Fractal and Smoothness Properties of Space-Time Gaussian Models

Yun Xue and Yimin Xiao *
Department of Statistics and Probability
Michigan State University

December 1, 2009

Abstract

Spatio-temporal models are widely used for inference in statistics and many applied areas. In such contexts interests are often in the fractal nature of the sample surfaces and in the rate of change of the spatial surface at a given location in a given direction. In this paper we apply the theory of Yaglom (1957) to construct a large class of space-time Gaussian models with stationary increments, establish bounds on the prediction errors and determine the smoothness properties and fractal properties of this class of Gaussian models. Our results can be applied directly to analyze the stationary space-time models introduced by Cressie and Huang (1999), Gneiting (2002) and Stein (2005), respectively.

Running head: Fractal and Smoothness Properties of Space-Time Gaussian Models.

2000 AMS Classification numbers: 62M40, 62M30, 60G15, 60G17, 60G60, 28A80.

Key words: Space-time models, anisotropic Gaussian fields, prediction error, mean square differentiability, sample path differentiability, Hausdorff dimension.

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,

*Research partially supported by NSF grant DMS-0706728.

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 space-time models whose variability changes with location and/or time. Henceforth, the construction of nonstationary space-time models has become an attractive topic and several approaches have been developed recently. These include to deform the coordinates of an isotropic and stationary random field to obtain a rich class of nonstationary random fields [see Schmidt and O'Hagan (2003), Anderes and Stein (2008)], or to use 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 paper, 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 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 paper are very different from those obtained by deformation from any isotropic random field. It is also worth to mention that the method in this paper 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 paper is organized as follows. In Section 2 we construct a class of space-time Gaussian models with stationary increments 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 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 proofs of our results.

We end the Introduction with some notation. Throughout this paper, instead of using space-time parameter space $\mathbb{R}^d \times \mathbb{R}$, we take the parameter space as $\mathbb{R}^N$ or $\mathbb{R}^N_+ = [0,\infty)^N$. 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 parameter, $t \in \mathbb{R}^N$ 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).

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.

2 Anisotropic Gaussian models with stationary increments

We consider a special class of *intrinsic random functions*; namely, space-time models with stationary increments. We will further restrict ourselves to Gaussian random fields for which powerful general Gaussian principles can be applied. Many of the results in this paper 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 paper. One can find some information for stable random fields in Xiao (2008).

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

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

where Q 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:

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

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 Q, $W(d\lambda)$ is a centered complex-valued Gaussian random measure which is independent of Y and satisfies

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

for all Borel sets $A, B \subseteq \mathbb{R}^N$. The spectral measure F is called the *control measure* of W. Since the linear term $\langle Y, t\rangle$ in (2.3) will not have any effect on the problems considered in this paper, we will from now on assume $Y = 0$. This is equivalent to assuming $Q = 0$ 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 $v(h)$, 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 Gaussian random fields with stationary increments 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 $Y = \{Y(t), t \in \mathbb{R}^N\}$ is a centered stationary Gaussian random field, it can be represented as $Y(t) = \int_{\mathbb{R}^N} e^{i\langle t,\lambda\rangle}\, W(d\lambda)$. Thus the random field X defined by

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

is Gaussian with stationary increments 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 Y.

In the following, we propose and investigate a class of centered, anisotropic Gaussian random fields with stationary increments, 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) ensures f is a legitimate spectral density function.

Proposition 2.1 *Assume that $f(\lambda)$ is a non-negative measurable function defined on $\mathbb{R}^N$. If*

$$\int_{|\lambda|\le 1} |\lambda|^2 f(\lambda) d\lambda < \infty$$

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 (2.6) are in the following.

Remark 2.2

- 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

$$c_4\left(\sum_{j=1}^{N} a_j\right)^q \le \sum_{j=1}^{N} a_j^q \le c_5\left(\sum_{j=1}^{N} a_j\right)^q$$

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

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

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

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

- It is also possible to consider Gaussian random fields with stationary increments 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, we will deal with them systematically in a subsequent paper.

- Nonstationary 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 the 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 paper.

3 Prediction error of 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 (1999, 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

$$X^*(u) = \mathbf{c}(u)^T \mathbf{\Sigma}^{-1} \mathbf{Z}, \tag{3.1}$$

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 (3.1) 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 $\mathbf{Z}$.

The main result of this section is Theorem 3.1 below, which gives lower and upper bounds for the mean square prediction error for Gaussian random fields with stationary increments which satisfy Condition (C). It shows that, similar to stationary Gaussian field models [cf. Stein (1999)], the prediction error of the models in this paper only depends on the high frequency behavior of the spectral density of X.

Theorem 3.1 *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a centered Gaussian random field valued in $\mathbb{R}$ with stationary increments and spectral density $f(\lambda)$ satisfying (2.6). Then there exist constants $c_6 > 0$ and $c_7 > 0$, such that for all integers $n \geq 1$ and all $u, t^1, \cdots, t^n \in \mathbb{R}^N$,*

$$c_6 \min_{0 \leq k \leq n} \sum_{j=1}^N |u_j - t_j^k|^{2H_j} \leq \mathrm{Var}\big(X(u) | X(t^1), \cdots, X(t^n)\big) \leq c_7 \min_{0 \leq k \leq n} \sum_{j=1}^N \sigma_j\big(|u_j - t_j^k|\big), \tag{3.2}$$

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

$$\sigma_j(r) = \begin{cases} r^{2H_j} & \text{if } 0 < H_j < 1, \\ r^2 |\log r| & \text{if } H_j = 1, \\ r^2 & \text{if } H_j > 1. \end{cases} \tag{3.3}$$

If $H_j < 1$, for $j = 1, \cdots, N$, then the two bounds in (3.2) match. When there is some $H_j > 1$, that means, the random field $X(t)$ is smoother in the j-th direction [see Corollaries 4.3, 4.7 and Theorem 4.8 below], then the upper and lower bounds are not the same any more. This suggests that the prediction error may become bigger as $X(t)$ gets smoother in some directions.

The proof of Theorem 3.1, as well as those of Theorems 4.9, 5.1 and 5.2 reply partially on the following lemma, which provides upper and lower bounds for the variogram of the model.

