Model-Checks Based on Least Squares Residual Partial Sums Processes

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

von der Fakultät für Mathematik der Universität Karlsruhe (TH) genehmigte

DISSERTATION

von

Wayan Somayasa, M.Si.

aus Jinengdalem, Buleleng (Indonesien)

Tag der mündlichen Prüfung: 18.07.2007

Referent: Prof. Dr. Wolfgang Bischoff

Korreferent: Prof. Dr. Norbert Henze

Karlsruhe 2007

Dedicated to my wife Sutiari, my son David and my father.
To the memories of my mother.

Acknowledgement

I would like to take this opportunity to express my thanks to Prof. Dr. Wolfgang Bischoff, my thesis advisor, who guided me meticulously through the research work.

I would like to express my special thanks to Prof. Dr. Norbert Henze, the second reader of this thesis, for the encouraging support and comments during this research. He was also very understanding and cooperative.

I acknowledge the financial support from the *German Academic Exchange Service* (DAAD), which has awarded me for the period October 2003 - March 2007.

I wish to thank my colleagues in the Institute of Stochastics for providing me with an excellent and friendly working environment. In particular, I would like to thank Dr. Bernhard Klar, Dipl.-Math. oec. Volker Baumstark, Dipl.-Math. oec. André Mundt and Dr. Matthias Heveling for sharing with me their knowledge in statistics and stochastics.

I would like to thank Apl. Prof. Dr. Rudolf Lohner, the chief of *Application and Software* division of the central computing center (RZ) of Karlsruhe university for providing me the scientific supercomputer for the simulations.

I express my deepest gratitude to my parents, my families in Indonesia, my wife and my son for their love and patience in guiding me during my stay in Germany. Without them this work would never have come into existence (literally).

Finally, I thank *Hyang Widi Wasa* (Almighty God) for all the grace and blessing that He has showered on me to reach this stage of my academic and personal life.

Karlsruhe, July, 2007

Table of Contents

Introduction

In mathematical statistics we learn about several statistical models which have many areas of application such as in Agriculture, Biology, Economic, Environmental science, Engineering, etc. In this work we specifically put our attention on a linear model for *spatial* data. In what follows we present a brief description by an example in meteorology to give a motivation how linear models (*empirical model building*) come into account of statistical methods for spatial data analysis and why we are interested in this subject.

Figure 1. U.S. Weather Stations. Source: National Climate Data Center.

In order to determine and quantify the functional relationship between geographical positions and the air temperatures in the USA, an experiment was conducted that consisted in recording independently air temperatures from $m \in \mathbb{N}$, say, weather stations installed in several different positions according to Figure 1. We denote the experimental region of this experiment as $\mathbf{D} \subset \mathbb{R}^2$. Suppose that the true unknown value of the air temperature recorded at a position with *cartesian* coordinate $(x, y) \in \mathbf{D}$, is given by an unknown function $f(x, y)$. This function may be either a first order or a second order polynomial or exponential in x and y. It is reasonable to treat the data obtained from this experiment as a *geostatistical* data since the experimental domain $\mathbf{D}$ is *continuous* and fixed, see Schabenberger and Gotway (2005), p. 7, for the notion of geostatistical data. Since it is clearly impossible to determine this function analytically (without knowing additional information about the physical characteristics of these phenomena), we may, based on data, approximate this function empirically. Let $Z(\mathbf{s}_i)$ be the air temperature recorded from the station with geographical position $\mathbf{s}_i \in \mathbf{D}$, $i = 1, \ldots, m$. We may regard these observations as a realization of a stochastic process (*random field*) $\{Z(x, y) : (x, y) \in \mathbf{D}\}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, say. Because experimental error is inherent in experiments involving measurement, we can assume that for every point $(x, y) \in \mathbf{D}$, the *response* variable $Z(x, y)$ can be decomposed as $Z(x, y) = f(x, y) + \varepsilon(x, y)$, i.e., the randomness of $Z(x, y)$ is contributed only by the random error $\varepsilon(x, y)$. Hence, the air temperature that is actually observed or measured at any particular position $\mathbf{s}_i \in \mathbf{D}$ can be written as $Z(\mathbf{s}_i) = f(\mathbf{s}_i) + \varepsilon(\mathbf{s}_i)$. The random error $\varepsilon(\mathbf{s}_i)$ represents the random difference between the observed and the true air temperatures at $\mathbf{s}_i$, $i = 1, \ldots, m$. Introducing the vectors $\mathbf{Z} := (Z(\mathbf{s}_1), \ldots, Z(\mathbf{s}_m))^\top$, $\mathbf{\Theta} := (f(\mathbf{s}_1), \ldots, f(\mathbf{s}_m))^\top$, and $\mathbf{E} := (\varepsilon(\mathbf{s}_1), \ldots, \varepsilon(\mathbf{s}_m))^\top$, the model may be written as $\mathbf{Z} = \mathbf{\Theta} + \mathbf{E}$.

We have to give reasonable assumptions to the model, so that further study and analysis can be conducted. In particular, reasonable assumptions muss be made on the unknown function $f(\cdot)$ and the random error $\varepsilon(\cdot)$. For example $f(\cdot)$ is sometimes assumed to be continuous and smooth function on $\mathbf{D}$ so that it can be approximated by a polynomial function of low order. If we assume that there exist functions

$g_1(\cdot), \ldots g_p(\cdot)$ defined on $\mathbf{D}$ such that $f(\mathbf{s}_i) = \sum_{j=1}^{p} \beta_j g_j(\mathbf{s}_i)$ for some unknown constants $\beta_1, \ldots, \beta_p$, we speak about a *linear regression* model (in a spatial data analysis context this model is called *universal kriging* model). In other word linear regression (universal kriging) model concerns on the assumption that $\mathbf{\Theta} \in \mathbf{V}$ for a subspace $\mathbf{V} \subset \mathbb{R}^m$, with $\mathbf{V} := [\mathbf{g}_1, \ldots, \mathbf{g}_p]$, i.e., $\mathbf{V}$ is a subspace of $\mathbb{R}^m$ spanned by the vectors $\{\mathbf{g}_1, \ldots, \mathbf{g}_p\}$, where $\mathbf{g}_i := (g_i(\mathbf{s}_1), \ldots, g_i(\mathbf{s}_m))^\top \in \mathbb{R}^m$, $i = 1, \ldots, p$. Moreover if we assume that $f(\mathbf{s}_i) = \mu$, for an unknown $\mu \in \mathbb{R}$ and $i = 1, \ldots, p$, the model is called a *constant* (*ordinary kriging*) model, see e.g., Christensen (1991), p. 263. Usually the random errors $\varepsilon(\mathbf{s}_i)$, $i = 1, \ldots, m$, are assumed to be independent and have some distribution on $\mathbb{R}$, e.g., a normal distribution with mean 0 and unknown finite variance $\sigma^2(\mathbf{s}_i)$. This means that the variance depends on the coordinate of the point $\mathbf{s}_i \in \mathbf{D}$. Indeed, this assumption is reasonable for the weather station experiment above since the air temperatures are recorded from station to station independently, so that under this assumption we investigate the model $\mathbf{Z} = \mathbf{\Theta} + \mathbf{E}$, with $\mathbb{E}(\mathbf{E}) = \mathbf{0}$ and $Cov(\mathbf{E}) = diag(\sigma^2(\mathbf{s}_1), \ldots, \sigma^2(\mathbf{s}_m))$. In many situations it may be unrealistic to assume that the observations are uncorrelated. For instance, we consider soil carbon regression model in Schabenberger and Gotway (2005), p. 321-352, for which it is assumed that $Cov(\varepsilon(\mathbf{s}_i), \varepsilon(\mathbf{s}_j)) = \sigma^2 \exp\{-\|\mathbf{s}_i - \mathbf{s}_j\| / \theta\}$ for $i \neq j$, where $\sigma^2 \in (0, \infty)$ and $\theta \in \mathbb{R}$ are unknown. Thus we get a more complicated model than before. Consequently, we need a more complicated statistical procedure for investigating such a model.

After fitting the model to the data, a further preliminary statistical analysis addressed to this model may include *model-checks* which are intended to check (based on the data) the adequateness of the model, i.e., whether or not our conjecture concerning validity of the model, e.g., linear regression (universal kriging) model, holds true. In many applications, e.g., *response surface methodology* in which one is mostly interested in polynomial models such a preliminary analysis plays an important role before one conducts further analysis such as predicting the future response at observed or unobserved locations or determining the optimum condition of the model, see e.g., Bisgaard and Ankenman (1996).

In studies concerning model-check or *change-point* problems for linear regression models one usually investigates the partial sums of least squares residuals, i.e., the partial sums of the components of the vector $\mathbf{Z} - \widehat{\mathbf{\Theta}}$, where $\widehat{\mathbf{\Theta}}$ is the *ordinary least squares estimator* of $\mathbf{\Theta}$ which is given by the orthogonal projection of $\mathbf{Z}$ onto $\mathbf{V}$. For example, MacNeill and Jandhyala (1993) and Xie and MacNeill (2004) propose a change-point method for spatial data based on cumulative sums (CUSUM) of least squares residuals. In the one-dimensional case MacNeill (1978a, b) proposed a test based on CUSUM of least squares residuals, and Jandhyala and Minogue (1993) derived Bayes-type change detection statistics based on partial sums of residuals and discussed their asymptotic distributions for general regression. Bischoff (1998) and Bischoff and Miller (2000) proposed an asymptotic test based on CUSUM of least squares residuals for polynomial regression models with one variable. In this work we propose a model-check method for the linear model confining the attention on polynomial regression models with two variables defined on the experimental region $[0, 1] \times [0, 1]$ by conducting tests of hypothesis.

