Statistical Methods & Applications (2026) 35:319–348
https://doi.org/10.1007/s10260-025-00826-9

ORIGINAL PAPER

(Bi)spectral analysis of Markov switching bilinear time series

Maddalena Cavicchioli[1] · **Ahmed Ghezal**[2] · **Imane Zemmouri**[3]

Received: 25 September 2024 / Accepted: 25 November 2025 / Published online: 10 December 2025

Abstract

We derive matrix expressions in closed form for the higher-order moments, the autocovariance function and the spectral and bispectral densities of Markov switching bilinear models and their powers. Under suitable assumptions, we prove that the sample estimators of the spectral and bispectral density matrices are consistent and asymptotically normally distributed. A simulation study confirms the validity of the asymptotic properties. These methods are also well suited for the analysis of time series in the frequency domain, as shown in some proposed real-world examples.

Keywords Markov switching model · Bilinear process · Higher-order moments · Autocovariance structure · (Bi)spectral density matrix · Asymptotic theory

JEL Classification C01 · C18 · C32

The paper has been presented by the first author at SIS 2024 Scientific Meeting, Bari, Italy.

✉ Maddalena Cavicchioli
maddalena.cavicchioli@unimore.it

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

Imane Zemmouri
zemmouri.imane@univ-annaba.org

1 Department of Economics Marco Biagi, University of Modena and Reggio Emilia, Viale Berengario 51, 41121 Modena, Italy

2 Department of Mathematics, Abdelhafid Boussouf University of Mila, Mila, Algeria

3 Department of Mathematics, University of Annaba, Annaba, Algeria

1 Introduction

Since the seminal works by Hamilton (1989, 1990) Markov switching (MS) time series models have received a growing interest in macroeconomics and continue to gain more popularity especially to adequately describe economic and financial data. For such models, the process is assumed to switch between different regimes, which are governed by an unobservable latent state variable following a Markov chain. This means that the probability of the current state depends only on the previous state. MS models have higher flexibility in capturing the persistence and/or the asymmetric effects on the shocks of volatility, and are useful in modeling time series which exhibit structural changes or distinct phases, such as periods of economic growth and recession in financial markets. Among them, the most popular being the Markov switching vector auto-regressive moving-average (MS VARMA) models and the Markov switching generalized autoregressive conditionally heteroskedastic (MS GARCH) models. For information concerning stationarity, estimation, consistency and asymptotic theory for MS VARMA and MS GARCH models see, e.g., Francq and Zakoïan (2001); Krolzig (1997); Hamilton (1994, 2016); Stelzer (2009); Cavicchioli (2014a, b, 2017a, b), and Alvarez et al. (2019) for the former, and Francq and Zakoïan (2005, 2008); Bauwens et al. (2010), and Cavicchioli (2021, 2022) for the latter.

An interesting class of Markov switching time series models, called Markov switching bilinear (MS BL) models, has been studied by Bibi and Ghezal (2015, 2016, 2018); Ghezal (2013, 2024), and Ghezal and Zemmouri (2023), in which the process follows locally from a bilinear representation in the sense of Subba Rao (1981, 1983); Subba Rao and Gabr (1984). For such models, the time series still switches between different regimes but, within each regime, the relationship between variables involves products of their lagged values. This non-linearity captures more complex patterns, such as volatility clustering, leverage effects, or sudden shifts in the data behavior. MS BL models cover many commonly used models in the literature, as e.g. those cited above, and give a general flexible and parsimonious framework for data-sets exhibiting occasional sharp spikes or involving high-amplitude oscillations (e.g. bull or bear markets). For these reasons, several authors have considered MS BL models to drive various economic and financial time series. For information concerning stationarity, existence of moments, estimation, geometric ergodicity and statistical inference for MS BL processes see, e.g., Aknouche and Rabehi (2010); Bibi and Ghezal (2015, 2018), and Maaziz and Kharfouchi (2018).

As benefits, MS BL models improve forecasting accuracy and flexibility in capturing a large variety of time series behaviors and non-linearities, as e.g. abrupt shifts, sudden jumps or volatility clustering in patterns and intricate relationships between lagged variables. This is common in financial data, economic cycles, and market volatility. Further, unlike conventional MS models, the MS BL ones allow non-linear dynamics to vary by regime, offering a greater modeling framework. This enables the model to better reflect the changing nature of time series data across different states. Finally, such models present advantages in terms of parsimony and ease of interpretation, coupled with lower computational requirements, while remaining effective in capturing non-linear characteristics of the data, thus providing a viable alternative to more complex approaches.

However, there are some limitations of the recent developments in MS BL models both in terms of their practical implications and the theoretical foundations, highlighted in the current literature. Issues like model complexity, estimation difficulties, overfitting, interpretability, and model validation make them challenging for both researchers and practitioners. Further advancements are required to fully exploit the potential of such models, and the present paper aims to be a proposal to address a number of the above limitations.

In this paper we consider univariate Markov switching bilinear (MS BL) models, and we concentrate on some problems that were not completely solved in the cited literature. Specifically, the contribution of this paper is threefold. First, we derive general and explicit matrix formulas in closed form for the (unconditional) higher-order moments and the autocovariance function of MS BL models, and describe tractable conditions to check the higher order stationarity of such models. Notice that Bibi and Ghezal (2015) already derive results on moments and autocovariance structure of such models albeit that their formulas to put it mildly are somewhat complicated, lacking simple closed forms. The difference to their approach is that we follow an indirect construction based on a suitable Markovian representation of the initial process. This allows to get much simpler matrix expressions in closed form which are essential to improve estimation efficiency and forecasting, for easier calculation of test statistics (e.g., heteroskedasticity, skewness and kurtosis), to help assess whether the model adequately captures the underlying data distribution (validation), providing deeper insight into model dynamics, and making the model more practical and computationally feasible for real-world applications. Second, based on the vector representation of the MS BL model, we compute the spectral and bispectral density matrices of the power vector process, and show the importance of the (bi)spectral analysis in distinguishing the linear and bilinear models. Such matrices are useful for analyzing MS BL models because they provide insights into the frequency-domain characteristics, capture non-linear interactions across different time lags, and serve to better understand the asymmetric responses of shocks. Sample estimates of the spectral and bispectral densities are then proposed, and their consistency and asymptotic normality are established (under suitable assumptions). To our knowledge, the proposed (bi)spectral analysis and the corresponding asymptotic results for general MS BL models cannot be found in the literature, and complete the works of Ghezal (2013, 2024), and Ghezal and Zemmouri (2023), where only special classes of MS BL processes have been treated. The main advantage of our approach is related to the simplicity of the mathematical expressions that eases the computational effort. Third, we propose numerical and real-world examples to illustrate the obtained results, and to describe how to use the practical formulas in (bi)spectral analysis. Specifically, we discuss the replication/comparison with some numerical outputs in Aknouche and Rabehi (2010), and Maaziz and Kharfouchi (2018), and analyze various processes related to empirical applications.

The paper is structured as follows. Section 2 formally states the MS BL model studied in the paper, and restates the important results on the existence of stationary MS BL processes from Bibi and Ghezal (2015, 2018). In Sect. 3 we provide matrix formulas in closed form for the higher-order moments and the autocovariance function of the MS BL model and its power. Section 4 uses the derived autocovariance functions to obtain expressions for the corresponding spectral and bispectral density matrices. Based on these expressions we define estimators for the spectral and bispectral density matrices and derive their limiting behavior. In Sects. 5 and 6 simulation studies and empirical examples are proposed to illustrate the theoretical results. Section 7 concludes. Proofs are given in the Appendix.

2 Markov switching bilinear models

An $\mathbb{R}$-valued univariate process $(y_t)_{t\in\mathbb{Z}}$, defined on some probability space $(\Omega, \mathcal{F}, Pr)$, has a *general Markov switching bilinear representation* of type $\mathrm{MS}(M)$ $\mathrm{BL}(p, q, P, Q)$ if it is a solution of the following equation

$$y_t = \sum_{i=1}^{p} a_i(s_t)\, y_{t-i} + \sum_{j=0}^{q} b_j(s_t)\, \eta_{t-j} + \sum_{i=1}^{P} \sum_{j=1}^{Q} c_{ij}(s_t)\, y_{t-i}\, \eta_{t-j} \tag{1}$$

where $\eta_t \sim \mathrm{IID}(0, 1)$, (s_t) is an irreducible and aperiodic Markov chain with a finite state space $\Xi = \{1, \ldots, M\}$, stationary transition probabilities $p_{ij} = Pr(s_t = j|s_{t-1} = i\}$ and unconditional (or steady state) probabilities $\pi_i = Pr(s_t = i)$, for $i \in \Xi$. Since the chain (s_t) is finite, assuming it irreducible and aperiodic is enough for it to be ergodic (i.e., regular). Let $\mathrm{P} = (p_{ij})$ denote the transition probability matrix of the chain. The Markov chain is *ergodic* if exactly one of the eigenvalues of P is unity and all other eigenvalues are inside the unit circle. Under this condition there exists a stationary probability distribution of the regimes, that is, $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_M)' \in \mathbb{R}^M$. The model coefficients $a_i(s_t)$, $b_j(s_t)$ and $c_{ij}(s_t)$ are functions of the unobserved Markov chain (s_t). Here $b_0(s_t) \neq 0$. In addition, assume that (s_t) is stationary, and that (η_t) is independent of (s_t).

Equation (1) includes many commonly used models in the literature, such as standard bilinear time series models, hidden Markov models, MS ARMA models, and some classes of MS GARCH and periodic models. See Bibi and Ghezal (2015) for more details and the references therein.

In what follows, we assume that $P = p$ in (1), without loss of generality, because otherwise zeros of $a_i(\cdot)$ and $c_{ij}(\cdot)$ can be filled in.

As remarked in (Bibi and Ghezal (2015), Sect. 2.2), it is difficult to handle the product terms in (1) because of the non-linear dependence between y_t and η_{t-k} for $k > 0$. So, in order to investigate the stationarity properties of the process y_t, the cited authors restrict their attention to the *subdiagonal bilinear model*, in short $\mathrm{MS}(M)$ $\mathrm{SBL}(p, q, p, Q)$, for which $c_{ij}(\cdot) = 0$ in (1) for $i < j$ and $Q \leq p$. That is, such a model is defined by

$$y_t = \sum_{i=1}^{p} a_i(s_t)\, y_{t-i} + \sum_{j=0}^{q} b_j(s_t)\, \eta_{t-j} + \sum_{j=1}^{Q} \sum_{i=j}^{p} c_{ij}(s_t)\, y_{t-i}\, \eta_{t-j}. \tag{2}$$

We briefly review the main results from Bibi and Ghezal (2015) to make the reading self-contained, but using slightly different constructions and notations. Consider a Markovian representation of (2) (set $n = p + q$ and $N = n\,(Q + 1)$) given by

$$\mathrm{z}_t \;=\; \mathbf{\Phi}_t\,\mathrm{z}_{t-1} \;+\; \boldsymbol{\omega}_t \tag{3}$$

where:

$$\begin{aligned}
\mathrm{z}_t \;&=\; (\mathrm{x}'_t, \mathrm{x}'_t\,\eta_t, \mathrm{x}'_{t-1}\,\eta_{t-1}, \ldots, \mathrm{x}'_{t-Q+1}\,\eta_{t-Q+1})' \in \mathbb{R}^N\\
\mathrm{x}_t \;&=\; (y_t, y_{t-1}, \ldots, y_{t-p+1}, \eta_t, \eta_{t-1}, \ldots, \eta_{t-q+1})' \in \mathbb{R}^n\\
\boldsymbol{\omega}_t \;&=\; \boldsymbol{\omega}(s_t, \eta_t) \;=\; (\mathrm{b}'_t\,\eta_t, \mathrm{b}'_t\,\eta_t^2, \mathbf{0}'_{1\times n}, \ldots, \mathbf{0}'_{1\times n})' \in \mathbb{R}^N\\
\mathrm{b}_t \;&=\; \mathrm{b}(s_t) \;=\; (b_0(s_t), 0, \ldots, 0, 1, 0, \ldots, 0)' \in \mathbb{R}^n
\end{aligned}$$

and $\mathbf{\Phi}_t = \mathbf{\Phi}(s_t, \eta_t)$ is the $N \times N$ matrix given by

$$\mathbf{\Phi}_t \;=\; \begin{pmatrix}
\mathbf{\Phi}_0(s_t) & \mathbf{\Phi}_1(s_t) & \cdots & \mathbf{\Phi}_{Q-1}(s_t) & \mathbf{\Phi}_Q(s_t)\\
\mathbf{\Phi}_0(s_t)\,\eta_t & \mathbf{\Phi}_1(s_t)\,\eta_t & \cdots & \mathbf{\Phi}_{Q-1}(s_t)\,\eta_t & \mathbf{\Phi}_Q(s_t)\,\eta_t\\
\mathbf{0}_n & \boldsymbol{I}_n & \cdots & \mathbf{0}_n & \mathbf{0}_n\\
\vdots & \vdots & & \vdots & \vdots\\
\mathbf{0}_n & \mathbf{0}_n & \cdots & \boldsymbol{I}_n & \mathbf{0}_n
\end{pmatrix}$$

with $n \times n$ matrices

$$\mathbf{\Phi}_0(s_t) \;=\; \begin{pmatrix}
a_1(s_t) & a_2(s_t) & \cdots & a_{p-1}(s_t) & a_p(s_t) & b_1(s_t) & b_2(s_t) & \cdots & b_{q-1}(s_t) & b_q(s_t)\\
1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0\\
0 & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0\\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\
0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 & 0\\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0\\
0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0\\
0 & 0 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0\\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 & 0
\end{pmatrix}$$

and

$$\mathbf{\Phi}_j(s_t) \;=\; \begin{pmatrix}
c_{jj}(s_t) & \cdots & c_{pj}(s_t) & 0 & \cdots & 0\\
\vdots & & \vdots & \vdots & & \vdots\\
0 & \cdots & 0 & 0 & \cdots & 0\\
0 & \cdots & 0 & 0 & \cdots & 0\\
\vdots & & \vdots & \vdots & & \vdots\\
0 & \cdots & 0 & 0 & \cdots & 0
\end{pmatrix}$$

for $j = 1, \ldots, Q$. Here $\mathbf{0}_{m\times n}$ denotes the $m \times n$ matrix whose entries are zeros, and I_n is the $n \times n$ identity matrix. For simplicity, we set $\mathbf{0}_n = \mathbf{0}_{n\times n}$. Note that z_{t-1} depends upon y_{t-1}, y_{t-2}, ..., η_{t-1}, η_{t-2}, ...and hence it is independent of η_t.

