REVSTAT – Statistical Journal
Volume 19, Number 3, July 2021, 443–462

PSEUDO-GAUSSIAN AND RANK-BASED TESTS FOR FIRST-ORDER SUPERDIAGONAL BILINEAR MODELS IN PANEL DATA

Authors: Aziz Lmakri
– Mathematics and Applications Laboratory, Faculty of Sciences and Techniques, Abdelmalek Essaadi University, Tangier, Morocco
lmakriaziz@gmail.com

Abdelhadi Akharif
– Mathematics and Applications Laboratory, Faculty of Sciences and Techniques, Abdelmalek Essaadi University, Tangier, Morocco
aakharif@gmail.com

Amal Mellouk
– Regional Center for Education and Training Trades, Tangier, Morocco
prmelloukamal@hotmail.com

Mohamed Fihri
– Laboratory of Mathematics, Computing and Applications – Information Security (LabMiA-SI), Department of Mathematics, Faculty of Sciences, Mohammed-V University in Rabat, Morocco
m.fihri@gmail.com

Received: March 2019 Revised: July 2019 Accepted: July 2019

Abstract:

- In this paper, locally asymptotically optimal (in the Hájek-Le Cam sense) parametric, pseudo-Gaussian and rank-based procedures are proposed for the problem of testing randomness against first-order superdiagonal bilinear panel dependence (in large n and small T panels). Local powers and asymptotic relative efficiencies are computed and show that the van der Waerden version of our rank-based tests uniformly dominates. Small-sample performances are investigated via simulations and confirm the theoretical findings, they also demonstrate the remarkable performances of rank procedures based on data-driven scores.

Keywords:

- *panel data; first-order superdiagonal bilinear model; local asymptotic normality; pseudo-Gaussian test; rank test; asymptotic relative efficiency.*

AMS Subject Classification:

- 62F03, 62G10.

1. INTRODUCTION

Nonlinear time series have attracted much attention in the last four decades. Many classes of models have been proposed and applied with great success in many important real-life problems; such as economics (Granger and Andersen [12]), demography (Subba Rao and Gabr [32]), environmental studies (Guegan [14]), etc. One of the most popular was the bilinear time series models $BL(p,q,P,Q)$[1]. In the first time, these models were proposed and developed by Granger and Andersen [12]; then becomes Phan and Tran [27], Subba Rao [31], Guegan [13], Liu and Brockwell [25]. Particularly to those models, we quote the first-order superdiagonal bilinear models $BL(0,0,2,1)$, who also recognized applications in many fields (see, for example, [26, 36, 5]).

This paper deals with the presence of a first-order superdiagonal bilinear model in ***panel data*** (a series of T observations made through time over a number n of individuals), denoted by $BLP(0,0,2,1)$ and defined, for $i=1,2,...,n$ and $t=1,2,...,T$, as:

$$X_{i,t} = bX_{i,t-2}\varepsilon_{i,t-1} + \varepsilon_{i,t}, \tag{1.1}$$

where $X_{i,t}$ is a panel observation (for individual i at time t) described by a nonlinear stochastic difference in time equation; $(\varepsilon_{i,t})$ is a white noise process, i.e. a sequence of independent, identically distributed random variables with mean zero, finite variance σ^2 and density distribution $\varepsilon \mapsto f(\varepsilon) := (1/\sigma)f_1(\varepsilon/\sigma)$ (where $f_1 \in \mathcal{F}_0$, see (2.1)) and b is a constant in $\mathbb{R}$. The probabilistic properties of a first-order superdiagonal time series model $BL(0,0,2,1)$ processes (such as invertibility and stationarity) have been studied by several references [28, 13]. These properties also remain valid under a *first-order superdiagonal panel model* $BLP(0,0,2,1)$. Let us denote by $\mathcal{F}_{i,t}(\varepsilon)$ and $\mathcal{F}_{i,t}(X)$ the σ-algebras generated by $\{\varepsilon_{i,s}|s \leqslant t\}$ and $\{X_{i,s}|s \leqslant t\}$, respectively. Then:

1. Equation (1.1) admits a unique stationary solution $(X_{i,t})$ (i.e., $\mathcal{F}_{i,t}(\varepsilon)$-measurable) iff $b^2\sigma^2 < 1$, in this case, one can write

$$X_{i,t} = \sum_{j=1}^{\infty} b^j \varepsilon_{i,t-2j} \prod_{k=1}^{j} \varepsilon_{i,t-2k+1} + \varepsilon_{i,t}; \tag{1.2}$$

2. Equation (1.1) is invertible (i.e., $\varepsilon_{i,t}$ is $\mathcal{F}_{i,t}(X)$-measurable) iff $2b^2\sigma^2 < 1$, in this case, one can write

$$\varepsilon_{i,t} = X_{i,t} + \sum_{j=1}^{\infty}(-b)^j X_{i,t-j} \prod_{k=1}^{j} X_{i,t-k-1}. \tag{1.3}$$

[1] These models are defined as:

$$X_t = \sum_{j=1}^{p} a_j X_{t-j} + \sum_{j=1}^{q} c_j \varepsilon_{t-j} + \sum_{j=1}^{P}\sum_{k=1}^{Q} b_{jk}\varepsilon_{t-j}X_{t-k} + \varepsilon_t.$$

Several methods — such as the method of moments, the least squares method and the repeated residual method — have been established in the literature for estimating the parameters of bilinear models, see, for example, Pham and Tran [28], Sesay and Subba Rao [30], Grahn [11], Bouzaachane [5] and Tan and Wang [34].

Before turning to the problem of estimating the parameters of model (1.1), it is very important to know if it is indeed a $BLP(0,0,2,1)$, and how the test proposed for testing randomness against first-order superdiagonal bilinear panel dependence is efficient. Note that if $b = 0$, $X_{i,t}$ reduces to white noise ($X_{i,t} = \varepsilon_{i,t}$), else $b \neq 0$, panel data follows a $BLP(0,0,2,1)$ (alternative hypothesis) — such a test is *bilateral.*

To start with, locally and asymptotically optimal parametric tests are constructed using the Local Asymptotic Normality LAN property. Then, the special case of the pseudo-Gaussian tests (optimal under Gaussian densities and valid under finite-variance non-Gaussian ones) is derived. Unfortunately, their local asymptotic power, under non-Gaussian g_1 (especially the skew and heavy-tailed ones), can be extremely poor. Which leads us to the construction of rank-based optimal tests (van der Waerden, Wilcoxon, Laplace, data-driven scores, etc.).

Asymptotic relative efficiencies with respect to the pseudo-Gaussian procedure show that the van der Waerden version of our rank-based tests uniformly dominates its pseudo-Gaussian countepart.

The paper is organized as follow: Section 2.1 provides the main definitions and assumptions. The local asymptotic normality, with respect to b and σ^2, in the vicinity of $b = 0$, of the family of distributions associated with (1.1) (with specified f_1), is established in Section 2.2. In Section 3.1, we propose (still, for specified f_1) the optimal parametric test. The particular case of the pseudo-Gaussian test is proposed in Section 3.2. Section 4 proposes rank-based procedures that remain valid irrespective of f_1. Particular cases (van der Warden, Wilcoxon, Laplace scores, ...) are considered in Section 4.3. Asymptotic relative efficiencies with respect to the pseudo-Gaussian test is derived in Section 5. Section 6 provides some simulation results assessing the finite-sample performance of the various tests proposed. Finally, Section 7 concludes.

2. LOCAL ASYMPTOTIC NORMALITY

2.1. Notations and main technical assumptions

Denote by $\mathbf{P}^{(n)}_{\sigma^2,0;f_1}$ the probability distribution under the null $X_{i,t} = \varepsilon_{i,t}$. Under the alternative, the probability distribution is denoted by $\mathbf{P}^{(n)}_{\sigma^2,b;f_1}$ ($b \neq 0$), the observations $X^{(n)} := (X_1^{(n)\prime}, X_2^{(n)\prime}, ..., X_n^{(n)\prime})'$ with $X_i^{(n)} := (X_{i,1}, ..., X_{i,T})'$ is generated by (1.1).

We suppose that the vector $X_0^{(n)} := \{(X^{(n)}_{i,-1}\varepsilon_{i,0}, X^{(n)}_{i,0}), i = 1, 2, ..., n\}$ is observable for each individual i, and admits a density $h_\theta(.)$ continuous in θ. The influence of these starting values is asymptotically negligible (see Hallin and Werker (1999) [20] for a detailed discussion).

Throughout, we consider the class of standardized densities

$$\mathcal{F}_0 := \left\{ f_1 : \int_{-\infty}^{0} f_1(u)du = 0.5 = \int_{-1}^{1} f_1(u)du \right\}. \tag{2.1}$$

Under $f_1 \in \mathcal{F}_0$, the median and median absolute deviation are 0 and σ respectively; this standardization avoids all moment assumptions and has no impact on subsequent results.

Our derivation of locally asymptotically optimal tests at density f_1 will be based on the local asymptotic normality, with respect to $(\sigma^2, b)'$, of the families of distributions

$$\mathcal{P}_{f_1}^{(n)} := \left\{ \mathbf{P}_{\sigma^2,b;f_1}^{(n)} | (\sigma^2, b)' \in \mathbb{R}_+^* \times \mathbb{R} \text{ and } 2b^2\sigma^2 < 1 \right\} \tag{2.2}$$

at any $\theta := (\sigma^2, 0)'$.

