SAMPLE PATH AND ASYMPTOTIC PROPERTIES OF SPACE-TIME MODELS

By

Yun Xue

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Statistics

2011

ABSTRACT

SAMPLE PATH AND ASYMPTOTIC PROPERTIES OF SPACE-TIME MODELS

By

Yun Xue

Spatio-temporal models are widely used for inference in statistics and many applied areas. In such contexts interests are often in the geometric nature (e.g. anisotropy), and the statistical properties of these models. This dissertation has two parts. The first part focuses on the sample path properties of space-time models. We apply the theory of Yaglom (1957) to construct a large class of space-time models with stationary increments (also called intrinsically stationary random fields) and study their statistical and geometric properties. We derive upper and lower bounds for the prediction errors, establish criteria for the mean-square and sample path differentiability, all in terms of the parameters of the models explicitly. Moreover, it is shown that when the random fields are not smooth, we can generate various kinds of random fractals and the related Hausdorff dimensions are computed. Our main results show that the statistical and geometric properties of the Gaussian random fields we propose are very different from those obtained by deformation from any isotropic random field; and they can be applied to analyze more general Gaussian intrinsic random functions, convolution-based space-time Gaussian models [Higdon (2002), Calder and Cressie (2007)] and the spatial processes in Fuentes (2002, 2005).

The second part of the dissertation pertains to equivalence of Gaussian measures and asymptotically optimal predictions of intrinsically stationary random fields. We extend the methods which Ibragimov and Rozanov (1978) use for stationary processes to study intrinsically stationary random fields. We describe the relationships among three corresponding

Hilbert spaces: the random variable space generated by the random field, the corresponding reproducing kernel Hilbert space, and the complex function space spanned by certain analytic functions using the spectral measure. Criteria for equivalence and orthogonality of intrinsically stationary Gaussian random fields are delivered in terms of their spectral measures and the structures of their reproducing kernel Hilbert spaces. Our results are different from those for stationary processes [see Ibragimov and Rozanov (1978)]. Given the equivalence of two Gaussian measures, the asymptotic optimality of linear predictions of intrinsically stationary random fields and the convergence rates are established in this part. Moreover, the asymptotic efficient prediction of non-stationary, anisotropic space-time models with a misspecified probability distribution is studied. The main results show that under the equivalence of two Gaussian measures, the prediction based on the incorrect distribution is asymptotically optimal and efficient relative to the prediction under the correct distribution, as the points of observations become increasingly dense in the study domain. Our results extend those of Stein (1988, 1990, 1999a, 1999b) which were concerned with isotropic and stationary Gaussian random fields.

This dissertation is dedicated to my parents for all their love and support.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my advisor Professor Yimin Xiao for his invaluable guidance, intellect, support and continuing encouragement throughout my research and the writing of this dissertation. This thesis could not have been finished without the uncountable number of hours he spent sharing his knowledge and discussing various ideas throughout the study. I would also like to express my heartfelt thanks to him for his patience.

In addition. I would like to thank all professors in my dissertation committee: Sasha Kravchenko, Chae Young Lim, V.S. Mandrekar, Mark Meerschaert. I appreciate their time taken out of busy schedules to serve on my committee and for their constructive feedback through the dissertation process.

I give my special thanks to Professor James Stapleton. He is always ready to help me like a father. Over the past five years, he has given me a lot of help and encouragement on my study and life. I also thank all the other professors and my friends for their care and help.

Last but not least, I would like to extend my thanks to my family. I thank my dear husband, Linkan, for his tremendous love and support. His patience and encouragement are highly appreciated. I thank my parents Xiuqin and Shoutuan, who have blessed me with their unending love and support- and their financial assistance to a poor student. I thank my parents-in-law Linying and Youxing for their love, encouragement and understandings. Without their support and love, this dissertation would not have been completed.

TABLE OF CONTENTS

Chapter 1

Introduction and preliminaries

1.1 Introduction

Spatio-temporal models are widely used for inference in statistics and many applied areas such as meteorology, climatology, geophysical science, agricultural sciences, environmental sciences, epidemiology, hydrology. Such models presume, on $\mathbb{R}^d \times \mathbb{R}$, where d is the spatial dimension, a collection of random variables $X(x,t)$ at location x and time t. The family $\{X(x,t) : (x,t) \in \mathbb{R}^d \times \mathbb{R}\}$ is referred to as a spatio-temporal random field or a space-time model.

Many authors have constructed various stationary space-time models and the topic has been under rapid development in recent years. See, for example, Jones and Zhang (1997), Cressie and Huang (1999), de Iaco, Myers and Posa (2001, 2002, 2003), Gneiting (2002), Gneiting, *et al.* (2009), Kolovos, *et al.* (2004), Kyriakidis and Journel (1999), Ma (2003a, 2003b, 2004, 2005a, 2005b, 2007, 2008), Stein (2005) and their combined references for further information on constructions of space-time models and their applications.

There has also been increasing demand for non-stationary space-time models. For example, in the analysis of spatio-temporal data of environmental studies, sometimes there is little reason to expect stationarity under the spatial covariance structures, and it is more advantageous to have a space-time model whose variability changes with location and/or time. Henceforth, the construction of non-stationary space-time models has become an attractive topic and several approaches have been developed recently. These include the deforming of the coordinates of an isotropic and stationary random field to obtain a rich class of non-stationary random fields [see Schmidt and O'Hagan (2003), Anderes and Stein (2008)], or the use of convolution-based methods [cf. Higdon, Swall and Kern (1999), Higdon (2002), Paciorek and Schervish (2006), Calder and Cressie (2007)] or spectral methods [Fuentes (2002, 2005)].

In this dissertation, firstly, we apply the theory of Yaglom (1957) to construct a class of space-time Gaussian models with stationary increments and study their statistical and geometric properties. The main feature of this class of space-time models is that they are anisotropic in time and space, and may have different smoothness and geometric properties along different directions. Such flexible properties make them potentially useful as stochastic models in various areas. By applying tools from Gaussian random fields, fractal geometry and Fourier analysis, we derive upper and lower bounds for the prediction errors, establish criteria for the mean-square and sample path differentiability and determine the Hausdorff dimensions of the sample surfaces, all in terms of the parameters of the models explicitly. Our main results show that the statistical and geometric properties of the Gaussian random fields in this dissertation are very different from those obtained by deformation from any isotropic random field. It is also worthwhile to mention that the methods developed in this dissertation may be applied to analyze more general Gaussian intrinsic random functions, convolution-based space-time Gaussian models [Higdon (2002), Calder and Cressie (2007)] and the spatial processes in Fuentes (2002, 2005).

On the other hand, optimal linear prediction has been widely used in spatial statistics and geostatistics, where it is known as kriging. In kriging, to guarantee good linear predictors based on an estimated Gaussian probability measure, it is of great value to be able to distinguish between two orthogonal probability measures and to determine when one can tell which measure is correct and which is not. Many authors have created various criteria for the equivalence and orthogonality of two Gaussian measures in a one-dimensional stochastic process or Gaussian random field. The references include Gihman and Skorohod (1974), Ibragimov and Rozanov (1978), Parzen (1963), Chatterji and Mandrekar (1978), Kallianpur and Oodaira (1963), Yadrenko (1983), Stein (1999b), Du (2009) and so on. In fact, Parzen (1963) developed an approach for equivalence of two Gaussian measures by using two concepts: the notion of probability spectral density function and the notion of a reproducing kernel Hilbert space (RKHS, for short) of a time series. Chatterji and Mandrekar (1978) also used the method of RKHS to find the sufficient and necessary conditions for the equivalence of two Gaussian measures in a general setting. It is worth noting that the approach which uses RKHS has no constraints like stationarity or isotropy on the underlying process, and the results are applicable to random fields. Ibragimov and Rozanov (1978) obtained necessary and sufficient conditions for equivalence of two Gaussian measures involving the entropy of distributions, and developed the conditions for stationary processes by associating a Hilbert space spanned by certain analytic functions. Moreover, given two equivalent Gaussian processes, Kallianpur and Oodaira (1963) defined the notion of a non-anticipative representation of one of the processes with respect to the other. Later, Yadrenko (1983) extended the results of Ibragimov and Rozanov (1978) to stationary and isotropic random fields. Du (2009) gave some reviews of the basic results for the equivalence and orthogonality of two Gaussian measures, and provided a detailed re-proof of Theorem 4 in Yadrenko (1983), page 156, under the setting of stationary and isotropic random fields. In the literature, there are few explicit results available for the equivalence of two Gaussian measures in a non-stationary random field, especially for anisotropic cases.

In the second part of this dissertation, we extend Ibragimov and Rozanov (1978)'s method to study intrinsically stationary random fields. We describe the relationships among three corresponding Hilbert spaces: the random variable space generated by the random field, the reproducing kernel Hilbert space corresponding to the covariance kernel, and the complex function space spanned by certain analytic functions. Criteria for equivalence and orthogonality of intrinsically stationary Gaussian random fields are given in terms of their probability spectral density functions and the structures of their reproducing kernel Hilbert spaces. The results we have obtained are different from those of stationary processes [see Ibragimov and Rozanov (1978)]. Moreover, given the equivalence of two random fields, we obtain a representation of one of the random fields with respect to the other. The advantage of our representation over the original one is that it is much simpler with respect to some prediction questions.

In practice, the true probability distribution of our Gaussian model is always unknown and must be estimated from the gathered data. To this end, it is of great value to investigate the effect of using a fixed but incorrect probability distribution, especially, when more sample data can be obtained by sampling the spatial or temporal domain increasingly densely (fix-domain asymptotics). We establish the asymptotic optimality of linear predictions of intrinsically stationary Gaussian models and the convergence rates in this dissertation. Moreover, the asymptotic efficient prediction of non-stationary, anisotropic space-time models with a misspecified probability distribution is studied. The main results show that under the equivalence of two Gaussian measures, the prediction based on the incorrect distribution is asymptotically optimal and efficient relative to the prediction under the correct distribution, as the points of observations become increasingly dense in the study domain. The results extend those of Stein (1988, 1990, 1999a, 1999b) which were concerned with isotropic and stationary Gaussian random fields.

The rest of this dissertation is organized as follows. In Section 2 of Chapter 1, we collect definitions and some properties of Gaussian random fields, equivalence and orthogonality of two measures, reproducing kernel Hilbert spaces and Hausdorff dimensions. Chapter 2 studies sample path properties of space-time models. We construct a class of space-time Gaussian models with stationary increments, establish bounds on the prediction errors and investigate smoothness properties and fractal properties of this class of Gaussian models. The results are applied directly to analyze the stationary space-time models in Section 5. Section 6 gives proofs of the main theorems and lemmas in this chapter. In Chapter 3, we investigate asymptotic properties of space-time models. We extend the methods in Ibragimov and Rozanov (1978) to study intrinsically stationary random fields. We obtain criteria for equivalence and orthogonality of this class of random fields in Section 2. The asymptotic optimality of linear predictions and the convergence rates are established in Sections 3 and 4. In Section 5, we present proofs of the main results in this chapter. Finally, we conclude by describing some of our ongoing projects and future work in Chapter 4.

1.2 Definition and preliminaries

This section contains basic definitions and facts of stationary and intrinsically stationary Gaussian random fields, equivalence and orthogonality of two measures, reproducing kernel Hilbert spaces and Hausdorff dimensions, which will be used in subsequent chapters.

Throughout this dissertation, for simplicity of notation, we use $\mathbb{R}^N$ ($\mathbb{R}^N_+ = [0,\infty)^N$) or $\mathbb{R}^d$ (in Chapter 3), instead of $\mathbb{R}^d \times \mathbb{R}$, as the index set for random fields. We use $|\cdot|$ to denote the Euclidean norm in $\mathbb{R}^N$. The inner product in $\mathbb{R}^N$ is denoted by $\langle \cdot, \cdot \rangle$. A typical coordinate, $t \in \mathbb{R}^N$ is written as $t = (t_1, \ldots, t_N)$. For any $s, t \in \mathbb{R}^N$ such that $s_j < t_j$ ($j = 1, \ldots, N$), $[s,t] = \prod_{j=1}^N [s_j, t_j]$ is called a closed interval (or a rectangle). For a positive number x, we use $\lfloor x \rfloor$ to denote the integer part of x.

We will use $c, c_1, c_2, \ldots,$ to denote unspecified positive and finite constants which may not be the same in each occurrence.

1.2.1 Random fields

A formal definition of random fields is as follows:

Definition 1.2.1. *Let a probability space, $(\Omega, \mathcal{U}, P)$, an integer $p \geq 1$, and an index set, T, be given. A random field indexed by T with values in $\mathbb{R}^p$ is then a $\mathbb{R}^p$-valued function $X(t,\omega)$ on $T \times \Omega$ such that for every fixed $t \in T$, $X(t,\cdot)$ is a random vector in $\mathbb{R}^p$.*

In this dissertation, we will take $T = \mathbb{R}^N$, the N-dimensional Euclidean space, or a subset of $\mathbb{R}^N$. In this case, X is simply referred to as an (N,p) random field. The dependency on the underlying probability space will usually be suppressed throughout the text, i.e. we write

$$X(t) = X(t,\omega), \qquad t \in \mathbb{R}^N.$$

For a fixed $\omega \in \Omega$, the function $X(t, \omega) : \mathbb{R}^N \to \mathbb{R}^p$ is a non-random function of t. This deterministic function is usually called a sample path or a realization of the random field. The variable t is called the coordinate or position by standard terminology. In this context, the formal definition of random field simply means:

An (N, p) random field $X(t)$ is a function whose values are random vectors in $\mathbb{R}^p$ for every $t \in \mathbb{R}^N$.

When $N = 1$, the random field is usually called a stochastic process. The term "random field" is usually used to stress that the dimension of the coordinate is higher than one. Random fields in two and three dimensions are widely used as spatial or spatio-temporal models in many applied areas such as meteorology, climatology, geophysical science, agricultural sciences, environmental sciences, epidemiology and hydrology.

In this dissertation, we mainly focus on Gaussian random fields, and study their properties and prediction problems. Gaussian random fields play an important role for several reasons: The specification of their finite-dimensional distributions is simple, and the model is determined by the mean and covariance functions only. Moreover, they are reasonable models for many natural phenomena.

Definition 1.2.2. *A Gaussian random field is a random field where all the finite dimensional distributions are multivariate normal distributions.*

In this dissertation, we construct a class of intrinsically stationary Gaussian random fields (i.e. Gaussian random fields with stationary increments). We study its nature and properties, and compare our results with that of stationary Gaussian random fields. The following are the definitions for stationary and intrinsically stationary Gaussian random fields.

Definition 1.2.3. *A Gaussian random field $X(t)$, $t \in \mathbb{R}^N$ is said to be stationary if its mean function $m(t)$ is constant, and the covariance function*

$$K(s,t) = \mathbb{E}\big[(X(s) - m(s))(X(t) - m(t))\big]$$

depends only on the difference $s - t$, for all $s, t \in \mathbb{R}^N$.

Definition 1.2.4. *A Gaussian random field $X(t)$, $t \in \mathbb{R}^N$ is said to be intrinsically stationary if the increment process $X(t+h) - X(t)$ is stationary for any fixed $h \in \mathbb{R}^N$. Or equivalently, $\forall\, h \in \mathbb{R}^N$*

$$\Big\{X(t+h) - X(t), t \in \mathbb{R}^N\Big\} \stackrel{d}{=} \Big\{X(t) - X(0), t \in \mathbb{R}^N\Big\},$$

where "$\stackrel{d}{=}$" means equality of all finite dimensional distributions.

In this dissertion, we call the pair (m, K) the second-order structure of the Gaussian random field $X(t)$. A random field X is said to be *isotropic* if, for all rotation R in $\mathbb{R}^N$, $X \circ R \stackrel{d}{=} X$. Otherwise, X is said to be *anisotropic*.

1.2.2 Equivalence and orthogonality of two measures

Let $\{X(t), t \in T\}$ be Gaussian random field on the probability space $(\Omega, \mathfrak{U}, P)$. Let P_1 be another Gaussian measure on the σ-algebra $\mathfrak{U}$. P_1 is said to be *absolutely continuous* with respect to P if for all $A \in \mathfrak{U}$, $P(A) = 0$ implies $P_1(A) = 0$. It is known that the absolutely continuous measure P_1 can be represented as

$$P_1(A) = \int_A p(\omega) P(d\omega), \qquad A \in \mathfrak{U},$$

where $p(\omega)$ is a nonnegative function on Ω, which is the Randon-Nikodym derivative of P_1 with respect to P, i.e. $p(\omega) = P_1(d\omega)/P(d\omega)$. We also call it a density of P_1 with respect to P. The two measures P and P_1 are said to be *equivalent* if they are mutually absolutely continuous. The measures P and P_1 are said to be *orthogonal* if there exists $A \in \mathfrak{U}$ such that $P(A) = 1$ and $P_1(A) = 0$. In this case, we also have $P(A^c) = 0$ and $P_1(A^c) = 1$.

Lemma 1.2.5. *Any two Gaussian measures P and P_1 are either equivalent or orthogonal.*

For the proof of this lemma, see page 77 of Ibragimov and Rozanov (1978) or page 117 of Stein (1999b). In Chapter 3 of this dissertation, we will provide some criteria for equivalency of two intrinsically stationary Gaussian random fields.

1.2.3 Reproducing kernel Hilbert spaces

Let $K(s,t)$ be the covariance function of a real-valued random field $X(t)$, $t \in T \subseteq \mathbb{R}^N$. For each $t \in T$, let $K(\cdot,t)$ be the function on T whose value at $s \in T$ is equal to $K(s,t)$. It may be shown [see Aronszajn (1950)] that there exists a unique Hilbert space, denoted as $R_K(T)$, with the following properties:

(1) The members of $R_K(T)$ are real-valued functions on T [if $K(s,t)$ were complex-valued, they would be complex-valued functions].

(2) For every $t \in T$, $K(\cdot,t) \in R_K(T)$.

(3) For every $t \in T$ and $f \in R_K(T)$,

$$f(t) = \langle f, K(\cdot,t)\rangle_{R_K(T)},$$

where the inner product between two functions f and g in $R_K(T)$ is written as $\langle f, g\rangle_{R_K(T)}$.

We call $R_K(T)$ the *reproducing kernel Hilbert space* (RKHS, for short) of the random field $X(t)$ with reproducing kernel $K(s,t)$, $s,t \in T$. In fact, the Hilbert space $R_K(T)$ is the closure of the subspace spanned by the functions $K(\cdot, t)$, $t \in T$.

In the end of Section 2 of Chapter 3, we will encounter another type of kernel. A kernel $b(s,t)$: $T \times T \to \mathbb{R}$ is of *Volterra type* if $b(s,t) \neq 0$ implies $t \leq s$, where $s,t \in T$. Sottinen and Tudor (2006) use this type of kernel to characterize a representation of a Gaussian sheet which is equivalent in law to the Brownian sheet. We can see that the kernel $b(s,t)$ is not a covariance function. But we can use this kernel to get a covariance function $K(s,t)$, $s,t \in T$, as follows:

$$K(s,t) = b(s,t) + b(t,s) - \int_T b(s,u)b(u,t)du.$$

See Sottinen and Tudor (2006) for more details on kernels of Volterra type.

1.2.4 Hausdorff dimension

Let Θ be the class of functions $\phi\colon\ (0,\delta) \to (0,1)$, which are right continuous, monotone increasing with $\phi(0+) = 0$ and satisfying the following "doubling" property:

There exists a finite constant $c > 0$, for which

$$\frac{\phi(2s)}{\phi(s)} \leq c, \qquad 0 < s < \frac{\delta}{2}.$$

For $\phi \in \Theta$, the ϕ-*Hausdorff measure* of $A \subseteq \mathbb{R}^N$ is defined as

$$\phi - m(A) = \liminf_{\epsilon \to 0} \left\{ \sum_j \phi(2r_j) : \quad A \subseteq \bigcup_{j=1}^{\infty} O(x_j, r_j),\ r_j < \epsilon \right\},$$

where $O(x, r)$ denotes the open ball of radius r, centered at x. $\phi - m$ is a metric outer measure and every Borel set in $\mathbb{R}^N$ is $\phi - m$ measurable [cf. Rogers (1970)].

The *Hausdorff dimension* of A is defined as

$$\begin{aligned} \dim_{\mathrm{H}} A &= \inf\{\alpha > 0 : \quad s^{\alpha} - m(A) = 0\} \\ &= \sup\{\alpha > 0 : \quad s^{\alpha} - m(A) = \infty\}. \end{aligned}$$

If $0 < s^{\alpha} - m(A) < \infty$, then A is called an α-set. If there exists $\phi \in \Theta$ with $0 < \phi - m(A) < \infty$, then ϕ is called an exact Hausdorff measure function for A.

The following are some basic properties of Hausdorff dimensions:

(1) Monotonicity: if $A \subseteq B$, then $\dim_{\mathrm{H}} A \leq \dim_{\mathrm{H}} B$.

(2) Hausdorff dimension is σ-stable: $\dim_{\mathrm{H}} \left(\bigcup_{n=1}^{\infty} A_n \right) = \sup_{n \geq 1} \dim_{\mathrm{H}} A_n$.

We refer to Falconer (1990) for more details in Hausdorff dimensions. In Chapter 2 of this dissertation, we write the Hausdorff dimension as dim, instead of $\dim_{\mathrm{H}}$ for simplicity.

Chapter 2

Sample path properties of space-time models

2.1 Introduction

Space-time models are widely used for inference in spatial statistics and geostatistics. Various stationary space-time models have been constructed in the literature, and the topic has been under rapid development in recent years. See, for example, Jones and Zhang (1997), Cressie and Huang (1999), de Iaco, Myers and Posa (2001, 2002, 2003), Gneiting (2002), Gneiting, *et al.* (2009), Kolovos, *et al.* (2004), Kyriakidis and Journel (1999), Ma (2003a, 2003b, 2004, 2005a, 2005b, 2007, 2008), Stein (2005) and their combined references for further information on constructions of space-time models and their applications.

In the meantime, there has also been increasing demand for non-stationary space-time models. For example, in the analysis of spatio-temporal data of environmental studies, sometimes there is little reason to expect stationarity under the spatial covariance structures, and it is more advantageous to have a space-time model whose variability changes with location and/or time. Henceforth, the construction of non-stationary space-time models has become an attractive topic and several approaches have been developed recently. These include the deforming of the coordinates of an isotropic and stationary random field to obtain a rich class of non-stationary random fields [see Schmidt and O'Hagan (2003), Anderes and Stein (2008)], or the use of convolution-based methods [cf. Higdon, Swall and Kern (1999), Higdon (2002), Paciorek and Schervish (2006), Calder and Cressie (2007)] or spectral methods [Fuentes (2002, 2005)].

In this chapter, we apply the theory of Yaglom (1957) to construct a class of space-time Gaussian models with stationary increments and study their statistical and geometric properties. The main feature of this class of space-time models is that they are anisotropic in time and space, and may have different smoothness and geometric properties along different directions. Such flexible properties make them potentially useful as stochastic models in various areas. By applying tools from Gaussian random fields, fractal geometry and Fourier analysis, we derive upper and lower bounds for the prediction errors, establish criteria for mean-square and sample path differentiability and determine the Hausdorff dimensions of the sample surfaces, all in terms of the parameters of the models explicitly. Our main results show that the statistical and geometric properties of the Gaussian random fields in this dissertation are very different from those obtained by deformation from any isotropic random field. It is also worth mentioning that the method in this dissertation may be applied to analyze more general Gaussian intrinsic random functions, convolution-based space-time Gaussian models [Higdon (2002), Calder and Cressie (2007)] and the spatial processes in Fuentes (2002, 2005).

The rest of this chapter is organized as follows. In Section 2 we construct a class of space-time intrinsically stationary Gaussian models by applying the theory of Yaglom (1957). Then we establish upper and lower bounds for the prediction errors of this class of models in Section 3. In Section 4 we consider smoothness properties of the models and establish explicit criteria for the existence of mean-square directional derivatives, mean-square differentiability and sample path continuity of partial derivatives. In Section 5 we look into the fractal properties of these models and determine the Hausdorff dimensions of the range, graph and level sets. In Section 6, we apply the main results of Section 5 to some stationary space-time models, such as those constructed by Cressie and Huang (1999), Gneiting (2002) and Stein (2005). Finally, in Section 7, we provide the proofs of the main results in this chapter.

2.2 Anisotropic Gaussian models with stationary increments

We consider a special class of *intrinsic random functions*; namely, space-time models with stationary increments (also called intrinsically stationary space-time models). We will further restrict ourselves to Gaussian random fields for which powerful general Gaussian principles can be applied. Many of the results in this chapter can be extended to non-Gaussian space-time models (such as stable or more general infinitely divisible random fields), but their proofs require different methods and go beyond the scope of this chapter. One can find some information for stable random fields in Xiao (2011).

Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a real-valued, centered Gaussian random field with $X(0) = 0$. We assume that X has stationary increments and continuous covariance function $K(s,t) = \mathbb{E}[X(s)X(t)]$. According to Yaglom (1957), $K(s,t)$ can be represented as

$$K(s,t) = \int_{\mathbb{R}^N} \big(e^{i\langle s,\lambda\rangle} - 1\big)\big(e^{-i\langle t,\lambda\rangle} - 1\big) F(d\lambda) + \langle s, Wt\rangle, \tag{2.1}$$

where W is an $N \times N$ non-negative definite matrix and $F(d\lambda)$ is a nonnegative symmetric measure on $\mathbb{R}^N \setminus \{0\}$ satisfying

$$\int_{\mathbb{R}^N} \frac{|\lambda|^2}{1+|\lambda|^2} F(d\lambda) < \infty. \tag{2.2}$$

In analogy to the stationary case, the measure F is called the *spectral measure* of X. If F is absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^N$, its density f will be called the *spectral density* of X.

It follows from (2.1) that X has the following stochastic integral representation:

$$\Big\{\mathbf{X}(t),\, t \in \mathbb{R}^N\Big\} \stackrel{d}{=} \Big\{\int_{\mathbb{R}^N} \big(e^{i\langle t,\lambda\rangle} - 1\big)\Phi(d\lambda) + \langle Y, t\rangle,\, t \in \mathbb{R}^N\Big\}, \tag{2.3}$$

where $X_1 \stackrel{d}{=} X_2$ means the processes X_1 and X_2 have the same finite dimensional distributions, Y is an N-dimensional Gaussian random vector with mean 0 and covariance matrix W, $\Phi(d\lambda)$ is a centered complex-valued Gaussian random measure which is independent of Y and satisfies

$$\mathbb{E}\Big(\Phi(A)\overline{\Phi(B)}\Big) = F(A \cap B) \quad \text{and} \quad \Phi(-A) = \overline{\Phi(A)}$$

for all Borel sets $A, B \subseteq \mathbb{R}^N$, with finite F-measure. The spectral measure F is called the *control measure* of Φ. Since the linear term $\langle Y, t\rangle$ in (2.3) will not have any effect on the problems considered in this dissertation, we will from now on assume $Y = 0$. This is equivalent to assuming $W = 0$ in (2.1). Consequently, we have

$$v(h) \triangleq \mathbb{E}\big(X(t+h) - X(t)\big)^2 = 2\int_{\mathbb{R}^N} \big(1 - \cos\langle h, \lambda\rangle\big) F(d\lambda). \tag{2.4}$$

It is important to note that the function $v(h)$, called *variogram* in spatial statistics, is a negative definite function in the sense of I. J. Schoenberg, which is determined by the spectral measure F. See Berg and Forst (1975) for more information on negative definite functions.

The above shows that various centered intrinsically stationary Gaussian random fields can be constructed by choosing appropriate spectral measures F. For the well known fractional Brownian motion $B^H = \{B^H(t), t \in \mathbb{R}^N\}$ of Hurst index $H \in (0, 1)$, its spectral measure has a density function

$$f_H(\lambda) = c(H, N)\frac{1}{|\lambda|^{2H+N}},$$

where $c(H, N) > 0$ is a normalizing constant such that $v(h) = |h|^{2H}$. Since $v(h)$ depends on $|h|$ only, B^H is isotropic. Other examples of isotropic Gaussian fields with stationary increments can be found in Xiao (2007). We also remark that all centered stationary Gaussian random fields can be treated using the above framework. In fact, if $Z = \{Z(t), t \in \mathbb{R}^N\}$ is a centered stationary Gaussian random field, it can be represented as $Z(t) = \int_{\mathbb{R}^N} e^{i\langle t,\lambda\rangle}\, \Phi(d\lambda)$. Thus the random field X defined by

$$X(t) = Z(t) - Z(0) = \int_{\mathbb{R}^N} \left(e^{i\langle t,\lambda\rangle} - 1\right) \Phi(d\lambda), \quad \forall\ t \in \mathbb{R}^N$$

is Gaussian with stationary increments (intrinsically stationary Gaussian random field) and $X(0) = 0$. Note that the spectral measure F of X in the sense of (2.4) is the same as the spectral measure [in the ordinary sense] of the stationary random field Z.

