with copulas

Fei Zheng, Stéphane Derrode, Wojciech Pieczynski

Fast Exact Filtering in Generalized Conditionally Observed Markov Switching Models with Copulas

Fei Zheng[1], Stéphane Derrode[1], and Wojciech Pieczynski[2]

[1] École Centrale de Lyon, Univ. de Lyon, LIRIS, CNRS UMR 5205, Écully, France.
fei.zheng@doctorant.ec-lyon.fr, stephane.derrode@ec-lyon.fr
[2] Telecom Sudparis, SAMOVAR, CNRS UMR 5157, Évry, France.
wojciech.piezynski@telecom-sudparis.eu

Abstract We deal with the problem of statistical filtering in the context of Markov switching models. For $\mathbf{X}_1^N$ hidden continuous process, $\mathbf{R}_1^N$ hidden finite Markov process, and $\mathbf{Y}_1^N$ observed continuous one, the problem is to sequentially estimate $\mathbf{X}_1^N$ and $\mathbf{R}_1^N$ from $\mathbf{Y}_1^N$. In the classical "conditional Gaussian Linear state space model" (CGLSSM), where $(\mathbf{R}_1^N, \mathbf{X}_1^N)$ is a hidden Gaussian Markov chain, fast exact filtering is not workable. Recently, "conditionally Gaussian observed Markov switching model" (CGOMSM) has been proposed, in which $(\mathbf{R}_1^N, \mathbf{Y}_1^N)$ is a hidden Gaussian Markov chain instead. This model allows fast exact filtering. In this paper, using copula, we extend CGOMSM to a more general one, in which $(\mathbf{R}_1^N, \mathbf{Y}_1^N)$ is a hidden Markov chain (HMC) with noise of any form and the regimes are no need to be all Gaussian, while the exact filtering is still workable. Experiments are conducted to show how the exact filtering results based on CGOMSM can be improved by the use of the new model.

Key words Markov switching model, CGLSSM, CGOMSM, GCOMSM, Copulas, Optimal filter, Triplet Markov chain.

1 Introduction

Consider three random processes $\mathbf{X}_1^N = (\mathbf{X}_1, \ldots, \mathbf{X}_N)$, $\mathbf{R}_1^N = (\mathbf{R}_1, \ldots, \mathbf{R}_N)$ and $\mathbf{Y}_1^N = (\mathbf{Y}_1, \ldots, \mathbf{Y}_N)$. Each $\mathbf{X}_n$, $\mathbf{R}_n$, $\mathbf{Y}_n$ takes their values in $\mathbb{R}^m$, $\boldsymbol{\Omega} = \{1, \ldots, K\}$ and $\mathbb{R}^q$ respectively. The problem that we deal with is to find the unobservable (or hidden) processes $(\mathbf{R}_1^N, \mathbf{X}_1^N)$ from the observation $\mathbf{Y}_1^N = \mathbf{y}_1^N$. In the model we propose, we assume, as it is usually made, that both triplet $(\mathbf{X}_1^N, \mathbf{R}_1^N, \mathbf{Y}_1^N)$ and $\mathbf{R}_1^N$ are Markov chains. The first markovianity then implies that the couple $(\mathbf{X}_1^N, \mathbf{Y}_1^N)$ is Markovian conditionally on $\mathbf{R}_1^N$. The distribution of $(\mathbf{X}_1^N, \mathbf{R}_1^N, \mathbf{Y}_1^N)$ is defined by the initial distribution $p(\mathbf{x}_1, \mathbf{r}_1, \mathbf{y}_1)$ and the transitions $p\left(\mathbf{x}_{n+1}, r_{n+1}, \mathbf{y}_{n+1} \left| \mathbf{x}_n, r_n, \mathbf{y}_n\right.\right)$ which will be taken of the form $p\left(r_{n+1} \left| r_n\right.\right) p\left(\mathbf{x}_{n+1}, \mathbf{y}_{n+1} \left| \mathbf{r}_n^{n+1}, \mathbf{x}_n, \mathbf{y}_n\right.\right)$ consistently with the Markovianity of $\mathbf{R}_1^N$. Here (r_n, r_{n+1}) is denoted by $(\mathbf{r}_n^{n+1})$ for short.

