People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
Mouloud Mammeri University, Tizi-Ouzou

Faculty of Sciences
Department of Mathematics

Doctoral Thesis

Specialty: Applied Mathematics-Random Process and Statistics

On The Prediction Of Stationary Random Fields

Presented by : Ouerdia Arezki

Jury :

President:	Fazia BEDOUHENE	Professor, Univ. Tizi-Ouzou
Supervisor:	Abdelghani HAMAZ	Senior Lecturer A, Univ. Tizi-Ouzou
Examiner:	Djamel HAMADOUCHE	Professor, Univ. Tizi-Ouzou
Examiner:	Abdelouahab BIBI	Professor , Univ. Constantine
Examiner:	Mohamed IBAZIZEN	Senior Lecturer, Univ. Poitiers
Invited:	Farida ACHEMINE	Senior Lecturer A, Univ. Tizi-Ouzou

Acknowledgments

This dissertation would not have been possible without the help and support of a great number of people :

I would first like to acknowledge my supervisor, Pr. Abdelghani Hamaz, whose expertise and guidance were invaluable in all the time of research and writing of this thesis. I am indebted and thankful for the opportunities he offered at several points during my thesis work and for confidence he has in my abilities.

I would like to thank Pr. Fazia Bedouhene for generously agreeing to head the committee. I would like to thank the dissertation committee, Pr. Djamel Hammadouche, Pr. Abdelouahab Bibi and Pr. Mohamed Ibazizen, not only for their time and patience, but for their insightful feedbacks which will help me to improve my work.

I would like also to thank Pr. Farida Achemine for accepting to take part of the jury.

My thanks and appreciation also go to all the Department of Mathematics, UMMTO University.

I am very grateful to my parents, my sisters and brothers for their support. Thank you for supporting me for everything, and especially I can't thank you enough for encouraging me throughout this experience. My adorable nephews, your messes always make me laugh !

Contents

List of Figures

List of publications

- International journals

  1. **Arezki, O.**, Hamaz, A., On linear prediction for stationary random fields with nonsymmetrical half-plane past, Communications in Statistics - Theory and Methods 1-12(2020).
  2. Hamaz, A., **Arezki, O.**, and Achemine, F. Impact of missing data on the prediction of random fields. Journal of Applied Statistics, 47(1), 132-149(2020).

- International conferences

  1. **Arezki, O.**, Hamaz, A. The first international conference on the "Evolution of Contemporary Mathematics and their Impact in Sciences and Technology", ECMI-SciTech 2017.
  2. Hamaz, A., **Arezki, O**. Influence des données manquantes sur la prédiction des champs aléatoires stationnaires. Modélisation Stochastique et Statistique, Book of Proceedings MSS'2019, 144-147(2019).

Introduction

Spatial statistics study phenomena whose observation is a random process $X = \{X(s), s \in \mathbb{K}\}$ indexed by a spatial set $\mathbb{K}$, The location of an observation site $s \in \mathbb{K}$ is either fixed and deterministic, or random. Classically, $\mathbb{K}$ is a two-dimensional subset, $\mathbb{K} \subseteq \mathbb{Z}^2$.

Spatial prediction methods are experiencing significant development due to strong demand from many fields of application such as agriculture, climatology, ecology, econometrics, geology, medical studies, oil prospecting and water pollution analysis, where data are available at specific locations. Most often, the multi-step prediction is accomplished with the kriging method, it is a geostatistical linear least squares estimation method in spatial statistics for the best linear prediction of a natural phenomenon at any unsampled arbitrary locations of interest as a weighted average of the neighboring observed values. The determination of unknown weights requires specification of a parametric model for the covariance structure with few parameters. It relies on the knowledge of the covariances between the observed and the interpolated locations. The key step lies in fitting the experimental variogram with a parametric variogram model function by choosing a suitable variogram model and estimating the corresponding parameters as for instance presented in [1]. However, the main disadvantage with this method is that one has to select a model for the covariance function which is a very subjective task ([2]). In practice, the model for the covariance is selected from a list of available models and then its parameters are estimated from the observed data $X(s)$. This presupposition that the data can be modelled by a specific covariance function is not universally accepted because a misspecified model can often lead to highly biased predictions. Some work has been done by using the nonparametric methods of estimating a multivariate covariance function ([3, 4]). Another drawback of this method is that the assumption of similar dependence structure at all the points may not hold true when the data is highly irregular.

Hence, developing a prediction method which makes minimal assumptions on the covariance structure and its data driven is a major challenge for prediction on random fields.

Thesis Statement

Prediction problem of random fields has been widely investigated in recent years ([5], [6], [7], [8], [9], [10]). In this thesis, we deal with two non-standard prediction problems of stationary random fields. Firstly, prediction of stationary random fields where a number of observations are missing to the quarter-plane past. Secondly, prediction of stationary random fields based on nonsymmetrical half-plane past. Toward this end, we consider the structure of random process when $\mathbf{S} \subseteq \mathbb{Z}^2$ (2-D discrete random fields). The analysis is performed by formulating the 2-D linear prediction problem in a similar manner to Wiener-Kolmogorov prediction analysis of 1-D discrete random processes. The basic step in the prediction theory of 1-D random processes is the orthogonal decomposition theorem, known as the Wold decomposition [11]. However, in the 2-D domain, contrary to the 1-D case, there is no natural order definition, and hence terms like "past" and "future" are with no significance unless defined with respect to a specific order. In general, different order definitions will lead to different orthogonal decompositions ([12], [13]), defined the predictor support to be a NSHP and derived a Wold-type decomposition for stationary random fields. They also studied the spectral theory which corresponds to this decomposition. [14] showed that by considering multiple total-order and nonsymmetrical half-plane definitions, a countably infinite Wold-type decomposition of the field can be obtained. It should be mentioned that other types of past have been considered in the litterature, the case of quarter-planes were studied . Thus, an important representation is the quarter-plane (QP) representation ([15], [16], [17], [18], [19], [7], [8], [9], [10]).

Thesis Contributions

The main contributions of this thesis are:

- The first contribution is presented in Section 2.5. We treat the prediction problems where a number of observations are missing to the quarter-plane past of a stationary random field. The purpose consists in quantifying the influence of missing values on the prediction by giving the simple bounds for the prediction error variance. These bounds allow to characterize the random fields for which the missing observations do not affect the prediction.
- The second contribution [20] is developed in Chapter 3. We investigate the problem of linear prediction of stationary random fields with nonsymmetrical half-plane past. We establish an explicit autoregressive series representation for the best multi-step ahead linear predictor of stationary random fields with nonsymmetrical half-plane past (NSHP). Furthermore, necessary and sufficient condition for the mean square convergence of these series is given.

Thesis Organization

This thesis is divided according to the objectives stated above and is therefore structured in three chapters.

- In Chapter 1, we have presented a certain number of structural properties of weakly stationary random fields. A spatial process may be considered as a generalisation of 1-D process. However, its characteristics make its analysis considerably more difficult. The key difference is that a 1-D process is unidirectional with a natural ordering from past to future. However, this ordering does not exist for a general spatial processes. There is no natural order definition, and hence terms like "past" and "future" are with no significance unless defined with respect to a specific order. In general, different order definitions will lead to different orthogonal decompositions. We list them as follows : Nonsymmetrical half-plane (NSHP) orthogonal decomposition : any stationary random fields is decomposed into two mutually orthogonal components: a purely indeterministic component and deterministic component ([14, 20]). Rational nonsymmetrical half-plane (RNSHP) decomposition : any stationary random fields is decomposed into two mutually orthogonal components: a purely indeterministic component and a deterministic component. The deterministic component is further decomposed into mutually orthogonal half-plane deterministic and generalized-evanescent components ([21]). Other pasts have been considered, namely quarter-plane pasts. Thus, an important representation is the quarter-plane (QP) representation;

- Chapter 2 contains two parts: In the first one, an overview of some basic results on stationary random fields is presented: Firstly, the authors in [7] studied in details some prediction problems for stationary random fields with quarter-plane past. They obtained the best predictor of $X(0,0)$ based on the quarter plane with finitely many observations added. Their solution expresses the error variance formulas in terms of the moving average (MA) parameters of the random fields. Secondly, the authors in [8] developed methods which get around the technical difficulties noted in [7] and solved the quarter plane prediction problem with finitely many missing values left open in [7] when the stationary Gaussian random fields possesses a onesided moving average representation. Two solutions have been presented for predicting a stationary Gaussian random fieldbased on a quarter-plane with finitely missing observations . One expresses the best predictor in terms of the moving average coefficients of the random field, and the other expresses the prediction error variance in terms of its autoregressive coefficients. Finally, in [9], the authors established an explicit formula for the prediction error of a future value of stationary random fields when the infinite past is altered by some missing observations.

  In the second part, we treat the prediction problems where a number of observations are missing to the quarter-plane past of a stationary random field. The aim is to quantify the impact of missing observations from the past. The central idea of our study consists in using the MA and AR representations of the random field. The obtained results highlight the important role of the AR parameters in forecasting. Indeed, we establish lower and upper bounds for the prediction error variance given in Theorem (2.5.1) which is the novelty of our work. This boundedness property of prediction error variance shows that the degradation of the prediction due to the missing data increases with the maximum value of the AR parameters of the missing data. They also allow to conclude that the larger the indices of missing values are, the better is the preciseness of the bound of the prediction error variance. Also, our results characterize the random fields for which the missing observations do not affect the prediction of $X(0,0)$. We conclude this with two illustrative examples.

- In Chapter 3, the problem of interest consists in finding the minimum-norm linear predictor $\hat{x}(s,t)$ of $x(s,t)$ as the projection of $x(s,t)$ on the Hilbert space spanned by all the field samples that precede the (s,t)-th sample, according to the defined orderwith. The main topics touched upon in this chapter are : an explicit autoregressive series representation for the best multi-step ahead linear predictor of stationary random fields with nonsymmetrical half-plane past (NSHP) is established. Necessary and sufficient condition for the mean square convergence of these series is given. Moreover, step recursive relations between the prediction coefficients for the infinite past predictor are provided. These relations are used to calculate explicitly the multi-step prediction coefficients.

- Finally, we conclude with the relevant conclusions of the work and some problems for further research.

CHAPTER 1

Properties and Orthogonal Decompositions of Random Fields

Contents

1.1 Introduction

This chapter covers essentially introductory and preparatory material and establishes the necessary notation and assumptions for this thesis. A random field X on $\mathbb{R}^d$means a collection $X = \{X(s), s \in \mathbb{R}^d\}$ of random variables indexed by $\mathbb{R}^d$.

A basic starting point in the prediction theory of 1-D random processes is the orthogonal decomposition theorem, known as the Wold decomposition [11]. The discussion about the Wold-like decomposition will be restricted to the case of two dimensional fields $X = \{X(s),\ s \in \mathbb{Z}^2\}$. In other words, we study the 2-D Wold-like decomposition for stationary random fields with respect to the defined support and total order.

The outline of this chapter is : Section 1.2 furnishes an overview of the needed theoretic background of random fields including some basic definitions and the special class of Gaussian random fields [22–24]. The main result presented in Section 1.3 is the spectral representation theorem of stationary random fields ([25]). To this end, some needed results, namely the mean and covariance functions of random fields; the notion of continuity and differentiability of random fields; and stochastic integration are presented beforehand. In Section 1.4, we detail the orthogonal decompositions of 2-D random fields.

1.2 Basic definitions and structural properties of random fields

The study of random fields is, by definition, the study of random functions defined over some Euclidean space [22–24]. Consequently, this study can cover an extremely wide area, since any question that can be asked about an ordinary non-random function, or class of functions, can just as readily be asked about their random counterparts. Hence, the general theory of random fields is certainly at least as large as the general theory of functions. Indeed, adding a random component to the theory of functions makes it a much larger, more interesting, and often more complex subject.

1.2.1 Some backgrounds in random fields

There are two virtually distinct approaches to defining random fields. One is essentially a measure theoretic approach and leads ultimately to a probabilistic setting, while the other starts probabilistically and can be eventually placed in a measure theoretic setting. Let us consider the measure theoretic approach first. Let $\mathbf{G}^{d,N}$ denote the set of all $\mathbb{R}^N-$valued functions on $\mathbb{R}^d, N, d \geq 1$, and $\mathscr{E}^{d,N}$ the $\sigma-$field containing all sets of the form $\{\mathbf{g} \in \mathbf{G}^{d,N}, \mathbf{g}(\mathbf{s}_j) \in \mathbf{B}_j, j = 1, \ldots, m\}$ where m is an arbitrary integer, the $\mathbf{s}_j$ are points of $\mathbb{R}^N$, and the $\mathbf{B}_j$ are half open intervals in $\mathbb{R}^N$. Then, much as we defined random variables, we define an (d,N) or $d-$dimensional, random field, or stochastic process, to be a measurable mapping X from $(\Omega, \mathscr{F})$ into $(\mathbf{G}^{d,N}, \mathscr{E}^{d,N})$. We use the notation $X(\mathbf{s}, \omega)$ to denote the value the function in $\mathbf{G}^{d,N}$ corresponding to ω takes at the point $\mathbf{t}$.

Given the existence of the probability measure $\mathbf{P}$ on $\mathscr{F}$ we can obtain from this definition a collection of measures $F_{\mathbf{s}_1,\ldots,\mathbf{s}_n}$ on $\mathscr{B}^{nN}$ defined by

$$F_{\mathbf{s}_1,\ldots,\mathbf{s}_n}(\mathbf{B}) = \mathbf{P}((X(\mathbf{s}_1), \ldots, X(\mathbf{s}_n)) \in \mathbf{B}) \tag{1.1}$$

for any $\mathbf{B} \in \mathscr{B}^{nN}$. The collection of all such measures, or, equivalently, the corresponding distribution functions, is known as the family of finite dimensional distributions for the field X. In general it is the finite dimensional distributions that we work with in the study of a random field. the finite dimensional distributions of any given random field uniquely define the $\mathbf{P}$ measure of all sets in the $\sigma-$field $\mathscr{E}^{d,N}$. Not all events of interest are in $\mathscr{E}^{d,N}$, however. For example, sets of the form

$$\{\mathbf{g},\ \mathbf{g}(\mathbf{s}) \in \mathbf{B},\ \text{for all}\ \ \mathbf{s} \in I\},$$

where $\mathbf{B} \in \mathscr{B}^d$ and I is an interval in $\mathbb{R}^d$, are not usually in $\mathscr{E}^{d,N}$, they are of obvious interest. To obtain the probability of such a set from knowledge of the finite dimensional distributions only it is necessary to assume some other condition, such as separability.

The second definition of a random field, which is more natural in modelling context, is to define it as a collection of random variables $X(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d$, together with a collection of measures or distribution functions of the form $F_{\mathbf{s}_1,\dots,\mathbf{s}_n}$ on $\mathbf{B} \in \mathscr{B}^{nN}, n = 1, \dots, \mathbf{s}_i \in \mathbb{R}^d$ which satisfies (1.1). The question to ask is whether or not one can always find a random field, according to the first definition, which posses theses measures as finite dimensional distributions. According to Kolmogorov's result is that a necessary and sufficient condition for the existence of such a field is that the given family of measures satisfies both conditions:

1. Symmetry : Writing F as a distribution function $F_{\mathbf{s}_1,\dots,\mathbf{s}_n}(X_1,\dots,X_n)$, F should remain invariant when the X_j and $\mathbf{s}_j$ are subjected to the same permutation.

2. Consistency : $F_{\mathbf{s}_1,\dots,\mathbf{s}_{n+n}}(B, \mathbb{R}^{mN}) = F_{\mathbf{s}_1,\dots,\mathbf{s}_n}(\mathbf{B})$ for every $n, m \geq 1$ and $\mathbf{B} \in \mathscr{B}^{nN}$.

Not every family of measures corresponds to a random field. However, we shall only deal with families that do. We conclude by introducing an powerful concept in the study of random fields called stationarity. We call a real- valued random field $X(\mathbf{s})$ strictly stationary if, for arbirary k, any real numbers $x_1, \dots, x_k$ and any $(k+1)$ points $\tau, \mathbf{s}_1, \dots, \mathbf{s}_k$ in $\mathbb{R}^d$ the following condition on its finite dimensional distributions is satisfied

$$\mathbf{P}(X(\mathbf{s}_1) \leq x_1, \dots, X(\mathbf{s}_k) \leq x_k) = \mathbf{P}(X(\mathbf{s}_1 + \tau) \leq x_1, \dots, X(\mathbf{s}_k + \tau) \leq x_k).$$

1.2.2 Gaussian random fields

An important special class of random fields is the class of Gaussian random fields. The main reason for this is the convenient analytic form of Gaussian density, which allows explicit results to be obtained for Gaussian fields which seem almost impossible to derive for more general processes.

Definition 1.2.1 ([22–24].). *X is a Gaussian process on $\mathbb{R}^d$ if for every finite subset $\Lambda \subset \mathbb{R}^d$ and real valued sequence $a = (a_s, s \in \Lambda)$, $\sum_{s \in \Lambda} a_s X(s)$ is a Gaussian random variable.*

If $m_\Lambda = E(X_\Lambda)$ is the mean of $X_\Lambda = (X(s), s \in \Lambda)$ and Σ_Λ its covariance, then if Σ_Λ is invertible, the density of (X_Λ) with respect to the Lebesgue measure on $\mathbb{R}^{*\Lambda}$ is

$$f_\Lambda(x_\Lambda) = (2\pi)^{\dfrac{-*\Lambda}{2}} (det\,\Sigma_\Lambda)^{\frac{1}{2}} e^{\dfrac{-1}{2}(x_\Lambda - m_\Lambda)^t \Sigma_\Lambda^{-1}(x_\Lambda - m_\Lambda)},$$

where $*U$ is the cardinality of U and x_Λ possible values of X_Λ. Such densities are defined and for any mean function m and p.d. covariance γ there exists a Gaussian random field with mean m and covarince γ.

1.3 Stationary random fields and their spectra

A full proof of the spectral representation Theorem 1.45 is given by [26] which only obtains the theorem under stronger conditions. However, his approach is not based on the Hilbert space theory that most fundamental results of random fields rely on. A spectral theory for processes defined on more general spaces than $\mathbb{R}^d$ is developed by [27] who study processes defined over complete, separable, metric spaces, while [28] studies processes defined over certain separable, locally compact groups. The spectral theory of stationary random fields was rigorously established in [25].

1.3.1 The mean and covariance functions

Although our primary interest lies in the study of real-valued random fields, it is mathematically more convenient, and often more useful, to develop spectral theory for complex valued processes. Hence, we shall assume that $X(\mathbf{s})$ takes values in the complex plane $\mathbb{C}$, while $\mathbf{s} \in \mathbb{R}^d$. We shall also assume that the expectation $E(X(\mathbf{s}))^2$ is finite for all $\mathbf{s} \in \mathbb{R}^d$, i.e. $E(X(\mathbf{s}))^2 < \infty$. With these assumptions the mean function

$$m(\mathbf{s}) = E(X(\mathbf{s}))$$

becomes a non random function from $\mathbb{R}^d$ to $\mathbb{C}$. We define the covariance function $\gamma : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$, defined for all $\mathbf{s}, \mathbf{t}$ by

$$\gamma(\mathbf{s},\mathbf{t}) = Cov((X(\mathbf{s}) - m(\mathbf{s}))(\overline{X(\mathbf{t})} - \overline{m(\mathbf{t})})), \tag{1.2}$$

where the bar denotes complex conjugation.
It is an immediate consequence of (1.2), the finiteness of $E(X(\mathbf{s}))^2$, and the Cauchy Schwartz inequality that $\gamma(\mathbf{s},\mathbf{t})$ is finite for all $\mathbf{s},\mathbf{t} \in \mathbb{R}^d$. Furthermore, it is easily seen that $\gamma(\mathbf{s},\mathbf{s})$, which is the variance of $X(\mathbf{s})$, is real and non-negative, and that the covariance function satisfies the relation

$$\gamma(\mathbf{s},\mathbf{t}) = \overline{\gamma(\mathbf{s},\mathbf{t})}$$

for all $\mathbf{s},\mathbf{t} \in \mathbb{R}^d$. Every covariance function is a non-negative definite function on $\mathbb{R}^{2d}$, i.e., if $\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_k$ is any collection of points of $\mathbb{R}^d$, and $z_1, z_2, ..., z_k$ are arbitrary complex numbers, then the Hermitian form

$$\sum_{i=1}^{k}\sum_{j=1}^{k} \gamma(\mathbf{s}_i, \mathbf{s}_j) z_i \overline{z_j} \tag{1.3}$$

is real and non-negative. This follows from the fact that this expression is equal to

$$E\left|\sum_{i=1}^{k}(X(\mathbf{s}_i) - m(\mathbf{s}_i))z_i\right|^2$$

which is real and non negative. The property of non negative definiteness actually characterizes covariance functions, so that given any function $m(\mathbf{s}) : \mathbb{R}^d \to \mathbb{C}$, and a non-negative definite $\gamma(\mathbf{s},\mathbf{t}) : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ it is always possible to construct a random field for which $m(\mathbf{s})$ and $\gamma(s,\mathbf{t})$ are the mean and covariance functions, respectively.
As we have already noted, a central concept in the study of random fields is that of stationarity. We say that a random field $X(\mathbf{s})$ is strictly stationary if its finite dimensional distributions are invariant under translations in the parameter s. That is, for any set of points $\tau, \mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_k \in \mathbb{R}^d$ the joint distribution of the k complex variables $(X(\mathbf{s}_1), X(\mathbf{s}_2), ..., X(\mathbf{s}_k))$ should be the same as for the variables

$$(X(\mathbf{s}_1 + \tau), X(\mathbf{s}_2 + \tau), ..., X(\mathbf{s}_k + \tau)).$$

An immediate consequence of this property is that the mean function $m(\mathbf{s})$ is identically equal to a constant, which we shall hereafter take to be zero. This assumption reduces notational complexity for no real loss of generality. A second consequence of strict stationarity is that the covariance function $\gamma(\mathbf{s},\mathbf{t})$ is forced to be a function of the difference $\mathbf{s} - \mathbf{t}$ only.
Very often, when dealing with random fields, it is not necessary to impose the restrictive condition of strict homogeneity. Hence, we shall introduce a class of fields which we shall call a second-order stationary, X is a second-order stationary process on $\mathbb{R}^d$ if it has constant mean and translation invariante covariance

$$\gamma:\ \forall \mathbf{s}, t \in \mathbb{R}^d : E(X(\mathbf{s})) = m \text{ and } \gamma(\mathbf{s},t) = Cov(X(\mathbf{s}), X(\mathbf{t})) = \gamma(\mathbf{s} - \mathbf{t}).$$

Example 1.3.1 (Strong White Noise (SWN) and Weak White Noise (WWN)). *X is a Strong White Noise if the variables $\{X(s), s \in \mathbb{R}^d\}$ are centered, independent and identically distributed (i.i.d.). X is a Weak White Noise if the variables $\{X(s), s \in \mathbb{R}^d\}$ are centered and uncorrelated with finite constant variance: if $s \neq t$, $cov(X(s), X(t)) = 0$ and $V(X(s)) = \sigma^2 < \infty$. A SWN on $\mathbb{R}^d$ is strictly stationary; a WWN on $\mathbb{R}^d$ is a stationary process in L^2.*

A strictly stationary field is clearly stationary, but in general the reverse is not true. However, it is true for certain Gaussian fields which satisfy the following condition.

Theorem 1.3.2. *In order that a stationary, Gaussian random field be strictly stationary it is necessary and sufficient that the function*

$$H(\mathbf{s}, \mathbf{t}) = E(X(\mathbf{s}), X(\mathbf{t}))$$

should be a function of $\mathbf{s} - \mathbf{t}$ only.

Note that if $X(\mathbf{s})$ is complex valued H is not its covariance function, since we do not take the complex conjugate of the second factor in the expectation. On the other hand, if $X(\mathbf{s})$ is real valued $H = \mathbb{R}$, so that as a corollary to the theorem, we have that every stationary Gaussian process with real values is also strictly stationary.
We conclude this section by noting that for stationarity fields the Hermitian form given in (1.3) simplifies to

$$\sum_{i=1}^{k} \sum_{j=1}^{k} \gamma(\mathbf{s}_i - \mathbf{s}_j) z_i \overline{z_j}. \tag{1.4}$$

Since this must be non-negative for any set $\mathbf{s}_i \in \mathbb{R}^d$ and complex numbers z_i, we have that the covariance function, considered now as a function on $\mathbb{R}^d$, is non-negative definite. It is useful to know that such functions can be characterized, as in the following result.

Theorem 1.3.3. *([29], [30, p. 209-211]) A continuous function $\gamma(\mathbf{s})$ from $\mathbb{R}^d$ to the complex plane is non-negative definite if and only if it can be represented in the form*

$$\gamma(\mathbf{s}) = \int_{\mathbb{R}^d} e^{i\langle \mathbf{s}, \lambda \rangle} dF(\lambda), \tag{1.5}$$

where $< \mathbf{s}, \lambda >$ denotes the inner product $\sum_{i=1}^{d} \mathbf{s}_i \lambda_i$ and $F(\lambda)$ is a bounded, real-valued function satisfying $\int_{\mathbf{A}} dF(\lambda) \geq 0$ for all measurable $\mathbf{A} \subset \mathbb{R}^d$.

Since every covariance function is non negative definite, each such function has a representation of the form (1.5). The function F of the theorem is only specified up to an additive constant. When we fix F by demanding that

$$F(-\infty, \ldots, -\infty) = 0$$

(thus implying that $F(\infty, \ldots, \infty) = \gamma(0)$, we call (1.5) the spectral representation of $\gamma(\mathbf{s})$. When $F(\lambda)$ is absolutely continuous we call the corresponding density, $f(\lambda)$ say, the spectral density and (1.5) becomes

$$\gamma(\mathbf{s}) = \int_{\mathbb{R}^d} e^{i\langle \mathbf{s}, \lambda \rangle} f(\lambda) d\lambda. \tag{1.6}$$

These representations of the covariance function are extremely helpful in investigating its properties.

1.3.2 Geometric properties: Continuity and differentiability of random fields

Questions about continuity and differentiability of a function $f(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d$ at point $\mathbf{s}^*$ boil down to questions about the convergence of sequences of the form $(f(\mathbf{s}_n))$ when $\|\mathbf{s}_n - \mathbf{s}^*\| \to \infty$. When the functions being studied is actually a random field, it follows that we are actually asking questions about the convergence of a sequence $(X(\mathbf{s}_n))$ of random variables. Hence, just as there are various modes of convergence, there are various types of continuity and differentiability for random fields. We shall consider only two of them.

The strongest form of stochastic convergence is almost sure convergence. Corresponding to this, we shall say that a random field X is almost surely continuous at $\mathbf{s}^*$ if for every sequence $(\mathbf{s}_n)$ for which $\|\mathbf{s}_n - \mathbf{s}^*\| \to \infty$ as $n \to \infty$, $X(\mathbf{s}_n) \stackrel{a.s.}{\to} X(\mathbf{s}^*)$. Alternatively, we could write this condition as

$$P\{\omega : \|X(\mathbf{s}_n, \omega) - X(\mathbf{s}^*, \omega)\| \to 0 \text{ as. } n \to \infty\} = 1.$$

We say that X is almost surely continuous through a set $A \subset \mathbb{R}^d$ if it is almost surely continuous at each $\mathbf{s} \in \mathbf{A}$.

Similarly, if we write f_j to denote the th first-order partial derivative $\dfrac{\partial f}{\partial t_j}$ of a function f , we say that a random field X is a.s. differentiable at a point $\mathbf{s}^*$ if $P\{\omega : \|X_j(\mathbf{s}^*, \omega)\| \to 0 \text{ exists } \} = 1$. Let $\delta_j, 1 \le j \le d$, denote the d vector, all of whose elements are zero, except for the j^{th}, which is one. Then if X is almost surely differentiable at $\mathbf{s}^*$ the a.s. limit

$$X_j(\mathbf{s}^*) = \lim_{h \to 0} \frac{X(\mathbf{s}^* + h\delta_j) - X(\mathbf{s}^*)}{h}$$

is called the function j^{th} partial derivative of X at $\mathbf{s}^*$. By allowing $\mathbf{s}^*$ to vary we obtain the partial derivative field $X_j(\mathbf{s})$.

We shall generally wish to use the information contained in the covariance function to investigate the behavior of random fields. Thus, because covariance functions are essentially only second-order moments, if often turms out that the most natural form of convergence to use in the theory of random fields is mean square convergence. We shall now see how this mode of convergence can also be used to develop notions of continuity and differentiablity for random fields, and how these relate to the covariance function.

