ALGERIAN REPUBLIC DEMOCRATIC AND POPULAR
MINISTRY OF HIGHER EDUCATION AND
SCIENTIFIC RESEARCH
MENTOURI UNIVERSITY
CONSTANTINE

FACULTY OF SCIENCE
DEPARTEMENT OF MATHEMATICS

THESIS

Presented to obtain the degree of

Doctor of science

In **: MATHEMATICS**

Option : Probability-Statistics

THEME

STATISTICAL INFERENCE IN SPATIAL BILINEAR PROCESSES

Presented by :

Mrs KHARFOUCHI SOUMIA

Jury Review:

M. F. Rahmani, Professor, U.M.C. President

M. A. Bibi, Professor, U.M.C. Supervisor

M. G. D'Aubigny, Professor, U.P.M.F- Grenoble II. Reviewer

Mrs. F. Messaci, Professor, U.M.C. Reviewer

Mrs. H. Guerbyenne, Professor, U.S.T.H.B. Reviewer

M. F. Hamdi, Assistant Professor, U.S.T.H.B. Reviewer

Remerciements

Mes travaux de recherche s'inscrivent dans le cadre du développement des méthodes de la modélisation non-linéaire. Au vu de ces vingt dernières années, je pense que mon parcours a été aussi non-linéaire que mes méthodes de modélisation. En effet, c'est par un chemin détourné que je me suis mise à la recherche et encore plus à la statistique spatiale. J'ai repris les études après une longue rupture mais j'ai eu la chance de croiser au cours de mes deux premières années de post graduation, Mme F. Messaci, professeur au département de mathématiques à l'Université Mentouri Constantine, période au cours de laquelle j'ai été admirablement initiée à la recherche, aujourd'hui encore, elle est là à me soutenir et me fait l'honneur d'accepter de participer à ma commission d'examen.

Mon thème de doctorat m'a été proposé par le professeur A. Bibi. Je tiens à lui exprimer mes profonds remerciements, pour m'avoir permis d'explorer l'univers merveilleux de la statistique spatiale. Il a su m'encourager et m'inciter à davantage de travail et de persévérance grâce à sa patience et à son assistance. J'espère qu'il acceptera encore longtemps de bien vouloir travailler avec moi.

Je remercie aussi tout particulièrement

Monsieur F. Rahmani, maître de conférences au département de mathématiques à l'Université Mentouri Constantine qui me fait l'honneur de présider le Jury.

Monsieur G. D'Aubigny, professeur à l'Université Pierre Mendès-France Grenoble 2, pour m'avoir fait l'honneur de lire mon travail et de le juger. Je lui exprime aussi mon profond respect pour la confiance qu'il m'a accordé.

Madame H. Guerbyenne et Monsieur F. Hamdi, maîtres de conférences à l'USTHB Alger, qui ont consacré du temps à examiner mon travail.

Enfin un grand merci à mes parents qui ont su me donner jeune le goût des études, ainsi qu'à mon mari et mes enfants pour leur patience face à une « maman chercheur » et à toute ma famille et mes amies.

TABLE OF CONTENTS

ABSTRACT . . . . . .

RESUME . . . . . .

INTRODUCTION **2**

0.1 Historical background 2

0.2 Formulation of the problem and motivation 4

0.3 Objectives and organization 7

0.4 Realized works 9

I GENERALITIES **11**

1.1 Concepts and definitions 12

1.1.1 Random fields 12

1.1.2 Stationarity 12

1.1.3 Total order and partial order 14

1.1.4 Indexing subsets 15

1.2 Models for spatial data 15

1.2.1 Spatial ARMA models 16

1.2.2 The 2-D GARCH model 18

1.3 Models for spatio-temporal data 18

1.3.1 The location dependent autoregressive (AR) process 18

1.3.2 A space-time bilinear model with spatial weighting matrix 19

II ON GENERAL MULTIDIMENSIONALLY INDEXED BILINEAR PROCESSES 22

2.1 Introduction 22

2.2 The model 23

2.3 Stationarity 25

2.4 Application 27

2.5 Limit theorem 28

III THE LAN FOR TWO DIMENSIONALLY INDEXED DISCRETE SPATIAL MODELS 33

3.1 Introduction 33

3.2 Locally asymptotic normality for discrete spatial models on a plane 34

3.3 Locally asymptotic normality for the spatial bilinear models 37

3.4 The Asymptotic Distribution of the Maximum Likelihood Estimator 45

IV GENERALIZED SPATIAL AUTOREGRESSIVE MODELS ON Z^2 50

4.1 Introduction 50

4.2 Second order stationarity 52
4.3 Strict stationarity 57
4.4 Higher order moments 62
4.5 Applications 63
4.5.1 The spatial RCA models 65

V A GENERALIZED MOMENTS ESTIMATOR FOR RANDOM COEFFICIENTS AUTOREGRESSIVE MODEL ON A PLANE 68
5.1 Introduction 68
5.2 Overview of the GMM methodology and assumptions 69
5.3 Application of GMM to the 2-D RCA model 71
VI BILINEAR SPATIO-TEMPORAL MODEL 79
6.1 Introduction 79
6.2 The model 81
6.3 Stationarity and causality of spatial bilinear processes 84
6.4 Central Limit Theorem for spatial bilinear process 88
6.5 Numerical Example 89
BIBIOGRAPHY

Notations

ARCH	AutoRegressive Conditional Heteroscedastic
ARMA	AutoRegressive and Moving Average
CAR	Conditional AutoRegressive
CLT	Central Limit Theorem
GARCH	Generalized AutoRegressive Conditional Heteroscedastic
GMM	Generalized Method of Moments
i.i.d.	Independent Identically Distributed
LAN	Local Asymptotic Normality
ML	Maximum Likelihood
OLS	Ordinary Least Squares
RCA	Random Coefficient Autoregressive
SAR	Simultaneous AutoRegressive
SBL_d	d-dimensional Spatial Bilinear
SGAR	Spatial Generalized AutoRegressive
$STBL_d$	d-dimensional Spatial Temporal Bilinear
$A, B, C\ldots$	Matrices
A'	The transpose of A
Vec(M)	The vector obtained by setting down the column of M underneath each other
$M^{\otimes 2} = M \otimes M$	The usual Kronecker product of M with itself
$\rho(M)$	The largest eigenvalues in modulus of the squared matrix M
$I_{(k)}$	The identity matrix of order k
$O_{(k)}$	The $k \times k$ matrix whose entries are zeros
$M = (m_{ij})$	The matrix whose components are m_{ij}
$\lvert M \rvert$	The matrix whose components are $\lvert m_{ij} \rvert$
$A \leq B$	Means $a_{ij} \leq b_{ij}$ for each i and j where a_{ij} and b_{ij} are the components of A

	and B respectively
$\preccurlyeq$	Usual partial order on $\mathbb{Z}^d$
$\prec$	Lexicographic order on $\mathbb{Z}^d$
$S[a,b]$	Subset of all $\boldsymbol{x} \in \mathbb{Z}^d$ such that $a \preccurlyeq \boldsymbol{x} \preccurlyeq b$ (or $a \prec \boldsymbol{x} \prec b$)
$S\langle a,b]$	The subset $S[a,b]\backslash\{a\}$
$S[a,b\rangle$	The subset $S[a,b]\backslash\{b\}$

INTRODUCTION

0.1 Historical background

It's so difficult to define the spatial statistics because of the complexity of the problems that arise in spatial analysis. These problems are neither well defined nor completely solved. Therefore, we are confronted with several alternatives when we have to select the content of an introduction that we hope to be of historical and technical order. We choose to begin this introduction from, both, historical and technical point of view.

Spatial data research has been involved since the beginning of the 20th century. In 1914, Students has suspected a possible relation between localized observations Students (1914). In 1919, some famous agricultural experiences have been developed at the Rothamsted's experimental station situated in the north of Londres Russel (1966). the spatial correlation was distinctly considered to be namely like an undesired property that hid the real process of interest. This latter was often considered as the rate of the plant's growth in different treatments. Because of randomization, ramming and reproduction, spatial correlation was neutralized, rather than being explored. However in many scientific domains as geology Davis (1973), ecology Kershaw (1964) and econometrics Fisher (1971), it's not possible, yet, to randomize, block and reproduce data in order to neutralize the spatial dependence. Besides it's sometimes the spatial nature itself which is the problem of interest Anders (1998).

Before the seventies, there were relatively not much in the literature concerning spatial or spatial-time processes, because we wrongly believed that the location had no impact on the data. P. Whittle (1954) was the first pioneer to integrate some spatial dependence in a regression model by proposing the study of SAR models on a plane. His results were unnoticed at the time. It is to Cliff and Ord that we are indebted, after a series of articles in the late sixties and early seventies, a book summarizing the current knowledge in spatial statistics and spatial econometrics Cliff and Ord(1973). This book is considered by most statisticians as a Bible spatial-time statistics.

The late seventies and eighties were marked by the refinement of the original framework of Cliff and Ord's analysis, particularly through the development of the theory of estimation and

testing. A number of books give a report of this development : Cliff and Ord (1981), Anselin (1988), Cressie (1993), Anselin and Florax (1995).

All these studies have shown that the space influences the natural configurations observed; thus, more the distance between locations was smaller, more the measures taken in two neighboring localities were often similar. This can be compared to the fact that the measures taken in 2 short intervals were generally more similar to those taken at 2 distant dates. The introduction of space in the statistical models is designed to take into account spatial autocorrelation, which refers to the dependence between geographical observations and diversity of behaviors. For example, a study on human health could be done according to the space where people prone to the study live or depending on the location where they work or the journey they conduct weekly; each choice has a direct impact on technics that can be used for spatial analysis and conclusions that can be obtained.

However, the question to be asked is : what is a spatial autocorrelation? And how is it addressed? In fact, there is no clear definition of spatial autocorrelation, its meaning is contextual and has several facets. In a purely geographical context, Griffith (1992) called spatial autocorrelation, the correlation of a variable with itself from geographical ordering. Anselin and Bera (1998) propose an intuitive definition of spatial autocorrelation. For them it would be positive if there is a trend towards concentration in the space of low or high values of a random variable. However, it would be negative if each location tends to be surrounded by neighboring locations for which the random variable takes very different values. Finally, the absence of spatial autocorrelation indicates that the spatial distribution of values of the variable is random. For others, it is also the presence of spatial effects which mean that there is a functional relationship between what happens at a point in the space and what happens elsewhere. its presence can be seen through a spatial model or using a statistic such as Moran's I statistic. This latter is the most widely used and is calculated from the residues, written in the form of a rate in which the numerator is interpreted as the weighted covariance between neighboring units while the denominator is the total variance.

The choice of the model of the spatial field is either an assumption done on the probability distribution that suits best the data among a group of other assumptions, or by a parametric

probability distribution fitted to the data. In other words, spatial dependence can be represented as a functional relationship between what happens at one point in space and what happens elsewhere. Spatial heterogeneity occurs when there is no uniformity of the observed effects across space.

Several models have been introduced to represent random fields. The models considered in practice were often of first order or of dimension at most equal to 3 because of the spatial data nature to be modelled and because that only the immediate neighbors of the site are taken into account. Only, spatial modelling is fundamentally different as indicated X. Guyon (1992) in his book "Random fields on a network" of temporal modelling that it is not causal. Indeed, unlike the temporal representation that has enough flexible properties of precedence and succession, the spatial representation has its limits and one must rely on the choice of neighboring values from which a point is supposed to depend in the space, to specify the type of past that we consider; this can be interpreted by the absence of a canonical order relation on the support. It is then necessary :

- To develop new tools,
- To introduce new methods,
- To find additional models.

They are often methods first introduced in the time series and then adapted to the spatial case. Hence in a sense, the spatial series can be regarded as a generalization of the time series, but have some particularities that make their analysis considerably more difficult.

0.2 Formulation of the problem and motivations

If there is autocorrelation, we must then take it into account by specifying a suitable model. The first models of spatial processes introduced are extensions of autoregressive or moving average models used in time series analysis. One large strand of the literature has been studied in the field of specification, estimation and testing of spatial data using Markov and Gibbs models, with as a main application the image restoration. Comes after, the class of simultaneous linear spatial autoregressive (SAR) models and/or the mixed regression spatial autoregressive (MRSAR)

models studied by Cliff and Ord (1973) and Whittle (1954). The class of linear SAR models, which extends time series autocorrelation to spatial dimensions, is of a particular interest in econometrics with the distinguishing feature of simultaneity in econometric equilibrium models.

The quadrantal-type AR random fields, as defined by Whittle (1954), were also examined in detail by Tjostheim (1978-1983), Basu and Reinsel (1992-1993) and Choi (1997). Later, another large strand of the literature has been considered in the field of spatial data using the class of conditional linear autoregression by specifying conditional distributions of the spatial variables; Besag (1974) and the book by Ripley (1981) have found usage in a variety of disciplines including image processing (Geman and Geman (1984); Chellappa (1985)) and economics (Bronars and Jansen (1987)). But as Basu and Reinsel (1993) pointed out, in spite of these varied uses and the further developments based on the CAR models, there are some difficulties with these models. Explicit spatial correlation properties of the CAR models are not easily determined, and the treatment of border cells is a problematic. The MA random fields were treated by Haining (1978) and Moore (1988). The study of ARMA random fields proved to be more complex, and many authors have studied the topic including Huang and Anh (1992) who considered spatial ARMA processes defined from the total lexicographic order. Other kind of spatial ARMA process has been defined without using a notion of order. For example, Etchinson, Pantula and Brownie (1994) considered two-dimensional lattice separable ARMA models, which are in fact "products" of 2 processes indexed over one-dimensional lattice set. The SAR and CAR models were originally used as models on the doubly infinite regular lattice. But as was noted by Wall (2004), when these models are applied to irregular lattices, the effect that the neighborhood structure has on the implied covariance structure is not well understood and has not been explicitly examined.

All these models are part of the linear modelling; however, the abundant technical literature devoted to linear models is restrictive and insufficient for a deep understanding and analysis of spatial phenomena exhibiting nonlinear behaviors. Nonlinear modelling is characterized by its wealth, especially in terms of interpretation and understanding of "abnormal" phenomena. In brief, spatial dynamics exhibit nonlinear features in a way that is similar to time series models.

Thus, recently a new component has emerged in the literature to explore the nonlinear and / or non-Gaussian relationship in the spatial data. We only look at the 2 following examples of

spatial modelling, in an attempt to illustrate some promising nonlinear models and to discuss the relative merits of our nonlinear approach :

- Spatial econometric modelling.

- Image processing.

In the first example, as Granger and Teräsvirta (1993) pointed out, the economic relationships are often nonlinear. Therefore, the class of nonlinear spatial models we are proposing through this research will provide a better universe for more precise modeling of the economic, environmental and social interactions in econometrics. In line with this, Fotheringham, Brunsdon, and Charlton (2002) and McMillen (2003) state that spatial relationships in housing prices data are highly nonlinear. Furthermore, Pede, Florax and Holt (2008) proposed a statistical methodology to formally test for nonlinearity in spatial processes models and allowed for a gradual regime switching structure in the form of a smooth transition autoregressive (STAR) process. They applied their methodology to per capita income data for 3074 U.S. counties provided by the Bureau of Economic Analysis (BEA) over the period 1969-2003 to determine convergence clubs. There is a possibility to extend the research on the spatial STAR models by adding some assumptions on the transition variable, in the transition function of the spatial STAR model, which allows to consider the spatial STAR model as one of the models studied in chapter 5 and 6 which is a spatial RCA model. The opportunity to express spatial STAR models as spatial RCA models would allow to apply the estimation method proposed for spatial RCA models in chapter 6, and to promote the approach by spatial STAR models (i.e. spatial models that incorporate nonlinearities in the form of regime switching) which remains not widely applied in spite of the popularity of STAR models in depicting nonlinear models.

The second example that is the image processing, and more specifically, the treatment of textural information such as images of geographical regions that allow the production of maps, and in general, almost all the images of the earth are naturally rich in texture, level of gray, ect. It is usual for the textures to consider the pixel values observed as a realization of a stationary Markov random field or an AR-2D model Bickel and Levina (2006) and Kashyap and Eom (1988). The objective is to estimate this random field and to simulate an approximate trajectory for reproducing a similar model but of larger texture. However, the observation of

certain corrupted images suggests the need for a stronger heteroscedasticity in order to decrease the random noise, while keeping the important image features. This leads to various researches in this field to issue the provocative claim that "*all image processing is nonlinear*". In this sense, Noiboar and Cohen (2005) introduced a two-dimensional Generalized Autoregressive Conditional Heteroscedastic (GARCH) model for clutter modeling and anomaly detection. Amirmazlaghani and Amindavar (2010) used the GARCH model for statistical modeling of wavelet coefficients in order to remove noise from digital images.

0.3 Objectives and organization

Often, the spatial sites on which data are collected are irregularly positioned, but with the increasing use of computer technology, data on a regular grid and measured on a continuous scale are becoming more and more common. As indicated by Tjostheim (1978), in some situations at least, irregulary spaced data may be replaced by data on a regular grid using interpolation techniques Delfiner and Delhomme (1975). Hence in this work, we are interested in nonlinear spatial processes on an infinite regular lattice, they are unilateral by construction, hence they are quite analogous to the well understood bilinear, RCAR and GARCH time series models defined on the integers. Furthermore, they are, also, simple extensions to the nonlinear case of the simultaneous autoregressive schemes. Our developments are not based on the CAR models, which are more general than the SAR models, because constructing conditional nonlinear autoregression models though specification of conditional distributions are difficult. For instance, in the spatial case, a conditional Gaussian Markov model essentially implies linearity (Gao, Lu and Tjostheim (2006)).

The objective of our work is to contribute in an effective development of new spatial models that can take into account the presence of two important effects :

- The effect of spatial dependence.

- The effect of nonlinear behavior of data from complex phenomena.

We try to show the richness of the nonlinear approach in the analysis of spatial data. Also, after the presentation of all these models and the necessary tools for their probabilistic study,

we develop some of them with a view to justify the necessity of proposing nonlinear spatial and sptio-temporal data models for modelling complex interactions that may not be modelled by using existing parametric linear models. Estimation procedures for these models can be borrowed from the regression literature because they can be viewed as regression models. After becoming familiar with these models, in future works, we can start thinking about possible practical problems where these models might be applied. It will be also an interesting problem to try to extend the theory of this work from the 'simultaneously' specified models here to conditional probability models.

Following this zipped introduction, six chapters make up this manuscript. The first one is devoted to the definition of various terms related to the whole study in this work, and an overview on various models of interest in the next chapters is given.

In the second chapter, we introduce our first nonlinear spatial model. This model generalizes to a spatial framework, the standard bilinear model introduced by Granger and Andersan (1978) in a temporal context. We first establish the necessary conditions of existence of stationary and ergodic solutions. We apply these conditions to some subclasses of spatial GARCH models which gives relatively weak stationary conditions for this subclass. Finally, we end up this chapter with a central limit theorem based on the CLT of Dedecker (1998) for stationary random fields of random variables belonging to $L^2(P)$.

In chapter 3, we use the central limit theorem of Huang (1992) for spatial martingale differences to establish the LAN property for a general two-dimensional discrete model. We then apply this result to the spatial bilinear model.

Chapter 4 aims to build up new model of random field on the plane of generalized autoregressive type that can include as special cases a variety of models of nonlinear random fields. We establish some results concerning in particular the second and strict stationarity, and the existence of higher order moments.

Chapter 5 proposes the GMM methodology to estimate the parameters of a spatial RCA process. The strong consistency and the asymptotic normality of the estimators are derived under optimal conditions.

The final chapter presents the space-time bilinear model. Under a minimal set of conditions,

we establish the existence of a stationary (over time) solution. We then deduce a central limit theorem for this new class of spatio-temporal models.

0.4 Realized works

Publications:

- **Kharfouchi S**., and Kimouche K., (2011). Central limit theorem associated with bilinear random fields. *Random Oper. Stoch. Equ.* Vol. 19, No. 2, 157-163. Germany

- **Kharfouchi S**., and Mehri H., (2012). Conditions for the existence of stationary and causal space-time bilinear model and Central Limit Theorem. *Int. J. Mathematics in Operational Research,* Vol. 4, No. 6, 638-650. USA

Submitted papers :

- **Kharfouchi S**. Stationarity of generalized Autoregressive processes on Z^2.

- **Kharfouchi S.** The local asymptotic normality for two dimensionally indexed bilinear models.

International conferences :

- **Kharfouchi S**. Sur des propriétés probabilistes des processus bilinéaires spatiaux. LAMOS5, Béjaia, 2007.

- **Kharfouchi S**. Etude d'un modèle bilinéaire spatial du premier ordre". COSI´2008, Tizi-ouzou, 2008.

- **Kharfouchi S**. Conditions d'existence et de stationnarité d'une classe de modèles bilinéaires et linéarisation. ISOR'2008, Bab-ezouar, 2008

- **Kharfouchi S**. Procédure d'identification de l'ordre d'un modèle bilinéaire spatial linéarisé. CISPA 2008, Constantine, 2008.

- **Kharfouchi S**. Estimation par la méthode des moindres carré conditionnelle d'un modèle bilinéaire spatial simple. CIMA'09, Annaba, 2009.

- **Kharfouchi S**. Central limit theorem associated with space-time bilinear process. CIRO'10, Marrakech, 2010.

- **Kharfouchi S**. Central limit theorem associated with bilinear random fields. SAAP2010,

Hammamet, 2010.

- **Kharfouchi S**. The LAN for two dimensionally indexed autoregressive models. TATAM11, Sousse, 2011.

- **Kharfouchi S**. Stationarity of Generalized Autoregressive Processes on Z^2. ICAAA 2011, Istanbull, 2011.

Chapter 1

Generalities

In this chapter, we want to present the main methodological concepts that have been developed to allow the study of random fields, the later being the probability tool for modeling spatial data. The field of spatial statistics is too vast to be covered by a simple thesis, we chose to restrict our study mainly to the analysis of discrete spatial stochastic processes. This chapter is organized into three sections. In Section 1, we give definitions, notations and tools that will be used throughout the thesis. We are particularly interested in the class of discrete random fields, thus in Section 2, we present two examples of discrete spatial models that will be useful for the introduction of the models that we will study in the following chapters. We devote Section 3 to the study of a situation described by the observations collected in different positions, but where time is not fixed, then we have observations that vary over time and space resulting in spatio-temporal data. We will emphasize the approach that builds a spatio-temporal model assuming that if the position is fixed, the observation at that position can be seen as a time series. The later approach has inspired some statisticians to define a spatio-temporal process in terms of its temporal dynamics. As an example, we present the location dependent autoregressive model studied by Suhasini Subba Rao (2007), and that we will generalize to a bilinear space-time model in Chapter 5. We will also present the bilinear space-time model studied and applied by Dai and Billard (1998) to the spatial spread of monthly surveillance data for mumps over 1971–1988 in 12 states of the USA.

1.1 Concepts and definitions

In this section, we introduce the definitions and notations that we will use throughout this thesis and which we have mainly taken from the books of Guyon (1995) and Geaton and Guyon (2008).

1.1.1 Random fields

Definition 1.1 *Let $(\Omega, \mathcal{A}, P)$ be a probability space and D a domain on $\mathbb{R}^d$. A random field on D is a collection of random variables on $(\Omega, \mathcal{A}, P)$, such that for all s of D is associated X_s, which is denoted $\{X_s, s \in D\}$.*

A random field is totally characterized by the joint distribution of any finite subset $(X_{s_1}, X_{s_2}, \ldots, X_{s_n})$ in a consistent way, requiring

$$F_{s_1,\ldots,s_n}(A_1 \times \ldots \times A_n) = P((X_{s_1}, \ldots, X_{s_n}) \in A_1 \times \ldots \times A_n);$$

where $A_1, \ldots, A_n$ are borelians of $\mathbb{R}$, such a specification is called the distribution of the process. A discrete random field X is a process taking random values on the lattice $\mathbb{Z}^d$.

1.1.2 Stationarity

In time series, inference of the parameters is done by independent repetition of the data. For spatial series, in practice, we must infer the properties of the later to a single realization. For example **an** episode of ozone pollution, **a** particular agricultural region, **a** plant epidemic..... In order to achieve the statistical inference for a single event, we must play on the repetition of certain characteristics from a point to another in space, instead of relying on independent repetitions from a moment to another. Then, we set the following stationarity and ergodicity assumptions.

Strict stationarity

Definition 1.2 *A random field is strictly stationary if all $finite-dimensional$ laws are invariant for any translation, i.e.,*
$(X_{s_1}, X_{s_2}, \ldots, X_{s_n})$ and $(X_{s_1+h}, X_{s_2+h}, \ldots, X_{s_n+h})$ have the same law for all n-tuple $s_1, \ldots, s_n$ and all $h \in D$.

This hypothesis hardly verifiable, is only used in some cases encountered in practice such as the non-square integrable fields. It is rarely used elsewhere because it has one major flaw: it says nothing about the existence of the moments, while the expectation and the variance are quantities needed for statistical analysis of data, and secondly it must be checked for n greater than the number of available data. We prefer the other hypothesis of second order stationarity which is lower.

Second order stationarity (or weak)

Definition 1.3 *A square integrable random field, X_s on $\mathbb{R}^d$ is second order stationary if the two first moments exist and are invariant by translation:*

$$\begin{aligned} E(X_s) &= m \ \forall s \in \mathbb{R}^d \\ Cov(X_u, X_v) &= \gamma(u, v) = Cov(X_{u+\tau}, X_{v+\tau}) \end{aligned}$$

for each u, v and $\tau \in \mathbb{R}^d$.

As $\gamma(u, v) = \gamma(u - v, 0)$ for all u, v, τ in $\mathbb{R}^d$ for a weak stationary random field, it will be convenient to redefine the autocovariance function as a function of one argument as fol*l*ows,

$$\gamma(h) = \gamma(h, 0).$$

If X is strictly stationary and if $X \in L^2$, then X is second order stationary. The reverse is not true in general but the two concepts coincide if X is Gaussian.

Ergodicity

The ergodicity is a property that reinforces the notion of stationarity, which overcomes the restriction of a unique realization.

Definition 1.4 *A stationary random field X_s on $\mathbb{R}^d$ is called ergodic for the mean if the spatial mean on a domain D_n ((D_n) is an increasing sequence of bounded convex such that $d(D_n) \to \infty$) converges to its mathematical expectation when $n \to \infty$. i.e., $\overline{X}_n = \frac{1}{|D_n|}\int_{D_n} X_s ds \to E(X_0) = \mu$ when $n \to \infty$, where $|D_n|$ denotes the volume of D (its Lebesque measure) and where the convergence is mean square convergence.*

Example 1.1 *$X = \{X_s, s \in \mathbb{Z}^d\}$, X_s i.i.d. is an example of ergodic process.*

After introducing the concepts of random fields, stationarity and ergodicity in the most general case (random fields indexed by $\mathbb{R}^d$), go back to the context of this doctoral work which is the area of random fields indexed by $\mathbb{Z}^d$. On $\mathbb{Z}^d$, we define in the next paragraph, the usual partial order and the lexicographic order. These notions of order will allow us, even artificially, to structure the space by a certain order that will play the role of past and present or more specifically of precedence and succession of the observations.

1.1.3 Total order and partial order

When the random processes are indexed by $\mathbb{Z}$, the notions of future and past are clear and natural. However when the random process is indexed by $\mathbb{Z}^d$ ($d \geq 2$), several formulations can be defined according to the order used on $\mathbb{Z}^d$. We give here the two orders that we will use, to define some useful sets to the description of the models that we introduce in the next chapters.

The usual partial order on $\mathbb{Z}^d$ is defined for two points $s = (s_1, \ldots, s_d)$ and $t = (t_1, \ldots, t_d)$ of $\mathbb{Z}^d$ by

$$s \preccurlyeq t \text{ (respectively } s \prec t)$$

if for all $i = 1, \ldots, d$

$$s_i \leq t_i \text{ (respectively } s_i < t_i).$$

This order is called the quarter plane order.

The lexicographic order on $\mathbb{Z}^d$ is defined for two points $s = (s_1, s')$ and $t = (t_1, t')$ of $\mathbb{Z}^d$ such that s_1 and t_1 belong to $\mathbb{Z}$ by

$$s \prec t$$

if and only if $s_1 < t_1$ and $s' \prec t'$ in $\mathbb{Z}^{d-1}$. This order is called the half plane order.

1.1.4 Indexing subsets

For two elements a, b of $\mathbb{Z}^d$, such that $a \lessdot b$ and $a \neq b$ ($\lessdot$ being one of the orders defined in the previous section), we consider now the following subsets of $\mathbb{Z}^d$:

$$\begin{aligned}
S\left[a,b\right] &= \left\{x \in \mathbb{Z}^d : a \lessdot x \lessdot b\right\}, \\
S\left[a,\infty\right] &= \left\{x \in \mathbb{Z}^d : a \lessdot x\right\}, \\
S\left[\infty,a\right] &= \left\{x \in \mathbb{Z}^d : x \lessdot a\right\}, \\
S\left\langle a,b\right] &= S\left[a,b\right] \diagdown \left\{a\right\}, \\
S\left[a,b\right\rangle &= S\left[a,b\right] \diagdown \left\{b\right\}, \\
S\left\langle a,\infty\right] &= S\left[a,\infty\right] \diagdown \left\{a\right\}.
\end{aligned}$$

$S\left[a,b\right]$ allows us to define a notion of spatial window which extends the classical notion of time interval used in the time frame, $S\left[a,\infty\right]$ generalizes the notion of future to a spatial framework, finally $S\left[\infty,a\right]$ allows to give an equivalent to the notion of the past of an observation time when it comes to an observed site a.

1.2 Models for spatial data

The purpose of this section and the next one is to provide an overview of the main models which we have used as a starting point for the introduction of the models that we study in the following chapters. The models that we will study are variants or natural extensions of the models we

present in this section, they are suggested by the constraints and needs of a theoretical and / or practice order.

1.2.1 Spatial ARMA models

First define a linear random field. A random field $(X_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^d}$ is linear if for all $\mathbf{t}$,

$$X_{\mathbf{t}} = \sum_{\mathbf{j}\in\mathbb{Z}^d} \omega_{\mathbf{j}}\epsilon_{\mathbf{t}-\mathbf{j}} \tag{1.2.1}$$

where $(\epsilon_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^d}$ is a family of independent and identically distributed (i.i.d) centred random variables with variance σ^2 and $\sum|\omega_{\mathbf{j}}| < \infty$. Linear random fields as defined above are strictly stationary. They are called MA random fields

Now, we propose to recall the results obtained for different ARMA fields. Several kind of spatial ARMA models may be defined according to the order chosen on the lattice $\mathbb{Z}^d$. With the usual partial order on $\mathbb{Z}^d$, we obtain the spatial autoregressive ARMA model having quadrantal representation where only the case of AR models has been widely studied, in particular by Tjostheim (1978, 83), Ha and Newton (1993), Guyon (1995) and Choi (1997).

Definition 1.5 *We say that a square integrable random field $(X_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^d}$ is a spatial unilateral ARMA(p,q) with parameters $p, q \in \mathbb{Z}^d$ if it satisfies the following equation*

$$X_{\mathbf{t}} - \sum_{\mathbf{j}\in S\langle\mathbf{0},\mathbf{p}]} \phi_{\mathbf{j}} X_{\mathbf{t}-\mathbf{j}} = \epsilon_{\mathbf{t}} + \sum_{\mathbf{k}\in S\langle\mathbf{0},\mathbf{q}]} \theta_{\mathbf{k}}\epsilon_{\mathbf{t}-\mathbf{k}}$$

where $(\epsilon_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^d}$ is a square integrable stationary field satisfying for all $\boldsymbol{s}, \boldsymbol{t} \in \mathbb{Z}^d$, $\mathbf{s} \neq \mathbf{t}$

$$E(\epsilon_{\mathbf{t}}) = 0,\ E(\epsilon_{\mathbf{t}}\epsilon_{\mathbf{s}}) = 0.$$

*If $\mathbf{q}$ (respectively $\mathbf{p}$) is **0**, the sum on $S\langle\mathbf{0},\mathbf{q}]$ (respectively $S\langle\mathbf{0},\mathbf{p}]$) is assumed zero and the random field is said AR(**p**) (respectively MA(**q**)).*
If $S\langle\mathbf{0},\mathbf{p}]$ and $S\langle\mathbf{0},\mathbf{q}]$ are defined from the lexicographic order (also called half-plane order) the ARMA model is called unilateral half-plane.