In the "conditionally Gaussian observed Markov switching model" (CGOMSM) proposed in [1] and applied to general non-linear systems in [7, 8], the transitions $p\left(\mathbf{x}_{n+1}, \mathbf{y}_{n+1} \left| \mathbf{r}_n^{n+1}, \mathbf{x}_n, \mathbf{y}_n\right.\right)$ are assumed to be Gaussian with linear regimes. The aim of the paper is to extend the CGOMSM, which allows exact filtering, to a more general one in which $p\left(\mathbf{y}_{n+1}\left|\mathbf{r}_n^{n+1},\mathbf{y}_n\right.\right)$ are no longer limited to be Gaussian and the regime $\mathcal{G}\left(\mathbf{x}_{n+1}\left|\mathbf{x}_n,\mathbf{r}_n^{n+1},\mathbf{y}_n^{n+1}\right.\right)$ are no longer necessarily to be linear on the observations. The new model, called "Generalized conditionally observed Markov switching model" (GCOMSM), benefits from the copulas, which has been widely used in statistical finance for dependence description [4,11]. Copulas were firstly introduced into hidden Markov chain (HMC) with dependent noise by [3], and importance of their role in segmentation efficiency is shown in [5,6]. However, to our best knowledge, no work considers them in switching state-space models. In the proposed GCOMSM, $p\left(\mathbf{y}_{n+1}\left|\mathbf{r}_n^{n+1},\mathbf{y}_n\right.\right)$ is much more flexible compare to the original CGOMSM making use of copula. Experiments are conducted to show the interest of the new model with comparison to the result given by traditional Gaussian linear assumptions.

The paper is organized as follows. In next Sections, we recall CGOMSM, specify the general GCOMSM and show how fast optimal filtering and smoothing runs in this new model. Experiments are displayed and analyzed in the third Section. Finally, the conclusion and perspectives are given in the last Section four.

2 Generalized conditionally observed Markov switching model (GCOMSM)

Let us consider a CGOMSM, in which the Markov triplet $(\mathbf{X}_1^N,\mathbf{R}_1^N,\mathbf{Y}_1^N)$ distribution is defined by $p\left(\mathbf{x}_1,\mathbf{r}_1,\mathbf{y}_1\right)=p\left(\mathbf{r}_1\right)p\left(\mathbf{x}_1,\mathbf{y}_1\left|\mathbf{r}_1\right.\right)$ and transitions of the form $p\left(\mathbf{x}_{n+1},r_{n+1},\mathbf{y}_{n+1}\left|\mathbf{x}_n,r_n,\mathbf{y}_n\right.\right)=p\left(r_{n+1}\left|r_n\right.\right)p\left(\mathbf{x}_{n+1},\mathbf{y}_{n+1}\left|\mathbf{x}_n,\mathbf{y}_n,\mathbf{r}_n^{n+1}\right.\right)$. Both $p\left(\mathbf{x}_1,\mathbf{y}_1\left|\mathbf{r}_1\right.\right)$ and $p\left(\mathbf{x}_{n+1},\mathbf{y}_{n+1}\left|\mathbf{x}_n,\mathbf{y}_n,\mathbf{r}_n^{n+1}\right.\right)$ are Gaussian. In [1], the CGOMSM is described by linear regime:

$$\begin{bmatrix}\mathbf{X}_{n+1}\\ \mathbf{Y}_{n+1}\end{bmatrix}=\underbrace{\begin{bmatrix}\boldsymbol{\mathcal{F}}^{xx}(\mathbf{R}_n^{n+1}) & \boldsymbol{\mathcal{F}}^{xy}(\mathbf{R}_n^{n+1})\\ \mathbf{0} & \boldsymbol{\mathcal{F}}^{yy}(\mathbf{R}_n^{n+1})\end{bmatrix}}_{\boldsymbol{\mathcal{F}}(\mathbf{R}_n^{n+1})}\begin{bmatrix}\mathbf{X}_n\\ \mathbf{Y}_n\end{bmatrix}+\begin{bmatrix}\mathbf{U}_{n+1}\\ \mathbf{V}_{n+1}\end{bmatrix}, \tag{1}$$

