Statistics and Its Interface Volume 11 (2018) 61–77

Testing the mean in multivariate regression using set-indexed Gaussian white noise

Wayan Somayasa* and Herdi Budiman

We propose an asymptotic test method for checking the validity of a multivariate spatial regression that utilizes the distribution model of set-indexed Gaussian white noise. The random set function is obtained as the limit of the partial sums of the vector of observations sampled according to a continuous probability measure (design). It is shown under relatively mild condition that the test which is defined as the integral with respect to the partial sums of the observation converges to an optimal test constructed based on the Cameron-Martin density of the multivariate shifted Gaussian white noise. The optimality of the design under which the experiment was performed is also investigated. We also study the application of the established test procedure to a multivariate real data obtained from a mining industry.

AMS 2000 subject classifications: 60G10, 60G15, 62G08, 62J05.
Keywords and phrases: Multivariate set-indexed Gaussian white noise, Multivariate spatial regression, Set-indexed partial sums process, Model-check, Cameron-Martin density, Signed measure, Optimal design.

1. INTRODUCTION

Modelling spatial data using multivariate regression analysis is extensively applied in earth and environmental sciences. Checking the appropriateness of an assumed regression model is important in the practice before using the model for prediction and uncertainty quantification. It can be conducted by investigating either the vector of the residuals directly, see e.g. Arnold [2], Seber and Lee [24], Christensen [13] and Johnson and Wichern [16], or the empirical processes of the residuals as proposed in Stute [30], Stute and et al. [30] and Stute [32].

The purpose of the present paper is to study the application of p-dimensional set-indexed Gaussian white noise in model check or lack of fit (LOF) test for the mean vector in multivariate spatial regression defined on high dimensional experimental region. In contrast to the classical methods of model diagnostic addressed in the textbooks on regression listed above, in the present paper we define a test statistics which is expressed as an integral with respect to p-dimensional set-indexed partial sums (Cumulative Sums=CUSUM) processes of the observations instead of the residuals, see Section 2. Our new approach will be shown to be more applicable in the practice.

The application of the set-indexed Gaussian white noise in statistical modelling of spatial data has been pioneered by MacNeill [20, 21] who firstly investigated the limit process of the sequence of least squares residuals partial sums processes of univariate polynomial regression. These famous approaches are generalized to the context of univariate spatial regression by MacNeill and Jandhyala [22], and Xie and MacNeill [37] who obtained the limit process as a functional of the set-indexed Brownian sheet. Bischoff and Somayasa [12] and Somayasa and et al. [26] derived the limit process in the spatial case by applying the geometric method proposed in Bischoff [7, 8]. These results can be used to establish asymptotic test of Kolmogorov-Smirnov and Cramér-von Mises type for model check and boundary detection problems. Recently, Wellner [35] established a likelihood ratio test based on the Cameron-Martin density formula of set-indexed Gaussian white noise derived in Lifshits [19].

The study has been extended to model-check for multivariate spatial regression with correlated responses by Somayasa and et al. [27, 28, 29] by considering the multidimensional partial sums process of the vector of least squares residuals. This technique is however restricted in the application since the limits of the Kolmogorov-Smirnov and Cramér-von Mises functionals of the processes are mathematically not tractable. Simulation must be developed for approximating the quantiles of the test statistics. In the present paper we show our test procedure not only distribution free but also asymptotically optimal in some sense.

Furthermore, to our knowledge the limit of the sequence of the partial sums processes of the residuals studied in the literatures mentioned above were obtained under an equidistance experimental design or a so-called regular lattice only. It is well known that regular lattice coincides asymptotically with Lebesgue measure, cf. [20, 21, 22, 37, 12]. Conversely, given a probability measure on a line, Bischoff [7] and Bischoff and Miller [10, 11] proposed a design technique based on that measure such that the sequence of the corresponding designs with finite sample converges in some sense to such a probability measure. This sampling strategy can be adopted in the practice in case the practitioners can not or will not sample equidistantly. Practical example of this problem is frequently encountered in mining industry in that the Engineers for economic, technical, ecological or geographical reasons avoid to sample equidistantly. In this work we

*Corresponding author.

propose sampling strategy by determining the design points in the spatial perspective based on a continuous probability measure defined on the experimental region by incorporating the technique proposed in [7, 10, 11]. However, we wonder whether such a design results in optimal decision in the comparison with regular lattice.

The organization of the rest paper is as follows. Section 2 formulates the model and the hypothesis under study. In Section 3 we investigate the limiting distribution of the test statistic when the hypotheses are true. In Section 4 we derive the Neyman-Pearson test and try to investigate the optimal property of the test. In Section 5 we investigate the optimality of a design of experiment by studying the limit power function of the associated test. The finite sample behavior of the test is studied by simulation in Section 6. Application of the proposed method to a real data is discussed in Section 7. At the end of this work we present conclusion and remark for future work. Proofs are postponed to the Appendices.

2. THE MODEL, DESIGN STRATEGY AND THE HYPOTHESES

To explain the problem in more detail let us consider a p-variate nonparametric spatial regression

$$\mathbf{Y}(\mathbf{x}) = \mathbf{g}(\mathbf{x}) + \mathcal{E}(\mathbf{x}),\ \mathbf{x} \in D := \times_{j=1}^{d}[a_j, b_j] \subset \mathbb{R}^d, \tag{2.1}$$

where $\mathbf{Y} := (Y_i)_{i=1}^p$ is the vector of random observations, $\mathbf{g} := (g_i)_{i=1}^p : D \to \mathbb{R}^p$ is the true-unknown regression function defined on D and $\mathcal{E} := (\varepsilon_i)_{i=1}^p$ is unobserved vector of random errors defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, say, with $\mathbf{E}(\mathcal{E}(\mathbf{x})) = \mathbf{0} \in \mathbb{R}^p$ and $Cov(\mathcal{E}(\mathbf{x})) = \mathbf{\Sigma}$, for every $\mathbf{x} \in D$, where $\mathbf{\Sigma}$ is assumed in this paper to be unknown and positive definite. For fixed $n_1 \geq 1, \dots, n_d \geq 1$, let $\mathbf{Y}$ be observed over an experimental design

$$\Xi_{\mu;n_1\cdots n_d} := \{\mathbf{t}_{n_1j_1\cdots n_dj_d} := (t_{n_1j_1}, \dots, t_{n_dj_d})^\top : 1 \leq j_k \leq n_k,\ 1 \leq k \leq d\} \subset D,$$

for a given probability measure μ on $(D, \mathcal{B}(D))$, where the points $\mathbf{t}_{n_1j_1\cdots n_dj_d} \in \Xi_{\mu;n_1\cdots n_d}$ are determined by generalizing the method due to [7, 10, 11] as follows. First we construct a partition $\{t_{n_11}, t_{n_12}, \dots, t_{n_1n_1}\}$ on $[a_1, b_1]$ based on the equation

$$F_\mu(t_{n_1j_1}, b_2, \dots, b_d) = j_1/n_1,\ 1 \leq j_1 \leq n_1,$$

where F_μ is the distribution function of μ defined on D which is assumed in this paper to be continuous, increasing on D and factorized as

$$F_\mu(t_1, t_2, \cdots, t_d) = \Pi_{k=1}^d F_{k\mu}(t_k),$$

where $F_{k\mu}$ is continuous and increasing on $[a_k, b_k]$, for $k = 1, \dots, d$. Next for a fixed j_1 we partition the interval $[a_2, b_2]$ by solving the equation

$$F_\mu(t_{n_1j_1}, t_{n_2j_2}, b_3, \dots, b_d) = j_1j_2/(n_1n_2),\ 1 \leq j_2 \leq n_2.$$

The obtained points $\{t_{n_21}, t_{n_22}, \dots, t_{n_2n_2}\}$ which constitutes a partition on $[a_2, b_2]$ that correspond with a fixed $t_{n_1j_1} \in [a_1, b_1]$ are uniquely determined. Similarly, for fixed j_1 and j_2, with $1 \leq j_1 \leq n_1$ and $1 \leq j_2 \leq n_2$ we move forward to develop the corresponding partition $\{t_{n_31}, t_{n_32} \dots, t_{n_3n_3}\}$ on $[a_3, b_3]$ by solving

$$F_\mu(t_{n_1j_1}, t_{n_2j_2}, t_{n_3j_3}, b_4, \dots, b_d) = j_1j_2j_3/(n_1n_2n_3),\ 1 \leq j_3 \leq n_3.$$

The similar manner is applied for the remaining intervals $[a_k, b_k]$, for $k = 4, 5, \dots, d$. In general we construct the partition $\{t_{n_k1}, t_{n_k2}, \dots, t_{n_kn_k}\}$ on $[a_k, b_k]$ that corresponds with the point $(t_{n_1j_1}, \dots, t_{n_{k-1}j_{k-1}})$, for $k = 4, 5, \dots, d$, by solving the equation

$$F_\mu(t_{n_1j_1}, \dots, t_{n_kj_k}, b_{k+1}, \dots, b_d) = \frac{\Pi_{u=1}^k j_u}{\Pi_{u=1}^k n_u}.$$

By this sampling method $\Xi_{\mu;n_1\cdots n_d}$ is not necessarily an equidistance experimental design or a regular lattice. If μ is the uniform probability measures on D with the distribution function

$$F_\mu(t_1, \dots, t_d) := \frac{1}{|D|}\Pi_{j=1}^d(t_j - a_j),\ (t_1, \dots, t_d) \in D,$$

where $|D|$ is the volume of D, then we get the equidistance experimental design with the experimental condition $(t_{n_1j_1}, \dots, t_{n_dj_d})$, where $t_{n_kj_k} = a_k + (b_k - a_k)\frac{j_k}{n_k}$, for $k = 1, \dots, d$.

Let $P_{n_1\cdots n_d}$ be a discrete probability measure on $\mathcal{B}(D)$ associated with $\Xi_{\mu;n_1\cdots n_d}$, defined by

$$P_{n_1\cdots n_d}(B) := \frac{1}{\Pi_{k=1}^d n_k}\sum_{j_1=1}^{n_1}\cdots\sum_{j_d=1}^{n_d}\delta_{\mathbf{t}_{n_1j_1\dots n_dj_d}}(B),\ B \in \mathcal{B}(D),$$

where $\delta_{t_{n_1j_1\dots n_dj_d}}$ is the one point measure in the point $(t_{n_1j_1}, \dots, t_{n_dj_d})$ which is frequently called Dirac measure in the literatures, defined on $\mathcal{B}(D)$ as

$$\delta_{t_{n_1j_1\dots n_dj_d}}(B) = \left\{\begin{array}{lll} 1 & ; & (t_{n_1j_1}, \dots, t_{n_dj_d}) \in B \\ 0 & ; & (t_{n_1j_1}, \dots, t_{n_dj_d}) \notin B \end{array}\right. .$$

We notice that $P_{n_1\cdots n_d}$ can also be written as

$$P_{n_1\cdots n_d}(B) := \frac{1}{\Pi_{k=1}^d n_k}\sum_{j_1=1}^{n_1}\cdots\sum_{j_d=1}^{n_d}\mathbf{1}_B(F_\mu^{-1}(\Pi_{k=1}^d j_k/\Pi_{k=1}^d n_k)),\ B \in \mathcal{B}(D),$$

where $\mathbf{1}_B$ is the indicator function of B. Let $F_{n_1\cdots n_d}$ be the associated distribution function of $P_{n_1\cdots n_d}$. Then by this sampling strategy we get the property that $F_{n_1\cdots n_d}$ converges uniformly to F_μ as $n_1, \dots, n_d$ simultaneously large. That is

$$\|F_{n_1\cdots n_d} - F_\mu\|_\infty := \sup_{(t_1,\ldots,t_d)\in D} |F_{n_1\cdots n_d}(t_1,\ldots,t_d) - F_\mu(t_1,\ldots,t_d)| \to 0,$$

which by the Portmanteau theorem (cf. Billingsley [6], pp. 18–19) immediately implies $P_{n_1\cdots n_d}$ converges in Distribution to μ. We note that all limits obtained in this work are for $n_1,\ldots,n_d$ simultaneously go to infinity, otherwise it will be stated in some way. Such a sampling scheme will be of our consideration throughout the work. The probability measure μ according to which $\Xi_{\mu;n_1\cdots n_d}$ is constructed will be called a design, cf. [7, 10, 11]. A design μ is called optimal for testing LOF if it maximizes the power of the test. In this paper we shall investigate the optimality property of a given design μ with the associated continuous and nondecreasing probability distribution F_μ in term of the limit power function of the proposed test. Under our assumption μ is absolutely continuous with respect to the Lebesque measure λ^d on D.

Let $\mathbf{V}$ be a finite dimensional spaces defined by $\mathbf{V} := [w_1,\ldots,w_q,w_{q+1},\ldots,w_m]$, and let $\mathbf{W} := [w_1,\ldots,w_q]$ be a subset of $\mathbf{V}$, where $w_1,\ldots,w_q,w_{q+1},\ldots,w_m$ are known regression functions which are assumed to be linearly independent as functions in $L_2(D,\mu)$, with $q \le m$, thereby $L_2(D,\mu)$ is the space of squared integrable functions on D with respect to μ. The product of p copies of $L_2(D,\mu)$ is denoted by $L_2^p(D,\mu)$, that is $L_2^p(D,\mu) := \times_{i=1}^p L_2(D,\mu)$. The common framework of LOF test for the mean of $\mathbf{Y}$ falls into the problem of testing the hypothesis that $\mathbf{g} \in \mathbf{W}^p$ while observing $\mathbf{g} \in \mathbf{V}^p$ (cf. [2, 13]), where $\mathbf{W}^p$ and $\mathbf{V}^p$ are the product of p copies of $\mathbf{W}$ and $\mathbf{V}$, respectively. Suppose that $\mathbf{g}$ can be decomposed as $\mathbf{g} = \mathbf{g}_1 + \mathbf{g}_2$ with $\mathbf{g}_1 := (g_{1i})_{i=1}^p \in \mathbf{W}^p$ and $\mathbf{g}_2 := (g_{2i})_{i=1}^p \in \mathbf{V}^p \cap (\mathbf{W}^p)^C$. Then the problem of testing $H_0 : \mathbf{g} \in \mathbf{W}^p$ while observing $\mathbf{g} \in \mathbf{V}^p$ can be handled by testing that of

(2.2)
$$H_0 : \mathbf{g}_2 \equiv \mathbf{0} \text{ against } H_1 : \mathbf{g}_2 \equiv \mathbf{f}_1, \text{ for some } \mathbf{f}_1 \in \mathbf{V}^p \cap (\mathbf{W}^p)^C.$$

Upon observing Model 2.1 over $\Xi_{\mu;n_1\cdots n_d}$ we get an array of independent p-dimensional vector of observations $\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}) = (\mathbf{Y}(\mathbf{t}_{n_1 j_1\cdots n_d j_d}))_{j_1=1,\cdots,j_d=1}^{n_1,\cdots,n_d}$ that satisfies the regression model

$$\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}) = \mathbf{g}(\Xi_{\mu;n_1\cdots n_d}) + \mathcal{E}(\Xi_{\mu;n_1\cdots n_d}), \tag{2.3}$$

where $\mathcal{E}(\Xi_{\mu;n_1\cdots n_d}) := (\mathcal{E}(\mathbf{t}_{n_1 j_1\cdots n_d j_d}))_{j_1=1,\cdots,j_d=1}^{n_1,\cdots,n_d}$ is an array of independent and identically distributed p-dimensional random errors such that (s.t.)

$$\mathbf{E}(\mathcal{E}(\mathbf{t}_{n_1 j_1\cdots n_d j_d})) = \mathbf{0} \text{ and } Cov(\mathcal{E}(\mathbf{t}_{n_1 j_1\cdots n_d j_d})) = \mathbf{\Sigma},$$

for $1 \le j_1 \le n_1,\ldots,1 \le j_d \le n_d$. For convenience we write $\mathbf{Y}(\mathbf{t}_{n_1 j_1\cdots n_d j_d})$, $\mathbf{g}(\mathbf{t}_{n_1 j_1\cdots n_d j_d})$ and $\mathcal{E}(\mathbf{t}_{n_1 j_1\cdots n_d j_d})$ throughout this paper simply as $\mathbf{Y}_{j_1\cdots j_d}$, $\mathbf{g}_{j_1\cdots j_d}$ and $\mathcal{E}_{j_1\cdots j_d}$, respectively. In contrast to the classical method studied in [2, 24, 13, 16] where the inference procedure was derived under normal distribution of $\mathcal{E}_{j_1\cdots j_d}$, in this work we aim to establish an asymptotic procedure for which the normality assumption can be ignored.

