Stochastic Models in Probability and Statistics
Vol. -, No. -, 2025, pp. -. Research Article

Higher-order moments of Markov switching bilinear models: theory and empirical evidence

Abdelouahab Bibi

Department of Mathematics, University Fréres Mentouri Constantine 1, Constantine, Algeria

Email(s): abd.bibi@gmail.com

Abstract. In this paper, we consider a *Markov-switching bilinear process* (*MS-BL*) that exhibits rich dynamic behavior and plays an important role in modeling non-Gaussian data characterized by structural breaks. In such models, the parameters depend on an unobservable (hidden) Markov chain with a finite state space. Although numerous recent studies have focused on the statistical aspects of Markov-switching models, systematic investigations of the probabilistic properties of this class of nonlinear models remain relatively scarce. So, we derive conditions for stationarity and compute the moments of the process up to the third order. Our analysis reveals that the conditions ensuring local stationarity within each regime of the observed process are neither sufficient nor necessary. Furthermore, we show that the second-order structure of the process is analogous to that of a Markov-switching *ARMA* (*MS-ARMA*) model with an additional uncorrelated white noise component. Therefore, the examination of higher-order moments becomes essential to distinguish between (locally) linear and nonlinear models. To illustrate the practical relevance of our theoretical results, we conduct Monte Carlo simulation studies and apply the proposed model to the exchange rate of the Algerian Dinar against the Euro. The empirical findings indicate that the proposed approach provides a better fit and demonstrates superior performance compared to alternative models.

Keywords: Higher-order moments; Markov-switching bilinear models; Stationarity.

1 Motivations

Markov-switching (*MS*) time series models proposed by Hamilton (1989) have recently received a growing interest in several areas of statistics because of their ability to describe adequately various financial time series and continue to gain more popularity, especially to model empirical macroeconomics and dynamic econometrics time series $(X_t)_{t\in\mathbb{Z}}$, $\mathbb{Z}=\{0,\pm1,\pm2,...\}$. The advantages of *MS* models are multiple, for instance, among others:
(i) *MS* models are nonlinear models (even locally) because linear models are not, in general, always suitable for use.
(ii) Higher flexibility in capturing the persistence and/or asymmetric effect in datasets.
(iii) In modeling time series which exhibit smooth or abrupt structure which occur frequently or occasionally depending on the transition probability of the chain.
So, there exist various different ways to model such a series exhibiting break changes through a finite number of "regimes". A rather general model is given by

$$X_t = f_{s_t}\left(X_{t-i}, e_{t-j}, 0 < i \le P, 0 < j \le Q\right) + e_t,$$

for some measurable function $f_{.}$ depending on some finite state Markov chain $(s_t)_t$ that control the change of regimes and some innovation process $(e_t)_t$. However, in literature, some locally (given s_t) linear or nonlinear models (determined by $f_{.}$) were investigated in order to study the probabilistic and statistical properties of such models. Indeed, the stationary $ARMA$ models in which the parameters are allowed to change through time according to Markov chain denoted by $MS-ARMA$ have, got considerable attention recently. For instance, Yang (2000), Francq and Zakoïan (2001), Stelzer (2009), Lee (2005), Yao and Attali (2000), Boubacar and Rabehasaina (2020) and the references therein, are references aims to describe the probabilistic and/or statistical properties of $MS-ARMA$ models. Namely, conditions assuring the strict stationarity, ergodicity, the existence of higher-order moments and spectral representation. On the other hand, to develop some appropriate statistical methods and their asymptotic inference. Francq and Zakoïan (2005), Bauwen et al. (2010), Alemohammad et al. (2020), Bibi and Ghezal (2015) and Wee et al. (2022) have proposed a $MS-GARCH$ model. So, the probabilistic structure, as in $MS-ARMA$ models were investigated and some procedures for estimating and forecasting the MS_GARCH model was studied. Recently, Bibi and Ghezal (2015) and Bibi and Hamdi (2025) and have introduced a new class of $MS-$bilinear ($MS-BL$ for short) model where several probabilistic properties were studied and explicit conditions ensuring the existence of a strictly stationary solution of such a model to belong in $\mathbb{L}_2$ are given. Moreover, it is also known that some bilinear processes have properties that are similar to those of an autoregressive conditionally heteroscedastic ($ARCH$) model, which plays important role in financial mathematics see subsection 3.1.

In this paper, a process $(X_t)_t$ defined on some probability space $(\Omega, \Im, P)$ is called $MS-BL(p,q,P,Q)$ if it generated by the following stochastic difference equation:

$$X_t = a_0(s_t) + \sum_{i=1}^{p} a_i(s_t)X_{t-i} + \sum_{j=0}^{q} b_j(s_t)e_{t-j} + \sum_{j=1}^{Q}\sum_{i=1}^{P} c_{ij}(s_t)X_{t-i}e_{t-j}. \quad (1)$$

In (1), the functions $a_i(s_t)$, $b_j(s_t)$ and $c_{ij}(s_t)$ depend upon a Markov chain $(s_t)_t$ that controls the dynamics of X_t and is subject to the following assumption:

The Markov chain $(s_t)_t$ is irreducible and aperiodic with a finite state space $\mathbb{S} = \{1, ..., d\}$, the $n-$step transition probability matrix, that determines the evolution in $\mathbb{S}$ is given by

$$P^{(n)} = \left(p_{ij}^{(n)}\right)'_{(i,j)\in\mathbb{S}\times\mathbb{S}},$$

where $p_{ij}^{(n)} = P(s_t = j|s_{t-n} = i)$, so by Chapman-Kolmogorov Equations $P^{(n)} = P^n$. The one-step transition probability matrix $P := (p_{ij})$ where $p_{ij} := p_{ij}^{(1)} = P(s_t = j|s_{t-1} = i)$ for $i, j \in \mathbb{S}$, such that $\sum_{j=1}^{d} p_{ij} = 1$, for all i. The stationary distribution of the Markov chain $(s_t)_t$ will be denoted by $\underline{\Pi} = (\pi(1), ..., \pi(d))'$ that solves the equation $P'\underline{\Pi} = \underline{\Pi}$ where $\pi(i) = P(s_0 = i)$, $i = 1, ..., d$. The chain $(s_t)_t$ is said to be independent if $p_{ij} = \pi(j)$ for all $i \in \mathbb{S}$.

The innovation process $(e_t)_t$ is assumed to be independent and identically distributed ($i.i.d$) with mean 0 and variance σ_2. In addition, we shall assume that e_t and $\{(X_{s-1}, s_t), s \leq t\}$ are independent.

As already pointed out by Bibi and Ghezal (2015), the $MS-BL(p,q,P,Q)$ includes as special cases several classes of interesting models having been investigated in the literature. It is worth noting that the key difference between $MS-BL$ and a threshold model is that the former assumes that the underlying state process that gives rise to the nonlinear dynamics (regime switching) is latent, whereas threshold models commonly allow the nonlinear effect to be driven by observable variables but assume the number of thresholds and the threshold values to be unknown. Before we proceed, we will first introduce some algebraic notations and definitions.

1.1 Algebraic notations and definitions

Some notations are used throughout the paper:

- For some specifications of the chain $(s_t)_t$, i.e., constant $(d=1)$, independent and dependent chain, in the sequel we shall indicate the corresponding models by $C-BL$, $I-BL$ and $D-BL$ respectively.
- $I_{(n)}$ is the $n\times n$ identity matrix, $\mathbb{I} := \left(I_{(s)} \vdots \ldots \vdots I_{(s)}\right)_{s\times ds}$, is the $d-$block matrix, and $\mathbf{1}_{(d)}$ (resp. $\underline{\mathbf{1}}_{(d)}$) is a $d\times d$ matrix (resp. $d\times 1$ vector) whose components are matrix unity (resp. 1).
- $O_{(k,l)}$ denotes the matrix of order $k\times l$ whose entries are zeros, for simplicity we set $O_{(k)} := O_{(k,k)}$ and $\underline{O}_{(k)} := O_{(k,1)}$.
- The spectral radius of a square matrix M is denoted by $\rho(M)$, $\|.\|$ denotes any operator norm on the set of $m\times n$ and $m\times 1$ matrices, $\otimes$ is the usual Kronecker product of matrices and $M^{\otimes r} = M\otimes M\otimes \ldots \otimes M$, $r-times$. If $(M(i), i\in I)$ is $n\times n$ matrices sequence, we shall note for any integer l and j, $\prod_{i=l}^{j} M(i) = M(l)M(l+1)\ldots M(j)$ if $l\leq j$ and $I_{(n)}$ otherwise.
- For any function $f:\mathbb{S}\to \mathscr{M}_{n\times m}(\mathbb{R})$, where $\mathscr{M}_{n\times m}(\mathbb{R})$ denotes the space of real $n\times m$ matrices, we write

$$\mathbb{P}(f^{\otimes r}) = \begin{pmatrix} p_{11}f^{\otimes r}(1).. & p_{d1}f^{\otimes r}(1) \\ \vdots & \vdots \\ p_{1d}f^{\otimes r}(d).. & p_{dd}f^{\otimes r}(d) \end{pmatrix}, \ \mathbb{P}^{(n)}(f^{\otimes r}) = \begin{pmatrix} p_{11}^{(n)}f^{\otimes r}(1).. & p_{d1}^{(n)}f^{\otimes r}(1) \\ \vdots & \vdots \\ p_{1d}^{(n)}f^{\otimes r}(d).. & p_{dd}^{(n)}f^{\otimes r}(d) \end{pmatrix}$$

  and

$$\underline{\Pi}(f^{\otimes r}) = \begin{pmatrix} \pi(1)f^{\otimes r}(1) \\ \vdots \\ \pi(d)f^{\otimes r}(d) \end{pmatrix}.$$

  In the sequel, we will use the following result due to Francq and Zakoïan (2005) stated in the next lemma.

Lemma 1. *1. For $i\geq 1$, if $(Z_{t-i})_t$ is an integrable random variable belonging to $\sigma(e_{t-s}, s\geq i)$, then*

$$\pi(k)E\{Z_{t-i}|s_t=k\} = \sum_{j=1}^{d} E\{Z_{t-i}|s_{t-i}=j\}\, p_{jk}^{(i)}\pi(j).$$

2. If $f:\mathbb{S}\to\mathscr{M}_{n\times n}(\mathbb{R})$ and $g:\mathbb{S}\to\mathscr{M}_{n\times 1}(\mathbb{R})$, then for any $k>0$ and $\tau>k$,

$$E\left\{f(s_t)f(s_{t-1})\ldots f(s_{t-k+1})\underline{g}(s_{t-k})|s_{t-k}\right\} = \mathbb{I}\{\mathbb{P}(f)\}^k\,\underline{\Pi}(\underline{g}).$$

Remark 1. *[Independent switching]. For the model $I-BL$, then $(\mathbb{P}(f))^k = \mathbb{P}\left(f(.)\,(E\{f(s_t)\})^{k-1}\right)$ and $\mathbb{I}\mathbb{P}(f) = [E\{f(s_t)\}, \ldots, E\{f(s_t)\}]$ where in the last equality, the matrix $E\{f(s_t)\}$ is duplicated $d-$times.*

We arrange the rest of the paper in the following manner. In the next section, we give the Markovian state-space representation, which is used to derive conditions for stationarity (in the strong and weak sense) and a recursive formula for the higher-order moments. Section 3 is devoted to giving explicit expressions of the second, third and fourth-order moments of some particular $MS-BL$ models. Section 4 provides results of a Monte Carlo experiment and section 5 concludes the paper.