For any function $\mathrm{f}:\Xi\to\mathbb{R}^{n\times m}$, where $\mathbb{R}^{n\times m}$ denotes the space of real $n\times m$ matrices, the following matrices are defined

$$\mathbb{P}_{\mathrm{f}^{\otimes r}} = \begin{pmatrix} p_{11}\,\mathrm{f}^{\otimes r}(1) & p_{21}\,\mathrm{f}^{\otimes r}(1) & \cdots & p_{M1}\,\mathrm{f}^{\otimes r}(1) \\ p_{12}\,\mathrm{f}^{\otimes r}(2) & p_{22}\,\mathrm{f}^{\otimes r}(2) & \cdots & p_{M2}\,\mathrm{f}^{\otimes r}(2) \\ \vdots & \vdots & & \vdots \\ p_{1M}\,\mathrm{f}^{\otimes r}(M) & p_{2M}\,\mathrm{f}^{\otimes r}(M) & \cdots & p_{MM}\,\mathrm{f}^{\otimes r}(M) \end{pmatrix}$$

which is $(Mn^r)\times(Mm^r)$, and

$$\mathbf{\Pi}_{\mathrm{f}^{\otimes r}} = \begin{pmatrix} \pi_1\,\mathrm{f}^{\otimes r}(1) \\ \pi_2\,\mathrm{f}^{\otimes r}(2) \\ \vdots \\ \pi_M\,\mathrm{f}^{\otimes r}(M) \end{pmatrix} \in \mathbb{R}^{(Mn^r)\times m^r}$$

where $\mathrm{f}^{\otimes r}(i)=\mathrm{f}(i)\otimes\cdots\otimes\mathrm{f}(i)$, r times, $r\geq 1$ (if $r=1$, then $\mathrm{f}^{\otimes r}(i)=\mathrm{f}(i)$).

Let $\mathbf{\Phi}^{(r)}(i)=E_\eta\left[\mathbf{\Phi}_t^{\otimes r}|s_t=i\right]$ denote the conditional expectation of $\mathbf{\Phi}_t^{\otimes r}$ given $s_t=i$, for all $i\in\Xi$. Collect them in a block matrix $\mathbf{\Phi}^{(r)}=(\mathbf{\Phi}^{(r)}(1),\ldots,\mathbf{\Phi}^{(r)}(M))'$.

Theorem 1 *(Bibi and Ghezal 2015). With the above notation, suppose that for some integer $r\geq 1$, it holds that $E[\eta_t^{r+2}]<\infty$ and $\rho\left(\mathbb{P}_{\mathbf{\Phi}^{(r)}}\right)<1$, where $\rho(\cdot)$ denotes the spectral radius. Then the $\mathrm{MS}(M)$ $\mathrm{SBL}(p,q,p,Q)$ model in (2) with Markovian representation (3) has a unique, causal, ergodic and strictly stationary solution (y_t), which satisfies $E[y_t^r]<\infty$.*

The existence of a *nonanticipative* (i.e., measurable with respect to the σ-field generated by (s_τ,η_τ) for all $\tau\leq t$) stationary solution of (2) is equivalent to the existence of a nonanticipative stationary solution of (3). Since (s_t,η_t) is an ergodic process, $(\mathbf{\Phi}_t,\boldsymbol{\omega}_t)$ is also a strictly stationary ergodic process. Suppose that (η_t) has finite higher-order moments. Then both $E[\log^+||\mathbf{\Phi}_t||]$ and $E[\log^+||\boldsymbol{\omega}_t||]$ are finite due to finiteness of the state space Ξ of $(s_t)_{t\in\mathbb{Z}}$ and finiteness of higher moments of $(\eta_t)_{t\in\mathbb{Z}}$. Here $\log^+ x=\max(\log x,0)$ for every positive real number x. Therefore, from Brandt (1986) and Bougerol and Picard (1992) the unique, causal, bounded in probability, strictly stationary and ergodic solution of (3) is given by almost surely (a.s.)

$$\mathrm{z}_t = \boldsymbol{\omega}_t + \sum_{k=1}^{\infty}\left(\prod_{i=0}^{k-1}\mathbf{\Phi}_{t-i}\right)\boldsymbol{\omega}_{t-k} \tag{4}$$

whenever the top Lyapunov exponent $\gamma_L(\mathbf{\Phi})$ associated with the sequence of stationary and ergodic random matrices $(\mathbf{\Phi}_t)_{t\in\mathbb{Z}}$ is strictly negative, i.e.,

$$\gamma_L(\mathbf{\Phi}) = \inf_{m\in\mathbb{N}} \left\{ E\left[\frac{1}{m} \log || \prod_{i=0}^{m-1} \mathbf{\Phi}_{t-i}|| \right] \right\} < 0. \tag{5}$$

Since all the norms are equivalent on a finite-dimensional vectorial space, the choice of the norm is unimportant in this definition.

In the general case, the Lyapunov exponent seems difficult to compute but it is easily estimated by Monte Carlo simulations since

$$\gamma_L(\mathbf{\Phi}) = \lim_{m\to\infty} \frac{1}{m} \log || \prod_{i=0}^{m-1} \mathbf{\Phi}_{t-i}|| \qquad a.s. \tag{6}$$

The following theorem provides a sufficient condition for the existence of a strictly stationary solution of (3), and hence of (2). The proof follows essentially the same arguments as in Francq and Zakoïan (2001) and Bibi and Ghezal (2015), hence it is omitted.

Theorem 2 . *Consider the* MS SBL(p, q, p, Q) *model in (2) with Markovian representation (3) and suppose* $\gamma_L(\mathbf{\Phi}) < 0$. *Then equation (3), and hence (2), has a strictly stationary solution. Moreover, the solution is unique, ergodic, nonanticipative and defined by (4).*

3 Higher-order moments and autocovariance structure

We give neat matrix expressions in closed form for higher-order moments and the autocovariance function of process (z_t) in (3), and hence (y_t) in (2). Such expressions are new and quite different to those obtained in Bibi and Ghezal (2015), who also derive results on moments and the autocovariance function but their formulas are somewhat complicated, lacking simple closed forms. Our matrix expressions have the advantage to be readily programmable in addition of greatly reducing the computational cost. Given this, we then illustrate tractable conditions to check the higher-order stationarity of MS SBL(p, q, p, Q) processes. The main advantage of our approach is related to the simplicity of the mathematical expressions that eases the computational effort.

3.1 Higher-order moments

The following results are proved in the Appendix:

Theorem 3 *Let us consider the* $\mathrm{MS}(M)$ $\mathrm{SBL}(p, q, p, Q)$ *model in (2) with Markovian representation (3). Under the conditions of Theorem 1 (case* $r = 1$*), the unconditional expectations of* (z_t) *in (3) and* (y_t) *in (2) are finite and are given by*

$$E(\mathrm{z}_t) = (\mathrm{e}' \otimes \mathrm{I}_N)\, \mathbf{\Pi}_{E[\mathrm{z}_t|s_t=\cdot]} \qquad E(y_t) = (\mathrm{e}' \otimes \boldsymbol{\delta}')\, \mathbf{\Pi}_{E[\mathrm{z}_t|s_t=\cdot]}$$

where

$$\mathbf{\Pi}_{E[\mathrm{z}_t|s_t=\cdot]} = [\mathrm{I}_{MN} - \mathbb{P}_{\mathbf{\Phi}}]^{-1}\, \mathbf{\Pi}_{\boldsymbol{\omega}}.$$

Here we set $\mathrm{e} = (1, \cdots, 1)' \in \mathbb{R}^M$, $\boldsymbol{\delta} = (1, 0, \cdots, 0)' \in \mathbb{R}^N$, $\mathbf{\Phi}(\cdot) = E_\eta[\mathbf{\Phi}_t|s_t = \cdot]$ *and* $\boldsymbol{\omega}(\cdot) = E_\eta[\boldsymbol{\omega}_t|s_t = \cdot]$.

Let $\mathbf{\Phi}^{(2)}(\cdot) = E_\eta[\mathbf{\Phi}_t^{\otimes 2}|s_t = \cdot]$, $\mathrm{D}(\cdot) = E_\eta[\mathbf{\Phi}_t \otimes \boldsymbol{\omega}_t + \boldsymbol{\omega}_t \otimes \mathbf{\Phi}_t|s_t = \cdot]$, and $\boldsymbol{\omega}^{(2)}(\cdot) = E_\eta[\boldsymbol{\omega}_t^{\otimes 2}|s_t = \cdot]$.

Theorem 4 *Let us consider the* $\mathrm{MS}(M)$ $\mathrm{SBL}(p, q, p, Q)$ *model in (2) with Markovian representation (3). Under the conditions of Theorem 1 (case* $r = 2$*), the second order moments of* (z_t) *in (3) and* (y_t) *in (2), are finite and are given by*

$$E[\mathrm{z}_t^{\otimes 2}] = (\mathrm{e}' \otimes \mathrm{I}_{N^2})\, \mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes 2}|s_t=\cdot]} \qquad E[y_t^2] = (\mathrm{e}' \otimes \boldsymbol{\delta}' \otimes \boldsymbol{\delta}')\, \mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes 2}|s_t=\cdot]}$$

where

$$\mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes 2}|s_t=\cdot]} = [\mathrm{I}_{MN^2} - \mathbb{P}_{\mathbf{\Phi}^{(2)}}]^{-1} \left[\mathbb{P}_{\mathrm{D}}\, \mathbf{\Pi}_{E[\mathrm{z}_t|s_t=\cdot]} + \mathbf{\Pi}_{\boldsymbol{\omega}^{(2)}}\right].$$

From Theorems 3 and 4 the variance of z_t (and hence y_t) is easily deduced.
Let $\mathbf{\Phi}^{(3)}(\cdot) = E_\eta[\mathbf{\Phi}_t^{\otimes 3}|s_t = \cdot]$, $\boldsymbol{\omega}^{(3)}(\cdot) = E_\eta[\boldsymbol{\omega}_t^{\otimes 3}|s_t = \cdot]$, and

$$\mathrm{D}_1^{(3)}(\cdot) = E_\eta[\boldsymbol{\omega}_t^{\otimes 2} \otimes \mathbf{\Phi}_t + \boldsymbol{\omega}_t \otimes \mathbf{\Phi}_t \otimes \boldsymbol{\omega}_t + \mathbf{\Phi}_t \otimes \boldsymbol{\omega}_t^{\otimes 2} \,|\, s_t = \cdot]$$
$$\mathrm{D}_2^{(3)}(\cdot) = E_\eta[\boldsymbol{\omega}_t \otimes \mathbf{\Phi}_t^{\otimes 2} + \mathbf{\Phi}_t \otimes \boldsymbol{\omega}_t \otimes \mathbf{\Phi}_t + \mathbf{\Phi}_t^{\otimes 2} \otimes \boldsymbol{\omega}_t \,|\, s_t = \cdot].$$

Theorem 5 *Let us consider the* $\mathrm{MS}(M)$ $\mathrm{SBL}(p, q, p, Q)$ *model in (2) with Markovian representation (3). Under the conditions of Theorem 1 (case* $r = 3$*), the third order moments of* (z_t) *in (3) and* (y_t) *in (2), are finite and are given by*

$$E[\mathrm{z}_t^{\otimes 3}] = (\mathrm{e}' \otimes \mathrm{I}_{N^3})\, \mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes 3}|s_t=\cdot]} \qquad E(y_t^3) = (\mathrm{e}' \otimes (\boldsymbol{\delta}')^{\otimes 3})\, \mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes 3}|s_t=\cdot]}$$

where

$$\mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes 3}|s_t=\cdot]} = [\mathrm{I}_{MN^3} - \mathbb{P}_{\mathbf{\Phi}^{(3)}}]^{-1} \left[\mathbb{P}_{\mathrm{D}_2^{(3)}}\, \mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes 2}|s_t=\cdot]} + \mathbb{P}_{\mathrm{D}_1^{(3)}}\, \mathbf{\Pi}_{E[\mathrm{z}_t|s_t=\cdot]} + \mathbf{\Pi}_{\boldsymbol{\omega}^{(3)}}\right].$$

The proof of the next general result is derived by iteration from the previous theorems using similar computations.