A reasonable statistic for testing (2.2) is defined by

$$\mathcal{J}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d})) := \int_D (\mathbf{\Sigma}^{-1/2}\mathbf{f}_1)^\top d\mathbf{\Sigma}^{-1/2}\mathbf{S}_{n_1\cdots n_d}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d})), \tag{2.4}$$

where for every $B \in \mathcal{B}(D)$,

$$\mathbf{S}_{n_1\cdots n_d}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}))(B) := \frac{1}{\sqrt{n_1\cdots n_d}} \sum_{j_1=1}^{n_1}\cdots\sum_{j_d=1}^{n_d} \mathbf{1}_B(\mathbf{t}_{j_1\cdots j_d})\mathbf{Y}_{j_1\cdots j_d},$$

is the p-dimensional partial sums process of the vector of observations indexed by $\mathcal{B}(D)$. This is actually the vectorial version of the univariate partial sums process defined e.g. in [1, 23, 17, 37]. It is well understood that the integral in (2.4) is defined as $\int_D \mathbf{u}^\top d\mathbf{v} := \sum_{i=1}^p \int_D u_i\, dv_i$, for every $\mathbf{u} := (u_i)_{i=1}^p$ and $\mathbf{v} := (v_i)_{i=1}^p$ provided the integrals are in some sense well defined. By the definition it can be immediately seen that the value of $\mathcal{J}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}))$ will be large when the observations appear as realization of the model with $\mathbf{g}_2 \not\equiv \mathbf{0}$. This is because of the appearance of a positive term which is asymptotically rational to $\|\mathbf{\Sigma}^{-1}\mathbf{f}_1\|^2_{L_2^p(D,\mu)}$ when the sample support H_1. In other word the larger the distance of the model from H_0 the greater the value of $\mathcal{J}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}))$. Therefore it is reasonable to reject H_0 for large value of $\mathcal{J}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}))$.

We notice that for each $\omega \in \Omega$, the set function $\mathbf{S}_{n_1\cdots n_d}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}))(\omega)$ constitutes a signed measure on $\mathcal{B}(D)$. Hence the integrals involved in the statistic $\mathcal{J}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}))$ can be interpreted path-wise as the integral of a function in $L_2(D,\mu)$ with respect to a signed measure. The reader is referred to [14], pp. 121–153 for the notion of the integral with respect to signed measure.

Remark 2.1. The problem of testing $H_0 : \mathbf{g} \in \mathbf{W}^p$ is equivalent with that of testing $H_0 : \frac{1}{n_1\cdots n_d}\mathbf{g} \in \mathbf{W}^p$ for all $n_k \ge 1$. On the other hand, the convergence componentwise of $\mathbf{\Sigma}^{-1/2}\mathbf{S}_{n_1\cdots n_d}(\frac{1}{\sqrt{n_1\cdots n_d}}\mathbf{g}(\Xi_{\mu;n_1\cdots n_d}))$ to $\mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{g}}(\cdot) \in \mathcal{C}^p(\mathcal{A}_0)$, where $\varphi_{\mathbf{g}}(A) := \int_A \mathbf{g}\, d\mu$, is useful for analyzing the limiting power function of the test. Hence without altering the test problem we observe the localized version of (2.3) defined by

(2.5)
$$\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}) := \frac{1}{\sqrt{n_1\cdots n_d}}\mathbf{g}(\Xi_{\mu;n_1\cdots n_d}) + \mathcal{E}(\Xi_{\mu;n_1\cdots n_d}).$$

Therefore rather than using the test statistics $\mathcal{J}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}))$, we consider its localized version $\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}))$ given by

$$\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d})) := \int_D \mathbf{f}_1^\top \mathbf{\Sigma}^{-1/2} d\mathbf{\Sigma}^{-1/2}\mathbf{S}_{n_1\cdots n_d}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}))$$

in defining a non randomized test for the hypothesis (2.2) as

$$\eta_{\mu;n_1\cdots n_d}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d})) := \begin{cases} 1\,; & \mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d})) \geq k \\ 0\,; & otherwise \end{cases},$$

where k satisfies $\mathbf{E}_{H_0}(\eta_{\mu;n_1\cdots n_d}) = \alpha$, for a pre-determined $\alpha \in (0,1)$. Power function plays important role for evaluating the performance of a test.

3. LIMITING DISTRIBUTIONS

In this section we derive the limit of the statistic $\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}))$ by applying the multivariate invariance principle presented in Theorem A.2 in the Appendix. Without loss of generality we assume $\{w_1,\ldots,w_q,w_{q+1},\ldots,w_m\}$ are orthogonal as functions in $L_2(D,\mu)$, so that $\mathbf{W} \perp (\mathbf{V}\cap\mathbf{W}^C)$ and $\mathbf{W}^p \perp (\mathbf{V}^p\cap(\mathbf{W}^p)^C)$ as well.

The following theorem gives the asymptotic test of size α for testing the more general hypotheses

$$(3.1)\qquad H_0 : \mathbf{g}_2 \equiv \mathbf{f}_0 \text{ vs. } H_1 : \mathbf{g}_2 \equiv \mathbf{f}_1, \quad \text{for some } \mathbf{f}_0,\mathbf{f}_1 \in \mathbf{V}^p\cap(\mathbf{W}^p)^C, \text{ with } \mathbf{f}_0 \neq \mathbf{f}_1.$$

Theorem 3.1. *Let $\mathbf{g} = \mathbf{g}_1 + \mathbf{g}_2$, with $\mathbf{g}_1 \in \mathbf{W}^p$ and $\mathbf{g}_2 \in \mathbf{V}^p\cap(\mathbf{W}^p)^C$. Suppose that for $i = 1,\ldots,p$, g_{1i} and g_{2i} are continuous with respect to (w.r.t.) the usual Euclidean distance and have bounded variation on D in the sense of Hardy (see Definition F.4 in the Appendix). Then an asymptotic test of size α for testing (3.1) will reject H_0, if and only if*

$$\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d})) \geq \Phi^{-1}(1-\alpha)\|\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0)\|_\mu^{(p)} + \langle\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0),\mathbf{\Sigma}^{-1/2}\mathbf{f}_0\rangle_\mu^{(p)},$$

where Φ is the cumulative distribution function of the standard normal distribution.

Remark 3.2. In the case of the hypothesis $H_0 : \mathbf{g}_2 \equiv \mathbf{0}$ against $H_1 : \mathbf{g}_2 \equiv \mathbf{f}_1$, both statistics $\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}))$ and $\mathcal{J}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}))$ converge under H_0 to the same distribution model, that is $N(0,(\|\mathbf{\Sigma}^{-1/2}\mathbf{f}_1\|_\mu^{(p)})^2)$. Suppose that $t^*_{\Xi_{\mu;n_1\cdots n_d}}$ is the value of the statistic $\mathcal{J}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}))$ computed on a given sample, then the p-value of the test for such simple hypothesis is calculated by using the equation

$$p - value = 1 - \Phi\left(\frac{t^*_{\Xi_{\mu;n_1\cdots n_d}}}{\|\mathbf{\Sigma}^{-1/2}\mathbf{f}_1\|_\mu^{(p)}}\right).$$

The asymptotic power function of the test is presented in the following corollary.

Corollary 3.3. *Let $\Psi_{\mu;n_1\cdots n_d} : \mathbf{V}^p\cap(\mathbf{W}^p)^C \to (0,1)$ be the power function of $\eta_{\mu;n_1\cdots n_d}$ at level α defined by*

$$\Psi_{\mu;n_1\cdots n_d}(\mathbf{f}) := P\{\omega\in\Omega : \mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d})(\omega)) \geq k | \mathbf{g}_2 \equiv \mathbf{f}\},$$

for $\mathbf{f} \in \mathbf{V}^p\cap(\mathbf{W}^p)^C$, where $k := \Phi^{-1}(1-\alpha)\|\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0)\|_\mu^{(p)} + \langle\mathbf{f}_1-\mathbf{f}_0,\mathbf{\Sigma}^{-1}\mathbf{f}_0\rangle_\mu^{(p)}$. Then under the assumption of Theorem 3.1, $\Psi_{\mu;n_1\cdots n_d}(\mathbf{f})$ converges point wise to

$$\Psi_\mu(\mathbf{f}) := 1 - \Phi\left(\Phi^{-1}(1-\alpha) - \frac{\langle\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0),\mathbf{\Sigma}^{-1/2}(\mathbf{f}-\mathbf{f}_0)\rangle_\mu^{(p)}}{\|\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0)\|_\mu^{(p)}}\right).$$

Remark 3.4. Let $\mathbf{Y}_{(n_1\cdots n_d)}$, and $\mathbf{X}_{(n_1\cdots n_d)}$ be the $(n_1\cdots n_d)\times p$-dimensional matrix of observations and the $(n_1\cdots n_d)\times m$-dimensional design matrix defined by

$$\mathbf{Y}_{(n_1\cdots n_d)} := (vec(Y_1(\Xi_{\mu;n_1\cdots n_d})),\ldots,vec(Y_p(\Xi_{\mu;n_1\cdots n_d})))$$
$$\mathbf{X}_{(n_1\cdots n_d)} := (vec(w_1(\Xi_{\mu;n_1\cdots n_d})),\ldots,vec(w_m(\Xi_{\mu;n_1\cdots n_d}))),$$

where *vec* denotes the well-known vec operator defined as follows. Let $\mathbf{M} : (\mathbf{m}_1,\cdots,\mathbf{m}_j,\cdots,\mathbf{m}_p)$ be a $q\times p$ dimensional matrix whose j-th column is $\mathbf{m}_j \in \mathbb{R}^q$. Then $vec(\mathbf{M}) := (\mathbf{m}_1^\top,\cdots,\mathbf{m}_j^\top,\cdots,\mathbf{m}_p^\top)^\top \in \mathbb{R}^{pq}$. A consistent estimator of $\mathbf{\Sigma}$ is given by the $p\times p$-dimensional matrix $\widehat{\mathbf{\Sigma}}_{n_1\cdots n_d}$, where

$$\widehat{\mathbf{\Sigma}}_{n_1\cdots n_d} := \frac{1}{n_1\cdots n_d}\mathbf{Y}^\top_{(n_1\cdots n_d)} pr_{\mathcal{C}(\mathbf{X}_{(n_1\cdots n_d)})^\perp}\mathbf{Y}_{(n_1\cdots n_d)},$$

cf. Arnold [3]. Thereby $pr_{\mathcal{C}(\mathbf{X}_{(n_1\cdots n_d)})^\perp}$ is the orthogonal projector onto the orthogonal complement of the column space of $\mathbf{X}_{(n_1\cdots n_d)}$.

Remark 3.5. It is important to note that computational difficulties appear in the practice for testing using our established test procedure, because the test statistic was expressed as the integral with respect to signed measure induced by $\mathbf{S}_{n_1\cdots n_d}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}))$ indexed by large family of subsets $\mathcal{A}$ which includes all open, closed as well as convex subsets of D. Fortunately as noted in [1, 23] the one dimensional invariance principle holds true for the much smaller family of subsets of D, that is $\mathcal{I}^d := \{\Pi_{j=1}^d[a_j,t_j] : a_j < t_j \leq b_j,\ j = 1,\ldots,d\}$. By the similar argument, Theorem D.1 can also be shown to hold true under the family $\mathcal{I}^d$. By this reason, the implementation of the test in the application is conducted by using a computer program that considers the family $\mathcal{I}^d$. This family belongs to the so-called Vapnik-Červonenkis Classes (VCC), which in general satisfies the prerequisites needed in the multivariate invariance principle. Further, by taking into account the family $\mathcal{I}^d$ as the index in the calculation of the test statistic, the random function $\mathbf{S}_{n_1\cdots n_d}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}))(\Pi_{j=1}^d[a_j,t_j])$ is written by $\mathbf{S}_{n_1\cdots n_d}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}))(t_1,\ldots,t_d)$ for brevity. It can be shown that the last can be regarded as a random continuous function on D with respect to the usual Euclidean distance leading us to the conclusion that the integral of any function of bounded variation with respect to $\mathbf{S}_{n_1\cdots n_d}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}))$ coincides path-wise with the Riemann-Stieltjes integral in the sense of Stroock [25], pp. 7–16.

The steps how to conduct asymptotic LOF test for the mean of multivariate spatial regression model in the practice according to our method is summarized in the following algorithm:

1. Determine the design μ under which the experimental design is constructed;
2. Construct the experimental conditions according to F_μ following the sampling strategy introduced in Section 2;
3. Compute the test statistic $\mathcal{J}(\mathbf{Y}(\Xi_{\mu;n_1\cdots n_d}))$;
4. Compute the p-value of the test by using the formula given in Remark 3.2;
5. Draw decision whether or not to reject H_0.

4. NEYMAN-PEARSON TEST

Our purpose in this section is to establish an optimal test and to use the optimality criterion defined in Bischoff and Miller [10] in order to show that the test $\eta_{\mu;n_1\cdots n_d}$ is asymptotically optimal.

Under the condition of Theorem 3.1 the sequence of the partial sums processes $\{\mathbf{\Sigma}^{-1/2}\mathbf{S}_{n_1\cdots n_d}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}))(A) : A \in \mathcal{A}_0\}$, $n_1 \geq 1, \ldots, n_d \geq 1$ converges to a signal plus noise model given by

$$\mathcal{Y} = \{\mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{g}}(A) + \mathbf{Z}_\mu(A) : \ A \in \mathcal{A}_0\},$$

with $\varphi_{\mathbf{g}} = \varphi_{\mathbf{g}_1} + \varphi_{\mathbf{g}_2}$ as the deterministic signal and the set-indexed Gaussian white noise $\mathbf{Z}_\mu$ as the random noise. We called $\mathcal{Y}$ asymptotic model that corresponds to $\{\mathbf{\Sigma}^{-1/2}\mathbf{S}_{n_1\cdots n_d}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}))(A) : \ A \in \mathcal{A}_0\}$. Hypothesis of the form

$$H_0 : \varphi_{\mathbf{g}_2} \equiv \varphi_{\mathbf{f}_0} \text{ against } H_1 : \varphi_{\mathbf{g}_2} \equiv \varphi_{\mathbf{f}_1}, \quad \text{for some } \varphi_{\mathbf{f}_0}, \varphi_{\mathbf{f}_1} \in \mathbf{V}_\varphi^p \cap (\mathbf{W}_\varphi^p)^C \tag{4.1}$$

is said the asymptotic hypothesis that corresponds to Hypothesis 3.1, where $\mathbf{V}_\varphi^p$ and $\mathbf{W}_\varphi^p$ are the product of p copies of $\mathbf{V}_\varphi$ and $\mathbf{W}_\varphi$, respectively, with $\mathbf{V}_\varphi := [\varphi_{w_1}, \ldots \varphi_{w_p}, \ldots, \varphi_{w_m}]$ and $\mathbf{W}_\varphi := [\varphi_{w_1}, \ldots \varphi_{w_p}]$. Furthermore, the problem of testing (4.1) when observing $\mathcal{Y}$ is called the asymptotic version of that of testing (3.1) when observing $\{\mathbf{\Sigma}^{-1/2}\mathbf{S}_{n_1\cdots n_d}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}))(A) : \ A \in \mathcal{A}_0\}$.