For the inference study of these models, Tjostheim (1983) needed to assume that the ARMA random field admits an unilateral representation of type (1.2.1). This leads to the notion of causal ARMA random field

Definition 1.6 *The ARMA random field is called causal if it has the following unilateral expansion*

$$X_{\mathbf{t}} = \sum_{\mathbf{j} \in S[\mathbf{0}, \infty]} \psi_{\mathbf{j}} \epsilon_{\mathbf{t}-\mathbf{j}} \ (\psi_{\mathbf{0}} = 1)$$

with

$$\sum_{\mathbf{j} \in S[\mathbf{0}, \infty]} \left| \psi_{\mathbf{j}} \right| < \infty.$$

For $z = (z_1, \ldots, z_d) \in \mathbb{C}^d$, denote $\Phi(z) = 1 - \sum_{\mathbf{j} \in S\langle \mathbf{0}, p]} \phi_{\mathbf{j}} z^{\mathbf{j}}$ the autoregressive polynomial and $\theta(z) = 1 + \sum_{\mathbf{j} \in S\langle \mathbf{0}, q]} \theta_{\mathbf{j}} z^{\mathbf{j}}$ the moving average polynomial. A sufficient condition for the random field to be causal was given in Tjostheim (1978). We recall this condition in the following proposition.

Proposition 1.1 *Let $(X_{\mathbf{t}})_{\mathbf{t} \in \mathbb{Z}^d}$ be an ARMA(p,q) random field. Then, a sufficient condition for the random field to be causal is that the autoregressive polynomial has no zeros in the closure of the open disc D^d in $\mathbb{C}^d$, i.e., the autoregressive polynomial $\Phi(z)$ is not zero if $|z_1| \leq 1$ or $|z_2| \leq 1$ or... $|z_d| \leq 1$.*

Using the lexicographic order, Huang and Anh (1992) considered nonsymmetric half-plane ARMA models. They gave a method based on an inverse model associated with the original one to estimate the orders and the parameters.

Other kind of spatial ARMA process, not depending on the order on $\mathbb{Z}^d$, can also be defined. Etchison, Pantula and Brownie (1994) were interested in two-dimensional lattice separable ARMA models (i.e., "products" of two ARMA processes indexed over one-dimensional). They discussed the use of the sample PACF (Partial AutoCorrelation Function) in identifying the orders of a spatial autoregressive process.

1.2.2 The 2-D GARCH model

The 2-D GARCH (p_1, p_2, q_1, q_2) process is defined as:

$$\begin{aligned} \varepsilon_{(i,j)} &= \sqrt{h_{(i,j)}}\eta_{(i,j)} \\ h_{(i,j)} &= \alpha_0 + \sum_{(k,l)\in S]\mathbf{0},(p_1,p_2)]} \alpha_{kl}\varepsilon^2_{(i-k,j-l)} + \sum_{(k,l)\in S]\mathbf{0},(q_1,q_2)]} \beta_{kl}h_{(i-k,j-l)}, \end{aligned} \tag{1.2.2}$$

$\alpha_0 > 0$, $\alpha_{kl} \geq 0$, $k, l \in S]\mathbf{0}, (p_1, p_2)]$ and $\beta_{kl} \geq 0$, $k, l \in S]\mathbf{0}, (q_1, q_2)]$ where $S]\mathbf{0}, (p_1, p_2)]$ and $S]\mathbf{0}, (q_1, q_2)]$ are defined by the partial order $\preceq$. And where $h_{(i,j)}$ is the conditional variance of $\varepsilon_{(i,j)}$(i.e., $\varepsilon_{(i,j)} \setminus \psi_{(i,j)} \sim \mathcal{N}\left(0, h_{(i,j)}\right)$ where $\psi_{(i,j)}$ denote all the information, namely $\psi_{(i,j)} = \left\{\left(\varepsilon_{(i-k,j-l)}\right)_{(k,l)\prec(p_1,p_2)}, \left(h_{(i-k,j-l)}\right)_{(k,l)\prec(q_1,q_2)}\right\}$, and $\eta_{(i,j)} \sim \mathcal{N}(0,1)$ is a stochastic 2D process independent of $h_{(i,j)}$, $\forall k \leq i, l \leq j$.

The 2-D GARCH model has been introduced by Noiboar and Cohen (2005) for clutter modeling and anomaly detection. They showed that the necessary and sufficient condition for the model (1.2.2) to be weakly stationary is

$$\sum_{(k,l)\in S]\mathbf{0},(p_1,p_2)]} \alpha_{kl} + \sum_{(k,l)\in S]\mathbf{0},(q_1,q_2)]} \beta_{kl} < 1.$$

They also proposed the method of maximum likelihood to estimate its parameters. Note finally, that the 2-D GARCH model has been used for statistical modeling of wavelet coefficients (see, for example, Amirmazlaghani and Amindavar 2010).

1.3 Models for spatio-temporal data

For observations through various locations, several types of spatial models have been developed to describe the spatial dependence. In the situation where time is not fixed, we obtain spatio-temporal data that are frequently used to model the data from the mapping of diseases, air pollution, satellite images, ect ...

1.3.1 The location dependent autoregressive (AR) process

The abundance of information from satellites, and automatic environmental data networks has produced an explosion of data with high time frequency and high space resolution. In such situations, many statisticians have observed that if we fix a location, the observations at that location can be considered as a time series. This inspires them to define a spatio-temporal process in terms of its time dynamics. For example, Suhasini Rao (2007) suppose that for every fixed location, the resulting time series has an AR representation, where the innovation are sample from a spatial process. More precisely, she defines the location dependent autoregressive (AR) process $\left\{X_t(u) : u \in [0,1]^2\right\}_t$, where $X_t(u)$ satisfies the equation

$$X_t(u) = \sum_{i=1}^{p} a_i(u) X_{t-i}(u) + \sigma(u)\xi_t(u) \qquad t = 1, \ldots, T,$$

with $u = (x, y) \in [0,1]^2$, and $\{a_i(u); i = 1, \ldots, p\}$ and $\sigma(.)$ are nonparametric functions. The innovations $\left\{\xi_t(u) : u \in [0,1]^2\right\}$ are supposed independent over time and are spatially stationary processes, with $E[\xi_t(u)] = 0$ and $var[\xi_t(u)] = 1$. She considered the prediction of observations at unknown locations using known neighboring observations. Further she proposed a local least squares based method to estimate the parameters at unobserved locations. This model has been used to fit ozone and house price data in Gilleland and Nychka (2005) and Gelfand et al (2003) respectively. We refer the interested reader to Chen, Fuentes and Davis (2008).

1.3.2 A space-time bilinear model with spatial weighting matrix

We consider here a collection of regions indexed by integers that form a lattice in the geographical area to which the study focuses. The observation on a site is modeled as a combination of observations from nearest neighbors of the considered site through past time. These are Dai and Billard (1998) who first introduced the spatio-temporal bilinear model with spatial weight matrix. They applied it to the spatial spread of monthly surveillance data for mumps over 1971-1988 in twelve states of the U.S.A and more recently Park and Heo (2009), have applied it to monthly precipitation data in South Korea.

The model was defined with the following objectives. First, the definition should contain the most general STARMA models, developed since 1975 by Cliff et al, (1975); Pfeifer and

Deutch (1980a, 1980b), as a special case when the nonlinear part is not present. Second, the definition should also contain such other case the general univariate bilinear model BL (p, q, r, s) corresponding to the case where the sites are not spatially dependent. Third, this definition should adapt to a multiple bilinear (MBL) model proposed by Stensholt and Tjostheim (1987), by making MBL models be a general case of STBL models. Finally, the proposed model should possess some specified properties in order to allow statistical inferences to be made.

Definition 1.7 *(definition of the STBL model)*
Let $\left\{\mathbf{z}(t) = [z_1(t), \ldots, z_n(t)]', \ t = 0, 1 \ldots\right\}$ *an* $n \times 1$ *stochastic vector process. The STBL model has the explicit scalar form*

$$\begin{aligned} z_k(t) &= \sum_{i=1}^{p}\sum_{m=0}^{\lambda_i} \Phi_m^i \left\{\sum_{u=1}^{n} w_{ku}^{(m)} z_u(t-i)\right\} + \sum_{j=1}^{q}\sum_{n=0}^{\eta_j} \theta_n^j \left\{\sum_{v=1}^{n} w_{kv}^{(n)} e(t-j)\right\} \\ &\quad + \sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{m=0}^{\zeta_i}\sum_{n=0}^{\mu_j} \beta_{mn}^{ij} \left\{\sum_{u=1}^{n} w_{ku}^{(m)} z_u(t-i)\right\}\left\{\sum_{v=1}^{n} w_{kv}^{(n)} e(t-j)\right\} + e_k(t), \end{aligned}$$

for $k = 1, \ldots, n$, *where* p *is the autoregressive order,* q *is the moving-average order,* r *is the autoregressive order in the bilinear part,* s *is the moving-average order in the bilinear part,* λ_i *is the spatial order of the autoregressive term at temporal lag* i, η_j *is the spatial order of the moving-average term at temporal lag* j, ζ_i *is the spatial order of the autoregressive term in the bilinear part at temporal lag* i, μ_j *is the spatial order of the moving-average term in the bilinear part at temporal lag* j. Φ_m^i, θ_n^j, *and* β_{mn}^{ij} *are the model's parameters,* $\boldsymbol{W}^{(m)} = \left(w_{ku}^{(m)}\right)$ *is a* $n \times n$ *weighting matrix at spatial order* m, *and* $e(t) = [e_1(t), \ldots, e_n(t)]^T$ *is a sequence of independent identically distributed vector random variables with*

$$\begin{aligned} E[e(t)] &= 0, \\ E\left[e(t)\, e(t+j)^T\right] &= \begin{cases} G & j = 0 \\ 0 & j \neq 0 \end{cases}, \\ E\left[z(t)\, e(t+j)^T\right] &= 0 \qquad j > 0. \end{aligned}$$

Notice here that the limitation using an ordering and weighting scheme is that the weights are determined before a model is identified; hence, this approach is subjective and relies upon the model builder's judgment and experience.

Chapter 2

On general multidimensionally indexed bilinear processes

2.1 Introduction

Unidimensional bilinear (BL) models are widely used for modeling time series $(X_t)_{t\in\mathbb{Z}}$ with sample paths of occasional sharp spikes which are often found in meteorology, oceanography, geology, biology and agriculture (see Subba Rao (1981) and the references therein for further discussion). Several authors has investigated in their probabilistic properties (see Bibi and Gautier (2010), Liu (1989), Pham and Tran (1981), Guegan (1987) and the references therein) and others has investigated in statistical domain as the asymptotic behavior of parameter estimation for instance see Hai-Bin (2005) among others. However, some crucial spatial series we come across in practice may be modeled by some random fields $(X_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^d}$. These series are often non Gaussian and thus nonlinear. During the last two decades there has been a growing interest in modeling some spatial time series via nonlinear models. These models, have been constructed by extending the series $(X_t)_{t\in\mathbb{Z}}$ to multidimensional one $(X_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^d}$. However, in our best knowledge, very few works have been presented on the multidimensional bilinear models. So, the main purpose of the chapter is to extend some probabilistic properties of one-dimensional BL to SBL_d models.

We arrange this chapter in the following manner. In the next section we introduce the general

spatial bilinear process and we impose some conditions on its polynomial representation. Section 3, describes the formal expansion of the process with respect on its Wienner's chaos, we then use this development to give sufficient conditions ensuring the existence of stationary and ergodic solutions. Some examples and particular subclasses are studied. In Section 4, we apply the conditions for the SBL_d model on some subclasses of $SGARCH_d$ models which leads to a weak conditions for stationarity. Finally, we establish the CLT for the general spatial bilinear model based on the CLT of Dedecker (1998).

2.2 The model

The bilinear time series model, introduced by Granger and Andersen (1978), has received particular interest. The behavior of this model is characterized by strong explosions followed by long areas of calm, because of this appealing feature the bilinear model has been widely used in seismological studies, control problems, economic modelling and physics. Let's cite some works that have used bilinear modelling, for example, Subba Rao and Gabr (1984) who have noted that this model gives a reasonably good approximation to seismological and meteorological data, Liu and Brockwell (1988), Liu (1989), Kristensen (2009). The general bilinear model $(X_t)_{t\in\mathbb{Z}}$, $\mathbb{Z} := \{0, \pm 1, \pm 2, ...\}$ is defined by

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

where $(e_t)_{t\in\mathbb{Z}}$ is a sequence of independent and identically distributed $(i.i.d.)$ random variables. By extending the series $(X_t)_{t\in\mathbb{Z}}$ to multidimensional one, we define the $\mathbb{R}-$valued spatial bilinear process $(X_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^d}$ on a probability space $(\Omega, \Im, P)$, denoted by $SBL_d(\mathbf{p}, \mathbf{q}, \mathbf{P}, \mathbf{Q})$, by the following stochastic difference equation

$$X(\mathbf{t}) = a_{\mathbf{0}} + \sum_{\mathbf{i}\in S[\mathbf{1},\mathbf{p}]} a_{\mathbf{i}} X(\mathbf{t}-\mathbf{i}) + \sum_{\mathbf{j}\in S[\mathbf{0},\mathbf{q}]} b_{\mathbf{j}} e(\mathbf{t}-\mathbf{j}) + \sum_{\mathbf{j}\in S[\mathbf{1},\mathbf{Q}]} \sum_{\mathbf{i}\in S[\mathbf{1},\mathbf{P}]} c_{\mathbf{ij}} X(\mathbf{t}-\mathbf{i}) e(\mathbf{t}-\mathbf{j}). \quad (2.2.1)$$

In (2.2.1) $e = (e(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$ is a Gaussian white noise defined on the same probability space $(\Omega, \Im, P)$ with zero mean and variance 1.The set $S[\mathbf{0}, \mathbf{s}]$ is defined using the quarter order over $\mathbb{Z}^d$. Note that different SBL_d representations appear to depend on the order chosen on the lattice $\mathbb{Z}^d$. The $SBL_d(\mathbf{p}, \mathbf{q}, \mathbf{P}, \mathbf{Q})$ includes as special cases several classes of models which have been investigated in the literature:

i. Standard bilinear models $BL(p, q; P, Q)$: these models are obtained by assuming $d = 1$.

ii. The spatial $ARMA_d(\mathbf{p}, \mathbf{q})$ models: these models can be deduced from (2.2.1) by setting $c_{\mathbf{ij}} = 0$ for all $\mathbf{i}$ and $\mathbf{j}$ (cf. Gaetan and Guyon (2008), Yao and Brockwell (2006), Tjostheim (1978, 1983)).

iii. Some spatial $GARCH_d(\mathbf{p}, \mathbf{q})$ models: these models are obtained as a special case from the spatial diagonal bilinear models in which $c_{\mathbf{ij}} = 0$ for all $\mathbf{i} \neq \mathbf{j}$ (cf. Amirzlaghani and Amindavar (2010)).

Throughout, we use the following notations: $\mathfrak{C}$ denote the Hilbert subspace generated by $(e(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$, $\Im = \Im(e) := \sigma\left\{e(\mathbf{t}), \mathbf{t} \in \mathbb{Z}^d\right\}$ the $\sigma-$algebra generated by all $e(\mathbf{t})$, $\Im_{\mathbf{t}}$ the sub $\sigma-$algebra of $\Im$ generated by $\{e(\mathbf{s}), \mathbf{s} \preceq \mathbf{t}\}$ and $\Im(\mathfrak{C}) = \bigoplus_{n\geq 0} \frac{1}{n!}\mathfrak{C}^{\odot n}$ the Fock space over the Hilbert space $\mathfrak{C}$ where $\bigoplus$ denotes the direct orthogonal sum and $\odot$ is the symmetric tensor product of Hilbert subspace with $\mathfrak{C}^{\odot 0} := \mathbb{R}$. The corresponding orthogonal decomposition is called Wiener's chaos. Similarly to one dimensional processes, The main technique to study the series that are nonlinear functionals of a Gaussian process is based on the Wiener's chaos. Indeed, it is well known (see Major (1981)) that the Hilbert space $\mathbb{L}_2(\Im)$ is isomorphic to the space $\Im(\mathfrak{C})$. Thus the algebraic tensor structure of the Hilbert space $\Im(\mathfrak{C})$, "abstract" algebraic construction, is transported to the "concrete" space $\mathbb{L}_2(\Im)$ on which we will work now on. To obtain the decomposition on the Wiener's chaos, we will use ordinary polynomials easier to obtain in practice, which then leads to the representation of volterra (for mor information about Volterra series expansion see Wiener (1958)). We need to write the Equation (2.2.1) in polynomial form as follows. Let $\mathbf{B} = (B_1, ..., B_d)$ be the back shift operator defined, for any $\mathbf{i} = (i_1, ..., i_d) \in \mathbb{Z}^d$, $\mathbf{z} = (z_1, ..., z_d) \in \mathbb{C}^d$, write $B^{\mathbf{i}}X(\mathbf{t}) = \prod_{j=1}^{d} B_j^{i_j} X(t_1, ..., t_d) = X(t_1 - i_1, ..., t_d - i_d) = X(\mathbf{t} - \mathbf{i})$ and $\mathbf{z}^{\mathbf{i}} = \prod_{j=1}^{d} z_j^{i_j}$. Define

$\Theta(\mathbf{z}) = 1 - \sum_{\mathbf{i}\in S[\mathbf{1},\mathbf{p}]} a_{\mathbf{i}}\mathbf{z}^{\mathbf{i}}, \Phi(\mathbf{z}) = \sum_{\mathbf{j}\in S[\mathbf{0},\mathbf{q}]} b_{\mathbf{j}}\mathbf{z}^{\mathbf{j}}$ and $\Psi(\mathbf{z}) = \sum_{\mathbf{j}\in S[\mathbf{1},\mathbf{Q}]}\sum_{\mathbf{i}\in S[\mathbf{1},\mathbf{P}]} c_{\mathbf{ij}}\left[\mathbf{z}^{\mathbf{i}}, \mathbf{z}^{\mathbf{j}}\right]$. Then the Model (2.2.1) can be written in a shorthand notation as:

$$\Theta(\mathbf{B})X(\mathbf{t}) = a_{\mathbf{0}} + \Phi(\mathbf{B})e(\mathbf{t}) + \Psi(\mathbf{B})[X(\mathbf{t}), e(\mathbf{t})] \tag{2.2.2}$$

with $\left[\mathbf{B}^{\mathbf{i}}, \mathbf{B}^{\mathbf{j}}\right][X(\mathbf{t}), e(\mathbf{t})] = X(\mathbf{t}-\mathbf{i})e(\mathbf{t}-\mathbf{j})$. Consider now the following condition

[Condition 1] The autoregressive polynomial $\Theta(\mathbf{z}) \neq 0$ for $|z_i| \leq 1$, $i = 1, ..., d$.

The Condition 1 ensures that the equation $\Theta(\mathbf{z}) = 0$ has its roots outside the circles $|z_i| \leq 1, i = 1, ..., d$. Other conditions ensuring the existence of the roots of polynomial $\Theta(\mathbf{z})$ outside the circles $|z_i| \leq 1, i = 1, ..., d$, can be found in Tjostheim (1983) and Yao and Brackwell (2006). Noting that a multivariable polynomial can be factored into factors which are themselves multivariable polynomials but which cannot be further factored, and these irreducible polynomials are unique to multiplicative constants (see Stetter 2004) Chap. 7, for further discussion).

2.3 Stationarity

Our main result is contained in the following theorem

Theorem 2.1 *For the spatial bilinear process represented by* (2.2.2)*, we assume that the Condition* 1 *hold. Then*

$$\left|\Theta^{-1}(\mathbf{1})\Psi(\mathbf{1})\right| < 1, \tag{2.3.1}$$

constitute a sufficient condition for the existence of stationary and ergodic solution of $(X(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$ *generated by* (2.2.1) *given by* $X(\mathbf{t}) := b_{\mathbf{0}} + \sum_{j=1}^{\infty} X_j(\mathbf{t})$ *where* $b_{\mathbf{0}} = \Theta^{-1}(\mathbf{1})a_{\mathbf{0}}$ *and where* $X_j(\mathbf{t})$ *are the* j*th-components of Wiener's chaos defined recursively by*

$$X_j(\mathbf{t}) = \begin{cases} \Theta^{-1}(\mathbf{B})\{\Phi(\mathbf{B})e(\mathbf{t}) + \Psi(\mathbf{B})[b_{\mathbf{0}}, e(\mathbf{t})]\} & \text{if } j = 1 \\ \Theta^{-1}(\mathbf{B})\Psi(\mathbf{B})[X_{j-1}(\mathbf{t}), e(\mathbf{t})] & \text{otherwise.} \end{cases}$$

Moreover $|X(\mathbf{t})| \leq M$ *for all* $\mathbf{t} \in \mathbb{Z}^d$ *where*

$$M = |b_{\mathbf{0}}| + \left|\Theta^{-1}(\mathbf{1})\right|\left(|\Phi(\mathbf{1})| + |b_{\mathbf{0}}\Psi(\mathbf{1})|\right)\left(1 - \left|\Theta^{-1}(\mathbf{1})\Psi(\mathbf{1})\right|\right)^{-1}$$

and the solution is unique.

Proof To establish this result, we use the so-called "reversion method" used by Subba Rao (1981) to show that the vector bilinear time series model can be written in the Volterra form. We replace $e(\mathbf{t})$ by $\lambda e(\mathbf{t})$ in (2.2.2) we obtain the following

$$\Theta(\mathbf{B})X(\mathbf{t}) = a_{\mathbf{0}} + \lambda\Phi(\mathbf{B})e(\mathbf{t}) + \Psi(\mathbf{B})[X(\mathbf{t}), \lambda e(\mathbf{t})] \tag{2.3.2}$$

where λ is a numerical parameter introduced to facilitate the solution, but ultimately λ is allowed to become unity. We want a solution in the form $X(\mathbf{t}) = b_{\mathbf{0}} + \sum_{i=1}^{\infty} \lambda^i X_i(\mathbf{t})$, we substitute the last series into (2.3.2) and then equating powers of λ on both sides we get $b_{\mathbf{0}} = \Theta^{-1}(\mathbf{1})\, a_{\mathbf{0}}$ and $\Theta(\mathbf{B})X_j(\mathbf{t}) = \Phi(\mathbf{B})e(\mathbf{t}) + \Psi(\mathbf{B})[b_{\mathbf{0}}, e(\mathbf{t})]$ if $j = 1$ and $\Psi(\mathbf{B})[X_{j-1}(\mathbf{t}), e(\mathbf{t})]$ otherwise. Hence

$$X_j(\mathbf{t}) = \begin{cases} \sum\limits_{\mathbf{s}\in S[\mathbf{0},\infty[} g(\mathbf{s})\left(\Phi(\mathbf{B})e(\mathbf{t}-\mathbf{s}) + \Psi(\mathbf{B})[b_{\mathbf{0}}, e(\mathbf{t}-\mathbf{s})]\right) \text{ if } j = 1 \\ \Theta^{-1}(\mathbf{B})\,\Psi(\mathbf{B})[X_{j-1}(\mathbf{t}), e(\mathbf{t})] \text{ otherwise.} \end{cases} \tag{2.3.3}$$

Furthermore, Condition 1 implies that the coefficients $(g(\mathbf{s}))_{\mathbf{s}\in S[\mathbf{0},\infty[}$ defined in (2.3.3) decay at an exponential rate, and in particular $|g(\mathbf{s})| \leq k\rho^{s_1+\cdots+s_d}$ for some $\rho \in]0,1[$ and $k > 0$. This implies that $X_1(\mathbf{t})$ defines a stationary and ergodic process, and hence for all $\mathbf{t} \in \mathbb{Z}^d$, $(X_j(\mathbf{t}))_{j\geq 1}$ is a stationary and ergodic sequence. Since $e(\mathbf{t}) = O_p(1)$, then $|X_1(\mathbf{t})| \leq M_0$ where $M_0 := |\Theta^{-1}(\mathbf{1})|\,(|\Phi(\mathbf{1})| + |b_{\mathbf{0}}\Psi(\mathbf{1})|)$ and $|X_2(\mathbf{t})| \leq |\Theta^{-1}(\mathbf{1})\,\Psi(\mathbf{1})|\,M_0$. By recurrence, it can be shown that $|X_k(\mathbf{t})| \leq |\Theta^{-1}(\mathbf{1})\,\Psi(\mathbf{1})|^{k-1} M_0$ for $k \geq 1$ and hence
$|X(\mathbf{t})| \leq |b_{\mathbf{0}}| + \sum_{k\geq 1} |X_k(\mathbf{t})| \leq |b_{\mathbf{0}}| + (1 - |\Theta^{-1}(\mathbf{1})\,\Psi(\mathbf{1})|)^{-1} M_0$. Now assume that $X^{(1)}(\mathbf{t})$ and $X^{(2)}(\mathbf{t})$ are two stationary solutions to (2.2.2). Define $Y(\mathbf{t}) = X^{(1)}(\mathbf{t}) - X^{(2)}(\mathbf{t})$. Then $(Y(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$ satisfies the equation $\Theta(\mathbf{B})Y(\mathbf{t}) = \Psi(\mathbf{B})[Y(\mathbf{t}), e(\mathbf{t})]$ and the sequence $(Y_j(\mathbf{t}))_{j\geq 1}$ of components of Wiener's chaos are almost surely $(a.s.)$ 0 for all $\mathbf{t} \in \mathbb{Z}^d$. Hence $Y(\mathbf{t}) = 0$. ■

It is seen from this theorem, as in one dimensional bilinear models, that the key element governing the stationarity is independent of the moving average part of the model and of the constant $a_{\mathbf{0}}$, so in what follows, we shall assume $a_{\mathbf{0}} = 0$. Noting that the final solution of $X(\mathbf{t})$

can be written in the form (see Major (1981))

$$X(\mathbf{t}) = \sum_{k=1}^{\infty} \left\{ \sum\nolimits_{(k)} V_k(\mathbf{u}_1, ..., \mathbf{u}_k) \prod_{j=1}^{k} e(\mathbf{t} - \mathbf{u}_j) \right\} \tag{2.3.4}$$

where $V_k(\mathbf{u}_1, ..., \mathbf{u}_k)$ are the kth-order Volterra kernels and $\sum_{(k)}$ is over all $\mathbf{u}_j \in S[\mathbf{0}, \infty[$, $j = 1, ..., k$.

2.4 Application

We here consider a class of spatial $GARCH$ processes ($SGARCH_d$) which can be written on the form of a spatial diagonal bilinear models. This enables us to give a sufficient conditions for stationarity. We consider thus the process $(\epsilon(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$ solving $\epsilon(\mathbf{t}) = \sigma(\mathbf{t})z(\mathbf{t})$ where $(z(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$ is an $i.i.d$ process. The volatility process $(\sigma(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$ is assumed to verify

$$f(\sigma(\mathbf{t})) = a_{\mathbf{0}} + \sum_{\mathbf{i}\in S\langle\mathbf{0},\mathbf{p}]} a_{\mathbf{i}} f(\sigma(\mathbf{t}-\mathbf{i})) + \sum_{\mathbf{i}\in S\langle\mathbf{0},\mathbf{q}]} c_{\mathbf{ii}} f(\sigma(\mathbf{t}-\mathbf{i}))\, h(z(\mathbf{t}-\mathbf{i})) \tag{2.4.1}$$

where $f : \mathbb{R}_+ \to \mathbb{R}$ is strictly monotonic function, $h : \mathbb{R} \to \mathbb{R}$ such that $h(\mathbf{z_t})$ is nondegenerate and bounded and where $a_0 > 0$, $a_{\mathbf{i}} \geq 0$ for all $\mathbf{i} \in S\langle\mathbf{0}, \mathbf{p}]$ and $c_{\mathbf{ii}} \geq 0$ for all $\mathbf{i} \in S\langle\mathbf{0}, \mathbf{q}]$ (see Kristensen 2009)). We observe that $X(\mathbf{t}) = f(\sigma(\mathbf{t}))$ and $e(\mathbf{t}) = h(z(\mathbf{t}))$ is a spatial diagonal bilinear process thus is a special case of (2.2.1). In this case the Condition (2.3.1) become

$$\sum_{\mathbf{i}\in S\langle\mathbf{0},\mathbf{p}]} a_{\mathbf{i}} + \sqrt{E\left|\{h(\mathbf{z_t})\}\right|} \sum_{\mathbf{i}\in S\langle\mathbf{0},\mathbf{q}]} c_{\mathbf{ii}} < 1 \tag{2.4.2}$$

and we have

Proposition 2.1 *Consider the following spatial GARCH models:*

(1): Linear $SGARCH_d$: for which $f(x) = x^2, h(x) = x^2$,

(2): Power $SGARCH_d$: for which $f(x) = x^r$, $h(x) = |x|^r$, $r > 0$,

(3): $L - SGARCH_d$: for which $f(x) = x, h(x) = x$.

and assume that the process $(h(z(\mathbf{t})))_{\mathbf{t}\in\mathbb{Z}^d}$ has a finite second order moments. Then the

Condition 1 *together 2.4.2 are sufficient conditions for that* $(\epsilon(\mathbf{t}))_{\mathbf{t}\in\mathbb{Z}^d}$ *to have a unique stationary and ergodic solution.*

2.5 Limit theorem

Let us consider a sequence of random variables on the probability space $(\Omega, \mathcal{F}, P)$. For k in $\mathbb{Z}^d$, T^k denote the translation operator from Ω to Ω which is defined by: $\left[T^k(\omega)\right]_i = \omega_{i+k}$. An element A of $\mathcal{F}$ is said to be invariant if $T^k(A) = A$ for any k in $\mathbb{Z}^d$. We denote by $\mathcal{I}$ the σ-algebra of all invariant sets.

For any finite subset Γ of $\mathbb{Z}^d$ and for any random variable Z defined on $(\Omega, \mathcal{F}, P)$ with $E(Z) = 0$, let S_Γ be the partial sum defined by :

$$S_\Gamma(Z) = \sum_{k\in\Gamma} Z \circ T^k.$$

Let the sets $\left\{V_i^l : i \in \mathbb{Z}^d,\ l \in \mathbb{N}^*\right\}$ be defined as follows:

$$V_i^1 = \left\{j \in \mathbb{Z}^d : j \prec i\right\},$$

and for $l \geq 2$:

$$V_i^l = V_i^1 \cap \left\{j \in \mathbb{Z}^d :\ |i - \ j| \geq l\right\} \text{ where } |i - \ j| = \max_{1\leq k\leq d} |i_k - \ j_k|.$$

Finally, for any $l \in \mathbb{N}^*$, let $\mathcal{F}_l$ be the σ-algebra defined by: $\mathcal{F}_l = \sigma\left(Z \circ T^j;\ j \in V_0^l\right)$. Throughout $(\Gamma_n)_{n\in\mathbb{N}}$ is a sequence of finite subsets of $\mathbb{Z}^d$ satisfying:

$$\lim_{n\to+\infty} |\Gamma_n| = +\infty \text{ and } \lim_{n\to+\infty} |\Gamma_n|^{-1} |\partial\Gamma_n| = 0,$$

where we denote by $|\Gamma_n|$ the cardinality of this set, and
$\partial\Gamma_n = \{i \in \Gamma_n : \exists j \notin \Gamma_n \text{ such that } |i - \ j| = 1\}$.

Theorem 2.2 *(Dedecker, 1998) If Z is a square integrable r.v. with $E(Z)=0$ such that*

$$\sum_{k\in V_0^1}\left\|Z\circ T^k E\left(Z\middle\backslash\mathcal{F}_{|k|}\right)\right\|_1<+\infty, \tag{2.5.1}$$

where for any k in $\mathbb{Z}^d$, we denote $|k|=\max_{1\leq i\leq d}k_i$.
Then, the sequence $|\Gamma_n|^{\frac{-1}{2}}S_{\Gamma_n}$ converges in distribution to $\sqrt{\eta}\epsilon$ where $\epsilon\sim\mathcal{N}(0,1)$, ϵ is independent of $\mathcal{I}$ and η is the nonnegative and $\mathcal{I}$-measurable random variable given by $\eta=\sum_{k\in\mathbb{Z}^d}E\left(Z\times Z\circ T^k\middle\backslash\mathcal{I}\right)$.

In the following, we apply the previous theorem to the SBL_d model. Let $(X(k))_{k\in\mathbb{Z}^d}$ be the spatial bilinear process satisfying the equation (2.2.1) and the stationary conditions of theorem 2.1.