Theorem 6 *Let us consider the* $\mathrm{MS}(M)$ $\mathrm{SBL}(p, q, p, Q)$ *model in (2) with Markovian representation (3). Under the conditions of Theorem 1 and for any integer* $r > 1$*, the r-th order unconditional moments of* (z_t) *in (3) and* (y_t) *in (2), are finite and are given by*

$$E[\mathrm{z}_t^{\otimes r}] = (\mathrm{e}' \otimes \mathrm{I}_{N^r})\, \mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes r}|s_t=\cdot]} \qquad E[y_t^r] = (\mathrm{e}' \otimes (\boldsymbol{\delta}')^{\otimes r})\, \mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes r}|s_t=\cdot]}$$

where

$$\mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes r}|s_t=\cdot]} = [\mathrm{I}_{MN^r} - \mathbb{P}_{\mathbf{\Phi}^{(r)}}]^{-1} \left[\sum_{\ell=1}^{r-1} \mathbb{P}_{\mathrm{D}_\ell^{(r)}}\, \mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes \ell}|s_t=\cdot]} + \mathbf{\Pi}_{\boldsymbol{\omega}^{(r)}} \right].$$

Here $\mathrm{D}_\ell^{(r)}(i)$*,* $i = 1, \ldots, M$*, are* $(MN^r) \times (MN^\ell)$ *well-specified matrices obtained by taking expectations (with respect to* η *and conditional on* $s_t = i$*) of sums of* $\binom{r}{\ell}$ *tensor products of r-tuples formed by* ℓ *components* $\mathbf{\Phi}_t$*'s and* $(r - \ell)$ *components* $\boldsymbol{\omega}_t$*'s (e.g.,* $\mathrm{D}_1^{(2)} = \mathrm{D}$ *as in Theorem 4;* $\mathrm{D}_1^{(3)}$ *and* $\mathrm{D}_2^{(3)}$ *are as in Theorem 5). Further, set* $\mathbf{\Phi}^{(r)}(\cdot) = E_\eta[\mathbf{\Phi}_t^{\otimes r}|s_t = \cdot]$ *and* $\boldsymbol{\omega}^{(r)}(\cdot) = E_\eta[\boldsymbol{\omega}_t^{\otimes r}|s_t = \cdot]$*, as usual.*

Theorem 6 directly implies the following consequence:

Corollary 7 Under the assumptions of Theorem 1, suppose that the Markov chain (s_t) is i.i.d. Then a sufficient condition for model (2) to have a unique *r*-th order stationary solution is that $\rho(\bar{\mathbf{\Phi}}^{(r)}) < 1$, where

$$\bar{\mathbf{\Phi}}^{(r)} := \sum_{i=1}^{M} \pi_i\, \mathbf{\Phi}^{(r)}(i) \in \mathbb{R}^{n^r \times n^r}.$$

This corollary generalizes Theorem 3 of Aknouche and Rabehi (2010), case $p = P = Q = 1$ and $q = 0$.

3.2 Autocovariance function

Once second-order stationarity is ensured, it can be useful to compute the autocovariance function of (z_t) and (y_t). Throughout the section, the conditions of Theorem 1 are assumed to hold.

For computational convenience, we consider the centred version of the vector process (z_t), that is,

$$\mathrm{z}_t^* = \mathbf{\Phi}_t\, \mathrm{z}_{t-1}^* + \boldsymbol{\omega}_t^* \tag{7}$$

where $\mathrm{z}_t^* = \mathrm{z}_t - \boldsymbol{\mu}_\mathrm{z}$ and $\boldsymbol{\mu}_\mathrm{z} = E[\mathrm{z}_t]$. Then $\boldsymbol{\omega}_t^*$ is the centred residual vector such that, given $s_t = i \in \Xi$, $\boldsymbol{\omega}_t^*(i, \eta_t)$ is orthogonal to z_τ^* for all $\tau < t$. By using (3) and (7), we get $\boldsymbol{\omega}_t^* = \boldsymbol{\omega}_t + (\mathbf{\Phi}_t - \mathrm{I}_N)\, \boldsymbol{\mu}_\mathrm{z}$.

Theorem 8 *Let us consider the* $\mathrm{MS}(M)$ $\mathrm{SBL}(p, q, p, Q)$ *model in (2) with Markovian representation (3). Under the conditions of Theorem 1* $(r = 2)$*, the autocovariance functions of* (z_t^*) *in (7) and its first component* (y_t^*) *are given by*

$$\operatorname{vec}\boldsymbol{\Gamma}_{z^*}(h) = (\mathrm{e}' \otimes \mathrm{I}_{N^2})\,\boldsymbol{\Pi}_{E[z_t^* \otimes z_{t-h}^* | s_t = \cdot]}$$
$$\gamma_{y^*}(h) = \left(\mathrm{e}' \otimes (\boldsymbol{\delta}')^{\otimes 2}\right)\boldsymbol{\Pi}_{E[z_t^* \otimes z_{t-h}^* | s_t = \cdot]}$$

where

$$\boldsymbol{\Pi}_{E[z_t^* \otimes z_{t-h}^* | s_t = \cdot]} = \left(\mathbb{P}_{\boldsymbol{\Phi}}^{h} \otimes \mathrm{I}_N\right)\left(\mathrm{I}_{MN^2} - \mathbb{P}_{\boldsymbol{\Phi}^{(2)}}\right)^{-1}\boldsymbol{\Pi}_{\boldsymbol{\omega}^{*(2)}}$$

for every $h \geq 0$ (here we set $\mathrm{A}^h = \mathrm{I}_m$ if $h = 0$ for any square $m \times m$ matrix A). Furthermore, $\boldsymbol{\Gamma}_{z^*}(h) = \boldsymbol{\Gamma}'_{z^*}(-h)$ and $\gamma_{y^*}(h) = \gamma_{y^*}(-h)$ for every $h < 0$.

Using Theorem 8 above and Theorem 4 from Cavicchioli (2014a) we have the following result whose proof is omitted because follows similar arguments as in the cited paper.

Theorem 9 *Under the conditions of Theorem 1 ($r = 2$), the $\mathrm{MS}(M)$ $\mathrm{SBL}(p, q, p, Q)$ process in (7) admits a stable $\mathrm{VARMA}(p^*, q^*)$ representation where $p^*, q^* \leq MN^2$ with $N = n(Q + 1)$ and $n = p + q$. If the lag polynomials of the AR and MA parts of the VARMA representation have no roots in common, then equalities hold in the above relations.*

This result is potentially useful for statistical applications and for model selection as, for example, the identification of the regime number and parameter estimation.

4 (Bi)spectral representation of MS SBL models

We compute the spectral and bispectral density matrices for the MS SBL process (2), and show the importance of the (bi)spectral analysis in distinguishing the linear and bilinear models. Sample estimates of the spectral and bispectral densities are then proposed, and their consistency and asymptotic normality are established (under suitable assumptions). The asymptotic theory for sample (bi)spectral density can be used for testing normality. The proposed (bi)spectral analysis for MS SBL models (2) are new and complete the works of Ghezal (2013, 2024), and Ghezal and Zemmouri (2023). More precisely, the last authors derive formulas for the third-order theoretical moments for the special MS BL model $y_t = c(s_t)y_{t-k}\,e_{t-\ell} + e_t$, for $k, \ell \in \mathbb{N}$, and an expression for the spectral and bispectral density functions, under the strong hypothesis that the innovations e_t are independent of $\{(y_{s-1}, s_t) : s \leq t\}$. In this section we need only the assumption that the innovations are independent of the Markov chain. The main novelty is that the (bi)spectral representation is provided for the general MS SBL model (2) together with the asymptotic theory of sample estimators.

4.1 Spectral density matrix

It might be useful to look at the frequency domain properties of the process in (7) compared to the time domain. Some features are more easily detectable in one domain than the other. For example, periodic behaviors and frequency components are easily observed in the frequency domain, but transient events and short-duration phenomena are clearer in the time domain. Further, the model specification process might compare the empirical spectral density with the theoretical counterpart of a fitted model, providing a useful validation test. In addition, the frequency domain gives a clearer view of how different frequencies contribute to the overall behavior, and helps to understand long-term cycles, which might be difficult to detect in the time domain due to the noise and complex transitions between regimes.

As usual (see Brockwell and Davis (1991), Sect. 11) the spectral density $\mathrm{F}_{\mathrm{z}^*}$ of (z_t^*) can be obtained from $\boldsymbol{\Gamma}_{\mathrm{z}^*}(h)$ by taking its Fourier transform, that is,

$$\operatorname{vec} \mathrm{F}_{\mathrm{z}^*}(\omega) \;=\; \sum_{h=-\infty}^{\infty} \operatorname{vec} \boldsymbol{\Gamma}_{\mathrm{z}^*}(h)\, e^{-i\omega h} \tag{8}$$

where the frequency ω belongs to $[-\pi, \pi]$.

As stated in (Pataracchia (2011), p.13), we can compute the diagonal terms of the spectral density matrix $\mathrm{F}_{\mathrm{z}^*}(\omega)$ by considering the summation (also denoted by the same symbol)

$$\operatorname{vec} \mathrm{F}_{\mathrm{z}^*}(\omega) \;=\; \sum_{h=-\infty}^{\infty} \operatorname{vec} \boldsymbol{\Gamma}_{\mathrm{z}^*}(|h|)\, e^{-i\omega h}$$

which can be expressed as

$$\begin{aligned} \operatorname{vec} \mathrm{F}_{\mathrm{z}^*}(\omega) \;&=\; (\mathrm{e}' \otimes \mathrm{I}_{N^2}) \left[\sum_{h=-\infty}^{\infty} \left(\mathbb{P}_{\boldsymbol{\Phi}}^{|h|} \otimes \mathrm{I}_N \right) e^{-i\omega h} \right] \\ &\quad \times \left(\mathrm{I}_{MN^2} - \mathbb{P}_{\boldsymbol{\Phi}^{(2)}} \right)^{-1} \boldsymbol{\Pi}_{\boldsymbol{\omega}^{*(2)}} \end{aligned} \tag{9}$$

from Theorem 8 above. This is a neat matrix expression in closed form that eases the computational effort to derive the spectral density. The above series converges as $\rho(\mathbb{P}_{\boldsymbol{\Phi}^{(2)}}) < 1$, which also implies that $\rho(\mathbb{P}_{\boldsymbol{\Phi}}) < 1$.

Theorem 10 *Under the condition of Theorem 1 ($r = 2$), the spectral density matrix of the centred process (z_t^*) in (7) is given by*

$$\operatorname{vec} \mathrm{F}_{\mathrm{z}^*}(\omega) \;=\; (\mathrm{e}' \otimes \mathrm{I}_{N^2})\, [-\mathrm{I}_{MN^2} + 2\, \operatorname{Re} \mathrm{X}(\omega)]\; \mathrm{W} \tag{10}$$

where

$$\begin{aligned} \mathrm{X}(\omega) \;&=\; \left(\mathrm{I}_{MN} - \mathbb{P}_{\boldsymbol{\Phi}}\, e^{-i\omega} \right)^{-1} \otimes \mathrm{I}_N \\ \mathrm{W} \;&=\; \left(\mathrm{I}_{MN^2} - \mathbb{P}_{\boldsymbol{\Phi}^{(2)}} \right)^{-1} \boldsymbol{\Pi}_{\boldsymbol{\omega}^{*(2)}}. \end{aligned}$$

For the centred process (y_t^*), defined as the first component of (z_t^*), we get

$$\gamma_{y^*}(\omega) \;=\; \left(\mathrm{e}' \otimes (\boldsymbol{\delta}')^{\otimes 2}\right) \left[-\mathrm{I}_{MN^2} \;+\; 2\,\mathrm{Re}\,\mathrm{X}(\omega)\right] \mathrm{W} \tag{11}$$

where the frequency ω *belongs to* $[-\pi, \pi]$.