Our concern is to establish Neyman-Pearson test procedure for (4.1) when the model $\mathcal{Y}$ is observed. For that we need the reproducing kernel Hilbert space (RKHS) of the components $\mathbf{Z}_\mu$, denoted by $\mathcal{H}_{Z_\mu^{(i)}}$, given by

$$\mathcal{H}_{Z_\mu^{(i)}} := \left\{h : \ h(A) = \int_A \ell \, d\mu, \ \ell \in L_2(D, \mu), \ A \in \mathcal{A}_0\right\}.$$

The space $\mathcal{H}_{Z_\mu^{(i)}}$ which is furnished with the inner product and norm defined by

$$\langle h_1, h_2\rangle_{\mathcal{H}_{Z_\mu^{(i)}}} := \langle \ell_1, \ell_2\rangle_\mu := \int_D \ell_1\ell_2 d\mu, \ \ell_1, \ell_2 \in L_2(D, \mu)$$

$$\|h\|_{\mathcal{H}_{Z_\mu^{(i)}}} := \|\ell\|_\mu := \sqrt{\int_D |\ell|^2 d\mu}, \ \ell \in L_2(D, \mu),$$

is decisive for our result, see also [26]. We get the result immediately by applying either Theorem 4.1 in [19] or Definition 1 in [5] the RKHS of $\mathbf{Z}_\mu$ which is given by $\mathcal{H}_{\mathbf{Z}_\mu} = \times_{i=1}^p \mathcal{H}_{Z_\mu^{(i)}}$. It becomes clear that $\mathbf{V}_\varphi^p$ as well as $\mathbf{W}_\varphi^p$ are subsets of $\mathcal{H}_{\mathbf{Z}_\mu}$. Analogously, inner product and norm on $\mathcal{H}_{\mathbf{Z}_\mu}$ are defined respectively by

$$\langle \varphi_{\mathbf{u}}, \varphi_{\mathbf{v}}\rangle_{\mathcal{H}_{\mathbf{Z}_\mu}} := \sum_{i=1}^p \langle \varphi_{v_i}, \varphi_{u_i}\rangle_{\mathcal{H}_{Z_\mu^{(i)}}} = \sum_{i=1}^p \langle v_i, u_i\rangle_\mu$$

$$\|\varphi_{\mathbf{u}}\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}} := \sum_{i=1}^p \|\varphi_{u_i}\|^2_{\mathcal{H}_{Z_\mu^{(i)}}} = \sum_{i=1}^p \|u_i\|^2_\mu = (\|\mathbf{u}\|^{(p)}_\mu)^2.$$

Definition 4.1. Let $\psi_0(\mathcal{Y})$ and $\psi_1(\mathcal{Y})$ be the density functions of $\mathcal{Y}$ when the samples are considered under H_0 and H_1, respectively. A non randomized Neyman-Pearson test for 4.1 when the asymptotic model $\mathcal{Y} = \mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{g}_1+\mathbf{g}_2} + \mathbf{Z}_\mu$ is observed is a function $\eta_\mu : \mathcal{C}^p(\mathcal{A}_0) \to \{0, 1\}$ defined by

$$\eta_\mu(\mathcal{Y}) = \begin{cases} 1 \ ; & \frac{\psi_0(\mathcal{Y})}{\psi_1(\mathcal{Y})} \leq k \\ 0 \ ; & otherwise \end{cases},$$

where k is a constant that satisfies $\mathbf{E}_{H_0}(\eta_\mu) = \alpha$, for a predetermined $\alpha \in (0, 1)$.

Since the Neyman-Pearson test by Theorem 3.2.1 in [18] leads us to an optimal test, Definition 4.1 suggest that in order to obtain an optimal test procedure for (4.1) we need to derive the density formulas of the process $\mathcal{Y}$ under H_0 as well as under H_1. Definition 4.2 below defines some important notations.

Definition 4.2. For any $\mathbf{h} := (h_i)_{i=1}^p \in \mathcal{C}^p(\mathcal{A}_0)$ the distribution of $\mathbf{h} + \mathbf{Z}_\mu$ on $(\mathcal{C}^p(\mathcal{A}_0), \mathcal{B}(\mathcal{C}^p(\mathcal{A}_0)))$ is denoted by $P^{\mathbf{h}}_{\mathbf{Z}_\mu}$, defined as $P^{\mathbf{h}}_{\mathbf{Z}_\mu}(B) := P_{\mathbf{Z}_\mu}(B - \mathbf{h})$, for every Borel set $B := \times_{i=1}^p B_i \in \mathcal{B}(\mathcal{C}^p(\mathcal{A}_0))$. The function $\mathbf{h}$ is called a shift. If $\mathbf{h} \in \mathcal{H}_{\mathbf{Z}_\mu}$, then $\mathbf{h}$ is called an admissible shift, cf. [19], pp. 34.

The following theorem presents a most powerful test of size α for testing (4.1). The rejection region is derived based on the Cameron-Martin density of the shifted measure $P^{\mathbf{h}}_{\mathbf{Z}_\mu}$ with respect to $P_{\mathbf{Z}_\mu}$, see Theorem D.1 in the appendix. As a comparison study, Wellner [35] applied the Cameron-Martin formula for signal plus Gaussian white noise on the space $\mathcal{C}([-c, c])$ with control measure the Lebesque measure and the signal a monotone function on $[-c, c]$ in establishing likelihood ratio test for testing the signal. The cameron-Martin formula for the Slepian processes on $\mathcal{C}([0, 1])$ has been investigated in [9].

Theorem 4.3. *A most powerful test of size* α *for the hypotheses* $H_0 : \varphi_{\mathbf{g}_2} \equiv \varphi_{\mathbf{f}_0}$ *against* $H_1 : \varphi_{\mathbf{g}_2} \equiv \varphi_{\mathbf{f}_1}$ *for some*

$\varphi_{\mathbf{f}_0}, \varphi_{\mathbf{f}_1} \in \mathbf{V}^p_\varphi \cap (\mathbf{W}^p_\varphi)^C$ *with* $\varphi_{\mathbf{f}_1} \not\equiv \varphi_{\mathbf{f}_0}$*, will reject* H_0 *if and only if*

$$\int_D (\mathbf{f}_1 - \mathbf{f}_0)^\top \mathbf{\Sigma}^{-1/2}\, d\mathcal{Y} \geq \Phi^{-1}(1-\alpha)\|\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0)\|^p_\mu + \langle \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0), \mathbf{\Sigma}^{-1/2}\mathbf{f}_0\rangle^{(p)}_\mu.$$

Furthermore, the corresponding power function of this test is given by

$$\Upsilon_\mu(\varphi_{\mathbf{f}}) = 1 - \Phi\left(\Phi^{-1}(1-\alpha) - \frac{\langle \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0), \mathbf{\Sigma}^{-1/2}(\mathbf{f} - \mathbf{f}_0)\rangle^{(p)}_\mu}{\|\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0)\|^{(p)}_\mu}\right),$$

for every $\varphi_{\mathbf{f}} \in \mathbf{V}^p_\varphi \cap (\mathbf{W}^p_\varphi)^C$*, where* Φ *is the cumulative distribution function of the standard normal distribution.*

Corollary 4.4. *The asymptotic test based on the statistics* $\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}))$ *defined in Theorem 3.1 is an asymptotically optimal test of size* α.

Proof. Since the sequence $\eta_{\mu;n_1\cdots n_d}$ converges in distribution to η_μ for the situation under H_0 as well as H_1, where the η_μ is an optimal test based on the statistic $\int_D \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0)^\top d\mathcal{Y}$, then by the optimality criterion defined in [10], $\eta_{\mu;n_1\cdots n_d}$ is an asymptotically optimal test based on the statistic $\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}))$. □

5. THE EXISTENCE OF AN ASYMPTOTICALLY OPTIMAL DESIGN

There are many criteria in defining optimal design of experiment for model check or testing LOF in regression, see e.g. [10, 11]. In this paper we define optimality criterion based on the power function of the test.

Definition 5.1. Let μ_1 and μ_2 be two designs that correspond to $\Xi_{\mu_1;n_1\cdots n_d}$ and $\Xi_{\mu_2;n_1\cdots n_d}$. The design μ_1 is said to be asymptotically more optimal than μ_2 for testing (2.2) or equivalently (4.1), if and only if $\Psi_{\mu_2}(\mathbf{f}) \leq \Psi_{\mu_1}(\mathbf{f})$ for every $\mathbf{f} \in \mathbf{V}^p \cap (\mathbf{W}^p)^C$. Further, let $\mathcal{G}$ be the set of designs μ s.t. F_μ is continuous and nondecreasing on D. A design $\mu_0 \in \mathcal{G}$ is called an asymptotically optimal design for testing (2.2), if and only if it holds

$$\sup_{\mu\in\mathcal{G}} \Psi_\mu(\mathbf{f}) = \Psi_{\mu_0}(\mathbf{f}),\ \forall \mathbf{f} \in \mathbf{V}^p \cap (\mathbf{W}^p)^C.$$

We notice that in this paper the space $\mathcal{G}$ is furnished with the uniform topology induced by the metric $d_{\mathcal{G}}$, defined by $d_{\mathcal{G}}(\mu_1, \mu_2) := \|F_{\mu_1} - F_{\mu_2}\|_\infty$.

Remark 5.2. Since Φ is nondecreasing on $\mathbb{R}$, a design μ_1 is asymptotically more optimal than μ_2 for testing $H_0 : \mathbf{g}_2 \equiv \mathbf{0}$ against $H_1 : \mathbf{g}_2 \equiv \mathbf{f}_1$, for $\mathbf{f}_1 \in \mathbf{V}^p \cap (\mathbf{W}^p)^C$, if and only if

$$\frac{\left\langle \mathbf{\Sigma}^{-1/2}\mathbf{f}_1, \mathbf{\Sigma}^{-1/2}\mathbf{f}\right\rangle^{(p)}_{\mu_2}}{\|\mathbf{\Sigma}^{-1/2}\mathbf{f}_1\|^{(p)}_{\mu_2}} \leq \frac{\left\langle \mathbf{\Sigma}^{-1/2}\mathbf{f}_1, \mathbf{\Sigma}^{-1/2}\mathbf{f}\right\rangle^{(p)}_{\mu_1}}{\|\mathbf{\Sigma}^{-1/2}\mathbf{f}_1\|^{(p)}_{\mu_1}},\ \forall \mathbf{f} \in \mathbf{V}^p \cap (\mathbf{W}^p)^C.$$

Let $\{\widetilde{w}_{1\mu}, \ldots, \widetilde{w}_{q\mu}, \widetilde{w}_{(q+1)\mu}, \widetilde{w}_{m\mu}\}$ be the Gram-Schmidt orthogonal versions of the original basis of $\mathbf{V}$ when the sampling is conducted under the designs μ, for arbitrary fixed $\mu \in \mathcal{G}$. Since $\mathbf{f}_1$ and $\mathbf{f}$ are arbitrary in $\mathbf{V}^p \cap (\mathbf{W}^p)^C$, there exist vectors of constants $\mathbf{a} := (a_1, \ldots, a_p)^\top, \mathbf{b} := (b_1, \ldots, b_p)^\top \in \mathbb{R}^p$, such that $\mathbf{f}_1 := \mathbf{a}(w_{(q+1)\mu} + \cdots + w_{m\mu})$ and $\mathbf{f} := \mathbf{b}(w_{(q+1)\mu} + \cdots + w_{m\mu})$. Then, it holds

$$\frac{\left\langle \mathbf{\Sigma}^{-1/2}\mathbf{f}_1, \mathbf{\Sigma}^{-1/2}\mathbf{f}\right\rangle^{(p)}_{\mu}}{\|\mathbf{\Sigma}^{-1/2}\mathbf{f}_1\|^{(p)}_{\mu}} = \frac{A_{\mathbf{ab}}}{\sqrt{A_{\mathbf{aa}}}}\sqrt{\|w_{(q+1)\mu}\|^2_\mu + \cdots + \|w_{m\mu}\|^2_\mu}$$

where $A_{\mathbf{ab}} := \mathbf{a}^\top\mathbf{\Sigma}^{-1}\mathbf{b}$ and $A_{\mathbf{aa}} := \mathbf{a}^\top\mathbf{\Sigma}^{-1}\mathbf{a}$. Thus by the orthogonality of the basis of $\mathbf{V}^p \cap (\mathbf{W}^p)^C$, the design μ_1 is asymptotically more optimal than μ_2, if and only if

$$\sqrt{\|w_{(q+1)\mu_1}\|^2_{\mu_1} + \cdots + \|w_{m\mu_1}\|^2_{\mu_1}} \leq \sqrt{\|w_{(q+1)\mu_2}\|^2_{\mu_2} + \cdots + \|w_{m\mu_2}\|^2_{\mu_2}}.$$

Hence by the preceding result, a design μ_0 is an asymptotically optimal design for testing LOF if and only if $\mathcal{W}_{\mathbf{f}}(\mu_0) = \sup_{\mu\in\mathcal{G}} \mathcal{W}_{\mathbf{f}}(\mu)$ for arbitrary fixed $\mathbf{f} \in \mathbf{V}^p \cap (\mathbf{W}^p)^C$, where

$$\mathcal{W}_{\mathbf{f}}(\mu) := \frac{A_{\mathbf{ab}}}{\sqrt{A_{\mathbf{aa}}}}\sqrt{\|w_{(q+1)\mu}\|^2_\mu + \cdots + \|w_{m\mu}\|^2_\mu}.$$

This means that the problem of finding an optimal design of testing LOF is now shifted to that of finding a design $\mu \in \mathcal{G}$ that maximizes the sum of the square norm of the basis of $\mathbf{V}^p \cap (\mathbf{W}^p)^C$ in $L_2(\mu, D)$. The set $\mathcal{G}$ is compact under the uniform topology, because $\mathcal{G}$ can be easily shown to be closed and bounded under such a topology. Furthermore, since integration with respect to μ involved in the computation of norm $\|\cdot\|^2_\mu$ is continuous function on $\mathcal{G}$, then as a composition of two continuous functions, $\mathcal{W}_{\mathbf{f}}(\cdot)$ is therefore continuous on $\mathcal{G}$, for every $\mathbf{f} \in \mathbf{V}^p \cap (\mathbf{W}^p)^C$. By a standard result in analysis there exists an asymptotically optimal design for testing (2.2).

Theorem 5.3. *Among the elements of* $\mathcal{G}$ *with the same sample size* $n_1 \times \cdots \times n_d$*, regular lattice is the asymptotically optimal design for testing LOF in multivariate regression.*

5.1 Example 1

Suppose we observe the following multivariate regression model

$$Y_i(t,s) = \beta_{0i} + \beta_{1i}t + \beta_{2i}s + \varepsilon_i(t,s),\ (t,s) \in [1,2]^2,\ i = 1,2,3,4,$$

with $\mathbf{E}(\mathcal{E}) = \mathbf{0}$ and the invertible covariance matrix $\mathbf{\Sigma}$ defined as

$$\boldsymbol{\Sigma} := \begin{pmatrix} 9 & 3 & -6 & 12 \\ 3 & 26 & -7 & -11 \\ -6 & -7 & 9 & 7 \\ 12 & -11 & 7 & 65 \end{pmatrix}.$$

Suppose we are given two probability measures μ_0 and the Lebesgue measure λ^2 for constructing the experimental design according to the design scheme introduced in Section 2, where $F_{\mu_0}(t,s) := 2(1-1/t)2(1-1/s)$, for $(t,s) \in D := [1,2]^2$. We determined the design points of $\Xi_{\mu_0;n_1\times n_2}$ for fixed $n_1 \geq 1$ and $n_2 \geq 1$ by the formula

$$t_{n_1j_1} = \frac{2n_1}{2n_1 - j_1} \text{ and } t_{n_2j_2} = \frac{4n_1n_2(t_{n_1j_1}-1)}{t_{n_1j_1}(4n_1n_2 - j_1j_2) - 4n_1n_2},$$

for $1 \leq j_1 \leq n_1$ and $1 \leq j_2 \leq n_2$. The design points of $\Xi_{\lambda^2;n_1\times n_2}$ is clearly given by the regular lattice with $t_{n_1j_1} = 1 + \frac{j_1}{n_1}$ and $t_{n_2j_2} = 1 + \frac{j_2}{n_2}$. We are interested in computing and comparing the asymptotic power functions when we are testing $H_0 : \mathbf{g} = (g^{(i)})_{i=1}^4 \in \mathbf{W}^4 := [w_1]^4$, while we are observing $\mathbf{g} \in \mathbf{V}^4 := [w_1,w_2,w_3]^4$, with $w_1(t,s) = 1$, $w_2(t,s) = t$ and $w_3(t,s) = s$, for $(t,s) \in D$. The Gram-Schmidt orthogonal version of these basis functions when considered as functions in $L_2(D,\mu_0)$ and $L_2(D,\lambda^2)$ are given by $\mathbf{V}^4_{\mu_0} := [\widetilde{w}_{1\mu_0}, \widetilde{w}_{2\mu_0}, \widetilde{w}_{3\mu_0}]$ and $\mathbf{V}^4_{\lambda^2} := [\widetilde{w}_{1\lambda^2}, \widetilde{w}_{2\lambda^2}, \widetilde{w}_{3\lambda^2}]$, respectively, where for $(t,s) \in D$,

$$\widetilde{w}_{1\mu_0}(t,s) = 1,\ \widetilde{w}_{2\mu_0}(t,s) = t - \ln 4,\ \widetilde{w}_{3\mu_0}(t,s) = s - \ln 4$$
$$\widetilde{w}_{1\lambda^2}(t,s) = 1,\ \widetilde{w}_{2\lambda^2}(t,s) = t - 3/2,\ \widetilde{w}_{3\lambda^2}(t,s) = s - 3/2.$$

