SeMA Journal (2024) 81:413–427
https://doi.org/10.1007/s40324-023-00329-1

On the Markov-switching autoregressive stochastic volatility processes

Ahmed Ghezal[1] · **Imane Zemmouri**[2]

Received: 11 March 2023 / Accepted: 9 May 2023 / Published online: 29 May 2023

Abstract

Regime switching models are able to capture clustering effects, nonlinearities in time series and jumps in volatility. In the present paper, we propose a broad class of Markov-switching AutoRegressive Stochastic Volatility ($MSAR - SV$) models, in which the log −volatility follows a $p^{th} - MS-$autoregression. So, it can be seen as a replacement of the general $MS - GARCH$ model. This parameterization draws a lot of attention in modeling structural changes in dependent data. The parameters of the log − volatility are expressed as a function of a homogeneous Markov chain with a finite state space. The primary goal of the proposed model is to confer it a change driven by a Markov chain in order to capture by the habitual changing behavior of volatility due to economic forces, as the discrete shift in volatility due to abrupt abnormal events. Several probabilistic properties of $MSAR - SV$ models have been obtained, especially, strictly (resp. second-order) stationary, causal and ergodic solution, geometric ergodicity, and computation of higher-order moments. Moreover, we derive the expression of the covariance function of the squared (resp. powers) process. Consequently, the logarithm squared (resp. powers) process admits an $ARMA$ representation. Then we provide the limit theory for quasi-maximum likelihood estimator ($QMLE$), and, in addition, establish the strong consistency of this estimator. Finally, we present a simulation study on the performance of the proposed estimation method.

Keywords Stochastic volatility · Markov-switching autoregression · Strictly (second-order) stationarity · QMLE · Geometric ergodicity

Mathematics Subject Classification 60G10 · 62F12

✉ Ahmed Ghezal
a.ghezal@centre-univ-mila.dz

Imane Zemmouri
imanezemmouri25@gmail.com

1 Department of Mathematics and Computer Sciences, Abdelhafid Boussouf University, Center of Mila, Mila, Algeria

2 Department of Mathematics, University of Annaba, Elhadjar 23, Annaba, Algeria

1 Introduction

Markov-switchingmodels *(MSM)* have attracted a lot of research attention and have become a robust tool for modeling and describing asymmetric business cycles in the econometric literature (see., Hamilton [14]). Such models go ahead to earn more and more popularity, especially in financial data. In the same context, these models have been chosen because of their high flexibility in capturing stability and/or asymmetric effects on volatility shocks and their competence in modeling time series. A linear or nonlinear MSM was developed by many authors (for example, $MS - ARMA$: Cavicchioli [3–5], nonlinear $MS - ARMA$: Stelzer [17], $MS - GARCH$: Hass et al. [13] and Cavicchioli [6], $MS - BL$: Ghezal et al. [1] and [12] among others). The various authors have pointed out that means identifying occasional switching in the parameter values may provide more appropriate modeling of volatility. In our paper, we alternatively present a $MS-$**A**uto**R**egressive **S**tochastic **V** olatility ($MSAR - SV$) process in which the process which follows locally (i.e., each regime) an autoregressive stochastic volatility ($AR - SV$) representation. In this model the log $-$volatility process follows a $p^{th} - MS-$autoregression, where the coefficients depend on a Markov chain. The latter was considered in the literature as the best alternative to the $MS - ARCH-$type models, where the volatility is driven by an exogenous innovation. The model presented in this paper is a natural extension of the $MS - SV$ model of So et al. [16], so that the observed process is described by heavy-tail innovations (see also Casarin [2], for a more qualitative discussion). The main reason for choosing a $MSAR - SV$ model is that it significantly promotes the predicting vigor of the $AR - SV$ model, does a perfect act in capturing the leading events affecting the oil market, and, in addition, simultaneously captures the usual changing behavior of volatility due to economic forces as well as the sudden discrete shift in volatility due to sudden abnormal events. (see So et al. [16], for a more qualitative discussion). To evaluate the $MSAR - SV$ model, its performance, in terms of goodness-of-fit and forecasting power, is compared to the standard $AR - SV$ model. Firstly, the probabilistic properties of the $MSAR-SV$ model are investigated. For this, we afford the sufficient and necessary assumptions to ensure the existence of a stationary solution, observing that the $MSAR - SV$ coefficients related to the Markov chain can breach the usual stationary assumptions of standard $AR - SV$ models. Secondly, this paper aims to analyse the strong consistency of the $QMLE$ of $MSAR - SV$ models. Before we go ahead, we present some symbols:

1.1 Symbols

Through the paper, the following symbols are used.

- $I_{(.)}$ is the square matrix whose every principal diagonal entry is equal to 1, and the remaining entries are equal to 0, $O_{(p,m)}$ is the $p \times m$ matrix such that all entries are zeros, $\underline{H}' := \left(I_{(1)}, \underline{O}'_{(p-1)}\right)$.
- $\log \underline{V}$, $\exp \underline{V}$ and $\underline{V}^{\frac{1}{2}}$ denote the vectors formed by the logarithm, exponential and square root of the entries of the vector $\underline{V}$, respectively, $diag\left(\underline{V}\right)$ denotes the diagonal matrix created by the entries of $\underline{V}$. $\rho(A)$ is the spectral radius of a square matrix A.
- $\|.\|$ is any norm on $m \times n$-(resp. $m \times 1-$) matrices (resp. vectors). $\otimes$ is the Kronecker product, and $A\,(1) = a_{11}$ of $A = (a_{ij})$.
- $(s_t, t \in \mathbb{Z})$ is a stationary, irreducible and aperiodic Markov chain.

- $\mathbb{P}^n = \left(p_{ij}^{(n)}, (i,j) \in \mathbb{S} \times \mathbb{S}\right)$ is the $n-$step transition probability matrix, where $p_{ij}^{(n)} = P(s_t = j | s_{t-n} = i)$ with one-step transition probability matrix $\mathbb{P} := \left(p_{ij}, (i,j) \in \mathbb{S} \times \mathbb{S}\right)$ where $p_{ij} := p_{ij}^{(1)} = P(s_t = j | s_{t-1} = i)$ for $i, j \in \mathbb{S} = \{1, ..., d\}$.
- $\underline{\Pi}' = (\pi(1), ..., \pi(d))$ is the initial stationary distribution, where $\pi(i) = P(s_0 = i)$, $i = 1, ..., d$, such that $\underline{\Pi}' = \underline{\Pi}'\mathbb{P}$.
- For any set of non-random matrices $A := \{A(i), i \in \mathbb{E}\}$, we note

$$\mathbb{P}^{(n)}(A) = \begin{pmatrix} p_{11}^{(n)} A(1) & \dots & p_{d1}^{(n)} A(1) \\ \vdots & \dots & \vdots \\ p_{1d}^{(n)} A(d) & \dots & p_{dd}^{(n)} A(d) \end{pmatrix}, \ \underline{\Pi}(A) = \begin{pmatrix} \pi(1)A(1) \\ \vdots \\ \pi(d)A(d) \end{pmatrix},$$

with $\mathbb{P}^{(1)}(A) = \mathbb{P}(A)$.