For a random sample of size T drawn from (z_t^*), the sample autocovariance function of (z_t^*) is given by

$$\widehat{\boldsymbol{\Gamma}}_{\mathrm{z}^*}(h) \;=\; \frac{1}{T} \sum_{t=1}^{T-|h|} \mathrm{z}_{t+|h|}^* \, \mathrm{z}_t^{*\prime}$$

where $h = 0, \pm 1, \ldots, \pm (T-1)$. Note that the estimation of the spectral density requires the observation of the disturbance (η_t) which are typically not observed. However, for T large we can approximate them by the standard normal distribution with zero mean and unit variance.

For the estimated autocovariance function, we can estimate the spectral density by

$$\mathrm{vec}\,\widehat{\mathrm{F}}_{\mathrm{z}^*}(\omega) = \frac{1}{2\pi} \sum_{h=-(T-1)}^{T-1} \lambda\left(\frac{h}{R}\right) \mathrm{vec}\,\widehat{\boldsymbol{\Gamma}}_{\mathrm{z}^*}(|h|)\, e^{-i\omega h} \tag{12}$$

where the following assumptions hold:

Assumption A The function $R = R(T)$ is chosen such that $R \to \infty$ and $R/T \to 0$ as T goes to infinity.

Assumption B The map $\lambda(\,\cdot\,)$ is a lag window generator, i.e., it is assumed to be bounded, even and square integrable with $\lambda(0) = 1$. See, for example, Subba Rao (1983) for details. (In *Sect.* 5 below we use the Bartlett-Priestly window).

We now provide distributional theory for the above spectral density estimator. These asymptotic results can be used for testing in empirical work.

Theorem 11 *Let the density function of* (η_t) *strictly positive, with support containing an open set centred on zero, and absolutely continuous everywhere in* $\mathbb{R}$ *with respect to the Lebesgue measure. Suppose that* z_t^* *is initialized in the infinite past. Under the conditions of Theorem 1 for all* $r \geq 1$ *the spectral density estimator in (12) is consistent and asymptotically normally distributed.*

From (Subba Rao (1983), Sect. 6) the variance-covariance matrix of the vector $\mathrm{vec}\,\widehat{\mathrm{F}}_{\mathrm{z}^*}(\omega)$ is given by

$$\mathrm{V}_T = T^{-1} R \left[\mathrm{vec}\,\mathrm{F}_{\mathrm{z}^*}(\omega)\right] \left[\mathrm{vec}\,\mathrm{F}_{\mathrm{z}^*}(\omega)\right]' \int_{-\infty}^{\infty} \lambda^2(u)\, du$$

for $\omega \neq 0$ and $\omega \neq \pi$. Thus we have

Theorem 12 *Under the hypothesis of Theorem 11, the asymptotic distribution of the spectral density estimator in (12) is given by*

$$\sqrt{T\,R^{-1}}\,[\mathrm{vec}\,\widehat{\mathrm{F}}_{\mathrm{z}^*}(\omega)\;-\;\mathrm{vec}\,\mathrm{F}_{\mathrm{z}^*}(\omega)] \xrightarrow{L} N(\mathbf{0}, \mathrm{V}(\omega))$$

where $\mathrm{V}(\omega) = T\,R^{-1}\,\mathrm{V}_T(\omega)$ *for* $\omega \neq 0$ *and* $\omega \neq \pi$.

In particular, from Theorem 12 and the obtained expression of the spectral density in Theorem 10, we get

$$\sqrt{T\,R^{-1}}\int_{-\pi}^{\pi}[\mathrm{vec}\,\widehat{\mathrm{F}}_{\mathrm{z}^*}(\omega)\;-\;\mathrm{vec}\,\mathrm{F}_{\mathrm{z}^*}(\omega)]\,d\omega \xrightarrow{L} N(\mathbf{0}, \mathrm{V})$$

where

$$\mathrm{V} = 2\pi\,(\mathrm{e}'\otimes\mathrm{I}_{N^2})\left\{\int_{-\pi}^{\pi}[-\mathrm{I}_{MN^2}\;+\;2\,\mathrm{Re}\,\mathrm{X}(\omega)]\;\mathrm{W}\mathrm{W}'\;\left[-\mathrm{I}_{MN^2}\;+\;2\,\mathrm{Re}\,\mathrm{X}'(\omega)\right]\,d\omega\right\}$$
$$\times\,(\mathrm{e}\otimes\mathrm{I}_{N^2})\int_{-\infty}^{\infty}\lambda^2(u)\,du.$$

4.2 Bispectral density matrix

Using the third ordinary cumulant, define

$$\mathrm{vec}\,\mathrm{C}_{\mathrm{z}^*}(h,k)\;=\;E[\mathrm{z}_t^*\;\otimes\;\mathrm{z}_{t-h}^*\otimes\mathrm{z}_{t-k}^*].$$

Taking the Fourier transform yields the bispectral density function of (z_t^*), that is,

$$\mathrm{vec}\,\mathrm{F}_{\mathrm{z}^*}(\omega_1,\omega_2)\;=\;\sum_{h=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\mathrm{vec}\,\mathrm{C}_{\mathrm{z}^*}(h,k)\,e^{-ih\omega_1-ik\omega_2}$$

for $\omega_1,\omega_2\in[-\pi,\pi]$. Reasoning as in **4.1** above, and to simplify computation we consider the following summation (also denoted by the same symbol):

$$\mathrm{vec}\,\mathrm{F}_{\mathrm{z}^*}(\omega_1,\omega_2)\;=\;\sum_{h=-\infty}^{\infty}\sum_{k=-\infty}^{\infty}\mathrm{vec}\,\mathrm{C}_{\mathrm{z}^*}(|h|,|k|)\,e^{-ih\omega_1-ik\omega_2}. \tag{13}$$

We now derive a matrix formula in closed form for the bispectral density function in (13). Using similar arguments as in **3.2**, for every $h,k>0$ with $h<k$, we get

$$\mathrm{vec}\,\mathrm{C}_{\mathrm{z}^*}(h,k)\;=\;(\mathrm{e}'\otimes\mathrm{I}_{N^3})\;\mathbf{\Pi}_{E[\mathrm{z}_t^*\otimes\mathrm{z}_{t-h}^*\otimes\mathrm{z}_{t-k}^*|s_t=\,\cdot]}$$

where

$$\mathbf{\Pi}_{E[\mathrm{z}_t^* \otimes \mathrm{z}_{t-h}^* \otimes \mathrm{z}_{t-k}^* | s_t = \cdot]} = \left(\mathbb{P}_{\mathbf{\Phi}}^h \otimes \mathrm{I}_{N^2}\right) \left(\mathbb{P}_{\mathbf{\Phi}^{(2)}}^{k-h} \otimes \mathrm{I}_N\right) \mathbf{\Pi}_{E[\mathrm{z}_t^{*\otimes 3} | s_t = \cdot]}$$

and $\mathbf{\Pi}_{E[\mathrm{z}_t^{*\otimes 3} | s_t = \cdot]}$ arises from Theorem 5 by obvious changes based on the centred power vector process (the case $h \geq k$ can be derived in a similar manner).

From Subba Rao (1983), p.306), a form of the bispectral estimate is given by

$$\mathrm{vec}\ \widehat{\mathrm{F}}_{\mathrm{z}^*}(\omega_1, \omega_2) = \sum_{h=-(T-1)}^{T-1} \sum_{k=-(T-1)}^{T-1} \Lambda(h/U, k/U)\ \mathrm{vec}\ \widehat{\mathrm{C}}_{\mathrm{z}^*}(h, k)\, e^{-ih\omega_1 - ik\omega_2} \quad (14)$$

where

$$\mathrm{vec}\ \widehat{\mathrm{C}}_{\mathrm{z}^*}(h, k) = \frac{1}{T} \sum_{t=1}^{T-\xi} \mathrm{z}_t^* \otimes \mathrm{z}_{t+|h|}^* \otimes \mathrm{z}_{t+|k|}^*$$

and $\xi = \max(0, |h|, |k|)$. Here $U = U(T)$ is a window parameter chosen such that $U^2/T \to 0$ as $T \to \infty$, $U \to \infty$. The two-dimensional lag window can be chosen as a product of the one-dimensional windows used in **4.1**, i.e., $\Lambda(x, y) = \lambda(x)\,\lambda(y)\,\lambda(x - y)$, where $\lambda(\,\cdot\,)$ is as above.

From Brillinger and Rosenblatt (1967a, 1967b), the variance of the bispectral estimate for (ω_1, ω_2) with $0 < \omega_2 < \omega_1$ is given by

$$\mathrm{var}\left(\mathrm{vec}\,\widehat{\mathrm{F}}_{\mathrm{z}^*}(\omega_1, \omega_2)\right) = \frac{U^2}{T}\, V\ \{\mathrm{vec}\,\mathrm{F}_{\mathrm{z}^*}(\omega_1)\, [\mathrm{vec}\,\mathrm{F}_{\mathrm{z}^*}(\omega_2)]'\} \otimes \mathrm{F}_{\mathrm{z}^*}(\omega_1, \omega_2)$$

where

$$V = \int_{-\infty}^{\infty} \int \Lambda^2(x, y)\, dx\, dy.$$

For the bispectral estimator in (14) the same results as in Theorems 11 and 12 maintain their validity.

5 Simulation studies

5.1 Moments

To check the correctness of the formulas proposed in **3.1**, we run some simulation experiments. Particularly, we first check if the proposed analytical moments approach their empirical counterparts. To do so, we simulate time series from the data generating process MS(*M*) SBL(*p*, *q*, *p*, *Q*), with $p = Q = 1$, $q = 0$, and $M = 2$, in (Maaziz and Kharfouchi (2018), Ex.4), that is, $y_t = a(s_t)\, y_{t-1} + b(s_t)\, \eta_t + c(s_t)\, y_{t-1}\eta_{t-1}$, where $s_t \in \{1, 2\}$, $\eta_t \sim \mathrm{NID}(0, 1)$ and the true parameter values are defined by

$a(s_t) = 0.1s_t + 0.05$, $b(s_t) = 1$, $c(s_t) = 0.2s_t - 0.1$, $p_{11} = 0.75$ and $p_{22} = 0.8$. This DGP is selected because it combines both regime-switching and bilinear features in the simplest possible structure, allowing us to directly verify the theoretical expressions without the confounding effects of higher-order dynamics. From a practical standpoint, such a process can mimic short-run non-linear dependencies typically encountered in financial time series (e.g., conditional heteroskedasticity and asymmetric reactions to shocks) while maintaining a relatively stable long-run regime structure.

The experiment simulates artificial time series of length $T + 50$ with $T = 100, 500, 1000$. The first 50 initial data points are discarded to minimize the effect of initial conditions. Then one thousand replications are carried out and a typical realization is plotted in Fig. 1. This figure provides an insight into how the underlying system switches between different regimes, and gives a sense of the randomness and variability inherent in the process. Even though the model is governed by deterministic rules (transition probabilities), the randomness of the state transitions introduces variability, reflecting abrupt regime-dependent changes in volatility as often found in financial returns.

In Table 1 we report the values of the true moments up to order four and their sample counterparts, together with the spectral radius of the matrices involved in

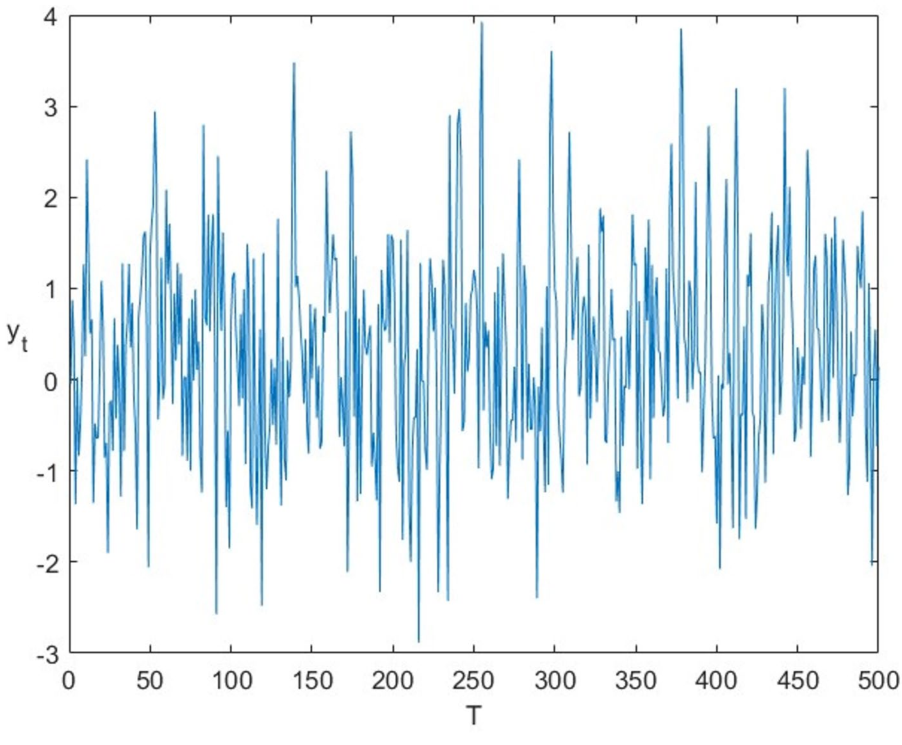


Fig. 1 A typical realization of the simulated univariate MS(2) SBL(1, 0, 1, 1) model with $T = 500$ as described in Subsection **5.1**

Table 1 True and empirical moments up to order four of the simulation exercise from the MS(2) SBL(1, 0, 1, 1) model described in Sect. 5.1

	True	$T = 100$	$T = 500$	$T = 1000$
1st moment	0.2701	0.2642	0.2757	0.2698
2nd moment	1.3058	1.2956	1.3151	1.3047
3rd moment	1.4053	1.3570	1.4481	1.4004
4th moment	13.5111	13.3641	13.6462	13.4877
$\rho(\mathbb{P}_{\mathbf{\Phi}})$	0.2178	0.2193	0.2173	0.2179
$\rho(\mathbb{P}_{\mathbf{\Phi}^{(2)}})$	0.1245	0.1223	0.1263	0.1241
$\rho(\mathbb{P}_{\mathbf{\Phi}^{(3)}})$	0.0670	0.0654	0.0683	0.0668
$\rho(\mathbb{P}_{\mathbf{\Phi}^{(4)}})$	0.0497	0.0475	0.0513	0.0494

their computations. Firstly, the assumptions of the theorems in **3.1** are all satisfied, as can be seen from the values of the computed spectral radius. Secondly, we clearly see that the empirical moments approach the true moments as the sample size increases. However, a certain discrepancy can be observed between $T = 100$ to $T = 500$ and it might be attributed to transient dynamics and finite sample bias. In fact, in MS models, the chain must "mix" well enough to ensure that it approximates its stationary distribution. If T is not large enough relative to the mixing time of the chain, the empirical moments can be biased or unstable. As the sample size increases, the mixing process has more time to take place and better reflect the true moments of the process, which is why moving from $T = 500$ to $T = 1000$ works well.