2 Markovian representation of $MS-BL$ and its properties

In the rest of the paper, we shall restrict ourselves to the case when $b_j(.)=0,\ j=1,...,q$ in (1), i.e., without moving average part

$$X_t = \sum_{i=1}^{p} a_i(s_t)X_{t-i} + \sum_{j=1}^{Q}\sum_{i=1}^{p} c_{ij}(s_t)X_{t-i}e_{t-j} + e_t, \tag{2}$$

which denoted also hereafter $MS-BL(p,0,p,Q)$. Because it is difficult to handle the product terms like $X_te_{t-j},\ j>0$, Liu (1992) when $d=1$, introduced the so-called lower triangular model for which $c_{ij}(.)=0$, if $i<j$. So in this paper we extend the lower tridiagonal model to include a $MS-BL$ one, i.e.,

$$X_t = \sum_{i=1}^{p} a_i(s_t)X_{t-i} + \sum_{j=1}^{Q}\sum_{i=j}^{p} c_{ij}(s_t)X_{t-i}e_{t-j} + e_t. \tag{3}$$

The representation (3) has however the advantage to admit the state-space representation $X_t = \underline{H}'\underline{Y}_t$ where

$$\underline{Y}_t = \Gamma_{s_t}(e_t)\underline{Y}_{t-1} + \underline{\eta}(e_t), \tag{4}$$

with $\Gamma_{s_t}(e_t) = \Gamma_0(s_t) + e_t\Gamma_1(s_t)$ where the matrices $\Gamma_0(s_t)$, $\Gamma_1(s_t)$ and the vectors $\underline{\eta}(e_t)$, $\underline{H}$ are explicitly given in Bibi and Ghezal (2015). Note that the representation (4) shows that the $MS-BL(p,0,p,Q)$ can be represented as a multidimensional first-order random coefficient Markov-switching Autoregressive ($MS-RCAR(1)$) model, and hence the extended process $\underline{Z}_t := (\underline{Y}_t', s_t)'$, is an aperiodic Markov chain on $(\mathscr{X}, \mathscr{B}(\mathscr{X}))$, where $\mathscr{X} = \mathbb{R}^s \times \mathbb{S}$. This compact representation allows us to provide a necessary and sufficient condition for the existence of strict stationery, ergodic, and $\Im_t = \sigma(e_{t-n}, s_{t-n}, n \geq 0)-$measurable (or causal) solutions for (3). These concepts are ensured under the strict negativeness of the Lyapunove exponent defined by $\gamma(\Gamma) := \inf\limits_{t\geq 1} \frac{1}{t} E\left\{\log\left\|\prod\limits_{i=0}^{t-1}\Gamma_{s_{t-i}}(e_{t-i})\right\|\right\}$ (the chosen of the norm is unimportant in this definition). However, it follows from Bibi and Ghezal (2015) that the unique strictly stationary and ergodic solution of (4) is given by the first component of

$$\underline{Y}_t = \sum_{k=0}^{\infty}\left\{\prod_{i=0}^{k-1}\Gamma_{s_{t-i}}(e_{t-i})\right\}\underline{\eta}(e_{t-k}),\ \text{a.s.}, \tag{5}$$

with the usual convention $\prod\limits_{i=k}^{k'} = 1$, whenever $k'<k$.

Remark 2. *For the process $(X_t)_t$ defined by $X_t = c_{11}(s_t)X_{t-1}e_{t-1} + e_t$, the Lyapunov exponent is*

$$\sum_{i=1}^{d}\pi(i)E\{\log(|c_{11}(i)e_0|)\},$$

so $\prod\limits_{i=1}^{d}|c_{11}(i)|^{\pi(i)} < e^{-E\{\log|e_0|\}}$ constitutes the necessary and sufficient condition for strict stationarity and ergodic solution. Note that when $(e_t)_t$ is Gaussian, the necessary and sufficient condition reduces to $\prod\limits_{i=1}^{d}|c_{11}(i)|^{\pi(i)} < 1.88736$. It is worth noting that this example shows the existence of explosive regimes (i.e. ,$|c_{11}(i)|>1$ compared with the standard case) does not preclude the strict stationarity.

Recalling here (interested readers are advised to see Bibi and Ghezal (2015) for more details) that if $\sigma_{2m+1} = E\left\{e_t^{2m+1}\right\} = 0$, $\sigma_{2m} = E\left\{e_t^{2m}\right\} < +\infty$ for any integer m, and if

$$\lambda_{(m)} := \rho\left(\mathbb{P}\left(\Gamma^{\otimes m}\right)\right) < 1, \tag{6}$$

where $\Gamma^{\otimes m} := (\Gamma^{\otimes m}(i), 1 \le i \le d)$ with $\Gamma^{\otimes m}(.) = E\left\{\Gamma_{s_t}^{\otimes m}(e_t)|s_t = .\right\}$, then equation (4) has a unique strictly stationary solution $(\underline{Y}_t)_t$ given by (5) and the process $(\underline{H}'\underline{Y}_t)_t$ is also a unique, strictly stationary, causal and ergodic solution of equation (3) which satisfies $X_t^{2m} \in \mathbb{L}_{2m}$.

Remark 3. *For the model $I-BL$ the condition for the existence a strictly stationary solution in $\mathbb{L}_m$ is that $\lambda_{(m)} := \rho\left(E\left\{\Gamma_{s_t}^{\otimes m}\right\}\right) < 1$ where $E\left\{\Gamma_{s_t}^{\otimes m}\right\} = \sum\limits_{k=1}^{d} \Gamma^{\otimes m}(k)\pi(k)$.*

2.1 Computation of the higher-order moments

Once the second-order stationarity condition is established, it is useful to compute the expectation and some cumulants of the process $(\underline{Y}_t)_t$. A property which will be heavily used in the sequel, associated with representation (4) is given in the following lemma.

Lemma 2. *Consider the representation (4), then for any integer $m \ge 0$, we have*

$$\underline{Y}_t^{\otimes m} = \sum_{i=0}^{m} \Psi_i^{(m)}(s_t, e_t)\,\underline{Y}_{t-1}^{\otimes i}, \tag{7}$$

where the matrices $\Psi_i^{(m)}(s_t, e_t)'s$ are uniquely determined by $\Gamma_{s_t}(e_t)$ and $\underline{\eta}(e_t)$ according to the following recursions

$$\Psi_0^{(m)}(s_t, e_t) = \underline{\eta}^{\otimes m}(e_t),\ \Psi_1^{(1)}(s_t, e_t) = \Gamma_{s_t}(e_t), \quad \textit{and for any } m > 0,$$
$$\Psi_i^{(m+1)}(s_t, e_t) = \underline{\eta}(e_t) \otimes \Psi_i^{(m)}(s_t, e_t) + \Gamma_{s_t}(e_t) \otimes \Psi_{i-1}^{(m)}(s_t, e_t),$$

with the convention $\Psi_i^{(m)}(s_t, e_t) = 0$ when $i > m$ or $i < 0$ and $\underline{Y}_t^{\otimes 0} = \Psi_0^{(0)}(s_t, e_t) = 1$.

Proof. The formula (7) for $m = 1$ is given by (4). Assuming that (7) holds true for some $m \ge 1$, then we have

$$\begin{aligned}
\underline{Y}_t^{\otimes(m+1)} &= \sum_{i=0}^{m} \left(\Gamma_{s_t}(e_t)\underline{Y}_{t-1} + \underline{\eta}(e_t)\right) \otimes \Psi_i^{(m)}(s_t, e_t)\,\underline{Y}_{t-1}^{\otimes i} \\
&= \sum_{i=0}^{m} \left\{\underline{\eta}(e_t) \otimes \left\{\Psi_i^{(m)}(s_t, e_t)\,\underline{Y}_{t-1}^{\otimes i}\right\} + \left(\Gamma_{s_t}(e_t) \otimes \Psi_i^{(m)}(s_t, e_t)\right)\left(\underline{Y}_{t-1} \otimes \underline{Y}_{t-1}^{\otimes i}\right)\right\} \\
&= \sum_{i=0}^{m+1} \left\{\left\{\underline{\eta}(e_t) \otimes \Psi_i^{(m)}(s_t, e_t)\right\}\underline{Y}_{t-1}^{\otimes i} + \left\{\Gamma_{s_t}(e_t) \otimes \Psi_{i-1}^{(m)}(s_t, e_t)\right\}\underline{Y}_{t-1}^{\otimes i}\right\}
\end{aligned}$$

and the result follows. □

Now, we note $\underline{U}^{(m)} = \left(\pi(k)E\left\{\underline{Y}_t'^{\otimes m}|s_t = k\right\}, k = 1, .., d\right)'$, then

$$\begin{aligned}
\pi(k)E\left\{\underline{Y}_t^{\otimes m}|s_t = k\right\} &= \sum_{i=0}^{m} \Psi_i^{(m)}(k)E\left\{\underline{Y}_{t-1}^{\otimes i}|s_t = k\right\}\pi(k) \\
&= \sum_{i=0}^{m} \Psi_i^{(m)}(k)\textstyle\sum_{j=1}^{d} E\left\{\underline{Y}_{t-1}^{\otimes i}|s_{t-1} = j\right\}p_{jk}\pi(j),
\end{aligned} \tag{8}$$

where $\Psi_i^{(m)}(k) = E\left\{\Psi_i^{(m)}(s_t, e_t)|s_t = k\right\}$, so $\underline{U}^{(m)} = \sum\limits_{i=0}^{m} \mathbb{P}\left(\Psi_i^{(m)}(k)\right)\underline{U}^{(i)}$, hence under the condition (6), $E\left\{\underline{Y}_t^{\otimes m}\right\} = \mathbb{I}\underline{U}^{(m)}$ and $\underline{U}^{(m)} = \left(I_{(ds^m)} - \mathbb{P}\left(\Psi_m^{(m)}(k)\right)\right)^{-1}\left\{\sum\limits_{i=0}^{m-1} \mathbb{P}\left(\Psi_i^{(m)}(k)\right)\underline{U}^{(i)}\right\}$.

Example 1. Explicit formulae for some power m are addressed below

$$\Psi_i^{(1)}(k) = \begin{cases} \underline{\eta}^{(0)}, & i=0, \\ \Gamma_0(k), & i=1. \end{cases}$$

$$\Psi_i^{(2)}(k) = \begin{cases} \underline{\eta}^{(0)\otimes 2}, & i=0, \\ \underline{\eta}^{(0)} \otimes \Gamma_0(k) + \underline{\eta}^{(1)} \otimes \Gamma_1(k) + \Gamma_0(k) \otimes \underline{\eta}^{(0)} + \Gamma_1(k) \otimes \underline{\eta}^{(1)}, & i=1, \\ \Gamma_0^{\otimes 2}(k) + \sigma_2 \Gamma_1^{\otimes 2}(k), & i=2. \end{cases}$$

For $m=3$, the non-zeros terms in $\Psi_i^{(3)}(k)$, $i=0,1,2$ and 3 are

$$\Psi_i^{(3)}(k) = \begin{cases} \underline{\eta}^{(0)\otimes 3}, & i=0, \\ \underline{\eta}^{(0)\otimes 2} \otimes \Gamma_0 k) + \underline{\eta}^{(1)\otimes 2} \otimes \Gamma_1 k) + \underline{\eta}^{(0)} \otimes \Gamma_0(k) \otimes \underline{\eta}^{(0)} + \underline{\eta}^{(0)} \otimes \Gamma_1(k) \otimes \underline{\eta}^{(1)(0)} & \\ +\Gamma_0(k) \otimes \underline{\eta}^{(0)} + \Gamma_1(k) \otimes \underline{\eta}^{(1)}, & i=1, \\ \underline{\eta}^{(0)} \otimes \Gamma_0^{\otimes 2}(k) + \underline{\eta}^{(2)} \otimes \Gamma_1^{\otimes 2}(k) + \underline{\eta}^{(1)} \otimes (\Gamma_0(k) \otimes \Gamma_1(k) + \Gamma_1(k) \otimes \Gamma_0(k)) & \\ +\Gamma_0(k) \otimes \underline{\eta}^{(0)} \otimes \Gamma_0(k) + \Gamma_0(k) \otimes \underline{\eta}^{(1)} \otimes \Gamma_1(k) + \Gamma_1(k) \otimes \underline{\eta}^{(1)} \otimes \Gamma_0(k) & \\ +\Gamma_1(k) \otimes \underline{\eta}^{(2)} \otimes \Gamma_1(k) & i=2, \\ \left(\Gamma_0^{\otimes 2}(k) + \sigma_2 \Gamma_1^{\otimes 2}(k)\right) \otimes \Gamma_0(k) + \sigma_2 (\Gamma_0(k) \otimes \Gamma_1(k) + \Gamma_1(k) \otimes \Gamma_0(k)) \otimes \Gamma_1(k), & i=3 \end{cases}$$

where $\underline{\eta}^{(i)} = E\left\{e_t^i \underline{\eta}(e_t)\right\}$, $i=0,1,\ldots..$ The fourth$-$order moments is also useful to study which may be deduced from the recursion $\underline{U}^{(m)}$.

