2D MA parameters identification using higher-order spatial cumulants

Soumia Kharfouchi*
Faculty of Medicine
Salah Boubnider University
Constantine
Algeria
s_kharfouchi@yahoo.fr

Hanane Beghriche
Faculty of Medicine
Salah Boubnider University
Constantine
Algeria
hanane_beghriche@hotmail.fr

Abstract. In this paper, the almost-sure convergence properties of sample estimates of higher-order spatial statistics are derived. As a practical framework, we address the problem of identification of 2D moving average (MA) models with non-Gaussian errors based on cumulants alone under a nonminimum phase assumption first and on a generalized method of moments approach after. A simulation study verifies the performance of the proposed methods.

Keywords: higher-order spatial statistics, nonminimum phase, 2D FIR, almost sure convergence.

1. Introduction

Most of natural (or real life) signals are signal processing for which the inputs are complex spatial data that exhibit non-Gaussian and / or non-linear behavior or characteristics. They need higher order correlation methods for their analyzing. Higher-order 2D statistics are able to measure periodicity between multiple points separated by given distances. They better refect geometry than the two-point correlation function (i.e., second-order statistics) (see for example; [6], [10], [29], [16]). Also, the extraction of new properties about the spatial data leads to a significant improvement of the predictive modeling performances. So in several cases, there is necessity to use higher-order spatial cumulants.

However, to take full advantage of the several virtues of experimental spatial cumulants, we need to be able to mistake the theoretical cumulants and the experimental (or sample) ones when dealing with a finite spatial data set, and

∗. Corresponding author

establish a strong convergence between them. For this purpose, we study in this work:

- the almost sure convergence of sample third- and forth-order spatial cumulants. Furthermore, for a spatial linear process we discuss the asymptotic normality of the sample estimates.

- as a practical framework of the proposed theoretical results, some methods for identification of one dimensional Moving Average systems of finite order (FIR MA systems) are extended to the 2-D case i.e., to the 2-D FIR MA system.

Our focus will be on higher-order cumulants based methods under non invertibility assumption. Indeed, recent studies in many science and engineering applications have shown that parameter values are in the nonminimum phase range.

A key finding in these studies is that higher order cumulants are necessary for the identification of nonminimum phase systems, implying that the assumption of Gaussian errors must be abandoned. In the 1-D case, theory and computational techniques of linear prediction have received extensive attention in the system theory and signal processing. However, despite the efforts of many researchers, most of this theory has not been extended to the 2-D case. 2D MA representation for random fields is relevant due to the great popularity of their Wold-type decomposition in the regular homogeneous case. Applications of the Wold-type decomposition and 2D MA modeling include modeling natural textures [12], image segmentation and restoration problems [22], image modeling and retrieval [24] and adaptive processing radar systems [13].

The identification problem is often recasted as one of finding a unique solution to a system of non-linear equations. Higher order cumulants are useful for identification of possibly nonminimum phase systems because the Fourier transform of the k-th order joint cumulants of a k-th order stationary random fieldnX is the k-th order polyspectrum of X : Indeed, provided that the k-th order cumulant of the errors exists and is non-zero for $k \geqslant 3$, one can recover the phase function from any k-th order spectrum, see [23] Lemma 1, [15], [14] for 1D case and [32] and [33] for the multidimensional one. Hence, two approaches for the identification of non-Gaussian 2D MA processes parameters are proposed. The first approach is based on experimental cumulants alone and without imposing minimum phase. While a generalized method of moments (GMM) is used in the second approach.

Throughout, we take $(Xt)_{t\in\mathbb{Z}^d}$ to be a real-valued k-th-order weakly stationary random field having mean zero. Its k-th-order joint moment is defined for $t, v^1, v^2, ..., v^{k-1}$ in $\mathbb{Z}^d$ by $r_k(v^1; v^2; ...; v^{k-1}) = EX_tX_{t+v^1}...X_{t+v^{k-1}}$, and its k-th-order joint cumulant is defined for $t, v^1, v^2, ..., v^{k-1}$ in $\mathbb{Z}^d$ by $C_{kX}(v^1; v^2; ...; v^{k-1})$ $= cum\left\{X_tX_{t+v^1}...X_{t+v^{k-1}},\right\}$.

Suppose that the stationary random field X_t is observed over the cube of lattice points $t_i = 1, ..., N$, $i = 1, ..., d$. Finally, define the lexicographic order $\prec$ as: For any $s = (s_1, s_2), t = (t_1, t_2) \in \mathbb{Z}^2$, we write $s \prec t$ if and only if $[(s_1 < t_1) \vee (s_1 = t_1 \text{ and } s_2 < t_2)]$.

The paper is organized as follows. In Section 2, Almost sure convergence of sample third and fourth spatial cumulants is discussed as well as the asymptotic normality of the sample estimates for a spatial linear process. Section 3 presents two approaches for the identification of 2D nonminimum Finite Impulse Response Moving Average systems (FIR MA systems). a Monte Carlo study is presented in section 4.

2. Asymptotic properties of spatial cumulants

We consider the sample third-order moment (cumulant) $\hat{r}_3\left(\nu^1;\nu^2\right)$ of $X_{\mathbf{t}}$, namely,

(2.1)
$$\hat{r}_{3X}\left(\nu^1;\nu^2\right) = \frac{1}{N^d} \sum_{\substack{t_i=1 \\ i=1,\ldots,d}}^{N} X_{\mathbf{t}} X_{\mathbf{t}+\nu^1} X_{\mathbf{t}+v^2}.$$

Our basic assumption on $X_{\mathbf{t}}$ is

(A) $C_{4X}\left(\nu^1;\nu^2;\nu^3\right) \in L_1\left(\mathbb{R}^3\right)$ and $C_{6X}\left(\nu^1;\nu^2;\nu^3;v^4;\nu^5\right) \in L_1\left(\mathbb{R}^5\right)$. Before presenting the almost sure convergence of $\hat{r}_3\left(\nu^1;\nu^2\right)$, we prove the following lemma

Lemma 1. *If* $(X_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^d}$, $EX_{\mathbf{t}} = 0$, *is a real 6-th-order weakly stationary process, and* $r_{2X}(\nu)$, $r_{3X}\left(\nu^1;\nu^2\right)$, $C_{4X}\left(\nu^1;\nu^2;\nu^3\right)$ *and* $C_{6X}\left(\nu^1;\nu^2;\nu^3;\nu^4;v^5\right) \in L_1$, *then for* $\mathbf{t} = (t_1, \ldots, t_d)$, $\nu^1 = \left(\nu_1^1, \nu_2^1, \ldots, \nu_d^1\right)$ *and* $\nu^2 = \left(\nu_1^2, \nu_2^2, \ldots, \nu_d^2\right)$

$$Y_{\mathbf{t}}^{\left(\nu^1,\nu^2\right)} = X_{\mathbf{t}} X_{\mathbf{t}+\nu^1} X_{\mathbf{t}+v^2} - r_{3X}\left(\nu^1;\nu^2\right)$$

is a stationary process, and the spectral density function $f_{Y^{\left(\nu^1,\nu^2\right)}}(\lambda)$ *satisfies*

$$\sup_{\lambda\in\mathbb{Z}^d} f_{Y^{\left(\nu^1,\nu^2\right)}}(\lambda) < \infty.$$

Proof. See Appendix A. □

The first important result of this section is the almost sure convergence of sample third-order moment (cumulant).

Theorem 1. *If* $(X_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^d}$, $EX_{\mathbf{t}} = 0$, *is a real 6-th-order weakly stationary process, and* $r_{2X}(\nu)$, $r_{3X}\left(\nu^1;\nu^2\right)$, $C_{4X}\left(\nu^1;\nu^2;\nu^3\right)$ *and* $C_{6X}\left(\nu^1;\nu^2;\nu^3;\nu^4;v^5\right) \in L_1$, *then for* $\mathbf{t} = (t_1, \ldots, t_d)$, $\nu^1 = \left(\nu_1^1, \nu_2^1, \ldots, \nu_d^1\right)$ *and* $\nu^2 = \left(\nu_1^2, \nu_2^2, \ldots, \nu_d^2\right)$

(1)

(2.2)
$$E\left\{\frac{1}{N^{2d}}\left|\sum_{\substack{t_i=1 \\ i=1,\ldots,d}}^{N} X_{\mathbf{t}} X_{\mathbf{t}+\nu^1} X_{\mathbf{t}+v^2} - r_3\left(\nu^1;\nu^2\right)\right|^2\right\} = O\left(\frac{1}{N^d}\right)$$

(2) *for* $k > \frac{3}{4}$,

$$\frac{1}{N^d}\left|\sum_{\substack{t_i=1\\ i=1,\ldots,d}}^{N} X_{\mathbf{t}}X_{\mathbf{t}+\nu^1}X_{\mathbf{t}+v^2} - r_3\left(\nu^1;\nu^2\right)\right| \overset{a.s}{=} o\left(\frac{1}{N^{d(1-k)}}\right), \tag{2.3}$$

where a.s represents almost sure convergence.