This LAN property requires some technical assumptions on the innovation density f_1. Denote by $\mathcal{F}_A$ the class of all densities f_1 satisfying the following technical assumptions:

(**A.1**) $f_1 \in \mathcal{F}_0$;

(**A.2**) $f_1(u) > 0, \quad \forall u \in \mathbb{R}$;

(**A.3**) f_1 is absolutely continuous on bounded intervals, i.e., there exists f_1' such that

$$f_1(b) - f_1(a) = \int_a^b f_1'(u)du \quad \text{for all} \quad a < b,$$

and, letting $\Phi_{f_1} = -f_1'/f_1$, assume that

$$I(f_1) := \int_{\mathbb{R}} \Phi_{f_1}^2(u) f_1(u)du \qquad \text{and} \qquad J(f_1) := \int_{\mathbb{R}} u^2 \Phi_{f_1}^2(u) f_1(u)du$$

are finite.

For instance, interesting special cases of f_1 are obtained:

- The double-exponential or Laplace distribution, with standardized density

$$f_1(u) = f_{\mathcal{L}}(u) := (1/2d)\exp(-|u|/d),$$

  with $I(f_1) = 1/d^2$ and $J(f_1) = 2$; the normalizing constant $d := 1/\ln(2) \simeq 1.4426$ is such that $f_{\mathcal{L}} \in \mathcal{F}_A$.
- The logistic distribution, with standardized density

$$f_1(u) = f_{Log}(u) := \sqrt{b}\exp(-\sqrt{b}u)/(1 + \exp(-\sqrt{b}u))^2,$$

  with $I(f_1) = b/3$ and $J(f_1) = (12 + \pi^2)/9$; the normalizing constant $b := (\ln 3)^2) \simeq 1.2069$ is such that $f_{\mathcal{L}} \in \mathcal{F}_A$.
- The Student distributions (with $\nu > 2$ degrees of freedom), with standardized density

$$f_1(u) = f_{t_\nu}(u) := \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)} \sqrt{a_\nu/\pi\nu}(1 + a_\nu u^2/\nu)^{-(\nu+1)/2},$$

  with $I(f_1) = a_\nu(\nu+1)/(\nu+3)$ and $J(f_1) = 3(\nu+1)/(\nu+3)$; the normalizing constant $a_\nu > 0$ is such that $f_{t_\nu} \in \mathcal{F}_A$.
- The Gaussian distribution, with standardized density (with mean zero and variance $1/a$)

$$f_1(u) = f_{\mathcal{N}}(u) := \sqrt{a/2\pi}\exp(-au^2/2),$$

  with $I(f_1) = a \simeq 0.4549$ and $J(f_1) = 3$.

2.2. LAN

Let us denote by $\theta^{(n)}$ the local sequences of perturbations of $\theta = (\sigma^2, 0)'$, where

$$\theta^{(n)} = \theta + n^{\frac{-1}{2}}\tau \quad \text{with} \quad \tau = \big(\tau_1, \tau_2\big)' \in \mathbb{R}^2.$$

The bilateral test is equivalent to:

$$\begin{cases} \mathbf{P}^{(n)}_{\theta;f_1} : \tau_2 = 0, \\ \mathbf{P}^{(n)}_{\theta^{(n)};f_1} : \tau_2 \neq 0. \end{cases}$$

Under the null, the likelihood function for $(X_0^{(n)}, X^{(n)})$ is

$$L_{\theta;f}(X_0^{(n)}, X^{(n)}) = h_\theta(X_0^{(n)}) \prod_{i=1}^{n}\prod_{t=1}^{T} f(X_{i,t}). \tag{2.3}$$

If $\tau_2 \neq 0$, the likelihood function for $(X_0^{(n)}, X^{(n)})$ in this case is

$$\begin{aligned} L_{\theta^{(n)};f}(X_0^{(n)}, X^{(n)}) &= h_{\theta^{(n)}}(X_0^{(n)}) \prod_{i=1}^{n}\prod_{t=1}^{T} f\big(X_{i,t} + \sum_{j=1}^{\infty}(-n^{\frac{-1}{2}}\tau_2)^j X_{i,t-j}\prod_{k=1}^{j} X_{i,t-k-1}\big) \\ &= h_{\theta^{(n)}}(X_0^{(n)}) \prod_{i=1}^{n}\prod_{t=1}^{T} f\big(X_{i,t} + \Upsilon_n(\tau_2)\big), \end{aligned} \tag{2.4}$$

where $\Upsilon_n(\tau_2) := \sum_{j=1}^{\infty}(-n^{\frac{-1}{2}}\tau_2)^j X_{i,t-j}\prod_{k=1}^{j} X_{i,t-k-1}$.

Denote by $\Lambda^{(n)}_{\theta^{(n)}/\theta;f}$ the logarithm of the likelihood ratio (conditional on $X_0^{(n)}$) for $\mathbf{P}^{(n)}_{\theta^{(n)};f}$ against $\mathbf{P}^{(n)}_{\theta;f}$:

$$\Lambda^{(n)}_{\theta^{(n)}/\theta;f} := \log\Big(L_{\theta^{(n)};f}(X_0^{(n)}, X^{(n)}) / L_{\theta;f}(X_0^{(n)}, X^{(n)})\Big). \tag{2.5}$$

It can be expressed as follows:

$$\Lambda^{(n)}_{\theta^{(n)}/\theta;f} = \sum_{i=1}^{n}\sum_{t=1}^{T}\Big(\log f(X_{i,t} + \Upsilon_n(\tau_2)) - \log f(X_{i,t})\Big) + o_p(1).$$

The $o_p(1)$ term (under $\mathbf{P}^{(n)}_{\theta;f}$, as $n \to \infty$) corresponds to the influence of the starting value $X_0^{(n)}$.

Write $Z_{i,t}$ for the standardized residual

$$Z_{i,t}(\sigma^2, b) := \sigma^{-1}\Bigg(X_{i,t} + \sum_{j=1}^{\infty}(-b)^j X_{i,t-j}\prod_{k=1}^{j} X_{i,t-k-1}\Bigg),$$

and note that, under $\mathbf{P}^{(n)}_{\theta;f_1}$, these residuals coincide with $\sigma^{-1}\varepsilon_{i,t}$. The local asymptotic normality result, with respect to σ^2 and the parameter of interest b for a fixed density f_1, is established in the next proposition.

Proposition 2.1. *Let $f_1 \in \mathcal{F}_A$. Then the family $\mathcal{P}_{f_1}^{(n)}$ is LAN at any $\theta = (\sigma^2, 0)'$, with central sequence*

$$\Delta_{f_1}^{(n)}(\theta) := \begin{pmatrix} \Delta_{f_1;1}^{(n)}(\theta) \\ \Delta_{f_1;2}^{(n)}(\theta) \end{pmatrix} := \begin{pmatrix} \frac{1}{2\sigma^2} n^{\frac{-1}{2}} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[\Phi_{f_1}(Z_{i,t}) Z_{i,t} - 1\right] \\ n^{\frac{-1}{2}} \sigma \sum_{i=1}^{n} \sum_{t=3}^{T} \Phi_{f_1}(Z_{i,t}) Z_{i,t-1} Z_{i,t-2} \end{pmatrix}, \tag{2.6}$$

and information matrix

$$\Gamma_{f_1}(\theta) := \big(\Gamma_{f_1;ij}(\theta)\big)_{1 \le i,j \le 2} := \begin{pmatrix} \frac{T}{4\sigma^4}\big(J(f_1) - 1\big) & 0 \\ 0 & \sigma^2 (T-2) I(f_1) \sigma_{f_1}^4 \end{pmatrix}. \tag{2.7}$$

More precisely, for any $\tau = (\tau_1, \tau_2)' \in \mathbb{R}^2$, under $\boldsymbol{P}_{\theta;f_1}^{(n)}$, as $n \to \infty$ and fixed T, we have

$$\Lambda_{\theta^{(n)}/\theta;f_1}^{(n)} = \tau' \Delta_{f_1}^{(n)}(\theta) - \frac{1}{2} \tau' \Gamma_{f_1}(\theta) \tau + o_p(1), \tag{2.8}$$

and $\Delta_{f_1}^{(n)}(\theta)$ is asymptotically normal, with mean zero under $\boldsymbol{P}_{\theta;f_1}^{(n)}$, mean $\Gamma_{f_1}(\theta)\tau$ under $\boldsymbol{P}_{\theta^{(n)};f_1}^{(n)}$ and variance $\Gamma_{f_1}(\theta)$ under both.

Proof: The proof relies on Swensen's conditions 1.2 to 1.7 of lemma 1 in [33]. More precisely, the only delicate one is the condition 1.2. The main point consists in showing that

$$(\sigma^2, b) \mapsto q_{\sigma^2,b;f_1}^{\frac{1}{2}}(z) := \left[\frac{1}{\sigma} f_1\left(\frac{z + \sum_{j=1}^{\infty} (-b)^j x_j \prod_{k=1}^{j} x_{k-1}}{\sigma}\right)\right]^{\frac{1}{2}}$$

is differentiable in mean quadratic. It is established in the following lemma.