Note here that the recursion $\underline{U}^{(m)}$ demonstrates the dependence of the $m-th$ moment terms on the $(m-1)th-$moment. Provided that the $(m-1)th-$moment.converges, and the summation in (8) is also stable, this term will converge. Moreover, we note that the matrices $\Psi_i^{(m)}(k)$ are not zero for any $k \in \mathbb{S}$.

3 Case studies

In this section we examine the following particular cases of $MS-BL$ models (3), i.e.,

$$\begin{cases} \text{(Superdiagonal model):} & \mathrm{X}_t = a(s_t) X_{t-l} e_{t-k} + e_t,\ 0<k<l, & \text{(MS-SBL)} \\ \text{(Diagonal model):} & \mathrm{X}_t = a(s_t) X_{t-k} e_{t-k} + e_t,\ 0<k, & \text{(MS-DBL)} \end{cases} \tag{9}$$

where for the sake of generality it is assumed that $(e_t)_t$ is $i.i.d.$, Gaussian sequence, and we set $\sigma_r = E\left\{e_t^r\right\}$. So, the even moments of $(e_t)_t$ up to eight order are $\sigma_2 = \sigma^2$, $\sigma_4 = 3\sigma_2^2$, $\sigma_6 = 15\sigma_2^3$, and $\sigma_8 = 105\sigma_2^4$ while all the odd moments are equal to zero. The above models was studied in non switching framework (i.e., $d=1$) by, among others, Gabr (1988) and Martin (1999) who gave some general discussions on the properties of such models including stationarity and moments properties. Moreover, Grange and Anderson (1978) studied the superdiagonal in detail and showed that this series might be mistaken as a white noise. In the following we give an explicit expression of the higher-order moments of the above models. Let us consider

$$\gamma_X(i) = E\left\{(X_t - \mu)(X_{t-i} - \mu)\right\} = \mu_X(i) - \mu^2, \tag{10}$$

$$\gamma_X(i,j) = E\left\{(X_t - \mu)(X_{t-i} - \mu)(X_{t-j} - \mu)\right\} = \mu_X(i,j) - \mu\left\{\mu_X(i) + \mu_X(j) + \mu_X(i-j)\right\} + 2\mu^3, \tag{11}$$

where $\mu = E\{X_t\}$, $\mu_X(i) = E\{X_t X_{t-i}\}$ and $\mu_X(i,j) = E\left\{X_t X_{t-i} X_{t-j}\right\}$. In this illustration, the autocorrelations of $\left(X_t^2\right)$ noted $\rho_2(i) = (\mu_{(2)}(i) - \mu_X^2(0))/Var\left\{X_t^2\right\}$ where $\mu_{(2)}(i) = E\left\{X_t^2 X_{t-i}^2\right\}$, it be use as a power criterion for bilinear model identification, which often replaced by a standardized third central moments, i.e.,

$$\begin{aligned}\rho_3(i,j) &= E\left\{(X_t-\mu)(X_{t-i}-\mu)(X_{t-j}-\mu)\right\}/\left\{Var\left\{X_t\right\}\right\}^{3/2}\\ &= \left\{\mu_X(i,j)-\mu\left\{\mu_X(j)+\mu_X(i)+\mu_X(i-j)\right\}+2\mu^3\right\}/\gamma_X^{3/2}(0).\end{aligned}$$

The last moment has several features with respect to $\rho_2(i)$. For instance, the coefficients of skewness of $(X_t)_t$ may be derived from $\rho_3(i,j)$. Moreover, if $\left(X_t^2\right)_t$ is generated by a (locally) linear model then $\rho_2(i)$ can be nonzero whereas $\rho_3(i,j)$ will be zero for all lags i and j. Also $\rho_3(i,j)$ should provide more information about the model under study than the autocorrelation of $\left(X_t^2\right)_t$. Note that from Gabr (1988), the following symmetry relations hold $\gamma_X(i,j)=\gamma_X(j,i)=\gamma_X(-i,j-i)=\gamma_X(i-j,-j)$ where $i,\ j\in\mathbb{Z}$. So, it is sufficient to calculate $\gamma_X(i,j)$ for $0\le i\le j$. In the sequel we shall note for any $i\ge 1$, $a^i=\left(a^i(j),j\in\mathbb{S}\right)$ and $\prod\limits_{j=1}^{i}\mathbb{P}^{(k)}(.)=\mathbb{P}^{i(k)}(.)$.

3.1 Superdiagonal model

Note that the $MS-SBL$ model may be written as $X_t=a(s_t)X_{t-k}e_{t-k+m}+e_t, 2\le k, 1\le m\le k-1$, which is conditionally heteroscedastic (but not a $MS-ARCH$ model). Indeed, for $k=2$, let $h_t(l):=E\left\{\pi(l)X_t^2 I_{\{s_t=l\}}|\Im_{t-2}\right\}$, $l\in\mathbb{S}$, then we have

$$\begin{aligned}h_t(l) &= \sigma_2 a^2(l)\pi(l)E\left\{X_{t-2}^2 I_{\{s_t=l\}}|\Im_{t-2}\right\}+\pi(l)\sigma_2\\ &= \sigma_2 a^2(l)\sum_{l'=1}^{d}E\left\{X_{t-2}^2 I_{\{s_{t-2}=l'\}}|\Im_{t-2}\right\}P_{ll'}^{(2)}\pi(l')+\pi(l)\sigma_2\\ &= \sigma_2 a^2(l)X_{t-2}^2\sum_{l'=1}^{d}I_{\{s_{t-2}=l'\}}P_{ll'}^{(2)}\pi(l')+\pi(l)\sigma_2.\end{aligned}$$

So, $h_t=\sum_{l=1}^{d}h_t(l)$ is given by $h_t=\left\{\underline{\mathbf{1}}\mathbb{P}^{(2)}\left(\sigma_2 a^2\right)\underline{Z}_{t-2}\right\}X_{t-2}^2+\sigma_2$ where $\underline{Z}_t=\left(I_{\{s_t=1\}},\ldots,I_{\{s_t=d\}}\right)'$, this show that $MS-SBL$ is an $ARCH$ model. From the discussion in subsection 2.1, it can be shown that a sufficient condition for the strict stationarity solution in $\mathbb{L}_{2m}$ is that $\rho\left(\mathbb{P}^{(k)}\left(\sigma_{2m}a^{2m}\right)\right)<1$, $m\ge 1$.

Lemma 3. [*First and second-order moments*] *For the $MS-SBL$, assume that $\rho\left(\mathbb{P}^{(k)}\left(\sigma_2 a^2\right)\right)<1$, which ensures the second-order stationarity, then we have*

1. $\mu_{(1)}(k)=E\left\{X_t\right\}=0$, $E\left\{X_t e_t^2\right\}=0$, $E\left\{X_t^2 e_t\right\}=0$, $E\left\{X_t^2 e_t^3\right\}=0$, $E\left\{X_t e_t\right\}=\sigma_2$, $E\left\{X_t e_t^3\right\}=\sigma_4$,
2. $\mu_{(2)}(k)=E\left\{X_t^2\right\}=\mathbf{1}\underline{\mu}_{(2)}(k)$ *where* $\underline{\mu}_{(2)}(k)=\left(I_{(d)}-\mathbb{P}^{(k)}\left(\sigma_2 a^2\right)\right)^{-1}\underline{\Pi}(\sigma_2)$,
3. $\widetilde{\mu}_{(2)}(k)=E\left\{X_t^2 e_t^2\right\}=\mathbf{1}\underline{\widetilde{\mu}}_{(2)}(k)$ *where* $\underline{\widetilde{\mu}}_{(2)}(k)=\sigma_2^2\mathbb{P}^{(k)}\left(a^2\right)\left(I_{(d)}-\mathbb{P}^{(k)}\left(\sigma_2 a^2\right)\right)^{-1}\underline{\Pi}(\sigma_2)+\underline{\Pi}(\sigma_4)$.
4. $\gamma_X(i)=\mu_{(2)}(k)\delta_{\{i=0\}}$.

Proof. The proof is straightforward and hence omitted. □

The nullity of $\gamma(i)$ for $i>0$, shows that the $MS-SBL$ appears as a sequence of weak white noise ($0-$dependent). So, in order to provide more informations, we need to investigate some moments of order greater than 2.

Lemma 4. *Under the condition of Lemma 3, all the third-order moments $\gamma(i,j)$ of the $MS-SBL$, are equal to zero except $i=k-m$ and $j=k$ for which we have*

$$\gamma_X(k-m,k)=\underline{\mathbf{1}}\mathbb{P}^{(k-m)}(a)\,\mathbb{P}^{(m)}(\sigma_2)\left(I_{(d)}-\mathbb{P}^{(k)}\left(\sigma_2a^2\right)\right)^{-1}\underline{\Pi}(\sigma_2).$$

Proof. It is easy to see that $\gamma_X(0,0)=0$. For $i=k-m, j=k$, we have

$$\begin{aligned}\pi(l)\,E\left\{X_tX_{t-k}X_{t-k+m}|s_t=l\right\}&=\pi(l)\,a(l)\,E\{X_{t-k+m}e_{t-k+m}X_{t-k}^2|s_t=l\}\\&=\underline{\mathbf{1}}\mathbb{P}^{(k-m)}(a)\,\underline{U},\end{aligned}$$

where

$$\underline{U}=\left(E\left\{X_{t-k+m}e_{t-k+m}X_{t-k}^2|s_{t-k+m}=l\right\}\pi(l),l=1,..,d\right)'=\mathbb{P}^{(m)}\left(\sigma^2\right)E\{X_t^2|s_t=u\}\pi(u),u=1,..,d)'.$$

So,

$$\gamma_X(k-m,k)=\underline{\mathbf{1}}\mathbb{P}^{(k-m)}(a)\,\mathbb{P}^{(m)}(\sigma_2)\left(I_{(d)}-\mathbb{P}^{(k)}\left(\sigma_2a^2\right)\right)^{-1}\underline{\Pi}(\sigma_2).$$

□

Lemma 5. *For the $MS-SBL$, we have under the condition $\rho\left(\mathbb{P}^{(k)}\left(\sigma_{2l}a^{2l}\right)\right)<1$, $l\geq1$.*

1. $\mu_{(r)}(k)=E\{X_t^r\}=\underline{\mathbf{1}}\underline{\mu}_{(r)}(k)\,I_{\{r=2l\}}$ *where*

$$\underline{\mu}_{(2l)}(k)=\left(I_{(d)}-\mathbb{P}^{(k)}\left(\sigma_{2l}a^{2l}\right)\right)^{-1}\left(\sum_{i=1}^{l-1}C_{2l}^{2i}\sigma_{2(l-i)}\mathbb{P}^{(k)}\left(\sigma_{2i}a^{2i}\right)\underline{\mu}_{(2i)}(k)+\underline{\Pi}(\sigma_{2l})\right).$$

2. $\widetilde{\mu}_{(r)}(k)=E\{(X_te_t)^r\}=\underline{\mathbf{1}}\underline{\widetilde{\mu}}_{(r)}(k)$ *where* $\underline{\widetilde{\mu}}_{(r)}(k)=\sigma_{2r}+\underline{\mathbf{1}}\left(\sum_{i=1}^{\left[\frac{r}{2}\right]}C_r^{2i}\mathbb{P}^{(k)}\left(\sigma_{2r}\,a^{2i}\right)\underline{\mu}_{(2i)}(k)\right)$, $[x]$ *is the integer part of x.*

Proof. 1. For the first assertion, we have for any integer $r\geq1$,

$$X_t^r=e_t^r+\sum_{i=1}^{r}C_r^ia^i(s_t)\,X_{t-k}^ie_{t-k+m}^ie_t^{r-i}.$$

So if $r=2l+1$, then due to Gaussianity of $(e_t)_t$, the expectation of the first term is 0 and because if i is even integer, $r-i$ is odd and vice versa. Therefore the expectation of the second sum is also 0. In contrast if $r=2l$ say, then

$$\begin{aligned}E\left\{X_t^{2l}|s_t=j\right\}\pi(j)&=E\left\{e_t^{2l}|s_t=j\right\}\pi(j)+\sum_{i=1}^{l}C_{2l}^{2i}a^{2i}(j)\,E\left\{X_{t-k}^{2i}|s_t=j\right\}\pi(j)\,\sigma_{2i}\sigma_{2(l-i)}\\&=\sigma_{2l}\pi(j)+\sum_{i=1}^{l}C_{2l}^{2i}a^{2i}(j)\sum_{u=1}^{d}E\left\{X_{t-k}^{2i}|s_{t-k}=u\right\}p_{uj}^{(k)}\pi(u)\,\sigma_{2i}\sigma_{2(l-i)}.\end{aligned}$$

