



On the existence of stationary threshold bilinear processes

Ahmed Ghezal¹ · Maddalena Cavicchioli²  · Imane Zemmouri³

Received: 3 April 2023 / Revised: 16 September 2023

© The Author(s) 2024

Abstract

This article investigates some statistical and probabilistic properties of general threshold bilinear processes. Sufficient conditions for the existence of a causal, strictly and weak stationary solution for the equation defining a self-exciting threshold superdiagonal bilinear (*SETBL*) process are derived. Then it is shown that under well-specified hypotheses the higher-order moments of the *SETBL* process are finite. As a result, the skewness and kurtosis indexes are explicitly computed. The exact autocorrelation function is derived with an arbitrarily fixed number of regimes. Also, the covariance functions of the process and its powers are evaluated and the second (respectively, higher)-order structure is shown to be similar to that of a linear process. This implies that the considered process admits an *ARMA* representation. Finally, necessary and sufficient conditions for the invertibility and geometric ergodicity of a *SETBL* model are established. Some examples illustrate the obtained theoretical results.

Keywords Bilinear model · *SETBL* model · *ARMA* representation · Stationarity · Invertibility · Geometric ergodicity

Mathematics Subject Classification Primary 62F12 · Secondary 62M05

✉ Maddalena Cavicchioli
maddalena.cavicchioli@unimore.it

Ahmed Ghezal
ahmedghezal25@yahoo.fr

Imane Zemmouri
imanezemmouri25@gmail.com

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

² Department of Economics “Marco Biagi”, University of Modena and Reggio E., Modena, Italy

³ Department of Mathematics, University of Annaba, Elhadjar 23, Annaba, Algeria

1 Introduction

An \mathbb{R} -valued univariate process $(X_t)_{t \in \mathbb{Z}}$ defined on some probability space $(\Omega, \mathfrak{F}, P)$ is said to be a *self-exciting threshold superdiagonal bilinear process of type* $(\ell; p, q, P, Q)$ and *threshold delay parameter* d , in short denoted by $SETBL_d(\ell; p, q, P, Q)$, if it is a solution of the following stochastic difference equation

$$X_t = \sum_{j=1}^{\ell} \left\{ a_0^{(j)} + \sum_{i=1}^p a_i^{(j)} X_{t-i} + \sum_{i=1}^q b_i^{(j)} e_{t-i} + \sum_{k=1}^Q \sum_{m=k}^P c_{mk}^{(j)} X_{t-m} e_{t-k} \right\} \times I_{\{X_{t-d} \in R_j\}} + e_t, \quad (1.1)$$

where $d \geq 1$, $R_j =]r_{j-1}, r_j] \subset \mathbb{R}$, $j = 1, \dots, \ell$, r_j is the threshold value such that $-\infty = r_0 < r_1 < \dots < r_\ell = +\infty$, hence $\bigcup_{j=1}^{\ell} R_j = \mathbb{R}$, and $I_{\{X_{t-d} \in R_j\}}$ is the indicator function of the set $\{X_{t-d} \in R_j\}$. The coefficients $(a_i^{(j)})_{0 \leq i \leq p}$, $(b_i^{(j)})_{1 \leq i \leq q}$ and $(c_{mk}^{(j)})_{1 \leq m, k \leq \max(P, Q)}$, $j = 1, \dots, \ell$, are constant. The innovation process $(e_t)_{t \in \mathbb{Z}}$ is an independent and identically distributed (*i.i.d.*) white noise with zero mean and unit variance. Assume that the higher-order moments of the process $(e_t)_{t \in \mathbb{Z}}$ are finite.

Note that the $SETBL_d$ model is characterized by a piecewise nonlinear structure which follows a bilinear model in each regime R_j , for all $j = 1, \dots, \ell$. This offers remarkably rich dynamics and complex behaviour to model non-Gaussian data-sets with structural changes or high-amplitude oscillations, which cannot be sufficiently explained by the theory of standard linear models.

A study on threshold first-order bilinear models can be found in Cappuccio et al. (1998) and Ferrante et al. (2003). These authors deal with the problems of the existence of ergodicity, regularity and geometric ergodicity for stationary solutions of the threshold first-order bilinear model. These questions are very old since Tong (1983) for the class of threshold models in nonlinear time series analysis, and Brockwell et al. (1992) for ARMA models. Recent developments on the theory of threshold models can be found in Tong (2011, 2012, 2015). The results proved in our paper extend the work of the cited authors.

The $SETBL_d(\ell; p, q, P, Q)$ model encompasses many commonly used models existing in the literature:

1. *Standard superdiagonal bilinear BL(p, q, P, Q) models.* They are obtained by assuming $\ell = 1$ in (1.1). See, for example, Granger and Anderson (1978), Davis and Resnick (1996), and Popovič and Bakouch (2020).
2. *SETARMA_d(ℓ; p, q) models.* They are obtained by setting $c_{mk}^{(j)} = 0$ for all m , k and j in (1.1). See, for example, Tong (1983). Gibson and Nur (2011) investigate the relative efficacy of two-regime threshold autoregressive (TAR) models, applied to econometric dynamics, in the finance domain. A Bayesian analysis of such models has been developed by Chen and Lee (1995). A portmanteau test to detect self-exciting threshold autoregressive-type nonlinearity in time series data has been proposed by Petruccielli and Davies (1986).

The literature on SETARMA models has recently increased. Chan and Goracci (2019) solve the long-standing problem regarding the irreducibility condition of a first-order threshold autoregressive moving-average (TARMA) model, and derive a set of necessary and sufficient conditions for the ergodicity of invertible first-order TARMA processes. Li and Li (2011) derive the asymptotic null distribution of a quaslikelihood ratio test statistic for an ARMA model against its threshold extension, and propose a novel bootstrap approximation of that distribution based on stochastic permutation (the null hypothesis is that of no threshold, and the error term could be dependent). Multiple change point detection and validation in autoregressive time series data have been investigated by Ma et al. (2020), and Mohr and Selk (2020). Goracci et al. (2021) present supremum Lagrange multiplier tests to compare a linear ARMA specification against its threshold ARMA extension. Then they prove the consistency of the tests, and derive their asymptotic distribution both under the null hypothesis and contiguous local alternatives. Chan et al. (2024) propose a Lagrange multiplier test for dynamic regulation within TARMA setting, and provide a rigorous proof of tightness in the context of testing for threshold nonlinearity against difference stationarity.

3. *Some classes of SET(G)ARCH_d(ℓ; p, q) models.* See Kristensen (2009) for the description of (G)ARCH(p, q) models as special cases of standard bilinear models. A class of threshold bilinear GARCH processes has been introduced by Choi et al. (2012) to study different asymmetries in volatilities, accommodating some existing asymmetric models.

Remark 1.1 The piecewise bilinear structure of the SETBL_d model allows to apply some well-known results, obtained by Granger and Anderson (1978) in the bilinear setting.

Remark 1.2 The SETBL_d(2; p, q, P, Q) model can be written as

$$X_t = X_t^{(1)} I_{\{X_{t-d} \in R_1\}} + X_t^{(2)} I_{\{X_{t-d} \in R_2\}}, \tag{1.2}$$

where $X_t^{(j)}$ is a standard superdiagonal bilinear BL(p, q, P, Q) model, $j = 1, 2$, and the indicator function

$$I_{\{X_{t-d} \in R_1\}} = \begin{cases} 1 & \text{if } X_{t-d} \leq r_1 \\ 0 & \text{if } X_{t-d} > r_1 \end{cases}, \tag{1.3}$$

is shortly denoted by I_{t-d} . It follows that $I_{\{X_{t-d} \in R_2\}} = 1 - I_{t-d}$. Of course, the switching between the two regimes in model (1.2) is governed by the process in (1.3).

Remark 1.3 If the process in (1.3) is second-order stationary and ergodic, then we have

$$E \{I_{t-d}\} = P(X_{t-d} \leq r_1) = p, \quad \text{Var}(I_{t-d}) = p(1 - p) \quad \text{and} \\ \gamma_I(j) = \text{Cov}(I_{t-d}, I_{t-d-j}) = pj - p^2,$$

where $p_j = E \{I_{t-d}I_{t-d-j}\} = P(X_{t-d} \leq r_1, X_{t-d-j} \leq r_1)$, for all $j \geq 0$ with $p_0 = p$.

In this case, the process in (1.3) is called a *Bernoulli process*.

The aim of the paper is to investigate some statistical and probabilistic properties of $SETBL_d(\ell; p, q, P, Q)$ models. For statistical purposes, it is desirable in practice that the solution processes $(X_t)_{t \in \mathbb{Z}}$ of equation (1.1) are stationary, ergodic, and satisfy $X_t = f(e_t, e_{t-1}, \dots)$ almost surely (*a.s.*), where f is a measurable function from \mathbb{R}^∞ to \mathbb{R} . Such solutions are called *causal*.

We now highlight our main contributions. First, using suitable Markovian state space representations of the $SETBL_d(\ell; p, q, P, Q)$ model, we derive sufficient conditions ensuring the existence of a unique, causal, strictly stationary and ergodic solution of the proposed model. Then we prove that the higher-order moments of the SETBL process and its powers are finite, under certain tractable matrix conditions. As a result, the skewness and kurtosis indexes are explicitly computed. Second, since the second-order structure gives a useful information to identify a time series process, we derive the autocorrelation function and the covariance functions of the considered process. As a consequence, we find that the second-order structure is similar to that of some linear processes, hence the SETBL models admit ARMA representations. This finding is useful for the estimation of the model parameters via the GMM procedure. Furthermore, the ARMA representation plays an important role in forecasting the initial process. Third, we establish the necessary and sufficient conditions ensuring the existence of geometrically ergodic solutions of the SETBL models and for the model invertibility. Examples are proposed to illustrate the usefulness of the proposed methodology.

Some notations are used throughout the paper: $I_{(n)}$ denotes the $n \times n$ identity matrix, I_Δ the indicator function of the set Δ , and $O_{(k,\ell)}$ the matrix of order $k \times \ell$ whose entries are zeros. For simplicity, we set $O_{(k)} := O_{(k,k)}$ and $\underline{O}_{(k)} := O_{(k,1)}$. Notation plim means the convergence in probability. The spectral radius of a square matrix M is denoted by $\rho(M)$. If $M = (m_{ij})$ is a $m \times n$ matrix and $\underline{X} = (x_i)$ is a $m \times 1$ vector, define $|M| := (|m_{ij}|)$ and $|\underline{X}| := (|x_i|)$. Then it is easy to see that $|\cdot|$ is submultiplicative, i.e., $|M_1M_2| \leq |M_1||M_2|$ and $|M\underline{X}| \leq |M||\underline{X}|$ for any appropriate vector \underline{X} . It is also subadditive, i.e., $|\sum_i M_i| \leq \sum_i |M_i|$, where the inequality $M \leq N$ denotes the element-wise relation $m_{ij} \leq n_{ij}$ for all i and j . If $M_i = (m_{jk}^{(i)})$ is a sequence of $m \times n$ matrices, define $\max_i \{|M_i|\}$ the $m \times n$ matrix whose (j, k) element is $\max_i \{|m_{jk}^{(i)}|\}$, i.e., $\max_i \{|M_i|\} = \left(\max_i \left\{ |m_{jk}^{(i)}| \right\} \right)$. The symbol \otimes denotes the usual Kronecker product of matrices. Set $M^{\otimes r} = M \otimes M \otimes \dots \otimes M$, r times, with $M^{\otimes r} = M$ if $r = 1$.