Let $\mathbf{s}_1, \mathbf{s}_2, \ldots$ be a sequence of points and $\mathbf{s}^*$ a fixed point in $\mathbb{R}^d$ for which $\|\mathbf{s}_n - \mathbf{s}^*\| \to \infty$ as $n \to \infty$. Then if

$$X(\mathbf{s}_n) \stackrel{m.s.}{\to} X(\mathbf{s}^*), \text{ as } n \to \infty$$

we say that X is continuous in mean square at $\mathbf{s}^*$. If this holds for all $\mathbf{s}^* \in A$, where A is some subset of $\mathbb{R}^d$, we say that $X(\mathbf{s})$ is continuous in mean square over A. It is interesting to note that whether or not a fields possesses this property can be red off from the covariance function of $X(\mathbf{s})$, as in the following theorem

Theorem 1.3.4 ([22]). *A random field $X(\mathbf{s})$ is continuous in mean square at the point $\mathbf{s}^* \in \mathbb{R}^d$, if and only if its covariance function $\gamma(\mathbf{s}, \mathbf{t})$ is continuous at the point $\mathbf{s} = \mathbf{t} = \mathbf{s}^*$. If $\gamma(\mathbf{s}, \mathbf{t})$ is continuous at every diagonal point $\mathbf{t} = \mathbf{s}$ then it is everywhere continuous.*

We now consider differentiability in mean square, for which we assume that the field X is real valued.

Theorem 1.3.5 ([22]). *If the derivative $\dfrac{\partial^2 \gamma(\mathbf{s}, \mathbf{t})}{\partial \mathbf{s}_i \partial \mathbf{t}_i}$ exists and is finite at the point $(\mathbf{s}, \mathbf{s}) \in \mathbb{R}^{2d}$, then the limit*

$$X_i(\mathbf{s}) = \lim_{h \to 0} \frac{X(\mathbf{s} + h\delta_i) - X(\mathbf{s})}{h} \tag{1.7}$$

exists, and $X_i(\mathbf{s})$ is called the mean square derivative of $X(\mathbf{s})$ at $\mathbf{s}$. If this exists for each $\mathbf{s} \in \mathbb{R}^d$ then $X(\mathbf{s})$ is said to possess a m.s. derivative. The covariance function of $X_i(\mathbf{s})$ is then given by $\dfrac{\partial^2 \gamma(\mathbf{s},\mathbf{t})}{\partial \mathbf{s}_i \partial \mathbf{t}_i}$.

We note that when the basic process X(t) is stationary, the conditions ensuring m.s. continuity and the existence of m.s. derivatives become particularly simple. For example, by Theorem 1.3.4, $X(\mathbf{s})$ will be m.s. continuous at $\mathbf{s}$ if its covariance function is continuous at $(\mathbf{s},\mathbf{s})$. But because of stationarity, this is equivalent to demanding that $\gamma(\mathbf{s})$ be continuous at $\mathbf{s}-\mathbf{s}=0$.

In these results, the fact that for a stationarity random field the behaviour of its covariance function in the neighbourhood of the origin may be a determining factor in regard to mean square local properties (continuity, differentiability, etc.) of the field is implicit.

1.3.3 Stochastic integration

In this section, we deal with the theory of stochastic integration on $\mathbb{R}^d$, we limit our treatment to the existence of certain Fourier type integrals. The theory of d-dimensional stochastic integration for is treated in detail in [26, 27] .

We start by defining a particular class of random fields. Given any complex-valued random field $Z(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d$, it is possible to define a random additive set function on the set of finite unions of intervals of $\mathbb{R}^d$ in the following manner. Let $I = (a_1, b_1] \times \ldots \times (a_d, b_d]$ be an interval in $\mathbb{R}^d$, and set

$$Z(I) = Z(b_1, \ldots, b_d) - [Z(b_1, \ldots, b_d) + \ldots + Z(b_1, \ldots, b_{d-1}, a_1)] + \ldots + (-1)^d Z(b_1, \ldots, b_d). \tag{1.8}$$

This defines $Z(I)$ for any interval $I \subset \mathbb{R}^d$. The definition is extended to the class $\mathscr{A}^d$ of sets $A \in \mathbb{R}^d$ which are finite unions of intervals of $\mathbb{R}^d$

$$Z(A \cup B) = Z(A) \cap Z(B)$$

if $A, B \mathscr{A}^d$ and $A \cap B = \varnothing$.

Now suppose that $X(\mathbf{s})$, considered as a random field, has finite variance for all $\mathbf{s} \in \mathbb{R}^d$. Then, by (1.8), $E|Z(I)|^2$ will also be finite for all $I \subset \mathbb{R}^d$. We then call $Z(\mathbf{s})$ a field with orthogonal increments if for every pair of disjoint intervals $I_1, I_2 \in \mathbb{R}^d$ by demanding additivity, i.e.

$$E\left(Z(I_1)\overline{Z(I_2)}\right) = 0. \tag{1.9}$$

An immediate consequence of (1.9) is that such a field determines a measure on $\mathscr{A}^d$, generally denoted by $F(I)$ and defined by

$$F(I) = E|Z(I)|^2. \tag{1.10}$$

Such a measure will, up to a constant, determine a point function on $\mathbb{R}^d$ that is non decreasing in each $s_i, 1 \le i \le d$. If we also use F to denote the point function, set $F(-\infty, \ldots, -\infty) = 0$, and demand that F is bounded over the whole of $\mathbb{R}^d$, then $F(\mathbf{s})$ has all the properties of a multidimensional distribution function multiplied by an arbitrary constant. The relationship between F as a point and set function is also of the form given in (1.8) for Z. Hereafter we shall assume, without explicit mention, that $F(\mathbf{s})$ is bounded for all $\mathbf{s} \in \mathbb{R}^d$.

For random fields of the above form it is possible to define a wide class of stochastic integrals known as mean square integrals, a subset of which is given in the following theorem.

Theorem 1.3.6 ([25]). *If $Z(\mathbf{s})$ is a complex-valued random field with orthogonal increments, then for every $\mathbf{s} \in \mathbb{R}^d$ the Riemann-Stieltjes integral*

$$X = \int_{\mathbb{R}^d} \mathbf{g}(\lambda) dZ(\lambda) \tag{1.11}$$

is well defined as

$$\lim_{\mathbf{A}\to\infty}\lim_{n\to\infty}\sum_{j=1}^{\mathbf{A}_n}\mathbf{g}(\lambda_{nj})Z(I_{nj}),$$

for all complex valued $\mathbf{g}$ *for which*

$$X = \int_{\mathbb{R}^d}|\mathbf{g}(\lambda)|^2 dF(\lambda) < \infty. \tag{1.12}$$

In defining the mean square limit we partition, for each $\mathbf{A} > 0$ and $n \geq 1$, the interval $(-\mathbf{A},\mathbf{A}]^d$ into $\mathbf{A}$ disjoint intervals $I_{nj}, 1 \leq j \leq \mathbf{A}_n$, in such a way that

$$\sup_{1\leq j\leq \mathbf{A}_n} \lambda_d(I_{nj}) \to 0 \quad \text{as} \quad n \to 0.$$

By setting $g(\lambda) = e^{i<\mathbf{s},\lambda>}$ we obtain from (1.11) a family of random variables

$$X(\mathbf{s}) = \int_{\mathbb{R}^d} e^{i<\mathbf{s},\lambda>} dZ(\lambda), \tag{1.13}$$

which is well defined for any Z for which

$$\sup_{\mathbf{A}_n} E|Z(I)|^2 < \infty.$$

The complex valued field X defined in this fashion has rather special properties, for if we set $F(\mathbf{s})$ to be the function defined by (1.10) and assume for the moment that $E(Z(\mathbf{s})) = 0$ for all $\mathbf{s}$, we have the following theorem.

Theorem 1.3.7 ([25]). *For the random field* $X(\mathbf{s})$ *determined by* (1.12)

$$E(X(\mathbf{s})) = 0 \;\; and \;\; \gamma(\mathbf{s},\mathbf{t}) = E\left(X(\mathbf{s})\overline{X(\mathbf{t})}\right) = \int_{\mathbb{R}^d} e^{i\langle \mathbf{s}-\mathbf{t},\lambda\rangle} dF(\lambda), \tag{1.14}$$

where the integral is a standard Riemann Stieltjes integral.

An immediate implication is that since the covariance function $\gamma(\mathbf{s},\mathbf{t})$ of $X(\mathbf{s})$ depends only on the difference $\mathbf{s}-\mathbf{t}$, the random field defined by the stochastic integral (1.13) is stationary. Furthermore, the spectral distribution function of $X(\mathbf{s})$, as defined in (1.5), is simply the function $F(\lambda)$ determined by the field $Z(\mathbf{s})$ in (1.10). This automatically leads us to question whether the fact that an integral such as (1.13) leads to a stationary field whose covariance function satisfies (1.14) implies that the converse holds; i.e. given a stationary field whose covariance function satisfies (1.14) can we represent it as a stochastic integral as in (1.13).

1.3.4 The spectral representation theorem

We now proceed to give a positive answer to the above questions. Suppose $X(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d$, is a stationary, complex valued random field for which $E(X(\mathbf{s})) = 0$ and $E|X(\mathbf{s})|^2 < \infty$. Furthermore, let $X(\mathbf{s})$ be continuous in mean square and let $\gamma(\mathbf{s})$ denote its covariance function. Then, by Theorem 1.3.4, $\gamma(\mathbf{s})$ is continuous for all $\mathbf{s} \in \mathbb{R}^d$ and so, by Theorem 1.5, has a representation of the form

$$\gamma(\mathbf{s}) = \int_{\mathbb{R}^d} e^{i\langle \mathbf{s},\lambda\rangle} dF(\lambda), \tag{1.15}$$

where $F(\lambda)$ is the spectral distribution function of $X(\mathbf{s})$. The following theorem is known as the spectral representation theorem.

Theorem 1.3.8 ([25]). *For every mean square continuous, zero mean, stationary random field $X(\mathbb{R}^d)$ there exists a field $Z(\mathbf{s})$ with orthogonal increments such that for each $\mathbf{s}$, $X(\mathbf{s})$ has the following representation as a mean square integral:*

$$X(\mathbf{s}) = \int_{\mathbb{R}^d} e^{i\langle \mathbf{s},\lambda\rangle} dZ(\lambda). \tag{1.16}$$

The field $Z(\mathbf{s})$ is defined up to an additive constant. If this is fixed by setting $Z(-\infty,\ldots,-\infty) = 0$, we have

$$E(Z(\mathbf{s})) = 0,\ E|Z(I)|^2 = F(I),\ E|Z(\lambda)|^2 = F(\lambda),$$

where I is any interval in $\mathbb{R}^d$ and F is determined by (1.15).

The representations of $\gamma(\mathbf{s})$ and $X(\mathbf{s})$ given, respectively, by (1.15) and (1.18) are fundamental to the study of stationary fields.
We shall now investigate some of their simpler implications for real valued random fields. The representation of the covariance function of $X(\mathbf{s})$ can be rewritten as

$$\gamma(\mathbf{s}) = \int_{\mathbf{s}} \cos < \mathbf{s},\lambda > dF(\lambda) + i\int_{\mathbb{R}^d} \sin < \mathbf{s},\lambda > dF(\lambda).$$

If the field $X(\mathbf{s})$ is real valued, then so is $\gamma(\mathbf{s})$, so that

$$\int_{\mathbb{R}^d} \sin < \mathbf{s},\lambda > dF(\lambda) = 0, \quad \text{for all } \ \mathbf{s} \in \mathbb{R}^d.$$

This fact immediately implies that the spectral distribution is symmetric about the origin of $\mathbb{R}^d$, in the sense that for any $\lambda_i, \nu_i, 1 \le i \le d$,

$$F\left((\lambda_1,\nu_1]\times\ldots\times(\lambda_d,\nu_d]\right) = F\left((-\lambda_1,-\nu_1]\times\ldots\times(-\lambda_d,-\nu_d]\right).$$

If a spectral density function $f(\lambda)$ exists, it is also symmetric, in the sense that $f(\lambda) = f(-\lambda)$ for all $\lambda \in \mathbb{R}^d$. Consequently, the odd-ordered moments of F, when they exist, are zero; i.e.

$$\int_{\mathbb{R}^d} \lambda_1^{i_1}\lambda_2^{i_2}\ldots\lambda_d^{i_d} dF(\lambda) = 0 \quad \text{if } \sum_{i=1}^{d} i_j \text{ is odd.} \tag{1.17}$$

Another useful consequence of (1.15) is the existence of relationships between the even ordered moments of F and the behaviour of $\gamma(\mathbf{s})$ near $\mathbf{s} = 0$. For example,

$$\int_{\mathbb{R}^d} dF(\lambda) = \gamma(0). \tag{1.18}$$

Furthermore, if we write λ_{ij} to denote the second order spectral moment

$$\int_{\mathbf{s}} \lambda_1^{i_1}\lambda_2^{i_2} dF(\lambda)$$

and set $\dfrac{\partial^2\gamma(\mathbf{s},\mathbf{t})}{\partial\mathbf{s}_i\partial\mathbf{t}_i}$, then

$$\lambda_{ij} = \left|\frac{\partial^2\gamma(\mathbf{s},\mathbf{t})}{\partial\mathbf{s}_i\partial\mathbf{t}_i}\right|_{\mathbf{s}=\mathbf{t}} = -\gamma_{ij}(0). \tag{1.19}$$

In general, the $2k^{th}$ order spectral moments are equal to the appropriate $2k^{th}$ partial derivative of $\gamma(\mathbf{s})$ at the origin times $(-1)^k$. From Section 1.3.2, the mean square derivative $X_i(\mathbf{s})$ of $X(\mathbf{s})$ if it exists, has the partial derivative $\gamma_{ii}(\mathbf{s})$ as its covariance function. Since the variance of $X_i(\mathbf{s})$ is then given by $\gamma_{ij}(0)$, in view of (1.19) we have

$$E|X_i(\mathbf{s})|^2 = \lambda_{ii} \text{ of } i = 1,\ldots,d. \tag{1.20}$$

Similarly, the variance of a k^{th} order mean square derivative of $X(\mathbf{s})$ will be given by the appropriate spectral moment of order 2K. Indeed, one can easily derive the following general result:

$$\left|\frac{\partial^{\alpha+\beta}X(\mathbf{s})}{\partial\mathbf{s}_i^{\alpha}\partial\mathbf{s}_j^{\beta}}.\frac{\partial^{\vartheta+\delta}X(\mathbf{s})}{\partial\mathbf{s}_i^{\vartheta}\delta\mathbf{s}_j^{\delta}}\right| = \left|\frac{\partial^{\alpha+\beta+\vartheta+\delta}X(\mathbf{s})}{\partial\mathbf{s}_i^{\alpha}\partial\mathbf{s}_j^{\beta}\partial\mathbf{s}_k^{\vartheta}\mathbf{s}_l^{\delta}}\gamma(\mathbf{s})\right|_{\mathbf{s}=0}$$

$$= (-1)^{\alpha+\beta_j\alpha+\beta+\vartheta+\delta}\int_{\mathbf{s}}\lambda_i^{\alpha}\lambda_j^{\beta}\lambda_k^{\vartheta}\lambda_l^{\delta}dF(\lambda).$$

Combining this with (1.17) it follows from an appropriate choice of $\alpha,\beta,\vartheta,\delta$ for a real-valued stationary field $X(\mathbf{s})$ for a real-valued stationary field all $1 \le i,k,l \le d$:

$$X_i(\mathbf{s}) \text{ and } X(\mathbf{s}) \text{ are uncorrelated} \quad \alpha = 1, \beta = \vartheta = \delta = 0, \tag{1.21}$$

$$X_i(\mathbf{s}) \text{ and } X_{kl}(\mathbf{s}) \text{ are uncorrelated} \quad \beta = 0, \alpha = \vartheta = \delta = 1. \tag{1.22}$$

If $X(\mathbf{s})$ is a Gaussian field than so are its mean square derivatives. (This is a simple consequence of the fact that derivatives are defined by taking limits of linear combinations of Gaussian variates. See equation (1.9)). Furthermore, the joint distribution of $X(\mathbf{s})$ with its derivatives is multivariate Gaussian. Thus in this case the variables considered in (1.21) and (1.22) are not only uncorrelated but also independent.

We conclude this section with a brief discussion of the heuristic meaning of the spectral representation of a random field. Firstly, note that if $X(\mathbf{s})$ is real valued we can rewrite (1.18) as

$$X(\mathbf{s}) = \int_{\mathbb{R}^d}\cos < \mathbf{s},\lambda > dU(\lambda) - \int_{\mathbb{R}^d}\sin < \mathbf{s},\lambda > dV(\lambda), \tag{1.23}$$

where $U(\mathbf{s})$ and $V(\mathbf{s})$ are real valued fields defined by

$$Z(\lambda) = U(\lambda) + iV(\lambda).$$

That is, U and V are the real and imaginary components of Z. Now approximate the integral in (1.24) by a sum of the form

$$X(\mathbf{s}) = \sum(\cos < \mathbf{s},\lambda_i > U(\Lambda_i) - \sin < \mathbf{s},\lambda_i > V(\Lambda_i)), \tag{1.24}$$

where the λ_i are disjoint intervals in $\mathbb{R}^d$ and $\lambda_i \in \Lambda_i, i = 1,2,\ldots$. Then it is clear from this approximation that the spectral representation of a random field is effectively a method for breaking it up into a large number of sinusoidal components.

In the one-dimensional situation the basic components in (1.24) are simple sine and cosine waves of random amplitudes $-V(\Lambda_i)$ and $U(\Lambda_i)$, respectively, and wavelengths equal to $2\pi/\lambda_i$. In higher dimensions the elementary components are harder to visualize. Consider the two-dimensional case. Dropping the subscript on $\boldsymbol{\lambda}_\mathbf{i}$ for the moment we have that an elementary cosine wave is of the form $cos(\lambda_1 s_1 + \lambda_2 t_2)$. The λ_i are fixed, and the point (s_1, s_2) takes values in $\mathbb{R}^d$. Elementary trigonometry shows that such a function actually forms a wave-pattern, in $\mathbb{R}^d$ of a sequence of waves travelling in a direction which makes an angle

$$\theta = \tan^{-1}\left(\frac{\lambda_1}{\lambda_2}\right)$$

with the t_1 axis, and having the wavelength

$$\lambda = \frac{2\pi}{\lambda_1^2 + \lambda_2^2}$$

as the distance between troughs or crests, as measured along the line perpendicular to the crests. A pictorial representation of such a surface is given in Figure 1.3.4

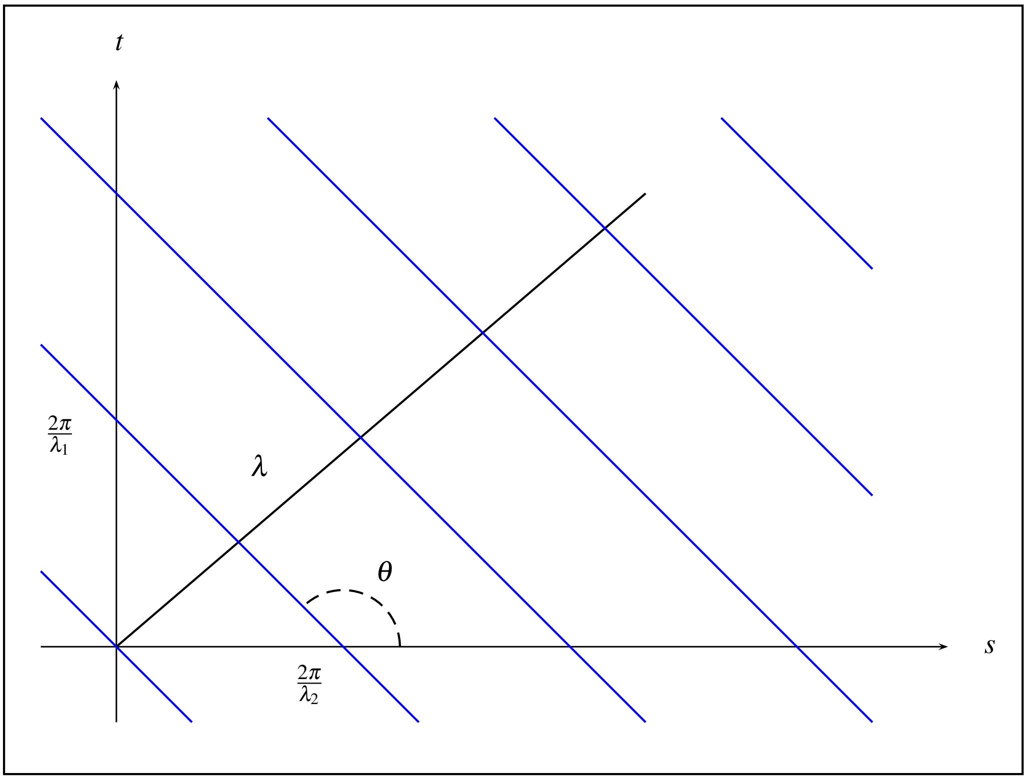


Figure 1.1: Elementary wave form $\cos(\lambda_1 t_1 + \lambda_2 t_2)$ in $\mathbb{R}^2$.

This concludes our study of the fundamental aspects of the spectral representation for random fields.

1.4 Orthogonal Decompositions of 2-D Random Fields

The discussion hereafter will be restricted to the case of real discrete 2-D random fields, $X = \{X(s,t),\ (s,t) \in \mathbb{Z}^2\}$. Let H be a Hilbert space (e.g, $H = L^2(\Omega, F, P)$) the space of all random variables on (Ω, F, P) with finite second order moments and zero mean, endowed by the inner product $< X, Y >= E(XY)$ with norm $\|X\| = \sqrt{EX^2}$ (that is, by convergence we will always mean convergence in mean square). The collection of all finite linear combinations of elements in the space and its closure are evenly included in the space [31].

The structure of 1-D stationnary process was rigorously formulated by the seminal results of Wiener-Kolmogorov prediction theory [32, 33] and [34]. The analysis of 2-D random fields is carried out by formulating the 2-D linear prediction problem in a manner similar to the one which is broadly used for the analysis of 1-D discrete random processes. However, there is no natural definition of the past and future when the index of the process is in two dimensions. In fact, the simplest way to extend the theory of the Wiener-Kolmogorov prediction to the 2-D case is to choose a past for which the prediction error is a white noise according to which the process must be able to be represented. Helson and Lowdenslager [12,13] have shown that pasts with this properly correspond to asymmetric half-planes. It should be mentioned that other types of past have been considered, the case of quarter-planes were studied in [15–17,35].

1.4.1 Nonsymmetrical half-plane orthogonal decomposition

The totally ordered NSHP support is a favorable type of support in the sense that it yields a natural extension to the 1-D results, this fact has been shown in [12, 13]. In what follows, we assume all definitions and theorems are stated with respect to the total order and nonsymmetrical half-plane (NSHP); and the 2-D random field $X(s,t)$. We first introduce some basic definitions related to the NSHP prediction scheme. Next, we state the 2-D Wold-type orthogonal decomposition theorem.

Definition 1.4.1 ([12]). *We call a nonsymmetrical half-plan past (NSHP) any subset S of $\mathbb{Z}^2$ satisfying*

1. *S stable under addition*
2. *$S \cup -S = \mathbb{Z}^2$*
3. *$S \cap -S = \{(0,0)\}$.*

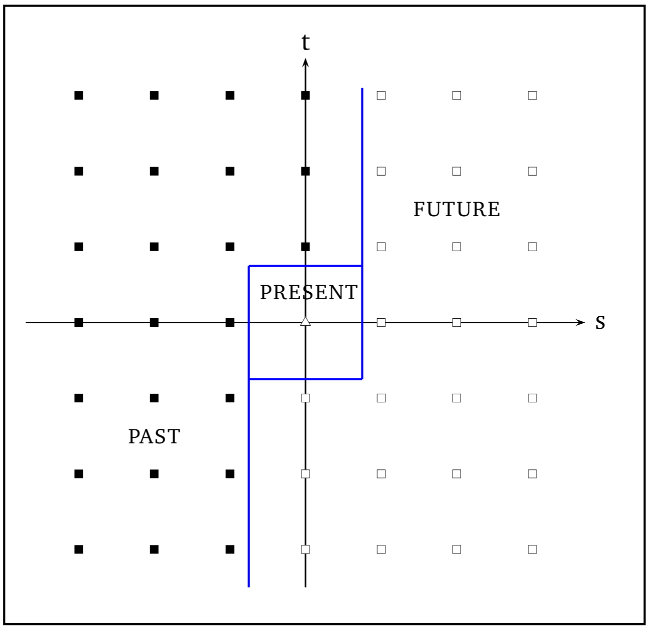


Figure 1.2: NSHP Total Order Definition

The NSHP support, S, results from the total-order definition (1.25a). The total order is defined as a raster-scan : top-to-bottom column-after-column from left to right. In mathematical terms: The order $<$ is defined by

$$(s_1,t_1) < (s_2,t_2) \text{ iff } (s_1,t_1) \in \{(k,l)|k = s_2, l < t_2\} \cup \{(k,l)|k < s_2, -\infty < l < \infty\}, \tag{1.25a}$$

and the order $\leq$ is naturally defined by

$$(s_1,t_1) \leq (s_2,t_2) \text{ iff } (s_1,t_1) < (s_2,t_2) \text{ or } (s_1,t_1) = (s_2,t_2), \tag{1.25b}$$

which coincides with the lexicographic order. Figure 1.2 illustrates this support graphically.
Because of the infinity in one of the scanning axes, we have an infinite number of samples between any two samples that are in different columns. We now present the definitions and the main theorems which result in the orthogonal decomposition of the 2-D random field. In the sequel we shall assume the 2-D random field $X(s,t)$ to be real, with zero mean. We shall also assume that the random field has finite second-order moments, i.e.,

$$\sup_{(s,t)\in\mathbb{Z}^2} E\left(X^2(s,t)\right) < \infty$$

and that $E\left(X^2(s,t)\right) > 0$ for at least one $(s,t) \in \mathbb{Z}^2$, i.e. $X \in H$. In what follows, the problem of prediction is developed in the case of a past defined by an nonsymmetrical half-plane resulting from the total order definition (1.25a), $S = \{(k,l) \in \mathbb{Z}^2, (k,l) < (s,t)\}$. We define $\mathscr{H}^X_{(s,t)}$ as the closed linear subspace spanned by $\{X(k,l), (k,l) \in S\}$ in the Hilbert space H, which represents the strict past of X at the point (s,t). We denote by $P_{\mathscr{H}^X}$ the orthogonal projection operator onto $\mathscr{H}^X_{(s,t)}$, i.e, $P_{\mathscr{H}^X_{(s,t)}} X(s,t))$ one-step minimum norm causal continuous support, linear predictor of $X(s,t)$. The innovation process $\{\varepsilon(s,t), (s,t) \in \mathbb{Z}^2\}$ with respect to the defined support and total order is given by

$$\varepsilon(s,t) = X(s,t) - P_{\mathscr{H}^X_{(s,t)}} X(s,t)).$$

Nonsymmetrical half-plane orthogonal decomposition of non-stationary random fields

We state the 2-D Wold decomposition theorem. This theorem is generalization of the well known 1-D Wold decomposition [11]. The 2-D Wold-like decomposition theorem was stated by

Theorem 1.4.2 ([12, 13]). *If $X(s,t)$, $(s,t) \in \mathbb{Z}^2$ is a regular random field, it can be represented uniquely by the following orthogonal decomposition:*

$$X(s,t) = W(s,t) + V(s,t), \tag{1.26}$$

where

$$w(s,t) = \sum_{k\geq 0}\sum_{l\geq 0} b_{s,t}(k,l)\varepsilon(s-k,t-l) \tag{1.27}$$

$b_{k,l}(m,n)$ is given by

$$b_{k,l}(m,n) = \begin{cases} \frac{E(X(k,l)\varepsilon(k-m,l-n))}{E(\varepsilon^2(k-m,l-n))}, & \text{if } E(\varepsilon^2(k-m,l-n)) > 0, \\ 0, & \text{if } E(\varepsilon^2(k-m,l-n)) = 0, \end{cases}$$

where if $E(\varepsilon^2(k-m,l-n)] = 0$, $b_{k,l}(m,n)$ is arbitrarily set to zero. Also,

(a) $\sum_{k\geq 0}\sum_{l\geq 0} a_{s,t}^2(k,l)E(\varepsilon^2(s-k,t-l)) < \infty.$

(b) $E(V(s,t)) = 0.$

(c) $E(\varepsilon(s,t)\varepsilon(m,n)) = 0, \quad (s,t) \neq (m,n).$

(d) $E(\varepsilon(s,t)V(m,n)) = 0, \quad \forall (s,t), (m,n).$

(e) $W(s,t) \in H_{(s,t)}^X$

(f) $V(s,t) \in H_{(s,-\infty)}^X$ *where the Hilbert space* $H_{(s,-\infty)}^X$ *is defined by* $H_{(s,-\infty)}^X = \bigcap_{t=-\infty}^{\infty} H_{(s,t)}^X$

(g) *If for all* $(s,t) \in \mathbb{Z}^2$, $E\left((\varepsilon(s,t))^2\right) > 0$, *the sequences* $\{\varepsilon(s,t)\}$ *and* $\{b_{(s,t)}\}$*are unique, i.e. there is only one 2-D sequence of random variables* $\{\varepsilon(s,t)\}$ *and only one 2-D sequence of constants* $\{b_{(s,t)}\}$ *satisfying the previously stated results. However, if there are* $(s,t) \in \mathbb{Z}^2$ *such that* $E\left((\varepsilon(s,t))^2\right) = 0$, *the uniqueness of the sequence* $\{b_{(s,t)}\}$ *is achieved by the arbitrary setting of the corresponding elements of the sequence* $\{b_{(s,t)}\}$ *to zero.*

From **(c)** we can immediately say that the corresponding innovation field $\{\varepsilon(s,t),(s,t) \in \mathbb{Z}^2\}$ defined by $\varepsilon(s,t) = X(s,t) - P_{\mathscr{H}_{(s,t)}^X} X(s,t)$ is a white noise field, whereas **(d)** implies that random fields are mutually orthogonal for all (s,t) and (m,n). In mathematical terms: $W(s,t) \perp V(m,n)$ for all (s,t) and (m,n).