By passing to vectorial representation the results follows. In particular,

$$\mu_{(4)}(k)=E\left\{X_t^4\right\}=\underline{\mathbf{1}}\left(I_{(d)}-\mathbb{P}^{(k)}\left(\sigma_4a^4\right)\right)^{-1}\left(6\sigma_2^2\mathbb{P}^{(k)}\left(a^2\right)\underline{\mu}_{(2)}(k)+\underline{\Pi}(\sigma_4)\right).$$

2. The proof of the second assertion is similar and is omitted.

□

Now, we carry out a second-order analysis of the squares ($Y_t=X_t^2$) of $MS-SBL$ which providing an alternate differentiation technique characterizing the processes 0−dependent

Lemma 6. *For the second-order stationary $MS-SBL$, consider the process $Y_t = X_t^2$ and assume (to facilitate the exposition) that $k = 2, m = 1$ and for any $j \geq 0$, we set*

$$\underline{\mu}_Y(j) = \left(\pi(u) E\left\{Y_t Y_{t-j} | s_t = u\right\}, u = 1, \ldots d\right)',$$

then under the condition $\rho\left(\mathbb{P}^{(2)}\left(\sigma_2 a^2\right)\right) < 1$ we have

$$\underline{\mu}_Y(j) = \begin{cases} \underline{\mu}_{(4)}(2) & \text{if } j = 0 \\ \left(I_{(d)} - \mathbb{P}\left(\sigma_2 a^2\right)\right)^{-1} \mathbb{P}(\sigma_2) \underline{\mu}_2(2) & \text{if } j = 1 \\ \mathbb{P}^{(2)}\left(\sigma_2 a^2\right) \underline{\mu}_Y(n-2) + \mathbb{P}^{(n)}(\sigma_2) \underline{\mu}_2(2) & \text{if } j \geq n \geq 2 \end{cases},$$

and hence $\gamma_Y(j) = \mathbf{1}\underline{\mu}_Y(j) - \mu_2^2(2)$ for any $j \geq 0$.

Proof. The first part is obtained from the Lemma 5. For the second, $\pi(u) E\{Y_t Y_{t-1} | s_t = u,\} = a^2(u) \pi(u) E\{X_{t-2}^2 X_{t-1}^2$ $e_{t-1}^2 | s_t = u\} + \sigma_2 \pi(u) E\left\{X_{t-1}^2 | s_t = u\right\}$, so

$$\underline{\mu}_Y(1) = \mathbb{P}\left(\sigma_2 a^2\right) \underline{\mu}_Y(1) + \mathbb{P}\left(a^2\right) \underline{\mu}_{(2)}(2).$$

The third one is immediate. □

Remark 4. *According to Brockwell and Davies (1987) Section 3.2, it follows that if $(X_t)_t$ is a is $q-$dependent and $(e_t)_t$ is i.i.d. then for any measurable functions of $(X_t)_t$ is $q-$dependent. So, since because $\gamma_Y(j) \neq 0$ implies that the process $(Y_t)_t$ is not an $0-$dependent and hence $(X_t)_t$ is fare from an $MA(1)$ process.*

The next corollary we summarize some results listed above when $(s_t)_t$ is independent sequence noted hereafter $I-SBL$. The proof of this corollary is omitted as it consists of straightforward, but tedious algebra.

Corollary 1. *[Switching-independent case] Set $\overline{a} = E\{a(s_t)\}$ and $\lambda = \overline{a^2}\sigma_2$, then, for the $I-SBL$, we have:*

1. $\mu = 0$.
2. $\gamma_X(i) = \sigma_2 (1-\lambda)^{-1} \delta_{\{i=0\}}$.
3. $\gamma_X(i,j) = \sigma_2^2\, \overline{a} (1-\lambda)^{-1} \delta_{\{i=k-m, j=k\}}$.
4. $E\{X_t^r\} = \left(\sigma_{2l} + \sum\limits_{i=1}^{l-1} C_{2l}^{2i} \sigma_{2i} \sigma_{2(l-i)} \overline{a^{2i}} E\left\{X_t^{2i}\right\}\right) \left(1 - \sigma_{2l}\, \overline{a^{2l}}\right)^{-1} \delta_{\{r=2l\}}$.
5. $\gamma_Y(j) = \begin{cases} \dfrac{2\sigma_2}{\left(1-\sigma_2 \overline{a^2}\right)\left(1-\sigma_4 \overline{a^2}\right)}, & j=0, \\ \dfrac{2\sigma_2^3 \overline{a^2}}{\left(1-\sigma_2 \overline{a^2}\right)^2}, & j=1, \\ \sigma_2 \overline{a^2} \gamma_Y(n-2), & j=n \geq 2. \end{cases}$

Remark 5. *The results of the fifth assertion of above corollary shows clearly that when $(s_t)_t$ is an independent sequence, the process $\left(X_t^2\right)_t$ associated to $k = 2$ and $m = 1$ has the same covariance structure of an $MS-ARMA(2,1)$, with an independent sequence $(s_t)_t$, i.e.,*

$$Z_t = \omega(s_t) + a_1(s_t) Z_{t-1} + a_2(s_t) Z_{t-2} + b(s_t) e_{t-1} + e_t \text{ with } a_1(.) = 0. \quad (12)$$

Indeed, the covariance structure of (12) *is given by* $\gamma_Z(j) = \mathbf{1}\underline{\mu}_Z(j) - \mu_Z^2$ *where*

$$\underline{\mu}_Z = \mathbf{1}\left(I_{(d)} - \mathbb{P}^{(2)}(a)\right)^{-1}\underline{\Pi}(\omega)$$

and

$$\underline{\mu}_Z(j) = \begin{cases} \left(I_{(d)} - \mathbb{P}^{(2)}(a_2^2)\right)^{-1}\left(\underline{\Pi}(\omega^2) + 2\mathbb{P}^{(2)}(\omega a_2)\underline{\mu}_Z + \underline{\Pi}(\sigma^2 b) + \underline{\Pi}(\sigma_2)\right), & j = 0, \\ \left(I_{(d)} - \mathbb{P}(a)\right)^{-1}\left(\mathbb{P}(b)\underline{\Pi}(\sigma_2) + \mathbb{P}(\omega)\underline{\mu}_Z\right), & j = 1, \\ \mathbb{P}(a_2)\underline{\mu}_Z(n-2) + \mathbb{P}^{(n)}(\omega)\underline{\mu}_Z, & j = n > 1, \end{cases}$$

which is reduces in $I-SBL$ *to* $\gamma_Z(j) = \mu_Z(j) - \mu_Z^2$ *where* $\mu_Z = (1-\overline{a_2})^{-1}\ \overline{\omega}$ *and*

$$\gamma_Z(j) = \begin{cases} \sigma_2\left(1 - \overline{a_2^2}\right)^{-1}(1 + \overline{b^2}), & j = 0, \\ \sigma_2\overline{b^2}(1 - \overline{a_2})^{-1}, & j = 1, \\ \overline{a_2}\gamma_Z(j-2), & j \geq 2. \end{cases}$$

Comparing $\gamma_Z(j)$ *and that given in fifth assertion of corollary* 1*, it can be shown that for* $\overline{\omega} = (1-\overline{a})\mu_{(2)}(k)$, $\overline{a_2} = \sigma_2\overline{a^2}$ *and* $\overline{b^2} = 1 + \sigma_4\overline{a^2}$, $\gamma_Z(j)$ *and* $\gamma_Y(j)$ *coincide. This finding excludes the* $I-SBL$ *to being a weak white noise.* ($0-$*dependent). Moreover, the results of parts* 1,2 *and* 3 *of the above Corollary are identical to the results given by Gabr (1988) when* $d = 1$. *Furthermore, it is easy to see that the results in part* 4 *of Corollary 1 are in agreement with the results given by Martin (1999).*

Example 2. Consider $d = 2$, the entries of the transition matrix P are $p_{11} = \alpha$, $p_{22} = \beta$, the ergodic distribution is then $\pi(1) = (1-\alpha)/(2-\alpha-\beta)$. The kurtosis (in terms of $a(1)$ and $a(2)$) of $MS-SBL$ are shown in the Figure 1. The kurtosis of $MS-SBL$ in above figure are strictly positive and hence their distributions are leptokurtic i.e., the tails are thicker than normal one. Note here that all calculations thereafter have been performed with native code under $MATLAB$'15 computation language.

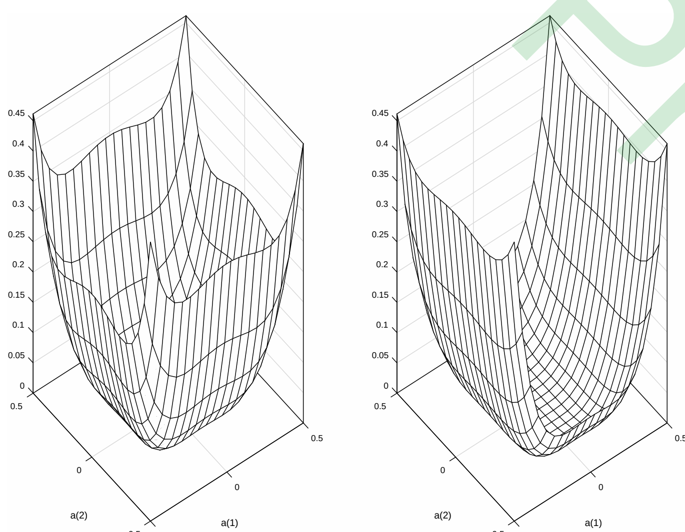

Figure 1: Kurtosis of the two-state $MS-SBL$. Left dependent $D-SBL$ and right $I-SBL$ with $\alpha = 0.75$ and $\beta = 0.95$.

3.2 Diagonal model

Lemma 7. *[Second-order moments] If $(X_t)_t$ is generated by a $MS-DBL$, then under the condition $\rho\left(\mathbb{P}^{(k)}\left(\sigma_2 a^2\right)\right)<1$ we have*

1. $\widetilde{\mu}_{(1)}(k)=E\{X_t e_t\}=\sigma_2$,
2. $\mu_{(1)}(k)=E\{X_t\}=\underline{\mathbf{1}}\underline{\mu}_{(1)}(k)$ *where* $\underline{\mu}_{(1)}(k)=\mathbb{P}^{(k)}(a)\underline{\Pi}(\sigma_2)$,
3. $\widetilde{\mu}_{(2)}(k)=E\left\{X_t^2 e_t^2\right\}=\underline{\mathbf{1}}\underline{\widetilde{\mu}}_{(2)}(k)$ *where* $\underline{\widetilde{\mu}}_2(k)=\left(I_{(d)}-\mathbb{P}^{(k)}\left(\sigma_2 a^2\right)\right)^{-1}\underline{\Pi}(\sigma_4)$ *and* $\mu_{(2)}(k)=E\left\{X_t^2\right\}=\underline{\mathbf{1}}\underline{\mu}_{(2)}(k)$ *where* $\underline{\mu}_{(2)}(k)=\mathbb{P}^{(k)}\left(a^2\right)\underline{\widetilde{\mu}}_{(2)}(k)+\underline{\Pi}(\sigma_2)$.
4. $$\underline{\mu}_X(j)=\begin{cases}\mathbb{P}^{(k)}(a^2)\left(I_{(d)}-\mathbb{P}^{(k)}\left(\sigma_2 a^2\right)\right)^{-1}\underline{\Pi}(\sigma_4)+\underline{\Pi}(\sigma_2), & j=0,\\ \mathbb{P}^{(k)}(\ a)\mathbb{P}^{(k)}(\ 2\sigma_2)\underline{\Pi}(\sigma_2), & j=k,\\ \mathbb{P}^{(k)}(a)\mathbb{P}^{(n-k)}(\sigma_2)\underline{\mu}_{(1)}(k), & j=n>k.\end{cases}$$

Proof. The proof is straightforward and hence omitted. □

Note that in standard case $(d=1)$ the process $(X_t)_t$ is an $k-$dependent and hence $\gamma(j)=0$ for all $j\neq k$, in contrast here $\gamma(j)\neq 0$ for all j. This finding could be used as a sufficient condition for testing the Markovianity switching with dependent chain. Similar results stated in Lemma 5 for $MS-SBL$ may be given in $MS-DBL$ one.