The remainder of the paper is organized as follows. The next section provides state-space representations of the $SETBL_d$ process generated by (1.1). Such representations will be used in Sect. 3 to derive sufficient conditions for the $SETBL_d$ model in (1.1) to have a unique stationary (in strong and weak sense), causal and ergodic solution. Section 4 is devoted to establish sufficient conditions for the finiteness of higher-order moments of the considered process. Then the autocorrelation function of a ℓ -th regime $SETBL_d$ model is derived analytically. In Sect. 5 the \mathbb{L}_2 -structure is analyzed and the covariance function is derived, which allows us to give an ARMA rep-

resentation. Extending the \mathbb{L}_2 -structure to \mathbb{L}_m yields that the power process $(X_t^m)_{t \in \mathbb{Z}}$ also admits an *ARMA* representation (Sect. 6). Section 7 is devoted to establish necessary and sufficient conditions for the invertibility and geometric ergodicity of the *SETBL_d* model. Section 8 concludes.

2 Markovian representations of the *SETBL_d* model

In what follows, we shall assume, without loss of generality, that in (1.1) $P = p$ and $d \leq r$, with $r = p + q$, since otherwise zeros of $a_i^{(j)}$ or $b_i^{(j)}$ or $c_{mk}^{(j)}$ can be filled in.

Define the r -dimensional state vector

$$\underline{X}_t := (X_t, \dots, X_{t-p+1}, e_t, \dots, e_{t-q+1})'$$

and the $r \times r$ block matrices $(A_i^{(j)}, 0 \leq i \leq Q)$, where

$$A_0^{(j)} = \begin{pmatrix} M_0^{(j)} & B^{(j)} \\ O_{(q,p)} & J \end{pmatrix} \quad A_i^{(j)} = \begin{pmatrix} M_i^{(j)} & O_{(p,q)} \\ O_{(q,p)} & O_{(q)} \end{pmatrix}$$

$$M_0^{(j)} = \begin{pmatrix} a_1^{(j)} & \dots & a_{p-1}^{(j)} & a_p^{(j)} \\ 1 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 1 & 0 \end{pmatrix} \in \mathbb{R}^{p \times p}$$

$$M_i^{(j)} = \begin{pmatrix} c_{ii}^{(j)} & \dots & c_{pi}^{(j)} & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} \in \mathbb{R}^{p \times p}$$

and

$$B^{(j)} = \begin{pmatrix} b_1^{(j)} & \dots & b_{p-1}^{(j)} & b_q^{(j)} \\ 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{pmatrix} \in \mathbb{R}^{p \times q} \quad J = \begin{pmatrix} 0 & \dots & 0 & 0 \\ 1 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 1 & 0 \end{pmatrix} \in \mathbb{R}^{q \times q}$$

for all $i = 1, \dots, Q$ and $j = 1, \dots, \ell$.

Set $\underline{H}_r := (1, \underline{O}'_{(r-1)})' \in \mathbb{R}^r$ and $\underline{F}_r := (\underline{H}'_p, \underline{H}'_q)' \in \mathbb{R}^r$. Then Eq. (1.1) can be expressed in the following state-space representation

$$X_t = \underline{H}'_r \underline{X}_t$$

$$\underline{X}_t = \sum_{j=1}^{\ell} \left(a_0^{(j)} \underline{H}_r + A_0^{(j)} \underline{X}_{t-1} + \sum_{k=1}^Q A_k^{(j)} \underline{X}_{t-k} e_{t-k} \right) I_{\{X_{t-1} \in R_j^{(d)}\}} + \underline{F}_r e_t \tag{2.1}$$

for all $t \in \mathbb{Z}$, where $R_j^{(d)}$ is the Cartesian product $\mathbb{R}^{d-1} \times R_j \times \mathbb{R}^{r-d}$, for all $j = 1, \dots, \ell$.

Now define

$$\underline{Y}_t := (\underline{X}'_t, \underline{X}'_t e_t, \dots, \underline{X}'_{t-Q+1} e_{t-Q+1})' \in \mathbb{R}^K$$

and

$$\underline{e}_t := (\underline{F}'_r e_t, \underline{F}'_r e_t^2, \underline{O}'_{(K-2r)})' \in \mathbb{R}^K,$$

where $K = r(Q + 1)$. Then (2.1) can be transformed into the following Markovian state-space representation

$$\begin{aligned} X_t &= \underline{H}'_K \underline{Y}_t \\ \underline{Y}_t &= \sum_{j=1}^{\ell} \left(\Gamma^{(j)}(e_t) \underline{Y}_{t-1} + \underline{\eta}^{(j)}(e_t) \right) I_{\{\underline{Y}_{t-1} \in R_j^{(d)}\}} + \underline{e}_t, \end{aligned} \tag{2.2}$$

where

$$\begin{aligned} \Gamma^{(j)}(e_t) &= \Gamma_0^{(j)} + e_t \Gamma_1^{(j)} \in \mathbb{R}^{K \times K} \\ \underline{\eta}^{(j)}(e_t) &= a_0^{(j)} (\underline{H}'_r, e_t \underline{H}'_r, \underline{O}'_{(K-2r)})' \in \mathbb{R}^K \end{aligned}$$

and

$$\begin{aligned} \Gamma_0^{(j)} &:= \begin{pmatrix} A_0^{(j)} & \dots & \dots & \dots & A_Q^{(j)} \\ O_{(r)} & \dots & \dots & \dots & O_{(r)} \\ O_{(r)} & I_{(r)} & \dots & \dots & O_{(r)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ O_{(r)} & \dots & O_{(r)} & I_{(r)} & O_{(r)} \end{pmatrix} \in \mathbb{R}^{K \times K} \\ \Gamma_1^{(j)} &:= \begin{pmatrix} O_{(r)} & \dots & \dots & O_{(r)} \\ A_0^{(j)} & \dots & \dots & A_Q^{(j)} \\ O_{(r)} & \dots & \dots & O_{(r)} \\ \vdots & \vdots & \vdots & \vdots \\ O_{(r)} & \dots & \dots & O_{(r)} \end{pmatrix} \in \mathbb{R}^{K \times K}, \end{aligned}$$

for all $j = 1, \dots, \ell$. It is evident that any solution (\underline{Y}_t) of (2.2) is a Markov chain with state space \mathbb{R}^K . Recall that a Markov chain is said to be *irreducible* if each state can be reached from any other state in a finite number of steps. For each Markov chain there exists a unique decomposition of the state space into a sequence of disjoint subsets in which the restriction of the chain results irreducible. So we can always assume that the considered Markov chain is irreducible (possibly, restricting the state space).

Irreducibility is a crucial point as it allows to deploy the theory developed by Tweedie (see Tweedie (1974a, b, 1975, 1976)).

Representation (2.2) can be rewritten as follows

$$\begin{aligned} X_t &= \underline{H}'_K \underline{Y}_t \\ \underline{Y}_t &= \Phi_t \underline{Y}_{t-1} + \underline{\omega}_t \end{aligned} \tag{2.3}$$

where

$$\begin{aligned} \Phi_t &= \sum_{j=1}^{\ell} \Gamma^{(j)}(e_t) I_{\{\underline{Y}_{t-1} \in R_j^{(d)}\}} \in \mathbb{R}^{K \times K} \\ \underline{\omega}_t &= \sum_{j=1}^{\ell} \eta_t^{(j)}(e_t) I_{\{\underline{Y}_{t-1} \in R_j^{(d)}\}} + \underline{e}_t \in \mathbb{R}^K. \end{aligned}$$

Compare with Eq. (2.1) from Bibi and Ghezal (2015b) for the class of Markov switching bilinear processes. A different Markovian representation for bilinear time series models has been proposed by Pham (see Pham (1985, 1986)). Such a representation is intriguing and constitutes an open issue, which will be allocated into a separate article with further discussions and several statistical properties.

Since $(e_t)_{t \in \mathbb{Z}}$ is stationary and ergodic, the process $\{(\Phi_t, \underline{\omega}_t)\}_{t \in \mathbb{Z}}$ is also stationary and ergodic. Let $\|\cdot\|$ denote any operator norm on the sets of matrices and vectors. It is clear that $E \{\log^+ \|\Phi_t\|\}$ and $E \{\log^+ \|\underline{\omega}_t\|\}$ are finite. Here $\log^+ x = \max(\log x, 0)$ for every positive real number x . By Brandt (1986) (see also Bougerol and Picard (1992)), the unique stationary solution of (2.3) is given by

$$\underline{Y}_t = \underline{\omega}_t + \sum_{k=1}^{\infty} \Phi_t \Phi_{t-1} \cdots \Phi_{t-k+1} \underline{\omega}_{t-k} \tag{2.4}$$

whenever the top Lyapunov exponent $\gamma_L(\Phi)$, associated to the sequence $(\Phi_t)_t$ and defined by

$$\gamma_L(\Phi) = \inf_{t > 1} \left\{ E \left[\frac{1}{t} \log \|\Phi_t \Phi_{t-1} \cdots \Phi_1\| \right] \right\},$$

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

In the general case the Lyapunov exponent seems difficult to compute but it is easily estimated by Monte Carlo simulation using (2.3) as

$$\gamma_L(\Phi) = \lim_{t \rightarrow \infty} \frac{1}{t} \log \left\| \prod_{i=0}^{t-1} \Phi_{t-i} \right\| \quad \text{a.s.}$$

Then we have

Proposition 2.1 Consider the threshold bilinear model $SETBL$ in (1.1) with Markovian representation (2.2) or (2.3), and suppose that $\gamma_L(\Phi) < 0$. Then the series (2.4) converges a.s., for all $t \in \mathbb{Z}$, and the process (X_t) , defined as the first component of \underline{Y}_t , is the unique causal strictly stationary and ergodic solution of (1.1).

The proof of this result follows essentially the same arguments as in Francq and Zakoïan (2001) for Markov switching (MS) ARMA models, and Bibi and Ghezal (2015b) for MS bilinear models.

3 Existence of a causal strictly stationary $SETBL_d$ process

In this section we provide a further sufficient condition for the existence of a causal strictly stationary and ergodic solution for the $SETBL_d$ process. Such a condition is based on the computation of the spectral radius of a well-specified matrix, hence it is easily tractable and programmable.

Let us consider the following assumptions:

[A.0] The probability distribution of (e_t) is absolutely continuous with respect to the Lebesgue measure λ .

[A.1] The conditional mean function $e(\cdot) : \mathbb{R}^K \rightarrow \mathbb{R}^K$, defined by

$$e(\underline{y}) = E \left\{ \underline{Y}_t \mid \underline{Y}_{t-1} = \underline{y} \right\} = \sum_{j=1}^{\ell} \Gamma_0^{(j)} \underline{y} I_{\{\underline{y} \in R_j^{(d)}\}}, \tag{3.1}$$

for all $\underline{y} \in \mathbb{R}^K$, is continuous.

Assumption [A.0] guarantees λ -irreducibility. Assumption [A.1] has a counterpart in the analysis of segmented regression models, as documented by Feder (1975). Stability of the considered process will be deduced via appropriate contraction conditions involving Lipschitz continuity.

