ISSN: 2832-4293

Pan-American Journal of Mathematics 2 (2023), 6
https://doi.org/10.28919/cpr-pajm/2-6

ESTIMATING $MS - BLGARCH$ MODELS USING RECURSIVE METHOD

AHMED GHEZAL* AND IMANE ZEMMOURI

ABSTRACT. In this paper a new class of models is proposed for modeling nonlinear and stationary time series. This new class of models is referred to as the Markov-switching bilinear $GARCH$ $(MS - BLGARCH)$ models. In these models, the parameters are allowed to depend on an unobservable time-homogeneous and stationary Markov chain with finite state space. The statistical inference for these models is rather difficult due to the dependence on the whole regime path. We propose a recursive algorithm for parameter estimation in $MS - BLGARCH$. The proposed method is useful for long time series as well as for data available in real time. The main idea is to use the maximum likelihood estimation (MLE) method and from this develop a recursive Expectation-Maximization (EM) algorithm.

1. INTRODUCTION

Since the seminal works by Hamilton [17], Markov-switching models (MSM) have received a growing interest and becomes an appealing tool for the modelling of business cycles (as originally proposed by Hamilton [17]) and continue to gain popularity especially in financial data because of their ability to model time series in which we can observe break, turning, or change points from where on the series seems to follow a different regime than before. A discrete-time MSM is a bivariate random process $((\varepsilon_t, s_t), t \in \mathbb{Z})$, $\mathbb{Z} := \{0, \pm 1, \pm 2, ...\}$ such that (i): $(s_t, t \in \mathbb{Z})$ is an unobservable (referred henceforth as "regime"), finite state space, discrete-time and homogeneous Markov chain and (ii): the conditional distribution of ε_k given $\{(\varepsilon_{t-1}, s_t), t \leq k\}$ depends on $\{(\varepsilon_{t-1}, s_k), t \leq k\}$ only. The changes in regime can be abrupt, and they occur frequently or occasionally depending on the transition probability of the chain. So some locally (i.e., in each "regime") linear or nonlinear representations were investigated in order to capture the probabilistic and statistical properties of such models. For instance, $MS - ARMA$: Francq and Zakoïan [10] and Stelzer [26]; $MS-$nonlinear $ARMA$: Lee [21] and Yao; $MS - GARCH$: Francq and Zakoïan [9]; Hass et *al.,* [16]; Liu [22]; $MS - DAR$: Ghezal [14] among others. The $MS - BLGARCH$ model generate series with a much more flexible dependence structure than in standard $BLGARCH$ specifications proposed by Storti and Vitale (c.f., [27]) is a generalization of the classical $MS - GARCH$ model studied by Cai (c.f., [3]), Hamilton and Susmel (c.f., [18]) and Francq et al (c.f., [11]), thereby explain the presence of leverage effects in the volatility dynamics. The so called leverage effect relates on the different effects that past returns of the same magnitude, but of opposite sign, may have on the current market volatility. In the last decades enormous empirical evidence has confirmed the existence of asymmetric effects in the conditional variance

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCES, UNIVERSITY CENTER OF MILA, ALGERIA
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ANNABA, ELHADJAR 23, ANNABA, ALGERIA
E-mail addresses: a.ghezal@centre-univ-mila.dz, imanezemmouri25@gmail.com.
Submitted on Feb. 05, 2023.
2020 *Mathematics Subject Classification.* 62F12, 62M05.
Key words and phrases. bilinear and $GARCH$ processes, recursive estimation, EM-algorithm, Markov-switching.
*Corresponding author.