Lemma 8. *For the $MS-DBL$ and under the condition $\rho\left(\mathbb{P}^{(k)}\left(\sigma_r a^r\right)\right)<1$, we have*

1. $\widetilde{\mu}_{(r)}(k)=E\{(X_t e_t)^r\}=\underline{\mathbf{1}}\underline{\widetilde{\mu}}_{(r)}(k)$ *where*
$$\underline{\widetilde{\mu}}_{(r)}(k)=\left(I_{(d)}-\mathbb{P}^{(k)}(\sigma_r\ a^r)\right)^{-1}\left(\sum_{i=1}^{\left[\frac{r}{2}\right]-1}C_r^{2i}\ \sigma_{2(r-i)}\mathbb{P}^{(k)}\left(\ \underline{a}^{2i}\right)\underline{\widetilde{\mu}}_{(2i)}(k)+\underline{\Pi}(\sigma_{2r})\right).$$
2. $$\mu_{(r)}(k)=E\{X_t^r\}=\underline{\mathbf{1}}\underline{\mu}_{(r)}(k)=\begin{cases}\underline{\mathbf{1}}\left(\underline{\Pi}(\sigma_{2n})\ +\sum_{i=1}^{n}C_{2n}^{2i}\sigma_{2(n-i)}\mathbb{P}^{(k)}\left(\ \underline{a}^{2i}\right)\underline{\widetilde{\mu}}_{(2i)}(k)\right), & r=2n,\\ \underline{\mathbf{1}}\sum_{i=1}^{n}C_{2n-1}^{2i-1}\ \mathbb{P}^{(k)}\left(\sigma_{2(n-i)}\ a^{2i-1}\right)\underline{\widetilde{\mu}}_{(2i-1)}(k), & r=2n-1.\end{cases}$$

Proof. For any integer $r\geq 1$, we have $(X_t e_t)^r=e_t^{2r}+\sum_{i=1}^{r}C_r^i a^i(s_t)(X_{t-k}e_{t-k})^i e_t^{2r-i}$, so

$$\begin{aligned}\pi(j)E\left\{(X_t e_t)^r|s_j=j\right\}&=E\left\{e_t^{2r}|s_t=j\right\}\pi(j)+\sum_{i=1}^{\left[\frac{r}{2}\right]}C_r^{2i}a^{2i}(j)E\left\{(X_{t-k}e_{t-k})^{2i}|s_t=j\right\}\pi(j)\sigma_{2(r-i)}\\ &=\sigma_{2r}\pi(j)+\sum_{i=1}^{\left[\frac{r}{2}\right]}C_r^{2i}a^{2i}(j)\sum_{u=1}^{d}E\left\{(X_{t-k}e_{t-k})^{2i}|s_{t-k}=u\right\}p_{uj}^{(k)}\pi(u)\sigma_{2(r-i)}\end{aligned}$$

and thus $\underline{\widetilde{\mu}}_{(r)}(k)=\left(I_{(d)}-\mathbb{P}^{(k)}(\ \sigma_r a^r)\right)^{-1}\left(\underline{\Pi}(\sigma_{2r})+\sum_{i=1}^{\left[\frac{r}{2}\right]-1}C_r^{2i}\ \sigma_{2(r-i)}\mathbb{P}^{(k)}\left(\ \underline{a}^{2i}\right)\underline{\widetilde{\mu}}_{(2i)}(k)\right)$. Therefore

$$E\{(X_t e_t)^r\}=\underline{\mathbf{1}}\underline{\widetilde{\mu}}_{(r)}(k).$$

For the second assertion, we have $X_t^r=e_t^r+\sum_{i=1}^{r}C_r^i a^i(s_t)(X_{t-k}e_{t-k})^i e_t^{r-i}$, so if $r=2n$, we get $\pi(l)E\left\{X_t^{2n}|s_t=l\right\}$ $=\sigma_{2n}\pi(l)+\sum\limits_{i=1}^{n}C_{2n}^{2i}\sigma_{2(n-i)}a^{2i}(l)E\left\{(X_{t-k}e_{t-k})^{2i}|s_t=l\right\}\pi(l)$. Hence

$$\underline{\mu}_{(2n)}(k)=\underline{\Pi}(\sigma_{2n})+\sum_{i=1}^{n}C_{2n}^{2i}\ \sigma_{2(r-i)}\mathbb{P}^{(k)}\left(\ a^{2i}\right)\underline{\widetilde{\mu}}_{(2i)}(k)$$

otherwise,

$$\underline{\mu}_{(2n-1)}(k)=\sum_{i=0}^{n-1}C_{2n-1}^{2i-1}\ \sigma_{2(r-i)}\mathbb{P}^{(k)}\left(\ a^{2i-1}\right)\widetilde{\underline{\mu}}_{(2i-1)}(k),$$

so $E\left\{X_t^r\right\}=\mathbf{1}\underline{\mu}_{(r)}(k)\,I_{\{r=2n\}}+\mathbf{1}\underline{\mu}_{(r)}(k)\,I_{\{r=2n-1\}}$. □

All the third-order moments of the $MS-DBL$ are summarized in the following lemma

Lemma 9 (Third-order moments). *If $(X_t)_t$ is generated by a $MS-DBL$, then under the condition of Lemma 8, the non centred third-order moments are*

$$\underline{\mu}_X(i,j)=\begin{cases}3\sigma_2\underline{\mu}_{(1)}(k)+\ \mathbb{P}\left(a^3\right)\widetilde{\underline{\mu}}_{(3)}(k), & i=j=0,\\ \mathbb{P}^{(k)}\left(a^2\right)\left\{3\sigma_4\underline{\mu}_{(1)}(k)+\mathbb{P}^{(k)}\left(\sigma_2a^3\right)\widetilde{\underline{\mu}}_{(3)}(k)\right\}+\mathbb{P}^{(k)}(\sigma_2)\,\underline{\mu}_{(1)}(k), & i=0,j=k,\\ \mathbb{P}\left(a^2\right)\left\{\mathbb{P}^{(k(n-1))}\left(2\sigma_2^2\right)\underline{\mu}_{(1)}(k)+\sigma_2\underline{\mu}_X(0,k(n-1))\right\} & \\ \qquad\qquad+\mathbb{P}^{(nk)}(\sigma_2)\,\underline{\mu}_{(1)}(k), & i=0,j=nk,n\geq 2\\ \mathbb{P}^{(k)}(a)\left\{\mathbb{P}^{(k)}\left(3\sigma_2a^2\right)\widetilde{\mu}_{(2)}(k)+\underline{\Pi}(\sigma_4)\right\}, & i=j=k,\\ 4\sigma_2^2\mathbb{P}^{2(k)}(a)\,\underline{\mu}_{(1)}(k), & i=k,j=2k,\\ 2\sigma_2\mathbb{P}^{2(k)}(a)\,\mathbb{P}^{(k(n-2))}(\sigma_2)\,\underline{\mu}_{(1)}(k), & i=k,j=nk,n\geq 3\\ \mathbb{P}^{(k)}(a)\,\mathbb{P}^{(k(l-1))}(\sigma_2)\,\underline{\mu}_X(k(n-l))), & i=lk,j=nk,l\geq 2,n\geq 3,\end{cases}$$

and hence $\gamma_Y(i,j)$ may be expressed from (11).

Proof. The first part corresponding to $i=j=0$ may be derived through Lemmas 7 and 8. For the case: $i=j=k>0$, then we have $\pi(u)E\left\{X_tX_{t-k}^2|s_t=u\right\}=\pi(u)a(u)E\left\{X_{t-k}^3e_{t-k}|s_t=u\right\}$, so $\underline{\mu}_X(k,k)=\mathbb{P}^{(k)}(a)\underline{U}(k)$ where the components of the vector $\underline{U}(k)$ are $\pi(u)E\left\{X_t^3e_t|s_t=u\right\}=\pi(u)\sigma_4+3\pi(k)\sigma_2a^2(u)E\left\{X_{t-k}^2e_{t-k}^2|s_t=u\right\}$ and we deduce that $\underline{U}(k)=\underline{\Pi}(\sigma_4)+\ \mathbb{P}^{(k)}\left(3\sigma_2a^2\right)\widetilde{\mu}_{(2)}(k)$ hence the results follows. Similar methodology can be used to obtain the expansion of the remainder cases. □

Lemma 10. *For the $MS-DBL$ consider the squared process $Y_t=X_t^2$ and assume that $\rho\left(\mathbb{P}^{(k)}(\sigma_4a4)\right)<1$, then we have*

$$\underline{\mu}_Y(j)=\begin{cases}\underline{\mu}_{(4)}(k), & j=0,\\ \mathbb{P}^{(k)}(a^2)\left\{\underline{\Pi}(\sigma_6)+6\sigma_4\mathbb{P}^{(k)}\left(a^2\right)\widetilde{\underline{\mu}}_{(2)}(k)+\sigma_2\mathbb{P}^{(k)}\left(a^4\right)\widetilde{\underline{\mu}}_{(4)}(k)\right\}+\mathbb{P}^{(k)}(\sigma_2)\,\underline{\mu}_{(2)}(k), & j=k,\\ \mathbb{P}^{(k)}(\sigma_2a^2)\underline{\mu}_Y(k(n-1))+\left(\mathbb{P}^{(k)}(a^2)\mathbb{P}^{(k(n-1))}\left(2\sigma_2^2\right)+\mathbb{P}^{(kn)}(\sigma_2)\right)\underline{\mu}_{(2)}(k), & j=nk,n\geq 2.\end{cases}$$

and hence $\gamma_Y(j)=\mathbf{1}\underline{\mu}_Y(j)-\mu_{(2)}^2$ for any $j\geq 0$.

Proof. The first part, for $j=0$ may be deduced from the Lemma 8. The second part, for $j=k$, let us define the vector $\underline{\mu}_Y(k)=\left(\pi(j)E\left\{X_t^2X_{t-k}^2|s_t=j\right\},1\leq j\leq d\right)'$, then its is not difficult to see that $\underline{\mu}_Y(k)=\mathbb{P}^{(k)}\left(a^2\right)\underline{W}(k)+\mathbb{P}^{(k)}(\sigma_2)\underline{\mu}_{(2)}(k)$ where $\underline{W}(k)=\left(\pi(j)E\left\{X_t^4e_t^2|s_t=j\right\},1\leq j\leq d\right)'$. Tedious calculation shows that $\underline{W}(k)=\underline{\Pi}(\sigma_6)+\mathbb{P}^{(k)}(6\sigma_4a^2)\widetilde{\underline{\mu}}_{(2)}(k)+\mathbb{P}^{(k)}(\sigma_2a^4)\widetilde{\underline{\mu}}_{(4)}(k)$ and hence the results follow. The third part, follow upon the observation that $\underline{\mu}_Y(kn)=\mathbb{P}^{(k)}(a^2)\underline{U}+\mathbb{P}^{(nk)}(\sigma_2\underline{\mu}_{(2)}(k)$, where

$$\underline{U}=(\pi(j)E\left\{X_t^2X_{t-k(n-1)}^2|s_t=j\right\},1\leq j\leq d)',$$

after some computations, we get $\underline{U}=\sigma_2\underline{\mu}_Y(k(n-1))+\mathbb{P}^{(k(n-1))}(2\sigma_2)\,\underline{\mu}_{(2)}(k)$ and hence the result follows.

□

The following corollary correspond to independent switching $MS-DBL$, we stated the moments properties without proofs. Details of the proof are of course available if it required.