We get the following lemma

Lemma 3.1 The conditional mean function $e(\cdot) : \mathbb{R}^K \rightarrow \mathbb{R}^K$ satisfies the Lipschitz condition

$$|e(\underline{y}) - e(\underline{z})| \leq \Lambda_0 |\underline{y} - \underline{z}|$$

where $\Lambda_0 = \max_j \left\{ |\Gamma_0^{(j)}| \right\} \in \mathbb{R}^{K \times K}$.

Proof Here we use some techniques from Brockwell et al. (1992). Assume that $\underline{y} \in R_j^{(d)}$ and $\underline{z} \in R_k^{(d)}$ with $j \leq k$.

Then there exists a sequence of $K \times 1$ vectors $\underline{y}_{h-1} := \underline{y}, \underline{y}_h, \dots, \underline{y}_{k+1} := \underline{z}$ on the boundaries of $R_{h-1}^{(d)}, R_h^{(d)}, \dots, R_{k+1}^{(d)}$ such that $|\underline{y} - \underline{z}| = |\underline{y}_{h-1} - \underline{y}_h| +$

$|\underline{y}_h - \underline{y}_{h+1}| + \dots + |\underline{y}_k - \underline{y}_{k+1}|$. So we have

$$\begin{aligned} |e(\underline{y}) - e(\underline{z})| &\leq \sum_{i=h}^{k+1} |e(\underline{y}_{i-1}) - e(\underline{y}_i)| \leq \sum_{i=h}^{k+1} \sum_{j=1}^{\ell} |\Gamma_0^{(j)} \underline{y}_{i-1} - \Gamma_0^{(j)} \underline{y}_i| \\ &\leq \sum_{i=h}^{k+1} \sum_{j=1}^{\ell} |\Gamma_0^{(j)}| |\underline{y}_{i-1} - \underline{y}_i| \\ &\leq \max_j \{|\Gamma_0^{(j)}|\} \sum_{i=h}^{k+1} |\underline{y}_{i-1} - \underline{y}_i| = \Lambda_0 |\underline{y} - \underline{z}|. \end{aligned}$$

This completes the proof. □

The next theorem provides sufficient conditions for the existence of strictly stationary, causal and ergodic solutions of model (1.1). Recall that the higher-order moments of the innovations are assumed to be finite. The matrix condition in (3.2) below is tractable and readily programmable, hence it is very simple to check. Assumptions [A.0], [A.1] and Eq. (3.2) below are common to investigate stationarity in the context of time series analysis. For example, conditions similar to (3.2) have been used by Francq and Zakoian (2001, 2005) to study stationarity of Markov switching ARMA and GARCH models, respectively.

Theorem 3.1 *Under Assumptions [A.0] and [A.1], suppose that*

$$\rho \left(E \left\{ \max_j \left\{ |\Gamma^{(j)}(e_t)| \right\} \right\} \right) < 1. \tag{3.2}$$

Then the SETBL_d model in (1.1) with Markovian representation (2.2) or (2.3) has a unique causal strictly stationary and ergodic solution $(\underline{Y}_t)_{t \in \mathbb{Z}}$ which belongs to \mathbb{L}_1 and it is the limit (a.s.) of the sequence defined recursively by

$$\underline{Y}_n(t) := \begin{cases} e_t & \text{if } n = 0 \\ \sum_{j=1}^{\ell} \left(\Gamma^{(j)}(e_t) \underline{Y}_{n-1}(t-1) + \eta_t^{(j)}(e_t) \right) I_{\{\underline{Y}_{n-1}(t-1) \in R_j^{(d)}\}} + e_t & \text{if } n > 0 \end{cases} \tag{3.3}$$

Proof For any $n > 2$, we have

$$\begin{aligned} |\underline{Y}_n(t) - \underline{Y}_{n-1}(t)| &\leq \max_j \left\{ |\Gamma^{(j)}(e_t)| \right\} |\underline{Y}_{n-1}(t-1) - \underline{Y}_{n-2}(t-1)| \\ &\quad \vdots \\ &\leq \left(\max_j \left\{ |\Gamma^{(j)}(e_t)| \right\} \right)^{n-2} |\underline{Y}_2(t-n+2) - \underline{Y}_1(t-n+2)|. \end{aligned}$$

As $E\{|e_t|\} < \infty$, $(\underline{Y}_n(t))$ is a Cauchy sequence in \mathbb{L}_1 for each fixed t . Then there exists an \mathbb{L}_1 limit, \underline{Y}_t say, which is also the a.s. limit of $\underline{Y}_n(t)$ as $n \rightarrow \infty$ because $E\{|\underline{Y}_n(t) - \underline{Y}_{n-1}(t)|\}$ is geometrically bounded. Moreover, we show that the process (\underline{Y}_t) is strictly stationary. Notice that the process $(\underline{Y}_n(t))$ is strictly stationary and ergodic, for any fixed n . Then, for any bounded continuous function g , we have

$$\begin{aligned} E\{g(\underline{Y}_t, \dots, \underline{Y}_{t+h})\} &= \lim_n E\{g(\underline{Y}_n(t), \dots, \underline{Y}_n(t+h))\} \\ &= \lim_n E\{g(\underline{Y}_n(t+s), \dots, \underline{Y}_n(t+h+s))\} \\ &= E\{g(\underline{Y}_{t+s}, \dots, \underline{Y}_{t+h+s})\}. \end{aligned}$$

Now we check that $\text{plim } I_{\{\underline{Y}_{n_k}(t) \in R_j^{(d)}\}} = I_{\{\underline{Y}_t \in R_j^{(d)}\}}$ for all $j \in \{1, \dots, \ell\}$ and some subsequence (n_k) of (n) . Note that, for any $\delta > 0$ and a.s. convergent subsequence $(\underline{Y}_{n_k}(t))$ of $(\underline{Y}_n(t)) := \left((Y_n^{(1)}(t), \dots, Y_n^{(K)}(t)) \right)'$, we have

$$\begin{aligned} &\limsup_{n_k} P\left(\left|I_{R_j^{(d)}}(\underline{Y}_{n_k}(t)) - I_{R_j^{(d)}}(\underline{Y}_t)\right| > \delta\right) \\ &\leq \limsup_{n_k} P\left(\left\{\underline{Y}_{n_k}(t) \in R_j^{(d)}, \underline{Y}_t \notin R_j^{(d)}\right\} \text{ or } \left\{\underline{Y}_{n_k}(t) \notin R_j^{(d)}, \underline{Y}_t \in R_j^{(d)}\right\}\right) \\ &\leq \limsup_{n_k} P\left(r_{j-1} \leq Y_{n_k}^{(d)}(t) \leq r_j, Y_t^{(d)} \notin]r_{j-1}, r_j]\right) \\ &\quad + \limsup_{n_k} P\left(Y_{n_k}^{(d)}(t) \leq r_{j-1}, X_{t-d} \in]r_{j-1}, r_j]\right) \\ &\quad + \limsup_{n_k} P\left(Y_{n_k}^{(d)}(t) \geq r_j, X_{t-d} \in]r_{j-1}, r_j]\right) \\ &\leq 2P(X_{t-d} = r_{j-1}) + P(X_{t-d} = r_j). \end{aligned}$$

It now suffices to note that $P(X_t = r_j) = 0$, for all $j \in \{1, \dots, \ell\}$. The rest is immediate. □

Remark 3.1 The process (X_t) , where $X_t = \underline{H}'_K \underline{Y}_t$, is also the unique causal strictly stationary and ergodic solution of (1.1).

Remark 3.2 Ling et al. (2007) investigate the first-order threshold moving-average model, denoted by TMA(1) and defined as

$$y_t = \phi + \psi I(y_{t-1} \leq r) \epsilon_{t-1} + \epsilon_t \quad (t \in \mathbb{Z}),$$

where ϕ, ψ and r are constant and ϵ_t is a sequence of independent and identically distributed random variables, with mean zero and a density function $f(x)$, for all $x \in X \subset \mathbb{R}$. These authors provide a sufficient condition for the ergodicity of the TMA(1) model, without the need to check irreducibility. Their main result (see Theorem 1 of Ling et al. (2007)) states that if $|\psi| \sup_{x \in X} |x f(x)| < 1$, then there exists a unique,

strictly stationary and ergodic solution of the TMA(1) model. It is interesting to see that the ergodicity of the TMA(1) model depends on the coefficient ψ , while the intercept ϕ is irrelevant. Since our SETBL model encompasses many nonlinear models, including the TMA models, Theorem 3.1 implies the ergodicity of TMA(1) by using a similar but different condition with respect to that proposed in the cited paper.

4 Higher-order moments of the SETBL_d model

In this section we are interested in conditions ensuring the existence of higher-order moments for a strictly stationary SETBL_d process $X = (X_t)_{t \in \mathbb{Z}}$. In particular, the exact form of the r -th moment of X is derived. The computation of higher-order moments for some classes of Markov switching (MS) bilinear models has been given in Bibi and Ghezal (2015b) and Bibi (2023). Matrix expressions in closed form for higher-order moments and asymptotic Fisher information matrix of general MS VARMA models have been provided in Cavicchioli (2017a, b), respectively.

Let $\mu_r(j)$ be the moment of order r for the j -th bilinear regime process, i.e., $\mu_r(j) = E \left\{ \left(X_t^{(j)} \right)^r \right\}$, $\tilde{\mu}_r(j)$ the r -th centered moment, $v_r(j) = E \left\{ \left(I_{t-d}^{(j)} - p_j \right)^r \right\}$, and $\gamma_{ij}(k)$ the autocovariance of X_t at lag k for the $i, j - th$ bilinear regime processes, i.e., $\gamma_{ij}(k) = Cov \left(X_t^{(j)}, X_{t-k}^{(i)} \right)$, for all $k > 0$. For convenience, model (1.1) can also be written as follows

$$X_t = \sum_{j=1}^{\ell} X_t^{(j)} I_{t-d}^{(j)}, \tag{4.1}$$

where $X_t^{(j)} \sim SBL(p, q, P, Q)$ is the standard superdiagonal bilinear process in regime j , for all $j = 1, \dots, \ell$, and

$$I_{t-d}^{(j)} := \begin{cases} 1 & \text{if } X_{t-d} \in R_j \\ 0 & \text{if } X_{t-d} \notin R_j \end{cases}.$$

We get the following lemma

Lemma 4.1 *According to representation (4.1), the processes $\left(I_{t-d}^{(j)} \right)_t$, for $j = 1, \dots, \ell$, satisfy the following properties:*

- (i) $E \left\{ I_{t-d}^{(j)} \right\} = P(X_{t-d} \in R_j) = p_j$.
- (ii) $\left(I_{t-d}^{(j)} \right)^m = I_{t-d}^{(j)}$ for all $j = 1, \dots, \ell$ and $m \geq 1$.
- (iii) $\left(I_{t-d}^{(j)} \right)^m \left(I_{t-d}^{(k)} \right)^n = 0$ for all $j \neq k \in \{1, \dots, \ell\}$ and $m, n \geq 1$.
- (iv) $I_{t-d}^{(j)} I_{t-d-k}^{(j)} = \begin{cases} 1 & \text{if } X_{t-d} \text{ and } X_{t-d-k} \in R_j \\ 0 & \text{otherwise} \end{cases}$
with $E \left\{ I_{t-d}^{(j)} I_{t-d-k}^{(i)} \right\} = P(X_{t-d} \in R_j, X_{t-d-k} \in R_i) = p_{ij}$.