Lemma 3.2 *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a centered Gaussian random field valued in $\mathbb{R}$ with stationary increments and spectral density $f(\lambda)$ satisfying (2.6). Then there exist constants $c_8 > 0$ and $c_9 > 0$ such that for $s, t \in \mathbb{R}^N$,*

$$c_8 \sum_{j=1}^N \sigma_j\big(|s_j - t_j|\big) \leq \mathbb{E}\big(X(s) - X(t)\big)^2 \leq c_9 \sum_{j=1}^N \sigma_j\big(|s_j - t_j|\big), \tag{3.4}$$

where the function σ_j is defined in (3.3).

The upper bound in (3.4) implies that X has a version whose sample functions are almost surely continuous. Throughout this paper, without loss of generality, we will assume that the sample function $t \mapsto X(t)$ is almost surely continuous.

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 fields. See Cramér and Leadbetter (1967), Alder (1981), Stein (1999), Banerjee and Gelfand (2003), Adler and Taylor (2007). In this section we provide explicit criteria for the mean square and sample path differentiability, for the models introduced in Section 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 random fields with stationary increments, let us first recall the definition of mean square directional derivatives.

Definition 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 $C(s,t) = \mathbb{E}\big[X(t)X(s)\big]$ 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

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

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

Theorem 4.2 *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a centered Gaussian random field valued in $\mathbb{R}$ with stationary increments, then the mean square partial derivative $X_j'(t)$ exists for all $t \in \mathbb{R}^N$ if and only if the limit*

$$\lim_{h,k\to 0} \frac{v(he_j) + v(ke_j) - v\big((h-k)e_j\big)}{hk} \tag{4.1}$$

exists. 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 4.3 *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a centered Gaussian random field valued in $\mathbb{R}$ with stationary increments and 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*

$$\beta_j\left(\gamma - \sum_{i=1}^{N} \frac{1}{\beta_i}\right) > 2, \tag{4.2}$$

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

Assume condition (4.1) of Theorem 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 $C_j^{(h)}(s,t)$ and $C_j(s,t)$ 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

$$C_j^{(h)}(s,t) = \frac{v(\Delta + he_j) + v(\Delta - he_j) - 2v(\Delta)}{2h^2}, \tag{4.3}$$

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

Theorem 4.4 *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a centered Gaussian random field valued in $\mathbb{R}$ with stationary increments. Suppose that all second-order partial derivatives of the variogram $v(t)$ exist. Then the covariance function of $X_j'(t)$ is given by*

$$C_j(s,t) = \frac{1}{2} v_j''(s-t), \tag{4.4}$$

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 (4.3). □

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

$$\begin{aligned}\mathrm{Cov}\big(X(s), X'_j(t)\big) &= \lim_{h\to 0}\frac{1}{2h}\Big\{v(t+he_j) - v(t) + v(\Delta) - v(\Delta - he_j)\Big\} \\ &= \frac{1}{2}\big(v'_j(t) + v'_j(\Delta)\big),\end{aligned} \tag{4.5}$$

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 $Y(t)$ is a stationary Gaussian field with mean square partial derivative $Y'_j(t)$, then $Y(t)$ and $Y'_j(t)$ are uncorrelated. That means, the level of the stationary random field at a particular location is uncorrelated with the partial derivative in any direction at that location. However, this is not always true for nonstationary random fields.

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

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

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

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

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 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 $u \in \mathcal{S}_N = \{t \in \mathbb{R}^N : |t| = 1\}$*

$$X(t + hu) = X(t) + hu^T \nabla_X(t) + r(t, hu), \tag{4.6}$$

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

In other words, for all vectors $u \in \mathcal{S}_N$, it is required that

$$\lim_{h \to 0} \mathbb{E}\left(\frac{X(t + hu) - X(t) - hu^T \nabla_X(t)}{h} \right)^2 = 0. \tag{4.7}$$

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 Gaussian random field with stationary increments to be mean square differentiable.

Theorem 4.6 *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a centered Gaussian random field valued in $\mathbb{R}$ with stationary increments. 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 4.6 we obtain

Corollary 4.7 *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a centered Gaussian random field valued in $\mathbb{R}$ with stationary increments and spectral density $f(\lambda)$ satisfying Condition (C). If*

$$\beta_j \left(\gamma - \sum_{i=1}^{N} \frac{1}{\beta_i} \right) > 2 \qquad \text{for every } j = 1, \ldots, N, \tag{4.8}$$

then X is mean square differentiable at every $t \in \mathbb{R}^N$.

4.3 Criterion for sample path differentiability

For many theoretical and applied purposes, one often needs to work with random fields with smooth sample paths. 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 paper, it turns out that under the same condition as Corollary 4.3, the partial derivatives of X are almost surely continuous.

Theorem 4.8 *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be a separable and centered Gaussian random field with values in $\mathbb{R}$. We assume that X has stationary increments and satisfies Condition (C).*
(i). *If*

$$\beta_j\left(\gamma - \sum_{i=1}^{N} \frac{1}{\beta_i}\right) > 2 \qquad (i.e., H_j > 1), \tag{4.9}$$

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 (4.9) 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,*

$$\lim_{h \to 0} \frac{\widetilde{X}(t+hu) - \widetilde{X}(t) - hu^T \nabla_{\widetilde{X}}(t)}{h} = 0 \qquad \text{for all } u \in \mathcal{S}_N \quad \text{and } t \in \mathbb{R}^N. \tag{4.10}$$

If condition (4.9) does not hold for some $j \in \{1, \cdots, N\}$, then $X(t)$ does not have mean square partial derivatives along those directions and $X(t)$ 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) and the methods there are applicable to X with little modification. For completeness, we state the following result which can be proved by using Lemma 3.2 and general Gaussian methods. We omit its proof.

Theorem 4.9 *Let $X = \{X(t), t \in \mathbb{R}^N\}$ be as in Theorem 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*

$$\limsup_{|\varepsilon| \to 0} \frac{\sup_{t \in I,\, s \in [0,\varepsilon]} |X(t+s) - X(t)|}{\sqrt{\varphi(\varepsilon) \log(1 + \varphi(\varepsilon)^{-1})}} \le c_{10}, \tag{4.11}$$

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

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 dimension has 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

$$\mathbf{X}(t) = \big(X_1(t), \ldots, X_p(t)\big), \qquad t \in \mathbb{R}^N, \tag{5.1}$$

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, let's 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 $p = 1$,

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

Also,

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

The fractal properties of fractional Brownian motion 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

$$0 < \overline{H}_1 \leq \overline{H}_2 \leq \cdots \leq \overline{H}_N \leq 1. \tag{5.2}$$

We assume the following conditions.