The rest of the paper is ordered as follows. In Sect. 2, we introduce the $MSAR - SV$ model and present several probabilistic properties of this model, specifically the strictly and second (resp. higher)-order stationary solution of $MSAR-SV$. Then the autocovariance functions of the squared and powers processes are derived. As a result, we find that the logarithm squared (resp. powers) process admits an $ARMA$ representation. We also provide here sufficient assumptions for the $MSAR - SV$ model to be geometrically ergodic and $\beta-$mixing. In Sect. 3, we propose the $QMLE$ for this model and derive the strong consistency. Simulation results are reported in Sect. 4.

2 *MSAR − SV* models

The Markov-switching autoregressive stochastic volatility model (denoted by $MSAR - SV(p)$) is given by

$$\begin{cases} \epsilon_t = \sqrt{h_t} e_t \\ \log h_t = a_0(s_t) + b_0(s_t)\eta_t + \sum_{i=1}^{p} a_i(s_t) \log h_{t-i} \end{cases}. \tag{2.1}$$

In Eq. (2.1), $\{(e_t, \eta_t), t \in \mathbb{Z}\}$ is an independent and identically distributed (**i.i.d.**) sequence of random vectors with mean $\underline{O}'_{(2)}$ and covariance matrix $I_{(2)}$. The functions $a_i(.)$, $i = 0, ..., p$ and $b_0(.)$ are related to the unobserved Markov chain $(s_t, t \in \mathbb{Z})$. We also suppose that (e_t, η_t) and $\{(\epsilon_{u-1}, s_t), u \leq t\}$ are independent. It is worth noting that h_t is conventionally called volatility. It is not the conditional variance of ϵ_t given its past informations up to time $t-1$ (this is justified by $E\left\{\epsilon_t^2 | \sigma - \{(\epsilon_u, s_u), u < t\}\right\} = E\{h_t | \sigma - \{(\epsilon_u, s_u), u < t\}\} \neq h_t$). The aim of this section is to show some of the most likely probabilistic properties of the $MSAR - SV$ model. As in much time-series, it is helpful to write Eq. (2.1) in an equivalent state-space representation to smooth the study. In this discussion we can write Eq. (2.1) in the MS multivariate stochastic volatility form

$$\begin{cases} \underline{\epsilon}_t = diag\left(\underline{H}_t^{\frac{1}{2}}\right)\underline{e}_t \\ \log \underline{H}_t = \Gamma(s_t) \log \underline{H}_{t-1} + \underline{\eta}_t(s_t) \end{cases}, \tag{2.2}$$

where

$$\underline{\epsilon}_t := \epsilon_t \underline{H}, \underline{e}_t := e_t \underline{H}, \underline{\eta}_t(s_t) := a_0(s_t)\underline{H} + b_0(s_t)\eta_t \underline{H}, \underline{H}'_t := \left(h_t, \dots, h_{t-p+1}\right),$$

and

$$\Gamma(s_t) := \begin{pmatrix} a_1(s_t) \dots a_p(s_t) \\ I_{(p-1)} \quad \underline{O}_{(p-1)} \end{pmatrix}.$$

So, the process $\left(\left(\log \underline{H}_t', s_t\right)', t \in \mathbb{Z}\right)$ is a Markov chain on $\mathbb{R}^p \times \mathbb{S}$. However, the study of the probabilistic properties of model (2.1) is easier and best through model (2.2). The second equation of Eq. (2.2) is the same as defined for the $D - MSAR$ model studied newly by Ghezal [11]. First, we get the following important result which implies strict stationarity.

Theorem 2.1 *Consider the MS-multivariate stochastic volatility model (2.2). Then*

1. *A sufficient condition for (2.2) to have a unique, strictly stationary, causal and ergodic solution; given by*

$$\epsilon_t = e_t \exp\left\{\frac{1}{2}\sum_{k=0}^{\infty}\left\{\prod_{j=0}^{k-1}\Gamma\left(s_{t-j}\right)\right\}(1)\left(a_0\left(s_{t-k}\right)+b_0\left(s_{t-k}\right)\eta_{t-k}\right)\right\} \tag{2.3}$$

which converges absolutely almost surely (a.s.) for all $t \in \mathbb{Z}$, is

$$\gamma_L(\Gamma) := \lim_{t\to\infty} E\left\{\frac{1}{t}\log\left\|\prod_{j=0}^{t-1}\Gamma\left(s_{t-j}\right)\right\|\right\} \overset{a.s}{=} \lim_{t\to\infty}\left\{\frac{1}{t}\log\left\|\prod_{j=0}^{t-1}\Gamma\left(s_{t-j}\right)\right\|\right\} < 0.$$

2. *Contrariwise, assume that $\left\{a_0(s_t)\underline{H}, b_0(s_t)\underline{H}, \Gamma(s_t)\right\}$ is controllable[1] and (2.2) has a strictly stationary solution. Then $\gamma_L(\Gamma) < 0$.*

Remark 2.1 Using Jensen's inequality, condition $E\left\{\left\|\prod_{j=0}^{t-1}\Gamma\left(s_{t-j}\right)\right\|\right\} < 1$ constitutes a sufficient condition for $\gamma_L(\Gamma) < 0$.

Example 2.1 Consider the $MSAR - SV(1)$ model. The sufficient condition is $\gamma_L(\Gamma) = \sum_{k=1}^{d}\pi(k)\log|a_1(k)| < 0$. In this case there exists a unique strictly stationary, causal and ergodic solution

$$\epsilon_t = e_t \prod_{k\geq 0}\sqrt{\exp\left\{\prod_{j=0}^{k-1} a_1\left(s_{t-j}\right)\left(a_0\left(s_{t-k}\right)+b_0\left(s_{t-k}\right)\eta_{t-k}\right)\right\}}.$$

Therefore, the local strict stationarity is not requisite, i.e., the presence of burst regimes (i.e., $\log|a_1(k)| > 0$) does not exclude the global strict stationarity. In the special case of $MSAR - SV(1)$ with two-regimes, we get

$$a_1(1) = a,\ a_1(2) = b \text{ and } \pi(1) = 3/4 \text{ (resp. } \pi(1) = 1/5).$$

The zone of strict stationarity is elucidated in Figure 1 below.

Other properties such as second-order stationarity and the existence of moments are clear and easy to obtain.

[1] The concept of controllability is defined in [1].

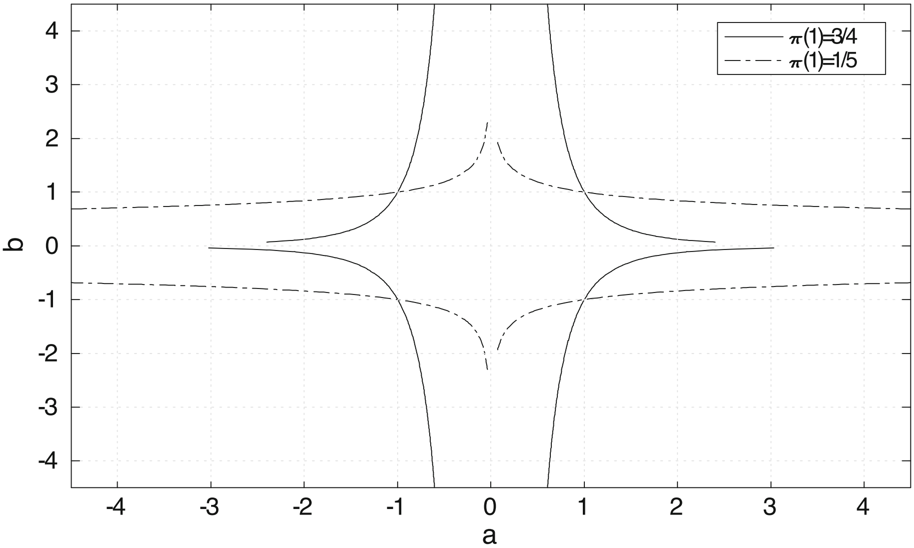

Fig. 1 The zones of strict stationarity for $MSAR-SV(1)$ model

Theorem 2.2 *Consider the MS-univariate stochastic volatility model (2.1) with MS-multivariate stochastic volatility model (2.2) and let $\Gamma^{(2)} := \left\{\Gamma^{\otimes 2}(k), k \in \mathbb{S}\right\}$. If*

$$\rho_{(2)} := \rho\left(\mathbb{P}(\Gamma^{(2)})\right) < 1. \tag{2.4}$$

Then Eq. (2.2) has a unique second-order stationary solution given by the Series (2.3), which converges absolutely a. s. and in $\mathbb{L}_2$. Furthermore, this solution is strictly stationary and ergodic.