with $\boldsymbol{\mathcal{F}}(\mathbf{R}_n^{n+1})$ an appropriate system transition matrix, and $\left[\mathbf{U}_{n+1}^{\intercal},\mathbf{V}_{n+1}^{\intercal}\right]^{\intercal}$ represents the independent Gaussian zero-mean noise which are independent from $\mathbf{T}_n=(\mathbf{X}_n,\mathbf{R}_n,\mathbf{Y}_n)$.

We see in CGOMSM, the pair $(\mathbf{R}_1^N,\mathbf{Y}_1^N)$ is a Markov chain, and

$$p\left(\mathbf{x}_{n+1},\mathbf{y}_{n+1}\left|\mathbf{r}_n^{n+1},\mathbf{x}_n,\mathbf{y}_n\right.\right)=p\left(\mathbf{y}_{n+1}\left|\mathbf{r}_n^{n+1},\mathbf{y}_n\right.\right)p\left(\mathbf{x}_{n+1}\left|\mathbf{x}_n,\mathbf{r}_n^{n+1},\mathbf{y}_n^{n+1}\right.\right), \tag{2}$$

which makes $p\left(\mathbf{r}_1^N\left|\mathbf{y}_1^N\right.\right)$ can be computed, thus the exact filtering is feasible.

Under CGOMSM, $p\left(\mathbf{y}_{n+1}\left|\mathbf{r}_n^{n+1},\mathbf{y}_n\right.\right)$ is Gaussian and $\mathcal{G}\left(\mathbf{x}_{n+1}\left|\mathbf{x}_n,\mathbf{r}_n^{n+1},\mathbf{y}_n^{n+1}\right.\right)$ is linear on $\mathbf{x}_n$, $\mathbf{y}_n$ and $\mathbf{y}_{n+1}$. However, to maintain the feasibility of exact filtering, the Gaussian setting and linear form are not necessary conditions.

2.1 Definition of Generalized conditionally observed Markov switching model (GCOMSM)

The Generalized conditionally observed Markov switching model (GCOMSM) which extends the CGOMSM considers still the triplet $\left(\mathbf{X}_1^N, \mathbf{R}_1^N, \mathbf{Y}_1^N\right)$ a Markov chain, defined by $p\left(\mathbf{x}_1, \mathbf{r}_1, \mathbf{y}_1\right)$ and transition of the form

$$\begin{aligned} p\left(\mathbf{x}_{n+1}, r_{n+1}, \mathbf{y}_{n+1} \left|\mathbf{x}_n, r_n, \mathbf{y}_n\right.\right) =& p\left(r_{n+1} \left|r_n\right.\right) \\ & p\left(\mathbf{y}_{n+1} \left|\mathbf{r}_n^{n+1}, \mathbf{y}_n\right.\right) p\left(\mathbf{x}_{n+1} \left|\mathbf{x}_n, \mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right.\right), \end{aligned} \tag{3}$$

Unlike in CGOMSM, $p\left(\mathbf{y}_{n+1} \left|\mathbf{r}_n^{n+1}, \mathbf{y}_n\right.\right)$ is enriched by copula represented as:

$$\begin{aligned} p\left(\mathbf{y}_{n+1} \left|\mathbf{r}_n^{n+1}, \mathbf{y}_n\right.\right) =& f_{n+1}^r\left(\mathbf{y}_{n+1} | \mathbf{r}_n^{n+1}\right) \\ & c_{n+1}\left(F_{n+1}^l\left(\mathbf{y}_n | \mathbf{r}_n^{n+1}\right), F_{n+1}^r\left(\mathbf{y}_{n+1} | \mathbf{r}_n^{n+1}\right) | \mathbf{r}_n^{n+1}\right), \end{aligned} \tag{4}$$