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

$$c_{11}^{-1} \sum_{j=1}^{N} |s_j - t_j|^{2\overline{H}_j} \leq \mathbb{E}\big[\big(X(t) - X(s)\big)^2\big] \leq c_{11} \sum_{j=1}^{N} |s_j - t_j|^{2\overline{H}_j}. \tag{5.3}$$

(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

$$\mathrm{Var}\,(X(u) \mid X(t)) \geq c_{12} \sum_{j=1}^{N} \big|u_j - t_j\big|^{2\overline{H}_j}. \tag{5.4}$$

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 5.1 and 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 5.1 *Let $\mathbf{X} = \{\mathbf{X}(t), t \in \mathbb{R}^N\}$ be an (N, p)-Gaussian random field defined by (5.1). If the coordinate process X satisfies Condition (D1), then, with probability 1,*

$$\dim \mathbf{X}\big([0,1]^N\big) = \min\left\{p;\, \sum_{j=1}^{N} \frac{1}{\overline{H}_j}\right\}, \tag{5.5}$$

and

$$\dim \mathrm{Gr}\mathbf{X}\big([0,1]^N\big) = \min_{1\le k\le N}\left\{\sum_{j=1}^{k}\frac{\overline{H}_k}{\overline{H}_j} + N - k + (1-\overline{H}_k)p;\, \sum_{j=1}^{N}\frac{1}{\overline{H}_j}\right\}, \tag{5.6}$$

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

Proof The right inequality in (5.3) and Theorem 4.9 show that $\mathbf{X}(t)$ satisfies a uniform Hölder condition on $[0,1]^N$ which, in turn, imply the desired upper bounds in (5.5) and (5.6).

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 (5.3) 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 5.2 *Let $\mathbf{X} = \{\mathbf{X}(t), t \in \mathbb{R}^N\}$ be an (N, p)-Gaussian random field defined by (5.1). 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*

$$\dim \mathbf{X}^{-1}(x) = \min_{1\le k\le N}\left\{\sum_{j=1}^{k}\frac{\overline{H}_k}{\overline{H}_j} + N - k - \overline{H}_k\, p\right\}. \tag{5.7}$$

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 Gaussian random field valued in $\mathbb{R}$ with stationary increments and spectral density $f(\lambda)$ satisfying (2.6). Let $H_1, \ldots, H_N$ be defined by (2.7). Then, by Theorem 3.1 and Lemma 3.2, we see that X satisfies (D1) for all $\overline{H}_j \in (0,\, 1 \wedge H_j)$ $(1 \le j \le N)$. 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 5.1 and 5.2 to derive the following result.

Corollary 5.3 *Let $\mathbf{X} = \{\mathbf{X}(t), t \in \mathbb{R}^N\}$ be a centered Gaussian random field valued in $\mathbb{R}^p$ defined by (5.1). We assume that its coordinate process X has stationary increments with*

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, (5.5) and (5.6) hold with $\overline{H}_j = 1 \wedge H_j$ $(1 \le j \le N)$.*
(ii). *If, in addition, $H_j \le 1$ for all $j = 1, \cdots, N$ and $\sum_{j=1}^{N} \frac{1}{H_j} > p$, then (5.7) 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 jth 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 (5.6) 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 (5.1).

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).

6.1 Stationary covariance models

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

$$C(x,t) = \frac{\sigma^2}{\big(1 + a|t|^{2\alpha}\big)^{\beta N/2}} \exp\left(- \frac{c|x|^{2\gamma}}{\big(1 + a|t|^{2\alpha}\big)^{\beta\gamma}}\right), \tag{6.1}$$

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 continuous in space x and discrete in time 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 5.1 and 5.2.

Proposition 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 (6.1). Then for any $M > 0$, there exist constants $c_{13} > 0$ and $c_{14} > 0$ such that*

$$c_{13}\big(|x-y|^{2\gamma} + |t-s|^{2\alpha}\big) \le \mathbb{E}\big(X(x,t) - X(y,s)\big)^2 \le c_{14}\big(|x-y|^{2\gamma} + |t-s|^{2\alpha}\big) \tag{6.2}$$

and

$$\mathrm{Var}\big(X(x,t)\big|X(y,s)\big) \ge c_{13}\left(|x-y|^{2\gamma} + |t-s|^{2\alpha}\right) \tag{6.3}$$

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

Proposition 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 (6.1), and let $\mathbf{X}$ be its associated (N,p)-random field defined by (5.1). Then, with probability 1,*

$$\dim \mathbf{X}([0,1]^{d+1}) = \min\left\{p;\ \frac{d}{\gamma} + \frac{1}{\alpha}\right\}. \tag{6.4}$$

And if $0 < \alpha \le \gamma < 1$, then

$$\dim \mathrm{Gr}\mathbf{X}([0,1]^{d+1}) = \begin{cases} d+1+(1-\alpha)p & \text{if } p < \frac{1}{\alpha}, \\ d+\frac{\gamma}{\alpha}+(1-\gamma)p & \text{if } \frac{1}{\alpha} \le p < \frac{1}{\alpha}+\frac{d}{\gamma}, \\ \frac{1}{\alpha}+\frac{d}{\gamma} & \text{if } p \ge \frac{1}{\alpha}+\frac{d}{\gamma}. \end{cases} \tag{6.5}$$

If $0 < \gamma \le \alpha < 1$, then

$$\dim \mathrm{Gr}\mathbf{X}([0,1]^{d+1}) = \begin{cases} d+1+(1-\gamma)p & \text{if } p < \frac{d}{\gamma}, \\ \frac{d\alpha}{\gamma}+1+(1-\alpha)p & \text{if } \frac{d}{\gamma} \le p < \frac{1}{\alpha}+\frac{d}{\gamma}, \\ \frac{1}{\alpha}+\frac{d}{\gamma} & \text{if } p \ge \frac{1}{\alpha}+\frac{d}{\gamma}. \end{cases} \tag{6.6}$$