Proof The outcome follows from second-order stationarity of the vector autoregression $\left(\log \underline{H}_t, t \in \mathbb{Z}\right)$ given by (2.2), which can be easily obtained by using the results of Ghezal et al. [1]. □

For this purpose, the explicit expressions of the moments up to second-order are shown in the following result

Proposition 2.1 *Consider the MS-univariate stochastic volatility model (2.1), if $\epsilon_t \in \mathbb{L}_2$, then*

1. $E\{\epsilon_t\} = 0$.
2. $\gamma_\epsilon(h) = E\{\epsilon_t \epsilon_{t-h}\}$

$$= \begin{cases} \sum\limits_{x_t, x_{t-1}, \ldots \in \mathbb{S}} \prod\limits_{k \geq 0} p_{x_{t-k-1} x_{t-k}} E\left\{\exp\left\{\left\{\prod\limits_{j=0}^{k-1} \Gamma\left(x_{t-j}\right)\right\}(1)\left(a_0\left(x_{t-k}\right)\right.\right.\right. \\ \left.\left.\left. +b_0\left(x_{t-k}\right)\eta_0\right)\right\}\right\} \text{ if } h = 0 \\ 0 \text{ otherwise} \end{cases}$$

. In the special case of normal innovations $(\eta_t, t \in \mathbb{Z})$, we obtain

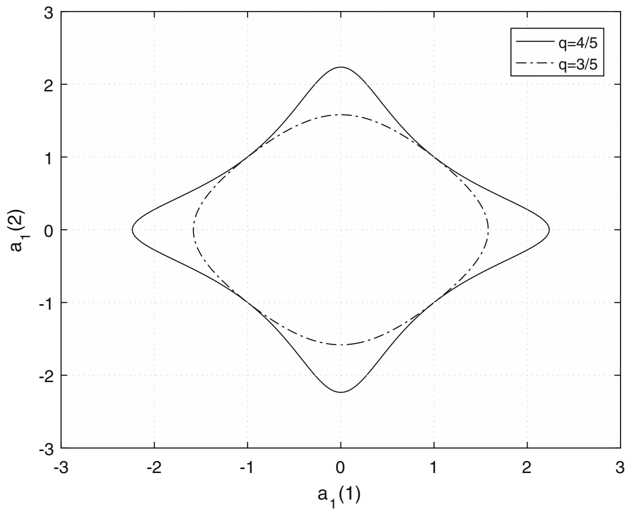

Fig. 2 Plots of the frontier curves $\rho\left(\mathbb{P}(\underline{a}_1^{(2)})\right) = 1$ for $MSAR - SV(1)$ model

$$\gamma_\epsilon(0) = \sum_{x_t, x_{t-1}, \ldots \in \mathbb{S}} \prod_{k \geq 0} p_{x_{t-k-1} x_{t-k}}$$
$$\exp\left\{\left\{\prod_{j=0}^{k-1} \Gamma\left(x_{t-j}\right)\right\}(1)\, a_0\left(x_{t-k}\right) + \frac{1}{2}\left(\left\{\prod_{j=0}^{k-1} \Gamma\left(x_{t-j}\right)\right\}(1)\right)^2 b_0^2\left(x_{t-k}\right)\right\}.$$

Proof Under the last condition, we can easily obtain the second-order moments and so the details are omitted. □

Example 2.2 Consider the $MSAR - SV(1)$ model. The condition (2.4) is reduced to $\rho\left(\mathbb{P}(\underline{a}_1^{(2)})\right) < 1$ where $\underline{a}_1^{(2)} := \left(a_1^2(k), k \in \mathbb{S}\right)'$. In particular, for two regimes with $p_{11} = 1 - q = p_{22}$, $p_{12} = q = p_{21}$, condition (2.4) is equivalent to the following two conditions

$$\begin{cases} (2q-1) \prod_{j=1}^{2} a_1^2(j) + (1-q) \sum_{j=1}^{2} a_1^2(j) < 1 \\ (1-q) \sum_{j=1}^{2} a_1^2(j) \leq 2 \end{cases}.$$

The region of second-order stationarity is shown in Figure 2 below.

Interesting now for assumptions guaranteeing the existence of higher-order moments for univariate-$MSAR - SV(p)$ having multivariate-$MSAR - SV(1)$ representation (2.2).

Remark 2.2 The odd-order moments of $(\epsilon_t, t \in \mathbb{Z})$ are null when they exist, while the existence of even-order moments of $(\epsilon_t, t \in \mathbb{Z})$ is summarized in the following theorem

Theorem 2.3 *Consider the MS-univariate stochastic volatility model (2.1) with MS-multivariate stochastic volatility model (2.2). For all integer* $m > 0$*, assume that* $E\left\{e_t^m\right\} < +\infty$, $E\left\{\eta_t^m\right\} < +\infty$ *and*

$$\rho_{(m)} := \rho\left(\mathbb{P}(\Gamma^{(m)})\right) < 1, \tag{2.5}$$

where $\Gamma^{(m)} := \left\{\Gamma^{\otimes m}(k), k \in \mathbb{S}\right\}$*. Then the* $MSAR - SV$ *defined by the state-space (2.2) has a unique, causal,ergodic and strictly stationary solution given by (2.3) having moment up to* $m-$ *order. Moreover, the closed form of the* $m-$*th moment of* ϵ_t *is given by*

$$E\left\{\epsilon_t^m\right\} = E\left\{e_t^m\right\} \sum_{x_t, x_{t-1}, \ldots \in \mathbb{S}} \prod_{k \geq 0} p_{x_{t-k-1} x_{t-k}}$$
$$E\left\{\exp\left\{\frac{m}{2}\left\{\prod_{j=0}^{k-1} \Gamma\left(x_{t-j}\right)\right\}(1)\left(a_0\left(x_{t-k}\right) + b_0\left(x_{t-k}\right)\eta_0\right)\right\}\right\}.$$

Proof We have used the same proof of the last theorem, the results obtained can be extended and hence omitted the details. □

Remark 2.3 Applying the normality of $(\eta_t, t \in \mathbb{Z})$ yields

$$E\left\{\epsilon_t^m\right\} = E\left\{e_t^m\right\} \sum_{x_t, x_{t-1}, \ldots \in \mathbb{S}} \prod_{k \geq 0} p_{x_{t-k-1} x_{t-k}}$$
$$\exp\left\{\frac{m}{2}\left\{\prod_{j=0}^{k-1} \Gamma\left(x_{t-j}\right)\right\}(1)\, a_0\left(x_{t-k}\right) + \frac{1}{2}\left(\frac{m}{2}\left\{\prod_{j=0}^{k-1} \Gamma\left(x_{t-j}\right)\right\}(1)\right)^2 b_0^2\left(x_{t-k}\right)\right\}.$$