The proof is straightforward, hence omitted.

Proposition 4.1 *Let us consider model (4.1). Suppose that the local bilinear process in each regime admits finite moments up to order r . Then the expected value of the process (X_t^r) is given by*

$$\mu_r = E \{X_t^r\} = \sum_{j=1}^{\ell} \mu_r(j) p_j \tag{4.2}$$

for all $r \geq 0$.

Proof Using Lemma 4.1, we have

$$X_t^r = \left(\sum_{j=1}^{\ell} X_t^{(j)} I_{t-d}^{(j)} \right)^r = \sum_{\substack{0 \leq x_j \leq x_{j-1} \\ 1 \leq j \leq \ell-1}} \left\{ \prod_{j=1}^{\ell-1} C_{x_{j-1}}^{x_j} \right\} \left\{ \prod_{j=1}^{\ell} \left(X_t^{(j)} I_{t-d}^{(j)} \right)^{x_{\ell-j} - x_{\ell-j+1}} \right\}.$$

By convention, set $x_0 = r$ and $x_{\ell} = 0$. Then the computation of the expected value of X_t^r gives formula (4.2). \square

Corollary 4.1 *Let us consider model (4.1). Under the assumptions of Proposition 4.1, we have*

$$\begin{aligned} Var(X_t) &= E \left\{ (X_t - \mu_1)^2 \right\} = \sum_{j=1}^{\ell} \left(\tilde{\mu}_2(j) p_j - \mu_1^2(j) v_2(j) \right) \\ &\quad - 2 \sum_{1 \leq i < j \leq \ell} \mu_1(j) \mu_1(i) p_j p_i. \end{aligned} \tag{4.3}$$

Proof The variance of (X_t) can be expressed by the following expansion

$$Var(X_t) = \sum_{j=1}^{\ell} Var \left(X_t^{(j)} I_{t-d}^{(j)} \right) + 2 \sum_{1 \leq i < j \leq \ell} Cov \left(X_t^{(j)} I_{t-d}^{(j)}, X_t^{(i)} I_{t-d}^{(i)} \right).$$

Now we can write $Var \left(X_t^{(j)} I_{t-d}^{(j)} \right) = \tilde{\mu}_2(j) p_j - \mu_1^2(j) v_2(j)$, for all $j = 1, \dots, \ell$, and

$$Cov \left(X_t^{(j)} I_{t-d}^{(j)}, X_t^{(i)} I_{t-d}^{(i)} \right) = -\mu_1(j) \mu_1(i) p_j p_i,$$

for $1 \leq i < j \leq \ell$. Thus formula (4.3) for the variance of X_t follows. \square

Corollary 4.2 *Let us consider model (4.1). Assume that $E \{e_t^4\} < \infty$. If $E \{X_t^4\} < \infty$, then the third and fourth central unconditional moments of X_t are given by*

$$E \left\{ (X_t - \mu_1)^3 \right\} = \sum_{j=1}^{\ell} J_1(j) - 6 \sum_{1 \leq i < j \leq \ell} J_2(i, j) + 12 \sum_{1 \leq i < j < k \leq \ell} J_3(i, j, k) \tag{4.4}$$

and

$$E \left\{ (X_t - \mu_1)^4 \right\} = \sum_{j=1}^{\ell} K_1(j) + 8 \sum_{1 \leq i < j \leq \ell} K_2(i, j) + 24 \sum_{1 \leq i < j < k \leq \ell} K_3(i, j, k) - 72 \sum_{1 \leq i < j < k < m \leq \ell} K_4(i, j, k, m), \tag{4.5}$$

where

$$\begin{aligned} J_0(j) &:= \tilde{\mu}_2(j) + \mu_1^2(j) (1 - 2p_j) \\ J_1(j) &:= \tilde{\mu}_3(j) p_j + \mu_1(j) v_2(j) (2\tilde{\mu}_2(j) + J_0(j)) \\ J_2(i, j) &:= \mu_1(j) p_j p_i J_0(i) \\ J_3(i, j, k) &:= \mu_1(i) p_i \mu_1(j) p_j \mu_1(k) p_k \\ K_0(j) &:= \mu_1^4(j) (3v_4(j) - 2v_2(j) p_j^2) \\ K_1(j) &:= \tilde{\mu}_4(j) p_j + K_0(j) + 2(1 - p_j) \mu_1(j) (2J_1(j) + \mu_1(j) v_2(j) J_0(j)) \\ K_2(i, j) &:= -\mu_1(j) p_j (J_1(i) + \mu_1^3(i) p_i^3) \\ K_3(i, j, k) &:= ((1 - 2p_i) (1 - \mu_1(i)) - p_i) J_3(i, j, k) + \mu_1(k) p_k J_2(i, j) \\ K_4(i, j, k, m) &:= J_3(i, j, k) \mu_1(m) p_m. \end{aligned}$$

Proof The proof is similar to that of Corollary 4.1. □

Now we derive expressions in closed-form to compute the kurtosis and the skewness of the stationary threshold bilinear process X_t .

Corollary 4.3 *Let us consider model (4.1). Under the assumptions of Corollary 4.2, the kurtosis and the skewness of X_t exist and are given by substituting the formulas in Corollaries 4.1 and 4.2 into the usual expressions*

$$\kappa = \frac{E \{ (X_t - \mu_1)^4 \}}{(E \{ (X_t - \mu_1)^2 \})^2}, \quad \tau = \frac{E \{ (X_t - \mu_1)^3 \}}{(E \{ (X_t - \mu_1)^2 \})^{\frac{3}{2}}}.$$

Proof Equations (4.3), (4.4), (4.5) linked to the unconditional variance, third and fourth central moments, respectively, are used to compute the skewness and the excess kurtosis. □

To complete the section we present some generic examples of threshold bilinear processes to illustrate the obtained theoretical results. In particular, some examples

show that the sufficient condition (3.2) for strict stationarity is not necessary in general. However, the pictures of the sample autocorrelation functions reveal that the considered processes are strictly stationary. In addition, we check the validity of the obtained expressions in closed-form for the higher-order moments and kurtosis. The proposed examples show that the values of the theoretical moments and kurtosis fit well with those of the empirical ones. We also compare sample and theoretical autocorrelation functions as goodness-of-fit criterion.

Example 4.1 Consider the $SETBL_d(2; 0, 0, 2, 1)$ model defined by

$$X_t = \begin{cases} c X_{t-2} e_{t-1} + e_t & \text{if } X_{t-d} \geq r_1 \\ a X_{t-2} e_{t-1} + e_t & \text{if } X_{t-d} < r_1 \end{cases} \text{ with } e_t \sim \mathcal{N}(0; 1).$$

Then the regime moments collected in the following table

	$j = 1$	$j = 2$
$\mu_1(j)$	0	0
$\mu_2(j)$	$\frac{1}{1-c^2}$	$\frac{1}{1-a^2}$
$\mu_3(j)$	0	0
$\mu_4(j)$	$3 \frac{1+c^2}{(1-3c^4)(1-c^2)}$	$3 \frac{1+a^2}{(1-3a^4)(1-a^2)}$

can be used to compute the skewness and the kurtosis of (X_t) . More precisely, the skewness of X_t is $\tau = 0$ and the kurtosis of X_t is given by

$$\kappa = 3 \frac{(1 - 3a^4) (1 - c^4) (1 - a^2)^2 p_1 + (1 - 3c^4) (1 - a^4) (1 - c^2)^2 p_2}{(1 - 3c^4) (1 - 3a^4) (a^2 p_1 + c^2 p_2 - 1)^2}.$$

The exact autocorrelation function (ACF) of the process (X_t) is derived. The ACF coefficient $\rho(k)$, defined by $\rho(k) = \frac{Cov(X_t, X_{t-k})}{Var(X_t)}$ with $\rho(k) = \rho(-k)$ accordingly to the symmetry of the autocovariance function, has the following expression:

$$\rho(k) = const + \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} \frac{\gamma_{ij}(k)}{Var(X_t)} p_{ij},$$

where $Var(X_t)$ is given in Eq. (4.3).

Set $a = 0.7, c = 0.5, d = 1, r_1 = 4,$ and $p_1 = 0.7$. Then a typical realization of such a process is depicted in Fig. 1 (on the left side). The second and fourth moments of this process are given by $\mu_2 = E(X_t^2) = 1.5216$ and $\mu_4 = E(X_t^4) = 13.7085,$ respectively, and the kurtosis is $\kappa = 5.9212$. The sample autocorrelation function is reported in Fig. 1 (on the right side). The matrix in the statement of Theorem 3.1 has spectral radius equal to 1. However, the sample ACF reveals that the given time series is strictly stationary. Thus the condition in Theorem 4.1 is sufficient but not necessary to guarantee strict stationarity.

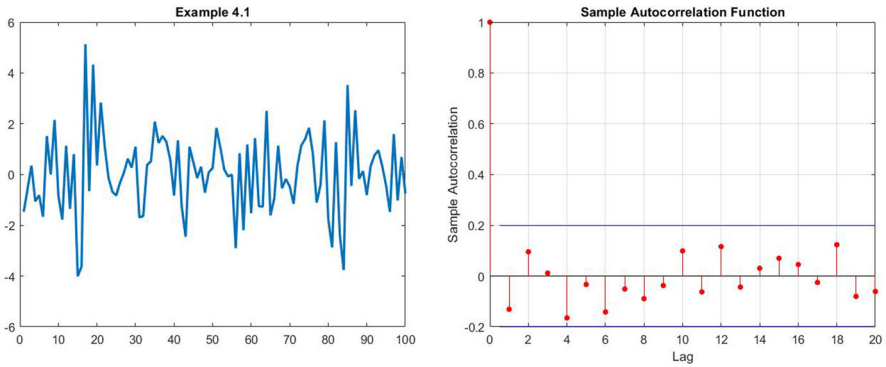


Fig. 1 Example 4.1: original data (left) and sample autocorrelation (ACF) (right)

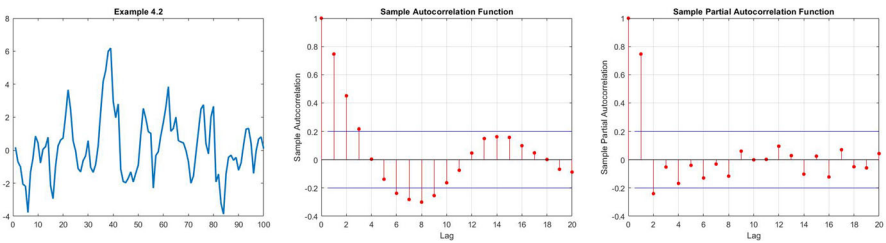


Fig. 2 Example 4.2: original data (left), sample ACF (middle), sample partial ACF (right)

Example 4.2 Consider the $SETBL_d(2; 1, 0, 1, 1)$ model defined by

$$X_t = \begin{cases} a X_{t-1} + b X_{t-1} e_{t-1} + e_t & \text{if } X_{t-d} \geq r_1 \\ c X_{t-1} + f X_{t-1} e_{t-1} + e_t & \text{if } X_{t-d} < r_1 \end{cases} \text{ with } e_t \sim \mathcal{N}(0; 1).$$

Set $a = 0.7, b = 0.5, c = 0.8, f = 0.3, d = 1, r_1 = 4$, and $p_1 = 0.7$. We report in Fig. 2 a typical realization of such a process (left), the sample ACF (middle) and the sample partial ACF (right). The sample second and fourth moments of this process are given by $\mu_2 = 0.2251$ and $\mu_4 = 3.6210$, respectively, the skewness is $\tau = 0.6404$ and the kurtosis is $\kappa = 3.7706$.