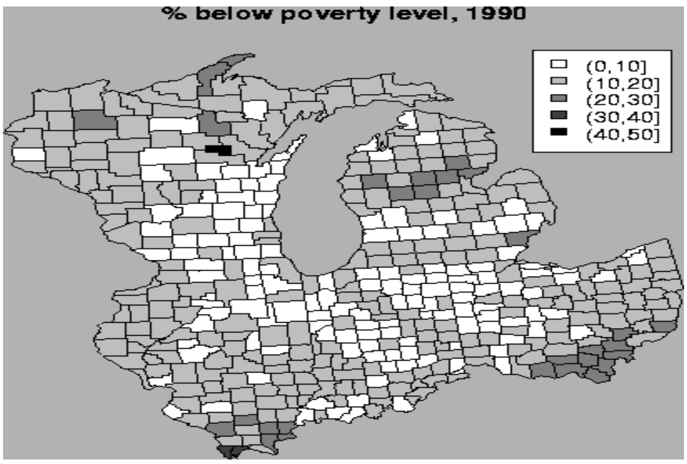

Figure 2. Percentage below poverty line, 5 US midwestern states, 1990 census.

In many areas of application the spatial data are often *lattice* data (*regional*) data, i.e., the experimental domain $\mathbf{D}$ is fixed (not random) and countable, see e.g., Cressie (1991), p. 383-384 or Schabenberger and Gotway (2005), p. 8-9, for the notion of lattice data. For example wheat-yield data measured on an agricultural plot, attributes collected by ZIP code, census tract, remotely sensed data reported by pixels, etc., are usually given by lattice data, see for example the data presented in Figure 2. These data are percentages below poverty level from 5 US midwestern states: Illinois, Indiana, Michigan, Ohio, and Wisconsin, recorded in a 1990 census. Throughout this work we consider our experimental domain as an $n \times n$ regular lattice on $[0,1] \times [0,1]$ in which it is not only fixed and countable, but the points are also equally spaced with the experimental condition $\{(\ell/n, k/n) : 1 \leq \ell, k \leq n\}$. A further important assumption made throughout this work is that the random errors $\{\varepsilon(\ell/n, k/n) : 1 \leq \ell, k \leq n\}$ are independent and identically distributed with unknown distribution having mean 0 and variance $\sigma^2 > 0$. So instead of assuming that the variance of $Z(\ell/n, k/n)$ may depend on ℓ/n and k/n, the so-called *heteroscedastic* linear regression model, we consider the case of a constant variance $Var(Z(\ell/n, k/n)) = \sigma^2 \in (0, \infty)$, the so-called *homoscedastic* linear regression model. Under these assumptions we are interested in the following tasks:

- to establish the limiting distribution of the sequence of the residual partial sums processes associated with the model for large sample size,
- to conduct asymptotic tests (model-checks) based on the residual partial sums process for detecting whether or not the model is a linear regression (universal kriging) model.

We now give an outline about the coverage and organization of our work in handling these problems. In *Chapter 1* we give the formal definition of the linear regression model with experimental condition an $n \times n$ regular lattice on $[0,1] \times [0,1]$.

To give a more convenient interpretation and for theoretical advance, we write the observations in both matrix and vector forms which, by using the *vec* operator, will turn out to be shown equivalent. We also study a *classical* asymptotic estimation procedure for the variance of the observations, see Arnold (1982), p. 147-148.

Chapter 2 introduces the theoretical background which is essential for future considerations. In this chapter we study *Wiener* measure, standard *Brownian (2) motion* and the standard *Brownian (2) bridge* on $(\mathcal{C}([0,1]^2), \mathcal{B}_{\mathcal{C}})$. The characteristics of weak convergence on metric space such as the continuous mapping theorem and its consequences are also discussed. We also present a basic result from classical integration theory (Elstrodt 2005, p. 63-65), that provides necessary and sufficient conditions so that both the Riemann-Stieltjes integral and the Lebesgue-Stieltjes integral coincide. The extended version of this important result to integration with respect to any function which has bounded variation on $[0,1] \times [0,1]$ in the sense of Vitali (Clarkson and Adams, 1933) and is right continuous on $[0,1)^2$ is also investigated. At the end of this chapter we apply this results to an extended version of the weak convergence in the sense of Högnäs (1977) and Johnson (1985) of a sequence of *signed* measures.

In *Chapter 3* we derive the limit process of the least squares residual partial sums process for the linear regression models defined in Chapter 1 by generalizing the approach of Bischoff (1998), Bischoff and Miller (2000), and Bischoff (2002), from the one-dimensional to a higher-dimensional case. We start this chapter with an investigation about the properties of the so-called *reproducing kernel Hilbert space (RKHS)* of the standard Brownian (2) motion. Our main result is stated in Theorem 3.2.6. We close this chapter with a discussion about examples of the residual partial sums limit process associated to several regression models and the generalization of Theorem 3.2.6 to a-regression model with experimental domain being an $n \times m$ regular lattice.

In *Chapter 4* we consider model-checks for spatial data in which we test the hypothesis that the model is universal kriging but we confine our discussion to polynomial models. For that we propose three test statistics based on residual partial sums process: a *Kolmogorov* type statistic, a *Kolmogorov-Smirnov* type statistic and a *Cramér-von Mises* type statistic. For each statistic the asymptotically size α critical region is approximated by simulation. We shall show that the tests are *asymptotically pointwise consistent*. We also investigate the *asymptotic power* of the tests under *localized alternatives* with *localizing rate* $1/n$ and conduct Monte Carlo simulations to approximate the limiting power of the tests for several functions under alternatives so that the behavior of the three tests can be compared. Generalizations of these three procedures to *weighted* tests are also discussed. In order to evaluate the performance of the proposed tests, we apply them by presenting an example of statistical analysis for spatial data in which we work with Mercer and Hall's data, see Section 4.7.

In *Chapter 5* we derive lower and upper bounds for the localized power of the Kolmogorov type test. More exactly, we derive bounds for the boundary crossing probability $\mathbb{P}\left\{\sup_{(t,s)\in[0,1]^2}\left(\rho\varphi(t,s)+B_{\mathbf{f}}(t,s)\right)\geq u(t,s)\right\}$, $\rho>0$, for a known trend $\varphi(\cdot)$ and boundary $u(\cdot)$. We confine our considerations to the standard Brownian (2) motion and the standard Brownian (2) bridge.

In *Chapter 6* we highlight major mathematical open problems which are encountered throughout this work. We also make suggestions for modifications and improvement, and we propose topics for future research.

In *Appendix A* and *Appendix B* we present several basic definitions and notations as well as theorems that are necessary for our work. Several important theorems discussed in Chapter 2 are proved in Appendix A and Appendix B.

Chapter 1

Linear regression models on $[0, 1]^2$

In this chapter we give a formal definition of the linear regression model defined on the unit square $[0, 1]^2$. We also present the point estimation procedures for the parameters of the model.

1.1 Definition of the models

To describe the model in more detail, let $[0, 1]^2 := [0, 1] \times [0, 1] \subset \mathbb{R}^2$ be the experimental domain (region), and let $f_1, \ldots, f_p : [0, 1]^2 \to \mathbb{R}$ be known real-valued regression functions defined on $[0, 1]^2$. We assume that the experiment is performed under $n \times n$ experimental conditions taken from a regular lattice, given by

$$\mathcal{E}_n := \{(\ell/n, k/n) : \; 1 \leq \ell, k \leq n, \; n \in \mathbb{N}\} \subset [0, 1]^2, \tag{1.1.1}$$

see also Figure 3 below for the geometrical visualization of $\mathcal{E}_n$. We give the experimental region $[0, 1]^2$ the topology induced by the *Euclidean metric*

$$\|\mathbf{x} - \mathbf{y}\| := \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}, \; \mathbf{x} := (x_1, x_2), \mathbf{y} := (y_1, y_2) \in [0, 1]^2.$$

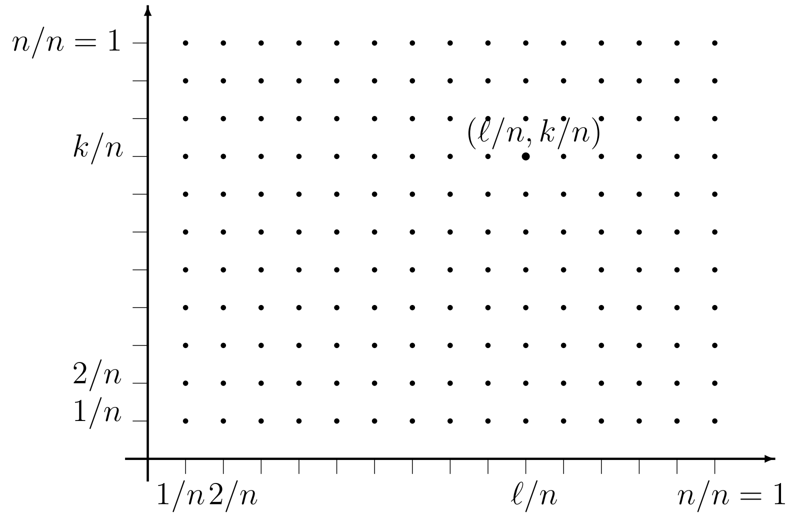

Figure 3. An $n \times n$ regular lattice

For convenience we take the observations over $\mathcal{E}_n$ row wise initializing at the point on the left-bottom corner, i.e., the point with coordinate $(1/n, 1/n)$. Hence we have an $n \times n$ dimensional matrix (array) of corresponding observable response variables