In the following, we propose and investigate a class of centered, anisotropic, intrinsically stationary Gaussian random fields, whose spectral measures are absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^N$. More precisely, we assume that the spectral measure

F of $X = \{X(t), t \in \mathbb{R}^N\}$ is absolutely continuous with density function $f(\lambda)$ which satisfies (2.2) and the following condition:

(C) There exist positive constants c_1, c_2, c_3, γ and $(\beta_1, \cdots, \beta_N) \in (0, \infty)^N$ such that

$$\gamma > \sum_{j=1}^{N} \frac{1}{\beta_j} \tag{2.5}$$

and

$$\frac{c_1}{\left(\sum_{j=1}^N |\lambda_j|^{\beta_j}\right)^\gamma} \le f(\lambda) \le \frac{c_2}{\left(\sum_{j=1}^N |\lambda_j|^{\beta_j}\right)^\gamma}, \qquad \forall\, \lambda \in \mathbb{R}^N \text{ with } |\lambda| \ge c_3. \tag{2.6}$$

The following proposition shows that (2.5) is needed to ensure f is a legitimate spectral density function.

Proposition 2.2.1. *Assume that $f(\lambda)$ is a non-negative measurable function defined on $\mathbb{R}^N$. If*

$$\int_{|\lambda| \le c_3} |\lambda|^2 f(\lambda) d\lambda < \infty$$

and (2.6) holds, then $f(\lambda)$ is a legitimate spectral density if and only if the parameters γ and β_j for $j = 1, \cdots, N$ satisfy (2.5).

Some remarks about Condition (C) are given in the following.

Remark 2.2.2

(1) There is an important connection between the random field models that satisfy Condition (C) and those considered in Xiao (2009). For $j = 1, \cdots, N$, let

$$H_j = \frac{\beta_j}{2}\left(\gamma - \sum_{i=1}^N \frac{1}{\beta_i}\right) \tag{2.7}$$

and let $Q = \sum_{j=1}^N \frac{1}{H_j}$. Then (2.6) can be rewritten as

$$\frac{c_4}{\left(\sum_{j=1}^N |\lambda_j|^{H_j}\right)^{2+Q}} \le f(\lambda) \le \frac{c_5}{\left(\sum_{j=1}^N |\lambda_j|^{H_j}\right)^{2+Q}}, \qquad \forall\, \lambda \in \mathbb{R}^N \text{ with } |\lambda| \ge c_3, \tag{2.8}$$

where the positive and finite constants c_4 and c_5 depend on N, c_1, c_2, β_j and γ only. To verify this claim, we will make use of the following elementary fact: For any positive numbers N and q, there exist positive and finite constants c_4 and c_5 such that

$$c_4\Big(\sum_{j=1}^N a_j\Big)^q \le \sum_{j=1}^N a_j^q \le c_5\Big(\sum_{j=1}^N a_j\Big)^q \tag{2.9}$$

for all non-negative numbers $a_1, \ldots, a_N$. Note that

$$\left(\sum_{j=1}^N |\lambda_j|^{H_j}\right)^{2+Q} = \left(\sum_{j=1}^N |\lambda_j|^{\beta_j \cdot \frac{1}{2}\left(\gamma - \sum_{i=1}^N \frac{1}{\beta_i}\right)}\right)^{2+Q}$$

and $\frac{1}{2}\big(\gamma - \sum_{i=1}^N \frac{1}{\beta_i}\big)(2+Q) = \gamma$. We apply (2.9) with $q = \frac{1}{2}\big(\gamma - \sum_{i=1}^N \frac{1}{\beta_i}\big)$ to see that (2.6) and (2.8) are equivalent.

In turns out that the expression (2.8) is essential in this chapter and will be used frequently. For simplicity of notation, from now on we take $c_3 = 1$.

(2) It is also possible to consider intrinsically stationary Gaussian random fields whose spectral measures are not absolutely continuous. Some examples of such covariance space-time models can be found in Cressie and Huang (1999), Gneiting (2002), Ma (2003a, 2003b). Since the mathematical tools for studying such random fields are quite different [see Luan and Xiao (2010)], we will deal with them systematically in the future.

(3) Non-stationary Gaussian random fields can be constructed through deformation of an isotropic Gaussian random field. Refer to Anderes and Stein (2008) for more details. One of the advantages of deformation is to closely connect a nonstationary and/or anisotropic random field to a stationary and isotropic one for which existing statistical techniques are available. However, there is also a disadvantage [from the point of view of flexibility] associated with deformation. Let $X(t) = Z(g^{-1}(t))$, where $\{Z(t), t \in \mathbb{R}^N\}$ is an isotropic Gaussian model and g is a smooth bijection of $\mathbb{R}^N$. Since the function g is bi-Lipschitz on compact intervals, the fractal dimensional properties of X are the same as those of Z. Hence deformation of isotropic Gaussian models will not generate anisotropic random fields with rich geometric structures as shown by the models introduced in this chapter.

2.3 Prediction error for anisotropic Gaussian models

Suppose we observe an anisotropic Gaussian random field X on $\mathbb{R}^N$ at $t^1, \ldots, t^n$ and wish to predict $X(u)$, for $u \in \mathbb{R}^N$. Then the inference about $X(u)$ will be based upon the conditional distribution of $X(u)$ given the observed values of $X(t^1), \ldots, X(t^n)$. Refer to Stein (1999b, Section 1.2) for the closed form of this conditional distribution. A statistical analysis typically aims at the optimal linear predictor of this unobserved $X(u)$, known as *simple kriging.* The simple kriging predictor of $X(u)$ is

$$X^*(u) = \mathbf{c}(u)^T \mathbf{\Sigma}^{-1} \mathbf{Z}, \tag{2.10}$$

where $\mathbf{Z} = \big(X(t^1), \ldots, X(t^n)\big)^T$, $\mathbf{c}(u)^T = \mathrm{Cov}\{X(u), \mathbf{Z}\}$ and $\mathbf{\Sigma} = \mathrm{Cov}(\mathbf{Z}, \mathbf{Z}^T)$. The form (2.10) minimizes the mean square prediction error, which then is given as $\mathrm{Var}(X(u)) - \mathbf{c}(u)^T \mathbf{\Sigma}^{-1} \mathbf{c}(u)$. Since X is Gaussian, the simple kriging is the conditional expectation of $X(u)$ given $\mathbf{Z}$, and the mean square prediction error is the conditional variance of $X(u)$ given the observations $\mathbf{Z}$.

The main result of this section is Theorem 2.3.1 below, which gives lower and upper bounds for the mean square prediction error for intrinsically stationary Gaussian random fields which satisfy Condition (C). It shows that, similar to stationary Gaussian field models [cf. Stein (1999b)], the prediction error of the models in this chapter only depends on the high frequency behavior of the spectral density of X.

Theorem 2.3.1. *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a centered intrinsically stationary Gaussian random field valued in $\mathbb{R}$ with spectral density $f(\lambda)$ satisfying (2.6). Then, for any given constant $M > 0$, there exist constants $c_6 > 0$ and $c_7 > 0$, such that for all integers $n \geq 1$ and all $u, t^1, \cdots, t^n \in [-M, M]^N$,*

$$c_6 \min_{0 \leq k \leq n} \sum_{j=1}^{N} |u_j - t_j^k|^{2H_j} \leq \mathrm{Var}\big(X(u) | X(t^1), \cdots, X(t^n)\big) \leq c_7 \min_{0 \leq k \leq n} \sum_{j=1}^{N} \sigma_j\big(|u_j - t_j^k|\big), \tag{2.11}$$

where H_j is given in (2.7), $t^0 = 0$ and $\sigma_j : \mathbb{R}_+ \to \mathbb{R}_+$ is defined by

$$\sigma_j(r) = \begin{cases} r^{2H_j} & \text{if } 0 < H_j < 1, \\ r^2 |\log r| & \text{if } H_j = 1, \\ r^2 & \text{if } H_j > 1. \end{cases} \tag{2.12}$$

If $H_j < 1$, for $j = 1, \cdots, N$, then the two bounds in (2.11) match. When there is some

$H_j > 1$, that means, the random field $X(t)$ is smoother in the j-th direction [see Corollaries 2.4.3, 2.4.7 and Theorem 2.4.8 below], then the upper and lower bounds are not the same any more. This suggests that the prediction error may become larger as $X(t)$ becomes smoother in some directions.

The proof of Theorem 2.3.1, as well as those of Theorems 2.4.9, 2.5.1 and 2.5.2 relies partially on the following lemma, which provides upper and lower bounds for the variogram of the model.

Lemma 2.3.2. *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a centered intrinsically stationary Gaussian random field valued in $\mathbb{R}$ with spectral density $f(\lambda)$ satisfying (2.6). Then, for any given constant $M > 0$, there exist constants $c_8 > 0$ and $c_9 > 0$, such that for $s, t \in [-M, M]^N$,*

$$c_8 \sum_{j=1}^{N} \sigma_j\big(|s_j - t_j|\big) \le \mathbb{E}\big(X(s) - X(t)\big)^2 \le c_9 \sum_{j=1}^{N} \sigma_j\big(|s_j - t_j|\big), \tag{2.13}$$

where the function σ_j is defined in (2.12).

The upper bound in (2.13) implies that X has a version whose sample functions are almost surely continuous. Throughout this dissertation, without loss of generality, we will assume that the sample function $t \mapsto X(t)$ is almost surely continuous.

2.4 Smoothness properties of anisotropic Gaussian models

Regularity properties of sample path of random fields are of fundamental importance in probability and statistics. Many authors have studied mean square and sample path continuity and differentiability of Gaussian processes and random fields. See Cramér and Leadbetter

(1967), Alder (1981), Stein (1999b), Banerjee and Gelfand (2003), Adler and Taylor (2007). In this section we provide explicit criteria for mean square and sample path differentiability for the models introduced in Section 2.

2.4.1 Distributional properties of mean square partial derivatives

Banerjee and Gelfand (2003) studied the smoothness properties of stationary random fields and some non-stationary relatives through directional derivative processes and their distributional properties. To apply their method to intrinsically stationary random fields, let us first recall the definition of mean square directional derivatives.

Definition 2.4.1. *Let $u \in \mathbb{R}^N$ be a unit vector. A second order random field $\{X(t), t \in \mathbb{R}^N\}$ has mean square directional derivative $X'_u(t)$ at $t \in \mathbb{R}^N$ in the direction u if, as $h \to 0$,*

$$X_{u,h}(t) = \frac{X(t+hu) - X(t)}{h}$$

converges to $X'_u(t)$ in the L_2-sense. In this case, we write $X'_u(t) = \text{l.i.m.}_{h\to 0} X_{u,h}(t)$.

Let $e_1, e_2, \cdots, e_N$ be an orthonormal basis for $\mathbb{R}^N$. If $u = e_j$, then $X'_{e_j}(t)$ is the mean square partial derivative in the j-th direction defined in Adler (1981), which will simply be written as $X'_j(t)$. We will also write $X_{e_j,h}(t)$ as $X_{j,h}(t)$.

For any second-order, centered random field $\{X(t), t \in \mathbb{R}^N\}$, similar to Theorem 2.2.2 in Adler (1981), one can easily establish a criterion in terms of the covariance function $K(s,t) = \mathbb{E}\big[X(s)X(t)\big]$ for the existence of mean square directional derivative $X'_u(t)$. Banerjee and

Gelfand (2003) further showed that the covariance function of $X'_u(t)$ is given by

$$
\begin{aligned}
K_u(s,t &= \lim_{h\to 0}\lim_{k\to 0}\mathbb{E}\big[X_{u,h}(s)X_{u,k}(t)\big]\\
&= \lim_{h\to 0}\lim_{k\to 0}\frac{K(s+hu,t+ku)-K(s+hu,t)-K(s,t+ku)+K(s,t)}{hk}.
\end{aligned}
$$

Extending their argument, one obtains the following theorem for intrinsically stationary Gaussian random fields.

Theorem 2.4.2. *Let $X=\{X(t),t\in\mathbb{R}^N\}$ be a centered intrinsically stationary Gaussian random field valued in $\mathbb{R}$, then the mean square partial derivative $X'_j(t)$ exists for all $t\in\mathbb{R}^N$ if and only if the limit*

$$
\lim_{h,k\to 0}\frac{v(he_j)+v(ke_j)-v\big((h-k)e_j\big)}{hk} \tag{2.14}
$$

exists, where $v(t)$ is defined in (2.4). Moreover, this later condition is equivalent to $v(t)$ has second-order partial derivatives at 0 in the j-th direction.

As a consequence, we obtain an explicit criterion for the existence of mean square partial derivatives of Gaussian random fields in Section 2.

Corollary 2.4.3. *Let $X=\{X(t),t\in\mathbb{R}^N\}$ be a centered intrinsically stationary Gaussian random field valued in $\mathbb{R}$ with spectral density $f(\lambda)$ satisfying Condition (C). Then for every $j=1,\cdots,N$, the mean square partial derivative $X'_j(t)$ exists if and only if*

$$
\beta_j\left(\gamma-\sum_{i=1}^{N}\frac{1}{\beta_i}\right)>2, \tag{2.15}
$$

or equivalently $H_j>1$ [cf. (2.7)].

Assume condition (2.14) of Theorem 2.4.2 holds so that the mean square partial derivative

$X_j'(t)$ exists for all $t \in \mathbb{R}^N$. We now consider the distributional properties of the random field $\{X_j'(t), t \in \mathbb{R}^N\}$.

Since $\mathbb{E}(X(t)) = 0$ for all $t \in \mathbb{R}^N$, we have $\mathbb{E}(X_{j,h}(t)) = 0$ and $\mathbb{E}(X_j'(t)) = 0$. Let $K_j^{(h)}(s,t)$ and $K_j(s,t)$ denote the covariance functions of the random fields $\{X_{j,h}(t), t \in \mathbb{R}^N\}$ and $\{X_j'(t), t \in \mathbb{R}^N\}$, respectively. Let $\Delta = s - t$, we immediately have

$$K_j^{(h)}(s,t) = \frac{v(\Delta + he_j) + v(\Delta - he_j) - 2v(\Delta)}{2h^2}, \tag{2.16}$$

and $\mathrm{Var}(X_{j,h}(t)) = v(he_j)/h^2$, which only depends on the scalar h.

Theorem 2.4.4. *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a centered intrinsically stationary Gaussian random field valued in $\mathbb{R}$. Suppose that all second-order partial derivatives of the variogram $v(t)$ exist. Then the covariance function of $X_j'(t)$ is given by*

$$K_j(s,t) = \frac{1}{2} v_j''(s-t), \tag{2.17}$$

where $v_j''(t)$ is the second-order partial derivative of v at t in the j-th direction. In particular, $\{X_j'(t), t \in \mathbb{R}^N\}$ is a stationary Gaussian random field.

Proof The desired result follows from (2.16). $\square$

It is also useful to determine the covariance of $X(s)$ and $X_j'(t)$ for all $s,t \in \mathbb{R}^N$. Since

$$\mathrm{Cov}\big(X(s), X_{j,h}(t)\big) = \frac{1}{2h}\Big\{v(t + he_j) - v(t) + v(\Delta) - v(\Delta - he_j)\Big\},$$

where $\Delta = s - t$, we obtain

$$\mathrm{Cov}\big(X(t), X_{j,h}(t)\big) = \frac{1}{2h}\Big\{v(t+he_j) - v(t) - v(he_j)\Big\}$$

and

$$\begin{aligned}\mathrm{Cov}\big(X(s), X_j'(t)\big) &= \lim_{h\to 0}\frac{1}{2h}\Big\{v(t+he_j) - v(t) + v(\Delta) - v(\Delta - he_j)\Big\} \\ &= \frac{1}{2}\big(v_j'(t) + v_j'(\Delta)\big),\end{aligned} \tag{2.18}$$

where $v_j'(t)$ is the partial derivative of v at t in the j-th direction.

In particular, $\mathrm{Cov}\big(X(t), X_j'(t)\big) = v_j'(t)/2$, which is different from the stationary case. Recall that if $Z(t)$ is a stationary Gaussian field with mean square partial derivative $Z_j'(t)$, then $Z(t)$ and $Z_j'(t)$ are uncorrelated (and thus independent). However, this is not always true for non-stationary Gaussian random fields, which is one of the reasons why non-stationary models are more difficult to study.

Next we consider the bivariate process

$$Y_j^{(h)}(t) = \begin{pmatrix} X(t) \\ X_{j,h}(t) \end{pmatrix}.$$

It can be verified that this process has mean 0 and cross-covariance matrix

$$\begin{aligned}&V_{j,h}(s,t) \\ &= \begin{pmatrix} \dfrac{v(s)+v(t)-v(\Delta)}{2} & \dfrac{v(t+he_j)-v(t)+v(\Delta)-v(\Delta-he_j)}{2h} \\ \dfrac{v(s+he_j)-v(s)+v(\Delta)-v(\Delta+he_j)}{2h} & \dfrac{v(\Delta+he_j)+v(\Delta-he_j)-2v(\Delta)}{2h^2} \end{pmatrix}.\end{aligned}$$

Because $Y_j^{(h)}(t)$ is obtained by linear transformation of $X(t)$, the above is a valid cross-covariance matrix in $\mathbb{R}^N$. Since this is true for every h, letting $h \to 0$ we see that

$$V_j(s,t) = \begin{pmatrix} \frac{1}{2}\{v(s)+v(t)-v(\Delta)\} & \frac{1}{2}\{v_j'(t)+v_j'(\Delta)\} \\ \frac{1}{2}\{v_j'(s)-v_j'(\Delta)\} & \frac{1}{2}v_j''(\Delta) \end{pmatrix}$$

is a valid cross-covariance matrix in $\mathbb{R}^N$. In fact, V_j is the cross-covariance matrix for the bivariate process

$$Y_j(t) = \begin{pmatrix} X(t) \\ X_j'(t) \end{pmatrix}.$$

2.4.2 Criterion for mean square differentiability

Benerjee and Gelfand (2003) pointed out that the existence of all mean square directional derivatives of a random field X does not even guarantee mean square continuity of X, and they introduced a notion of mean square differentiability which has analogous properties of total differentiability of a function in $\mathbb{R}^N$ in the non-stochastic setting. We first recall their definition.

Definition 2.4.5. *A random field $\{X(t), t \in \mathbb{R}^N\}$ is mean square differentiable at $t \in \mathbb{R}^N$ if there exists a (random) vector $\nabla_X(t) \in \mathbb{R}^N$ such that for all scalar $h > 0$, all vectors $u \in \mathcal{S}_N = \{t \in \mathbb{R}^N : |t| = 1\}$*

$$X(t+hu) = X(t) + hu^T \nabla_X(t) + r(t,hu), \tag{2.19}$$

where $r(t,hu)/h \to 0$ in the L_2-sense as $h \to 0$.

Refer to Definition 2.1 of Potthoff (2010) for more details on the definition of mean square differentiability. In other words, for all vectors $u \in \mathcal{S}_N$, it is required that

$$\lim_{h \to 0} \mathbb{E}\left(\frac{X(t+hu) - X(t) - hu^T \nabla_X(t)}{h}\right)^2 = 0. \tag{2.20}$$

It can be seen that if X is mean square differentiable at t, then for all unit vectors $u \in \mathcal{S}_N$

$$\begin{aligned} X'_u(t) &= \underset{h \to 0}{\text{l.i.m.}} \frac{X(t+hu) - X(t)}{h} = \underset{h \to 0}{\text{l.i.m.}} \frac{hu^T \nabla_X(t) + r(t, hu)}{h} \\ &= u^T \nabla_X(t). \end{aligned}$$

Hence it is necessary that $\nabla_X(t) = (X'_1(t), \ldots, X'_N(t))$.

The next theorem provides a sufficient condition for a intrinsically stationary Gaussian random field to be mean square differentiable.

Theorem 2.4.6. *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a centered intrinsically stationary Gaussian random field valued in $\mathbb{R}$. If all the second-order partial and mixed derivatives of the variogram $v(t)$ exist and are continuous, then X is mean square differentiable at every $t \in \mathbb{R}^N$.*

As a consequence of Theorem 2.4.6 and Corollary 2.4.3 we obtain

Corollary 2.4.7. *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a centered intrinsically stationary Gaussian random field valued in $\mathbb{R}$ with spectral density $f(\lambda)$ satisfying Condition (C). Then X is mean square differentiable at every $t \in \mathbb{R}^N$ if and only if*

$$\beta_j \left(\gamma - \sum_{i=1}^{N} \frac{1}{\beta_i}\right) > 2 \qquad \text{for every } j = 1, \ldots, N. \tag{2.21}$$

2.4.3 Criterion for sample path differentiability

For many theoretical and applied purposes, one often needs to work with random fields which have smooth sample functions. Refer to Adler (1981), Adler and Taylor (2007) and the reference therein for more information. Since in general mean square differentiability does not imply almost sure sample path differentiability, it is of interest to provide convenient criteria for the latter.

For Gaussian random fields considered in this chapter, it turns out that under the same condition as Corollary 2.4.3, the partial derivatives of X are almost surely continuous.

Theorem 2.4.8. *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a separable and centered intrinsically stationary Gaussian random field with values in $\mathbb{R}$. We assume that X satisfies Condition (C).*

(i). *If*

$$\beta_j\left(\gamma - \sum_{i=1}^{N} \frac{1}{\beta_i}\right) > 2 \qquad (i.e., H_j > 1), \tag{2.22}$$

for some $j \in \{1, \cdots, N\}$, then X has a version $\widetilde{X}$ with continuous sample functions such that its jth partial derivative $\widetilde{X}'_j(t)$ is continuous almost surely.

(ii). *If (2.22) holds for all $j \in \{1, \cdots, N\}$, then X has a version $\widetilde{X}$ which is continuously differentiable in the following sense: with probability 1,*

$$\lim_{h\to 0} \frac{\widetilde{X}(t+hu) - \widetilde{X}(t) - hu^T \nabla_{\widetilde{X}}(t)}{h} = 0 \qquad \text{for all } u \in \mathcal{S}_N \ \text{ and } t \in \mathbb{R}^N. \tag{2.23}$$

If condition (2.22) does not hold for some $j \in \{1, \cdots, N\}$, then $X(t)$ does not have mean square partial derivatives along the j-th direction and the sample path of $X(t)$ is usually a random fractal. In this case, it is of interest to characterize the asymptotic behavior of $X(t)$ by its local and uniform moduli of continuity.

These problems for anisotropic Gaussian random fields have been considered in Xiao (2009), Meerschaert, Wang and Xiao (2010). The methods there are applicable to X with a little modification. For completeness, we state the following result which can be proved by using Lemma 2.3.2 and general Gaussian methods. We omit its proof.

Theorem 2.4.9. *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be as in Theorem 2.4.8. Then for every compact interval $I \subset \mathbb{R}^N$, there exists a positive and finite constant c_{10}, depending only on I and H_j, $(j = 1, \ldots, N)$ such that*

$$\limsup_{|\varepsilon| \to 0} \frac{\sup_{t \in I,\, s \in [0,\varepsilon]} |X(t+s) - X(t)|}{\sqrt{\varphi(\varepsilon) \log(1 + \varphi(\varepsilon)^{-1})}} \leq c_{10}, \tag{2.24}$$

where $\varphi(\varepsilon) = \sum_{j=1}^{N} \sigma_j(|\varepsilon_j|)$ for all $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_N) \in \mathbb{R}^N$, and the function σ_j is defined in (2.12).

2.5 Fractal properties of anisotropic Gaussian models

The variations of soil, landform and geology are usually highly non-regular in form and can be better approximated by a stochastic fractal. Hausdorff dimensions have been extensively used in describing fractals. We refer to Kahane (1985) or Falconer (1990) for their definitions and properties.

Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a real-valued, centered Gaussian random field. For any integer $p \geq 1$, we define an (N, p)-Gaussian random field $\mathbf{X} = \{\mathbf{X}(t), t \in \mathbb{R}^N\}$ by

$$\mathbf{X}(t) = \big(X_1(t), \ldots, X_p(t)\big), \qquad t \in \mathbb{R}^N, \tag{2.25}$$

where $X_1, \ldots, X_p$ are independent copies of X.

In this section, under more general conditions on X than those in Sections 2–4, we study the Hausdorff dimensions of the range $\mathbf{X}([0,1]^N) = \{\mathbf{X}(t) : t \in [0,1]^N\}$, the graph $\mathrm{Gr}\mathbf{X}([0,1]^N) = \{(t, \mathbf{X}(t)) : t \in [0,1]^N\}$ and the level set $\mathbf{X}^{-1}(x) = \{t \in \mathbb{R}^N : \mathbf{X}(t) = x\}$ ($x \in \mathbb{R}^p$). The results in this section can be applied to wide classes of Gaussian spatial or space-time models (with or without stationary increments).

First, consider fractional Brownian motion $B^H = \{B^H(t), t \in \mathbb{R}^N\}$ valued in $\mathbb{R}^p$ with Hurst index $H \in (0,1)$. $B^H(t)$ is a special example of our model which, however, has isotropic spectral density. It is known [cf. Kahane (1985)] that

$$\dim \mathrm{Gr}\, B^H\big([0,1]^N\big) = \min\left\{N + (1-H)p,\ \frac{N}{H}\right\} \qquad \text{a.s.}$$

Especially, when $p = 1$,

$$\dim \mathrm{Gr}\, B^H\big([0,1]^N\big) = N + 1 - H \qquad \text{a.s.}$$

and moreover, for every $x \in \mathbb{R}$,

$$\dim\,(B^H)^{-1}(x) = N - H, \qquad \text{a.s.}$$

The fractal properties of B^H have been applied by many statisticians to estimate the Hurst index H and it is sufficient to choose $p = 1$. Refer to Hall and Wood (1993), Constantine and Hall (1994), Kent and Wood (1997), Davis and Hall (1999), Chan and Wood (2000, 2004), Zhu and Stein (2002).

Let $(\overline{H}_1, \ldots, \overline{H}_N) \in (0,1]^N$ be a constant vector. Without loss of generality, we assume that they are ordered as

$$0 < \overline{H}_1 \leq \overline{H}_2 \leq \cdots \leq \overline{H}_N \leq 1. \tag{2.26}$$

We assume the following conditions.

(D1) For any $\eta > 0$, there exist positive constants δ_0, $c_{11} \geq 1$ such that for all $s, t \in [0,1]^N$ with $|s-t| \leq \delta_0$

$$c_{11}^{-1} \sum_{j=1}^{N} |s_j - t_j|^{2\overline{H}_j + \eta} \leq \mathbb{E}\big[\big(X(t) - X(s)\big)^2\big] \leq c_{11} \sum_{j=1}^{N} |s_j - t_j|^{2\overline{H}_j - \eta}. \tag{2.27}$$

(D2) For any constant $\varepsilon \in (0,1)$, there exists a positive constant c_{12} such that for all $u, t \in [\varepsilon, 1]^N$, we have

$$\mathrm{Var}\,(X(u) \mid X(t)) \geq c_{12} \sum_{j=1}^{N} \big|u_j - t_j\big|^{2\overline{H}_j + \eta}. \tag{2.28}$$

The following theorems determine the Hausdorff dimensions of range, graph and level sets of $\mathbf{X}$. Because of anisotropy, these results are significantly different from the aforementioned results for fractional Brownian motion or other isotropic random fields [cf. Xiao (2007)]. Even though Theorems 2.5.1 and 2.5.2 below are similar to Theorems 6.1 and 7.1 in Xiao (2009), they have wider applicability. In particular, they can be applied to a random field X which may be smooth in certain (or all) directions.

Theorem 2.5.1. *Let $\mathbf{X} = \{\mathbf{X}(t), t \in \mathbb{R}^N\}$ be an (N,p)-Gaussian random field defined by (2.25). If the coordinate process X satisfies Condition (D1), then, with probability 1,*

$$\dim \mathbf{X}\big([0,1]^N\big) = \min\left\{p; \sum_{j=1}^{N} \frac{1}{\overline{H}_j}\right\}, \tag{2.29}$$

and

$$\dim \mathrm{Gr}\mathbf{X}\big([0,1]^N\big) = \min_{1\le k\le N} \left\{ \sum_{j=1}^{k} \frac{\overline{H}_k}{\overline{H}_j} + N - k + (1-\overline{H}_k)p;\ \sum_{j=1}^{N} \frac{1}{\overline{H}_j} \right\}, \tag{2.30}$$

where $\sum_{j=1}^{0} \frac{1}{\overline{H}_j} := 0$.