Lemma 2.1. *Let $f_1 \in \mathcal{F}_A$. Define, for $z \in \mathbb{R}$,*

$$\begin{aligned} D_{\sigma^2} q_{\sigma^2,0;f_1}^{\frac{1}{2}}(z) &= \frac{1}{4\sigma^2} q_{\sigma^2,0;f_1}^{\frac{1}{2}}(z) \big(\tfrac{z}{\sigma} \Phi_{f_1}\big(\tfrac{z}{\sigma}\big) - 1\big), \\ D_b q_{\sigma^2,b;f_1}^{\frac{1}{2}}(z)_{|b=0} &= \frac{1}{2\sigma} q_{\sigma^2,0;f_1}^{\frac{1}{2}}(z) \Phi_{f_1}\big(\tfrac{z}{\sigma}\big) x_1 x_0. \end{aligned}$$

Then, as s and $l \to 0$,

1. $\displaystyle\int_{\mathbb{R}} \left[q_{\sigma^2+s,l;f_1}^{\frac{1}{2}}(z) - q_{\sigma^2+s,0;f_1}^{\frac{1}{2}}(z) - l D_b q_{\sigma^2+s,b;f_1}^{\frac{1}{2}}(z)_{|b=0}\right]^2 dz = o(l^2),$

2. $\displaystyle\int_{\mathbb{R}} \left[q_{\sigma^2+s,0;f_1}^{\frac{1}{2}}(z) - q_{\sigma^2,0;f_1}^{\frac{1}{2}}(z) - s D_{\sigma^2} q_{\sigma^2,0;f_1}^{\frac{1}{2}}(z)\right]^2 dz = o(s^2),$

3. $\displaystyle\int_{\mathbb{R}} \left[q_{\sigma^2+s,l;f_1}^{\frac{1}{2}}(z) - q_{\sigma^2,0;f_1}^{\frac{1}{2}}(z) - (s,l) \begin{pmatrix} D_{\sigma^2} q_{\sigma^2,0;f_1}^{\frac{1}{2}}(z) \\ D_b q_{\sigma^2,b;f_1}^{\frac{1}{2}}(z)_{|b=0} \end{pmatrix}\right]^2 dz = o(\| (s,l)' \|^2).$

Proof of Lemma 2.1:

1. Let $\Upsilon(b) = \sum_{j=1}^{\infty}(-b)^j x_j \prod_{k=1}^{j} x_{k-1}$. Then 1 takes the form

$$\int_{\mathbb{R}} \left[\frac{1}{\sqrt{\sigma^2+s}} f_1^{\frac{1}{2}}\left(\frac{z+\Upsilon(l)}{\sqrt{\sigma^2+s}}\right) - \frac{1}{\sqrt{\sigma^2+s}} f_1^{\frac{1}{2}}\left(\frac{z}{\sqrt{\sigma^2+s}}\right) - l\frac{1}{2\sqrt{\sigma^2+s}} q^{\frac{1}{2}}_{\sigma^2+s,0;f_1}(z)\Phi_{f_1}\left(\frac{z}{\sqrt{\sigma^2+s}}\right)x_1x_0\right]^2 dz = o(l^2),$$

is equivalent to

$$\int_{\mathbb{R}} \left[f^{\frac{1}{2}}\big(z+\Upsilon(l)\big) - f^{\frac{1}{2}}(z) - \frac{l}{2} f^{\frac{1}{2}}(z)\Phi_f(z)x_1x_0\right]^2 dz = o(l^2),$$

which is equivalent to

$$\int_{\mathbb{R}} l^2 \left[\frac{f^{\frac{1}{2}}\big(z+\Upsilon(l)\big) - f^{\frac{1}{2}}(z)}{l} + \frac{1}{2}\frac{f'(z)}{f^{\frac{1}{2}}(z)}x_1x_0\right]^2 dz = o(l^2),$$

hence, for proving that, it is sufficient to prove that

$$\lim_{l\to 0}\int_{\mathbb{R}} \left[\frac{f^{\frac{1}{2}}\big(z+\Upsilon(l)\big) - f^{\frac{1}{2}}(z)}{l} + \frac{1}{2}\frac{f'(z)}{f^{\frac{1}{2}}(z)}x_1x_0\right]^2 dz = 0.$$

We have

$$\begin{aligned}
\lim_{l\to 0}\frac{f^{\frac{1}{2}}\big(z+\Upsilon(l)\big) - f^{\frac{1}{2}}(z)}{l} &= \lim_{l\to 0}\frac{f^{\frac{1}{2}}\big(z+\Upsilon(l)\big) - f^{\frac{1}{2}}(z)}{\Upsilon(l)} \times \frac{\Upsilon(l)}{l} \\
&= \big(f^{\frac{1}{2}}(z)\big)' \times (-x_1x_0) \\
&= -\tfrac{1}{2}\frac{f'(z)}{f^{\frac{1}{2}}(z)}x_1x_0.
\end{aligned}$$

And just show that $\int_{\mathbb{R}} \left[\frac{f^{\frac{1}{2}}\big(z+\Upsilon(l)\big) - f^{\frac{1}{2}}(z)}{l}\right]^2 dz \leqslant \int_{\mathbb{R}}\left[\frac{-1}{2}\frac{f'(z)}{f^{\frac{1}{2}}(z)}x_1x_0\right]^2 dz < \infty.$

We know that $f^{\frac{1}{2}}\big(z+\Upsilon(l)\big) - f^{\frac{1}{2}}(z) = \int_z^{z+\Upsilon(l)} \frac{1}{2} f'(t) f^{\frac{-1}{2}}(t)dt$, then

$$\begin{aligned}
\int_{z=-\infty}^{+\infty}\left[\frac{f^{\frac{1}{2}}\big(z+\Upsilon(l)\big) - f^{\frac{1}{2}}(z)}{l}\right]^2 dz &= \int_{z=-\infty}^{+\infty}\frac{1}{l^2}\left[\int_{t=z}^{z+\Upsilon(l)}\frac{1}{2}f'(t)f^{\frac{-1}{2}}(t)dt\right]^2 dz \\
&\leqslant \frac{\Upsilon(l)}{l^2}\int_{z=-\infty}^{+\infty}\int_{t=z}^{z+\Upsilon(l)}\left[\frac{1}{2}f'(t)f^{\frac{-1}{2}}(t)\right]^2 dt\,dz \\
&\leqslant \frac{\Upsilon(l)}{l^2}\int_{t=-\infty}^{+\infty}\int_{z=t-\Upsilon(l)}^{t}\left[\frac{1}{2}f'(t)f^{\frac{-1}{2}}(t)\right]^2 dt\,dz \\
&\leqslant \left[\frac{\Upsilon(l)}{l}\right]^2\int_{t=-\infty}^{+\infty}\left[\frac{1}{2}f'(t)f^{\frac{-1}{2}}(t)\right]^2 dt \\
&\leqslant (-x_1x_0)^2\int_{t=-\infty}^{+\infty}\left[\frac{1}{2}f'(t)f^{\frac{-1}{2}}(t)\right]^2 dt \\
&\leqslant \int_{\mathbb{R}}\left[\frac{-1}{2}f'(t)f^{\frac{-1}{2}}(t)x_1x_0\right]^2 dt.
\end{aligned}$$

This completes the proof of part 1 of Lemma 2.1.

2. The problem here is reduced to the classical case of linear models considered by Swensen (1985) [33].

3. The result here follows from 1 and 2 above. This completes the proof of Lemma 2.1. □

The diagonal form of the information matrix confirms that σ^2 and b are not related, in the parametric family (2.2). They play distinct and well separated roles.

The Gaussian versions ($f_1 = f_{\mathcal{N}}$) of (2.6) and (2.7) are

$$\Delta_{\mathcal{N}}^{(n)}(\theta) = \begin{pmatrix} \dfrac{1}{2\sigma^2} n^{\frac{-1}{2}} \displaystyle\sum_{i=1}^{n}\sum_{t=1}^{T}\left[aZ_{i,t}^2 - 1\right] \\ n^{\frac{-1}{2}}\sigma a \displaystyle\sum_{i=1}^{n}\sum_{t=3}^{T} Z_{i,t}Z_{i,t-1}Z_{i,t-2} \end{pmatrix} \quad \text{and} \quad \Gamma_{\mathcal{N}}(\theta) = \begin{pmatrix} \dfrac{T}{2\sigma^4} & 0 \\ 0 & \dfrac{\sigma^2}{a}(T-2) \end{pmatrix},$$

respectively.

The result of Proposition 2.1, implies that, under assumptions $\mathcal{F}_A$, as $n \to \infty$, the family of *first-order superdiagonal panel models* $BLP(0,0,2,1)$ possesses the LAN property in a neighbourhood of white noise. This result leads us to construct asymptotically optimal parametric tests under a specified f_1. Note that these tests are valid under a specified f_1, and thereafter we will propose more general tests such as Pseudo-Gaussian and Rank-based procedures which are valid under general densities.

3. OPTIMAL PARAMETRIC AND PSEUDO-GAUSSIAN TESTS

As mentioned above, the Le Cam theory of LAN experiments allows for constructing tests which are locally asymptotically optimal (namely, *most stringent*). The basic idea is the weak convergence concept of the sequence of local experiments to the Gaussian shift two-dimensional model $\boldsymbol{\Delta} \sim \mathcal{N}(\boldsymbol{\Gamma}\tau, \boldsymbol{\Gamma})$. For a general theory on locally asymptotically optimal testing in LAN families, the reader is referred to Le Cam (1986) [23] and van der Vaart (1998) [35].