$$\mathbf{Y}_{n\times n} := \begin{pmatrix} Y_{11} & Y_{21} & \cdots & Y_{n1} \\ \vdots & \vdots & \cdots & \vdots \\ Y_{1k} & Y_{2k} & \cdots & Y_{nk} \\ \vdots & \vdots & \cdots & \vdots \\ Y_{1n} & Y_{2n} & \cdots & Y_{nn} \end{pmatrix} \in \mathbb{R}^{n\times n},$$

where the kth row of this matrix represents the observations at the points $\{(\ell/n, k/n) : 1 \leq \ell \leq n\}$ of $\mathcal{E}_n$. Let $\mathbf{M}_i := (f_i(\ell/n, k/n))_{k=1,\ell=1}^{n,\ n}$ be an $n \times n$ dimensional matrix generated by assigning the regression function $f_i(\cdot)$ to the regular lattice $\mathcal{E}_n$, $i = 1, \ldots, p$. We assume that $\mathbf{Y}_{n\times n}$ can be decomposed as

$$\mathbf{Y}_{n\times n} = \mathbf{M}_{n\times n} + \mathbf{E}_{n\times n}, \tag{1.1.2}$$

for some unknown matrix $\mathbf{M}_{n\times n} \in \mathbf{V}_n \subset \mathbb{R}^{n\times n}$, where $\mathbf{V}_n$ is a subspace of $\mathbb{R}^{n\times n}$ generated by the set $\{\mathbf{M}_1, \ldots, \mathbf{M}_p\}$, and $\mathbf{E}_{n\times n} := (\varepsilon_{\ell k})_{\ell=1,k=1}^{n,\ n}$ is an $n \times n$ dimensional *random matrix* with components $\varepsilon_{\ell k}$, $1 \le \ell, k \le n$, are independent and identically distributed real-valued random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with $\mathbb{E}(\varepsilon_{\ell k}) = 0$ and $Var(\varepsilon_{\ell k}) = \sigma^2 \in (0, \infty)$. Furthermore, (1.1.2) is called the *ordinary* linear model, see Arnold, 1981, p. 55. We refer the reader to Muirhead (1982), p. 75-79, for the definition of random matrix. By assumption, there exists some unknown vector of parameters $\beta = (\beta_1, \ldots, \beta_p)^\top \in \mathbb{R}^p$, such that $\mathbf{M}_{n\times n} = \sum_{i=1}^{p} \beta_i \mathbf{M}_i$. Hence, the model can be equivalently represented as

$$\mathbf{Y}_{n\times n} = \sum_{i=1}^{p} \beta_i \mathbf{M}_i + \mathbf{E}_{n\times n}. \tag{1.1.3}$$

In the sequel we interpret all vectors as column vectors unless otherwise stated. For any $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_k, \ldots, \mathbf{a}_n) \in \mathbb{R}^{n\times n}$, where $\mathbf{a}_k \in \mathbb{R}^n$ is the kth column of $\mathbf{A}$, the *vec* operator defined on $\mathbb{R}^{n\times n}$ gives $\mathbf{A}$ the value $vec(\mathbf{A}) := (\mathbf{a}_1^\top, \ldots, \mathbf{a}_k^\top, \ldots, \mathbf{a}_n^\top)^\top \in \mathbb{R}^{n^2}$, see Harville (1997), p. 340-343. Furthermore, let $\mathbf{X}_n := (vec(\mathbf{M}_1), \ldots, vec(\mathbf{M}_p)) \in \mathbb{R}^{n^2\times p}$ be the *design matrix* of the model, i.e., $\mathbf{X}_n$ is an $n^2 \times p$ matrix whose (i, j)th element is given by $f_j(\ell/n, k/n)$, such that $n(k-1) + \ell = i$, for $1 \le k, \ell \le n$, where $1 \le j \le p$, and $1 \le i \le n^2$. By using the *vec* operator, (1.1.3) can be equivalently expressed in the form

$$vec(\mathbf{Y}_{n\times n}) = \mathbf{X}_n\beta + vec(\mathbf{E}_{n\times n}), \tag{1.1.4}$$

with

$$\mathbb{E}(vec(\mathbf{Y}_{n\times n})) = \mathbf{X}_n\beta \text{ and } Cov(vec(\mathbf{Y}_{n\times n})) = \sigma^2\mathbf{I}_{n^2\times n^2},\ \sigma^2 \in (0, \infty).$$

Here $\mathbf{I}_{n^2\times n^2}$ is the $n^2 \times\ n^2$ identity matrix. Model (1.1.3) and (1.1.4) are called the *coordinate* version of the linear model, see Arnold, 1981, p. 55.

It is worth mentioning that for our model we do not assume any specific distribution for the random errors $\varepsilon_{\ell k}$, $1 \le \ell, k \le n$. The only assumption is that these errors are independent and identically distributed with zero means and finite second moments and defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

1.2 Parameter estimation

We furnish the vector space $\mathbb{R}^{n\times n}$ with an inner product $\langle\cdot,\cdot\rangle_{\mathbb{R}^{n\times n}}$ and the corresponding norm $\|\cdot\|_{\mathbb{R}^{n\times n}}$, defined as follows. For any $\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_k, \ldots, \mathbf{a}_n)$, and $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_k, \ldots, \mathbf{b}_n) \in \mathbb{R}^{n\times n}$,

$$\langle \mathbf{A}, \mathbf{B}\rangle_{\mathbb{R}^{n\times n}} := \sum_{k=1}^{p} \langle \mathbf{a}_k, \mathbf{b}_k\rangle_{\mathbb{R}^n} = \langle vec(\mathbf{A}), vec(\mathbf{B})\rangle_{\mathbb{R}^{n^2}} = trace(\mathbf{A}^\top \mathbf{B}),$$

$$\|\mathbf{A}\|^2_{\mathbb{R}^{n\times n}} := \langle \mathbf{A}, \mathbf{A}\rangle_{\mathbb{R}^{n\times n}} = \sum_{k=1}^{p} \|\mathbf{a}_k\|^2_{\mathbb{R}^n} = \|vec(\mathbf{A})\|^2_{\mathbb{R}^{n^2}} = trace(\mathbf{A}^\top \mathbf{A}).$$

Here $\langle\cdot,\cdot\rangle_{\mathbb{R}^n}$ and $\|\cdot\|_{\mathbb{R}^n}$ are the Euclidean inner product and the associated norm on the vector space $\mathbb{R}^n$.

Let $pr_{\mathbf{V}_n}$ and $pr_{\mathbf{V}_n^\perp}$ denote the orthogonal projectors onto the subspace $\mathbf{V}_n$ and onto the orthogonal complement of $\mathbf{V}_n$, respectively. Since the components of the random matrix $\mathbf{E}_{n\times n}$ satisfy the *Gauss-Markov conditions*, then, by the *Gauss-Markov theorem*, see e.g., Arnold (1981), p. 75 or Stapleton (1995), p. 88-94, the *best linear unbiased estimator* (BLUE) for the unknown matrix $\mathbf{M}_{n\times n}$ in (1.1.2) coincides with the *least squares estimator* given by $\widehat{\mathbf{M}}_{n\times n} := pr_{\mathbf{V}_n}\mathbf{Y}_{n\times n}$. The corresponding matrix of the *least squares residuals* is given by

$$\mathbf{R}_{n\times n} := (r_{\ell k})_{k=1,\ell=1}^{n,\ n} := pr_{\mathbf{V}_n^\perp}\mathbf{Y}_{n\times n} = pr_{\mathbf{V}_n^\perp}\mathbf{E}_{n\times n}. \tag{1.2.1}$$

Equation (1.2.1) will be important for our theoretical purposes.

Let us consider Model (1.1.4), and let $pr_{\mathbf{X}_n}$ and $pr_{\mathbf{X}_n^\perp}$ be the orthogonal projectors onto the column space of $\mathbf{X}_n$ and onto the orthogonal complement of the column space of $\mathbf{X}_n$, respectively. Analogous as before, the least squares estimator of $\mathbf{X}_n\beta$ is given by $pr_{\mathbf{X}_n} vec(\mathbf{Y}_{n\times n})$. The corresponding vector of residuals is

$$\mathbf{r}_n := pr_{\mathbf{X}_n^\perp} vec(\mathbf{Y}_{n\times n}) = pr_{\mathbf{X}_n^\perp} vec(\mathbf{E}_{n\times n}) = vec(\mathbf{R}_{n\times n}) \in \mathbb{R}^{n^2}. \tag{1.2.2}$$

For our problems it is no restriction to assume $rank(\mathbf{X}_n) = p$. Hence, the set of matrices $\{\mathbf{M}_1, \ldots, \mathbf{M}_p\}$ is a basis of $\mathbf{V}_n$. Suppose that $\{\mathbf{M}_1, \ldots, \mathbf{M}_p\}$ is an orthogonal basis of $\mathbf{V}_n$, then by the elementary linear algebra, we further get

$$\begin{aligned}
\mathbf{R}_{n\times n} &= \mathbf{Y}_{n\times n} - \sum_{i=1}^{p} \langle \mathbf{M}_i / \|\mathbf{M}_i\|_{\mathbb{R}^{n\times n}}, \mathbf{Y}_{n\times n}\rangle_{\mathbb{R}^{n\times n}} \mathbf{M}_i / \|\mathbf{M}_i\|_{\mathbb{R}^{n\times n}} \\
&= \mathbf{E}_{n\times n} - \sum_{i=1}^{p} \langle \mathbf{M}_i / \|\mathbf{M}_i\|_{\mathbb{R}^{n\times n}}, \mathbf{E}_{n\times n}\rangle_{\mathbb{R}^{n\times n}} \mathbf{M}_i / \|\mathbf{M}_i\|_{\mathbb{R}^{n\times n}}, \\
\mathbf{r}_n &= vec(\mathbf{E}_{n\times n}) - \mathbf{X}_n(\mathbf{X}_n^\top \mathbf{X}_n)^{-1}\mathbf{X}_n^\top vec(\mathbf{Y}_{n\times n}) \\
&= vec(\mathbf{E}_{n\times n}) - \mathbf{X}_n(\mathbf{X}_n^\top \mathbf{X}_n)^{-1}\mathbf{X}_n^\top vec(\mathbf{E}_{n\times n}).
\end{aligned}$$