Proposition 6.3 *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 (6.1), 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*

$$\dim \mathbf{X}^{-1}(x) = \begin{cases} d+1-\alpha p & \text{if } p < \frac{1}{\alpha}, \\ d+\frac{\gamma}{\alpha}-\gamma p & \text{if } p \ge \frac{1}{\alpha}, \end{cases} \tag{6.7}$$

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

$$\dim \mathbf{X}^{-1}(x) = \begin{cases} d+1-\gamma p & \text{if } p < \frac{d}{\gamma}, \\ \frac{d\alpha}{\gamma}+1-\alpha p & \text{if } p \ge \frac{d}{\gamma}. \end{cases} \tag{6.8}$$

6.2 Stationary spectral density models

In Section 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 spectral densities of the form

$$f(\lambda) = \left(\sum_{j=1}^{d+1} c_j \left(a_j + |\lambda_j|^2\right)^{\alpha_j}\right)^{-\nu}, \qquad \forall \lambda \in \mathbb{R}^d \times \mathbb{R}, \tag{6.9}$$

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 (6.9) 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 the smoothness properties and the fractal properties.

Proposition 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 spectral density as (6.9).*
(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 models can also be computed using Corollary 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.

7 Proofs

Proof of Proposition 2.1

Since (2.2) is equivalent to $\int_{\mathbb{R}^N} \big(1 \wedge |\lambda|^2\big) f(\lambda) d\lambda < \infty$, and $\int_{|\lambda|\le 1} |\lambda|^2 f(\lambda) d\lambda < \infty$ is given, it is enough for us to show

$$\int_{|\lambda|>1} \frac{d\lambda}{\big(\sum_{j=1}^N |\lambda_j|^{\beta_j}\big)^\gamma} < \infty$$

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 $a > 0$,

$$\int_0^\infty \frac{dx}{(a + x^\beta)^\gamma} = \begin{cases} c_{15}\, a^{-(\gamma - \frac{1}{\beta})} & \text{if } \beta\gamma > 1, \\ +\infty & \text{if } \beta\gamma \le 1. \end{cases} \tag{7.1}$$

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 (7.1) 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 (7.1) $(N-1)$ 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 $\tau \in \{1, \dots, N\}$ 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 (7.1) and integrating $d\lambda_1 \cdots d\lambda_\tau$, we see that the last integral is divergent. This finishes the proof. □

Proof of Lemma 3.2

For any $s, t \in \mathbb{R}^N$, 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

$$\begin{aligned}
\mathbb{E}\big(X(s) - X(t)\big)^2 &\le N \sum_{k=1}^N \mathbb{E}\big(X(\hat{s}_k) - X(\hat{s}_{k-1})\big)^2 \\
&= 2N \sum_{k=1}^N \int_{\mathbb{R}^N} \big(1 - \cos(h_k\lambda_k)\big) f(\lambda) d\lambda \\
&\le 2N \sum_{k=1}^N \int_{|\lambda|\le 1} \big(1 - \cos(h_k\lambda_k)\big) f(\lambda) d\lambda \\
&\qquad + 2Nc_5 \sum_{k=1}^N \int_{|\lambda|>1} \big(1 - \cos(h_k\lambda_k)\big) \frac{d\lambda}{\big(\sum_{i=1}^N |\lambda_i|^{H_i}\big)^{Q+2}} \\
&\triangleq I_1 + I_2.
\end{aligned} \tag{7.2}$$

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

$$I_1 \le 2N\Big(\sum_{k=1}^N h_k^2\Big) \int_{|\lambda|\le 1} |\lambda|^2 f(\lambda) d\lambda \le c_{16} |s-t|^2 \tag{7.3}$$

for some positive and finite constant c_{16}.

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.

$$\begin{aligned}
&\int_{|\lambda|>1} \big(1 - \cos(h_k\lambda_k)\big) \frac{d\lambda}{\big(\sum_{i=1}^N |\lambda_i|^{H_i}\big)^{Q+2}} \\
&\le 2 \int_{\frac{1}{\sqrt{N}}}^\infty \big(1 - \cos(h_k\lambda_k)\big) d\lambda_k \int_{\mathbb{R}^{N-1}} \frac{d\lambda_1 \cdots d\lambda_{k-1} d\lambda_{k+1} \cdots d\lambda_N}{\big(\sum_{i=1}^N |\lambda_i|^{H_i}\big)^{Q+2}} \\
&\qquad + 4 \int_0^1 \big(1 - \cos(h_k\lambda_k)\big) d\lambda_k \int_{\frac{1}{\sqrt{N}}}^\infty d\lambda_{j_0} \int_{\mathbb{R}^{N-2}} \frac{d\lambda_{k,j_0}^\vee}{\big(\sum_{i=1}^N |\lambda_i|^{H_i}\big)^{Q+2}} \\
&\triangleq I_3 + I_4,
\end{aligned} \tag{7.4}$$

where $d\lambda_{k,j_0}^{\vee}$ denotes integration in λ_i $(i \neq k, j_0)$.

By using (7.1) repeatedly [$N-1$ times], we obtain

$$\begin{aligned} I_3 &\le c \int_{\frac{1}{\sqrt{N}}}^{\infty} \frac{1-\cos(h_k\lambda_k)}{|\lambda_k|^{2H_k+1}}\, d\lambda_k \\ &\le c \left(\int_{\frac{1}{\sqrt{N}}}^{\frac{1}{|h_k|}} \frac{h_k^2\lambda_k^2}{\lambda_k^{2H_k+1}} d\lambda_k + \int_{\frac{1}{|h_k|}}^{\infty} \frac{1}{\lambda_k^{2H_k+1}} d\lambda_k \right) \\ &\le c\, \sigma_k(|h_k|), \end{aligned} \tag{7.5}$$

where σ_k is defined as in (3.3).