We are interested in testing the null hypothesis $b = 0$ of randomness in (1.1), with unspecified standardized error density in $\mathcal{F}_0$. To do, let us start with the case when $f_1 \in \mathcal{F}_0$ is specified, i.e., the null hypothesis is such that

$$\mathcal{H}_0^{(n)}(f_1) := \bigcup_{\sigma^2>0} \{\mathbf{P}_{\sigma^2,0;f_1}^{(n)}\},$$

and parametric alternatives take the form

$$\mathcal{H}_1^{(n)}(f_1) := \bigcup_{\sigma^2>0}\bigcup_{b\in\mathbb{R}} \{\mathbf{P}_{\sigma^2,b;f_1}^{(n)}\}.$$

3.1. Optimal parametric tests

Since $\theta=(\sigma^2,0)'=(1,0)'\sigma^2=:\Omega\sigma^2$, then $\theta\in\mathcal{M}(\Omega)$, where $\mathcal{M}(\Omega)$ is the linear subspace of dimension 1 of $\mathbb{R}^2$ generated by the vector $\Omega:=(1,0)'$. Recall that we are testing $\tau_2=0$ against $\tau_2\neq 0$, which is equivalent to testing $\tau\in\mathcal{M}(\Omega)$ against $\tau\notin\mathcal{M}(\Omega)$. Such tests should be based on the asymptotically chi-square distribution (see S. Ghosh (1999) [10]) and therefore the test statistic takes the form

$$Q_{f_1}(\theta):=\Delta^{(n)'}_{f_1}(\theta)\Big[\Gamma^{-1}_{f_1}(\theta)-\Omega\big(\Omega'\Gamma_{f_1}(\theta)\Omega\big)^{-1}\Omega'\Big]\Delta^{(n)}_{f_1}(\theta). \tag{3.1}$$

By algebra calculations, one can write

$$Q_{f_1}(\theta)=\Gamma^{-1}_{f_1;22}(\theta)\Delta^{(n)^2}_{f_1;2}(\theta)=\underline{\Delta}^{(n)^2}_{f_1}/((T-2)I(f_1)\sigma^4_{f_1})=:\underline{Q}_{f_1}, \tag{3.2}$$

with $\underline{\Delta}^{(n)}_{f_1}=n^{\frac{-1}{2}}\sum_{i=1}^{n}\sum_{t=3}^{T}\Phi_{f_1}(Z_{i,t})Z_{i,t-1}Z_{i,t-2}$.

The test based on (3.2) is locally asymptotically *most stringent* for the problem of detecting the $BLP(0,0,2,1)$ dependance in white noise process. The application of Le Cam's Third Lemma provides the asymptotic law of $\underline{Q}_{f_1}$ under $\mathbf{P}^{(n)}_{\theta^{(n)};f_1}$, so we have the following proposition.

Proposition 3.1. *Let $f_1\in\mathcal{F}_A$. Then, for any $\tau=(\tau_1,\tau_2)'\in\mathbb{R}^2$,*

- **(i)** *$\underline{Q}_{f_1}$ is asymptotically central chi-square with 1 degree of freedom under $\boldsymbol{P}^{(n)}_{\theta;f_1}$, and asymptotically noncentral chi-square, still with 1 degrees of freedom and with noncentrality parameter $\lambda_{f_1}:=(T-2)I(f_1)\sigma^2\sigma^4_{f_1}\tau_2^2$ under $\mathbf{P}^{(n)}_{\theta^{(n)};f_1}$;*
- **(ii)** *The sequence of tests rejecting the null hypothesis $\boldsymbol{P}^{(n)}_{\theta;f_1}$ whenever $\underline{Q}_{f_1}>\chi^2_{1,1-\alpha}$,* [2] *is locally asymptotically most stringent, at asymptotic level α, for $\bigcup_{\sigma^2}\{\boldsymbol{P}^{(n)}_{\sigma^2,0;f_1}\}$ against $\bigcup_{\sigma^2\in\mathbb{R}^*_+}\bigcup_{b\in\mathbb{R}}\left\{\boldsymbol{P}^{(n)}_{\sigma^2,b;f_1}\right\}$;*
- **(iii)** *The asymptotic power under $\boldsymbol{P}^{(n)}_{\theta^{(n)};f_1}$ is $1-F(\chi^2_{1,1-\alpha},\lambda_{f_1})$.* [3]

Proof:

(i) From Proposition 2.1, one can write

$$Q_{f_1}(\theta)=\Gamma^{-1}_{f_1;22}(\theta)\Delta^{(n)^2}_{f_1;2}(\theta), \tag{3.3}$$

with

$$\Delta^{(n)}_{f_1;2}(\theta)=n^{\frac{-1}{2}}\sigma\sum_{i=1}^{n}\sum_{t=3}^{T}\Phi_{f_1}(Z_{i,t})Z_{i,t-1}Z_{i,t-2}=\sigma\underline{\Delta}^{(n)}_{f_1},$$

[2] $\chi^2_{1,1-\alpha}$ is the $(1-\alpha)$-quantile of the central chi-square distribution with one degree of freedom.
[3] $F(.,\lambda_{f_1})$ is the noncentral chi-square distribution function with one degree of freedom and noncentrality parameter λ_{f_1}.

where $\underline{\Delta}_{f_1}^{(n)} := n^{\frac{-1}{2}} \sum_{i=1}^{n}\sum_{t=3}^{T} \Phi_{f_1}(Z_{i,t})Z_{i,t-1}Z_{i,t-2}$, then

$$Q_{f_1}(\theta) = \left[\sigma^2(T-2)I(f_1)\sigma_{f_1}^4\right]^{-1}\left[\sigma\underline{\Delta}_{f_1}^{(n)}\right]^2 = \underline{\Delta}_{f_1}^{(n)^2}/((T-2)I(f_1)\sigma_{f_1}^4) = \underline{Q}_{f_1}.$$

(ii) Under $\mathbf{P}_{\theta;f_1}^{(n)} : \underline{\Delta}_{f_1}^{(n)} \sim \mathcal{N}\big(0, (T-2)I(f_1)\sigma_{f_1}^4\big)$, then $\underline{\Delta}_{f_1}^{(n)^2}/((T-2)I(f_1)\sigma_{f_1}^4) = \underline{Q}_{f_1} \sim \chi_1^2$.

Under $\mathbf{P}_{\theta^{(n)};f_1}^{(n)}$, from Le Cam's Third Lemma, we have

$$\underline{\Delta}_{f_1}^{(n)} \sim \mathcal{N}\big((T-2)I(f_1)\sigma\sigma_{f_1}^4\tau_2, (T-2)I(f_1)\sigma_{f_1}^4\big),$$

hence $\underline{\Delta}_{f_1}^{(n)^2}/((T-2)I(f_1)\sigma_{f_1}^4) = \underline{Q}_{f_1} \sim \chi_1^2(\lambda_{f_1})$: noncentral chi-square of one degree of freedom and non-centrality parameter

$$\lambda_{f_1} := \left(\sqrt{(T-2)I(f_1)\sigma_{f_1}^4}\,\sigma\tau_2\right)^2 = (T-2)I(f_1)\sigma^2\sigma_{f_1}^4\tau_2^2.$$

(iii) We know that the power of the test is defined by

$$1-\beta \;:=\; \text{Prob}\left[\text{rejecting } \;\mathcal{H}_f^{(n)}(\theta) \;/\; \mathcal{H}_f^{(n)}(\theta^{(n)})\right] \;=\; \text{Prob}\left[\underline{Q}_{f_1} > \chi_{1,1-\alpha}^2 \;/\; \tau_2 \neq 0\right]$$

where β is the second species risk and defined by

$$\text{Prob}\left[\underline{Q}_{f_1} < \chi_{1,1-\alpha}^2 \;/\; \tau_2 \neq 0\right] = F(\chi_{1,1-\alpha}^2, \lambda_{f_1}).$$

Hence, the power of the test is $1 - F(\chi_{1,1-\alpha}^2, \lambda_{f_1})$. □

The Gaussian versions of $\underline{Q}_{f_1}$ is

$$\underline{Q}_{\mathcal{N}} = \frac{a^3}{T-2}\left[n^{\frac{-1}{2}} \sum_{i=1}^{n}\sum_{t=3}^{T} Z_{i,t}Z_{i,t-1}Z_{i,t-2}\right]^2. \tag{3.4}$$

Unfortunately, this test statistic needs f_1 to be specified as a standardized Gaussian one, so the parameter a also has to be given. In the next, we will show that an appropriate version remains asymptotically valid under arbitrary f_1 with finite variance and optimal under Gaussian one (pseudo-Gaussian test).

3.2. Pseudo-Gaussian tests

The Gaussian central sequence $\Delta_{\mathcal{N};2}^{(n)}(\theta)$ allows obtaining asymptotically optimal tests under $f_1 = f_{\mathcal{N}}$, as well as efficient detection of panel bilinear models, in the parametric Gaussian model characterized by Gaussian disturbances. Extending the validity of the Gaussian optimal test to general densities g_1 in a broad class of densities is of course highly desirable.