where we use $f_{n+1}^l\left(\mathbf{y}_n | \mathbf{r}_n^{n+1}\right)$ and $f_{n+1}^r\left(\mathbf{y}_{n+1} | \mathbf{r}_n^{n+1}\right)$ to denote the probability density function (PDF) of the left and right margins respectively. Similarly, $F_{n+1}^l\left(\mathbf{y}_n | \mathbf{r}_n^{n+1}\right)$, $F_{n+1}^r\left(\mathbf{y}_{n+1} | \mathbf{r}_n^{n+1}\right)$ are their associated cumulative distribution function (CDF), while $c_{n+1}\left(\cdot, \cdot | \mathbf{r}_n^{n+1}\right)$ represents the density of the two-dimensional copula conditionally on switches. The copula above then completes the two margins to form a joint distribution $p\left(\mathbf{y}_n^{n+1} \left|\mathbf{r}_n^{n+1}\right.\right)$ which can theoretically embrace any distribution form.

Moreover, the simple linear regime $\mathcal{G}\left(\mathbf{x}_{n+1} \left|\mathbf{x}_n, \mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right.\right)$ which corresponds to $p\left(\mathbf{x}_{n+1} \left|\mathbf{x}_n, \mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right.\right)$ in CGOMSM is extend to

$$\mathbf{x}_{n+1} = \mathbf{A}_{n+1}\left(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right) \mathbf{x}_n + \mathbf{B}_{n+1}\left(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right) + \boldsymbol{\nu}_{n+1} \tag{5}$$

in GCOMSM, in which $\mathbf{A}_{n+1}(\cdot)$ and $\mathbf{B}_{n+1}(\cdot)$ can be any function forms of r_n, r_{n+1}, $\mathbf{y}_n$, $\mathbf{y}_{n+1}$. $\boldsymbol{\nu}_{n+1} \sim \mathcal{N}\left(0, \boldsymbol{\mathcal{V}}_{n+1}\left(\mathbf{r}_n^{n+1}\right)\right)$. Integrally, they can be also written as

$$\mathbf{x}_{n+1} \sim \mathcal{N}\left\{\mathbf{A}_{n+1}\left(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right) \mathbf{x}_n + \mathbf{B}_{n+1}\left(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right), \boldsymbol{\mathcal{V}}_{n+1}\left(\mathbf{r}_n^{n+1}\right)\right\}. \tag{6}$$

2.2 Fast exact filtering in GCOMSM

The Markov property of $\left(\mathbf{R}_1^N, \mathbf{Y}_1^N\right)$ in GCOMSM leads to $p\left(\mathbf{x}_n \left|\mathbf{r}_n^{n+1}, \mathbf{y}_1^{n+1}\right.\right) = p\left(\mathbf{x}_n \left|r_n, \mathbf{y}_1^n\right.\right)$. Besides, since $p\left(\mathbf{x}_{n+1} \left|\mathbf{x}_n, \mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right.\right)$ is Gaussian defined as (6), we have

$$\mathbb{E}\left[\mathbf{X}_{n+1} \left|\mathbf{x}_n, \mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right.\right] = \mathbf{A}_{n+1}\left(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right) \mathbb{E}\left[\mathbf{X}_n \left|r_n, \mathbf{y}_1^n\right.\right] + \mathbf{B}_{n+1}\left(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right). \tag{7}$$