Set $Y(t)=X(t)-\mu$, where $\mu=EX(t)$. Then $(Y(t))_{t\in\mathbb{Z}^d}$ is a zero mean, strictly stationary process, and $Y(\mathbf{0})$ is a square integrable r.v. We can write $Y(k)=Y(\mathbf{0})\circ T^k$ for any $k\in\mathbb{Z}^d$. For $\Gamma_N=S[\mathbf{1},\mathbf{n}]$; $|\Gamma_N|=N=n_1\times n_2\times\ldots\times n_d$, the partial sum of interest is

$$S_N=\sum_{k\in S[\mathbf{1},\mathbf{n}]}(X(k)-\mu).$$

Theorem 2.3 *Under the stationary conditions of theorem 1, the sequence of random variable $\frac{1}{\sqrt{N}}\sum_{k\in S[\mathbf{1},\mathbf{n}]}(X(k)-\mu)$ converges in distribution to $\sqrt{\eta}\epsilon$, where $\eta=\sum_{k\in\mathbb{Z}^d}E\left(Y(\mathbf{0})\times Y(\mathbf{0})\circ T^k\middle\backslash\mathcal{I}\right)$.*

Proof We just need to verify the following condition

$$\sum_{m\in V_0^1}\left\|Y(m)\,E\left(Y(0)\middle\backslash\mathcal{F}_{|m|}\right)\right\|_1<+\infty,$$

where $|m|=\max_{1\leq i\leq d}|m_i|$ and $\|Z\|_1=E|Z|$ for all random variable Z. In the proof of Theorem 1, it's shown that, under the stationarity condition, for all $\mathbf{t}\in\mathbb{Z}^d$, $X(\mathbf{t})=\sum_{j\geq1}X_j(\mathbf{t})$ where $X_j(\mathbf{t})$ is

given by (2.3.3),without loss of generality we take $q = Q$, hence from (2.3.3) we can deduce that

$$\begin{aligned} X_1(\mathbf{t}) &= f_1\left(e\left(\mathbf{t}-s\right),\ s \in [\mathbf{0},\mathbf{q}]\right) \\ X_j(\mathbf{t}) &= f_j\left(e\left(\mathbf{t}-s\right),\ s \in]\mathbf{0},\mathbf{q}-\mathbf{j}]\right),\ \text{for all } j>1, \end{aligned}$$

where $\mathbf{j} = (0,\ldots,0,j)$ and $f_1(.)$ and $f_j(.)$ are both measurable, hence $\sum_{j=1}^{|\mathbf{m}|-|\mathbf{q}|-1} X_j(\mathbf{0})$ is independent of $\mathcal{F}_{|m|}$, furthermore we can write $Y(\mathbf{0})$ as

$$\begin{aligned} Y(\mathbf{0}) &= \sum_{j=1}^{|\mathbf{m}|-|\mathbf{q}|-1} X_j(\mathbf{0}) + \sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty} X_j(\mathbf{0}) - \mu \\ &= \sum_{j=1}^{|\mathbf{m}|-|\mathbf{q}|-1} X_j(\mathbf{0}) - E\left\{\sum_{j=1}^{|\mathbf{m}|-|\mathbf{q}|-1} X_j(\mathbf{0})\right\} + \sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty} X_j(\mathbf{0}) - E\left\{\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty} X_j(\mathbf{0})\right\}, \end{aligned}$$

for all $\mathbf{m} = (m_1,\ldots,m_d) \succ q$ and $m_d > q_d$, so

$$\begin{aligned} &\sum_{\mathbf{m}\in V_0^1} \left\| Y(\mathbf{m})\, E\left(Y(0) \diagdown \mathcal{F}_{|\mathbf{m}|}\right)\right\|_1 \\ \leq\ &\sum_{\mathbf{m}\in V_0^1} E\left|Y(\mathbf{m})\, E\left\{\left(\sum_{j=1}^{|\mathbf{m}|-|\mathbf{q}|-1} X_j(\mathbf{0}) - E\left\{\sum_{j=1}^{|\mathbf{m}|-|\mathbf{q}|-1} X_j(\mathbf{0})\right\}\right) \diagdown \mathcal{F}_{|\mathbf{m}|}\right\}\right| + \\ &E\left|Y(\mathbf{m})\, E\left\{\left(\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty} X_j(\mathbf{0}) - E\left\{\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty} X_j(\mathbf{0})\right\}\right) \diagdown \mathcal{F}_{|\mathbf{m}|}\right\}\right|. \end{aligned}$$

As $\sum_{j=1}^{|\mathbf{m}|-|\mathbf{q}|-1} X_j(\mathbf{0})$ is independent of $\mathcal{F}_{|\mathbf{m}|}$, we get

$$E\left|Y(m)\, E\left(\sum_{j=1}^{|m|-|q|-1} X_j(\mathbf{0}) - E\left\{\sum_{j=1}^{|m|-|q|-1} X_j(\mathbf{0})\right\}\right) \diagdown \mathcal{F}_{|m|}\right| = 0. \tag{2.5.2}$$

Also by using Schwartz inequality , we get

$$E\left|Y(\mathbf{m})E\left\{\left(\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty}X_j(\mathbf{0})-E\left\{\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty}X_j(\mathbf{0})\right\}\right)\diagdown\mathcal{F}_{|\mathbf{m}|}\right\}\right|$$
$$\leq \left\{E\left[Y(\mathbf{m})\right]^2\right\}^{\frac{1}{2}}\left\{E\left[E\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty}X_j(\mathbf{0})\diagdown\mathcal{F}_{|\mathbf{m}|}-E\left\{\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty}X_j(\mathbf{0})\right\}\right]^2\right\}^{\frac{1}{2}},$$

by using the inequality $E\left\{E\left(Z\diagdown\mathcal{F}\right)-E\left(Z\right)\right\}^2\leq V\left(Z\right)$ and the property $V\left(\sum_k Z\left(k\right)\right)\leq\sum_k V\left(Z\left(k\right)\right)+\left[\sum_k\sqrt{V\left(Z\left(k\right)\right)}\right]^2$, we get

$$E\left|Y(\mathbf{m})E\left(\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty}X_j(\mathbf{0})-E\left\{\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty}X_j(\mathbf{0})\right\}\right)\diagdown\mathcal{F}_{|\mathbf{m}|}\right|$$
$$\leq \left\{E\left[Y(\mathbf{m})\right]^2\right\}^{\frac{1}{2}}V\left(\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty}X_j(\mathbf{0})\right)^{\frac{1}{2}},$$
$$\leq \left\{E\left[Y(\mathbf{m})\right]^2\right\}^{\frac{1}{2}}\left\{\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty}V\left(X_j(\mathbf{0})\right)+\left(\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty}\sqrt{V\left(X_j(\mathbf{0})\right)}\right)^2\right\}^{\frac{1}{2}}$$

since $|X_k(\mathbf{t})|\leq\left|\Theta^{-1}(\mathbf{1})\Psi(\mathbf{1})\right|^{k-1}M_0$ for all $k\geq 1$, and using the property $V\left(Z\right)\leq EZ^2$

$$E\left|Y(m)E\left(\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty}X_j(\mathbf{0})-E\left\{\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty}X_j(\mathbf{0})\right\}\right)\diagdown\mathcal{F}_{|\mathbf{m}|}\right|$$
$$\leq \left\{E\left[Y(\mathbf{m})\right]^2\right\}^{\frac{1}{2}}M_0\left\{2\sum_{j=|\mathbf{m}|-|\mathbf{q}|}^{\infty}\left|\Theta^{-1}(\mathbf{1})\Psi(\mathbf{1})\right|^{j-1}\right\}^{\frac{1}{2}}$$

by the stationary condition $E\left[Y(m)\right]^2$ is a constant independent of m and we have $\sum_{j=|m|-|q|}^{\infty}\left|\Theta^{-1}(\mathbf{1})\Psi(\mathbf{1})\right|^{j-1}=\frac{\left|\Theta^{-1}(\mathbf{1})\Psi(\mathbf{1})\right|^{|m|-|q|-1}}{1-\left|\Theta^{-1}(\mathbf{1})\Psi(\mathbf{1})\right|}$, from (2.5.2) we obtain

$$\sum_{m\in V_0^1}\left\|Y(m)E\left(Y(0)\diagdown\mathcal{F}_{|m|}\right)\right\|_1\leq const.\sum_{m\in V_0^1}\left|\Theta^{-1}(\mathbf{1})\Psi(\mathbf{1})\right|^{|m|}<\infty,$$

which completes the proof. ■

Chapter 3

The LAN for two dimensionally indexed discrete spatial models

3.1 Introduction

The existence of spatial dependence creates difficulties for building the parameter estimates of the regression models presented in this thesis. Despite apparent similarities, the temporal and spatial contexts do not induce the same consequences on the behavior of regression models. One of the feature which distinguishes high-dimensionally indexed processes from one-dimensional time series is illustrated in the following example: For the temporal case, the OLS estimator remains consistent, even for a model with autoregressive dependent variables, as residuals are not autocorrelated. We lose optimality properties such as lack of bias in finite horizon, but these are checked in asymptotic situation. In a spatial context, this result is not true, regardless of the assumptions made on the distribution law of residuals, the OLS estimator in the spatial autoregressive model is not consistent.

Local asymptotic normality consists a theoretical framework which has become a standard tool for proving efficiency of tests and estimators (in particular of the maximum likelihood estimator). Hence, we adopt the LAN approach for a general discrete spatial model with Gaussian errors. We show in section 2, that using Huang central limit theorem for weak martingale-differences, relative to the lexicographic order, the log-likelihood ratios exhibit the typical LAN behavior, so

that the local experiments converge weakly to a Gaussian shift experiment. In section 3, we use this result to establish the LAN property for the spatial bilinear model. Finally, we derive the limit distribution of the maximum likelihood estimators of the parameters of the spatial bilinear models.

3.2 Locally Asymptotic Normality for Discrete Spatial Models on a Plane

Let $(X_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}_+^2}$ be a random field , i.e., a collection of random variables indexed by $2-$dimensional integer lattice $\mathbb{Z}_+^2$ and defined on some parametric probability space $(\Omega, \Im, P_{\underline{\theta}})$ in which $P_{\underline{\theta}}$ is a probability measure indexed by an unknown $\underline{\theta} \in \Theta \subset \mathbb{R}^k$. For any $\mathbf{n} = (n_1, n_2) \in \mathbb{Z}_+^2$, let $(\Omega_{\mathbf{n}}, \Im_{\mathbf{n}}, P_{\mathbf{n},\underline{\theta}})$ be the restriction of $(\Omega, \Im, P_{\underline{\theta}})$ to the sample $X(\mathbf{n}) = (X_{\mathbf{t}})_{\mathbf{t}\in S[\mathbf{1},\mathbf{n}]}$. We can note here that we observe $n_1 n_2$ random variables such that $\left(X_{s_1}, \ldots, X_{s_{n_1 n_2}}\right) \sim P_{\mathbf{n},\underline{\theta}})$.
Let $\underline{\theta}_{\mathbf{n}} = \underline{\theta} + \underline{h}\,(n_1 n_2)^{-\frac{1}{2}}$ where $\underline{h} \in \mathbb{R}^k$ is a local parameter and consider the log-likelihood ratio $\Lambda_{\mathbf{n}}(\underline{\theta}) = \log\left\{f_{\mathbf{n},\underline{\theta}_{\mathbf{n}}}/f_{\mathbf{n},\underline{\theta}}\right\}$ where $f_{\mathbf{n},\underline{\theta}}$ is the density corresponding to $P_{\mathbf{n},\underline{\theta}}$ with respect to some positive measure $\mu_{\mathbf{n}}$ where $\mu_{\mathbf{n}}$ can be one of the elements of the family $\mathcal{P}_{\mathbf{n}} = \{P_{\mathbf{n},\underline{\theta}},\ \underline{\theta} \in \Theta\}$ which dominates it. For all $\mathbf{k} = (k_1, k_2) \in S[\mathbf{0}, \mathbf{n}]$, let's define the conditional density :

$$f_{\mathbf{k},\underline{\theta}}\left(X_{\mathbf{k}} | X_{\mathbf{s}},\ \mathbf{s} \in S[\mathbf{1}, \mathbf{k}[\right) = f_{(k_1,k_2),\underline{\theta}}/f_{(k_1,k_2-1),\underline{\theta}}\ ,$$

and denote $g_{\mathbf{nk}}(\underline{\theta}) = \left\{f_{\mathbf{k},\underline{\theta}_{\mathbf{n}}}\left(X_{\mathbf{k}} | X_{\mathbf{s}},\ \mathbf{s} \in S[\mathbf{0}, \mathbf{k}[\right)\right\}\left\{f_{\mathbf{k},\underline{\theta}}\left(X_{\mathbf{k}} | X_{\mathbf{s}},\ \mathbf{s} \in S[\mathbf{1}, \mathbf{k}[\right)\right\}^{-1}$. Then**,** we have **:**

$$\Lambda_{\mathbf{n}}(\underline{\theta}) = \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} \log g_{\mathbf{nk}}(\underline{\theta})\,.$$

Throughout this work, we consider only sample vectors $X(\mathbf{n}) = \left(X_{s_1}, \ldots, X_{s_{n^2}}\right)$ for all $\mathbf{n} = (n, n)$. With these notations, the one-dimensional regularity LeCam's conditions can be extended to $2-$dimensional one. The tool we use to establish the LAN for general 2**-**dimensionally indexed processes is based on spatial martingale differences of Huang whose definition is recalled below.

Definition 3.1 *A random field* $\{X_{\mathbf{t}},\ \mathbf{t} \in S[\mathbf{1},\mathbf{N}]\}$ *is a spatial martingales differences of Huang if :*

$$\forall \mathbf{t},\ E\left(X_{\mathbf{t}} \backslash X_{\mathbf{s}},\ \mathbf{s} \preceq \mathbf{t},\ \mathbf{s} \neq \mathbf{t}\right) = 0\ a.s.$$

For such martingale, we quote from Huang (1992) the following central limit theorem.

Theorem 3.1 *For* $\mathbf{N} = (N, N)$, *let* $\{X_{\mathbf{t}},\ \mathbf{t} \in S[\mathbf{1},\mathbf{N}]\}$ *be a spatial martingale differences of Huang with :*

$$\frac{1}{N^2} \sum_{\mathbf{t} \in S[\mathbf{1},\mathbf{N}]} X_{\mathbf{t}}^2 \xrightarrow{P} \eta^2 \ \textit{and} \sup_{\mathbf{t}} E\left(|X_{\mathbf{t}}|^2 1_{\{|X_{\mathbf{t}}| > C\}}\right) \to 0 \ \textit{as}\ C \to \infty.$$

Then $\frac{1}{N} \sum_{\mathbf{t} \in S[\mathbf{1},\mathbf{N}]} X_{\mathbf{t}} \underset{N \to \infty}{\xrightarrow{\mathcal{L}}} Z$, *where* $E\left\{\exp\left(itZ\right)\right\} = E\left\{\exp\left(-\frac{1}{2}\eta^2 z^2\right)\right\}$.

Assumptions Under $P_{\mathbf{n},\underline{\theta}}$ **,** we assume that

C1 $\underset{\mathbf{k} \in S[\mathbf{1},\mathbf{n}]}{max} |g_{\mathbf{nk}}(\underline{\theta}) - 1| \xrightarrow{P} 0$;

C2 $\sum_{\mathbf{k} \in S[\mathbf{1},\mathbf{n}]} \left(g_{\mathbf{nk}}(\underline{\theta}) - 1\right)^2 \xrightarrow{P} \tau^2(\underline{\theta})$, where $\tau^2(\underline{\theta})$ is non-random and positive;

C3 There exists a sequence of random variables $\left(Z_{\mathbf{k}}(\underline{\theta}), \mathbf{k} \in S[\mathbf{1},\mathbf{n}]\right)_{\mathbf{n} \in \mathbb{Z}^2}$ such that :

$$\begin{aligned} \sum_{\mathbf{k} \in S[\mathbf{1},\mathbf{n}]} \left(g_{\mathbf{nk}}(\underline{\theta}) - 1\right) &= \frac{1}{n} \sum_{\mathbf{k} \in S[\mathbf{1},\mathbf{n}]} Z_{\mathbf{k}}(\underline{\theta}) + o_p(1) \ \text{and almost surely}\ (a.s), \\ E\left\{Z_{\mathbf{k}}(\underline{\theta}) | Z_{\mathbf{s}}(\underline{\theta}), \mathbf{s} \preceq \mathbf{k},\ \mathbf{s} \neq \mathbf{k}\right\} &= 0 \text{ for any } \mathbf{k} \in S[\mathbf{1},\mathbf{n}]. \end{aligned}$$

C4 $\frac{1}{n^2} \sum_{\mathbf{k} \in S[\mathbf{1},\mathbf{n}]} Z_{\mathbf{k}}^2(\underline{\theta}) \xrightarrow{P} \tau^2(\underline{\theta})$ as $n \to \infty$;

C5 $\sup_{\mathbf{k}} E\left|Z_{\mathbf{k}}(\underline{\theta})\right|^2 1_{\{|Z_{\mathbf{k}}(\underline{\theta})| > C\}} \to 0$ as $C \to \infty$.

Remark 3.1 *Conditions C 1 and C 2 are used to obtain a quadratic approximation for* $\Lambda_{\mathbf{n}}(\underline{\theta})$ *and are the same as the conditions imposed by LeCam (1986) Chap. 10, Sec. 5. Condition C 3 states that the* $\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]}(g_{\mathbf{n}k}(\underline{\theta})-1)$ *can be approximated by an other sum* $\frac{1}{n}\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} Z_{\mathbf{k}}(\underline{\theta})$ *which in turn is a Huang martingale differences. C 4 and C 5 are the regularity conditions for the martingale central limit theorem of Huang.*

Theorem 3.2 *We have the following assertions:*

1. *Under C 1 and C 2, we have* $\Lambda_{\mathbf{n}}(\underline{\theta}) \xrightarrow{P} V_{\mathbf{n}}(\underline{\theta}) - \frac{1}{2}\tau^2(\underline{\theta})$ *where :*

$$V_{\mathbf{n}}(\underline{\theta}) = \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]}(g_{\mathbf{nk}}(\underline{\theta})-1) \tag{3.2.1}$$

2. *Under* $C\ 3 - C\ 5$, $V_{\mathbf{n}}(\underline{\theta}) \overset{d}{\rightsquigarrow} N(0,\tau^2(\underline{\theta}))$,

3. *Under* $C\ 1 - C\ 5$, $\Lambda_{\mathbf{n}}(\underline{\theta}) \overset{d}{\rightsquigarrow} N\left(-\frac{1}{2}\tau^2(\underline{\theta}),\tau^2(\underline{\theta})\right)$.

   All the limits are under $P_{\mathbf{n},\underline{\theta}}$, *as* $n \to \infty$.

Proof (i) We have $\Lambda_{\mathbf{n}}(\underline{\theta}) = \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} \log g_{\mathbf{nk}}(\underline{\theta})$, condition $C\ 1$ ensures that for n large, $(g_{\mathbf{nk}}(\underline{\theta})-1)$ belongs to a neighborhood not too large, of 0. Thus, the Taylor expansion of the function $x \to log(x+1)$ at the point $(g_{\mathbf{nk}}(\underline{\theta})-1)$, to the second order gives

$$\Lambda_{\mathbf{n}}(\underline{\theta}) = \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]}(g_{\mathbf{nk}}(\underline{\theta})-1) - \frac{1}{2}\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]}(g_{\mathbf{nk}}(\underline{\theta})-1)^2 + \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]}(g_{\mathbf{nk}}(\underline{\theta})-1)^2 R(g_{\mathbf{nk}}(\underline{\theta})-1);$$

where the remainder $R(.)$ is such that $R(x) \to 0$ as $x \to 0$. Consequently :

$$\left|\Lambda_{\mathbf{n}}(\underline{\theta}) - \left[V_{\mathbf{n}}(\underline{\theta}) - \frac{1}{2}W^2_{\mathbf{n}}(\underline{\theta})\right]\right| \leq \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]}(g_{\mathbf{nk}}(\underline{\theta})-1)^2 \max_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]}|R(g_{\mathbf{nk}}(\underline{\theta})-1)|,$$

where $V_{\mathbf{n}}(\underline{\theta})$ is as defined in 3.2.1, and $W^2_{\mathbf{n}}(\underline{\theta}) = \sum_{k\in S[\mathbf{1},\mathbf{n}]}(g_{\mathbf{nk}}(\underline{\theta})-1)^2$.

The result follows from $C\ 1$ and $C\ 2$.

(ii) From $(C\ 3)$ we have $V_{\mathbf{n}}(\underline{\theta}) = \frac{1}{n}\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} Z_{\mathbf{k}}(\underline{\theta}) + o_p(1)$ where $Z_{\mathbf{k}}(\underline{\theta})$ are spatial martingale differences of Huang. $(C\ 4)$ and $(C\ 5)$ are sufficient for the central limit theorem of Huang to hold. Consequently :

$$\frac{1}{n}\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} Z_{\mathbf{k}}(\underline{\theta}) \xrightarrow{\mathcal{L}} N\left(0, \tau^2(\underline{\theta})\right),$$

thus :

$$V_{\mathbf{n}}(\underline{\theta}) \xrightarrow{\mathcal{L}} N\left(0, \tau^2(\underline{\theta})\right).$$

(iii) Using (i), $\Lambda_{\mathbf{n}}(\underline{\theta}) = V_{\mathbf{n}}(\underline{\theta}) - \frac{1}{2}\tau^2 + o_p(1)$, and by (ii) $V_{\mathbf{n}}(\underline{\theta}) \xrightarrow{\mathcal{L}} N(0, \tau^2(\underline{\theta}))$, consequently : $\Lambda_{\mathbf{n}}(\underline{\theta}) \xrightarrow{\mathcal{L}} N\left(-\frac{1}{2}\tau^2(\underline{\theta}), \tau^2(\underline{\theta})\right)$. ■

3.3 Locally Asymptotic Normality for the Spatial Bilinear Models

Consider the spatial bilinear model defined by :

$$\forall \mathbf{t} \in \mathbb{Z}^d:\ X_{\mathbf{t}} = a_{\mathbf{0}} + \sum_{\mathbf{i}\in S[\mathbf{1},\mathbf{p}]} a_{\mathbf{i}} X_{\mathbf{t}-\mathbf{i}} + \sum_{\mathbf{j}\in S[\mathbf{0},\mathbf{q}]} b_{\mathbf{j}} e_{\mathbf{t}-\mathbf{j}} + \sum_{\mathbf{j}\in S[\mathbf{1},\mathbf{Q}]}\sum_{\mathbf{i}\in S[\mathbf{1},\mathbf{P}]} c_{\mathbf{ij}} X_{\mathbf{t}-\mathbf{i}} e_{\mathbf{t}-\mathbf{j}} \tag{3.3.1}$$

where $(e_{\mathbf{t}})_{\mathbf{t}\in\mathbb{Z}^d}$ is a Gaussian random field defined on the same probability space $(\Omega, \Im, P)$ with zero mean and variance σ^2. Without loss of generality we put $\mathbf{p} = \mathbf{P}$ and $\mathbf{q} = \mathbf{0}$, Then as in one-dimensional case, the $2D - BL(\mathbf{p}, \mathbf{0}, \mathbf{p}, \mathbf{Q})$ model can be represented in several different state space forms with time-varying matrices. For instance :

$$\underline{X}_{\mathbf{t}} = \underline{e}_{\mathbf{t}} + A_1(\mathbf{t})\underline{X}_{\mathbf{t}-\mathbf{u}_0} + A_2(\mathbf{t})\underline{X}_{\mathbf{t}-\mathbf{u}_1}.$$

where $\underline{X}_{\mathbf{t}}$, $A_1(\mathbf{t})$, $A_2(\mathbf{t})$ and $\underline{e}_{\mathbf{t}}$ are defined in the appendix1.1 and $u_0 = (0,1)$ and $u_1 = (1,0)$.

Lemma 3.1 *1) Let $X_{\mathbf{t}}$, $\mathbf{t} \in \mathbb{Z}^2$ be a random field which satisfies* (3.3.1), *suppose that the following conditions are satisfied*
(i) $\rho(|\Gamma_1|) < 1$ *and* $\rho\left((I - |\Gamma_1|)^{-1}|\Gamma_2|\right) < 1$,
(ii) $\rho(|\Gamma_2|) < 1$ *and* $\rho\left((I - |\Gamma_2|)^{-1}|\Gamma_1|\right) < 1$,

where Γ_1, Γ_2 *are defined in Appendix 1.2, then* $X_{\mathbf{t}}$ *is second order stationary.*

2) Furthermore, the bilinear process in (3.3.1) *is forth order stationary if the following conditions are satisfied*

(*a*) $\rho\left(\underline{\Lambda}_1\right) < 1$ *and* $\rho\left(\left(I - \underline{\Lambda}_1\right)^{-1}\underline{\Lambda}_2\right) < 1$,

(*b*) $\rho\left(\underline{\Lambda}_2\right) < 1$ *and* $\rho\left(\left(I - \underline{\Lambda}_2\right)^{-1}\underline{\Lambda}_1\right) < 1$,

where $\underline{\Lambda}_1$ *and* $\underline{\Lambda}_2$ *are defined as in the proof.*

3) Finally, the invertibility condition for the bilinear process in (3.3.1) *can be obtained by interchanging the position of* $X_{\mathbf{t}}$ *and* $e_{\mathbf{t}}$ *in the stationary conditions, provided that there already exists a strictly stationary solution of the equation* (3.3.1).

Proof *1) To establish the stastionary condition for the models defined by* (3.3.1), *we use a similar approach to that used by several authors (for instance Liu and Brockwell (1988)). To make our presentation simple, we shall consider the particular case where there is only 4 matrices* $B_{(0,1)}$, $B_{(0,2)}$, $B_{(1,0)}$ *and* $B_{(2,0)}$ *which are different from* $O_{((p_1+1)(p_2+1))}$. *Consider now for* $t = (i,j)$ *the stochastic process :*

$$S_{n,m}(i,j) = \begin{cases} 0 \text{ if } n < 0 \text{ or } m < 0 \\ C\underline{e}_{(i,j)} \text{ if } n = 0 \text{ or } m = 0 \\ C\underline{e}_{(i,j)} + \left(A_1 + B_{(0,1)}e_{(i;j-1)} + B_{(0,2)}e_{(i;j-2)}\right) S_{n,m-1}(i,j-1) + \\ \left(A_2 + B_{(1,0)}e_{(i-1,j)} + B_{(2,0)}e_{(i-2,j)}\right) S_{n-1,m}(i-1,j) \text{ if not.} \end{cases}.$$

Put $\Delta_{n,m}(i,j) = S_{n,m}(i,j) - S_{n,m-1}(i,j)$ *(respec.* $\Delta^*_{n,m}(i,j) = S_{n,m}(i,j) - S_{n-1,m}(i,j)$*), we shall consider only the recursion relations* $\Delta_{n,m}(i,j)$ *(the same work can be done for* $\Delta^*_{n,m}(i,j)$*) become :*

$$\Delta_{n,m}(i,j) = \begin{cases} 0 \text{ if } n < 0 \text{ or } m < 0 \\ C\underline{e}_{(i,j)} \text{ if } n = 0 \text{ or } m = 0 \\ \left(A_1 + B_{(0,1)}e_{(i;j-1)} + B_{(0,2)}e_{(i;j-2)}\right) \Delta_{n,m-1}(i,j-1) + \\ \left(A_2 + B_{(1,0)}e_{(i-1,j)} + B_{(2,0)}e_{(i-2,j)}\right) \Delta_{n-1,m}(i-1,j) \text{ if not.} \end{cases} \tag{3.3.2}$$

The L^2 convergence of $\{S_{n,m}(i,j)\}$ can be obtained by the Cauchy criterion, hence a sufficient condition for this convergence is :

$$E\left\|\Delta_{n,m}(i,j)\right\|^2 \to 0 \text{ as } n,\ m \to \infty.$$

If we use a finite number of step-by-step recursion of (3.3.2)*, we get if we note :*

$$\begin{cases} V_{n,m} = E\left(\Delta_{n,m}(i,j)\,\Delta'_{n,m}(i,j)\right) \\ D_{n,m} = E\left(e_{(i,j-1)}\Delta_{n,m}(i,j)\,\Delta'_{n,m}(i,j)\right) \\ F_{n,m} = E\left(e^2_{(i,j-1)}\Delta_{n,m}(i,j)\,\Delta'_{n,m}(i,j)\right) \\ D^*_{n,m} = E\left(e_{(i-1,j)}\Delta_{n,m}(i,j)\,\Delta'_{n,m}(i,j)\right) \\ F^*_{n,m} = E\left(e^2_{(i,j-1)}\Delta_{n,m}(i,j)\,\Delta'_{n,m}(i,j)\right) \end{cases}$$

■

.

and

$$U_{n,m} = \begin{bmatrix} \overrightarrow{V}_{n,m} \\ \overrightarrow{D}_{n,m} \\ \overrightarrow{D}^*_{n,m} \\ \overrightarrow{F}_{n,m} \\ \overrightarrow{F}^*_{n,m} \end{bmatrix},\ n \geq 0, m \geq 3,$$

the 2 - D system described by Fornasini-Marchesini second model with the boundary conditions $U_{0,m}$ and $U_{n,3}$,

$$U_{n,m} = \Gamma_1 U_{n,m-1} + \Gamma_2 U_{n-1,m} \tag{3.3.3}$$

where Γ_1 and Γ_2 are defined in Appendix1.2.

From (3.3.3)*, we obtain*

$$|U_{n,m}| \leq |\Gamma_1|\,|U_{n,m-1}| + |\Gamma_2|\,|U_{n-1,m}|\,. \tag{3.3.4}$$

The right side of (3.3.4) *can be calculated in the following way :*

$$|\Gamma_1|\,|U_{n,m-1}| + |\Gamma_2|\,|U_{n-1,m}| = \sum_{k=1}^{n} T^{n-k,m-2}\,|\Gamma_1|\,|U_{k,3}| + \sum_{l=1}^{m} T^{n-1,m-l}\,|\Gamma_2|\,|U_{0,l}|$$

where the state transition matrix $T^{n,m}$ *is defined as follows :*

1) $T^{0,1} = |\Gamma_1|$ *and* $T^{1,0} = |\Gamma_2|$

2) $T^{n,m} = |\Gamma_2|\,T^{n-1,m} + |\Gamma_1|\,T^{n,m-1}$

3) $T^{0,0} = I$

4) $T^{-n,m} = T^{n,-m} = 0.$

Since matrices $|\Gamma_1|$ *and* $|\Gamma_2|$ *are nonnegative, we have*

$$\sum_{k=1}^{n} T^{n-k,m-2} \leq \sum_{k=0}^{\infty} T^{k,m-2} \text{ and } \sum_{l=1}^{m} T^{n-1,m-l} \leq \sum_{l=0}^{\infty} T^{n-1,l}$$

Then, if $\rho\left(|\Gamma_1|\right) < 1$ *and if* $\rho\left(|\Gamma_2|\right) < 1$ *one can calculate*

$$\begin{aligned}\sum_{k=0}^{\infty} T^{k,m-2} &= \left(I - T^{0,1}\right)^{-1}\left[\left(I - |\Gamma_1|\right)^{-1}|\Gamma_2|\right]^{m-2} \\ &= \left(I - |\Gamma_1|\right)^{-1}\left[\left(I - |\Gamma_1|\right)^{-1}|\Gamma_2|\right]^{m-2},\end{aligned}$$

and

$$\sum_{l=0}^{\infty} T^{n-1,l} = \left(I - |\Gamma_2|\right)^{-1}\left[\left(I - |\Gamma_2|\right)^{-1}|\Gamma_1|\right]^{n-1}.$$

Now, if $\widetilde{M}$ *denotes the sum of the absolute values of all components of* $|M|$*, then using Proposition 2.1 of Liu and Brockwell (1988), if*

$$\lambda_1 = \rho\left(\left(I - |\Gamma_1|\right)^{-1}|\Gamma_2|\right) < 1 \text{ and } \lambda_2 = \rho\left(\left(I - |\Gamma_2|\right)^{-1}|\Gamma_1|\right) < 1,$$

we have the inequality

$$\widetilde{U}_{n,m} \leq c_1\lambda_1^{\frac{m}{2}} + c_2\lambda_2^{\frac{n}{2}},$$

for all $n \in \{0,1,\ldots\}$ *and* $m \in \{3,4,\ldots\}$ *and for some constants* c_1 *and* c_2. *Thus it is easy to see that* $\widetilde{U}_{n,m} \to 0$ *as* $n,m \to \infty$ *if the following conditions are satisfied*

(i) $\rho\left(|\Gamma_1|\right) < 1$ *and* $\rho\left(\left(I - |\Gamma_1|\right)^{-1}|\Gamma_2|\right) < 1$,

(ii) $\rho\left(|\Gamma_2|\right) < 1$ *and* $\rho\left(\left(I - |\Gamma_2|\right)^{-1}|\Gamma_1|\right) < 1$.