5.2 (Bi)spectral density

Consider the univariate MS(2) SBL(1, 0, 1, 1) model in 5.1. The frequency domain properties of this model are first investigated by plotting the spectral density function in Fig. 2. Here we can clearly note long term memory of the process, where the composite lagged effects play an important role. Specifically, initial value 0.72 at frequency 0 suggests that the process has a significant low-frequency (or long-run) trend component. This implies that the series exhibits persistence or "trend-like" behavior, meaning that it has a strong autocorrelation at low frequencies. Values between 0.72 and 0.78 in the frequency range 0 to 0.5 indicates that there is a moderate level of variability at low frequencies, but this variability is not increasing significantly. This suggests that the data has some cyclical but smooth behavior, with less volatility in the short run. The fact that the spectral density increases with negative convexity means that as the frequency increases, the variance at higher frequencies rises at a decreasing rate. Overall, this spectral density function indicates a system that has a dominant long-term trend with moderate periodic fluctuations at low frequencies, and increasing short-term variability that slows down as the frequency increases. Such a structure is beneficial (in practical terms) to model financial data where both long-term stability and short-term shocks or adjustments are present.

The bispectral density function of the same model is depicted in Fig. 3. This graph reveals a main peak at high frequencies and two other peaks at longer time horizons, indicating a system with nonlinear interactions across different frequency bands (which can be critical for making accurate predictions). Thus, the bispectral analysis may be useful to suggest departure from linearity identifying mixed features of the considered processes. Moreover, it may guide the researcher in the model selection

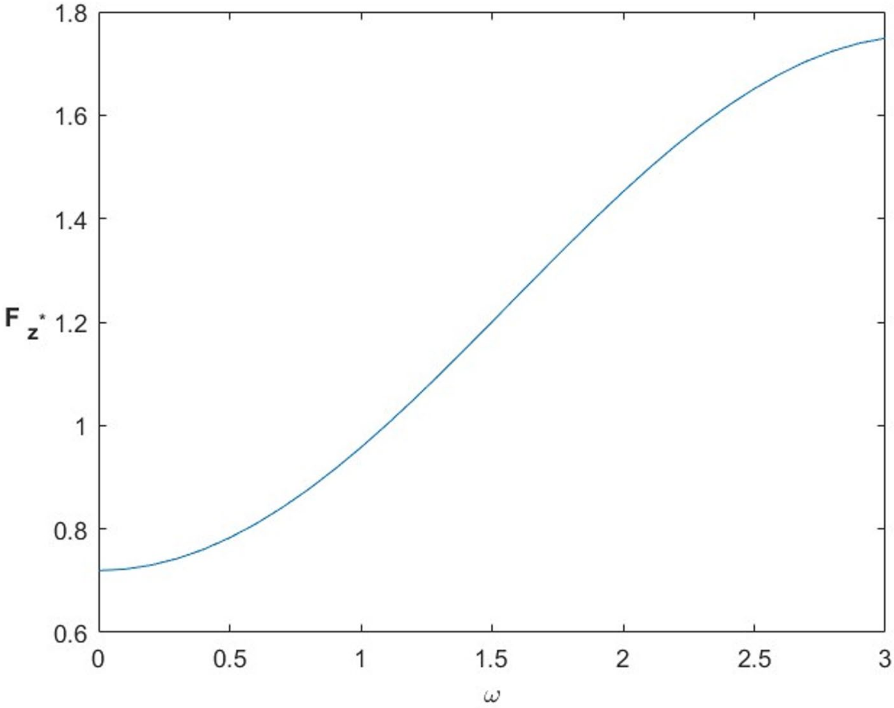


Fig. 2 Spectral density function of the univariate MS(2) SBL(1, 0, 1, 1) model described in Sect. 5.2, where ω is the frequency

phase for the identification of nonlinear characteristics, such as bilinear features (the memory of the process) and switching features (the high-frequency changes of the process), which can be beneficial for more accurate modeling of complex signals.

5.3 (Bi)spectral comparison

To shed some light on the behavior of MS SBL models under different parameterizations, we now present an high-persistence design which generates smoother regime durations and stronger autocorrelation, obtained by setting $a(s_t) = 0.45s_t$ and $p_{11} = p_{22} = 0.9$ (other parameters as before). This configuration may represent macroeconomic time series with slow regime transitions (e.g., business cycles), as can be seen from a typical realization in Fig. 4, which is substantially different from Fig. 1 (much volatile). The spectral and bispectral density functions of such DGP are depicted in Fig. 5.

While the spectrum is unable to detect differences between the two DGPs, the bispectrum clearly differentiates their features. In particular, the sharp and dominant spike at the origin indicates strong correlation among low-frequency components, which is typical of non-linear yet stationary processes, accompanied by a symmetric cyclical decay away from the origin. This pattern is consistent with business-cycle or trend interactions (e.g., growth/trend mixing with slow cyclical components). The

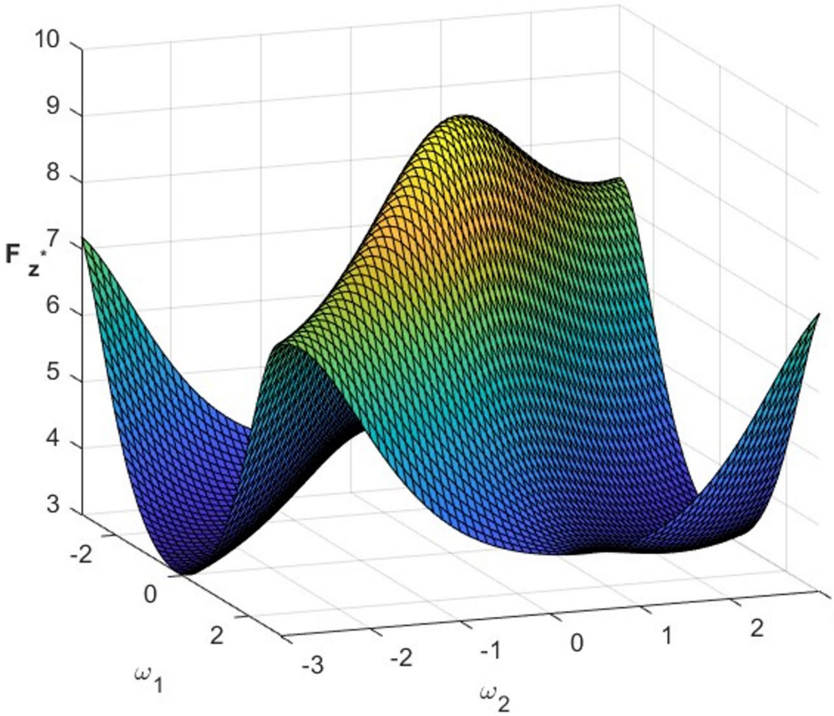
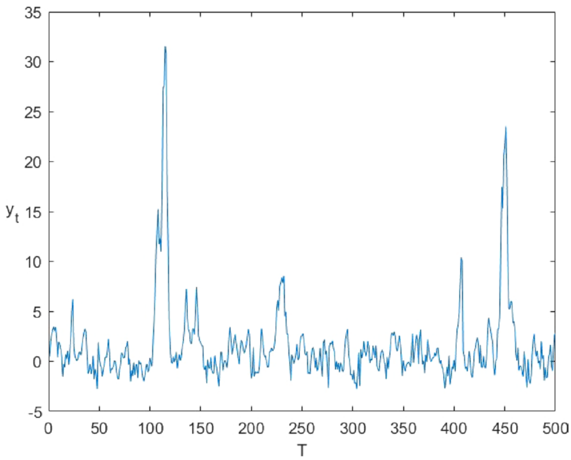


Fig. 3 Bispectral density function of the univariate MS(2) SBL(1, 0, 1, 1) model described in Sect. 5.2, where ω_1 and ω_2 are the frequencies

Fig. 4 A typical realization of the simulated univariate MS(2) SBL(1, 0, 1, 1) model with $T = 500$ as described in Sect. 5.3



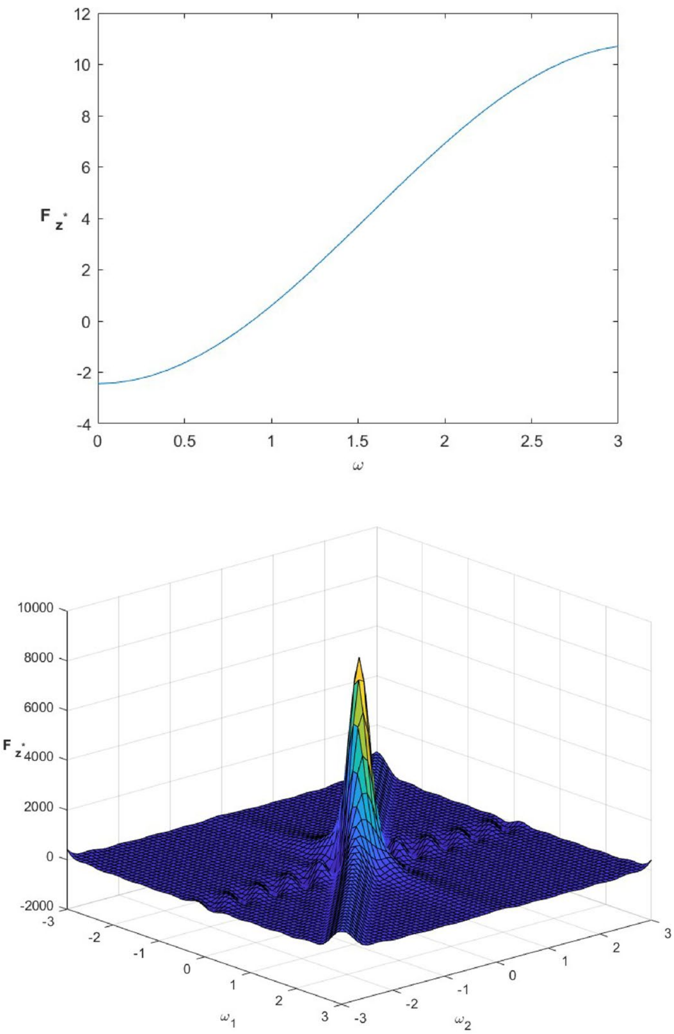


Fig. 5 Spectral and bispectral density function of the univariate MS(2) SBL(1, 0, 1, 1) model described in Sect. 5.3, where ωs are the frequencies

elevated central peak is likely influenced by a persistent level (trend), which can inflate the bispectral magnitude at the origin and suggest the presence of regime-dependent dynamics (e.g., expansions versus recessions).

The simulation results across the different DGPs consistently demonstrate that bispectral properties of MS SBL models are reliable and robust to changes in persistence, non-linearity, and volatility structure. In fact, the bispectral analyses reveal how distinct parameterizations can be associated with meaningful empirical behaviors: smooth, persistent cycles typical of macroeconomic indicators, or rapid, high-frequency adjustments as observed in financial data.

As a final note, we point out that simulations and empirical exercises were performed using Matlab(2024a) in a desktop (Intel Core i7CPU with 32 GB Ram). Codes are given in the Supplementary materials.