Proof. See Appendix B. □

Remark 1. It is not difficult to generalize Theorem 1 to k-th-order moment and cumulant estimates, provided that we take $Y_{\mathbf{t}}^{\left(\nu^1;\ldots;\nu^{k-1}\right)} = X_{\mathbf{t}}X_{\mathbf{t}+\nu^1}\ldots X_{\mathbf{t}+v^{k-1}} - r_{kX}\left(\nu^1;\ldots;\nu^{k-1}\right)$, and that the cumulants up to $2k$-th-order are absolutely summable.

Now, we show the asymptotic normality of $\hat{r}_{3X}\left(\nu^1,\nu^2\right)$ for spatial linear processes.

Theorem 2. *Let* $X_{\mathbf{t}} = \sum_{\mathbf{j}\in\mathbb{Z}^d}\psi_{\mathbf{j}}\varepsilon_{\mathbf{t}-\mathbf{j}}$, *where* $\varepsilon_{\mathbf{j}}$ *are independent and identically distributed with* $E(\varepsilon_{\mathbf{j}}) = 0$, $E(\varepsilon_{\mathbf{j}}^2) = 1$, $E(\varepsilon_{\mathbf{j}}^3) = \gamma_3$, $E(\varepsilon_{\mathbf{j}}^6) < \infty$ *and* $\sum_{\mathbf{j}\in\mathbb{Z}^d}|\psi_{\mathbf{j}}| < \infty$, *then*

$$N^{\frac{d}{2}}\left(\hat{r}_{3X}\left(\nu^1;\nu^2\right) - r_{3X}\left(\nu^1;\nu^2\right)\right) \to \mathcal{N}\left(0,\sigma_X^2\left(\nu^1;\nu^2\right)\right), \tag{2.4}$$

where $\sigma_X^2(\nu^1;\nu^2) = \sum_{\substack{\tau_i=-\infty\\ i=1,\ldots,d}}^{+\infty}\{C_{6X}(\nu^1;\nu^2;\tau;\tau+\nu^1;\tau+v^2). +\{r_{2X}(\nu^1)C_{4X}(\tau-\nu^2;\tau+\nu^1-\nu^2;\nu^2)\}_{15} .+\{r_{2X}(\nu^1)r_{2X}(\tau-\nu^2)r_{2X}(\nu^2-\nu^1)\}_{15}\}$.

Proof. See Appendix C. □

3. 2D FIR moving average identification

In the sequel, we propose two approaches for the identification of non-Gaussian 2D MA processes parameters. The first approach is based on experimental cumulants alone and without imposing minimum phase. While a generalized method of moments (GMM) is used in the second approach.

3.1 2D FIR MA identification using third order cumulants alone

In this subsection, we address the problem of identification of 2D moving average (MA) models with non-Gaussian errors based on experimental cumulants alone and without imposing minimum phase. The proposed algorithm is an extension to the 2D case of the one proposed by authors in [40] in the 1D case.

Consider the 2-D finite impulse response (FIR) signal process given by the following equation :

$$x(n,m) = \sum_{i=0}^{q_1}\sum_{j=0}^{q_2} b(i,j)\,w(n-i,m-j), \tag{3.1}$$

where $b(i,j)$, $i = 1, \ldots, q_1, j = 1, \ldots, q_2$ are unknown MA coefficients, and the signal $x(n,m)$ is observed in additive noise :

$$y(n,m) = x(n,m) + v(n,m)\,. \tag{3.2}$$

The following conditions are assumed to hold:

i The driving noise sequence $w(n,m)$ is unobservable and is a zero-mean i.i.d. non-Gaussian process with at least one cumulant $0 < |\gamma_{m,w}| < \infty$, $m > 2$.

ii The system is of nonminimum phase (i.e. $x(n,m)$ is not invertible).

iii The additive noise $v(n,m)$ is a Gaussian ARMA process with unknown power spectrum and is independent of the input $w(n,m)$ and hence of the output $x(n,m)$.

To obtain conditions of minimum phase for model (3.1), we need to write it in a polynomial form as follows:

Let $\mathbf{B} = (B_1, B_2)$ the back shift operator, for any $\mathbf{i} = (i_1, i_2) \in \mathbb{Z}^2$, $\mathbf{z} = (z_1, z_2) \in \mathbb{C}^2$, write $B^{\mathbf{i}} X(t_1,t_2) = B_1^{i_1} B_2^{i_2} X(t_1,t_2) = X(t_1 - i_1, t_2 - i_2)$ and $\mathbf{z}^{\mathbf{i}} = z_1^{i_1} z_2^{i_2}$. Define $\Theta(\mathbf{z}) = \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} b(i,j)\, z_1^i z_2^j$, then the Model (3.1) can be written in a shorthand notation as:

$$x(n,m) = \Theta(\mathbf{B})\, w(n,m).$$

(3.1) is of minimum phase (or invertible) if the following condition is satisfied:

$\mathbf{C}_0$. none of the roots of $\Theta(z_1, z_2)$ lie within the closed unit polydisc $(|z_1| \le 1, |z_2| \le 1)$.

If we consider a semi-infinite prediction support from (3.1) with known order (q_1, q_2), we can derive a relationship between the MA parameters $b(i,j), i = 0, \ldots, q_1, j = 0, \ldots, q_2$, by arranging its terms by the lexicographic order, and some functions of the statistics. Indeed, it's not difficult to see that under condition (iii),

$$C_{ky}\left(\nu^1; \ldots; \nu^{k-1}\right) = C_{kx}\left(\nu^1; \ldots; \nu^{k-1}\right), \tag{3.3}$$

and that the k th order-cumulants are related to the impulse response by

$$\begin{aligned} C_{ky}\left(\nu^1; \ldots; \nu^{k-1}\right) &= \gamma_{k,w} \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} h(i,j) \times h\left[(i,j) + \nu^1\right] \\ &\quad \times \ldots \times h\left[(i,j) + \nu^{k-1}\right], \end{aligned} \tag{3.4}$$

with $\nu^1 = (t_1, t_2)$, $\nu^2 = \ldots = \nu^{k-1} = \mathbf{0}$ (where $\mathbf{0} = (0,0)$) and $k = 3$ in (3.4) we have,

$$C_{3y}\left((t_1, t_2); \mathbf{0}\right) = \gamma_{3,w} \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} h^2(i,j)\, h(i + t_1, j + t_2)\,. \tag{3.5}$$

We substitute $h(i+t_1, j+t_2) = b(i+t_1, j+t_2)$ in (3.5) to get a new normal equation that is,

$$C_{3y}((t_1,t_2),\mathbf{0}) = \gamma_{3,w} \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} b(i,j) \times h^2(i-t_1, j-t_2), \text{ for any } (t_1,t_2), \tag{3.6}$$

where we have used $h(i,j) = b(i,j) = 0$ for $i < 0$ or $j < 0$ or $i > q_1$ or $j > q_2$. Recall [32] that

$$h(i,j) = b(i,j) = \frac{C_{3y}((q_1,q_2);(i,j))}{C_{3y}((q_1,q_2);\mathbf{0})}. \tag{3.7}$$

Substituting (3.7) into (3.6) yields

$$\begin{aligned}&\gamma_{3,w} \sum_{i=0}^{q_1} \sum_{j=0}^{q_2} b(i,j)\, C_{3y}^2((q_1,q_2);(i-t_1, j-t_2)) \\ &= C_{3y}((t_1,t_2);\mathbf{0}) \times C_{3y}^2((q_1,q_2);\mathbf{0}), \text{ for any } (t_1,t_2).\end{aligned} \tag{3.8}$$

We refer to (3.8) as the spatial cumulant-based normal equation for pure 2D MA model. To remove $\gamma_{3,w}$ from equation (3.8) we use again (3.5) and (3.7) to get

$$\gamma_{3,w} = \frac{C_{3y}^2((q_1,q_2);\mathbf{0})}{C_{3y}((q_1,q_2);(q_1,q_2))}, \tag{3.9}$$

where we have used $b(0,0) = 1$ and $C_{3y}((q_1,q_2);(t_1,t_2)) = 0$ for $t_1 < 0$ or $t_2 < 0$ or $t_1 > q_1$ or $t_2 > q_2$, now substituting (3.9) into (3.8), we obtain

$$\begin{aligned}&\sum_{i=0}^{q_1} \sum_{j=0}^{q_2} b(i,j)\, C_{3y}^2((q_1,q_2);(i-t_1, j-t_2)) \\ &= C_{3y}((t_1,t_2);\mathbf{0}) \times C_{3y}((q_1,q_2);(q_1,q_2)).\end{aligned} \tag{3.10}$$

If we arrange vectors and matrices in the lexicographic order, concatenating the normal equation (3.10) for $-q_1 \le t_1 \le q_1$ and $-q_2 \le t_2 \le q_2$ we obtain the following single matricial equation

$$\begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \end{bmatrix} \mathbf{b} = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix}, \tag{3.11}$$

(see Appendix D for the expressions of $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{c}_1$ and $\mathbf{c}_2$)

The following theorem guarantees the identification for 2D FIR MA parameters (without imposing minimum phase) by solving (3.11).