Example 4.3 Consider the $SETARMA_d(2; 1, 1)$ model defined by

$$X_t = \begin{cases} a X_{t-1} + e_t & \text{if } X_{t-d} \geq r_1 \\ b X_{t-1} + e_{t-1} + e_t & \text{if } X_{t-d} < r_1 \end{cases} \text{ with } e_t \sim \mathcal{N}(0; 1).$$

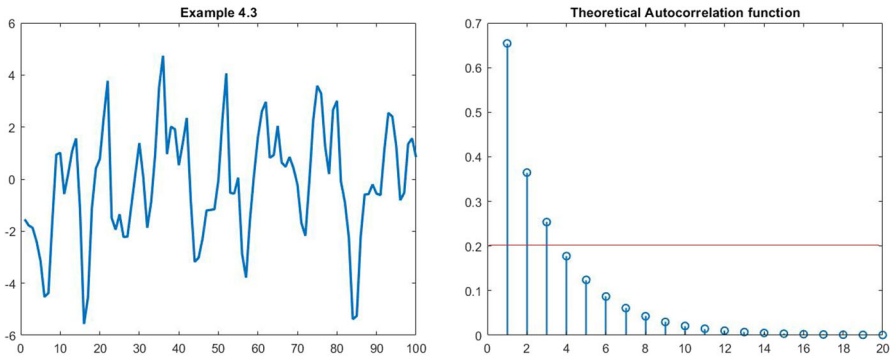


Fig. 3 Example 4.3: original data (left) and theoretical autocorrelation function (ACF) (right)

Then the regime moments in the following table

	$j = 1$	$j = 2$
$\mu_1(j)$	0	0
$\mu_2(j)$	$\frac{1}{1-a^2}$	$\frac{2}{1-b}$
$\mu_3(j)$	0	0
$\mu_4(j)$	$\frac{3}{(1-a^2)^2}$	$4 \frac{(b-2)(1-b)^3+(b+5)}{(1-b)(1-b^4)}$

can be used to compute the skewness and the kurtosis of (X_t) . More precisely, the skewness of X_t is $\tau = 0$ and the kurtosis of X_t is given by

$$\kappa = \frac{3(1-b)(1-b^4)p_1 + 4(1-a^2)^2((b-2)(1-b)^3 + (b+5))p_2}{(1+b)(1+b^2)((1-b)p_1 + 2(1-a^2)p_2)^2}$$

Furthermore, the autocorrelation of X_t is given by

$$\rho(k) = \frac{a^k(1-b)p_1 + 2b^k(1-a^2)p_2 + (1-b)(1-a^2)p_2I_{\{k=1\}}}{(1-b)p_1 + 2(1-a^2)p_2}$$

Set $a = 0.1, b = 0.7, d = 1, r_1 = 4,$ and $p_1 = 0.7$. Then a typical realization of such a process is depicted in Fig. 3 (left) together with the theoretical autocorrelation function (right). The ACF declines to near zero rapidly at lag 20 with exponential decay. This means that the given time series is strictly stationary. The second and fourth moments of this process are given by $\mu_2 = E(X_t^2) = 2.7071$ and $\mu_4 = E(X_t^4) = 9.4589,$ respectively, and the kurtosis is $\kappa = 4.3615.$

Example 4.4 Consider the $SETGARCH_d(2; 1, 1)$ model defined as $X_t = \sqrt{h_t}e_t,$ where the volatility process h_t (at time t) is as follows

$$h_t = \begin{cases} a + c h_{t-1} e_{t-1}^2 & \text{if } X_{t-d} \geq r_1 \\ a + b h_{t-1} & \text{if } X_{t-d} < r_1 \end{cases} \text{ with } e_t \sim \mathcal{N}(0; 1).$$

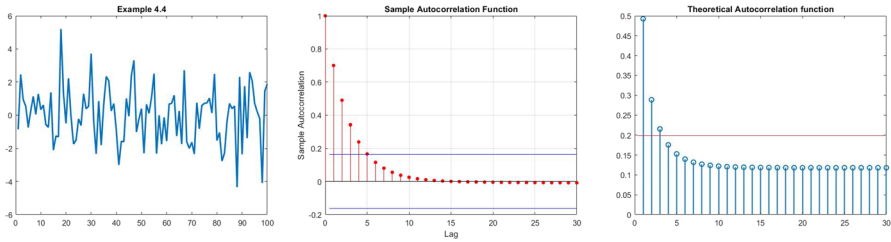


Fig. 4 Example 4.4: original data (left), sample autocorrelation function (in the middle) and the theoretical autocorrelation function (right)

Then the regime moments in the following table

	$j = 1$	$j = 2$
$\mu_1(j)$	0	0
$\mu_2(j)$	$\frac{a}{1-c}$	$\frac{a}{1-b}$
$\mu_3(j)$	0	0
$\mu_4(j)$	$3 \frac{a^2(1+c)}{(1-c)(1-3c^2)}$	$3 \frac{a^2}{(1-b)^2}$

can be used to compute the skewness and the kurtosis of (X_t) . Indeed, the skewness of X_t is $\tau = 0$ and the kurtosis of X_t has the following expression

$$\kappa = 3 \frac{(1 - c^2)(1 - b)^2 p_1 + (1 - c)^2(1 - 3c^2) p_2}{(1 - 3c^2)(1 - bp_1 - cp_2)^2}.$$

Furthermore, the autocorrelation of X_t^2 is given by

$$\rho(k) = \begin{cases} \frac{1}{3} \frac{(1 + 3c)(1 - b)^2 p_1 + (1 - c)(1 - 3c^2) p_2}{(1 + c)(1 - b)^2 p_1 + (1 - c)(1 - 3c^2) p_2} & \text{if } k = 1 \\ \frac{1}{3} \frac{(1 + 2c^k - 3c^2)(1 - b)^2 p_1 + (1 + 2b^k)(1 - c)^2(1 - 3c^2) p_2}{(1 - c^2)(1 - b)^2 p_1 + (1 - c)^2(1 - 3c^2) p_2} & \text{if } k > 1 \end{cases}$$

Set $a = 1, b = 0.7, c = 0.5, d = 1, r_1 = 4$, and $p_1 = 0.7$. Then a typical realization of such a process is depicted in Fig. 4 (left) together with the sample ACF (middle) and the theoretical ACF (right). The second and fourth moments of this process are given by $\mu_2 = E(X_t^2) = 2.4000$ and $\mu_4 = E(X_t^4) = 35.2000$, respectively, and the kurtosis is $\kappa = 6.1111$. This time series is strictly stationary.

5 Second-order stationarity and ARMA representation

(5.1) *Computation of the second-order moments.* Once second-order stationarity is established, it can be useful to compute the second-order moment of the process

$(X_t)_{t \in \mathbb{Z}}$. We shall consider the centered version of the state vector \underline{Y}_t , i.e.,

$$\widehat{\underline{Y}}_t = \sum_{j=1}^{\ell} \left(\Gamma^{(j)}(e_t) \widehat{\underline{Y}}_{t-1} + \widehat{\underline{\eta}}_t^{(j)}(e_t) \right) I_{\{\widehat{\underline{Y}}_{t-1} \in R_j^{(d)}\}} + \widehat{\underline{e}}_t, \tag{5.1}$$

where $\widehat{\underline{Y}}_t := \underline{Y}_t - E\{\underline{Y}_t\}$ and $\widehat{\underline{\eta}}_t^{(j)}(e_t)$ and $\widehat{\underline{e}}_t$ are centered residual vectors.
 Let

$$\begin{aligned} \widehat{\underline{\Sigma}}_{\underline{\eta}}^{(2)} &:= \sum_{j=1}^{\ell} E \left\{ \left(\widehat{\underline{\eta}}_t^{(j)}(e_t) \right)^{\otimes 2} \right\} p_j, & \widehat{\underline{\Sigma}}_{\underline{e}}^{(2)} &:= \ell^2 E \left\{ \widehat{\underline{e}}_t^{\otimes 2} \right\}, \\ \widehat{\underline{\Xi}}(h) &:= E \left\{ \widehat{\underline{Y}}_t \otimes \widehat{\underline{Y}}_{t-h} \right\}, \\ \widehat{\underline{\Lambda}}^{(2)} &:= \ell \sum_{j=1}^{\ell} E \left\{ \widehat{\underline{\eta}}_t^{(j)}(e_t) \otimes \widehat{\underline{e}}_t + \widehat{\underline{e}}_t \otimes \widehat{\underline{\eta}}_t^{(j)}(e_t) \right\} p_j. \end{aligned}$$

Proposition 5.1 Consider the SETBL_d($\ell; p, q, p, Q$) process in (1.1) or (2.2) with centered state-space representation (5.1). Suppose that $E\{e_t^4\} < \infty$ and

$$\rho \left(\max_j \left\{ \left| \Gamma^{(2)}(j) \right| \right\} \right) < 1,$$

where

$$\Gamma^{(2)}(j) := \left(\Gamma_0^{(j)} \right)^{\otimes 2} + \left(\Gamma_1^{(j)} \right)^{\otimes 2}.$$

Then we have

$$\begin{aligned} \Xi_X(h) &= \text{vec Cov}(\underline{X}_t, \underline{X}_{t-h}) \\ &= \begin{cases} \left(H_K^{\otimes 2} \right)' \left(I_{(K^2)} - \Gamma^{(2)} \right)^{-1} \left(\widehat{\underline{\Sigma}}_{\underline{e}}^{(2)} + \widehat{\underline{\Sigma}}_{\underline{\eta}}^{(2)} + \widehat{\underline{\Lambda}}^{(2)} \right) & \text{if } h = 0 \\ \left(H_K^{\otimes 2} \right)' \Gamma_0^h \left(I_{(K^2)} - \Gamma^{(2)} \right)^{-1} \left(\widehat{\underline{\Sigma}}_{\underline{e}}^{(2)} + \widehat{\underline{\Sigma}}_{\underline{\eta}}^{(2)} + \widehat{\underline{\Lambda}}^{(2)} \right) & \text{if } h > 0 \end{cases} \end{aligned}$$

where

$$\Gamma^{(2)} := \sum_{j=1}^{\ell} \Gamma^{(2)}(j) p_j \quad \text{and} \quad \Gamma_0 := \left(\sum_{j=1}^{\ell} \Gamma_0^{(j)} p_j \right) \otimes I_{(K)}.$$

Proof Starting from (5.1), we have:

(a) If $h = 0$, then

$$\widehat{\underline{\Xi}}(0) = E \left\{ \left(\sum_{j=1}^{\ell} \left(\Gamma^{(j)}(e_t) \widehat{\underline{Y}}_{t-1} + \widehat{\underline{\eta}}_t^{(j)}(e_t) \right) I_{\{\widehat{\underline{Y}}_{t-1} \in R_j^{(d)}\}} + \widehat{\underline{e}}_t \right)^{\otimes 2} \right\}$$

$$= \underline{\widehat{\Sigma}}_{\underline{e}}^{(2)} + \underline{\widehat{\Sigma}}_{\underline{\eta}}^{(2)} + \underline{\widehat{\Lambda}}^{(2)} + \Gamma^{(2)} \underline{\widehat{\Xi}}(0),$$

where $\Gamma^{(2)}$ is as above. Thus it follows that

$$\underline{\widehat{\Xi}}(0) = \left(I_{(K^2)} - \Gamma^{(2)} \right)^{-1} \left(\underline{\widehat{\Sigma}}_{\underline{e}}^{(2)} + \underline{\widehat{\Sigma}}_{\underline{\eta}}^{(2)} + \underline{\widehat{\Lambda}}^{(2)} \right)$$

because $I_{(K)} - \Gamma^{(2)}$ is invertible as $\rho(\Gamma^{(2)}) < 1$.