The autocovariance function of the squared process $\left(\epsilon_t^2, t \in \mathbb{Z}\right)$ is summarized in the following theorem

Theorem 2.4 *Under the assumptions of the last theorem, we have*

1. *If* $(\epsilon_t, t \in \mathbb{Z})$ *follows the* MS*-univariate stochastic volatility model (2.1) and* $\epsilon_t \in \mathbb{L}_4$*, then*

$$\gamma_{\epsilon^2}(0) = E\left\{e_t^4\right\} \sum_{x_t, x_{t-1}, \ldots \in \mathbb{S}} \prod_{k \geq 0} p_{x_{t-k-1} x_{t-k}}$$
$$E\left\{\exp\left\{2\left\{\prod_{j=0}^{k-1} \Gamma\left(x_{t-j}\right)\right\}(1)\left(a_0\left(x_{t-k}\right) + b_0\left(x_{t-k}\right)\eta_0\right)\right\}\right\} - \gamma_\epsilon^2(0),$$

and $\gamma_{\epsilon^2}(h) = 0$ *otherwise.*

2. *If* $(\epsilon_t, t \in \mathbb{Z})$ *follows the* MS*-univariate stochastic volatility model (2.1) and* $\epsilon_t \in \mathbb{L}_{2m}$*, then*

$$\gamma_{\epsilon^m}(0) = E\left\{e_t^{2m}\right\} \sum_{x_t, x_{t-1}, \ldots \in \mathbb{S}} \prod_{k \geq 0} p_{x_{t-k-1} x_{t-k}}$$
$$E\left\{\exp\left\{m\left\{\prod_{j=0}^{k-1} \Gamma\left(x_{t-j}\right)\right\}(1)\left(a_0\left(x_{t-k}\right) + b_0\left(x_{t-k}\right)\eta_0\right)\right\}\right\} - \left(E\left\{\epsilon_t^m\right\}\right)^2,$$

and $\gamma_{\epsilon^m}(h) = 0$ *otherwise.*

Proof It is enough to remark that the processes (ϵ_t^2) and (ϵ_t^{2m}) are both white noise processes.

□

Remark 2.4 It is clear that the process (ϵ_t^2) and its power do not admit an $ARMA$ representation. However, the logarithm process $(\log \epsilon_t^2)$ has an $ARMA$ autocovariance structure.

From the previous remark, we can obtain the following representation

$$\log \epsilon_t^2 = \widetilde{a}_0 (s_t) + \sum_{i=1}^{p} a_i (s_t) \log \epsilon_{t-i}^2 - \sum_{i=1}^{p} a_i (s_t) \omega_{t-i} + \omega_t + b_0 (s_t) \eta_t , \qquad (2.6)$$

with $\omega_t := \log e_t^2 - E\left\{\log e_t^2\right\}$ and $\widetilde{a}_0 (.)$ is an intercept, easily obtained. So Eq. (2.6) can be written in the following vectorial representation,

$$\underline{M}_t = A (s_t) \underline{M}_{t-1} + \underline{V}_t (s_t) , \qquad (2.7)$$

where $\underline{M}_t' := \left(\log \epsilon_t^2, ..., \log \epsilon_{t-p+1}^2, \omega_t, ..., \omega_{t-p+1}\right)$ and $\underline{V}_t (s_t) := \underline{v}_0 + \underline{v}_1 \omega_t + \underline{v}_2 (s_t) \eta_t$, whereas $\underline{v}_0,\ \underline{v}_1,\ \underline{v}_2 (s_t),\ A (s_t)$ are appropriate vectors and matrix easily obtained and uniquely determined by $\{a_i (s_t) , b_0 (s_t) , 0 \le i \le p\}$. The feature of the vectorial representation (2.7) when $s_t = k$, the vector $\underline{M}_t$ is independent of $\underline{V}_u (k)$ for $u > t$. For appropriateness, we will treat the centered version of the vector $\underline{M}_t$,

$$\underline{\widetilde{M}}_t = A (s_t) \underline{\widetilde{M}}_{t-1} + \underline{\widetilde{V}}_t (s_t) , \qquad (2.8)$$

where $\underline{\widetilde{M}}_t = \underline{M}_t - E\left\{\underline{M}_t\right\}$ and $\underline{\widetilde{V}}_t (s_t)$ is the centered residual vector such that $s_t = k$, $\underline{\widetilde{V}}_t (k) \perp \underline{\widetilde{M}}_u$ for $t > u$.

Proposition 2.2 *Consider the $MSAR - SV (p)$ process (2.1) with vectorial representation (2.8). Then under the assumptions of Theorem 2.2, we get*

$$Cov\left(\log \epsilon_t^2, \log \epsilon_{t-h}^2\right) = \begin{cases} \left(\underline{1} \otimes \underline{H}_0^{\otimes 2}\right)' \left(I_{(4dp^2)} - \mathbb{P}\left(A^{(2)}\right)\right)^{-1} \underline{\Pi}\left(\underline{\widetilde{V}}^{(2)}\right) & if\, h = 0 \\ \left(\underline{1} \otimes \underline{H}_0^{\otimes 2}\right)' \mathbb{P}^h \left(A^{(1)} \otimes I_{(2p)}\right) \left(I_{(4dp^2)} - \mathbb{P}\left(A^{(2)}\right)\right)^{-1} \underline{\Pi}\left(\underline{\widetilde{V}}^{(2)}\right) \\ if\, h > 0 \end{cases},$$

where $\underline{\widetilde{V}}^{(2)} := \left(\underline{\widetilde{V}}_k^{(2)} = E\left\{\underline{\widetilde{V}}_t^{\otimes 2} (k)\right\}, k \in \mathbb{S}\right)$, $A^{(n)} := \left(A^{(n)}(k) = A^{\otimes n} (k) , k \in \mathbb{S}\right)$; $n = 1, 2$, $\underline{1} = (1, ..., 1)' \in \mathbb{R}^d$, $\underline{H}_0' := \left(I_{(1)}, \underline{O}_{(2p-1)}'\right)$.

Proof Starting from (2.8), for $h = 0$ we get

$$\pi(k) E\left\{\underline{\widetilde{M}}_t^{\otimes 2} | s_t = k\right\} = \pi(k) \underline{\widetilde{V}}_k^{(2)} + A^{(2)}(k) \sum_{j \in \mathbb{S}} E\left\{\underline{\widetilde{M}}_{t-1}^{\otimes 2} | s_{t-1} = j\right\} p_{jk} \pi(j),$$

and when $h > 0$,

$$\pi(k) E\left\{\underline{\widetilde{M}}_t \otimes \underline{\widetilde{M}}_{t-h} \middle| s_t = k\right\} = \left(A(k) \otimes I_{(2p)}\right) \sum_{j \in \mathbb{S}} E\left\{\underline{\widetilde{M}}_{t-1} \otimes \underline{\widetilde{M}}_{t-1-(h-1)} \middle| s_{t-1} = j\right\} p_{jk} \pi(j).$$

Let $\underline{\Lambda} (h) = \left(E\left\{\underline{\widetilde{M}}_t \otimes \underline{\widetilde{M}}_{t-h} | s_t = k\right\}, k \in \mathbb{S}\right)$. Then

$$\underline{\Pi}\left(\underline{\Lambda}(h)\right) = \begin{cases} \mathbb{P}\left(A^{(2)}\right) \underline{\Pi}\left(\underline{\Lambda}(0)\right) + \underline{\Pi}\left(\underline{\widetilde{V}}^{(2)}\right) & \text{if } h = 0 \\ \mathbb{P}\left(A^{(1)} \otimes I_{(2p)}\right) \underline{\Pi}\left(\underline{\Lambda}(h-1)\right) = \mathbb{P}^h \left(A^{(1)} \otimes I_{(2p)}\right) \underline{\Pi}\left(\underline{\Lambda}(0)\right) & \text{if } h > 0 \end{cases}.$$

□

The following result gives an $ARMA$ representation for the $MSAR - SV (p)$ model.