Let us show that this is indeed possible and that a slight modification, $\Delta^{*(n)}_{\mathcal{N};2}$, say, of the efficient central sequence $\Delta^{(n)}_{\mathcal{N};2}$ leads to a *pseudo-Gaussian test* which remaining valid when the actual density g_1 belongs to the class $\mathcal{F}^{(2)}_A$ of all densities in $\mathcal{F}_A$ with *finite* variance. Define

$$\Delta^{*(n)}_{\mathcal{N};2}(\theta) = n^{\frac{-1}{2}}\sigma a \sum_{i=1}^{n}\sum_{t=3}^{T}(Z_{i,t}-m_1^{(n)})(Z_{i,t-1}-m_1^{(n)})(Z_{i,t-2}-m_1^{(n)}),$$

where $m_1^{(n)} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T} Z_{i,t}$ is a $\sqrt{n}$-consistent estimator, under $\mathbf{P}^{(n)}_{\theta;g_1}$, of $\mu_1(g_1) := \int_{\mathbb{R}} z g_1(z)dz$. Decomposing $Z_{i,t}-m_1^{(n)}$ into $(Z_{i,t}-\mu_1(g_1)) + (\mu_1(g_1)-m_1^{(n)})$, then, it is easy to check that under $\mathbf{P}^{(n)}_{\theta;g_1}$, as $n\to\infty$,

$$\Delta^{*(n)}_{\mathcal{N};2}(\theta) = n^{\frac{-1}{2}}\sigma a \sum_{i=1}^{n}\sum_{t=3}^{T}(Z_{i,t}-\mu_1(g_1))(Z_{i,t-1}-\mu_1(g_1))(Z_{i,t-2}-\mu_1(g_1)) + o_p(1).$$

Then, still under $\mathbf{P}^{(n)}_{\theta;g_1}$, $\Delta^{*(n)}_{\mathcal{N};2}(\theta)$ is asymptotically normal with zero mean and variance

$$\Gamma^{*}_{\mathcal{N};g_1;22} = a^2\sigma^2(T-2)\sigma^6_{g_1},$$

where $\sigma^2_{g_1} := \int_{\mathbb{R}}(z-\mu_1(g_1))^2 g_1(z)dz.$

On the other hand, it is easy to see that, under $\mathbf{P}^{(n)}_{\theta^{(n)};g_1}$, $\Delta^{*(n)}_{\mathcal{N};2}(\theta)$ and the log-likelihood $\Lambda^{(n)}_{\theta^{(n)}/\theta;g_1}$ are jointly binormal; the desired result then follows from a routine application of Le Cam's Third Lemma.

A pseudo-Gaussian test may then be based on a statistic of the form

$$\begin{aligned} Q^{*}_{\mathcal{N};g_1}(\theta) &:= (\Gamma^{*}_{\mathcal{N};g_1;22}(\theta))^{-1}\Delta^{*(n)^2}_{\mathcal{N};2}(\theta) \\ &:= \frac{1}{(T-2)\sigma^6_{g_1}}\left[n^{\frac{-1}{2}}\sum_{i=1}^{n}\sum_{t=3}^{T}(Z_{i,t}-m_1^{(n)})(Z_{i,t-1}-m_1^{(n)})(Z_{i,t-2}-m_1^{(n)})\right]^2. \end{aligned} \tag{3.5}$$

In practice, the pseudo-Gaussian test will be based on

$$Q^{\dagger}_{\mathcal{N}} := \frac{1}{(T-2)s^6}\left[n^{\frac{-1}{2}}\sum_{i=1}^{n}\sum_{t=3}^{T}(Z_{i,t}-m_1^{(n)})(Z_{i,t-1}-m_1^{(n)})(Z_{i,t-2}-m_1^{(n)})\right]^2,$$

where $s^2 = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(Z_{i,t}-m_1^{(n)})^2$ is the empirical variance of the $(Z_{i,t}-m_1^{(n)})$'s.

Showing that, under $\mathbf{P}^{(n)}_{\theta;g_1}$, $Q^{\dagger}_{\mathcal{N}} - Q^{*}_{\mathcal{N};g_1}(\theta) = o_p(1).$, as $n\to\infty$, we thus have the following result.

Proposition 3.2. *Let* $g_1 \in \mathcal{F}^{(2)}_A$. *Then,*

(i) $Q^{\dagger}_{\mathcal{N}}$ *is asymptotically central chi-square with* 1 *degree of freedom under* $\boldsymbol{P}^{(n)}_{\theta;g_1}$, *and asymptotically noncentral chi-square, still with* 1 *degree of freedom and with noncentrality parameter* $\lambda_{\mathcal{N}} := (T-2)\sigma^2_g\tau^2_2$ *under* $\boldsymbol{P}^{(n)}_{\theta^{(n)};g_1}$;

(ii) *The sequence of tests rejecting the null hypothesis* $\bigcup_{g_1 \in \mathcal{F}_A^{(2)}} \bigcup_{\sigma \in \mathbb{R}_+^*} \left\{ \boldsymbol{P}^{(n)}_{\sigma^2,0;g_1} \right\}$ *whenever* $Q^{\dagger}_{\mathcal{N}} > \chi^2_{1,1-\alpha}$, *is locally asymptotically most stringent, at asymptotic level* α, *against alternatives of the form* $\bigcup_{\sigma \in \mathbb{R}_+^*} \bigcup_{b \in \mathbb{R}} \left\{ \boldsymbol{P}^{(n)}_{\sigma^2,b;f_{\mathcal{N}}} \right\}$;

(iii) *The asymptotic power under* $\boldsymbol{P}^{(n)}_{\theta^{(n)};g_1}$ *is* $1 - F(\chi^2_{1,1-\alpha}, \lambda_{\mathcal{N}})$.

4. OPTIMAL RANK TESTS

We start by describing the group invariance structure of the testing problem considered. Then we introduce (and study the properties of) rank-based versions of the central sequences. This will allow us to develop the resulting (optimal) rank tests and to derive their asymptotic properties. The general results of Hallin and Werker (2003) indicate that semiparametrically efficient and rank-based procedures have been established in relation with ranks that are being maximal invariants under model-generating groups of transformations. It is clearly that the null hypothesis $\mathcal{H}_0^{(n)}$ is invariant under the group $(\mathcal{G}^{(nT)}, \star)$, such as for any transformation $\mathcal{G}_h$ of $\mathbb{R}^{nT}$ we define $\mathcal{G}_h(Y_{11}, ..., Y_{nT}) := (h(Y_{11}), ..., h(Y_{nT}))$, where $y \mapsto h(y)$ is continuous and monotone increasing and $\lim_{y \to \pm\infty} h(y) = \pm\infty$. The invariance principle therefore suggests restricting to tests that are invariant with respect to this group. The maximal invariant associated with $(\mathcal{G}^{(nT)}, \star)$ is the rank $R^{(n)} := \left(R^{(n)}_{1,1}, ..., R^{(n)}_{n,T}\right)$, where $R^{(n)}_{i,t}$ denotes the rank of $Z^{(n)}_{i,t}$ among $\left(Z^{(n)}_{1,1}, ..., Z^{(n)}_{n,T}\right)$. It is easy to check that $(\mathcal{G}^{(nT)}, \star)$ is actually a generating group for the null hypothesis $\mathcal{H}_0^{(n)}$. As a direct corollary, rank tests are distribution-free under the whole null hypothesis. This explains why rank tests will be validity-robust.

4.1. Rank-based versions of central sequences

According to Hallin and Werker (2003) [21] and under the LAN property with efficient central sequence $\Delta^{(n)}_{f_1;2}$, an efficient semiparametric inference obtained conditioning $\Delta^{(n)}_{f_1;2}$ by the rank vector $R^{(n)}$, under the null hypothesis

$$\underset{\sim}{\Delta}^{(n)}_{f_1;2} := E\left[\Delta^{(n)}_{f_1;2} \mid R^{(n)}\right]. \tag{4.1}$$

The conditional definition (4.1) of $\Delta^{(n)}_{f_1;2}$ gives a statistic based on the ranks of exact scores, thus *Hájek's projection theorem* establishes the asymptotic equivalence between a non-parametric statistic and its parametric counterpart (for more details, consult the book of Hájek, Šidák and Sen (1999) [16]).

To combine validity-robustness/invariance with Le Cam optimality at density f_1, we introduce rank-based versions of the central sequence that appear in the LAN property above (Proposition 2.1).

$$\underset{\sim}{\Delta}^{(n)}_{f_1;2} := n^{\frac{-1}{2}} \sigma \sum_{i=1}^{n} \sum_{t=3}^{T} \left\{ \varphi_{f_1}\left(\frac{R^{(n)}_{i,t}}{N+1}\right) F_1^{-1}\left(\frac{R^{(n)}_{i,t-1}}{N+1}\right) F_1^{-1}\left(\frac{R^{(n)}_{i,t-2}}{N+1}\right) - \overline{m}_{f_1} \right\} \tag{4.2}$$

with $N = n(T-2)$, $\varphi_{f_1} := \Phi_{f_1} \circ F_1^{-1}$ and

$$\overline{m}_{f_1} := \frac{1}{N(N-1)(N-2)} \sum\sum\sum_{1 \leqslant t_1 \neq t_2 \neq t_3 \leqslant N} \varphi_{f_1}\Big(\frac{t_1}{N+1}\Big) F_1^{-1}\Big(\frac{t_2}{N+1}\Big) F_1^{-1}\Big(\frac{t_3}{N+1}\Big).$$