(b) If $h > 0$, then

$$\underline{\widehat{\Xi}}(h) = \sum_{j=1}^{\ell} \left(\Gamma_0^{(j)} \otimes I_{(K)} \right) E \left\{ \widehat{Y}_{t-1} \otimes \widehat{Y}_{t-1-(h-1)} \right\} p_j.$$

Then we get

$$\underline{\widehat{\Xi}}(h) = \Gamma_0 \underline{\widehat{\Xi}}(h-1) = \Gamma_0^h \underline{\widehat{\Xi}}(0), \tag{5.2}$$

where Γ_0 is as above. Hence $\Xi_X(h) = \left(H_K^{\otimes 2} \right)' \underline{\widehat{\Xi}}(h) = \left(H_K^{\otimes 2} \right)' \Gamma_0^h \underline{\widehat{\Xi}}(0)$. The result follows. \square

(5.2) *ARMA representation.* ARMA representations play an important role in forecasting and model identification. For these reasons, certain nonlinear processes are already represented as ARMA models. Indeed, Bibi (2003), Bibi and Ghezal (2015b) showed that superdiagonal bilinear processes with time-varying coefficients and Markov-switching bilinear processes admit weak ARMA representations. Weak VARMA representations of multivariate Markov switching (MS) AR and MA models and MS state-space models can be found in Cavicchioli (2014, 2016), respectively.

The following proposition establishes an ARMA representation for the SETBL_d (ℓ; p, q, P, Q) model.

Proposition 5.2 *Under the conditions of Proposition 5.1, the SETBL_d (ℓ; p, q, P, Q) process with state-space representation (5.1) is an ARMA process.*

Proof Use the same approach as in Bibi and Ghezal (2015b). It is important to note that the ARMA representation naturally arises from the Markovian nature of the model. This was demonstrated by Pham (1985), Theorem 4.4, for the class of bilinear processes. \square

6 Covariance structure of higher-powers of a SETBL_d model

For the identification purpose it is necessary to look at higher-powers of the process in order to distinguish between different ARMA representations. For this purpose, we first establish the following lemma based on Lemma 3.1 in Bibi and Ghezal (2016), which is proved for the class of Markov switching bilinear processes.

Lemma 6.1 Consider the SETBL_d (ℓ; p, q, P, Q) model in (1.1) or (2.2) with centered state space representation (5.1). Let us define the following matrices B_j^(k)(i, e_t), j = 0, ..., k = 0, ..., m and i = 1, ..., ℓ, with appropriate dimension such that

$$\forall k \in \mathbb{N} : \widehat{Y}_t^{\otimes k} = \sum_{x_1, \dots, x_k=1}^{\ell} \bigotimes_{i=1}^k \left(\left(\Gamma^{(x_i)}(e_t) \widehat{Y}_{t-1} + \widehat{\eta}_t^{(x_i)}(e_t) \right) I_{\{\widehat{Y}_{t-1} \in R_{x_i}^{(d)}\}} + \widehat{e}_t \right) \tag{6.1}$$

and

$$\left(\Gamma^{(x_i)}(e_t) \widehat{Y}_{t-1} + \widehat{\eta}_t^{(x_i)}(e_t) + \widehat{e}_t \right)^{\otimes k} = \sum_{j=0}^k B_j^{(k)}(x_i, e_t) \widehat{Y}_{t-1}^{\otimes j},$$

where by convention B_j^(k)(., .) = 0 if j > k or j < 0, and $\widehat{Y}_t^{\otimes 0} = B_0^{(0)}(., .) = 1$. Then B_j^(k)(i, e_t) are uniquely determined by the following recursive formulas

$$B_0^{(1)}(i, e_t) = \widehat{\eta}_t^{(i)}(e_t) + \widehat{e}_t, \quad B_1^{(1)}(i, e_t) = \Gamma_0^{(i)} + e_t \Gamma_1^{(i)} = \Gamma^{(i)}(e_t),$$

$$B_j^{(k+1)}(i, e_t) = B_0^{(1)}(i, e_t) \otimes B_j^{(k)}(i, e_t) + B_1^{(1)}(i, e_t) \otimes B_{j-1}^{(k)}(i, e_t) \text{ for } k > 1.$$

Proof The proof follows by using an approach which is similar to that employed by Bibi and Ghezal (2015b), Lemma 3.1, for the class of Markov-switching subdiagonal bilinear processes. □

Now, set

$$\Xi_{\underline{Y}}^{(m)} = E \left\{ \widehat{Y}_t^{\otimes m} \right\}, \quad \Xi_{\underline{Y}}^{(k,m)}(h) = E \left\{ \widehat{Y}_t^{\otimes k} \otimes \widehat{Y}_{t-h}^{\otimes m} \right\} \text{ and } B_j^{(m,i)} = E \left\{ B_j^{(m)}(i, e_t) \right\},$$

for all i = 1, ..., ℓ. Then the following relations hold:

$$\Xi_{\underline{Y}}^{(m)} = \sum_{i=1}^{\ell} \sum_{j=0}^m B_j^{(m,i)} \Xi_{\underline{Y}}^{(j)} p_i \text{ and } \Xi_{\underline{Y}}^{(k,m)}(h) = \sum_{i=1}^{\ell} \sum_{j=0}^k \tilde{B}_j^{(k,i)} \Xi_{\underline{Y}}^{(k-j,m)}(h-1) p_i,$$

for all k > 1, where $\tilde{B}_j^{(m,i)} = B_j^{(m,i)} \otimes I_{(K^m)}$. Moreover, we have

$$W^{(m)}(h) = \begin{pmatrix} \Xi_{\underline{Y}}^{(m,m)}(h) \\ \Xi_{\underline{Y}}^{(m-1,m)}(h) \\ \vdots \\ \Xi_{\underline{Y}}^{(0,m)}(h) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{\ell} \tilde{B}_0^{(m,i)} p_i & \sum_{i=1}^{\ell} \tilde{B}_1^{(m,i)} p_i & \dots & \sum_{i=1}^{\ell} \tilde{B}_m^{(m,i)} p_i \\ O & \sum_{i=1}^{\ell} \tilde{B}_0^{(m-1,i)} p_i & \dots & \sum_{i=1}^{\ell} \tilde{B}_{m-1}^{(m-1,i)} p_i \\ \vdots & \ddots & \dots & \vdots \\ \vdots & \ddots & \dots & \vdots \\ O & \dots & \sum_{i=1}^{\ell} \tilde{B}_0^{(1,i)} p_i & \sum_{i=1}^{\ell} \tilde{B}_1^{(1,i)} p_i \end{pmatrix} W^{(m)}(h-1)$$

$$= \Lambda(m) W^{(m)}(h-1),$$

in which O is the null matrix with appropriate dimension.

Proposition 6.1 Consider the SETBL_d ($\ell; p, q, P, Q$) model (1.1) or (2.2) with centered state–space representation (5.1). If $E \left\{ e_t^{2(m+1)} \right\} < +\infty$ and

$$\lambda_{(2m)} := \rho \left(\max_j \left\{ \left| \Gamma^{(2m)}(j) \right| \right\} \right) < 1, \tag{6.2}$$

with $\Gamma^{(2m)}(j) := E \left\{ \left(\Gamma_0^{(j)} + e_t \Gamma_1^{(j)} \right)^{\otimes 2m} \right\}$, then $(\widehat{Y}_t^{\otimes m})_{t \in \mathbb{Z}}$ is a second-order stationary process, and the following relations hold:

$$\begin{aligned} \underline{\Xi}_Y^{(m)} &= \begin{cases} 1 & \text{if } m = 0 \\ \left(I_{(K^m)} - \sum_{i=1}^{\ell} B_m^{(m,i)} p_i \right)^{-1} \sum_{j=0}^{m-1} \left(\sum_{i=1}^{\ell} B_j^{(m,i)} p_i \right) \underline{\Xi}_Y^{(j)} & \text{if } m > 0 \end{cases}, \\ \underline{\Xi}_Y^{(k,m)}(1) &= \sum_{j=0}^k \left(\sum_{i=1}^{\ell} \tilde{B}_j^{(k,i)} p_i \right) \underline{\Xi}_Y^{(j+m)} \quad \text{and} \\ \underline{W}^{(m)}(h) &= \Lambda(m) \underline{W}^{(m)}(h-1) \quad \text{for any } h > 1. \end{aligned}$$

Remark 6.1 Proposition 6.1 allows to compute $W^{(m)}(h)$ recursively for all $h \geq 0$. The unconditional mean and the covariance structure of $(\underline{X}_t^{\otimes m})_{t \in \mathbb{Z}}$ are given by

$$\begin{aligned} \underline{\mu}_X^{(m)} &= E \left\{ \widehat{X}_t^{\otimes m} \right\} = \underline{F}^{\otimes m'} \underline{\Xi}_Y^{(m)}, \\ \Sigma_X(h) &= E \left\{ \widehat{X}_t^{\otimes m} \otimes \widehat{X}_{t-h}^{\otimes m} \right\} - E \left\{ \widehat{X}_t^{\otimes m} \right\} \otimes E \left\{ \widehat{X}_{t-h}^{\otimes m} \right\} \\ &= \underline{F}^{\otimes 2m'} \underline{\Xi}_Y^{(m,m)}(h) - E \left\{ \widehat{X}_t^{\otimes m} \right\} \otimes \underline{F}^{\otimes m'} \underline{\Xi}_Y^{(0,m)}(h) \\ &= \underline{v}' W^{(m)}(h), \end{aligned}$$

where $\underline{v}' = (\underline{F}^{\otimes 2m'}; O; \dots; O; -E \left\{ \widehat{X}_t^{\otimes m} \right\} \otimes \underline{F}^{\otimes m'})$.

Corollary 6.1 [The SETARMA_d model] For the SETARMA_d model, whose coefficients $c_{mk}^{(j)}$ in (1.1) are all zeros, Condition (6.2) reduces to $\rho \left(\max_j \left\{ \left| \left(A_0^{(j)} \right)^{\otimes 2m} \right| \right\} \right) < 1$.