Proposition 2.3 *Under the conditions of Theorem 2.2, the $MSAR-SV(p)$ process with vectorial representation (2.8) is a $ARMA$ process.*

Proof To demonstrate proposition 2.3, we utilize the same technique as Ghezal et al. [1]. □

For this purpose, it is requisite to behold a higher-power of the log −squared observed process admits a $ARMA$ representation. Let the vector $\underline{Z}_t' := \left(\log \epsilon_t^2, ..., \log \epsilon_{t-p+1}^2, \omega_t, ..., \omega_{t-p+1}\right)$. Then the next result can be shown by a straightforward modification of Ghezal et al. [1]

Proposition 2.4 *Consider the model (2.1), and suppose that* $\log e_t^2, \eta_t \in \mathbb{L}_{2m}$ *for any positive integer m. Then $\left(\underline{Z}_t^{\otimes m}\right)$ is solution of a $ARMA(n_1, n_2)$ equation of the form*

$$\underline{Z}_t^{\otimes m} - E\left\{\underline{Z}_t^{\otimes m}\right\} = \sum_{i=1}^{n_1} \Psi_i \left(\underline{Z}_{t-i}^{\otimes m} - E\left\{\underline{Z}_t^{\otimes m}\right\}\right) + \sum_{j=1}^{n_2} \Lambda_j \underline{v}_{t-j} + \underline{v}_t, \tag{2.9}$$

where $\left(\underline{v}_t\right)$ is a white noise and (Ψ_i), $\left(\Lambda_j\right)$ are sequences of $(2p)^m \times (2p)^m$ matrices.

At the end of this section, the geometric ergodicity and $\beta-$ mixing are manifest.

Theorem 2.5 *Consider the model (2.1). Under the condition (2.4), $\left(\left(\log \underline{H}_t\right)', s_t\right)'$ is a geometrically ergodic Markovian chain. If it is initialized from the invariant measure, then (ϵ_t) and $(\log h_t)$ are strict stationarity and $\beta-$mixing with exponential rate.*

Proof The result follows from geometric ergodicity of the process $(\log \underline{H}_t)$, which can be easily created using Ghezal et al. [1]. □

3 QML estimation

The estimation of Markov-switching models is rather complex. So, some specific models were considered in the literature (see for example, Francq and Zakoian [9], Ghezal [11] for further discussions). There is already established Markov Chain Monte Carlo ($MCMC$) procedure in the literature for estimating a few particular states of Eq. (2.1) including [16], [18] among others. Now, we consider a given realization $(\epsilon_1, \epsilon_2, ..., \epsilon_n)$ created from the unique, causal and strictly stationary $MSAR-SV$ model, and suppose that p and d are known and (e_t) is standard Gaussian. The unknown parameters $a_i(.)$, $b_0(.)$, $i = 0, ..., p$ and $\left(p_{i,j}, i, j = 1, ..., s, i \neq j\right)$ collected in a vector $\underline{\theta}$ belonging to the parameter space Θ, while $\underline{\theta}_0$ is the true values. Xie [19] advocated the $QMLE$ and established its strong consistency for $MS-GARCH_s(p, q)$, Ghezal [11] imposes some assumptions under which the strong consistency of $QMLE$ for the doubly $MS-AR$ model is satisfied. The Gaussian likelihood function is given

$$L_n\left(\underline{\theta}\right) = \sum_{s_1,...,s_n \in \mathbb{S}} \pi(s_1) \left\{\prod_{i=2}^{n} p_{s_{i-1},s_i}\right\} \left\{\prod_{i=1}^{n} f_{s_i}(\epsilon_1, ..., \epsilon_i)\right\}, \tag{3.1}$$

where

$$f_{s_i}(\epsilon_1, ..., \epsilon_i) = \frac{1}{\left(2\pi h_{s_i}(\epsilon_1, ..., \epsilon_{i-1})\right)^{1/2}} \exp\left\{-\frac{\epsilon_i^2}{2h_{s_i}(\epsilon_1, ..., \epsilon_{i-1})}\right\},$$

with the log −transformed conditional stochastic variance process is $\log h_{s_i}(\epsilon_1, ..., \epsilon_{i-1})$ defined by the second equation in (2.1). This likelihood function is also written in the following form

$$L_n\left(\underline{\theta}\right) = \underline{\mathbf{1}}'_{(s)}\left\{\prod_{i=1}^{n}\mathbb{P}_{\underline{\theta}}\left(f(\epsilon_1, ...\epsilon_i)\right)\right\}\underline{\Pi}(f\left(\epsilon_1\right)). \tag{3.2}$$

A $QMLE$ of $\underline{\theta}_0$ is defined as any measurable solution $\widehat{\underline{\theta}}_n$ of

$$\widehat{\underline{\theta}}_n = \arg\max_{\theta\in\Theta} L_n\left(\underline{\theta}\right). \tag{3.3}$$

In this section, let $f_{s_t}\left(\epsilon_t | \underset{\leftarrow}{\epsilon}_{t-1}\right)$ (resp. $f_{s_t}\left(\epsilon_t | \underline{\epsilon}_1\right)$) be the density of ϵ_t given the all past observations (resp. past observations unto ε_1) and let $g_{\underline{\theta}}\left(\epsilon_t | \underset{\leftarrow}{\epsilon}_{t-1}\right)$ (resp. $g_{\underline{\theta}}\left(\epsilon_t | \underline{\epsilon}_1\right)$) be the corresponding logarithm conditional density of ϵ_t given $\{\epsilon_{t-1}, \epsilon_{t-2}, ...\}$ (resp. $\{\epsilon_{t-1}, \epsilon_{t-2}, ...\epsilon_1\}$). Now, we determine the likelihood function $\widetilde{L}_n\left(\underline{\theta}\right)$ based on all past observations which is defined as $L_n\left(\underline{\theta}\right)$ in (3.1) except changing the density $f_{s_t}(\epsilon_1, ...\epsilon_t)$ by $f_{s_t}\left(\epsilon_t | \underset{\leftarrow}{\epsilon}_{t-1}\right)$. Furthermore, we can write $\widetilde{L}_n\left(\underline{\theta}\right)$ as

$$\widetilde{L}_n\left(\underline{\theta}\right) = \underline{\mathbf{1}}'_{(s)}\left\{\prod_{t=2}^{n}\mathbb{P}_{\underline{\theta}}\left(f\left(\epsilon_t | \underset{\leftarrow}{\epsilon}_{t-1}\right)\right)\right\}\underline{\Pi}(f\left(\epsilon_1 | \underset{\leftarrow}{\epsilon}_0\right)), \tag{3.4}$$

where the matrix $\mathbb{P}_{\underline{\theta}}\left(f\left(\epsilon_i | \underset{\leftarrow}{\epsilon}_{i-1}\right)\right)$ (resp. the vector $\underline{\Pi}(f\left(\epsilon_1 | \underset{\leftarrow}{\epsilon}_0\right))$) replaces $f_{s_i}(\epsilon_1, ...\epsilon_i)$ by $f_{s_i}\left(\epsilon_i | \underset{\leftarrow}{\epsilon}_{i-1}\right)$, $i = 1, .., n$ in $\mathbb{P}_{\underline{\theta}}\left(f(\epsilon_1, ...\epsilon_i)\right)$ (resp. $\underline{\Pi}(f\left(\epsilon_1\right))$).