Proof: We first prove **(c)**. From the orthogonal projection theorem, $\varepsilon(s,t)$ is orthogonal to every vector in $\mathscr{H}_{(s,t)}^X$, we deduce that $\varepsilon(n,m) \perp H_{(s,t)}^X$ for all $(n,m) < (s,t)$. Because $\varepsilon(s,t) \in H_{(s,t)}^X$ for all (s,t), we conclude that $\varepsilon(n,m) \perp \varepsilon(s,t)$ for all $(s,t) < (n,m)$. By interchanging the role of indices we also have that $\varepsilon(s,t) \perp \varepsilon(n,m)$ for all $(n,m) < (s,t)$, and this completes the proof of **(c)**.

We shall now prove **(a)**. Let the support S' be defined by $S' = S_{N,M} \cup \{0;0)\}$, where
$S_{N,M} = \{(k,l), k = 0,\ 0 < l < M\} \cup \{(k,l), k = 0,\ -M < l < M\}$
We shall at the following expression:

$$\begin{aligned}
0 &< E\left(X(s,t) - \sum_{S'}\sum_{S'} b_{k,l}\varepsilon(s-k,t-l)\right)^2 \\
&= E\left(X(s,t)^2\right) - 2\sum_{S'}\sum_{S'} b_{k,l}E(X(s,t)\varepsilon(s-k,t-l)) \\
&\quad + \sum_{S'}\sum_{S'} b_{k,l}b_{n,m}E(\varepsilon(s-k,t-l)\varepsilon(s-n,t-m)) \\
&= E\left(X(s,t)^2\right) - 2\sum_{S'}\sum_{S'} b_{k,l}^2 E(\varepsilon(s-k,t-l)^2) + \sum_{S'}\sum_{S'} b_{k,l}^2 E\left(\varepsilon(s-k,t-l)^2\right) \\
&= E\left(X(s,t)^2\right) - \sum_{S'}\sum_{S'} b_{k,l}^2 E\left(\varepsilon(s-k,t-l)^2\right).
\end{aligned}$$

By assumption $\sup_{(s,t)\in\mathbb{Z}^2} E\left(X(s,t)^2\right) < \infty$. We therefore conclude that

$$\sum_{S'}\sum_{S'} b_{k,l}^2 E[\varepsilon^2(s-k,t-l)] \leq \sup_{(s,t)\in\mathbb{Z}^2} E\left(X(s,t)^2\right) < \infty.$$

This sum is bounded for any N and M by an expression which is neither a function of N nor M. Therefore, the positive series converges since the sequence of its partial sums is bounded. This completes the proof of **(a)**.

From (1.26) $W(s,t)$ is in the linear manifold spanned by $\varepsilon(i,j)$ such that $(i,j) \leq (s,t)$. Because $\varepsilon(n,m) \in H^X_{(n,m)}$ for all (n,m) and $H^X_{(n,m)} \subset H^X_{(s,t)}$ for all $(n,m) \leq (s,t)$, $W(s,t)$ is a linear combination of elements in $H^X_{(s,t)}$. Since its second moment is finite, $W(s,t) \in H^X_{(s,t)}$ as stated in **(e)**.

The proof of **(b)** follows immediately from the definition of $V(s,t)$ as

$$V(s,t) = X(s,t) - W(s,t) \tag{1.28}$$

We shall now turn to prove **(d)**. For every $(n,m) \leq (s,t)$ such that $E\left(\varepsilon(s,t)^2\right) > 0$ we have

$$E(V(s,t)\varepsilon(n,m)) = E(X(s,t)\varepsilon(n,m)) - \sum_{k>0}\sum_{l>0} b_{k,l} E(\varepsilon(s-k,t-l)\varepsilon(n,m))).$$

Since

$$\sum_{k>0}\sum_{l>0} b_{k,l} E[\varepsilon(s-k,t-l)\varepsilon(n,m)) = b_{s-n,t-m} E\left((\varepsilon(n,m))^2\right),$$

we have that $E[V(s,t)\varepsilon(n,m)] = 0$.

For the case in which $E\left((\varepsilon(n,m))^2\right)$, we have using the Cauchy-Schwarz inequality that

$$0 \leq |E(V(s,t)\varepsilon(n,m))|^2 \leq E\left((V(s,t))^2) E((\varepsilon(n,m))^2\right) = 0.$$

Hence, $E(V(s,t)\varepsilon(n,m)) = 0$ in this case as well.

For $(s,t) \leq (n,m)$ we have by using (1.28), and since both $W(s,t) \perp \mathscr{H}^X_{(s,t)}$ and $X(s,t) \in H^X_{(s,t)}$, that $V(s,t) \in H^X_{(s,t)}$. Since $\varepsilon(s,t) \perp \mathscr{H}^X_{(s,t)}$ and since $H^X_{(s,t)} \subset \mathscr{H}^X_{(n,m)}$ whenever $(s,t) < (n,m)$, we have that for every $(s,t) < (n,m)$, $\varepsilon(n,m) \perp V(s,t)$. Combining the two cases, we conclude that for every two pairs of indices (n,m) and (s,t), $\varepsilon(s,t) \perp V(n,m)$.

In order to prove **(f)**, define $\overline{sp}\{\varepsilon(s,t)\}$ as the subspace of $H^X_{(s,t)}$ spanned by the vector $\varepsilon(s,t)$. From the orthogonal projection theorem, $\varepsilon(s,t) \perp \mathscr{H}^X_{(s,t)}$ and therefore $H^X_{(s,t)} = \mathscr{H}^X_{(s,t)} \oplus \varepsilon(s,t)$. Since $V(s,t) \perp \varepsilon(s,t)$ and $V(s,t) \in H^X_{(s,t)}$ it follows that $V(s,t) \in \mathscr{H}^X_{(s,t)}$. By induction $V(s,t) \in H^X_{(s,-\infty)}$

Let us now prove **(g)**. From the orthogonal projection theorem it follows that $\varepsilon(s,t)$ is unique. This holds for every (s,t) and therefore the field $\{\varepsilon(s,t)\}$ is unique.

If for all $(s,t) \in \mathbb{Z}^2$, $E\left((\varepsilon(n,m))^2\right) > 0$, then since for every (s,t) and (n,m) such that $(n,m) \geq (0,0)$, $X(s,t)$ and $\varepsilon(s-n,t-m)$ are elements in the Hilbert space $H^X_{(s,t)}$ where the inner product is defined as $E[xy]$, the uniqueness of $\varepsilon(s,t)$ implies the uniqueness of $b_{k,l}$. However, if there are $(s,t) \in \mathbb{Z}^2$, such that $\left\{\varepsilon(n,m))^2\right\}$, the uniqueness of the sequence $b_{s,t}$ may be achieved by the arbitrary setting of the corresponding elements of the sequence $b_{s,t}$ to zero. $\square$

Properties Of The 2-D Wold-Like Decomposition

In order to further investigate the properties of the decomposition components we need additional definitions.

Definition 1.4.3. *A field $\{X(s,t)\}$ is deterministic if for all (s,t). $E[X(s,t) - P_{\mathscr{H}^X_{(s,t)}} X(s,t)]^2 = \sigma^2 = 0$. This means that for all (s,t), $\{X(s,t)\}$ can be perfectly predicted as a linear combination of elements of its past (or as a limit of such), i.e., elements of $H^X_{(s,t)}$. Note that the deterministic field is a random field and it is deterministic only in the mean square sense.*

Definition 1.4.4. *A regular field $\{X(s,t)\}$ is called purely indterministic if for all $(s,t), H^X_{(s,t)} = H^{\varepsilon}_{(s,t)}$, i.e, if its deterministic component $V(s,t)$ vanishes, so that $\{X(s,t)\}$ can be represented completely by the white innovation driven moving average term in equation* (1.26)*:*

$$\sum_{k\geq 0}\sum_{l\geq 0} b(s,t)(k,l)\varepsilon(s-k,t-l).$$

For any integers k, l, we define the subspaces of Hilbert space H :

$$\begin{aligned} H^{\varepsilon}_{(s,t)} &= \overline{sp}\{\varepsilon(k,l), (k,l)\in \mathbf{S}\}, \quad \mathscr{H}^{\varepsilon}_{(s,t)} = \overline{sp}\{\varepsilon(k,l), (k,l)\in S\};\\ H^{W}_{(s,t)} &= \overline{sp}\{W(k,l), (k,l)\in \mathbf{S}\}, \quad \mathscr{H}^{\varepsilon}_{(s,t)} = \overline{sp}\{W(k,l), (k,l)\in S\};\\ H^{V}_{(s,t)} &= \overline{sp}\{V(k,l), (k,l)\in \mathbf{S}\}, \quad \mathscr{H}^{V}_{(s,t)} = \overline{sp}\{V(k,l), (k,l)\in S\}; \end{aligned}$$

where $\mathbf{S}$ and S are defined as follows $\mathbf{S} = \{(k,l)\in \mathbb{Z}^2, (k,l)\leq (s,t)\}$, $S = \{(k,l)\in \mathbb{Z}^2, (k,l)\leq (s,t)\}$.

Theorem 1.4.5. *Let $\{X(s,t)\}$ be a 2-D regular random field. Its component $W(s,t)$ is purely indeterministic and regular.*

Proof: Let us rewrite (1.27) as

$$W(s,t) = b_{0,0} + \sum_{k>0}\sum_{l>0} b_{k,l}\varepsilon(s-k,t-l). \tag{1.29}$$

If $E\left((\varepsilon(n,m))^2\right) = 0$, then $b_{0,0} = 1$. We have already proved that $W(s,t)$ and $\varepsilon(s,t)\in H^X_{(s,t)}$, that $\varepsilon(s,t)\perp \mathscr{H}^X_{(s,t)}$, and that $\sum_{k>0}\sum_{l>0} b_{k,l}\varepsilon(s-k,t-l)\in \mathscr{H}^X_{(s,t)}$. Therefore, the orthogonal projection theorem and the uniqueness of both the projection and the residual, together with the above representation of $W(s,t)$, imply that $P_{\mathscr{H}^X_{(s,t)}}W(s,t)$, which is the projection of $W(s,t)$ on $\mathscr{H}^X_{(s,t)}$, is given by

$$P_{\mathscr{H}^X_{(s,t)}}W(s,t) = \sum_{k>0}\sum_{l>0} b_{k,l}\varepsilon(s-k,t-l). \tag{1.30}$$

Clearly (1.30) holds also if $E\left((\varepsilon(s,t))^2\right) = 0$, since in that case $W(s,t) = P_{\mathscr{H}^X_{(s,t)}}W(s,t)$, and both are elements of $\mathscr{H}^X_{(s,t)}$.

In order to prove that $\{W(s,t)\}$ is a purely indeterministic random field we show that $H^W_{(s,t)} = H^{\varepsilon}_{(s,t)}$. As we have $W(s,t)$ is a linear combination of the elements $\varepsilon(k,l)$ where $(k,l)\leq (s,t)$, Therefore $H^W_{(s,t)}\subset H^{\varepsilon}_{(s,t)}$. On the other hand, $P_{\mathscr{H}^X_{(s,t)}}W(s,t)\in \mathscr{H}^X_{(s,t)}$. Hence, there exists a sequence of constants $\{c_{k,l}\}$ such that $P_{\mathscr{H}^X_{(s,t)}}W(s,t)$ is represented by

$$P_{\mathscr{H}^X_{(s,t)}}W(s,t) = \sum_{k>0}\sum_{l>0} c_{k,l}X(s-k,t-l), \tag{1.31}$$

or by a limit of such expression. Using (1.26), we can rewrite (1.31) as

$$P_{\mathscr{H}^X_{(s,t)}}W(s,t) = \sum_{k>0}\sum_{l>0} c_{k,l}W(s-k,t-l) + \sum_{k>0}\sum_{l>0} c_{k,l}V(s-k,t-l). \tag{1.32}$$

From (1.30), $P_{\mathscr{H}^X_{(s,t)}}W(s,t)\in \mathscr{H}^{\varepsilon}_{(s,t)}$. Also, since $\mathscr{H}^W_{(s,t)}\subset \mathscr{H}^{\varepsilon}_{(s,t)}$, we have that $\sum_{k>0}\sum_{l>0} c_{k,l}W(s-k,t-l)\in \mathscr{H}^{\varepsilon}_{(s,t)}$. Theorem 1.4.2 **(d)** implies that for all $(0,0)<(k,l)$, $V(s-k,t-l)\perp \mathscr{H}^{\varepsilon}_{(s,t)}$. Hence, (1.32) holds if and only if for all $(0,0)<(k,l)$, $V(s-k,t-l) = 0$. Therefore, $P_{\mathscr{H}^X_{(s,t)}}W(s,t) = \sum_{k>0}\sum_{l>0} c_{k,l}W(s-$

$k, t-l)$. This implies that $P_{\mathscr{H}^X_{(s,t)}} W(s,t) \in \mathscr{H}^W_{(s,t)} \subset H^W_{(s,t)}$. From (1.29), $\varepsilon(s,t) = W(s,t) - P_{\mathscr{H}^X_{(s,t)}} W(s,t)$ and therefore $\varepsilon(s,t) \in H^W_{(s,t)}$ for all (s,t), so that $H^\varepsilon(s,t) \subset H^W_{(s,t)}$, we finally conclude that $H^\varepsilon_{(s,t)} = H^W_{(s,t)}$. Since $P_{\mathscr{H}^X_{(s,t)}} W(s,t) \in H^W_{(s,t)}$, we conclude that $\{\varepsilon(s,t)\}$ is the innovation field of $\{W(s,t)\}$, as well. Therefore if $\{X(s,t)\}$ is a regular field, then $\{W(s,t)\}$ is also a regular field. □

Corollary 1.4.6. *$H^X(s,t)$ has a direct sum representation of*

$$H^X_{(s,t)} = H^\varepsilon_{(s,t)} \oplus H^V_{(s,t)}. \tag{1.33}$$

Proof: The definition of $\{X(s,t)\}$ (1.26), and Theorem (1.4.2) (**d**) imply that $V(s,t) \perp \varepsilon(s,t)$ for all (s,t) and (n,m). By Theorem (1.4.2), for all (s,t), $X(s,t)$ can be represented uniquely as $X(s,t) = W(s,t) + V(s,t)$, where $W(s,t) \in H^\varepsilon_{(s,t)}$ and $V(s,t) \in H^V_{(s,t)}$. Since the two subspace $H^\varepsilon_{(s,t)}$ and $H^V_{(s,t)}$ are orthogonal, it follows that $H^X_{(s,t)} = H^\varepsilon_{(s,t)} \oplus H^V_{(s,t)}$.

□

Theorem 1.4.7. *Let $\{X(s,t)\}$ be a 2-D regular random field. Its component $\{V(s,t)\}$ is a deterministic random field.*

Proof: The direct sum representation(1.33), implies in particular that $\mathscr{H}^X_{(s,t)} = \mathscr{H}^\varepsilon_{(s,t)} \oplus \mathscr{H}^V_{(s,t)}$. By Theorem 1.4.2 (**f**), $V(s,t) \in H^X_{(s,-\infty)} \subset \mathscr{H}^X_{(s,t)}$. Since $V(s,t) - \varepsilon(n,m)$ for all (s,t) and (n,m), it follows that $V(s,t) \perp \mathscr{H}^\varepsilon_{(s,t)}$. Finally, because $V(s,t) \in \mathscr{H}^\varepsilon_{(s,t)}$ and $V(s,t) \perp \mathscr{H}^\varepsilon_{(s,t)}$, we conclude that $V(s,t) \in \mathscr{H}^V_{(s,t)}$, i.e., $\{V(s,t)\}$ is a deterministic random field. □

Nonsymmetrical half-plane orthogonal decomposition of stationary random fields

We shall restrict our attention to the class of stationary random finds. Clearly, The results obtained above are applicable to this special class of 2D random fields. For stationary random fields, frequency domain analysis is applicable since there exists for these fields a spectral representation in the form of a Fourier-Stieltjes integral, both for the field variables and for the associated covariance functions. In the following, we For stationary 2-D random fields frequency domain analysis is applicable due to the existence of spectral representations both for the field variables and for the associated covariance functions.

Definition 1.4.8. *A family of real random variables $\{X(s,t)\}$, $(s,t) \in \mathbb{Z}^2$ is called a second order stationary random field. If*

$$\sup_{(s,t)\in\mathbb{Z}^2} E\left(X^2(s,t)\right) < \infty$$

and for all integers s_1, s_2, t_1 and t_2, the covariance of the $X(s_1,t_1)$ and $X(s_2,t_2)$ depends on the lags$(s_1 - s_2, t_1 - t_2)$, namely,

$$cov(X(s_1,t_1), X(s_2,t_2)) = \gamma(s_1 - s_2, t_1 - t_2)$$

We first describe these spectral representations. From (1.5) of Theorem 1.3.3, there exists a unique spectral distribution function $F(.,.)$ on the rectangular region $\left[-\frac{1}{2}, \frac{1}{2}\right] \times \left[-\frac{1}{2}, \frac{1}{2}\right]$ such that

$$\gamma(s,t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i(s\lambda_1 + t\lambda_2)} dF(\lambda_1, \lambda_2); \quad (s,t) \in \mathbb{Z}^2.$$

From (1.18) of Theorem 1.45, there exists a spectral representation of the field $\{X(s,t)\}$ such that:

$$X(s,t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i(s\lambda_1 + t\lambda_2)} dZ(\lambda_1, \lambda_2);$$

where $Z(\lambda_1,\lambda_2)$ is a doubly orthogonal increment process, i.e.,

$$E\left[dZ(\lambda_1,\lambda_2)\, dZ'(\lambda_3,\lambda_4)\right] = 0, \quad \lambda_1 \neq \lambda_3,\ \lambda_2 \neq \lambda_4,$$

which is related to $F(\lambda_1,\lambda_2)$ by

$$dF(\lambda_1,\lambda_2) = E[dZ(\lambda_1,\lambda_2)\, dZ'(\lambda_3,\lambda_4)].$$

$f(\lambda_1,\lambda_2)$ is the corresponding spectral density function, which is the Lebesgue 2-D derivative of $F(\lambda_1,\lambda_2)$

$$f(\lambda_1,\lambda_2) = \frac{\partial^2 F(\lambda_1,\lambda_2)}{\partial\lambda_1\,\partial\lambda_2}. \tag{1.34}$$

Let $F^s(\lambda_1,\lambda_2)$ be the singular part of $F(\lambda_1,\lambda_2)$ (with respect to the singularity of the measure associated with $F(\lambda_1,\lambda_2)$ and the Lebesgue measure). Note that for real valued random fields $\gamma(k,l) = \gamma(-k,-l)$ and hence $F(\lambda_1,\lambda_2) = F(-\lambda_1,-\lambda_2)$.

Theorem 1.4.9 ([12,13])**.** *A stationary 2-D random field $\{X(s,t)\}$ is regular i.e, purely non deterministic (PND) if and only if $f(\lambda_1,\lambda_2) > 0$ almost everywhere in K (Lebesgue measure) and*

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \log f(\lambda_1,\lambda_2) d\lambda_1 d\lambda_2 > -\infty. \tag{1.35}$$

If $\{X(s,t)\}$ is a stationary 2-D random field, then $\{\varepsilon(s,t)\}$ is wide sense stationary, and so are $\{w(s,t)\}$, $\{v(s,t)\}$. variance of the innovation field $\{\varepsilon(s,t)\}$ is constant and it will be denoted by σ^2. We then say that the field $\{X(s,t)\}$ is if non deterministic $E[X(s,t) - \widehat{X}(s,t)]^2 = \sigma^2 > 0$, in other words a discrete 2-D random field is regular if its innovation field $\varepsilon(s,t)$ does not vanish. Contrary to the non-stationary case, the innovations variance of a regular stationary 2-D random field is strictly positive and equals σ^2 for all $(s,t) \in \mathbb{Z}^2$. Therefore, the 2-D Wold decomposition for stationary random fields is unique, whereas in the case non-stationary this uniqueness is not always guaranteed.

Theorem 1.4.10 ([12, 13])**.** *Let $\{X(s,t),(s,t) \in \mathbb{Z}^2\}$ be a regular field. The variance of the variance of the innovations field $\{\varepsilon(s,t)\}$ with respect to the defined support S and the total order, is given by*

$$\sigma^2(S) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \log f(\lambda_1,\, \lambda_2) d\lambda_1 d\lambda_2.$$

In the regular case the variance of the innovations field $\{\varepsilon(s,t)\}$ is given by

$$\sigma^2 = e^{\left(\int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \log f(\lambda_1,\lambda_2) d\lambda_1 d\lambda_2\right)} \tag{1.36}$$

A stationary 2-D random field $\{X(s,t)\}$ is deterministic if an only if

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} \log f(\lambda_1,\lambda_2) d\lambda_1 d\lambda_2 = -\infty, \tag{1.37}$$

σ^2 should be interpreted as zero if $\int\int \log f(\omega,\lambda_2) d\omega d\lambda_2 = -\infty$.

The decomposition of the regular random field into two mutually orthogonal components, the purely indeterministic field and the deterministic field (Theorems 1.4.2, 1.4.5), results in the following spectral decomposition.

Theorem 1.4.11 ([12, 13]). *Let $F_W(\lambda_1,\lambda_2)$ be the spectral distribution function of the purely indeterministic component of a regular and homogeneous random field, and let $F_V(\lambda_1,\lambda_2)$ be the spectral distribution function of the deterministic component. The spectral distribution function of $F(\lambda_1,\lambda_2)$ the regular field is uniquely represented by $F(\lambda_1,\lambda_2) = F_W(\lambda_1,\lambda_2) + F_V(\lambda_1,\lambda_2)$, where the spectral distribution function $F_W(\lambda_1,\lambda_2)$ of the purely indeterministic component $W(s,t)$ absolutely continuous and $F_V(\lambda_1,\lambda_2) = F^s(\lambda_1,\lambda_2)$ Hence, the spectral measure induced by $F_V(\lambda_1,\lambda_2)$ is singular with respect to the Lebesgue measure, and therefore it is concentrated on a Borel set L of Lebesgue measure zero in $K = \left[-\frac{1}{2},\frac{1}{2}\right] \times \left[-\frac{1}{2},\frac{1}{2}\right]$. The derivative of $F_V(\lambda_1,\lambda_2)$ is zero, except on the set L. The spectral representations of the purely indeterministic and deterministic random fields are given by*

$$
\begin{aligned}
W(s,t) &= \int_{K|L} e^{i(k\lambda_1+l\lambda_2)} dZ(\lambda_1,\lambda_2), \\
V(s,t) &= \int_{L} e^{i(k\lambda_1+l\lambda_2)} dZ(\lambda_1,\lambda_2).
\end{aligned}
\tag{1.38}
$$

Since the random field is stationary, the coefficients a_{kl} in the moving-average representation (**??**), of the purely indeterministic component are space invariant, i.e., for all (s,t), $a_{st}(k,l) = a(k,l)$. Hence (**??**) implies that $f(\lambda_1,\lambda_2)$ is given by

$$
f(\lambda_1,\lambda_2) = \sigma^2 \left| \sum_{k\geq 0} \sum_{l\geq 0} a(k,l) e^{i(k\lambda_1+l\lambda_2)} \right|^2 .
$$

We therefore conclude that a necessary and sufficient condition for a stationary random field $\{X(s,t)\}$ to be purely indeterministic is that $F(\lambda_1,\lambda_2)$ is absolutely continuous, $f(\lambda_1,\lambda_2) > 0$ almost everywhere in K (Lebesgue measure), and $\log f(\lambda_1,\lambda_2)$ is Lebesgue integrable. At the other extreme, if (1.35) is false the stationary random field $\{X(s,t)\}$ is nonregular and deterministic even if its spectral distribution function is absolutely continuous. Hence for example, any band-limited 2-D process is nonregular and deterministic. However, it is interesting to observe that the spectral measure of the deterministic component resulting from the orthogonal decomposition of a regular field (Theorems 1.4.2, 1.4.5) is always concentrated on a set of Lebesgue measure zero in K and it is therefore a special case of the class of nonregular fields.

Thus, the decomposition of the stationary regular field into purely indeterministic and deterministic components corresponds in terms of spectral measures to the decomposition (1.38) of the spectral measure of the regular field into a sum of two mutually singular spectral measures. In terms of spectral distributions it corresponds to the representation of F as the sum of its absolutely continuous and singular components, F_W and F_V respectively.

1.4.2 Multiple order definitions and the decomposition of the deterministic component

The NSHP support definition which results from the total order definition in (1.25) is not the only possible one of that type on the 2-D lattice. In the following we generalize the results derived based on the total order definition (1.25), by introducing a family of total-order-and NSHP-support definitions in which the boundary line of the NSHP is of rational slope (These order definitions are obtained by rotating the NSHP support by angles of rational tangent, rather than considering only the vertical and horizontal orientations). Note that it is only the total order imposed on the random field that is changed, but not the 2-D discrete grid itself.

Definition 1.4.12. *Let α, β be two coprime integers, such that $\alpha \neq 0$. Let us define a new total-order and NSHP support by rotating the NSHP support which was defined with respect to* (1.25)*, through a counterclockwise angle about the origin of its coordinate system, such that* $\tan\theta = \frac{\beta}{\alpha}$.

Let the coordinates (s^*, t^*) be defined by

$$\begin{pmatrix} s^* \\ t^* \end{pmatrix} = \begin{pmatrix} \sqrt{\alpha^2+\beta^2} & 0 \\ 0 & 1/\sqrt{\alpha^2+\beta^2} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix},$$

where (s,t) are the sample original coordinates, and

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad \begin{pmatrix} \sqrt{\alpha^2+\beta^2} & 0 \\ 0 & 1/\sqrt{\alpha^2+\beta^2} \end{pmatrix}$$

are the rotation transformation matrix, and the normalization matrix, respectively. The normalization matrix is such that the indices of the columns are consecutive integers and the distance between two neighboring samples on the same column is one. Thus, the new coordinates $(s^{\alpha,\beta}, t^{\alpha,\beta})$ of the original (s,t)th sample are given by

$$s^{\alpha,\beta} = s^*,$$
$$t^{\alpha,\beta} = t^* - c(s^{\alpha,\beta})$$

$c(s^{\alpha,\beta})$ is a correction term which guarantees that $t^{\alpha,\beta}$ is an integer as well. For each fixed column index $s^{\alpha,\beta}$ of the new total order, $c(t^{\alpha,\beta})$ is determined by $c(s^{\alpha,\beta}) = argmin_{(s^*,t^*)}\{|t^*|\}$, that is, $c(t^{\alpha,\beta})$ is set equal to the t^* of the lowest absolute value in the $s^{\alpha,\beta}$ column. For $\theta = \frac{\pi}{2}$ the transformation is obtained by interchanging the roles of columns and rows. The total order in the rotated system is defined similarly to (1.25a), that is

$$(s_1^{\alpha,\beta}, t_1^{\alpha,\beta}) < (s_2^{\alpha,\beta}, t_2^{\alpha,\beta}) \text{ iff } (s_1^{\alpha,\beta}, t_1^{\alpha,\beta}) \in \{(k,l)|k = s_2^{\alpha,\beta}, l < t_2^{\alpha,\beta}\} \cup \{(k,l)|k < s_2^{\alpha,\beta}, -\infty < l < \infty\}. \quad (1.39)$$

Let us denote by O the set of all possible total order and NSHP definitions on the 2-D lattice, in which the boundary line of the NSHP is of rational slope, that is $O = \{\alpha, \beta | \alpha, \beta$ are coprime integers $\}$. We shall call such support rational nonsymmetrical half-plane (RNSHP). An example of one such total order and RNSHP definition is illustrated in Figure 1.3. Note the way the "column" is defined.