of financial asset. For shares these effects are usually linked to the so called leverage effect. The experimental results have motivated a remarkably long series of papers, suggest amendments of the $MS-GARCH$ model in order to allow for asymmetric volatility dynamics. Recently, evidence has been provided suggesting the presence of similar effects in the dynamics of conditional correlations. Similarly to the conditional variance, the asymmetry is due to the presence of a significant relationship linking the level of the conditional correlation between two assets to the signs of past returns on both assets. $MS-BLGARCH$ models are currently applied in many fields, in both macroeconomics and finance, can also be considered as an extension of the hidden Markov model (HMM), which is popular in various fields such as engineering, genetic biology and statics. We say that a second order $\mathbb{R}$-valued process $(\varepsilon_t, t\in\mathbb{Z})$ defined on some probability space $(\Omega,\Im,P)$ has a general Markov-switching bilinear $GARCH$ representation (denoted by $MS-BLGARCH\,(p,q,s)$) if it is a solution of the following stochastic difference equation

$$\left\{\begin{array}{l}\varepsilon_t=h_te_t\\ h_t^2=\alpha_0(s_t)+\sum\limits_{i=1}^{q}\alpha_i(s_t)\varepsilon_{t-i}^2+\sum\limits_{i=1}^{p}\beta_i(s_t)h_{t-i}^2+\sum\limits_{i=1}^{s}\gamma_i(s_t)\varepsilon_{t-i}h_{t-i},\ t\in\mathbb{Z}\end{array}\right., \tag{1.1}$$

where ε_t is an observed time series, $(e_t, t\in\mathbb{Z})$ is an independent and identically distributed $(i.i.d.)$ sequence of random variables defined on the same probability space $(\Omega,\Im,P)$ with $E\left\{\log^+|e_t|\right\}<+\infty$ where $\log^+x=\max\{0,\log x\}$, $x>0$ and p,q,s are non negative integers with $s=\min(p,q)$. The functions $\alpha_i(s_t)$, $\beta_i(s_t)$ and $\gamma_i(s_t)$* depend upon a Markov chain $(s_t, t\in\mathbb{Z})$ subject to the following assumption:

Assumption 1. *The Markov chain $(s_t,t\in\mathbb{Z})$ is stationary, irreducible, finite state space $\mathbb{S}=\{1,...,d\}$, $n-$step transition probabilities matrix $\mathbb{P}^n=\left(p_{ij}^{(n)},(i,j)\in\mathbb{S}\times\mathbb{S}\right)$ where $p_{ij}^{(n)}=P\left(s_t=j|s_{t-n}=i\right)$ with one-step transition probability matrix $\mathbb{P}:=(p_{ij},(i,j)\in\mathbb{S}\times\mathbb{S})$ where $p_{ij}:=p_{ij}^{(1)}=P\left(s_t=j|s_{t-1}=i\right)$ for $i,j\in\mathbb{S}$, and initial stationary distribution $\Pi=(\pi(1),...,\pi(d))'$ where $\pi(i)=P\left(s_t=i\right)$, $i=1,...,d$ such that $\Pi'=\Pi'\mathbb{P}$. In addition, we assume that e_t and $\{(\varepsilon_{s-1},s_t),s\le t\}$ are independent.*

The estimation of parameters of finite parameter time series models is an important part of time series analysis. MLE criterion offers as an indicator in estimation when the unknown parameters are deterministic. However in a lot of cases the received data does not provide complete information necessary for such maximizing. In particular, parameter estimation is generally performed through the iterative EM-algorithm introduced by Dempster and his coauthors [6], is a popular tool for MLE. The common strand to problems where this approach is applicable is a notion of incomplete data, which includes the conventional sense of missing data but is much broader than that. The EM algorithm demonstrates its strength in situations where some hypothetical experiments yields complete data that are related to the parameters more conveniently than the measurements are. However, for fairly large series or when data are sequentially available in real time, the latter algorithm entails a large bookkeeping due to its iterative nature and then may be very computationally expensive. Many applications of the EM algorithm to receiver design are also presented in Nelson [23] and Georghiades [12]. The model parameters are estimated using a novel technique based on the recursive (also sometimes called adaptive or online) MLE. The adjective recursive refers to the idea of computing estimates of model parameters, without storing the data and by continuously updating the estimates as more observations become available. In such a situation, recursive procedures that update parameter estimates so that the computational complexity does not depend on the sample size are appealing. The proposed method which is useful for long time series as well as for data available in real time. Traditional applications of recursive algorithms involve situations in which the data cannot be stored, due to its volume and rate of sampling as in real time signal processing or stream mining. Recursive algorithms are