The least squares estimator of the unknown vector of parameters β is obtained by solving the system of linear equations $(\mathbf{X}_n^\top \mathbf{X}_n)\beta = \mathbf{X}_n^\top vec(\mathbf{Y}_{n\times n})$ for β, which is given by

$$\widehat{\beta}_n := (\widehat{\beta}_{n1}, \ldots, \widehat{\beta}_{np})^\top = (\mathbf{X}_n^\top \mathbf{X}_n)^{-1}\mathbf{X}_n^\top vec(\mathbf{Y}_{n\times n}).$$

By combining (1.1.3) and (1.2.1), we can further write the matrix of least squares residuals $\mathbf{R}_{n\times n}$ in terms of the components of $\widehat{\beta}_n$ as follows:

$$\mathbf{R}_{n\times n} = \mathbf{Y}_{n\times n} - \widehat{\mathbf{M}}_{n\times n} = \mathbf{Y}_{n\times n} - \sum_{i=1}^{p} \widehat{\beta}_{ni}\mathbf{M}_i.$$

For $1 \le i \le n$, let $\mathbf{e}_i$ be a unit vector in $\mathbb{R}^n$ whose ith component is 1, while the others are zero. Then for $1 \le k, \ell \le n$, the (k, ℓ)th component of $\mathbf{R}_{n\times n}$ can be computed by using the equation

$$r_{\ell k} = \mathbf{e}_k^\top \mathbf{Y}_{n\times n}\mathbf{e}_\ell - \sum_{i=1}^{p} \widehat{\beta}_{ni}\mathbf{e}_k^\top \mathbf{M}_i \mathbf{e}_\ell = Y_{\ell k} - \sum_{i=1}^{p} f_i(\ell/n, k/n)\widehat{\beta}_{ni}.$$

If the variance σ^2 is unknown, we use the estimator

$$\hat{\sigma}_n^2 := \frac{\left\|pr_{\mathbf{V}_n^\perp}\mathbf{Y}_{n\times n}\right\|^2_{\mathbb{R}^{n\times n}}}{n^2-p} = \frac{\left\|pr_{\mathbf{X}_n^\perp}vec(\mathbf{Y}_{n\times n})\right\|^2_{\mathbb{R}^{n^2}}}{n^2-p}$$
$$= \frac{vec^\top(\mathbf{Y}_{n\times n})\left(\mathbf{I}_{n^2\times n^2}-\mathbf{X}_n(\mathbf{X}_n^\top\mathbf{X}_n)^{-1}\mathbf{X}_n^\top\right)vec(\mathbf{Y}_{n\times n})}{n^2-p}, \tag{1.2.3}$$

which by (1.2.1) and (1.2.2) is equivalent to

$$\hat{\sigma}_n^2 = \frac{\left\|pr_{\mathbf{V}_n^\perp}\mathbf{E}_{n\times n}\right\|^2_{\mathbb{R}^{n\times n}}}{n^2-p} = \frac{\left\|pr_{\mathbf{X}_n^\perp}vec(\mathbf{E}_{n\times n})\right\|^2_{\mathbb{R}^{n^2}}}{n^2-p}$$
$$= \frac{vec^\top(\mathbf{E}_{n\times n})\left(\mathbf{I}_{n^2\times n^2}-\mathbf{X}_n(\mathbf{X}_n^\top\mathbf{X}_n)^{-1}\mathbf{X}_n^\top\right)vec(\mathbf{E}_{n\times n})}{n^2-p}.$$

It is clear that $\hat{\sigma}_n^2$ is an unbiased estimator for σ^2 in the sense that $\mathbb{E}_{\sigma^2}(\hat{\sigma}_n^2) = \sigma^2$. Furthermore, it can be shown that $\hat{\sigma}_n^2$ is a *consistent* estimator for σ^2, i.e., $\hat{\sigma}_n^2$ *converges in probability* to σ^2 as $n \to \infty$, denoted by $\hat{\sigma}_n^2 \xrightarrow{P} \sigma^2$. Consequently, as $n \to \infty$, we have

$$\frac{\hat{\sigma}_n^2}{\sigma^2} \xrightarrow{P} 1. \tag{1.2.4}$$

We refer the reader to Arnold (1981), p. 147-148, for the preceding results.

Corresponding to the $n\times n$ dimensional matrix of the least squares residuals (1.2.1) we define, for a fixed $n \in \mathbb{N}$, the partial sums process $\{S_n(z_1,z_2) : (z_1,z_2) \in [0,1]^2\}$, where

$$S_n(z_1,z_2) := \sum_{k=1}^{[nz_2]}\sum_{\ell=1}^{[nz_1]} r_{\ell k} + (nz_1 - [nz_1])\sum_{k=1}^{[nz_2]} r_{[nz_1]+1,k} + (nz_2 - [nz_2])\sum_{\ell=1}^{[nz_1]} r_{\ell,[nz_2]+1}$$
$$+ (nz_1 - [nz_1])(nz_2 - [nz_2])r_{[nz_1]+1,[nz_2]+1},$$

which is called the *least squares residual partial sums process* or the *residual partial sums process*, for short. Here $[t] := \max\{n \in \mathbb{Z} : n \le t\}$, $t \in \mathbb{R}$. Our aim is to find the limit process as $n \to \infty$, for the sequence of residual partial sums processes

$$\{S_n(z_1,z_2) : (z_1,z_2) \in [0,1]^2\}_{n\ge 1},$$

and use the result for establishing some asymptotic tests which can be applied to *model-checks* and *change-point* problems.

Chapter 2

Gaussian processes on $\mathcal{C}([0,1]^2)$ and comparison of Riemann-Stieltjes and Lebesgue-Stieltjes integral

To establish asymptotic tests based on the residual partial sums processes for checking the regression model or detecting changes in the regression functions defined on the unit square $[0,1]^2$, we need *standard Brownian (2) motion* with sample paths in $\mathcal{C}([0,1]^2)$, i.e., the space of continuous functions on $[0,1]^2$. The paths of the process additionally fulfill $x(t,0) = x(0,s) = 0$, for $t, s \in [0,1]$. Our aim in this chapter is to define the necessary theoretical background for developing our results described in Chapter 3, Chapter 4 and Chapter 5. We also observe the constructions of the standard Brownian (2) motion as a limit process of a sequence of partial sums processes $\{\mathcal{E}_n(z_1,z_2) : (z_1,z_2) \in [0,1]^2\}_{n\geq 1}$, for $n \to \infty$, where

$$\mathcal{E}_n(z_1,z_2) := \sum_{k=1}^{[nz_2]}\sum_{\ell=1}^{[nz_1]} \varepsilon_{\ell k} + (nz_1 - [nz_1]) \sum_{k=1}^{[nz_2]} \varepsilon_{[nz_1]+1,k} + (nz_2 - [nz_2]) \sum_{\ell=1}^{[nz_1]} \varepsilon_{\ell,[nz_2]+1}$$

$$+(nz_1 - [nz_1])(nz_2 - [nz_2])\varepsilon_{[nz_1]+1,[nz_2]+1},\ (z_1, z_2) \in [0,1]^2,$$

and ε_{ij}, $1 \le i, j \le n$ are independent and identically random variables having distribution with mean zero and finite variances.

As usual, the metric space $\mathcal{C}([0,1]^2)$ is furnished with the uniform topology induced by the *sup-metric*

$$\rho(h, g) := \sup_{(x,y)\in[0,1]^2} \mid h(x,y) - g(x,y) \mid,\ \ h, g \in \mathcal{C}([0,1]^2),$$

with the corresponding *sup-norm*

$$\|g\|_\infty := \sup_{(x,y)\in[0,1]^2} |g(x,y)|\,,\ \ g \in \mathcal{C}([0,1]^2).$$

By using analogous arguments as in the space $\mathcal{C}([0,1])$, the metric space $\mathcal{C}([0,1]^2)$ is *separable* and *complete* with respect to the topology generated by the sup-metric, see e.g., Werner (2005), p. 5 and p. 33, or Billingsley (1999), p. 11-12.

2.1 Standard Brownian (2) motion

Starting with the definition of the *Wiener* measure on the space $\mathcal{C}([0,1]^2)$, we now give the formal definition of the Standard Brownian (2) motion. For a discussion about the existence of the Wiener measure on $\mathcal{C}([0,1]^2)$, see, e.g., Yeh (1960) and Kuelbs (1968) and the references cited there. For fixed $m, n \in \mathbb{N}$, let $\{t_h\}_{0\le h\le m+1}$ and $\{s_k\}_{0\le k\le n+1}$ be preassigned points in the interval $[0,1]$ satisfying $0 = t_0 < t_1 < \cdots < t_m < t_{m+1} = 1$, $0 = s_0 < s_1 < \cdots < s_n < s_{n+1} = 1$. We define the mapping

$$\pi_{(t_1,s_1),\ldots,(t_m,s_n)} : \begin{cases} \mathcal{C}([0,1]^2) & \rightarrow & \mathbb{R}^{mn} \\ \\ w(\cdot) & \mapsto & (w(t_1,s_1),\ldots,w(t_m,s_n))^\top. \end{cases}$$

Let $\mathcal{B}^{mn}$ be the *Borel* $\sigma-$algebra of subsets of $\mathbb{R}^{mn}$, and let $\mathcal{J}_{(t_1,s_1),\ldots,(t_m,s_n)}$ be the collection of subsets of $\mathcal{C}([0,1]^2)$, defined by

$$\mathcal{J}_{(t_1,s_1),\ldots,(t_m,s_n)} := \pi^{-1}_{(t_1,s_1),\ldots,(t_m,s_n)}(\mathcal{B}^{mn}) := \{\pi^{-1}_{(t_1,s_1),\ldots,(t_m,s_n)}(B) : \ B \in \mathcal{B}^{mn}\}.$$

Furthermore, let

$$\mathcal{Z}_{\mathcal{C}([0,1]^2)} := \sigma\left(\bigcup_{m,n\in\mathbb{N}} \bigcup_{\substack{0=t_0<\cdots<t_{m+1}=1 \\ 0=s_0<\cdots<s_{n+1}=1}} \mathcal{J}_{(t_1,s_1),\ldots,(t_m,s_n)}\right)$$

denote the *cylindrical* σ-algebra.