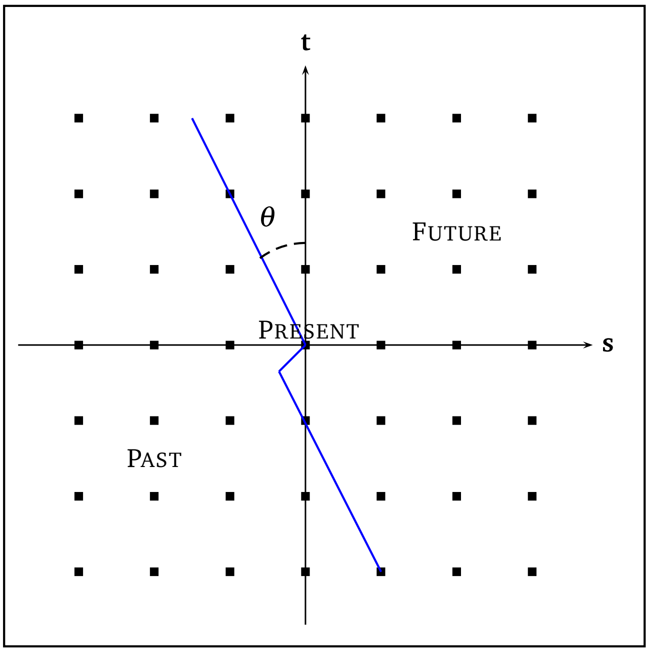


Figure 1.3: RNSHP Total-Order Definition

Definition 1.4.13. *A field $\{G(s,t)\}$ is called generalized evanescent if it can be represented as a linear infinite combination of evanescent fields. Each of the evanescent fields generates column to column innovation with respect a different total order and RNSHP support.*

We study the family of total orders defined by Definition 1.4.12 to better understand the structure of the deterministic component of the decomposition (1.26). Define $\mathscr{H}^{\varepsilon}_{(\infty,\infty)} = \overline{sp}\{\varepsilon(k,l), (k,l) \in \mathbb{Z}^2\}$, to be the Hilbert space spanned by the purely indeterministic component of the regular field and similarly define $\mathscr{H}^{V}_{(\infty,\infty)} = \overline{sp}\{V(k,l), (k,l) \in \mathbb{Z}^2\}$ to be the Hilbert space spanned by the deterministic component of the regular field.

Lemma 1.4.14. *Interchanging of roles of past, and future is any RNSHP total-ordering definition imposed on a regular field results in identical evanescent components.*

Proof: Let P be some RNSHP total order. The Hilbert space spanned by the corresponding evanescent component is given by

$$\bigoplus_{s=-\infty}^{\infty} \mathscr{H}^{P,V}_{s} = \bigoplus_{s=-\infty}^{\infty} \overline{Sp}\left\{V | V \in \mathscr{H}^{P,V}_{(s,-\infty)}, V \perp \mathscr{H}^{P,V}_{(s-1,-\infty)}\right\}.$$

Let F denote the total order obtained by rotating the order P by $\theta = \pi$. The Hilbert space spanned by the corresponding evanescent component is given by

$$\bigoplus_{s'=-\infty}^{\infty} \mathscr{H}^{P,V}_{s'} = \bigoplus_{s'=-\infty}^{\infty} \overline{Sp}\left\{V | V \in \mathscr{H}^{F,V}_{(s',-\infty)}, V \perp \mathscr{H}^{F,V}_{(s'-1,-\infty)}\right\}.$$

Let (s,t) and (s_2,t_2) be the indices of the same grid point under the two different order definitions. The properties of the deterministic random field imply that for any (s_1,t_1) and for any fixed $o \in O$, $\mathscr{H}^{o,P}_{s_1,t_1} = \mathscr{H}^{o,P}_{(s_1,-\infty)} = \mathscr{H}^{o,P}_{(s_1,\infty)}$.

Because for any $o' \in O$, $\mathscr{H}^{o',V}_{s',t'} = \overline{Sp}\{V(s_2,t_2)|(s_2,t_2) \leq (s',t')\}$ and because the deterministic component of the random field is unique and NSHP total-ordering invariant,

$$\mathscr{H}^{F,V}_{(s',-\infty)} = \mathscr{H}^{F,V}_{(s',t')} = \mathscr{H}^{V}_{(\infty,\infty)} \ominus \mathscr{H}^{P,V}_{(s,t-1)} = \mathscr{H}^{V}_{(\infty,\infty)} \ominus \mathscr{H}^{P,V}_{(s,-\infty)}.$$

Due to the reversed order of the indexing induced by the two total order definitions P and F, when $t \to \infty$, $t' \to -\infty$. By the same argument,

$$\mathscr{H}^{F,v}_{(s'-1,-\infty)} = \mathscr{H}^{v}_{(\infty,\infty)} \ominus \mathscr{H}^{P,v}_{(s+1,-\infty)}.$$

Hence,

$$\begin{aligned}
\bigoplus_{s'=-\infty}^{\infty} \mathscr{H}^{F,V}_{s'} &= \bigoplus_{s=-\infty}^{\infty} \overline{Sp}\left\{V|V \in \mathscr{H}^{V}_{(\infty,\infty)} \ominus \mathscr{H}^{P,V}_{(s,-\infty)}, V \perp \mathscr{H}^{V}_{(\infty,\infty)} \ominus \mathscr{H}^{P,V}_{(s+1,-\infty)}\right\} \\
&= \bigoplus_{s=-\infty}^{\infty} \overline{Sp}\left\{V|V \in \mathscr{H}^{P,V}_{(s+1,-\infty)}, V \in \mathscr{H}^{P,V}_{(s,-\infty)}\right\} = \bigoplus_{s=-\infty}^{\infty} \mathscr{H}^{P,V}_{s}.
\end{aligned}$$

□

Define $\mathscr{H}^{X}_{(-\infty,-\infty)} = \cup_{o \in O} \mathscr{H}^{o,X}_{(-\infty,-\infty)}$. Note that $\mathscr{H}^{X}_{(-\infty,-\infty)}$ is the Hilbert space spanned by the intersection of all Hilbert spaces spanned by the regular field samples $\{X(s,t)\}$ for all (s,t) and with respect to all possible RNSHP total-order definitions.

Definition 1.4.15. *A 2-D deterministic random field $\{P(s,t)\}$ is called half-plane deterministic if it has no column-to-column innovations with respect to any RNSHP total ordering definition.*

The deterministic field $\{V(s,t)\}$ can be further decomposed as follows:

Theorem 1.4.16 ([14, 21])**.** *The deterministic component $\{V(s,t)\}$ of a 2-D regular and stationary random field can be represented uniquely by the following countable infinite orthogonal decomposition*

$$V(s,t) = P(s,t) + G(s,t) = P(s,t) + \sum_{\alpha,\beta \in O}\sum e_{\alpha,\beta}(s,t).$$

The random field $\{P(s,t)\}$ is half-plane deterministic. The field $\{e_{\alpha,\beta}\}$ is the evanescent component that generates the column-to-column innovations of the deterministic field with respect to the RNSHP total-ordering definition $\alpha, \beta \in O$, and $G(s,t)$ is called generalized evanescent random field. The two random fields $P(s,t)$ and $G(s_1,t_1)$ are mutually orthogonal for all (s,t) and (s_1,t_1). The generalized evanescent field is a linear combination of a countable number of mutually orthogonal evanescent fields.

Proof: We first show that for any pair of RNSHP total-ordering definitions $o, o' \in O$ such that o' is not obtained by rotating o by $\theta = \pi$.

$$\left(\bigoplus_{s=-\infty}^{\infty} \mathscr{H}^{o,V}_{s}\right) \perp \left(\bigoplus_{s'=-\infty}^{\infty} \mathscr{H}^{o',V}_{s'}\right). \tag{1.40}$$

Using Lemma 1.4.14, we conclude that it is sufficient to consider only $-\pi < \theta < 0$. Let $V(s,t)$ be the deterministic component of the (s,t)th sample of the regular field, where the indexing is with respect to the order definition o. Let also $\mathscr{H}^{o}_{s}(q) = \overline{Sp}\{V(k,l)|k = s,\ l \leq q\}$, the indexing is with respect to o. Let u be a vector such that $u \in \mathscr{H}^{o,V}_{s}$ for some fixed n. Using the definition $\mathscr{H}^{o,V}_{s}$ and because for all m, $\mathscr{H}^{o,V}_{s,t} = \mathscr{H}^{o,V}_{s,-\infty}$, we conclude that for all m,

$$\mathscr{H}^{o,V}_{s} = \mathscr{H}^{o,V}_{s,t} \ominus \mathscr{H}^{o,V}_{s-1,\infty} \subset \mathscr{H}^{o}_{s}(t). \tag{1.41}$$

Assume $u \in \mathscr{H}_{s'}^{o',V}$ for some fixed s'. Therefore, $u \in \mathscr{H}_{s',-\infty}^{o',V}$, $u \perp \mathscr{H}_{s'-1,-\infty}^{o',V}$.
Because any support o' considered here will contain an infinite number of samples $\{V(s,t)\}_{t=-\infty}^{t_1}$, from sth column defined with respect tothe RNSHP order definition o, we have that for any o', $\mathscr{H}_s^{o,V} \subset \mathscr{H}_{s'-1,-\infty}^{o'}$. Because (1.41) holds for all m, we have $\mathscr{H}_s^{o,V} \subset \mathscr{H}_s^{o}(m)$.Hence, $\mathscr{H}_s^{o,v} \subset \mathscr{H}_{s'-1,-\infty}^{o',v}$ and $u = 0$.
Because the preceding argument holds for all s', we conclude that $\mathscr{H}_s^{o,v} \perp \bigoplus_{s'=-\infty}^{\infty} \mathscr{H}_{s'}^{o',v}$. Repeating the same arguments for each s, we obtain (1.40). Hence, the evanescent fields are mutually orthogonal.
The deterministic component of the random field is unique and NSHP total ordering invariant. We can therefore rewrite (**??**) for any total-order $o \in O$, while letting $s,t \to \infty$.

$$\mathscr{H}_{(\infty,\infty)}^{V} = \mathscr{H}_{(\infty,\infty)}^{o,V} = \mathscr{H}_{(-\infty,-\infty)}^{o,X} \oplus \bigoplus_{s'=-\infty}^{\infty} \mathscr{H}_{l}^{o,V}. \tag{1.42}$$

For any $o \in O, \oplus_{l=-\infty}^{\infty} \mathscr{H}_l^{o,V} \subset \mathscr{H}_{(\infty,\infty)}^{V}$. Also, for any two total-order definitions $o, o' \in O$ such that o' is not obtained by rotating o by $\theta = \pi$, $(\oplus_{l=-\infty}^{\infty} \mathscr{H}_l^{o,V}) \perp (\oplus_{k'=-\infty}^{\infty} \mathscr{H}_{k'}^{o',V})$. Hence, we conclude using (1.42) that for any two such total order definitions $o, o' \in O$,

$$\left(\bigoplus_{k'=-\infty}^{\infty} \mathscr{H}_{k'}^{o',V} \right) \subset \mathscr{H}_{(-\infty,-\infty)}^{o,X}. \tag{1.43}$$

Using (1.42) together with the uniqueness and NSHP total ordering invariance of the deterministic component, we conclude that

$$\begin{aligned} \mathscr{H}_{(\infty,\infty)}^{V} &= \bigcap_{o \in O} \mathscr{H}_{(\infty,\infty)}^{o,V} = \bigcap_{o \in O} \left(\mathscr{H}_{(-\infty,-\infty)}^{o,X} \oplus \bigoplus_{k'=-\infty}^{\infty} \mathscr{H}_{l}^{o,V} \right) \\ &= \mathscr{H}_{(-\infty,-\infty)}^{X} \oplus \bigoplus_{o \in O} \bigoplus_{l=-\infty}^{\infty} \mathscr{H}_{l}^{o,V}, \end{aligned} \tag{1.44}$$

where the last equality results from the definition of $\mathscr{H}_{(-\infty,-\infty)}^{X}$ from (1.40), which results in the elimination of the cross terms that involve the intersection of more than one Hilbert space of the type $\oplus_{l=-\infty}^{\infty} \mathscr{H}_l^{o,V}$ and from (1.43). $\oplus_{o \in O} \oplus_{l=-\infty}^{\infty} \mathscr{H}_l^{o,V}$ is the Hilbert space spanned by all the evanescent components of the regular field. Because O is a countable set, the number of evanescent components of a regular field is countable. By (1.44) and Definition 1.4.15, $\mathscr{H}_{(-\infty,-\infty)}^{X}$ is spanned by a half plan deterministic field. □

The result in (1.43) implies that for each RNSHP total-order definition $o \in O$, all subspaces spanned by the evanescent components $e_{o'}$, where $o' \neq o$, remain in the remote past space $\mathscr{H}_{(-\infty,-\infty)}^{o,X}$, which corresponds to the definition o. Hence, from (1.44) we conclude that in order to resolve all the evanescent components of a regular field, the field has to be tested against all the possible RNSHP total-ordering definitions in O. Note also that because column-to column innovations are found only when RNSHP total-ordering definitions are imposed on the field, and because the half-plane deterministic component of the decomposition is deterministic by definition, we conclude that the half-plane deterministic field has no innovations nor column-to-column innovations with respect to any NSHP total ordering.

1.4.3 Quarter-plane representations

The problem of predicting random fields can be posed in several ways, depending on the considered past. With quarter-plane Q as the past, the authors in [36] obtained a four-fold Wold decomposition in the time-domain for a stationary random field. For the precise statement of the corresponding MA representation and the necessary and sufficient spectral conditions for it, we recall that a stationary random field $\{X(s,t)\}$ is said to have a one-sided (unilateral) MA(∞) representation on Q if there exists a white noise process $\varepsilon(s,t)$ such that

$$X(s,t) = \varepsilon(s,t) + \sum\sum_{(k,l)\in Q} b_{k,l}\varepsilon(s-k,t-l), \tag{1.45}$$

where the sequence $\{b_{k,l};(k,l)\in\mathbb{Z}^2\}$ consists of the MA parameters of the process, with $b_{0,0}=1$, $b_{k,l}=0$, when $k<0$ or $l<0$ and

$$\sum\sum_{(k,l)\in Q} b_{k,l} < \infty.$$

In view of the equality of the two subspaces involving the pasts of the two process, we call $\varepsilon(s,t)$ the innovation process of $X(s,t)$ and the $V(\varepsilon(s,t)) \equiv \sigma^2 = \sigma^2(Q)$ is referred to as the innovation variance. The following result from [17] gives the spectral characterization of stationary random fields having a one-sided MA(∞) representation on Q.

Theorem 1.4.17 ([17]). *A stationary random field $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ with spectral density function $f(\lambda_1,\lambda_2)$ has a one-sided* MA(∞) *representation* (1.45) *on the quarter-plane, if and only if*

(i) $\log(f)\in L^1$,

(ii) *the Fourier coefficients of* $\log(f)$ *vanish outside of* $\{-Q\}\cup Q\cup\{0\}$,

(iii) $\mathscr{H}_f^{00} = \mathscr{H}_f^{0\infty}\cap\mathscr{H}^{\infty 0}$,

where H_f^M is defined as before with M:

$$\begin{aligned} M &= \{(s,t);s\le 0,t\le 0\}, \textit{ for } \mathscr{H}_f^{00},\\ &= \{(s,t);s\le 0,t\in\mathbb{Z}\}, \textit{ for } \mathscr{H}_f^{0\infty},\\ &= \{(s,t);s\in\mathbb{Z},t\le 0\}, \textit{ for } \mathscr{H}_f^{00}. \end{aligned}$$

The spectral factorization of the spectral density function on Q provides an analytic $\phi(z_1;z_2)$ of complex variables $z_1=e^{i\lambda_1}$ and $z_2=e^{i\lambda_2}$ such that the spectral satisfies $f(\lambda_1,\lambda_2)=|\phi(z_1,z_2)|^2$, where $\phi(0,0)=1$ and

$$\phi(z_1,z_2) = e^{\left(\sum_{k=0}^{+\infty}\sum_{l=0}^{+\infty} c_{k,l} z_1^k z_2^l\right)},$$

where $c_{k,l}$ are the two-dimensional cepstral coefficients of $f(\lambda_1,\lambda_2)$.

Recursive Formulas for the AR and MA Coefficients

Recursive formulas which relate the MA and AR parameters in terms of the cepstral or Fourier coefficients of the logarithm of the spectral density function are given hereafter, (see [37]). Consider the Taylor expansions of the optimal factor ϕ and its inverse:

$$\begin{aligned} \phi(z_1,z_2) &= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} b_{k,l} z_1^k z_2^l,\\ \phi^{-1}(z_1,z_2) &= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_{k,l} z_1^k z_2^l, \end{aligned} \tag{1.46}$$

from which it follows that the MA and AR parameters of the random field are related to each other via the recursions

$$\begin{aligned} b_{0,0} &= a_{0,0} = 1, \\ b_{i,j} &= \sum_{\substack{k=0 \\ (k,l)\neq(i,j)}}^{i} \sum_{l=0}^{j} b_{k,l} a_{i-k,j-l}. \end{aligned} \tag{1.47}$$

The corresponding autoregressive representation of $\{X(s,t);\, (s,t) \in \mathbb{Z}^2\}$ is given by

$$X(s,t) = \varepsilon(s,t) + \sum\sum_{(k,l)\in Q} a_{k,l} X(s-k, t-l). \tag{1.48}$$

1.5 Conclusion

We have mainly presented a Wold-like orthogonal decomposition for 2-D discrete random fields. In the stationary case, the regular random field is decomposed into a purely indeterministic component, a countable number of mutually orthogonal evanescent components and a half-plane deterministic component. The resulting decomposition of the spectral distribution function of the regular and stationary random field into absolutely continuous and singular spectral distributions was presented. Finally, we concluded by giving quarter-plane representations.

CHAPTER

2

Prediction of Stationary 2-D Random Fields with Quarter-Plane Past

Contents

2.1 Introduction

Prediction theory of stationary 2-D random fields with quarter-plane past has attracted much attention in recent years. Firstly, we give a small overview of some works on random fields multi-step prediction with quarter-plane past [7], [8], [9]. Secondly, we detail our first contribution, namely, we treat the prediction problems where a number of observations are missing to the quarter-plane past of a stationary random field.

In what follows, the problem of prediction is developed in the case when the past is generated by a quarter-plane Q, defined by

$$Q = \{(k,l),\ (k,l) \in \mathbb{Z}^2,\ k \leq 0, l \leq 0, (k,l) \neq (0,0)\}.$$

We define H_Q^X as the closed linear subspace spanned by $\{X(k,l), (k,l) \in Q\}$ in the Hilbert space H, which represents the strict past of X at the point $(0,0)$. We denote by $P_{H_Q^X}$ the orthogonal projection operator onto H_Q^X, i.e, $P_{H_Q^X}X(s,t)$ one-step minimum norm causal continuous support, linear predictor of $X(s,t)$. The innovation process $\{\varepsilon(s,t), (s,t) \in \mathbb{Z}^2\}$ with respect to the defined support and total order is given by

$$\varepsilon(s,t) = X(s,t) - P_{H_Q^X}X(s,t)$$

The outline of this chapter is as follows: In Section 2.2, some prediction problems for stationary random fields with quarter-plane past are given. Indeed, [7] obtained the best predictor based on the quarter-plane with finitely many observations added. Furthermore, solution expresses the error variance formulas in terms of the moving average (MA) parameters of the random fields are obtained. However, expressing the solution in terms of the autoregressive (AR) parameters depends on a mysterious projection operator(see [38]). In section 2.3, the prediction problem is solved when the stationary Gaussian random field possesses a one-sided moving average representation in a certain strict sense [8]. In section 2.4, prediction error of a future value of stationary random fields when the infinite past is altered by some missing observations is established in [9]. Furthermore, an explicit formula for $X(0,0) - P_{H_{Q_1}^X}X(0,0)$ is given in Theorem 2.4.1. This formula leads to the above-mentioned expression of $var\left(X(0,0) - P_{H_{Q_1}^X}X(0,0)\right)$ and allows to derive the AR representation of $P_{H_{Q_1}^X}X(0,0)$ in Theorem 2.4.2. In Theorem 2.4.3, we characterize the processes for which the loss of observations in the past does not affect the prediction of $X(0,0)$, i.e., $P_{H_{Q_1}^X}X(0,0) = P_{H_Q^X}X(0,0)$. This representation can be seen as an alternative solution to the problem posed in [7].

In Section 2.5, we detail our first contribution. The aim is to quantify the impact of missing observations from the past. The central idea of our study consists in using the MA and AR representations of the random field. The obtained results highlight the important role of the AR parameters in forecasting. Indeed, we establish lower and upper bounds for the prediction error variance given in Theorem (2.5.1) which is the novelty of our work. This boundedness property of prediction error variance shows that the degradation of the prediction due to the missing data increases with the maximum value of the AR parameters of the missing data. They also allow to conclude that the larger the indices of missing values are, the better is the preciseness of the bound of the prediction error variance. Also, our results characterize the random fields for which the missing observations do not affect the prediction of $X(0,0)$. Finally, note that $X(0,0)$ is chosen without loss of generality, and the main conclusions extend, naturally, if we consider the prediction of $X(h_1,h_2)$, $h_1,h_2 > 0$. To summarize, we establish lower and upper bounds for the prediction error variance, see Theorem (2.5.1). In Corollary (2.5.2), some properties of the predictor are presented, and in Corollary (2.5.3) , asymptotic behaviors for the prediction error variance are obtained. We conclude this section with two illustrative examples.

2.2 Multi-step ahead prediction with quarter-plane past

2.2.1 Introduction

Predicting future values other than $X(0,0)$ is important in the theory and applications of stationary random fields, this problem have been studied in [7,39]. In this section, the multi-step ahead predictor and its prediction error variance when Q is used as the past are derived. For more details on the results and the proofs, we refer the reader to [7,39].

Theorem 2.2.1 ([7]). *Let $X(s,t)$ be a PND stationary random field with the spectral density function $f(\lambda_1,\lambda_2)$ satisfying the conditions of Lemma 1.4.17. Then for any $((h_1,h_2)\in Q^c$, where Q^c is the complement of Q, the (h_1,h_2)-step ahead linear least-squares predictor of $X(h_1,h_2)$ based on the past Q is given by*

$$X(h_1,h_2) = \sum_{\substack{k=h_1 \\ (k,l)\neq(h_1,h_2)}}^{\infty}\sum_{l=h_2}^{\infty} b_{k,l}\varepsilon(h_1-k,h_2-l),$$

with the corresponding prediction error

$$\begin{aligned} X(h_1,h_2)-P_{H_Q^X}X(h_1,h_2) &= b_{h_1,h_2}\varepsilon(0,0)+\sum_{k=0}^{h_1-1}\sum_{l=h_2}^{\infty} b_{k,l}\varepsilon(h_1-k,h_2-l) \\ &+ \sum_{k=h_1}^{\infty}\sum_{l=0}^{h_2-1} b_{k,l}\varepsilon(h_1-k,h_2-l)+\sum_{k=0}^{h_1-1}\sum_{l=0}^{h_2-1} b_{k,l}\varepsilon(h_1-k,h_2-l), \end{aligned} \tag{2.1}$$

and the prediction error variance

$$V\left\{X(h_1,h_2)-P_{H_Q^X}X(h_1,h_2)\right\} = \sigma^2\left(b^2_{h_1,h_2}+\sum_{k=0}^{h_1-1}\sum_{l=h_2}^{\infty}|b_{k,l}|^2+\sum_{k=h_1}^{\infty}\sum_{l=0}^{h_2-1}|b_{k,l}|^2+\sum_{k=0}^{h_1-1}\sum_{l=0}^{h_2-1}|b_{k,l}|^2\right), \tag{2.2}$$

where $\sigma=\sigma^2(Q)$ is the innovation variance.

Remark 1. *Depending on $(h_1;h_2)$, some terms in* (2.1) *and* (2.2) *become zero for either of $h_i<0; i=1,2$, but they always involve the infinite sums which makes them quite different from their counterparts in* 1*-D process in the sense that the latter are always finite sums, (see [40, p. 181]).*

2.2.2 Prediction with augmented quarter-plane Past

We consider the problem of prediction of $X(0,0)$ when a finite number of observations are added to the quarter-plane Q and the past is modified to:

$$I = Q\cup K,$$

where K represents a finite dimensional space spanned by the additional observations. This is the first natural step in generalizing the 1-D to the 2-D processes. It turns out that computing such predictors and prediction error variances are closely related to finding multi-step ahead predictions using the MA(∞) representation (1.45) of a PND random field with Q as its past. The orthogonal projection of $X(0,0)$ onto the linear subspace generated by the random variables with indices from the modified past, $I=Q\cup K$ provides the best linear predictor of $X(0,0)$. However, due to non orthogonality of Q and K computing this predictor is not easy even though it is known how to compute the projections of $X(0,0)$ onto Q and K separately. A natural way to alleviate this problem is to re-express I as the orthogonal sum of Q and a finite-dimensional subspace orthogonal to it. To this end, define

$$A = sp\{X(i,j)-P_{H_Q^X}(i,j);\,(i,j)\in K\}, \tag{2.3}$$

Lemma 2.2.2 ([7,39]). *Let $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ be a PND stationary random field satisfying the conditions of Theorem 1.4.17. Suppose the MA parameters of $\{X(s,t)\}$ are $\{b_{k,l};(k,l)\in\mathbb{Z}^2\}$, with the modified past $I=Q\cup \mathrm{K}$,*

(a) *$X(0,0)$ is not in the modified set I, that is, $X(0,0)\notin I$.*

(b) *Q and I are orthogonal subspaces spanning I such that, $I=Q\oplus A$, where A is given in* (2.3).

(c) *The orthogonal projection of $X(0,0)$ onto $I=Q\oplus A$ is given by $P_I X(0,0)=P_Q(0,0)+P_{H_Q^X}X(0,0)$*

(d) *The orthogonal projection of $X(0,0)$ onto the finite-dimensional subspace A is given by*

$$P_{H_Q^X}X(0,0)=\sum\sum_{(i,j)\in K}\beta_{ij}\{X(i,j)-P_{H_Q^X}(i,j)\}=\beta'(X_K-\widehat{X}_K), \tag{2.4}$$

where $\beta=\{\beta_{i,j},(i,j)\in K\}$ arranged using the lexicographic order of $\mathbb{Z}^2$ is given by $\beta=\mathrm{C}^{-1}b$, and C is the variance-covariance matrix of the prediction errors in A, and $b=\{b_{i,j},(i,j)\in K\}$ is a vector of MA parameters, then the following normal equations hold true $C\beta=b$.

The results in Lemma 2.2.2 are the counterparts of those for prediction problems in 1-D in stationary processes with modified past, (see [41,42]). We now establish the prediction based on two different configurations of the set K in Lemma 2.2.4 and Theorem 2.2.3.

A Single Additional Observation

When the modified past has only the single additional observation $X(h_1,h_2)$, that is,

$$I_1=Q\cup\{X(h_1,h_2)\}=Q\cup K$$

with $h_i\geq 0, i=1,2$ and $(h_1,h_2)\neq(0,0)$, then computing the prediction error variance for $P_{I_1}X(0,0)$ involves projecting onto a one-dimensional subspace.

Theorem 2.2.3 ([7,39]). *Let $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ be a PND stationary random field with the innovation variance σ^2 and MA parameters $\{b_{k,l};(k,l)\in\mathbb{Z}^2\}$. Then, the best linear predictor of $X(0,0)$ based on I_1 is given by*

$$P_{I_1}X(0,0)=P_{H_Q^X}X(0,0)+\beta_{h_1,h_2}(X(h_1,h_2)-P_{H_Q^X}X(h_1,h_2)),$$

where

$$\beta_{h_1,h_2}=\frac{b_{h_1,h_2}^2}{b_{h_1,h_2}^2+\sum_{k=0}^{h_1-1}\sum_{l=h_2}^{\infty}b_{k,l}^2+\sum_{k=h_1}^{\infty}\sum_{l=0}^{h_2-1}b_{k,l}^2+\sum_{k=0}^{h_1-1}\sum_{l=0}^{h_2-1}b_{k,l}^2}.$$

The corresponding prediction error variance is

$$\sigma^2(I_1)=\sigma^2\left(1-\frac{b_{h_1,h_2}^2}{b_{h_1,h_2}^2+\sum_{k=0}^{h_1-1}\sum_{l=h_2}^{\infty}b_{k,l}^2+\sum_{k=h_1}^{\infty}\sum_{l=0}^{h_2-1}b_{k,l}^2+\sum_{k=0}^{h_1-1}\sum_{l=0}^{h_2-1}b_{k,l}^2}\right). \tag{2.5}$$

It is clear from (2.5), that for a stationary random field with $b_{h_1,h_2}=0$ adding $X(h_1,h_2)$ to Q will have no effect on the prediction error variance of $X(0,0)$. However, unlike the 1-D case the prediction error variance for 2-D processes involves infinite number of terms.