Let

$$\begin{aligned} s_{f_1}^{(n)^2} := {} & \frac{1}{N(N-1)(N-2)} \sum\sum\sum_{1 \leqslant t_1 \neq t_2 \neq t_3 \leqslant N} \Big[\varphi_{f_1}\Big(\frac{t_1}{N+1}\Big) F_1^{-1}\Big(\frac{t_2}{N+1}\Big) F_1^{-1}\Big(\frac{t_3}{N+1}\Big)\Big]^2 \\ & + \frac{2}{N(N-1)(N-2)(N-3)} \\ & \times \sum\sum\sum\sum_{1 \leqslant t_1 \neq t_2 \neq t_3 \neq t_4 \leqslant N} \varphi_{f_1}\Big(\frac{t_1}{N+1}\Big) \varphi_{f_1}\Big(\frac{t_2}{N+1}\Big) F_1^{-1}\Big(\frac{t_2}{N+1}\Big) \Big[F_1^{-1}\Big(\frac{t_3}{N+1}\Big)\Big]^2 F_1^{-1}\Big(\frac{t_4}{N+1}\Big) \\ & + \frac{2}{N(N-1)(N-2)(N-3)(N-4)} \\ & \times \sum\sum\sum\sum\sum_{1 \leqslant t_1 \neq t_2 \neq t_3 \neq t_4 \neq t_5 \leqslant N} \varphi_{f_1}\Big(\frac{t_1}{N+1}\Big) F_1^{-1}\Big(\frac{t_2}{N+1}\Big) F_1^{-1}\Big(\frac{t_3}{N+1}\Big) \varphi_{f_1}\Big(\frac{t_3}{N+1}\Big) \\ & \times F_1^{-1}\Big(\frac{t_4}{N+1}\Big) F_1^{-1}\Big(\frac{t_5}{N+1}\Big) \\ & + \frac{N-5}{N(N-1)(N-2)(N-3)(N-4)(N-5)} \\ & \times \sum\sum\sum\sum\sum\sum_{1 \leqslant t_1 \neq t_2 \neq t_3 \neq t_4 \neq t_5 \neq t_6 \leqslant N} \varphi_{f_1}\Big(\frac{t_1}{N+1}\Big) F_1^{-1}\Big(\frac{t_2}{N+1}\Big) F_1^{-1}\Big(\frac{t_3}{N+1}\Big) \\ & \times \varphi_{f_1}\Big(\frac{t_4}{N+1}\Big) F_1^{-1}\Big(\frac{t_5}{N+1}\Big) F_1^{-1}\Big(\frac{t_6}{N+1}\Big) - N\overline{m}_{f_1}^2. \end{aligned}$$

Define the cross-information coefficients $\mathcal{I}(f_1, g_1)$ and $\sigma(f_1, g_1)$ as

$$\mathcal{I}(f_1, g_1) := \int_0^1 \Phi_{f_1}(F_1^{-1}(u)) \Phi_{g_1}(G_1^{-1}(u)) du \quad \text{and} \quad \sigma(f_1, g_1) := \int_0^1 F_1^{-1}(v) G_1^{-1}(v) dv,$$

we then have, for the rank-based $\underset{\sim}{\Delta}{}^{(n)}_{f_1;2}$, the following asymptotic representation result.

Proposition 4.1. *Let f_1 and $g_1 \in \mathcal{F}_A$. Then, as $n \to \infty$ and fixed T,*

(i) *Under $\boldsymbol{P}^{(n)}_{\theta;g_1}$,*

$$\begin{aligned} \underset{\sim}{\Delta}{}^{(n)}_{f_1;2} & := E^{(n)}_{g_1}\big[\Delta^{(n)}_{f_1;2} \mid R^{(n)}_{1,1}, \ldots, R^{(n)}_{n,T}\big] + o_{L^2}(1) \\ & = \Delta^{*(n)}_{f_1,g_1;2} + o_{L^2}(1), \end{aligned} \tag{4.3}$$

with (denoting by G_1 the distribution function associated with g_1)

$$\Delta^{*(n)}_{f_1,g_1;2} := n^{\frac{-1}{2}} \sigma \sum_{i=1}^{n} \sum_{t=3}^{T} \varphi_{f_1}\big(G_1(Z_{i,t})\big) F_1^{-1}\big(G_1(Z_{i,t-1})\big) F_1^{-1}\big(G_1(Z_{i,t-2})\big); \tag{4.4}$$

(ii) *Still under $\boldsymbol{P}^{(n)}_{\theta;g_1}$, $\underset{\sim}{\Delta}{}^{(n)}_{f_1;2}$ has zero mean and variance $\Gamma^{*(n)}_{f_1;22} := \sigma^2(T-2)s_{f_1}^{(n)^2} = \Gamma^{*}_{f_1;22} + o(1)$, where $\Gamma^{*}_{f_1;22} := (T-2)I(f_1)\sigma^2\sigma^4_{f_1}$;*

(iii) *$\Delta^{*(n)}_{f_1,g_1;2}$ is asymptotically normal, with zero mean under $\boldsymbol{P}^{(n)}_{\theta;g_1}$, mean $(T-2)\mathcal{I}(f_1,g_1)\sigma^2(f_1,g_1)\sigma^2\tau_2$ under $\boldsymbol{P}^{(n)}_{\theta^{(n)};g_1}$ and variance $\Gamma^{*}_{f_1;22}$ under both.*

Proof: The proof of Part (i) of the proposition follows along the same lines as in Hallin *et al.* (1985) [18], and therefore is omitted. Part (ii) is obtained by direct computation. As for Part (iii), under $\mathbf{P}^{(n)}_{\theta;g_1}$, the result straightforwardly follows from classical central limit theorem. On the other hand, it is easy to see that, still under $\mathbf{P}^{(n)}_{\theta^{(n)};g_1}$, $\Delta^{*(n)}_{f_1,g_1;2}$ and the log-likelihood $\Lambda^{(n)}_{\theta^{(n)}/\theta;g_1}$ are jointly binormal; the desired result then follows from a routine application of Le Cam's Third Lemma. □

4.2. Optimal rank tests

The rank-based version of the quadratic statistic is given by

$$
\begin{aligned}
\underset{\sim}{Q}_{f_1} &:= (\Gamma^{*(n)}_{f_1;22})^{-1}\underset{\sim}{\Delta}^{(n)^2}_{f_1;2} \\
&= \frac{1}{(T-2)s^{(n)^2}_{f_1}}\left[n^{\frac{-1}{2}}\sum_{i=1}^{n}\sum_{t=3}^{T}\left\{\varphi_{f_1}\Big(\frac{R^{(n)}_{i,t}}{N+1}\Big)F_1^{-1}\Big(\frac{R^{(n)}_{i,t-1}}{N+1}\Big)F_1^{-1}\Big(\frac{R^{(n)}_{i,t-2}}{N+1}\Big)-\overline{m}_{f_1}\right\}\right]^2,
\end{aligned}
\tag{4.5}
$$

we then have the following general result.

Proposition 4.2. *Let f_1 and $g_1 \in \mathcal{F}_A$. Then, for any $\tau=(\tau_1,\tau_2)' \in \mathbb{R}^2$, as $n\to\infty$ and for all fixed T,*

- **(i)** $\underset{\sim}{Q}_{f_1}$ *is asymptotically central chi-square with 1 degree of freedom under $\boldsymbol{P}^{(n)}_{\theta;g_1}$, and asymptotically noncentral chi-square, still with 1 degree of freedom and with noncentrality parameter*

$$\lambda_{f_1,g_1} := (T-2)\mathcal{I}^2(f_1,g_1)\sigma^4(f_1,g_1)\sigma^2\tau_2^2/I(f_1)\sigma^4(f_1)$$

  under $\boldsymbol{P}^{(n)}_{\theta^{(n)};g_1}$;

- **(ii)** *The sequence of tests rejecting the null hypothesis $\bigcup_{g_1\in\mathcal{F}_A}\bigcup_{\sigma^2}\left\{\boldsymbol{P}^{(n)}_{\sigma^2,0;g_1}\right\}$ whenever $\underset{\sim}{Q}_{f_1} > \chi^2_{1,1-\alpha}$, is locally asymptotically most stringent, at asymptotic level α, against alternatives of the form $\bigcup_{\sigma\in\mathbb{R}^*_+}\bigcup_{b\in\mathbb{R}}\left\{\boldsymbol{P}^{(n)}_{\sigma^2,b;f_1}\right\}$;*
- **(iii)** *The asymptotic power under $\boldsymbol{P}^{(n)}_{\theta^{(n)};f_1}$ is $1-F(\chi^2_{1,1-\alpha},\lambda_{f_1,g_1})$.*

4.3. Examples of non-parametric statistics

The quadratic statistic $\underset{\sim}{Q}_{f_1}$ is a non-parametric statistic that depends only on the determining of the score function f_1 and provides general form for the optimal rank tests of the null hypothesis of randomness.

The three most important particular cases for the rank test statistic presented are the *van der Waerden* (normal score), the *Wilcoxon* (logistic score) and the *Laplace* (double-exponential score) test statistics, which are respectively optimal at normal, logistic and double-exponential distributions.

(**i**) The ***van der Waerden's*** test statistic is given by

$$\underset{\sim vdW}{Q} := \frac{a^2}{(T-2)s^{(n)^2}_{f_{\mathcal{N}}}} \underset{\sim vdW}{\Delta}^{(n)^2},$$

with

$$\underset{\sim vdW}{\Delta^{(n)}} = n^{\frac{-1}{2}} \sum_{i=1}^{n}\sum_{t=3}^{T}\left\{\Psi^{-1}\left(\frac{R^{(n)}_{i,t}}{N+1}\right)\Psi^{-1}\left(\frac{R^{(n)}_{i,t-1}}{N+1}\right)\Psi^{-1}\left(\frac{R^{(n)}_{i,t-2}}{N+1}\right) - \overline{m}_{vdW}\right\} \tag{4.6}$$

and

$$\overline{m}_{f_{\mathcal{N}}} = \frac{1}{N(N-1)(N-2)} \sum\sum\sum_{1\leqslant t_1\neq t_2\neq t_3\leqslant N} \Psi^{-1}\left(\frac{t_1}{N+1}\right)\Psi^{-1}\left(\frac{t_2}{N+1}\right)\Psi^{-1}\left(\frac{t_3}{N+1}\right),$$

where Ψ is the standard normal distribution function.