Corollary 2. [*Switching-independent case*] *For the $I-DBL$, we set $\lambda = \sigma_2\overline{a^2}$, then we have:*

1. $\mu = E\{X_t\} = \sigma_2\overline{a}$ *and* $Var(X) = \sigma_2 + \dfrac{\overline{a^2}\sigma_4}{1-\lambda} - \mu^2.$

2. $\gamma_X(i) = \lambda^2\ \delta_{\{i=k\}}.$

3. $$\gamma_X(i,j) = \begin{cases} \frac{3\overline{a^3}\sigma_4}{(1-\lambda)}\left(\overline{a^2}\sigma_4 - \sigma_2\right) + \overline{a^3}\left(\ \sigma_6 - 2\sigma_2^3\right), & i=j=0, \\ \frac{\overline{a^3}\sigma_2}{\left(1-\sigma_2\overline{a^2}\right)} - \beta - 2\overline{a^3}\sigma_2^3, & i=0, j=k \\ \frac{\overline{a^{2n+1}}\sigma_2^{2n}}{\left(1-\sigma_2\overline{a^2}\right)}\beta, & i=0, j=nk, n=2,3,... \\ \frac{\overline{a}}{\left(1-\sigma_2\overline{a^2}\right)}\left(\ \sigma_4 - \sigma_2^2 + \overline{a^2}\ \left(\sigma_4 - \sigma_2^2 + 2\overline{a^2}\sigma_2^3\right)\right), & i=j=k,,. \\ \overline{a^3}\sigma_2^3, & i=k, j=2k, \\ 0, & otherwise \end{cases}$$

   where $\beta = \overline{a^4}\left(3\sigma_4^2 - \sigma_2\sigma_6\right) + \overline{a^2}\left(\sigma_6 + \sigma_2\sigma_4\right) + 2\sigma_4.$

4. $$\gamma_Y(j) = \begin{cases} E\left\{X_t^4\right\} -\ \mu_{(2)}^2, & j=0 \\ \overline{a^2}\left\{\sigma_6 + 6\sigma_4\overline{a^2}\widetilde{\mu_{(2)}} + \sigma_2\overline{a^4}\widetilde{\mu}_{(4)}\right\} + \sigma_2\mu_{(2)} - \mu_{(2)}^2, & j=k, \\ \sigma_2\overline{a^2}\gamma_Y(k(n-1)) + \left(2\sigma_2\overline{a^2} + \sigma_2^2\right)\mu_{(2)} + \sigma_2\overline{a^2}\mu_{(2)}^2\ \text{-}\mu_{(2)}^2\ \ if\ j=nk, & n\geq 2. \end{cases}$$

Remark 6. *The results of second assertion of above corollary shows clearly that $I-DBL$ is a $k-$dependent process and hence it has the same covariance structure of an $MS-MA(k)$ model, in other hand the result of the fourth assertion clearly indicate that $I\ -DBL$ is not an $MS-MA(k)$.*

Remark 7. *The results obtained for $MS-BL$ can be useful in modeling nonlinear time series, particularly in the choice of the lags k and l of some simple bilinear models for which the innovation is Gaussian. Note here that the results obtained in independent switching case are similar to standard ones $(d=1)$ when the parameters in moment properties are replaced by the moments of their coefficients. Despite the difficulty and the complexity of computation, especially in dependent switching, the results obtained here constitute an alternative for the study of nonlinear series.*

Example 3. Consider the same distribution as in Example 2, then the kurtosis and the skewness of $MS-DBL$ are shown in the Figure 2 terms of $a(1)$ and $a(2)$) of $MS-SBL$ are shown in the Figure 1. The kurtosis of $MS-DBL$ in top panels of the Figure 2 are strictly positive and hence their distributions are leptokurtic. The skewness of $MS-DBL$ in bottom panels are negative for some values of $a(1)$ and $a(2)$ and hence the distribution has a long tail to the left, i.e., mode>median>mean.

3.3 Simulation study

In order to examine the problems associated with the identification of a $MS-BL$ model with the help of the third or fourth order moment, we consider an $MS-SBL$ (resp $MS-DBL$) with $(e_t)_t$ is a Gaussian sequence $\mathcal{N}(0,1)$ and $(s_t)_t$ is a 2$-$state Markov chain with $p_{11}=0.75$, $p_{22}=0.95$ and $a(s_t)\in\{-0.75, 0.15\}$. Now we simulate a series of length $n=1000$, according to the models already mentioned, and calculate the second, third and fourth order moments $\widehat{\gamma}_X(i)$, $\widehat{\gamma}_X(i,j)$, and $\widehat{\gamma}_{X^2}(i)$, $i,j=1,...,\max lag$. We repeated the experiment 500 times and the mean values of $\widehat{\gamma}_X(i)$, $\widehat{\gamma}_X(i,j)$ and $\widehat{\gamma}_{X^2}(i)$ are obtained. The simulation experiment has

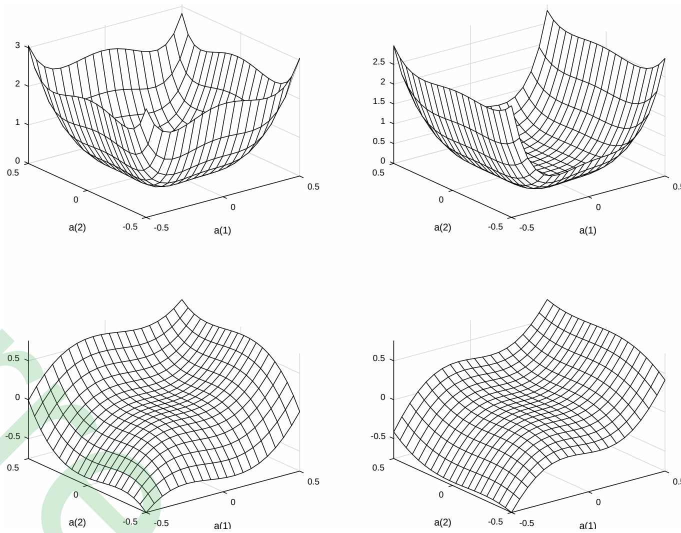

Figure 2: Kurtosis and skewness of the two state $MS-DBL$. The left of top panel displayed the kurtosis of $D-DBL$ followed by $I-DBL$. Bottom panel displayed the skewness of $D-DBL$ in left followed by $I-DBL$.

been made in order to compare the difference between the theoretical values of $\gamma_X(i)$, $\gamma_X(i,j)$ and $\gamma_{X^2}(i)$ and those obtained via a Monte Carlo experiment. To avoid biased simulation's intervention, we delete first 100 observations to make sure the randomness of the model. The different plots of models according to the variability of the chain $(s_t)_t$ are presented in Figure 3. In Figure 3, we can see that the graphic associated to $MS-SBL$, presents more volatility than the $MS-DBL$. Moreover, it can also be observed that there is a closed similarity between the $C-$. and $I-$. models.

3.4 Second-order properties

The sample $\widehat{\gamma}_X(s)$ up to lag 25, according to the variability of the chain $(s_t)_t$ are given in Figure 4. As expected theoretically, the covariances in all cases decay to zero as the lag increases, though the rate of decay in the case of the $MS-DBL$ is slower. The results of the simulation of the second-order moment are given in Table 1.

Observations of the series simulated from $MS-SBL$ (resp. $MS-DBL$) are plotted in right (resp. left) side panel of Figure 3 for different specification of the chain $(s_t)_t$. The covariance $\gamma_X(h)$ of these series are shown in Table 1, for $h \in \{0,1,2,3,4,5\}$ and their plots are in Figure 4. So, we can see that the covariance function are zero beyond 0 (resp. 1) for $MS-SBL$(resp. $MS-DBL$). This finding confirm the theoretical results summarized in Figure 4.

3.5 Third-order properties

Now, the third$-$order moments are concerned. The sample third-order moments $\widehat{\gamma}_X(0,s)$ up to lag 25, according to the variability of the chain $(s_t)_t$ are given in Figure 5. It can be observed that for the $MS-SBL$, $\widehat{\gamma}_X(0,s)$ are approximately zero, contrary to the $MS-DBL$, $\widehat{\gamma}_X(0,s)$ are not zero even beyond 1. This means that the white noise structure of $MS-SBL$ is preserved by the third-order moment properties and the structure $MA(1)$ of $MS-DBL$ is violated. The results of simulation of the third-order moments are shown in Table 2 for $MS-SBL$ when $(s_t)_t$ is an independent (resp. dependent) chain, and in Table 2 for $MS-DBL$ with the same structure chain followed by their true values (results between parentheses). Table 2 gives theoretical values (in the parenthesis) of $\gamma(i,j)$ for different values of i,j and the corresponding

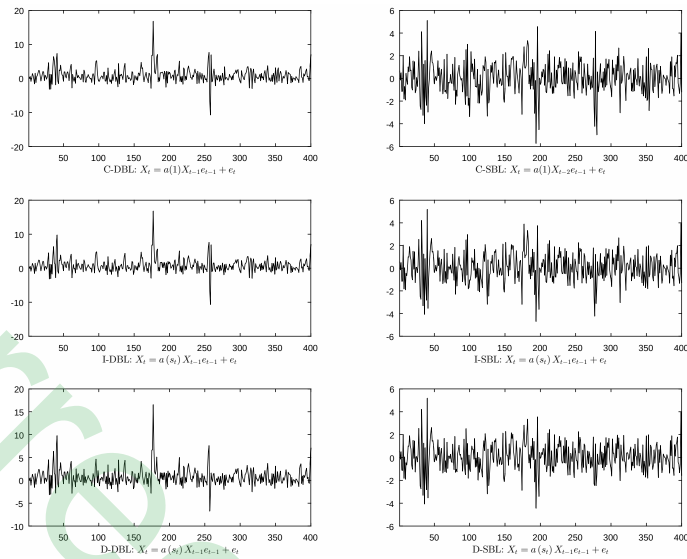

Figure 3: Displayed are in top panel $C-DBL$ followed by $C-SBL$ in second panel the $I-DBL$ followed by $I-SBL$ and in the third $D-DBL$ followed by $D-SBL$.

simulated values $\widehat{\gamma}(i,j)$ for $MS-DBL$.

Observations of the first series simulated from $MS-SBL$ is plotted in first panel of Figure 3. Its second-order $\gamma_X(m)$ and third-order $\gamma_X(i,j)$ moments are shown respectively in first column of Table 1 and in Table 2 with $(i,j) \in \{0,1,2,3,4\} \times \{0,1,2,3,4\}$. It can be seen that simulation results agree well with the results reported in in Figure 5 and with the theoretical results. Moreover, the order of $\gamma_X(i,j)$ and $\gamma_X(m)$ are quite different, we believe that the simple bilinear structure are not obvious and hence, an *AIC* criterion should be discussed further.

Observations of this series (when $(s_t)_t$ is independent) simulated from $MS-SBL$ is plotted in second panel of Figure 1. Its second-order $\widehat{\gamma}_X(m)$ and third-order $\widehat{\gamma}_X(i,j)$ moments are shown respectively in second block of Table 1 and in Table 2. In this case also the simulation results are quite closely with the theoretical results. Moreover, the values corresponding to the cell $(0,0)$ are found to be absolutely dominant over other values. Since $\gamma_X(2)$ is zero, we believe that $\gamma_X(2,2)$ should be wrong message.

3.6 Covariance analysis of squared $MS-SBL$ and $MS-DBL$

The purpose of this subsection is to carry out a second-order analysis on the squared $MS-SBL$ and $MS-DBL$ noted $Y_t = X_t^2$ with the help of the Lemmas 6 and 10 to providing an alternate differentiation technique from their linear representations

Observations of this series simulated from the squared version of $MS-SBL$ and $MS-DBL$ their covariance $\widehat{\gamma}_Y(m)$ is plotted in Figure 6. and there values are shown in Table 4.

As a consequence of the results reported in Table 4 is that the second-order structure of white-noise for $MS-SBL$ model (resp. $MA(1)$ for $MS-DBL$) are not preserved by the process $(Y_t)_t$. So, the squared process maybe constituted a powerful alternative for the differentiation between linear and/or nonlinear models. In the other words, in view of Figure 6, it can be see that the monotony of $\widehat{\gamma}_X(h)$ is also preserved by $\widehat{\gamma}_{X^2}(h)$ for $h = 0,1,....$ Note here that in spite the large values of $\gamma_{X^2}(h)$ for $C-DBL$ and $I-DBL$, the values of $\gamma_{X^2}(h)$ corresponding to $D-DBL$ are significantly relaxed.