*For $\gamma_i(.)<0$, a positive quantity will be added to h_t^2 if $\varepsilon_{t-i}<0$ while the same quantity will be subtracted if $\varepsilon_{t-i}>0$.

often more efficient i.e., converging faster towards the target parameter value. The main motivation for this restriction is to stick to computationally simple iterations which is an essential requirement of successful recursive methods. A common approach to the recursive estimation problem is to base the estimation on suboptimal modifications of Kalman filtering techniques. The basic idea in the research is to use the MLE method and from this develop a recursive EM algorithm. Recursive EM algorithm is more attractive when it's maximization step can be done analytically in a recursive manner. Recursive EM algorithm have been proposed by many authors in different context of missing data problems and latent variable models ones: Titterington [28], Sato [25], Aknouche [2], Cappé [4] and Holst [20]. In particular, we mention the words of Holst and Lindgren [19] and Collings et al [5] for hidden Markov models.

The $MS-BLGARCH\,(p,q,s)$ model includes as special cases several classes of interesting models having been investigated in the literature, indeed:

- **(i):** Standard $BLGARCH\,(p,q,s)$ models: These models are obtained by assuming constant the functions $\alpha_i(.)$, $\beta_i(.)$ and $\gamma_i(.)$ in (1.1) or equivalently by assuming that the chain (s_t) has a single regime (e.g., G.Storti, C. Vitale., [27]).
- **(ii):** Some classes of $MS-GARCH\,(p,q)$: These models are obtained by assuming $\gamma_i(s_t)=0,\forall i=1,...,s$ (e.g., Abramson and Cohen [1], Francq and Zakoïan [9] and Liu [22]).
- **(iii):** Markov-switching $ARMA$ models $(MS-ARMA\,(\max(p,q),p))$: These models are obtained by assuming $\gamma_i(s_t)=0,\forall i=1,...,s$ (e.g., Francq and Zakoïan [10]).
- **(iv):** Some classes of periodic $GARCH$ model: These models are obtained by setting $s_t=\sum\limits_{k=1}^{d}k\mathbb{I}_{\triangle(k)}(t)$ and $\gamma_i(s_t)=0,\forall i=1,...,s$ where $\triangle(k)$ denoting the set of indices corresponding to regime k at time t (see, for example, the following periodic models [13], [15]).

An overview of the paper is as follows. bilinear $GARCH$ models with Markov switching coefficients are introduced in first section. Second, recursive MLE is discussed and from this develop a recursive EM algorithm.