Proof The right inequality in (2.27) and Theorem 2.4.9 show that $\mathbf{X}(t)$ satisfies a uniform Hölder condition on $[0,1]^N$ which, in turn, implies the desired upper bounds in (2.29) and (2.30).

The lower bounds for $\dim \mathbf{X}\big([0,1]^N\big)$ and $\dim \mathrm{Gr}\mathbf{X}\big([0,1]^N\big)$ can be derived from the left inequality in (2.27) and a capacity argument. See the proof of Theorem 6.1 in Xiao (2009) for details. □

For the level sets of $\mathbf{X}$, we have

Theorem 2.5.2. *Let* $\mathbf{X} = \{\mathbf{X}(t), t \in \mathbb{R}^N\}$ *be an* (N,p)*-Gaussian random field defined by (2.25). If the coordinate process* X *satisfies Conditions (D1) and (D2), then the following statements hold:*

(i) *If* $\sum_{j=1}^{N} \frac{1}{\overline{H}_j} < p$, *then for every* $x \in \mathbb{R}^p \backslash \{0\}$, $\mathbf{X}^{-1}(x) = \emptyset$ *a.s.*

(ii) *If* $\sum_{j=1}^{N} \frac{1}{\overline{H}_j} > p$, *then for any* $x \in \mathbb{R}^p$, *with positive probability*

$$\dim \mathbf{X}^{-1}(x) = \min_{1\le k\le N} \left\{ \sum_{j=1}^{k} \frac{\overline{H}_k}{\overline{H}_j} + N - k - \overline{H}_k\, p \right\}. \tag{2.31}$$

Proof The results (i) and (ii) follow from the proof of Theorem 7.1 in Xiao (2009). □

Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a centered intrinsically stationary Gaussian random field valued in $\mathbb{R}$ with spectral density $f(\lambda)$ satisfying (2.6). Let $H_1, \ldots, H_N$ be defined by (2.7).

Then, by Theorem 2.3.1 and Lemma 2.3.2, we see that X satisfies (D1) for all $\overline{H}_j = 1 \wedge H_j$ $(1 \le j \le N)$. It also satisfies Condition (D2) with $\overline{H}_j = H_j$ provided $H_j \le 1$ for all $j = 1, \cdots, N$. Hence one can apply Theorems 2.5.1 and 2.5.2 to derive the following result.

Corollary 2.5.3. *Let* $\mathbf{X} = \{\mathbf{X}(t), t \in \mathbb{R}^N\}$ *be a centered intrinsically stationary Gaussian random field valued in* $\mathbb{R}^p$ *defined by (2.25). We assume that its coordinate process* X *has spectral density* $f(\lambda)$ *satisfying (2.6) and* H_j $(j = 1, \cdots, N)$ *defined by (2.7) are ordered as* $H_1 \le H_2 \le \cdots \le H_N$. *We have*

(i) *With probability 1, (2.29) and (2.30) hold with* $\overline{H}_j = 1 \wedge H_j$ $(1 \le j \le N)$.

(ii) *If, in addition,* $H_j \le 1$ *(so* $\overline{H}_j = H_j$*) for all* $j = 1, \cdots, N$ *and* $\sum_{j=1}^{N} \frac{1}{H_j} > p$, *then (2.31) holds with positive probability.*

We believe that the above fractal properties can also be useful for estimating the parameters $H_1, \ldots, H_N$ of our model. However this will be more subtle than the isotropic case, where only the single parameter is involved, for the following two reasons. First, if a parameter $H_j > 1$, then the sample function $X(t)$ is smooth in the j-th direction and the Hausdorff dimensions of $\mathbf{X}$ has nothing to do with H_j. In other words, based on fractal dimensions, a parameter H_j can be explicitly estimated only when $H_j < 1$.

Secondly, if we let $p = 1$, then (2.30) gives $\dim \mathrm{Gr} X\big([0,1]^N\big) = N + 1 - \overline{H}_1$, which does not give any information about the other parameters $H_2, \ldots, H_N$. This suggests that, in order to estimate all the parameters of an anisotropic random field model, one has to work with a *multivariate* random field $\mathbf{X}$ as defined by (2.25).

2.6 Applications to some stationary space-time models

The above results can be applied to the stationary space-time Gaussian fields constructed by Cressie and Huang (1999), Gneiting (2002), de Iaco, Myers, and Posa (2002), Ma (2003a, 2003b) and Stein (2005).

2.6.1 Stationary covariance models

Extending the results of Cressie and Huang (1999), Gneiting (2002) showed that, for $(x, t) \in \mathbb{R}^d \times \mathbb{R}$,

$$K(x,t) = \frac{\sigma^2}{\big(1 + a|t|^{2\alpha}\big)^{\beta d/2}} \exp\bigg(- \frac{c|x|^{2\gamma}}{\big(1 + a|t|^{2\alpha}\big)^{\beta\gamma}} \bigg), \tag{2.32}$$

is a stationary space-time covariance function, where $\sigma > 0$, $a > 0$, $c > 0$, $\alpha \in (0, 1]$, $\beta \in (0, 1]$ and $\gamma \in (0, 1]$ are constants. It can be verified that the corresponding spectral measure is absolutely continuous in the space-variable x and discrete in the time-variable t. See Ma (2003a, 2003b) for more examples of stationary covariance models.

In the following, we verify that the sample functions of these space-time models are fractals. We will check Conditions (D1) and (D2) first, and then obtain the corresponding Hausdorff dimension results from Theorems 2.5.1 and 2.5.2.

Proposition 2.6.1. *Let $X = \{X(x, t), (x, t) \in \mathbb{R}^d \times \mathbb{R}\}$ be a centered stationary Gaussian random field in $\mathbb{R}$ with covariance function as (2.32). Then for any $M > 0$, there exist constants $c_{13} > 0$ and $c_{14} > 0$ such that*

$$c_{13}\big(|x - y|^{2\gamma} + |t - s|^{2\alpha}\big) \le \mathbb{E}\big(X(x,t) - X(y,s)\big)^2 \le c_{14}\big(|x - y|^{2\gamma} + |t - s|^{2\alpha}\big) \tag{2.33}$$

and

$$\mathrm{Var}\big(X(x,t)\big|X(y,s)\big) \geq c_{13}\left(|x-y|^{2\gamma}+|t-s|^{2\alpha}\right) \tag{2.34}$$

for all (x,t) *and* $(y,s) \in [-M,M]^{d+1}$.

Proposition 2.6.2. *Let* $X = \{X(x,t),(x,t) \in \mathbb{R}^d \times \mathbb{R}\}$ *be a centered stationary Gaussian random field in* $\mathbb{R}$ *with covariance function as (2.32), and let* $\mathbf{X}$ *be its associated* (N,p)*-random field defined by (2.25). Then, with probability 1,*

$$\dim \mathbf{X}\big([0,1]^{d+1}\big) = \min\left\{p\,;\ \frac{d}{\gamma}+\frac{1}{\alpha}\right\}. \tag{2.35}$$

And if $0<\alpha\leq\gamma<1$, *then*

$$\dim \mathrm{Gr}\mathbf{X}\big([0,1]^{d+1}\big) = \begin{cases} d+1+(1-\alpha)p & \text{if } p<\frac{1}{\alpha}, \\ d+\frac{\gamma}{\alpha}+(1-\gamma)p & \text{if } \frac{1}{\alpha}\leq p<\frac{1}{\alpha}+\frac{d}{\gamma}, \\ \frac{1}{\alpha}+\frac{d}{\gamma} & \text{if } p\geq\frac{1}{\alpha}+\frac{d}{\gamma}. \end{cases} \tag{2.36}$$

If $0<\gamma\leq\alpha<1$, *then*

$$\dim \mathrm{Gr}\mathbf{X}\big([0,1]^{d+1}\big) = \begin{cases} d+1+(1-\gamma)p & \text{if } p<\frac{d}{\gamma}, \\ \frac{d\alpha}{\gamma}+1+(1-\alpha)p & \text{if } \frac{d}{\gamma}\leq p<\frac{1}{\alpha}+\frac{d}{\gamma}, \\ \frac{1}{\alpha}+\frac{d}{\gamma} & \text{if } p\geq\frac{1}{\alpha}+\frac{d}{\gamma}. \end{cases} \tag{2.37}$$

Remark 2.6.3 Applying the method in Luan and Xiao (2011), it is possible to further determine the exact Hausdorff measure function for $\mathbf{X}\big([0,1]^{d+1}\big)$.

Proposition 2.6.4. *Let* $X = \{X(x,t),(x,t) \in \mathbb{R}^d \times \mathbb{R}\}$ *be a centered stationary Gaussian random field in* $\mathbb{R}$ *with covariance function as (2.32), and let* $\mathbf{X}$ *be its associated* (N,p)*-*

random field.

(i) *When* $\frac{1}{\alpha} + \frac{d}{\gamma} < p$*, then for every* $x \in \mathbb{R}^p$, $\mathbf{X}^{-1}(x) = \emptyset$ *a.s.*

(ii) *When* $\frac{1}{\alpha} + \frac{d}{\gamma} > p$*, if* $0 < \alpha \le \gamma \le 1$*, then for any* $x \in \mathbb{R}^p$*, with positive probability*

$$\dim \mathbf{X}^{-1}(x) = \begin{cases} d + 1 - \alpha p & \text{if } p < \frac{1}{\alpha}, \\ d + \frac{\gamma}{\alpha} - \gamma p & \text{if } p \ge \frac{1}{\alpha}, \end{cases} \tag{2.38}$$

and if $0 < \gamma \le \alpha \le 1$*, then for any* $x \in \mathbb{R}^p$*, with positive probability*

$$\dim \mathbf{X}^{-1}(x) = \begin{cases} d + 1 - \gamma p & \text{if } p < \frac{d}{\gamma}, \\ \frac{d\alpha}{\gamma} + 1 - \alpha p & \text{if } p \ge \frac{d}{\gamma}. \end{cases} \tag{2.39}$$

2.6.2 Stationary spectral density models

In Section 2.6.1, the stationary space-time models are constructed directly by covariance functions, which are isotropic in the space variable. Stein (2005) showed that stationary covariance functions which are anisotropic in space can be constructed by choosing

$$f(\lambda) = \left(\sum_{j=1}^{d+1} c_j \big(a_j + |\lambda_j|^2\big)^{\alpha_j} \right)^{-\nu}, \qquad \forall \lambda \in \mathbb{R}^d \times \mathbb{R}, \tag{2.40}$$

where $\nu > 0$, $c_j > 0$, $a_j > 0$ and $\alpha_j \in \mathbb{N}$ for $j = 1, \cdots, d+1$ are constants such that

$$\sum_{j=1}^{d+1} \frac{1}{\alpha_j} < 2\nu.$$

This last condition guarantees $f \in L^1(\mathbb{R}^{d+1})$. Clearly $f(\lambda)$ in (2.40) satisfies (2.6) with $\beta_j = \alpha_j$ and $\gamma = 2\nu$. Hence we may apply our results to analyze this class of models, through their smoothness and fractal properties.

Proposition 2.6.5. *Let $X = \{X(x,t), (x,t) \in \mathbb{R}^d \times \mathbb{R}\}$ be a centered stationary Gaussian random field in $\mathbb{R}$ with spectral density as (2.40).*

(i) *If*

$$2\nu > \sum_{j=1}^{d+1} \frac{1}{\alpha_j} + \frac{2}{\min_{1\le\ell\le d+1} \alpha_\ell},$$

then $X(x,t)$ is mean square differentiable and has a version $\widetilde{X}(x,t)$ which is sample path differentiable almost surely.

(ii) *X is a fractal [i.e. the sample path of X may have fractional Hausdorff dimension] if and only if*

$$\sum_{j=1}^{d+1} \frac{1}{\alpha_j} < 2\nu \le \sum_{j=1}^{d+1} \frac{1}{\alpha_j} + \frac{2}{\min_{1\le\ell\le d+1} \alpha_\ell}.$$

The Hausdorff dimensions of various fractals generated by this kind of model can also be computed using Corollary 2.5.3, with $H_j = \alpha_j\big(\nu - \sum_{\ell=1}^{d+1} \frac{1}{2\alpha_\ell}\big)$, $\overline{H}_j = 1 \wedge H_j$ for $j = 1, \cdots, d+1$. We leave the details to an interested reader.

2.7 Proofs

Proof of Proposition 2.2.1

Note that (2.2) is equivalent to $\int_{\mathbb{R}^N} \big(1 \wedge |\lambda|^2\big) f(\lambda) d\lambda < \infty$. Since $\int_{|\lambda|\le 1} |\lambda|^2 f(\lambda) d\lambda < \infty$ is given, it is enough for us to show $\int_{|\lambda|>1} \frac{d\lambda}{\left(\sum_{j=1}^N |\lambda_j|^{\beta_j}\right)^\gamma} < \infty$ is equivalent to (2.5).

For this purpose, we appeal to the following fact: Given positive constants β and γ, there exists a finite constant c_{15} such that for all $a > 0$,

$$\int_0^\infty \frac{dx}{(a + x^\beta)^\gamma} = \begin{cases} c_{15}\, a^{-(\gamma - \frac{1}{\beta})} & \text{if } \beta\gamma > 1, \\ +\infty & \text{if } \beta\gamma \le 1. \end{cases} \tag{2.41}$$

To verify this, we make a change of variable $x = a^{\frac{1}{\beta}} y$ to obtain

$$\int_0^\infty \frac{dx}{(a + x^\beta)^\gamma} = a^{-(\gamma - \frac{1}{\beta})} \int_0^\infty \frac{dy}{(1 + y^\beta)^\gamma}.$$

Thus (2.41) follows.

First we assume (2.5) holds. Since $|\lambda| > 1$ implies that $|\lambda_{j_0}| > \frac{1}{\sqrt{N}}$ for some $j_0 \in \{1, \ldots, N\}$. Without loss of generality we assume $j_0 = 1$. Then by using (2.41) $(N-1)$ times we obtain

$$\begin{aligned} \int_{|\lambda|>1} \frac{d\lambda}{\big(\sum_{j=1}^N |\lambda_j|^{\beta_j}\big)^\gamma} &\le 2^N \int_{\frac{1}{\sqrt{N}}}^\infty d\lambda_1 \underbrace{\int_0^\infty \cdots \int_0^\infty}_{N-2} \frac{d\lambda_2 \cdots d\lambda_{N-1}}{\big(\sum_{j=1}^{N-1} |\lambda_j|^{\beta_j}\big)^{\gamma - \frac{1}{\beta_N}}} \\ &\le c \int_{\frac{1}{\sqrt{N}}}^\infty \frac{d\lambda_1}{\big(|\lambda_1|^{\beta_1}\big)^{\gamma - \sum_{j=2}^N \frac{1}{\beta_j}}} < \infty, \end{aligned}$$

because $\beta_1\big(\gamma - \sum_{j=2}^N \frac{1}{\beta_j}\big) > 1$. This proves the sufficiency of (2.5).

To prove the converse, we assume (2.5) does not hold. Then there is a unique integer $\tau \in \{1, \ldots, N\}$ such that $\sum_{i=1}^{\tau-1} \frac{1}{\beta_i} < \gamma \le \sum_{i=1}^{\tau} \frac{1}{\beta_i}$. Note that

$$\int_{|\lambda|>1} \frac{d\lambda}{\big(\sum_{j=1}^N |\lambda_j|^{\beta_j}\big)^\gamma} \ge \underbrace{\int_0^\infty \cdots \int_0^\infty}_{N-1} \int_1^\infty \frac{d\lambda_1 \cdots d\lambda_N}{\big(\sum_{j=1}^N |\lambda_j|^{\beta_j}\big)^\gamma}.$$

By using (2.41) and integrating $d\lambda_1 \cdots d\lambda_\tau$, we see that the last integral is divergent. This completes the proof. $\square$

Proof of Lemma 2.3.2

For any $s, t \in [-M,\, M]^N$, denote $\hat{s}_0 = t$, $\hat{s}_1 = (s_1, t_2, \cdots, t_N)$, $\hat{s}_2 = (s_1, s_2, t_3, \cdots, t_N)$, $\cdots$, $\hat{s}_{N-1} = (s_1, \cdots, s_{N-1}, t_N)$ and $\hat{s}_N = s$. Let $h = s - t \triangleq (h_1, \cdots, h_N)$. By Jensen's inequality, (2.4) and (2.8) we can write

$$\begin{aligned}
\mathbb{E}\big(X(s) - X(t)\big)^2 &= \mathbb{E}\Big[\sum_{k=1}^{N} \big(X(\hat{s}_k) - X(\hat{s}_{k-1})\big)\Big]^2 \\
&\le N \sum_{k=1}^{N} \mathbb{E}\big(X(\hat{s}_k) - X(\hat{s}_{k-1})\big)^2 \\
&= 2N \sum_{k=1}^{N} \int_{\mathbb{R}^N} \big(1 - \cos(h_k \lambda_k)\big) f(\lambda) d\lambda \\
&\le 2N \sum_{k=1}^{N} \int_{|\lambda| \le 1} \big(1 - \cos(h_k \lambda_k)\big) f(\lambda) d\lambda \\
&\qquad + 2N c_5 \sum_{k=1}^{N} \int_{|\lambda| > 1} \big(1 - \cos(h_k \lambda_k)\big) \frac{d\lambda}{\big(\sum_{i=1}^{N} |\lambda_i|^{H_i}\big)^{Q+2}} \\
&\triangleq I_1 + I_2.
\end{aligned} \tag{2.42}$$

By using the inequality $1 - \cos x \le x^2$ we have

$$\begin{aligned}
I_1 &\le 2N \Big(\sum_{k=1}^{N} h_k^2\Big) \int_{|\lambda| \le 1} |\lambda|^2 f(\lambda) d\lambda \\
&\le c_{16}\, |s - t|^2
\end{aligned} \tag{2.43}$$

for some positive and finite constant c_{16}, which depends on M. To bound the kth integral in I_2, we note that, when $|\lambda| > 1$, either $|\lambda_k| > \frac{1}{\sqrt{N}}$ or there is $j_0 \ne k$ such that $|\lambda_{j_0}| > \frac{1}{\sqrt{N}}$.

We break the integral according to these two possibilities.

$$\begin{aligned}
&\int_{|\lambda|>1} \big(1-\cos(h_k\lambda_k)\big)\frac{d\lambda}{\big(\sum_{i=1}^N |\lambda_i|^{H_i}\big)^{Q+2}} \\
&\le 2\int_{\frac{1}{\sqrt{N}}}^{\infty} \big(1-\cos(h_k\lambda_k)\big)d\lambda_k \int_{\mathbb{R}^{N-1}} \frac{d\lambda_1\cdots d\lambda_{k-1}d\lambda_{k+1}\cdots d\lambda_N}{\big(\sum_{i=1}^N |\lambda_i|^{H_i}\big)^{Q+2}} \\
&\qquad + 4\int_0^1 \big(1-\cos(h_k\lambda_k)\big)d\lambda_k \int_{\frac{1}{\sqrt{N}}}^{\infty} d\lambda_{j_0}\int_{\mathbb{R}^{N-2}} \frac{d\lambda^{\vee}_{k,j_0}}{\big(\sum_{i=1}^N |\lambda_i|^{H_i}\big)^{Q+2}} \\
&\triangleq I_3 + I_4,
\end{aligned} \tag{2.44}$$

where $d\lambda^{\vee}_{k,j_0}$ denotes integration in λ_i $(i \ne k, j_0)$.

By using (2.41) repeatedly [$N-1$ times], we obtain

$$\begin{aligned}
I_3 &\le c\int_{\frac{1}{\sqrt{N}}}^{\infty} \frac{1-\cos(h_k\lambda_k)}{|\lambda_k|^{2H_k+1}}\, d\lambda_k \\
&\le c\left(\int_{\frac{1}{\sqrt{N}}}^{\frac{1}{|h_k|}} \frac{h_k^2\lambda_k^2}{\lambda_k^{2H_k+1}} d\lambda_k + \int_{\frac{1}{|h_k|}}^{\infty} \frac{1}{\lambda_k^{2H_k+1}} d\lambda_k\right) \\
&\le c\,\sigma_k(|h_k|),
\end{aligned} \tag{2.45}$$

where σ_k is defined as in (2.12).

Similarly, we use (2.41) $N-2$ times to get

$$\begin{aligned}
I_4 &\le c\int_0^1 \big(1-\cos(h_k\lambda_k)\big)d\lambda_k \int_{\frac{1}{\sqrt{N}}}^{\infty} \frac{d\lambda_{j_0}}{\big(\lambda_k^{H_k}+\lambda_{j_0}^{H_{j_0}}\big)^{2+\frac{1}{H_k}+\frac{1}{H_{j_0}}}} \\
&\le c\int_0^1 \big(1-\cos(h_k\lambda_k)\big)d\lambda_k \int_{\frac{1}{\sqrt{N}}}^{\infty} \frac{d\lambda_{j_0}}{\lambda_{j_0}^{2H_{j_0}+1+\frac{H_{j_0}}{H_k}}} \le c\,|h_k|^2.
\end{aligned} \tag{2.46}$$

Combining (2.42)–(2.46) yields the upper bound in (2.13) holds for all $s, t \in [-M, M]^N$.

Next we prove the lower bound in (2.13). By (2.4) and (2.8) we have

$$\mathbb{E}\big(X(s)-X(t)\big)^2 \geq c_4 \int_{|\lambda|>1} \big(1-\cos\langle s-t,\lambda\rangle\big)\frac{d\lambda}{\rho(\lambda)^{Q+2}}, \tag{2.47}$$

where $\rho(\lambda)=\sum_{j=1}^N |\lambda_j|^{H_j}$, $\lambda\in\mathbb{R}^N$. So, for the lower bound of $\mathbb{E}\big(X(s)-X(t)\big)^2$, it is enough to show that for every $j=1,\cdots,N$ and all $h\in\mathbb{R}^N$, we have

$$\int_{|\lambda|>1} \big(1-\cos\langle h,\lambda\rangle\big)\frac{d\lambda}{\rho(\lambda)^{Q+2}} \geq c\,\sigma_j(|h_j|), \tag{2.48}$$

where c is a positive constant.

We only prove (2.48) for $j=1$, and the other cases are similar. Fix $h\in\mathbb{R}^N$ with $|h_1|>0$ [otherwise there is nothing to prove] and we make a change of variables

$$y_\ell=\rho(h)^{H_\ell^{-1}}\lambda_\ell, \qquad \forall\, \ell=1,\cdots,N.$$

We consider a subset of the integration region defined by

$$D(h)=\Big\{y\in\mathbb{R}^N : |y_1|\in[\rho(h)^{H_1^{-1}},1],\ |y_\ell|\leq 1 \ \text{ and } y_\ell h_\ell>0 \text{ for } 1\leq\ell\leq N\Big\}.$$

Since $\rho(\lambda)=\rho(y)/\rho(h)$, we have

$$\int_{|\lambda|>1} \big(1-\cos\langle h,\lambda\rangle\big)\frac{d\lambda}{\rho(\lambda)^{Q+2}} \geq \rho(h)^2 \int_{D(h)} \frac{1-\cos\Big(\sum_{\ell=1}^N h_\ell \rho(h)^{-H_\ell^{-1}} y_\ell\Big)}{\big(\sum_{\ell=1}^N |y_\ell|^{H_\ell}\big)^{Q+2}}\,dy. \tag{2.49}$$

By using the inequality $1-\cos x\geq c\,x^2$ for all $|x|\leq N$, where $c>0$ is a constant, and the fact that $h_\ell y_\ell>0$ for all $1\leq\ell\leq N$, we derive that the last integral is at least [up to a constant]

$$
\begin{aligned}
&\rho(h)^2 \int_{D(h)} \frac{\Big(\sum_{\ell=1}^N h_\ell \rho(h)^{-H_\ell^{-1}} y_\ell\Big)^2}{\big(\sum_{\ell=1}^N |y_\ell|^{H_\ell}\big)^{Q+2}}\, dy \\
&\geq \rho(h)^2 \int_{\rho(h)^{H_1^{-1}}}^1 h_1^2 \rho(h)^{-\frac{2}{H_1}} y_1^2 dy_1 \underbrace{\int_0^1 \cdots \int_0^1}_{N-1} \frac{dy_2 \cdots dy_N}{\big(\sum_{\ell=1}^N |y_\ell|^{H_\ell}\big)^{Q+2}} \\
&\geq c\, \rho(h)^{2-\frac{2}{H_1}} h_1^2 \int_{\rho(h)^{H_1^{-1}}}^1 \frac{y_1^2 dy_1}{\big(y_1^{H_1}\big)^{\frac{1}{H_1}+2}} \\
&= c\, \sigma_1(|h_1|).
\end{aligned}
\tag{2.50}
$$

This proves (2.48) and hence Lemma 2.3.2. □

In order to prove Theorem 2.3.1, we need use the following lemma which implies that the prediction error of X is determined by the behavior of the spectral density $f(\lambda)$ at infinity.

Lemma 2.7.1. *Assume (2.6) is satisfied, then for any fixed constant $M > 0$, there exists a positive and finite constant c_{17} such that for all functions g of the form*

$$
g(\lambda) = \sum_{k=1}^n a_k \big(e^{i\langle t^k, \lambda\rangle} - 1\big),
$$

where $a_k \in \mathbb{R}$ and $t^k \in [-M, M]^N$, we have

$$
|g(\lambda)| \leq c_{17}\, |\lambda| \bigg(\int_{\mathbb{R}^N} |g(\xi)|^2 f(\xi) d\xi \bigg)^{1/2} \tag{2.51}
$$

for all $\lambda \in \mathbb{R}^N$ that satisfy $|\lambda| \leq 1$.

Proof By (2.6), we can find positive constants C and η, such that

$$f(\lambda) \geq \frac{C}{|\lambda|^{\eta}}, \quad \forall \lambda \in \mathbb{R}^N \text{ with } |\lambda| \text{ large enough.}$$

Then the desired result follows from the proof of Lemma 2.2 in Xiao (2007). □

Proof of Theorem 2.3.1

First, let's prove the upper bound in (2.11). By Lemma 2.3.2 we have

$$\begin{aligned} \mathrm{Var}\Big(X(u)|X(t^1), \cdots, X(t^n)\Big) &\leq \min_{0 \leq k \leq n} \mathbb{E}\big(X(u) - X(t^k)\big)^2 \\ &\leq c_9 \min_{0 \leq k \leq n} \sum_{j=1}^{N} \sigma_j\big(|u_j - t_j^k|\big). \end{aligned} \tag{2.52}$$

In order to prove the lower bound for the conditional variance in (2.11), we assume that $u, t^1, \cdots, t^n \in [-M, M]^N$ are arbitrary and denote $r \equiv \min_{0 \leq k \leq n} \sum_{j=1}^N |u_j - t_j^k|^{H_j}$. Working in the Hilbert space setting, the conditional variance is just the square of $L^2(\mathbb{P})$-distance of $X(u)$ from the subspace generated by $\{X(t^1), \cdots, X(t^n)\}$, so it is sufficient to prove that for all $a_k \in \mathbb{R}$, $1 \leq k \leq n$,

$$\mathbb{E}\bigg(X(u) - \sum_{k=1}^{n} a_k X(t^k)\bigg)^2 \geq c_6\, r^2, \tag{2.53}$$

where c_6 is a positive constant which may only depend on $H_1, \ldots, H_N$ and N.

By using the stochastic integral representation (2.3) of X, the left hand side of (2.53) can be written as

$$\mathbb{E}\bigg(X(u) - \sum_{k=1}^{n} a_k X(t^k)\bigg)^2 = \int_{\mathbb{R}^N} \bigg| e^{i\langle u, \lambda\rangle} - 1 - \sum_{k=1}^{n} a_k (e^{i\langle t^k, \lambda\rangle} - 1)\bigg|^2 f(\lambda)\, d\lambda. \tag{2.54}$$

Hence, we only need to show

$$\int_{\mathbb{R}^N} \left| e^{i\langle u,\lambda\rangle} - \sum_{k=0}^{n} a_k e^{i\langle t^k,\lambda\rangle} \right|^2 f(\lambda)\, d\lambda \geq c_6\, r^2, \tag{2.55}$$

where $t^0 = 0$ and $a_0 = 1 - \sum_{k=1}^n a_k$.