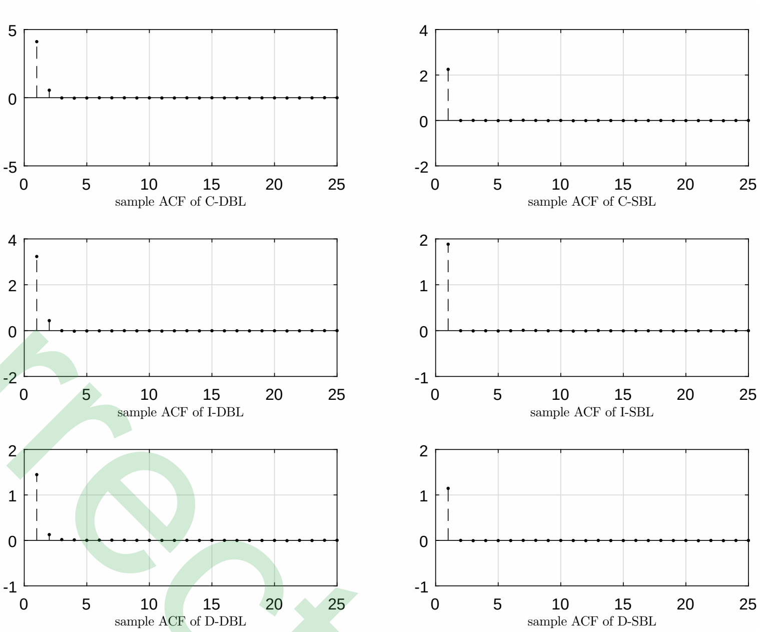

Figure 4: From top to bottom: The covariance functions associated respectively to $(.)-DBL$ followed in right side by $(.)-SBL$

4 Application to investments

In this section, we use our theoretical results to analyze certain series of asset returns via $MS-BL$ models (9) aimed to predict the future values of such series. The resort to regime switching in asset returns and their pronounced implications for investments have been widely documented through several models according some accommodates features. (see Marcucci (2005) for more discussions). Based on the findings in this literature, we analyze the exchange rates of the Algerian Dinar against the single European currency (Euro) (EUR/DZD). We investigate some descriptive statistics and the impact of such series for the different regimes. Our base model is an $MS-BL$ models (9) with two regimes are considered. We compare its implications to those of restricted models to determine the impact of regime switching models.

4.1 Sample data and Preliminary Analysis

The proposed models $MS-SBL$ and $MS-DBL$ are investigated to model the series of EUR/DZD already studied by Bibi (2021), observed from January 3, 2000 to September 29, 2011. Since there is some weeks comprise less than five observations (due to legal holidays), we remove the entire week with less than five available. So, the final length of the series is $T = 3055$ observations uniformly disturbed on 611 weeks. The elementary statistics are tabulated in Table ($Stat$) below which shows some description of the series EUR/DZD. The kurtosis is less than the normal value of 3 indicating that the distribution has lighter tails and a flatter peak than a normal distribution (platykurtic). The skewness is significantly negative, showing that the tail on the left side of the histogram is longer than the right, with most of the data clustered on the right side of the mean.

Note: The $Jbstat$-test which has a χ^2 distribution with 2 degrees of freedom under the null hypothesis of normally distributed errors. The 5% critical value is therefore 6.4795. The $LM(12)$ statistic is the $ARCHLM$ test up to the twelfth lag and under the null hypothesis of no $ARCH$ effects it has a $\chi^2(q)$

Table 1: The Monte Carlo simulations and theoretical values (in parenthesis) of $\gamma_X(s)$ in different lags of MS-SBL and MS-DBL model.

Lags\Models	$MS-SBL$			$MS-DBL$		
	$C-SBL$	$I-SBL$	$D-SBL$	$C-DBL$	$I-DBL$	$D-DBL$
0	2.2878 (2.2857)	1.8946 (1.8957)	1.4419 (1.4269)	4.0592 (4.0946)	3.2296 (3.2639)	1.4608 (1.4594)
1	−0.0042 (0.0000)	−0.0022 (0.0000)	−0.0006 (0.0000)	0.5219 (0.5191)	0.4200 (0.4192)	0.1216 (0.0984)
2	0.0017 (0.0000)	−0.0068 (0.0000)	−0.0059 (0.0000)	−0.0400 (0.0000)	−0.0249 (0.0000)	0.0053 (0.0045)
3	−0.0031 (0.0000)	−0.0038 (0.0000)	−0.0038 (0.0000)	−0.0583 (0.0000)	−0.0498 (0.0000)	−0.0010 (0.0006)
4	−0.0095 (0.0000)	−0.0082 (0.0000)	−0.0042 (0.0000)	−0.0453 (0.0000)	−0.0378. (0.0000)	−0.0091 (0.0080)
5	−0.0026 (0.0000)	−0.0031 (0.0000)	−0.0036 (0.0000)	−0.0305 (0.0000)	−0.0315 (0.0000)	−0.0057 (0.0045)

distribution, where q is the number of lags. The $Q^2(12)$ statistic is the Ljung-Box test on the squared residuals of the conditional mean regression up to the twelfth order. Under the null hypothesis of no serial correlation, the test is also distributed as a $\chi^2(q)$ where q is the number of lags. Thus, for both tests the 5% critical value are 6.5597 (resp.31.4104). At a confidence level of 5% both skewness and kurtosis are significant, since the standard errors under the null of normality are $6/T = 0.0019$ and $24/T = 0.0077$ respectively.
The plot of the trajectory of EUR/DZD series followed by its second and third-order cumulants are given in the following Figure 7

It is clear that the series displayed in Figure 7 exhibit a structural break changes. More precisely, from the results of Table 6 we can see that the descriptive statistics show that the negativity skewed, indicate that the most data points are clustered on the right side (higher values), with a few extreme low values stretching out to the left. So, we have mean < median < mode. Moreover, the positivity of the kurtosis shows that the data are Leptokurtic and hence the null hypothesis of normality is rejected. Additionally, the Jarque-Bera test statistics ($Jbstat$) is larger than the critical value (6.5597) which is significant of 1% level which confirm the rejection of null hypothesis. The plot of third-order cumulant displayed in Figure 7 performs the rejection.

4.2 Fit $MS-BL$ to daily EUR/DZD

Based on the above description of EUR/DZD series, the $MS-ARMA$ model did not reflect the behavior of such data, and hence $MS-BL$ models may be called for modelling and forecasting this series. Indeed, for modelling purpose, we are firstly using a quasi-maximum likelihood (QML) procedure for fitting the series by $MS-BL$. Table 6 some goodness-of-fit statistics are reported according to C_BL. These statistics are used as model selection criteria. The Akaike information criterion (AIC) and the Schwarz criterion (BIC) (results not reported here) both indicate that the best model is the $MS-DBL$. Another property of $MS-DBL$ models that emerges from Table 6 is the high persistence showed by small parameters estimates. The results of the parameter estimation followed by the mean squares errors (results between parentheses), according to $C-SBL$ and $C-DBL$ models are reported in the column (Parameters) of Table 6. The results gathered in (statistics column) of Table 6 are the elementary statistics associated with the adjusted series

Table 2: The third-order moments of I-SBL and D-SBL models.

$\frac{Models}{Lags}$	$I-SBL$ model					$10^{-3}\times D-SBL$ models				
	0	1	2	3	4	0	1	2	3	4
0	−0.0007 (−0.0006)	0.0002 (0.0000)	−0.0009 (−0.0007)	0.0003 (0.0005)	0.0002 (0.0000)	−0.0363 (−0.0291)	0.1553 (0.1522)	−0.0265 (−0.0255)	0.0154 (0.0164)	−0.1085 (−0.1001)
1		0.0003 (0.0000)	0.0019 (0.0020)	0.0001 (0.0000)	−0.0001 (0.0000)		0.0820 (0.0789)	0.5082 (0.4811)	−0.1040 (−0.1003)	−0.0599 (−0.0411)
2			−0.0007 (−0.0006)	0.0001 (0.0000)	0.0003 (0.0001)			0.0641 (0.0788)	0.0369 (0.0335)	−0.0614 (0.0059)
3				−0.0002 (−0.0005)	−0.0002 (0.0000)				−0.0310 (0.0325)	0.0295 (0.0198)
4					−0.0003 (0.0000)					−0.0252 (−0.0250)

Table 3: The third-order moments of I-DBL and D-DBL models

$\frac{Models}{Lags}$	$I-DBL$ model					$D-DBL$ model				
	0	1	2	3	4	0	1	2	3	4
0	−8.1885 (−8.09871)	3.7328 (3.7294)	−0.4178 (−0.3987)	−0.3624 (−0.3592)	−0.2654 (−25904)	1.9461 (1.9392)	1.0668 (1.0590)	−0.0452 (−0.0450)	−0.0528 (−0.0492)	0.0540 (0.0601)
1		5.9508 (5.8972)	0.2878 (0.2798)	−0.1515 (−0.1405)	−0.2007 (−0.2001)		1.2645 (1.2551)	0.0826 (0.0802)	0.0283 (0.0279)	−0.1077 (−0.1009)
2			3.0620 (3.1001)	−0.0706 (−0.0710)	−0.0334 (0.0335)			0.6148 (0.5941)	0.0414 (0.0092)	−0.0323 (0.0325)
3				1.2265 (1.2155)	−0.0571 (0.0493)				0.2780 (0.2689)	−0.0415 (−0.0491)
4					0.5082 (0.4905)					0.1569 (0.1601)

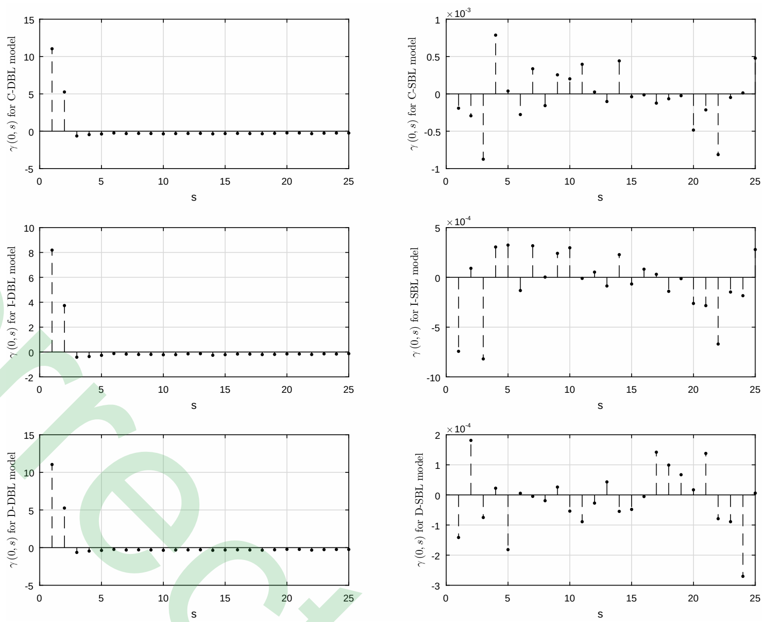

Figure 5: Displayed are the sample third-order moments $\widehat{\gamma}_X(0,s)$ of the $MS-DBL$ in left panel and of $MS-SBL$ in right.

Table 4: The simulate values of γ_{X^2} for different values of lags for the MS-SBL and MS-DBL.

$\frac{Models}{Lags}$	Squared $MS-SBL$			Squared $MS-DBL$		
	$C-DBL$	$I-DBL$	$D-DBL$	$C-DBL$	$I-DBL$	$D-DBL$
0	46.2211	19.4792	4.3941	405.3164	206.3111	40.9966
1	5.9235	3.4578	0.6353	222.7964	95.6614	18.5253
2	23.4193	8.2800	1.0314	113.4164	38.2723	7.6013
3	3.6595	1.7531	0.2424	56.5737	15.7618	4.0707
4	11.4029	3.4330	0.3205	32.3876	7.0561.	3.8038
5	1.8171	0.7900	0.0889	15.4928	3.3257	3.2314

This table shows the parameter estimates of adjusted EUR/DZD by $MS-BL$ under $d=1$ and their elementary statistics. The main observation is that the values of elementary statistics of the series adjusted by $C-DBL$ model are better than $C-SBL$ comparing with the original series. This is not surprising observation because the model $C-SBL$ is a weak white noise whereas the EUR/DZD series is very far to be considered as a white noise. The second approach for modelling the EUR/DZD is by $MS-\ BL$ under $d=2$ with independent and dependent chain, are summarized in Table 7.