Corollary 6.2 [The SETGARCH_d model] For the SETGARCH_d model, whose volatility process can be regarded as a diagonal bilinear model without moving average part, Condition (6.2) reduces to

$\rho \left(\max_j \left\{ E \left\{ \left| \left(\tilde{A}_0^{(j)} + e_t^2 \tilde{A}_1^{(j)} \right)^{\otimes 2m} \right| \right\} \right\} \right) < 1$, where the matrices $(\tilde{A}_0^{(j)})_j, (\tilde{A}_1^{(j)})_j$ are uniquely determined and easily obtained.

We are now in a position to state the following result

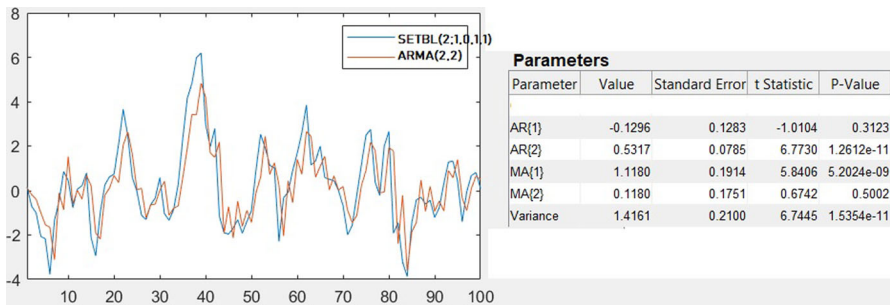


Fig. 5 The SETBL process in Example 4.2, and its ARMA(2, 2) representation

Proposition 6.2 Let $(X_t)_{t \in \mathbb{Z}}$ be the causal second-order stationary solution of the SETBL_d($\ell; p, q, P, Q$) model with centered state-space representation (5.1). Then under the conditions of Proposition 6.1, the power process $(X_t^m)_{t \in \mathbb{Z}}$ is an ARMA process.

Proof The proof follows essentially the same arguments as in Proposition 5.2. □

Remark 6.2 It is often necessary to look into higher-order cumulants in order to distinguish between linear and nonlinear models.

Example 6.1 The characteristics of the ARMA representations for the process X_t^m driven by SETBL_d($\ell; 1, 0, 1, 1$) models are collected in the following table

Specification		Representation
	$m = 1$	
Standard		ARMA(1, 1)
SETGARCh _d		ARMA(2 ℓ - 1, 2 ℓ - 1)
SETBL _d		ARMA(ℓ - 1, ℓ - 1)
	$m > 1$	
Standard		ARMA(m - 1, m - 1)
SETGARCh _d		ARMA((m + 1) ℓ - 1, (m + 1) ℓ - 1)
SETBL _d		ARMA((m + 1) ℓ - 1, (m + 1) ℓ - 1)

Example 6.2 Let us consider the SETBL process in Example 4.2. The process admits an ARMA(2, 2) representation $(1 - \phi_1 L - \phi_2 L^2)X_t = (1 + \theta_1 L + \theta_2 L^2)\epsilon_t$, whose coefficient estimates are reported aside Fig. 5.

Example 6.3 The SETARMA process in Example 4.3 admits an ARMA(1, 1) representation

$$(1 - \phi L) X_t = (1 + \theta L) \epsilon_t,$$

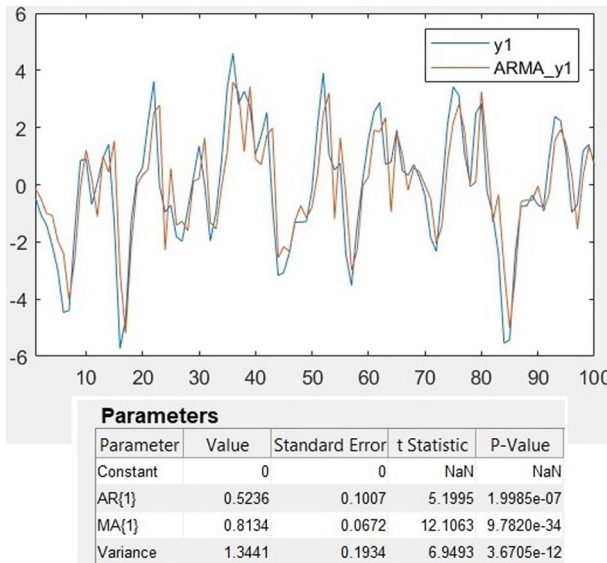


Fig. 6 The SETARMA process in Example 4.3, and its ARMA(1, 1) representation

whose coefficient estimates are reported in Fig. 6. Indeed, the AR and MA coefficients ϕ and θ are solutions of the equations

$$1 + \theta^2 + 2\phi\theta = \mu_2(1 - \phi^2), \quad \theta = [\rho(1) - \phi]\mu_2$$

and

$$\frac{\theta}{1 + \theta^2} = \frac{\text{cov}(X_t - \phi X_{t-1}, X_{t-1} - \phi X_{t-2})}{\text{var}(X_t - \phi X_{t-1})} = \frac{\text{cov}(\epsilon_t + \theta\epsilon_{t-1}, \epsilon_{t-1} + \theta\epsilon_{t-2})}{\text{var}(\epsilon_t + \theta\epsilon_{t-1})}.$$

Substituting the numerical values from Example 4.3, we have to solve the second-order equations

$$2.736\phi^2 - 9.0824\phi + 4.0058 = 0 \quad \text{and} \quad \frac{\theta}{1 + \theta^2} = 0.4895.$$

Choosing the solutions of module less than 1 yields $\phi = 0.5236$ and $\theta = 0.8134$.

7 Invertibility and geometric ergodicity of SETBL_d processes

(7.1) *Invertibility.* The concept of invertibility plays a fundamental role in the analysis of time series and statistical applications. Various definitions of this concept have been proposed in the literature. Granger and Anderson (1978), Guégan and Pham (1987), Pham and Tran (1981), Subba Rao and Gabr (1984), Liu (1990), and Bibi and

Ghezal (2015a) have derived invertibility conditions for some particular stationary (Markov Switching-)bilinear models.

The SETBL_d(ℓ; 0, 0, p, 1) process $(X_t)_{t \in \mathbb{Z}}$ defined by the stochastic difference equation

$$X_t = \sum_{j=1}^{\ell} \sum_{m=1}^p c_{m1}^{(j)} X_{t-m} e_{t-1} I_{\{X_{t-d} \in R_j\}} + e_t, \quad (7.1)$$

is said to be *invertible* if for all $t \in \mathbb{Z}$, the innovation process e_t is $\mathcal{F}_t(X)$ -measurable.

Let $\Lambda_t := - \sum_{j=1}^{\ell} \sum_{m=1}^p c_{m1}^{(j)} X_{t-m} I_{\{X_{t-d} \in R_j\}}$. Then we get

$$e_t = X_t + \Lambda_t e_{t-1} = X_t + \sum_{k=1}^m \left\{ \prod_{j=0}^{k-1} \Lambda_{t-j} \right\} X_{t-k} + \left\{ \prod_{j=0}^m \Lambda_{t-j} \right\} e_{t-m-1}. \quad (7.2)$$

Assume that (X_t) is strictly stationary and ergodic. By the ergodic theorem, if $E \{\log |X_t|\} < \infty$, then

$$\frac{1}{m} \left(\sum_{t=0}^m \log |\Lambda_t| + \log |X_{t-m}| \right) \longrightarrow E \{\log |\Lambda_t|\} \quad \text{a.s.}$$

as $m \rightarrow \infty$. If $E \{\log |\Lambda_t|\} < 0$, then

$$\left\{ \prod_{j=0}^m |\Lambda_{t-j} X_{t-m}| \right\}^{\frac{1}{m}} \longrightarrow \exp E \{\log |\Lambda_t|\} < 1 \quad \text{a.s.}$$

as $m \rightarrow \infty$. For any $\omega \in \Omega$ such that the last inequality holds, there exists a natural

number n_ω such that $\left\{ \prod_{j=0}^m |\Lambda_{t-j} X_{t-m}| \right\}^{\frac{1}{m}}(\omega) \leq \delta_\omega < 1$, for any $m > n_\omega$. Thus,

$\left\{ \sum_{k=1}^m \left\{ \prod_{j=0}^{k-1} \Lambda_{t-j} \right\} X_{t-k} \right\}(\omega)$ converges as m goes to infinity, that is,

$$\sum_{k=1}^m \left\{ \prod_{j=0}^{k-1} \Lambda_{t-j} \right\} X_{t-k} \xrightarrow{m \rightarrow \infty} \sum_{k \geq 1} \left\{ \prod_{j=0}^{k-1} \Lambda_{t-j} \right\} X_{t-k} \quad \text{a.s.} \quad (7.3)$$

In the same manner, we can show that the second term in (7.2) converges to zero a.s. In this case, we get

$$e_t = X_t + \sum_{k \geq 1} \left\{ \prod_{j=0}^{k-1} \Lambda_{t-j} \right\} X_{t-k} \quad \text{a.s.}, \tag{7.4}$$

hence model (7.1) is invertible.

We are now in a position to state the following results.

Theorem 7.1 *Let (X_t) be the unique strictly stationary and ergodic solution of model (7.1) with $E \{\log |X_t|\} < \infty$. Then model (7.1) is invertible if*

$$\lambda := E \left\{ \log \left(\prod_{j=1}^{\ell} \left| \sum_{m=1}^p c_{m1}^{(j)} X_{t-m} \right|^{F_X(r_j) - F_X(r_{j-1})} \right) \right\} < 1,$$

where $F_X(x)$ is the distribution of X_t . Furthermore, model (7.1) is not invertible if $\lambda > 1$.

The above arguments can be generalized to the case of a $SETBL_d(\ell; 0, 0, p, q)$ model as follows.

Let us consider the $SETBL_d(\ell; 0, 0, p, q)$ model defined by

$$X_t = \sum_{j=1}^{\ell} \sum_{m=1}^p \sum_{k=1}^q c_{mk}^{(j)} X_{t-m} e_{t-k} I_{\{X_{t-d} \in R_j\}} + e_t, \tag{7.5}$$

Let $e_t := (e_t, e_{t-1}, \dots, e_{t-q+1})' \in \mathbb{R}^q$. Then we get

$$\begin{aligned} e_t &= H_q X_t + \Lambda_t e_{t-1} = H_q X_t \\ &+ \sum_{k=1}^m \left\{ \prod_{j=0}^{k-1} \Lambda_{t-j} \right\} H_q X_{t-k} + \left\{ \prod_{j=0}^m \Lambda_{t-j} \right\} e_{t-m-1}, \end{aligned} \tag{7.6}$$

with $(\Lambda_t)_t$ defined as

$$\Lambda_t := \begin{pmatrix} - \sum_{j=1}^{\ell} \sum_{m=1}^p c_{m1}^{(j)} X_{t-m} I_{\{X_{t-d} \in R_j\}} & \cdots & \cdots & - \sum_{j=1}^{\ell} \sum_{m=1}^p c_{mq}^{(j)} X_{t-m} I_{\{X_{t-d} \in R_j\}} \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}_{q \times q}.$$

Similarly as above, we can show that the second term in (7.6) converges to zero a.s. In this case, we get

$$e_t = X_t + \sum_{k \geq 1} H'_q \left\{ \prod_{j=0}^{k-1} \Lambda_{t-j} \right\} H_q X_{t-k} \quad \text{a.s.}, \tag{7.7}$$

hence model (7.5) is invertible. Finally, we obtain the following result.