Also $E\left\|S_{n,m}(i,j) - S_{n,m-1}(i,j)\right\|^2 = E\left\|\Delta_{n,m}(i,j)\right\|^2 = trace\left(V_{n,m}\right) \leq \tilde{V}_{n,m} \leq \widetilde{U}_{n,m} \to 0$ *as* $n,m \to \infty$ *if* (i) *and* (ii) *are satisfied (note that* $E\left\|\Delta^*_{n,m}(i,j)\right\|^2 \to 0$ *as* $n,m \to \infty$ *if also* (i) *and* (ii) *are satisfied) . From this, it follows that for each* $(i,j) \in \mathbb{Z}^2$, $\{S_{n,m}(i,j)\}$ *converges in* L^2 *as* $n,m \to \infty$ *if the conditions* (i) *and* (ii) *are satisfied. For the general case, the same above arguments can be applied. We just need to replace the sequences* $\{S_{n,m}(i,j)\}$ *by :*

$$\breve{S}_{n,m}(i,j) = \begin{cases} 0 \text{ if } n < 0 \text{ or } m < 0 \\ C\underline{e}_{(i,j)} \text{ if } n = 0 \text{ or } m = 0 \\ C\underline{e}(i,j) + A_1(i,j)\,\breve{S}_{n,m-1}(i,j-1) + A_2(i,j)\,\breve{S}_{n-1,m}(i-1,j) \text{ if not.} \end{cases}$$

Where $A_1(i,j) = A_1 + \sum_{\substack{(l_1,l_2)\in S]\mathbf{0},\mathbf{Q}] \\ l_2 \neq 0}} B_{(l_1,l_2)} e_{(i-l_1,j-l_2)}$ *and* $A_2(i,j) = A_2 + \sum_{1\leq k\leq Q_1} B_{(k,0)} e_{(i-k,j)}$.
2) To establish forth order stationarity, we must establish, under stationary condition, the L^4 *convergence of* $\{S_n(i,j)\}$. *The same argument used in the previous point still applies if the recursion defining* $\{V_{n,m}\}$ *is now replaced by :*

$$W_{n,m} = E\left\{\left(\Delta_{n,m}(i,j)^{\otimes 4}\right)\right\}$$

To obtain a recursive representation for $W_{n,m}$, *we need to evaluate the following vector moments :*

$$\begin{aligned} D^{(t)}_{n,m} &= E\left(e_{\mathbf{u}-\mathbf{q}_t}\Delta_{n,m}(\mathbf{u})^{\otimes 4}\right), \quad t = 1,\ldots,s-1, \\ F^{(lk)}_{n,m} &= E\left(e_{\mathbf{u}-\mathbf{q}_l}e_{\mathbf{u}-\mathbf{q}_k}\Delta_{n,m}(\mathbf{u})^{\otimes 4}\right), \ 1 \leq l \leq k \leq s, \\ G^{(tlk)}_{n,m} &= E\left(e_{\mathbf{u}-\mathbf{q}_t}e_{\mathbf{u}-\mathbf{q}_l}e_{\mathbf{u}-\mathbf{q}_k}\Delta_{n,m}(\mathbf{u})^{\otimes 4}\right), \ 1 \leq t \leq l \leq k \leq s, \\ G^{(tlkh)}_{n,m} &= E\left(e_{\mathbf{u}-\mathbf{q}_t}e_{\mathbf{u}-\mathbf{q}_l}e_{\mathbf{u}-\mathbf{q}_k}e_{\mathbf{u}-\mathbf{q}_h}\Delta_{n,m}(\mathbf{u})^{\otimes 4}\right), \ 1 \leq t \leq l \leq k \leq h \leq s, \end{aligned}$$

where $u \in \mathbb{Z}^2$, $S[\mathbf{0},\mathbf{Q}] = \{\mathbf{0},\mathbf{Q}_1,\ldots,\mathbf{Q}_s = \mathbf{Q}\}$ *i.e.*$\sharp S[0,\mathbf{Q}] = s+1$, *after we define the vector* $U'_{n,m}$ *consisting of the above moment vectors, then we can write the matrix* Λ_1 *and* Λ_2 *such that*

:

$$\left|U'_{n,m}\right| \leq |\Lambda_1| \left|U'_{n,m-1}\right| + |\Lambda_2| \left|U'_{n-1,m}\right|.$$

Thus, the following conditions will ensure the existence of the fourth order stationarity.

a) $\rho(|\Lambda_1|) < 1$ *and* $\rho\left((I - |\Lambda_1|)^{-1} |\Lambda_2|\right) < 1$,

b) $\rho(|\Lambda_2|) < 1$ *and* $\rho\left((I - |\Lambda_2|)^{-1} |\Lambda_1|\right) < 1$.

3) Finally, the invertibility condition for the bilinear process in (3.3.1) *can be obtained by interchanging the position of* $X_{i,j}$ *and* $e_{i,j}$ *in the stationary condition, provided that there already exists a strictly stationary solution of the equation* (3.3.1).

In order to apply Theorem 4.2 to the spatial diagonal bilinear model, we must specify $\underline{\theta}$, $g_{\mathbf{n}k}$ and verify conditions C 1 to C 5 of section 2. We point out here that, $\underline{\theta} = \left(\underline{a}'_0, \ldots, \underline{a}'_{p_1}, \underline{c}'_0, \ldots, \underline{c}'_{p_1}\right)'$, where $\underline{a}_0 = \left(a_{(0,1)}, \ldots, a_{(0,p_2)}\right)'$, $\underline{c}_0 = \left(c_{(0,1)}, \ldots, c_{(0,p_2)}\right)'$, $\underline{a}_k = \left(a_{(k,0)}, \ldots, a_{(k,p_2)}\right)'$, $\underline{c}_k = \left(c_{(k,0)}, \ldots, c_{(k,p_2)}\right)'$ for all $1 \leq k \leq p_1$ and $\underline{h} = \left(\left(h_{\mathbf{s}}\right)_{\mathbf{s}\in]\mathbf{0},\mathbf{p}]}, \left(\underline{h}_{\mathbf{l}}\right)_{\mathbf{l}\in]0,\mathbf{Q}]}\right)'$ with $\underline{h}_{\mathbf{l}} = \left(h_{(p_1,p_2+1)+\mathbf{l}}, \ldots, h_{(p_1,2p_2)+\mathbf{l}}, h_{(p_1+1,p_2)+\mathbf{l}}, h_{(p_1+1,p_2+1)+\mathbf{l}}, \ldots, h_{(p_1+1,2p_2)+\mathbf{l}}, \ldots, h_{(2p_1,2p_2)+\mathbf{l}}\right)$.
Then, it is easily verified that $g_{\mathbf{nk}}(\underline{\theta}) = \exp\left\{H_{\mathbf{nk}}(\underline{\theta})\right\}$ where
$H_{\mathbf{nk}}(\underline{\theta}) = n^{-1}\sigma^{-2} e_{\mathbf{k}} B_{\mathbf{k}}(\underline{\theta}) - \left(2n^2\sigma^2\right)^{-1} B^2_{\mathbf{k}}(\underline{\theta})$ with $B_{\mathbf{k}}(\underline{\theta}) = \sum_{\mathbf{s}\in]\mathbf{0},\mathbf{p}]} h_{\mathbf{s}} X_{\mathbf{k}-\mathbf{s}} + \sum_{\mathbf{s}\in]\mathbf{0},\mathbf{Q}]} [\underline{h}_{\mathbf{s}} \underline{X}_{\mathbf{k}}]\, e_{\mathbf{k}-\mathbf{s}}$.
We shall assume that $X_{\mathbf{0}}$ is fixed at $x_{\mathbf{0}}$ and $e_{\mathbf{0}} = 0$. It will be assumed that throughout, the spatial bilinear model in (3.3.1) is second and forth order stationary (this condition is necessary for the existence of finite variances of the observed process), and also invertible.
Conditions C 1 to C 5 of section 2 will be verified in the following lemma.

Lemma 3.2 *Under second and forth order stationary conditions, we have:*

(a) $\max_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} |H_{\mathbf{nk}}(\underline{\theta})| = o_p(1)$;

(b) $\max_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} |g_{\mathbf{nk}}(\underline{\theta}) - 1| = o_p(1)$.

(c) $\left(n^2\sigma^2\right)^{-1} \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} B^2_{\mathbf{k}}(\underline{\theta}) \xrightarrow{P} \tau(\underline{\theta})^2$, *under* $P_{\mathbf{n},\underline{\theta}}$ *where* $\tau(\underline{\theta})^2 = EB^2_{1,1}(\underline{\theta})/\sigma^2$.

(d) $\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} H^2_{\mathbf{nk}}(\underline{\theta}) \xrightarrow{P} \tau(\underline{\theta})^2$ *and* $\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} \left(g_{\mathbf{nk}}(\underline{\theta}) - 1\right)^2 \xrightarrow{P} \tau(\underline{\theta})^2$.

Proof (a) and (b). The Taylor expansion to the first order of the function exp applied at the point $H_{\mathbf{n}k}$ gives,

$$g_{\mathbf{nk}}(\underline{\theta}) = \exp(H_{\mathbf{nk}}(\underline{\theta})) - 1 = H_{\mathbf{nk}}(\underline{\theta})(1 + R(H_{\mathbf{nk}}(\underline{\theta}))),$$

where $R(x) \to 0$ as $x \to 0$. From this, we have :

$$\begin{aligned} \max_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} |g_{\mathbf{nk}}(\underline{\theta}) - 1| &= \max_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} |H_{\mathbf{nk}}(\underline{\theta})(1 + R(H_{\mathbf{nk}}(\underline{\theta})))| \\ &\leq \left\{ \max_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} |H_{\mathbf{n}k}(\underline{\theta})| \right\} \left\{ \max_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} |1 + R(H_{\mathbf{nk}}(\underline{\theta}))| \right\}. \end{aligned}$$

Then (b) is obtained if we verify (a).

(a) follows from second and forth stationarity and the law of large numbers.

(c) and (d). The results here can be shown using (a), (b) and the same arguments in Basawa and Huang (1997) Lemma 3.2. and Lemma 3.3.

■

Theorem 3.3 *For the model in* (3.3.1)*, under second and forth order stationary conditions, the log-likelihood ratio is locally asymptotically normal in the sense of Theorem 4.2.*

Proof *Lemma 4.2 ensures that the conditions* $(C\ 1)$ *and* $(C\ 2)$ *are satisfied. Consequently, we can write* $\Lambda_{\mathbf{n}}(\underline{\theta}) = V_{\mathbf{n}}(\underline{\theta}) - \frac{1}{2}\tau^2 + o_p(1)$ *where* $V_{\mathbf{n}}(\underline{\theta}) = \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} (g_{\mathbf{nk}}(\underline{\theta}) - 1)$.
It remains to verify conditions $(C\ 3)$ *to* $(C\ 5)$. *Consider :*

$$\begin{aligned} V_{\mathbf{n}}(\underline{\theta}) &= \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} (g_{\mathbf{nk}}(\underline{\theta}) - 1) = \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} (\exp H_{\mathbf{nk}}(\underline{\theta}) - 1) \\ &= \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} H_{\mathbf{nk}}(\underline{\theta}) + \frac{1}{2} \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} (H_{\mathbf{n}k}(\underline{\theta}))^2 + \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} (H_{\mathbf{nk}}(\underline{\theta}))^2 R(H_{\mathbf{nk}}(\underline{\theta})) \end{aligned}$$

where $R(x) \to 0$ *as* $x \to 0$. *We have :*

$$\left| V_{\mathbf{n}}(\underline{\theta}) - \left(\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} H_{\mathbf{nk}}(\underline{\theta}) + \frac{1}{2} \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} (H_{\mathbf{nk}}(\underline{\theta}))^2 \right) \right| \leq \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} (H_{\mathbf{nk}}(\underline{\theta}))^2 \max_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} R(H_{\mathbf{nk}}(\underline{\theta}))$$

but $\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} \left(H_{\mathbf{n}k}\left(\underline{\theta}\right)\right)^2 \max_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} R\left(H_{\mathbf{nk}}\left(\underline{\theta}\right)\right) \xrightarrow{P} 0$ *under* $P_{\mathbf{n},\underline{\theta}}$ *by lemma 4.2, so :*

$$V_{\mathbf{n}}\left(\underline{\theta}\right) = \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} H_{\mathbf{nk}}\left(\underline{\theta}\right) + \frac{1}{2}\tau^2\left(\underline{\theta}\right) + o_p\left(1\right).$$

Now, recall that

$$\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} H_{\mathbf{nk}}\left(\underline{\theta}\right) = n^{-1}\sigma^{-2}\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} e_{\mathbf{k}}B_{\mathbf{k}}\left(\underline{\theta}\right) - \left(2n^2\sigma^2\right)^{-1}\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} B^2_{\mathbf{k}}\left(\underline{\theta}\right),$$

so :

$$\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} H_{\mathbf{nk}}\left(\underline{\theta}\right) = \quad n^{-1}\sigma^{-2}\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} e_{\mathbf{k}}B_{\mathbf{k}}\left(\underline{\theta}\right) \quad - \frac{1}{2}\tau^2\left(\underline{\theta}\right) + o_p\left(1\right),$$

thus, we get $V_{\mathbf{n}}\left(\underline{\theta}\right) = \frac{1}{n}\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} Z_{\mathbf{k}}\left(\underline{\theta}\right) + o_p\left(1\right)$ *where* $Z_{\mathbf{k}}\left(\underline{\theta}\right) = \sigma^{-2}e_{\mathbf{k}}B_{\mathbf{k}}$.

Also :

$$E\left(Z_{\mathbf{k}}\left(\underline{\theta}\right)\backslash X_{\mathbf{s}}\left(\underline{\theta}\right),\ \mathbf{s}\prec\mathbf{k}\right) = \sigma^{-2}B_{\mathbf{k}}Ee_{\mathbf{k}} = 0,$$

then, we obtain $(C\ 3)$.

Now let's verify $(C\ 4)$. *Consider :*

$$\begin{aligned}
& P\left[\left|\frac{1}{n^2}\sum_{\mathbf{k}\in S\in[\mathbf{1},\mathbf{n}]}\left(\left(\sigma^2\right)^{-2}e^2_{\mathbf{k}}B^2_{\mathbf{k}} - \sigma^{-2}B^2_{\mathbf{k}}\right)\right| > \epsilon\right] \\
\leq\ & E\left[\sum_{\mathbf{k}\in S\in[\mathbf{1},\mathbf{n}]}\left(\sigma^{-4}e^2_{\mathbf{k}}B^2_{\mathbf{k}} - \sigma^{-2}B^2_{\mathbf{k}}\right)\right]^2\left(n^2\epsilon\right)^{-2} \\
=\ & \frac{1}{n^4\epsilon^2}\sum_{\mathbf{k}\in S\in[\mathbf{1},\mathbf{n}]} E\left(\sigma^{-4}e^2_{\mathbf{k}} - \sigma^{-2}\right)^2 B^4_{\mathbf{k}} \\
=\ & \frac{1}{n^2\epsilon^2}\sigma^{-4}E\left(e^2_{\mathbf{k}} - 1\right)^2 EB^4_{\mathbf{k}} \to 0 \text{ as } n\to\infty.
\end{aligned}$$

Consequently :

$$\frac{1}{n^2}\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} Z^2_{\mathbf{k}}\left(\underline{\theta}\right) \xrightarrow{P} \tau^2\left(\underline{\theta}\right),\ \text{as } n\to\infty.$$

It now remains to verify $(C5)$. *Since* $EZ_{\mathbf{k}}^2(\underline{\theta}) = \tau^2(\underline{\theta}) < \infty$, *we can state that* $P\left\{\sup_k EZ_{\mathbf{k}}^2(\underline{\theta}) < \infty\right\} = 1$, *furthermore the sequence of random variables indexed by* C, $\sup_k EZ_k^2(\underline{\theta}) 1_{\{|Z_k|>C\}}$ *(we can take* $C \in \mathbb{N}$*) is decreasing, so*

$$\sup_{\mathbf{k}} EZ_{\mathbf{k}}^2(\underline{\theta}) 1_{\{|Z_{\mathbf{k}}|>C\}} \searrow 0 \text{ as } C \to \infty,$$

on the set $\left\{\sup_{\mathbf{k}} EZ_{\mathbf{k}}^2(\underline{\theta}) < \infty\right\}$.

■

3.4 The Asymptotic Distribution of the Maximum Likelihood Estimator

Theorem 3.4 *For the process defined by* (3.3.1), *we have under* $P_{\mathbf{n},\underline{\theta}}$

$$i)\ \Lambda_{\mathbf{n}}(\underline{\theta}) = \frac{1}{n}\underline{h}' S_{\mathbf{n}}(\underline{\theta})\,\underline{h} - \frac{1}{2}\underline{h}'\Gamma\left(\underline{\theta},\sigma^2\right)\underline{h} + o_p(1),\ ii)\ n^{-1}\left(\hat{\underline{\theta}}_{\mathbf{n}} - \underline{\theta}\right) \rightsquigarrow N\left(0, \Gamma^{-1}\left(\underline{\theta},\sigma^2\right)\right)$$

where :

$$S'_{\mathbf{n}}(\underline{\theta}) = \sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{n}]} \left(\underline{X}'_{\mathbf{k}}(1), \underline{X}'_{\mathbf{k}}(2)\right),$$

with $\underline{X}'_{\mathbf{k}}(1) = \left(X_{(k_1,k_2-1)}e_{\mathbf{k}}, \ldots, X_{(k_1,k_2-p_2)}e_{\mathbf{k}}, \ldots, X_{(k_1-p_1,k_2)}e_{\mathbf{k}}, \ldots, X_{(k_1-p_1,k_2-p_2)}e_{\mathbf{k}}\right)$,
$\underline{X}'_{\mathbf{k}}(2) = \left(X_{(k_1,k_2-1)}e_{(k_1,k_2-1)}e_{\mathbf{k}}, \ldots, X_{(k_1,k_2-p_2)}e_{(k_1,k_2-p_2)}e_{\mathbf{k}}, \ldots, X_{(k_1-p_1,k_2)}e_{(k_1-p_1,k_2)}e_{\mathbf{k}},\right.$
$\left.\ldots, X_{(k_1-p_1,k_2-p_2)}e_{(k_1-p_1,k_2-p_2)}e_{\mathbf{k}}\right)$,
and where :

$$\Gamma\left(\underline{\theta},\sigma^2\right) = \sigma^{-2}\begin{pmatrix} E\underline{X}_{\mathbf{k}}(1)\,\underline{X}'_{\mathbf{k}}(1) & E\underline{X}_{\mathbf{k}}(1)\,\underline{X}'_{\mathbf{k}}(2) \\ E\underline{X}_{\mathbf{k}}(2)\,\underline{X}'_{\mathbf{k}}(1) & E\underline{X}_{\mathbf{k}}(2)\,\underline{X}'_{\mathbf{k}}(2) \end{pmatrix},$$

And $\hat{\underline{\theta}}_{\mathbf{n}}$ *is defined by :*

$$\hat{\underline{\theta}}_{\mathbf{n}} = \hat{T}_{\mathbf{n}} + \left\{n^2\Gamma\left(\hat{T}_{\mathbf{n}}, \hat{\sigma}^2_{\mathbf{n}}\right)\right\}^{-1} S_{\mathbf{n}}\left(\hat{T}_{\mathbf{n}}\right),$$

$\hat{T}_{\mathbf{n}}$ is any n-consistent preliminary estimator of $\underline{\theta}$ and $\hat{\sigma}^2_{\mathbf{n}} = \frac{1}{n^2} \sum_{\mathbf{k} \in S[\mathbf{1},\mathbf{n}]} e^2_{\mathbf{k}}\left(\hat{T}_{\mathbf{n}}\right)$ where $e^2_{\mathbf{k}}\left(\hat{T}_{\mathbf{n}}\right)$ are defined recursively from (3.3.1) .

Proof *From Theorem 4.3, we have :*

$$\Lambda_{\mathbf{n}}\left(\underline{\theta}\right) = V_{\mathbf{n}}\left(\underline{\theta}\right) - \frac{1}{2}\tau\left(\underline{\theta}\right)^2 + o_p\left(1\right),$$

where $V_{\mathbf{n}}\left(\underline{\theta}\right) = \frac{1}{n} \sum_{\mathbf{k} \in S[\mathbf{1},\mathbf{n}]} Z_{\mathbf{k}}\left(\underline{\theta}\right) + o_p\left(1\right)$, *by setting* $S'_{\mathbf{n}}\left(\underline{\theta}\right) = \sum_{\mathbf{k} \in S[\mathbf{1},\mathbf{n}]} \left(\underline{X}'_{\mathbf{k}}\left(1\right), \underline{X}'_{\mathbf{k}}\left(2\right)\right)$ *we directly obtain :* $\underline{h}' S_{\mathbf{n}}\left(\underline{\theta}\right) = \sum_{\mathbf{k} \in S[\mathbf{1},\mathbf{n}]} Z_{\mathbf{k}}\left(\underline{\theta}\right)$.

Now,

$$\tau\left(\underline{\theta}\right)^2 = EB^2_{1,1}\left(\underline{\theta}\right)/\sigma^2.$$

Setting $\underline{h}'\Gamma\left(\underline{\theta}, \sigma^2\right)\underline{h} = \tau\left(\underline{\theta}\right)^2$, *from the relation* $n^{-1}\underline{h}' S_{\mathbf{n}}\left(\underline{\theta}\right) = \frac{1}{n} \sum_{\mathbf{k} \in S[\mathbf{1},\mathbf{n}]} Z_{\mathbf{k}}\left(\underline{\theta}\right)$ *and the result* $\frac{1}{n} \sum_{\mathbf{k} \in S[\mathbf{1},\mathbf{n}]} Z_{\mathbf{k}}\left(\underline{\theta}\right) \xrightarrow{\mathcal{L}} N\left(0, \tau^2\left(\underline{\theta}\right)\right)$, *we can deduce that :*

$$\frac{1}{n} S_{\mathbf{n}}\left(\underline{\theta}\right) \xrightarrow{d} N_2\left(0, \Gamma\left(\underline{\theta}, \sigma^2\right)\right).$$

Hence :

$$n^{-1}\left(\hat{\underline{\theta}}_{\mathbf{n}} - \underline{\theta}\right) = \frac{1}{n}\Gamma^{-1}\left(\underline{\theta}, \sigma^2\right) S_{\mathbf{n}}\left(\underline{\theta}\right) + o_p\left(1\right).$$

Which gives the result in ii). ■

Appendix 1.1

For $\mathbf{t} = (i, j)$

$A_1(i,j) = A_1 + \sum_{(l_1,l_2) \in S]\mathbf{0},\mathbf{Q}]} B_{(l_1,l_2)} e_{(i-l_1,j-l_2)}$, $A_2(i,j) = A_2 + \sum_{1 \le k \le Q_1} B_{(k,0)} e_{(i-k,j)}$,
$l_2 \neq 0$

$\underline{X}_{(i,j)} = \left(\underline{x}\left(i,j\right), \underline{x}\left(i-1,j\right), \ldots, \underline{x}\left(i-p_1,j\right)\right)$

with $\underline{x}\left(i,j\right) = \left(X_{(i,j)}, X_{(i,j-1)}, \ldots, X_{(i,j-p_2)}\right)$,

$$A_1 = \begin{pmatrix} \underline{a}_0 & \underline{a}_1 & \dots & \underline{a}_{p_1} \\ I' & O_{(p_2+1)} & \dots & O_{(p_2+1)} \\ 0 & \dots & & 0 \\ O_{(p_2+1)} & I' & & \\ \vdots & O_{(p_2+1)} & \ddots & O_{(p_2+1)} \\ \vdots & \vdots & & \vdots \\ 0 & \dots & & 0 \\ O_{(p_2+1)} & \dots & O_{(p_2+1)} & I' \end{pmatrix},$$

$$A_2 = \begin{pmatrix} (a_{10}, 0, \dots, 0) & \dots & & (a_{p_1 0}, 0, \dots, 0) & \underline{0} \\ O_{(p_2+1)} & I'' & & O_{(p_2+1)} & \dots \\ \vdots & \ddots & & \ddots & \\ \vdots & \dots & O_{(p_2+1)} & I'' & O_{(p_2+1)} \\ O_{p_2 \times (p_2+1)} & & \dots & \dots & O_{p_2 \times (p_2+1)} \end{pmatrix},$$

and :

$$B_l = \begin{pmatrix} 0 & \dots & 0 & c_l & 0 & \dots & 0 \\ \vdots & & & 0 & & & \vdots \\ & & & & \vdots & & \\ 0 & \dots & 0 & \vdots & 0 & \cdots & 0 \end{pmatrix}_{((p_1+1)(p_2+1))},$$

c_l being at position $l_1 (p_2 + 1) + l_2$,

and $\underline{e}(i,j) = (e(i,j) + a_0, 0, \dots, 0)'$ is a $(p_1+1)(p_2+1) \times 1$ vector, $\underline{0} = (0, \dots, 0) \in \mathbb{R}^{p_1+1}$ where :

$$I' = \begin{pmatrix} & & 0 \\ & I_{p_2} & \vdots \\ & & 0 \end{pmatrix}_{p_2 \times (p_2+1)}, \quad I'' = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \\ 0 & & & 0 \end{pmatrix}_{(p_2+1)\times(p_2+1)}$$

$\underline{a}_s = \left(a_{(s,1)}, \dots, a_{(s,p_2)}, 0\right)$, for all $0 \le s \le p_1$.

Appendix 1.2

For $m \ge 3$,

$V_{n,m} = A_1 V_{n,m-1} A_1^{'} + B_{(0,2)} D_{n,m-1} A_1^{'} + A_1 D_{n,m-1} \left(B_{(0,2)}\right)^{'} + \sigma^2 B_{(0,1)} V_{n,m-1} \left(B_{(0,1)}\right)^{'} +$
$B_{(0,2)} F_{n,m-1} \left(B_{(0,2)}\right)^{'} + A_2 V_{n-1,m} A_2^{'} + B_{(2,0)} D^{*}_{n-1,m} A_2^{'} + A_2 D^{*}_{n-1,m} \left(B_{(2,0)}\right)^{'} +$
$\sigma^2 B_{(1,0)} V_{n-1,m} \left(B_{(1,0)}\right)^{'} + B_{(2,0)} F^{*}_{n-1,m} \left(B_{(2,0)}\right)^{'}$.

so,

$\overrightarrow{V}_{n,m} = \left(A_1 \otimes A_1 + \sigma^2 B_{(0,1)} \otimes B_{(0,1)}\right) \overrightarrow{V}_{n,m-1} + \left(A_1 \otimes B_{(0,2)} + B_{(0,2)} \otimes A_1\right) \overrightarrow{D}_{n,m-1} +$
$\left(B_{(0,2)} \otimes B_{(0,2)}\right) \overrightarrow{F}_{n,m-1} + \left(A_2 \otimes A_2 + \sigma^2 B_{(1,0)} \otimes B_{(1,0)}\right) \overrightarrow{V}_{n-1,m} +$
$\left(A_2 \otimes B_{(2,0)} + B_{(2,0)} \otimes A_2\right) \overrightarrow{D}^{*}_{n-1,m} + \left(B_{(2,0)} \otimes B_{(2,0)}\right) \overrightarrow{F}^{*}_{n-1,m}$.

$\overrightarrow{D}_{n,m} = \sigma^2 \left(A_1 \otimes B_{(0,1)} + B_{(0,1)} \otimes A_1\right) \overrightarrow{V}_{n,m-1} + \sigma^2 \left(B_{(0,2)} \otimes B_{(0,1)} + B_{(0,1)} \otimes B_{(0,2)}\right) \overrightarrow{D}_{n,m-1}$.

$\overrightarrow{D}^{*}_{n,m} = \sigma^2 \left(A_2 \otimes B_{(1,0)} + B_{(1,0)} \otimes A_1\right) \overrightarrow{V}_{n-1,m} + \sigma^2 \left(B_{(2,0)} \otimes B_{(1,0)} + B_{(1,0)} \otimes B_{(2,0)}\right) \overrightarrow{D}^{*}_{n-1,m}$.

$\overrightarrow{F}_{n,m} = \left(\sigma^2 A_1 \otimes A_1 + \gamma^4 B_{(0,1)} \otimes B_{(0,1)}\right) \overrightarrow{V}_{n,m-1} + \sigma^2 \left(A_1 \otimes B_{(0,2)} + B_{(0,2)} \otimes A_1\right) \overrightarrow{D}_{n,m-1}$
$+\sigma^2 \left(B_{(0,2)} \otimes B_{(0,2)}\right) \overrightarrow{F}_{n,m-1}$.

$\overrightarrow{F}^{*}_{n,m} = \left(\sigma^2 A_2 \otimes A_2 + \gamma^4 B_{(1,0)} \otimes B_{(1,0)}\right) \overrightarrow{V}_{n-1,m} + \sigma^2 \left(A_2 \otimes B_{(2,0)} + B_{(2,0)} \otimes A_2\right) \overrightarrow{D}^{*}_{n-1,m}$
$+\sigma^2 \left(B_{(2,0)} \otimes B_{(2,0)}\right) \overrightarrow{F}^{*}_{n-1,m}$.

if we put :

$$U_{n,m} = \begin{bmatrix} \overrightarrow{V}_{n,m} \\ \overrightarrow{D}_{n,m} \\ \overrightarrow{D}^{*}_{n,m} \\ \overrightarrow{F}_{n,m} \\ \overrightarrow{F}^{*}_{n,m} \end{bmatrix}, \ n \geq 0, m \geq 3.$$

and

$\alpha_1 = A_1 \otimes A_1 + \sigma^2 B_{(0,1)} \otimes B_{(0,1)}$, $\alpha_2 = A_2 \otimes A_2 + \sigma^2 B_{(1,0)} \otimes B_{(1,0)}$, $\beta_1 = A_1 \otimes B_{(0,2)} + B_{(0,2)} \otimes A_1$,
$\beta_2 = A_2 \otimes B_{(2,0)} + B_{(2,0)} \otimes A_2$, $\delta_1 = \sigma^2 \left(A_1 \otimes B_{(0,1)} + B_{(0,1)} \otimes A_1\right)$, $\delta_2 = \sigma^2 \left(A_2 \otimes B_{(1,0)} + B_{(1,0)} \otimes A_2\right)$,
$\gamma_1 = \sigma^2 \left(B_{(0,2)} \otimes B_{(0,1)} + B_{(0,1)} \otimes B_{(0,2)}\right)$, $\gamma_2 = \sigma^2 \left(B_{(2,0)} \otimes B_{(1,0)} + B_{(1,0)} \otimes B_{(2,0)}\right)$
$\eta_1 = \sigma^2 A_1 \otimes A_1 + \gamma^4 B_{(0,1)} \otimes B_{(0,2)}$, $\eta_2 = \sigma^2 A_2 \otimes A_2 + \gamma^4 B_{(1,0)} \otimes B_{(2,0)}$, $\theta_1 = \sigma^2 \left(A_1 \otimes B_{(0,2)} + B_{(0,2)} \otimes A_1\right)$,
$\theta_2 = \sigma^2 \left(A_2 \otimes B_{(2,0)} + B_{(2,0)} \otimes A_2\right)$, $\lambda_1 = B_{(0,2)} \otimes B_{(0,2)}$, $\lambda_2 = B_{(2,0)} \otimes B_{(2,0)}$.

$$\Gamma_1=\begin{pmatrix} \alpha_1 & \beta_1 & 0 & \lambda_1 & 0 \\ \delta_1 & \gamma_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \eta_1 & \theta_1 & 0 & \sigma^2\lambda_1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \Gamma_2=\begin{pmatrix} \alpha_2 & 0 & \beta_2 & 0 & \lambda_2 \\ 0 & 0 & 0 & 0 & 0 \\ \delta_2 & 0 & \gamma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \eta_2 & 0 & \theta_2 & 0 & \sigma^2\lambda_2 \end{pmatrix},$$

Chapter 4

Generalized spatial autoregressive models on $\mathbb{Z}^2$

4.1 Introduction

One of the most famous spatial process which has attracted considerable attention in the literature is the spatial autoregressive model (SAR) which has been investigated in many works (see, for example Anslin (1988); Haining (1990); Cressie (1993); Besag and Higdon (1999)). It was put by Cliff and Ord (1973, 1981) and it was a variant of the model considered by Whittle (1954). Spatial autoregressive models have been used extensively in various domain, for instance, they have been used to represent images and real sceneries in Bennett and Khotanzad (1999) and Kashyap and Eom (1988) respectively.