Definition 2.1.1. *Wiener measure* $\mathbb{W}$ *is a probability measure defined on the measurable space* $(\mathcal{C}([0,1]^2), \mathcal{Z}_{\mathcal{C}([0,1]^2)})$ *that satisfies the following conditions.*

(1) $\mathbb{W}\{w \in \mathcal{C}([0,1]^2) : \pi_{(t,s)}(w) = 0, \ \text{if } t = 0 \text{ or } s = 0\} = 1,$

(2) The increments of the stochastic process $\pi := \{\pi_{(t,s)} : \ (t,s) \in [0,1]^2\}$ *are normally distributed with respect to* $\mathbb{W}$*, i.e., for every* $(t_1,s_1),(t_2,s_2) \in [0,1]^2$*, with* $t_1 < t_2$ *and* $s_1 < s_2$*, we have*

$$\Delta_{[t_1,t_2]\times[s_1,s_2]}\pi := \pi_{(t_2,s_2)} - \pi_{(t_2,s_1)} - \pi_{(t_1,s_2)} + \pi_{(t_1,s_1)} \sim \mathcal{N}(0,(t_2-t_1)(s_2-s_1)),$$

(3) The stochastic process $\pi = \{\pi_{(t,s)} : (t,s) \in [0,1]^2\}$ *has independent increments with respect to* $\mathbb{W}$*, i.e., for any family*

$$\{\mathbf{I}_{ij} := [t_{i-1},t_i] \times [s_{j-1},s_j] : \ 1 \le i \le m, 1 \le j \le q\}$$

of rectangles in $[0,1]^2$ *with*

$$0 \le t_0 < \ldots < t_{i-1} < t_i < \ldots < t_m \le 1,$$

$$0 \le s_0 < \ldots < s_{j-1} < s_j < \ldots < s_q \le 1,$$

we have

$$\begin{aligned}&\mathbb{W}\{w\in\mathcal{C}([0,1]^2):(\Delta_{\mathbf{I}_{ij}}\pi)(w)\le\alpha_{ij},\ 1\le i\le m,\ 1\le j\le q\}\\&=\prod_{i=1}^{m}\prod_{j=1}^{q}\mathbb{W}\{w\in\mathcal{C}([0,1]^2):(\Delta_{\mathbf{I}_{ij}}\pi)(w)\le\alpha_{ij}\},\ \alpha_{ij}\in\mathbb{R}.\end{aligned}$$

Remark 2.1.2. *Let $\mathcal{B}_{\mathcal{C}}$ be the Borel σ-algebra over $\mathcal{C}([0,1]^2)$, i.e., the smallest σ-algebra containing all open sets with respect to the uniform topology (topology induced by the sup-metric). Since $\mathcal{C}([0,1]^2)$ is a separable and complete metric space with respect to this topology, by using the analogous argument as in Billingsley (1999), p. 12, and Yeh (1972), p. 449-452, we have $\mathcal{Z}_{\mathcal{C}([0,1]^2)}=\mathcal{B}_{\mathcal{C}}$. Hence, Wiener measure just defined can be regarded as a probability measure on the measurable space $(\mathcal{C}([0,1]^2),\mathcal{B}_{\mathcal{C}})$. The existence of $\mathbb{W}$ on $(\mathcal{C}([0,1]^2),\mathcal{B}_{\mathcal{C}})$ was studied intensively by Kuelbs (1968), Park (1970), and Park (1971).*

Remark 2.1.3. *Under the condition (1) of Definition 2.1.1, it can be shown that conditions (2) and (3) of Definition 2.1.1 are equivalent to the following one: For $d\in\mathbb{N}$ and any $(t_1,s_1),\ldots,(t_d,s_d)\in[0,1]^2$, with respect to $\mathbb{W}$ the vector $(\pi_{(t_1,s_1)},\ldots,\pi_{(t_d,s_d)})^\top$ has a d-variate normal distribution with mean zero and covariance matrix*

$$\mathbf{\Sigma}:=\begin{bmatrix}t_1s_1&(t_1\wedge t_2)(s_1\wedge s_2)&\cdots&(t_1\wedge t_d)(s_1\wedge s_d)\\(t_1\wedge t_2)(s_1\wedge s_2)&t_2s_2&\cdots&(t_2\wedge t_d)(s_2\wedge s_d)\\\vdots&\vdots&\vdots&\vdots\\(t_1\wedge t_d)(s_1\wedge s_d)&(t_2\wedge t_d)(s_2\wedge s_d)&\cdots&t_ds_d\end{bmatrix}. \tag{2.1.1}$$

Here $x\wedge y$ stands for the minimum between x and y.

Definition 2.1.4. *A real-valued stochastic process $X:=\{X(t,s):\ (t,s)\in[0,1]^2\}$ is said to have stationary increments, if and only if for any choice of $h\ge0,\nu\ge0$, and* *any choice of finitely many rectangles*

$$\Gamma := \{\mathbf{I}_{ij} := [t_{i-1}, t_i] \times [s_{j-1}, s_j] : \ 1 \le i \le m, 1 \le j \le q\},$$
$$\Gamma^{h\nu} := \left\{\mathbf{I}_{ij}^{h\nu} := [t_{i-1}+h, t_i+h] \times [s_{j-1}+\nu, s_j+\nu] : \ 1 \le i \le m, 1 \le j \le q\right\}$$

such that

$$0 \le t_0 < \ldots < t_{i-1} < t_i < \ldots < t_m \le 1, \ t_m + h \le 1$$
$$0 \le s_0 < \ldots < s_{j-1} < s_j < \ldots < s_q \le 1, \ s_q + \nu \le 1,$$

the distributions of $(\Delta_{\mathbf{I}_{11}}X, \Delta_{\mathbf{I}_{12}}X, \ldots, \Delta_{\mathbf{I}_{mq}}X)$ *and* $(\Delta_{\mathbf{I}_{11}^{h\nu}}X, \Delta_{\mathbf{I}_{12}^{h\nu}}X, \ldots, \Delta_{\mathbf{I}_{mq}^{h\nu}}X)$ *are identical, denoted by* $(\Delta_{\mathbf{I}_{11}}X, \Delta_{\mathbf{I}_{12}}X, \ldots, \Delta_{\mathbf{I}_{mq}}X) \sim (\Delta_{\mathbf{I}_{11}^{h\nu}}X, \Delta_{\mathbf{I}_{12}^{h\nu}}X, \ldots, \Delta_{\mathbf{I}_{mq}^{h\nu}}X)$.

Definition 2.1.5. *Let* $B_2 = \{B_2(t,s) : (t,s) \in [0,1]^2\}$ *be a real-valued stochastic process defined on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *induced by the* $\mathcal{F}$*-*$\mathcal{B}_{\mathcal{C}}$*-measurable mapping*

$$B_2 : \begin{cases} \Omega & \to \mathcal{C}([0,1]^2) \\ \omega & \mapsto B_2(\omega)(\cdot). \end{cases}$$

We say that the stochastic process $B_2 = \{B_2(t,s) : \ (t,s) \in [0,1]^2\}$ *is a standard Brownian (2) motion (Wiener process) if and only if the distribution of the random function* B_2 *is the Wiener measure on* $(\mathcal{C}([0,1]^2), \mathcal{B}_{\mathcal{C}})$*, i.e.,* $\mathbb{P} \circ B_2^{-1} = \mathbb{W}$*, where* $\mathbb{P} \circ B_2^{-1}(A) := \mathbb{P}\{B_2 \in A\}$*, for each* $A \in \mathcal{B}_{\mathcal{C}}$*.*

Remark 2.1.6. *By conditions (2) and (3) of Definition 2.1.1, for any family of rectangles* $\Gamma, \Gamma^{h\nu}$ *of* $[0,1]^2$*,* $h \ge 0, \nu \ge 0$*, defined in Definition 2.1.4, the random vectors* $(\Delta_{\mathbf{I}_{11}}B_2, \Delta_{\mathbf{I}_{12}}B_2, \ldots, \Delta_{\mathbf{I}_{mq}}B_2)$ *and* $(\Delta_{\mathbf{I}_{11}^{h\nu}}B_2, \Delta_{\mathbf{I}_{12}^{h\nu}}B_2, \ldots, \Delta_{\mathbf{I}_{mq}^{h\nu}}B_2)$ *have the same normal distribution*

$$\mathcal{N}_{mq}(\mathbf{0}, diag(\lambda^2(\mathbf{I}_{11}), \lambda^2(\mathbf{I}_{12}), \ldots, \lambda^2(\mathbf{I}_{mq}))).$$