2. The estimation in models $MS-BLGARCH$

Let $m_{x,h}=max\{p;q;s\}$, we assume $\left\{\varepsilon^2_{-m_{x,h}+1},...,\varepsilon^2_0,\varepsilon^2_1,...,\varepsilon^2_n\right\}$ observation process generated from model (1.1) stationary, causal. Therefore suppose that $(e_t,t\in\mathbb{Z})$ is normally distributed with mean zero and variance σ^2. Note $\underline{\theta}:=(\underline{\alpha}',\underline{\beta}',\underline{\gamma}',\underline{p}',\sigma)'$ with $\underline{\alpha}':=\left(\underline{\alpha}'_0,\underline{\alpha}'_1,...,\underline{\alpha}'_{m_{x,h}}\right)$, $\underline{\beta}':=\left(\underline{\beta}'_1,...,\underline{\beta}'_{m_{x,h}}\right)$, $\underline{\gamma}':=\left(\underline{\gamma}'_1,...,\underline{\gamma}'_{m_{x,h}}\right)$, $\underline{p}':=\left(\underline{p}'_1,...,\underline{p}'_d\right)$ and $\underline{\alpha}'_i:=(\alpha_i(j),1\le j\le d)$, $\underline{\beta}'_i:=(\beta_i(j),1\le j\le d)$, $\underline{\gamma}'_i:=(\gamma_i(j),1\le j\le d)$, $\underline{p}'_i:=(p_{ij},1\le j\le d)$, $\underline{\theta}$ which belong to parameter space Θ, a subset of the Euclidean space. The true parameter value will be denoted $\underline{\theta}_0$ and assumed to belong to Θ. Given $\underline{\varepsilon}_0:=\left\{\varepsilon^2_{-m_{x,h}+1},...,\varepsilon^2_0,h^2_{-m_{x,h}+1},...,h^2_0\right\}$, conditional likelihood with respect to the measurement $d(\lambda\otimes N)$ (where λ denotes the Lebesgue measure and N is the counting of $\mathbb{S}$) is defined for all $\underline{\theta}\in\Theta$ by

$$L_n(\underline{\theta})=p\left(\underline{Z}_n=(\varepsilon^2_1,...,\varepsilon^2_n)\middle|\,\underline{\varepsilon}_0,s_0;\underline{\theta}\right)=\sum_{x\in\mathbb{S}}p_{s_0x}p\left(\underline{Z}_n\middle|\,\underline{\varepsilon}_0,s_1=x;\underline{\theta}\right).$$

2.1. **Recursive** MLE.

A recursive estimation $\underline{\theta}_n$ of the parameter $\underline{\theta}$ based on the first n observations of $(\varepsilon^2_t,t\in\mathbb{Z})$, as an iterative approach, is given, using the Newton-Raphson method, by

$$\underline{\theta}_n=\underline{\theta}_{n-1}+\frac{1}{n}K_{n-1}\nabla_{\underline{\theta}_{n-1}}(\varepsilon_n),\tag{2.1}$$

where $\nabla_{\underline{\theta}}(\varepsilon)$ is a score vector function, defined by $\nabla_{\underline{\theta}}(\varepsilon)=\left(\frac{\partial\ln p(\varepsilon;\theta)}{\partial\theta_j},\ j=1,...,s\right)$ with $s=d(3m_{x,h}+d+1)+1$ and K_n is the inverse of the information matrix, i.e., $K^{-1}_{n-1}=I(\underline{\theta}_{n-1})=\nabla^2_{\underline{\theta}_{n-1}}$, determined by

$$I\left(\underline{\theta}_{n-1}\right)=E\left(\nabla_{\underline{\theta}_{n-1}}\left(\varepsilon_{n}\right)\nabla'_{\underline{\theta}_{n-1}}\left(\varepsilon_{n}\right)\right).$$

Recursive MLE with this $I\left(\underline{\theta}_{n-1}\right)$ are treated by Fabian [7] and they are proven to be asymptotically normal under suitable regularity conditions

$$\sqrt{n}\left(\underline{\theta}_{n}-\underline{\theta}\right)\xrightarrow{d}\mathcal{N}\left(\underline{0},I^{-1}\left(\underline{\theta}\right)\right).$$

Computation of $I\left(\underline{\theta}_{n-1}\right)$ often requires numerical integration and it is thus cumbersome. Moreover, the prohibitive calculations. So we shall instead use the inverse of the observed information matrix, i.e.,

$$K_{n-1}^{-1}=\frac{1}{n-1}\sum_{k=1}^{n-1}\nabla_{\underline{\theta}_{k-1}}\left(\varepsilon_{k}\right)\nabla'_{\underline{\theta}_{k-1}}\left(\varepsilon_{k}\right).$$