We choose a function $\delta(\cdot) : \mathbb{R}^N \to [0,1]$ in $C^\infty(\mathbb{R}^N)$ [the space of all infinitely differentiable functions defined on $\mathbb{R}^N$] such that $\delta(0) = 1$ and it vanishes outside the open set $\big\{t \in \mathbb{R}^N : \sum_{j=1}^N |t_j|^{H_j} < 1\big\}$. Denote by $\hat{\delta}$ the Fourier transform of δ. Then one can verify that $\hat{\delta}(\cdot) \in C^\infty(\mathbb{R}^N)$ as well and $\hat{\delta}(\lambda)$ decays rapidly as $|\lambda| \to \infty$.

Let E be the $N \times N$ diagonal matrix with $H_1^{-1}, \cdots, H_N^{-1}$ on its diagonal and let $\delta_r(t) = r^{-Q}\, \delta(r^{-E}t)$ for all $t \in \mathbb{R}^N$. Then the inverse Fourier transformation and a change of variables yield

$$\delta_r(t) = (2\pi)^{-N} \int_{\mathbb{R}^N} e^{-i\langle t,\lambda\rangle} \hat{\delta}(r^E\lambda)\, d\lambda. \tag{2.56}$$

Since $\min\big\{ \sum_{j=1}^N |u_j - t_j^k|^{H_j} : 0 \leq k \leq n\big\} \geq r$, we have $\delta_r(u - t^k) = 0$ for $k = 0, 1, \cdots, n$. This and (2.56) together imply that

$$\begin{aligned} J :&= \int_{\mathbb{R}^N} \left(e^{i\langle u,\lambda\rangle} - \sum_{k=0}^{n} a_k e^{i\langle t^k,\lambda\rangle} \right) e^{-i\langle u,\lambda\rangle} \hat{\delta}(r^E\lambda)\, d\lambda \\ &= (2\pi)^N \left(\delta_r(0) - \sum_{k=0}^{n} a_k \delta_r(u - t^k) \right) \\ &= (2\pi)^N r^{-Q}. \end{aligned} \tag{2.57}$$

Now we split the integral in (2.57) over $\{\lambda : |\lambda| < 1\}$ and $\{\lambda : |\lambda| \geq 1\}$ and denote the two integrals by I_1 and I_2, respectively. It follows from Lemma 2.7.1 that

$$
\begin{aligned}
I_1 &\leq \int_{|\lambda|<1} \Big| e^{i\langle u,\lambda\rangle} - \sum_{k=0}^{n} a_k e^{i\langle t^k,\lambda\rangle} \Big| |\hat{\delta}(r^E\lambda)| d\lambda \\
&\leq c_{17} \left[\int_{\mathbb{R}^N} \Big| e^{i\langle u,\lambda\rangle} - \sum_{k=0}^{n} a_k e^{i\langle t^k,\lambda\rangle} \Big|^2 f(\lambda)\, d\lambda \right]^{1/2} \int_{|\lambda|<1} |\lambda||\hat{\delta}(r^E\lambda)| d\lambda \qquad (2.58) \\
&\leq c_{18} \left[\mathbb{E}\Big(X(u) - \sum_{k=1}^{n} a_k X(t^k) \Big)^2 \right]^{1/2},
\end{aligned}
$$

where the last inequality follows from (2.54) and the boundedness of $\hat{\delta}$.

On the other hand, by the Cauchy-Schwarz inequality and (2.54), we have

$$
\begin{aligned}
I_2^2 &\leq \int_{|\lambda|\geq 1} \Big| e^{i\langle u,\lambda\rangle} - \sum_{k=0}^{n} a_k e^{i\langle t^k,\lambda\rangle} \Big|^2 f(\lambda) d\lambda \int_{|\lambda|\geq 1} \frac{1}{f(\lambda)} |\hat{\delta}(r^E\lambda)|^2\, d\lambda \\
&\leq \mathbb{E}\Big(X(u) - \sum_{k=1}^{n} a_k X(t^k) \Big)^2 r^{-Q} \int_{|\lambda|\geq 1} \frac{1}{f(r^{-E}\lambda)} |\hat{\delta}(\lambda)|^2\, d\lambda \qquad (2.59) \\
&= \mathbb{E}\Big(X(u) - \sum_{k=1}^{n} a_k X(t^k) \Big)^2 r^{-2Q-2} \int_{|\lambda|\geq 1} \frac{1}{f(\lambda)} |\hat{\delta}(\lambda)|^2\, d\lambda.
\end{aligned}
$$

The last integral is convergent thanks to the fast decay of $\hat{\delta}(\lambda)$. Finally, combining (2.57), (2.58) and (2.59), we get

$$
(2\pi)^N r^{-Q} \leq c_{19} \left[\mathbb{E}\Big(X(u) - \sum_{k=1}^{n} a_k X(t^k) \Big)^2 \right]^{1/2} r^{-Q-1}.
$$

Henceforth (2.53) follows, and the theorem was proved because of (2.52) and (2.53). □

Proof of Theorem 2.4.2

For $t \in \mathbb{R}^N$, it is known that $X_{j,h} = \frac{X(t+he_j)-X(t)}{h}$ converges in L^2-sense, as $h \to 0$, if and only if

$$D_{h,k} \triangleq \frac{1}{hk}\mathbb{E}\Big\{\big(X(t+he_j)-X(t)\big)\big(X(t+ke_j)-X(t)\big)\Big\}$$

converges to a constant as $h, k \to 0$. However,

$$\begin{aligned} D_{h,k} &= \frac{1}{hk}\Big\{C(t+he_j, t+ke_j) - C(t, t+ke_j) - C(t+he_j, t) + C(t,t)\Big\} \\ &= \frac{1}{2hk}\Big\{v(he_j) + v(ke_j) - v\big((h-k)e_j\big)\Big\}. \end{aligned} \tag{2.60}$$

So the first part of the theorem is proved. For the second part, it is clear that if $v(t)$ has second-order partial derivatives at 0 in the j-th direction then (2.14) holds [thanks to Taylor's theorem]. On the other hand, if (2.14) holds, then by taking $h = k \to 0$ in (2.60) we see that $\partial v/\partial t_j(0) = 0$. This fact, together with (2.14), implies that

$$\begin{aligned} \frac{\partial^2 v}{\partial t_j^2}(0) &= \lim_{k\to 0}\frac{1}{k}\lim_{h\to 0}\frac{v((k+h)e_j) - v(ke_j)}{h} \\ &= \lim_{k\to 0}\lim_{h\to 0}\frac{v((k+h)e_j) - v(ke_j) + v(he_j)}{hk} \end{aligned}$$

exists. This completes the proof of Theorem 2.4.2. □

Proof of Corollary 2.4.3

By Theorem 2.4.2 it suffices to show that $\lim_{h,k\to 0} D_{h,k}$ exists if and only if (2.15) holds, i.e.,

$$\beta_j\big(\gamma - \sum_{i=1}^{N}\frac{1}{\beta_i}\big) > 2.$$

It follows from (2.60) and (2.4) that

$$D_{h,k} = \int_{\mathbb{R}^N}\frac{1 - \cos\langle he_j, \lambda\rangle - \cos\langle ke_j, \lambda\rangle + \cos\langle (h-k)e_j, \lambda\rangle}{hk}\, f(\lambda)\, d\lambda. \tag{2.61}$$

To prove the sufficiency of (2.15), we note that for each fixed $\lambda \in \mathbb{R}^N$,

$$\lim_{h,k\to 0} \frac{1-\cos(h\lambda_j)-\cos(k\lambda_j)+\cos((h-k)\lambda_j)}{hk} = \lambda_j^2 \tag{2.62}$$

and by the mean value theorem,

$$\left|\frac{1-\cos(h\lambda_j)-\cos(k\lambda_j)+\cos((h-k)\lambda_j)}{hk}\right| \le \lambda_j^2.$$

Now we assume (2.15) holds. Then, as in the proof of Proposition 2.2.1, we have

$$\int_{\lambda\in\mathbb{R}^N:|\lambda_j|>1} \frac{\lambda_j^2\, d\lambda}{\left(\sum_{i=1}^N |\lambda_i|^{\beta_i}\right)^{\gamma}} \le c\int_1^\infty \frac{\lambda_j^2\, d\lambda_j}{\lambda_j^{\beta_j(\gamma-\sum_{i\neq j}\frac{1}{\beta_i})}} < \infty.$$

This implies $\int_{\mathbb{R}^N} \lambda_j^2 f(\lambda)d\lambda < \infty$. By (2.61), (2.62) and the dominated convergence theorem, we obtain

$$\lim_{h,k\to 0} D_{h,k} = \int_{\mathbb{R}^N} \lambda_j^2 f(\lambda)d\lambda.$$

To prove the necessity of (2.15), we assume $\beta_j\big(\gamma - \sum_{i=1}^N \frac{1}{\beta_i}\big) \le 2$. Then, as in the proof of Proposition 2.2.1, we have

$$\int_{\lambda\in\mathbb{R}^N:|\lambda_j|>1} \frac{\lambda_j^2\, d\lambda}{\left(\sum_{i=1}^N |\lambda_i|^{\beta_i}\right)^{\gamma}} = \infty. \tag{2.63}$$

We let $h = k \downarrow 0$ and use Fatou's lemma to (2.61) [note the integrand is non-negative] to derive

$$\liminf_{h=k\downarrow 0} D_{h,k} \ge \int_{\mathbb{R}^N} \lambda_j^2 f(\lambda)d\lambda = \infty,$$

where the last equality follows from (2.63). So $\lim_{h,k\to 0} D_{h,k}$ does not exist and the proof is finished. $\square$

Proof of Theorem 2.4.6

If $v(t)$ has continuous second-order partial derivatives, then Theorem 2.4.2 implies that X has mean square partial derivatives in all N directions. Let $\nabla_X(t) = \big(X_1'(t), \cdots, X_N'(t)\big)^T$ and we show that it satisfies (2.20).

For any unit vector u in $\mathbb{R}^N$, we can write it as $u = \sum_{j=1}^N u_j e_j$ and $\sum_{j=1}^N u_j^2 = 1$. So $u^T \nabla_X(t) = \sum_{j=1}^N u_j X_j'(t)$. Hence

$$\begin{aligned}
&\mathbb{E}\left(\frac{X(t+hu)-X(t)}{h} - u^T \nabla_X(t)\right)^2 \\
&= \mathbb{E}\left(\frac{X(t+hu)-X(t)}{h} - \sum_{j=1}^N u_j X_j'(t)\right)^2 \\
&= \frac{1}{h^2} v(hu) + \mathbb{E}\left(\sum_{j=1}^N u_j X_j'(t)\right)^2 - \frac{2}{h}\sum_{j=1}^N u_j \Big(\mathbb{E}X(t+hu)X_j'(t) - \mathbb{E}X(t)X_j'(t)\Big) \\
&= \frac{1}{h^2} v(hu) + \mathbb{E}\left(\sum_{j=1}^N u_j X_j'(t)\right)^2 - \frac{1}{h}\sum_{j=1}^N u_j v_j'(hu).
\end{aligned} \tag{2.64}$$

The last equality in (2.64) follows from (2.18).

Since $v(t)$ is an even function with $v(0) = 0$ and has continuous second-order partial and mixed partial derivatives, Taylor's theorem implies

$$\lim_{h\to 0} \frac{1}{h^2} v(hu) = \lim_{h\to 0} \frac{v(hu) + v(-hu) - 2v(0)}{2h^2} = \frac{1}{2} u^T \Omega(0) u, \tag{2.65}$$

where $\Omega(0)$ is an $N \times N$ matrix, with $\big(\Omega(0)\big)_{ij} = v_{ij}''(0)$ for $i \neq j$, and $\big(\Omega(0)\big)_{ii} = v_i''(0)$.

For the second term in the last line of (2.64), note that for any i, $j = 1, \cdots, N$ and $i \neq j$, and any $l > 0$, $m > 0$,

$$
\begin{aligned}
&\mathbb{E}\left(\frac{X(t+le_i)-X(t)}{l}\frac{X(t+me_j)-X(t)}{m}\right) \\
=&\frac{1}{lm}\mathbb{E}\Big(X(t+le_i)X(t+me_j)-X(t)X(t+me_j)-X(t+le_i)X(t)+X^2(t)\Big) \\
=&\frac{1}{2lm}\Big(v(le_i)+v(-me_j)-v(le_i-me_j)\Big).
\end{aligned}
\tag{2.66}
$$

Let $l \to 0$, $m \to 0$, then the last term in (2.66) goes to $\frac{1}{2}v''_{ij}(0)$, where $v''_{ij}(0)$ is the second-order mixed partial derivative of v at 0 in the i-th and j-th directions. By Theorem 2.4.4, we have $\mathbb{E}\big(X'_j(t)\big)^2 = \frac{1}{2}v''_j(0)$, for $j = 1, \cdots, N$. Hence

$$
\mathbb{E}\bigg(\sum_{j=1}^N u_j X'_j(t)\bigg)^2 = \frac{1}{2}u^T\Omega(0)u.
$$

Finally for the last term in (2.64), we use Taylor's theorem again to derive

$$
\lim_{h\to 0}\frac{1}{h}\sum_{j=1}^N u_j v'_j(hu) = u^T\Omega(0)u.
$$

Combining this with (2.65) and (2.66) shows that (2.64) goes to 0, as $h \to 0$. This completes the proof. □

Proof of Theorem 2.4.8

Under (2.15), Corollary 2.4.3 ensures that the mean square partial derivative $X'_j(t)$ exists. In order to show that $X'_j(t)$ has a continuous version, by Kolmogorov's continuity theorem or general Gaussian theory [cf. Adler (1981), Adler and Taylor (2007)], it is enough to show there exist constants $c_{20} > 0$ and $\eta > 0$ such that

$$\mathbb{E}\big[X_j'(s) - X_j'(t)\big]^2 \le c_{20}\,|s-t|^{\eta}, \quad \forall\, s, t \in [-M, M]^N. \tag{2.67}$$

Recall that

$$\begin{aligned} K(s,t) &= \int_{\mathbb{R}^N} \big(e^{i\langle s,\lambda\rangle} - 1\big)\big(e^{-i\langle t,\lambda\rangle} - 1\big) f(\lambda) d\lambda \\ &= \int_{\mathbb{R}^N} \Big[\cos\langle s-t,\lambda\rangle - \cos\langle t,\lambda\rangle - \cos\langle s,\lambda\rangle + 1\Big] f(\lambda) d\lambda. \end{aligned}$$

Thanks to (2.15), we derive

$$\frac{\partial K(s,t)}{\partial s_j} = \int_{\mathbb{R}^N} \Big[-\lambda_j \sin\langle s-t,\lambda\rangle + \lambda_j \sin\langle s,\lambda\rangle\Big] f(\lambda)\, d\lambda$$

and

$$\frac{\partial K(s,t)}{\partial s_j \partial t_j} = \int_{\mathbb{R}^N} \lambda_j^2 \cos\langle s-t,\lambda\rangle f(\lambda)\, d\lambda.$$

So

$$\begin{aligned} \mathbb{E}\big(X_j'(s) - X_j'(t)\big)^2 &= \mathbb{E}\big(X_j'(s)\big)^2 + \mathbb{E}\big(X_j'(t)\big)^2 - 2\mathbb{E}\big(X_j'(s)X_j'(t)\big) \\ &= 2\int_{\mathbb{R}^N} \lambda_j^2 \big(1 - \cos\langle s-t,\lambda\rangle\big) f(\lambda) d\lambda. \end{aligned}$$

The rest of the proof is similar to that of Lemma 2.3.2. Denote $\hat{s}_0 = t$, $\hat{s}_1 = (s_1, t_2, \cdots, t_N)$, $\hat{s}_2 = (s_1, s_2, t_3, \cdots, t_N)$, $\cdots$, $\hat{s}_{N-1} = (s_1, \cdots, s_{N-1}, t_N)$ and $\hat{s}_N = s$. Then

$$\mathbb{E}\big(X_j'(s) - X_j'(t)\big)^2 \le N \sum_{k=1}^{N} \mathbb{E}\big(X_j'(\hat{s}_k) - X_j'(\hat{s}_{k-1})\big)^2$$

$$
\begin{aligned}
&= 2N \sum_{k=1}^{N} \Bigg\{ \int_{|\lambda|\le 1} \lambda_j^2 \big(1-\cos(s_k-t_k)\lambda_k\big) f(\lambda) d\lambda \\
&\qquad\qquad + \int_{|\lambda|>1} \lambda_j^2 \big(1-\cos(s_k-t_k)\lambda_k\big) f(\lambda) d\lambda \Bigg\} \qquad (2.68)\\
&\le c_{21}\, |s-t|^2 + c_{22} \sum_{k=1}^{N} \int_{|\lambda|>1} \frac{\big(1-\cos(s_k-t_k)\lambda_k\big)\lambda_j^2}{\big(\sum_{i=1}^N |\lambda_i|^{H_i}\big)^{Q+2}}\, d\lambda.
\end{aligned}
$$

Now we estimate the last N integrals in (2.68). For simplicity of notation, we only consider the case when $k = j$ [the cases of $k \ne j$ are similar]. Denote $h_k = s_k - t_k$ and, similar to (2.44), (2.45) and (2.46), we derive

$$
\begin{aligned}
&\int_{|\lambda|>1} \frac{\big(1-\cos(s_k-t_k)\lambda_k\big)\,\lambda_k^2}{\big(\sum_{i=1}^N |\lambda_i|^{H_i}\big)^{Q+2}}\, d\lambda \\
&\le 2 \int_{\frac{1}{\sqrt{N}}}^{\infty} \big(1-\cos(h_k\lambda_k)\big)\lambda_k^2\, d\lambda_k \int_{\mathbb{R}^{N-1}} \frac{d\lambda_1\cdots d\lambda_{k-1} d\lambda_{k+1}\cdots d\lambda_N}{\big(\sum_{i=1}^N |\lambda_i|^{H_i}\big)^{Q+2}} \\
&\quad + 2 \int_0^1 \big(1-\cos(h_k\lambda_k)\big)\lambda_k^2\, d\lambda_k \int_{\frac{1}{\sqrt{N}}}^{\infty} d\lambda_{j_0} \int_{\mathbb{R}^{N-2}} \frac{d\lambda^{\vee}_{k,j_0}}{\big(\sum_{i=1}^N |\lambda_i|^{H_i}\big)^{Q+2}} \\
&\le c \int_{\frac{1}{\sqrt{N}}}^{\infty} \frac{\big(1-\cos(h_k\lambda_k)\big)\lambda_k^2}{\lambda_k^{2H_k+1}}\, d\lambda_k + c\int_0^1 \lambda_k^2 \big(1-\cos(h_k\lambda_k)\big) d\lambda_k \\
&\le c_{23}\Big(|h_k|^{2(H_k-1)} \log\frac{1}{|h_k|} + |h_k|^2\Big),
\end{aligned}
$$

thanks to $H_k > 1$. Combining this with (2.68) proves (2.67).

It follows from (2.67)) that the Gaussian field $X_j' = \{X_j'(t), t \in \mathbb{R}^N\}$ has a continuous version [which will still be denoted by X_j']. Now we define a new Gaussian random field $\widetilde{X} = \{\widetilde{X}(t), t \in \mathbb{R}^N\}$ by

$$
\widetilde{X}(t) = X(t_1, \cdots, t_{j-1}, 0, t_{j+1}, \cdots, t_N) + \int_0^{t_j} X_j'(t_1, \cdots, t_{j-1}, s_j, t_{j+1}, \cdots, t_N)\, ds_j. \quad (2.69)
$$

Then we can verify that $\widetilde{X}$ is a continuous version of X and, for every $t \in \mathbb{R}^N$, $\widetilde{X}_j'(t) = X_j'(t)$

almost surely. This amounts to verify that for every $t \in \mathbb{R}^N$,

$$\mathbb{E}\big(\widetilde{X}(t)^2\big) = v(t) \quad \text{and} \quad \mathbb{E}\Big[\big(\widetilde{X}(t) - X(t)\big)^2\Big] = 0,$$

which can be proved by using (2.69), Theorem 2.4.4 and (2.18). Since the verification is elementary, we omit the details. This proves Part (i) of Theorem 2.4.8.

It remains to prove Part (ii) of Theorem 2.4.8. By applying Part (i) to $j = 1$, we obtain a continuous version $\widetilde{X^{(1)}}$ of X such that $\frac{\partial \widetilde{X^{(1)}}}{\partial t_1}(t)$ is continuous. Then we apply Part (i) to $\widetilde{X^{(1)}}$ with $j = 2$ and obtain a version $\widetilde{X^{(2)}}$ of $\widetilde{X^{(1)}}$ defined by

$$\widetilde{X^{(2)}}(t) = \widetilde{X^{(1)}}(t_1, 0, t_3, \cdots, t_N) + \int_0^{t_j} \widetilde{X^{(1)}}'_2(t_1, s_2, t_3, \cdots, t_N)\, ds_2. \tag{2.70}$$

Then $\frac{\partial \widetilde{X^{(2)}}}{\partial t_1}(t)$ and $\frac{\partial \widetilde{X^{(2)}}}{\partial t_2}(t)$ are almost surely continuous. Repeating this "updating" procedure for $j = 3, \cdots, N$, we obtain a continuous version $\widetilde{X^{(N)}}$ of X such that all first-order partial derivatives of $\widetilde{X^{(N)}}$ are continuous almost surely. Hence the sample function of $\widetilde{X^{(N)}}$ is almost surely differentiable in the sense of (2.23). The proof of Theorem 2.4.8 is complete. □

Proof of Proposition 2.6.1

By stationarity, we'll have

$$\begin{aligned}
&\mathbb{E}\big(X(x,t) - X(y,s)\big)^2 \\
&= \mathbb{E}\big(X(x,t)\big)^2 + \mathbb{E}\big(X(y,s)\big)^2 - 2\mathbb{E}\big(X(x,t)X(y,s)\big) \\
&= 2K(0,0) - 2K(x-y, t-s),
\end{aligned}$$

as well as

$$
\begin{aligned}
&2K(0,0) - 2K(x,t) \\
&= 2\sigma^2 - \frac{2\sigma^2}{(1+a|t|^{2\alpha})^{\beta d/2}} \exp\Big(-\frac{c|x|^{2\gamma}}{(1+a|t|^{2\alpha})^{\beta\gamma}}\Big) \\
&= 2\sigma^2 \frac{(1+a|t|^{2\alpha})^{\beta d/2} - \exp\Big(-\frac{c|x|^{2\gamma}}{(1+a|t|^{2\alpha})^{\beta\gamma}}\Big)}{(1+a|t|^{2\alpha})^{\beta d/2}}.
\end{aligned}
\tag{2.71}
$$

By using Taylor expansion, we can write (2.71) as

$$
\begin{aligned}
&2\sigma^2 \frac{1 + \frac{\beta d}{2} a|t|^{2\alpha} + o(|t|^{2\alpha}) - 1 + \frac{c|x|^{2\gamma}}{(1+a|t|^{2\alpha})^{\beta\gamma}} - o\bigg(\frac{c|x|^{2\gamma}}{(1+a|t|^{2\alpha})^{\beta\gamma}}\bigg)}{(1+a|t|^{2\alpha})^{\beta d/2}} \\
&= 2\sigma^2 \frac{\frac{\beta d}{2} a|t|^{2\alpha} + \frac{c|x|^{2\gamma}}{(1+a|t|^{2\alpha})^{\beta\gamma}} + o(|t|^{2\alpha}) - o\bigg(\frac{c|x|^{2\gamma}}{(1+a|t|^{2\alpha})^{\beta\gamma}}\bigg)}{(1+a|t|^{2\alpha})^{\beta d/2}}.
\end{aligned}
$$

Hence we can find positive constants $c_{24} \le c_{25}$ such that

$$
c_{24}\big(|x|^{2\gamma} + |t|^{2\alpha}\big) \le 2K(0,0) - 2K(x,t) \le c_{25}\big(|x|^{2\gamma} + |t|^{2\alpha}\big) \tag{2.72}
$$

for all $(x,t) \in \mathbb{R}^{d+1}$ with $|x|$ and $|t|$ small. Replace x and t in (2.72) by $x-y$ and $t-s$ respectively; (2.33) follows.

To prove (2.34), we make use of the fact that for any Gaussian random vector (U,V) with mean 0,

$$
\mathrm{Var}(U|V) = \frac{\big(\rho_{U,V}^2 - (\sigma_U - \sigma_V)^2\big)\big((\sigma_U + \sigma_V)^2 - \rho_{U,V}^2\big)}{4\sigma_V^2},
$$

where $\rho_{U,V}^2 = \mathbb{E}\big[(U-V)^2\big]$, $\sigma_U^2 = \mathbb{E}(U^2)$ and $\sigma_V^2 = \mathbb{E}(V^2)$. Let $U = X(x,t)$ and $V = X(y,s)$, we derive

$$
\begin{aligned}
\mathrm{Var}\big(X(x,t)|X(y,s)\big) &= \frac{\big[K(0,0)-K(x-y,t-s)\big]\big[K(0,0)+K(x-y,t-s)\big]}{K(0,0)} \\
&\geq c_{24}\big(|x-y|^{2\gamma}+|t-s|^{2\alpha}\big).
\end{aligned}
$$

This proves (2.34). □

Proof of Proposition 2.6.2

Eq. (2.35) follows from Proposition 2.6.1 and Theorem 2.5.1. Then let's prove (2.36), where $0 < \alpha \leq \gamma \leq 1$. By Proposition 2.6.1 and Theorem 2.5.1, we get

$$
\dim\mathrm{Gr}\mathbf{X}([0,1]^{d+1}) = \min_{1\leq k\leq d+1}\left\{\sum_{j=1}^{k}\frac{\overline{H}_k}{\overline{H}_j}+d+1-k+(1-\overline{H}_k)p;\ \sum_{j=1}^{d+1}\frac{1}{\overline{H}_j}\right\},
$$

where $\overline{H}_1=\alpha$, $\overline{H}_2=\cdots=\overline{H}_{d+1}=\gamma$. Denote

$$
S(k)=\sum_{j=1}^{k}\frac{\overline{H}_k}{\overline{H}_j}+d+1-k+(1-\overline{H}_k)p.
$$

We have $S(1)=d+1+(1-\alpha)p$, $S(k)=d+\frac{\gamma}{\alpha}+(1-\gamma)p \triangleq S$, for $2\leq k\leq d+1$. Also $\sum_{j=1}^{d+1}\frac{1}{\overline{H}_j}=\frac{1}{\alpha}+\frac{d}{\gamma}$. We can verify directly that if $p<\frac{1}{\alpha}$, then

$$
S(1) < S < \sum_{j=1}^{d+1}\frac{1}{\overline{H}_j},
$$

which yields $\dim\mathrm{Gr}\mathbf{X}([0,1]^{d+1})=S(1)$. The verifications for the cases $\frac{1}{\alpha}\leq d<\frac{1}{\alpha}+d$ and $p\geq\frac{1}{\alpha}+d$ are similar. We omit the details. □

Proof of Proposition 2.6.4

By Proposition 2.6.1 and Theorem 2.5.2 we find that when $\frac{1}{\alpha} + \frac{d}{\gamma} < p$, for every $x \in \mathbb{R}^p$, $\mathbf{X}^{-1}(x) = \emptyset$ a.s. Also, when $\frac{1}{\alpha} + \frac{d}{\gamma} > p$, for any $x \in \mathbb{R}^p$, with positive probability

$$\dim\big(\mathbf{X}^{-1}(x)\big) = \min_{1 \le k \le d+1} \left\{ \sum_{j=1}^{k} \frac{\overline{H}_k}{\overline{H}_j} + d + 1 - k - \overline{H}_k\, p \right\}.$$

If $0 < \alpha \le \gamma < 1$, we have $\overline{H}_1 = \alpha$, $\overline{H}_2 = \cdots = \overline{H}_{d+1} = \gamma$. Denote

$$T(k) = \sum_{j=1}^{k} \frac{\overline{H}_k}{\overline{H}_j} + d + 1 - k - \overline{H}_k\, p,$$

then $T(1) = d + 1 - \alpha p$, $T(k) = d + \frac{\gamma}{\alpha} - \gamma p \triangleq T$, for $2 \le k \le d+1$. Since $T(1) < T$, if and only if $p < \frac{1}{\alpha}$, (2.38) follows.