Theorem 3. *Under the conditions i)-iii), given the true output cumulants* $C_{3y}(\mathbf{q};\mathbf{t})$ *and* $C_{3y}(\mathbf{r};\mathbf{0})$ *of model* (3.1) *and* (3.2) *for lags* $(0 \le t_1 \le q_1$ *and* $0 \le t_2 \le q_2)$ *and* $(-q_1 \le \tau_1 \le q_1$ *and* $-q_2 \le \tau_2 \le q_2)$ *respectively, the* $(q_1+1)(q_2+1)$ *unknowns* $b(i,j), i = 0,\ldots,q_1, j = 0,\ldots,q_2$ *can be recovered uniquely by the solution to* (3.11).

Proof. The proof is quite similar to the proof of Theorem 1 of [40] in the 1D FIR MA case and hence omitted. □

In practice , only sample estimates of $C_{3y}\left((q_1,q_2)\,;(t_1,t_2)\right)$ and $C_{3y}\left((\tau_1,\tau_2)\,;\mathbf{0}\right)$ are available from the finite sample $\{y\,(n,m)\,, 1 \le n \le N_1, 1 \le m \le N_2\}$. However, it may be possible to replace the theoretical cumulants by experimental ones when solving (3.11) for $\mathbf{b}$ which would help to identify it. So, the algorithm for 2D MA parameter identification is summarized in Algorithm I as follows.

Algorithm I.

Step 1. From the output $y\,(n,m)\,, 1 \le n \le N_1, 1 \le m \le N_2$, form the estimates

$$\hat{C}_{3y}((q_1,q_2);(t_1,t_2)) = \frac{1}{N_1N_2}\sum_{i=1}^{N_1-q_1}\sum_{j=1}^{N_2-q_2} y(i,j)\times y(i+q_1,j+q_2)y(i+t_1,j+t_2);$$

for $0 \le t_1 \le q_1$ and $0 \le t_2 \le q_2$,

and

$$\hat{C}_{3y}((\tau_1,\tau_2);\mathbf{0}) = \frac{1}{N_1N_2}\sum_{i=\max(1,-\tau_1)}^{\min(N_1,N_1-\tau_1)}\sum_{j=\max(1,-\tau_2)}^{\min(N_2,N_2-\tau_2)} y^2(i,j)\times y(i+\tau_1,j+\tau_2),$$

for $-q_1 \le \tau_1 \le q_1$ and $-q_2 \le \tau_2 \le q_2$.

Step 2. Using the estimates $\hat{C}_{3y}((q_1,q_2);(t_1,t_2))$ and $\hat{C}_{3y}((\tau_1,\tau_2);\mathbf{0})$, iteratively solve the system (3.11) to obtain the estimates $\hat{b}(i,j)$ of $b(i,j)$ $i = 0,\ldots,q_1, j = 0,\ldots,q_2$.

Step 3. Finally to satisfy the normalized condition $b(0,0) = 1$, we take $\hat{b}_1(i,j) = \frac{\hat{b}(i,j)}{\hat{b}(0,0)}$ as the final estimate of $b(i,j)$ $0 \le i,j \le 1$, $(i,j) \ne (0,0)$.

Remark 2. Theorem 1 and Theorem 2 established the almost sure convergence and the asymptotic normality of the estimates $\hat{C}_{3y}((q_1,q_2);(t_1,t_2))$ and $\hat{C}_{3y}((\tau_1,\tau_2);\mathbf{0})$.

This strong consistency guarantees that estimating the parameters $b(i,j), i = 0,\ldots,q_1, j = 0,\ldots,q_2$ results in strong consistent estimators provided that the output process $y(i,j)$ satisfies the stationary conditions of Theorem 1.

Indeed, both 2D MA parameters are identified as measurable functions of $C_{3y}((q_1,q_2),(t_1,t_2))$ and $C_{3y}((\tau_1,\tau_2);\mathbf{0})$, and $\hat{C}_{3y}((q_1,q_2);(t_1,t_2))$, $\hat{C}_{3y}((\tau_1,\tau_2);\mathbf{0})$ are consistent estimators of $C_{3y}((q_1,q_2);(t_1,t_2))$, $C_{3y}((\tau_1,\tau_2);\mathbf{0})$, respectively. Hence the 2D MA parameters will themselves be consistent. Also, the asymptotic distribution of our parameters estimators will be Gaussian, because they are functions of the asymptotically Gaussian $\hat{C}_{3y}((q_1,q_2);(t_1,t_2))$, $\hat{C}_{3y}((\tau_1,\tau_2);\mathbf{0})$ estimators.

Remark 3. In practical applications, true order (q_1,q_2) is not known a priory. In the proof of Theorem 3 it is shown that the rank of the cumulant matrix $\begin{bmatrix}\mathbf{C}_1\\ \mathbf{C}_2\end{bmatrix}$ is $R=(q_1+1)(q_2+1)$. Hence, (q_1,q_2) can be determined using the singular value decomposition (SVD) by testing for

$$\hat{C}_{3y}^{R}((q_1,q_2);\mathbf{0})\neq 0, \hat{C}_{3y}^{\varkappa+1}((k_1,k_2);\mathbf{0})\approx 0, \tag{3.12}$$

where $\varkappa=(k_1+1)(k_2+1)$. See [39], for details, of the SVD based on (3.12).

3.2 2D FIR MA identification using GMM estimators

The results in Subsection 2.2 suggest a second estimation within the method of moments framework. In the case of 2D FIR MA(1) model given by (3.1) and (3.2), we can derive orthogonality conditions to estimate

$$\gamma=(\frac{b^2(0,0)}{b(1,1)},\frac{b^2(0,1)}{b(1,1)},\frac{b^2(1,0)}{b(1,1)},b(1,1))'$$

from equations (3.4), (3.7) and (3.9). Indeed, for $k=3$, $\nu^1=(-t_1,-t_2)$ and $\nu^2=\mathbf{0}$ equation (3.4) gives,

$$C_{3y}((-t_1,-t_2);\mathbf{0})=\gamma_{3,w}\sum_{i=0}^{1}\sum_{j=0}^{1}h^2(i,j)\,h(i-t_1,j-t_2),$$

since $C_{3y}((-t_1,-t_2);\mathbf{0})=C_{3y}((t_1,t_2);(t_1,t_2))$, and using (3.7) and (3.9) we get,

$$\begin{aligned}C_{3y}((t_1,t_2);(t_1,t_2)) &= \frac{b^2(0,0)}{b(1,1)}C_{3y}((q_1,q_2);(-t_1,-t_2))\\ &+\frac{b^2(0,1)}{b(1,1)}C_{3y}((q_1,q_2);(-t_1,1-t_2))\\ &+\frac{b^2(1,0)}{b(1,1)}C_{3y}((q_1,q_2);(1-t_1,-t_2))\\ &+b(1,1)\,C_{3y}((q_1,q_2);(1-t_1,1-t_2)),\end{aligned}$$

for $0\leq t_1\leq 1$ and $0\leq t_2\leq 1$. This represents a system of four orthogonality conditions given by

$$\begin{cases}h(\gamma,(i,j))=A(i,j)\times\gamma-H(i,j),\\ E(h(\gamma,(i,j)))=0\end{cases}$$

where

$$A\left(i,j\right)=y(i,j)y(i+1,j+1) \times \begin{pmatrix} y(i,j) & y(i,j+1) & y(i+1,j) & y(i+1,j+1) \\ y(i,j-1) & y(i,j) & y(i+1,j-1) & y(i+1,j) \\ y(i-1,j) & y(i-1,j+1) & y(i,j) & y(i,j+1) \\ y(i-1,j-1) & y(i-1,j) & y(i,j-1) & y(i,j) \end{pmatrix},$$

and

$$H\left(i,j\right)=y(i,j)\times\begin{pmatrix} y^2(i,j) \\ y^2(i,j+1) \\ y^2(i+1,j) \\ y^2(i+1,j+1) \end{pmatrix}.$$

For given data $\{y(n,m), 1 \leq n \leq N_1, 1 \leq m \leq N_2\}$, define $g_{\mathbf{N}}(\gamma,(i,j)) = \frac{1}{N_1 N_2}\sum_{i=1}^{N_1}\sum_{i=1}^{N_2} h(\gamma,(i,j))$, where $\mathbf{N} = (N_1, N_2)$, and assume that γ_0, the true value of γ, is in the interior of a compact parameter space Γ. The GMM estimator of γ is then defined by solving the minimization problem,

$$(3.13) \qquad \hat{\gamma}_{\mathbf{N}}^{GMM} = \arg\min_{\gamma\in\mathbf{\Omega}} \left[g_{\mathbf{N}}\left(\gamma,(i,j)\right)\right]' W_{\mathbf{N}} \left[g_{\mathbf{N}}\left(\gamma,(i,j)\right)\right],$$

where the weight matrix $W_{\mathbf{N}}$ is positive definite. Since there are four parameters for four conditions of orthogonality, the system is exactly identified and the GMM estimator $\gamma_{\mathbf{N}}^{GMM}$ of γ based on the N_1N_2 observations is defined by the equation $g_{\mathbf{N}}\left(\gamma_{\mathbf{N}}^{GMM},(i,j)\right)=0$.