Several Additional Observations

Next we turn our attention to the prediction of $X(0,0)$ based on the knowledge of

$$I_2 = Q \cup \{X(i,j); 0 \leq i \leq h_1,\ 0 \leq j \leq h_2,\ (i,j) \neq (0,0)\} = Q \cup \mathrm{K},$$

where K is a finite set of n future observations in the first quadrant. Here the set K in the $h_1 \times h_2$ rectangle or any subset of it. This is the simplest form of the problem of finding the best linear predictor of $X(0,0)$ and its prediction error variance when a finite number of observations are added to Q. It is analogous to the interpolation problem in stationary 1-D processes to predict $X(0)$ see [41,43]. In the next lemma, the vector of prediction errors with entries from the set A in (2.3) is written as a linear transformation of the innovation process.

Lemma 2.2.4 ([7,39]). *Let $X(s,t)$ be a PND stationary random field satisfying the conditions of Theorem 2.2.1. Suppose the innovations for $X(s,t)$ are $\varepsilon(k,l)$, $(k,l) \in \mathbb{Z}$, the MA and AR parameters are $\{b_{k,l}\}$ and $\{a_{k,l}\}$, respectively. Then,*

(a) *The vector of prediction errors for the observations in K based on the knowledge of Q can be expressed as:*

$$X_K - \hat{X}_K = b_K \varepsilon(0,0) + T^t \varepsilon_K, \tag{2.6}$$

where $b_K = \{b_{k,l}, (k,l) \in K\}$ is a vector of MA parameters corresponding to observations in the set K. ε_K is a vector with entries indexed by the lattice points in the finite set K future observations with all points in the $X(h_1,h_2)$ rectangle.Theterms in both b_K and ε_K are arranged lexicographically, T is a rectangular matrix with infinite number of rows and n columns. The MA parameters appearing in the prediction errors of the n observations in K correspond to the n columns in T .

(b) *The variance-covariance matrix of the vector of prediction errors $X_K - \hat{X}_K = b_K$ is*

$$C = \sigma^2(T^t\overline{T} + b_K^t \overline{b}_K),$$

where $\overline{T} = (T^t)^*$ is the entry wise complex conjugate of T .

(c) *Let (a_K) be a vector of AR parameters with indices arranged as in ε_K. Then, $b_K = T^t a_K$.*

For 1-D stationary processes the matrix T involved in the prediction errors for additional observations, is a lower triangular, square and Toeplitz matrix of MA parameters, see [40, p.271]. For 2-D processes, if the set K consists of observations of the form $\{X(i,j), i > 0\}$, then the matrix T in (2.6) reduces to a rectangular block lower triangular and Toeplitz matrix. Furthermore, for K finite the matrix $G = T^t\overline{T}$, as a scalar multiple of the covariance matrix of $T^t\varepsilon_K$, is invertible since ε_K is a segment of the innovation process.

Theorem 2.2.5 ([7,39]). *Let $\{X(s,t), (s,t) \in \mathbb{Z}^2\}$ be a PND stationary random field satisfying the conditions of Theorem 2.2.1. Suppose the innovations for $X(s,t)$ are $\varepsilon(k,l)$, $(k,l) \in \mathbb{Z}$, the MA and AR parameters are $\{b_{k,l}\}$ and $\{a_{k,l}\}$, respectively. Then, the best linear predictor of $X(0,0)$ based on I_2 is*

$$P_{I_2}X(0,0) = P_{H_Q^X}X(0,0) + \beta^t(X_K - P_{H_Q^X}X_K),$$

where $P_{H_Q^X}X(0,0)$ is the orthogonal projection of $X(0,0)$ onto Q and

$$\beta = G^{-1}b_k(1 + b_k^* G^{-1} b_k)^{-1},$$

where $G = T^t\overline{T}$. The corresponding prediction error variance in terms of the MA parameters is

$$\sigma^2(I_2) = \sigma^2(1 + b_k^* G^{-1} b_k)^{-1}.$$

The prediction error variance can also be expressed in terms of the AR parameters as:

$$\sigma^2(I_2) = \sigma^2(1 + a_k^* P a_k)^{-1}, \tag{2.7}$$

where $P = \overline{T}(T^t\overline{T})^{-1}T^t$ is a projection matrix.

The prediction error variance in (2.7) seems to be expressible only in terms of the MA parameters. An attempt to express (2.7) in terms of the AR parameters runs into the projection operator P which makes the situation for 1-D and 2-D processes to differ considerably.

2.3 Prediction of Gaussian random fields with incomplete quarter-plane past

2.3.1 Introduction

In [17], the basic quarter-plane prediction problem is solved when the stationary Gaussian random field possesses a one-sided moving average representation in a certain strict sense. Based on these results, [7] expresses the error variance predictors formulas in terms of the moving average (MA) parameters of the random field. However, they found that expressing the solution in terms of the autoregressive (AR) parameters depends on a mysterious projection operator (see [38]). [8] is concerned with solving the quarter-plane prediction of stationary Gaussian random field based on a quarter-plane with finitely missing observations which was left open in [7]. The methods used here sidestep, but do not solve, the technical obstacles noted in [7]. Two solutions have been presented in [8]. The first solution, representing a direct approach, expresses the best predictor in terms of the MA coefficients of the random field. The second solution requires more stringent spectral conditions, and employs a modified duality argument to express the prediction error variance in terms of the AR coefficients of the random field.

Throughout this section, the problem of prediction is developed in the case when the past is generated by a quarter-plane $\mathbf{Q}$, defined by

$$\mathbf{Q} = \{(s,t),\ \ s \geq 0,\ t \geq 0, (s,t) \neq (0,0)\}. \tag{2.8}$$

Let λ be normalized Lebesgue measure on the unit circle $\mathbf{T}$ of the complex plane. Throughout this section we take $\{X(s,t),\ (s,t) \in \mathbb{Z}^2\}$ to be a complex valued, centered, stationary Gaussian random field. We assume that $\{X(s,t)\}$ has a continuous spectral measure with respect to normalized Lebesgue measure $d\lambda^2 = d\lambda \times \lambda$ on the torus $\mathbf{T}^2$. Its spectral density function $f(e^{i\lambda_1}, e^{i\lambda_2})$, defined in (1.34), is assumed to factorize as $f = |\phi|^2$.

In this situation, the trigonometric isomorphism $X(s,t) \longmapsto e^{i(s\lambda_1 + t\lambda_2)}$ then brings in the theory of functions on $\mathbf{T}^2$. We identify the Hardy class $H^2(\mathbf{T}^2)$ of functions on the torus with that subspace of $\mathbf{T}^2$. For analytical reasons, the choice of quarter-plane past enables us to investigate the prediction problem by analogy with the one-parameter case, by drawing upon the rich function theory on $\mathbf{T}^2$, the condition of f being outer is merely necessary, but not sufficient, for the property

$$\overline{sp}\{f(e^{i\lambda_1}, e^{i\lambda_2})e^{i(s\lambda_1 + t\lambda_2)},\ (s,t) \in \mathbf{Q} \cap \{(0,0)\}\} = H^2(\mathbf{T}^2). \tag{2.9}$$

With that in mind, we adopt the terminology of [44, 45] and speak of f as being strongly outer in $H^2(\mathbf{T}^2)$ if it enjoys the property (2.9). The property of the spectral factor $\phi(e^{i\lambda_1}, e^{i\lambda_2})$ being strongly outer is equivalent to the condition

$$\overline{sp}\{X(s,t), s \geq 0, t \geq 0\} = \overline{sp}\{\varepsilon(s,t), s \geq 0, t \geq 0\}.$$

[17] solved the basic quarter-plane prediction problem under the assumption that the spectral measure of the random field has a strongly outer factorization. To solve the corresponding problem with missing observations, we wil need to develop further the theory of strongly outer functions, and that is the subject of the next section.

If $\mathscr{S}$ is a subset of $\mathbb{Z}^2 \backslash (0,0)$, let $\mathscr{S}^c$ be its complement in $\mathbb{Z}^2 \backslash (0,0)$. Let us stress that the origin is excluded in this complementation. By $-\mathscr{S}$ we understand the set $\{(-s,-t), (s,t) \in \mathscr{S}\}$. We adopt the notation $f \in \mathscr{S}$ to mean that $f(e^{i\lambda_1}, e^{i\lambda_2})$ has a square summable Fourier Series, and its coefficients $\hat{f}^{s,t}$ vanish for all (s,t) outside of $\mathscr{S}$. Let us say that $\mathscr{S}$ is **Q**-invariant if $\mathscr{S} + (s,t) \subseteq \mathscr{S}$ for all $(s,t) \in \mathbf{Q}$). It is obvious that **Q** itself is **Q** -invariant. Our results for **Q** will in many cases extend to parameter sets $\mathscr{S}$ that are **Q** -invariant.

2.3.2 Strongly outer functions

The terms strongly outer and weakly outer were introduced in [46] in connection with the behavior of certain projection operators in the prediction of stationary random fields. The importance of the strong outer condition to quarter-plane prediction problems was previously established in [15, 17], and the associated function theory was further investigated in [44, 46]. This section presents some additional theory of strongly outer functions. The assertions below enable us to extend many of our results from the quarter-plane **Q** to more general parameter sets. The following characterizations of **Q**-invariance are obvious and will be useful.

Proposition 2.3.1. *Let S be a subset of $\mathbb{Z}$. The following are equivalent:*

1. *$\mathscr{S}$ is **Q** -invariant;*
2. *$\mathscr{S} + (1,0) \subseteq \mathscr{S}$ and $\mathscr{S} \subseteq \mathscr{S} + (0,1)$;*
3. *The parameter set $-(\mathscr{S}^c)$ is **Q** -invariant.*

The importance of this idea is that parameter sets $\mathscr{S}$ that are **Q**-invariant enjoy of property (2.9) in connection with strongly outer functions. We have already seen that this is an important tool in solving prediction problems.

Proposition 2.3.2. *If φ is strongly outer in $H^2(\mathbf{T}^2)$, and $\mathscr{S}$ is a **Q**-invariant subset of $\mathbb{Z}^2 \backslash (0,0)$, then*

$$\{\phi(e^{i\lambda_1}, e^{i\lambda_2}) e^{i(s\lambda_1 + t\lambda_2)} : (s,t) \in \mathscr{S}\} = \{e^{i(s\lambda_1 + t\lambda_2)} : (s,t) \in \mathscr{S}\}.$$

This underscores the importance of the parameter set **Q** , even though it may appear to be an artificial and limited special case.

The remaining results of this section provide ways to generate additional examples of strongly outer functions.

Proposition 2.3.3. *If ϕ is strongly outer in $H^2(\mathbf{T}^2)$, then ϕ^α is strongly outer in $H^2(\mathbf{T}^2)$ for all $\alpha, 0 < \alpha < 1$.*

The theorem below is an immediate consequence of Proposition (2.3.3). It assures that there are many,examples of spectral factors ϕ meeting the requirements of the sections that follow.

Theorem 2.3.4. *Let ϕ be an outer function on the torus $\mathbf{T}^2$. If ϕ and $\frac{1}{\phi}$ belong to $H^{2+\varepsilon}(\mathbf{T}^2)$ for some $\varepsilon > 0$, then ϕ and $\frac{1}{\phi}$ are strongly outer in $H^2(\mathbf{T}^2)$.*

As an illustration, suppose that for some constants K_1 and K_2 the spectral density function $f(e^{i\lambda_1}, e^{i\lambda_2})$ satisfies

$$0 < K_1 \leq f(e^{i\lambda_1}, e^{i\lambda_2}) \leq K_2 < \infty, \tag{2.10}$$

throughout $\mathbf{T}^2$. Then $\log f$ is integrable, and it factorizes as $f = |\phi|^2$, where the outer function ϕ has analytic extension given by

$$\mathbf{f}(z,f) = c \exp \frac{1}{2} \int_{\mathbf{T}^2} \frac{e^{i\lambda_1} + ze^{i\lambda_2} + f}{e^{i\lambda_1} - ze^{i\lambda_2} - f} \log f(e^{i\lambda_1}, e^{i\lambda_2}) d\lambda(e^{i\lambda_1}) \times d\lambda(e^{i\lambda_2}).$$

By Theorem 2.3.4, ϕ is strongly outer, as was already found in [17, Theorem 4.5].

2.3.3 Missing values

Let us now consider the problem of predicting $X(0,0)$ based on the parameter set $\mathbf{Q}'$, where $\mathbf{Q}'$ is the quarter-plane $\mathbf{Q}$ with finitely many points missing. For some fixed $(h_1, h_2) \in \mathbf{Q}$, we can express $\mathbf{Q}'$as the disjoint union

$$\mathbf{Q}' = J \cup K \tag{2.11}$$

where $J = \{(s,t) \in \mathbf{Q} : s \geq h_1$ or $n \geq h_2\}$ and K is a subset of the finite set $\{(s,t) : 0 \leq s < h_1$ and $0 \leq s < h_2 \backslash (0,0)\}$.

Our objective is to obtain explicit formulas for the best prediction and prediction error variance. Similar to [7], our strategy will be first to project $X(0,0)$ onto the space generated by J; this is made possible by our previous work on $\mathbf{Q}$-invariance. Then we project $X(0,0)$ onto the span of the finitely many co-projections (relative to J) of the observations from K. The (orthogonal) sum of these two projections will be the best estimate of $X(0,0)$ based on $\mathbf{Q}' = J \cup K$. Figure 2.1 illustrates this support definition graphically.

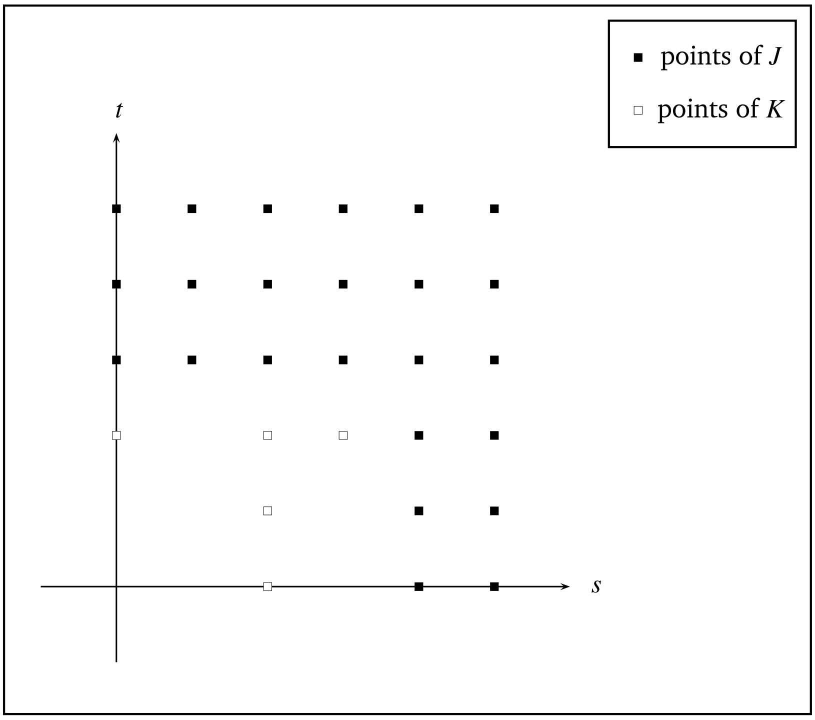


Figure 2.1: Incomplete quarter-plane support $\mathbf{Q}' = J \cup K$.

The projection of $X(0,0)$ onto $\overline{sp} = \{X(s,t) : (s,t) \in J\}$ is isomorphic to the projection of ϕ onto $\overline{sp} = \{\phi(e^{i\lambda_2}, e^{i\lambda_1})e^{i(s\lambda_1+t\lambda_2)}, (s,t) \in J\}$. But J is $\mathbf{Q}$-invariant, and thus by Proposition (2.3.2), this last span is equal to $\overline{sp} = \{e^{i(s\lambda_1+t\lambda_2)}, (s,t) \in J\}$.
The projection of ϕ on this space is simply $\sum\sum_{(s,t)\in J} c_{s,t} e^{i(s\lambda_1+t\lambda_2)}$. This tells us that the projection of $X(0,0)$ onto the subspace generated by J is

$$\sum\sum_{(s,t)\in J} b_{s,t}\varepsilon_{s,t} \tag{2.12}$$

where again $\{b_{s,t}\}$ are the moving average coefficients of $\{X(s,t)\}$ defined by (1.45), and $\varepsilon(s,t)$ is a Gaussian white noise.
Similarly, for each $(s,t) \in K$, the co-projection of $X(s,t)$ relative to the subspace generated by J is given by

$$Y(s,t) = \sum\sum_{\substack{s \le p < h_1 \\ t \le q < h_2}} b_{p-s,q-t}\varepsilon(p,q). \tag{2.13}$$

Then the projection of $X(0,0)$ onto the span of $\{Y(s,t) : (s,t) \in K\}$ is some linear combination

$$\sum\sum_{(s,t)\in K} c_{s,t} Y(s,t).$$

The coefficients $c_{s,t}$ can be determined as follows. Let A be the finite column vector indexed by K, with its (s,t) entry being $E(X(0,0)\overline{Y}(s,t))$. Let C be the finite square matrix with rows and columns each indexed by K, and with its entry in the (s,t)-row and (p,q)-column given by $E(Y(p,q)\overline{Y}(s,t))$. Take B to be the column vector, indexed by K, with entries $b_{p,q}$. Then $A = CB$.

We claim that C is full rank. To see this, let us suppose for the sake of argument that the vectors $\{Y(s,t) : (s,t) \in K\}$ are linearly dependent. Then let the elements of K be placed in lexicographical order $<$ by their coordinates, and let (M,N) be the smallest index under $<$ for which $Y(M,N)$ lies in the linear span of its successors. Then it follows that $X(M,N)$ lies in the subspace generated by $\{(p,q) \in \mathbb{Z}^2, (M,N) < (p,q)\}$. But this parameter set is the NSHP support, results from the total-order definition (1.25a) in the sense of Helson and Lowdenslager (see [12]). It would then follow Theorem 1.4.17, that $\log|\varphi|$ fails to be integrable, contradicting the outer property of φ. This affirms that $\{Y(s,t),\ (s,t) \in K\}$ are linearly independent.

Next, let $\{v(s,t),\ (s,t) \in K\}$ be any orthonormalization of $\{Y(s,t),\ (s,t) \in K\}$; thus each v_k can be written

$$v_{s,t} = \sum\sum_{(p,q)\in K} Y_{(s,t)(p,q)} Y_{(p,q)}.$$

The matrix T is necessarily full rank, and indeed, $\{Y(s,t),\ (s,t) \in K\}$ is isomorphic to the column vectors of T^{-1}. Finally, we have $C = (T^{-1})^* T^{-1}$, confirming that C is full rank. We may therefore compute the coefficient vector B by

$$B = C^{-1} A. \tag{2.14}$$

Let us summarize these findings

Theorem 2.3.5. *Let $X(s,t)$ be a centered, stationary Gaussian random field with strongly outer spectral factor ϕ. Let $\mathbf{Q}' = J \cup K$ be a quarter-plane with finitely many missing values, expressed as in* (2.11)*. Then the projection of $X(0,0)$ onto the subspace generated by $\mathbf{Q}'$ is given by*

$$\sum\sum_{(s,t)\in J} b_{s,t}\varepsilon_{s,t} \oplus \sum\sum_{(s,t)\in K} c_{s,t} Y(s,t).$$

where $\{b_{s,t}\}$ are the MA coefficients of $\{X(s,t)\}$ defined by (1.45)*, $\varepsilon_{s,t}$ is a Gaussian white noise, $Y(s,t)$ are the co-projections* (2.13)*, and the coefficients $c_{s,t}$ arise from* (2.14)*. The resulting prediction error variance is*

$$\left(\sum\sum_{\substack{0\le s<k\\0\le t<l}} |b_{s,t}|^2\right) - A^* C^{-1} A, \tag{2.15}$$

where A^* is the Hermitian conjugate of A. To see this, note that the co-projection of $X(0,0)$ onto the space for $\mathbf{Q}'$ is given by

$$\begin{aligned} X(0,0) - \left[\sum\sum_{(s,t)\in J} b_{s,t}\varepsilon(s,t) \oplus \sum\sum_{(s,t)\in K} c_{s,t} Y(s,t)\right] &= \sum\sum_{\mathbf{Q}\cup\{(0,0)\}} b_{s,t}\varepsilon(s,t) - \left[\sum\sum_{(s,t)\in J} b_{s,t}\varepsilon(s,t) \oplus \sum\sum_{(s,t)\in K} c_{s,t} Y(s,t)\right] \\ &= \sum\sum_{\substack{0\le s<k\\0\le t<l}} b_{s,t}\varepsilon(s,t) - \sum\sum_{(s,t)\in K} c_{s,t} Y(s,t). \end{aligned}$$

The variance of this expression is

$$\begin{aligned} E\left|\sum\sum_{\substack{0\le s<k\\0\le t<l}} b_{s,t}\varepsilon(s,t) - \sum\sum_{(s,t)\in K} c_{s,t} Y(s,t)\right|^2 &= \sum\sum_{\substack{0\le s<k\\0\le t<l}} |b_{s,t}|^2 - 2\mathscr{R}\left(\sum\sum_{\substack{0\le s<k\\0\le t<l}} \sum\sum_{(i,j)\in K} b_{s,t}\bar{b}_{i,j} E(\varepsilon(s,t)\bar{Y}(i,j))\right) \\ &+ \sum\sum_{(s,t)\in K} \sum\sum_{(i,j)\in K} c_{s,t}\bar{b}_{i,j}, \end{aligned}$$

where $\mathscr{R}(z)$ stands for the real part of z. Notice that the middle term is unchanged if the first sum is taken over all $(s,t) \in \mathbf{Q} \cup (0,0)$. This is because each $Y(i,j)$ is orthogonal to all the additional terms. Also, the *expected values appearing in the sums can be expressed in terms of A, B and C. The variance calculation can then continue as*

$$\begin{aligned}
&\sum_{\substack{0\le s<k\\0\le t<l}}\sum |b_{s,t}|^2 - 2\mathscr{R}\left(\sum_{\substack{0\le s<k\\0\le t<l}}\sum \sum_{(i,j)\in K}\sum b_{s,t}\bar{b}_{i,j}E(\varepsilon_{s,t}\bar{Y}_{i,j})\right) + \sum_{(s,t)\in K}\sum \sum_{(i,j)\in K}\sum c_{s,t}\bar{b}_{i,j}\\
&\quad= \sum_{\substack{0\le s<k\\0\le t<l}}\sum |b_{s,t}|^2 2\mathscr{R}\left(\sum_{(i,j)\in K}\sum \bar{b}_{i,j}E(X_{0,0}\bar{Y}_{i,j})\right) + B^*CB\\
&\quad= \sum_{\substack{0\le s<k\\0\le t<l}}\sum |b_{s,t}|^2 - 2\mathscr{R}(B^*A) + (A^*C^{-1})C(C^{-1}A)\\
&\quad= \sum_{\substack{0\le s<k\\0\le t<l}}\sum |b_{s,t}|^2 - 2\mathscr{R}(A^*C^{-1}A) + A^*C^{-1}A\\
&\quad= \sum_{\substack{0\le s<k\\0\le t<l}}\sum |b_{s,t}|^2 - A^*C^{-1}A
\end{aligned}$$

which confirms (2.15).

2.3.4 Duality

The second solution to the problem of prediction with respect to a quarter-plane with missing values is presented here. A related "dual" prediction problem with added observations is solved, and it is shown that the resulting prediction error variances are reciprocals of each other. A particular challenge in this situation is that the complement of a quarter-plane is not a quarter-plane, and so duality arguments from one variable do not carry over directly. This solution requires both the spectral factor Ï† and its reciprocal to be strongly outer in $H^2(\mathbf{T}^2)$. As noted in [7], the duality argument from the one-parameter case (see [47]) does not extend directly, since the complement of quarter-plane $\mathbf{Q}$ does not coincide with its reflection over the origin. Suitably modified, however, this approach yields a useful duality relationship for prediction variance formulas, similar to that in [47]. A further result is a prediction error variance formula in terms of the AR coefficients of the random field.

Theorem 2.3.6. *Let ϕ be outer and let $\frac{1}{\phi}$ be strongly outer in $H^2(\mathbf{T}^2)$, and let the parameter set $\mathbf{Q}'$ be the quarter-plane with finitely missing values. Then,*

$$\inf\left\{\|1+P\|^2_{L^2(f)} : P \in \mathbf{Q}'\right\} = \left[\inf\left\{\|1+P\|^2_{L^2(\frac{1}{f})} : P \in \left(\mathbf{Q}'\right)^c\right\}\right]^{-1}, \tag{2.16}$$

where $f = \phi^2$.

Let us now turn to the dual problem of computing the right hand side of equation (2.16). Assume that $\frac{1}{\phi}$ is strongly outer in $H^2(\mathbf{T}^2)$, and let $\{Z(s,t)\}$ be the random field

$$\sum_{(j,k)\in-\mathbf{Q}}\sum \bar{a}_{j,k}\varepsilon(j+m,k+n)$$

Then the mapping $Z(s,t) \mapsto e^{i(s\lambda_1+t\lambda_2)}/\bar{\phi}(e^{i\lambda_1},e^{i\lambda_2})$ induces a Hilbert space isomorphism from $L^2(\mathbf{P})$ to $L^2(d\lambda^2)$. The functions $\{e^{i(s\lambda_1+t\lambda_2)}, (s,t)\in\mathbb{Z}^2\}$ correspond to a Gaussian white noise. Since $\frac{1}{\phi}$ is strongly outer, and $-\mathbf{O}$ is $\mathbf{O}$-invariant Proposition 2.3.2 provides that

$$\overline{sp}\{e^{i(s\lambda_1+t\lambda_2)}\backslash\bar{\phi}(e^{i\lambda_1},e^{i\lambda_2}) : (s,t)\in\mathbf{Q}^c\} = \overline{sp}\{e^{i(s\lambda_1+t\lambda_2)} : (s,t)\in\mathbf{Q}\}.$$

The expression $\inf\left\{\|1+P\|^2_{L^2(\frac{1}{f})} : P \in (\mathbf{Q}')^c\right\}$ is the prediction error variance of estimating $Z(0,0)$ based on the disjoint union $\mathbf{Q}\cup\{(0,0)$. Proceeding as in the derivation of Theorem 2.3.5, let us project $Z(0,0)$ separately onto the subspace induced by $\mathbf{Q}^c$, and onto the span of the co-projections $W(s,t)$ of $Z(s,t)$ with respect to $\mathbf{Q}^c$, for all $(s,t)\in I$. Thus, for each

$$\sum\sum_{\substack{0\le p\le s\\ 0\le q\le t}} a_{s-p,t-q}\varepsilon(p,q) \tag{2.17}$$

The projection of $Z(0,0)$ onto $\{Z(s,t) : (s,t)\in \mathbf{Q}^c\}$ is given simply by

$$\sum\sum_{(j,k)\in\mathbf{Q}} \bar{a}_{-j,-k}\varepsilon(p,q).$$

Since the coefficients of $\frac{1}{\phi}$ vanish outside of $-\mathbf{Q}\cup\{(0,0)$. The projection of $Z(0,0)$ onto the span of $\{W(s,t) : (s,t)\in I\}$ can be expressed as

$$\sum\sum_{(s,t)\in I}\beta_{s,t}W(s,t).$$

Again the coefficients $\beta_{s,t}$ can be obtained from the equation $A = CB$; in this case, B is the column vector of the coefficients $\beta_{s,t}$; A is the column vector with entries $A_{s,t} = E(Z(0,0)W(s,t))$, and C is the square matrix with entries

$$C_{(m,n),(p,q)} = E(W(p,q)W(s,t)).$$

These vectors and matrices are all finite, with respective rows and columns indexed by I. Again, the outer property of $\frac{1}{\phi}$ ensures that the matrix C is full rank, and we may obtain B via

$$B = C^{-1}A. \tag{2.18}$$

We thus obtain the following result

Theorem 2.3.7. *Let $\{Z(s,t)\}$ be a centered, stationary Gaussian random field with strongly outer spectral factor $\frac{1}{\phi}$. Let I be a subset of $\{(s,t) : 0\le s<h, 0\le t<k\}$. Then the projection of $Z(0,0)$ onto the subspace generated by $\mathbf{Q}^c\cup I$ is given by*

$$\sum\sum_{(j,k)\in\mathbf{Q}}\bar{a}_{-j,-k}\varepsilon(j,k)\bigoplus\sum\sum_{(s,t)\in I}\beta_{s,t}W(s,t) \tag{2.19}$$

where $\{a_{k,l}\}$ are the AR coefficients of $\{X(s,t)\}$ define by (1.46)*, $\{\varepsilon(s,t)\}$ is a Gaussian white noise, $\{W(s,t)\}$ are the co-projections* (2.17)*, and the coefficients $\beta_{s,t}$ arise from* (2.18)*. The associated prediction error variance is given by $|a^2_{0,0}A^*a^{-1}A|$. To see this, start with the prediction error vector*

$$\begin{aligned} Z_{0,0} &- \sum\sum_{(j,k)\in\mathbf{Q}}\bar{a}_{-j,-k}\varepsilon(j,k)\bigoplus\sum\sum_{(s,t)\in I}\beta_{s,t}W_{s,t}\\ &= \sum\sum_{(j,k)\in-\mathbf{Q}\cup\{(0,0)\}}\bar{a}_{-j,-k}\varepsilon(j,k) - \left[\sum\sum_{(j,k)\in\mathbf{Q}}\bar{a}_{-j,-k}\varepsilon(j,k)\bigoplus\sum\sum_{(s,t)\in I}\beta_{s,t}W_{s,t}\right]\\ &= \bar{a}_{0,0}\varepsilon(0,0) - \sum\sum_{(s,t)\in I}\beta_{s,t}W_{s,t}. \end{aligned}$$

This has variance given by

$$E|\bar{a}_{0,0}\varepsilon(0,0)-\sum\sum_{(s,t)\in I}\beta_{s,t}W_{s,t}|^2 = |\bar{a}_{0,0}|^2 - 2\mathscr{R}\left(\sum\sum_{(s,t)\in I}\bar{a}_{0,0}\bar{\beta}_{s,t}E(\varepsilon(0,0)\bar{W}(s,t))\right)+\sum\sum_{(s,t)\in I}\sum\sum_{(i,j)\in I}\beta_{s,t}\bar{\beta}_{i,j}E(W_{s,t}\bar{W}_{i,j}).$$

2.4 Prediction with incomplete quarter-plane past

2.4.1 Introduction

Prediction of stationary random fields with incomplete quarter-plane past has been studied in [9]. An explicit formula for the prediction error variance of a future value of a weakly stationary random field, when the infinite past is altered by some missing observations is given. This explicit formula allows us to derive the AR representation of $P_{H^X_{Q_1}}X(0,0)$. Let us introduce some notations. Let $\{X(s,t),\,(s,t)\in\mathbb{Z}^2\}$ be a zero mean weakly stationary random field with spectral density f define by (1.34). We denote by $P_{H^X_Q}X(0,0)$ the best linear mean-square predictor of $X(0,0)$ based on the quarter-plane Q. Assume that the data $X(-n_1,-m_1),X(-n_2,-m_2),...,X(-n_N,-m_N)$ are missing in the past, and we set

$$\mathscr{M}=\{(n_1,m_1),...,(n_N,m_N),\,n_i\geq 0,m_i\geq 0,(n_i,m_i)\neq(0,0)\}.$$

We denote by $-\mathscr{M}$ the set of indices of the missing variables in the quadrant Q and by Q_1 the set of indices points constituting the observed past, namely, $Q_1=Q\backslash\{-\mathscr{M}\}$. Figure 2.2 illustrates this support definition graphically.