The matrix K_{n-1} can be computed recursively by means of the matrix inversion lemma (see for example Rao [24] Section 1b). Writing $\nabla_{n}=\nabla_{\underline{\theta}_{n-1}}\left(\varepsilon_{n}\right)$ we have

$$K_{n-1}=\frac{n-1}{n-2}\left(K_{n-2}^{-1}+\frac{1}{n-2}\nabla_{n-1}\nabla'_{n-1}\right)^{-1}=\frac{n-1}{n-2}\left(K_{n-2}-\frac{K_{n-2}\nabla_{n-1}\nabla'_{n-1}K_{n-2}}{(n-2)+\nabla'_{n-1}K_{n-2}\nabla_{n-1}}\right).$$

2.2. **Recursive MLE in $MS-BLGARCH$.** We shall now examine the structure of the recursive algorithm (2.1) when $\left(\varepsilon_{t}^{2},t\in\mathbb{Z}\right)$ is a $MS-BLGARCH$. Define the filtering probabilities

$$\pi_{j}\left(n\right)=p\left(s_{n}=j|\,\underline{Z}_{n}\right)\text{ and }\pi_{ij}\left(n\right)=p\left(s_{n-1}=i,\ s_{n}=j|\,\underline{Z}_{n}\right),n\geq1\text{ with }\pi_{j}\left(n\right)=\sum_{i=1}^{d}\pi_{ij}\left(n\right),$$

and the indicator function $1_{ij}\left(n\right)=1_{\{s_{n-1}=i,\,s_{n}=j\}}$, $\underline{s}_{n}=\left(s_{1},...,s_{n}\right)$ and write the conditional density of

$\left(X_{n},s_{n}\right)$ given $\left(\underline{X}_{n-1},\underline{s}_{n-1}\right)$, i.e.,

$$p\left(\varepsilon_{n},s_{n}|\,\underline{\varepsilon}_{n-1},\underline{s}_{n-1},s_{0};\underline{\theta}\right)=p\left(s_{n}|\,s_{n-1;}\underline{\theta}\right)\ p\left(\varepsilon_{n}|\,\underline{\varepsilon}_{n-1},s_{n};\underline{\theta}\right).\tag{2.2}$$

Let $\mathcal{L}_{n}:=p\left(\underline{Z}_{n},\underline{s}_{n}|\,\underline{\varepsilon}_{0},s_{0};\underline{\theta}\right)$ denote the simultaneous density of the complete sequence $\left(\underline{Z}_{n},\underline{s}_{n}\right)$, then $\log\mathcal{L}_{n}=\log\mathcal{L}_{n-1}+l_{n}$ where

$$l_{n}=\sum_{i,j=1}^{d}1_{ij}\left(n\right)\left(\log p_{ij}+\log p\left(\varepsilon_{n}|\,\underline{\varepsilon}_{n-1},s_{n}=j;\underline{\theta}\right)\right).$$

A nature generalization of the recursive MLE procedure (2.1) to a recursive EM algorithm is to use the conditional expectation of the derivative of l_{n} given the observed $\underline{Z}_{n}$ as score function, i.e., $\nabla_{\underline{\theta}}\left(\varepsilon_{n}|\,\underline{\varepsilon}_{n}\right)=E\left\{\left.\frac{\partial l_{n}}{\partial\underline{\theta}}\right|\underline{\varepsilon}_{n}\right\}$.

Theorem 2.1. *In a $MS-BLGARCH$ model the score function for an observation $\underline{Z}_{n}$ is*

$$\nabla_{\underline{\theta}}\left(\varepsilon_{n}|\,\underline{\varepsilon}_{n}\right)=\sum_{i,j=1}^{d}\pi_{ij}\left(n\right)\frac{\partial p\left(\varepsilon_{n},s_{n}=j|\,\underline{\varepsilon}_{n-1},s_{n-1}=i,\underline{s}_{n-2},s_{0};\underline{\theta}\right)}{\partial\underline{\theta}},\tag{2.3}$$

and

$$\pi_{ij}\left(n\right)=\pi_{i}\left(n-1\right)p_{ij}\ p\left(\varepsilon_{n}|\,\underline{Z}_{n-1},s_{n-1}=j;\underline{\theta}\right)\frac{p\left(\underline{Z}_{n-1}|\,\underline{\theta}\right)}{p\left(\underline{Z}_{n}|\,\underline{\theta}\right)}.$$