First we consider the design μ_0 in testing the hypothesis $H_0 : \mathbf{g}_2 \equiv \mathbf{0}$ against an alternative $H_1 : \mathbf{g}_2 \equiv \mathbf{f}_{1\mu_0}$, for a vector $\mathbf{f}_{1\mu_0} \in \mathbf{V}^4_{\mu_0} \cap (\mathbf{W}^4_{\mu_0})^C$ defined by

$$\mathbf{f}_{1\mu_0}(t,s) := (18,11,13,73)^\top(\widetilde{w}_{2\mu_0}(t,s) + \widetilde{w}_{3\mu_0}(t,s))\ (t,s) \in D,$$

where $\mathbf{g}_2$ satisfies the decomposition $\mathbf{g} = \mathbf{g}_1 \oplus \mathbf{g}_2$, with $\mathbf{g}_1 \in \mathbf{W}^4_{\mu_0}$ and $\mathbf{g}_2 \in \mathbf{V}^4_{\mu_0} \cap (\mathbf{W}^4_{\mu_0})^C$. The intention is to compute the limit of $\Psi_{\mu_0;n_1n_2}(\mathbf{f})$ when it is evaluated at $\mathbf{f} \equiv \rho\mathbf{f}_{1\mu_0}$, for $\rho \in \mathbb{R}$. After conducting a little computation we get

$$\begin{aligned}\mathcal{W}_{\mathbf{f}}(\mu_0) &= \rho\sqrt{(18,11,13,73)\boldsymbol{\Sigma}^{-1}(18,11,13,73)^\top} \\ &\quad \times \sqrt{\|\widetilde{w}_{2\mu_0}\|^2_{\mu_0} + \|\widetilde{w}_{3\mu_0}\|^2_{\mu_0}} \\ &= 19.36492\sqrt{0.078188 + 0.078188}\ \rho = 7.657738\rho.\end{aligned}$$

Hence, by Corollary 3.3 we obtain the limiting power function of the test of size α based on the statistic $\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu_0;n_1\times n_2}))$ as

$$\begin{aligned}\Psi_{\mu_0}(\rho) &:= \Psi_{\mu_0}(\rho\mathbf{f}_{1\mu_0}) \\ &= 1 - \Phi(\Phi^{-1}(1-\alpha) - 7.657738\rho),\ \text{for } \rho \in \mathbb{R}.\end{aligned}$$

Next we test the similar hypothesis $H_0 : \mathbf{g}_2 \equiv \mathbf{0}$ against an alternative $H_1 : \mathbf{g}_2 \equiv \mathbf{f}_{1\lambda^2}$, for a vector $\mathbf{f}_{1\lambda^2} \in \mathbf{V}^4_{\lambda^2} \cap (\mathbf{W}^4_{\lambda^2})^C$ defined by

$$\mathbf{f}_{1\lambda^2}(t,s) := (18,11,13,73)^\top(\widetilde{w}_{2\lambda^2}(t,s) + \widetilde{w}_{3\lambda^2}(t,s))\ (t,s) \in D,$$

where $\mathbf{g}_2$ is obtained from the orthogonal decomposition $\mathbf{g} = \mathbf{g}_1 \oplus \mathbf{g}_2$, with $\mathbf{g}_1 \in \mathbf{W}^4_{\lambda^2}$ and $\mathbf{g}_2 \in \mathbf{V}^4_{\lambda^2} \cap (\mathbf{W}^4_{\lambda^2})^C$. By applying the analogous computation steps as before, the power function $\Psi_{\lambda^2;n_1n_2}(\mathbf{f})$ when it is evaluated at $\mathbf{f} \equiv \rho\mathbf{f}_{1\lambda^2}$ converges point wise to

$$\begin{aligned}\Psi_{\lambda^2}(\lambda) &:= \Psi_{\lambda^2}(\rho\mathbf{f}_{1\lambda^2}) \\ &= 1 - \Phi(\Phi^{-1}(1-\alpha) - 7.905694\rho),\ \text{for } \rho \in \mathbb{R}.\end{aligned}$$

It is seen that $\Psi_{\lambda^2}(\rho)$ is slightly larger than $\Psi_{\mu_0}(\rho)$ for all ρ. This means that the design λ^2 is asymptotically more optimal than that of μ_0 suggesting us to choose the regular lattice as the design instead of that obtained using the distribution function of μ_0 in order to get asymptotically more optimal test.

5.2 Example 2

In the second example we investigate the behavior of the designs μ_0 and λ^2 defined above for testing the hypothesis $H_0 : \mathbf{g} \in \mathbf{W}^4$ against $H_1 : \mathbf{g} \in \mathbf{V}^4$ when we observe the multivariate model

$$Y_i(t,s) = \sum_{k=1}^5 \beta_{ki}w_k(t,s) + \varepsilon_i(t,s),\ i = 1,2,3,4,$$

with $\mathbf{E}(\mathcal{E}) = \mathbf{0}$ and the covariance matrix given as in the preceding example, where $\mathbf{W} := [w_1,w_2,w_3]$ and $\mathbf{V} := [w_1,w_2,w_3,w_4,w_5]$ built by the regression functions $w_1(t,s) = 1$, $w_2(t,s) = t$, $w_3(t,s) = s$, $w_4(t,s) = t^2$, and $w_5(t,s) = s^2$, for $(t,s) \in [1,2]^2$ which are clearly linearly independent as functions in $L_2(D,\mu_0)$ as well as in $L_2(D,\lambda^2)$. In other word we want to test whether or not a first-order polynomial is adequate for representing the model under the test procedure based on $\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu_0;n_1\times n_2}))$ and $\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\lambda^2;n_1\times n_2}))$. The orthogonal representation of g_i in $L_2(D,\mu_0)$ and $L_2(D,\lambda^2)$ is respectively given by $g_i = \sum_{k=1}^5 \beta_{ki}\widetilde{w}_{k\mu_0}(t,s)$ and $g_i = \sum_{k=1}^5 \beta_{ki}\widetilde{w}_{k\lambda^2}(t,s)$, for $i = 1,2,3,4$, where

$$\widetilde{w}_{1\mu_0}(t,s) = 1, \widetilde{w}_{2\mu_0}(t,s) = t - \ln(4), \widetilde{w}_{3\mu_0}(t,s) = s - \ln(4),$$
$$\widetilde{w}_{4\mu_0}(t,s) = t^2 - K_1t + K_2, \widetilde{w}_{5\mu_0}(t,s) = s^2 - K_1s + K_2$$
$$\widetilde{w}_{1\lambda^2}(t,s) = 1, \widetilde{w}_{2\lambda^2}(t,s) = t - 3/2, \widetilde{w}_{3\lambda^2}(t,s) = s - 3/2,$$
$$\widetilde{w}_{4\lambda^2}(t,s) = t^2 - t/4 - 47/24, \widetilde{w}_{5\lambda^2}(t,s) = s^2 - s/4 - 47/24,$$

where $K_1 := 3 - 2\ln(4)$ and $K_2 := [3 - 2\ln(4)]\ln(4) - 2$. If we consider μ_0 as the design, the limiting power function of $\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu_0;n_1\times n_2}))$ for testing $H_0 : \mathbf{g}_2 \equiv \mathbf{0}$ against $H_1 : \mathbf{g}_2 \equiv \mathbf{f}_{1\mu_0}$, where $\mathbf{f}_{1\mu_0} := (\widetilde{w}_{4\mu_0} + \widetilde{w}_{5\mu_0})\mathbf{1} \in \mathbf{V}^4 \cap (\mathbf{W}^4)^C$, when it is evaluated at $\mathbf{f} \equiv \theta\mathbf{f}_{1\mu_0}$ is given by

$$\Psi_{\mu_0}(\mathbf{f}) = \Psi_{\mu_0}(\theta \mathbf{f}_{1\mu_0})$$
$$= 1 - \Phi\left(\Phi^{-1}(1-\alpha) - 3.022116\ \theta\right),\ \theta \in \mathbb{R},$$

where $\mathbf{1} := (1,1,1,1)^\top \in \mathbb{R}^4$. Analogously, under the design λ^2 we get the limiting power function of the test based on $\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\lambda_0^2;n_1\times n_2}))$ for testing $H_0 : \mathbf{g}_2 \equiv \mathbf{0}$ against $H_1 : \mathbf{g}_2 \equiv \mathbf{f}_{1\lambda^2}$, where $\mathbf{f}_{1\lambda^2} := (\widetilde{w}_{4\lambda^2} + \widetilde{w}_{5\lambda^2})\mathbf{1} \in \mathbf{V}^4 \cap (\mathbf{W}^4)^C$, when it is evaluated at $\mathbf{f} \equiv \theta \mathbf{f}_{1\lambda^2}$ is given by

$$\Psi_{\lambda^2}(\mathbf{f}) := \Psi_{\lambda^2}(\theta \mathbf{f}_{1\lambda^2})$$
$$= 1 - \Phi\left(\Phi^{-1}(1-\alpha) - 6.41156\ \theta\right),\ \theta \in \mathbb{R}.$$

By these results it can be stated that λ^2 is asymptotically more optimal than μ_0.

6. SIMULATION STUDY

In this section we simulate the finite-sample performance of the test for the hypotheses studied in Example 1 and Example 2 of Section 5. In the simulation we build the graphs of the power functions of the tests for various sample-sizes, where the observations are sampled using the experimental designs μ_0 and λ^2 defined on the experimental region D. In each case the vector of random errors is generated independently from the centered 4-variate normal distribution with the covariance matrix $\mathbf{\Sigma}$ defined in the example. We will demonstrate for the case of $d = 2$ that independent to the choice of the design strategies and the level α, the power functions $\Psi_{\mu;n_1\times n_2}(\mathbf{f})$ converges to $\Psi_\mu(\mathbf{f})$ for every $\mathbf{f} \in \mathbf{V}^4 \cap (\mathbf{W}^4)^C$ as the sample sizes get large.

For computational reason we consider the class $\{[1,t] \times [1,s] : 1 \le t, s \le 2\}$ of $[1,2]^2$ as the index sets. Hence $\mathbf{S}_{n_1\times n_2}(\mathbf{Y}(\Xi_{\mu;n_1\times n_2}))([1,t] \times [1,s])$ is written by $\mathbf{S}_{n_1\times n_2}(\mathbf{Y}(\Xi_{\mu;n_1\times n_2}))(t,s)$ for brevity where the last can be regarded as a random continuous function on $[1,2]^2$ with respect to the usual Euclidean norm. This leads us to the conclusion that the integral of any function of bounded variation with respect to $\mathbf{S}_{n_1\times n_2}(\mathbf{Y}(\Xi_{\mu;n_1\times n_2}))$ coincides pathwise with the Riemann-Stieltjes integral, cf. Stroock [25], pp. 7–16. By considering the partition on $[1,2]^2$ built up by the design points of $\Xi_{\mu;n_1\times n_2}$ we get by the definition of the Riemann-Stieltjes integral

$$\mathcal{J}(\mathbf{Y}(\Xi_{\mu;n_1\times n_2}))$$
$$= \int_{[1,2]^2}^{(R)} \mathbf{f}_1^\top \mathbf{\Sigma}^{-1/2} d\mathbf{\Sigma}^{-1/2} \mathbf{S}_{n_1\times n_2}(\mathbf{Y}(\Xi_{\mu;n_1\times n_2}))$$
$$\approx \sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2} \mathbf{f}_1^\top(t_{n_1j_1}, t_{n_2j_2}) \mathbf{\Sigma}^{-1} \Delta_{j_1,j_2} \mathbf{S}_{n_1\times n_2}(\mathbf{Y}(\Xi_{\mu;n_1\times n_2}))$$
$$= \frac{1}{\sqrt{n_1n_2}} \sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_2} \mathbf{f}_1^\top(t_{n_1j_1}, t_{n_2j_2}) \mathbf{\Sigma}^{-1} \mathbf{Y}(t_{n_1j_1}, t_{n_2j_2}),$$

where for any function $\mathbf{u} : [1,2]^2 \to \mathbb{R}^4$, $1 \le j_1 \le n_1$ and $1 \le j_2 \le n_2$,

$$\mathbf{\Delta}_{j_1,j_2}\mathbf{u} := \mathbf{u}(t_{n_1j_1}, t_{n_2j_2}) - \mathbf{u}(t_{n_1j_1-1}, t_{n_2j_2})$$
$$- \mathbf{u}(t_{n_1j_1}, t_{n_2j_2-1}) + \mathbf{u}(t_{n_1j_1-1}, t_{n_2j_2-1}).$$

Here the notation $\int^{(R)}$ stands for the integration in the sense of Riemann-Stieltjes.

6.1 Simulation 1

In the first simulation we consider the situation described in Example 1 of Section 5. Using the design μ_0 we generate the observations independently from the model

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = \frac{1}{\sqrt{n_1n_2}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \rho \begin{pmatrix} 18 \\ 11 \\ 13 \\ 73 \end{pmatrix} \frac{(t_{n_1j_1} - \ln 4 + t_{n_2j_2} - \ln 4)}{\sqrt{n_1n_2}} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix},$$

likewise when the design is λ^2, the observations are generated independently from the model

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = \frac{1}{\sqrt{n_1n_2}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \rho \begin{pmatrix} 18 \\ 11 \\ 13 \\ 73 \end{pmatrix} \frac{(t_{n_1j_1} - 3/2 + t_{n_2j_2} - 3/2)}{\sqrt{n_1n_2}} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix},$$

where in both cases $\mathcal{E}$ are generated independently from the distribution $N_4(\mathbf{0}, \mathbf{\Sigma})$. It is clear that the observations are from H_0 if and only if $\rho = 0$, otherwise they are from H_1. In the case $\rho = 0$ the power must attain the presigned level of significance α. The simulation results for $\alpha = 0.05$ under the designs μ_0 and λ^2 are exhibited in in Figure 1 and Figure 2, respectively. Figure 1 compares between $\Psi_{\mu_0;n_1\times n_2}(\rho \mathbf{f}_1)$ (dashed line) generated for the sample size (i) 40×40, (ii) 50×50, (iii) 60×60 and (iv) 70×70, respectively and $\Psi_{\mu_0}(\rho \mathbf{f}_1)$ (smooth line). Figure 2 presents $\Psi_{\lambda^2;n_1\times n_2}(\rho \mathbf{f}_1)$ scattered by dashed lines based on the sample sizes (i) 60×60 and (ii) 70×70 and those of $\Psi_{\lambda^2}(\rho \mathbf{f}_1)$ represented by smooth lines. The simulation results show that for testing constant model, the power functions $\Psi_{\mu_0;n_1\times n_2}(\mathbf{f})$ gives a good approximation to $\Psi_{\mu_0}(\mathbf{f})$, whereas $\Psi_{\lambda^2;n_1\times n_2}(\mathbf{f})$ approximate very well its limiting power function $\Psi_{\lambda^2}(\mathbf{f})$ achieving the level of significance $\alpha = 0.05$ when ρ is set equal to 0.