Here and throughout λ^2 stands for Lebesgue measure on $[0,1]^2$. Hence, by Definition 2.1.4, $\{B_2(t,s) : (t,s) \in [0,1]^2\}$ has stationary and independent increments. Moreover, by Remark 2.1.3, for any rectangles $\mathbf{I}_1$, $\mathbf{I}_2 \subseteq [0,1]^2$, with $\mathbf{I}_1 \cap \mathbf{I}_2 \neq \emptyset$, we have

$$Cov(\Delta_{\mathbf{I}_1} B_2, \Delta_{\mathbf{I}_2} B_2) = \lambda^2(\mathbf{I}_1 \cap \mathbf{I}_2). \tag{2.1.2}$$

Remark 2.1.7. *The Karhunen-Loève expansion of B_2 is given by*

$$B_2(t,s) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \sqrt{\lambda_{ij}}\ Z_{ij} \psi_{ij}(t,s), (t,s) \in [0,1]^2, \tag{2.1.3}$$

where $\{Z_{ij} : i,j \in \mathbb{N}\}$ is a sequence of i.i.d. standard normal random variables, $\lambda_{ij} = [(2i+1)(2j+1)\pi^2/4]^{-2}$, and $\psi_{ij}(t,s) = 2\sin[(2i+1)\pi t/2]\sin[(2j+1)\pi s/2]$, for $i,j \in \mathbb{N}$, see also MacNeill and Jandhyala (1993). Expansion (2.1.3) can be derived directly from the Karhunen-Loève expansion of B_1 (the standard Brownian motion on the unit interval $[0,1]$) due to Yeh (1973), p. 268-279, by the fact that the covariance function of B_2 can be expressed as a multiplication of the covariance functions of B_1.

2.2 Standard Brownian (2) bridge

Corresponding to the standard Brownian (2) motion we now define a stochastic process called a *standard Brownian (2) bridge.*

Definition 2.2.1. *A real-valued stochastic process $B_2^0 = \{B_2^0(t,s) : (t,s) \in [0,1]^2\}$, defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that*

$$B_2^0(t,s) = B_2(t,s) - tsB_2(1,1),\ (t,s) \in [0,1]^2, \tag{2.2.1}$$

is called a standard Brownian (2) bridge.

The following are several important properties of the standard Brownian (2) bridge just defined.

Corollary 2.2.2. (1) $\mathbb{P}\{B_2^0(t,s) = \ 0,\ \textit{if}\ t=0\ \textit{or}\ s=0\} = 1.$

(2) *For any (t,s), and $(t',s') \in [0,1]^2$, we have*

$$Cov(B_2^0(t,s), B_2^0(t',s')) = \ (t \wedge t')(s \wedge s') - tt'ss'. \tag{2.2.2}$$

(3) *For any $(t,s) \in [0,1]^2$, $B_2^0(t,s)$ is normally distributed with mean zero and variance $ts(1-ts)$.*

(4) *For any rectangles $\mathbf{I}_1$, $\mathbf{I}_2 \subseteq [0,1]^2$, with $\mathbf{I}_1 \neq \emptyset \neq \mathbf{I}_2$, we have*

$$Cov(\Delta_{\mathbf{I}_1} B_2^0, \Delta_{\mathbf{I}_2} B_2^0) = \ \lambda^2(\mathbf{I}_1 \cap \mathbf{I}_2) - \lambda^2(\mathbf{I}_1)\lambda^2(\mathbf{I}_2). \tag{2.2.3}$$

Proof. Properties (1), (2), (3) and (4) follow immediately from the properties of the standard Brownian (2) motion and Equation (2.2.1). As an example let us show (2.2.3). By (2.2.1) and by the linearity of Δ, for $i = 1,2$, we get

$$\Delta_{\mathbf{I}_i} B_2^0 = \Delta_{\mathbf{I}_i} B_2 - B_2(1,1)\Delta_{\mathbf{I}_i}(ts) = \Delta_{\mathbf{I}_i} B_2 - B_2(1,1)\lambda^2(\mathbf{I}_i). \tag{2.2.4}$$

Hence, Equation (2.1.2) yields

$$Cov(\Delta_{\mathbf{I}_1} B_2^0, \Delta_{\mathbf{I}_2} B_2^0) = \lambda^2(\mathbf{I}_1 \cap \mathbf{I}_2) - \lambda^2(\mathbf{I}_1)\lambda^2(\mathbf{I}_2).$$

□

As special cases of (2.2.3), if $\mathbf{I}_1 \cap \mathbf{I}_2 = \emptyset$, $\mathbf{I}_i \neq \emptyset$, $i = 1,2$, we have

$$Cov(\Delta_{\mathbf{I}_1} B_2^0, \Delta_{\mathbf{I}_2} B_2^0) = -\lambda^2(\mathbf{I}_1)\lambda^2(\mathbf{I}_2) \neq \ 0. \tag{2.2.5}$$

If $\mathbf{I}_1 = \mathbf{I}_2 =: \mathbf{I}$, $\mathbf{I} \neq \emptyset$, we obtain

$$Var(\Delta_{\mathbf{I}} B_2^0) = \lambda^2(\mathbf{I})[1 - \lambda^2(\mathbf{I})]. \tag{2.2.6}$$

Corollary 2.2.3. *In contrast to B_2, the standard Brownian (2) bridge B_2^0 does not have independent increments. Like B_2, the process B_2^0 has stationary increments.*

Proof. The first assertion follows by (2.2.5). To show the second assertion, let Γ and $\Gamma^{h\nu}$, $h > 0$, $\nu > 0$ be the set of rectangles given in Definition 2.1.4. Then from (2.2.5) and (2.2.6), the random vectors

$$(\Delta_{\mathbf{I}_{11}} B_2^0, \Delta_{\mathbf{I}_{12}} B_2^0, \ldots, \Delta_{\mathbf{I}_{mq}} B_2^0) \text{ and } (\Delta_{\mathbf{I}_{11}^{h\nu}} B_2^0, \Delta_{\mathbf{I}_{12}^{h\nu}} B_2^0, \ldots, \Delta_{\mathbf{I}_{mq}^{h\nu}} B_2^0)$$

have the same centered mq-dimensional normal distribution with covariance matrix

$$\begin{bmatrix} \lambda^2(\mathbf{I}_{11})[1-\lambda^2(\mathbf{I}_{11})] & -\lambda^2(\mathbf{I}_{11})\lambda^2(\mathbf{I}_{12}) & \cdots & -\lambda^2(\mathbf{I}_{11})\lambda^2(\mathbf{I}_{mq}) \\ -\lambda^2(\mathbf{I}_{12})\lambda^2(\mathbf{I}_{11}) & \lambda^2(\mathbf{I}_{12})[1-\lambda^2(\mathbf{I}_{12})] & \cdots & -\lambda^2(\mathbf{I}_{12})\lambda^2(\mathbf{I}_{mq}) \\ \vdots & \vdots & \vdots & \vdots \\ -\lambda^2(\mathbf{I}_{mq})\lambda^2(\mathbf{I}_{11}) & -\lambda^2(\mathbf{I}_{mq})\lambda^2(\mathbf{I}_{12}) & \cdots & \lambda^2(\mathbf{I}_{mq})[1-\lambda^2(\mathbf{I}_{mq})] \end{bmatrix}.$$

Thus, according to Definition 2.1.4, the standard Brownian (2) bridge B_2^0 is a process with stationary increments. □

Corollary 2.2.4. *By combining Equation (2.1.3) and (2.2.1), for every $(t,s) \in [0,1]^2$, the standard Brownian (2) bridge can be expressed as*

$$B_2^0(t,s) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \sqrt{\lambda_{ij}}\; Z_{ij}\left(\psi_{ij}(t,s) - ts\psi_{ij}(1,1)\right), \tag{2.2.7}$$

where $\{Z_{ij} : i,j \in \mathbb{N}\}$ is a sequence of independent and identically distributed standard normal random variables, $\lambda_{ij} = [(2i+1)(2j+1)\pi^2/4]^{-2}$ and $\psi_{ij}(t,s) = 2\sin[(2i+1)\pi t/2]\sin[(2j+1)\pi s/2]$, for $i,j \in \mathbb{N}$.

2.3 Invariance principle and the construction of the standard Brownian (2) motion

In this section we generalize Donsker's theorem for the metric space $\mathcal{C}([0,1])$ (Billingsley, 1968, p. 68-70 and Billingsley, 1999, p. 90-91) to the metric space $\mathcal{C}([0,1]^2)$ and show that standard Brownian (2) motion is a limit process of some sequence of partial sums processes, see Definition 2.3.1 below.

Definition 2.3.1. *For every $n \in \mathbb{N}$, let us define a mapping*

$$\mathbf{T}_n : \begin{cases} \mathbb{R}^{n\times n} & \to \mathcal{C}([0,1]^2) \\ \\ \mathbf{A} & \mapsto \mathbf{T}_n(\mathbf{A})(\cdot) \end{cases}$$

such that for $(z_1, z_2) \in [0,1]^2$, and $\mathbf{A} := (a_{\ell k})_{\ell=1,k=1}^{n,\ n} \in \mathbb{R}^{n\times n}$,

$$\begin{aligned} \mathbf{T}_n(\mathbf{A})(z_1,z_2) = & \frac{1}{n}\sum_{k=1}^{[nz_2]}\sum_{\ell=1}^{[nz_1]} a_{\ell k} + \frac{1}{n}(nz_1 - [nz_1])\sum_{k=1}^{[nz_2]} a_{[nz_1]+1,k} \\ & + \frac{1}{n}(nz_2 - [nz_2])\sum_{\ell=1}^{[nz_1]} a_{\ell,[nz_2]+1} \\ & + \frac{1}{n}(nz_1 - [nz_1])(nz_2 - [nz_2])a_{[nz_1]+1,[nz_2]+1}, \end{aligned} \tag{2.3.1}$$

where $\sum_{k=1}^{j}\sum_{\ell=1}^{i} a_{\ell k} = 0$, for $j = 0$ or $i = 0$. In what follows $\mathbf{T}_n$ will be called the partial sums operator.