However, in many situations the mentioned spatial autoregressive model need to be extend for describing the underlying random mechanism. For instance, a first situation is the case where at every spatial location, both the neighboring sample variances and the neighboring conditional variances play a role in the current conditional variance. Indeed, in the one-dimensional GARCH model, the conditional variance change over time, however in the two dimensional GARCH model, this conditional variance change over space. Specifically, conditioning on the past is replaced by conditioning on the nearest neighbors of the considered location. Notice here that the two-dimensional GARCH models have been used in clutter modeling and anomaly detection (see, for

example Nerboar and Cohen (2005), they have been also used for statistical modeling of wavelet coefficients (see, for example Amirmazlaghani and Amindavar (2010)). The second situation is the case where parameters of this model can slowly change from location to another and one can propose a varying coefficients spatial autoregressive model (RCSAR) which have the following equation

$$X(i,j) = \sum_{k=1}^{p_1} \sum_{l=1}^{p_2} a_{kl}(i,j) X(i-k, j-l) + \varepsilon(i,j),$$

the parameters $a_{kl}(i,j)$, $1 \leq k \leq p_1$ and $1 \leq l \leq p_2$ are fixed (non-random) functions of (i,j). This model is a spatially varying coefficients process, early contributions in this respect are due to Gelfand et al (2003) and Fotheringham et al (2002) who gave an overview on the basics technical details underlying spatially varying coefficients processes. In an other situation the coefficients arise as a result of complicated processes and actions which usually have many random position. Thus we are naturally led to another kind of spatial AR model having the same previous equation but where the parameters are random variables. These models will be studied in more detail in Section 5.

All these models are generalizations of spatial autoregressive models and our aim is to formulate suitable and simple model to explain the evolution space series. One can associate with each of these processes a multivariate stochastic difference equation of the form

$$\underline{X}(i,j) = A_1(i,j)\,\underline{X}(i,j-1) + A_2(i,j)\,\underline{X}(i-1,j) + \underline{\eta}(i,j), \quad (i,j) \in \mathbb{Z}^2 \tag{4.1.1}$$

where $\underline{X}(i,j)$ and $\underline{\eta}(i,j)$ are random vectors in $\mathbb{R}^d$, the $A_1(i,j)$'s and $A_2(i,j)$'s are $d \times d$ random matrices. The order chosen on the lattice $\mathbb{Z}^2$ is the lexicographic order noted "$\prec$" that is for $\mathbf{s} = (s_1, s_2)$, $\mathbf{t} = (t_1, t_2) \in \mathbb{Z}^2$, we write $\mathbf{s} \prec \mathbf{t}$ if $[(s_1 < t_1) \vee (s_1 = t_1 \text{ and } s_2 < t_2)]$.

We argue that the generalized spatial model given by the equation (1.1) is deserved to be one of the tools of statisticians. There is several arguments supporting this claim.

First, the stationarity properties of models which describe the different situations mentioned above are directly related to the stationarity properties of the solution of (4.1.1).

Second, the model and the corresponding stationary results are natural generalizations of the

well-known Brandt (1986), Bougerol and Picard (1992) as well as Nicholls and Quinn (1982) results for 1-D generalized autoregressive and RCA models.

Second order properties are very required in statistical inference, hence most of the literature that exists was concerned with weak stationarity, but often financial series exhibit parameters which are not in the second-order stationarity region (see Engle and Bollerslev (1986) and references therein). These series may be strictly stationary even though they are not square integrable. Therefore, in this chapter, we look more closely at a large class of two-dimensional generalized autoregressive models (this class includes, for instance, two-dimensional RCA, GARCH and bilinear models). Our goal is to investigate strict and second order stationarity to cover both situations, with and without square-integrable assumption. Section 2 is devoted to the study of the necessary and sufficient conditions for the existence and uniqueness of a second order stationary solution of (4.1.1). In section 3, we establish our second result on the existence of strictly stationary solution of (4.1.1). Section 4 discusses the existence of higher order moments. Section 5 contains two applications.

4.2 Second order stationarity

In this section, we are looking for a solution of (4.1.1) which may be considered as a sequence of random vectors on the same probability space as that generated by $A_1(i,j)$, $A_2(i,j)$ and $\underline{\eta}(i,j)$. It is even more interesting to find a solution for which $\{A_1(i,j),\ (s_1,s_2) \prec (i,j)\}$, $\{A_2(i,j),\ (s_1,s_2) \prec (i,j)\}$ and $\{\underline{\eta}(i,j),\ (s_1,s_2) \prec (i,j)\}$ are independent of $\{\underline{X}(i,j),\ (i,j) \preceq (s_1,s_2)\ \}$. Let $\mathcal{F}_{(i,j)}$ be the σ-field generated by $\{\underline{\eta}(s_1,s_2),\ A_1(s_1,s_2),\ A_2(s_1,s_2),\ (s_1,s_2) \preceq (i,j)\ \}$. The above mentioned properties hold for an $\mathcal{F}_{(i,j)}-$measurable solution $\{\underline{X}(i,j)\}$ of (4.1.1).

For checking the second order stationarity, the following assumptions are made.

(i) $\{\underline{\eta}(i,j);\ (i,j) \in \mathbb{Z}^2\}$ is an independent second order stationary sequence of d-variate random vectors with mean zero and covariance matrix G.

(ii) $\{A_1(i,j);\ (i,j) \in \mathbb{Z}^2\}$ and $\{A_2(i,j);\ (i,j) \in \mathbb{Z}^2\}$ are independent sequences of $d \times d$ random matrices with $E[A_1(i,j)] = A_1$, $E[A_2(i,j)] = A_2$ and $E[A_1(i,j) \otimes A_1(i,j)] = C_1$ and

$E\left[A_2\left(i,j\right)\otimes A_2\left(i,j\right)\right]=C_2$. Also $A_1\left(i,j\right)$, $A_2\left(i,j\right)$ and $\underline{\eta}\left(i,j\right)$ are mutually independent. Now, we state our general result which is given in the following theorem.

Theorem 4.1 *Suppose that the spatial generalized autoregressive model (4.1.1) satisfies assumptions (i)-(ii). When (C_1+C_2) does not have a unit eigenvalue, (4.1.1) has a unique stationary $\mathcal{F}_{(i,j)}-$measurable solution if and only if*

$$\sum_{k=0}^{r}\sum_{l=0}^{k}E\left[T^{l,k-l}\left(i,j\right)\otimes T^{l,k-l}\left(i,j\right)\right] vecG \text{ converges as } r\to\infty$$

This solution is given by

$$\underline{X}\left(i,j\right)=\sum_{k\geq 0}\sum_{l=0}^{k}T^{l,k-l}\left(i,j\right)\underline{\eta}\left(i-l,j-k+l\right), \tag{4.2.2}$$

where the transition matrix $T^{n,m}\left(i,j\right)$ is defined as follows:

1) $T^{0,1}\left(i,j\right)=A_1\left(i,j\right)$ *and* $T^{1,0}\left(i,j\right)=A_2\left(i,j\right)$

2) $T^{n,m}\left(i,j\right)=A_1\left(i,j\right)T^{n,m-1}\left(i,j-1\right)+A_2\left(i,j\right)T^{n-1,m}\left(i-1,j\right)$

3) $T^{0,0}\left(i,j\right)=I_{d\times d}$, *and*

4) $T^{-n,m}\left(i,j\right)=T^{n,-m}\left(i,j\right)=O_{d\times d}$.

Proof Assume that (C_1+C_2) does not have a unit eigenvalue and let's show that the necessary and sufficient condition for (4.1.1) to have a unique stationary $\mathcal{F}_{(i,j)}-$measurable solution is given by

$$\sum_{k=0}^{r}\sum_{l=0}^{k}E\left[T^{l,k-l}\left(i,j\right)\otimes T^{l,k-l}\left(i,j\right)\right] vecG \text{ converges as } r\to\infty.$$

First suppose $\underline{X}\left(i,j\right)$ satisfies (4.1.1) and is $\mathcal{F}_{(i,j)}-$measurable and stationary. By iterating (4.1.1) we obtain

$$\underline{X}\left(i,j\right)=\sum_{k=0}^{r}\sum_{l=0}^{k}T^{l,k-l}\left(i,j\right)\underline{\eta}\left(i-l,j-k+l\right)+\sum_{l=0}^{r+1}T^{l,r+1-l}\left(i,j\right)\underline{X}\left(i-l,j-r-1+l\right),$$

where the transition matrix $T^{n,m}(i,j)$ is defined as follows:

1) $T^{0,1}(i,j) = A_1(i,j)$ and $T^{1,0}(i,j) = A_2(i,j)$

2) $T^{n,m}(i,j) = A_1(i,j)\,T^{n,m-1}(i,j-1) + A_2(i,j)\,T^{n-1,m}(i-1,j)$

3) $T^{0,0}(i,j) = I_{d\times d}$, and

4) $T^{-n,m}(i,j) = T^{n,-m}(i,j) = O_{d\times d}$.

Letting $V = E\left(\underline{X}(i,j)\,\underline{X}'(i,j)\right)$, and denoting $W_r(i,j) = \sum_{k=0}^{r}\sum_{l=0}^{k} T^{l,k-l}(i,j)\,\underline{\eta}(i-l,j-k+l)$ and $R_r(i,j) = \sum_{l=0}^{r+1} T^{l,r+1-l}(i,j)\,\underline{X}(i-l,j-r-1+l)$, we have

$$\begin{aligned} vecV &= vec\left[E\left(W_r(i,j)+R_r(i,j)\right)\left(W_r(i,j)+R_r(i,j)\right)'\right] \\ &= vec\left[E\left(W_r(i,j)\,W_r'(i,j)\right)+E\left(R_r(i,j)\,R_r'(i,j)\right)+\right. \\ &\quad \left. E\left(R_r(i,j)\,W_r'(i,j)\right)+E\left(W_r(i,j)\,R_r'(i,j)\right)\right]. \end{aligned}$$

Now, by (i) and (ii), and noting that $E\left[\underline{\eta}(i-n,j-m)\,\underline{\eta}'(i-p,j-q)\right] = 0$ if $(n,m)\neq(p,q)$, we have

$$\begin{aligned} vecE\left(W_r(i,j)\,W_r'(i,j)\right) &= \sum_{k=0}^{r}\sum_{l=0}^{k}\left\{E\left[T^{l,k-l}(i,j)\otimes T^{l,k-l}(i,j)\right]\times\right. \\ &\quad \left. vecE\left[\underline{\eta}(i-l,j-k+l)\,\underline{\eta}'(i-l,j-k+l)\right]\right\}, \end{aligned}$$

by recurrence, we can say that

$$vecE\left(W_r(i,j)\,W_r'(i,j)\right) = \sum_{k=0}^{r}\sum_{l=0}^{k} E\left[T^{l,k-l}(i,j)\otimes T^{l,k-l}(i,j)\right] vecG.$$

Also,

$$vecE\left(R_r(i,j)\,R_r'(i,j)\right) = \sum_{l=0}^{r+1}\left[T^{l,r+1-l}(i,j)\otimes T^{l,r+1-l}(i,j)\right] vecV,$$

and

$$vecE\left(W_r(i,j)\,R_r'(i,j)\right) = vecE\left(R_r(i,j)\,W_r'(i,j)\right) = 0$$

since $E\left[\underline{\eta}\left(i-n,j-m\right)\underline{X}'\left(i-p,j-q\right)\right]=0$ for $(n,m)\prec(p,q)$. Thus

$$\begin{aligned} vecV &= \sum_{k=0}^{r}\sum_{l=0}^{k}E\left[T^{l,k-l}\left(i,j\right)\otimes T^{l,k-l}\left(i,j\right)\right]vecG+ \\ &\quad \sum_{l=0}^{r+1}\left[T^{l,r+1-l}\left(i,j\right)\otimes T^{l,r+1-l}\left(i,j\right)\right]vecV. \end{aligned}$$

Following Nicholls and Quinn (1982) in the proof of Theorem 2.2. p 21, let

$$\begin{aligned} Q_0 &= G;\ Q_k=E\left\{\left[A_1\left(i,j\right)+A_2\left(i,j\right)\right]Q_{k-1}\left[A_1'\left(i,j\right)+A_2'\left(i,j\right)\right]\right\} \quad k=1,2,\ldots; \\ R_0 &= V;\ R_k=E\left\{\left[A_1\left(i,j\right)+A_2\left(i,j\right)\right]R_{k-1}\left[A_1'\left(i,j\right)+A_2'\left(i,j\right)\right]\right\} \quad k=1,2,\ldots; \end{aligned}$$

We can see that each of Q_k, R_k are non-negative definite. Also

$$V=\sum_{k=0}^{r}\left[Q_k\right]+R_{r+1},\ r=0,1,2,\ldots.$$

If z is any $d\times 1$ fixed vector, then

$$z'Vz=\sum_{k=0}^{r}z'Q_kz+z'R_{r+1}z,\ r=0,1,2,\ldots. \tag{4.2.3}$$

$z'Vz$, $z'R_{r+1}z$ are non-negative, hence $\sum_{k=0}^{r}z'Q_kz$ is nondecreasing in r. Since (4.2.3) holds for $r=0,1,2\ldots$, it follows that $\sum_{k=0}^{r}z'\left[Q_k\right]z$ is bounded above by $z'Vz$ and is therefore convergent for every vector z. Thus $\sum_{k=0}^{r}\left[Q_k\right]$ converges, as $r\to\infty$ to a non-negative matrix and so $\sum_{k=0}^{r}\sum_{l=0}^{k}E\left[T^{l,k-l}\left(i,j\right)\otimes T^{l,k-l}\left(i,j\right)\right]vecG$ converges as $r\to\infty$, this completes the proof of necessary condition.

Suppose now that $\sum_{k=0}^{r}\sum_{l=0}^{k}E\left[T^{l,k-l}\left(i,j\right)\otimes T^{l,k-l}\left(i,j\right)\right]vecG$ converges as $r\to\infty$, in the first step of the proof, it has been shown that the limit $W\left(i,j\right)$ of $W_r\left(i,j\right)$ as $r\to\infty$ exists in mean

square, and thus in probability. Furthermore, $\{W(i,j)\}$ satisfies (4.1.1), indeed

$$\begin{aligned} A_1(i,j)W(i,j-1)+A_2(i,j)W(i-1,j) &= \sum_{n+m=1}^{\infty} T^{n,m}(i,j)\,\underline{\eta}(i-n,j-m) \\ &= W(i,j)-\underline{\eta}(i,j). \end{aligned}$$

The expression of $W(i,j)$ clearly indicates that, $W(i,j)$ is $\mathcal{F}_{(i,j)}-$measurable and also stationary since $vecE\left(W(i,j)W'(i,j)\right)=\sum_{k=0}^{\infty}\sum_{l=0}^{k}E\left[T^{l,k-l}(i,j)\otimes T^{l,k-l}(i,j)\right]vecG$ which is finite.
Let's now prove the unicity, suppose there are two solutions to (4.1.1), $W(i,j)$ and $Y(i,j)$ such that $W(i,j)=g\left(\underline{\eta}(i-s_1,j-s_2),\ A_1(i-s_1,j-s_2),\ A_2(i-s_1,j-s_2),\ (0,0)\preceq(s_1,s_2)\right)$ for some measurable function $g:\mathbb{R}^{\infty}\to\mathbb{R}^d$ and $Y(i,j)$ is the solution found in (4.2.2), then if we put $U(i,j)=W(i,j)-Y(i,j)$, $U(i,j)$ is also $\mathcal{F}_{(i,j)}-$measurable, stationary and satisfies

$$U(i,j)=A_1(i,j)U(i,j-1)+A_2(i,j)U(i-1,j),$$

so we must have

$$vecEU(i,j)U'(i,j)=C_1vecEU(i,j-1)U'(i,j-1)+C_2vecEU(i-1,j)U'(i-1,j).$$

However, since $\{U(i,j)\}$ is also stationary we must have

$$vecEU(i,j)U'(i,j)=(C_1+C_2)\,vecEU(i,j)U'(i,j),$$

since (C_1+C_2) has not unit eigenvalues, $EU(i,j)U'(i,j)=0$. Thus $U(i,j)=0$ and $W(i,j)=Y(i,j)$ almost everywhere. ■

Remark 4.1 *If we suppose that $\underline{\eta}(i,j)$, $A_1(i,j)$ and $A_2(i,j)$ are also identically distributed, it is possible to infer stronger properties for the $\mathcal{F}_{(i,j)}-$measurable solution $\{\underline{X}(i,j)\}$ to (4.1.1), so if $\underline{\eta}(i,j)$, $A_1(i,j)$ and $A_2(i,j)$ satisfies assumptions (i) and (ii) and are also identically distributed sequences. Then, if a unique $\mathcal{F}_{(i,j)}-$measurable second order stationary solution $\{\underline{X}(i,j)\}$ exists*

to (4.1.1), $\underline{X}(i,j)$ is also strictly stationary and ergodic (The result follows from the functional form of the solution $\underline{X}(i,j)$ which remains the same for all (i,j)).

4.3 Strict stationarity

In the previous section, conditions ensuring the existence of second-order stationary solutions have been established. We have seen that such solutions can also be strictly stationary. In this section we are interested in strictly stationary solutions that are not necessarily square-integrable. The usefulness of such solutions has been asserted by many authors, in particular in the spatial $ARCH$ and $GARCH$ models. We will use transition matrix $T^{n,m}(i,j)$ defined in Theorem 5.1 and the theory of products of random matrices to show that the necessary and sufficient condition for strict stationarity is the negativity of a coefficient associated with a suitable sequence of stationary and ergodic matrices.

Let us recall that for each $m \times m$ matrix $A = (a_{ij})$, the norm $\|A\|$ of A is defined by $\|A\| = \max_i \sum_j |a_{ij}|$ and for all $x > 0$, $\log^+ x = \max\{\log x, 0\}$.

In order to prove the strict stationarity, we quote from J. Cohen (1988) the following theorem which suggests an application of the subadditive ergodic theorem.

Theorem 4.2 *Let $\{M_t\}_{t\in T}$ be a stationary sequence of random matrices from $C^{k\times k}$ and let*

$$S_{vt} = M_{v+1} + \ldots + M_{t-1} + M_t, \quad v, t \in T,\ v < t.$$

If $E\|M_1\| < \infty$, then

$$\lim_{t\to\infty} \frac{\|S_{0t}\|}{t} = \xi,$$

exists with probability one, is finite, and has an expectation $E(\xi)$ that satisfies

$$E(\xi) = \lim_{t\to\infty} \frac{E\{\|S_{0t}\|\}}{t}.$$

If, in addition, the random matrices $\{M_t\}$ are independently and identically distributed (i.i.d.),

then

$$\xi = E(\xi) \text{ with probability one.}$$

First, the following assumptions are made.

(i) $\left\{\underline{\eta}(i,j);\ (i,j) \in \mathbb{Z}^2\right\}$ is an independent strict stationary sequence of d-variate random vectors with mean zero.

(ii) $\{A_1(i,j);\ (i,j) \in \mathbb{Z}^2\}$ and $\{A_2(i,j);\ (i,j) \in \mathbb{Z}^2\}$ are independent sequences of i.i.d. $d \times d$ random matrices.

Theorem 4.3 *We assume that* (4.1.1) *holds, assume also that*

$$E\log^+\{\|A_1(i,j)\| + \|A_2(i,j)\|\} < \infty,$$

and let M_0 *such that, almost surely* $\left\|\underline{\eta}(i,j)\right\| \leq M_0$. *If*

$$\varrho = \limsup_k \frac{1}{k}\log\prod_{l=0}^{k-1}\left\{\sum_{s=0}^{l}\left[\|A_1(i-s,j-l+s)\| + \|A_2(i-s,j-l+s)\|\right]\right\} < 0, \tag{4.3.0}$$

Then

$$\underline{X}(i,j) = \sum_{k\geq 0}\sum_{l=0}^{k} T^{l,k-l}(i,j)\,\underline{\eta}(i-l,j-k+l) \tag{4.3.1}$$

converges absolutely a.s. and the process $\{X(i,j);\ (i,j) \in \mathbb{Z}^2\}$ *is the unique stationary solution of* (4.1.1). $T^{n,m}(i,j)$ *is defined as in Theorem 4.1.*

Proof *We shall follow arguments similar to those in Brandt (1986) who found out conditions that ensure the existence of a uniquely determined stationary solution of the stochastic equation* $Y_{n+1} = A_{n+1}Y_n + B_{n+1}$ *with stationary coefficients. We first establish the absolutely convergence of the series of random vectors* $\sum_{k\geq 0}\sum_{l=0}^{k} T^{l,k-l}(i,j)\,\underline{\eta}(i-l,j-k+l)$, *we use the Cauchy criterion applied to the series* $\sum_{k=0}^{\infty}\|W_k(i,j)\|$, *where* $W_k(i,j) = \sum_{l=0}^{k} T^{l,k-l}(i,j)\,\underline{\eta}(i-l,j-k+l)$. *We can see that*

$\|W_1(i,j)\| \leq [\|A_1(i,j)\| + \|A_2(i,j)\|]\,M_0$ *a.s., and*

$\|W_2(i,j)\| \leq [\|A_1(i,j)\| + \|A_2(i,j)\|] \times$
$[\|A_1(i,j-1)\| + \|A_2(i,j-1)\| + \|A_1(i-1,j)\| + \|A_2(i-1,j)\|] M_0$ *a.s. By recurrence, it can be shown that a.s.,*

$$\|W_k(i,j)\| \leq M_0 \prod_{l=0}^{k-1} \left\{ \sum_{s=0}^{l} [\|A_1(i-s,j-l+s)\| + \|A_2(i-s,j-l+s)\|] \right\}$$

So

$$\begin{aligned} & \limsup_k \frac{1}{k} \log \|W_k(i,j)\| \\ \leq & \limsup_k \frac{1}{k} \sum_{l=0}^{k-1} \log \left\{ \sum_{s=0}^{l} [\|A_1(i-s,j-l+s)\| + \|A_2(i-s,j-l+s)\|] \right\} + \limsup_k \frac{1}{k} \log M_0 \\ \leq & \limsup_k \frac{1}{k} \sum_{l=0}^{k-1} \log \left\{ \sum_{s=0}^{l} [\|A_1(i-s,j-l+s)\| + \|A_2(i-s,j-l+s)\|] \right\}, \end{aligned}$$

Put
$\varrho = \limsup_k \frac{1}{k} \log \prod_{l=0}^{k-1} \left\{ \sum_{s=0}^{l} [\|A_1(i-s,j-l+s)\| + \|A_2(i-s,j-l+s)\|] \right\}$, *the existence of the non-random quantity* ϱ *is guaranteed by Theorem 4.2. indeed, by adopting the notations of Theorem 4.2, consider the case where the stationary sequence from* $C^{m\times m}$, $\{M_t\}$, *is such that* $m = 1$ *and* $M_t = \log \left\{ \sum_{s=1}^{t} [\|A_1(i-s+1,j-t+s+1)\| + \|A_2(i-s+1,j-t+s+1)\|] \right\}$, *and put*

$$S_{vk} = \sum_{l=v+1}^{k} M_l,$$

note first that
$\limsup_k \frac{1}{k} \|S_{0k}\| = \limsup_k \frac{1}{k} \left| \log \prod_{l=0}^{k-1} \left\{ \sum_{s=0}^{l} [\|A_1(i-s,j-l+s)\| + \|A_2(i-s,j-l+s)\|] \right\} \right| =$
$|\rho|$. *Since* $E\|M_1\| = E\log^+ \{\|A_1(i,j)\| + \|A_2(i,j)\|\} < \infty$ *and* $\{M_t\}$ *are (i.i.d.), then by Theorem 4.2* $\lim_{t\to\infty} \frac{\|S_{0t}\|}{t}$ *exists with probability one, is finite, and has an expectation that satisfies* $E\left(\lim_{t\to\infty} \frac{\|S_{0t}\|}{t}\right) = \lim_{t\to\infty} \frac{\|S_{0t}\|}{t}$ *which means that* $|\rho| = E|\rho|$. *Therefore, when* $\varrho < 0$, *the*

series

$$\underline{X}(i,j) = \sum_{k\geq 0}\sum_{l=0}^{k} T^{l,k-l}(i,j)\,\underline{\eta}(i-l,j-k+l)$$

converges a.s. This sequence is also clearly a strictly stationary solution of (4.1.1).

Let's now prove the unicity, suppose there are two stationary solutions to (4.1.1), $X(i,j)$ *and* $Y(i,j)$ *and put* $U(i,j) = X(i,j) - Y(i,j)$. *Thus* $U(i,j)$ *satisfies*

$$U(i,j) = A_1(i,j)\,U(i,j-1) + A_2(i,j)\,U(i-1,j),$$

repeating this formula we get

$$U(i,j) = \sum_{l=0}^{k} T^{l,k-l}(i,j)\,U(i-l,j-k+l),$$

hence

$$\begin{aligned} \|U(i,j)\| &\leq \sum_{l=0}^{k} \left\|T^{l,k-l}(i,j)\,U(i-l,j-k+l)\right\| \\ &\leq \sum_{l=0}^{k} \left\|T^{l,k-l}(i,j)\right\| \|U(i-l,j-k+l)\|. \end{aligned} \tag{4.3.2}$$

Note first that

$$\begin{aligned} \sum_{l=0}^{k} \left\|T^{l,k-l}(i,j)\right\| \|U(i-l,j-k+l)\| &\leq \left\{\sum_{l=0}^{k} \left\|T^{l,k-l}(i,j)\right\|\right\} \sum_{l=0}^{k} \|U(i-l,j-k+l)\| \\ &\leq \left\{\sum_{l=0}^{k} \left\|T^{l,k-l}(i,j)\right\|\right\} \sum_{l=0}^{k} \|X(i-l,j-k+l)\| + \\ &\quad \left\{\sum_{l=0}^{k} \left\|T^{l,k-l}(i,j)\right\|\right\} \sum_{l=0}^{k} \|Y(i-l,j-k+l)\|, \end{aligned}$$

Now let's prove that $\sum_{l=0}^{k} \left\| T^{l,k-l}(i,j) \right\|$ *converges a.s. to 0 as* $k \to \infty$. *It has been shown above that*

$$\sum_{l=0}^{k} \left\| T^{l,k-l}(i,j) \right\| \leq \prod_{l=0}^{k-1} \left\{ \sum_{s=0}^{l} \left[\|A_1(i-s, j-l+s)\| + \|A_2(i-s, j-l+s)\| \right] \right\},$$

note also that

$$\begin{aligned} & \prod_{l=0}^{k-1} \left\{ \sum_{s=0}^{l} \left[\|A_1(i-s, j-l+s)\| + \|A_2(i-s, j-l+s)\| \right] \right\} \\ = & \left[\exp\left(\frac{1}{k} \log \prod_{l=0}^{k-1} \left\{ \sum_{s=0}^{l} \left[\|A_1(i-s, j-l+s)\| + \|A_2(i-s, j-l+s)\| \right] \right\} \right) \right]^k \end{aligned}$$

hence, when $\varrho < 0$, *according to the strong law of large numbers*

$$\left[\exp\left(\frac{1}{k} \log \prod_{l=0}^{k-1} \left\{ \sum_{s=0}^{l} \left[\|A_1(i-s, j-l+s)\| + \|A_2(i-s, j-l+s)\| \right] \right\} \right) \right]^k \to 0 \text{ as } k \to \infty \text{ a.s.} \tag{4.3.3}$$

since the laws of $\|X(i-l, j-k+l)\|$ *and* $\|Y(i-l, j-k+l)\|$ *remain constant, and due to (4.3.3) we have* $\left(\sum_{l=0}^{k} \left\| T^{l,k-l}(i,j) \right\| \right) \sum_{l=0}^{k} \|X(i-l, j-k+l)\| \xrightarrow{P} 0$ *as* $k \to \infty$ *and* $\left\{ \sum_{l=0}^{k} \left\| T^{l,k-l}(i,j) \right\| \right\} \sum_{l=0}^{k} \|Y(i-l, j-k+l)\| \xrightarrow{P} 0$ *as* $k \to \infty$, *thus by (4.3.2)* $U(i,j) = 0$ *as* $k \to \infty$ *a.s. which means* $X(i,j) = Y(i,j)$.

Remark 4.2 *The Condition* (4.3.0) *seems difficult to obtain explicitly, however, a potential method to verify whether or not* $\rho < 0$ *is via Monte-Carlo simulation using Equation (4.1.1). On the other hand and for some statistical applications, the solution process* needs *to have some moments to make an estimation theory possible and the Criterion* (4.3.0) *does not guarantee the existence of such moments. In the previous section, we gave conditions ensuring the existence of square integrable strict stationary solution for Equation (4.1.1).*

Example 4.1 *Consider a simple spatial bilinear process of first order,*

$$X(i,j) = aX(i,j-1)+bX(i-1,j)+\alpha X(i,j-1)\,\varepsilon(i-1,j)+\beta X(i-1,j)\,\varepsilon(i,j-1)+\varepsilon(i,j), \tag{4.3.4}$$

where $X(i,j)$ and $\varepsilon(i,j)$ are random variables, a, b, α and β are constants and assume that $\{\varepsilon(i,j)\}$ is a Gaussian white noise with zero mean and variance σ^2 (in this case, M_0 could be roughly chosen to be 4). Here $\|A_1(i,j)\| = |a+\alpha\varepsilon(i-1,j)|$, $\|A_2(i,j)\| = |b+\beta\varepsilon(i,j-1)|$, Hence

$$\begin{aligned} E\log\{\|A_1(i,j)\|+\|A_2(i,j)\|\} &\leq \log\{E(|a+\alpha\varepsilon(i-1,j)|+|b+\beta\varepsilon(i,j-1)|)\} \\ &\leq \frac{1}{2}\log E(|a+\alpha\varepsilon(i-1,j)|+|b+\beta\varepsilon(i,j-1)|)^2 \\ &\leq \frac{1}{2}\log\left(a^2+b^2+\left(\alpha^2+\beta^2\right)\sigma^2+2|ab|\right) \\ &< \infty. \end{aligned}$$

Furthermore,

$$\begin{aligned} \varrho &= \limsup_k \frac{1}{k}\log\prod_{l=0}^{k-1}\left\{\sum_{s=0}^{l}\left[\|A_1(i-s,j-l+s)\|+\|A_2(i-s,j-l+s)\|\right]\right\} \\ &\leq E\log\left(\|A_1(i,j)\|+\|A_2(i,j)\|\right), \ \textit{a.s.,} \end{aligned}$$

hence $\varrho<0$, if $\frac{1}{2}\log\left((|a|+|b|)^2+\left(\alpha^2+\beta^2\right)\sigma^2\right)<0$ and we get the following simple condition for the existence of a unique strictly stationary solution of (4.3.4)

$$(|a|+|b|)^2+\left(\alpha^2+\beta^2\right)\sigma^2<1.$$

Example 4.2 *Consider the spatial RCA process of first order defined by the equation,*

$$X(\mathbf{t}) = a_1(\mathbf{t})X(\mathbf{t}-\mathbf{e}_1)+a_2(\mathbf{t})X(\mathbf{t}-\mathbf{e}_2)-a_1(\mathbf{t})a_2(\mathbf{t})X(\mathbf{t}-\mathbf{e}_0)+\varepsilon(\mathbf{t}), \tag{4.3.5}$$

where $\{\varepsilon(\mathbf{t}),\mathbf{t}\in\mathbb{Z}^2\}$ is an i.i.d. sequence with zero mean and variance σ_e^2, $(a_l(\mathbf{t}),\mathbf{t}\in\mathbb{Z}^2)$ are

i.i.d. random variables such that $E\left\{a_l^2\left(\mathbf{t}\right)\right\} = \mu_{_l}$, $l = 1, 2$. *In this case,* $\underline{X}\left(i,j\right) = \left(X\left(i,j\right), X\left(i,j-1\right), X\left(i-1\right.\right.$

$$A_1\left(i,j\right) = \begin{pmatrix} a_1\left(i,j\right) & 0 & -a_1\left(i,j\right)a_2\left(i,j\right) \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ A_2\left(i,j\right) = \begin{pmatrix} a_1\left(i,j\right) & 0 & -a_1\left(i,j\right)a_2\left(i,j\right) \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Let's now consider the max norm defined for $\underline{X}\left(i,j\right) = \left(X_1\left(i,j\right), X_2\left(i,j\right), X_3\left(i,j\right)\right)$ *by* $\left\|\underline{X}\left(i,j\right)\right\|_\infty = \max\left\{\left|X_1\left(i,j\right)\right|, \left|a_1\left(i,j\right)\right|\left|X_2\left(i,j\right)\right|, \left|a_2\left(i,j\right)\right|\left|X_3\left(i,j\right)\right|\right\}$ *and the matrix norm induced by* $\left\|.\right\|_\infty$ *i.e.,* $\left\|A_l\left(i,j\right)\right\| = \sup_{\left\|\underline{X}(i,j)\right\|_\infty=1} \frac{\left\|A_l(i,j)\underline{X}(i,j)\right\|_\infty}{\left\|\underline{X}(i,j)\right\|_\infty}$, $l = 1, 2$. *Then we have*

$$\begin{aligned} \left\|A_1\left(i,j\right)\underline{X}\left(i,j\right)\right\|_\infty &= \max\left\{\left|a_1\left(i,j\right)X_1\left(i,j\right) - a_1\left(i,j\right)a_2\left(i,j\right)X_3\left(i,j\right)\right|, \left|a_1\left(i,j\right)\right|\left|X_1\left(i,j\right)\right|\right\} \\ &\leq \max\left\{\left|a_1\left(i,j\right)\right|\left\|\underline{X}\left(i,j\right)\right\|_\infty + \left|a_1\left(i,j\right)\right|\left\|\underline{X}\left(i,j\right)\right\|_\infty, \left|a_1\left(i,j\right)\right|\left\|\underline{X}\left(i,j\right)\right\|_\infty\right\} \end{aligned}$$

so

$$\left\|A_1\left(i,j\right)\right\| \leq 2\left|a_1\left(i,j\right)\right|.$$

Also,

$$\left\|A_2\left(i,j\right)\underline{X}\left(i,j\right)\right\|_\infty = \max\left\{\left|a_2\left(i,j\right)X_1\left(i,j\right)\right|, \left|a_2\left(i,j\right)X_1\left(i,j\right)\right|\right\}$$

so

$$\left\|A_2\left(i,j\right)\right\| \leq \left|a_2\left(i,j\right)\right|.$$

Hence $E\left(\log\left\{\left\|A_1\left(i,j\right)\right\| + \left\|A_2\left(i,j\right)\right\|\right\}\right) \leq E\left(\log\left\{2\left|a_1\left(i,j\right)\right| + \left|a_2\left(i,j\right)\right|\right\}\right) \leq \log\left(2\sqrt{\mu_1} + \sqrt{\mu_2}\right)$. *So* $2\sqrt{\mu_1} + \sqrt{\mu_2} < 1$ *is a sufficient condition for the existence and the uniqueness of strictly stationary solution of (4.3.5).*

■

4.4 Higher order moments

In this section, we focus our attention on deriving conditions governing existence of r-th order moments. The key quantity of interest in determining this is

$$E\underline{X}(i,j)\left(\underline{X}'(i,j)\right)^{\otimes(r-1)},$$

using the same notations as used in the proof of Theorem 4.1, we can prove the following result

Theorem 4.4 *Let $\{\underline{X}(i,j)\}$ an $\mathcal{F}_{(i,j)}-$measurable and second order stationary solution of (4.1.1). Suppose that*

$$\left\{E\underline{\eta}(i,j)\left(\underline{\eta}'(i,j)\right)^{\otimes(r-1)} = G_r < \infty\right.,$$

and

$$\sum_{k\geq 0}\sum_{l=0}^{k} E\left[T^{l,k-l}(i,j)^{\otimes r}\right] vecG_r \text{ converges.}$$

Then the r-th order moment of $\underline{X}(i,j)$ is finite.