Summarizing, in the recursive procedure

$$\underline{\theta}_{n}=\underline{\theta}_{n-1}+\frac{1}{n}K_{n-1}E\left\{\left.\frac{\partial l_{n}}{\partial\underline{\theta}}\right|\underline{\varepsilon}_{n}\right\},$$

we shall use the score function (2.3) where the filtering probabilities are calculated using the parameter estimate $\underline{\theta}_{n-1}$.

Example 2.1. *In this example we shall investigate the performance of the recursive EM algorithm for $MS-BLGARCH$, we shall take*

$$\begin{cases} \varepsilon_t = h_t e_t \\ h_t^2 = 1,5\varepsilon_{t-1}^2 + 0,72h_{t-1}^2 + 0,2\varepsilon_{t-1}h_{t-1},\ \textit{if}\ s_t = 1 \\ h_t^2 = 1,7\varepsilon_{t-1}^2 + 0,52h_{t-1}^2 + 1,3h_{t-2}^2 + 0,15\varepsilon_{t-1}h_{t-1},\ \textit{if}\ s_t = 2 \end{cases},$$

and $(e_t, t \in \mathbb{Z})$ sequence of iid $\mathcal{N}(0;1)$, denote the transition matrix $\mathbb{P} = \begin{pmatrix} 1-p_{12} & p_{12} \\ 1-p_{22} & p_{22} \end{pmatrix}$. In these example $p_{12} = 0,08$ and $p_{22} = 0,92$ $(p_{12} = p_{21})$. The model parameters are denoted by $\underline{\theta} = \left(\underline{\theta}_1', \underline{\theta}_2', p_{12}, p_{22}\right)'$ where $\underline{\theta}_i = (\alpha(i), \beta_1(i), \beta_2(i), \gamma(i))'$, $i = 1, 2$, we have

$$\log p\left(\varepsilon_n | \underline{\varepsilon}_{n-1}, s_n = j; \underline{\theta}\right) = -\frac{1}{2}\log 2\pi - \frac{1}{2}\left(\log h_t^2(j) + \frac{\varepsilon_t^2(j)}{h_t^2(j)}\right),$$

where $\varepsilon_t(j) = h_t(j)e_t$, $h_t^2(j) = \alpha(j)\varepsilon_{t-1}^2 + \beta_1(j)h_{t-1}^2 + \beta_2(j)h_{t-2}^2 + \gamma(j)\varepsilon_{t-1}h_{t-1}$, and

$$\frac{\partial \log p\left(\varepsilon_n | \underline{\varepsilon}_{n-1}, s_n = j; \underline{\theta}\right)}{\partial \underline{\theta}_j} = \begin{pmatrix} -\frac{\varepsilon_{t-1}^2}{h_t(j)} \\ -\frac{h_{t-1}^2}{h_t(j)} \\ -\frac{h_{t-2}^2}{h_t(j)} \\ -\frac{\varepsilon_{t-1}h_{t-1}}{h_t(j)} \end{pmatrix}.$$

From theorem previous the score function is given by

$$\nabla_{\underline{\theta}}\left(\varepsilon_n | \underline{\varepsilon}_n\right) = \begin{pmatrix} \pi_1(n)\frac{\partial \log p\left(\varepsilon_n | \underline{\varepsilon}_{n-1}, s_n=1; \underline{\theta}\right)}{\partial \underline{\theta}_1} \\ \pi_2(n)\frac{\partial \log p\left(\varepsilon_n | \underline{\varepsilon}_{n-1}, s_n=2; \underline{\theta}\right)}{\partial \underline{\theta}_2} \\ \pi_{12}(n)p_{12}^{-1} - \pi_{11}(n)(1-p_{12})^{-1} \\ \pi_{22}(n)p_{22}^{-1} - \pi_{21}(n)(1-p_{22})^{-1} \end{pmatrix},$$

which is less convenient due to the unpleasant denominator being unbounded near the boundaries $p_{ij} = 0$ or 1.