Remark 2.3.2. *By definition, for a fixed $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{T}_n(\mathbf{A})(\cdot)$ is a continuous function on $[0,1]^2$, whereas for fixed $(z_1,z_2) \in [0,1]^2$, $\mathbf{T}_n(\cdot)(z_1,z_2)$ constitutes a continuous functional on $\mathbb{R}^{n\times n}$.*

For several values of n, Figure 4 shows the graph of $\mathbf{T}_n(\mathbf{A})(z_1,z_2)$, where $\mathbf{A} \in \mathbb{R}^{n\times n}$ is a realization of an $n \times n$ dimensional *random matrix* with independent standard normally distributed components. It will be shown in Theorem 2.3.4 that Figure 4 illustrates approximated sample paths of standard Brownian (2) motion.

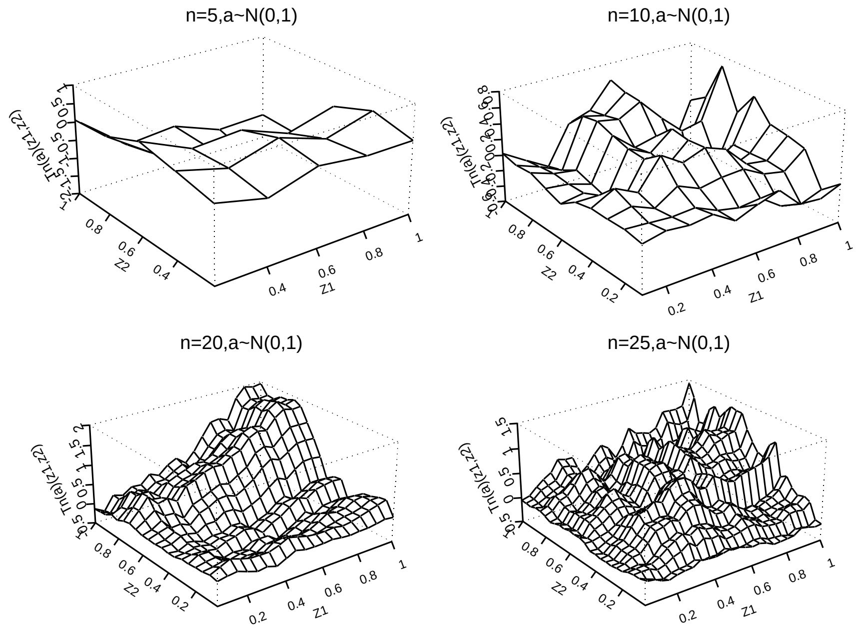

Figure 4. A geometrical visualization of $\mathbf{T}_n(\mathbf{A})(\cdot)$.

Let $\mathbf{E}_{n\times n} := (\varepsilon_{\ell k})_{\ell=1,k=1}^{n,\ n} \in \mathbb{R}^{n\times n}$ be an $n \times n$ dimensional random matrix, with $\mathbb{E}(\varepsilon_{\ell k}) = 0$ and $Var(\varepsilon_{\ell k}) = \sigma^2 \in (0,\infty)$, defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. By using the partial sums operator $\mathbf{T}_n$ defined above, we embed the sequence $(\mathbf{E}_{n\times n})_{n\geq 1}$ into a sequence of stochastic processes

$$\{\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2) : (z_1,z_2) \in [0,1]^2\}_{n\geq 1}, \tag{2.3.2}$$

where

$$\mathbf{T}_n(\mathbf{E}_{n\times n}(\omega))(z_1,z_2) = \frac{1}{n} S_{[nz_1][nz_2]}(\omega) + \frac{1}{n}\psi_{n,z_1,z_2}(\omega),\ \omega \in \Omega,$$

$$S_{[nz_1][nz_2]} := \sum_{k=1}^{[nz_2]} \sum_{\ell=1}^{[nz_1]} \varepsilon_{\ell k} \tag{2.3.3}$$

and

$$\psi_{n,z_1,z_2} := (nz_1 - [nz_1]) \sum_{k=1}^{[nz_2]} \varepsilon_{[nz_1]+1,k} + (nz_2 - [nz_2]) \sum_{\ell=1}^{[nz_1]} \varepsilon_{\ell,[nz_2]+1} + (nz_1 - [nz_1])(nz_2 - [nz_2])\varepsilon_{[nz_1]+1,[nz_2]+1}. \tag{2.3.4}$$

It is clear that the sequence of stochastic processes (2.3.2) has *sample paths* (*trajectories*) in the metric space $\mathcal{C}([0,1]^2)$. We denote the distribution of the sequence $(\mathbf{T}_n(\mathbf{E}_{n\times n}))_{n\geq 1}$ on the probability space $(\mathcal{C}([0,1]^2), \mathcal{B}_{\mathcal{C}})$ by $(\mu_n)_{n\geq 1}$, where $\mu_n := \mathbb{P} \circ \mathbf{T}_n^{-1}(\mathbf{E}_{n\times n})$. Thus, the corresponding sequence of the finite-dimensional distributions of the sequence of stochastic processes $\{\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2) : (z_1,z_2) \in [0,1]^2\}_{n\geq 1}$ are $\left(\mu_n \circ \pi^{-1}_{(t_1,s_1),\ldots,(t_p,s_q)}\right)_{n\geq 1}$, $0 < t_1 < \cdots < t_p \leq 1$, $0 < s_1 < \cdots < s_q \leq 1$, $p, q \in \mathbb{N}$, where for $n \in \mathbb{N}$ and $A \in \mathcal{B}^{pq}$,

$$\begin{aligned}\mu_n \circ \pi^{-1}_{(t_1,s_1),\ldots,(t_p,s_q)}(A) &= \mathbb{P} \circ \mathbf{T}_n^{-1}(\mathbf{E}_{n\times n}) \circ \pi^{-1}_{(t_1,s_1),\ldots,(t_p,s_q)}(A) \\ &= \mathbb{P}\{(\mathbf{T}_n(\mathbf{E}_{n\times n})(t_1,s_1), \ldots, \mathbf{T}_n(\mathbf{E}_{n\times n})(t_p,s_q)) \in A\}.\end{aligned}$$

Since the random variables $\varepsilon_{\ell k}$, $1 \leq \ell, k \leq n$ are independent, the following results can be directly verified by the definition of $\mathbf{T}_n$.

Corollary 2.3.3. (1) *For every* (z_1,z_2) *and* $(z_1',z_2') \in [0,1]^2$, *we have*

$$\mathbb{E}\left(\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2)\right) = 0,$$
$$Cov\left(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2), \frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1',z_2')\right) \to (z_1 \wedge z_1')(z_2 \wedge z_2'), \ \text{as } n \to \infty.$$

(2) *For any* $(z_1,z_2) \in [0,1]^2$, *we have*

$$\begin{aligned}\frac{[nz_1]}{n}\frac{[nz_2]}{n} &\leq Var(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2)) \\ &= \frac{[nz_1][nz_2]}{n^2} + \frac{(nz_1-[nz_1])^2[nz_2]}{n^2} \\ &\quad + \frac{(nz_2-[nz_2])^2[nz_1]}{n^2} + \frac{(nz_1-[nz_1])^2(nz_2-nz_2)^2}{n^2} \\ &\leq \frac{[nz_1]}{n}\frac{[nz_2]}{n} + \frac{2}{n} + \frac{1}{n^2}.\end{aligned} \tag{2.3.5}$$

Hence, $Var(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2)) \to z_1z_2$, *as* $n \to \infty$.

(3) *By the strong law of large numbers, we have* $\frac{1}{n\sigma}\psi_{n,z_1,z_2} \xrightarrow{a.s} 0$, *for every* $(z_1,z_2) \in [0,1]^2$. *Moreover, by the Lindeberg-Lévy central limit theorem and Slutzky's theorem, we have*

$$\frac{1}{n\sigma}S_{[nz_1][nz_2]} \xrightarrow{\mathcal{D}} \mathcal{N}(0, z_1z_2), \ \ as\ n \to \infty.$$

Hence, for $n \to \infty$, *we get*

$$\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2) \xrightarrow{\mathcal{D}} \mathcal{N}(0, z_1z_2), (z_1,z_2) \in [0,1]^2.$$

The following result describes the weak convergence of the sequence of partial sums processes (2.3.2) to the standard Brownian (2) motion. It generalizes Donsker's theorem (*invariance principle*) in the metric space $\mathcal{C}([0,1])$ to the space $\mathcal{C}([0,1]^2)$.

Theorem 2.3.4. ***(Invariance principle)***
Let $(\mathbf{E}_{n\times n})_{n\geq 1}$, $\mathbf{E}_{n\times n} := (\varepsilon_{\ell k})_{\ell=1,k=1}^{n,\ n}$ *be a sequence of independent* $n \times n$ *dimensional random matrices whose components are independent and identically distributed random variables with* $\mathbb{E}(\varepsilon_{\ell k}) = 0$ *and* $Var(\varepsilon_{\ell k}) = \sigma^2 \in (0,\infty)$, *defined on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$, $1 \leq \ell, k \leq n$, $n \geq 1$. *Then* $\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n}) \xrightarrow{\mathcal{D}} B_2$ *as* $n \to \infty$.