Proof *We shall restrict attention to the moments up to 4-th order, higher order moments can be deduced in a similar way. $\{\underline{X}(i,j)\}$ is an $\mathcal{F}_{(i,j)}-$measurable and second order stationary solution of (4.1.1), it has the form $\underline{X}(i,j) = \sum_{k\geq 0}\sum_{l=0}^{k} T^{l,k-l}(i,j)\,\underline{\eta}(i-l,j-k+l)$. Note that $E\{\underline{X}(i,j)\} = 0$ and we have*

$$\begin{aligned} & E\underline{X}(i,j)\left(\underline{X}'(i,j)\right)^{\otimes 3} \\ = & \sum_{k\geq 0}\sum_{l=0}^{k} E\left\{\left[T^{l,k-l}(i,j)\right]\underline{\eta}(i-l,j-k+l)\left(\underline{\eta}'(i-l,j-k+l)\right)^{\otimes 3}\left[T^{l,k-l}(i,j)'\right]^{\otimes 3}\right\} \end{aligned}$$

since $E\left[T^{k,l}(i,j)\,\underline{\eta}(i-n,j-m)\right] = 0$ for all (n,m) and (k,l). Via the vec operator, we get

$$vecE\underline{X}(i,j)\left(\underline{X}'(i,j)\right)^{\otimes 3} = \sum_{k\geq 0}\sum_{l=0}^{k} E\left[T^{l,k-l}(i,j)^{\otimes 4}\right] vecG_4.$$

Hence the 4-th order moment of $\underline{X}(i,j)$ is finite if $\sum_{k\geq 0}\sum_{l=0}^{k} E\left[T^{l,k-l}(i,j)^{\otimes 4}\right] vecG_4$ converges.

■

4.5 Application

4.5.1 The spatial RCA models

Random coefficients autoregressive (RCA) models are a natural extension of classical autoregressive (AR) models very important in linear time series analysis. Pioneering investigations in this direction are due to Conlisk (1974, 1976) who dealt with econometric modeling questions and Andel (1976) who argued that many engineering applications can be adequately described by RCA models. Extending RCA time series models by taking into account spatial heterogeneity, we define the spatial 2-D RCA model $X(i,j)$ which satisfies the representation

$$X(i,j) = \sum_{(k,l)\in S]\mathbf{0},\mathbf{p}]} [\beta_{kl} + \beta_{kl}(i,j)] X(i-k, j-l) + e(i,j), \tag{4.5.1}$$

where $\mathbf{p} = (p_1, p_2)$, $\mathbf{0} = (0,0)$ and the indexed set $S]\mathbf{0},\mathbf{p}]$ is defined as
$S]\mathbf{0},\mathbf{p}] = \{(l,m) \in \mathbb{Z}^2,\ (l,m) \prec (p_1,p_2) \text{ or } (l,m) = (p_1,p_2)\} \setminus \{(0,0)\}$, and $\prec$ denote the lexicographic order on $\mathbb{Z}^2$.

For this model we need the following assumptions

(a) $\{\varepsilon(i,j);\ (i,j) \in \mathbb{Z}^2\}$ is an independent second order stationary sequence of d-variate random variable with mean zero and variance σ^2.

(b) The β_{kl}, $(k,l) \in S]\mathbf{0},\mathbf{p}]$, are real constants.

(c) $\{\beta_{kl}(i,j);\ (i,j) \in \mathbb{Z}^2,\ (k,l) \in S]\mathbf{0},\mathbf{p}]\}$ is an independent sequence of random variables with mean zero. $\{\beta_{kl}(i,j);\ (k,l) \in S]\mathbf{0},\mathbf{p}]\}$ is independent of $\{\varepsilon(i,j)\}$ for all $(i,j) \in \mathbb{Z}^2$.

(d) Letting

$$\mathcal{B}_1(i,j) = \begin{pmatrix} \underline{\beta}_0(i,j) & \dots & \underline{\beta}_{p_1}(i,j) \\ 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix},\ \mathcal{B}_2(i,j) = \begin{pmatrix} \overline{\beta}_1(i,j) & \dots & \overline{\beta}_{p_1}(i,j) & \overline{0} \\ 0 & \dots & 0 & \vdots \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & \overline{0} \end{pmatrix}$$

where $\underline{\beta}_k(i,j)$, $\overline{\beta}_s(i,j)$ and $\overline{0}$ are (p_2+1)-vectors such that, $\underline{\beta}_k(i,j) = \left(\beta_{k1}(i,j), \dots, \beta_{kp_2}(i,j), 0\right)$,

and $\overline{\beta}_s(i,j) = (\beta_{s0}(i,j), 0, \ldots, 0)$, $0 \leq k \leq p_1$, $1 \leq s \leq p_1$and $\overline{0} = (0, \ldots, 0)$, we suppose that

$$E\{\mathcal{B}_1(i,j) \otimes \mathcal{B}_1(i,j)\} = \mathcal{I}_1 \text{ and } E\{\mathcal{B}_2(i,j) \otimes \mathcal{B}_2(i,j)\} = \mathcal{I}_2.$$

To apply Theorem 4.1 to the SRCA model, we need to represent it as a state-space version of equation (4.5.1), this can be done by defining the $(p_1+1)p_2 \times 1$ vectors

$$\underline{X}(i,j) = (x_0(i,j), \ldots, x_{p_1}(i,j))' \text{ and } \underline{\varepsilon}(i,j) = (\varepsilon(i,j), 0, \ldots, 0)$$

where $x_k(i,j) = (X(i-k,j), X(i-k,j-1), \ldots, X(i-k,j-p_2))$, $k = 0, \ldots, p_1$ we also need the $(p_1+1)(p_2+1) \times (p_1+1)(p_2+1)$ matrix

$$\mathcal{M}_1 = \begin{pmatrix} \underline{\beta}_0 & \underline{\beta}_1 & \cdots & \cdots & \underline{\beta}_{p_1} \\ \mathcal{J} & O_{p_2\times(p_2+1)} & \cdots & \cdots & O_{p_2\times(p_2+1)} \\ O_{1\times(p_2+1)} & \cdots & \cdots & & O_{1\times(p_2+1)} \\ O_{p_2\times(p_2+1)} & \mathcal{J} & O_{p_2\times(p_2+1)} & \cdots & O_{p_2\times(p_2+1)} \\ O_{1\times(p_2+1)} & \cdots & \ddots & & O_{1\times(p_2+1)} \\ \vdots & & & \ddots & \vdots \\ O_{p_2\times(p_2+1)} & \cdots & & O_{p_2\times(p_2+1)} & \mathcal{J} \end{pmatrix},$$

$$\mathcal{M}_2 = \begin{pmatrix} \overline{\beta}_1 & \cdots & \overline{\beta}_{p_1} & \overline{0} \\ O_{p_2\times(p_2+1)} & & & \\ \mathcal{T}_{(p_2+1)\times(p_2+1)} & \cdots & O_{(p_2+1)\times(p_2+1)} & \vdots \\ \vdots & \ddots & \vdots & \vdots \\ O_{(p_2+1)\times(p_2+1)} & \cdots & \mathcal{T}_{(p_2+1)\times(p_2+1)} & O_{(p_2+1)\times(p_2+1)} \end{pmatrix}$$

where $\underline{\beta}_k = (\beta_{k1}, \ldots, \beta_{kp_2}, 0)$, $\overline{\beta}_s = (\beta_{s0}, 0, \ldots, 0)$, $0 \leq k \leq p_1$, $1 \leq s \leq p_1$, $O_{p\times q}$ denotes the

null matrix of order $p \times q$ and

$$\mathcal{J} = \begin{pmatrix} 1 & 0 & \dots & & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}_{p_2 \times (p_2+1)} \text{, and } \mathcal{T}_{r\times s} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & & \\ \vdots & & \ddots & \\ 0 & & & 0 \end{pmatrix}_{r\times s}$$

.

With this notations, equation (4.5.1) can be rewritten in the form

$$\underline{X}(i,j) = [\mathcal{M}_1 + \mathcal{B}_1(i,j)]\,\underline{X}(i,j-1) + [\mathcal{M}_2 + \mathcal{B}_2(i,j)]\,\underline{X}(i-1,j) + \underline{\varepsilon}(i,j).$$

In this case,

$$G = E\underline{\varepsilon}(i,j)\,\underline{\varepsilon}'(i,j) = \sigma^2\mathcal{T}_{(p_1+1)p_2\times(p_1+1)p_2},\ C_1 = \mathcal{M}_1 \otimes \mathcal{M}_1 + \mathcal{I}_1 \text{ and } C_2 = \mathcal{M}_2 \otimes \mathcal{M}_2 + \mathcal{I}_2.$$

Hence we deduce from Theorem 4.1 that when $[\mathcal{M}_1 \otimes \mathcal{M}_1 + \mathcal{I}_1 + \mathcal{M}_2 \otimes \mathcal{M}_2 + \mathcal{I}_2]$ has not a unit eigenvalues, (4.5.1) has a unique stationary $\mathcal{F}_{(i,j)}-$measurable solution if and only if the condition $\sum_{k=0}^{r}\sum_{l=0}^{k} E\left[T^{l,k-l}(i,j) \otimes T^{l,k-l}(i,j)\right] \times vecG$ converges as $r \to \infty$ is satisfied

Chapter 5

A generalized moments estimator for random coefficients autoregressive model on a plane

5.1 Introduction

We have previously seen in chapter 4 that 2-D RCA models are a natural extension of RCA time series, conditions were found for the existence of stationary and ergodic solutions to equations of the form (4.5.1). Estimation procedures for RCA time series are well established, and the asymptotic properties of these estimates are well known (see, for example Niccolls and Quinn (1982)). For the spatial and nonlinear 2-D RCA model, estimation procedures remain totally unexplored. In part because of this, in this chapter we will consider generalized methods of moments (GMM) estimation which provide convenient and efficient estimators. As an alternative motivation, GMM estimators are often available where a likelihood analysis is extremely difficult or even impossible (for example, if there is much uncertainty on the distributional form). The GMM estimator is derived directly from a set of minimal assumptions, the so-called moment conditions that the model should satisfy. On the other hand, a practical difficulty with the usual method of ML is its computational cost on medium to large sample size. Moreover, the GMM estimation includes some well-known estimators as special cases, so that they can

be analyzed conveniently on a theoretical level. In a spatial framework, Kelejian and Prucha (1999) introduced a generalized moments estimator for the autoregressive parameter of the spatial disturbance process that is simple to compute and remains computationally feasible even for large sample sizes. Recently (2010), they introduced a new class of GMM estimators for the autoregressive parameter of a spatially autoregressive disturbance process that allows for heteroskedastic innovations. They proved that their GMM estimators are again computationally simple even in large samples. Hence our aim is to propose a GMM estimators for a 2-D RCA model parameters which are simple to compute and which have desirable sample properties.
The chapter is organized as follows: In Section 2 we summarize the properties of Generalized Method of Moments (GMM) estimators and discuss the underlying assumptions. Section 3 defines and establishes the GMM estimators for parameters of the spatial RCA model defined by equation (4.5.1).

5.2 Overview of the GMM methodology and assumptions

In this section we will consider GMM methodology. To introduce the class of GMM estimators let $\underline{w}(i,j)$ be an $N_1N_2 \times 1$ vector of random variables that are observed at location (i,j), let $\underline{\theta} \in \Omega$ denotes an unknown $a \times 1$ vector of coefficients, and let $h(\underline{\theta}, \underline{w}(i,j))$ be an $r \times 1$ vector-valued function. Define for $\mathbf{N} = (N_1, N_2)$, $g_{\mathbf{N}}(\underline{\theta}) = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} h(\underline{\theta}, \underline{w}(i,j))$. Moreover assume that $\underline{\theta}_0$ is the true value of $\underline{\theta}$ and that the moment conditions can be written as $E_{\underline{\theta}_0}[h(\underline{\theta}, \underline{w}(i,j))] = 0$. The GMM estimator is defined by solving the minimization problem

$$\hat{\underline{\theta}}_{\mathbf{N}} = \arg\min_{\underline{\theta} \in \Omega} [g_{\mathbf{N}}(\underline{\theta})]' W_{\mathbf{N}} [g_{\mathbf{N}}(\underline{\theta})], \tag{5.2.1}$$

where the weight matrix $W_{\mathbf{N}}$ is positive definite.
As a convenient framework the following set of assumptions are met.

I.1 The process $(\underline{w}(i,j), \ (i,j) \in \mathbb{Z}^2)$ is strictly stationary, ergodic and admits fourth order moments.

I.2 The parameter space Ω is compact and $\underline{\theta}_0$ is in its interior.

I.3 For all $\underline{\theta} \in \Omega$, $h(\underline{\theta}, .)$ is a measurable function.

I.4 $h\left(\underline{\theta}, \underline{w}\left(i,j\right)\right)$ is continuous and continuously differentiable in $\underline{\theta}$ on a neighborhood of $\underline{\theta}_0$.

I.5 $W_{\mathbf{N}}$ a sequence of positive definite $r \times r$ matrices which converges in probability to the deterministic positive definite matrix W.

I.6 $\left|h_k\left(\underline{\theta}, \underline{w}\left(i,j\right)\right)\right| \leq f\left(\underline{w}\left(i,j\right)\right)$ for all $\underline{\theta} \in \Omega$ and $k = 1, \ldots, r$ where $f\left(.\right)$ is a non negative function such that $E_{\underline{\theta}_0}\left\{f\left(\underline{w}\left(i,j\right)\right)\right\} < \infty$.

I.7 $\underline{\theta}_0$ is the unique solution of $E_{\underline{\theta}_0}\left[h\left(\underline{\theta}, \underline{w}\left(i,j\right)\right)\right] = 0$ i.e., $E_{\underline{\theta}_0}\left[h\left(\underline{\theta}, \underline{w}\left(i,j\right)\right)\right] = 0$ if and only if $\underline{\theta} = \underline{\theta}_0$.

I.8 $\sqrt{N_1 N_2} g_{\mathbf{N}}\left(\underline{\theta}_0\right) \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}\left(0, \mathcal{S}\right)$.

I.9 For all sequence $\left(\underline{\theta}^*_{\mathbf{N}},\ \mathbf{N} \in \mathbb{N}^2\right)$ such that $\underline{\theta}^*_{\mathbf{N}} \xrightarrow{P} \underline{\theta}_0$, $\left.\frac{\partial g_N(\theta)}{\partial \underline{\theta}'}\right|_{\underline{\theta}^*_{\mathbf{N}}}$ converges in probability to $\mathcal{D}'$, where $\mathcal{D}' = E\left.\frac{\partial h(\underline{\theta}, \underline{w}(i,j))}{\partial \underline{\theta}'}\right|_{\underline{\theta}_0}$ with the columns of $\mathcal{D}'$ linearly independent.

Let $Dg_{\mathbf{N}}\left(\underline{\theta}\right)$ be the $r \times a$ dimensional matrix of derivatives with typical element $Dg_{\mathbf{N}}\left(\underline{\theta}\right)_{ij} = \frac{\partial g_{\mathbf{N}}(\theta)_i}{\partial \underline{\theta}_j}$. Then $\hat{\underline{\theta}}_{\mathbf{N}}$ can also be defined in terms of the solution to the first order conditions of the minimization problem

$$Dg_{\mathbf{N}}\left(\hat{\underline{\theta}}_{\mathbf{N}}\right)' W_{\mathbf{N}}\left[g_{\mathbf{N}}\left(\hat{\underline{\theta}}_{\mathbf{N}}\right)\right] = 0.$$

When $r = a$, the model is exactly identified, the weighting matrix becomes irrelevant and $\hat{\underline{\theta}}_{\mathbf{N}}$ is defined by the equation $g_{\mathbf{N}}\left(\hat{\underline{\theta}}_{\mathbf{N}}\right) = 0$.

Given these assumptions the GMM estimator is consistent and asymptotically normally distributed with

$$\sqrt{N_1 N_2}\left(\hat{\underline{\theta}}_{\mathbf{N}} - \underline{\theta}_0\right) \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}\left(0, \mathcal{V}\right), \tag{5.2.2}$$

where $\mathcal{V} = \left(\mathcal{D}\mathcal{S}^{-1}\mathcal{D}'\right)^{-1}$.

Before applying this result to the 2-D RCA model, it is useful to comment on the meaning and the role of regularity conditions. Assumption (**I.1**) is necessary for checking the TCL given in (**I.8**) and to apply the ergodic theorem in several times. Assumption (**I.2**) restricts the parameters to a closed and bounded subset of Euclidian space, so when it holds the estimator can be characterized in terms of its first order condition. Assumptions (**I.3**) and (**I.4**) are mathematical regularity conditions that guarantee that the moment functions are continuous and have finite variances, and that in a neighborhood of $\underline{\theta}_0$ they can be differentiated. Assumption (**I.7**) is a key

identification condition that rules out both local identification failure. Assumptions (**I.9**) and (**I.8**) give (5.2.2).

5.3 Application of GMM to the 2-D RCA model

Notice first that the 2-D RCA model (4.5.1) can be rewritten as

$$X(i,j) = \underline{x}'(i,j)\,\underline{\theta} + U(i,j), \tag{5.3.1}$$

where $\underline{\theta} = \left(\beta_{01}, \ldots, \beta_{0p_2}, \beta_{10}, \ldots \beta_{1p_2}, \ldots, \beta_{p_10}, \ldots \beta_{p_1p_2}\right)'$, $U(i,j) = \underline{\beta}(i,j)\,\underline{x}(i,j) + \varepsilon(i,j)$, $\underline{\beta}(i,j) = \left(\beta_{01}(i,j), \ldots, \beta_{0p_2}(i,j), \beta_{10}(i,j), \ldots \beta_{1p_2}(i,j), \ldots, \beta_{p_1}(i,j), \ldots \beta_{p_1p_2}(i,j)\right)$ and where $\underline{x}(i,j)$ equal to

$(X(i,j-1), \ldots, X(i,j-p_2), X(i-1,j), \ldots, X(i-1,j-p_2), \ldots,$

$X(i-p_1,j), \ldots, X(i-p_1,j-p_2))'$.

Assumptions (a), (b) and (c) give these moment conditions

$$E\left\{\underline{x}(i,j)\left[X(i,j) - \underline{x}'(i,j)\,\underline{\theta}_0\right]\right\} = 0,$$

which represents a system of $[(p_1+1)(p_2+1)-1]$ orthogonality conditions.

Put $\underline{w}(i,j) = X(i,j)\,\underline{x}(i,j)$, we get

$$\begin{cases} h\left(\underline{\theta}, \underline{w}(i,j)\right) = \underline{x}(i,j)\left[X(i,j) - \underline{x}'(i,j)\,\underline{\theta}\right] \\ E_{\underline{\theta}_0}\left\{h\left(\underline{\theta}, \underline{w}(i,j)\right)\right\} = 0 \end{cases}. \tag{5.3.2}$$

Since there is $a = [(p_1+1)(p_2+1)-1]$ parameters for $r = [(p_1+1)(p_2+1)-1]$ conditions of orthogonality, the system is exactly identified and the GMM estimator $\hat{\underline{\theta}}_{\mathbf{N}}$ of $\underline{\theta}$ based on a sample of N^2 observations is defined by the equation $g_{\mathbf{N}}\left(\hat{\underline{\theta}}_{\mathbf{N}}\right) = 0$, where

$$g_{\mathbf{N}}(\underline{\theta}) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\underline{x}(i,j)\left[X(i,j) - \underline{x}'(i,j)\,\underline{\theta}\right], \tag{5.3.3}$$

hence we just solve the following system,

$$\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\underline{x}\left(i,j\right)\left[X\left(i,j\right)-\underline{x}'\left(i,j\right)\hat{\underline{\theta}}_{\mathbf{N}}\right]=0,$$

which can be rewritten as

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\underline{x}\left(i,j\right)X\left(i,j\right)=\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\underline{x}\left(i,j\right)\underline{x}'\left(i,j\right)\right)\hat{\underline{\theta}}_{\mathbf{N}},$$

consequently

$$\hat{\underline{\theta}}_{\mathbf{N}}=\left(\sum_{i=1}^{N}\sum_{j=1}^{N}\underline{x}\left(i,j\right)\underline{x}'\left(i,j\right)\right)^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\underline{x}\left(i,j\right)X\left(i,j\right).$$

In order to obtain consistency and asymptotic normality of $\hat{\underline{\theta}}_{\mathbf{N}}$, we must verify the regularity conditions Assumption *1.* to *9.* for the 2-D RCA model. In addition of assumptions a), b) and c), assume that also the parameters β_{kl}, $(k,l)\in S\,]\mathbf{0},\mathbf{p}]$ and γ^2 are such that a unique second order stationary solution $X\left(i,j\right)$ to (4.5.1) exists and that this solution admits fourth order moments.

Since $X\left(i,j\right)$ is strictly stationary, ergodic and admits fourth order moments, so is $\underline{w}\left(i,j\right)$, this gives Assumption *1.*

The conditions of stationarity and existence of fourth order moments define the parameter space Ω which will be in this case compact. So, we verify Assumption *2.*

Considering the expression of the function $h\left(\underline{\theta},\underline{w}\left(i,j\right)\right)$ in (5.3.2), for all $\underline{\theta}\in\Omega$, $h\left(\underline{\theta},.\right)$ is a measurable function and $h\left(\underline{\theta},\underline{w}\left(i,j\right)\right)$ is continuous and continuously differentiable in $\underline{\theta}$ on a neighborhood of $\underline{\theta}_0$ which leads to Assumption *3* and *4.*

Moreover, basic calculations show that $h\left(\underline{\theta},\underline{w}\left(i,j\right)\right)=\underline{x}\left(i,j\right)\left[X\left(i,j\right)-\underline{x}'\left(i,j\right)\underline{\theta}\right]$ are serially uncorrelated, a good candidate for the weighting matrix $W_{\mathbf{N}}$ for minimizing (5.2.1) is then $W_{\mathbf{N}}=\left[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}h\left(\tilde{\underline{\theta}}_{\mathbf{N}},\underline{w}\left(i,j\right)\right)h'\left(\tilde{\underline{\theta}}_{\mathbf{N}},\underline{w}\left(i,j\right)\right)\right]^{-1}$,where $\tilde{\underline{\theta}}_{\mathbf{N}}$ is any consistent estimate of $\underline{\theta}_0$. Note that one good way to get a consistent preliminary estimator $\tilde{\underline{\theta}}_{\mathbf{N}}$ is to minimize (5.2.1) with an arbitrary weithing matrix such as $W_{\mathbf{N}}=I_r$. Since $h\left(\underline{\theta},\underline{w}\left(i,j\right)\right)$ is strictly stationary and ergodic (it's a measurable function of $X\left(i-k,j-l\right)$, $\beta_{kl}\left(i,j\right)$ and $e\left(i,j\right)$, $(k,l)\preceq(p_1,p_2)$), by

the ergodic theorem and the fact that $\tilde{\underline{\theta}}_{\mathbf{N}} \xrightarrow{P} \underline{\theta}_0$ we get

$$\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} h\left(\tilde{\underline{\theta}}_{\mathbf{N}}, \underline{w}(i,j)\right) h'\left(\tilde{\underline{\theta}}_{\mathbf{N}}, \underline{w}(i,j)\right) \xrightarrow{P} \mathcal{S} = Eh\left(\underline{\theta}_0, \underline{w}(i,j)\right) h'\left(\underline{\theta}_0, \underline{w}(i,j)\right).$$

Hence $W_{\mathbf{N}} \xrightarrow{P} W = \mathcal{S}^{-1}$, it's evident that $W_{\mathbf{N}}$ and W are both definite positive. So, we get Assumption *5.*

Also, $|h_k(\underline{\theta}, \underline{w}(i,j))| = |X(i-i_k, j-j_k)[X(i,j) - \underline{x}'(i,j)\underline{\theta}]|$

$\leq |X(i-i_k, j-j_k)|\{|X(i,j)| + \|\underline{x}'(i,j)\|\,\|\underline{\theta}\|\}$, since the parameter space is compact, $\|\underline{\theta}\|$ is bounded and we verify Assumption *6*.

Assumption *7.* is easily verified.

Now, the following lemma ensures Assumption *8.*

Lemma 5.1 *Let $X(i,j)$ be a strictly stationary and ergodic solution of equation (4.5.1) which admits fourth order moments.*

Then

$$\sqrt{N_1 N_2}\, g_{\mathbf{N}}(\underline{\theta}_0) \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}(0, \mathcal{S}),$$

where $g_{\mathbf{N}}(\underline{\theta}_0)$ is given by (5.3.3) and $\mathcal{S} = E_{\underline{\theta}}\left\{\underline{x}(i,j)\,\underline{x}'(i,j)\left(\sigma^2 + \left[\underline{x}(i,j)^{\otimes 2}\right]'\underline{\Sigma}\right)\right\}$, where $\underline{\Sigma} = VecE_{\underline{\theta}}\{\underline{\theta}'(i,j)\,\underline{\theta}(i,j)\}$.

Proof To prove this lemma we will use the Huang martingale central limit theorem (Theorem A.1.1). Now if α is any r-component vector, we first show that $z(i,j) = \alpha'\underline{x}(i,j)\,U(i,j)$ are spatial martingales differences of Huang (Definition A.1.1). It's easily verified that $E_{\underline{\theta}}(z(i,j)\backslash z(k,l),\ (k,l) \prec (i,j)) = 0$, this means that $z(i,j) = \alpha'\underline{x}(i,j)\,U(i,j)$ are spatial martingales differences of Huang. To apply the CLT of Huang to $z(i,j)$, we must verify the regularity conditions of this CLT. The first condition we must verify is $\frac{1}{N^2}\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{N}]}[z(i,j)]^2 \xrightarrow{P} \eta^2$.