2.3. **The choice of initial values.** To begin a recursive procedure, one needs initial values for each of the estimates $\underline{\theta}_0$ and for the adaptive matrix K_0. By Holst et al [20], depending upon the a priori knowledge about the true values of the parameters, the algorithm should be initiated in different ways. In this article we consider only estimation situations where there exists fairly good knowledge about the true values of the parameters and in particular the adaptive matrix is treated in a conservative way during the first steps. The choice of K_{n-1} seems to be great importance, since the optimal, true information matrix can be almost singular. In order to calculate an approximate value of the information matrix in the model defined by the initial value $\underline{\theta}_0$ we simulated $N = 10000$ steps in the $MS-BLGARCH$ and estimated the information matrix by

$$I_N^{-1}\left(\underline{\theta}_0\right) = \frac{1}{N}\sum_{k=1}^{N} \nabla_{\underline{\theta}_0}\left(\varepsilon_k | \underline{\varepsilon}_k\right) \nabla_{\underline{\theta}_0}'\left(\varepsilon_k | \underline{\varepsilon}_k\right).$$

For the ergodic Markov regime this will lead to an approximately correct information matrix for the initial model. In order to avoid wild fluctuations of the estimates at the early stages we modified the procedure in the following ways.

(i) The initial matrix K_0 was controlled by a ridge parameter α. if $\frac{\gamma_{\min}}{\gamma_{\max}} < \alpha$ ($\gamma_{\min}$ and $\gamma_{\max}$ is the smallest and largest eigenvalue for the information matrix) we added a fixed proportion α of the maximal eigenvalue to the diagonal elements of $I\left(\underline{\theta}_0\right)$, i.e., we used as an initial K_0 matrix

$$K_0 = \left(I\left(\underline{\theta}_0\right) + \alpha\gamma_{\max}\mathbf{1}\right)^{-1},$$

where **1** is the unit matrix.

(ii) The updating of the adaptive matrix K_{n-1} by means of formula (2.2) was modified by addition of a constant n_0 to the number of observations, i.e.,

$$K_{n-1} = \frac{n+n_0-1}{n+n_0-2}\left(K_{n-2} - \frac{K_{n-2}\nabla_{n-1}\nabla'_{n-1}K_{n-2}}{(n+n_0-2)+\nabla'_{n-1}K_{n-2}\nabla_{n-1}}\right).$$

(iii) During the early stages of the recursion, $n \leq n_r$, the adaptive matrix K_{n-1} was held constant and equal to K_0. In the meantime K_{n-1} was updated and for $n > n_r$ used as the adaptive matrices.

Acknowledgements We thank the editor-in-chief of the journal, the associate editor, and the two anonymous referees for their constructive comments and very useful suggestions and remarks which were most valuable for improvement of the final version of the paper.

Conflicts of interest We have no conflicts of interest to disclose.

References

[1] A. Abramson, I. Cohen, On the stationarity of Markov-switching $GARCH$ processes, Econometric Theory, 23 (2007), 485-500.

[2] A. Aknouche, Recursive online EM estimation of mixture autoregressions, J. Stat. Comp. Sim. 83 (2013), 370-383.

[3] J. Cai, A Markov model of switching regime $ARCH$. J. Bus. Econ. Stat. 12 (1994), 309-316.

[4] O. Cappé, E. Moulines, Online expectation maximization algorithm for latent data models, J. R. Stat. Soc. B. 71 (2009), 593-613.