Hence, we just solve the system $\frac{1}{N_1 N_2}\sum_{i=1}^{N_1}\sum_{i=1}^{N_2} h\left(\gamma_{\mathbf{N}}^{GMM},(i,j)\right)=0$, that can be rewritten as $\left[\sum_{i=1}^{N_1}\sum_{i=1}^{N_2} A\left(i,j\right)\right]\times\gamma_{\mathbf{N}}^{GMM}=\sum_{i=1}^{N_1}\sum_{i=1}^{N_2} H\left(i,j\right)$, consequently

$$(3.14) \qquad \gamma_{\mathbf{N}}^{GMM}=\left[\sum_{i=1}^{N_1}\sum_{i=1}^{N_2} A\left(i,j\right)\right]^{-1}\times\sum_{i=1}^{N_1}\sum_{i=1}^{N_2} H\left(i,j\right).$$

The consistency and asymptotic normality of $\gamma_{\mathbf{N}}^{GMM}$ can be obtained after the check of some regularity conditions. First, a good way for the weighting matrix $W_{\mathbf{N}}$ to solve the minimization problem (3.13) is by putting $W_{\mathbf{N}}\equiv\tilde{W}_{\mathbf{N}}=[\frac{1}{N_1N_2}\sum_{i=1}^{N_1}\sum_{i=1}^{N_2} h(\gamma_{\mathbf{N}}^{(1)},(i,j))h'(\gamma_{\mathbf{N}}^{(1)},(i,j))]^{-1}$ where

$$\gamma_{\mathbf{N}}^{(1)}=\left(\frac{\hat{b}_1^2\left(0,0\right)}{\hat{b}_1\left(1,1\right)},\frac{\hat{b}_1^2\left(0,1\right)}{\hat{b}_1\left(1,1\right)},\frac{\hat{b}_1^2\left(1,0\right)}{\hat{b}_1\left(1,1\right)},\hat{b}_1\left(1,1\right)\right)'$$

with $\hat{b}_1\left(0,0\right), \hat{b}_1\left(0,1\right), \hat{b}_1\left(1,0\right), \hat{b}_1\left(1,1\right)$ being the consistent estimates of $b\left(0,0\right)$, $b\left(0,1\right)$, $b\left(1,0\right)$, $b\left(1,1\right)$, respectively, derived in Algorithm I. Since $h\left(\gamma,(i,j)\right)$ is a measurable function of $X\left(i,j\right), 0\leq i,j\leq 1$ and γ is assumed to be such that conditions in Theorem 3 hold, we get $\tilde{W}_{\mathbf{N}}\xrightarrow{P}\tilde{W}=\left[Eh\left(\gamma_0,(i,j)\right)h'\left(\gamma_0,(i,j)\right)\right]^{-1}$, it is clear that $\tilde{W}_{\mathbf{N}}$ and $\tilde{W}$ are both definite and positive. Hence, the consistency and asymptotic normality of $\gamma_{\mathbf{N}}^{GMM}$ can be established in the following theorem:

Theorem 4. *Consider a 2D FIR MA(1) model and suppose that the conditions in Theorem 3 hold. Then,*

$$\sqrt{N_1N_2}\left(\gamma_{\mathbf{N}}^{GMM}-\gamma_0\right)\overset{\mathcal{L}}{\rightsquigarrow}\mathcal{N}\left(\mathbf{0},\Sigma\right),$$

where $\Sigma = V^{-1}\Omega V^{-1}$ *with* $V = E\left(A\left(i,j\right)\right)$ *and* $\Omega = \tilde{W}^{-1}$.

Proof. In view of (3.14) we can write,

$$\begin{aligned}\left(\gamma_{\mathbf{N}}^{GMM}-\gamma_0\right) &= \left[\frac{1}{\sqrt{N_1N_2}}\sum_{i=1}^{N_1}\sum_{i=1}^{N_2}A\left(i,j\right)\right]^{-1}\\ &\times\frac{1}{\sqrt{N_1N_2}}\sum_{i=1}^{N_1}\sum_{i=1}^{N_2}h\left(\gamma_0,(i,j)\right),\end{aligned}\tag{3.15}$$

so it is not difficult to show that, under the conditions of Theorem 3, $h\left(\gamma_0,(i,j)\right)$ is a 2D martingale difference with finite variance $\Omega = \tilde{W}^{-1}$. Hence, by Huang martingale difference central limit theorem $\frac{1}{\sqrt{N_1N_2}}\sum_{i=1}^{N_1}\sum_{i=1}^{N_2}h\left(\gamma_0,(i,j)\right)$ converges to the multivariate normal distribution with mean zero and covariance matrix Ω. From (3.15) and by Cramer-Wold device, the result follows. □

Now, to derive the estimate of the parameter of interest $\mathbf{b}$ from $\gamma_{\mathbf{N}}^{GMM}$, notice first that $\gamma_{\mathbf{N}}^{GMM}(i)$ and $\gamma_{\mathbf{N}}^{GMM}(4)$ $i = 1,2,3$ have the same sign so we simply compute directly $\hat{b}_{GMM}(1,1) = \gamma_{\mathbf{N}}^{GMM}(4)$; $|\hat{b}_{GMM}(0,0)| = [\gamma_{\mathbf{N}}^{GMM}(1)\gamma_{\mathbf{N}}^{GMM}(4)]^{\frac{1}{2}}$; $|\hat{b}_{GMM}(0,1)| = [\gamma_{\mathbf{N}}^{GMM}(2)\gamma_{\mathbf{N}}^{GMM}(4)]^{\frac{1}{2}}$ and $|\hat{b}_{GMM}(1,0)| = [\gamma_{\mathbf{N}}^{GMM}(3)\gamma_{\mathbf{N}}^{GMM}(4)]^{\frac{1}{2}}$ by keeping the sign of the estimates of $b(0,0)$, $b(0,1)$ and $b(1,0)$ given in Algorithm I. Finally to satisfy the normalized condition $b(0,0) = 1$ we take $\frac{\hat{b}_{GMM}(i,j)}{\hat{b}_{GMM}(0,0)}$ as the final estimate of $b(i,j)$ $0 \le i,j \le 1$, $(i,j) \ne (0,0)$.

3.3 Simulation results

In this section, we evaluate the proximity of the parameter estimators of $\mathbf{b}$ to their expected values by a Monte Carlo study and compare the performances of the two proposed estimators. This study allows to check and to corroborate, first the efficiency of the previous identification algorithm and second the validity of the proposed theoretical results presented in the previous subsections. Consider the 2D MA model of first order,

$$\begin{aligned}x\left(n,m\right) = {} & b_1w\left(n,m-1\right)+b_2w\left(n-1,m\right)\\ &+b_3w\left(n,m-2\right)+w\left(n,m\right),\end{aligned}\tag{4.1}$$

The additive noise $v\left(n,m\right)$ is a Gaussian ARMA process given by

$$\begin{aligned}v\left(n,m\right) &= 0.3v\left(n,m-1\right)-0.08v\left(n-2,m-1\right)\\ &\quad+0.54+e\left(n,m\right)-1.25e\left(n,m-1\right),\end{aligned}$$

the signal $x(n,m)$ is observed in additive noise:

$$y(n,m) = x(n,m) + v(n,m),$$

where $e(n,m) \sim \mathcal{N}(0, 0.5)$ and $w(n,m)$ are i.i.d $(0,1)$ and follows a generalized lambda distribution (GLD) with a skewness $\kappa_3 = 0$ and a kurtosis $\kappa_4 = 3$. The true values of the parameters are and $(b_1 = 1, b_2 = 0.25, b_3 = -1, 65)$ (nonminimum phase case) and $(b_1 = 0, 25, b_2 = 0, 35, b_3 = 0, 15)$ (minimum phase case), Proposition 1 in [2] gives a convenient set of conditions on the parameters b_1, b_2 and b_3 to satisfy $\mathbf{C}_0$). Notice that in the proposed identification algorithm the minimum phase is not imposed.

Furthermore, errors with nonGaussian features are the key to identification. It is why we choose a GLD distribution errors considered in [28] that can be simulated in a flexible fashion and can accommodate a wide range of values for skewness and excess kurtosis parameters. Table 1 and 2 present the averages mean columns, mean square errors columns, kurtosis columns and skewness columns of b_1, b_2 and b_3 estimators over 500 Monte Carlo replications for the sample sizes $(100, 100)$, $(500, 500)$ and $(1000, 1000)$ in both the minimum and nominimum phase respectively. The average values of the estimator given by Algorithm I $(\hat{b}_i)$ and the GMM estimator $(\hat{b}_i^{gmm})$, $i = 1, 2, 3$, in the two cases, over the 500 replications were always close to the true values for all grid sizes. Furthermore and as expected, the increase in the sample size substantially improves the results in terms of bias reduction. Also, the averages of the coefficients of kurtosis and skweness over the 500 replications tends to be closer to the reference values 3 and 0 respectively when the sample size is increasing for the two estimators. The observations from Table 1 and 2 confirm the consistency and the asymptotic normality of the proposed estimators as indicated in Remark 2 and Theorem 4. They also suggest that the GMM estimator produces estimates of $\mathbf{b}$ that are somewhat more precise than those given by Algorithm I. However, despite the reasonable computational burden of the GMM estimator of $\mathbf{b}$, a major advantage of the estimator given by Algorithm I is that it doesn't require a preliminary consistent estimator.