Put $\mathcal{F}_{\mathbf{k}} = \sigma\left\{\left(\underline{\theta}\left(\mathbf{k}-\mathbf{s}\right), \varepsilon\left(\mathbf{k}-\mathbf{s}\right)\right),\ \mathbf{0} \prec \mathbf{s}\right\}$,

$$\begin{aligned} E_{\underline{\theta}}\left(z\left(i,j\right)^2\right) &= E_{\underline{\theta}}\left\{E_{\underline{\theta}}\left(\left(\alpha'\underline{x}\left(i,j\right)U\left(i,j\right)\right)^2\right)\backslash\mathcal{F}_{(i,j)}\right\} \\ &= E_{\underline{\theta}}\left\{\left[\alpha'\underline{x}\left(i,j\right)\right]^2 E_{\underline{\theta}}\left(U\left(i,j\right)^2\backslash\mathcal{F}_{(i,j)}\right)\right\} \\ &= E_{\underline{\theta}}\left\{\left[\alpha'\underline{x}\left(i,j\right)\right]^2 E_{\underline{\theta}}\left(U\left(i,j\right)^2\backslash\mathcal{F}_{(i,j)}\right)\right\}, \end{aligned}$$

we have also

$$\begin{aligned} E_{\underline{\theta}}\left(U\left(i,j\right)^2\backslash\mathcal{F}_{(i,j)}\right) &= E_{\underline{\theta}}\left(\left[\underline{\theta}\left(i,j\right)\underline{x}\left(i,j\right)+\varepsilon\left(i,j\right)\right]^2\backslash\mathcal{F}_{(i,j)}\right) \\ &= E_{\underline{\theta}}\left(\left[\underline{\theta}\left(i,j\right)\underline{x}\left(i,j\right)\right]^2\backslash\mathcal{F}_{(i,j)}\right)+2E_{\underline{\theta}}\left(\left\{\varepsilon\left(i,j\right)\underline{\theta}\left(i,j\right)\underline{x}\left(i,j\right)\right\}\backslash\mathcal{F}_{(i,j)}\right) \\ &\quad +E_{\underline{\theta}}\left\{\varepsilon^2\left(i,j\right)\right\} \\ &= E_{\underline{\theta}}\left(\left[\underline{x}'\left(i,j\right)\underline{\theta}'\left(i,j\right)\underline{\theta}\left(i,j\right)\underline{x}\left(i,j\right)\right]\backslash\mathcal{F}_{(i,j)}\right) \\ &\quad +2E_{\underline{\theta}}\left\{\varepsilon\left(i,j\right)\right\}E_{\underline{\theta}}\left(\left\{\underline{\theta}\left(i,j\right)\underline{x}\left(i,j\right)\right\}\backslash\mathcal{F}_{(i,j)}\right)+\sigma^2 \\ &= \sigma^2+\underline{x}'\left(i,j\right)E_{\underline{\theta}}\left\{\underline{\theta}'\left(i,j\right)\underline{\theta}\left(i,j\right)\right\}\underline{x}\left(i,j\right) \\ &= \sigma^2+\left[\underline{x}\left(i,j\right)^{\otimes 2}\right]'\underline{\Sigma}, \end{aligned}$$

where $\underline{\Sigma} = VecE_{\underline{\theta}}\left\{\underline{\theta}'\left(i,j\right)\underline{\theta}\left(i,j\right)\right\}$. Hence

$$E_{\underline{\theta}}\left[z\left(i,j\right)\right]^2 = E_{\underline{\theta}}\left\{\left[\alpha'\underline{x}\left(i,j\right)\right]^2\left(\sigma^2+\left[\underline{x}\left(i,j\right)^{\otimes 2}\right]'\underline{\Sigma}\right)\right\}.$$

This expectation exists if $E_{\underline{\theta}}\left\{X^4\left(i,j\right)\right\} < \infty$ since the components of $z\left(i,j\right)z'\left(i,j\right)\left[\underline{x}\left(i,j\right)^{\otimes 2}\right]'\underline{\Sigma}$ are quartic in $X\left(i,j\right)$. Now $z\left(i,j\right)$ is strict stationary and ergodic (it's a measurable function of $X\left(i-k,j-l\right)$, $\beta_{kl}\left(i,j\right)$ and $e\left(i,j\right)$, $\left(k,l\right) \preceq \left(p_1,p_2\right)$) , by the ergodic theorem

$$\frac{1}{N^2}\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{N}]}\left[z\left(i,j\right)\right]^2 \overset{a.s}{\to} E_{\underline{\theta}}\left[z\left(i,j\right)\right]^2.$$

It remains to verify that $\sup_{\mathbf{k}} E_{\underline{\theta}}\left[z\left(i,j\right)\right]^2 1_{\{|z(i,j)|>C\}} \searrow 0$ as $C \to \infty$.
Put $\eta^2 = E_{\underline{\theta}}\left[z\left(i,j\right)\right]^2 = \alpha' E_{\underline{\theta}}\left\{\underline{x}\left(i,j\right)\underline{x}'\left(i,j\right)\left(\sigma^2+\left[\underline{x}\left(i,j\right)^{\otimes 2}\right]'\underline{\Sigma}\right)\right\}\alpha$, since $E_{\underline{\theta}}\left[z\left(i,j\right)\right]^2 < \infty$, we can state that

$\mathrm{P}\left\{\sup_k E_{\underline{\theta}}\left[z\left(i,j\right)\right]^2 < \infty\right\} = 1$, furthermore the sequence of random variables indexed by C, $\sup_k E_{\underline{\theta}}\left\{\left[z\left(i,j\right)\right]^2 1_{\{|Z_{\mathbf{k}}|>C\}}\right\}$ (we can take $C \in \mathbb{N}$) is decreasing, so

$$\sup_{\mathbf{k}} E_{\underline{\theta}}\left\{\left[z\left(i,j\right)\right]^2 1_{\{|Z_{\mathbf{k}}|>C\}}\right\} \searrow 0 \text{ as } C \to \infty,$$

on the set $\left\{\sup_{\mathbf{k}} E_{\underline{\theta}}\left[z\left(i,j\right)\right]^2 < \infty\right\}$. Hence $\frac{1}{N}\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{N}]} z\left(i,j\right) \underset{N\to\infty}{\overset{\mathcal{L}}{\to}} \mathcal{N}\left(0,\eta^2\right)$. Thus $\frac{1}{N}\sum_{\mathbf{k}\in S[\mathbf{1},\mathbf{N}]} \underline{x}\left(i,j\right) U\left(i,j\right)$ converges in distribution to the multivariate normal distribution with mean zero and covariance matrix

$$\mathcal{S} = E_{\underline{\theta}}\left\{\underline{x}\left(i,j\right)\underline{x}'\left(i,j\right)\left(\sigma^2 + \left[\underline{x}\left(i,j\right)^{\otimes 2}\right]'\underline{\Sigma}\right)\right\}.$$

■

Finally, Let's verify Assumption ***9.,*** we have

$$\begin{aligned}\frac{\partial g_{\mathbf{N}}\left(\underline{\theta}\right)}{\partial \underline{\theta}'} &= -\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\partial h\left(\underline{\theta},\underline{w}\left(i,j\right)\right)}{\partial \underline{\theta}'} \\ &= -\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\underline{x}\left(i,j\right)\underline{x}'\left(i,j\right),\end{aligned}$$

by the ergodic theorem and the fact that $\left\|\frac{\partial h(\underline{\theta},\underline{w}(i,j))}{\partial \underline{\theta}'}\right\| \leq E\underline{x}\left(i,j\right)\underline{x}'\left(i,j\right)$, $\frac{\partial g_{\mathbf{N}}(\underline{\theta})}{\partial \underline{\theta}'} \xrightarrow{P} E\frac{\partial h(\underline{\theta},\underline{w}(i,j))}{\partial \underline{\theta}'} = -E\underline{x}\left(i,j\right)\underline{x}'\left(i,j\right)$ uniformly for $\underline{\theta} \in \Omega$, furthermore $E\frac{\partial h(\underline{\theta},\underline{w}(i,j))}{\partial \underline{\theta}'}$ is almost surely continuous at $\underline{\theta}_0$, hence by the continuous mapping theorem, for all sequence $\left(\underline{\theta}^*_{\mathbf{N}},\ \mathbf{N} \in \mathbb{N}^2\right)$ such that $\underline{\theta}^*_{\mathbf{N}} \xrightarrow{P} \underline{\theta}_0$, $\left.\frac{\partial g_N(\underline{\theta})}{\partial \underline{\theta}'}\right|_{\underline{\theta}^*_{\mathbf{N}}}$ converges in probability to $\mathcal{D} = E\left.\frac{\partial h(\underline{\theta},\underline{w}(i,j))}{\partial \underline{\theta}'}\right|_{\underline{\theta}_0}$.
With all the assumptions verified, we can derive the asymptotic properties of the most important parameter associated with (4.5.1) as

$$N\left(\hat{\underline{\theta}}_{\mathbf{N}} - \underline{\theta}_0\right) \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}\left(0, \mathcal{D}^{-1}\mathcal{S}\mathcal{D}^{-1}\right),$$

where $\mathcal{D} = E\left.\frac{\partial h(\underline{\theta},\underline{w}(i,j))}{\partial \underline{\theta}'}\right|_{\underline{\theta}_0} = -E\underline{x}\left(i,j\right)\underline{x}'\left(i,j\right)$, and $\mathcal{S} = E_{\underline{\theta}}\left\{\underline{x}\left(i,j\right)\underline{x}'\left(i,j\right)\left(\sigma^2 + \left[\underline{x}\left(i,j\right)^{\otimes 2}\right]'\underline{\Sigma}\right)\right\}$.

In fact, $\mathcal{S} = \sigma^2_{U(i,j)} E_{\underline{\theta}} \left\{ \underline{x}(i,j) \underline{x}'(i,j) \right\}$ where $\sigma^2_{U(i,j)} = E_{\underline{\theta}} U^2(i,j) = \left[\sigma^2 + \left[\underline{x}(i,j)^{\otimes 2} \right]' \underline{\Sigma} \right]$. To complete the estimation procedure, we need to have an estimate of $\sigma^2_{U(i,j)}$. By using (5.3.1), we form the residuals $\hat{U}(i,j) = X(i,j) - \underline{x}'(i,j) \tilde{\underline{\theta}}_{\mathbf{N}}$, $(i,j) \in \mathbb{Z}^2$ where $\tilde{\underline{\theta}}_{\mathbf{N}}$ is any preliminary consistent estimator of $\underline{\theta}$. Consequently a consistent estimate of $\sigma^2_{U(i,j)}$ can be obtained by

$$\hat{\sigma}^2_{U(i,j)} = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \hat{U}^2(i,j).$$

Example 5.1 *Consider a simple 2-D RCA model of first order defined by*

$$X(i,j) = \alpha X(i,j-1) + \beta X(i-1,j) + \gamma X(i-1,j-1) + \beta(i,j) X(i-1,j) + \varepsilon(i,j),$$

where $\{\varepsilon(i,j);\ (i,j) \in \mathbb{Z}^2\}$ *is an i.i.d sequence of random variables with mean zero, variance* 1 *and* $E\varepsilon^4(i,j) < \infty$, α *and* β *are constants and* $\{\beta(i,j);\ (i,j) \in \mathbb{Z}^2\}$ *is an other i.i.d sequence of random variables with mean zero,* $E\left[\beta^2(i,j)\right] = \frac{1}{2}$ *and* $E\beta^4(i,j) < \infty$. $\beta(i,j)$ *is independent of* $\{\varepsilon(i,j);\ (i,j) \in \mathbb{Z}^2\}$. *In this case*

$$\mathcal{M}_1 = \begin{pmatrix} \alpha & 0 & \gamma & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \mathcal{M}_2 = \begin{pmatrix} \beta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\mathcal{B}_1(i,j) = O_{4\times 4}, \ \mathcal{B}_2(i,j) = \begin{pmatrix} \beta(i,j) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

hence

$$\mathcal{I}_1 = E\left[\mathcal{B}_1(i,j) \otimes \mathcal{B}_1(i,j)\right] = O_{16\times 16},$$

$$\mathcal{I}_2 = E\left[\mathcal{B}_2(i,j) \otimes \mathcal{B}_2(i,j)\right] = \begin{pmatrix} \frac{1}{2} & 0 & \dots & 0 \\ & O_{15\times 16} & & \end{pmatrix}$$

and the condition of stationarity is that $\sum_{k\geq 0}\sum_{l=0}^{k} E\left[T^{l,k-l}(i,j)^{\otimes 2}\right]$ *vecG converges. Condition for the existence of fourth order moment is given by*

$$\sum_{k\geq 0}\sum_{l=0}^{k} E\left[T^{l,k-l}(i,j)^{\otimes 4}\right] vecG_4 \text{ converges.}$$

where

1) $T^{0,1}(i,j) = \begin{pmatrix} \alpha & 0 & \gamma & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ *and* $T^{1,0}(i,j) = \begin{pmatrix} \beta+\beta(i,j) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

2) $T^{n,m}(i,j) = A_1(i,j)\,T^{n,m-1}(i,j-1) + A_2(i,j)\,T^{n-1,m}(i-1,j)$

3) $T^{0,0}(i,j) = I_{4\times 4}$, *and*

4) $T^{-n,m}(i,j) = T^{n,-m}(i,j) = O_{4\times 4}$.

After some calculations we find that $\sum_{k\geq 0}\sum_{l=0}^{k} E\left[T^{l,k-l}(i,j)^{\otimes 4}\right] vecG_4$ *converges as* $k\to\infty$ *if* $\alpha^4 < 1$, *provided that stationary conditions are satisfied..*

$r = a = 3$, $\underline{\theta} = (\alpha,\beta,\gamma)$, $\underline{x} = (X(i,j-1), X(i-1,j), X(i-1,j-1))$,

$\underline{w}(i,j) = X(i,j)\,\underline{x}(i,j-1)$, *and*

$$\begin{aligned} h(\underline{\theta},\underline{w}(i,j)) &= \underline{x}(i,j)\left[X(i,j) - \underline{x}'(i,j)\,\underline{\theta}\right] \\ &= \left(X(i,j-1)\,U(i,j), X(i-1,j)\,U(i,j), X(i-1,j-1)\,U(i,j)\right), \end{aligned}$$

where $U(i,j) = X(i,j) - \alpha X(i,j-1) - \beta X(i-1,j) - \gamma X(i-1,j-1)$. *Assume that* $\underline{\theta}_0 = (\alpha_0,\beta_0,\gamma_0)$ *is the true value of* $\underline{\theta}$, *our GMM estimator of* $\underline{\theta}$ *is*

$$\hat{\underline{\theta}}_{\mathbf{N}} = \left(\sum_{i=1}^{N}\sum_{j=1}^{N}\left[X^2(i,j-1) + X^2(i-1,j) + X^2(i-1,j-1)\right]\right)^{-1} \sum_{i=1}^{N}\sum_{j=1}^{N}\underline{x}(i,j)\,X(i,j).$$

with

$$N\left(\hat{\underline{\theta}}_{\mathbf{N}} - \underline{\theta}_0\right) \overset{\mathcal{L}}{\rightsquigarrow} \mathcal{N}\left(0, \mathcal{D}^{-1}\mathcal{S}\mathcal{D}^{-1}\right),$$

where $\mathcal{D} = -E\left[X^2(i,j-1) + X^2(i-1,j) + X^2(i-1,j-1)\right],$

$\mathcal{S} = E_{\underline{\theta}}\left\{\left[X^2(i,j-1) + X^2(i-1,j) + X^2(i-1,j-1)\right]\left(1 + \left[\underline{x}(i,j)^{\otimes 2}\right]'\right)\right\}.$

Chapter 6

Bilinear spatio-temporal model

6.1 Introduction

For observations over various location, several kind of spatial models have been defined to describe the dependence over space. In situation where time is not fixed, then we have spatio-temporal data which are frequently used in many context to model situations as disease mapping, air-pollution monitoring, house price, satellite images of parts of the earth, election results for voting districts. They seem to provide "good" models for these situations. This type of data can be model in several ways. For instance, often it is assumed that the observations are Gaussian, therefor to model the dependence a covariance is often fitted (see, for example, Aline et al 2005). It is also generally supposed that the process is spatially stationary and the covariance has a particular structure and belongs to a certain family. In which case likelihood methods are often used to estimate the parameters (see, for example, Zhang (2004), June and Stein (2007)). An other primary interests in analyzing such data the study of spatial nonstationarity because many factors could cause the process to be spatially nonstationary (see, for example, Subba Rao (2007)). We note that many new spatio-temporal model classes have been developed in recent years, and this remains an active area of research (e.g., Cressie and Huang (1999); Gneiting (2002) ; Ma (2003); Stein (2005); Fuentes et al. (2008); Gregori et al. (2008); Zastavnyi and Porcu (2009)). Cressie and Huang (1999); Gneiting (2002); Stein (2005) present examples of non-separable stationary covariance functions for space-time processes. Cressie and Huang

(1999) propose classes of covariances that are obtained by inverting spectral densities of the form $h(\omega, v) = \rho(\omega, v)\, k(\omega)$, where ω denotes the spatial frequency and v the temporal one. Gneiting (2002) defines a class of covariances that provides a direct construction of valid covariance functions, without the need of inverting the spectral density. Stein (2005) considers classes of covariance functions that are smooth everywhere, except possibly at the origin. Fuentes et al. (2008) introduce a general and flexible parametric class of spatial temporal covariance models, that allows for lack of stationarity and separability by using a spectral representation of the process. Gregori et al. (2008) propose an easy-to-implement and interpretable spatio-temporal covariance models allowing for negative values or oscillations from positive to negative values. They show particular analytical examples that may be of interest in the biometrical context. Zas-tavnyi and Porcu (2009) characterize completely the Gneiting class of space-time covariance functions and give more relaxed conditions on the involved functions. They then show necessary conditions for the construction of compactly supported functions of the Gneiting type. For recent discussion of development in spatio-temporal models see Turkman et al (2010) and the references therein.

Many statisticians notice that if we fix a location, the observations at that location can be considered as a time series. This inspires them to define a spatio-temporal process in terms of its time dynamics. For instance, Suhasini Rao (2007) assumed that for every fixed location, the resulting time series has an AR(p) representation, where the innovations are samples from a spatial process. However, in many situations, linear models are found to be inappropriate to describe the underlying random mechanism. In the unidimensional case, the most famous nonlinear models which have attracted considerable attention in the literature are bilinear time series which have been investigated in many works (see, for example, Subba Rao (1981); Quinn (1982); Guégan and Pham (1989) and Bibi and Gautier (2010)). Extending the bilinear time series models by taking into account various locations, Jun and Stein (2007) and Dai and Billard (1998, 2003) propose a class of space-time bilinear models based on the ordering and weighting scheme to model spacial time series data which have certain nonlinear behaviors. This model was fitted in the spatial spread of monthly surveillance data for mumps over 1971-1988 in twelve states of the U.S.A. The limitation of the approach using an ordering and weithing scheme is that the weights

are determined before a model is identified; hence, this approach relies upon the model builder's judgement and experience. In an attempt to formulate suitable nonlinear but simple models to explain the evolution space-time series and Following Suhasini Subba Rao (2007) , we add to the literature above, a new clutter models which are location dependent bilinear processes, where for each fixed location, the observations at that location can be considered as a bilinear time series. The organization of this chapter is as follows : In section 2, we introduce some definitions and recall some results about space-time bilinear models. In section 3, we establish, under a minimal set of assumptions, the existence and time stationarity of $X_t(u)$. Our main result is given in section 4, where a central limit theorem is obtained. In section 5, we give some simple examples.

6.2 The model

As we noted in the introduction, our model is a location dependent bilinear process $\left\{X_t(u) : u \in [0,1]^2\right\}_t$ denoted $STBL_2(p,q,P,Q)$, where for each fixed location, the observations at that location can be considered as bilinear time series models, then $X_t(u)$ satisfies the representation

$$X_t(u) = \sum_{i=1}^{p} a_i(u) X_{t-i}(u) + \sum_{j=0}^{q} b_j(u) e_{t-j}(u) + \sum_{i=1}^{P}\sum_{j=1}^{Q} c_{ij}(u) X_{t-i}(u) e_{t-j}(u). \tag{6.2.1}$$

In (6.2.1) the innovations $e_t(u)$ are independent over time and spatially, strictly stationary and ergodic over time, with $E[e_t(u)] = 0$ and $var[e_t(u)] = 1$. $\{a_i(u);\ i = 1,\dots,p\}$, $\{b_i(u);\ i = 1,\dots,q\}$ and

$\{c_{ij}(u);\ i = 1,\dots,p \text{ and } j = 1,\dots,Q\}$ are nonparametric functions. The $STBL_2(p,q,P,Q)$ includes as special cases several classes of models which have been investigated in the literature. For instance, The location dependent autoregressive AR processes : these models can be deduced from (6.2.1) by setting $b_j(u) = 0$ and $c_{kl}(u) = 0$ over $[0,1]^2$ for all $1 \le j \le q$, $1 \le k \le P$ and $1 \le l \le Q$. They have been used by Gilleland and Nychka (2005) and Gelfand, Kim Siiamans and Banergee (2003) to fit ozone and house price data respectively.

Our first aim is to prove that, under some weak regularity conditions, there exists a measurable function $g : \mathbb{R}^{\infty} \to \mathbb{R}$ such that, for each fixed $u \in [0,1]^2$,
$X_t(u) = g(e_t(u), e_{t-1}(u), e_{t-2}(u)\ldots)$ all most surely for all $t \in \mathbb{Z}$. It's then easy to see that $\{X_t(u)\}$ is both strictly stationary and ergodic over time.
Without loss of generality we take $p = P$. So model (6.2.1) can be written in the eqivalent matrix form

$$\underline{X}_t(u) = C(u)\,\underline{e}_t(u) + A(u)\,\underline{X}_{t-1}(u) + \sum_{j=1}^{Q} B_j(u)\, e_{t-j}(u)\underline{X}_{t-1}(u), \tag{6.2.2}$$

where

$$\underline{X}_t(u) = \begin{bmatrix} X_t(u) \\ X_{t-1}(u) \\ \vdots \\ X_{t-p+1}(u) \end{bmatrix}_{p\times 1}, \qquad \underline{e}_t(u) = \begin{bmatrix} e_t(u) \\ \vdots \\ e_{t-q}(u) \end{bmatrix}_{(q+1)\times 1}$$

$$C(u) = \begin{bmatrix} 1 & b_1(u) & \ldots & b_q(u) \\ 0 & \ldots & & 0 \\ \vdots & & & \vdots \\ 0 & \ldots & & 0 \end{bmatrix}_{p\times(q+1)}, \; A(u) = \begin{bmatrix} a_1(u) & a_2(u) & \ldots & a_p(u) \\ 1 & 0 & \ldots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \ldots & 1 & 0 \end{bmatrix}_{p\times p}$$

$$B_j(u) = \begin{bmatrix} c_{1j}(u) & c_{2j}(u) & \ldots & c_{pj} \\ 0 & 0 & \ldots & 0 \\ \vdots & & & \vdots \\ 0 & \ldots & 0 & 0 \end{bmatrix}_{p\times p}.$$

Let $\|.\|_{\infty}$ denote the sup-norm of a vector. Suppose A a $p \times p$ matrix, A is said to be *spectral radius diagonable* if it is similar to a matrix of the form

$$\begin{bmatrix} G & \mathbf{0} \\ \mathbf{0} & H \end{bmatrix}$$

where G is diagonal with all diagonal elements equal to $\rho(A)$ in modulus and $\rho(H) < \rho(G)$. In

particular, if A has all distinct eigenvalues on the Jordan representation of A is diagonal, then A is spectral radius diagonable. For any $m \times m$ matrix $G = [g_{ij}]$ we define

$$\|G\|_\infty = \max_{1 \le i \le m} \sum_{j=1}^{m} |g_{ij}| .$$

We now state the following lemma, the proof of which is available in Ortega (1987)(see Exercise 5.3.7).

Lemma 6.1 *Let A be spectral radius diagonable. Then there exists a $p \times p$ nonsingular matrix P such that $\|P^{-1}AP\|_\infty = \rho_0$ where $\rho_0 = \rho(A)$.*

Define now $C_d = C\left([0,1]^2, \mathbb{R}^d\right)$ to be the collection of all continuous functions $f : [0,1]^2 \to \mathbb{R}^d$ (it is metrized by $d_{C_d}(f,g) = \sup_{t \in [0,1]^2} |f(t) - g(t)|$, f, $g \in C_d$). Recall that a functional is a real-valued function on C_d, and a functional Λ is continuous if for all f, $f_n \in C_d$, $\lim_{n\to\infty} d_{C_d}(f_n, f) = 0$ implies $\lim_{n\to\infty} \Lambda(f_n) = \Lambda(f)$.
For such functional, we quote from Kholshnevisan (2002) (see Theorem 3.4.1) the following theorem.

Theorem 6.1 *Suppose that $X_n \Rightarrow X_\infty$, as elements of C_d. Then for any continuous functional Λ on C_d, $\Lambda(X_n)$ converges in distribution to $\Lambda(X_\infty)$.*

We also present the following result which is available in Brockwell and Davis (1991) (Problem 6.16 p 216).

Lemma 6.2 *Suppose that $X_n \sim N(\mu_n, \upsilon_n)$ where $\mu_n \to \mu$, $\upsilon_n \to \upsilon$ and $0 < \upsilon < \infty$, so $X_n \Rightarrow X$, where $X \sim N(\mu, \upsilon)$.*

6.3 Stationarity and causality of spatial bilinear processes

In the case of standard bilinear models BL(p, q, P, Q), existence and stationarity problems have been partially solved by Pham and Tran (1981), Quinn (1982), Liu and Brockwell (1988) and Chanda (1989). This later has established such conditions substantially weaker than those proposed by the earlier researchers. The goal of the present section is to extend Chanda results to the space-time bilinear model defined by equation (6.2.1). Chanda (1989) considers the model

$$X_t = C\epsilon_t + AX_{t-1} + \sum_{j=1}^{Q} B_j e_{t-j} X_{t-1},$$

where $\{X_t,\ t = 0,\ \pm 1, \ldots\}$ is a p-variate process. $C\,(p \times (s+1))$, $A\,(p \times p)$, $B_j\,(p \times p)$ $(1 \leq j \leq q)$ are arbitrary matrices of constants. $\epsilon_t^T = (e_t, \ldots, e_{t-q+1})$ and $\{e_t,\ t = 0,\ \pm 1, \ldots\}$ is a strictly stationary ergodic sequence of random variables. He investigates a set of minimal regularity conditions (on C, A, B_j and $\{e_t\}$) under which, the existence and causality of X_t is established.

Notice that, since the coefficients of the $STBL_2\,(p, q, P, Q)$ vary with u, the process is not spatially stationary. Also, there is no separability between the space and the time components. That is, the joint space-time covariance function can not be factorized into a space covariance multiplied by a time covariance. Fortunately, our aim is to obtain a stationary process at each location u, because we only need time-stationarity for checking the CLT.

Remark 3.1 *It is easy to show that the location dependent* $STBL_2\,(p, q, P, Q)$ *process is only space-time spatially stationary and separable if* $a_i\,(.)$, $b_j\,(.)$, $c_{kl}\,(.)$ *and* $\sigma\,(.)$ *are constant over* $[0, 1]^2$, *for all* $1 \leq i \leq p$, $1 \leq j \leq q$, $1 \leq k \leq P$ *and* $1 \leq l \leq Q$.

Now, we shall consider a vector process described by equation (6.2.2), where $A\,(u)$, $C\,(u)$ and $B_j\,(u)$, $(1 \leq j \leq Q)$ are arbitrary matrices not necessarily of the type given in (6.2.2).

Under these assumptions, note that (6.2.2) can be written as

$$\underline{X}_t\,(u) = \underline{X}_{t,m}\,(u) + \underline{X}^*_{t,m}\,(u)$$

for every $m \geq 1$ where

$$\underline{X}_{t,m}(u) = C(u)\,\underline{e}_t(u) + \sum_{r=1}^{m}\prod_{k=0}^{r-1}\left(A(u) + \sum_{j=1}^{Q} B_j(u)\,e_{t-j-k}(u)\right) C(u)\,\underline{e}_{t-r}(u)$$

and

$$\underline{X}^*_{t,m}(u) = \prod_{k=0}^{m}\left(A(u) + \sum_{j=1}^{Q} B_j(u)\,e_{t-j-k}(u)\right)\underline{X}_{t-m-1}(u).$$

Let's prove now the following theorem which guarantees time stationarity at each point in the spatial domain.

Theorem 6.2 *Let* $\left\{\underline{X}_t(u) : u \in [0,1]^2\right\}_t$ *be as defined in* (6.2.2), $C(u)$, $A(u)$ *and* $B_j(u)$ *are arbitrary and not necessary of the type described in* (6.2.2). *Assume that* $A(u)$ *is spectral radius diagonable matrix with* $\|P^{-1}(u)\,A(u)\,P(u)\|_\infty = \rho_0(u)$, *set* $\delta_j(u) = \|P^{-1}(u)\,B_j(u)\,P(u)\|_\infty$. *Assume further that*

$$\sup_u E\log\left(\rho_0(u) + \sum_{j=1}^{Q}\delta_j(u)\,|e_j(u)|\right) < 0.$$

Then for all fixed u, *there exists a unique strictly stationary (over time) process* $\underline{X}_t(u)$, *satisfying* (6.2.2), *given by*

$$\underline{X}_t(u) = C(u)\,\underline{e}_t(u) + \sum_{r=1}^{\infty}\prod_{k=0}^{r-1}\left(A(u) + \sum_{j=1}^{Q} B_j(u)\,e_{t-j-k}(u)\right) C(u)\,\underline{e}_{t-r}(u), \tag{6.3.1}$$

with the infinite series being almost surely convergent.

Proof *Suppose there exists a strictly stationary process* $\underline{X}_t(u)$ *satisfying* (6.2.2), *then*

$$\underline{X}_t(u) = C\underline{e}_t(u) + A\underline{X}_{t-1}(u) + \sum_{j=1}^{Q} B_j(u)\,e_{t-j}(u)\underline{X}_{t-1}(u),$$

by repeating this formula

$$\begin{aligned}\underline{X}_t(u) &= C(u)\,\underline{e}_t(u) + \sum_{r=1}^{m}\prod_{k=0}^{r-1}\left(A(u) + \sum_{j=1}^{Q} B_j(u)\, e_{t-j-k}(u)\right) C(u)\,\underline{e}_{t-r}(u) \\ &\quad + \prod_{k=0}^{m}\left(A(u) + \sum_{j=1}^{Q} B_j(u)\, e_{t-j-k}(u)\right)\underline{X}_{t-m-1}(u),\end{aligned}$$

for all $m \geq 1$.

Put $P_{t,n}(u) = \prod_{k=0}^{n}\left(A(u) + \sum_{j=1}^{Q} B_j(u)\, e_{t-j-k}(u)\right)$, *and let* γ *be an arbitrary* $p \times 1$ *vector of real numbers. Then*

$$|\gamma' P_{t,n}(u)| \leq M \|P(u)\|_\infty \prod_{k=0}^{n}\left\|P^{-1}(u)\left(A(u) + \sum_{j=1}^{Q} B_j(u)\, e_{t-j-k}(u)\right)P(u)\right\|_\infty \left\|P^{-1}(u)\right\|_\infty,$$

where $M = \|\gamma'\|_\infty$ *which is a finite positive constant independent of* n. *Hence*

$$|\gamma' P_{t,n}(u)| \leq M \prod_{k=0}^{n}\left(\rho_0(u) + \sum_{j=1}^{Q}\delta_j(u)\,|e_{t-j-k}(u)|\right).$$

Let us consider

$$\lim \frac{1}{n}\log|\gamma' P_{t,n}(u)| \leq \lim \frac{1}{n}\log M + \lim \frac{1}{n}\sum_{k=0}^{n}\log\left(\rho_0(u) + \sum_{j=1}^{Q}\delta_j(u)\,|e_{t-j-k}(u)|\right),$$

by the strong law of large numbers

$$\begin{aligned}\lim \frac{1}{n}\log|\gamma' P(t,n)| &\leq E\log\left(\rho_0(u) + \sum_{j=1}^{Q}\delta_j(u)\,|e_{t-j}(u)|\right) \\ &\leq \sup_u E\log\left(\rho_0(u) + \sum_{j=1}^{Q}\delta_j(u)\,|e_{t-j}(u)|\right) \qquad a.s.\end{aligned}$$

which implies $P_{t,n}(u) \to 0$ *a.s. as* $n \to \infty$.

By strict stationarity of $\underline{X}_t(u)$, $\underline{X}_{t-m-1}(u)$ *have the same distribution for all* $m \geq 1$, *this*

fact together with $P_{t,n}(u) \to 0$ *a.s., imply* $\prod_{k=0}^{m}\left(A(u)+\sum_{j=1}^{Q}B_j(u)\,e_{t-j-k}(u)\right)\underline{X}_{t-m-1}(u) \to 0$ *in probability as* $m \to \infty$. *Thus*

$$\underline{X}_t(u) = C(u)\,\underline{e}_t(u) + \sum_{r=1}^{\infty}\prod_{k=0}^{r-1}\left(A(u)+\sum_{j=1}^{Q}B_j(u)\,e_{t-j-k}(u)\right)C(u)\,\underline{e}_{t-r}(u).$$

Conversely, suppose that $\underline{X}_t(u)$ *satisfies* (6.3.1). *Let's first prove that the series of random vectors*

$$\sum_{r=1}^{\infty}\prod_{k=0}^{r-1}\left(A(u)+\sum_{j=1}^{Q}B_j(u)\,e_{t-j-k}(u)\right)C(u)\,\underline{e}_{t-r}(u),$$

converges absolutely a.s. for every fixed $t \in \mathbb{Z}$ *and* $u \in [0,1]^2$. *We can establish this if we can show that* $\sum_{r\geq 1}W_{rt}(u)$ *converges absolutely a.s. where*

$$W_{rt}(u) = \gamma'\prod_{k=0}^{r-1}\left(A(u)+\sum_{j=1}^{Q}B_j(u)\,e_{t-j-k}(u)\right)C(u)\,\underline{e}_{t-r}(u).$$

Note here that

$$|W_{rt}(u)| \leq M\prod_{k=0}^{r-1}\left\|P^{-1}(u)\left(A(u)+\sum_{j=1}^{Q}B_j(u)\,e_{t-j-k}(u)\right)P(u)\right\|_{\infty}\left\|\underline{e}_{t-r}(u)\right\|_{\infty}, \tag{6.3.2}$$

where for any $m\times 1$ *vector* $\alpha' = (\alpha_1,\ldots,\alpha_m)$, $\|\alpha\|_{\infty} = \max_{1\leq i\leq m}|\alpha_i|$.
From (6.3.2) *we have*

$$\frac{1}{r}\log|W_{rt}(u)| \leq \frac{1}{r}\log M + \frac{1}{r}\sum_{k=0}^{r}\log\left(\rho_0(u)+\sum_{j=1}^{Q}\delta_j(u)\,|e_{t-j-k}(u)|\right) + \frac{1}{r}\sum_{v=0}^{q}\log|e_{t-r-v}(u)|, \tag{6.3.3}$$

where the right-hand of (6.3.3) *tends to* $E\log\left(\rho_0(u)+\sum_{j=1}^{Q}\delta_j(u)\,|e_{t-j-k}(u)|\right)$ *a.s. as* $r\to\infty$, *it follows immediately that, for every* t *and* u, $limsup|W_{rt}(u)|^{\frac{1}{r}} < 1$ *a.s., which imply that* $\sum_{r\geq 1}W_{rt}(u)$ *converges absolutely a.s., then* $\underline{X}_t(u)$ *is strictly stationary, ergodic and satisfies* (6.2.2). ■

6.4 Central Limit Theorem for spatial bilinear process

In this section we shall investigate the asymptotic normality of the sample mean derived from a multivariate space time bilinear process $\underline{X}_t(u)$ defined in 6.2.2. Our result provides a useful tool for statistical inference for this process. We shall also focus our interest to the first component (call it $X_t(u)$) of $\underline{X}_t(u)$ defined in (6.2.2) with arbitrary $A(u)$, $C(u)$ and $B_j(u)$.