6.2 Simulation 2

In the second simulation we consider the test exhibited in Example 2 of Section 5. The observations are generated from the models

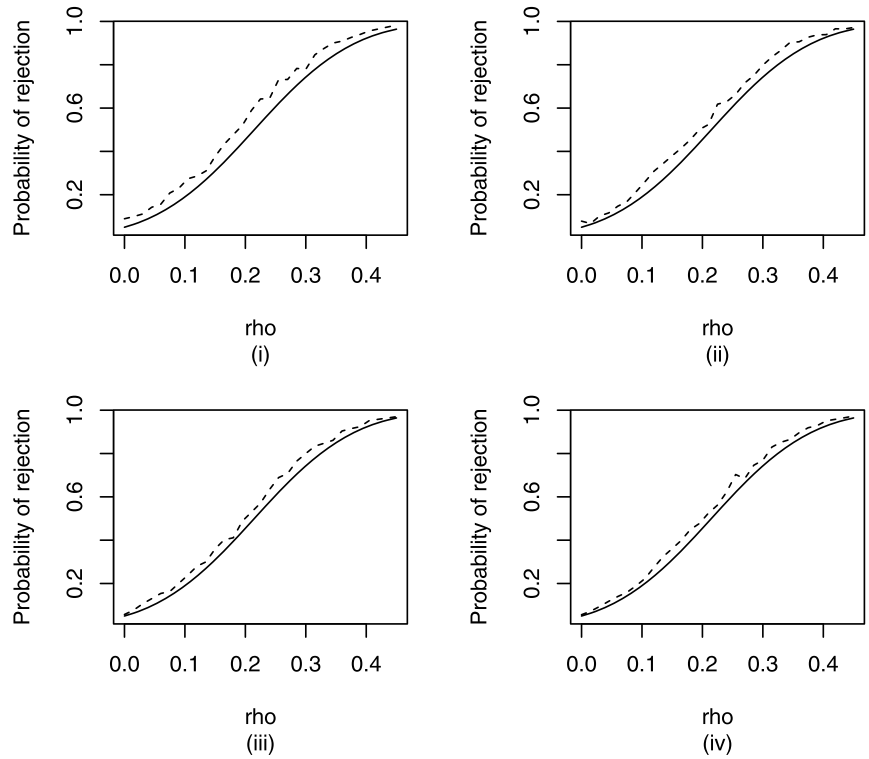

Figure 1. The graphs of $\Psi_{\mu_0;n_1\times n_2}(\rho\mathbf{f}_1)$ *(dashed lines) and* $\Psi_{\mu_0}(\rho\mathbf{f}_1)$ *(smooth lines) for* $\alpha = 0.05$. *The sample sizes are chosen for* (i) 40×40, (ii) 50×50, (iii) 60×60 *and* (iv) 70×70. *The graphs are generated under 1000 runs using R.*

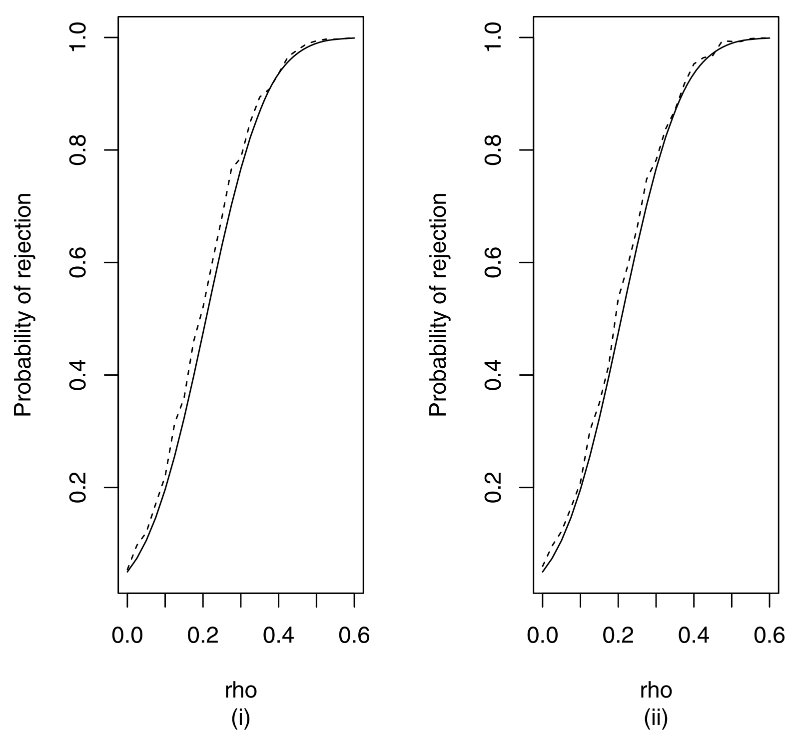

Figure 2. The graphs of $\Psi_{\lambda^2;n_1\times n_2}(\rho\mathbf{f}_1)$ *(dashed lines) and* $\Psi_{\lambda^2}(\rho\mathbf{f}_1)$ *(smooth lines) for* $\alpha = 0.05$. *The simulations are based on the sample sizes* (i) 60×60 *and* (ii) 70×70 *generated under 1000 runs.*

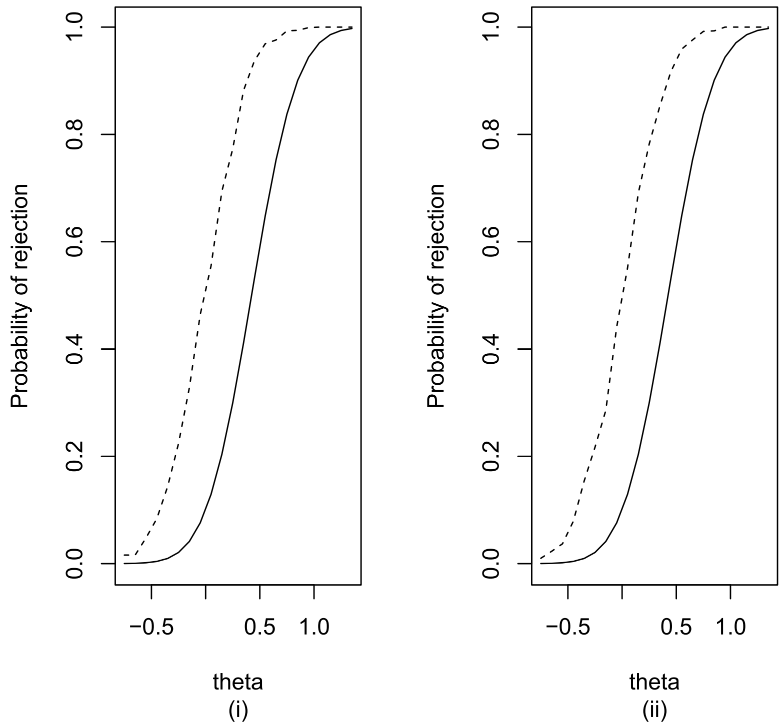

Figure 3. The graphs of $\Psi_{\mu_0;n_1\times n_2}(\theta\mathbf{f}_1)$ (dashed lines) and $\Psi_{\mu_0}(\theta\mathbf{f}_1)$ (smooth lines) for $\alpha = 0.10$. The sample sizes are (i) 50×50, and (ii) 60×60. The graphs are generated under 1000 runs using R.

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \frac{(1 + t_{n_1j_1} + t_{n_2j_2} - 2\ln 4)}{\sqrt{n_1n_2}} + \theta \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \times \frac{(t^2_{n_1j_1} + t^2_{n_2j_2} - K_1(t_{n_1j_1} + t_{n_2j_2}) + 2K_2)}{\sqrt{n_1n_2}} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix},$$

when μ_0 is chosen as the design, whereas under the regular lattice we generate the observations using the following one

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \frac{(1 + t_{n_1j_1} + t_{n_2j_2} - 3)}{\sqrt{n_1n_2}} + \theta \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \times \frac{(t^2_{n_1j_1} + t^2_{n_2j_2} - (t_{n_1j_1} + t_{n_2j_2})/4 - 47/12)}{\sqrt{n_1n_2}} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix},$$

In both cases we generate the vector of random errors independently from the centered 4-dimensional normal distribution with the covariance matrix $\mathbf{\Sigma}$. The simulation results are exhibited in Figure 3 and Figure 4, respectively. It can be seen from both figures that the distance between the curves representing the empirical power functions of the tests and their limits are still large although the sample sizes are large enough. This means that for testing polynomial model of higher degree, the method based on the partial sums of the vector of observations fails to give a satisfactory approximation result.

7. APPLICATION

Our intention in this section is to demonstrate the application of the proposed method in real data in which we investigate the mining data as studied in Tahir [33]. The sample was obtained by drilling bores over the exploration region of the company positioned according to a 7×14 dimensional lattice with 7 equidistance rows running west to east and 14 equidistance column running south to north, see also Somayasa and et al. [27, 28, 29]. Figure 4 exhibits the pairs plot among the percentages of Ni, CaO, Co, the logarithm of the percentages of SiO_2 ($LogSiO_2$), MgO (LogMgO), and Fe (LogFe) measured simultaneously. The plot indicates the existence of positive as well as negative correlations among the measured variables. For example LogMgO and $LogSiO_2$, LogFe and Co, CaO and LogMgO, CaO and $LogSiO_2$ seem positively correlated. Further, the plot also shows negative

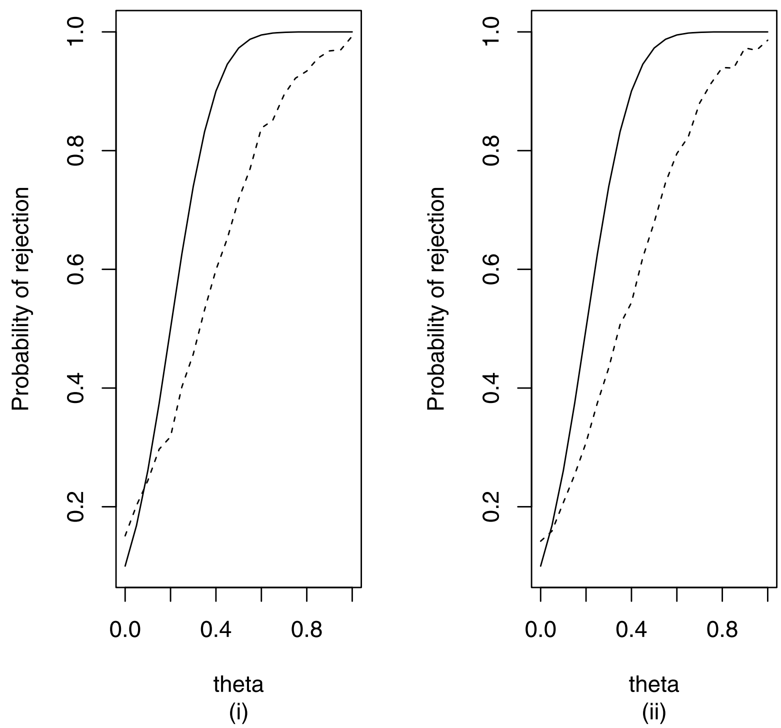

Figure 4. The graphs of $\Psi_{\lambda^2;n_1\times n_2}(\rho\mathbf{f}_1)$ (dashed lines) and $\Psi_{\lambda^2}(\rho\mathbf{f}_1)$ (smooth lines) for $\alpha = 0.10$. The simulations are based on the sample sizes (i) 50×50 and (ii) 60×60 generated under 1000 runs.

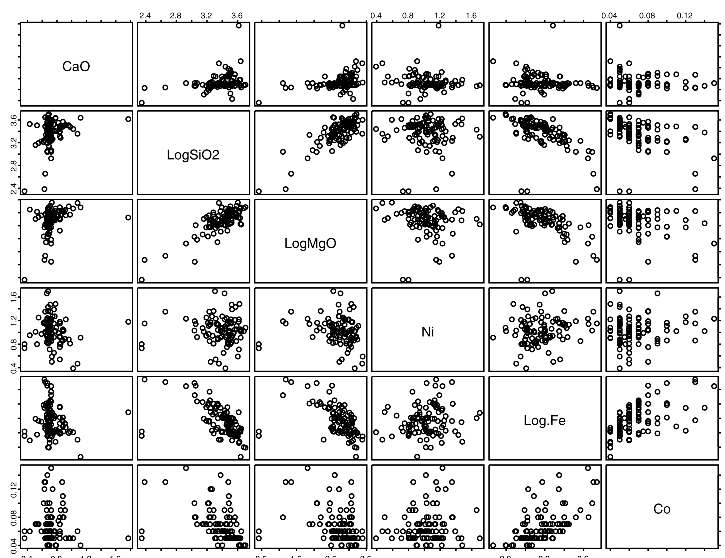

Figure 5. The pairs plot of the percentages of Ni, CaO, Co, SiO_2, MgO, and Fe observed over an 7×14 lattice points showing the correlations among the variables. Source of data: Tahir [33].

correlation between LogFe and $\mathrm{LogSiO_2}$, LogFe and LogMgO, CO and LogMgO, and CO and $\mathrm{LogSiO_2}$. By this reason a multivariate analysis must be conducted in the statistical modelling taking into account the unknown covariance matrix of the vector of the variables. Furthermore based on the individual scatter plot of each variable presented in [27], we can give a conjecture that polynomials of lower order is appropriate to approximate the population model. More precisely let $\mathbf{Y} := (Y_1, Y_2, Y_3, Y_4, Y_5, Y_6)^\top$ be the vector of observations representing the observed percentages of CaO, $\mathrm{LogSiO_2}$, LogMgO, Co, Ni, and LogFe, respectively. It seems that the common model for describing the functional relationship between the vector of observations and the position of the observations on the earth is a multivariate first-order model. By defining a suitable transformation we can assume that the experiment was conducted on the experimental region given by the unit rectangle $[0,1]\times[0,1]$ and the experimental design $\Xi_{\lambda^2;7\times 14}$ corresponds to the Lebesgue measure λ^2 defined on $\mathcal{B}([0,1]\times[0,1])$. The point where the observation was started is regarded as the origin $(0,0)$ which is laid on the south-west corner, whereas the point where the observation was ended is regarded as the point $(1,1)$ which is put on the north-east corner. Consequently, for $1\leq \ell\leq 7$ and $1\leq k\leq 14$, the observation $\mathbf{Y}_{\ell k}$ that stands for the measurement on the point $(\ell/7, k/14)$ in the unit rectangle is actually the ℓ^{th} observation on the west to east and the k^{th} observation on the south to north direction.

We test the hypothesis H_0 that the true model is first-order model while observing second-order model. The critical region of this test is determined by computing $\mathcal{J}(\mathbf{Y}(\Xi_{\lambda^2;7\times 14}))$ using the following approximation formula:

$$\mathcal{J}(\mathbf{Y}(\Xi_{\lambda^2;7\times 14})) \approx \frac{1}{\sqrt{98}}\sum_{\ell=1}^{7}\sum_{k=1}^{14}\mathbf{f}_1^{\top}(\frac{\ell}{7},\frac{k}{14})\widehat{\mathbf{\Sigma}}_{6\times 6}^{-1}\mathbf{Y}_{\ell k}.$$

This formula is actually the value of the component-wise Riemann-Stieltjes sum of $\mathbf{f}_1$ with respect $\mathbf{S}_{7\times 14}(\Xi_{\lambda^2;7\times 14})$ over the partition given by $\Xi_{\lambda^2;7\times 14}$. When all outliers are ignored we get for the mining data the value of the statistic test as $\mathcal{J}(\mathbf{Y}(\Xi_{\lambda^2;7\times 14})) = 275.43467$, when under H_1 we assume the function

$$\begin{aligned}\mathbf{f}_1(t,s) := [&\sqrt{5}(6t^2-6t+1)+(4ts-2t-2s+1)/3\\ &+\sqrt{5}(6s^2-6s+1)]\mathbf{1}\in\mathbf{V}^6\cap(\mathbf{W}^6)^C,\end{aligned}$$

where $\mathbf{1} := (1,1,1,1,1,1)^{\top}\in\mathbb{R}^6$, giving $\|\widehat{\mathbf{\Sigma}}^{-1/2}\mathbf{f}_1\|_{\lambda^2} = 106.84680$ and p-value 0.497%. This means that H_0 will never be rejected for all level of significance α less than 0.497%. However, from the practical view-point this value is too small for H_0 being rejected. Therefore, we conclude that first-order model is simultaneously appropriate for describing the model.

8. CONCLUSION

We have developed an asymptotic method for testing lack of fit of the mean vector in multivariate spatial regression by considering the multidimensional set-indexed partial sums of the vector of observations. The experimental design is constructed by incorporating a sampling technique according to a probability measure having a continuous-increasing distribution function. The limit process is given by a type of multidimensional signal plus noise model with the multivariate set-indexed Gaussian white noise as the signal. Observing the limit process we can formulate an optimum test by utilizing the Cameron-Martin density formula of the limit process. The optimality of the test depends also on the choice of the design. In this paper we have shown that regular lattice is an asymptotically optimal design. In the future we will establish likelihood ratio test for more general hypothesis then that studied in this paper. The application of the method to the mining data shows that multivariate first-order model is plausible for α less then 0.497%.

ACKNOWLEDGEMENTS

This work was supported by the KLN and Publikasi Internasional research fund 2016 provided by the ministry of RISTEK-DIKTI. The first author thanks Professor Wolfgang Bischoff for his nice discussion during the preparation of the manuscript. The authors are also grateful to the anonymous referees for their valuable comments in improving the exposition of the paper.