3.1 Strong consistency of the *QMLE*

To prove the strong consistency of the $QMLE$, we use the following assumptions

A1. Θ is compact subset of $\mathbb{R}^s$ and the true value $\underline{\theta}_0$ belongs to Θ.
A2. $\gamma_L\left(\Gamma^0\right) < 0$ for any $\underline{\theta} \in \Theta$ where Γ^0 is the sequence $(\Gamma(s_t), t \in \mathbb{Z})$ when the parameters $\underline{\theta}$ are changed by $\underline{\theta}_0$.
A3. For any $\underline{\theta}, \underline{\theta}^* \in \Theta$, if almost surely $g_{\underline{\theta}}\left(\epsilon_t | \underset{\leftarrow}{\epsilon}_{t-1}\right) = g_{\underline{\theta}^*}\left(\epsilon_t | \underset{\leftarrow}{\epsilon}_{t-1}\right)$ then $\underline{\theta} = \underline{\theta}^*$.

Assumption **A1** is a standard assumption and it is used in many results of real analysis. Assumption **A2** ensures the strict stationarity of the process $(\epsilon_t, t \in \mathbb{Z})$. Assumption **A3** ensures that the parameter $\underline{\theta}$ is identifiable. First, we present the following key lemmas.

Lemma 3.1 *Under Assumptions **A2** and **A3**, almost surely, we have*

$$\lim_{n\longrightarrow\infty}\frac{1}{n}\log L_n\left(\underline{\theta}\right) = \lim_{n\longrightarrow\infty}\frac{1}{n}\log\widetilde{L}_n\left(\underline{\theta}\right) = E_{\underline{\theta}_0}\left\{g_{\underline{\theta}}\left(\epsilon_t | \underset{\leftarrow}{\epsilon}_{t-1}\right)\right\}$$

Proof Using the logarithmic function, we have $\log\widetilde{L}_n\left(\underline{\theta}\right) = \sum_{t=1}^{n} g_{\underline{\theta}}\left(\epsilon_t | \underset{\leftarrow}{\epsilon}_{t-1}\right)$ and $\log L_n\left(\underline{\theta}\right) = \sum_{t=1}^{n} g_{\underline{\theta}}\left(\epsilon_t | \underline{\epsilon}_1\right)$. Then

$$\frac{1}{n}\sum_{t=1}^{n} g_{\underline{\theta}}\left(\epsilon_t | \underline{\epsilon}_1\right) = \frac{1}{n}\sum_{t=1}^{n} g_{\underline{\theta}}\left(\epsilon_t | \underset{\leftarrow}{\epsilon}_{t-1}\right) + \frac{1}{n}\sum_{t=1}^{n}\left(g_{\underline{\theta}}\left(\epsilon_t | \underline{\epsilon}_1\right) - g_{\underline{\theta}}\left(\epsilon_t | \underset{\leftarrow}{\epsilon}_{t-1}\right)\right).$$

Now, for all $\varsigma \geq 0$, the process $(N_t(m))$ is defined as $N_t(m) = \sup_{\varsigma \geq m} \left| g_{\underline{\theta}}\left(\epsilon_t | \underline{\epsilon}_{t-\varsigma}\right) - g_{\underline{\theta}}\left(\epsilon_t | \underleftarrow{\epsilon}_{t-1}\right)\right|$. Then for fixed m, the process $(N_t(m))$ is also strictly stationary and ergodic with $E_{\underline{\theta}_0}\{N_t(m)\} < +\infty$. We have

$$\begin{aligned}
&\lim \sup_{n \longrightarrow \infty} \left| \frac{1}{n} \sum_{t=1}^{n} \left(g_{\underline{\theta}}\left(\epsilon_t | \underline{\epsilon}_1\right) - g_{\underline{\theta}}\left(\epsilon_t | \underleftarrow{\epsilon}_{t-1}\right)\right)\right| \\
&\leq \lim \sup_{n \longrightarrow \infty} \frac{1}{n} \sum_{t=1}^{n} \left| g_{\underline{\theta}}\left(\epsilon_t | \underline{\epsilon}_1\right) - g_{\underline{\theta}}\left(\epsilon_t | \underleftarrow{\epsilon}_{t-1}\right)\right| \\
&\leq \lim \sup_{n \longrightarrow \infty} \frac{1}{n} \sum_{t=m+1}^{n} N_t(m) = E_{\underline{\theta}_0}\{N_0(m)\},
\end{aligned}$$

the result follows. □

The next lemma compares the ratios $\frac{L_n(\underline{\theta})}{L_n(\underline{\theta}_0)}$ and $\frac{\widetilde{L}_n(\underline{\theta})}{\widetilde{L}_n(\underline{\theta}_0)}$. Let $Z_n\left(\underline{\theta}\right) = \frac{1}{n} \log\left(\frac{L_n(\underline{\theta})}{L_n(\underline{\theta}_0)}\right)$. Then, we have

Lemma 3.2 *Under Assumptions **A1–A3**, we have*

$$\lim_{n \to \infty} \frac{1}{n} \log\left(\frac{\widetilde{L}_n\left(\underline{\theta}\right)}{\widetilde{L}_n\left(\underline{\theta}_0\right)}\right) = \lim_{n \longrightarrow \infty} Z_n\left(\underline{\theta}\right) \leq 0,$$

with $\lim_{n \longrightarrow \infty} Z_n\left(\underline{\theta}\right) = 0$ *iff* $\underline{\theta} = \underline{\theta}_0$ *for all* $\underline{\theta} \in \Theta$.

Proof Under assumptions **A1–A3**, the function $Z_n\left(\underline{\theta}\right)$ is well defined. Moreover by lemma 3.1 and Jensen's inequality, we get

$$\lim_{n \longrightarrow \infty} Z_n\left(\underline{\theta}\right) = E_{\underline{\theta}_0}\left\{\log \frac{g_{\underline{\theta}}\left(\epsilon_t | \underleftarrow{\epsilon}_{t-1}\right)}{g_{\underline{\theta}_0}\left(\epsilon_t | \underleftarrow{\epsilon}_{t-1}\right)}\right\} \leq \log E_{\underline{\theta}_0}\left\{\frac{g_{\underline{\theta}}\left(\epsilon_t | \underleftarrow{\epsilon}_{t-1}\right)}{g_{\underline{\theta}_0}\left(\epsilon_t | \underleftarrow{\epsilon}_{t-1}\right)}\right\} = 0.$$

Under the Assumption **A3,** $Z_n\left(\underline{\theta}\right)$ converges to Kullback-Leinbler information which equals zero iff $\underline{\theta} = \underline{\theta}_0$. □