Similarly, we use (7.1) $N-2$ times to get

$$\begin{aligned} I_4 &\le c \int_0^1 \big(1-\cos(h_k\lambda_k)\big) d\lambda_k \int_{\frac{1}{\sqrt{N}}}^{\infty} \frac{d\lambda_{j_0}}{\big(\lambda_k^{H_k} + \lambda_{j_0}^{H_{j_0}}\big)^{2+\frac{1}{H_k}+\frac{1}{H_{j_0}}}} \\ &\le c \int_0^1 \big(1-\cos(h_k\lambda_k)\big) d\lambda_k \int_{\frac{1}{\sqrt{N}}}^{\infty} \frac{d\lambda_{j_0}}{\lambda_{j_0}^{2H_{j_0}+1+\frac{H_{j_0}}{H_k}}} \\ &\le c\, |h_k|^2. \end{aligned} \tag{7.6}$$

Combining (7.2)–(7.6) yields the upper bound in (3.4).

Next we prove the lower bound in (3.4). By (2.4) and (2.8) we have

$$\mathbb{E}\big(X(s)-X(t)\big)^2 \ge c_4 \int_{|\lambda|>1} \big(1-\cos\langle s-t,\lambda\rangle\big) \frac{d\lambda}{\rho(\lambda)^{Q+2}}, \tag{7.7}$$

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

$$\int_{|\lambda|>1} \big(1-\cos\langle h,\lambda\rangle\big) \frac{d\lambda}{\rho(\lambda)^{Q+2}} \ge c\sigma_j(|h_j|), \tag{7.8}$$

where c is a positive constant.

We only prove (7.8) 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

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

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

$$\int_{|\lambda|>1} \big(1-\cos\langle h,\lambda\rangle\big) \frac{d\lambda}{\rho(\lambda)^{Q+2}} \ge \rho(h)^2 \int_{D(h)} \frac{1-\cos\Big(\sum_{\ell=1}^N h_\ell \rho(h)^{-H_\ell^{-1}} y_\ell\Big)}{\big(\sum_{\ell=1}^N |y_\ell|^{H_\ell}\big)^{Q+2}}\, dy. \tag{7.9}$$

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]

$$\begin{aligned}
&\rho(h)^2 \int_{D(h)} \frac{\Big(\sum_{\ell=1}^N h_\ell \rho(h)^{-H_\ell^{-1}} y_\ell\Big)^2}{\big(\sum_{\ell=1}^N |y_\ell|^{H_\ell}\big)^{Q+2}}\, dy \\
&\geq \rho(h)^2 \int_{\rho(h)^{H_1^{-1}}}^1 h_1^2 \rho(h)^{-\frac{2}{H_1}} y_1^2 dy_1 \underbrace{\int_0^1 \cdots \int_0^1}_{N-1} \frac{dy_2 \cdots dy_N}{\big(\sum_{\ell=1}^N |y_\ell|^{H_\ell}\big)^{Q+2}} \\
&\geq c\,\rho(h)^{2-\frac{2}{H_1}}\, h_1^2 \int_{\rho(h)^{H_1^{-1}}}^1 \frac{y_1^2 dy_1}{\big(y_1^{H_1}\big)^{\frac{1}{H_1}+2}} \\
&= c\,\sigma_1(|h_1|).
\end{aligned} \tag{7.10}$$

This proves (7.8) and hence Lemma 3.2. □

In order to prove Theorem 3.1, we will make use of 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 7.1 *Assume (2.6) is satisfied, then for any fixed constant $T>0$, 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 $t^k \in [-T,T]^N$, we have

$$|g(\lambda)| \leq c_{17}\,|\lambda| \bigg(\int_{\mathbb{R}^N} |g(\xi)|^2 f(\xi) d\xi \bigg)^{1/2} \tag{7.11}$$

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 3.1

First, let's prove the upper bound in (3.2). By Lemma 3.2 we have

$$\begin{aligned}
\mathrm{Var}\Big(X(u)|X(t^1), \cdots, X(t^n)\Big) &\leq \min_{0\leq k\leq n} \mathbb{E}\big(X(u) - X(t^k)\big)^2 \\
&\leq c_9 \min_{0\leq k\leq n} \sum_{j=1}^N \sigma_j\big(|u_j - t_j^k|\big).
\end{aligned} \tag{7.12}$$

In order to prove the lower bound for the conditional variance in (3.2), we denote $r \equiv \min_{0\le k\le n} \sum_{j=1}^N |u_j - t_j^k|^{H_j}$. 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 $a_k \in \mathbb{R}$, $1 \le k \le n$,

$$\mathbb{E}\bigg(X(u) - \sum_{k=1}^n a_k X(t^k)\bigg)^2 \ge c_6\, r^2, \tag{7.13}$$

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 (7.13) can be written as

$$\mathbb{E}\bigg(X(u) - \sum_{k=1}^n a_k X(t^k)\bigg)^2 = \int_{\mathbb{R}^N} \bigg|e^{i\langle u,\lambda\rangle} - 1 - \sum_{k=1}^n a_k (e^{i\langle t^k,\lambda\rangle} - 1)\bigg|^2 f(\lambda)\, d\lambda. \tag{7.14}$$

Hence, we only need to show

$$\int_{\mathbb{R}^N} \bigg|e^{i\langle u,\lambda\rangle} - \sum_{k=0}^n a_k e^{i\langle t^k,\lambda\rangle}\bigg|^2 f(\lambda)\, d\lambda \ge c_6\, r^2, \tag{7.15}$$

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

$$\delta_r(t) = (2\pi)^{-N} \int_{\mathbb{R}^N} e^{-i\langle t,\lambda\rangle} \hat\delta(r^E\lambda)\, d\lambda. \tag{7.16}$$

Since $\min\big\{\sum_{j=1}^N |u_j - t_j^k|^{H_j} : 0 \le k \le n\big\} \ge r$, we have $\delta_r(u - t^k) = 0$ for $k = 0, 1, \cdots, n$. This and (7.16) together imply that

$$\begin{aligned} J :&= \int_{\mathbb{R}^N} \bigg(e^{i\langle u,\lambda\rangle} - \sum_{k=0}^n a_k e^{i\langle t^k,\lambda\rangle}\bigg) e^{-i\langle u,\lambda\rangle} \hat\delta(r^E\lambda)\, d\lambda \\ &= (2\pi)^N \bigg(\delta_r(0) - \sum_{k=0}^n a_k \delta_r(u - t^k)\bigg) \\ &= (2\pi)^N r^{-Q}. \end{aligned} \tag{7.17}$$