If $0 < \gamma \le \alpha < 1$, then $\overline{H}_1 = \cdots = \overline{H}_d = \gamma$ and $\overline{H}_{d+1} = \alpha$. It follows that $T(d+1) = \frac{d\alpha}{\gamma} + 1 - \alpha p$ and $T(k) = d + 1 - \gamma p \triangleq \tilde{T}$, for $1 \le k \le d$. Since $\tilde{T} < T(d+1)$, if and only if $p < \frac{d}{\gamma}$, we obtain (2.39). The proof is complete. $\square$

Chapter 3

Criteria for equivalence and asymptotically optimal predictions

3.1 Introduction

Optimal linear prediction has been widely used in spatial statistics and geostatistics, where it is known as kriging. In kriging, to guarantee good linear predictors based on an estimated Gaussian probability measure, it is of great value to be able to distinguish between two orthogonal probability measures and to determine when one can tell which measure is correct and which is not. Many authors have created various criteria for the equivalence and orthogonality of two Gaussian measures corresponding to one-dimensional Gaussian processes or Gaussian random fields. The references include Gihman and Skorohod (1974), Ibragimov and Rozanov (1978), Parzen (1963), Chatterji and Mandrekar (1978), Kallianpur and Oodaira (1963), Yadrenko (1983), Stein (1999b) and so on. In fact, Parzen (1963) developed an approach for equivalence of two Gaussian measures by using two concepts: the notion of probability spectral density function and the notion of a reproducing kernel Hilbert space of a time series. Chatterji and Mandrekar (1978) also used the method of RKHS to find sufficient and necessary conditions for the equivalence of two Gaussian measures in a general setting. It is worth noting that the approach which uses RKHS has no constrains like stationarity or isotropy on the underlying process, and the results are applicable to random fields. Ibragimov and Rozanov (1978) obtained the conditions for equivalence of two Gaussian measures involving the entropy of distributions, and developed the conditions for stationary processes by associating a Hilbert space spanned by analytic functions. Moreover, given two equivalent Gaussian processes, Kallianpur and Oodaira (1973) defined the notion of a non-anticipative representation of one of the processes with respect to the other, for one dimensional case. Later, Yadrenko (1983) extended Ibragimov and Rozanov's results to stationary and isotropic random fields. Du (2009) reviewed of the basic results for the equivalence and orthogonality of two Gaussian measures, and provided a detailed re-proof of Theorem 4 in Yadrenko (1983), page 156, under the setting of stationary and isotropic random fields. However, in the literature, there are few explicit results available for the equivalence of two Gaussian measures in a non-stationary random field, especially for anisotropic cases.

In this chapter, we extend Ibragimov and Rozanov's method to study intrinsically stationary random fields. We determine the relationships among three corresponding Hilbert spaces: the random variable space generated by the random field, the reproducing kernel Hilbert space corresponding to the covariance kernel and the complex function space spanned by the analytic functions of the form $\lambda \mapsto e^{i\langle \lambda, t\rangle} - 1$, $t \in D$. Criteria for equivalence and orthogonality of intrinsically stationary Gaussian random fields are given in terms of their probability spectral density functions and the structures of their reproducing kernel Hilbert spaces. The results we have obtained are different from those for stationary processes [see Ibragimov and Rozanov (1978)]. Moreover, given the equivalence of two random fields, we obtain a representation of one of the random fields with respect to the other. The advantage is that we can use the equivalent representation instead of the original one whenever the representation is simpler with respect to some prediction questions.

As we know, in practice, the true probability distribution of our Gaussian model is always unknown and must be estimated from the gathered data. To this end, it is of great value to investigate the effect of using a fixed but incorrect probability distribution, especially, when more sample data can be obtained by sampling the spatial or temporal domain increasingly densely (fix-domain asymptotics). Actually, the asymptotic optimality of linear predictions of intrinsically stationary Gaussian models and the convergence rates are established in this chapter. Moreover, the asymptotic efficient prediction of non-stationary, anisotropic space-time models with a misspecified probability distribution is studied. The main results show that under the equivalence of two Gaussian measures, the prediction based on the incorrect distribution is asymptotically optimal and efficient relative to the prediction under the correct distribution, as the points of observations become increasingly dense in the study domain. Our results extend those of Stein (1988, 1990, 1999a, 1999b) which were concerned with isotropic and stationary Gaussian random fields.

The rest of this chapter is organized as follows. Section 2 studies the relationships among the three Hilbert spaces we have constructed. In Section 3 we obtain criteria for equivalence and orthogonality of two Gaussian measures in the intrinsically stationary random fields. We study the asymptotic optimality of linear predictions in Section 4 and the convergence rates of the predictors are established in Section 5. In Section 6, we show the proofs of the main results in this chapter.

In the spatial statistics contexts, one would feel more comfortable using the space model as $X = \{X(t), t \in \mathbb{R}^d\}$ and space-time model as $X = \{X(x,t), (x,t) \in \mathbb{R}^d \times \mathbb{R}\}$, where d

denotes the dimension for the space variable.

In this chapter, we study the asymptotic and prediction properties of the random field $X = \{X(t), t \in \mathbb{R}^d\}$.

3.2 Three corresponding Hilbert spaces and equivalence

Let $X = \{X(t), t \in \mathbb{R}^d\}$ be a real-valued, centered intrinsically stationary Gaussian random field (i.e. Gaussian random field with stationary increments) with $X(0) = 0$. We assume that X has continuous covariance function $K(s,t) = \mathbb{E}[X(s)X(t)]$. As in Chapter 2, $K(s,t)$ can be represented as

$$K(s,t) = \int_{\mathbb{R}^d} \big(e^{i\langle s,\lambda\rangle} - 1\big)\big(e^{-i\langle t,\lambda\rangle} - 1\big) F(d\lambda), \tag{3.1}$$

where $F(d\lambda)$ is a nonnegative symmetric measure on $\mathbb{R}^d \setminus \{0\}$ satisfying

$$\int_{\mathbb{R}^d} \frac{|\lambda|^2}{1+|\lambda|^2} F(d\lambda) < \infty. \tag{3.2}$$

Moreover, X has the following stochastic integral representation:

$$X(t) = \int_{\mathbb{R}^d} \big(e^{i\langle t,\lambda\rangle} - 1\big) \Phi(d\lambda), \tag{3.3}$$

where $\Phi(d\lambda)$ is a centered complex-valued Gaussian random measure which satisfies

$$\mathbb{E}\Big(\Phi(A)\overline{\Phi(B)}\Big) = F(A \cap B) \quad \text{and} \quad \Phi(-A) = \overline{\Phi(A)}$$

for all Borel sets $A, B \subseteq \mathbb{R}^d$ with finite F-measure.

Let D be a bounded region in $\mathbb{R}^d$. Without loss of generality, we assume $0 \in D$. Let L_D^0 be the linear hull of the complex exponential functions $\lambda \mapsto e^{i\langle \lambda, t\rangle} - 1$, $t \in D$ and take $L_F(D)$ to be the closure of L_D^0 under the inner product

$$\langle \varphi_1, \varphi_2 \rangle_F \triangleq \langle \varphi_1, \varphi_2 \rangle_{L_F(D)} = \int_{\mathbb{R}^d} \varphi_1(\lambda) \overline{\varphi_2(\lambda)} F(d\lambda),$$

where $\varphi_1, \varphi_2 \in L_D^0$. Let $H_F(D)$ be the closed linear hull of the random variables $X(t)$, $t \in D$ with respect to the inner product

$$\langle X(s), X(t) \rangle_{H_F(D)} = K(s,t) = \int_{\mathbb{R}^d} \big(e^{i\langle \lambda, s\rangle} - 1\big)\big(e^{-i\langle \lambda, t\rangle} - 1\big) F(d\lambda).$$

On $H_F(D)$, there exist mean and covariance operators, which we also call m and K, such that for $\eta_1, \eta_2 \in H_F(D)$, $\mathbb{E}(\eta_1) = m(\eta_1)$ and $\mathrm{Cov}(\eta_1, \eta_2) = K(\eta_1, \eta_2)$. We will freely switch between the functions m and K and the operators m and K in the rest of this chapter, the meaning being apparent from context.

We denote by $R_K(D)$ the reproducing kernel Hilbert space (RKHS, for short) of the random field $X(t)$ with reproducing kernel $K(s,t)$, $s, t \in D$. That is, for every real function $f \in R_K(D)$, we have

$$\langle f, K(\cdot, t) \rangle_{R_K(D)} = f(t), \quad \forall\, t \in D.$$

In fact, the Hilbert space $R_K(D)$ is the closure of the subspace spanned by real functions $K(\cdot, t)$, $t \in D$, with respect to the inner product $\langle \cdot, \cdot \rangle_{R_K(D)}$. Note that $R_K(D)$ is separable because of the continuity of $K(s,t)$.

Define a mapping ρ from $H_F(D)$ onto the RKHS $R_K(D)$ such that, for every $t \in D$,

$$\rho\left(X(t)\right) = K(\cdot, t). \tag{3.4}$$

It can be proved that ρ is a linear, isometric, one to one mapping. First, we obtain the following lemma, which gives a representation for random variables in $H_F(D)$ with respect to the analytic functions in $L_F(D)$.

Lemma 3.2.1. *Each random variable $\eta \in H_F(D)$ can be represented as*

$$\eta(\varphi) = \int_{\mathbb{R}^d} \varphi(\lambda)\Phi(d\lambda) \tag{3.5}$$

for some $\varphi(\lambda) \in L_F(D)$. For every function $\varphi(\lambda) \in L_F(D)$, (3.5) is well defined and $\eta \in H_F(D)$. The mapping $\phi : \eta \mapsto \varphi$ is linear, isometric and one to one.

Regarding (3.4), for every $\eta \in H_F(D)$ and $t \in D$,

$$\rho(\eta)(t) = \mathbb{E}\left(\eta X(t)\right). \tag{3.6}$$

The function $t \mapsto \rho(\eta)(t)$ in (3.6) belongs to $R_K(D)$.

Since $X(t) = \int_{\mathbb{R}^d} \big(e^{i\langle\lambda,t\rangle} - 1\big)\Phi(d\lambda)$, and $\eta = \eta(\varphi) = \int_{\mathbb{R}^d} \varphi(\lambda)\Phi(d\lambda)$ by Lemma 3.2.1, (3.6) can be rewritten as

$$\rho(\eta)(t) = \int_{\mathbb{R}^d} \varphi(\lambda)\big(e^{-i\langle\lambda,t\rangle} - 1\big)F(d\lambda).$$

Hence, there is a linear, isometric, one to one mapping θ from the Hilbert space $L_F(D)$ onto the RKHS $R_K(D)$ such that, for every $\varphi \in L_F(D)$,

$$\theta(\varphi)(\cdot) = \int_{\mathbb{R}^d} \varphi(\lambda)\big(e^{-i\langle \lambda, \cdot \rangle} - 1\big) F(d\lambda). \tag{3.7}$$

Some remarks about the intrinsically stationary random field follow.

Remark 3.2.2

(1) Not like the case of stationary random fields, the covariance function $K(s,t)$ for an intrinsically stationary random field can not be represented as the Fourier transform of the spectral measure F.

(2) The Hilbert space $L_F(D)$ for the intrinsically stationary random field is different from that for stationary case; the latter contains real constants as members.

Let P_0 be the Gaussian probability measure on the σ-algebra $\mathcal{U}(D)$ generated by $X(t)$ for $t \in D$, with the second-order structure $(0, K_0)$ and the spectral measure $F_0(d\lambda)$. Let P_1 be the Gaussian probability measure on $\mathcal{U}(D)$ for a random field $X_1(t)$, with the second-order structure $(0, K_1)$, which has the form

$$X_1(t) = X(t) - m_1(t),$$

where $m_1(t) = \mathbb{E}_1(X(t))$. Denote the spectral measure of $X_1(t)$ as $F_1(d\lambda)$, then we have the following lemma, which is similar to (1.27) of Ibragimov and Rozanov (1978), page 70.

Lemma 3.2.3. *Suppose the Gaussian measures P_0 and P_1 are equivalent on the σ-algebra*

$\mathcal{U}(D)$, *then*[1]

$$\|\varphi\|_{F_0} \asymp \|\varphi\|_{F_1}, \qquad \varphi \in L_D^0. \tag{3.8}$$

Lemma 3.2.3 implies that when the Gaussian measures P_0 and P_1 are equivalent on $\mathcal{U}(D)$, we have $L_{F_0}(D) = L_{F_1}(D)$, $H_{F_0}(D) = H_{F_1}(D)$ and $R_{K_0}(D) = R_{K_1}(D)$.

3.3 Some conditions for equivalence of two Gaussian measures

Many authors have created various criteria for the equivalence and orthogonality of two Gaussian measures. Parzen (1963) studied the equivalence and orthogonality of two Gaussian measures using the tools of RKHS $R_K(D)$ and determined the corresponding Randon-Nikodym derivative under two different cases: sure signal case (with two Gaussian measures having the same covariance function) and stochastic signal case (with two Gaussian measures having the same mean function), respectively. Chatterji and Mandrekar (1978) also used the method of RKHS to find sufficient and necessary conditions for the equivalence of two Gaussian measures in a general setting. It is worth noting that the approach which uses RKHS has no constraints like stationarity or isotropy on the underlying process, and the results are applicable to random fields. Kallianpur and Oodaira (1963) gave necessary and sufficient conditions for equivalence of two Gaussian measures by defining an operator between the corresponding reproducing kernels (covariance functions), and obtained a non-anticipative representation of one Gaussian process by another. Sottinen and Tudor (2006) applied Kallianpur and Oodaira (1963)'s idea to investigate the equivalence in law of mul-

[1] $\|\varphi\|_{F_0} \asymp \|\varphi\|_{F_1}$ means $0 < c_1 \le \|\varphi\|_{F_0}/\|\varphi\|_{F_1} \le c_2 < \infty$, where c_1 and c_2 are constants.

tiparameter Gaussian processes, i.e. Gaussian random fields, with a Brownian sheet and a fractional Brownian sheet. They surveyed multiparameter analogous of Hitsuda, Girsonov and Shepp representations. On the other hand, Ibragimov and Rozanov (1978) investigated equivalence of stationary processes by using analytic tools, namely, the Hilbert space $L_F(D)$. In this section, we apply Ibragimov and Rozanov's method to study the criteria for the equivalence of two Gaussian measures in an intrinsically stationary random field and compare our results with the existing ones.

Let P_0 and P_1 be Gaussian measures on the σ-algebra $\mathcal{U}(D)$ generated by all random variables of the intrinsically stationary random field $X(t)$, $t \in D$. The second-order structures of P_0 and P_1 are $(0, K_0)$ and (m_1, K_1), respectively. It is known that Gaussian measures have the following property. See page 77 and 78 of Ibragimov and Rozanov (1978) for more details and proofs.

Lemma 3.3.1. *The Gaussian measures P_0 and P_1 are equivalent if and only if there exists a Gaussian measure P such that the pairs P_0 and P, and P_1 and P are equivalent; for the equivalent measures P_0 and P_1, the density $P_1(d\omega)/P_0(d\omega)$ is such that*

$$\frac{P_1(d\omega)}{P_0(d\omega)} = \frac{P_1(d\omega)}{P(d\omega)} \frac{P(d\omega)}{P_0(d\omega)}.$$

Based on Lemma 3.3.1, in this section, we may consider two cases. In case one, the covariance function K_0 coincides with K_1, such that $\langle \varphi_1, \varphi_2 \rangle_{F_0} = \langle \varphi_1, \varphi_2 \rangle_{F_1}$, for all $\varphi_1, \varphi_2 \in L_D^0$. From Lemma 1 of Bonami and Estrade (2003), we know that $K_0 = K_1$ implies $F_0 = F_1$. We write $F_0 = F_1 = F$ in this case. In case two, the mean function $m_1(t) \equiv 0$, but the covariance functions K_0 and K_1 are different. We obtain necessary and sufficient conditions for P_0 and P_1 to be equivalent under the two cases, respectively. Let us first consider the case where the two measures differ only in the mean functions.

3.3.1 Case I: Same covariance function

By Lemma 3.2.1, each random variable $\eta \in H_F(D)$ can be expressed as $\eta(\varphi) = \int_{\mathbb{R}^d} \varphi(\lambda)\Phi(d\lambda)$, for some $\varphi(\lambda) \in L_F(D)$. Denote $m_1(\varphi)$ as the mean of $\eta(\varphi)$ under the second-order structure (m_1, K_1). We will then have the following extension of Theorem 3 of Ibragimov and Rozanov (1978), page 78, where the case of stationary processes is considered.

Theorem 3.3.2. *Suppose $K_0 = K_1$, the Gaussian measures P_0 and P_1 are equivalent on $\mathfrak{U}(D)$ if and only if, the mean value $m_1(\varphi)$ is a linear continuous functional on the Hilbert space $L_F(D)$:*

$$m_1(\varphi) = \langle \varphi, \psi \rangle_F, \qquad \varphi \in L_D^0, \tag{3.9}$$

for some $\psi(\lambda) \in L_F(D)$.

As a consequence, we obtain a more explicit necessary and sufficient condition for the equivalence of two measures which differ only in the mean functions.

Theorem 3.3.3. *Suppose $K_0 = K_1$, the Gaussian measures P_0 and P_1 are equivalent on $\mathfrak{U}(D)$ if and only if the mean function $m_1(t)$, $t \in D$, permits a representation as*

$$m_1(t) = \int_{\mathbb{R}^d} \left(e^{-i\langle \lambda, t\rangle} - 1\right)\varphi(\lambda)F(d\lambda), \tag{3.10}$$

for some $\varphi(\lambda) \in L_F(D)$. And in the latter case, the Randon-Nikodym derivative $p(\omega) = P_1(d\omega)/P_0(d\omega)$ on the σ-algebra $\mathfrak{U}(D)$ can be expressed as

$$p(\omega) = \exp\left\{\int_{\mathbb{R}^d} \varphi(\lambda)\Phi(d\lambda) - \frac{1}{2}\|\varphi\|_F^2\right\}. \tag{3.11}$$

Corollary 3.3.4. *Under the conditions of Theorem 3.3.3, the Gaussian measures P_0 and P_1 are equivalent on $\mathcal{U}(D)$ if and only if, the mean function $m_1(t)$, $t \in D$, is in the RKHS $R_K(D)$.*

Proof The proof of Corollary 3.3.4 follows directly from Theorem 3.3.3 and (3.7). □

We need to mention that Corollary 3.3.4 is consistent with the results obtained by Parzen (1963) and Chatterji and Mandrekar (1978), where the tools of RKHS are applied to study the equivalence and orthogonality of two Gaussian measures. This criterion obtained by using the method of RKHS is general, has no constraints like stationarity or isotropy on the underlying process, and the result can be applied to any multi-dimensional case.

We now assume P_0 and P_1 have spectral densities f_0 and f_1. We have known that $K_0 = K_1$ implies $f_0 = f_1$.

Corollary 3.3.5. *If $K_0 = K_1$ has a bounded density function $f(\lambda)$, then P_0 and P_1 are equivalent on $\mathcal{U}(D)$ if and only if, $m_1(t)$, $t \in D$ can be extended to all $t \in \mathbb{R}^d$ and there exists a square-integrable function ψ on $\mathbb{R}^d$ such that*

$$m_1(t) = \int_{\mathbb{R}^d} \left(e^{-i\langle t,\lambda\rangle} - 1\right)\psi(\lambda)d\lambda \tag{3.12}$$

and

$$\int_{\mathbb{R}^d} \frac{|\psi(\lambda)|^2}{f(\lambda)} d\lambda < \infty. \tag{3.13}$$

Gaussian random fields whose spectral densities are described by a power law model provide a simple and flexible class of models for inferences. This class includes fractional Brownian fields as a special case. Ibragimov and Rozanov (1978) obtained necessary and sufficient conditions for the equivalence of two Gaussian measures with power law densities, under the setting of stationary processes [see Theorems 10 in Chapter III of Ibragimov and Rozanov (1978)]. Michael Stein stated in a SAMSI workshop in 2010 that Ibragimov and Rozanov (1978)'s results for stationary random fields might be extendable to certain nonstationary processes. In the following, we give a necessary condition for the equivalence under the setting of intrinsically stationary random fields. For the sufficient condition, we restrict to the one-dimensional case.

Corollary 3.3.6. *If $f(\lambda)$ is bounded and satisfies*

$$f(\lambda) \leq \frac{K}{(1+|\lambda|^2)^n}, \tag{3.14}$$

for some constants $K > 0$ and $n \geq 1$, then a necessary condition for Gaussian measures P_0 and P_1 to be equivalent on $\mathcal{U}(D)$ is that $m_1(t)$ must have partial derivatives in each variable up to the order $\left\lfloor n - \frac{d+1}{2} \right\rfloor$. Equivalently, $\forall j = 1, 2, \cdots, d$

$$\frac{\partial^k}{\partial t_j^k} m_1(t) = \int_{\mathbb{R}^d} \big(-i\lambda_j\big)^k e^{-i\langle t,\lambda\rangle} \psi(\lambda) d\lambda,$$

for all $k = 1, 2, \cdots, \left\lfloor n - \frac{d+1}{2} \right\rfloor$.

To obtain a sufficient condition for P_0 and P_1 to be equivalent, we restrict ourselves to $d = 1$.

Corollary 3.3.7. *Suppose $f(\lambda)$ is bounded and the mean function $m_1(t)$ is differentiable on $D = [0, \tau]$ and $m_1'(t)$ can be extended to be a mean-square integrable function on $\mathbb{R}$. If the Fourier transform $\psi(\lambda)$ of $m_1'(t)$ satisfies $\psi \in L^1(\mathbb{R})$ and*

$$\int_{\mathbb{R}} \frac{|\psi(\lambda)|^2}{|\lambda|^2 f(\lambda)} d\lambda < \infty,$$

then P_0 and P_1 are equivalent on the σ-algebra $\mathcal{U}([0,\tau])$.

3.3.2 Case II: Same mean function

In this subsection, we consider the case where the two Gaussian measures differ only in the covariance functions. Assume $m_1(t) \equiv 0$, for all $t \in D$. In analogy to L_D^0 defined before, let $L_{D\times D}^0$ be the linear hull of the functions $\big(e^{i\langle\lambda,s\rangle}-1\big)\big(e^{-i\langle\mu,t\rangle}-1\big)$ of $s,t \in D$ and $\lambda,\mu \in \mathbb{R}^d$. Take $L_{F\times F}(D\times D)$ to be the closure of $L_{D\times D}^0$ under the inner product

$$\langle\varphi_1,\,\varphi_2\rangle_{F\times F} = \iint_{\mathbb{R}^d\times\mathbb{R}^d} \varphi_1(\lambda,\mu)\,\overline{\varphi_2(\lambda,\mu)}F(d\lambda)F(d\mu),$$

where $\varphi_1,\,\varphi_2 \in L_{D\times D}^0$. Let $H_{F\times F}(D\times D)$ be the closed linear hull of functions $X(s)X(t)-K(s,t)$, $s,t\in D$, under the second-order structure $(0,K)$. Let us consider the linear space everywhere dense in $H_{F\times F}(D\times D)$ of all variables represented in the symmetric form

$$\eta = \sum_{k,j} c_{kj}\left[X(t_k)X(t_j) - K(t_k,t_j)\right] \tag{3.15}$$

with symmetric real coefficients $c_{kj} = c_{jk}$, $k,j = 1,2,\cdots$.

We recall the general formula for products of Gaussians [see Ibragimov and Rozanov (1978), page 16]:

$$\begin{aligned}&\mathbb{E}\big(X(t_1)X(t_2)X(t_3)X(t_4)\big)\\ =&K(t_1,t_2)K(t_3,t_4)+K(t_1,t_3)K(t_2,t_4)+K(t_1,t_4)K(t_2,t_3).\end{aligned}$$

For any variables η_1, η_2 of the given type in (3.15), such that

$$\eta_1 = \sum_{k,j} c'_{kj} \left[X(t_k)X(t_j) - K(t_k, t_j) \right]$$

and

$$\eta_2 = \sum_{k,j} c''_{kj} \left[X(t_k)X(t_j) - K(t_k, t_j) \right],$$

we derive that

$$\begin{aligned}
&\mathbb{E}(\eta_1\eta_2) \\
=&\sum_{k,j}\sum_{m,n} c'_{kj}\, c''_{mn} K(t_k, t_m)K(t_j, t_n) + \sum_{k,j}\sum_{m,n} c'_{kj}\, c''_{mn} K(t_k, t_n)K(t_j, t_m) \\
=&2\sum_{k,j}\sum_{m,n} c'_{kj}\, c''_{mn} K(t_k, t_m)K(t_j, t_n).
\end{aligned} \tag{3.16}$$

Let us define a new random measure $\Psi(d\lambda, d\mu)$ as

$$\Psi(A \times B) = \Phi(A)\overline{\Phi(B)} - F(A \cap B),$$

for all Borel sets $A, B \subseteq \mathbb{R}^d$ with finite F-measure, where $F(A \cap B) = \mathbb{E}\big(\Phi(A)\overline{\Phi(B)}\big)$. [see Section 2.2]

We can see that each variable of the type given in (3.15) can be expressed as

$$\eta(\varphi) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(\lambda, \mu)\Psi(d\lambda, d\mu), \tag{3.17}$$

where

$$\varphi(\lambda, \mu) = \sum_{k,j} c_{kj}\big(e^{i\langle \lambda, t_k\rangle} - 1\big)\big(e^{-i\langle \mu, t_j\rangle} - 1\big). \tag{3.18}$$

For more details on multiple stochastic integrals, see Major (1981), where systematic accounts on multiple integrals of Gaussian measures are given.

From (3.16), we obtain

$$\mathbb{E}(\eta_1\eta_2) = 2\iint_{\mathbb{R}^d\times\mathbb{R}^d} \varphi'(\lambda,\mu)\overline{\varphi''(\lambda,\mu)}F(d\lambda)F(d\mu) = 2\langle\varphi',\varphi''\rangle_{F\times F}, \tag{3.19}$$

where

$$\varphi'(\lambda,\mu) = \sum_{k,j} c'_{k,j}\left(e^{i\langle\lambda,t_k\rangle}-1\right)\left(e^{-i\langle\mu,t_j\rangle}-1\right)$$

and

$$\varphi''(\lambda,\mu) = \sum_{k,j} c''_{k,j}\left(e^{i\langle\lambda,t_k\rangle}-1\right)\left(e^{-i\langle\mu,t_j\rangle}-1\right).$$

It is seen from (3.17)–(3.19) that the convergent sequence $\{\eta_n \in H_{F\times F}(D\times D)\}$ is associated with a sequence of functions $\{\varphi_n \in L_{F\times F}(D\times D)\}$, and the double stochastic integral in (3.17) can be extended using L^2-convergence to all $\varphi \in H_{F\times F}(D\times D)$. Moreover, any variable $\eta \in H_{F\times F}(D\times D)$ as the limit of a sequence η_n of the type given in (3.15) can be represented by (3.17), where the function $\varphi(\lambda,\mu) \in L_{F\times F}(D\times D)$ is the limit of the corresponding functions φ_n of the type given in (3.18). Given any function $\varphi(\lambda,\mu)$, (3.17) defines a certain variable $\eta \in H_{F\times F}(D\times D)$. So we have finished the proof of the following Lemma.

Lemma 3.3.8. *Each random variable $\eta \in H_{F\times F}(D\times D)$ can be represented as*

$$\eta(\varphi) = \iint_{\mathbb{R}^d\times\mathbb{R}^d} \varphi(\lambda,\mu)\Psi(d\lambda,d\mu), \tag{3.20}$$

for some $\varphi(\lambda,\mu)\in L_{F\times F}(D\times D)$*. Especially, for any* $s,t\in D$*,*

$$X(s)X(t)-K(s,t)=\iint_{\mathbb{R}^d\times\mathbb{R}^d}\big(e^{i\langle\lambda,s\rangle}-1\big)\big(e^{-i\langle\mu,t\rangle}-1\big)\Psi(d\lambda,d\mu). \qquad (3.21)$$

For every function $\varphi(\lambda,\mu)\in L_{F\times F}(D\times D)$*, (3.20) is well defined and* $\eta\in H_{F\times F}(D\times D)$*.*

Similar to the definition of $L_{F\times F}(D\times D)$, let us take $L_{F_0\times F_1}(D\times D)$ to be the closure of $L^0_{D\times D}$ with respect to the inner product

$$\langle\varphi_1,\,\varphi_2\rangle_{F_0\times F_1}=\iint_{\mathbb{R}^d\times\mathbb{R}^d}\varphi_1(\lambda,\mu)\,\overline{\varphi_2(\lambda,\mu)}F_0(d\lambda)F_1(d\mu),$$

where $\varphi_1,\,\varphi_2\in L^0_{D\times D}$. For random variables $\eta(\varphi),\eta(\psi)\in H_F(D)$, denote

$$b(\varphi,\psi)=K_0\big(\eta(\varphi),\eta(\psi)\big)-K_1\big(\eta(\varphi),\eta(\psi)\big),$$

where $K_0,\ K_1$ are covariance operators, which are defined in Section 2. The following theorem gives a criterion for the equivalence of two Gaussian measures with the same mean function, which is an extension of Theorem 5 of Ibragimov and Rozanov (1978), page 84, where stationary processes are considered.