Then $\mathbb{E}\left[\mathbf{X}_{n+1} \left|r_{n+1}, \mathbf{y}_1^{n+1}\right.\right]$ is computable from $\mathbb{E}\left[\mathbf{X}_n \left|r_n, \mathbf{y}_1^n\right.\right]$ with

$$\begin{aligned} \mathbb{E}\left[\mathbf{X}_{n+1} \left|r_{n+1}, \mathbf{y}_1^{n+1}\right.\right] =& \sum_{r_n} p\left(r_n \left|r_{n+1}, \mathbf{y}_1^{n+1}\right.\right) \left\{\mathbf{A}_{n+1}\left(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right) \mathbb{E}\left[\mathbf{X}_n \left|r_n, \mathbf{y}_1^n\right.\right]\right. \\ & \left. + \mathbf{B}_{n+1}\left(\mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right)\right\} \end{aligned} \tag{8}$$

in which $p\left(r_n \left| r_{n+1}, \mathbf{y}_1^{n+1}\right.\right)$ is computable because of the Markovianity of $\left(\mathbf{R}_1^N, \mathbf{Y}_1^N\right)$. More precisely, we can write

$$p\left(r_n \left| r_{n+1}, \mathbf{y}_1^{n+1}\right.\right) = \frac{p\left(\mathbf{r}_n^{n+1}, \mathbf{y}_1^{n+1}\right)}{\sum_{r_n} p\left(\mathbf{r}_n^{n+1}, \mathbf{y}_1^{n+1}\right)}, \tag{9}$$

and $p\left(\mathbf{r}_n^{n+1}, \mathbf{y}_1^{n+1}\right)$ can be calculated recursively with

$$\begin{aligned} p\left(\mathbf{r}_n^{n+1}, \mathbf{y}_1^{n+1}\right) &= \sum_{\mathbf{r}_{n-1}} p\left(\mathbf{r}_{n-1}, \mathbf{r}_n^{n+1}, \mathbf{y}_1^{n+1}\right) \\ &= \sum_{\mathbf{r}_{n-1}} p\left(\mathbf{r}_{n-1}^n, \mathbf{y}_1^n\right) p\left(r_{n+1} \left| r_n\right.\right) p\left(\mathbf{y}_{n+1} \left| \mathbf{r}_n^{n+1}, \mathbf{y}_n\right.\right), \end{aligned} \tag{10}$$

Finally, the filtering is given by

$$\mathbb{E}\left[\mathbf{X}_{n+1} \left| \mathbf{y}_1^{n+1}\right.\right] = \sum_{r_{n+1}} p\left(r_{n+1} \left| \mathbf{y}_1^{n+1}\right.\right) \mathbb{E}\left[\mathbf{X}_{n+1} \left| r_{n+1}, \mathbf{y}_1^{n+1}\right.\right]. \tag{11}$$

3 Example of GCOMSM and experiment on the matched exact filtering

We present here an example to show the flexibility of the proposed GCOMSM as well as the performance of the matched exact filtering. We focus on the time-independent case of the general GCOMSM, which means that the parameters depend only on the switches $(\mathbf{r}_n^{n+1})$. For simplification, we assume that $(\mathbf{R}_1^N, \mathbf{Y}_1^N)$ is stationary reversible, which means that $p\left(\mathbf{y}_{n+1} \left| \mathbf{r}_n^{n+1}\right.\right) = p\left(\mathbf{y}_{n+1} \left| r_{n+1}\right.\right)$, therefore the "left" and "right" margins in (4) are equal. Under these assumptions the equation (4) and (6) can be written as

$$p\left(\mathbf{y}_{n+1} \left| \mathbf{r}_n^{n+1}, \mathbf{y}_n\right.\right) = f_{r_{n+1}}\left(\mathbf{y}_{n+1}\right) c_{\mathbf{r}_n^{n+1}}\left(F_{r_n}\left(\mathbf{y}_n\right), F_{r_{n+1}}\left(\mathbf{y}_{n+1}\right)\right) \tag{12}$$

$$\mathbf{x}_{n+1} \sim \mathcal{N}\left\{\mathbf{A}_{\mathbf{r}_n^{n+1}}\left(\mathbf{y}_n^{n+1}\right)\mathbf{x}_n + \mathbf{B}_{\mathbf{r}_n^{n+1}}\left(\mathbf{y}_n^{n+1}\right), \boldsymbol{\mathcal{V}}_{\mathbf{r}_n^{n+1}}\right\}. \tag{13}$$