(**ii**) The ***Wilcoxon's*** test statistic is given by

$$\underset{\sim W}{Q} := \frac{1}{(T-2)bs^{(n)^2}_{l}} \underset{\sim W}{\Delta}^{(n)^2},$$

with

$$\underset{\sim W}{\Delta^{(n)}} = n^{\frac{-1}{2}} \sum_{i=1}^{n}\sum_{t=3}^{T}\left\{\left(2\frac{R^{(n)}_{i,t}}{N+1}-1\right)\log\left(\frac{R^{(n)}_{i,t-1}}{N+1-R^{(n)}_{i,t-1}}\right)\log\left(\frac{R^{(n)}_{i,t-2}}{N+1-R^{(n)}_{i,t-2}}\right) - \overline{m}_{l}\right\} \tag{4.7}$$

and

$$\overline{m}_{l} = \frac{1}{N(N-1)(N-2)} \sum\sum\sum_{1\leqslant t_1\neq t_2\neq t_3\leqslant N} \left(\frac{2t_1}{N+1}-1\right)\log\left(\frac{t_2}{N+1-t_2}\right)\log\left(\frac{t_3}{N+1-t_3}\right).$$

(**iii**) The ***Laplace's*** test statistic is given by

$$\underset{\sim \mathcal{L}}{Q} := \frac{d^2}{(T-2)s^{(n)^2}_{\mathcal{D}e}} \underset{\sim \mathcal{L}}{\Delta}^{(n)^2},$$

with

$$\underset{\sim \mathcal{L}}{\Delta^{(n)}} = n^{\frac{-1}{2}} \sum_{i=1}^{n}\sum_{t=3}^{T}\left\{sign\left(F_1^{-1}\left(\frac{R^{(n)}_{i,t}}{N+1}\right)\right)F_1^{-1}\left(\frac{R^{(n)}_{i,t-1}}{N+1}\right)F_1^{-1}\left(\frac{R^{(n)}_{i,t-2}}{N+1}\right) - \overline{m}_{\mathcal{D}e}\right\} \tag{4.8}$$

and

$$\overline{m}_{\mathcal{D}e} = \frac{1}{N(N-1)(N-2)} \sum\sum\sum_{1\leqslant t_1\neq t_2\neq t_3\leqslant N} sign\left(F_1^{-1}\left(\frac{t_1}{N+1}\right)\right)F_1^{-1}\left(\frac{t_2}{N+1}\right)F_1^{-1}\left(\frac{t_3}{N+1}\right),$$

where F_1 is the distribution function of the double-exponential and

$$F_1^{-1}(u) = \begin{cases} d\log(2u) & if\ \ 0<u\leqslant \frac{1}{2} \\ -d\log(2-2u) & if\ \ \frac{1}{2}\leqslant u<1. \end{cases}$$

5. ASYMPTOTIC RELATIVE EFFICIENCIES

In order to compare the performance of the parametric and non-parametric tests presented, we calculate the Asymptotic Relative Efficiencies (AREs) of rank based tests with respect to the Pseudo-Gaussian one. The results obtained are satisfactory. Hence, under $\mathbf{P}^{(n)}_{\theta^{(n)};g_1}$, for any g_1 and for different scores f_1, the asymptotic relative efficiencies of $\underset{\sim f_1}{Q}$ with respect to $\underline{Q}_{\mathcal{N}}$ are

$$\text{(5.1)} \qquad \begin{aligned} ARE_{g_1}(\underset{\sim f_1}{Q}/\underline{Q}_{\mathcal{N}}) &= \left(\frac{\lambda_{f_1,g_1}}{\lambda_{\mathcal{N}}}\right)^2 \\ &= \left(\frac{\mathcal{I}^2(f_1,g_1)\sigma^4(f_1,g_1)}{\sigma^2_{g_1}\sigma^4_{f_1}I(f_1)}\right)^2. \end{aligned}$$

Table 1 gives the numerical values of (5.1) for $\underset{\sim f_1}{Q} = \underset{\sim vdW}{Q}, \underset{\sim W}{Q}, \underset{\sim La}{Q}, \underset{\sim t_5}{Q}, \underset{\sim s\mathcal{N}(2)}{Q}$ and $\underset{\sim st_5(2)}{Q}$ under densities g_1 that are normal, Logistic, Double exponential, Student-t_5, Skew-normal $s\mathcal{N}(2)$ and Skew-Student $st_5(2)$.

Note that for $f_1 = vdW$ these values are always greater than one, i.e., the van der Waerden test (vdW) always has an efficiency greater than or equal to one, the equality being realized only if the density underlying g is itself a Gaussian density ($\mathcal{N}$), which means that rank based tests are asymptotically more powerful than Gaussian tests (this result is proved in many cases, see for example, Chernoff and Savage (1958) [7] and Hallin (1993) [17] for ARMA models). Note also that each value is maximum in its corresponding column. Thus, at each of the densities, non-parametric tests perform better, compared to the Pseudo-Gaussian test.

Table 1: Asymptotic relative efficiencies of some rank tests compared to their Pseudo-Gaussian counterpart.

Scores f_1 \ Actual density g_1	$\mathcal{N}$	l	$\mathcal{D}e$	t_5	$s\mathcal{N}(2)$	$st_5(2)$
Van der Waerden	1.0000	1.1723	1.5244	1.3435	1.6328	1.7262
Wilcoxon	0.9347	1.2026	2.3421	1.5002	1.9782	1.7822
Laplace	0.4275	1.1337	4.0000	1.0349	1.5433	1.6889
Student-t_5	0.8160	1.1569	2.7812	1.5625	1.8922	1.9501
Skew-normal $s\mathcal{N}(2)$	0.9520	1.0989	1.5633	1.1490	2.2301	2.3301
Skew-Student $st_5(2)$	0.5179	0.9734	1.9331	1.2150	1.7325	3.0133

6. SIMULATION

To enhance the interpretation and validity of the theoretical results of the previous sections, we present a simulation experiment using R-programming. The purpose of this section is to evaluate the performance of the proposed tests, at asymptotic level $\alpha = 5\%$.

We simulated several $BLP(0,0,2,1)$ panel data described by

$$X_{i,t} = bX_{i,t-2}\varepsilon_{i,t-1} + \varepsilon_{i,t} \qquad i = 1,2,...,100, \quad t = 1,2,...,12, \tag{6.1}$$

where:

- $b = 0$ for null hypothesis, and $b = 0.05, 0.1, 0.15, 0.2$ for increasingly severe alternatives;
- The $(\varepsilon_{i,t})$'s are i.i.d. with a symmetric density - Gaussian ($\mathcal{N}$), logistic (l), double exponential ($\mathcal{D}e$), Student with $\nu = 5$ degrees of freedom (t_5) — or with an asymmetric density — the skew-normal $s\mathcal{N}(\delta)$ and skew-Student $st_5(\delta)$ densities [4] (both with skewness parameter value $\delta = 2$).

We performed the simulations for $n = 100$ and $T = 12$. In each case we generated 2500 independent samples of size $N = n(T-2) = 1000$ from (6.1).

For each replication, we performed the following tests at asymptotic level $\alpha = 5\%$: the pseudo-Gaussian test based on $Q^{\dagger}_{\mathcal{N}}$, the van der Waerden test based on $\underset{\sim}{Q}_{vdW}$, the Wilcoxon test based on $\underset{\sim}{Q}_{W}$, the Laplace test based on $\underset{\sim}{Q}_{L}$, the rank tests based on Student with 5 degrees of freedom and data-driven skew-Student $st_{\hat{\nu}}(\hat{\delta})$ scores.

Rejection frequencies are reported in Table 2 and they amply confirm the excellent overall performances of our rank-based procedure with data-driven scores. It also appears from the skew-normal and skew-Student simulations that asymmetry significantly improves the superiority of rank tests over the pseudo-Gaussian one.

7. CONCLUSION

The problem of testing the null hypothesis of a randomness against *first-order superdiagonal panel model* $BLP(0,0,2,1)$ (in large n and small T) is considered for specified and unspecified error density. Optimal parametric and pseudo-Gaussian procedures are derived based on the Local Asymptotic Normality property. Moreover, the pseudo-Gaussian test appears to have quite poor performances under skewed and heavy-tailed distributions. Therefore a rank-based version of the test is considered. Particular cases such as van der Waerden, Wilcoxon, Laplace and data-driven scores are given. These tests exhibit remarkably high ARE values with respect to their pseudo-Gaussian counterpart. Simulations confirm the excellent overall performances of the proposed tests.

[4] See, for instance, Azzalini and Capitanio (2003) [2] for a definition of skew-normal and skew-Student densities.

Table 2: Rejection frequencies (out of 2500 replications), for $b = 0$ (null hypothesis) and various non-zero values of b (alternative hypotheses), with error density g_1 that is Gaussian ($\mathcal{N}$), logistic (l), double exponential ($\mathcal{D}e$), Student (t_5), skew-normal ($s\mathcal{N}(2)$) and skew-Student t5 ($st_5(2)$) of the pseudo-Gaussian and rank based (based on van der Waerden, Wilcoxon, Laplace, Student-t_5 and data-driven scores) procedures.