6 Empirical applications

6.1 Canadian lynx

We apply our new tools to the well-known dataset on the Canadian Lynx data. It comprises annual records of lynx trappings in the Mackenzie River district of North-West Canada from 1821 to 1934. These data are publicly available and widely used in time series analysis. Such series has been studied in different works: (Subba Rao and Gabr (1984), p.93) estimate a standard bilinear model (with 12 lags), Aknouche and Rabehi (2010) apply a 2-state mixture bilinear model, while (Maaziz and Kharfouchi 2018, Sect. 5.1) consider a 2-state Markov switching bilinear model. Since the first two models are nested in the third, we are able to provide a comparison between the graphical representations of spectral and bispectral density functions. In Fig. 6 we report the spectra density functions of the three estimated models: bilinear (BL), mixture BL and MS BL. The spectral density of the MS BL model differs from those of the BL and mixture BL models because the former specification introduces distinct regimes, each with its own bilinear dynamics. The unconditional spectrum is then obtained as a weighted combination of the regime-specific spectra according to the stationary distribution of the Markov chain. Such an averaging effect smooths out the peaks at the lower frequency (say, ω less than 1.5) that are characteristic of the BL-type dynamics and results in highlighting the high frequency explosion of the series (say, ω greater than 1.5). Nevertheless, despite these differences, the second-order spectra alone are not sufficient to capture and fully discriminate the non-linear features of the data.

On the contrary, from the bispectral densities reported in Fig. 7, we can clearly differenciate the different behavior of the estimated models. The first and second bispectral densities of Fig. 7 relate to the standard bilinear and mixture bilinear models, respectively. Here the dominant spikes at high frequencies do not capture

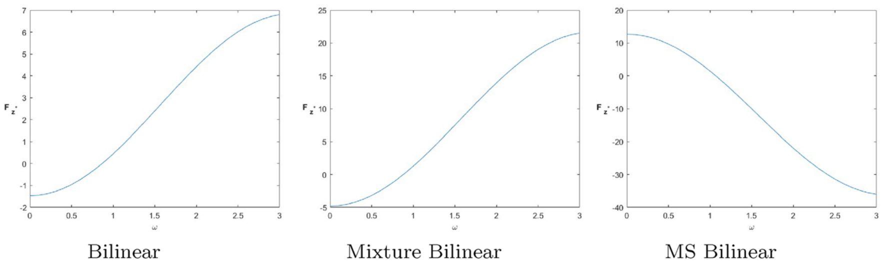


Fig. 6 Spectral density functions of the standard bilinear model by Subba Rao and Gabr (1984), p.93), the mixture bilinear model by Aknouche and Rabehi (2010), and the MS bilinear model by (Maaziz and Kharfouchi (2018), Sect. 5.1)

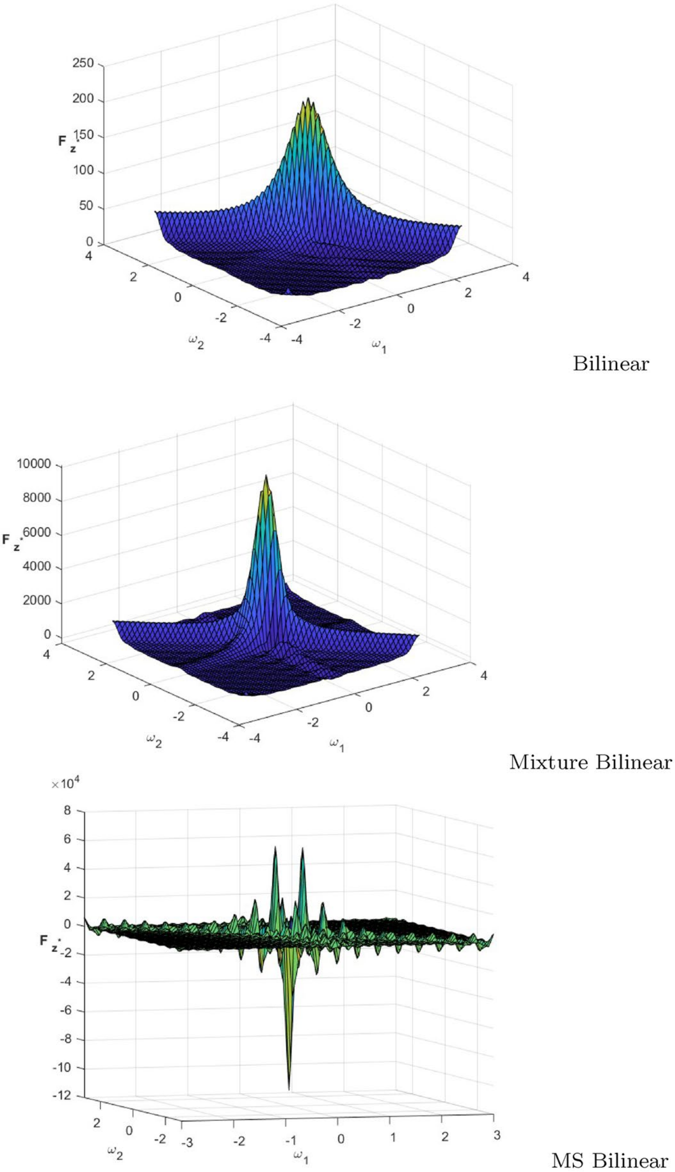


Fig. 7 Bilinear Mixture Bilinear MS Bilinear Bispectral density functions of the standard bilinear model by Subba Rao and Gabr (1984), p.93), the mixture bilinear model by Aknouche and Rabehi (2010), and the MS bilinear model by (Maaziz and Kharfouchi (2018), Sect. 5.1)

any recurrent behavior at lower frequencies. The bottom plot, instead, relates to the Markov Switching bilinear model that is able to capture the bimodality supported also by Chan and Tong (2004). Moreover, the plot shows regular jumps at various frequencies, highlighting the different states of the world as well as their progressive decreases show the typical cyclical behavior of the data. Finally, from the latter bispectral density function, the presence of high-magnitude peaks and structured spectral interactions strongly suggests that the underlying data exhibits leptokurtic behavior. This means that the data likely contains heavy-tailed distributions, where extreme values occur more frequently than in a normal distribution.

6.2 Stock prices

Firstly, as a worked example, we apply our findings to a well-known dataset previously used in Aknouche and Rabehi (2010), i.e. IBM stock closing price data (source: Box et al. 1994). We preliminary analyze this time series because the key characteristics of the series are already known to us. In fact, such data are subject to multimodality and non-linear features which are detected in the bispectral representation. The bispectral density is reported in Fig. 8 where bimodality is actually present at lower frequencies indicating recurrent long cycles. In particular, sharp peaks and valleys in the graph show strong nonlinear interactions between frequencies and the presence of multiple lobes suggests periodic or recurrent market behaviors. Such characterization via spectral analysis retraces the common features identified by the mentioned researchers. Furthermore, the bispectral density function suggests a moderately leptokurtic distribution due to higher spectral kurtosis values with respect to the Gaussian ones and structured peaks.

Next, we analyze a more recent dataset, i.e. daily log returns of the S&P500 stock price index from June 3, 2013 to December 30, 2022 (total number of observations: 2,414). Data are publicly available in Yahoo Finance. These data exhibit several stylized facts that suggest the need for nonlinear models, including the absence of serial correlation, a heavy-tailed marginal distribution, and volatility clustering. Francq and Zakoïan (2008) applied a two-state MS-GARCH(1,1) and their results demonstrated that this model outperforms the standard GARCH approach. We estimate a MS bilinear model of type MS(2) SBL(1,0,1,1) using the EM algorithm proposed by Maaziz and Kharfouchi (2018), and the bispectral density function for such a series is

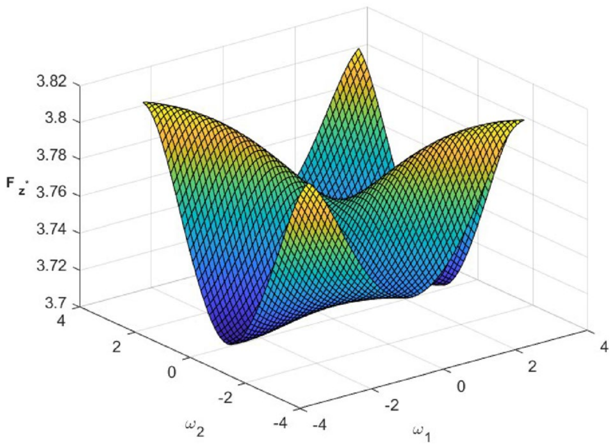


Fig. 8 Bispectral density function of the IBM stock prices mixture bilinear model in Aknouche and Rabehi (2010)

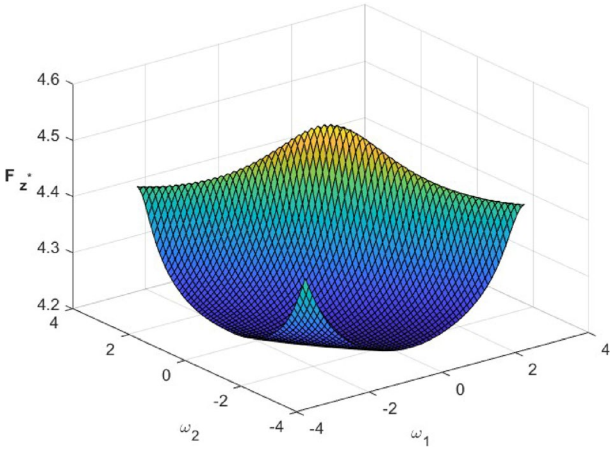


Fig. 9 Bispectral density function of the S&P500 stock price index modelled as a MS SBL model

reported in Fig. 9. Here the central peak suggests a dominant non-linear effect rather than multiple interacting regions, with higher movements at high frequencies of the business cycle. Moreover, the S&P500 bispectrum exhibits a more global non-linear dependency, possibly due to broader market trends influencing higher frequencies. In addition, the bispectral density function suggests that the data exhibits strong leptokurtic behavior, as seen from the high spectral kurtosis values. Here the absence of extreme peaks suggests a more continuous heavy-tailed distribution, rather than one dominated by extreme spikes. This could indicate financial return data or other real-world signals where extreme deviations occur more frequently than in a Gaussian process.

To conclude, the comparison of the bispectral plots of two stock prices reveal different non-linear behaviors. The first IBM plot exhibits more localized and distinct non-linearity, while the S&P500 plot suggests a more uniform and globally non-linear structure. This could imply that individual stocks experience more pronounced, event-driven market shocks, whereas an index like the S&P500 reflects overall market sentiment with less distinct non-linear frequency interactions and higher movements at the business cycle frequencies.

7 Conclusions

We analyze the probabilistic structure of Markov switching bilinear models, and derive explicit matrix expressions in closed form for higher-order moments and autocovariances of the process and its powers. These matrix expressions improve computational performance since they are readily programmable and greatly reduce the computational cost. We also give sufficient conditions for ARMA representations of these processes. These results generalize various methods that have been proposed in the literature for standard bilinear models. Then we discuss spectral and bispectral representations of MS subdiagonal bilinear models, and provide distributional theory for sample spectral and bispectral density estimators. This serves to derive statistical tests for linearity and Gaussianity of such general processes. To our knowledge, the proposed (bi)spectral analysis and the corresponding asymptotic results for the general MS SBL model (2) cannot be found in the literature. Some authors partially develop such a theory in different manner and only for special classes of MS SBL models. Deriving closed-form matrix expressions for higher-order moments, autocovariance and (bi)spectral density of MS SBL models provides numerous practical benefits, ranging from computational efficiency in estimation to improved model understanding (interpretability, model diagnostics and validation, frequency-domain behavior) and predictive capabilities (forecasting and uncertainty quantification), which are critical for making informed decisions in various applied fields. Numerical and empirical examples suggest that the switching recurrence plays an important role in the frequency decomposition of regime switching bilinear models.

To complete our proposal, we plan to extend our results by exploring alternative directions. First, in the context of MS bilinear models, we aim to study forecasting accuracy to evaluate model performance over short-, medium-, and long-term forecast horizons. Second, it would be of interest to analyze spectral and bispectral characterizations for other classes of non-linear models (e.g., NN-GARCH, MS-FIGARCH, MS-ARMA-GARCH, TAR, or TR-GARCH). Third, we intend to further investigate the integration of the MS framework in bilinear models and its contribution in terms of structural, predictive, and computational parsimony.