APPENDIX A. MULTIVARIATE INVARIANCE PRINCIPLE

Definition A.1. Let $(D,\mathcal{B}(D),\mu)$ be a probability space on the d-dimensional closed rectangle $D := \times_{j=1}^{d}[a_j,b_j]\subset\mathbb{R}^d$, with $a_j<b_j$, where $\mathcal{B}(D)$ is Borel σ-algebra over D. Let $\mathcal{A}_0$ be the family of subsets in $\mathcal{B}(D)$ which has finite measure under μ. A centered Gaussian process $\mathbf{Z}_\mu := \{(Z_\mu^{(i)}(A))_{i=1}^p : A\in\mathcal{A}_0\}$ is called a p-dimensional set-indexed Gaussian white noise with the control measure μ, if and only if

$$Cov(\mathbf{Z}_\mu(A_1),\mathbf{Z}_\mu(A_2)) = \mu(A_1\cap A_2)\mathbf{I}_p,\ \forall A_1,A_2\in\mathcal{A}_0$$

where $\mathbf{I}_p$ is the $p\times p$ dimensional identity matrix. Let $P_{\mathbf{Z}_\mu}$ be the probability distributions of $\mathbf{Z}_\mu$ on $(\mathcal{C}^p(\mathcal{A}_0),\mathcal{B}(\mathcal{C}^p(\mathcal{A}_0)))$ and let $P_{Z_\mu^{(i)}}$ be the marginal distribution of $\mathbf{Z}_\mu$ on $(\mathcal{C}(\mathcal{A}_0),\mathcal{B}(\mathcal{C}(\mathcal{A}_0)))$, for $i=1,\ldots,p$. Then by the definition it is clear that $\mathbf{Z}_\mu$ has mutually independent components in the sense

$$P_{\mathbf{Z}_\mu}(\times_{i=1}^{p}C_i) = \Pi_{i=1}^{p}P_{Z_\mu^{(i)}}(C_i),\ \forall\times_{i=1}^{p}C_i\subset\mathcal{C}^p(\mathcal{A}_0).$$

We notice that when the control measure μ is the Lebesgue measure on $\mathcal{B}(D)$, we obtain the well known p-dimensional set-indexed Brwonian sheet, see e.g. [27, 28]. If the index is the family $\left\{\times_{j=1}^{d}[a_j,t_j]:\ a_j\leq t_j\leq b_j,\ j=1,\ldots,d\right\}$ of subsets in $\mathcal{B}(D)$, then we get a Gaussian random field which is commonly called the multi parameter p-dimensional Brwonian motion.

Theorem A.2. *Let $\mathcal{E}(\Xi_{\mu;n_1\cdots n_d}) := \{\mathcal{E}_{j_1\cdots j_d} : 1\leq j_1\leq n_1,\ldots,1\leq j_d\leq n_d\}$ be an arrays of independent and identically distributed p-dimensional random vectors with $\mathbf{E}(\mathcal{E}_{1\cdots 1}) = \mathbf{0}$ and $Cov(\mathcal{E}_{1\cdots 1}) = \mathbf{\Sigma}$, where $\mathbf{\Sigma}$ is assumed to be positive definite. Let $\mathbf{Z}_\mu := (Z_\mu^{(i)})_{i=1}^p$ be the set-indexed Gaussian white noise with the control measure μ defined on the measurable space $(D,\mathcal{B}(D))$. Then*

$$\mathbf{\Sigma}^{-1/2}\mathbf{S}_{n_1\cdots n_d}(\mathcal{E}(\Xi_{\mu;n_1\cdots n_d}))\xrightarrow{\mathcal{D}}\mathbf{Z}_\mu.$$

Proof. (1) By the well-known multivariate Lindeberg-Feller central limit theorem (cf. [34], p. 20) we show that for any Borel subset $B_1, \ldots, B_m$ in $\mathcal{B}(D)$ and the constants $a_1, \ldots, a_m$, the general linear combination

$$\begin{aligned}\mathbf{F}_{n_1\cdots n_d} &:= \sum_{k=1}^{m} a_k \mathbf{\Sigma}^{-1/2}\mathbf{S}_{n_1\cdots n_d}(\mathcal{E}(\Xi_{\mu;n_1\cdots n_d}))(B_k)\\ &= \sum_{k=1}^{m} a_k \frac{1}{\sqrt{n_1\cdots n_d}}\mathbf{\Sigma}^{-1/2}\sum_{j_1=1}^{n_1}\cdots\sum_{j_d=1}^{n_d}\mathbf{1}_{B_k}(\mathbf{t}_{j_1\cdots j_d})\mathcal{E}_{j_1\cdots j_d}\end{aligned}$$

converges in distribution to $\sum_{j=k}^{m} a_k \mathbf{Z}_\mu(B_k)$ which can be easily shown to follow a p-variate normal distribution having zero mean and the covariance matrix

$$Cov\left(\sum_{k=1}^{m} a_k \mathbf{Z}_\mu(B_k)\right) = \sum_{k=1}^{m}\sum_{\ell=1}^{m} a_k a_\ell \mu(B_k\cap B_\ell)\mathbf{I}_p.$$

It suffices to show the covariance of $\mathbf{F}_{n_1\cdots n_d}$ converges to that of $\sum_{k=1}^{m} a_k\mathbf{Z}_\mu(B_k)$, and it satisfies Lindeberg condition. Since $\mathbf{E}(\mathbf{F}_{n_1\cdots n_d}) = \mathbf{0}$ and $\mathcal{E}_{j_1\cdots j_d}$ are independent and identically distributed, we have

$$\begin{aligned}&Cov(\mathbf{F}_{n_1\cdots n_d})\\ &= \sum_{k=1}^{m}\sum_{\ell=1}^{m} a_k a_\ell \frac{1}{n_1\cdots n_d}\sum_{j_1=1}^{n_1}\cdots\sum_{j_d=1}^{n_d}\mathbf{1}_{B_k\cap B_\ell}(\mathbf{t}_{j_1\cdots j_d})\mathbf{I}_p\\ &= \sum_{k=1}^{m}\sum_{\ell=1}^{m} a_k a_\ell P_n(B_k\cap B_\ell)\mathbf{I}_p \to \sum_{k=1}^{m}\sum_{\ell=1}^{m} a_k a_\ell \mu(B_k\cap B_\ell)\mathbf{I}_p\end{aligned}$$

where the last is the covariance of $\sum_{k=1}^{m} a_k\mathbf{Z}_\mu(B_k)$. Next we show Lindeberg condition. For $\varepsilon > 0$, let

$$\mathcal{L}(\varepsilon) := \sum_{j_1}^{n_1}\cdots\sum_{j_d=1}^{n_d}\mathbf{E}\left(\left\|\sum_{k=1}^{m}\frac{a_k\mathbf{\Sigma}^{-1/2}\mathbf{1}_{B_k}(\mathbf{t}_{j_1\cdots j_d})\mathcal{E}_{j_1\cdots j_d}}{\sqrt{n_1\cdots n_d}}\right\|^2 \times \mathbf{1}_{\left\{\left\|\sum_{k=1}^{m}\frac{a_k\mathbf{\Sigma}^{-1/2}\mathbf{1}_{B_k}(\mathbf{t}_{j_1\cdots j_d})\mathcal{E}_{j_1\cdots j_d}}{\sqrt{n_1\cdots n_d}}\right\|>\varepsilon\right\}}\right).$$

Let $M := \max_{1\le k\le m}|a_k|$, then by the property of the Euclidean norm, we get the following inequality

$$\left\|\sum_{k=1}^{m}\frac{a_k\mathbf{\Sigma}^{-1/2}\mathbf{1}_{B_k}(\mathbf{t}_{j_1\cdots j_d})\mathcal{E}_{j_1\cdots j_d}}{\sqrt{n_1\cdots n_d}}\right\|^2 \le \frac{M^2\|\mathbf{\Sigma}^{-1/2}\mathcal{E}_{j_1\cdots j_d}\|^2}{n_1\cdots n_d}.$$

Also $\left\|\sum_{k=1}^{m}\frac{a_k\mathbf{\Sigma}^{-1/2}\mathbf{1}_{B_k}(\mathbf{t}_{j_1\cdots j_d})\mathcal{E}_{j_1\cdots j_d}}{\sqrt{n_1\cdots n_d}}\right\| > \varepsilon$ implies $\|\mathbf{\Sigma}^{-1/2}\mathcal{E}_{j_1\cdots j_d}\| \ge \frac{\varepsilon\sqrt{n_1\cdots n_d}}{M}$. Therefor we get

$$\mathcal{L}(\varepsilon) \le M^2 E\left(\left\|\mathbf{\Sigma}^{-1/2}\mathcal{E}_{1\cdots 1}\right\|^2 \mathbf{1}_{\left\{\|\mathbf{\Sigma}^{-1/2}\mathcal{E}_{1\cdots 1}\|>\frac{\varepsilon\sqrt{n_1\cdots n_d}}{M}\right\}}\right)$$

which by the bounded convergence theorem (cf. Corollary 2.3.13 in [4]), the right-hand side converges to zero.

(2) We show that $\{\mathbf{S}_{n_1\cdots n_d}(\mathcal{E}(\Xi_{\mu;n_1\cdots n_d}))(B) : B \in \mathcal{B}(D)\}$ is tight. The suitable definition of the modulus of continuity of $\mathbf{S}_{n_1\cdots n_d}(\mathcal{E}(\Xi_{\mu;n_1\cdots n_d}))$ is

$$w(\mathbf{S}_{n_1\cdots n_d}(\mathcal{E}(\Xi_{\mu;n_1\cdots n_d}));\delta) := \sup_{\{A,B\in\mathcal{A}:\ d_\mu(A,B)<\delta\}}\Gamma(A,B),$$

where $\Gamma(A,B) := \|\mathbf{S}_{n_1\cdots n_d}(\mathcal{E}(\Xi_{\mu;n_1\cdots n_d}))(A) - \mathbf{S}_{n_1\cdots n_d}(\mathcal{E}(\Xi_{\mu;n_1\cdots n_d}))(B)\|$ for any $A, B \in \mathcal{B}(D)$ whose i^{th} component is given by

$$\begin{aligned}&\mathbf{S}_{n_1\cdots n_d}(\varepsilon_i(\Xi_{\mu;n_1\cdots n_d}))(B)\\ &\quad := \frac{\sum_{j_1=1,\ldots,j_d=1}^{n_1,\ldots,n_d}\mathbf{1}_B(\mathbf{t}_{j_1\cdots j_d})\varepsilon_{i,j_1\cdots j_d}}{\sqrt{n_1\cdots n_d}},\ i = 1,\ldots,p.\end{aligned}$$

By the definition it clearly holds

$$\begin{aligned}&w(\mathbf{S}_{n_1\cdots n_d}(\mathcal{E}(\Xi_{\mu;n_1\cdots n_d}));\delta)\\ &\quad\le \sum_{i=1}^{p} w(\mathbf{S}_{n_1\cdots n_d}(\varepsilon_i(\Xi_{\mu;n_1\cdots n_d}));\delta)\end{aligned}$$

Hence, to establish the tightness of $\{\mathbf{V}_n(\mathbf{S}_{n_1\cdots n_d}(\mathcal{E}(\Xi_{\mu;n_1\cdots n_d}))(B) : B \in \mathcal{B}(D)\}$ it suffices to show the component $\{\mathbf{S}_{n_1\cdots n_d}(\varepsilon_i(\Xi_{\mu;n_1\cdots n_d}))(B) : B \in \mathcal{B}(D)\}$ is tight, for $i = 1,\ldots,p$. The uniform central limit theorem investigated in [1] and [23] established the proof since convergence in distribution implies tightness. □

APPENDIX B. PROOF OF THEOREM 3.1

By the definition of $\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d})$ and by the linearity of the integral on the space $\mathcal{C}^p(\mathcal{A}_0)$, if H_0 is true we have

$$\begin{aligned}&\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}))\\ &= \int_D(\mathbf{f}_1-\mathbf{f}_0)^\top\mathbf{\Sigma}^{-1/2}d\mathbf{\Sigma}^{-1/2}\mathbf{S}_{n_1\cdots n_d}\\ &\quad(\frac{1}{\sqrt{n_1\cdots n_k}}\mathbf{g}_1(\Xi_{\mu;n_1\cdots n_d}))\\ &\quad+\int_D(\mathbf{f}_1-\mathbf{f}_0)^\top\mathbf{\Sigma}^{-1/2}d\mathbf{\Sigma}^{-1/2}\mathbf{S}_{n_1\cdots n_d}\\ &\quad(\frac{1}{\sqrt{n_1\cdots n_k}}\mathbf{f}_0(\Xi_{\mu;n_1\cdots n_d}))\\ &\quad+\int_D(\mathbf{f}_1-\mathbf{f}_0)^\top\mathbf{\Sigma}^{-1/2}d\mathbf{\Sigma}^{-1/2}\mathbf{S}_{n_1\cdots n_d}(\mathcal{E}(\Xi_{\mu;n_1\cdots n_d})).\end{aligned}$$

Since g_{1i} and f_{0i} are continuous and of bounded variation on D, for $i = 1,\ldots,p$, it holds $\mathbf{S}_{n_1\cdots n_d}(\frac{1}{\sqrt{n_1\cdots n_k}}\mathbf{g}_1(\Xi_{\mu;n_1\cdots n_d}))$ and $\mathbf{S}_{n_1\cdots n_d}(\frac{1}{\sqrt{n_1\cdots n_k}}\mathbf{f}_1(\Xi_{\mu;n_1\cdots n_d}))$ converge respectively to the set functions $\varphi_{\mathbf{g}_1}$ and $\varphi_{\mathbf{f}_1}$. The multivariate invariance principle (Theorem A.2) gives the result that $\mathbf{\Sigma}^{-1/2}\mathbf{S}_{n_1\cdots n_d}(\mathcal{E}(\Xi_{\mu;n_1\cdots n_d}))$ converges in distribution to $\mathbf{Z}_\mu$.

Hence, by applying the well-known continuous mapping theorem (cf. Theorem 27 in [6]), $\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}))$ converges in distribution to

$$\begin{aligned}&\int_D (\mathbf{f}_1-\mathbf{f}_0)^\top \mathbf{\Sigma}^{-1/2} d\mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{g}_1}\\&\quad+\int_D (\mathbf{f}_1-\mathbf{f}_0)^\top \mathbf{\Sigma}^{-1/2} d\mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{f}_0}\\&\quad+\int_D (\mathbf{f}_1-\mathbf{f}_0)^\top \mathbf{\Sigma}^{-1/2} d\mathbf{Z}_\mu\\&\qquad=0+\langle \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0),\mathbf{\Sigma}^{-1/2}\mathbf{f}_0\rangle_\mu^{(p)}\\&\qquad\quad+\int_D (\mathbf{f}_1-\mathbf{f}_0)^\top \mathbf{\Sigma}^{-1/2} d\mathbf{Z}_\mu.\end{aligned}$$

We notice that all integrals involved should be interpreted path wise as integral with respect to signed measure. The first term in the right hand side of the last equality follows from the fact that $\mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{g}_1}$ and $\mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{f}_0}$ are absolutely continuous with the $L_2^p(D,\mu)$-densities $\mathbf{\Sigma}^{-1/2}\mathbf{g}_1$ and $\mathbf{\Sigma}^{-1/2}\mathbf{f}_0$, respectively and $\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0)\perp \mathbf{\Sigma}^{-1/2}\mathbf{g}_1$. Furthermore, since

$$\int_D (\mathbf{f}_1-\mathbf{f}_0)^\top \mathbf{\Sigma}^{-1/2} d\mathbf{Z}_\mu \sim N\left(0,(\|\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0)\|_\mu^{(p)})^2\right)$$

then by applying the well known Lindeberg-Levy central limit theorem we get

$$\frac{\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}))-\langle \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0),\mathbf{\Sigma}^{-1/2}\mathbf{f}_0\rangle_\mu^{(p)}}{\|\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0)\|_\mu^{(p)}}\xrightarrow{\mathcal{D}} N(0,1).$$

Hence, by setting the constant k for

$$\begin{aligned}k:=\Phi^{-1}(1-\alpha)\|\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0)\|_\mu^{(p)}\\+\langle \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0),\mathbf{\Sigma}^{-1/2}\mathbf{f}_0\rangle_\mu^{(p)}\end{aligned}$$

we obtain an asymptotically size α test asserted in the theorem establishing the proof of the theorem.