Table 1: Sample distribution of the estimators $\widehat{b_1}$, $\widehat{b_2}$ and $\widehat{b_3}$ and $\widehat{b}_1^{gmm}$, $\widehat{b}_2^{gmm}$ and $\widehat{b}_3^{gmm}$ (between brackets) in the minimum phase case with $\kappa_3 = 0$ and $\kappa_4 = 3$.

	$N =$	$(100, 100)$	
true b_i	0.25	0.35	0.15
	$\widehat{b}_1$ ($\widehat{b}_1^{gmm}$)	$\widehat{b}_2$ ($\widehat{b}_2^{gmm}$)	$\widehat{b}_3$ ($\widehat{b}_3^{gmm}$)
mean	0.2501 (0.2498)	0.3502 (0.3501)	0.150 (0.153)
rmse	0.18×10^{-4} (0.12×10^{-4})	0.0029 (0.002)	0.0038 (0.0025)
kurt	2.5910 (3.2985)	3.1432 (3.176)	2.879 (2.764)
skew	0.0151 (0.0099)	0.0092 (0.0111)	0.0098 (0.0089)

	$N=$	$(500,500)$	
true b_i	0.25	0.35	0.15
	$\widehat{b_1}$ ($\widehat{b_1^{gmm}}$)	$\widehat{b_2}$ ($\widehat{b_2^{gmm}}$)	$\widehat{b_3}$ ($\widehat{b_3^{gmm}}$)
mean	0.25009 (0.2500)	0.3501 (0.3497)	0.15006 (0.1502)
rmse	0.119×10^{-4} (0.099×10^{-4})	0.0015 (0.0014)	0.0029 (0.0025)
kurt	2.7910 (3.321)	2.9143 (3.156)	3.1787 (2.8691)
skew	0.0105 (0.0098)	0.0100 (0.0113)	0.0091 (0.0079)

	$N=$	$(1000,1000)$	
true b_i	0.25	0.35	0.15
	$\widehat{b_1}$ ($\widehat{b_1^{gmm}}$)	$\widehat{b_2}$ ($\widehat{b_2^{gmm}}$)	$\widehat{b_3}$ ($\widehat{b_3^{gmm}}$)
mean	0.25001 (0.2489)	0.35005 (0.35005)	0.150 (0.1499)
rmse	0.119×10^{-4} (0.074×10^{-4})	0.0015 (0.0013)	0.0029 (0.0025)
kurt	2.7910 (3.146)	2.9143 (3.09)	3.1787 (3.234)
skew	0.0105 (0.0098)	0.0100 (0.0054)	0.0091 (0.0123)

Table 2: Sample distribution of the estimators $\widehat{b_1}$, $\widehat{b_2}$ and $\widehat{b_3}$ and $\widehat{b_1^{gmm}}$, $\widehat{b_2^{gmm}}$ and $\widehat{b_3^{gmm}}$ (between brackets) in the nonminimum case with $\kappa_3=0$ and $\kappa_4=3$.

	$N=$	$(100,100)$	
true b_i	1	0.25	−1.65
	$\widehat{b_1}$ ($\widehat{b_1^{gmm}}$)	$\widehat{b_2}$ ($\widehat{b_2^{gmm}}$)	$\widehat{b_3}$ ($\widehat{b_3^{gmm}}$)
mean	1.0003 (1.0005)	0.2499 (0.2503)	−1.6501 (−1.6498)
rmse	0.9229×10^{-4} (0.9×10^{-4})	0.00218 (0.002)	0.00079 (0.00062)
kurt	3.2110 (2.763)	2.8942 (3.245)	3.1787 (3.0985)
skew	0.0099 (0.0134)	0.0210 (0.0125)	0.0081 (0.0099)

	$N=$	$(500,500)$	
true b_i	1	0.25	-1.65
	$\widehat{b_1}$ ($\widehat{b_1^{gmm}}$)	$\widehat{b_2}$ ($\widehat{b_2^{gmm}}$)	$\widehat{b_3}$ ($\widehat{b_3^{gmm}}$)
mean	1 (0.999)	0.2499 (0.2508)	-1.6502 (−1.65)
rmse	0.8589×10^{-4} (0.812×10^{-4})	0.00191 (0.0019)	0.00059 (0.0005)
kurt	2.9902 (3.146)	2.6942 (2.885)	3.0840 (3.1328)
skew	0.0089 (0.0121)	0.0110 (0.0085)	0.0071 (0.0099)

	$N=$	$(1000,1000)$	
true b_i	1	0.25	-1.65
	$\widehat{b_1}$ ($\widehat{b_1^{gmm}}$)	$\widehat{b_2}$ ($\widehat{b_2^{gmm}}$)	$\widehat{b_3}$ ($\widehat{b_3^{gmm}}$)
mean	1 (1)	0.25005 (0.2481)	-1.65004 (−1.65)
rmse	0.8143×10^{-4} (0.835×10^{-4})	0.00132 (0.0011)	0.00055 (0.00043)
kurt	3.002 (3.0904)	2.9941 (3.281)	3.1840 (2.871)
skew	0.0009 (0.0012)	0.0120 (0.0088)	0.0069 (0.005)

4. Conclusion

This paper presented new approaches to take into account complex non-linear and /or non-Gaussian signal processing data. The originality of this work was the investigation of asymptotic properties of high-order statistics in a spatial context. Convergence properties of third- and fourth-order moments and cumulants estimates have been described and proven. The study also presented new ways to identify a linear 2-D nonminimum phase system using higher order cumulants alone and with GMM estimators. A Monte Carlo study has been conduced.

Appendix A (proof of Lemma 1)

Proof. It is obvious that $Y_{\mathbf{t}}^{(\nu^1,\nu^2)}$ is a stationary process. Put for all $\tau \in \mathbb{Z}^d$, $R_{Y^{(\nu^1,\nu^2)}}(\tau) = EY_{\mathbf{t}}^{(\nu^1,\nu^2)} Y_{\mathbf{t}+\tau}^{(\nu^1,\nu^2)} = r_{6X}(\nu^1,\nu^2,\tau,\tau+\nu^1,\tau+v^2) - [r_{3X}(\nu^1,\nu^2)]^2$. Using the conversion relationship between higher-order moments and cumulants we have (see proof of Lemma1 of [18])

$$\begin{aligned} R_{Y^{(\nu^1,\nu^2)}}(\tau) &= C_{6X}\left(\nu^1,\nu^2,\tau,\tau+\nu^1,\tau+v^2\right) \\ &\quad + \left\{r_{2X}\left(\nu^1\right) C_{4X}\left(\tau-\nu^2,\tau+\nu^1-\nu^2,\nu^2\right)\right\}_{15} \\ &\quad + \left\{r_{2X}\left(\nu^1\right) r_{2X}\left(\tau-\nu^2\right) r_{2X}\left(\nu^2-\nu^1\right)\right\}_{15}, \end{aligned}$$

where the notation $\{.\}_j$ denotes the sum of all j different terms obtained by interchanging the arguments of the terms in brackets, and by the absolutely summable properties of $r_{2X}(\nu)$, $r_{3X}(\nu^1,\nu^2)$, $C_{4X}(\nu^1,\nu^2,\nu^3)$ and $C_{6X}(\nu^1,\nu^2,\nu^3,\nu^4,v^5)$, we have

$$\sum_{\tau\in\mathbb{Z}^d} \left| R_{Y^{(\nu^1,\nu^2)}}(\tau) \right| < +\infty.$$

If $f_{Y^{(\nu^1,\nu^2)}}(.)$ is the spectral density function of $Y_{\mathbf{t}}^{(\nu^1,\nu^2)}$, by spectral theory (see [17]), $f_{Y^{(\nu^1,\nu^2)}}(\lambda)$ has the representation

$$f_{Y^{(\nu^1,\nu^2)}}(\lambda) = \frac{1}{(2\pi)^d} \sum_{\mathbf{n}\in\mathbb{Z}^d} R_{Y^{(\nu^1,\nu^2)}}(\mathbf{n})\, e - i\left\langle \lambda,\mathbf{n}\right\rangle,$$

where $\langle \lambda,\mathbf{n}\rangle = \sum_{i=1}^d \lambda_i n_i$, hence

$$\sup_{\lambda\in\mathbb{Z}^d} f_{Y^{(\nu^1,\nu^2)}}(\lambda) \le \frac{1}{(2\pi)^d} \sum_{\mathbf{n}\in\mathbb{Z}^d} \left| R_{Y^{(\nu^1,\nu^2)}}(\mathbf{n}) \right| < +\infty. \qquad \square$$

Appendix B (proof of Theorem 1)