Underlying densities g_1	Test	b				
		0	0.05	0.1	0.15	0.2
Normal	Pseudo Gaussien	0.0520	0.2236	0.7224	0.9680	0.9996
	Van der Waerden	0.0512	0.2448	0.6844	0.9564	1.0000
	Wilcoxon	0.0508	0.2280	0.7400	0.9640	1.0000
	Laplace	0.0512	0.2160	0.6928	0.8840	0.9992
	Student-t_5	0.0496	0.2360	0.6560	0.9760	1.0000
	Data-Driven	0.0524	0.2800	0.7400	0.9760	1.0000
Logistic	Pseudo Gaussien	0.0464	0.2400	0.7144	0.9632	0.9992
	Van der Waerden	0.0488	0.2688	0.7204	0.9844	0.9996
	Wilcoxon	0.0520	0.3044	0.7880	0.9620	0.9980
	Laplace	0.0496	0.2960	0.7320	0.9840	0.9980
	Student-t_5	0.0560	0.2488	0.7640	0.9840	0.9996
	Data-Driven	0.0500	0.3240	0.8360	0.9920	1.0000
Double exponential	Pseudo Gaussien	0.0524	0.2236	0.6908	0.9544	0.9972
	Van der Waerden	0.0476	0.2324	0.7820	0.9956	0.9888
	Wilcoxon	0.0492	0.3720	0.8412	0.9884	0.9992
	Laplace	0.0520	0.4924	0.9080	0.9960	1.0000
	Student-t_5	0.0484	0.3920	0.8800	0.9920	1.0000
	Data-Driven	0.0480	0.3760	0.8760	0.9520	1.0000
Student-t_5	Pseudo Gaussien	0.0496	0.3248	0.8768	0.9932	0.9996
	Van der Waerden	0.0488	0.3044	0.8660	0.9924	1.0000
	Wilcoxon	0.0492	0.4964	0.9248	0.9732	0.9989
	Laplace	0.0488	0.4560	0.8840	0.9880	0.9996
	Student-t_5	0.0476	0.4640	0.9560	0.9960	1.0000
	Data-Driven	0.0540	0.4960	0.9720	1.0000	1.0000
Skew-normal $s\mathcal{N}(2)$	Pseudo Gaussien	0.0496	0.1264	0.4572	0.7900	0.9612
	Van der Waerden	0.0464	0.1328	0.4112	0.8084	0.9488
	Wilcoxon	0.0468	0.1440	0.4560	0.8240	0.9440
	Laplace	0.0492	0.2120	0.4824	0.7244	0.8680
	Student-t_5	0.0432	0.1760	0.4120	0.7360	0.9240
	Data-Driven	0.0460	0.2080	0.5360	0.8080	0.9400
Skew-Student $st_5(2)$	Pseudo Gaussien	0.0480	0.2240	0.6800	0.9392	0.9904
	Van der Waerden	0.0524	0.2368	0.7200	0.9240	0.9888
	Wilcoxon	0.0488	0.3120	0.7284	0.9688	0.9992
	Laplace	0.0540	0.3160	0.6800	0.9124	0.9640
	Student-t_5	0.0504	0.2840	0.7280	0.9440	0.9920
	Data-Driven	0.0484	0.3480	0.8360	0.9720	0.9960

ACKNOWLEDGMENTS

The authors wish to express their special thanks to the reviewers for their contribution to improve the quality of this paper.

REFERENCES

[1] Akharif, A. and Hallin, M. (2003). Efficient detection of random coefficients in autoregressive models, *Annals of Statistics*, **31**(2), 675–704.

[2] Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution, *Journal of the Royal Statistical Society, Series B*, **65**, 367–389.

[3] Benghabrit, Y. and Hallin, M. (1992). Optimal rank-based tests for against first-order superdiagonal bilinear dependence against diagonal bilinear dependence, *Journal of Statistical Planning and Inference*, **32**(1), 45–61.

[4] Bennala, N.; Hallin, M. and Paindaveine, D. (2012). Pseudo-Gaussian and rank-based optimal tests for random individual effects in large n small T panels, *Journal of Econometrics*, **170**, 50–67.

[5] Bouzaachane, K. (2006). *Modèles de séries chronologiques bilinéaires: estimations, algorithmes et applications*, PhD thesis, University of Mohammed V-Agdal, Morocco.

[6] Cassart, D.; Hallin, M. and Paindaveine, D. (2011). A class of optimal tests for symmetry based on local Edgeworth approximations, *Bernoulli*, **17**, 1063–1094.

[7] Chernoff, H. and Savage, I.R. (1958). Asymptotic normality and efficiency of certain nonparametric tests, *Annals of Mathematical Statistics*, **29**, 972–994.

[8] El Melhaoui, S. and Allal, J. (2006). Optimal rank-based detection of exponential component in autoregressive models, *Journal of Nonparametric Statistics*, **18**(7), 431–447.

[9] Fihri, M.; Mellouk, A. and Akharif, A. (2016). Rank and signed-rank tests for random coefficient regression model, *Journal of Statistics Applications and Probability*, **5**, 233–247.

[10] Ghosh, S. (1999). *Asymptotics, Nonparametrics, and Time Series*, CRC Press, 1st Edition.

[11] Grahn, T. (1995). A conditional least squares approach to bilinear time series, *Journal of Time Series Analysis*, **16**, 509–529.

[12] Granger, C.W.J. and Andersen, A.P. (1978). *An Introduction to Bilinear Time Series Models*, Vandenhoeck and Ruprecht, Gettingen.

[13] Guegan, D. (1981). Étude d'un modèle non linéaire, le modèle superdiagonal d'ordre 1, *C.R. Acad. Sci. Paris, Série 1, Math.*, **293**, 95–99.

[14] Guegan, D. (1988). *Modèles bilinéaires et polynomiaux de sèries chronologiques: étude probabiliste et analyse statistique*, Thèse d'État, Universitè Joseph Fourier, Grenoble.

[15] Guegan, D. and Pham, D.T. (1992). Power of the score test against bilinear time series models, *Statistica Sinica*, **2**, 157–170.

[16] Hájek, J.; Šidák, Z. and Sen, P.K. (1999). *Theory of Rank Tests*, Academic Press, New York – London, 2nd edition.

[17] Hallin, M. (1993). On the Pitman non-admissibility of correlogram-based methods, *Journal of Time Series Analysis*, **16**, 607–612.

[18] Hallin, M.; Ingenbleek, J.F. and Puri, M.L. (1985). Linear serial rank tests for randomness against ARMA alternatives, *The Annals of Statistics*, **13**, 1156–1181.

[19] Hallin, M. and Mèlard, G. (1988). Rank-Based Tests for Randomness against First-Order Serial Dependence, *Journal of the American Statistical Association*, **83**, 1117–1128.

[20] Hallin, M. and Werker, B.J.M. (1999). *Optimal testing for semiparametric AR models: from Gaussian Lagarange multipliers to autoregression rank scores and adaptive tests.* In "Asymptotics, Nonparametrics and Time Series" (S. Ghosh, Ed.), Dekker, New York, 295–350.

[21] Hallin, M. and Werker, B.J.M. (2003). Semi-parametric efficiency, distribution-freeness and invariance, *Bernoulli*, **9**(1), 137–165.

[22] Hwang, S.Y. and Basawa, I.V. (1993). Asymptotic optimal inference for a class of nonlinear time series models, *Stochastic Processes and their Applications*, **46**, 91–113.

[23] Le Cam, L.M. (1986). *Asymptotic Methods in Statistical Decision Theory*, Springer Verlag, New York.

[24] Le Cam, L.M. and Yang, G.L. (2000). *Asymptotics in Statistics: Some Basic Concepts*, Springer Verlag, New York, 2nd edition.

[25] Liu, J. and Brockwell, P.J. (1988). On the general bilinear time series model, *Journal of Applied Probability*, **25**, 553–564.

[26] Maravall, A. (1983). An application of nonlinear time series forecasting, *Journal of Business and Economic Statistics*, **1**, 66–74.

[27] Phan, T.D. and Tran, L.T. (1980). Quelques résultats sur les modèles bilinéaires de séries chronologiques, *C.R. Acad. Sci. Paris*, **290**, 335–338.

[28] Phan, T.D. and Tran, L.T. (1981). On the first-order bilinear time series model, *Journal of Applied Probability*, **18**, 617–627.

[29] Saikkonen, P. and Luukkonen, R. (1991). Power properties of a time series linearity test against some simple bilinear alternatives, *Statistica Sinica*, **1**, 453–464.

[30] Sesay, S.A.O. and Subba Rao, T. (1991). Difference equations for higher order moments and cumulant for the bilinear time series model BL$(p, 0, p, 1)$, *Journal of Time Series Analysis*, **10**, 159–169.

[31] Subba Rao, T. (1981). On the theory of bilinear time series models, *Journal of the Royal Statistical Society, Series B*, **43**, 244–255.

[32] Subba Rao, T. and Gabr, M.M. (1984). *An Introduclion to Bispectral Analysis and Bilinear Time Series Models*, Lecture Notes in Statistics No. 24, Springer, Berlin.

[33] Swensen, A.R. (1985). The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend, *Journal of Multivariate Analysis*, **16**, 54–70.

[34] Tan, L. and Wang, L. (2015). The LASSO method for bilinear time series models, *Communications in Statistics – Simulation and Computation*, **45**, 1–11.

[35] Van der Vaart, A.W. (1998). *Asymptotic Statistics*, Cambridge University Press, Cambridge.

[36] Weiss, A.A. (1986). ARCH and bilinear time series models: Comparison and combination, *Journal of Business and Economic Statistics*, **4**, 59–70.