Theorem 7.2 *Let (X_t) be the unique strictly stationary and ergodic solution of model (7.5) with $E \{\log |X_t|\} < \infty$. Then model (7.5) is invertible if*

$$\lambda := E \{\log |\Lambda_t|\} < 0,$$

and it is not invertible if $\lambda > 0$.

(7.2) *Geometric ergodicity.* The final result concerns the geometric ergodicity of the process (\underline{Y}_t) driven by the SETBL_d model in (2.2). The geometric ergodicity of a threshold first-order bilinear process has been studied by Cappuccio et al. (1998). In our article we study the geometric ergodicity of general threshold bilinear models. Let \underline{Y}_0 be an arbitrarily specified random vector in \mathbb{R}^K . The process (\underline{Y}_t) is a Markov chain with a state space \mathbb{R}^K and with n -step transition probability $P^{(n)}(\underline{y}, C) = P(\underline{Y}_n \in C | \underline{Y}_0 = \underline{y})$ and invariant probability measure $\pi(C) = \int P(\underline{y}, C) \pi(d\underline{y})$ for any Borelean set $C \in \mathcal{B}_{\mathbb{R}^K}$. The chain is said to be ϕ -irreducible if, for some non trivial measure ϕ on $(\mathbb{R}^K, \mathcal{B}_{\mathbb{R}^K})$ such that $\forall C \in \mathcal{B}_{\mathbb{R}^K}$, $\phi(C) > 0 \implies \exists n > 0, P^{(n)}(\underline{y}, C) > 0$ for every \underline{y} . It is called a *Feller Markov chain* if the function $E \left\{ g(\underline{Y}_t) | \underline{Y}_{t-1} = \underline{y} \right\}$ is continuous for every bounded and continuous function g defined on \mathbb{R}^K . The chain is said to be *geometrically ergodic* if there exist some probability measure π on $\mathcal{B}_{\mathbb{R}^K}$ and a positive real number $c \in]0, 1]$ such that $c^{-n} \left\| P^{(n)}(\underline{y}, \cdot) - \pi(\cdot) \right\|_V \rightarrow 0$ as $n \rightarrow +\infty$, where $\|\cdot\|_V$ denotes the total variation norm.

Let us make the following assumption

[A.2] The marginal distribution of e_t is absolutely continuous with respect to the Lebesgue measure λ on \mathbb{R} . The support of e_t , defined by its strictly positive density f_e , contains an open set around zero.

Proposition 7.1 *Under Assumptions [A0] – [A2], the Markov chain (\underline{Y}_t) is a λ -irreducible, aperiodic and strong Feller chain. Moreover, if (\underline{Y}_t) is starting with the stationary distribution, then its first component process is strictly stationary and geometrically ergodic.*

The proof of this result follows using the same arguments from Theorem 4.1 in Bibi and Ghezal (2015b) for the class of Markov switching bilinear models.

8 Conclusion

In this paper we analyse some statistical and probabilistic properties of the self-exciting threshold superdiagonal bilinear (SETBL) time series model, which are useful in statistics and econometrics. Using a suitable Markovian representation of such a model, we derive sufficient conditions to have a unique, strictly stationary, causal and ergodic

solution. Under neat matrix conditions, we prove that the higher-order moments of the SETBL process and its powers are finite. Then we obtain matrix expressions in closed-form for the computations of such moments. As a result, the skewness and kurtosis indexes are explicitly computed. These theoretical results are useful for the diagnostic analysis of the process and for model identification. Furthermore, we derive the autocorrelation function and the covariance functions of the SETBL process, and provide its ARMA representations. Such representations play an important role in forecasting and to estimate model parameters via GMM method. Finally, we examine necessary and sufficient conditions for the invertibility and geometric ergodicity of the proposed model. Several examples illustrate the obtained theoretical results.

In the paper we concentrate on superdiagonal bilinear models as the crucial step of our method is to represent them by a Markovian state-space form. As pointed out by one of the referees, the general model may be also represented by an elegant ℓ -Markovian form. This is an interesting problem, which can be solved by a direct modification of the proposed techniques. In this setting, and particularly to compute explicitly the higher-order moments, we need to consider the additional assumption that the innovation e_t and the sequence $\{(\underline{Y}_{\tau-1}, s_\tau)\}_{\tau \leq t}$ are independent, where \underline{Y}_t is the Markov chain with state space \mathbb{R}^K from Sect. 2 and s_t denotes the corresponding state function. After that, all the main results of the paper maintain their validity for the general model.

Acknowledgements Maddalena Cavicchioli acknowledges financial support by FAR 2022 research grant of the University of Modena and Reggio E., Italy. The authors thank the Editor in Chief of the journal, Professor Carsten Jentsch, and the three anonymous referees for their constructive comments and very useful suggestions and remarks which were most valuable for improvement of the final version of the paper.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bibi A (2003) On the covariance structure of time-varying bilinear models. *Stoch Anal Appl* 21:25–60
- Bibi A (2023) Higher-order moments of Markov switching bilinear models. *Stat Pap* (Manuscript)
- Bibi A, Ghezal A (2015a) Consistency of quasi-maximum likelihood estimator for Markov-switching bilinear time series models. *Stat Probab Lett* 100:192–202
- Bibi A, Ghezal A (2015b) On the Markov-switching bilinear processes: stationarity, higher-order moments and β -mixing. *Stochastics* 1–27
- Bibi A, Ghezal A (2016) Minimum distance estimation of Markov-switching bilinear processes. *Statistics* 50(6):1290–1309
- 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 Probab* 18:221–254
- Brockwell PJ, Liu J, Tweedie RL (1992) On the existence of stationary threshold autoregressive moving-average processes. *J Time Ser Anal* 13(2):95–107
- Cappuccio N, Ferrante M, Fonseca G (1998) A note on the stationarity of a threshold first-order bilinear process. *Stat Probab Lett* 40:379–384
- Cavicchioli M (2014) Determining the number of regimes in Markov switching *VAR* and *VMA* models. *J Time Ser Anal* 35(2):173–186
- Cavicchioli M (2016) Weak *VARMA* representations of regime-switching state-space models. *Stat Pap* 57(3):705–720
- Cavicchioli M (2017a) Higher order moments of Markov switching *VARMA* models. *Economet Theory* 33(6):1502–1515
- Cavicchioli M (2017b) Asymptotic Fisher information matrix of Markov switching *VARMA* models. *J Multivar Anal* 157:124–135
- Chan K-S, Goracci G (2019) On the ergodicity of first-order threshold autoregressive moving-average processes. *J Times Ser Anal* 40(2):256–264
- Chan K-S, Giannerini S, Goracci G, Tong H (2024) Testing threshold regulation in presence of measurement error. *Stat Sin* 34(3):1–40
- Chen CWS, Lee JC (1995) Bayesian inference of threshold autoregressive models. *J Times Ser Anal* 16(5):483–492
- Choi MS, Park JA, Hwang SY (2012) Asymmetric GARCH processes featuring both threshold effect and bilinear structure. *Stat Probab Lett* 82:419–426
- Davis RA, Resnick SI (1996) Limit theory for bilinear processes with heavy-tailed noise. *Ann Appl Probab* 6(4):1191–1210
- Feder PI (1975) On asymptotic distribution theory in segmented regression problems-identified case. *Ann Stat* 3(1):49–83
- Ferrante M, Fonseca G, Vidoni P (2003) Geometric ergodicity, regularity of the invariant distribution and inference for a threshold bilinear Markov process. *Stat Sin* 13:367–384
- Franq C, Zakoian JM (2001) Stationarity of multivariate Markov-switching ARMA models. *J Econometr* 102:339–364
- Franq C, Zakoian JM (2005) L^2 -structures of standard and switching regime GARCH models. *Stoch Process Appl* 115:1557–1582
- Gibson D, Nur D (2011) Threshold autoregressive models in finance: a comparative approach. *Applied Statistics Education and Research Collaboration (ASEARC)—Conference paper 26*. <https://ro.uow.edu.au/asearc/26>
- Goracci G, Giannerini S, Chan K-S, Tong H (2021) Testing for threshold effects in the TARMA framework. arXiv preprint [arXiv:2103.13977](https://arxiv.org/abs/2103.13977)
- Granger CWJ, Anderson A (1978) An introduction to bilinear time series models. Vandenhoeck and Ruprecht, Göttingen
- Guégan D, Pham DT (1987) Minimalité et Inversibilité des modèles bilinéaires à temps discret. *C R Acad Sci Paris Serie I Math* 448:159–162
- Kristensen D (2009) On stationarity and ergodicity of the bilinear model with applications to the *GARCH* models. *J Time Ser Anal* 30:125–144
- Li G, Li W (2011) Testing a linear time series model against its threshold extension. *Biometrika* 98(1):243–250
- Ling S, Tong H, Li D (2007) Ergodicity and invertibility of threshold moving-average models. *Bernoulli* 13(1):161–168
- Liu J (1990) A note on causality and invertibility of a general bilinear time series model. *Adv Appl Probab* 22:247–250
- Ma L, Grant AJ, Sofronov G (2020) Multiple change point detection and validation in autoregressive time series data. *Stat Pap* 61:1507–1528
- Mohr M, Selk L (2020) Estimating change points in nonparametric time series regression models. *Stat Pap* 61:1437–1463
- Petrucelli JD, Davies N (1986) A portmanteau test for self-exciting threshold autoregressive-type nonlinearity in time series. *Biometrika* 73(3):687–694
- Pham DT (1985) Bilinear Markovian representation and bilinear models. *Stoch Process Appl* 20(2):295–306

- Pham DT (1986) The mixing property of bilinear and generalized random coefficient autoregressive models. *Stoch Process Appl* 23(2):291–300
- Pham DT, Tran LT (1981) On the first order bilinear time series models. *J Appl Probab* 18(3):617–627
- Popovič PM, Bakouch HS (2020) A bivariate integer-valued bilinear autoregressive model with random coefficients. *Stat Pap* 61:1819–1840
- Subba Rao T, Gabr MM (1984) An introduction to bispectral analysis and bilinear time series models, vol 24. *Lecture Notes in statistics*. Springer, Berlin
- Tong H (1983) Threshold models in non-linear time series analysis. *Lecture Notes in statistics*, vol 21. Springer, Berlin, pp 59–121
- Tong H (2011) Threshold models in time series analysis—30 years on. *Stat Interface* 4(2):107–118
- Tong H (2012) Threshold models in non-linear time series analysis, vol 21. Springer, Berlin
- Tong H (2015) Threshold models in time series analysis—some reflections. *J Econometr* 189(2):485–491
- Tweedie RL (1974a) R-Theory for Markov chains on a general state space I. Solidarity properties and r -recurrent chains. *Ann Probab* 2(5):840–864
- Tweedie RL (1974b) R-Theory for Markov chains on a general state space II. r -subinvariant measures for r -transient chains. *Ann Probab* 2(5):865–878
- Tweedie RL (1975) Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space. *Stoch Process Appl* 3(4):385–403
- Tweedie RL (1976) Criteria for classifying general Markov chains. *Adv Appl Probab* 8(4):737–771

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.