In place of the time dependence in original definition, the dependence on switches are moved to subscript of all functions. For this example, we assume that $\mathbf{R}_1^N$ has two component values $\boldsymbol{\Omega} = \{1, 2\}$. And for each $j, k \in \boldsymbol{\Omega}$, $f_j\left(\mathbf{y}_n\right) = f_{r_n=j}\left(\mathbf{y}_n\right)$, $c_{j,k}\left(F_j\left(\mathbf{y}_n\right), F_k\left(\mathbf{y}_{n+1}\right)\right) = c_{r_n=j, r_{n+1}=k}\left(F_j\left(\mathbf{y}_n\right), F_k\left(\mathbf{y}_{n+1}\right)\right)$ with F_j, $C_{j,k}$ the associated CDF in (12). In (13), the abbreviation is taken in the same way: $\mathbf{A}_{j,k}\left(\mathbf{y}_n^{n+1}\right) = \mathbf{A}_{r_n=j, r_{n+1}=k,}\left(\mathbf{y}_n^{n+1}\right)$, so as $\mathbf{B}_{j,k}\left(\mathbf{y}_n^{n+1}\right)$ and $\boldsymbol{\mathcal{V}}_{j,k}$.

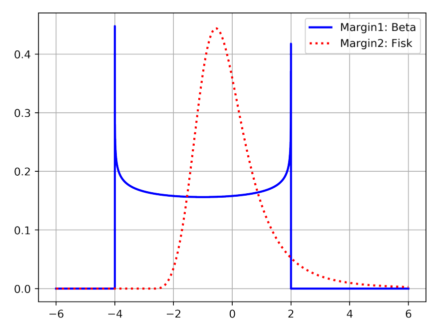
The parameters of $p\left(\mathbf{y}_{n+1} \left| \mathbf{r}_n^{n+1}, \mathbf{y}_n\right.\right)$ which are set to be non-Gaussian as

- Margins: $f_1(\mathbf{y}_n) = \text{Beta}^{\,1}\{\alpha_1 = 0.9, \beta_1 = 0.9, loc_1 = -4, scale_1 = 6\}$,
  $f_2(\mathbf{y}_n) = \text{Fisk}^{\,2}\{\beta_2 = 4, loc_2 = -2.7, scale_2 = 2.4\}$.
- Copulas: $c_{1,1}\{\cdot,\cdot\} = \text{Arch14}^{\,3}\{\cdot,\cdot|\alpha_{1,1} = 3\}$, $c_{2,2}\{\cdot,\cdot\} = \text{FGM}\{\cdot,\cdot|\alpha_{2,2} = 0.5\}$,
  $c_{1,2}\{\cdot,\cdot\} = c_{2,1}\{\cdot,\cdot\} = \text{Clayton}\{\cdot,\cdot|\alpha_{1,2} = 4.7\}$.

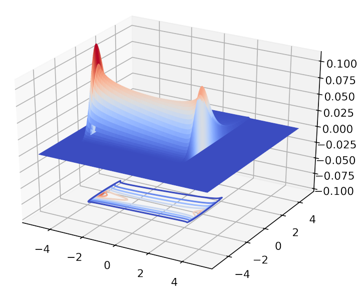
The marginal and joint distribution are displayed in Figure 1a, 1b.

$p\left(\mathbf{x}_{n+1}\left|\mathbf{x}_n, \mathbf{r}_n^{n+1}, \mathbf{y}_n^{n+1}\right.\right)$ is set with $\mathbf{A}_{j,k}\left(\mathbf{y}_n^{n+1}\right) = a_{j,k}\mathbf{x}_n$, simple non-linear function on $\mathbf{y}_n$, $\mathbf{y}_{n+1}$ that $\mathbf{B}_{j,k}\left(\mathbf{y}_n^{n+1}\right) = b_{j,k}\mathbf{y}_n\mathbf{y}_{n+1}$, and in which the parameters are assigned as