Lemma 3.3 *Under **A1–A3.** For all $\underline{\theta}^* \neq \underline{\theta}_0$, there exists a neighborhood $\mathcal{V}\left(\underline{\theta}^*\right)$ of $\underline{\theta}^*$ such that*

$$\lim \sup_{n \to +\infty} \sup_{\underline{\theta} \in \mathcal{V}(\underline{\theta}^*)} Z_n\left(\underline{\theta}\right) < 0 \; a.s.$$

Proof In Eq. (3.4), we obtain

$$\begin{aligned}
&\min_j \pi(j) f_j\left(\epsilon_1 | \underleftarrow{\epsilon}_0\right) \left\| \left\{ \prod_{t=2}^{n} \mathbb{P}_{\underline{\theta}}\left(f\left(\epsilon_t | \underleftarrow{\epsilon}_{t-1}\right)\right)\right\}\right\| \\
&\leq \widetilde{L}_n\left(\underline{\theta}\right) \leq \max_j \pi(j) f_j\left(\epsilon_1 | \underleftarrow{\epsilon}_0\right) \left\| \left\{ \prod_{t=2}^{n} \mathbb{P}_{\underline{\theta}}\left(f\left(\epsilon_t | \underleftarrow{\epsilon}_{t-1}\right)\right)\right\}\right\|.
\end{aligned}$$

So we obtain

$$\begin{aligned}&\lim_{n\to\infty}\frac{1}{n}\log\widetilde{L}_n\left(\underline{\theta}\right)\\&=\lim_{n\to\infty}\log\frac{1}{n}\left\|\left\{\prod_{t=2}^{n}\mathbb{P}_{\underline{\theta}}\left(f\left(\epsilon_t|\underleftarrow{\epsilon}_{t-1}\right)\right)\right\}\right\|=E_{\underline{\theta}_0}\left\{g_{\underline{\theta}}\left(\epsilon_t|\underleftarrow{\epsilon}_{t-1}\right)\right\}.\end{aligned}$$

Let $\mathcal{V}_m\left(\underline{\theta}^*\right)=\left\{\underline{\theta}:\left\|\underline{\theta}-\underline{\theta}^*\right\|\leq\frac{1}{m}\right\}$ and $\Sigma_{2:n}^m=\sup_{\underline{\theta}\in\mathcal{V}_m\left(\underline{\theta}^*\right)}\left\|\prod_{t=2}^{n}\mathbb{P}_{\underline{\theta}}\left(\underline{f}\left(\epsilon_t|\underleftarrow{\epsilon}_{t-1}\right)\right)\right\|$. Because the norm is multiplicative, we obtain on $\mathcal{V}_m\left(\underline{\theta}^*\right)$

$$\begin{aligned}&\sup_{\underline{\theta}}\left\|\prod_{t=2}^{n+k}\mathbb{P}_{\underline{\theta}}\left(f\left(\epsilon_t|\underleftarrow{\epsilon}_{t-1}\right)\right)\right\|\\&\leq\sup_{\underline{\theta}}\left\|\prod_{t=2}^{n}\mathbb{P}_{\underline{\theta}}\left(f\left(\epsilon_t|\underleftarrow{\epsilon}_{t-1}\right)\right)\right\|.\sup_{\underline{\theta}}\left\|\prod_{t=n+1}^{n+k}\mathbb{P}_{\underline{\theta}}\left(f\left(\epsilon_t|\underleftarrow{\epsilon}_{t-1}\right)\right)\right\|,\end{aligned}$$

that implies

$$\log\Sigma_{2:n+k}^m\leq\log\Sigma_{2:n}^m+\log\Sigma_{n+1:n+k}^m\ \forall n,k.$$

Now $\left(\log\Sigma_{2:n}^m\right)$ is a strictly stationary and ergodic process with $E_{\underline{\theta}_0}\left\{\log\Sigma_{2:n}^m\right\}$ is finite. Then we get

$$\xi_m\left(\underline{\theta}^*\right)=\lim_{n\longrightarrow\infty}\frac{1}{n}\log\Sigma_{2:n}^m=\inf_{n>1}\frac{1}{n}E_{\underline{\theta}_0}\left\{\log\Sigma_{2:n}^m\right\}a.s.,$$

where $\xi\left(\underline{\theta}\right)$ is the Lyapunov exponent of the sequence $\left(\mathbb{P}_{\underline{\theta}_0}\left(f\left(\epsilon_t|\underleftarrow{\epsilon}_{t-1}\right)\right),t\in\mathbb{Z}\right)$ i,e., $\xi\left(\underline{\theta}\right)=\inf_{n>1}\frac{1}{n}E_{\underline{\theta}_0}\left\{\log\left\|\prod_{t=2}^{n}\mathbb{P}_{\underline{\theta}_0}\left(f\left(\epsilon_t|\underleftarrow{\epsilon}_{t-1}\right)\right)\right\|\right\}\overset{a.s.}{=}\lim_{n\longrightarrow\infty}\frac{1}{n}\log\left\|\prod_{t=2}^{n}\mathbb{P}_{\underline{\theta}_0}\left(f\left(\epsilon_t|\underleftarrow{\epsilon}_{t-1}\right)\right)\right\|$.

Hence, using Lemma 3.2, there exist $\varepsilon>0$ and $n_\varepsilon\in\mathbb{N}$ such that $\frac{1}{n_\varepsilon}E_{\underline{\theta}_0}\left\{\log\left\|\prod_{t=2}^{n_\varepsilon}\mathbb{P}_{\underline{\theta}^*}\left(f\left(\epsilon_t|\underleftarrow{\epsilon}_{t-1}\right)\right)\right\|\right\}<\xi\left(\underline{\theta}_0\right)-\varepsilon$. From the *DCT* theorem, it follows that for m large enough we get

$$\xi_m\left(\underline{\theta}^*\right)\leq\frac{1}{n_\varepsilon}E_{\underline{\theta}_0}\left\{\log\left\|\prod_{t=2}^{n_\varepsilon}\mathbb{P}_{\underline{\theta}^*}\left(f\left(\epsilon_t|\underleftarrow{\epsilon}_{t-1}\right)\right)\right\|\right\}+\frac{\varepsilon}{2}<\xi\left(\underline{\theta}_0\right)-\frac{\varepsilon}{2}.$$

The result follows by Lemma 3.1. □

Second, we present the following main theorem.

Theorem 3.1 *Under* $\mathbf{A}1-\mathbf{A}3$, *the sequence of* QML *estimators* $\left(\widehat{\underline{\theta}}_n\right)_n$ *satisfying (3.3) is strong consistency, i.e.,*

$$\widehat{\underline{\theta}}_n\rightarrow\underline{\theta}_0\text{ almost surely when }n\rightarrow+\infty.$$

Proof Assume that $\hat{\underline{\theta}}_n$ does not converge to $\underline{\theta}_0$ a.s., i.e.,

$$\forall n,\ \exists\delta>0,\ N>n,\ \text{such that}\ \left\|\hat{\underline{\theta}}_N-\underline{\theta}^0\right\|\geq\delta.$$

Table 1 Average and RMSE of $QMLE$ for Gaussian $MSAR-SV(1)$ models with different values of the sample size