Appendix

Proof of Theorem 3 Starting from (3), for $i = 1, \ldots, M$, we have

$$\begin{aligned}
&\pi_i \, E[\mathrm{z}_t \,|\, s_t = i] \;=\; \pi_i \, E[\boldsymbol{\omega}_t \;+\; \boldsymbol{\Phi}_t \, \mathrm{z}_{t-1} \,|\, s_t = i] \\
&= \pi_i \, E_\eta[\boldsymbol{\omega}_t | s_t = i] \;+\; \pi_i \, E_\eta[\boldsymbol{\Phi}_t \, \mathrm{z}_{t-1} \,|\, s_t = i] \\
&= \pi_i \, \boldsymbol{\omega}(i) \;+\; \pi_i \, E_\eta[\boldsymbol{\Phi}_t | s_t = i] \, E[\mathrm{z}_{t-1} \,|\, s_t = i] \\
&= \pi_i \, \boldsymbol{\omega}(i) \;+\; \pi_i \, \boldsymbol{\Phi}(i) \sum_{j=1}^{M} E[\mathrm{z}_{t-1} \,|\, s_t = i, s_{t-1} = j] \, p_{ji} \, \pi_j \\
&= \pi_i \, \boldsymbol{\omega}(i) \;+\; \sum_{j=1}^{M} p_{ji} \, \boldsymbol{\Phi}(i) \quad \pi_j \, E[\mathrm{z}_{t-1} \,|\, s_{t-1} = j]
\end{aligned}$$

as $\boldsymbol{\Phi}_t$ is is a function of s_t and η_t which is independent of z_{t-1}. In vector form, we have

$$\boldsymbol{\Pi}_{E[\mathrm{z}_t | s_t = \cdot]} \;=\; \mathbb{P}_{\boldsymbol{\Phi}} \, \boldsymbol{\Pi}_{E[\mathrm{z}_t | s_t = \cdot]} \;+\; \boldsymbol{\Pi}_{\boldsymbol{\omega}}$$

hence

$$(\mathrm{I}_{MN} - \mathbb{P}_{\mathbf{\Phi}})\,\mathbf{\Pi}_{E[\mathrm{z}_t\,|\,s_t=\,\cdot\,]} = \mathbf{\Pi}_{\boldsymbol{\omega}}.$$

Since $\rho(\mathbb{P}(\mathbf{\Phi})) < 1$ (see Theorem 1 for $r = 1$), the matrix $\mathrm{I}_{MN} - \mathbb{P}(\mathbf{\Phi})$ is invertible. This implies the last equation of the statement of Theorem 3.1. Finally, we have $E[\mathrm{z}_t] = \sum_{i=1}^{M} \pi_i\, E[\mathrm{z}_t \,|\, s_t = i] = (\mathrm{e}' \otimes \mathrm{I}_N)\,\mathbf{\Pi}_{E[\mathrm{z}_t|s_t=\,\cdot\,]}$. The unconditional expectation of (y_t) is now easily deduced as y_t is the first component of z_t. □

Proof of Theorem 4 Starting from (3) and reasoning as above, we have

$$\begin{aligned}
\pi_i\, E[\mathrm{z}_t^{\otimes 2} \,|\, s_t = i] &= \pi_i\, E\left[(\boldsymbol{\omega}_t + \mathbf{\Phi}_t\, \mathrm{z}_{t-1})^{\otimes 2} \,|\, s_t = i\right] \\
&= \pi_i\, E_\eta[\boldsymbol{\omega}_t^{\otimes 2} | s_t = i] + \pi_i\, E_\eta[\boldsymbol{\omega}_t \otimes \mathbf{\Phi}_t + \mathbf{\Phi}_t \otimes \boldsymbol{\omega}_t \,|\, s_t = i]\, E[\mathrm{z}_{t-1} \,|\, s_t = i] \\
&\quad + \pi_i\, E_\eta\left[\mathbf{\Phi}_t^{\otimes 2} \,|\, s_t = i\right]\, E[\mathrm{z}_{t-1}^{\otimes 2} \,|\, s_t = i] \\
&= \pi_i\, \boldsymbol{\omega}^{(2)}(i) + \pi_i\, \mathrm{D}(i)\, E[\mathrm{z}_{t-1} \,|\, s_t = i] + \pi_i\, \mathbf{\Phi}^{(2)}(i)\, E[\mathrm{z}_{t-1}^{\otimes 2} \,|\, s_t = i] \\
&= \pi_i\, \boldsymbol{\omega}^{(2)}(i) + \pi_i\, \mathrm{D}(i) \sum_{j=1}^{M} E[\mathrm{z}_{t-1} \,|\, s_t = i, s_{t-1} = j]\, p_{ji}\, \pi_j \\
&\quad + \pi_i\, \mathbf{\Phi}^{(eq2)}(i) \sum_{j=1}^{M} E[\mathrm{z}_{t-1}^{\otimes 2} \,|\, s_t = i, s_{t-1} = j]\, p_{ji}\, \pi_j \\
&= \pi_i\, \boldsymbol{\omega}^{(2)}(i) + \sum_{j=1}^{M} p_{ji}\, \mathrm{D}(i) \quad \pi_j\, E[\mathrm{z}_{t-1} \,|\, s_{t-1} = j] \\
&\quad + \sum_{j=1}^{M} p_{ji}\, \mathbf{\Phi}^{(2)}(i) \quad \pi_j\, E[\mathrm{z}_{t-1}^{\otimes 2} \,|\, s_{t-1} = j]
\end{aligned}$$

for $i = 1, \ldots, M$. For this computations, the fact that the matrices $\mathbf{\Phi}_t$ and the vectors $\boldsymbol{\omega}_t$ are independent of z_{t-1} conditional s_t is crucial.

In vector form

$$\mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes 2}|s_t=\,\cdot\,]} = \mathbb{P}_{\mathbf{\Phi}^{(2)}}\, \mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes 2}|s_t=\,\cdot\,]} + \mathbb{P}_{\mathrm{D}}\, \mathbf{\Pi}_{E[\mathrm{z}_t|s_t=\,\cdot\,]} + \mathbf{\Pi}_{\boldsymbol{\omega}^{(2)}}.$$

Since the spectral radius of $\mathbb{P}_{\mathbf{\Phi}^{(2)}}$ is less than one (see Theorem 1 for $r = 2$), the matrix $\mathrm{I}_{MN^2} - \mathbb{P}_{\mathbf{\Phi}^{(2)}}$ is invertible. This implies the last equation of the statement. Then the result follows as above. □

Proof of Theorem 5 Starting from (3) and reasoning as above, for $i = 1, \ldots, M$ we have

$$\pi_i\, E[\mathrm{z}_t^{\otimes 3}\,|\, s_t = i] \;=\; \pi_i\, E[(\boldsymbol{\omega}_t \,+\, \boldsymbol{\Phi}_t\, \mathrm{z}_{t-1})^{\otimes 3}\,|\, s_t = i]$$
$$= \pi_i\, E_\eta[\boldsymbol{\omega}_t^{\otimes 3}\,|\, s_t = i] \,+\, \pi_i\, E_\eta[\boldsymbol{\omega}_t^{\otimes 2} \otimes \boldsymbol{\Phi}_t \,+\, \boldsymbol{\omega}_t \otimes \boldsymbol{\Phi}_t \otimes \boldsymbol{\omega}_t \,+\, \boldsymbol{\Phi}_t \otimes \boldsymbol{\omega}_t^{\otimes 2})\,|\, s_t = i]$$
$$\times\; E[\mathrm{z}_{t-1}\,|\, s_t = i]$$
$$+\, \pi_i\, E_\eta[\boldsymbol{\omega}_t \otimes \boldsymbol{\Phi}_t^{\otimes 2} \,+\, \boldsymbol{\Phi}_t \otimes \boldsymbol{\omega}_t \otimes \boldsymbol{\Phi}_t \,+\, \boldsymbol{\Phi}_t^{\otimes 2} \otimes \boldsymbol{\omega}_t\,|\, s_t = i]\, E[\mathrm{z}_{t-1}^{\otimes 2}\,|\, s_t = i]$$
$$+\, \pi_i\, E_\eta[\boldsymbol{\Phi}_t^{\otimes 3}\,|\, s_t = i]\, E[\mathrm{z}_{t-1}^{\otimes 3}\,|\, s_t = i]$$
$$=\; \pi_i\, \boldsymbol{\omega}^{(3)}(i) \,+\, \pi_i\, \mathrm{D}_1^{(3)}(i)\, E[\mathrm{z}_{t-1}\,|\, s_t = i] \,+\, \pi_i\, \mathrm{D}_2^{(3)}(i)\, E[\mathrm{z}_{t-1}^{\otimes 2}\,|\, s_t = i]$$
$$+\, \pi_i\, \boldsymbol{\Phi}^{(3)}(i)\, E[\mathrm{z}_{t-1}^{\otimes 3}\,|\, s_t = i]$$
$$=\; \pi_i\, \boldsymbol{\omega}^{(3)}(i) \,+\, \sum_{j=1}^{M} p_{ji}\, \mathrm{D}_1^{(3)}(i)\; \pi_j\, E[\mathrm{z}_{t-1}\,|\, s_{t-1} = j]$$
$$+\, \sum_{j=1}^{M} p_{ji}\, \mathrm{D}_2^{(3)}(i)\; \pi_j\, E[\mathrm{z}_{t-1}^{\otimes 2}\,|\, s_{t-1} = j] \,+\, \sum_{j=1}^{M} p_{ji}\, \boldsymbol{\Phi}^{(3)}(i)\, \pi_j\, E[\mathrm{z}_{t-1}^{\otimes 3}\,|\, s_{t-1} = j].$$

In vector form

$$\boldsymbol{\Pi}_{E[\mathrm{z}_t^{\otimes 3}|s_t=\,\cdot]} = \mathbb{P}_{\boldsymbol{\Phi}^{(3)}}\, \boldsymbol{\Pi}_{E[\mathrm{z}_t^{\otimes 3}|s_t=\,\cdot]} \,+\, \mathbb{P}_{\mathrm{D}_2^{(3)}}\, \boldsymbol{\Pi}_{E[\mathrm{z}_t^{\otimes 2}|s_t=\,\cdot]} \,+\, \mathbb{P}_{\mathrm{D}_1^{(3)}}\, \boldsymbol{\Pi}_{E[\mathrm{z}_t|s_t=\,\cdot]} \,+\, \boldsymbol{\Pi}_{\boldsymbol{\omega}^{(3)}}.$$

Since the spectral radius of $\mathbb{P}_{\boldsymbol{\Phi}^{(3)}}$ is less than one (see Theorem 1 for $r = 3$), the matrix $\mathrm{I}_{MN^3} - \mathbb{P}_{\boldsymbol{\Phi}^{(3)}}$ is invertible. This implies the last equation of the statement. Then the result follows. □

Proof of Theorem 6 Starting from (3) and reasoning as above, for $i = 1, \ldots, M$ we have

$$\pi_i\, E[\mathrm{z}_t^{\otimes r}\,|\, s_t = i] \;=\; \pi_i\, E\left[(\boldsymbol{\omega}_t \,+\, \boldsymbol{\Phi}_t\, \mathrm{z}_{t-1})^{\otimes r}\,|\, s_t = i\right]$$
$$= \pi_i\, E\left[\boldsymbol{\omega}_t^{\otimes r} \,+\, \sum_{\ell=1}^{r-1} \mathrm{B}_\ell^{(r)}\, \mathrm{z}_{t-1}^{\otimes \ell} \,+\, \boldsymbol{\Phi}_t^{\otimes r}\, \mathrm{z}_{t-1}^{\otimes r}\,|\, s_t = i\right]$$
$$= \pi_i\, E_\eta[\boldsymbol{\omega}_t^{\otimes r}\,|\, s_t = i] \,+\, \pi_i \sum_{\ell=1}^{r-1} E_\eta[\mathrm{B}_\ell^{(r)}\,|\, s_t = i]\, E[\mathrm{z}_{t-1}^{\otimes \ell}\,|\, s_t = i]$$
$$+\, \pi_i\, E_\eta[\boldsymbol{\Phi}_t^{\otimes r}\,|\, s_t = i]\, E[\mathrm{z}_{t-1}^{\otimes r}\,|\, s_t = i]$$
$$=\; \pi_i\, \boldsymbol{\omega}^{(r)}(i) \,+\, \pi_i \sum_{\ell=1}^{r-1} \mathrm{D}_\ell^{(r)}(i)\, E[\mathrm{z}_{t-1}^{\otimes \ell}\,|\, s_t = i] \,+\, \pi_i\, \boldsymbol{\Phi}^{(r)}(i)\, E[\mathrm{z}_{t-1}^{\otimes r}\,|\, s_t = i]$$
$$=\; \pi_i\, \boldsymbol{\omega}^{(r)}(i) \,+\, \sum_{\ell=1}^{r-1} \sum_{j=1}^{M} p_{ji}\, \mathrm{D}_\ell^{(r)}(i)\; \pi_j\, E[\mathrm{z}_{t-1}^{\otimes \ell}\,|\, s_{t-1} = j]$$
$$+\, \sum_{j=1}^{M} p_{ji}\, \boldsymbol{\Phi}^{(r)}(i)\; \pi_j\, E[\mathrm{z}_{t-1}^{\otimes r}\,|\, s_{t-1} = j]$$

where $\mathrm{D}_\ell^{(r)}(i) = E_\eta[\mathrm{B}_\ell^{(r)} \mid s_t = i]$. In vector form, the last relation becomes

$$\mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes r}|s_t=\cdot]} = \mathbb{P}_{\mathbf{\Phi}^{(r)}} \mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes r}|s_t=\cdot]} + \sum_{\ell=1}^{r-1} \mathbb{P}_{\mathrm{D}_\ell^{(r)}} \mathbf{\Pi}_{E[\mathrm{z}_t^{\otimes \ell}|s_t=\cdot]} + \mathbf{\Pi}_{\boldsymbol{\omega}^{(r)}}.$$