Theorem 3.3.9. *Gaussian measures* P_0 *and* P_1 *with 0 mean values are equivalent on* $\mathcal{U}(D)$ *if and only if,* $b(\varphi,\psi)$ *being a functional on the class of functions* $\varphi(\lambda)\overline{\psi(\mu)}\in L^0_{D\times D}$*, can be extended to a linear continuous functional on* $L_{F_0\times F_1}(D\times D)$*.*

Proof The proof is similar to that of Theorem 5 of Ibragimov and Rozanov (1978), page 84. It should also be based on the entropy of Gaussian distribution and the definition of the $L_{F_0\times F_1}(D\times D)$. We omit the proof, and leave it to interested readers. □

As a consequence, we obtain a more explicit necessary and sufficient condition for the equivalence of two Gaussian measures which differ only in the covariance functions.

Theorem 3.3.10. *Gaussian measures P_0 and P_1 with 0 mean values are equivalent on $\mathcal{U}(D)$ if and only if, the difference of the two covariance functions $b(s,t) = K_0(s,t) - K_1(s,t)$ can be expressed as*

$$b(s,t) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \left(e^{-i\langle \lambda, s\rangle} - 1\right)\left(e^{i\langle \mu, t\rangle} - 1\right)\varphi(\lambda,\mu) F_0(d\lambda) F_1(d\mu) \tag{3.22}$$

for all $s, t \in D$, where $\varphi(\lambda,\mu) \in L_{F_0 \times F_1}(D \times D)$. Moreover, the Randon-Nikodym derivative $p(\omega) = P_1(d\omega)/P_0(d\omega)$ on the σ-algebra $\mathcal{U}(D)$ can be represented as

$$p(\omega) = C \exp\left\{-\frac{1}{2}\iint_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(\lambda,\mu)\Psi(d\lambda, d\mu)\right\}, \tag{3.23}$$

where C is a normalizing multiplier, and the definition of the double integral in (3.23) is the same as (3.20).

Let $R_{K_0 \times K_1}(D \times D)$ be the reproducing kernel Hilbert space corresponding to the kernel $K_0 \times K_1$, which is a function of four variables (s, s_1, t, t_1) defined by

$$K_0 \times K_1(s, s_1, t, t_1) = K_0(s,t) K_1(s_1, t_1).$$

Similar to Corollary 3.3.4, we have the following result, which is consistent with the results in Parzen (1963) and Chatterji and Mandrekar (1978), where the tools of RKHS are applied to study the equivalence and orthogonality of two Gaussian measures. It is worth noting that the criterion obtained by using the method of RKHS is general, which has no constraints like stationarity or isotropy on the underlying process, and the result are applicable to any multi-dimensional case.

Corollary 3.3.11. *Under the conditions of Theorem 3.3.10, the Gaussian measures P_0 and P_1 are equivalent on $\mathcal{U}(D)$, if and only if $b(s,t) = K_0(s,t) - K_1(s,t)$ is in the RKHS $R_{K_0 \times K_1}(D \times D)$.*

Proof The proof is similar to that of Corollary 3.3.4, and it follows directly from Theorem 3.3.10. □

We now assume P_0 and P_1 have spectral densities f_0 and f_1, respectively.

Theorem 3.3.12. *Gaussian measures P_0 and P_1 with 0 mean are equivalent on $\mathcal{U}(D)$ if and only if, $b(s,t)$ can be represented as*

$$b(s,t) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \left(e^{-i\langle \lambda, s\rangle} - 1\right)\left(e^{i\langle \mu, t\rangle} - 1\right) g(\lambda, \mu) d\lambda \, d\mu \tag{3.24}$$

for all $s, t \in \mathbb{R}^d$ (i.e. $b(s,t)$ is extendable to be a function on $\mathbb{R}^d \times \mathbb{R}^d$) and $g(\lambda, \mu)$ satisfies

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} \frac{|g(\lambda, \mu)|^2}{f_0(\lambda) f_1(\mu)} \, d\lambda \, d\mu < \infty. \tag{3.25}$$

Remark 3.3.13 If $f_0(\lambda) \leq \frac{K}{(1+|\lambda|^2)^n}$, for $|\lambda|$ large, then $g(\lambda, \mu)$ satisfies

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} \left(1 + |\lambda|^2\right)^n \left(1 + |\mu|^2\right)^n |g(\lambda, \mu)|^2 < \infty.$$

This implies that for any $k, m = 0, 1, \cdots, \left\lfloor n - \frac{d+1}{2} \right\rfloor$,

$$
\begin{aligned}
&\iint_{\mathbb{R}^d \times \mathbb{R}^d} |\lambda|^k |\mu|^m |g(\lambda, \mu)| d\lambda\, d\mu \\
&\leq \left[\int_{\mathbb{R}^d} \left(\frac{|\lambda|^k}{1 + |\lambda|^n} \right)^2 d\lambda \cdot \int_{\mathbb{R}^d} \left(\frac{|\mu|^m}{1 + |\mu|^n} \right)^2 d\mu \right]^{1/2} \\
&\quad \cdot \left[\iint_{\mathbb{R}^d \times \mathbb{R}^d} (1 + |\lambda|^n)^2 (1 + |\mu|^n)^2 |g(\lambda, \mu)|^2 d\lambda\, d\mu \right]^{1/2} < \infty.
\end{aligned}
$$

Therefore, the function $b(s,t)$ has all partial derivatives in each variable up to the order $\left\lfloor n - \frac{d+1}{2} \right\rfloor$: e.g. $\forall\, k, m = 0, 1, \cdots, \left\lfloor n - \frac{d+1}{2} \right\rfloor$,

$$
\frac{\partial^{k+m}}{\partial s_j^k\, \partial t_\ell^m} b(s,t) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} (-i\lambda_j)^k (-i\mu_\ell)^m e^{-i(\langle \lambda, s \rangle - \langle \mu, t \rangle)} g(\lambda, \mu) d\lambda\, d\mu
$$

for all $j, \ell = 1, 2, \cdots, d$.

When $d = 1$ and the processes are stationary, Ibragimov and Rozanov (1978) gave a necessary and sufficient condition for P_0 and P_1 to be equivalent on $\mathcal{U}([0,\tau])$ in terms of the $(2n)$th-order derivative [see Theorem 13 of Ibragimov and Rozanov (1978), page 99]. It seems to be an open problem whether analogous results still hold for intrinsically stationary Gaussian random fields (i.e. Gaussian random fields with stationary increments).

In the following, we prove a sufficient condition for the equivalence of the Gaussian measure P_0 and P_1 on $\mathcal{U}([0,\tau])$ when $d = 1$.

Corollary 3.3.14. *Assume $d = 1$, $\frac{\partial^2 b(s,t)}{\partial s \partial t}$ is the Fourier transform of the form*

$$
\frac{\partial^2 b(s,t)}{\partial s \partial t} = \iint_{\mathbb{R} \times \mathbb{R}} e^{-i(\lambda s - \mu t)} \psi(\lambda, \mu) d\lambda d\mu
$$

for some $\psi(\lambda,\mu) \in L^1(\mathbb{R}^2)$*, which satisfies*

$$\iint_{\mathbb{R}\times\mathbb{R}} \frac{|\psi(\lambda,\mu)|^2}{\lambda^2\mu^2 f_0(\lambda) f_1(\mu)} d\lambda d\mu. \tag{3.26}$$

Then P_0 *and* P_1 *are equivalent on the* σ*-algebra* $\mathcal{U}([0,\tau])$*.*

For our conjecture, the following is another sufficient condition for P_0 and P_1 to be equivalent on the σ-algebra $\mathcal{U}([0,\tau])$, which extends Theorem 17 in Ibragimov and Rozanov (1978), page 104.

Conjecture 3.3.15 We assume $d = 1$ and the spectral densities f_0 and f_1 satisfy the condition

$$f_0(\lambda) \asymp f_1(\lambda) \asymp (1+\lambda^2)^{-n}.$$

If

$$\int_{\mathbb{R}} \frac{\big(f_0(\lambda) - f_1(\lambda)\big)^2}{f_0^2(\lambda)} d\lambda < \infty,$$

then P_0 and P_1 are equivalent on the σ-algebra $\mathcal{U}([0,\tau])$. □

We have listed several criteria for two Gaussian measures to be equivalent in the intrinsically stationary random fields on the above. Then, with these conditions for the equivalence of two Gaussian measures, you may ask "What if the two measures are equivalent?" Here is an answer. Theorem 3.2 of Sottinen and Tudor (2006) states that every mean square continuous Gaussian random field $\{X(t), P_1\}$ which is equivalent to a given Gaussian random field $\{X(t), P_0\}$ admits a non-anticipative representation with respect to $\{X(t), P_0\}$. Now, we work to derive an explicit representation under the equivalence of two intrinsically stationary random fields.

Theorem 3.3.16. *Suppose Gaussian measures P_0 and P_1 are equivalent, then $\{X(t), P_1\}$ has a representation $x(t)$ with respect to $\{X(t), P_0\}$, such that*

$$x(t) = X(t) + \int_{\mathbb{R}^d} \int_{[-\infty,\lambda]} b(\mu,\lambda)\Phi(d\mu)\big(e^{i\langle\lambda,t\rangle} - 1\big)d\lambda,$$

where b is a square integrable Volterra kernel.

We can find more consequences of the equivalence of two Gaussian measures from the next section.

3.4 Asymptotic optimality of linear predictions

In practice, the true probability distribution of our Gaussian model is always unknown and must be estimated from the gathered data. To this end, it is of great value to investigate the effect of using a fixed but incorrect probability distribution, especially, when more sample data can be obtained by sampling the spatial or temporal domain increasingly densely. This section studies the effect of misspecifying the mean and covariance function of a random field on optimal linear predictions of the random field.

Suppose P_0 and P_1 are two equivalent Gaussian measures. Write $H_{F_0}(D)$ as $H_0(D)$ for short in this section. Let $h_1, h_2, \cdots$ be a complete system of linearly independent elements from $H_0(D)$, and take $\psi_1, \psi_2, \cdots$ to be the Gram-Schmidt orthogonalization of $h_1, h_2, \cdots$ under $(0, K_0)$, such that

$$K_0(\psi_j,\psi_k) = \begin{cases} 1, & j = k, \\ 0, & j \neq k, \end{cases} \quad \text{and} \quad K_1(\psi_j,\psi_k) = \begin{cases} \sigma_k^2, & j = k, \\ 0, & j \neq k. \end{cases}$$

Of course, the closed linear hull of $\psi_1, \psi_2, \cdots$ under the inner product defined by $(0, K_0)$ is $H_0(D)$. Let $\psi \in H_0(D)$, then the best linear predictor of ψ given $\psi_1, \cdots, \psi_n$ under $(0, K_0)$ is $\hat{\psi}_n = k_n' \Psi_n$, where $\Psi_n = (\psi_1, \cdots, \psi_n)'$ and $k_n = (K_0(\psi, \psi_1), \cdots, K_0(\psi, \psi_n))'$. Let $e_0(\psi, n) = \psi - \hat{\psi}_n$ be the prediction error under $(0, K_0)$. Similarly, define $e_1(\psi, n)$ to be the error of the best linear prediction with respect to (m_1, K_1). In the following, we suppose (m_1, K_1) to be the presumed second-order structure when in fact $(0, K_0)$ is the actual second-order structure. We will then consider the behavior of the best linear predictor as $n \to \infty$. Conventionally, we assume $0/0 = 0$ throughout this section.

First, any $\psi \in H_0(D)$ can be written as $\psi = \sum_{i=1}^{\infty} c_i \psi_i$, where $c_i = \langle \psi, \psi_i \rangle_{K_0} = K_0(\psi, \psi_i)$. So $\sum_{i=1}^{\infty} c_i^2 < \infty$. We can then write

$$e_0(\psi, n) = \sum_{i=n+1}^{\infty} c_i \psi_i.$$

Define

$$\mu_j = \mathbb{E}_1 \psi_j, \quad \text{for } j = 1, 2, \cdots$$

and

$$b_{jk} = K_1(\psi_j, \psi_k) - K_0(\psi_j, \psi_k), \quad \text{for } j, k = 1, 2, \cdots.$$

The following results of asymptotic theory are from Stein (1988, 1990, 1999a and 1999b), which hold for any Gaussian random field, including both stationary and intrinsically stationary ones.

Theorem 3.4.1. *Suppose P_0 and P_1 are two equivalent Gaussian measures. As $n \to \infty$,*

$$\sup_{\psi \in H_0(D)} \frac{\mathbb{E}_1 e_0(\psi, n)^2 - \mathbb{E}_0 e_0(\psi, n)^2}{\mathbb{E}_0 e_0(\psi, n)^2} = \Lambda_n \downarrow 0$$

and

$$\inf_{\psi \in H_0(D)} \frac{\mathbb{E}_1 e_0(\psi, n)^2 - \mathbb{E}_0 e_0(\psi, n)^2}{\mathbb{E}_0 e_0(\psi, n)^2} = \lambda_n \uparrow 0,$$

where Λ_n and λ_n are, respectively, the largest and smallest eigenvalues of the infinite matrix $(b_{jk} + \mu_j \mu_k)_{jk=n+1}^{\infty}$.

Switching the roles of $(0, K_0)$ and (m_1, K_1), we can define the corresponding largest and smallest eigenvalues as $\widetilde{\Lambda}_n$ and $\widetilde{\lambda}_n$, respectively, such that

$$\sup_{\psi \in H_0(D)} \frac{\mathbb{E}_0 e_1(\psi, n)^2 - \mathbb{E}_1 e_1(\psi, n)^2}{\mathbb{E}_1 e_1(\psi, n)^2} = \widetilde{\Lambda}_n \downarrow 0$$

and

$$\inf_{\psi \in H_0(D)} \frac{\mathbb{E}_0 e_1(\psi, n)^2 - \mathbb{E}_1 e_1(\psi, n)^2}{\mathbb{E}_1 e_1(\psi, n)^2} = \widetilde{\lambda}_n \uparrow 0.$$

The above theorem comes from Stein (1990), page 855. Moreover, using some elementary results, we obtain the following results, see Stein (1999b), page 130.

Corollary 3.4.2. *Suppose P_0 and P_1 are two equivalent Gaussian measures. Then*

$$\lim_{n\to\infty} \sup_{\psi \in H_0(D)} \left| \frac{\mathbb{E}_1 e_0(\psi, n)^2 - \mathbb{E}_0 e_0(\psi, n)^2}{\mathbb{E}_0 e_0(\psi, n)^2} \right| = 0,$$

$$\lim_{n\to\infty} \sup_{\psi \in H_0(D)} \frac{\mathbb{E}_0 e_1(\psi, n)^2 - \mathbb{E}_0 e_0(\psi, n)^2}{\mathbb{E}_0 e_0(\psi, n)^2} = 0$$

and

$$\lim_{n\to\infty} \sup_{\psi \in H_0(D)} \frac{\mathbb{E}_0 \big(e_1(\psi, n) - e_0(\psi, n)\big)^2}{\mathbb{E}_0 e_0(\psi, n)^2} = 0.$$

Switching the roles of $(0, K_0)$ and (m_1, K_1), then

$$\lim_{n\to\infty} \sup_{\psi\in H_0(D)} \left| \frac{\mathbb{E}_0 e_1(\psi,n)^2 - \mathbb{E}_1 e_1(\psi,n)^2}{\mathbb{E}_1 e_1(\psi,n)^2} \right| = 0,$$

$$\lim_{n\to\infty} \sup_{\psi\in H_0(D)} \frac{\mathbb{E}_1 e_0(\psi,n)^2 - \mathbb{E}_1 e_1(\psi,n)^2}{\mathbb{E}_1 e_1(\psi,n)^2} = 0$$

and

$$\lim_{n\to\infty} \sup_{\psi\in H_0(D)} \frac{\mathbb{E}_1 \big(e_0(\psi,n) - e_1(\psi,n)\big)^2}{\mathbb{E}_1 e_1(\psi,n)^2} = 0.$$

Taking the observations as $\psi_1, \psi_2, \cdots$, which form a basis of the Hilbert space $H_0(D)$ is convenient mathematically, but in fact excludes some common and interesting applications in the asymptotics. In real life, we care more about the prediction for an unknown value $X(t)$, $t \in D$, based on the observations $X(t_1), \cdots, X(t_n)$, where $t_1, \cdots, t_n \in D$ but different from t. Let us take $\hat{X}_i(n)$ for $i = 0, 1$ to denote the best linear predictor of $X(t)$, using (m_i, K_i) as the second-order structure, where $m_0 \equiv 0$. Define $e_i(n) = X(t) - \hat{X}_i(n)$, the error of the corresponding prediction. We obtain the following results which are directly related to Corollary 3.4.2.

Corollary 3.4.3. *Suppose P_0 and P_1 are two equivalent Gaussian measures. Let $t \in D$ and $\{t_i\}_{i=1}^{\infty}$ be a sequence in D not containing t but having t as its limit point, such that $\mathbb{E}_0 e_0(n)^2 > 0$. Then*

$$\lim_{n\to\infty} \left| \frac{\mathbb{E}_1 e_0(n)^2 - \mathbb{E}_0 e_0(n)^2}{\mathbb{E}_0 e_0(n)^2} \right| = 0, \tag{3.27}$$

$$\lim_{n\to\infty} \frac{\mathbb{E}_0 e_1(n)^2 - \mathbb{E}_0 e_0(n)^2}{\mathbb{E}_0 e_0(n)^2} = 0 \tag{3.28}$$

and

$$\lim_{n\to\infty} \frac{\mathbb{E}_0 \big(e_1(n) - e_0(n)\big)^2}{\mathbb{E}_0 e_0(n)^2} = 0. \tag{3.29}$$

The corollary above follows directly from Theorem 10 [Stein (1999b), page 132], and switching the roles of $(0, K_0)$ and (m_1, K_1) is also feasible. Note that, the assumption of $\mathbb{E}_0 e_0(n)^2 \to 0$ as $n \to 0$ in Theorem 10 [Stein (1999b), page 132] is guaranteed by mean-square continuity of $X(t), t \in D$, under P_0. We obtain Equation (3.28), saying there is an asymptotically efficient predictor $\hat{X}_1(n)$ under the presumed second-order structure (m_1, K_1) when, in fact, $(0, K_0)$ is the correct second-order structure, as long as P_0 and P_1 are equivalent. Moreover, the predictions obtained under those two second-order structures are asymptotically close to each other [see (3.29)], and the discrepancy between the presumed mean-squared prediction error and the actual mean-squared prediction error is asymptotically 0 [see (3.27)].

3.5 Explicit bounds with equal covariance functions

In this section, we want to obtain the bounds on Λ_n and λ_n of Theorem 3.4.1 for intrinsically stationary Gaussian random fields. The bounds can be obtained by approximating an element of a Hilbert space by an element of a finite-dimensional subspace. This problem has been considered as it applies to optimal design for estimating the regression coefficients of a stochastic process. The references include Sacks and Ylvisaker (1966, 1968, 1970), Wahba (1971, 1974) and Eubank, Smith and Smith (1981). Stein (1990) obtained results on these bounds for less smooth mean functions than those considered in previous work for stationary, second-order random fields. We extend Stein (1990)'s method to investigate the bounds for intrinsically stationary random fields. Actually, the general case appears to be rather difficult, however, it simplifies considerably under equal covariance functions, like the case I in Section 3.3.1. From Stein (1990), page 857, we have

$$\Lambda_n = \sum_{j=n+1}^{\infty} \mu_j^2 \quad \text{and} \quad \lambda_n = 0.$$

Moreover, Λ_n has an upper bound as follows:

$$\Lambda_n \leq \mathbb{E}_0(\nu - \nu_n)^2, \tag{3.30}$$

for some $\nu_n \in H_n(D)$ (the subspace of $H_F(D)$ generated by $\psi_1, \cdots, \psi_n$), where ν is the Randon-Nikodym derivative of P_1 with respect to P_0. In the following, we derive bounds on Λ_n under certain conditions on F, by using the characteristics of the associated function space $L_F(D)$. Let us start from the one-dimensional process.

3.5.1 One-dimensional Processes

Let $D = [0, \tau]$, for $\tau > 0$. Suppose $F(d\lambda) = f(\lambda)d\lambda$, and $f(\lambda)$ satisfies $\int_{\mathbb{R}}(1 \wedge |\lambda|^2)f(\lambda)d\lambda < \infty$ and for a positive integer m,

$$f(\lambda) \asymp (1 + \lambda^2)^{-m}. \tag{3.31}$$

Theorem 3.5.1. *Under the condition (3.31), all elements of the function space $L_F(D)$ can be expressed as*

$$\varphi(\lambda) = P(i\lambda) + (1 + i\lambda)^{m-1} \int_0^{\tau} (e^{i\lambda t} - 1)c(t)dt, \tag{3.32}$$

with

$$P(i\lambda) = \sum_{k=1}^{m-1} c_k (i\lambda)^k, \tag{3.33}$$

where c_k's are real and $c(t)$ is a square-integrable real function on $D = [0, \tau]$.

Remark 3.5.2

(1) If the condition (3.31) changes with m as a positive non-integer number, then Theorem 3.5.1 still holds, by replacing m with its integer part $\lfloor m \rfloor$ in (3.32) and (3.33).

(2) The conclusion of Theorem 3.5.1 still holds well under the following weaker condition than (3.31):

$$f(\lambda) \leq \beta \lambda^{-2m}, \qquad \text{as } |\lambda| \to \infty,$$

where $\beta > 0$, $m > 0$ are constant. We can derive this statement by applying Lemma 2.7.1 to the proof of Theorem 3.5.1.

(3) The analytic function space $L_F(D)$ for an intrinsically stationary random field is different from that of a stationary random field. All elements in $L_F(D)$ of a stationary random field under condition (3.31) are given as

$$\sum_{k=0}^{m-1} c_k (i\lambda)^k + (1 + i\lambda)^m \int_0^\tau e^{i\lambda t} c(t) dt, \tag{3.34}$$

where c_k's and $c(t)$ are the same in (3.32) [see Stein (1990)]. As we can see from (3.34) the Hilbert space $L_F(D)$ for the stationary case contains real constants as members.

In fact, if the spectral density $f(\lambda)$ satisfies (3.31), the Gaussian process $X(t)$ has $(m-1)$th mean-square derivative. Besides, Theorem 2.4.8 in Chapter 2 shows that the sample functions are differentiable up to $(m-1)$ orders. Without loss of generality, we will take $\tau = 1$ in the following. Let $H_{n,p}$ be the subspace generated by $X^{(j)}(t_k)$ for $j = 0, \cdots, p$ with $p \leq m - 1$ and $0 = t_0 < \cdots < t_n = 1$. Let $L_{n,p}$ be the subspace of $L_F(D)$ isomorphic to $H_{n,p}$, and let

$P_{n,p}$ be the operator that projects elements of $L_F(D)$ onto $L_{n,p}$, so that

$$\inf_{\varphi_n \in L_{n,p}} \|\varphi - \varphi_n\|_F^2 = \|\varphi - P_{n,p}\varphi\|_F^2 \tag{3.35}$$

for all $\varphi \in L_F(D)$. From Theorem 3.4.1 and (3.30), for any $\varphi_n \in L_{n,p}$, $\|\varphi - \varphi_n\|_F^2$ is an uniform bound for Λ_n. Define $\Delta_k = t_k - t_{k-1}$, for $k = 1, \cdots, n$. Assume $f(\lambda)$ satisfies (3.31). From Theorem 3.3.3, we know that a necessary and sufficient condition for the equivalence of Gaussian measures P_0 and P_1 with the same covariance function is

$$m_1(t) = \int_0^1 \left(e^{-i\lambda t} - 1\right)\varphi(\lambda) f(\lambda) d\lambda, \tag{3.36}$$

for some $\varphi(\lambda) \in L_F(D)$, which can be written as (3.32). In the rest of this subsection, we assume $m_1(t)$ can be represented as (3.36). Then we obtain the following upper bounds, which extend Theorem 4.1 and Theorem 4.2 of Stein (1990), page 859.

Proposition 3.5.3. *Suppose there exists $\ell \le m$, such that $c(t)$ given in (3.32) has an absolutely continuous $(\ell - 1)$th derivative and $c^{(\ell)}(t)$ is square-integrable on $[0,1]$. Let $h(t) = c(t)e^{-t}$, then*

$$\|\varphi - P_{n,m-1}\varphi\|_F^2 \le c\big[(\ell-1)!\big]^{-2} \sum_{k=1}^{n} \left(\frac{\Delta_k}{2}\right)^{2\ell} \int_{t_{k-1}}^{t_k} h^{(\ell)}(t)^2 dt, \tag{3.37}$$

where c is a positive constant. Moreover, if $\Delta_k = 1/n$ for all k, then the upper bound in (3.37) can be written as

$$c\, n^{-2\ell} 2^{-2\ell} \big[(\ell-1)!\big]^{-2} \int_0^1 h^{(\ell)}(t)^2 dt.$$

Remark 3.5.4 If we specify (3.31) as that there exist two positive constants α and β, such that

$$\alpha(1+\lambda^2)^{-m} \leq f(\lambda) \leq \beta(1+\lambda^2)^{-m}, \tag{3.38}$$

then the constant c in (3.37) can be expressed as $4\pi\beta e^2$.

Proposition 3.5.5. *Suppose $\varphi(\lambda)$ given in (3.36) is of the form*

$$\varphi(\lambda) = \int_0^1 c(t)\big(e^{i\lambda t} - 1\big)dt,$$

where $|c(t)|$ is uniformly bounded by C on $[0,1]$. Let $\rho = n\max\{\Delta_k, 1 \leq k \leq n\}$. Then for $m > 1$,

$$\|\varphi - P_{n,0}\varphi\|_F^2 \leq \frac{16\beta n C^2}{2n-1}\big[2\rho(m-1)\big]^{2m} \max\left\{1, \Big(\frac{4\rho(m-1)}{m!}\Big)^2\right\} n^{-2m+1}.$$

3.5.2 Two-dimensional random fields

We now give an analogue to Theorem 3.5.1 for two-dimensional random fields. The extension to high dimensions is not difficult. For convenience, let us introduce the separable random field first.

Suppose $D = [0,\tau]\times[0,\tau]$ for $\tau > 0$, the spectral density satisfies that for $\lambda = (\lambda_1, \lambda_2) \in \mathbb{R}^2$ and positive integers m_1, m_2,

$$f(\lambda) \asymp \big(1+|\lambda_1|^2\big)^{-m_1}\big(1+|\lambda_2|^2\big)^{-m_2}. \tag{3.39}$$

Theorem 3.5.6. *Under the condition given by (3.39), all elements of the function space*

$L_F(D)$ can be expressed as

$$\varphi(\lambda) = P(i\lambda) + Q(\lambda) + (1+i\lambda_1)^{m_1-1}(1+i\lambda_2)^{m_2-1}\int_D (e^{i\lambda' t}-1)c(t)dt, \tag{3.40}$$

for

$$P(i\lambda) = \sum_{j=0}^{m_1-1}\sum_{k=0}^{m_2-1} a_{jk} i^{j+k}\lambda_1^j\lambda_2^k \tag{3.41}$$

and

$$\begin{aligned} Q(\lambda) = & \sum_{k=0}^{m_2-1} a_k (i\lambda_2)^k (1+i\lambda_1)^{m_1-1}\int_0^\tau (e^{i\lambda_1 t_1}-1)b_1(t_1)dt_1 \\ & + \sum_{j=0}^{m_1-1} \tilde{a}_k (i\lambda_1)^j (1+i\lambda_2)^{m_2-1}\int_0^\tau (e^{i\lambda_2 t_2}-1)b_2(t_2)dt_2 \Big], \end{aligned} \tag{3.42}$$

where $a_{00}=0$, a_{jk}'s, a_k's and $\tilde{a}_k$'s are real, $b_1(t_1)$, $b_2(t_2)$ are square-integrable real functions on $[0,\tau]$ and $c(t)$ is a square-integrable real function on $D=[0,\tau]\times[0,\tau]$.