[5] I. Conllings, V. Krishnamurthy, J. B. Moore, Online identification of hidden Markov models via recursive prediction error techniques, IEEE Trans. Signal Process. 42 (1994), 3535-3539.

[6] A. P. Dempster, N. M. Laid, D. B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, Stat. Soc. B. 39 (1977), 1-38.

[7] V. Fabian, On asymptotically efficient recursive estimation, Ann. Stat. 6 (1978), 854-866.

[8] C. Francq, M. Roussignol, Ergodicity of autoregressive processes with Markov-switching and consistency of the maximum likelihood estimator, Statistics, 32 (1998), 151-173.

[9] C. Francq, J. M. Zakoïan, $\mathbb{L}^2$-structures of standard and switching-regime $GARCH$ models, Stoch. Processes Appl. 115 (2005), 1557-1582.

[10] C. Francq, J. M. Zakoïan, Stationarity of multivariate Markov-switching $ARMA$ models, J. Econometrics. 102 (2001), 339-364.

[11] C. Francq, M. Roussignol, On white noises driven by hidden Markov chains, J. Time Series Anal. 18 (1997), 553-578.

[12] C. N. Georghiades, J. C. Han, Sequence estimation in presence of random parameters via the EM algorithm, IEEE trans. Commun. 45 (1997), 300-308.

[13] A. Ghezal, $QMLE$ for periodic time-varying asymmetric $\log GARCH$ models, Commun. Math. Stat. 9 (2021), 273-297.

[14] A. Ghezal, A doubly Markov switching AR model: Some probabilistic properties and strong consistency, J. Math. Sci. (2023). https://doi.org/10.1007/s10958-023-06262-y.

[15] A. Ghezal, Asymptotic inference for periodic time-varying bivariate Poisson $INGARCH(1,1)$ processes, J. Stat. Appl. Prob. Lett. 10 (1) (2023), 77-82.

[16] M. Haas, S. Mittnik, M. S. Paolella, A new approach to Markov-switching $GARCH$ models, J. Financial Econometrics. 2 (2004), 493-530.

[17] J. D. Hamilton, A new approach to the economic analysis of nonstationary time series and the business cycle, Econometrica. 57 (1989), 357-384.

[18] J. D. Hamilton, R. Susmel, Autoregressive conditional heteroskedasticity and changes in regime, J. Econ. 64 (1994), 307-333.

[19] U. Holst, G. Lindgren, Recursive estimation in mixture models with Markov regime, IEEE Trans. Inform. Theory. 37 (1991), 1683-1690.

[20] U. Holst, G. Lindgren, J. Holst, M. Thuvesholmen, Recursive estimation in switching autoregressions with a Markov regime, J. Time Ser. Anal. 15 (1994), 489-506.

[21] O. Lee, Probabilistic properties of a nonlinear $ARMA$ process with Markov switching, Comm. Stat. Theory Meth. 34 (2005), $193-204$.

[22] J.C. Liu, Stationarity of a Markov-switching $GARCH$ model, J. Financial Econometrics. 4 (2006), 573-593.

[23] L. N. Nelson, H. V. Poor, Iterative multiuser receivers for CDMA channels: An EM-based apprpach, IEEE trans. Comm. 44 (1996), 1700-1710.

[24] C. Rao, Linear Statistical inference and its applications, New York: Wiley, 1973.

[25] M. Sato, S. Ishii, Online EM algorithm for the normalized Gaussian network, Neural Comp. 12 (2000), 407-432.

[26] R. Stelzer, On Markov-switching $ARMA$ processes: Stationarity, existence of moments and geometric ergodicity, Econometric Theory. 25 (2009), 43-62.

[27] G. Storti, Vitale, $BL-GARCH$ models and asymmetries in volatility, Stat. Meth. Appl. 12 (2003), 19-39.

[28] D. M. Titterington, Recursive parameter estimation using incomplete data, J. R. Stat. Soc. B. 46 (1984), 257-267.