Now we split the integral in (7.17) over $\{\lambda : |\lambda| < 1\}$ and $\{\lambda : |\lambda| \ge 1\}$ and denote the two

integrals by I_1 and I_2, respectively. It follows from Lemma 7.1 that

$$\begin{aligned} I_1 &\le \int_{|\lambda|<1} \Big| e^{i\langle u,\lambda\rangle} - \sum_{k=0}^{n} a_k e^{i\langle t^k,\lambda\rangle} \Big| |\hat{\delta}(r^E\lambda)| d\lambda \\ &\le c_{17} \left[\int_{\mathbb{R}^N} \Big| e^{i\langle u,\lambda\rangle} - \sum_{k=0}^{n} a_k e^{i\langle t^k,\lambda\rangle} \Big|^2 f(\lambda)\, d\lambda \right]^{1/2} \int_{|\lambda|<1} |\lambda| |\hat{\delta}(r^E\lambda)| d\lambda \qquad (7.18) \\ &\le c_{18} \left[\mathbb{E}\Big(X(u) - \sum_{k=1}^{n} a_k X(t^k) \Big)^2 \right]^{1/2}, \end{aligned}$$

where the last inequality follows from (7.14) and the boundedness of $\hat{\delta}$.

On the other hand, by the Cauchy-Schwarz inequality and (7.14), we have

$$\begin{aligned} I_2^2 &\le \int_{|\lambda|\ge 1} \Big| e^{i\langle u,\lambda\rangle} - \sum_{k=0}^{n} a_k e^{i\langle t^k,\lambda\rangle} \Big|^2 f(\lambda) d\lambda \int_{|\lambda|\ge 1} \frac{1}{f(\lambda)} |\hat{\delta}(r^E\lambda)|^2\, d\lambda \\ &\le \mathbb{E}\Big(X(u) - \sum_{k=1}^{n} a_k X(t^k) \Big)^2 r^{-Q} \int_{|\lambda|\ge 1} \frac{1}{f(r^{-E}\lambda)} |\hat{\delta}(\lambda)|^2\, d\lambda \qquad (7.19) \\ &= \mathbb{E}\Big(X(u) - \sum_{k=1}^{n} a_k X(t^k) \Big)^2 r^{-2Q-2} \int_{|\lambda|\ge 1} \frac{1}{f(\lambda)} |\hat{\delta}(\lambda)|^2\, d\lambda. \end{aligned}$$

The last integral is convergent thanks to the fast decay of $\hat{\delta}(\lambda)$. Finally, combining (7.17), (7.18) and (7.19), we get

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

Henceforth (7.13) follows, and the theorem was proved because of (7.12) and (7.13). □

Proof of Theorem 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, 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,

$$\begin{aligned} D_{h,k} &= \frac{1}{hk}\Big\{ C(t+he_j, t+ke_j) - C(t, t+ke_j) - C(t+he_j, t) + C(t,t) \Big\} \\ &= \frac{1}{2hk}\Big\{ v(he_j) + v(ke_j) - v\big((h-k)e_j\big) \Big\}. \end{aligned} \qquad (7.20)$$

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 (4.1) holds [thanks to Taylor's

theorem]. On the other hand, if (4.1) holds, then by taking $h = k \to 0$ in (7.20) we see that $\partial v/\partial t_j(0) = 0$. This fact, together with (4.1), 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 finishes the proof of Theorem 4.2. □

Proof of Corollary 4.3

By Theorem 4.2 it amounts to show that $\lim_{h,k\to 0} D_{h,k}$ exists if and only if (4.2) holds. [i.e., $\beta_j\big(\gamma - \sum_{i=1}^N \frac{1}{\beta_i}\big) > 2$.] It follows from (7.20) and (2.4) that

$$D_{h,k} = \int_{\mathbb{R}^N} \frac{1 - \cos\langle he_j, \lambda\rangle - \cos\langle ke_j, \lambda\rangle + \cos\langle (h-k)e_j, \lambda\rangle}{hk} f(\lambda)\, d\lambda. \tag{7.21}$$

To prove the sufficiency of (4.2), we note that for each fixed $\lambda \in \mathbb{R}^N$,

$$\lim_{h,k\to 0} \frac{1 - \cos(h\lambda_j) - \cos(k\lambda_j) + \cos((h-k)\lambda_j)}{hk} = \lambda_j^2 \tag{7.22}$$

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 (4.2) holds. Then, as in the proof of Proposition 2.1, we have

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

This implies $\int_{\mathbb{R}^N} \lambda_j^2 f(\lambda) d\lambda < \infty$. By (7.21), (7.22) 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 (4.2), we assume $\beta_j\big(\gamma - \sum_{i=1}^N \frac{1}{\beta_i}\big) \le 2$. Then, as in the proof of Proposition 2.1, we have

$$\int_{\lambda\in\mathbb{R}^N:|\lambda_j|>1} \frac{\lambda_j^2\, d\lambda}{\big(\sum_{i=1}^N |\lambda_i|^{\beta_i}\big)^\gamma} = \infty. \tag{7.23}$$

We let $h = k \downarrow 0$ and use Fatou's lemma to (7.21) [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 (7.23). So $\lim_{h,k\to 0} D_{h,k}$ does not exist and the proof is finished. □

Proof of Theorem 4.6

If $v(t)$ has continuous second-order partial derivatives, then Theorem 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 (4.7).

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

$$\begin{aligned}
&\mathbb{E}\bigg(\frac{X(t+hu)-X(t)}{h} - u^T\nabla_X(t)\bigg)^2 \\
&= \mathbb{E}\bigg(\frac{X(t+hu)-X(t)}{h} - \sum_{j=1}^N u_j X_j'(t)\bigg)^2 \\
&= \frac{1}{h^2}v(hu) + \mathbb{E}\bigg(\sum_{j=1}^N u_j X_j'(t)\bigg)^2 - \frac{2}{h}\sum_{j=1}^N u_j\Big(\mathbb{E}X(t+hu)X_j'(t) - \mathbb{E}X(t)X_j'(t)\Big) \\
&= \frac{1}{h^2}v(hu) + \mathbb{E}\bigg(\sum_{j=1}^N u_j X_j'(t)\bigg)^2 - \frac{1}{h}\sum_{j=1}^N u_j v_j'(hu).
\end{aligned} \tag{7.24}$$

The last equality in (7.24) follows from (4.5).