Since the spectral radius of $\mathbb{P}_{\mathbf{\Phi}^{(r)}}$ is less than one, the matrix $\mathrm{I}_{MN^r} - \mathbb{P}_{\mathbf{\Phi}^{(r)}}$ is invertible. This proves the last equation of the statement. Then the result follows as above. □

Proof of Theorem 8 For every $h > 0$ we have

$$\begin{aligned}
\pi_i\, E[\mathrm{z}_t^* \otimes \mathrm{z}_{t-h}^* | s_t = i] &= \pi_i\, E[(\mathbf{\Phi}_t\, \mathrm{z}_{t-1}^* + \boldsymbol{\omega}_t^*) \otimes \mathrm{z}_{t-h}^* | s_t = i] \\
&= \pi_i\, (E_\eta[\mathbf{\Phi}_t | s_t = i] \otimes \mathrm{I}_N)\; E[\mathrm{z}_{t-1}^* \otimes \mathrm{z}_{t-h}^* | s_t = i] + \pi_i\, E[\boldsymbol{\omega}_t^* \otimes \mathrm{z}_{t-h}^* | s_t = i] \\
&= \pi_i\, [\mathbf{\Phi}(i) \otimes \mathrm{I}_N] \sum_{j=1}^{M} E[\mathrm{z}_{t-1}^* \otimes \mathrm{z}_{t-1-(h-1)}^* | s_t = i, s_{t-1} = j]\, p_{ji}\, \pi_j \\
&= \sum_{j=1}^{M} p_{ji}\, [\mathbf{\Phi}(i) \otimes \mathrm{I}_N]\; \pi_j\, E[\mathrm{z}_{t-1}^* \otimes \mathrm{z}_{t-1-(h-1)}^* | s_{t-1} = j].
\end{aligned}$$

In vector form

$$\mathbf{\Pi}_{E[\mathrm{z}_t^* \otimes \mathrm{z}_{t-h}^* | s_t=\cdot]} = (\mathbb{P}_{\mathbf{\Phi}} \otimes \mathrm{I}_N)\, \mathbf{\Pi}_{E[\mathrm{z}_t^* \otimes \mathrm{z}_{t-(h-1)}^* | s_t=\cdot]}$$

hence by iteration

$$\mathbf{\Pi}_{E[\mathrm{z}_t^* \otimes \mathrm{z}_{t-h}^* | s_t=\cdot]} = \left(\mathbb{P}_{\mathbf{\Phi}}^h \otimes \mathrm{I}_N\right) \mathbf{\Pi}_{E[\mathrm{z}_t^* \otimes \mathrm{z}_t^* | s_t=\cdot]}$$

for every $h \geq 0$. Now the matrix expression in closed form of the term $\mathbf{\Pi}_{E[\mathrm{z}_t^* \otimes \mathrm{z}_t^* | s_t=\cdot]}$ is given by Theorem 4. □

Proof of Theorem 10 We have

$$\begin{aligned}
\operatorname{vec} \mathrm{F}_{\mathrm{z}^*}(\omega) &= (\mathrm{e}' \otimes \mathrm{I}_{N^2}) \left[\sum_{h=0}^{\infty} \left(\mathbb{P}_{\mathbf{\Phi}}\, e^{-i\omega}\right)^h \otimes \mathrm{I}_N + \sum_{k=1}^{\infty} \left(\mathbb{P}_{\mathbf{\Phi}}\, e^{i\omega}\right)^k \otimes \mathrm{I}_N\right] \\
&\quad \times (\mathrm{I}_{MN^2} - \mathbb{P}_{\mathbf{\Phi}^{(2)}})^{-1}\, \mathbf{\Pi}_{\boldsymbol{\omega}^{*(2)}} \\
&= (\mathrm{e}' \otimes \mathrm{I}_{N^2}) \left[-\mathrm{I}_{MN^2} + \sum_{h=0}^{\infty} \left(\mathbb{P}_{\mathbf{\Phi}}\, e^{-i\omega}\right)^h \otimes \mathrm{I}_N + \sum_{k=0}^{\infty} \left(\mathbb{P}_{\mathbf{\Phi}}\, e^{i\omega}\right)^k \otimes \mathrm{I}_N\right] \mathrm{W} \\
&= (\mathrm{e}' \otimes \mathrm{I}_{N^2}) \left\{-\mathrm{I}_{MN^2} + \left[\left(\mathrm{I}_{MN} - \mathbb{P}_{\mathbf{\Phi}}\, e^{-i\omega}\right)^{-1} + \left(\mathrm{I}_{MN} - \mathbb{P}_{\mathbf{\Phi}}\, e^{i\omega}\right)^{-1}\right] \otimes \mathrm{I}_N\right\} \mathrm{W}
\end{aligned}$$

as $\rho(\mathbb{P}_{\mathbf{\Phi}}) < 1$. Here W is defined as in the statement. It is a finite matrix as $\rho(\mathbb{P}_{\mathbf{\Phi}^{(2)}}) < 1$. □

Proof of Theorem 11 As (s_t) is stationary and ergodic, the joint random sequence (η_t, s_t) is stationary and ergodic. Because (y_t) is a function of the infinite past of the Markov chain (s_t) and the noise (η_t), the process (y_t) is ergodic. By the existence of the higher order moments and the stationarity and ergodicity of (z_t^*) resp. (y_t^*) and their powers, we can apply the ergodic theorem to establish the consistency of the empirical autocovariances. Then the sample autocovariance estimator converges to the true autocovariance with probability one as $T \to \infty$. The consistency of the sample spectral density estimator follows. The assumptions of the theorem imply the geometric ergodicity of (z_t^*, s_t), and hence (y_t^*, s_t), which are therefore strongly mixing with geometric rate. Such properties hold for the MS SBL(p, q, p, Q) process (y_t) by Theorem 4.1 from Bibi and Ghezal (2015). The geometric ergodicity ensures not only that a unique stationary probability measure for the process (z_t^*, s_t) exists, but also that the Markov chain converges to it at a geometric rate with respect to the total variation norm. Markov switching processes with such properties satisfy conventional limit theorems such as the law of large numbers and the central limit theorem (CLT) given the existence of higher order moments. See, for example, Jones (2004), Theorem 1 and Corollary 5. Geometric ergodicity implies that if the process is initiated from its infinite past, it is regular mixing with exponential decaying mixing numbers. See (Doukhan (1994), Sect. 1.1). By (Zivot (2013), Remark 3 after Theorem 28), any transformation $g(\cdot)$ of a stationary and geometric ergodic process (z_t^*) is also stationary and geometric ergodic whenever the moments of $g(\mathrm{z}_t^*)$ exist. Thus, the existence of moments and the geometric ergodicity of (z_t^*, s_t) imply the geometric ergodicity of the tensor product $(\mathrm{z}_t^{*\otimes r}, s_t)$ for every $r > 1$. See also (Jones (2004), Corollary 2 and Theorem 9). This gives the asymptotic normality of the empirical spectral density. □

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10260-025-00826-9.

Acknowledgements We thank the Editor-in-Chief of the journal, Professor Brunero Liseo, and the anonymous referees for their constructive comments and very useful suggestions and remarks which were most valuable for the improvement of the final version of the paper.

Funding Open access funding provided by Università degli Studi di Modena e Reggio Emilia within the CRUI-CARE Agreement.

References

Aknouche A, Rabehi N (2010) On an independent and identically distributed mixture bilinear time series model. J Time Ser Anal 31(2):113–131
Alvarez R, Camacho M, Ruiz M (2019) Inference on filtered and smoothed probabilities in Markov-switching autoregressive models. J Bus Econ Stat 37(3):484–495
Bauwens L, Preminger A, Rombouts JVK (2010) Theory and inference for a Markov switching GARCH model. Econom J 13:218–244
Bibi A, Ghezal A (2015) On the Markov-switching bilinear processes: stationarity, higher-order moments and β-mixing Stochastics: Int J Prob Stoch Processes 87(6):919–945
Bibi A, Ghezal A (2016) On periodic time-varying bilinear processes: structure and asymptotic inference. Stat Methods Appl 25:395–420
Bibi A, Ghezal A (2018) Markov-switching bilinear GARCH models: structure and estimation. Commun Stat Theory Methods 47(2):307–323
Bougerol P, Picard N (1992) Strict stationarity of generalized autoregressive processes. Ann Probab 20:1714–1729
Brandt A (1986) The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. Adv Appl Prob 18:221–254
Brillinger DR, Rosenblatt M (1967) Asymptotic theory of estimates of kth order spectra. In: Harris B (ed) Spectral analysis of time series. Wiley, New York, pp 153–188
Brillinger DR, Rosenblatt M (1967) Computation and interpretation of kth order spectra. In: Harris B (ed) Spectral analysis of time series. Wiley, New York, pp 189–232
Brockwell PJ, Davis R (1991) Time series: theory and methods. Springer, New York
Cavicchioli M (2014a) Determining the number of regimes in Markov switching VAR and VMA models. J Time Ser Anal 35(2):173–186
Cavicchioli M (2014b) Analysis of the likelihood function for Markov switching VAR(CH) models. J Time Ser Anal 35(6):624–639
Cavicchioli M (2017a) Higher order moments of Markov switching VARMA models. Economet Theor 33(6):1502–1515
Cavicchioli M (2017b) Asymptotic Fisher information matrix of Markov switching VARMA models. J Multivar Anal 157:124–135
Cavicchioli M (2021) Fourth moment structure of Markov switching multivariate GARCH models. J Financ Economet 19(4):565–582
Cavicchioli M (2022) Markov switching GARCH models: higher order moments, kurtosis measures, and volatility evaluation in recessions and pandemic. J Bus Econ Stat 40(4):1772–1783
Chan KS, Tong H (2004) Testing for multimodality with dependent data. Biometrika 91(1):113–123
Doukhan P (1994) Mixing: properties and examples. Springer, Berlin-Heidelberg-New York
Francq C, Zakoïan JM (2001) Stationarity of multivariate Markov switching ARMA models. J Econ 102:339–364
Francq C, Zakoïan JM (2005) The L^2-structures of standard and switching-regime GARCH models. Stoch Processes Appl 115:1557–1582
Francq C, Zakoïan JM (2008) Deriving the autocovariances of powers of Markov-switching GARCH models, with applications to statistical inference. Comput Stat Data Anal 52:3027–3046
Ghezal A (2013) The bispectral representation of Markov switching BL models. Séminaire Mathématique de Béjaïa (La MOS) 12:107–113
Ghezal A (2024) Spectral representation of Markov-switching bilinear processes. Sao Paulo J Math Sci 18(1):459–479
Ghezal A, Zemmouri I (2023) The bispectral representation of Markov switching bilinear models. Commun Fac Sci Univ Ankara Ser A1 Math Stat 72(3):857–866
Hamilton JD (1989) A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57:357–384
Hamilton JD (1990) Analysis of time series subject to changes in regime. J Econo 45:39–70
Hamilton JD (1994) Time series analysis. Princeton University Press, Princeton
Hamilton JD (2016) Chapter 3- Macroeconomic regimes and regime shifts. In: Handbook of macroeconomics vol.2. Elsevier, Holland
Jones GL (2004) On the Markov chain central limit theorem. Probab Surv 1:299–320

Krolzig HM (1997) Markov-switching vector autoregressions: modelling, statistical inference, and application to business cycle analysis. In: Lecture Notes in Economics and Mathematical Systems, Springer, Verlag, Berlin-Heidelberg-New York

Maaziz M, Kharfouchi S (2018) Parameter estimation of Markov switching bilinear model using the (EM) algorithm. J Stat Plan Inference 192:35–44

Pataracchia B (2011) The Spectral Representation of Markov Switching ARMA Models. Econ Lett 112:11–15

Stelzer R (2009) On Markov-switching ARMA processes-stationarity, existence of moments and geometric ergodicity. Economet Theor 25(1):43–62

Subba Rao T (1981) On the theory of bilinear time series model. J R Stat Soc Ser B (Stat Methodol) 43(2):244–255

Subba Rao T (1983) The bispectral analysis of nonlinear stationary time series with reference to bilinear time series models. In: Brillinger DR, Krishnaiah PR (eds) Handbook of Statistics, vol 3. North-Holland, Amsterdam-New York-Oxford, pp 293–319

Subba Rao T, Gabr MM (1984) An introduction to bispectral analysis and bilinear time series models. In: Lecture Note in Statistics vol.24. Springer, New York-Berlin-Heidelberg-Tokyo

Zivot E (2013) A primer on asymptotics. Dept. of Economics, Univ. of Washington