Proof. By spectral theory (see [17]), we have

$$R_{Y^{(\nu^1,\nu^2)}}(\tau) = \int_{[0,2\pi]^d} e^{i\langle\lambda,\tau\rangle} f_{Y^{(\nu^1,\nu^2)}}(\lambda)\, d\lambda,$$

furthermore

$$E\left\{\left|\sum_{\substack{t_i=1\\ i=1,\ldots,d}}^{N} X_{\mathbf{t}}X_{\mathbf{t}+\nu^1}X_{\mathbf{t}+v^2}-r_3\left(\nu^1,\nu^2\right)\right|^2\right\}$$

$$=E\left\{\left|\sum_{\substack{t_i=1\\ i=1,\ldots,d}}^{N} Y_{\mathbf{t}}^{\left(\nu^1,\nu^2\right)}\right|^2\right\}$$

$$=E\left\{\sum_{\substack{t_i,s_i=1\\ i=1,\ldots,d}}^{N} Y_{\mathbf{t}}^{\left(\nu^1,\nu^2\right)}Y_{\mathbf{s}}^{\left(\nu^1,\nu^2\right)}\right\}$$

$$=\sum_{\substack{t_i,s_i=1\\ i=1,\ldots,d}}^{N} R_{Y^{\left(\nu^1,\nu^2\right)}}\left(\mathbf{t}-\mathbf{s}\right),$$

hence

$$E\left\{\left|\sum_{\substack{t_i=1\\ i=1,\ldots,d}}^{N} X_{\mathbf{t}}X_{\mathbf{t}+\nu^1}X_{\mathbf{t}+v^2}-r_3\left(\nu^1,\nu^2\right)\right|^2\right\}$$

$$=\sum_{\substack{t_i,s_i=1\\ i=1,\ldots,d}}^{N}\int_{[0,2\pi]^d} e^{i\langle\lambda,\mathbf{t}-\mathbf{s}\rangle} f_{Y^{\left(\nu^1,\nu^2\right)}}\left(\lambda\right)d\lambda$$

$$\leq \sup_{\lambda\in\mathbb{Z}^d} f_{Y^{\left(\nu^1,\nu^2\right)}}\left(\lambda\right)\sum_{\substack{t_i,s_i=1\\ i=1,\ldots,d}}^{N}\int_{[0,2\pi]^d} e^{i\langle\lambda,\mathbf{t}-\mathbf{s}\rangle}d\lambda\leq C\times N^d,$$

where C represents a constant which is independent of N, then

$$E\left\{\frac{1}{N^{2d}}\left|\sum_{\substack{t_i=1\\ i=1,\ldots,d}}^{N} X_{\mathbf{t}}X_{\mathbf{t}+\nu^1}X_{\mathbf{t}+v^2}-r_3\left(\nu^1,\nu^2\right)\right|^2\right\}\leq C\times\frac{N^d}{N^{2d}}=\frac{C}{N^d},$$

which is (2.2). Following [18], consider (2.3) on the basis of (2.2) and let

$$S\left(N\right)=\frac{1}{N^{kd}}\left|\sum_{\substack{t_i=1\\ i=1,\ldots,d}}^{N} Y_{\mathbf{t}}^{\left(\nu^1,\nu^2\right)}\right|.$$

For $k > \frac{3}{4}$, there is a β which satisfies $\beta > 1$ and,

$$k - \frac{1}{2} > \frac{1}{2\beta} > \frac{1}{4}.$$

Now, consider the two case where $N = M^{\beta}$ and for general N.

(I) When $N = M^{\beta}$, we have

$$\begin{aligned} E\left|S(N)\right|^2 &= E\frac{1}{N^{2kd}}\left|\sum_{\substack{t_i=1\\ i=1,\ldots,d}}^{N} Y_{\mathbf{t}}^{\left(\nu^1,\nu^2\right)}\right|^2 \\ &= \frac{1}{N^{2kd-2d}}E\frac{1}{N^{2d}}\left|\sum_{\substack{t_i=1\\ i=1,\ldots,d}}^{N} Y_{\mathbf{t}}^{\left(\nu^1,\nu^2\right)}\right|^2 \leq \frac{1}{N^{2kd-2d}} \times \frac{C}{N^d} = CM^{\beta d(1-2k)}, \end{aligned}$$

since $\beta d\left(1-2k\right) < -d \leq -1$, hence $\sum_{M=1}^{+\infty} M^{\beta d(1-2k)}$ is a convergent Riemman series, consequently

$$\sum_{M=1}^{+\infty} E\left|S\left(M^{\beta}\right)\right|^2 < +\infty.$$

It follows from the Chebychev inequality and the Borel-Cantelli lemma, that

$$\lim_{M\to\infty} S\left(M^{\beta}\right) \overset{a.s}{=} 0. \tag{2.5}$$

(II) For general N, and $N_M \leq N \leq N_{M+1}$, where $N_M = M^{\beta}$, we have

$$\left|S(N) - N^{-kd}M^{kd\beta}S\left(M^{\beta}\right)\right| \leq M^{-\beta kd}\left|\sum_{\substack{t_i=M^{\beta}+1\\ i=1,\ldots,d}}^{N} Y_{\mathbf{t}}^{\left(\nu^1,\nu^2\right)}\right|,$$

and

$$\begin{aligned} &E\left\{\sup_{N_M\leq N\leq N_{M+1}}\left|S(N) - N^{-kd}M^{kd\beta}S\left(M^{\beta}\right)\right|^2\right\} \\ &\leq M^{-2\beta kd}\sup_{N_M\leq N\leq N_{M+1}} E\left|\sum_{\substack{t_i=M^{\beta}+1\\ i=1,\ldots,d}}^{N} Y_{\mathbf{t}}^{\left(\nu^1,\nu^2\right)}\right| \\ &\leq CM^{-2\beta kd}\sup_{N_M\leq N\leq N_{M+1}}\left[N - M^{\beta}\right]^d \leq CM^{-2\beta kd}\left[(M+1)^{\beta} - M^{\beta}\right]^d, \end{aligned}$$

noting that $(M+1)^{\beta}-M^{\beta}\leq(2M)^{\beta}$, we deduce that

$$E\sup_{N_M\leq N\leq N_{M+1}}\left|S\left(N\right)-N^{-kd}M^{kd\beta}S\left(M^{\beta}\right)\right|^{2}\leq C_1M^{\beta d(1-2k)},$$

where C_1 represents a constant which is independent of M. Since $\beta d\left(1-2k\right)<-d\leq-1$, we obtain

$$\sum_{M=1}^{+\infty}E\sup_{N_M\leq N\leq N_{M+1}}\left|S\left(N\right)-N^{-kd}M^{kd\beta}S\left(M^{\beta}\right)\right|^{2}<\infty.$$

It follows from the Chebychev inequality and the Borel-Cantelli lemma, that

$$\lim_{M\to\infty}\sup_{N_M\leq N\leq N_{M+1}}\left|S\left(N\right)-N^{-kd}M^{kd\beta}S\left(M^{\beta}\right)\right|^{2}\overset{a.s}{=}0. \tag{2.6}$$

By (2.5), (2.6) and

$$S(N)\leq\sup_{N_M\leq N\leq N_{M+1}}|S(N)-N^{-kd}M^{kd\beta}S(M^{\beta})|+N^{-kd}M^{kd\beta}S(M^{\beta}),$$

we have, for $k>\frac{3}{4}$,

$$S\left(N\right)=\frac{1}{N^{kd}}\left|\sum_{\substack{t_i=1\\i=1,\ldots,d}}^{N}Y_{\mathbf{t}}^{\left(\nu^1,\nu^2\right)}\right|\overset{a.s}{\rightarrow}0,$$

which completes the proof. □

Appendix C: Proof of Theorem 2

Proof. For $\mathbf{k}=(k,\ldots,k)\in\mathbb{Z}^d$ such that $k>0$, let

$$X_{\mathbf{t},\mathbf{k}}=\sum_{\substack{s_i=-k\\i=1,\ldots,d}}^{+k}\psi_{\mathbf{s}}\varepsilon_{\mathbf{t}-\mathbf{s}},$$

$$\hat{r}_{3,\mathbf{k}\mathbf{X}}\left(\nu^1,\nu^2\right)=\frac{1}{N^d}\sum_{\substack{t_i=1\\i=1,\ldots,d}}^{N}X_{\mathbf{t},\mathbf{k}}X_{\mathbf{t}+\nu^1,\mathbf{k}}X_{\mathbf{t}+v^2,\mathbf{k}},$$

$$E\left(X_{\mathbf{t},\mathbf{k}}X_{\mathbf{t}+\nu^1,\mathbf{k}}X_{\mathbf{t}+v^2,\mathbf{k}}\right)=r_{3,\mathbf{k}\mathbf{X}}\left(\nu^1,\nu^2\right).$$

By the symmetries of $r_{3X}\left(\nu^1,\nu^2\right)$, we can assume $0\leq\nu_i^1\leq\nu_i^2$ for all $i=1,\ldots,d$, then

$$E\left(X_{\mathbf{t},\mathbf{k}}X_{\mathbf{t}+\nu^1,\mathbf{k}}X_{\mathbf{t}+v^2,\mathbf{k}}\right)=\begin{cases}\gamma_3\displaystyle\sum_{\substack{s_i=-k\\i=1,\ldots,d}}^{k-\nu_i^2}\psi_{\mathbf{s}}\psi_{\nu^1+\mathbf{s}}\psi_{\nu^2+\mathbf{s}},\\ \nu_i^2\leq 2k, & \text{for at least one } i\\ 0, & \text{if not}\end{cases}.$$