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

$$\lim_{h\to 0}\frac{1}{h^2}v(hu) = \lim_{h\to 0}\frac{v(hu)+v(-hu)-2v(0)}{2h^2} = \frac{1}{2}u^T\Omega(0)u, \tag{7.25}$$

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 (7.24), note that for any i, $j = 1, \cdots, N$ and $i\neq j$, and any $l>0$, $m>0$,

$$\begin{aligned}
&\mathbb{E}\bigg(\frac{X(t+le_i)-X(t)}{l}\frac{X(t+me_j)-X(t)}{m}\bigg) \\
=&\frac{1}{lm}\mathbb{E}\Big(X(t+le_i)X(t+me_j) - X(t)X(t+me_j) - X(t+le_i)X(t) + X^2(t)\Big) \\
=&\frac{1}{2lm}\Big(v(le_i)+v(-me_j)-v(le_i-me_j)\Big).
\end{aligned} \tag{7.26}$$

Let $l\to 0$, $m\to 0$, then the last term in (7.26) 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 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 (7.24), 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 (7.25) and (7.26) shows that (7.24) goes to 0, as $h \to 0$. This finishes the proof. □

Proof of Theorem 4.8

Under (4.2), Corollary 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 Kolmorogov'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

$$\mathbb{E}\big[X_j'(s) - X_j'(t)\big]^2 \le c_{20}\,|s-t|^{\eta}. \tag{7.27}$$

Recall that

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

Thanks to (4.2), we derive

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

and

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

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 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

$$\begin{aligned} \mathbb{E}\big(X_j'(s) - X_j'(t)\big)^2 &\le N\sum_{k=1}^{N} \mathbb{E}\big(X_j'(\hat{s}_k) - X_j'(\hat{s}_{k-1})\big)^2 \\ &= 2N\sum_{k=1}^{N}\Bigg\{\int_{|\lambda|\le 1} \lambda_j^2\big(1-\cos(s_k-t_k)\lambda_k\big) f(\lambda) d\lambda \\ &\qquad\qquad + \int_{|\lambda|>1} \lambda_j^2\big(1-\cos(s_k-t_k)\lambda_k\big) f(\lambda) d\lambda\Bigg\} \\ &\le c_{21}\,|s-t|^2 + c_{22}\sum_{k=1}^{N}\int_{|\lambda|>1} \frac{\big(1-\cos(s_k-t_k)\lambda_k\big)\lambda_j^2}{\big(\sum_{i=1}^{N}|\lambda_i|^{H_i}\big)^{Q+2}}\, d\lambda. \end{aligned} \tag{7.28}$$

Now we estimate the last N integrals in (7.28). For simplicity of notation, we only consider the case when $k = j$ [the cases of $k \neq j$ are similar]. Denote $h_k = s_k - t_k$ and, similar to (7.4), (7.5) and (7.6), we derive

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

thanks to $H_k > 1$. Combining this with (7.28) proves (7.27).

It follows from (7.27)) 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

$$\widetilde{X}(t) = X(t_1,\cdots,t_{j-1},0,t_{j+1},\cdots,t_N) + \int_0^{t_j} X_j'(t_1,\cdots,t_{j-1},s_j,t_{j+1},\cdots,t_N)\,ds_j. \tag{7.29}$$

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 (7.29), Theorem 4.4 and (4.5). Since the verification is elementary, we omit the details. This proves Part (i) of Theorem 4.8.

It remains to prove Part (ii) of Theorem 4.8. By applying Part (i) to $j = 1$, we derive that there is 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

$$\widetilde{X^{(2)}}(t) = \widetilde{X^{(1)}}(t_1,0,t_3,\cdots,t_N) + \int_0^{t_j} \widetilde{X^{(1)}}_2'(t_1,s_2,t_3,\cdots,t_N)\,ds_2. \tag{7.30}$$

Then $\frac{\partial\widetilde{X^{(2)}}}{\partial t_1}(t)$ and $\frac{\partial\widetilde{X^{(2)}}}{\partial t_2}(t)$ are almost surely continuous. Repeat 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 (4.10). The proof of Theorem 4.8 is finished.

Proof of Proposition 6.1

By stationarity, we'll have

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

And

$$\begin{aligned}2C(0,0)-2C(x,t) &= 2\sigma^2-\frac{2\sigma^2}{(1+a|t|^{2\alpha})^{\beta N/2}}\exp\Big(-\frac{c|x|^{2\gamma}}{(1+a|t|^{2\alpha})^{\beta\gamma}}\Big)\\ &= 2\sigma^2\frac{(1+a|t|^{2\alpha})^{\beta N/2}-\exp\Big(-\frac{c|x|^{2\gamma}}{(1+a|t|^{2\alpha})^{\beta\gamma}}\Big)}{(1+a|t|^{2\alpha})^{\beta N/2}}.\end{aligned}\tag{7.31}$$

By using Taylor expansion, we can write (7.31) as

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

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

$$c_{24}\big(|x|^{2\gamma}+|t|^{2\alpha}\big)\le 2C(0,0)-2C(x,t)\le c_{25}\big(|x|^{2\gamma}+|t|^{2\alpha}\big)\tag{7.32}$$

for all $(x,t)\in\mathbb{R}^{d+1}$ with $|x|$ and $|t|$ small. Replace x and t in (7.32) with $x-y$ and $t-s$ respectively, then (6.2) follows.