Remark 3.5.7

(1) If the condition (3.39) changes with m_1 and m_2 as positive non-integer numbers, then Theorem 3.5.6 still holds, by replacing m_1 (m_2) with its integer part $[m_1]$ ($[m_1]$) in (3.40), (3.41) and (3.42).

(2) The conclusion of Theorem 3.5.6 still holds under the following weaker condition than (3.39):

$$f(\lambda) \le \beta\lambda_1^{-2m_1}\lambda_2^{-2m_2}, \qquad \text{as } |\lambda_1|, |\lambda_2| \to \infty,$$

where $\beta > 0$ is constant. We can derive this statement by applying Lemma 2.7.1 to the proof of Theorem 3.5.6.

For the two-dimensional case, we can have similar results to Propositions 3.5.3 and 3.5.5. We leave it to interested readers.

3.6 Proofs

Proof of Lemma 3.2.1

Let us prove the representation of $\eta \in H_F(D)$ first. By (2.3), it is obvious that $\forall\, t \in D$, $\eta = X(t)$ satisfies (3.5), with $\varphi(\lambda) = e^{i\langle \lambda, t\rangle} - 1 \in L_F(D)$. Then for any positive integer m, the linear combination of $X(t)$, such as $\eta = \sum_{k=1}^{m} c_k X(t_k)$, where c_k's are real and t_k's $\in D$, can be written as

$$\sum_{k=1}^{m} c_k \int_{\mathbb{R}^d} \big(e^{i\langle \lambda, t_k\rangle} - 1\big)\Phi(d\lambda)$$
$$= \int_{\mathbb{R}^d} \sum_{k=1}^{m} c_k \big(e^{i\langle \lambda, t_k\rangle} - 1\big)\Phi(d\lambda).$$

Denote $\varphi(\lambda) = \sum_{k=1}^{m} c_k\big(e^{i\langle \lambda, t_k\rangle} - 1\big) \in L_F(D)$, then (3.5) is satisfied. We can also verify that for any $\eta_1 = \sum_{k=1}^{m} c_k X(t_k) \in H_F(D)$, $\eta_2 = \sum_{\ell=1}^{n} d_\ell X(t_\ell) \in H_F(D)$,

$$\begin{aligned}
\langle \eta_1, \eta_2\rangle &= \sum_{k=1}^{m}\sum_{\ell=1}^{n} c_k d_\ell \int_{\mathbb{R}^d} \big(e^{i\langle \lambda, t_k\rangle} - 1\big)\overline{\big(e^{i\langle \lambda, t_\ell\rangle} - 1\big)} F(d\lambda) \\
&= \int_{\mathbb{R}^d} \Big[\sum_{k=1}^{m} c_k\big(e^{i\langle \lambda, t_k\rangle} - 1\big)\Big]\overline{\Big[\sum_{\ell=1}^{n} d_\ell\big(e^{i\langle \lambda, t_\ell\rangle} - 1\big)\Big]} F(d\lambda) \qquad (3.43)\\
&= \langle \varphi_1, \varphi_2\rangle_F,
\end{aligned}$$

where $\varphi_1(\lambda) = \sum_{k=1}^{m} c_k\big(e^{i\langle \lambda, t_k\rangle} - 1\big) \in L_F(D)$, $\varphi_2(\lambda) = \sum_{\ell=1}^{n} d_\ell\big(e^{i\langle \lambda, t_\ell\rangle} - 1\big) \in L_F(D)$. It follows at once that if there is a sequence $\{\eta_n\} \in H_F(D)$ of the form

$$\eta_n = \sum_{k=1}^{m_n} c_{kn} X(t_{kn}),$$

with a limit point $\eta \in H_F(D)$, i.e.,

$$\mathbb{E}|\eta_n - \eta|^2 \to 0, \quad \text{as} \quad n \to \infty,$$

there should exist a corresponding sequence $\{\varphi_n\} \in L_F(D)$, which can be written as

$$\varphi_n = \sum_{k=1}^{m_n} c_{kn} \big(e^{i\langle \lambda, t_{kn}\rangle} - 1\big),$$

such that

$$\|\varphi_n - \varphi\|_F^2 = \int_{\mathbb{R}^d} |\varphi_n(\lambda) - \varphi(\lambda)|^2 F(d\lambda) \to 0, \quad \text{as} \quad n \to \infty,$$

where $\varphi \in L_F(D)$. In fact, φ depends only on η instead of the sequence $\{\eta_n\} \in L_F(D)$, such as $\eta = \int_{\mathbb{R}^d} \varphi(\lambda)\Phi(d\lambda)$. So (3.5) holds for any $\eta \in H_F(D)$.

By (3.43) and a similar limiting argument, we see that for any function $\varphi \in L_F(D)$, (3.5) is well defined, which then yields $\eta \in H_F(D)$. □

Proof of Lemma 3.2.3

It is clear that if $\|\varphi\|_{F_0} = 0$ and $\|\varphi\|_{F_1} \neq 0$ for a function $\varphi(\lambda) \in L_D^0$, the measures P_0 and P_1 are orthogonal, since the corresponding random variable $\eta(\varphi) \in H_F(D)$ from (3.5) satisfies

$$P_0\{\eta(\varphi) = 0\} = 1 \quad \text{and} \quad P_1\{\eta(\varphi) = 0\} = 0.$$

Furthermore, if there exists a sequence $\{\varphi_n(\lambda)\} \in L_D^0$, such that

$$\|\varphi_n\|_{F_0} = 1 \text{ and } \sigma_n = \|\varphi_n\|_{F_1} \to 0, \quad \text{as } n \to \infty,$$

then for $m_1(\varphi_n) = \mathbb{E}_1(\eta(\varphi_n))$, we show that as $n \to \infty$,

$$P_0 \{|\eta(\varphi_n) - m_1(\varphi_n)| < \sqrt{\sigma_n}\} = \int_{|x - m_1(\varphi_n)| < \sqrt{\sigma_n}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \to 0,$$

$$P_1 \{|\eta(\varphi_n) - m_1(\varphi_n)| < \sqrt{\sigma_n}\} = \int_{|x| < (1/\sqrt{\sigma_n})} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \to 1.$$

Similar relations hold true if

$$\|\varphi_n\|_{F_1} = 1 \text{ and } \|\varphi_n\|_{F_0} \to 0, \quad \text{as } n \to \infty.$$

Hence, the desired result follows. □

Proof of Theorem 3.3.2

First of all, suppose P_0 and P_1 are two equivalent Gaussian measures. We first prove that the linear functional $m_1(\cdot)$ is bounded. Let $\{\varphi_n(\lambda)\}$ be a sequence in L_D^0, such that $\sigma_n = \|\varphi_n\|_{F_1} \asymp \|\varphi_n\|_{F_0} = 1$. Suppose $m_1(\varphi_n) \to \infty$ as $n \to \infty$, then

$$P_0 \left\{\eta(\varphi_n) > \sqrt{m_1(\varphi_n)}\right\} = \int_{\sqrt{m_1(\varphi_n)}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \to 0,$$

$$P_1 \left\{\eta(\varphi_n) > \sqrt{m_1(\varphi_n)}\right\} = \int_{-m_1(\varphi_n) + \sqrt{m_1(\varphi_n)}}^{\infty} \frac{1}{\sqrt{2\pi}\, \sigma_n} e^{-x^2/2\sigma_n^2} dx \to 1,$$

which imply a contradiction to the equivalence of P_0 and P_1. So $m_1(\varphi)$ is a linear bounded functional on the Hilbert space $L_F(D)$, which is equivalent to saying that the linear functional $m_1(\varphi)$ is continuous on $L_F(D)$. Hence, $\exists\ \psi(\lambda) \in L_F(D)$, such that $m_1(\varphi) = \langle \varphi, \psi \rangle_F$.

To prove the converse, suppose the mean value $m_1(\varphi)$ is a continuous linear functional on the Hilbert space $L_F(D)$, then there exists a unique $\psi \in L_F(D)$ such that $m_1(\varphi) = \langle \varphi, \psi \rangle_F$, for all $\psi \in L_F(D)$. Let $\{\varphi_k\} \in L_D^0$ be a complete orthonormal system in $L_F(D)$. It is known that the entropy distance between P_0 and P_1 on the σ-algebra $\mathfrak{U}_n$ generated by the variables $\eta(\varphi_k)$, $k = 1, \cdots, n$, is

$$r_n = \sum_{k=1}^{n} m_1(\varphi_k)^2 = \sum_{k=1}^{n} \langle \varphi_k, \psi \rangle_F^2;$$

see (2.9) of Ibragimov and Rozanov (1978), page 76. So

$$\lim_{n\to\infty} r_n = \sum_{k=1}^{\infty} \langle \varphi_k, \psi \rangle_F^2 = \|\psi\|_F^2 < \infty.$$

The equivalence of the Gaussian measures P_0 and P_1 follows now from Lemma 3 of Ibragimov and Rozanov (1978), page 77. □

Proof of Theorem 3.3.3

Since the system of functions $\varphi(\lambda) = e^{i\langle \lambda, t\rangle} - 1$, $t \in D$, is complete in $L_F(D)$, a necessary and sufficient condition for P_0 and P_1 to be equivalent is (3.10), which follows from Theorem 3.3.2.

Now suppose P_0, P_1 are two equivalent Gaussian measures, and the Randon-Nikodym derivative is $p(\omega) = P_1(d\omega)/P_0(d\omega)$ on $\mathfrak{U}(D)$. Choose a complete orthonormal system $\varphi_1(\lambda)$, $\varphi_2(\lambda), \cdots \in L_D^0$. First, consider the density $p_n(\omega) = P_1(d\omega)/P_0(d\omega)$ on the σ-algebra $\mathfrak{U}_n$, each of which is generated by the variables $\eta(\varphi_k)$, $k = 1, \cdots, n$. Actually, $p_n(\omega) =$

$\mathbb{E}\left(p(\omega)|\mathcal{U}_n\right)$. By the martingale convergence theorem,

$$p(\omega) = \lim_{n\to\infty} p_n(\omega).$$

Let $a_k = m_1(\varphi_k)$, $k = 1, 2, \cdots$, and $\psi_n(\lambda) = \sum_{k=1}^n a_k\varphi_k(\lambda)$. By Theorem 3.3.2, there exists $\varphi(\lambda) \in L_F(D)$ such that

$$a_k = \langle\varphi_k,\, \varphi\rangle_F, \quad \forall\, k \geq 1.$$

So

$$\psi_n(\lambda) = \sum_{k=1}^n \langle\varphi_k,\, \varphi\rangle_F\, \varphi_k(\lambda).$$

According to formula (2.2) of Ibragimov and Rozanov (1978), page 75, we have

$$\begin{aligned} p_n(\omega) &= \exp\left\{\sum_{k=1}^n a_k\eta(\varphi_k) - \frac{1}{2}\sum_{k=1}^n a_k^2\right\} \\ &= \exp\left\{\sum_{k=1}^n a_k \int_{\mathbb{R}^d} \varphi_k(\lambda)\Phi(d\lambda) - \frac{1}{2}\|\psi_n\|_F^2\right\} \\ &= \exp\left\{\int_{\mathbb{R}^d} \psi_n(\lambda)\Phi(d\lambda) - \frac{1}{2}\|\psi_n\|_F^2\right\}. \end{aligned}$$

Moreover,

$$\lim_{n\to\infty} \psi_n(\lambda) = \lim_{n\to\infty} \sum_{k=1}^n \langle\varphi_k,\, \varphi\rangle_F\, \varphi_k(\lambda) = \varphi(\lambda),$$

$$\lim_{n\to\infty} \|\psi_n\|_F^2 = \|\varphi\|_F^2.$$

Therefore, (3.11) holds. □

Proof of Corollary 3.3.5

On one hand, if P_0 and P_1 are equivalent on $\mathcal{U}(D)$, then by Theorem 3.3.3, there exists $\varphi \in L_F(D)$ such that (3.10) holds. We define $\psi(\lambda) = \varphi(\lambda)f(\lambda)$, then

$$\int_{\mathbb{R}^d} |\psi(\lambda)|^2 d\lambda = \int_{\mathbb{R}^d} |\varphi(\lambda)|^2 f^2(\lambda) d\lambda \leq c \int_{\mathbb{R}^d} |\varphi(\lambda)|^2 f(\lambda) d\lambda < \infty,$$

where c is a positive constant. So $\psi \in L^2(\mathbb{R}^d)$, and (3.10) can be rewritten as (3.12). Moreover, ψ satisfies (3.13).

On the other hand, suppose there exists $\psi \in L^2(\mathbb{R}^d)$, such that (3.12) and (3.13) hold. We take $\varphi(\lambda) = \psi(\lambda)/f(\lambda)$, then (3.13) implies

$$\int_{\mathbb{R}^d} |\varphi(\lambda)|^2 f(\lambda) d\lambda = \int_{\mathbb{R}^d} \frac{|\psi(\lambda)|^2}{f(\lambda)} d\lambda < \infty.$$

Let $\widetilde{\varphi}$ be the projection of φ into $L_F(D)$, then (3.12) implies that

$$m_1(t) = \int_{\mathbb{R}^d} \left(e^{-i\langle t,\lambda\rangle} - 1\right) \widetilde{\varphi}(\lambda) f(\lambda) d\lambda.$$

Hence, P_0 and P_1 are equivalent by Theorem 3.3.3. □

Proof of Corollary 3.3.6

Suppose P_0 and P_1 are equivalent, then there exists $\psi \in L^2(\mathbb{R}^d)$ such that (3.12) and (3.13) hold. By (3.13) and (3.14), we have

$$\int_{\mathbb{R}^d} \left(1 + |\lambda|^2\right)^n |\psi(\lambda)|^2 d\lambda < \infty.$$

This and Hölder's inequality imply that $\forall\, k \leq \left\lfloor n - \frac{d+1}{2} \right\rfloor$,

$$\begin{aligned}
&\int_{\mathbb{R}^d} |\lambda|^k |\psi(\lambda)| d\lambda \\
\leq &\left[\int_{\mathbb{R}^d} \frac{|\lambda|^{2k}}{\left(1 + |\lambda|^n\right)^2} d\lambda \right]^{1/2} \left[\int_{\mathbb{R}^d} \left(1 + |\lambda|^n\right)^2 |\psi(\lambda)|^2 d\lambda \right]^{1/2} < \infty.
\end{aligned}$$

Hence the conclusion follows from the dominated convergence theorem. □

Proof of Corollary 3.3.7

Since we can write

$$m_1'(t) = \int_{\mathbb{R}} e^{-i\lambda s} \psi(\lambda) d\lambda \qquad \forall\, t \in \mathbb{R}.$$

Fubini's theorem gives

$$\begin{aligned}
m_1(t) = \int_0^t m_1'(s) ds &= \int_0^t \int_{\mathbb{R}} e^{-i\lambda s} \psi(\lambda) d\lambda ds \\
&= \int_{\mathbb{R}} \left(\int_0^t e^{-i\lambda s} ds \right) \psi(\lambda) d\lambda \\
&= - \int_{\mathbb{R}} \left(e^{-i\lambda s} - 1\right) \frac{\psi(\lambda)}{i\lambda} d\lambda.
\end{aligned}$$

Since

$$\int_{\mathbb{R}} \left| \frac{\psi(\lambda)}{i\lambda} \right|^2 \frac{1}{f(\lambda)} d\lambda = \int_{\mathbb{R}} \frac{|\psi(\lambda)|^2}{\lambda^2 f(\lambda)} d\lambda < \infty,$$

the conclusion follows from Corollary 3.3.5 □

Proof of Theorem 3.3.10

Since the functions $\varphi(\lambda, \mu) = \left(e^{i\langle\lambda, s\rangle} - 1\right)\left(e^{-i\langle\mu, t\rangle} - 1\right)$, $s, t \in D$ form a complete system in $L_{F_0 \times F_1}(D \times D)$, (3.22) is equivalent to the condition given in Theorem 3.3.9.

Now, assume P_0, P_1 are equivalent and choose a sequence $t_1, t_2, \cdots$ everywhere dense in D. Then $\{X(t_k)\}$ forms a complete system in $H_F(D)$. Let us first consider the density $p_n(\omega) = P_1(d\omega)/P_0(d\omega)$ on the σ-algebra $\mathfrak{U}_n$, generated by the variables $X(t_k)$, $k = 1, \cdots, n$. Analogous to (3.11) of Ibragimov and Rozanov (1978), page 89, we have

$$\log p_n - \mathbb{E} \log p_n = -\frac{1}{2} \sum_{k,j=1}^{n} c_{kj} \left[X(t_k) X(t_j) - K(t_k, t_j) \right],$$

where $(c_{kj}) = \big(K_1(t_k, t_j)\big)^{-1} - \big(K_0(t_k, t_j)\big)^{-1}$, difference between the two matrix inverses.

Let

$$\eta_n = \sum_{k,j=1}^{n} c_{kj} \big[X(t_k) X(t_j) - K(t_k, t_j) \big].$$

By Lemma 3.3.8, we can also write

$$\eta_n = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \varphi_n(\lambda, \mu) \Psi(d\lambda, d\mu),$$

where

$$\varphi_n(\lambda, \mu) = \sum_{k,j=1}^{n} c_{kj} \big(e^{i\langle \lambda, t_k \rangle} - 1\big)\big(e^{-i\langle \mu, t_j \rangle} - 1\big).$$

Since

$$\big(K_0(t_k, t_j)\big)(c_{kj})\big(K_1(t_k, t_j)\big) = b(s, t),$$

for $s, t = t_1, \cdots, t_n$, it can be verified that

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} \big(e^{-i\langle \lambda, s \rangle} - 1\big)\big(e^{i\langle \mu, t \rangle} - 1\big) \varphi_n(\lambda, \mu) F_0(d\lambda) F_1(d\mu) = b(s, t),$$

which is rewritten as

$$\langle \varphi_n,\, \varphi_0 \rangle_{F_0 \times F_1} = b(s,t), \qquad s, t \in T_n = \{t_1, \cdots, t_n\},$$

where

$$\varphi_0(\lambda, \mu) = \big(e^{i\langle \lambda, s\rangle} - 1\big)\big(e^{-i\langle \mu, t\rangle} - 1\big).$$

For any $m \leq n$, $\varphi_m(\lambda, \mu)$ coincides with the projection of $\varphi_n(\lambda, \mu) \in L_{F_0 \times F_1}(T_n \times T_n)$ onto the subspace $L_{F_0 \times F_1}(T_m \times T_m)$, so that

$$\|\varphi_n - \varphi_m\|^2_{F_0 \times F_1} = \|\varphi_n\|^2_{F_0 \times F_1} - \|\varphi_m\|^2_{F_0 \times F_1} \geq 0.$$

So $\|\varphi_n\|^2_{F_0 \times F_1}$ is nondecreasing. Since $\sum_{j,k} b(t_j, t_k)^2 < \infty$, which follows from the equivalence of P_0 and P_1 [see (2.20) of Ibragimov and Rozanov (1978), page 81], then $\lim_{n \to \infty} \|\varphi_n\|^2_{F_0 \times F_1}$ exists. So

$$\lim_{n \to \infty} \varphi_n(\lambda, \mu) = \varphi(\lambda, \mu) \in L_{F_0 \times F_1}(D \times D).$$

Hence, by the bounded convergence theorem,

$$\eta_n \to \eta = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(\lambda, \mu) \Psi(d\lambda, d\mu), \quad \text{as } n \to \infty.$$

It follows from (1.33) of Ibragimov and Rozanov (1978), page 73 that the limit of $\mathbb{E} \log p_n$ exists for equivalent Gaussian measures, so denote the limit as $\log C$. Also, analogous to the proof of Theorem 3.3.3, we have $p(\omega) = \lim_{n \to \infty} p_n(\omega)$. Hence

$$\log p(\omega) = \log C - \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(\lambda, \mu) \Psi(d\lambda, d\mu),$$

and the desired result follows. □

Proof of Theorem 3.3.12

On one hand, suppose P_0 and P_1 are equivalent. Then by Theorem 3.3.10, $b(s,t)$ can be written as: $\forall\, s,t \in D$,

$$b(s,t) = \iint_{\mathbb{R}^d\times\mathbb{R}^d} \big(e^{-i\langle\lambda,s\rangle} - 1\big)\big(e^{i\langle\mu,t\rangle} - 1\big)\varphi(\lambda,\mu)f_0(\lambda)f_1(d\mu)d\lambda d\mu$$

for some $\varphi \in L_{F_0\times F_1}(D\times D)$.

Define

$$g(\lambda,\mu) = \varphi(\lambda,\mu)f_0(\lambda)f_1(\mu),$$

then

$$\begin{aligned}&\iint_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|g(\lambda,\mu)|^2}{f_0(\lambda)f_1(\mu)}\,d\lambda\,d\mu\\ &= \iint_{\mathbb{R}^d\times\mathbb{R}^d} |\varphi(\lambda,\mu)|^2 f_0(\lambda)f_1(\mu)d\lambda\,d\mu < \infty.\end{aligned}$$

i.e. g satisfies (3.25).

On the other hand, suppose that (3.24) and (3.25) hold. We take

$$\varphi(\lambda,\mu) = \frac{g(\lambda,\mu)}{f_0(\lambda)f_1(\mu)}.$$

By (3.25), we have

$$\begin{aligned}&\iint_{\mathbb{R}^d\times\mathbb{R}^d} |\varphi(\lambda,\mu)|^2 F_0(d\lambda)F_1(d\mu)\\ &= \iint_{\mathbb{R}^d\times\mathbb{R}^d} \frac{|g(\lambda,\mu)|^2}{f_0(\lambda)f_1(\mu)}\,d\lambda\,d\mu < \infty.\end{aligned}$$

i.e. $\varphi \in L_{F_0 \times F_1}(D \times D)$ (or we may take the projection of φ onto $L_{F_0 \times F_1}(D \times D)$).

Note that we can rewrite $b(s,t)$ as

$$b(s,t) = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \big(e^{-i\langle \lambda, s\rangle} - 1\big)\big(e^{i\langle \mu, t\rangle} - 1\big)\varphi(\lambda,\mu) f_0(\lambda) f_1(\mu) d\lambda\, d\mu.$$

That is, (3.22) holds. It follows from Theorem 3.3.10 that P_0 and P_1 are equivalent on the σ-algebra $\mathcal{U}(D)$. □

Proof of Corollary 3.3.14

We write

$$\begin{aligned} b(s,t) &= \int_0^s \int_0^t \frac{\partial^2 b(u,v)}{\partial u \partial v} du\, dv \\ &= \int_0^s \int_0^t \iint_{\mathbb{R}\times\mathbb{R}} e^{-i(\lambda u - \mu v)} \psi(\lambda,\mu)\, d\lambda\, d\mu\, du\, dv \\ &= \iint_{\mathbb{R}\times\mathbb{R}} \big(e^{-i\lambda s} - 1\big)\big(e^{-i\mu t} - 1\big) \frac{\psi(\lambda,\mu)}{\lambda\mu}\, d\lambda\, d\mu. \end{aligned}$$

By (3.26), the function

$$g(\lambda,\mu) = \frac{\psi(\lambda,\mu)}{\lambda\mu}$$

satisfies (3.25) in Theorem 3.3.12; the conclusion follows. □

Proof of Theorem 3.3.16

By Theorem 3.2 of Sottinen and Tudor (2006), we know that under the equivalence of P_0 and P_1, $\{X(t), P_1\}$ has a non-anticipative representation $x(t)$ with respect to $\{X(t), P_0\}$ and $x(t) \in H_F(D)$. According to Lemma 3.2.1, there is an isometric isomorphism ζ from $H_F(D)$ to $L_F(D)$, such that $\zeta X(t) = e^{i\langle \cdot, t\rangle} - 1$, for $t \in D$. By the sample path continuity of $X(t)$, it follows from (4.3) of Kallianpur and Oodaira (1973) that $\zeta x(t)$ can be written as

$$\zeta x(t) = (I+B)\zeta X(t) = (I+B)\big(e^{i\langle \cdot, t\rangle} - 1\big),$$

where I is an identity operator and B is a Volterra operator on $L_F(D)$. Therefore, we can write

$$(I+B)\big(e^{i\langle \cdot, t\rangle} - 1\big) = e^{i\langle \cdot, t\rangle} - 1 + \int_{\mathbb{R}^d} b(\cdot, \lambda)\big(e^{i\langle \lambda, t\rangle} - 1\big)d\lambda,$$

where b is the Volterra kernel corresponding to the Volterra operator B. We thus obtain the representation

$$\begin{aligned} x(t) &= \int_{\mathbb{R}^d} \big(e^{i\langle \lambda, t\rangle} - 1\big)\Phi(d\lambda) + \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} b(\mu, \lambda)\big(e^{i\langle \lambda, t\rangle} - 1\big)d\lambda\Phi(d\mu) \\ &= X(t) + \int_{\mathbb{R}^d}\int_{[-\infty, \lambda]} b(\mu, \lambda)\Phi(d\mu)\big(e^{i\langle \lambda, t\rangle} - 1\big)d\lambda. \end{aligned}$$

□

Proof of Theorem 3.5.1

On one hand, we show that the function satisfying (3.32) belongs to $L_F(D)$. For any $s \in [0, \tau]$, by the bounded convergence theorem, $i\lambda\, e^{i\lambda s}$ is the limit of

$$\frac{\big(e^{i\lambda(s+h)} - 1\big) - \big(e^{i\lambda s} - 1\big)}{h},$$

as $h \to 0$, under the inner product defined on $L_F(D)$; and for $k = 2, \cdots, m-1$, $(i\lambda)^k e^{i\lambda s}$ being limits of the form

$$\lim_{h\to 0} (i\lambda)^{k-1}\, \frac{e^{i\lambda(s+h)} - e^{i\lambda s}}{h}$$

belong to $L_F(D)$. Take $s = 0$, we can show that each polynomial $P(i\lambda) = \sum_{k=1}^{m-1} c_k (i\lambda)^k \in$

$L_F(D)$, where c_k's are real. We can show that $(1+i\lambda)^{m-1}\int_0^\tau(e^{i\lambda t}-1)c_s(t)dt$ is also contained in $L_F(D)$, where

$$c_s(t)=\begin{cases} e^t, & 0\le t\le s,\\ 0, & s<t\le\tau. \end{cases}$$

In fact,

$$\begin{aligned}(1+i\lambda)^{m-1}\int_0^\tau(e^{i\lambda t}-1)c_s(t)dt &= (1+i\lambda)^{m-1}\left[\int_0^s e^{(1+i\lambda)t}dt-\int_0^s e^t dt\right]\\ &= (1+i\lambda)^{m-2}\big[(e^{i\lambda s}-1)e^s+i\lambda(1-e^s)\big]\end{aligned}$$

is contained in $L_F(D)$.

It can also be seen that the linear hull of "step" functions $c_s(t)$, for $s\in[0,\tau]$ is everywhere dense in the space $L^2([0,\tau])$, which is the space of all square-integrable functions $c(t)$, $0\le t\le\tau$ [see Ibragimov and Rozanov (1978), page 30]. Moreover, for any $\varphi_j(\lambda)$, $j=1,2$ of the form

$$(1+i\lambda)^{m-1}\int_0^\tau(e^{i\lambda t}-1)c_j(t)dt,$$

where $c_j(t)$ is a linear combination of step functions $c_s(t)$, we derive from (3.31) that

$$\begin{aligned}\|\varphi_1(\lambda)-\varphi_2(\lambda)\|_F^2 &= \int_{-\infty}^{\infty}|\varphi_1(\lambda)-\varphi_2(\lambda)|^2 f(\lambda)d\lambda\\ &\le c\int_{-\infty}^{\infty}\frac{1}{1+\lambda^2}\left|\int_0^\tau(e^{i\lambda t}-1)(c_1(t)-c_2(t))dt\right|^2 d\lambda\\ &\le c\int_0^\tau|c_1(t)-c_2(t)|^2dt,\end{aligned}$$

where c is a positive constant.