In above Table 7 the column "Matrix P" means the estimate of transition matrix of the chain. Now, a few comments can be made. Indeed, it is clear that the Table 7 demonstrates that the $MS-DBL$ model applied to the EUR/DZD series performs as effectively as the applications to EUR/DZD concerning classification certainty for the mean. At least 96% of the fitted observations are considered with low uncertainty. In terms of variance, the model classify at most 95% of the observations as decisive. The $MS-DBL$ with a greater number of states, applied to EUR/DZD is the least effective in classifying variance, with a percentage of "decisive" between 99% and 100%. Regarding the skewness and the kurtosis are the same as for original data with non-significant difference The $Jbstat$-test is also significantly

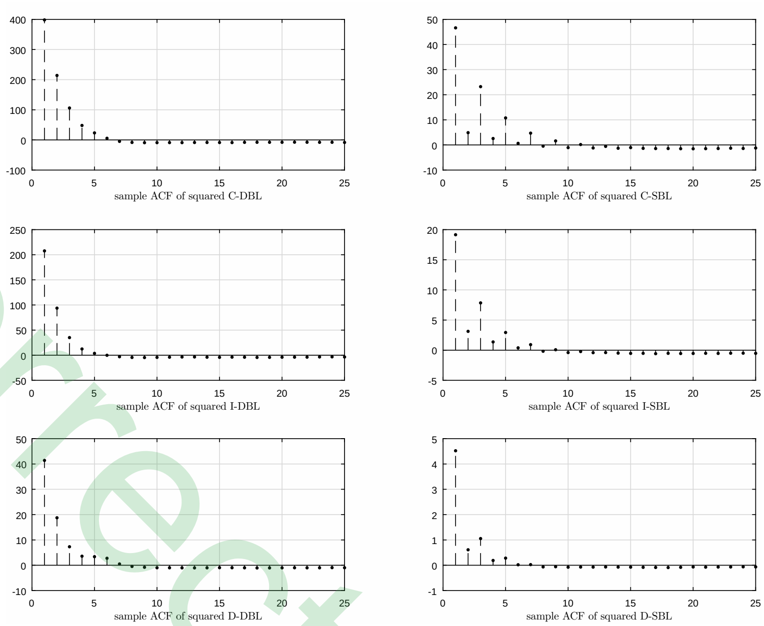

Figure 6: Displayed are the $\widehat{\gamma}(s)$ of the squared $MS-DBL$ in left panel and of squared $MS-SBL$ in right one.

Table 5: Elementary statistics of daily EUR/DZD series.

Statistics:	$Mean$	Std	$Skew$	Kur	$Jbstat$	$LM(12)$	$Q^2(12)$
10^4*Results	0.0088	0.0011	-0.0001	0.0002	0.0232	0.2688	5.8687

Table 6: The elementary statistics of adjusted $MS-BL$ under $d=1$.

		$10^4*Statistics$						
Models	Parameters	$Mean$	Std	$skew$	kur	$Jbstat$	$LM(12)$	$Q^2(12)$
$C-SBL$	0.099 (0.045)	0.0088	0.0011	-0.0008	0.0004	0.0548	0.2720	0.8480
$C-DBL$	0.096 (0.031)	0.0089	0.0012	-0.0001	0.0003	0.0211	0.2697	1.6060

to reject the null. In end, the $ARCH$ effect tests indicate clearly that no serial correlation up to the twelfth lag. Nonetheless, the model is indicated as good alternatives in terms of goodness− of− fit. According to the study, the $MS-DBL$ with $d>2$, is a reasonable choice for the application. This conclusion is obtained when accounting for both ”classification power” and goodness-of -fit.

Table 7: The elementary statistics of adjusted $MS-BL$ under $d=2$.

Models	Parameters	$10^4 * Statistics$							Matrix P
		$Mean$	Std	$skew$	kur	$Jbstat$	$LM(12)$	$Q^2(12)$	
$I-SBL$	0.0103 0.0097 (0.021 0.014)	0.0088	0.0011	−0.0008	0.0004	0.0550	0.2720	0.8480	$\begin{pmatrix} 0.5420 & 0.4580 \\ 0.5416 & 0.4584 \end{pmatrix}$
$D-SBL$	0.0080 0.0117 (0.037 0.024)	0.0088	0.0011	−0.0009	0.0004	0.0521	0.2718	0.8573	$\begin{pmatrix} 0.3132 & 0.6868 \\ 0.3021 & 0.6979 \end{pmatrix}$
$I-DBL$	0.0115 0.0086 (0.011 0.015)	0.0089	0.0012	−0.0001	0.0003	0.0211	0.2697	1.6060	$\begin{pmatrix} 0.9979 & 0.0021 \\ 0.9979 & 0.0021 \end{pmatrix}$
$D-DBL$	0.0090 0.0121 (0.087 0.064)	0.0088	0.0019	−0.0008	0.0004	0.0521	0.2718	0.8573	$\begin{pmatrix} 0.3132 & 0.6868 \\ 0.6320 & 0.3680 \end{pmatrix}$

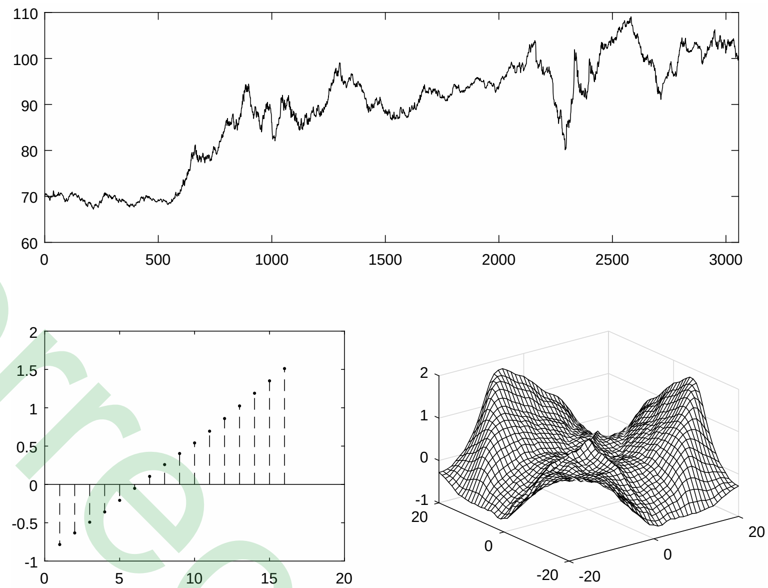

Figure 7: Top panel: the EUR/DZD series. Bottom panels: the second (left) and third (right) order cumulant of EUR/DZD series.

5 Conclusion

In this paper, we propose a unified framework that enables the calculation of higher-order moments and the diagnostic analysis of an $MS\text{-}BL(p,0,P,Q)$ model under both independent and dependent switching structures. As a crucial first step, we represent the model in a Markovian state-space form, which allows us to derive the conditions for the existence of strict stationary and ergodic solutions in $\mathbb{L}_p$, $p \geq 1$, under which explicit expressions for the first-, second-, and third-order moments are obtained. The second step of our framework focuses on analyzing specific cases of $MS\text{-}BL(p,0,P,Q)$ models. More precisely, the first-order *MS-SBL* and *MS-DBL* models are examined in detail.

For the *MS-SBL* model, the first- and second-order moments are derived under both (in-)dependent switching schemes, which exhibit structural similarities to those of an $MS\text{-}MA(1)$ model. Consequently, studying the second-order structure of the squared process allows us to distinguish between *MS-SBL* and $MS\text{-}MA(1)$ models. The same methodology is applied to the *MS-DBL* model. Our main theoretical contribution lies in demonstrating how this generalized formulation can be used to derive the first and second moments conditional on dependent regime switching. In contrast, for the *MS-DBL* model with independent regime switching, the first- and second-order moments coincide with those of the standard model and can be identified under covariance with the $MS\text{-}MA(1)$ model. However, this equivalence does not extend to the third-order moment structure, which breaks this covariance-based similarity. Nonetheless, examining the covariance structure of squared processes can help resolve this ambiguity.

In this study, we have made the following assumptions:
(i) $d = 2$, since the implications of the transition matrix prevent us from considering higher-order models, even in cases such as $d = 3$ and $m = 4$ (in MS-SBL).
(ii) A lower-triangular model is adopted to obtain a Markovian representation that simplifies the analytical derivations.
(iii) The restriction to a model without a moving average component is justified by the fact that the stationarity and ergodicity conditions are independent of this part.
(iv) Gaussian innovations are assumed to simplify otherwise cumbersome computations.

(v) The estimation of cumulants is known to be biased.

These specifications introduce some limitations to our analysis. Nonetheless, the empirical results demonstrate that ***MS-DBL*** models significantly outperform standard ***BL*** models according to a broad range of elementary statistical criteria. This strong conclusion holds regardless of whether the differences in performance are statistically significant. We conclude that our theoretical findings are useful for model identification, and the tools we propose can be directly applied to investigate processes that exhibit structural changes.

Acknowledgement

The author would like to extend his gratitude to the editor of the journal and to the reviewers for their insightful comments and valuable suggestions.

References

Alemohammad, N., Rezakhah, S., and Alizadeh, S. H. (2020). Markov switching asymmetric GARCH model: stability and forecasting. *Statistical Papers*, **61**, 1309–1333. (Cited on page 2.)

Bauwens, L., Preminger, A. and Rombouts, J., V., K. (2010). Theory and inference for a Markov switching GARCH model. *Econometrics Journal*, **13**, 218–244. (Cited on page 2.)

Bibi, A. (2021). Asymptotic properties of QMLE for seasonal threshold GARCH model with periodic coefficients. *Statistical Methods and Applications*, **30**, 477–514. (Cited on page 16.)

Bibi, A., and Ghezal, A. (2015). On the Markov-switching bilinear processes: stationarity, higher-order moments and β-mixing. *Stochastics An International Journal of Probability and Stochastic Processes*, **87**, 919–945. (Cited on pages 2 and 4.)

Bibi, A., and Hamdi, F. (2025). Covariance analysis and GMM estimation of Markov switching bilinear processes. *Statistics*, **59**, 354–387. (Cited on page 2.)

Boubacar M., Y., and Rabehasaina, L. (2020). Estimation of weak ARMA models with regime changes. *Statistical Inference for Stochastic Processes*, **23**, 1–52. (Cited on page 2.)

Brockwell, P. J. and Davis, R. A. (1987). *Time Series: Theory and Methods*. Springer-Verlag, New York. (Cited on page 9.)

Francq, C., and Zakoïan, J. M. (2001). Stationarity of multivariate Markov-switching ARMA models. *Journal of Econometrics*, **102**, 339–364. (Cited on page 2.)

Francq, C., and Zakoïan, J. M. (2005). L_2-structures of standard and switching-regime GARCH models. *Stochastic Processes and Their Applications*, **115**, 1557–1582. (Cited on pages 2 and 3.)

Gabr, M. M. (1988). On the third-order moment structure and bispectral analysis of some bilinear time series. *Journal of Time Series Analysis*, **9**, 11–20. (Cited on pages 6, 7, and 10.)

Granger, C. W. J. and Anderson, A. (1978). *An Introduction to Bilinear Time Series Models*. Vandenhoeck and Ruprecht, Gottingen. (Cited on page 6.)

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, **57**, 357–384. (Cited on page 1.)

Lee, O .(2005). Probabilistic properties of a nonlinear ARMA process with Markov switching. *Communications in Statistics: Theory and Methods*, **34**, 193–204. (Cited on page 2.)

Liu, J. (1992). On stationarity and asymptotic inference of bilinear time series models. *Statistica Sinica*, **2**, 479–494. (Cited on page 4.)

Marcucci, J. (2005). Forecasting stock market volatility with regime-switching GARCH models. *Studies in Nonlinear Dynamics and Econometrics*, **9**, 1–53. (Cited on page 16.)

Martin, C. M. (1999). A note on the third-order moments structure of a bilinear model. *Portugaliae Mathematica*, **16**, 115–125. (Cited on pages 6 and 10.)

Stelzer, R. (2009). On Markov-switching ARMA processes-stationarity, existence of moments, and geometric ergodicity. *Econometric Theory*, **25**, 43–62. (Cited on page 2.)

Wee, D. C., Chen, F., and Dunsmuir, W. T. (2022). Likelihood inference for Markov switching GARCH (1, 1) models using sequential Monte Carlo. *Econometrics and Statistics*, **21**, 50–68. (Cited on page 2.)

Yang, M. (2000). Some properties of vector autoregressive processes with Markov-switching coefficients. *Econometric Theory*, **16**, 23–43. (Cited on page 2.)

Yao, J. F., and Attali, J. G. (2000). On stability of nonlinear AR processes with Markov switching. *Advances in Applied Probability*, **32**, 394–407. (Cited on page 2.)