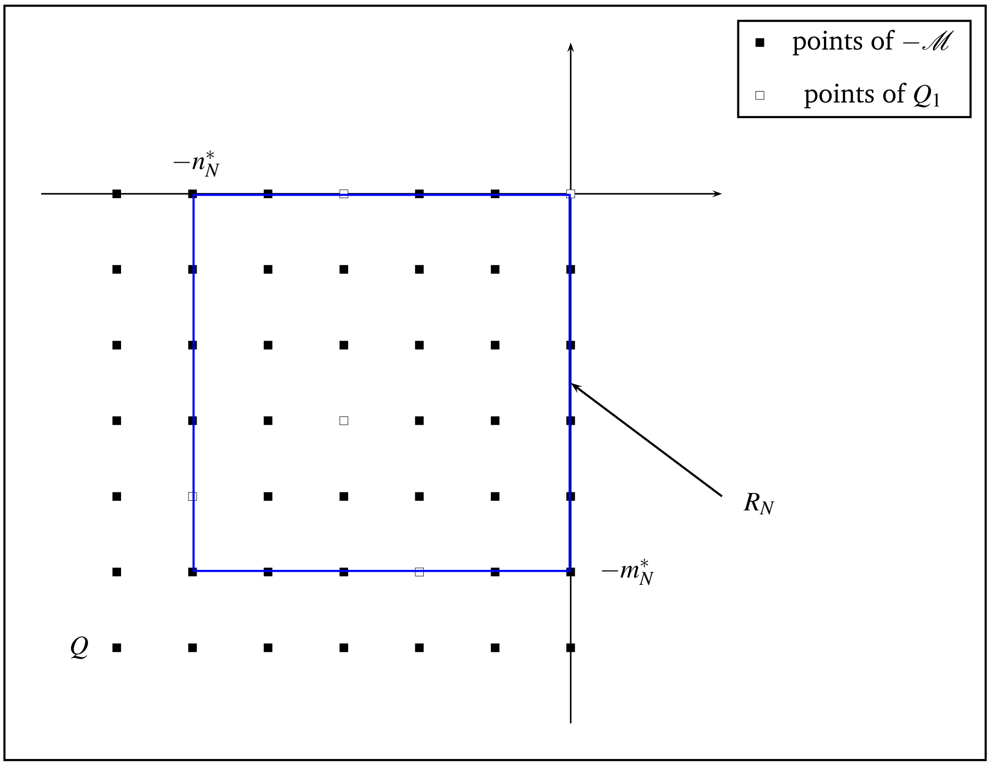


Figure 2.2: Incomplete quarter-plane support $Q_1=Q\backslash\{-\mathscr{M}\}$

Let $n_N^*=\max\limits_{1\leq i\leq N}n_i, m_N^*=\max\limits_{1\leq i\leq N}m_i$. Let $P_{H^X_{Q_1}}X(0,0)$ be the best linear mean square predictor of $X(0,0)$ based on the incomplete past $\{X(s,t),\,(s,t)\in Q_1=Q\backslash\{-\mathscr{M}\}\}$.

2.4.2 Prediction with incomplete past

Using the above notations, we are able to establish an explicit formula for the prediction error of a future value of weakly stationary random fields when the infinite past is altered by some missing observations.

Theorem 2.4.1 ([9]). *Let $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ be a zero mean weakly stationary random field satisfying the conditions of Theorem* (1.4.17) *with the innovation process $\{\varepsilon(s,t);(s,t)\in\mathbb{Z}^2\}$ and the AR parameters $\{a_{k,l}\}$ define by* (1.46). *Then*

$$X(0,0)-P_{H^X_{Q_1}}X(0,0)=-\sum_{p=0}^{N}\sum_{q=0}^{N}\psi_{p,q}\sum_{i=0}^{n_p}\sum_{j=0}^{n_q}a_{n_p-i,n_q-j}\varepsilon(-i,-j), \tag{2.20}$$

where the coefficients $\psi_{p,q}$ satisfy the matrix equation

$$U(\psi_{0,0},\,\psi_{0,1},\ldots.,\,\psi_{1,0},\,\ldots,\,\psi_{N,N})^t=(1,0,\ldots,0)^t. \tag{2.21}$$

U being the nonsingular $(N+1)\times(N+1)$ matrix with elements

$$U_{p,q}=\sum_{i=0}^{n_p\wedge n_q}\sum_{j=0}^{n'_p\wedge n'_q}a_{n_p-i,n_q-j}\varepsilon(-i,-j)a_{n_p-i,n_q-j}\varepsilon(-i,-j)\ \ p,p',q,q'=0,\ldots,N \tag{2.22}$$

where $k\wedge n$ stands for the minimum of k and n. The predictor error variance is

$$var\left(X(0,0)-P_{H^X_{Q_1}}X(0,0)\right)=\sigma^2\psi_{0,0}. \tag{2.23}$$

Remark 2. *Assume that $0\le n_i\le N;0\le m_j\le N,\ (n_i,m_j)\neq(0,0)$,so that $P_{H^X_{Q_1}}X(0,0)$ is the $(N+1)\times(N+1)$-step predictor of $X(0,0)$. It follows from Equation* (2.22) *that $U=A'A$, and since $A^{-1}=-C$, Equation* (2.21) *is equivalent to*

$$A(\psi_{00},\ldots,\psi_{NN})^t=-(b_{00},\ldots,b_{NN})^t,$$

we deduce from Equations (2.20) *and* (2) *that*

$$X(0,0)-P_{H^X_{Q_1}}X(0,0)=-\sum_{i=0}^{N}\sum_{j=0}^{N}\left(\sum_{p=0}^{N}\sum_{q=0}^{N}\psi_{pq}a_{n_p-i,n_q-j}\right)\varepsilon(-i,-j)=\sum_{i=0}^{N}\sum_{j=0}^{N}b_{ij}\varepsilon(-i,-j) \tag{2.24}$$

Since the innovation process $\{\varepsilon(n,m)\}$ is not directly observable, Equation (2.20) *can not be used to calculate $P_{H^X_{Q_1}}X(0,0)$ unless one can express the innovation $\varepsilon(k,l)$ in term of the observations $X(k-i,l-j);i\ge 0,j\ge 0$. This is equivalent to finding a mean square convergent AR series representation for $P_{H^X_{Q_1}}X(0,0)$. In the following theorem, we establish such representation.*

Theorem 2.4.2 ([9]). *Let $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ be a non deterministic zero mean stationary random field satisfying the conditions of Theorem 1.4.17 with AR parameters $\{a_{s,t}\}$. The predictor $P_{H^X_{Q_1}}X(0,0)$ has an AR representation for any finite set of missing data iff $\{X(s,t)\}$ has the AR representation* (1.48). *In this case, the AR representation of $P_{H^X_{Q_1}}X(0,0)$ is unique and is given by*

$$P_{H^X_{Q_1}}X(0,0)=\sum\sum_{\substack{(k,l)\notin\mathbb{N}^2\setminus\mathscr{M}\\(k,l)\neq(0,0)}}h_{k,l}X(-k,-l), \tag{2.25}$$

where

$$h_{k,l} = \delta_{k,l} - \sum_{p=0}^{N}\sum_{q=0}^{N}\psi_{p,q}\sum_{i=0}^{n_p\wedge k}\sum_{j=0}^{n_q\wedge l} a_{n_p-i,n_q-j}\varepsilon(-i,-j)a_{n_p-i,n_q-j}\varepsilon(-i,-j),\ (k,l)\in\mathbb{N}^2. \tag{2.26}$$

and the coefficients $\psi_{p,q}$ are defined in Theorem 2.4.1.

Remark 3. *According to Equation* (2.26)*, h_{00}, and for any $(k,l), k\geq 0, l\geq 0, (k,l)\neq(0,0))$, we have*

$$h_{k,l} = -\sum_{\substack{i=0\\(i,j)\neq(0,0)}}^{n_N^*}\sum_{j=0}^{m_N^*}\alpha_{ij}a_{k-i,l-j}, \tag{2.27}$$

where $\alpha_{ij} = \sum_{p=0}^{N}\sum_{q=0}^{N}\psi_{p,q}a_{n_p-i,n_q-j}$. Therefore,

$$h_{k,l}^2 \leq \left(\sum_{i=0}^{n_N^*}\sum_{j=0}^{m_N^*}\alpha_{ij}^2\right)\left(\sum_{i=0}^{n_N^*}\sum_{j=0}^{m_N^*}a_{k-i,l-j}^2\right), \tag{2.28}$$

for any $(k,l)\in\mathbb{N}^2$.

According to Equation (2.23), the increase in variance of the prediction error of $X(0,0)$ due to the missing data $X(-n_1,-m_1),\ldots,X(-n_N,-m_N)$, is equal to $\sigma^2(\psi_{0,0}-1)$. In the following theorem, we characterize the processes for which the missing observations do not affect the prediction of $X(0,0)$.

Theorem 2.4.3 ([9])**.** *Let $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$ be a non deterministic weakly stationary process with the AR parameters $(a_{k,l})$. Then*

$$P_{H_{Q_1}^X}X(0,0) = P_{H_Q^X}X(0,0) \text{ iff } a_{n_i,m_i}=0 \quad \text{for } (n_i,m_i)\in\mathscr{M}.$$

Remark 4. *Under the assumption that $\{X(s,t),(s,t)\in\mathbb{Z}^2\}$ has the AR representation Equation* (1.48)*, Theorem 2.4.3 is easily proved as follows. If $P_{H_{Q_1}^X}X(0,0) = P_{H_Q^X}X(0,0)$, we deduce from Equations* (2.25) *and* (1.48) *that*

$$\sum_{\substack{(k,l)\notin\mathbb{N}\setminus\mathscr{M}\\(k,l)\neq(0,0)}}\sum h_{k,l}X(-k,-l) = \sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{\infty}\sum_{l=0}^{\infty}h_{k,l}X(-k,-l), \tag{2.29}$$

which implies that $a_{k,l}=0$ for all $(k,l)\in\mathscr{M}$. Conversely, if $a_{k,l}=0$ for all $(k,l)\in\mathscr{M}\setminus\{(0,0)\}$,

$$P_{H_Q^X}X(0,0) = \sum_{(k,l)\notin\mathbb{N}\setminus\mathscr{M}}\sum a_{k,l}X(-k,-l).$$

and thus, $P_{H_Q^X}X(0,0)\in\overline{sp} = \{X(s,t),(s,t)\in Q_1 = Q\setminus\{-\mathscr{M}\}\}$ and $X(0,0)-P_{H_Q^X}X(0,0)\perp\overline{sp} = \{X(s,t),(s,t)\in Q\}\supset\overline{sp} = \{X(s,t),(s,t)\in Q_1 = Q\setminus\{-\mathscr{M}\}\}$. Therefore, $P_{H_Q^X}X(0,0) = P_{H_{Q_1}^X}X(0,0)$.

To summarize, an explicit formula for the prediction error variance of a future value of a weakly stationary random field is established in [9], when the infinite past, is altered by some missing observations. This explicit formula allows us to derive the AR representation of $P_{H_{Q_1}^X}X(0,0)$. It should be emphasized that this representation can be seen as an alternative solution to the problem posed in [7].

2.5 Impact of missing data on the prediction of random fields

2.5.1 Introduction

Our first contribution (see [10]) is detailed in this section. We aim to quantify the impact of missing observations from the past. The central idea of our study consists in using the MA and AR representations of the random field. The obtained results highlight the important role of the AR parameters in forecasting. Indeed, we establish lower and upper bounds for the prediction error variance given in Theorem 2.1 which is the novelty of our work. This boundedness property of prediction error variance shows that the degradation of the prediction due to the missing data increases with the maximum value of the AR parameters of the missing data. They also allow to conclude that the larger the indices of missing values are, the better is the preciseness of the bound of the prediction error variance. Also, our results characterize the random fields for which the missing observations do not affect the prediction of $X(0,0)$. Finally, note that $X(0,0)$ is chosen without loss of generality, and the main conclusions extend, naturally, if we consider the prediction of $X(h_1,h_2), h_1,h_2 > 0$.

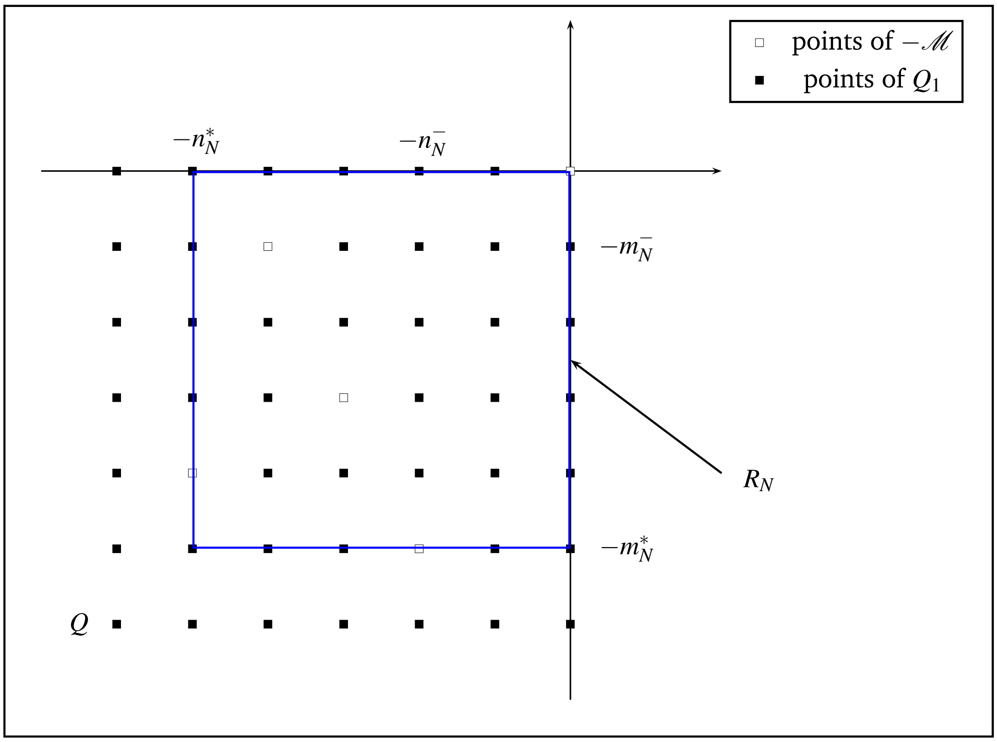
In order to clarify our contribution with respect to previous works, we begin by collecting some needed material and establish the necessary notation and assumptions on weakly stationary random fields for this section. Assume that the data $X(-n_1,-m_1), X(-n_2,-m_2), ..., X(-n_N,-m_N)$ are missing in the past, and we set

$$\mathscr{M} = \{(n_1,m_1),...,(n_N,m_N); n_i \geq 0, m_i \geq 0, (n_i,m_i) \neq (0,0)\}.$$

We denote by $-\mathscr{M}$ the set of indices of the missing variables in the quadrant Q and by Q_1 the set of indices points constituting the observed past, namely,

$$Q_1 = Q \backslash \{-\mathscr{M}\},$$

and $n_N^* = \max_{1 \leq i \leq N} n_i$, $m_N^* = \max_{1 \leq i \leq N} m_i$, $n_{\bar{N}} = \min_{1 \leq i \leq N} n_i$, $m_{\bar{N}} = \min_{1 \leq i \leq N} m_i$. The following figure illustrates this support definition graphically.



In what follows, if S is an arbitrary nonempty family of random fields in L_2, the subspace of all (finite) linear combinations of elements of S is denoted by $sp\{S\}$ and its closure in L_2 by $\overline{sp}\{S\}$. For any $(k,l) \in \mathbb{Z}^2$, we set

$$\begin{aligned} H_{k,l} &= \overline{sp}\{X(i,j); i \le k, j \le l\} \\ H_{-\infty,-\infty} &= \bigcap_{(k,l)\in\mathbb{Z}^2} H_{k,l}. \end{aligned}$$

Let $P_{H_Q^X} X(0,0)$ be the best linear mean-square predictor of $X(0,0)$ based on $H_{0,0} = sp\{X(i,j);\ (i,j) \in Q\}$ and $P_{H_{Q_1}^X} X(0,0)$ the best linear mean square predictor of $X(0,0)$ based on the incomplete past $H'_{0,0} = \overline{sp}\{X(n,m);\ (n,m) \in Q_1 = Q \backslash \{-\mathscr{M}\}\}$. In the case, where $0 \le n_i, m_i = N$, $(n_i, m_i) \ne (0,0)$, $P_{H_{Q_1}^X} X(0,0)$ is the $(N+1, N+1)$ is the $(N+1,N+1)-$step predictor of $X(0,0)$ which is easily calculated from the moving average representation of $X(k,l)$, (see [7]). In [8], two solutions have been presented for predicting a stationary Gaussian random field based on a quater-plane with finitely missing observations. One expresses the best predictor in terms of the moving average coefficients, and the other expresses the prediction error variance in terms of its autoregressive coefficients. When $(n_1,m_1),\ldots,(n_N,m_N)$ are arbitrary in $\mathbb{N}^2|(0,0)$, let K be the number of the observed values located in the rectangle R_N delimited by the points $(0,0)$, $(-n_N^*, 0)$, $(0, -m_N^*)$ and. $(-n_N, -m_N^*)$. We have $K = (n_N^* + 1) \times (m_N^* + 1) - N - 1$. Let $((-k_1, -l_1), \ldots, (-k_K, -l_K))$ be the indices corresponding to the K observed values in the rectangle R_N. The observation space is decomposed as the orthogonal sum

$$\overline{sp}\{X(k,l); (k,l) \in Q_1\} \quad = \quad \overline{sp}\{X(k,l); k \le -n_N^*, l \le -m_N^*;\ (k,l) \ne (-n_N^*, -m_N^*)\} + sp\{Y_1, \ldots, Y_K\},$$

where $Y_i = X(-n_i, -m_i) - \tilde{X}(-n_i, -m_i)$, $i = 1, \ldots K$ and $\tilde{X}(-n_i, -m_i) = P_{H_{-n_N^*, -m_N^*}} X(-n_i, -m_i)$ is the orthogonal projection of $\tilde{X}(-n_i, -m_i)$ onto

$$H_{-n_N^*, -m_N^*} = \overline{sp}\{X(k,l); k \le -n_N^*, l \le -m_N^*;\ (k,l) \ne (-n_N^*, -m_N^*)\}.$$

Next, the space $sp\{Y(i) = X(n_i,m_i) - \tilde{X}(ni,mi); (ni,mi) \in Q_1\}$ is decomposed into the sum of K orthogonal one dimensional subspaces by means of the Gram-Schmidt procedure. Hence, the observation space is decomposed as the $K+1$ orthogonal subspaces, and $\hat{X}'(0,0)$ is the sum of the projections of $X(0,0)$ onto each subspace. For $(h_1,h_2) \in Q^c$, where Q^c is the complement of Q, the (h_1,h_2)-step ahead mean square predictor of $X(h_1,h_2)$ based on the past $H_{0,0}$ is given by

$$\hat{X}(h_1,h_2) = \sum_{\substack{k=h_1 \\ (k,l)\neq(h_1,h_2)}}^{\infty} \sum_{l=h_2}^{\infty} b_{k,l}\varepsilon(h_1-k,h_2-l), \tag{2.30}$$

with the corresponding error

$$\begin{aligned} X(h_1,h_2) - \hat{X}(h_1,h_2) &= b_{h_1,h_2}\varepsilon(0,0) + \sum_{k=0}^{h_1-1}\sum_{l=h_2}^{\infty} b_{k,l}\varepsilon(h_1-k,h_2-l) \\ &\quad + \sum_{k=h_1}^{\infty}\sum_{l=0}^{h_2-1} b_{k,l}\varepsilon(h_1-k,h_2-l) + \sum_{k=0}^{h_1-1}\sum_{l=0}^{h_2-1} b_{k,l}\varepsilon(h_1-k,h_2-l). \end{aligned}$$

The predictor error variance of $X(h_1,h_2)$ is $Var\left(X(h_1,h_2) - \hat{X}(h_1,h_2)\right) = \sigma^2 V$ with

$$V = |b_{h_1,h_2}|^2 + \sum_{k=0}^{h_1-1}\sum_{l=h_2}^{\infty} |b_{k,l}|^2 + \sum_{k=h_1}^{\infty}\sum_{l=0}^{h_2-1} |b_{k,l}|^2 + \sum_{k=0}^{h_1-1}\sum_{l=0}^{h_2-1} |b_{k,l}|^2. \tag{2.31}$$

For $(s,t) \in \mathbb{Z}^2$, we set

$$I_{s,t} = \overline{sp}\{X(i,j); (i,j) \neq (s,t)\},$$

the past and the future of the random field $\{X(s,t)\}$ at (s,t). $P_{I_{s,t}}$ denotes the orthogonal projection onto $I_{s,t}$, and $\eta(s,t) = X(s,t) - P_{I_{s,t}}X(s,t)$.. The random field $\{X(s,t)\}$ is zero mean, weakly stationary, and is called the interpolation error or the two-sided innovation of $X(s,t)$. We set $\sigma_n = \|\eta(s,t)\|$, $\sigma_\varepsilon = \|\varepsilon(s,t)\|$ and $\sigma_x = \|X(s,t)\|$. The random field $X(s,t), (s,t) \in \mathbb{Z}^2$ is said to be minimal if $\eta(s,t) \neq 0$ for some $(s,t) \in \mathbb{Z}^2$. Since $H_{s,t} \subset I_{s,t}$, every minimal stationary random field is nondeterministic. In this case, we have

$$\sigma_\eta^2 = \frac{\sigma_\varepsilon}{\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_{i,j}^2}.$$

2.5.2 Lower and upper bounds for the prediction error variance

The aim of this section is to quantify the influence of the missing data $X(-n_1,-m_1),(-n_2,-m_2),$ $...,X(-n_N,-m_N)$ on the linear prediction of $X(0,0)$. Minimum and maximum values of the prediction error variance are obtained. These values allow us to characterize the lower and upper bounds for $\|X(0,0) - \hat{X}'(0,0)\|$ given in 2.5.1 Note that $P_{H^X_{Q_1}}X(0,0) \in H'_{0,0} \subset H_{0,0}$ and $\varepsilon(0,0) \perp H_{0,0}$. Consequently, $P_{H^X_Q}X(0,0) - P_{H^X_{Q_1}}X(0,0) \perp \varepsilon(0,0)$, which entails

$$\begin{aligned} \|X(0,0) - P_{H^X_{Q_1}}X(0,0)\|^2 &= \|\varepsilon(0,0) - X(0,0) - P_{H^X_{Q_1}}X(0,0)\|^2 \\ &= \sigma_\varepsilon^2 \|X(0,0) - P_{H^X_{Q_1}}X(0,0)\|^2. \end{aligned}$$

Theorem 2.5.1 ([10]). *Let $\{X(s,t);\,(s,t)\in\mathbb{Z}^2\}$ be a weakly stationary random field with the innovation process $\{\varepsilon(s,t)\}$, the MA parameters $\{b_{s,t}\}$ and the AR parameters $\{a_{s,t}\}$ Equation (1.46) .Then,*

$$\begin{aligned}\sigma_\eta \max_{(i,j)\in M}|a_{i,j}| &\leq \|P_{H_Q^X}X(0,0)-P_{H_{Q_1}^X}X(0,0)\| \\ &\leq \sigma_\varepsilon \sum\sum_{(i,j)\in\mathscr{M}}\varphi_{i,j}|a_{i,j}| \leq \sigma_X \sum\sum_{(i,j)\in\mathscr{M}}|a_{i,j}|, \end{aligned} \tag{2.32}$$

where

$$\varphi_{s,t}=\left(\sum_{k=0}^{-n_N^*+s}\sum_{l=0}^{-m_N^*+t}b_{k,l}^2+\sum_{k=0}^{-n_N^*+s}\sum_{l=-m_N^*+t+1}^{+\infty}b_{k,l}^2+\sum_{k=n_N^*+s+1}^{+\infty}\sum_{l=0}^{m^*+1}b_{k,l}^2\right)^{\frac{1}{2}}.$$

Proof: According to (1.45), we have

$$\begin{aligned}X(0,0)-\sum_{\substack{i=0\\(i,j)\neq(0,0)}}^{n_N^*}\sum_{j=0}^{m_N^*}a_{i,j}X(-i,-j) &= -\sum_{i=0}^{n_N^*}\sum_{j=0}^{m_N^*}a_{i,j}X(-i,-j)\\ &= -\sum_{i=0}^{n_N^*}\sum_{j=0}^{m_N^*}a_{i,j}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}b_{p,q}\varepsilon(-i-p,-j-q)\\ &= \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}S_{k,l}\varepsilon(-k,-l),\end{aligned}$$

with

$$S_{k,l}=\sum_{i=0}^{\min(n_N^*,k)}\sum_{j=0}^{\min(m_N^*,l)}a_{i,j}b_{k-i,l-j}.$$

According to (1.47), $S_{0,0}=-1$ and $S_{i,j}=0$ for $(i,j)\in R_{k,l}=\{(i,j)\in\mathbb{Z}^2;\,0\leq i\leq k, 0\leq j\leq l;\,(i,j)\neq(0,0)\}$. Therefore,

$$X(0,0)=\sum_{\substack{i=0\\(i,j)\neq(0,0)}}^{n_N^*}\sum_{j=0}^{m_N^*}a_{i,j}X(-i,-j)+\varepsilon(0,0)+U_{0,0},$$

where

$$U_{0,0}=-\sum_{\substack{k=n_N^*\\(i,j)\neq(0,0)}}^{\infty}S_{k,l}\varepsilon(-k,-l)\in H_{n_N^*,m_N^*}=\overline{sp}\{X(i,j);\,i\leq -n_N^*;\,j\leq -m_N^*,\,(i,j)\neq(n_N^*,m_N^*)\}.$$

Consequently,

$$\begin{aligned}P_{H_Q^X}X(0,0) &= \sum_{\substack{i=0\\(i,j)\neq(0,0)}}^{n_N^*}\sum_{j=0}^{m_N^*}a_{i,j}X(-i,-j)+U_{0,0}\\ P_{H_{Q_1}^X}X(0,0) &= \sum_{\substack{i=0\\(i,j)\neq(0,0)}}^{n_N^*}\sum_{j=0}^{m_N^*}a_{i,j}X(-i,-j)+\sum\sum_{(i,j)\in\mathscr{M}}a_{i,j}P_{H_{Q_1}^X}X(-i,-j)+U_{0,0}\end{aligned}$$

and

$$P_{H_Q^X}X(0,0)-P_{H_{Q_1}^X}X(0,0)=\sum\sum_{(i,j)\in\mathscr{M}}a_{i,j}\left(P_{H_Q^X}X(-i,-j)-P_{H_{Q_1}^X}X(-i,-j)\right). \tag{2.33}$$