On the other hand, we can get $e^{i\lambda s}-1$ by means of iterated integration from $(1+$

$i\lambda)^{m-2}\big(e^{(1+i\lambda)s} - e^s\big)$ and $\sum_{k=1}^{m-1} c_k (i\lambda)^k$, since

$$\begin{aligned}
&\int_0^t (1+i\lambda)^{m-2}\big(e^{(1+i\lambda)s} - e^s\big)ds \\
&= (1+i\lambda)^{m-3}\big[e^{(1+i\lambda)t} - 1 - (1+i\lambda)(e^t - 1)\big] \\
&= (1+i\lambda)^{m-3}\big(e^{(1+i\lambda)t} - e^t\big) + (1+i\lambda)^{m-3}\big(i\lambda\big)(1 - e^t).
\end{aligned}$$

So it is proved that the closed linear hull of functions $(1+i\lambda)^{m-1}\int_0^\tau (e^{i\lambda t} - 1)c_s(t)dt$ and $(i\lambda)^k$, $1 \le k \le m-1$ forms the space $L_F(D)$. Hence the theorem is proved. □

Proof of Proposition 3.5.3

First of all, we need to prove the following fact: For $l = 0, 1, \cdots, m-2$,

$$(1+i\lambda)^{m-1}\int_0^{t_k} t^l \big(e^{(1+i\lambda)t} - e^t\big)dt, \qquad k = 1, \cdots, n, \tag{3.44}$$

belong to $L_{n,m-1}$. To verify (3.44), let us first check, for $l = 0, 1, \cdots, m-2$,

$$\begin{aligned}
&(1+i\lambda)^{m-1}\int_0^{t_k} t^l e^{(1+i\lambda)t}dt \\
=&(1+i\lambda)^{m-2}\Big[t_k^l e^{(1+i\lambda)t_k} - l\int_0^{t_k} t^{l-1}e^{(1+i\lambda)t}dt\Big] \\
=&(1+i\lambda)^{m-2}t_k^l e^{(1+i\lambda)t_k} - l\Big[(1+i\lambda)^{m-3}t_k^{l-1}e^{(1+i\lambda)t_k} \\
&\quad - (1+i\lambda)^{m-3}(l-1)\int_0^{t_k} t^{l-2}e^{(1+i\lambda)t}dt\Big] \\
=&e^{(1+i\lambda)t_k}\sum_{j=0}^{l}(1+i\lambda)^{m-j-2}\, t_k^{l-j}(-1)^j\, l!/(l-j)! + (-1)^{l+1}l!(1+i\lambda)^{m-l-2}.
\end{aligned} \tag{3.45}$$

As a consequence, we show that

$$
\begin{aligned}
&(1+i\lambda)^{m-1}\int_0^{t_k} t^l\big(e^{(1+i\lambda)t}-e^t\big)dt\\
=&e^{(1+i\lambda)t_k}\sum_{j=0}^{l}(1+i\lambda)^{m-j-2}\,t_k^{l-j}(-1)^j\,l!/(l-j)!+(-1)^{l+1}l!(1+i\lambda)^{m-l-2}\\
&-(1+i\lambda)^{m-1}\int_0^{t_k} t^le^tdt \qquad (3.46)\\
=&e^{t_k}\big(e^{i\lambda t_k}-1\big)\sum_{j=0}^{l}(1+i\lambda)^{m-j-2}\,t_k^{l-j}(-1)^j\,l!/(l-j)!+(-1)^{l+1}l!(1+i\lambda)^{m-l-2}\\
&+e^{t_k}\sum_{j=0}^{l}(1+i\lambda)^{m-j-2}\,t_k^{l-j}(-1)^j\,l!/(l-j)!-(1+i\lambda)^{m-1}\int_0^{t_k}t^le^tdt.
\end{aligned}
$$

It can be seen that the first term on the second equality of (3.46) is in $L_{n,m-1}$. In order to get the result of (3.44), we just need to check

$$
e^{t_k}\sum_{j=0}^{l}t_k^{l-j}(-1)^j\,l!/(l-j)!+(-1)^{l+1}l!=\int_0^{t_k}t^le^tdt,
$$

which is the same as (3.45), with $\lambda=0$. Hence, (3.44) is proved.

Secondly, based on the fact in (3.44), we know that for any constant b_{jk}, $j=0,\cdots,m-2$ and $k=1,\cdots,n$

$$
(1+i\lambda)^{m-1}\sum_{k=1}^{n}\sum_{j=0}^{m-2}b_{jk}\int_0^{t_k}t^j\big(e^{(1+i\lambda)t}-e^t\big)dt\in L_{n,m-1}.
$$

And obviously, $P(i\lambda) = \sum_{k=1}^{m-1} c_k (i\lambda)^k$ is in $L_{n,m-1}$. Then by (3.32), (3.35) and (3.38)

$$
\begin{aligned}
&\|\varphi - P_{n,m-1}\varphi\|_F^2 \\
&\leq \int_{-\infty}^{\infty} f(\lambda) \bigg| (1+i\lambda)^{m-1} \int_0^1 \big(e^{(1+i\lambda)t} - e^t\big) h(t) dt \\
&\quad - (1+i\lambda)^{m-1} \sum_{k=1}^{n} \sum_{j=0}^{m-2} b_{jk} \int_0^{t_k} t^j \big(e^{(1+i\lambda)t} - e^t\big) dt \bigg|^2 d\lambda \\
&\leq \beta \int_{-\infty}^{\infty} \frac{1}{(1+\lambda^2)} \bigg| \bigg(\int_0^1 e^{(1+i\lambda)t} h(t) - \sum_{k=1}^{n} \sum_{j=0}^{m-2} b_{jk} t^j e^{(1+i\lambda)t} I_{\{t \leq t_k\}} dt \bigg) \\
&\quad - \bigg(\int_0^1 e^t h(t) - \sum_{k=1}^{n} \sum_{j=0}^{m-2} b_{jk} t^j e^t I_{\{t \leq t_k\}} dt \bigg) \bigg|^2 d\lambda \\
&\leq 4\beta\pi \int_0^1 \bigg\{ e^t h(t) - \sum_{k=1}^{n} \sum_{j=0}^{m-2} b_{jk} t^j e^t I_{\{t \leq t_k\}} \bigg\}^2 dt \\
&= 4\beta\pi \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} e^{2t} \bigg\{ h(t) - \sum_{l=k}^{n} \sum_{j=0}^{m-2} b_{jl} t^j \bigg\}^2 dt.
\end{aligned}
$$

Referring to Stein (1990), page 861, we get the desired result. □

Proof of Proposition 3.5.5

For $m > 1$, with any constant a_k, $k = 0, \cdots, n$, we have

$$
\|\varphi - P_{n,0}\varphi\|_F^2 \leq \beta \int_{-\infty}^{\infty} (1+\lambda^2)^{-m} \left| \int_0^1 c(t)(e^{i\lambda t} - 1) dt - \sum_{k=0}^{n} a_k (e^{i\lambda t_k} - 1) \right|^2 d\lambda. \quad (3.47)
$$

Similarly to page 862 of Stein (1990), the right side of (3.47) can be obtained by polynomial interpolating quadratures [Krylov (1962), Chapter 6]. Denote

$$
b_k = \int_{t_0}^{t_{m-1}} c(t) \frac{u(t) dt}{(t - t_k) u'(t_k)},
$$

where $k = 0, \cdots, m-1$ and $u(t) = (t - t_0) \cdots (t - t_{m-1})$. From Equation (6.1.9) of Krylov (1962), page 81, we know that for any real function $h(t)$, whose mth derivative is bounded by M,

$$\left| \int_{t_0}^{t_{m-1}} c(t)h(t)dt - \sum_{k=0}^{m-1} b_k h(t_k) \right| \leq \frac{M}{m!} \int_{t_0}^{t_{m-1}} |c(t)h(t)|\, dt$$
$$\leq \frac{MC}{m!}(t_{m-1} - t_0)^{m+1}.$$

Applying this bound to the real and imaginary parts of $e^{i\lambda t} - 1$ separately, we will have

$$\left| \int_{t_0}^{t_{m-1}} c(t)\big(e^{i\lambda t} - 1\big)dt - \sum_{k=0}^{m-1} b_k\big(e^{i\lambda t_k} - 1\big) \right| \leq \frac{2C}{m!}|\lambda|^m \,(t_{m-1} - t_0)^{m+1}.$$

The rest of the proof is the same as on page 863 of Stein (1990) □

Proof of Theorem 3.5.6

Similarly to the proof of Theorem 3.5.1, we first show that any function which satisfies (3.40) belongs to $L_F(D)$. For any $s \in D$, by the bounded convergence theorem, $i\lambda_1 e^{i\lambda' s}$ and $i\lambda_2 e^{i\lambda' s}$ are the limits of

$$\frac{1}{h}\left[\big(e^{i(\lambda' s+\lambda_1 h)} - 1\big) - (e^{i\lambda' s} - 1)\right] \quad \text{and} \quad \frac{1}{h}\left[\big(e^{i(\lambda' s+\lambda_2 h)} - 1\big) - (e^{i\lambda' s} - 1)\right],$$

respectively, as $h \to 0$, under the inner product defined on $L_F(D)$. For $k = 0, \cdots, m_1 - 1$, $j = 0, \cdots, m_2 - 1$ and $k = j = 0$ does not hold, $i^{j+k}\lambda_1^j \lambda_2^k e^{i\lambda' s}$, being the limits of the form

$$\lim_{h\to 0} \frac{1}{h}\, i^{j+k-1}\lambda_1^{j-1}\lambda_2^k\big(e^{i(\lambda' s+\lambda_1 h)} - e^{i\lambda' s}\big)$$

or

$$\lim_{h\to 0}\frac{1}{h}\, i^{j+k-1}\lambda_1^j\lambda_2^{k-1}\big(e^{i(\lambda' s+\lambda_2 h)} - e^{i\lambda' s}\big)$$

belong to $L_F(D)$. So each polynomial $P(i\lambda) = \sum_{j=0}^{m_1-1}\sum_{k=0}^{m_2-1} a_{jk} i^{j+k}\lambda_1^j\lambda_2^k \in L_F(D)$, where $a_{00}=0$, a_{jk}'s are real. From the proof of Theorem 6.1, we know that $Q(\lambda)\in L_F(D)$. We can show that

$$(1+i\lambda_1)^{m_1-1}(1+i\lambda_2)^{m_2-1}\int_D (e^{i\lambda' t}-1)c_s(t)dt$$

is also contained in $L_F(D)$, where for $t=(t_1,t_2)\in D$ and $s=(s_1,s_2)\in D$,

$$c_s(t) = \begin{cases} e^{t_1}e^{t_2}, & 0\le t_1\le s_1 \text{ and } 0\le t_2\le s_2, \\ 0, & \text{otherwise.} \end{cases}$$

In fact, for $m_1, m_2 > 1$,

$$\begin{aligned}
&(1+i\lambda_1)^{m_1-1}(1+i\lambda_2)^{m_2-1}\int_D (e^{i\lambda' t}-1)c_s(t)dt \\
&= (1+i\lambda_1)^{m_1-1}(1+i\lambda_2)^{m_2-1}\Big[\int_0^{s_1} e^{(1+i\lambda_1)t_1}dt_1\int_0^{s_2} e^{(1+i\lambda_2)t_2}dt_2 \\
&\qquad - \int_0^{s_1}\int_0^{s_2} e^{t_1}e^{t_2}dt_1dt_2\Big] \\
&= (1+i\lambda_1)^{m_1-2}(1+i\lambda_2)^{m_2-2}\Big[\big(e^{(1+i\lambda_1)s_1}-1\big)\big(e^{(1+i\lambda_2)s_2}-1\big) \\
&\qquad - (1+i\lambda_1)(1+i\lambda_2)(e^{s_1}-1)(e^{s_2}-1)\Big] \\
&= (1+i\lambda_1)^{m_1-2}(1+i\lambda_2)^{m_2-2}\Big[e^{s_1+s_2}(e^{i\lambda' s}-1) - e^{s_1}(e^{i\lambda_1 s_1}-1) - e^{s_2}(e^{i\lambda_2 s_2}-1) \\
&\qquad - (i\lambda_1+i\lambda_2+i^2\lambda_1\lambda_2)(e^{s_1}e^{s_2}-e^{s_1}-e^{s_2}+1)\Big]
\end{aligned}$$

is contained in $L_F(D)$. It can also be seen that the linear hull of "step" functions $c_s(t)$, for $s \in D$ is everywhere dense in the space $L^2([0,\tau]\times[0,\tau])$, generated by square-integrable functions $c(t)$, $t \in [0,\tau]\times[0,\tau]$. Moreover, for any $\varphi_i(\lambda)$, $i = 1, 2$ of the form

$$(1+i\lambda_1)^{m_1-1}(1+i\lambda_2)^{m_2-1}\int_D (e^{i\lambda' t}-1)c_i(t)dt,$$

where $c_i(t)$ is a linear combination of step functions $c_s(t)$,

$$\begin{aligned}
\|\varphi_1(\lambda)-\varphi_2(\lambda)\|_F^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|\varphi_1(\lambda)-\varphi_2(\lambda)|^2 f(\lambda)d\lambda \\
&\le \beta\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{(1+\lambda_1^2)(1+\lambda_2^2)}\Big|\int_D (e^{i\lambda' t}-1)(c_1(t)-c_2(t))dt\Big|^2 d\lambda_1 d\lambda_2 \\
&\le c\int_D |c_1(t)-c_2(t)|^2 dt,
\end{aligned}$$

where c is a positive constant.

On the other hand, we can get $e^{i\lambda' s}-1$ by means of iterated integration from $Q(\lambda)$, $P(i\lambda)$ and $(1+i\lambda_1)^{m_1-2}(1+i\lambda_2)^{m_2-2}e^{s_1+s_2}(e^{i\lambda' s}-1)$. So it is proved that the closed linear hull of functions expressed as (3.40) forms the space $L_F(D)$. Hence the theorem is proved. □

Chapter 4

Conclusion and future work

In this dissertation, we propose a family of anisotropic space-time intrinsically stationary Gaussian model. We study the smoothness and fractal properties of the model, all in terms of the parameters of the models explicitly, and obtain the criteria for two Gaussian measures to be equivalent in intrinsically stationary random fields. We derive upper and lower bounds for the prediction errors of the model, and investigate its asymptotically optimal predictions. This work is of importance in studying the statistical properties of non-stationary Gaussian random fields.

There are some open problems for future work. First, how to estimate the parameter (now a vector $(H_1, \cdots .H_N)$). Guo, Lim and Meerschaert (2009) develop the local Whittle method to simultaneously estimate the Hurst index $H = (H_1, H_2)$ of self-similarity, based on the asymptotic properties of the spectral density of a stationary and anisotropic random field near the origin. They prove the consistency of the local Whittle estimators of the long memory parameters and obtain the asymptotic distribution of the local Whittle estimators. The main goal here is to construct consistent estimators for $(H_1, \cdots .H_N)$ that are applicable in various space-time modeling and to study their asymptotic normality. It is a great challenge for the realization, since multiple smoothness parameters have to be estimated simultaneously in our model which has non-stationary and anisotropic properties. In order to estimate all the parameters of an anisotropic random field model, one has to work with a multivariate random field **X** as defined by (2.25).

Second, Gaussian random fields whose spectral densities are described by a power law model provide a simple and flexible class of models for inferences. Because most of these random fields are nonstationary, the extensive results available on equivalence of Gaussian measures for stationary models [see Theorems 10 and 13 in Chapter III of Ibragimov and Rozanov (1978)] do not apply to them. Basically, the result of Theorem 13 states that for mean 0 stationary Gaussian processes on the interval [0, T] with two possible spectral densities f_0 and f_1, if $f_0(w)(1+w^2)^n$ is bounded away from 0 and ∞ and b is the difference of the two covariance functions viewed as a function on $[0,T]^2$, then the measures are equivalent if and only if

$$\int_0^T \int_0^T \left\{ \frac{\partial^{2n} b(s,t)}{\partial s^n \partial t^n} \right\}^2 ds\, dt < \infty.$$

It seems to be an open problem whether analogous results still hold for Gaussian random fields with stationary increments.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Adler, R. J. (1981), *The Geometry of Random Fields.* Wiley, New York.

[2] Adler, R. J. and Taylor, J. E. (2007), *Random Fields and Geometry.* Springer, New York.

[3] Anderes, E. B. and Stein, M. L. (2008), Estimating deformations of isotropic Gaussian random fields on the plane. *Ann. Statist.* **36**, 719–741.

[4] Aronszajn, N. (1950), Theory of reproducing kernels. *Trans. Am. Math. Soc.,* **68**, 337–404.

[5] Banerjee, S. and Gelfand, A. E. (2003), On smoothness properties of spatial processes. *J. Multi. Anal.* **84**, 85–100.

[6] Banerjee, S., Gelfand, A. E. and Sirmans, C. F. (2003), Directional rates of change under spatial process models. *J. Amer. Statistical Assoc.* **98**, 946–954.

[7] Berg, C. and Forst, G. (1975), *Potential Theory on Locally Compact Abelian Groups.* Springer-Verlag, New York-Heidelberg.

[8] Biermé, H., Lacaux, C. and Xiao, Y. (2009), Hitting probabilities and the Hausdordorf dimension of the inverse images of anisotropic Gaussian random fields. *Bull. London Math. Soc.* **41**, 253–273.

[9] Bonami, A. and Estrade, A. (2003), Anisotropic analysis of some Gaussian models. *J. Fourier Anal. Appl.* **9**, 215–236.

[10] Calder, C. A. and Cressie, N. (2007), Some topics in convolution-based spatial modeling. In: *Proceedings of the 56th Session of the International Statistics Institute.* Lisbon, Portugal.

[11] Chan, G. and Wood, A. T. A. (2000), Increment-based estimators of fractal dimension for two-dimensional surface data. *Statist. Sinica* **10**, 343–376.

[12] Chan, G. and Wood, A. T. A. (2004), Estimation of fractal dimension for a class of non-Gaussian stationary processes and fields. *Ann. Statist.* **32**, 1222–1260.

[13] Chatterji, S. D. and Mandrekar, V. S. (1978), Equivalence and singularity of Gaussian measures and applications. *Probabilistic Analysis and Related Topics* (Ed. A.T. Barucha Ried), **1**, AP, 169–197.

[14] Constantine, A. G. and Hall, P. (1994), Characterizing surface smoothness via estimation of effective fractal dimension. *J. Roy. Statist. Soc. Ser. B* **56**, 97–113.

[15] Cramér, H. and Leadbetter, M. R. (1967), *Stationary and Related Stochastic Processes.* John Wiley & Sons, Inc., New York.

[16] Cressie, N. (1993), *Statistics for Spatial Data* (rev. ed.). Wiley, New York.

[17] Cressie, N. and Huang, H.-C. (1999), Classes of nonseparable, spatiotemporal stationary covariance functions. *J. Amer. Statist. Assoc.* **94**, 1330–1340.

[18] Davies, S. and Hall, P. (1999), Fractal analysis of surface roughness by using spatial data (with discussion). *J. Roy. Statist. Soc. Ser. B* **61**, 3–37.

[19] de Iaco, S., Meyers, D. E. and Posa, D. (2001), Space-Time analysis using a general product-sum model. *Statist. Probab. Letters* **52**, 21–28.

[20] de Iaco, S., Myers, D. E. and Posa, D. (2002), Nonseparable space-time covariance models: some parametric families. *Math. Geology* **34**, 23–42.

[21] de Iaco, S., Myers, D. E. and Posa, D. (2003), The linear coregionalization model and the product-sum space-time variogram. *Math. Geology* **35**, 25–38.

[22] Du, J (2009), Asymptotic and computational methods in spatial statistics. Ph.D. thesis, Michigan State University.

[23] Eubank, R. L., Smith, P.L. and Smith, P.W. (1981), Uniqueness and eventual uniqueness of optimal designs in some time series models. *Ann. Statist.* **9**, 486–493..

[24] Falconer, K. J. (1990), *Fractal Geometry – Mathematical Foundations and Applications*. Wiley & Sons, New York.

[25] Fuentes, M. (2002), Spectral methods for nonstationary spatial processes. *Biometrika* **89**, 197–210.

[26] Fuentes, M. (2005), A formal test for nonstationarity of spatial stochastic processes. *J. Multi. Anal.* **96**, 30–54.

[27] Gihman, I. I. and Skorohod, A. V. (1974), *The Theory of Stochastic Processes,* vol. 1. Springer-Verlag, Berlin.

[28] Gneiting, T. (2002), Nonseparable, stationary covariance functions for space-time data. *J. Amer. Statist. Assoc.* **97**, 590–600.

[29] Gneiting, T., Kleiber, W. and Schlather, M. (2009), Matérn cross-covariance functions for multivariate random fields. *Preprint*.

[30] Guo, H., Lim, C. and Meerschaert, M. M. (2009), Local whittle estimator for anisotropic random fields. *J. Multi. Anal.* **100**, 993–1028.

[31] Hall, P. and Wood, A. T. A. (1993), On the performance of box-counting estimators of fractal dimension. *Biometrika* **80**, 246–252.

[32] Higdon, D. (2002), Space and space-time modeling using process convolutions. In: *Quantitative Methods for Current Environmental Issues* (Anderson, C., Barnett, V., Chatwin, P.C., El-Shaarawi, A.H., editors), pp. 37–56. Springer-Verlag, New York.

[33] Higdon, D., Swall, J., and Kern, J. (1999), Nonstationary spatial modeling. In: *Bayesian Statistics*, (J. M. Bernardo et al, editors), Vol **6**, pp. 761–768, Oxford University Press, Oxford, U.K.

[34] Ibragimov, I. A. and Rozanov, Y. A. (1978), *Gaussian Random Processes*. Springer, Verlag.

[35] Jones, R. H. and Zhang, Y. (1997), Models for continuous stationary space-time processes. In: *Modelling Longitudinal and Spatially Correlated Data*, (T. G. Gregoire,

D. R. Brillinger, P. J. Diggle, E. Russek-Cohen, W. G. Warren, and R. D. Wolfinger, editors), Lecture Notes in Statist. No. **122**, pp. 289–298, Springer, New York.

[36] Kahane, J.-P. (1985), *Some Random Series of Functions.* 2nd edition, Cambridge University Press, Cambridge.

[37] Kallianpur, G. and Oodaira, H. (1963), The equivalence and singularity of Gaussian measures. *Proc. Sympos. Time Series Analysis.* Wiley, New York, 279–291.

[38] Kallianpur, G. and Oodaira, H. (1973), Non-anticipative representations of equivalent Gaussian processes. *Ann. Probab.* **1**, 104–122.

[39] Kent, J. T. and Wood, A. T. A. (1997), Estimating the fractal dimension of a locally self-similar Gaussian process by using increments. *J. Roy. Statist. Soc. Ser. B* **59**, 679–699.

[40] Kolovos, A., Christakos, G., Hristopulos, D. T. and Serre, M. L. (2004), Methods for generating non-separable spatiotemporal covariance models with potential environmental applications. *Adv. Water Resour.* **27**, 815–830.

[41] Krylov, V. I. (1962), *Approximate Calculation of Integrals.* Macmillan, New York.

[42] Kyriakidis, P. C. and Journel, A. G. (1999), Geostatistical space-time models: a review. *Math. Geology* **31**, 651–684.

[43] Li, Y. and Xiao, Y. (2010), Multivariate operator-self-similar random fields. *Stoch. Process. Appl.*, to appear.

[44] Luan, N. and Xiao, Y. (2010), Spectral conditions for strong local nondeterminism and exact Hausdorff measure of ranges of Gaussian random fields. Submitted.

[45] Ma, C. (2003a), Families of spatio-temporal stationary covariance models. *J. Statist. Plan. Infer.* **116**, 489–501.

[46] Ma, C. (2003b), Spatio-temporal stationary covariance models. *J. Multivariate Anal.* **86**, 97–107.

[47] Ma, C. (2004), Spatial autoregression and related spatio-temporal models. *J. Multivariate Anal.* **88**, 152–162.

[48] Ma, C. (2005a), Spatio-temporal variograms and covariance models. *Adv. in Appl. Probab.* **37**, 706–725.

[49] Ma, C. (2005b), A class of stationary random fields with a simple correlation structure. *J. Multivariate Anal.* **94**, 313–327.

[50] Ma, C. (2007), Stationary random fields in space and time with rational spectral densities. *IEEE Trans. Inform. Th.* **53**, 1019–1029.

[51] Ma, C. (2008), Recent developments on the construction of spatio-temporal covariance models. *Stoch. Environ. Res. Risk Assess.* **22** suppl. 1, 39–47.

[52] Major, P. (1981), *Multiple Wiener-Itô Integrals*. Lecture Notes in Math. **849**, Springer-Verlag, Berlin.

[53] Meerschaert, M. M., Wang, W. and Xiao, Y. (2010), Fernique-type inequalities and moduli of continuity of anisotropic Gaussian random fields. *Submitted*.

[54] Paciorek, C. J. and Schervish, M. J. (2006), Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics* **17**, 483–506.

[55] Parzen, E. (1963), Probability density functionals and reproducing kernel Hilbert spaces. *Proc. Sympos. Time Series Analysis*. John Wiley & Sons, Inc., New York, 155–169.

[56] Potthoff, J. (2010), Sample properties of random fields III: differentiability. *Comm. Stoch. Anal.* **4**, 335–353.

[57] Rogers, C. A. (1970), *Hausdorff Measures*, Cambridge University Press.

[58] Sacks, J. and Ylvisaker, D. (1966), Designs for regression problems with correlated errors. *Ann. Math. Statist* **37**, 66–89.

[59] Sacks, J. and Ylvisaker, D. (1968), Designs for regression problems with correlated errors: Many parameters. *Ann. Math. Statist* **39**, 49–69.

[60] Sacks, J. and Ylvisaker, D. (1970), Designs for regression problems with correlated errors. III. *Ann. Math. Statist* **41**, 2057–2074.

[61] Schmidt, A. and O'Hagan, A. (2003), Bayesian inference for nonstationary spatial covariance structure via spatial deformation. *J. Roy. Statist. Soc. Ser. B* **65**, 745–758.

[62] Sottinen, T. and Tudor, C. A. (2006), On the equivalence of multiparameter Gaussian processes. *J. Theoret. Probab.* **19**, 461–485.

[63] Stein, M. L. (1988), Asymptotically efficient prediction of a random field with a misspecified covariance function. *Ann. Statist.* **16**, 55–63.

[64] Stein, M. L. (1990), Uniform asymptotic optimality of linear predictions of a random field using an incorrect second-order structure. *Ann. Statist.* **18**, 850–872.

[65] Stein, M. L. (1999a), Predicting random fields with increasing dense observations. *Ann. Appl. Probab.* **9**, 242–273.

[66] Stein, M. L. (1999b), *Interpolation of Spatial Data: Some Theory for Kriging.* Springer, New York.

[67] Stein, M. L. (2005), Space-time covariance functions. *J. Amer. Statist. Assoc.* **100**, 310–321.

[68] Wahba, G. (1971), On the regression problem of Sacks and Ylvisaker. *Ann. Math. Statist.* **42**, 1035–1053.

[69] Wahba, G. (1974), Regression design for some equivalence classes of kernels. *Ann. Statist.* **2**, 925–934.

[70] Xiao, Y. (2007), Strong local nondeterminism of Gaussian random fields and its applications. In: *Asymptotic Theory in Probability and Statistics with Applications*, (T.-L. Lai, Q.-M. Shao and L. Qian, editors), pp. 136–176, Higher Education Press, Beijing.

[71] Xiao, Y. (2009), Sample path properties of anisotropic Gaussian random fields. In: *A Minicourse on Stochastic Partial Differential Equations,* (D. Khoshnevisan and F. Rassoul-Agha, editors), Lecture Notes in Math. **1962**, pp. 145–212, Springer, New York.

[72] Xiao, Y. (2011), Properties of strong local nondeterminism and local times of stable random fields. In: *Seminar on Stochastic Analysis, Random Fields and Applications VI*, pp. 279–310. Progr. Probab., 63, Birkhäuser, Basel.

[73] Xue, Y. and Xiao, Y. (2011a), Fractal and smoothness properties of anisotropic Gaussian models. *Frontiers Math. China,* to appear.

[74] Xue, Y. and Xiao, Y. (2011b), Criteria for equivalence and asymptotically optimal predictions of intrinsically stationary random fields. To be submitted.

[75] Yadrenko, M. I. (1983), *Spectral Theory of Random Fields.* Optimization Software, New York.

[76] Yaglom, A. M. (1957), Some classes of random fields in n-dimensional space, related to stationary random processes. *Th. Probab. Appl.* **2**, 273–320.

[77] Yaglom, A. M. (1987), *Correlation Theory of Stationary and Related Random Functions*, Vol. 1. Springer-Verlag, New York.

[78] Zhu, Z. and Stein, M. L. (2002), Parameter estimation for fractional Brownian surfaces. *Statist. Sinica* **12**, 863–883.