We shall assume throughout that

(i) $e_t(u)$ is an i.i.d. sequence with $E(e_t(u)) = 0$, $E\left(e_t(u)^2\right) = 1$, and $E(e_t(u)^r) = \mu_r(u)$ are finite for $r = 3, 4$.

(ii) $\{a_i(u);\ i = 1, \ldots, p\}$, $\{b_i(u);\ i = 1, \ldots, q\}$ and $\{c_{ij}(u);\ i = 1, \ldots, p$ and $j = 1, \ldots, Q\}$ are continuous functions.

(iii) $\sup_u \rho(A(u) \otimes A(u)) = \rho_1 < 1$ and $\sup_u \rho(G(u)) = \rho_2 < 1$, where $G(u) = A(u) \otimes A(u) + \sum_{j=1}^{Q} B_j(u) \otimes B_j(u)$.

First put

$$S_{T,n} = \sum_{t=1}^{n}\sum_{s=1}^{T} \frac{X_t(u_s) - \mu}{(nT)^{\frac{1}{2}}}.$$

Assumptions (i,iii) allow us to apply Chanda CLT associated with bilinear time series, to the sum $\Sigma_n(u) = \sum_{t=1}^{n} \frac{X_t(u)-\mu}{(n)^{\frac{1}{2}}}$. We use assumption (ii) to have $\Sigma_n(.) \in C_1$.

The following theorem establishes the asymptotic normality of the quantity $\sum_{t=1}^{n}\sum_{s=1}^{T} \frac{X_t(u_s)-\mu}{(nT)^{\frac{1}{2}}}$ which correspond to the sample mean derived from the space time bilinear process (6.2.1).

Theorem 6.3 *Let*

$$S_{T,n} = \sum_{t=1}^{n}\sum_{s=1}^{T} \frac{X_t(u_s) - \mu}{(nT)^{\frac{1}{2}}},$$

and assume that the conditions i), ii) and iii) hold. Then

$$\mathcal{L}(S_{T,n}) \to \mathcal{N}\left(0, \eta^2\right),$$

as $n, T \to \infty$, *where*

$$\eta^2 = \lim_{T\to\infty} \frac{1}{T} \sum_{s=1}^{T} \left[\lim_{m\to\infty} \left\{ EX_{1,m}^2 (u_s) + 2 \sum_{\upsilon=1}^{m} \left| X_{1,m} (u_s) X_{1+\upsilon,m} (u_s) \right| \right\} \right].$$

Proof We first consider the sum $\Sigma_n (u) = \sum_{t=1}^{n} \frac{X_t(u)-\mu}{(n)^{\frac{1}{2}}}$. Using the CLT given by Chanda (1989) (Theorem 4.1) for general bilinear time series, we can claim that for each fixed $u \in [0,1]^2$, $\mathcal{L}(\Sigma_n (u_s)) \to \mathcal{N}(0, \sigma^2 (u_s))$, as $n \to \infty$, where

$$\sigma^2 (u_s) = \lim_{m\to\infty} \left\{ EX_{1,m}^2 (u_s) + 2 \sum_{\upsilon=1}^{m} \left| X_{1,m} (u_s) X_{1+\upsilon,m} (u_s) \right| \right\}.$$

We consider a continuous function $\theta : \left\{[0,1]^2\right\}^T \to \mathbb{R}$, for any fixed $u_1, \ldots, u_T \in [0,1]^2$ and all $f \in C_1$, define $\Lambda(f) = \theta(f(u_1), \ldots, f(u_T)) = \frac{1}{T^{\frac{1}{2}}} \sum_{j=1}^{T} f(u_j)$, $\Lambda(f)$ is a continuous functional (see example 2, p199 of Kholshnevisan (2002)). By condition ii) and (6.3.1) (there exists a measurable function $h : \mathbb{R}^\infty \to \mathbb{R}$, $\Sigma_n (u) = \sum_{t=1}^{n} h(e_t (u), e_{t-1} (u), \ldots)$) we can claim that $\Sigma_n (.) \in C_1$. As $\Sigma_n (u_s) \Rightarrow \Sigma_\infty (u_s)$ for $s = 1, \ldots, T$, where $\Sigma_\infty (u_s) \sim \mathcal{N}(0, \sigma^2 (s))$, by Theorem 6.1 $\Lambda(\Sigma_n (u_s)) \Rightarrow \Lambda(\Sigma_\infty (u_s))$ hence $\mathcal{L}(S_{T,n}) \to \mathcal{L}\left(\frac{1}{T^{\frac{1}{2}}} \sum_{j=1}^{T} \Sigma_\infty (u_j) \right)$. $e_t (u)$ are independent over time and spatially, hence $\Sigma_\infty (u_1)$, $\Sigma_\infty (u_2), \ldots, \Sigma_\infty (u_T)$ are independent, as $\Sigma_\infty (u_j) \sim \mathcal{N}(0, \sigma^2 (u_j))$ for all $j = 1, \ldots, T$, so $\frac{1}{T^{\frac{1}{2}}} \sum_{j=1}^{T} \Sigma_\infty (u_j) \sim \mathcal{N}\left(0, \frac{1}{T} \left[\sum_{j=1}^{T} \sigma^2 (u_j) \right]\right)$. Also it is easy to see that $\eta_T^2 = \frac{1}{T} \left[\sum_{j=1}^{T} \sigma^2 (u_j) \right]$ tends to a finite constant $\eta^2 = \lim_{T\to\infty} \frac{1}{T} \sum_{s=1}^{T} \left[\lim_{m\to\infty} \left\{ EX_{1,m}^2 (u_s) + 2 \sum_{\upsilon=1}^{m} \left| X_{1,m} (u_s) X_{1+\upsilon,m} (u_s) \right| \right\} \right]$, consequently by Lemma 6.2 $\mathcal{L}(S_{T,n}) \to \mathcal{N}(0, \eta^2)$. ■

6.5 Numerical Example

Consider the $STBL_2 (1, 0, 1, 1)$ process

$$X_t(u) = a(u) X_{t-1}(u) + c(u) X_{t-1}(u) e_{t-1} (u) + e_t (u),$$

where $a : [0,1]^2 \to [-1,1]$, $c : [0,1]^2 \to [-1,1]$ with for $u = (x,y)$

$$\begin{aligned} a(x,y) &= a_1 \sin x \cos y, \\ c(x,y) &= a_2 (5-x) \cos y \end{aligned}$$

and the innovations $\{e_t(u)\}$ are independent over time and spatially, strictly stationary and ergodic over time, with $E[e_t(u)] = 0$ and $var[e_t(u)] = 1$. In this case $\underline{X}_t(u) = X_t(u)$, $\underline{e}_t(u) = e_t(u)$, $A(u) = a(u)$, $C(u) = 1$ and $B_1(u) = c(u)$. It is clear that $A(u)$ is a spectral radius diagonable 1×1 matrix with $\rho_0(u) = |a(u)|$ and $\delta_1(u) = |c(u)|$. The sufficient condition for the existence of a strictly stationary (over time) process $X_t(u)$ is

$$\sup_u E \log \left(|a(u)| + |c(u)| \, |e_j(u)| \right) < 0,$$

which becomes

$$|a_1| + 5 |a_2| < 1.$$

The solution is given by

$$X_t(u) = e_t(u) + \sum_{r=1}^{\infty} \prod_{k=0}^{r-1} \left[a(u) + c(u) \, e_{t-1-k}(u) \right] e_{t-r}(u).$$

Bibliography

[1] Aline, A. Nobre, Alexandra, M. Schmidt, and Hedibert, F. Lopes. (2005) Spatio-temporal models for mapping the incidence of malaria in Pará. *Environmetrics*. **16**, 291-304.

[2] Amirmazlaghani, M., Amindavar, H. (2010) Image denoising using two-dimensional GARCH model. *Systems, Signals and image processing*. 397-400.

[3] Andel, J. (1976) Autoregressive series with random parameters. *Math. Operations for sch. Statist*. **7**, 735–41.

[4] Andel, J. (1989) On nonlinear models for time series. *Statistics.* **4**, 615-632.

[5] Anders, H. (1998) *A History of Mathematical Statistics.* New York: Wiley.

[6] Anselin, L. (1988) *Spatial Econometrics: Methods and Models.* Dordrecht: Kluwer Academic Publishers.

[7] Anselin, L. and Florax, R. (1995) *New Directions in Spatial Econometrics*. London: Springer.

[8] Anselin, L. and Bera, A. (1998) Spatial dependence in linear regression models with an introduction to spatial econometrics. In A. Ullah and D. Giles (Eds.). *Handbook of Applied Economic Statistics*, 237-289.

[9] Barlett, M. S. (1971) Physical nearest neighbor models and nonlinear time series. *J. Appl. Prob.* **8**, 222-232.

[10] Barlett, M. S. (1975) *The Statistical Analysis of Spatial Pattern.* Chapman and Hall, London.

[11] Basawa, I. V. and Hwang, S. Y. (1997) The local asymptotic normality of a first order bilinear time series model. *Statistics and prob. letters.* **34**(2), 165-170.

[12] Basu, S., Reinsel, G.C. (1992) A note on properties of spatial Yule–Walker estimators. *J. Statist. Comput. Simulation* **41**, 243–255.

[13] Basu, S., Reinsel, G.C. (1993) Properties of the spatial unilateral first-order ARMA model. *Adv. Appl. Probab.* **25**, 631–648.

[14] Bennett, J. and Khotanzad, A. (1999) Maximum likelihood estimation methods for multispectral autoregressive image models. *IEEE Trans. Pattern Anal. mach. Intell.* **21**, 537-543.

[15] Bezag, J. E. (1972) On the correlation structure of some two dimensional stationary processes. *Biometrika.* **59,** 43-48**.**

[16] Besag, J. (1974) Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B* **36**, 192–236

[17] Besag, J. and Higdon, D. (1999) Bayesian analysis of agricultural field experiments. *J. Royal. Statis. Soc. Series B.* **61**(4), 691-746.

[18] Biau, G. and Cadre, B. (2004) Non parametric spatial prediction. *Statistical Inference for Stochastic Processes.* **7**, 327-349.

[19] Bibi, A. and Gautier, A. (2010) Consistent and asymptotically normal estimators for periodic bilinear models. *Bull. Korean Math. Soc.* **47**(5), 889-905.

[20] Bickel, P. J. and Levina, E. (2004) Some theory for Fisher's linear discriminant function, "naive Bayes", and some alternatives when there are many more variables than observations. *Bernoulli*, **10**(6):989–1010.

[21] Bougerol, P. and Picard, N. (1992) Strict stationarity of generalized autoregressive processes. *The Annals of Probability.* **20**, 1714-1730.

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

[23] Brandt, A. (1986) The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Adv. Appl. Prob.* **18** (1), 211-220.

[24] Brockwell, P. and Davis, R. A. (1991) *Time series. Theory and methods (second edition).* New York, Springer.

[25] Bronars, S. and D. Jansen. (1987) The geographic distribution of unemployment rates in the US. *Journal of Econometrics.* **36**, 251-79.

[26] Chanda, C. (1989) Stationarity and Central Limit Theorem associated with bilinear time series models. *J. Time Ser. Analysis.* **12**(4), 301-313.

[27] Cliff, A. D. and Haggett, P. Ord, J. K. Basset, K. A. and Davis, R. B. (1975) *Elements of statistical structure :A quantitative approach.* New York :Cambridge University Press.

[28] Cliff, A. D. and Ord, J. (1973) *Spatial Autocorrelation.* Pion, London.

[29] Cliff, A. D. and Ord, J. (1981) *Spatial Processes, Models and Applications.* Pion, London.

[30] Chils, J. P. and Delfiner, P. (1999) *Geostatistics: Modeling Spatial Uncertainty.* Wiley, New York.

[31] Choi, B. S. (1992) *ARMA models Identification.* Springer, New York.

[32] Choi, B. S. (1997) A recursive algorithm for solving the spatial Yule-Walker equations of causal spatial AR models. *Statist. Probab. Lett.* **33**(3), 241-251.

[33] Choi, B. S. (2000) On the asymptotic distributions of mean, autocovariance, autocorrelation, crosscovariance and impulse response estimators of a stationary multidimensional random field. *Commun. Statist. Theory Meth.* **29**, 1703-1724.

[34] Conlisk, J. (1974) Stability in a Random Coefficient Model. *International Economic Review.* **15**, 529-533.

[35] Conlisk, J. (1976) A Further Note on Stability in a Random Coefficient Model. *International Economic Review.* **17**, 759-764.

[36] Cressie, N. A. C. (1993) *Statistics for Spatial Data.* John Wiley & Sons, New York.

[37] Cressie, N. and Huang, N. C. (1999) Classes of nonseparable, spatio temporal stationary covariance functions. *Journal of the American Statistical Association.* **94**, 1330–1340.

[38] Dai, Y. and Billard, L. (1998) A space time bilinear model and its identification. *J. Time Ser. Analysis.* **19**, 657-79.

[39] Dai, Y. and Billard, L. (2003) Maximum likelihood estimation in space time bilinear models. *J. Time Ser. Analysis.* **24**, 26-44.

[40] Davis, J. C. (1973) *Statistics and Data Analysis in Geology.* Wiley, New York.

[41] Dedecker, J. (1998) A central limit theorem for stationary random fields. *Probab, Theory Relat. Fields.* **110**, 397-426.

[42] Delfiner, P. and Delhomme, J. P. (1975) Optimum interpolation by Kriging. *In Nato Advanced Study Institute on Display and Analysis of Spatial Data*, 96-115.

[43] Engle R. F. and Bollerslev T. (1986) Modelling the Persistence of Conditional Variances. *Econometric Reviews.* **5**,1-50.

[44] Etchison, T. Pantula, S. G. and Browni, C. (1994) Partial autocorrelation function for spatial processes. *Statist. Probab. Lett.* **21**(1), 9-19.

[45] Fisher, W. D. (1971) Econometric estimation with spatial dependence. *Regional and Urban Economics.* **1**, 40-64.

[46] Fornasini, E. and Marchesini, G. (1978) Doubly indexed dynamical systems: state space models and structural properties. *math. syst. theory.* **12,** 59-72.

[47] Fotheringham, A. S. Brunsdon, C. Charlton, M. (2002) *Geographically weighted regression : the analysis of spatially varying relation ships.* Wiley, West Sussex.

[48] Fuentes, M. Chen, L. and Davis, J. M. (2008) A class of nonseparable and nonstationary spatio temporal covariance functions. *Environmetrics.* **19**, 487-507.

[49] Gaetan, C. and Guyon, X. (2008) *Modélisation et statistique spatiale*. Springer.

[50] Gao, J. Lu, Z. and Tjostheim, D. (2006) Estimation in semiparametric spatial regression. *Annals of Statistics*. **34**, 1395-1435.

[51] Gelfand, AE. Kim, H. Sirmans CF. and Banerjee, S. (2003) Spatial modeling with spatially varying coefficient processes. *J Am Stat Assoc*, **98**, 387–396.

[52] Geman, D., Geman, S. (1984) Stochastic relaxation, Gibbs distributions and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721–741.

[53] Gilleland, E, & Nychka, D, (2005) Statistical models for the monitoring and regulating ground-level ozone. *Environmetrics*. **16**, 535-546.

[54] Gneiting, T. (2002) Nonseparable, stationary covariance functions for space-time data. *Journal of the American Statistical Association*. **97**, 590-600.

[55] Granger, C. W. and Anderson, A. P. (1978) *An introduction to bilinear time series models.* Göttingen. Vandenhœck and Ruprecht.

[56] Granger, C. W. J. and Teräsvirta, T. (1993) *Modelling Nonlinear Economic Relationships*. Oxford University Press, Oxford.

[57] Gregori, P. Porcu, E. Mateu, J. and Sasvari, Z. (2008) On potentially negative space time covariances obtained as sum of products of marginal ones. *Annals of the Institute of Statistical Mathematics*. **60**, 865-882.

[58] Griffith, D.A. (1992) *What is Spatial Autocorrelation? Reflections on the Past 25 Years of Spatial Statistics*. L'Espace Geographique, **21**(3), 265-280.

[59] Guegan, D. (1987) Different representations for bilinear models. *J. Time Series Anal.* **8**(4), 389-408.

[60] Guegan, D. and Pham, D. T. (1989) A note on the estimation of the parameters of the diagonal bilinear model by the method of least squares. *Scand J Statist.* **16**, 129-136.

[61] Guo, J. H. and Billard, L. (1998) Some inference results for causal autoregressive processes on a plan. *J. Time Ser. Anal.* **19**(6), 681-691.

[62] Guyon, X. (1982) Parameter estimation for a stationary process on a d-dimensional lattice. *Biometrika.* **69**, 95-105.

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

[64] Ha, E. and Newton. H. J. (1993) The bias of estimators of causal spatial autoregressive processes. *Biometrika.* **80**(1), 242-245.

[65] Hai-Bin, W. (2005) Parameter estimation and subset selection for separable lower triangular bilinear models. *J. Time Ser. Anal.* **26**(5), 743-757.

[66] Haining, R. P. (1978) Estimation spatial-interaction models. *Env. Planning A.* **10**, 305-320.

[67] Haining, R. (1990) *Spatial data analysis in the social and environmental sciences*. Cam bridge : Cambridge University Press.

[68] Huang, D. W. (1992) Central limit theorem for two-parameter martingale differences with application to stationary random fields. *Sci. China Ser. A.* **35**(4), 413-425.

[69] Huang, D. W., and V. V. Anh. (1992) Estimation of spatial ARMA models. *Austral. J. Statist.* **34**(3), 513-530.

[70] Jain, A. K. (1981) Advances in mathematical models for image processing. *Proc. IEEE.* ***69***, 502-528.

[71] Jun, M. and Stein, ML. (2007) An approach to producing space-time covariance functions on spheres. *Technometrics.* **49**(4), 468-479.

[72] Kashyap, R. and Eom, K. (1988) Robust images techniques with an image restoration application. *IEEE Trans. Acoust. Speech Signal Process.* **36**, 1313-1325.

[73] Kelejian, H. H. and Prucha, I. (1999) A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model. *International Economic Review.* **40**, 509-533.

[74] Kendall, M. G. (1953) The analysis of economic time series. Part I: Prices. *J. Roy. Statist. Soc, A.* **96**, 11-25.

[75] Kershaw, K. A. (1964) *Quantitative and Dynamic Ecology.* Arnold, London.

[76] Kharfouchi S., and Kimouche K. (2011) Central limit theorem associated with bilinear random fields. *Random Oper. Stoch. Equ.* **19**(2), 157-163.

[77] Kharfouchi S., and Mehri H., (2012). Conditions for the existence of stationary and causal space-time bilinear model and Central Limit Theorem. *Int. J. Mathematics in Operational Research,* **4**(6), 638-650.

[78] Kholshnevisan, D. (2002) *Multiparameter resources: An introduction to random fields.* Springer.

[79] Kristensen, D. (2009) On stationarity and ergodicity of the bilinear model with applications to GARCH models. *J. Time Ser. Anal.* **30**, 125-144.

[80] Le Cam, L. (1986) *Asymptotic Methods in Statistical Decision Theory.* Springer-Verlag.

[81] Liu, J. and Brockwell, P. J. (1988) On the general bilinear time series models. *J. Appl Prob.* **25**, 553-564.

[82] Liu, J. (1989) On the existence of a general multiple bilinear time series. *J. Time Ser. Anal.* **10**(4), 341-55.

[83] Lu, Z., Tjostheim, D. and Yao, Q. (2007) Adaptive Varying-Coefficient Linear Models for Stochastic Processes: Asymptotic Theory. *Statistica Sinica.* **17**, 177-197.

[84] Ma, C. (2003) Families of spatio-temporal stationary covariance models' *Journal of Statistical Planning and Inference*, **116**, 489–501.

[85] Major, P. (1981) *Multiple Vienner-Itô integrals.* Lecture Notes in mathematics. **849**, Springer .

[86] Matheron, G. (1976) A simple substitute for conditional expectation: the disjunctive kriging. *Proceedings, NATO ASI: Advanced Geostatistics in the Mining Industry* (Rome,Oct.1975), 221-236.

[87] McMillen, P. (2003) Neighborhood house price indexes in Chicago: a Fourier repeat sales approach. *Journal of Economic Geography, Oxford University Press*, **3**(1), 57-73.

[88] Moore, H. (1988) Incomplete Contracts and Renegotiation. *Econometrica*, **56**(4), 755-785.

[89] Nicholls, D. F. and Quinn, B. G. (1982) *Random coefficient autoregressive models: An introduction.* Lecture notes in statist. Springer, Berlin.

[90] Noiboar, A. Cohen, I. (2005) *Two-Dimensional GARCH Model with Application to Anomaly Detection.* Proceedings of ICASSP.

[91] Ortega, J. M. (1987) *Matrix Theory: A Second Course.* New York: Plenum.

[92] Park, S and Heo, Y. (2009) Seasonal Spatial-temporal Model for Rainfall Data of South Korea. *Journal of Applied Sciences Research*, **5**(5), 565-572.

[93] Pede, O., Florax, J.G.M. and Holt, T. (2008) Modeling nonlinear Spatial Dynamics: A Family of Spatial STAR Models and an Application to U.S. Economic Growth. 2008 Annual Meeting, July 27-29, 2008, Orlando, Florida 6518, American Agricultural Economics Association.

[94] Pfeifer, P. E. and Deutsch, S J. (1980a) A three-stage iterative procedure for space-time modeling. *Technometrics.* **22**, 35-47.

[95] Pfeifer, P. E. and Deutsch, S J. (1980b) Identification and interpretation of first order space-time ARMA models. *Technometrics.* **22**, 397-408.

[96] Pham, D. T. and Tran, LT. (1981) On the first order bilinear time series model. *J. Appl. Prob.* **18**, 617-27.

[97] Pham, D. and Tuan, D. (1986) The mixing property of bilinear and generalized random coefficient autoregressive models. *Stoch. Processes and their App.* ***23****, 291-300.*

[98] Quinn, B. G. (1982) Stationarity and invertibility of simple bilinear models. *Stoch. Proc. Appl.* **12**, 225-30.

[99] Ripley, B.D. (1981) *Spatial Statistics*. John Wiley & Sons, New York.

[100] Rivoirard, J. (1994) *Introduction to Disjunctive Kriging and nonlinear Geostatistics*. Clarendon Press, Oxford.

[101] Russel, E. J. (1966) *A History of Agricultural Science in Great Britain: 1620-1954.* London: G. Allen and Unwin.

[102] Stein, M. (2005) Space-time covariance functions. *Journal of the American Statistical Association*. **100**, 310–321.

[103] Stensholt, B. K. and Tjostheim, D. (1987) Multiple bilinear time series models. *J. Time Ser. Anal.* **8**, 221-23.

[104] Stetter, H. J. (2004) *Numerical polynomial algebra*. Siam.

[105] Storvik, G. Frigessi, A. and Hirst, D. (2002) Space Time Gaussian fields and their autoregressive representation. *Statistical Modelling*. **2**, 139-61.

[106] Student (1914). "The Elimination of Spurious Correlation Due to Position in Time or Space." *Biometrika*, **10**, 179-80.

[107] Subba Rao, T. (1981) On the theory of bilinear time series model. *J. R. Statist. Soc., Ser. B.* **43**, 244-255.

[108] Subba Rao, T. and Gabr, M. M. (1984) *An Introduction to Bispectral Analysis and Bilinear Time Series Models*. Lecture Notes in Statistics **24**, Springer-Verlag, New York.

[109] Subba Rao, T. and Terdik, G. (2003) On the theory of discrete and continuous bilinear time series models. *Handbook of Statistics (D.N. Shanbhag and C.R. Rao, Eds).* **21**, 827-870.

[110] Subba Rao, S. (2007) Statistical analysis of a spatio-temporal model with location dependent parameters and a test for spatial stationarity, Technical Report. ⟨http://www.stat.tamu.edu/»suhasini/papers/spatial temporal technical report.pdf⟩.

[111] Tjostheim, D. (1978) Statistical spatial series modelling I. *Adv. Appl. Probab.* **10**, 130-154.

[112] Tjostheim, D. (1983) Statistical spatial series modelling II. Some further results on unilateral lattice processes. *Adv. Appl. Probab.* **15**, 562-584.

[113] Turkman, K. F. Amaral Turkman, M. A. and Pereira, J. M. (2010) Asymptotic models and inference for extremes of spatio-temporal data. *Extremes.* **13,** 375-397.

[114] Wall, M. (2004) A close look at the spatial structure implied by the CAR and SAR models. *Journal of Statistical Planning and Inference.* **121**, 311 – 324.

[115] Wegman, E.J. Thomas, J.B. and Schwartz, S.C. (1988) *Topics in non-Gaussian signal processing.* Springer.

[116] Whittle, P. (1954). On stationary processes in the plane. *Biometrika.* **41**, 434-49.

[117] Wiener, N. (1958). *Nonlinear problems in random theory.* New York: Wiley.

[118] Yao, Q. and Brockwell, P. (2006) Gaussian Maximum likelihood estimation for ARMA models II: Spatial processes. *Bernoulli.* **12**(3), 403-429.

[119] Noiboar and I. Cohen (2007). Anomaly detection based on wavelet domain GARCH random field modeling. IEEE, Trans. Geos. Romote Sensing, **45**(5). 1361-1373.

[120] Zaharim, A, Ahmad, I, Mohamed, IB, and Yahaya, M.S. (2007) Detection Procedure for Single Outlier and Innovational Outlier in a Bilinear Model. *Pakistan Journal of Statistics and Operation Research.* **3**(1), 1-5.

[121] Zastavnyi, V. P. and Porcu, E. (2009) Space-time covariance functions with compact support.

http://arxiv.org/abs/0902.3656v1

[122] Zhang, J. (2004) Frequent false detection of positive selection by the likelihood method with branch-site models. *Mol Biol Evol,* **21**, 1332-1339.

NAME : Soumia KHARFOUCHI

Title: STATISTICAL INFERENCE IN SPATIAL BILINEAR PROCESSES

Abstract. The spatial treatment of spatial data exhibiting non-linear behavior is the subject of this thesis. This study is motivated by the fact that non-linear modeling is characterized by its wealth, especially in terms of interpretation and understanding of "abnormal" phenomena. In fact, spatial dynamics exhibit non-linear features in a way that is similar to time series models. Attention is focused mainly on extending the bilinear time series model to a spatial (or spatio-temporal) framework, but several useful extensions of the basic spatial autoregressive (SAR) model, which concern random coefficients spatial models, are also considered. We start by deriving conditions for the stationarity of the general spatial bilinear model and establishing the central limit theorem under these conditions. Next, the local asymptotic normality (LAN) based on the loglikelihood ratio is derived for two dimensionally indexed diagonal bilinear processes driven by a Gaussian random field. The joint limit distribution of the maximum likelihood estimators (MLE) is established. A new model of random field on the plane of generalized autoregressive type, that can include as special cases a variety of models of non-linear random fields, is also built. For such models, we establish some results concerning in particular the second and strict stationarity, and the existence of higher order moments.

Estimation procedures for non-linear two-dimensional processes remain relatively poorly explored, that’s why we partly explore the large sample properties of GMM estimator of a two-dimensional random coefficients autoregressive model. Finally, we present the space-time bilinear model. Under a minimal set of conditions, we establish the existence of a stationary (over time) solution. We then deduce a central limit theorem for this new class of spatio-temporal models.

Key Words: Spatial statistic; Spatial autocorrelation; Non-linear modeling; Stationarity; Estimation; CLT.

NOM : Soumia KHARFOUCHI

Titre : INFERENCE STATISTIQUE DANS LES MODELES BILINEAIRES SPATIAUX

Résumé. Le Traitement spatial des données spatiales ayant un comportement non-linéaire fait l'objet de cette thèse. Cette étude est motivée par le fait que la modélisation non linéaire est caractérisée par sa richesse, en particulier en termes d'interprétation et de compréhension de phénomènes «anormaux». En fait, la dynamique spatiale présente des caractéristiques non-linéaires d'une façon similaire aux modèles de séries chronologiques. L'attention est portée principalement sur l'extension du modèle bilinéaire de série temporelle à un cadre spatial (ou spatio-temporel), mais plusieurs autres extensions utiles du modèle autorégressif spatial (SAR) de base, qui concernent les modèles à coefficients aléatoires spatiaux, sont également pris en considération. Nous commençons par dériver les conditions de stationnarité du modèle bilinéaire spatial général et par établir le théorème central limite sous ces conditions. En suite, nous déduisons la normalité asymptotique locale (LAN) basée sur le rapport du log-vraisemblance pour les processus bilinéaires diagonaux à deux dimensions engendrés par un champ Gaussien. La loi limite jointe des estimateurs du maximum de vraisemblance (MV) est établie. Un nouveau modèle de champ aléatoire dans le plan de type autorégressif généralisé, qui peut inclure comme cas particuliers une variété de champs aléatoires non-linéaires, est également construit. Pour ces modèles, nous établissons quelques résultats concernant en particulier la stationnarité du second ordre et stricte, et l'existence des moments d'ordres supérieurs.

Les procédures d'estimation pour les modèles à deux dimensions non-linéaires restent relativement peu explorées, c'est en partie pour ça que nous explorons les propriétés empiriques de l'estimateur GMM du modèle autorégressif à coefficients aléatoires bidimensionnel. Finalement, nous présentons le modèle bilinéaire spatio-temporel. Sous un ensemble minimal de conditions, nous établissons l'existence d'une solution stationnaire (à travers le temps). Nous déduisons alors un théorème central limite pour cette nouvelle classe de modèles spatio-temporels.

Mots Clés : Statistique spatiale; Autocorrélation spatiale ; Modélisation non-linéaire; Stationnarité; Estimation; TCL.

الإسم : سمية خرفوشي

العنوان : الإستدلال الإحصائي في النماذج الثنائية الخطية الفضائية

الملخص

المعالجة الفضائية للبيانات الفضائية التي لها سلوك غير خطي هو موضوع هذه الرسالة. والدافع وراء هذه الدراسة هو ثراء النمذجة غير الخطية ، ولا سيما من حيث تفسير وفهم الظواهر "غير الطبيعية " . في الواقع ، الديناميكيات الفضائية لديها خصائص غير خطية مماثلة لنماذج السلاسل الزمنية. يتم التركيز بشكل رئيسي على تمديد النموذج الثنائي الخطي للسلاسل الزمنية إلى الإطار الفضائي ، لكن العديد من الامتدادات الاخرى لنموذج الانحدار الذاتي الفضائي العام التي تنطوي على النماذج ذات العوامل العشوائية ، تم اعتبارها أيضا. نبدأ باستخراج شروط استقرار النموذج الثنائي الخطي الفضائي العام و تأسيس نظرية النهاية المركزية في ظل هذه الظروف. بعد ذلك نستنتج التقارب الخطي العمودي على أساس نسبة لوغاريثم الاحتمالية للنماذج الثنائية الخطية ذات البعدين التي تم انشاؤها بواسطة حقل غوسي . نموذج جديد للحقل العشوائي في المستوي من نوع الانحدار الذاتي المعمم ، و الذي قد يشمل كحالات خاصة مجموعة متنوعة من الحقول العشوائية غير الخطية ، تم انشاؤه أيضا. من أجل هذه الحقول ، نضع بعض النتائج ، خاصة المتعلقة بالاستقرار من الدرجة الثانية و الصارم و وجود العزوم من الدرجة العليا.

اجراءات التقدير للنماذج غير الخطية ذات بعدين لا تزال غير مستكشفة نسبيا ، من أجل هذا نستكشف الخصائص التجريبية لتقديرات طريقة العزوم المعممة لنماذج الانحدار الذاتي ذات المعاملات العشوائية ذات بعدين. أخيرا نقدم النموذج الثنائي الخطي الزمني الفضائي. تحت حد أدنى من الشروط ، نثبت وجود حل مستقر.

كلمات مفتاحية : الاحصاء الفضائي ؛ الارتباط الذاتي الفضائي ؛ الاستقرار ؛ التقدير ؛ نظرية النهاية المركزية.