For $\nu_i^2 \leq 2k$ for at least one $i = 1, \ldots, d$, we have

$$r_{3,\mathbf{kX}}\left(\nu^1, \nu^2\right) = E\hat{r}_{3,\mathbf{kX}}\left(\nu^1, \nu^2\right).$$

We deduce that $X_{\mathbf{t},\mathbf{k}} X_{\mathbf{t}+\nu^1,\mathbf{k}} X_{\mathbf{t}+v^2,\mathbf{k}}$ is independent of $X_{\mathbf{s},\mathbf{k}} X_{\mathbf{s}+\nu^1,\mathbf{k}} X_{\mathbf{s}+v^2,\mathbf{k}}$ provided that $|s_i - t_i| > 2k + \nu_i^2$ for at least one $i = 1, \ldots, d$ via the independence of $\varepsilon_{\mathbf{t}}$. Hence, if we put $\mathbf{m} = 2\mathbf{k} + \nu^2$, $X_{\mathbf{t},\mathbf{k}} X_{\mathbf{t}+\nu^1,\mathbf{k}} X_{\mathbf{t}+v^2,\mathbf{k}} - r_{3,\mathbf{kX}}\left(\nu^1, \nu^2\right)$ is a **m**-dependent stationary random field with mean zero, it follows that

$$N^{-\frac{d}{2}} \left(\sum_{\substack{t_i=1 \\ i=1,\ldots,d}}^{N} \left\{ X_{\mathbf{t},\mathbf{k}} X_{\mathbf{t}+\nu^1,\mathbf{k}} X_{\mathbf{t}+v^2,\mathbf{k}} - r_{3,\mathbf{kX}}\left(\nu^1, \nu^2\right) \right\} \right)$$
$$\to \mathcal{N}\left(0, \sigma_{X,\mathbf{k}}^2\left(\nu^1, \nu^2\right)\right), \tag{2.7}$$

where

$$\sigma_{X,\mathbf{k}}^2(\nu^1, \nu^2) = \lim_{N\to\infty} N^d \times Var\left(N^{-d} \sum_{\substack{t_i=1 \\ i=1,\ldots,d}}^{N} \left\{ X_{\mathbf{t},\mathbf{k}} X_{\mathbf{t}+\nu^1,\mathbf{k}} X_{\mathbf{t}+v^2,\mathbf{k}} - r_{3,\mathbf{kX}}(\nu^1, \nu^2) \right\} \right),$$

it is easy to show that

$$\begin{aligned}
\sigma_{X,\mathbf{k}}^2(\nu^1, \nu^2) = \sum_{\substack{\tau_i=-\infty \\ i=1,\ldots,d}}^{+\infty} &\{ C_{6,\mathbf{kX}}(\nu^1, \nu^2, \tau, \tau + \nu^1, \tau + v^2) \\
&+ \{ r_{2,\mathbf{kX}}(\nu^1) C_{4,\mathbf{kX}}(\tau - \nu^2, \tau + \nu^1 - \nu^2, \nu^2) \}_{15} \\
&+ \{ r_{2,\mathbf{kX}}(\nu^1) r_{2,\mathbf{kX}}(\tau - \nu^2) r_{2,\mathbf{kX}}(\nu^2 - \nu^1) \}_{15} \}.
\end{aligned}$$

Now, to prove (2.4) we use Proposition 6.3.9 of [7]. For this, observe that;

$$\begin{aligned}
&N^{\frac{d}{2}}[\hat{r}_{3X}(\nu^1, \nu^2) - r_{3X}(\nu^1, \nu^2)] = N^{\frac{d}{2}}\{\hat{r}_{3X}(\nu^1, \nu^2) - \hat{r}_{3,\mathbf{kX}}(\nu^1, \nu^2) \\
&- E[\hat{r}_{3X}(\nu^1, \nu^2) - \hat{r}_{3,\mathbf{kX}}(\nu^1, \nu^2)]\} + N^{\frac{d}{2}}[\hat{r}_{3,\mathbf{kX}}(\nu^1, \nu^2) - r_{3,\mathbf{kX}}(\nu^1, \nu^2)] \\
&= T_{k,N} + Z_{k,N},
\end{aligned}$$

where $T_{k,N} = N^{-\frac{d}{2}}\{\hat{r}_{3X}(\nu^1, \nu^2) - \hat{r}_{3,\mathbf{kX}}(\nu^1, \nu^2) - E[\hat{r}_{3X}(\nu^1, \nu^2) - \hat{r}_{3,\mathbf{kX}}(\nu^1, \nu^2)]\}$ and $Z_{k,N} = N^{-\frac{d}{2}}[\hat{r}_{3,\mathbf{kX}}(\nu^1, \nu^2) - r_{3,\mathbf{kX}}(\nu^1, \nu^2)]$. But according to Proposition 6.3.9 of [7], it remains to prove that:

(i) $Z_{k,N} \Rightarrow Z_k$ as $N \to \infty$ for each $k = 1, 2, \ldots$

(ii) $Z_k \Rightarrow Z$ as $k \to \infty$,

(iii) $\lim_{k\to\infty} \lim_{N\to\infty} \sup P(|T_{k,N}| > \epsilon) = 0$ for every $\epsilon > 0$, after which we can immediately deduce (2.4). First, (2.7) establish (i). Now let's prove (iii). According to Chebychev's inequality, to show (iii) it suffices to show that

$$\lim_{k\to\infty}\lim_{N\to\infty}\sup Var\left(T_{k,N}\right)=0. \tag{2.8}$$

By some calculations we obtain

$$N^{\frac{d}{2}}\left\{\hat{r}_{3X}\left(\nu^1,\nu^2\right)-\hat{r}_{3,\mathbf{k}\mathbf{X}}\left(\nu^1,\nu^2\right)\right\}$$
$$=N^{-\frac{d}{2}}\sum_{\substack{t_i=1\\ i=1,\dots,d}}^{N}\left\{\underbrace{X_{\mathbf{t},\mathbf{k}}X_{\mathbf{t}+\nu^1,\mathbf{k}}u_{\mathbf{t}+v^2,\mathbf{k}}}_{S_1}+\underbrace{X_{\mathbf{t},\mathbf{k}}u_{\mathbf{t}+v^1,\mathbf{k}}X_{\mathbf{t}+\nu^2,\mathbf{k}}}_{S_2}\right.$$
$$\left.+\underbrace{u_{\mathbf{t},\mathbf{k}}X_{\mathbf{t}+\nu^1,\mathbf{k}}X_{\mathbf{t}+\nu^2,\mathbf{k}}}_{S_3}\right\}=S_1+S_2+S_3,$$

where

$$u_{\mathbf{t},\mathbf{k}}=X_{\mathbf{t}}-X_{\mathbf{t},\mathbf{k}}=\sum_{\substack{|s_i|>k\\ i=1,\dots,d}}\psi_{\mathbf{s}}\varepsilon_{\mathbf{t}-\mathbf{s}}.$$

It is easy to show that as $k\to\infty$,

$$Var\left(S_1+S_2+S_3\right)\le 3\left\{VarS_1+VarS_2+VarS_3\right\}\to 0,$$

which gives (2.8) and hence gives (iii).

It remains to show (ii). By (2.5), $Z_k\sim\mathcal{N}(0,\sigma^2_{X,\mathbf{k}}(\nu^1,\nu^2))$. Observe that if we put $\sigma^2_X(\nu^1,\nu^2)=\lim_{k\to\infty}\sigma^2_{X,\mathbf{k}}(\nu^1,\nu^2)$, it is not difficult to show that

$$\sigma^2_X(\nu^1,\nu^2)=\sum_{\substack{\tau_i=-\infty\\ i=1,\dots,d}}^{+\infty}\{C_{6X}(\nu^1,\nu^2,\tau,\tau+\nu^1,\tau+v^2)$$
$$+\{r_{2X}(\nu^1)r_{2X}(\tau-\nu^2)r_{2X}(\nu^2-\nu^1)\}_{15}\}<\infty,$$

then, using the result of Problem 6.16 of [7] $Z_k\Rightarrow Z$, where $Z\sim\mathcal{N}(0,\sigma^2_X(\nu^1,\nu^2))$. From (i), (ii) and (iii) $N^{\frac{d}{2}}(\hat{r}_{3X}(\nu^1,\nu^2)-r_{3X}(\nu^1,\nu^2))\to\mathcal{N}(0,\sigma^2_X(\nu^1,\nu^2))$. □