Since $H_{-n_N^*,-m_N^*} \subset H'_{0,0}$, we have

$$\begin{aligned}\left\| \sum\sum_{(i,j)\in\mathscr{M}} a_{i,j}\left(\hat{X}(-i,-j) - P_{H^X_{Q_1}} X(-i,-j)\right)\right\| &\leq \sum\sum_{(i,j)\in\mathscr{M}} |a_{i,j}| \left\|\hat{X}(-i,-j) - P_{H^X_{Q_1}} X(-i,-j)\right\| \\ &\leq \sum\sum_{(i,j)\in\mathscr{M}} |a_{i,j}| \left\|\hat{X}(-i,-j) - P_{H_{-n_N^*,-m_N^*}} P_{H^X_{Q_1}} X(-i,-j)\right\|. \quad (2.34)\end{aligned}$$

According to (1.47) again, for any $(s,t)\in\mathbb{Z}^2$, we have

$$P_{H_{-n_N^*,-m_N^*}} X(s,t) \sum_{\substack{i=n_N^*+s\\(k,l)\neq(n_N^*+s,m_N^*+t)}}^{\infty} \sum_{j=m_N^*+t}^{\infty} b_{k,l}\varepsilon(s-k,t-l)$$

$$\begin{aligned}X(s,t) - P_{H_{-n_N^*,-m_N^*}} X(s,t) = &\sum_{k=0}^{n_N^*+s}\sum_{l=0}^{m_N^*+t} b_{k,l}\varepsilon(s-k,t-l) + \sum_{k=0}^{n_N^*+s}\sum_{l=m_N^*+t+1}^{+\infty} b_{k,l}\varepsilon(s-k,t-l) \\ &+ \sum_{k=n_N^*+s+1}^{+\infty}\sum_{l=0}^{m_N^*+t} b_{k,l}\varepsilon(s-k,t-l).\end{aligned}$$

Hence, $\left\|\hat{X}(-i,-j) - P_{H_{-n_N^*,-m_N^*}} X(-i,-j)\right\| = \sigma_\varepsilon \varphi_{i,j}$, and the first upper bound in (2.32) follows by using (2.33) and (2.34). Since $\varphi_{i,j} \leq \left(\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} b_{k,l}^2\right)^{\frac{1}{2}}$, and (1.47) implies that $\sigma_X^2 = \sigma_\varepsilon^2 \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} b_{i,j}^2$, the second upper bound in (2.32) follows immediately from the first one. In order to prove the lower bound, let $(i,j),(k,l)\in\mathscr{M}$, we have

$$\langle X(-i,-j), \eta(-k,-l)\rangle = \begin{cases} \sigma_\eta^2, & \text{if } (i,j) = (k,l), \\ 0, & \text{otherwise}. \end{cases}$$

Moreover, since $H'_{0,0} \subset I_{-i,-j}$ and $\eta(-i,-j) \perp I_{-i,-j}$, we have $\eta(-i,-j) \perp H'_{0,0}$. Then

$$\left\langle \sum_{(i,j)\in\mathscr{M}} a_{i,j}\left(X(-i,-j) - P_{H^X_{Q_1}} X(-i,-j), \eta(-k,-l)\right)\right\rangle = a_{k,l}\sigma_\eta^2$$

and using Schwarz inequality, we get

$$|a_{i,j}|\sigma_\eta \leq \left\| \sum_{(i,j)\in\mathscr{M}} \left(X(-i,-j) - P_{H^X_{Q_1}} X(-i,-j)\right)\right\|$$

which, using (2.33), gives the lower bound in (2.32). □

Remark 5. *For a weakly stationary random field $\{X(s,t);(s,t)\in\mathbb{Z}^2\}$, with the AR parameters, $\{a_{k,l}\}$ satisfying $\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_{i,j}^2 < \infty$, the lower bound in* (2.32) *hows that the degradation of the prediction due to the missing data increases with* $\max_{(i,j)\in M}$. *In the particular case, where $\mathscr{M} = \{(n_N,m_N), n_N>0, m_N>0\}$, inequalities* (2.32) *give*

$$\frac{\sigma_\varepsilon |a_{n_N,m_N}|}{\left(\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} a_{i,j}^2\right)^{\frac{1}{2}}} \leq \left\| P_{H^X_Q} X(0,0) - P_{H^X_{Q_1}} X(0,0)\right) \| \leq \sigma_\varepsilon |a_{n_N,m_N}|, \quad (2.35)$$

and it follows that

$$\left\| P_{H_Q^X} X(0,0) - P_{H_{Q_1}^X} X(0,0) \right\| = \frac{\sigma_\varepsilon |a_{n_N,m_N}|}{\left(\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_{i,j}^2 \right)^{\frac{1}{2}}}. \tag{2.36}$$

Therefore, the larger the values of n_N and m_N are, the better is the preciseness of the lower bound in (2.35).

In the following corollary, we characterize the random fields for which the missing observations do not affect the prediction of $X(0,0)$.

Corollary 2.5.2 ([10]). *Let $\{X(s,t);\ (s,t) \in \mathbb{Z}^2\}$ be a weakly stationary random field with the innovation process $\{\varepsilon(s,t)\}$, the MA parameters, $\{b_{s,t}\}$ and the AR parameters $\{a_{s,t}\}$. Then*

(a) $P_{H_{Q_1}^X} X(0,0) = P_{H_Q^X} X(0,0)$ *if and only if* $a_{i,j} = 0, \forall (i,j) \in \mathscr{M}$,

(b) $P_{H_{Q_1}^X} X(0,0) \to P_{H_Q^X} X(0,0)$ *in* L_2 *as* $\bar{n}_N, \bar{m}_N \to \infty$.

Proof: **(a)** follows from inequality (2.32). To show **(b)**, note that $\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_{i,j}^2$ implies that $a_{i,j} \to\ 0$, as $i,\ j \to \infty$. and the upper bound in (2.32) implies **(b)**. It is important to note that Theorem 2.5.1 gives rates of convergence $P_{H_{Q_1}^X} X(0,0)$ to $P_{H_Q^X}(0,0)$ as $\bar{n}_N, \bar{m}_N$ tend to infinity.

□

The following corollary provides a typical asymptotic behavior result.

Corollary 2.5.3 ([10]). *Let $\{X(s,t);\ (s,t) \in \mathbb{Z}^2\}$ be a weakly stationary random field with the innovation process $\{\varepsilon(s,t)\}$, the MA parameters, $\{b_{s,t}\}$ and the AR parameters, $\{a_{s,t}\}$. Then $\bar{n}_N, \bar{m}_N \to\ \infty$*

$$\left\| P_{H_Q^X}(0,0) - P'_{H_Q^X}(0,0) \right\| = O\left(\alpha^{\bar{n}_N + \bar{m}_N}\right) \text{ is } O(\alpha^{i+j}) \text{ as } i,j \to \text{ for some } \alpha \in (0,1).$$

Proof: The result follows from $\sum\sum_{(i,j)\in\mathscr{M}} |a_{i,j}| = O(\alpha^{\bar{n}_N + \bar{n}_M})$ and from the upper bound in (2.32). □

2.5.3 Illustrative examples

To evaluate the impact of missing values on the prediction of $X(0,0)$, we present here two examples carried out using the statistical software R 3.0.2. The steps involved in the computation of the estimates for the linear predictor, the missing data and their impact are summarized as follows. Data are generated as $n \times m$ rectangular grid from a spatial models of the form (2.37) and (2.38), where $\{\varepsilon(s,t)\}$ is white noise process with mean 0 and variance $\sigma^2 = 25$. Different sample sizes n and n are considered. After the suppression of $X(n-1,m-1)$, we calculate the value of $P_{H_{Q_1}^X}(n,m)$ by replacing 'this missing data' by its orthogonal projection onto $H_{n-1,m-1}$.The criterion used to quantify the impact of missing data on the predictor is the mean square prediction impact defined by

$$MSPI = \sqrt{\frac{1}{rep} \sum_{1}^{rep} \left(P_{H_Q^X}(0,0) - P_{H_{Q_1}^X}(0,0) \right)^{\frac{1}{2}}}$$

which corresponds to the quantity of $\left\| P_{H_Q^X}(0,0) - P_{H_{Q_1}^X}(0,0) \right\|$ defined previously.

Note that this criterion depends on the parameters of the model, so it has to be computed for several values. Moreover, for our models, the values of the parameters used in the simulation satisfy the stationarity conditions (i.e. [48, Lemma1]).

For obtaining a suitable estimate of mean square prediction impact of missing data, the simulation replicated 500 times for each sample size and pair of parameters (α,β).The results are provided in Tables 1 and 2

Example 2.5.4. *Consider the stationary first order multiplicative spatial autoregressive model (MSAR(1)) defined by*

$$X(s,t) = \alpha X(s-1,t) + \beta X(s,t-1) - \alpha\beta X(s-1,t-1) + \varepsilon(s,t), \tag{2.37}$$

where $\{(s,t);(s,t) \in \mathbb{Z}^2\}$*are independent random variables with* $E(\varepsilon(s,t)) = 0$*,* $Var(\varepsilon(s,t)) = \sigma^2, |a| < 1$ *and* $\beta < 1$.

By using the recursions given by (1.47) and the fact that $a_{10} = \alpha, a_{10} = \beta, a_{11} = -\alpha\beta$ and $a_{ij} = 0$ if $(i, j) \notin \{(1,0),(0,1),(1,1)\}$, it can be shown that the MA representation of the MSAR(1) model is

$$b_{k,l} = \begin{cases} \alpha^k\beta^l, & \text{if } k \geq 0, l \geq 0, \\ 0, & \text{if } k < 0, \text{ or } l < 0. \end{cases}$$

Note that the method used to obtain the best linear predictor of $X(0,0)$ is similar to prediction in time series AR(p) models, where the prediction is linear combination of $p-$nearest past observations.

$$P_{H_Q^X}(0,0) = \alpha X(-1,0) + \beta X(0,-1) - \alpha\beta X(-1,-1).$$

[48] obtained the prediction of $X(0,0)$ in the MSAR(1) model. Their prediction is a linear combination of data points in the nearest neighborhood to the prediction point:

$$\begin{aligned} \tilde{X}(0,0) &= \frac{1}{1+\alpha^2+\beta^2+\alpha^2\beta^2}\{(\alpha-\alpha\beta^2)(X(-1,0)+X(1,0)+(\beta+\alpha^2\beta)\times X(0,-1) \\ &\quad +X(0,1)) - \alpha\beta(X(-1,-1)+X(1,1)) - \alpha\beta\times(X(-1,1)+X(1,-1))\} \end{aligned}$$

and its mean squared error is

$$\sigma_\eta^2 = \|\eta(k,l)\|^2 = \left\|\tilde{X}(0,0) - X(0,0)\right\| = \frac{1}{1+\alpha^2+\beta^2+\alpha^2\beta^2} = \frac{1}{(1+\alpha^2)(1+\beta^2)}.$$

According to Remark 5, it is clear that if the indices of missing data $(n_i,m_i) \notin \{(-1,0),(0,-1),(-1,-1)\}$, we , we have $P_{H_Q^X}(0,0) = P_{H_{Q_1}^X}(0,0)$. Also, if we have one missing data with the indices in $\{(1,0),(0,1),(1,1)\}$, for example if $(n_i,m_i) = (1,0) \Rightarrow a_{1,0} = a$, and (2.35) can be simplified and reduces to

$$\frac{\sigma^2\alpha^2}{(1+\alpha^2)(1+\beta^2)} \leq \left\|P_{H_Q^X}(0,0) - P_{H_{Q_1}^X}(0,0)\right\|^2 \leq \sigma^2\alpha^2,$$

and it follows from (2.36) that $\left\|P_{H_Q^X}(0,0) - P_{H_{Q_1}^X}(0,0)\right\| = \sigma|\alpha|$. Table 2.1 gives the simulated values of MSPI when the data are generated by the model (2.37). We consider the parameters for which the stationarity conditions are verified. Using these parameters values, we generated the data for several grid sizes. Note that the MSPI is obtained after filling the missing value of $X(-1,-1)$ by its orthogonal projection onto $H_{-1,-1}$for each iteration.

From Table 2.1, it can be seen that the MSPI is smaller for values of α and β for which the product $\alpha.\beta$ is small. This confirms the validity of the theory because the coefficient of $X(-1,-1)$ is $\alpha.\beta$.

	(m,n)				
(α,β)	(25,25)	(25,50)	(50,50)	(100,100)	(150,150)
(0.2,0.2)	0.253	0.246	0.238	0.202	0.199
(0.2,0.5)	0.538	0.540	0.528	0.512	0.510
(0.2,0.9)	1.115	1.027	1.001	0.904	0.901
(0.5,0.2)	0536	0.538	0.531	0.514	0.509
(0.5,0.5)	1.892	1.882	1.863	1.139	1.136
(0.5,0.9)	2.892	2.843	2.841	2.228	2.225
(0.7,0.2)	0.812	0.801	0.795	0.722	0.723
(0.7,0.5)	2.114	2.038	1.987	1.829	1.825
(0.9,0.2)	1.113	1.031	0.996	0.910	0.908
(0.9,0.5)	2.889	2.850	2.837	2.232	2.228
(0.9,0.9)	3.586	3.218	3.106	2.943	2.939

Table 2.1: Mean Square Prediction Impact when the value of $X(-1,-1)$ is missing

In addition, it seems that for all values of (α,β), the MSPI decreases as the grid sizes increase. In addition, the simulations highlight a certain symmetry of the results. Indeed, the MSPI for the values $\alpha = a$ and $\beta = b$ are very close to $\alpha = b$ to $\beta = a$, that is, by permuting the values of α and β, the MSPI remains almost identical.

Example 2.5.5. *Assume that $\{X(s,t)\}$ is a spatial autoregressive model with parameters $\alpha_1,\alpha_2,\alpha_3$, α_4 and α_5 instead of $a_{1,0},a_{0,1},a_{1,1},a_{2,0}$ and $a_{2,1}$:*

$$X(s,t) = \alpha_1 X(s-1,t)+\alpha_2 X(s,t-1)+\alpha_3 X(s-1,t-1)+\alpha_4 X(s-2,t-1)+\alpha_5 X(s-2,t-1)+\varepsilon(s,t). \tag{2.38}$$

The extension of the prediction method developed in [49] was applied to this model in [50] and they obtained

$$P_{H_Q^X}(0,0) = \alpha_1 X(-1,0)+\alpha_2(0,-1)+\alpha_3 X(-1,-1)+\alpha_4 X(-2,0)+\alpha_5 X(-2,-1),$$

$$\begin{aligned}\tilde{X}(0,0) = &\frac{1}{1+\alpha_1^2+\alpha_2^2+\alpha_3^2+\alpha_4^2+\alpha_5^2}\times((\alpha_1-\alpha_1\alpha_4-\alpha_2\alpha_3-\alpha_3\alpha_5)(X(-1,0)+X(1,0))\\ &+((\alpha_2-\alpha_1\alpha_3-\alpha_4\alpha_5)(X(0,-1)+X(0,1))-(\alpha_1\alpha_2+\alpha_3\alpha_4)(X(-1,1)+X(1,-1)\\ &+(\alpha_3-\alpha_1\alpha_5)((X(-1,-1)+X(1,1)))+(\alpha_4-\alpha_2\alpha_5)(X((-2,0)+X(2,0))\\ &-\alpha_2\alpha_4(X(-2,1)+X(2,-1)),\end{aligned}$$

and its variance is

$$\sigma_\eta^2 = \|\eta(0,0)\|^2 = \frac{1}{1+\alpha_1^2+\alpha_2^2+\alpha_3^2+\alpha_4^2+\alpha_5^2}.$$

Of course, when the indices of missing data $(n_i,m_i)\notin\{(-1,0),(0,-1),(-1,-1),(-2,0),(-2,-1)\}$ we have $P_{H_Q^X}(0,0) = P_{H_{Q_1}^X}(0,0)\|$. From (2.35) *if we have one missing data with the indices in $\{(-1,0),$ $(0,-1),(-1,-1),(-2,0),(-2,-1)\}$, for example $(n_i,m_i) = (-1,-1)$, we have*

$$\frac{\sigma^2\alpha_3^2}{1+\alpha_1^2+\alpha_2^2+\alpha_3^2+\alpha_4^2+\alpha_5^2} \le \left\|P_{H_Q^X}(0,0)-P_{H_{Q_1}^X}(0,0)\right\|^2 \le \sigma^2\alpha_3^2.$$

	(m,n)					
$\alpha=(\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5)$	(25,25)	(25,50)	(50,50)	(100,100)	(150,150)	(200,200)
(0.1,0.2,0.3,0.4,0.5)	1.410	1.125	1.119	0.998	0.860	0.858
(0.1,0.3,0.4,0.5,0.6)	1.687	1.589	1.420	1.267	1.107	1.103
(0.1,0.4,0.5,0.6,0.7)	1.759	1.602	1.499	1.382	1.202	1.117
(0.1,0.5,0.6,0.7,0.8)	2.101	1.892	1.643	1.497	1.421	1.387
(0.1,0.6,0.7,0.8,0.9)	2.101	1.892	1.643	1.497	1.421	1.387
(0.2,0.2,0.3,0.4,0.5)	1.724	1.663	1.597	1.118	1.105	1.098
(0.2,0.3,0.4,0.5,0.6)	1.943	1.813	1.739	1.576	1.548	1.442
(0.2,0.4,0.5,0.6,0.7)	2.118	2.107	2.065	2.003	1.847	1.720
(0.2,0.5,0.6,0.7,0.8)	2.807	2.764	2.669	2.501	2.357	2.582
(0.2,0.6,0.7,0.8,0.9)	3.179	3.109	2.964	2.869	2.776	2.693
(0.5,0.2,0.3,0.4,0.5)	2.013	2.007	1.986	1.876	1.789	1.781
(0.5,0.3,0.4,0.5,0.6)	2.716	2.683	2.547	2.443	2.222	2.113
(0.5,0.4,0.5,0.6,0.7)	2.946	2.938	2.920	2.863	2.768	2.616
(0.5,0.5,0.6,0.7,0.8)	3.210	3.117	3.113	3.002	2.981	2.867
(0.5,0.6,0.7,0.8,0.9)	3.332	3.326	3.119	3.107	2.987	2.991
(0.7,0.2,0.3,0.4,0.5)	2.121	2.111	2.103	2.058	1.977	1.807
(0.7,0.3,0.4,0.5,0.6)	2.458	2.441	2.396	2.356	2.241	2.123
(0.7,0.4,0.5,0.6,0.7)	3.114	3.012	2.888	2.769	2.563	2.453
(0.7,0.5,0.6,0.7,0.8)	3.389	3.427	3.415	3.333	3.119	3.008
(0.7,0.6,0.7,0.8,0.9)	3.651	3.558	3.447	3.339	3.247	3.219
(0.9,0.2,0.3,0.4,0.5)	3.476	3.339	3.332	3.221	3.117	3.101
(0.9,0.3,0.4,0.5,0.6)	4.101	3.876	3.776	3.655	3.549	3.247
(0.9,0.4,0.5,0.6,0.7)	4.997	4.919	4.837	4.773	4.486	4.369
(0.9,0.5,0.6,0.7,0.8)	5.553	5.106	4.992	4.917	4.873	4.796
(0.9,0.6,0.7,0.8,0.9)	5.646	5.428	5.001	4.891	4.763	4.668

Table 2.2: Mean square prediction impact when the values of $X(-1,0)$ and $X(-1,-1)$ are missing.

Using (2.36)*, we deduce*

$$\left\|P_{H_Q^X}(0,0)-P_{H_{Q_1}^X}(0,0)\right\|=\frac{\sigma|\alpha_3^2|}{1+\alpha_1^2+\alpha_2^2}.$$

In what follows, we apply the same idea of simulation used previously to the model (2.38). We use the MSPI criterion to quantify the impact of the missing values $X(-1,0)$ and $X(-1,-1)$. In order to highlight the role of the autoregressive parameter, we vary all the coefficients of the vector $\alpha=(\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5)$. In view of Table 2.2, the same conclusions apply to the model (2.38). Indeed, the values of MSPI increase as α_1 and α_3 increase. Also, it appears from Table 2.2, the larger the sum $\sum_{i=1}^{5}\alpha_i^2$, the greater the impact of the missing data on the prediction, which highlights the role of the coefficients of the autoregressive representation. As for the symmetry of results, it is also confirmed in this model. Indeed, by permuting the values of the coefficients of α_i, the MSPI's values remain very close.

2.6 Conclusion

Section 2.2 focus on prediction of random fields has been on moving average (MA) representation and the multi-step ahead predictors and their prediction error variances. The purpose of the second section was to use these relations to solve several nonstandard prediction problems when the third-quadrant Q is used as the past of a stationary random field and some observations are added to the past. More precisely, we are interested in predicting $X(0,0)$ when a finite number of observations are added to the past Q. It turns out that solutions of such nonstandard prediction problems are not straightforward extensions of their stationary 1-D process counterparts. In fact, it was found in in [7] that one can express the prediction error variances in terms of the MA parameters for the random field case. However, attempts to express them in terms of the autoregressive (AR) parameters depend on a new projection operator, which seems intrinsic to the random field situation. This task is riddled with technical challenges and was left as an open problem in [7] .

In Section 2.3, two solutions have been presented for predicting a stationary Gaussian random field based on a quarter-plane with finitely missing observations . One expresses the best predictor in terms of the moving average coefficients of the random field, and the other expresses the prediction error variance in terms of its autoregressive coefficients. In both cases, the solutions depend on spectral conditions associated with the notion of a strongly outer function on the torus. Needed properties of strongly outer functions were obtained.

In Section 2.4, an explicit formula for the prediction error variance of a future value of a weakly stationary random field is established, when the infinite past, is altered by some missing observations.It should be emphasized that this representation can be seen as an alternative solution to the problem posed by [7].

In Section 2.5, we treat the prediction problems where a number of observations are missing to the quarter-plane past of a stationary random field. Our aim is to quantify the influence of missing values on the prediction by giving the simple bounds for the prediction error variance. These bounds allow to characterize the random fields for which the missing observations do not affect the prediction.

CHAPTER 3

Prediction for Stationary Random Fields with Nonsymmetrical Half-Plane Past

Contents

3.1 Introduction

Our second contribution (see [20]) is detailed in this chapter, we investigate the problem of linear prediction of stationary random fields with nonsymmetrical half-plane past. We aim at finding an explicit formula of the mean square convergent autoregressive series representation for all (h_1, h_2)-step ahead linear predictors, $(h_1, h_2) \geq (0, 0)$. In order to calculate explicitly the prediction coefficients of our new expression, we provide recursive relations between the infinite multi-step prediction coefficients which are a generalization of the relations between the prediction coefficients of infinite multi-step ahead linear predictors of stationary time series established by [51].

The outline of this chapter is : In section 3.2, we give necessary and sufficient condition for the existence of the predictor proposed herein, when the considered past is the strict past of X at the point (s, t). In section 3.3, we provide recursive relations between the prediction coefficients of multi-step ahead linear predictor of stationary random fields. Some specific examples to validate the applicability of our relations are presented.

Under the assumption of stationarity of the random field, [7] developed a framework for computing the predictors of random fields by extending some works of stationary time series cited in their bibliography such as [52] and [42]. By following the lines in [7], we establish a formal basis for the analysis of the problem of finding the autoregressive representation for all (h_1,h_2)-step ahead linear predictors, $(h_1,h_2) \geq (0,0)$ of $X(s+h_1,t+h_2)$ based on the strict past of X at the point (s,t). Unlike the 1-D case, there is no natural order definition in the 2-D domain. In what follows, The problem of prediction is developed in the case of a past defined by an nonsymmetrical half-plane resulting from the total order definition (1.25a), $S = \{(k,l) \in \mathbb{Z}^2, (k,l) < (s,t)\}$. We define H^S as the closed linear subspace spanned by $\{X(k,l), (k,l) \in S\}$ in the Hilbert space H, which represents the strict past of X at the point (s,t). We denote by P_{H^S} the orthogonal projection operator onto H^S. Then, from the above we can immediately say that the corresponding innovation process $\{\varepsilon(s,t), (s,t) \in \mathbb{Z}^2\}$ defined by

$$\varepsilon(s,t) = X(s,t) - P_{H^S}X(s,t)$$

is a white noise. If $E(X(s,t) - P_{H^S}X(s,t))^2 > 0$, the random field is called non deterministic. In this case the variance of the innovations field $\{\varepsilon(s,t)\}$ is given by Equation (1.36). The generalisation of the Wold decomposition theorem to the 2-D case was established by [14]. This generalisation states that every non deterministic stationary process $X(s,t)$ may be decomposed into two stationary and orthogonal components: the purely-nondeterministic process (that produces the innovations) and the deterministic process see Theorem 1.4.2, namely

$$X(s,t) = \sum_{k=0}^{+\infty}\sum_{l=0}^{+\infty} b_{k,l}\varepsilon(s-k,t-l) + V(s,t),$$

where the sequence $\{b_{k,l}, (k,l) \in \mathbb{Z}^2\}$ satisfies $\sum_{k=0}^{+\infty}\sum_{l=0}^{+\infty} |b_{k,l}|^2 < \infty$, with $b_{0,0} = 1$. We note that $\{\varepsilon(s,t), (s,t) \in \mathbb{Z}^2\}$ is the innovation process (one step predictor error) of $\{X(s,t); (s,t) \in \mathbb{Z}^2\}$ with mean zero and variance σ^2, which is uncorrelated with $X(u,v)$, $(u,v) < (s,t)$.
For a purely non deterministic stationary (PND) random field $X(s,t)$ (the case when $V(s,t) = 0$, i.e $X(s,t) \in \overline{sp}\{\varepsilon(s,t), (s,t) \in S\}$). $X(s,t)$ have a mean square convergent infinite moving average $MA(\infty)$ representation [7]

$$X(s,t) = \sum_{k=0}^{+\infty}\sum_{l=0}^{+\infty} b_{kl}\varepsilon(s-k,t-l), \tag{3.1}$$

the sequence $\{b_{k,l}, (k,l) \in \mathbb{Z}^2\}$ is called the MA(∞) parameters.
The author in [53] extended the solution of the 1-D spectral factorization problem to the 2-D case, and showed that the purely indeterministic field has an innovations driven autoregressive (AR) representation, in addition to the existing MA representation (3.1), if the MA representation is invertible, then it is inverted to give

$$\varepsilon(s,t) = \sum_{k=0}^{+\infty}\sum_{l=0}^{+\infty} a_{k,l}X(s-k,t-l). \tag{3.2}$$

In the remainder of this chapter, convergence of representation (3.2) will always be in mean square.

3.2 Autoregressive representation of the Multi-step ahead linear predictor

The multi-step ahead prediction problem of stationary random fields has been studied by [7] when the third quadrant is used as the past. We extends their pioneer work to random fields with nonsymmetrical half-plane past. Let $\{X(s,t); (s,t) \in \mathbb{Z}^2\}$ is a PND stationary random field. The procedure for solving the $(h_1, h_2)-$step ahead linear prediction problem with respect to the total order and nonsymmerical half-plane (NSHP) support defined by (1.25a) involves the construction of predictor of future values as a linear combination of $\{X(k,l), (k,l) \in S\}$ which are close to $X(s+h_1, t+h_2)$, $(h_1, h_2) \geq (0,0)$ in the sense of mean squared error. The collection of all finite linear combinations of elements in the space and its closure are also included in the space. At first we fix our attention on the problem of finding convergent representation for the one-step ahead linear predictor $P_{H^S}X(s,t)$, i.e the minimum norm linear causal and continuous support predictor of $X(s,t)$. We show that when (3.2) converges, such a representation exists.

Theorem 3.2.1 ([20]). *Let $\{X(T); T \in \mathbb{Z}^2\}$ be a PND stationary random field. The one step ahead linear predictor $P_{H^S}X(s,t)$ of $X(s,t)$ possesses a convergent serie representation given by*

$$P_{H^S}X(s,t) = -\sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{+\infty}\sum_{l=0}^{+\infty} a_{k,l}X(s-k, t-l), \tag{3.3}$$

if and only if $\varepsilon(s,t)$ has the series representation (3.2).