APPENDIX C. PROOF OF COROLLARY 3.3

If $\mathbf{g}_2\equiv\mathbf{f}$ for any $\mathbf{f}\in\mathbf{V}^p\cap(\mathbf{W}^p)^C$, then we have

$$\begin{aligned}\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d})&=\frac{1}{\sqrt{n_1\cdots n_d}}\mathbf{g}_1(\Xi_{\mu;n_1\cdots n_d})\\&\quad+\frac{1}{\sqrt{n_1\cdots n_d}}\mathbf{f}(\Xi_{\mu;n_1\cdots n_d})+\mathcal{E}(\Xi_{\mu;n_1\cdots n_d}).\end{aligned}$$

By applying the similar argument as in the proof of Theorem 3.1, $\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}))$ converges in distribution to

$$\begin{aligned}&\int_D (\mathbf{f}_1-\mathbf{f}_0)^\top \mathbf{\Sigma}^{-1/2} d\mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{f}_0}+\int_D (\mathbf{f}_1-\mathbf{f}_0)^\top \mathbf{\Sigma}^{-1/2} d\mathbf{Z}_\mu\\&=\langle \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0),\mathbf{\Sigma}^{-1/2}\mathbf{f}\rangle_\mu^{(p)}+\int_D (\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0))^\top d\mathbf{Z}_\mu.\end{aligned}$$

Let k^* be a constant defined by $k^*:=\frac{k-\langle \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0),\mathbf{\Sigma}^{-1/2}\mathbf{f}\rangle_\mu^{(p)}}{\|\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0)\|_\mu^{(p)}}$, then by the Lindeberg-Levy central limit theorem, we get

$$\begin{aligned}&\Psi_\mu(\mathbf{f})\\&\quad=\lim_{n_1,\cdots,n_d\to\infty}P\{\omega\in\Omega:\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d})(\omega))\geq k|\mathbf{g}_2\equiv\mathbf{f}\}\\&\quad=\lim_{n_1,\cdots,n_d\to\infty}P\Big\{\frac{1}{\|\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0)\|_\mu^{(p)}}\\&\qquad\times[\mathcal{J}(\mathbf{Y}^{loc}(\Xi_{\mu;n_1\cdots n_d}))-\langle \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0),\mathbf{\Sigma}^{-1/2}\mathbf{f}\rangle_\mu^{(p)}]\geq k^*\Big\}\\&\quad=1-\Phi\left(\Phi^{-1}(1-\alpha)-\frac{\langle \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0),\mathbf{\Sigma}^{-1/2}(\mathbf{f}-\mathbf{f}_0)\rangle_\mu^{(p)}}{\|\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1-\mathbf{f}_0)\|_\mu^{(p)}}\right).\end{aligned}$$

We are done.

APPENDIX D. MULTIVARIATE CAMERON-MARTIN DENSITY

Theorem D.1. *The shifted probability distribution $P_{\mathbf{Z}_\mu}^{\mathbf{h}}$ is absolutely continuous with respect to $P_{\mathbf{Z}_\mu}$ on $\mathcal{C}^p(\mathcal{A}_0)$, if and only if $\mathbf{h}\in\mathcal{H}_{\mathbf{Z}_\mu}$. If $\mathbf{h}\in\mathcal{H}_{\mathbf{Z}_\mu}$ with $\mathbf{h}(A):=\int_A \ell\, d\mu$ for $A\in\mathcal{A}$ and $\ell:=(\ell_i)_{i=1}^p\in L_2^p(D,\mu)$, then*

$$\frac{dP_{\mathbf{Z}_\mu}^{\mathbf{h}}}{dP_{\mathbf{Z}_\mu}}(\mathbf{x})=\exp\left\{\int_D \ell^\top d\mathbf{x}-\frac{1}{2}\|\mathbf{h}\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}}\right\},$$
$$\mathbf{x}:=(x_1,\ldots,x_p)^\top\in\mathcal{C}^p(\mathcal{A}_0)$$

where $\int_D \ell^\top d\mathbf{x}:=\sum_{i=1}^p\int_D \ell_i dx_i$ is interpreted path wise as the integral with respect to signed measure. See [19], pp. 13–15 for the definition of integral involving signed measure.

Proof. The proof of the theorem is mainly based on the univariate case presented in Theorem 5.1 of [19] and the stochastically independence between the components of $\mathbf{Z}_\mu$. To the first assertion, suppose $\mathbf{h}\in\mathcal{H}_{\mathbf{Z}_\mu}$. Then by the definition of $\mathcal{H}_{\mathbf{Z}_\mu}$, it holds $h_i\in\mathcal{H}_{Z_\mu^{(i)}}$ for all $i=1,\ldots,p$, which implies $P^{h_i}_{Z_\mu^{(i)}}$ is absolutely continuous w.r.t. $P_{Z_\mu^{(i)}}$ on $\mathcal{C}(\mathcal{A}_0)$ (c.f. Theorem 5.1 of [19]). There exists a $\gamma_i\in L_2(\mathcal{C}(\mathcal{A}_0),P_{Z_\mu^{(i)}})$ s.t. $P^{h_i}_{Z_\mu^{(i)}}(B_i)=\int_{B_i}\gamma_i\, dP_{Z_\mu^{(i)}}$, for $B_i\subset\mathcal{C}(\mathcal{A}_0)$. Hence, by Definition 4.1 and the independence of $Z_\mu^{(i)}$ we have

$$\begin{aligned}&P_{\mathbf{Z}_\mu}^{\mathbf{h}}(\times_{i=1}^p B_i)\\&\quad=P_{\mathbf{Z}_\mu}(\times_{i=1}^p B_i-\mathbf{h})=P_{\mathbf{Z}_\mu}(\times_{i=1}^p(B_i-h_i))\\&\quad=\Pi_{i=1}^p P^{h_i}_{Z_\mu^{(i)}}(B_i)=\int_B(\Pi_{i=1}^p\gamma_i)\, dP_{Z_\mu^{(1)}}\cdots\, dP_{Z_\mu^{(p)}}.\end{aligned}$$

Conversely, let $P_{\mathbf{Z}_\mu}^{\mathbf{h}}$ be absolutely continuous w.r.t. $P_{\mathbf{Z}_\mu}$. Since we can write $P^{h_i}_{Z_\mu^{(i)}}(B_i)$ for all $i=1,\ldots,p$, as

$$\begin{aligned}P^{h_i}_{Z_\mu^{(i)}}(B_i)=P_{\mathbf{Z}_\mu}^{\mathbf{h}}(\mathcal{C}(\mathcal{A}_0)\times\cdots\times\mathcal{C}(\mathcal{A}_0)\times B_i\\\times\mathcal{C}(\mathcal{A}_0)\times\cdots\times\mathcal{C}(\mathcal{A}_0)),\end{aligned}$$

for $\mathbf{h} = (0, \ldots, 0, h_i, 0, \ldots, 0)^\top$, it can be concluded that $P^{h_i}_{Z^{(i)}_\mu}$ is absolutely continuous w.r.t. $P_{Z^{(i)}_\mu}$, for all $i = 1, \ldots, p$. Hence by Theorem 5.1 of [19] $h_i \in \mathcal{H}_{Z^{(i)}_\mu}$, for all $i = 1, \ldots, p$, establishing the first statement. To proof the second one we move further by recalling Definition 4.2 and the independence of $Z^{(i)}_\mu$ that

$$
\begin{aligned}
&P^{\mathbf{h}}_{\mathbf{Z}_\mu}(\times_{i=1}^p B_i) \\
&\quad = \Pi_{i=1}^p P^{h_i}_{Z^{(i)}_\mu}(B_i) \\
&\quad = \Pi_{i=1}^p \int_{B_i} \exp\left\{ \int_D \ell_i \, dx_i - \frac{1}{2}\|h_i\|^2_{\mathcal{H}_{Z^{(i)}_\mu}} \right\} \, dP_{Z^{(i)}_\mu} \\
&\quad = \int_B \exp\left\{ \int_D \ell^\top \, d\mathbf{x} - \frac{1}{2}\|\mathbf{h}\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}} \right\} \, dP_{\mathbf{Z}_\mu}
\end{aligned}
$$

completing the proof of the theorem. □

APPENDIX E. PROOF OF THEOREM 4.3

Since $\mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{g}} = \varphi_{\mathbf{\Sigma}^{-1/2}\mathbf{g}}$, then by directly applying Theorem D.1, under H_0 and H_1 we respectively have the following density formulas

$$
\begin{aligned}
\psi_0(\mathcal{Y}) = \exp\Bigg\{ &\int_D (\mathbf{g}_1 + \mathbf{f}_0)^\top \mathbf{\Sigma}^{-1/2} d\mathcal{Y} \\
&- \frac{1}{2}\|\mathbf{\Sigma}^{-1/2}\varphi_{(\mathbf{g}_1+\mathbf{f}_0)}\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}} \Bigg\}
\end{aligned}
$$

and

$$
\begin{aligned}
\psi_1(\mathcal{Y}) = \exp\Bigg\{ &\int_D (\mathbf{g}_1 + \mathbf{f}_1)^\top \mathbf{\Sigma}^{-1/2} d\mathcal{Y} \\
&- \frac{1}{2}\|\mathbf{\Sigma}^{-1/2}\varphi_{(\mathbf{g}_1+\mathbf{f}_1)}\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}} \Bigg\}.
\end{aligned}
$$

By further recalling the assumption that $\varphi_{\mathbf{g}_1}$ is orthogonal to $\varphi_{\mathbf{f}_0}$ and $\varphi_{\mathbf{f}_1}$ in $\mathcal{H}_{\mathbf{Z}_\mu}$, we get the equations $\|\mathbf{\Sigma}^{-1/2}\varphi_{(\mathbf{g}_1+\mathbf{f}_0)}\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}} = \|\mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{g}_1}\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}} + \|\mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{f}_0}\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}}$ and $\|\mathbf{\Sigma}^{-1/2}\varphi_{(\mathbf{g}_1+\mathbf{f}_1)}\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}} = \|\mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{g}_1}\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}} + \|\mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{f}_1}\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}}$. Hence the ratio between $\psi_0(\mathcal{Y})$ and $\psi_1(\mathcal{Y})$ can be further simplified and we obtain the expression

$$
\begin{aligned}
\frac{\psi_0(\mathcal{Y})}{\psi_1(\mathcal{Y})} = \exp\Bigg\{ &\int_D (\mathbf{f}_0 - \mathbf{f}_1)^\top \mathbf{\Sigma}^{-1/2} d\mathcal{Y} \\
&+ \frac{1}{2}\left(\|\mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{f}_1}\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}} - \|\mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{f}_0}\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}} \right) \Bigg\}.
\end{aligned}
$$

Neyman-Pearson theorem (cf. Theorem 3.2.1 in [18]) guarantees that η_μ will become a most powerful test of size $\alpha \in (0, 1)$, if there exists a constant k that satisfies

$$
\mathbb{P}\left\{ \omega \in \Omega : \frac{\psi_0(\mathcal{Y}(\omega))}{\psi_1(\mathcal{Y}(\omega))} \leq k | H_0 \right\} = \alpha.
$$

However, since $\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0) \perp \mathbf{\Sigma}^{-1/2}\mathbf{g}_1$ in $L_2^p(D, \mu)$, when H_0 is true, the integral involved in the formula of the ratio between $\psi_0(\mathcal{Y})$ and $\psi_1(\mathcal{Y})$ can be written as

$$
\begin{aligned}
\int_D (\mathbf{f}_0 - \mathbf{f}_1)^\top \mathbf{\Sigma}^{-1/2} d\mathcal{Y} &= \int_D (\mathbf{f}_0 - \mathbf{f}_1)^\top \mathbf{\Sigma}^{-1/2} d\varphi_{\mathbf{\Sigma}^{-1/2}\mathbf{g}_1} \\
+ \int_D (\mathbf{f}_0 - \mathbf{f}_1)^\top \mathbf{\Sigma}^{-1/2} &d\varphi_{\mathbf{\Sigma}^{-1/2}\mathbf{f}_0} + \int_D (\mathbf{f}_0 - \mathbf{f}_1)^\top \mathbf{\Sigma}^{-1/2} d\mathbf{Z}_\mu \\
= -\int_D (\mathbf{f}_1 - \mathbf{f}_0)^\top \mathbf{\Sigma}^{-1/2} &d\mathbf{Z}_\mu - \langle \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0), \mathbf{\Sigma}^{-1/2}\mathbf{f}_0 \rangle^{(p)}_\mu.
\end{aligned}
$$

This leads us to get

$$
\begin{aligned}
&\mathbb{P}\left\{ \frac{\psi_0(\mathcal{Y})}{\psi_1(\mathcal{Y})} \leq k | H_0 \right\} = \alpha \\
&\Leftrightarrow \mathbb{P}\left\{ \int_D (\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0))^\top \, d\mathcal{Y} \geq \kappa | H_0 \right\} = \alpha \\
&\Leftrightarrow \mathbb{P}\left\{ \int_D (\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0))^\top d\mathbf{Z}_\mu \geq \kappa^* \right\} = \alpha,
\end{aligned}
$$

where κ and κ^* are constants defined respectively as

$$
\begin{aligned}
\kappa &:= -\ln(k) + \frac{1}{2}\left\| \mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{f}_1} \right\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}} - \frac{1}{2}\left\| \mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{f}_0} \right\|^2_{\mathcal{H}_{\mathbf{Z}_\mu}} \\
\kappa^* &:= \kappa - \langle \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0), \mathbf{\Sigma}^{-1/2}\mathbf{f}_0 \rangle^{(p)}_\mu.
\end{aligned}
$$

Moreover, since

$$
\int_D (\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0))^\top d\mathbf{Z}_\mu \sim N\left(0, (\|\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0)\|^{(p)}_\mu)^2\right),
$$

then we must set the lower bound presented above for the $(1-\alpha)$-th quantile of the standard normal distribution. In other word we have the equation

$$
\begin{aligned}
&\frac{\kappa - \langle \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0), \mathbf{\Sigma}^{-1/2}\mathbf{f}_0 \rangle^{(p)}_\mu}{\left\| \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0) \right\|^{(p)}_\mu} = \Phi^{-1}(1-\alpha) \\
&\Leftrightarrow \kappa = \Phi^{-1}(1-\alpha) \left\| \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0) \right\|^{(p)}_\mu \\
&\qquad + \langle \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0), \mathbf{\Sigma}^{-1/2}\mathbf{f}_0 \rangle^{(p)}_\mu.
\end{aligned}
$$

We notice that the original constant κ could be computed if desired, but it is not necessary in order to perform the tests. The power function of the Neyman-Pearson test of size α evaluated at $\varphi_{\mathbf{f}}$ is defined as

$$
\Upsilon_\mu(\varphi_{\mathbf{f}}) := P\left\{ \int_D (\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0))^\top d\mathcal{Y} \geq \kappa | \ \varphi_{\mathbf{g}_2} \equiv \varphi_{\mathbf{f}} \right\}.
$$

If the true model is $\mathcal{Y} = \mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{g}_1} + \mathbf{\Sigma}^{-1/2}\varphi_{\mathbf{f}} + \mathbf{Z}_\mu$, then we get

$$
\begin{aligned}
\Upsilon_\mu(\varphi_{\mathbf{f}}) = P\Bigg\{ &\frac{\int_D (\mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0))^\top d\mathbf{Z}_\mu}{\left\| \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0) \right\|^{(p)}_\mu} \\
&\geq \frac{\kappa - \langle \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0), \mathbf{\Sigma}^{-1/2}\mathbf{f} \rangle^{(p)}_\mu}{\left\| \mathbf{\Sigma}^{-1/2}(\mathbf{f}_1 - \mathbf{f}_0) \right\|^{(p)}_\mu} \Bigg\}.
\end{aligned}
$$

By substituting κ we obtain the expression stated in the theorem.