Appendix D: Chapter head

Setting $\mathbf{q}=(q_1,q_2)$ and $R=(q_1+1)(q_2+1)$, the expressions of $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{c}_1$ and $\mathbf{c}_2$ are:

$$\mathbf{c}_1=\begin{pmatrix} C_{3y}\left(\left(-q_1,-q_2\right),\mathbf{0}\right)\times C_{3y}\left(\mathbf{q},\mathbf{q}\right)\\ C_{3y}\left(\left(-q_1,-q_2+1\right),\mathbf{0}\right)\times C_{3y}\left(\mathbf{q},\mathbf{q}\right)\\ \vdots\\ C_{3y}\left(\left(0,-1\right),\mathbf{0}\right)\times C_{3y}\left(\mathbf{q},\mathbf{q}\right)\end{pmatrix}_{(R-1)\times 1}$$

$$\mathbf{c}_2 = \begin{pmatrix} C_{3y}\left((0,0),\mathbf{0}\right) \times C_{3y}\left(\mathbf{q},\mathbf{q}\right) \\ C_{3y}\left((0,1),\mathbf{0}\right) \times C_{3y}\left(\mathbf{q},\mathbf{q}\right) \\ \vdots \\ C_{3y}\left(\mathbf{q},\mathbf{0}\right) \times C_{3y}\left(\mathbf{q},\mathbf{q}\right) \end{pmatrix}_{R\times 1}$$

$$\mathbf{C}_1 = \begin{pmatrix} C_{3y}^2\left(\mathbf{q},\mathbf{q}\right) & 0 & \ldots & 0 & 0 \\ \vdots & C_{3y}^2\left(\mathbf{q},\mathbf{q}\right) & & & \\ \vdots & & \ddots & & \\ C_{3y}^2\left(\mathbf{q},(0,1)\right) & \ldots & \ldots & C_{3y}^2\left(\mathbf{q},\mathbf{q}\right) & 0 \end{pmatrix}_{(R-1)\times R}$$

$$\mathbf{C}_2 = \begin{pmatrix} C_{3y}^2\left(\mathbf{q},\mathbf{0}\right) & C_{3y}^2\left(\mathbf{q},(0,1)\right) & \ldots & C_{3y}^2\left(\mathbf{q},\mathbf{q}\right) \\ 0 & C_{3y}^2\left(\mathbf{q},\mathbf{0}\right) & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & \ldots & C_{3y}^2\left(\mathbf{q},\mathbf{0}\right) \end{pmatrix}_{R\times R}.$$

References

[1] M. Bashkansky, J. Reintjes, *Statistics and reduction of speckle in optical coherence tomography*, Optics Letters, 25 (2000), 545-547.

[2] R. Basu, G. C. Reinsel, *Properties of the spatial unilateral first-order ARMA model*, Adv. Apll. Prob, 25 (1993), 631-648.

[3] J. E. Besag, *On the correlation structure of some two dimensional stationary processes*, Biometrika, 59 (1973), 43-48.

[4] J. E. Besag, *Spatia linteraction and the statistical analysis of lattice systems*, Journal of the Royal Statistical Society, Series B, 36 (1974), 192-236.

[5] D. R. Billinger, M. Rosenblatt, *Asymptotic theory of kth-order spectra.* In: Harris B (ed) Spectral analysis of time series, Wiley, New York, 1966, 189-232.

[6] A. Boucher, *Considering complex training images with search tree partitioning*, Comput Geosci, 35 (2009), 1151-1158.

[7] Brockwell ????????????????????????????

[8] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, New York: Dover Publications, (1966).

[9] C. B. Burckhardt, *Speckle in Ultrasound B Mode Scans*, IEEE Trans. Sonics Ultrasonics, 25 (1978), 1-6.

[10] TL. Chugunova, LY. Hu, *Multiple-point simulations constrained by continuous auxiliary data*, Math Geosci, 40 (2008), 133-146.

[11] R. Dimitrakopoulos, H. Mustapha, E.Gloaguen, *High-order statistics of spatial random fields: exploring spatial cumulants for modeling complex non-gaussian and non-linear phenomena*, Math Geosci, 42 (2010), 65-99.

[12] J. M. Francos, A. Narasimhan, J. W. Woods, *Maximum likelihood parameter estimation of textures using a Wold-decomposition based model*, IEEE Transactions on Image Processing, 4 (1995), 1655-1666.

[13] J. M. Francos, A. Nehorai, *Interference mitigation in STAP using the two-dimensional wold decomposition model*, IEEE Tran. Signal Proces, 51 (2003), 2461-2470.

[14] G. B. Giannakis, J. M. Mendel, *Cumulant-based order determination of non-Gaussian ARMA models*, IEEE Transactions on Acoustics, Speech, and Signal Processing, 38 (1993), 1411-1423.

[15] G. B. Giannakis, A. Swami, *Identifiability of general ARMA processes using linear cumulant-based estimators*, Automatica, 28 (1992), 771-779.

[16] E. Gloaguen, R. Dimitrakopoulos, *Two-dimensional conditional simulations based on the wavelet decomposition of training images*, Math Geosci, 41 (2009), 679-701.

[17] X. Guyon, *Random fields on a net work. Probability and its applications*, Springer, New York, 1995.

[18] L. Hongwei, C. Qiansheng, *Convergence properties of higher-order moment and cumulant estimates*, Journal of Electronics, 15 (1998), 242-247.

[19] D. Huang, E.A. Swanson, C.P. Lin, J.S. Schuman, W.G. Stinson, W. Chang, M.R. Hee, T. Flotte, K. Gregory, C.A. Pulifito, J.G. Fujimoto, *Optical coherence tomography*, Science, 254 (1991), 1178-1181.

[20] A. Kourgly, A. Belhadj-Aissa, *Segmentation texturale des images urbaines par le biais de l'analyse variographique*, Télédétection, 3 (2003), 337-348.

[21] A. Kourgly, *Analyse variographique et synthèse de textures appliquées aux images numériques*, Thèse de Doctorat, Université des Sciences et de la Technologie Houari Boumediene, Algérie, (2007), 181 pages.

[22] R. Krishnamurthy, J. W. Woods, J. M. Francos, *Adaptive restoration of textured images with mixed spectra*, IEEE Trans. Image Process, 5 (1996), 648-52.

[23] K. S. Lii, M. Rosenblatt, *Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes*, Ann. Star, 10 (1982),1195-1208.

[24] F. Liu, R. W. Picard, *Detecting and segmenting periodic motion Media lab vision and modelling tr-400*, MIT, 1996.

[25] J. M. Mendel, *Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications*, Proc. IEEE, 79 (1991), 278-305.

[26] C. L. Nikias, J. M. Mendel, *Signal processing with higher-order spectra*, IEEE Signal Processing Magazine, 10 (1993), 10-37.

[27] P. W. Mirowski, D. M. Trtzlaff, R. C. Davies, D. S. McCormick, N. Williams, C. Signer, *Stationary scores on training images for multipoint geostatistics*, Math Geosci, 41 (2008), 447-474.

[28] J. Ramberg, B. Schmeiser, *An approximate method for generating asymmetric random variables*, Communications of the ACM, 17 (1975), 78-82.

[29] N. Remy, A. Boucher, J. Wu, *Applied geostatistics with SGeMS: A user's guide*, Cambridge University Press, Cambridge, 2007.

[30] B. D. Ripley, *Spatial statistics*, John Wiley & Sons, New York, 1981.

[31] M. Rosenblatt, *Stationary sequences and random fields*, Birkhäuser, Boston, MR0885090, 1985.

[32] A. Swami, *Modeling of multi-dimensional non-Gaussian processes using cumulants*, J. Multidimensional Signals and Systems, 1 (1990), 1-37.

[33] A. Swami, G. B. Giannakis, J. M. Mendel, *Modeling and parameter estimation of multidimensional non-gaussian processes using cumulants,* Multidimensional Systems and Signal Processing, 1 (1991), 11-37.

[34] P. Treitz, P. Howarth, *Hight spatial remote sensing data for forest ecosystem classification: an examination of spatial scale*, Remote sensing of Environment, 72 (2000), 268-289.

[35] J. M. Schmitt, S. H. Xiang, K. M. Yung, *Speckle in optical coherence tomography*, J. Biomed. Opt, 4 (1999), 95-105.

[36] V. Vapnik, *The nature of statistical learning theory*, New York, NY, USA: Springer-Verla, 1995.

[37] G. Wollstein, L. A. KO. T. H. Paunescu, J. G. Fujimoto, A. Kowaleviez, I. Hartl, S. Beaton, H. Ishikawa, C. Mttox, O. Singh, J. Duker, W. Drexler, J. S. Schuman, *Ultrahigh-resolution optical coherence tomography in glaucoma*, Ophtalmology, 112 (2005), 229-237.

[38] J. Yuan, *Testing Gaussianity and linearity for random fields in the frequency domain*, J. Times Series Analysis, 21 (1999), 723-737.

[39] X. D. Zhang, Y. S. Zhang, *Singular value decompositionbased MA order determination of non-Gaussian ARMA models*, IEEE Trans. Signal Processing, 41, 2657-2664.

[40] X. D. Zhang, Y. S. Zhang, *FIR system identification using higher-order statistics alone*, IEEE Trans. Signal Processing, 42 (1994), 2854-2858.

[41] R. C. Zeimer, J. T. Wilensky, D. K. Gieser, M. A. Viana, *Association between intraocular pressure peaks and progression of visual field loss*, Ophtolmology, 98 (1991), 64-69.

Accepted: July 22, 2020