Proof: We have (3.1) implies that

$$E(\varepsilon(s,t)X(s,t)) = E(\varepsilon(s,t))^2.$$

From (3.2) we deduce that

$$E(\varepsilon(s,t))^2 = a_{00}E(\varepsilon(s,t)X(s,t)),$$

and necessarily $a_{00} = 1$. Thus, (3.2) may be rewritten as

$$X(s,t) = \varepsilon(s,t) - \sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{+\infty}\sum_{l=0}^{+\infty} a_{k,l}X(s-k, t-l).$$

Since $\varepsilon(s,t)$ is uncorrelated with $X(u,v)$, $(u,v) < (s,t)$, we deduce that the one-step predictor of $X(s,t)$ is given by

$$P_{H^S}X(s,t) = -\sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{+\infty}\sum_{l=0}^{+\infty} a_{k,l}X(s-k, t-l). \tag{3.4}$$

The existence of the convergent representation (3.3) for the one step predictor is assured by the convergence of (3.2). Conversely, if the one-step predictor $P_{H^S_{s,t}}X(s,t)$ has the mean square representation (3.3), then the one-step prediction error satisfies

$$\varepsilon(s,t) = X(s,t) - P_{H^S}X(s,t) = \sum_{k=0}^{+\infty}\sum_{l=0}^{+\infty} a_{k,l}X(s-k, t-l),$$

with $a_{0,0} = 1$ and the sum convergent in mean square. Thus, we have shown that a necessary and sufficient condition for the existence of (3.3) as a mean square limit is the existence and the convergence of (3.2). $\square$

Now, we are interested in predicting future values other than $X(s,t)$ which is greatly important in the theory and applications of stationary random fields. The next lemma is useful for computing the predictor.

Lemma 3.2.2 ([20]). *Let $\{X(T); T \in \mathbb{Z}^2\}$ be a PND stationary random field, the MA and the AR parameters are $\{b_{k,l}, (k,l) \in \mathbb{Z}^2\}$ and $\{a_{k,l}, (k,l) \in \mathbb{Z}^2\}$, respectively, then the following equation is satisfied for all $(k,l) \geq (0,1)$*

$$\sum_{i=0}^{k}\sum_{j=0}^{l} a_{ij} b_{k-i,l-j} = 0. \tag{3.5}$$

Proof: By substituting (3.2) we obtain for all $(k,l) \geq (0,1)$

$$\begin{aligned} 0 = E(\varepsilon(s-k,t-l)\varepsilon(s,t)) &= E(\varepsilon(s-k,t-l)\sum_{i=0}^{+\infty}\sum_{j=0}^{+\infty} a_{ij} X(s-i,t-j)) \\ &= E(\varepsilon(s-k,t-l)\sum_{i=0}^{k}\sum_{j=0}^{l} a_{ij} X(s-i,t-j)) \\ &= \sum_{i=0}^{k}\sum_{j=0}^{l} a_{ij} E(\varepsilon(s-k,t-l)X(s-i,t-j)) \\ &= \sum_{i=0}^{k}\sum_{j=0}^{l} a_{ij} b_{k-i,l-j} E(\varepsilon(s-k,t-l))^2. \end{aligned}$$

Accordingly

$$\sum_{i=0}^{k}\sum_{j=0}^{l} a_{ij} b_{k-i,l-j} = 0, \quad (k,l) \geq (0,1).$$

□

In the following theorem, we show that when (3.2) converges, we can find convergent representation for all (h_1,h_2)-step ahead linear predictor, $(h_1,h_2) \geq (0,0)$, $(h_1,h_2) \neq (0,0)$.

Theorem 3.2.3. *Let $\{X(T); T \in \mathbb{Z}^2\}$ be a PND stationary random field. Then, for any $(s+h_1,t+h_2) \in S^c$, where S^c is the complement of S. The $(h_1,h_2)-$step ahead predictor of $X(s+h_1,t+h_2)$, $(h_1,h_2) \geq (0,0)$, $(h_1,h_2) \neq (0,0)$ based on the past S possesses a convergent series representation given by*

$$P_{H^S} X(s+h_1,t+h_2) = \sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{+\infty}\sum_{l=0}^{+\infty} a_{k,l}^{(h_1,h_2)} X(s-k,t-l), \tag{3.6}$$

where

$$a_{k,l}^{(h_1,h_2)} = -\sum_{p=0}^{h_1}\sum_{q=0}^{h_2} b_{h_1-p,h_2-q} a_{p+k,q+l}, \quad (k,l) \geq (0,1) \tag{3.7}$$

if and only if $\varepsilon(s,t)$ has the convergent series representation (3.2).

Proof: The necessity is obvious, we show sufficiency. Using the convergent representation (3.2)

$$\begin{aligned}
\sum_{p=0}^{h_1} \sum_{q=0}^{h_2} b_{h_1-p,h_2-q}\varepsilon(s+p,t+q) &= \sum_{p=0}^{h_1}\sum_{q=0}^{h_2} b_{h_1-p,h_2-q}\left(\sum_{i=0}^{+\infty}\sum_{j=0}^{+\infty} a_{ij}X(s+p-i,t+q-j)\right)\\
&= \sum_{p=0}^{h_1}\sum_{q=0}^{h_2} b_{h_1-p,h_2-q}\Bigg(\sum_{i=0}^{p-1}\sum_{j=0}^{q-1} a_{ij}X(s+p-i,t+q-j)+\sum_{i=p}^{+\infty}\sum_{j=q}^{+\infty} a_{ij}X(s+p-i,t+q-j)\\
&\quad +\sum_{i=0}^{p-1}\sum_{j=q}^{+\infty} a_{ij}X(s+p-i,t+q-j)+\sum_{i=p}^{+\infty}\sum_{j=0}^{q-1} a_{ij}X(s+p-i,t+q-j)\Bigg)\\
&= \sum_{k=1}^{p}\sum_{l=1}^{q} X(s+k,t+l)\sum_{p=0}^{h_1}\sum_{q=0}^{h_2} b_{h_1-p,h_2-q}a_{p-k,q-l}+\sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{+\infty}\sum_{l=0}^{+\infty} X(s-k,t-l)\sum_{p=0}^{h_1}\sum_{q=0}^{h_2} b_{h_1-p,h_2-q}a_{p+k,q+l}\\
&= \sum_{k=1}^{h_1}\sum_{l=1}^{h_2} X(s+k,t+l)\sum_{p=k}^{h_1}\sum_{q=l}^{h_2} b_{h_1-p,h_2-q}a_{p-k,q-l}+\sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{+\infty}\sum_{l=0}^{+\infty} X(s-k,t-l)\sum_{p=0}^{h_1}\sum_{q=0}^{h_2} b_{h_1-p,h_2-q}a_{p+k,q+l}.
\end{aligned}$$

Now, from Lemma 3.2.2 we have

$$\sum_{p=k}^{h_1}\sum_{q=l}^{h_2} b_{h_1-p,h_2-q}a_{p-k,q-l} = \sum_{i'=0}^{h_1-k}\sum_{j'=0}^{h_2-l} b_{h_1-k-i',h_2-l-j'}a_{i',j'} = \begin{cases} 1 & \text{for } (h_1,h_2)=(k,l),\\ 0 & \text{for } (h_1,h_2)>(k,l).\end{cases}$$

From this,

$$\sum_{p=0}^{h_1}\sum_{q=0}^{h_2} b_{h_1-p,h_2-q}\varepsilon(s+p,t+q) = X(s+h_1,t+h_2)-\sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{+\infty}\sum_{l=0}^{+\infty} a_{k,l}^{(h_1,h_2)}X(s-k,t-l),$$

where

$$a_{k,l}^{(h_1,h_2)} = -\sum_{p=0}^{h_1}\sum_{q=0}^{h_2} b_{h_1-p,h_2-q}a_{p+k,q+l}.$$

Since $\{\varepsilon(s,t),(s,t)\in\mathbb{Z}^2\}$ is uncorrelated with $\{X(u,v),(u,v)<(s,t)\}$, it follows that the $(h_1,h_2)-$step ahead predictor of $X(s+h_1,t+h_2)$ possesses a series representation given by

$$P_{H^s}X(s+h_1,t+h_2) = \sum_{\substack{k=0\\(k,l)\neq(0,0)}}^{+\infty}\sum_{l=0}^{+\infty} a_{k,l}^{(h_1,h_2)}X(s-k,t-l).$$

□

The convergence of (3.6) stems from its construction as a finite linear combination of tails sums of convergent series.

3.3 Recursive relations between multi-step prediction coefficients

In this section, firstly we give infinite recursive relations between the prediction coefficients for multi-step ahead linear predictor of stationary random fields. We focus on how this relations can be used for computing efficiently the (h_1,h_2)-step prediction coefficients. Secondly, we give some illustrative examples to verify the validity of our relations.

3.3.1 Infinite step recursive relations

According to the previous theorem, the best $(h_1, h_2)-$step ahead predictor based on the infinite past S, chosen so that the prediction error is a white noise according to which the process $X(s,t)$ must be able to be represented, is

$$P_{H^S}X(s+h_1, t+h_2) = \sum_{\substack{k=0 \\ (k,l)\neq(0,0)}}^{+\infty}\sum_{l=0}^{+\infty} a_{k,l}^{(h_1,h_2)}X(s-k,t-l), \quad (h_1, h_2) \geq (0, 0).$$

The following relations are analogous to the relations (2.2) in [51] in the one dimensional case. From (3.7), we have

$$a_{k,l}^{(h_1,h_2)} = \sum_{p=0}^{h_1}\sum_{q=0}^{h_2} b_{p,q}a_{k+h_1-p,l+h_2-q}, \quad (k, l) \geq (0, 1), \tag{3.8}$$

and using Lemma 3.2.2, we obtain

$$a_{k,l}^{(h_1,h_2)} = -\sum_{p=h_1}^{h_1+k}\sum_{q=h_2+1}^{h_2+l} b_{p,q}a_{k+h_1-p,l+h_2-q}, \quad (k, l) \geq (0, 1). \tag{3.9}$$

Step recursive relations when the prediction is based on infinite past can be immediately deduced from relations (3.7) and (3.8) as follows,

$$a_{k,l}^{(h_1,h_2)} = a_{k,l+1}^{(h_1,h_2-1)} + \sum_{p=0}^{h_1} b_{p,h_2}a_{k+h_1-p,l}, \quad (k, l) \geq (0, 1). \tag{3.10}$$

We shall notice that $a_{k,l}^{(0,0)} = -a_{k,l}$ for all $(k, l) \geq (0, 1)$. The parameters $a_{k,l}^{(h_1,h_2)}, (h_1, h_2) > (0, 0)$, can be calculated recursively using (3.10). These relations are generalization of the recursive relations (2.4) between the prediction coefficients of infinite multi-step ahead linear predictor of stationary time series in [51].

3.3.2 Some Specific examples

In order to evaluate the operation of our innovative results, we present in this section two theoretical examples in detail.

Example 3.3.1. *The stationary first order multiplicative spatial autoregressive model (MSAR(1)) defined by*

$$X(s, t) = \alpha X(s-1, t) + \beta X(s, t-1) - \alpha\beta X(s-1, t-1) + \varepsilon(s, t), (s, t) \in \mathbb{Z}^2, \tag{3.11}$$

where $\varepsilon(s, t)$ are independent identically distributed random variables, this model is stationary if $|\alpha| < 1$ and $|\beta| < 1$ (see for instance [54]. It can be shown that the MA representation (MSAR(1)) model is

$$b_{i,j} = \begin{cases} \alpha^i\beta^j & , \text{ if } i \geq 0, j \geq 0 \\ 0 & , \text{ if } i < 0 \text{ or } j < 0. \end{cases}$$

Let $a_{1,0} = \alpha$, $a_{0,1} = \beta$ and $a_{1,1} = -\alpha\beta$, the coefficients $a_{k,l}^{(h_1,h_2)}$, $(h_1, h_2) \geq (0, 0)$ and $(k, l) \in \xi = \{(1, 0), (0, 1), (1, 1)\}$ are calculated by using the recursive relation (3.10) *and are given by*

1. *For $(h_1, h_2) = (0, 0)$, $a_{k,l}^{(0,0)} = -a_{k,l}$, $(k,l) \in \xi$,*

2. *For* $h_1 = 0,\ h_2 \neq 0,$

$$a_{k,l}^{(0,h_2)} = \begin{cases} a_{k,l+1}^{(0,h_2-1)} + \alpha^k \beta^{l+h_2} & ,\ (k,l) \in \{(0,1),(1,0)\} \\ a_{k,l+1}^{(0,h_2-1)} - \alpha^k \beta^{l+h_2} & ,\ (k,l) = (1,1). \end{cases}$$

3. *For* $h_1 \neq 0,\ h_2 = 0,\ a_{k,l}^{(h_1,0)} = k\alpha^{h_1+k}(-\beta)^l,\ (k,l) \in \xi,$

4. *For* $h_1 \geq 1, h_2 \geq 1,$

$$a_{k,l}^{(h_1,h_2)} = \begin{cases} a_{k,l+1}^{h_1,h_2-1} + \alpha^{h_1+k}\beta^{h_2+l} & ,\ (k,l) = (1,0), \\ a_{k,l+1}^{h_1,h_2-1} & ,\ (k,l) = (0,1), \\ a_{k,l+1}^{h_1,h_2-1} - \alpha^{h_1+k}\beta^{h_2+l} & ,\ (k,l) = (1,1). \end{cases}$$

Example 3.3.2. *One of the most studied two-dimensional autoregression models is given by*

$$X(s,t) = a_{1,0}X(s-1,t) + a_{0,1}X(s,t-1) + \varepsilon(s,t), \quad (s,t) \in \mathbb{Z}^2. \tag{3.12}$$

This model is stationary , if $|a_{1,0}| + |a_{0,1}| < 1$ *(see [48]).*

a) *Let* $a_{1,0} = \alpha,\ a_{0,1} = \beta$*, the MA representation is*

$$b_{i,j} = \begin{cases} \alpha^i \beta^j & \text{if } i = 0 \text{ or } j = 0, \\ (i+j)\alpha^i\beta^j & \text{if } i \neq 0, j \neq 0. \end{cases}$$

The coefficients $a_{k,l}^{(h_1,h_2)}$, $(h_1,h_2) \geq (0,0)$ *and* $(k,l) \in \rho = \{(1,0),(0,1)\}$ *are calculated by using the recursive relation* (3.10) *and are given by*

1. *For* $(h_1,h_2) = (0,0),\ a_{k,l}^{(0,0)} = a_{k,l},\ (k,l) \in \rho,$
2. *For* $h_1 = 0,\ h_2 \neq 0,\ a_{k,l}^{(0,h_2)} = a_{k,l+1}^{(0,h_2-1)} + \alpha^k\beta^{h_2+l},\ (k,l) \in \rho,$
3. *For* $h_1 \neq 0,\ h_2 = 0,\ a_{k,l}^{(h_1,0)} = \alpha^{k+h_1}\beta^l,\ (k,l) \in \rho,$
4. *For* $h_1 \geq 1, h_2 \geq 1,\ a_{k,l}^{(h_1,h_2)} = a_{k,l+1}^{(h_1,h_2-1)} + (h_1+h_2)\alpha^{h_1+k}\beta^{h_2+l},\ \ (k,l) \in \rho.$

b) *The case of* $a_{1,0} = a_{0,1} = \alpha$*, was considered by [55]. The coefficients* $a_{k,l}^{(h_1,h_2)}$, $(h_1,h_2) \geq (0,0)$ *and* $(k,l) \in \rho = \{(1,0),(0,1)\}$ *are calculated by using the recursive relation* (3.10) *and are given by*

1. *For* $(h_1,h_2) = (0,0),\ a_{k,l}^{(0,0)} = a_{k,l},\ (k,l) \in \rho,$
2. *For* $h_1 = 0,\ h_2 \neq 0,\ a_{k,l}^{(0,h_2)} = a_{k,l+1}^{(0,h_2-1)} + \alpha^{k+h_2+l},\ (k,l) \in \rho,$
3. *For* $h_1 \neq 0,\ h_2 = 0,\ a_{k,l}^{(h_1,0)} = \alpha^{k+h_1+l},\ (k,l) \in \rho,$
4. *For* $h_1 \geq 1, h_2 \geq 1,\ a_{k,l}^{(h_1,h_2)} = a_{k,l+1}^{(h_1,h_2-1)} + (h_1+h_2)\alpha^{h_1+h_2+k+l}, (k,l) \in \rho.$

The multi-step prediction coefficients are explicitly calculated by using the recursive relations given (3.10). *The relations are tractable and present great practical importance, since they efficiently facilitates the computation of predictors.*

Conclusion and perspectives

The purpose of this thesis was to study the prediction problem of stationary random fields. It turns out that the solutions are not straightforward extensions of their 1-D counterparts. We gave solutions to several nonstandard prediction problem. Firstly, we treat the prediction problems where a number of observations are missing to the quarter-plane past of a stationary random field. The change in the impact due to the missing of the observations provides a way to measure the worth of observations in prediction. This measure of worth depends on the AR parameters of the random fields and in any case does not exceed the upper value of the bound established by Theorem 2.5.1. Secondly, we establish a formal basis for the analysis of the problem of finding the autoregressive representation for all (h_1, h_2)-step ahead linear predictors of stationary random fields with nonsymmetrical half-plane past (NSHP). Firstly, we gave an explicit autoregressive series representation for the best multi-step ahead linear predictor of stationary random fields with nonsymmetrical half-plane past (NSHP). Secondly, Necessary and sufficient condition for the mean square convergence of these series is given. Moreover, step recursive relations between the prediction coefficients for the infinite past predictor are provided, these relations are used to calculate explicitly the multi-step prediction coefficients.

Research prospects

Many questions still open in this area. As future perspectives on the results presented in this manuscript.

1. In practical problems, we have only a finite number of observations from which to construct the predictor. Therefore, we aim at further to provide step recursive relations between the prediction coefficients when the considered past is finite. These relations would to be a generalization of those presented in Chapter 3.

2. Another practical avenue of research is to develop various algorithms allowing the calculation of the coefficients of the predictors with incomplete past.

3. In a future work, we would apply the theoretical results presented here in practice. The relations (3.10) are tractable and present great practical importance, since they efficiently facilitates the

computation of predictors.

4. We note that these nonstandard prediction problems are closely related to the optimal network site selection problem in environmental studies ([7], [56], [57], [58]). Since observations at different locations have different effect on prediction, a measure of worth for the observations can be defined in terms of the relative change in the prediction error variance, due to inclusion (exclusion) of a set of observations. In future works, we discuss the details of network selection applications and how to apply the theoretical results presented here in practice.

5. Two dimensionally indexed Random Coefficients AutoRegressive models,2D-RCAR for short and the corresponding statistical inference are important tools for the analysis of spatial lattice data. The study of such models is motivated by their broad range of applications in many areas and their second order properties that are similar to those of 2D(G)ARCH which plays an important role in spatial econometric.

Bibliography

Bibliography

[1] N. Cressie. *Statistics for spatial data*. John Wiley & Sons, 2015.

[2] D. J. Gorsich and M. G. Genton. Variogram model selection via nonparametric derivative estimation. *Mathematical geology*, 32(3):249–270, 2000.

[3] P. Hall and P. Patil. Properties of nonparametric estimators of autocovariance for stationary random fields. *Probability Theory and Related Fields*, 99(3):399–424, 1994.

[4] H. K. Im, L. M. Stein, and Z. Zhu. Semiparametric estimation of spectral density with irregular observations. *Journal of the American Statistical Association*, 102(478):726–735, 2007.

[5] A. Kizilkaya. On the parameter estimation of 2-d moving average random fields. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 54(11):989–993, 2007.

[6] A. Kizilkaya. Computation of the exact cramer–rao lower bound for the parameters of a non-symmetric half-plane 2-D ARMA model. *Digital Signal Processing*, 18(5):835–843, 2008.

[7] P. Kohli and M. Pourahmadi. Some prediction problems for stationary random fields with quarter-plane past. *Journal of Multivariate Analysis*, 127:112–125, 2014.

[8] R. Cheng. Prediction of stationary gaussian random fields with incomplete quarter-plane past. *Journal of Multivariate Analysis*, 139:245–258, 2015.

[9] A. Hamaz. Prediction of random fields with incomplete quarter-plane past. *Communications in Statistics-Theory and Methods*, 48(11):2707–2716, 2019.

[10] A. Hamaz, O. Arezki, and F. Achemine. Impact of missing data on the prediction of random fields. *Journal of Applied Statistics*, 47(1):132–149, 2020.

[11] T. W. Anderson. The statistical analysis of time series. hohn wiley & sons. *New York*, 1971.

[12] H. Helson and D. Lowdenslager. Prediction theory and fourier series in several variables. *Acta mathematica*, 99(1):165–202, 1958.

[13] H. Helson and D. Lowdenslager. Prediction theory and fourier series in several variables. ii. *Acta Mathematica*, 106(3):175–213, 1961.

[14] J. M. Francos, A. Z. Meiri, and B. Porat. A unified texture model based on a 2-D Wold-like decomposition. *IEEE transactions on Signal Processing*, 41(8):2665–2678, 1993.

[15] G. Kallianpur and V. Mandrekar. *Nondeterministic random fields and Wold and Halmos decompositions for commuting isometries*. University of North Carolina, 1982.

[16] D. Tjøstheim. Statistical spatial series modelling ii: some further results on unilateral lattice processes. *Advances in Applied Probability*, 15(3):562–584, 1983.

[17] A. R. Soltani. Extrapolation and moving average representation for stationary random fields and beurling's theorem. *The Annals of Probability*, pages 120–132, 1984.

[18] Tjøstheim. Statistical spatial series modelling.

[19] P. Loubaton. Champs stationnaires au sens large sur $\mathbb{Z}^2$: Propriétés structurelles et mod èles paramétriques. 6(41):223–247, 1989.

[20] O. Arezki and A. Hamaz. On linear prediction for stationary random fields with nonsymmetrical half-plane past. *Communications in Statistics-Theory and Methods*, pages 1–12, 2020.

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

[22] X. Guyon. *Random fields on a network: modeling, statistics, and applications*. Springer Science & Business Media, 1995.

[23] M. Rosenblatt. *Stationary sequences and random fields*. Springer Science & Business Media, 2012.

[24] V. S. Mandrekar and D. A. Redett. *Weakly Stationary Random Fields, Invariant Subspaces and Applications*. Chapman and Hall/CRC, 2017.

[25] A. M. Yaglom. Some classes of random fields in n-dimensional space, related to stationary random processes. *Theory of Probability & Its Applications*, 2(3):273–320, 1957.

[26] K. S. Miller. Complex random fields. *Information Sciences*, 9(3):185–225, 1975.

[27] I. Gihman and A. Skorohod. The theory of stochastic processes I. *Vols. I–III, Grundlagen der Mathematischen Wissenschaften, Vols*, 210:218–323, 1974.

[28] A. M. Yaglom. Second-order homogeneous random fields. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Contributions to Probability Theory*. The Regents of the University of California, 1961.

[29] S. Bochner. Monotone funktionen, stieltjessche integrale und harmonische analyse. *Mathematische Annalen*, 108(1):378–410, 1933.

[30] I. I. Gikhman and A. V. Skorokhod. *The theory of stochastic processes II*. Springer Science & Business Media, 2004.

[31] J. L. Doob. *Stochastic processes*, volume 101. New York Wiley, 1953.

[32] A. N. Kolmogorov. Sur l'interpolation et extrapolation des suites stationnaires. *Comptes Rendus de l'Académie des Sciences,*, 208:2043–2045, 1939.

[33] A. N. Kolmogorov. *On inequalities between the upper bounds of the successive derivatives of an arbitrary function on an infinite interval*. Number 4. American Mathematical Society, 1949.

[34] N. Wiener. Extrapolation, interpolation, and smoothing of stationary time series, 1949.

[35] H. Korezlioglu and P. Loubaton. Spectral factorization of wide sense stationary processes on $\mathbb{Z}^2$. *Journal of Multivariate Analysis*, 19(1):24–47, 1986.

[36] G. Kallianpur and V. Mandrekar. *Nondeterministic random fields and Wold and Halmos decompositions for commuting isometries*. University of North Carolina, 1982.

[37] M. Pourahmadi. Taylor expansion of and some applications. *The American Mathematical Monthly*, 91(5):303–307, 1984.

[38] X. Guyon. Parameter estimation for a stationary process on a d-dimensional lattice. *Biometrika*, 69(1):95–105, 1982.

[39] P. Kohli. *Prediction and Estimation of Random Fields*. PhD thesis, Texas AM University, 2012.

[40] M. Pourahmadi. *Foundations of time series analysis and prediction theory*, volume 379. John Wiley & Sons, 2001.

[41] M. Pourahmadi. Estimation and interpolation of missing values of a stationary time series. *Journal of Time Series Analysis*, 10(2):149–169, 1989.

[42] M. Pourahmadi, A. Inoue, and Y. Kasahara. A prediction problem in $l^2(\omega)$. *Proceedings of the American Mathematical Society*, 135(4):1233–1239, 2007.

[43] T. Nakazi. Two problems in prediction theory. *Studia Mathematica*, 1(78):7–14, 1984.

[44] R. Cheng and S. Seubert. Weakly outer polynomials. *The Michigan Mathematical Journal*, 41(2):235–245, 1994.

[45] R. Cheng. Weakly and strongly outer functions on the bidisc. *The Michigan Mathematical Journal*, 39(1):99–109, 1992.

[46] R. Cheng. The spectral measure of a regular stationary random field with the weak or strong commutation property. *The Annals of Probability*, 21(3):1263–1274, 1993.

[47] R. Cheng, A.G. Miamee, and M. Pourahmadi. Some extermal problems in $l^p(\omega)$. *Proceedings of the American Mathematical Society*, pages 2333–2340, 1998.

[48] S. Basu and G. C. Reinsel. Properties of the spatial unilateral first-order ARMA model. *Advances in applied Probability*, pages 631–648, 1993.

[49] S. Basu and G. C. Reinsel. Properties of the spatial unilateral first-order ARMA model. *Advances in applied Probability*, 25(3):631–648, 1993.

[50] A. Mojiri, H. R. Y. Waghei, Nili Sani, and G. R. M. Borzadaran. Comparison of predictions by kriging and spatial autoregressive models. *Communications in Statistics - Simulation and Computation*, 47(6):1785–1795, 2018.

[51] P. Bondon. Recursive relations for multistep prediction of a stationary time series. *Journal of Time Series Analysis*, 22(4):67 –76, 2002.

[52] R Cheng and M Pourahmadi. Prediction with incomplete past and interpolation of missing values. *Statistics & probability letters*, 33(4):341–346, 1997.

[53] P. Whittle. On stationary processes in the plane. *Biometrika*, pages 434–449, 1954.

[54] V. B. Goryainov. Least-modules estimates for spatial autoregression coefficients. *Journal of Computer and Systems Sciences International*, 50(4):565, 2011.

[55] S. Baran, G. Pap, and M. C. van Zuijlen. Asymptotic inference for an unstable spatial ar model. *Statistics*, 38(6):465–482, 2004.

[56] N. Cressie, C. A. Gotway, and M. O. Grondona. Spatial prediction from networks. *Chemometrics and Intelligent Laboratory Systems*, 7(3):251–271, 1990.

[57] Z. Zhu and M. L. Stein. Spatial sampling design for parameter estimation of the covariance function. *Journal of Statistical Planning and Inference*, 134(2):583–603, 2005.

[58] D. L. Zimmerman. Optimal network design for spatial prediction, covariance parameter estimation, and empirical prediction. *Environmetrics: The official journal of the International Environmetrics Society*, 17(6):635–652, 2006.

Abstract

In this work, we investigate the problem of linear prediction of stationary random fields : We treat the prediction problems where a number of observations are missing to the quarter-plane past of a stationary random field. Our aim is to quantify the influence of missing values on the prediction by giving the simple bounds for the prediction error variance. These bounds allow to characterize the random fields for which the missing observations do not affect the prediction. Furthermore, An explicit autoregressive series representation for the best multi-step ahead linear predictor of stationary random fields with nonsymmetrical half-plane past (NSHP) is established. Necessary and sufficient condition for the mean square convergence of these series is given. Moreover, step recursive relations between the prediction coefficients for the infinite past predictor are provided, these relations are used to calculate explicitly the multi-step prediction coefficients.

Keywords: stationnary random fields ; autoregessive representation ; moving average representation ; linear prediction; multi-step prediction ; mean square convergence ; missing values.

Résumé

Dans ce travail, nous étudions le problème de prédiction linéaire des champs aléatoires stationnaires : Nous traitons le problème de prédiction d'un champ aléatoire stationnaire basé sur un quart du plan altéré par un nombre d'observations manquantes. L'objectif consiste à quantifier l'influence des valeurs manquantes sur la prédiction des champs aléatoires en donnant les bornes de la variance de l'erreur de prédiction. Ces bornes permettent de caractériser les champs aléatoires pour lesquels les observations manquantes n'affectent pas la prédiction. D'autre part, une représentation autorégressive explicite pour le meilleur prédicteur linéaire à multi-pas des champs aléatoires stationnaires basé sur un demi-plan non symétrique (NSHP) est établie. La condition nécessaire et suffisante pour la convergence en moyenne quadratique de cette représentation est donnée. De plus, des relations récursives entre les coefficients de prédiction à multi-pas pour le prédicteur basé sur un passé infini sont fournies, ces relations sont utilisées pour calculer explicitement les coefficients de prédiction à multi-pas. D'autre part,