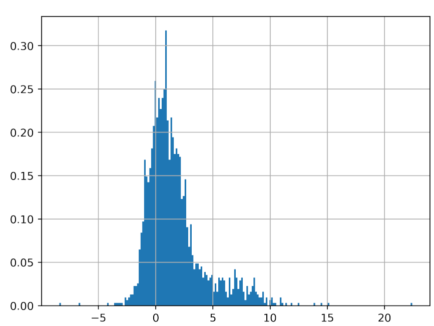
- $a_{j,k}$: $a_{1,1} = 0.2$, $a_{1,2} = 0.4$, $a_{2,1} = 0.6$, $a_{2,2} = 0.8$,
- $b_{j,k}$: $b_{1,1} = 0.7$, $b_{1,2} = 0.5$, $b_{2,1} = 0.6$, $b_{2,2} = 0.9$,
- $\boldsymbol{\mathcal{V}}_{j,k}$: $\boldsymbol{\mathcal{V}}_{1,1} = \boldsymbol{\mathcal{V}}_{2,2} = 1.0$, $\boldsymbol{\mathcal{V}}_{1,2} = \boldsymbol{\mathcal{V}}_{2,1} = 0.8$.



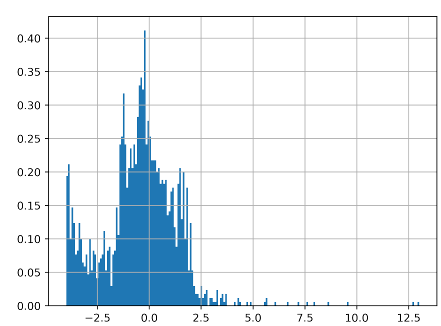
(a) Margins ($f_1(\mathbf{y}_n)$:Beta, $f_2(\mathbf{y}_n)$:Fisk).



(b) joint distribution of $f_{1,2}\left(\mathbf{y}_n, \mathbf{y}_{n+1}\right)$.



(c) Histogram of $\mathbf{x}_1^N$.



(d) Histogram of $\mathbf{y}_1^N$.

Figure 1: Distributions and histograms of simulated GCOMSM data.

1. α_1 and β_1 are the shape parameters, loc_1 and $scale_1$ are short for location and scale.
2. β_2 represents the shape parameter, loc_2 and $scale_2$ for location and scale.
3. Short for Archimiedean copula, order: 14.

2000 samples are simulated according to the above setting of GCOMSM. We see from the histograms of the simulated data illustrated in Figure 1c, 1d that they are hardly to be approximated by Gaussian mixtures with small component number. Exact filtering for GCOMSM is applied on $\mathbf{y}_1^N$ to restore the hidden $\mathbf{r}_1^N$ (decided by maximum posterior mode criterion from $p\left(r_n\left|\mathbf{y}_1^n\right.\right)$) and $\mathbf{x}_1^N$. For comparison, we conducted also the filtering based on Gaussian assumptions (both margins and copulas are assumed to be Gaussian) by using Maximum likelihood (ML) and Pseudo-Likelihood Maximization (PLM) [9] for Gaussian parameter estimation of margins and copulas applied on data. Restoration results are average of 100 independent experiments, illustrated in Table 1. We can see that the exact filtering performs well on restoring the hidden switches and states for GCOMSM, while the Gaussian assumption is obviously inferior comparing to the exact filter.

Table 1: Restoration result.

Observation	Exact filtering		Filtering (Gaussian)	
MSE	Error Ratio	MSE	Error Ratio	MSE
16.48	13.96%	1.18	35.91%	2.16

Their performance can also be told from the trajectories. Figure 2 illustrates an trajectory example from one instance among the 100 experiment.

4 Conclusion

In this work, copula is introduced in the recent "conditionally Gaussian observed Markov switching model" (CGOMSM), and fuse to a more general one called "generalized conditionally observed Markov switching model" (GCOMSM). Experiments verify the capability of GCOMSM to work on data under flexible distributions. The fast exact filtering for GCOMSM can be much less time consuming comparing to using other non-Gaussian models which do not allow exact filtering and Monte-Carlo methods are needed to be applied. The future work may contain the model identification (identifying the margins, copulas [2, 6, 12], and also the conditional regime functions [10]) of GCOMSM, and application of the model on non-Gaussian non-linear data restoration. In addition, smoothing can also be a perspective of interest.