APPENDIX F. FUNCTION OF BOUNDED VARIATION ON D

Definition F.1. Let $f : D := \times_{k=1}^{d}[a_k, b_k] \to \mathbb{R}$ be a real valued function with d variables. For a_k, b_k, let $\Delta_{a_k}^{b_k} f$ be a real-value function defined on D, given by

$$\Delta_{a_k}^{b_k} f := f(x_1, \ldots, x_{k-1}, b_k, x_{k+1}, \ldots, x_d) - f(x_1, \ldots, x_{k-1}, a_k, x_{k+1}, \ldots, x_d),$$

for $k = 1, \ldots, d$. Furthermore, for $\mathbf{a} := (a_k)_{k=1}^{d}, \mathbf{b} := (b_k)_{k=1}^{d} \in D$, $\Delta_{\mathbf{a}}^{\mathbf{b}} f$ is defined on D recursively starting from the last components of $\mathbf{a}$ and $\mathbf{b}$. More precisely,

$$\Delta_{\mathbf{a}}^{\mathbf{b}} f := \Delta_{a_1}^{b_1}(\cdots(\Delta_{a_{d-1}}^{b_{d-1}}(\Delta_{a_d}^{b_d} f))\cdots).$$

Let $\{j_1, \cdots, j_d\}$ be a permutation of $\{1, 2, \ldots, d\}$, then it holds

$$\Delta_{\mathbf{a}}^{\mathbf{b}} f = \Delta_{a_{j_1}}^{b_{j_1}}(\cdots(\Delta_{a_{j_{d-1}}}^{b_{j_{d-1}}}(\Delta_{a_{j_d}}^{b_{j_d}} f))\cdots) = \Delta_{a_1}^{b_1}(\cdots(\Delta_{a_{d-1}}^{b_{d-1}}(\Delta_{a_d}^{b_d} f))\cdots).$$

This means that the operation of $\Delta_{\mathbf{a}}^{\mathbf{b}} f$ does not depend on the order. By this reason we write $\Delta_{\mathbf{a}}^{\mathbf{b}} f$ by $\Delta_{a_1}^{b_1} \cdots \Delta_{a_{d-1}}^{b_{d-1}} \Delta_{a_d}^{b_d} f$ ignoring the brackets. The reader is referred to [36] and [15], pp. 44–45.

Definition F.2. (Yeh [36]) Let $\Gamma_k := \{[x_{k_0}, x_{k_1}], [x_{k_1}, x_{k_2}], \ldots, [x_{k_{M_k-1}}, x_{k_{M_k}}]\}$ be a collection of M_k rectangles on the closed interval $[a_k, b_k]$ with $a_k = x_{k_0} \le x_{k_1} \le \ldots \le x_{k_{M_k}} = b_k$, for $k = 1, \ldots, d$. The Cartesian product $\mathcal{K} := \times_{k=1}^{d}\Gamma_k$ which consists of $\Pi_{k=1}^{d} M_k$ closed rectangles is called a non-overlapping finite exact cover of D. The family of all non-overlapping finite exact cover of D is denoted by $\mathcal{J}(\mathcal{K})$.

Definition F.3. (Yeh [36]) For $1 \le w_k \le M_k$, with $k = 1, \ldots, d$, let $\mathbf{J}_{w_1 \cdots w_d}$ be the element of $\mathcal{K}$ defined by $\mathbf{J}_{w_1 \cdots w_d} := \times_{k=1}^{d}[x_{k_{w_k-1}}, x_{k_{w_k}}]$. Let $\psi : D \to \mathbb{R}$ be a real valued function on D. Operator $\Delta_{\mathbf{J}_{w_1 \cdots w_d}}$ acting on a function ψ is defined by

$$\Delta_{\mathbf{J}_{w_1 \cdots w_d}} \psi := \Delta_{x_{1_{w_1-1}}}^{x_{1_{w_1}}} \Delta_{x_{2_{w_2-1}}}^{x_{2_{w_2}}} \cdots \Delta_{x_{d_{w_d-1}}}^{x_{d_{w_d}}} \psi.$$

The variation of ψ over the finite exact cover $\mathcal{K}$ is defined by

$$v(\psi; \mathcal{K}) := \sum_{w_1=1}^{M_1} \cdots \sum_{w_d=1}^{M_d} \left| \Delta_{\mathbf{J}_{w_1 \cdots w_d}} \psi \right|.$$

Accordingly, the total variation of ψ over D is defined by

$$V(\psi; D) := \sup_{\mathcal{K} \in \mathcal{J}(\mathcal{K})} v(\psi; \mathcal{K}).$$

Furthermore, the function ψ is said to have bounded variation in the sense of Vitaly on D if there exists a real number $M > 0$ s.t. $V(\psi; D) \le M$ for some real number $M > 0$. The class of such functions is denoted by $BVV(\mathbf{I}^d)$.

Definition F.4. (Yeh [36]) Let $(x_k)_{k=1}^{d}$ be a variable in D. For a fixed k, let D^k be a k-dimensional closed rectangle constructed in the following way. We choose $d-k$ components of the variable $(x_k)_{k=1}^{d}$. For each choice from all possible elements of the set C_{d-k}^{d}, we set each x_i with a_i or b_i and let the remaining k variables to satisfy $a_i \le x_i \le b_i$. Then for each k we get $2^{d-k}|C_{d-k}^{d}|$ k-dimensional closed rectangles D^k. For convention we denote the collection of all $2^{d-k}|C_{d-k}^{d}|$ of closed rectangles D^k by $\mathcal{B}^k$ and the j-th element of $\mathcal{B}^k$ will be denoted by D_j^k. A function ψ is said to have bounded variation in the sense of Hardy on D, if and only if for each $k = 1, \ldots, d$ and $j = 1, \ldots, 2^{d-k}|C_{d-k}^{d}|$, there exists a real number $M_{jk} > 0$ s.t. $V(\psi_{D_j^k}(\cdot); D_j^k) \le M_{jk}$, where for $k = 1, \ldots, d$ and $j = 1, \ldots, 2^{d-k}|C_{d-k}^{d}|$, $\psi_{D_j^k}(\cdot)$ is a function with k variables obtained from the function $\psi(x_1, x_2 \ldots, x_d)$ by setting the $d-k$ selected variables with a_j or b_j, whereas the remaining k variables lies in the interval $[a_k, b_k]$. The class of such functions will be denoted by $BV_H(D)$.

APPENDIX G. PROOF OF THEOREM 5.3

Suppose we consider a p-variate model $Y_i(\mathbf{x}) = \sum_{j=1}^{m} \beta_{ij} w_j(\mathbf{x}) + \mathcal{E}_i(\mathbf{x})$, for $\mathbf{x} \in D \subset \mathbb{R}^d$, $i = 1, \ldots, p$, where $\{w_1, \ldots, w_q, w_{(q+1)}, \ldots, w_m\}$ forms a basis of $\mathbf{V}$ which can be considered as functions in $L_2(D, \lambda^d)$ as well as in $L_2(D, \mu)$ for arbitrary fixed $\mu \in \mathcal{G}$. Let $\{\widetilde{w}_{1\lambda^d}, \ldots, \widetilde{w}_{q\lambda^d}, \ldots, \widetilde{w}_{m\lambda^d}\}$ and $\{\widetilde{w}_{1\mu}, \ldots, \widetilde{w}_{q\mu}, \ldots, \widetilde{w}_{m\mu}\}$ be the Gram-Schmidt orthogonal versions of the original bases when the sampling is conducted under the designs λ^d and μ, respectively. It can be easily shown there exist constants $c_{1k\mu}, \ldots, c_{mk\mu}$ which depend on μ, s.t. $\widetilde{w}_{k\mu} = \sum_{j=1}^{m} c_{jk\mu} \widetilde{w}_{j\lambda^d}$, for $k = 1, \ldots, m$. Let $\hat{f}_\mu$ be the density of μ w.r.t. to λ^d. Note that by the assumption there exists $F_{k\mu}$, such that $\hat{f}_\mu(t_k) := \Pi_{k=1}^{d} F'_{k\mu}(t_k) > 0$ and it holds $\|\hat{f}_\mu\|_{\lambda^d} \le \|\hat{f}_\mu\|_\infty |D| =: C$. Then by the definition of $\|\cdot\|_\mu$ and Holder's inequality, we get

$$\begin{aligned}
&\|\widetilde{w}_{(q+1)\mu}\|_\mu^2 + \cdots + \|\widetilde{w}_{m\mu}\|_\mu^2 \\
&\quad = \int_D (\widetilde{w}_{(q+1)\mu}^2 \hat{f}_\mu) d\lambda^d + \cdots + \int_D (\widetilde{w}_{m\mu}^2 \hat{f}_\mu) d\lambda^d \\
&\quad \le \left(\|\widetilde{w}_{(q+1)\mu}^2\|_{\lambda^d} + \cdots + \|\widetilde{w}_{m\mu}^2\|_{\lambda^d} \right) \|\hat{f}_\mu\|_{\lambda^d} \\
&\quad \le \sum_{j,\ell=1}^{m} (\|c_{j(q+1)\mu} c_{\ell(q+1)\mu} \widetilde{w}_{j\lambda^d} \widetilde{w}_{\ell\lambda^d}\|_{\lambda^d} + \cdots \\
&\qquad + \|c_{jm\mu} c_{\ell m\mu} \widetilde{w}_{j\lambda^d} \widetilde{w}_{\ell\lambda^d}\|_{\lambda^d}) C \\
&\quad \le \sum_{j=1}^{m} \sum_{\ell=1}^{m} \left(M^2 \left\| \widetilde{w}_{j\lambda^d} \widetilde{w}_{\ell\lambda^d} \right\|_{\lambda^d} + \cdots \right.
\end{aligned}$$

$$+ M^2 \left\| \widetilde{w}_{j\lambda^d} \widetilde{w}_{\ell\lambda^d} \right\|_{\lambda^d}) C$$

$$\leq M^2 (m-q) C \left(\sum_{j=1}^{m} \sum_{\ell=1}^{m} \left\| \widetilde{w}_{j\lambda^d} \widetilde{w}_{\ell\lambda^d} \right\|_{\lambda^d} \right),$$

where $M := \sup_{\mu \in \mathcal{G}} \max_{1 \leq j \leq m;\ q+1 \leq k \leq m} |c_{jk\mu}|$. The constant M is well-defined, since for $j = 1, \ldots, m$ and $k = q+1, \ldots, m$, $c_{jk\mu}$ is a continuous function of μ on $\mathcal{G}$ w.r.t. the uniform topology. The upper bound on the right-hand side of the last inequality is calculated using the integral w.r.t. λ^d only, therefore the Lebsesque measure λ^d is the design such that $\sup_{\mu \in \mathcal{G}} \mathcal{W}_{\mathbf{f}}(\mu) = \mathcal{W}_{\mathbf{f}}(\lambda^d)$, finishing the proof.

Received 22 August 2016

REFERENCES

[1] Alexander, K. S. and Pyke, R. (1968). A uniform central limit theorem for set-indexed partial-sum processes with finite variance, *Annals Probab.* **14** 582–597. MR0832025

[2] Arnold, S. F. (1981). *The Theory of Linear Models and Multivarite Analysis*, John Wiley & Sons, Inc., New York. MR0606011

[3] Arnold, S. F. (1984). Asymptotic validity of invariant procedures for the repeated measured model and multivariate linear model, *Journal of multivariate analysis* **15** 325–335. MR0768501

[4] Athreya, K. B. and Lahiri, S. N. (2006). *Measure Theory and Probability Theory*, Springer Science+Business Media, LLC., New York. MR2247694

[5] Berlinet, A. and Agnan, C. T. (2003). *Reproducing Kernel Hilbert Spaces in Probability and Statistics*, Springer Verlag New York.

[6] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. Wiley, New York.

[7] Bischoff, W. (1998). A functional central limit theorem for regression models, *Ann. Stat.* **6** 1398–1410.

[8] Bischoff, W. (2002). The structure of residual partial sums limit processes of linear regression models, *Theory of Stochastic Processes* **2** 23–28. MR2026252

[9] Bischoff, W. and Gegg, A. (2016). The Cameron-Martin theorem for ($p-$)Slepian processes, *J Theor Probab.* **29** 707–715. MR3500417

[10] Bischoff, W. and Miller, F. (2000). Asymptotically optimal test and optimal designs for testing the mean in regression model with application to change-point problem, *Ann. Inst. Statist. Math.* **52**(4) 658–679. MR1820743

[11] Bischoff, W. and Miller, F. (2006). Optimal design which are efficient for lack of fit tests, *Ann. Statis.* **34**(2) 2015–2025.

[12] Bischoff, W. and Somayasa, W. (2009). The limit of the partial sums process of spatial least squares residuals, *J. Multivariate Analysis* **100** 2167–2177. MR2560361

[13] Christensen, R. (2001). *Linear Models for Multivariate, Time Series, and Spatial Data*, Springer-Verlag New York, Inc., New York.

[14] Cohn, D. L. (1980). *Measure Theory*, Birkhäuser, Inc., Boston.

[15] Elstrodt, J. (2011). *Maß- und Inegrationstheorie*, 7., korregierte und aktualisierte Auflage. Berlin: Springer.

[16] Johnson, R. A. and Wichern, D. W. (1992). *Applied Multivariate Statistical Analysis* (3rd edition), Prentice-Hall, Inc., New York.

[17] Gaenssler, P. (1993). On recent development in the theory of set-indexed processes (A unified approach to empirical and partial-sum processes) in *Asymptotic Statistics*, Springer, Berlin.

[18] Lehmann, E. L. and Romano, J. P. (2005). *Testing Statitical Hypotheses*, 3rd edn., Springer, New York.

[19] Lifshits, M. (2012). *Lectures on Gaussian Processes*, Springer Briefs in Mathematics, Springer, Berlin.

[20] MacNeill, I. B. (1978). Properties of partial sums of polynomial regression residuals with applications to test for change of regression at unknown times, *Ann. Statist.* **6** 422–433.

[21] MacNeill, I. B. (1978). Limit processes for sequences partial sums of regression residuals, *Ann. Probab.* **6** 695–698.

[22] MacNeill, I. B. and Jandhyala, V. K. (1993). *Change-point methods for spatial data, Multivariate Environmental Statistics eds. by G. P. Patil and C. R. Rao*, Elsevier Science Publishers B.V., pp. 298–306.

[23] Pyke, R. (1983). A uniform central limit theorem for partial sum processes indexed by sets, *Ann. Probab.* **79** 219–240.

[24] Seber, G. A. F. and Lee, A. J. (2003). *Linear Regression Analysis (2nd ed.)*, John Wiley & Sons, New Jersey. MR1958247

[25] Stroock, D. W. (1994). *A Concise Introduction to the Theory of Integration* (2nd edition), Birkhäuser, Berlin.

[26] Somayasa, W., Ruslan, Cahyono, E. and Engkoimani, L. O. (2015). Cheking adequateness of spatial regressions using set-indexed partial sums technique, *Fareast Journal of Mathematical Sciences* **96**(8) 933–966.

[27] Somayasa, W., Adi Wibawa, G. N. and Pasolon, Y. B. (2015). Multidimensional set-indexed partial sums method for checking the appropriateness of a multivariate spatial regression *Int. J. Mathematical Model and Method in Applied Sciences* **9** 700–713.

[28] Somayasa, W. and Adi Wibawa, G. N. (2015). Asymptotic model-check for multivariate spatial regression with correlated responses *Fareast Journal of Mathematical Sciences* **98**(5) 613–939.

[29] Somayasa, W., Adi Wibawa, G. N., Hamimu, La. and Ngkoimani, L. O. (2016). Asymptotic theory in model diagnostic for general multivariate spatial regression *Int. J. Mathematics and Mathematical Sciences* **2016** 1–16. MR3548419

[30] Stute, W. (1997). Nonparametric model checks for regression, *Ann. Statist.* **25** 613–641.

[31] Stute, W., Gonzalez Manteiga, W. and Presedo Quindimil, M. (1998). Bootstrap approximations in model checks for regression, *J. Amer. Statist. Assoc.* **93**(441) 141–149.

[32] Stute, W., Xu, W. L. and Zhu, L. X. (2008). Model diagnostic for parametric regression in high-dimensional space, *Biometrika*, **95**(2) 451–467.

[33] Tahir, M. (2010). Prediction of the amount of nickel deposit based on the results of drilling bores on several points (case study: south mining region of PT. Aneka Tambang Tbk., Pomalaa, Southeast Sulawesi), research report, Halu Oleo University, Kendari.

[34] Van der Vaart, A. W. (1998). *Asymptotic Statistics*, Cambridge University Press, Cambridge. MR1652247

[35] Wellner, J. A. (2003). Gaussian white noise models: Some results for monotone functions. Crossing boundaries: Statistical essays in honor of Jack Hall, Institute of Mathematical Statistics, pp. 87–104.

[36] Yeh, J. (1963). Cameron-Martin translation theorem in the Wiener space of functions of two variables, *Trans. Amer. Math. Soc.* **107**(3) 409–420.

[37] Xie, L. and MacNeill, I. B. (2006). Spatial residual processes and boundary detection, *South African Statist. J.* **4** 33–53.

Wayan Somayasa
The Department of Mathematics
Halu Oleo University
Indonesia
E-mail address: wayan.somayasa@uho.ac.id

Herdi Budiman
Department of Mathematics
Halu Oleo University
Indonesia
E-mail address: herdi.budiman@uho.ac.id