To prove (6.3), 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

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

This proves (6.3). □

Proof of Proposition 6.2

Eq. (6.4) follows from Proposition 6.1 and Theorem 5.1. Then let's prove (6.5), where $0<\alpha\le\gamma\le 1$. By Proposition 6.1 and Theorem 5.1, we get

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

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 $2 \le k \le d+1$. Also $\sum_{j=1}^{d+1} \frac{1}{\overline{H}_j} = \frac{1}{\alpha} + \frac{d}{\gamma}$. Since

$$\begin{aligned} p < \frac{1}{\alpha} \quad &\Longleftrightarrow \quad (\gamma - \alpha)p < \frac{\gamma - \alpha}{\alpha} \quad \Longleftrightarrow \quad S(1) < S \\ &\Longleftrightarrow \quad d + \frac{\gamma}{\alpha} + (1-\gamma)p < d + \frac{1}{\alpha} \quad \Longrightarrow \quad S < \sum_{j=1}^{d+1} \frac{1}{\overline{H}_j}, \end{aligned} \tag{7.33}$$

and

$$p < \frac{1}{\alpha} + \frac{d}{\gamma} \quad \Longleftrightarrow \quad (1-\gamma)p < \frac{1-\gamma}{\alpha} + \frac{d(1-\gamma)}{\gamma} \quad \Longleftrightarrow \quad S < \sum_{j=1}^{d+1} \frac{1}{\overline{H}_j}, \tag{7.34}$$

one can see that (6.5) follows from (7.33) and (7.34).

If $0 < \gamma \le \alpha \le 1$, then $\overline{H}_1 = \cdots = \overline{H}_d = \gamma$ and $\overline{H}_{d+1} = \alpha$. So $S(d+1) = \frac{d\alpha}{\gamma} + 1 + (1-\alpha)p$, $S(k) = d + 1 + (1-\gamma)p \triangleq \tilde{S}$, for $1 \le k \le d$ and $\sum_{j=1}^{d+1} \frac{1}{\overline{H}_j} = \frac{1}{\alpha} + \frac{d}{\gamma}$. Similarly, by comparing these three terms, we will get (6.6). □

Proof of Proposition 6.3

By Proposition 6.1 and Theorem 5.2 we get that when $\frac{1}{\alpha} + \frac{d}{\gamma} < p$, for every $x \in \mathbb{R}^p$, $\mathbf{X}^{-1}(x) = \emptyset$ a.s. And also, when $\frac{1}{\alpha} + \frac{d}{\gamma} > p$, then for any $x \in \mathbb{R}^p$, with positive probability

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

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 $T(1) < T \Leftrightarrow p < \frac{1}{\alpha}$, (6.7) 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

$$\tilde{T} < T(d+1) \iff p < \frac{d}{\gamma},$$

we derive (6.8). The proof is finished. □

References

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

[5] 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.

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

[7] 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.

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

[9] 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.

[10] 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.

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

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

[13] Cressie, N. and Huang, H.-C. (1999), Classes of nonseparable, spatiotemporal stationary covariance functions. *J. Amer. Statist. Assoc.* **94**, 1330–1340.

[14] 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.

[15] 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.

[16] de Iaco, S., Myers, D. E. and Posa, D. (2002), Nonseparable space-time covariance models: some parametric families. *Mathematical Geology* **34**, 23–42.

[17] 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.

[18] Falconer, K. J. (1990), *Fractal Geometry – Mathematical Foundations and Applications*. Wiley & Sons, New York.

[19] Fuentes, M. (2002), Spectral methods for nonstatuionary spatial processes. *Biometrika* **89**, 197–210.

[20] Fuentes, M. (2005), A formal test for nonstationarity of spatial stochastic processes. *J. Multi. Anal.* **96**, 30–54.

[21] Gneiting, T. (2002), Nonseparable, stationary covariance functions for space-time data. *J. Amer. Statist. Assoc.* **97**, 590–600.

[22] Gneiting, T., Kleiber, W. and Schlather, M. (2009), Matérn cross-covariance functions for multivariate random fields. *Preprint*.

[23] Hall, P. and Wood, A. T. A. (1993), On the performance of box-counting estimators of fractal dimension. *Biometrika* **80**, 246–252.

[24] Higdon, D. (2002), Space and space-time modeling using process convolutions. In: *Quantative Methods for Current Environmental Issues* (Anderson, C., Barnett, V., Chatwin, P.C., El-Shaarawi, A.H., editors), pp. 37–56. Springer-Verlag, New York.

[25] 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.

[26] 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.

[27] Kahane, J.-P. (1985), *Some Random Series of Functions*. 2nd edition, Cambridge University Press, Cambridge.

[28] 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.

[29] 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.

[30] Kyriakidis, P. C. and Journe, A. G. (1999), Geostatistical space-time models: a review. *Math. Geology* **31**, 651–684.

[31] Ma, C. (2003a), Families of spatio-temporal stationary covariance models. *J. Statist. Plan. Infer.* **116**, 489–501.

[32] Ma, C. (2003b), Spatio-temporal stationary covariance models. *J. Multivariate Anal.* **86**, 97–107.

[33] Ma, C. (2004), Spatial autoregression and related spatio-temporal models. *J. Multivariate Anal.* **88**, 152–162.

[34] Ma, C. (2005a), Spatio-temporal variograms and covariance models. *Adv. in Appl. Probab.* **37**, 706–725.

[35] Ma, C. (2005b), A class of stationary random fields with a simple correlation structure. *J. Multivariate Anal.* **94**, 313–327.

[36] Ma, C. (2007), Stationary random fields in space and time with rational spectral densities. *IEEE Trans. Inform. Th.* **53**, 1019–1029.

[37] Ma, C. (2008), Recent developments on the construction of spatio-temporal covariance models. *Stoch. Environ. Res. Risk Assess.* **22** suppl. 1, 39–47.

[38] Paciorek, C. J. and Schervish, M. J. (2006), Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics* **17**, 483–506.

[39] 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.

[40] Stein, M. L. (1999), *Interpolation of Spatial Data: Some Theory for Kriging*. Springer, New York.

[41] Stein, M. L. (2005), Space-time covariance functions. *J. Amer. Statist. Assoc.* **100**, 310–321.

[42] 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.

[43] Xiao, Y. (2008), Properties of strong local nondeterminism and local times of stable random fields. To appear in: *Seminar on Stochastic Analysis, Random Fields and Applications VI.*

[44] 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.

[45] Yaglom, A. M. (1957), Some classes of random fields in n-dimensional space, related to stationary random processes. *Th. Probab. Appl.* **2**, 273–320.

[46] Zhu, Z. and Stein, M. L. (2002), Parameter estimation for fractional Brownian surfaces. *Statist. Sinica* **12**, 863–883.