	Tv\ n	500	1000	2000
p_{12}	0.75	0.7409 (0.1154)	0.7411 (0.1125)	0.7423 (0.0892)
p_{21}	0.50	0.4931 (0.1027)	0.4953 (0.0928)	0.5089 (0.0724)
a_0 (1)	1.00	1.1207 (0.535)	1.0183 (0.365)	1.0052 (0.276)
a_0 (2)	0.00	−0.001 (0.031)	0.000 (0.016)	0.000 (0.009)
a_1 (1)	−0.50	−0.5231 (0.0734)	−0.5257 (0.0395)	−0.5256 (0.0375)
a_1 (2)	1.20	1.2546 (0.0753)	1.2524 (0.0651)	1.1873 (0.0613)
b_0 (1)	0.35	0.365 (0.222)	0.361 (0.152)	0.356 (0.089)
b_0 (2)	0.20	0.215 (0.062)	0.209 (0.049)	0.206 (0.038)

Table 2 Average and RMSE of $QMLE$for Gaussian MSAR-SV(2) models with different values of the sample size

	Tv\n	500	1000	2000
p_{12}	0.85	0.8107 (0.1589)	0.8287 (0.0917)	0.8354 (0.0580)
p_{21}	0.50	0.4598 (0.1530)	0.4663 (0.0891)	0.4814 (0.0675)
a_0 (1)	1.50	1.5239 (0.0241)	1.5084 (0.0238)	1.5056 (0.0232)
a_0 (2)	1.00	1.0021 (0.0181)	0.9978 (0.0165)	0.9993 (0.0149)
a_1 (1)	−0.98	−0.9823 (0.0174)	−0.9805 (0.0168)	−0.9802 (0.0153)
a_1 (2)	0.50	0.5101 (0.0291)	0.5083 (0.0279)	0.5049 (0.0248)
a_2 (1)	0.99	0.9710 (0.0166)	0.9791 (0.0121)	0.9863 (0.0098)
a_2 (2)	−0.50	−0.4948 (0.0284)	−0.5033 (0.0202)	−0.5003 (0.0128)
b_0 (1)	0.10	0.1166 (0.0135)	0.1009 (0.116)	0.1003 (0.0102)
b_0 (2)	0.44	0.4483 (0.0249)	0.4452 (0.0243)	0.4417 (0.0221)

Using the Lemma 3.3, we have $L_n\left(\hat{\underline{\theta}}_n\right) < L_n\left(\underline{\theta}_0\right)$. However, by the $QMLE$ provided in (3.3), we get

$$L_n\left(\hat{\underline{\theta}}_n\right) = \sup_{\underline{\theta}\in\Theta^*} L_n\left(\underline{\theta}\right) \geq L_n\left(\underline{\theta}_0\right)$$

for any compact subset Θ^* of Θ containing $\underline{\theta}_0$. This discrepancy confers the result. □

4 Simulation study

In order to evaluate the performance of the QML method for parameters estimation, we carried out a simulation study based on the Gaussian $MSAR-SV(p)$ model with $d=2$. We simulated 1000 data samples with different lengths. The sample sizes to be examined in this simulation study are $n \in \{500, 1000, 2000\}$. The corresponding parameter values are chosen to satisfy the stationarity condition $\gamma_L(\Gamma) < 0$. For each trajectory the vector $\underline{\theta}$ of parameters of interest has been estimated with $QMLE$ noted as $\widehat{\underline{\theta}}$. The $QMLE$ algorithm has been executed for these series under the $MATLAB8$ using "$fminsearch.m$" as a minimizer function. In Tables below, the root mean square errors ($RMSE$) of $\widehat{\theta}$, are displayed in parenthesis in each table, the true values (TV) of the parameters of each of the considered data-generating process are reported.

The roots mean square errors are the main focus in this study. The results provide some preliminary evidence with respect to the finite sample properties of the $QMLE$ in the $MSAR-SV$ framework. It can be observed that the parameters are quite well estimated by the $QMLE$ method. Now let us devote a few comments. Table 1 shows that the strong consistency of $QMLE$ of $MS-$models is fairly satisfying and the associated $RMSE$ decreases closely as the sample size increases. Regarding outcomes associated with $MS-$models reported in Table 2, it is obvious that the strong consistency is fully approved. Furthermore, it can be seen that even with a relatively small sample size, the procedure of estimation gives a good result.

Acknowledgements We should like to thank the Editor in Chief of the journal, an Associate Editor and the anonymous referees for their constructive comments and very useful suggestions and remarks which were most valuable for improvement in the final version of the paper. We also would like to thank our colleagues, Khadidja Himour and Dr. Khalil Zerari, who encouraged us a lot.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Bibi, A., Ghezal, A.: On the Markov-switching bilinear processes: stationarity, higher-order moments and $\beta-$mixing. Stochas. Int. J. Probab. Stochast. Process. **87**(6), 919–945 (2015)
2. Casarin, R.: Bayesian inference for generalised Markov switching stochastic volatility models. Conference materials at the 4th International Workshop on Objective Bayesian Methodology, CNRS, Aussois (2003)
3. Cavicchioli, M.: Spectral density of Markov-switching $VARMA$ models. Econ. Lett. **121**, 218–220 (2013)
4. Cavicchioli, M.: Higher order moments of Markov switching $VARMA$ models. Economet. Theory **33**(6), 1502–1515 (2017)
5. Cavicchioli, M.: Asymptotic Fisher information matrix of Markov switching $VARMA$ models. J. Multivar. Anal. **157**, 124–135 (2017)
6. Cavicchioli, M.: Markov switching $GARCH$ models: higher order moments, kurtosis measures and volatility evaluation in recessions and pandemic. J. Bus. Econ. Stat. **40**(4), 1772–1783 (2022)
7. Douc, R., Moulines, E., Stoffer, D.: Nonlinear Time Series Theory, Methods, and Applications with R Examples. CRC Press, New York (2014)
8. Francq, C., Zakoïan, J.-M.: Linear representation based estimation of stochastic volatility models. Scand. J. Stat. **33**(4), 785–806 (2006)
9. Francq, C., Zakoïan, J.-M.: Deriving the autocovariances of powers of Markov-switching $GARCH$ models, with applications to statistical inference. Comput. Stat. Data Anal. **52**(6), 3027–3046 (2008)
10. Francq, C., Zakoïan, J.-M.: QML estimation of a class of multivariate asymmetric $GARCH$ models. Economet. Theor. **28**(1), 179–206 (2012)
11. Ghezal, A.: 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
12. Ghezal, A., Zemmouri, I.: Estimating $MS-BLGARCH$ models using recursive method. Pan-Am. J. Math. **2**, 1–7 (2023)
13. Haas, M., Mittnik, S., Paolella, M.S.: A new approach to Markov-switching $GARCH$ models. J. Financ. Economet. **2**(4), 493–530 (2004)
14. Hamilton, J.D.: A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica **57**(2), 357–384 (1989)
15. Leroux, B.G.: Maximum-likelihood estimation for hidden Markov models. Stochas. Process. Appl. **40**, 127–143 (1992)

16. So, M.E.C.P., Lam, K., Li, W.K.: A stochastic volatility model with Markov switching. J. Bus. Econ. Stat. **16**(2), 244–253 (1998)
17. Stelzer, R.: On Markov-switching $ARMA$ processes: stationarity, existence of moments and geometric ergodicity. Economet. Theory **25**(1), 43–62 (2009)
18. Vo, M.T.: Regime-switching stochastic volatility: evidence from the crude oil market. Energy Econ. **31**(5), 779–788 (2009)
19. Xie, Y.: Consistency of maximum likelihood estimators for the regimes witching $GARCH$ model. Stat. A J. Theor. Appl. Stat. **43**(2), 153–165 (2009)