References

1. Noufel Abbassi, Dalila Benboudjema, Stéphane Derrode, and Wojciech Pieczynski. Optimal filter approximations in conditionally Gaussian pairwise Markov switching models. *IEEE trans. on Automatic*

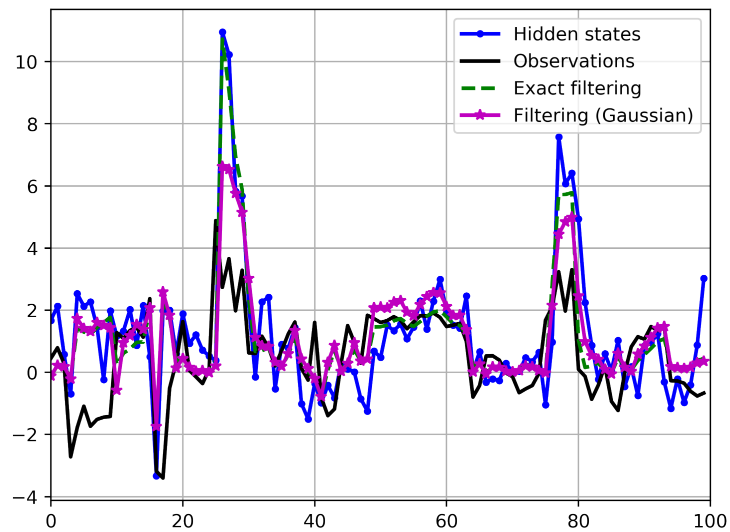


Figure 2: Trajectory example (100 samples).

Control, 60(4):1104–1109, 2015.

2. Armand Kodjo Atiampo and Georges Laussane Loum. Unsupervised image segmentation with pairwise Markov chains based on nonparametric estimation of copula using orthogonal polynomials. *International Journal of Image and Graphics*, 16(4):1650020, 2016.
3. Nicolas Brunel and Wojciech Pieczynski. Unsupervised signal restoration using hidden Markov chains with copulas. *Signal processing*, 85(12):2304–2315, 2005.
4. Barbara Choroś, Rustam Ibragimov, and Elena Permiakova. Copula estimation. *Copula theory and its applications*, pages 77–91, 2010.
5. Stéphane Derrode and Wojciech Pieczynski. Unsupervised data classification using pairwise Markov chains with automatic copulas selection. *Computational Statistics & Data Analysis*, 63:81–98, 2013.
6. Stéphane Derrode and Wojciech Pieczynski. Unsupervised classification using hidden Markov chain with unknown noise copulas and margins. *Signal Processing*, 128:8–17, 2016.
7. Ivan Gorynin, Stéphane Derrode, Emmanuel Monfrini, and Wojciech Pieczynski. Exact fast smoothing in switching models with application to stochastic volatility. In *Signal Processing Conference (EUSIPCO), 2015 23rd European*, pages 924–928. IEEE, 2015.
8. Ivan Gorynin, Stéphane Derrode, Emmanuel Monfrini, and Wojciech Pieczynski. Fast filtering in switching approximations of nonlinear Markov systems with applications to stochastic volatility. *IEEE trans. on Automatic Control*, 62(2):853–862, 2017.
9. Gunky Kim, Mervyn J Silvapulle, and Paramsothy Silvapulle. Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics and Data Analysis*, 51(6):2836–2850, 2007.
10. Kenneth Levenberg. A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics*, 2(2):164–168, 1944.
11. Roger B Nelsen. *An introduction to copulas*. Springer Science & Business Media, 2007.
12. Hideatsu Tsukahara. Semiparametric estimation in copula models. *Canadian Journal of Statistics*, 33(3):357–375, 2005.