Proof. The proof is given in Appendix B where the result in Billingsley (1999), p. 90-91, is extended to the metric space $\mathcal{C}([0,1]^2)$. We refer the reader to Park (1971) for a different proof of this theorem. □

Remark 2.3.5. *By the preceding theorem we can approximate the sample paths of* B_2 *with those of the partial sums process* $\{\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2) : (z_1,z_2) \in [0,1]^2\}_{n\geq 1}$. *Hence, as it was mentioned before, if we put* $\sigma = 1$ *in Theorem 2.3.4, then Figure 4 gives approximations to sample paths of* B_2.

A direct consequence of the *continuous mapping theorem* (Billingsley, 1968, p. 29-31, and Billingsley, 1999, p. 20-22) and Theorem 2.3.4 is described next. We refer the reader to Bader (1997), p. 68-76, for the generalization of the following results to any finite positive measure and any metric space.

Corollary 2.3.6. *Let $(\mathcal{S}, d)$ be a metric space with metric d, and let $\mathcal{B}_{\mathcal{S}}$ be the Borel σ-algebra of $\mathcal{S}$. If the mapping $h : (\mathcal{C}([0,1]^2), \mathcal{B}_{\mathcal{C}}, \mathbb{W}) \to (\mathcal{S}, \mathcal{B}_{\mathcal{S}})$ is continuous on $\mathcal{C}([0,1]^2)$, then $h(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})) \xrightarrow{\mathcal{D}} h(B_2)$ as $n \to \infty$. Thus, the sequence of probability measures $(\mu_n \circ h^{-1})_{n\geq 1}$ converges weakly to the probability measure $\mathbb{W}\circ h^{-1}$ on $(\mathcal{S}, \mathcal{B}_{\mathcal{S}})$, as $n \to \infty$, where $\mathbb{W}$ is Wiener measure on $(\mathcal{C}([0,1]^2), \mathcal{B}_{\mathcal{C}})$. However, the continuity assumption on h can be weakened as follows. Let D_h be the set of discontinuity points of h. If h is $\mathcal{B}_{\mathcal{C}}$-$\mathcal{B}_{\mathcal{S}}$- measurable and $\mathbb{W}(D_h) = 0$, then*

$$h\left(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})\right) \xrightarrow{\mathcal{D}} h(B_2) \text{ as } n \to \infty.$$

Theorem 2.3.7. *Let h, $h_n : (\mathcal{C}([0,1]^2), \mathcal{B}_{\mathcal{C}}, \mathbb{W}) \to (\mathcal{S}, \mathcal{B}_{\mathcal{S}})$, $n \geq 1$, be $\mathcal{B}_{\mathcal{C}}$-$\mathcal{B}_{\mathcal{S}}$-measurable mappings. Let $G := \{x \in \mathcal{C}([0,1]^2) : \exists (x_n)_{n\geq 1},\ x_n \overset{n\to\infty}{\longrightarrow} x,\ \text{but } h_n(x_n) \overset{n\to\infty}{\nrightarrow} h(x)\}$. If $G \subseteq N$ with $\mathbb{W}(N) = 0$, then $h_n(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})) \xrightarrow{\mathcal{D}} h(B_2)$ as $n \to \infty$.*

Remark 2.3.8. *If we put $h_n := h$, for the sequence $(h_n)_{n\geq 1}$ in Theorem 2.3.7, then $G = D_h$, hence this result reduces to Corollary 2.3.6. If $(h_n)_{n\geq 1}$ is a sequence such that $h_n(x_n)$ converges to $h(x)$ whenever $(x_n)_{n\geq 1}$ is a sequence in $\mathcal{C}([0,1]^2)$ which converges to x, then $G = \emptyset$. Hence, we get $h_n(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})) \xrightarrow{\mathcal{D}} h(B_2)$ as $n \to \infty$. We remark here that Theorem 2.3.7 will take an important role for deriving our results in Chapter 3 and Chapter 4.*

2.4 Comparison of Lebesgue-Stieltjes and Riemann-Stieltjes integrals on $[0,1]^2$

In this section we study the comparison between the Lebesgue-Stieltjes and the Riemann-Stieltjes integrals on $[0,1]^2$ defined in Appendix A. We refer the reader to Elstrodt (2005), p. 120-136, for the notion of the Lebesgue-Stieltjes integral considered in this section. A necessary and sufficient conditions for the two integrals to coincide is described in Theorem 2.4.4.

Definition 2.4.1. *A function $F(\cdot)$ defined on $[0,1]^2$ is said to be right continuous on the half-open rectangle $[0,1)^2$, if and only if for every $(c_1,c_2) \in [0,1)^2$ and every sequence $((x_n,y_n))_{n\geq 1} \subseteq [0,1]^2$ such that $(x_n,y_n) \downarrow (c_1,c_2)$, we have $F(x_n,y_n) \to F(c_1,c_2)$ as $n \to \infty$. Analogously, $F(\cdot)$ is said to be left continuous on the half-open rectangle $(0,1]^2$, if and only if for every $(c_1,c_2) \in (0,1]^2$ and sequence $((x_n,y_n))_{n\geq 1} \subseteq [0,1]^2$, with $(x_n,y_n) \uparrow (c_1,c_2)$, we have $F(x_n,y_n) \to F(c_1,c_2)$ as $n \to \infty$. In the sequel the set of functions $F(\cdot)$ defined on $[0,1]^2$ which is right continuous on $[0,1)^2$ and type I non decreasing (see Definition A.1.2) will be denoted by $\mathcal{R}_c([0,1]^2)$.*

Definition 2.4.2. *A Lebesgue-Stieltjes measure μ on $((0,1]^2, \mathcal{B}^2 \cap (0,1]^2)$ is said to be finite, if it satisfies the condition $\mu((0,1] \times (0,1]) < \infty$, see Elstrodt (2005), p. 47.*

Remark 2.4.3. *Let $\mathcal{R}_c^0([0,1]^2)$ be the set of all functions $\psi(\cdot) \in \mathcal{R}_c([0,1]^2)$ with $\psi(x,y) = 0$, if $x = 0$ or $y = 0$. Let $\mathcal{M}((0,1]^2, \mathcal{B}^2 \cap (0,1]^2)$ be the set of finite Lebesgue-Stieltjes measures on $((0,1]^2, \mathcal{B}^2 \cap (0,1]^2)$. Then the mapping $\mathcal{R}_c^0([0,1]^2) \ni \psi(\cdot) \mapsto \mu_\psi(\cdot) \in \mathcal{M}((0,1]^2, \mathcal{B}^2 \cap (0,1]^2)$, where*

$$\mu_\psi((t_1,t_2] \times (s_1,s_2]) := \Delta_{[t_1,t_2]\times[s_1,s_2]}\psi,\ 0 \leq t_1 \leq t_2 \leq 1,\ 0 \leq s_1 \leq s_2 \leq 1, \quad (2.4.1)$$

is one to one. The uniqueness of $\psi(\cdot) \in \mathcal{R}_c^0([0,1]^2)$ *follows from the fact that two functions* $\psi_1(\cdot), \psi_2(\cdot) \in \mathcal{R}_c^0([0,1]^2)$ *that satisfy* $\Delta_{[t_1,t_2]\times[s_1,s_2]}\psi_1 = \Delta_{[t_1,t_2]\times[s_1,s_2]}\psi_2$, *for* $0 \leq t_1 \leq t_2 \leq 1,\ 0 \leq s_1 \leq s_2 \leq 1$, *agree, i.e., we have* $\psi_1(\cdot) = \psi_2(\cdot)$.

The proof of the following result is given in Appendix A. A one-dimensional version of this results was observed by Young (1914), and Stroock (1994), p. 81-82. We refer the reader to Elstrodt (2005), p. 151-153, for the comparison between the Lebesgue-Stieltjes and the Riemann integral on any p-dimensional compact cube $[a_1, b_1] \times \cdots \times [a_p, b_p]$, $a_i < b_i$, $i = 1, \ldots, p$, $p \geq 1$.

Theorem 2.4.4. *Let* ν_ψ *be a finite Lebesgue-Stieltjes measure on* $((0,1]^2, \mathcal{B}^2 \cap (0,1]^2)$ *that corresponds to a function* $\psi(\cdot) \in \mathcal{R}_c([0,1]^2)$ *according to (2.4.1). Suppose that* $\varphi(\cdot) : [0,1]^2 \to \mathbb{R}$ *is bounded on* $[0,1]^2$. *Then* $\varphi(\cdot)$ *is Riemann-Stieltjes integrable on* $[0,1]^2$ *with respect to* $\psi(\cdot)$, *if and only if* $\varphi(\cdot)$ *is continuous* $\bar{\nu}_\psi$ *a.e. on* $(0,1]^2$, *where* $\bar{\nu}_\psi$ *is the completion of the measure* ν_ψ. *Moreover,*

$$\int_{(0,1]^2} \varphi(t,s)\bar{\nu}_\psi(dt,ds) = \int_{[0,1]^2}^{R} \varphi(t,s)d\psi(t,s). \tag{2.4.2}$$

Here and throughout the work, $\int^R$ *denotes the Riemann-Stieltjes integral. We refer the reader to Elstrodt (2005), p.63-65, for the notion of a complete measure.*

Proposition 2.4.5. *Let* $(\psi_n)_{n\geq 1}$ *be a sequence of functions which have bounded variation on* $[0,1]^2$ *in the sense of Vitali, see Definition A.1.5, and let* $(V(\psi_n;[0,1]^2))_{n\geq 1}$ *be the corresponding sequence of the total variations of* $(\psi_n)_{n\geq 1}$. *If* $\|\psi_n\|_\infty \stackrel{n\to\infty}{\longrightarrow} 0$ *and* $(V(\psi_n;[0,1]^2))_{n\geq 1}$ *is bounded, i.e., there exists a positive constant* M *such that for* $n \geq 1$, $V(\psi_n;[0,1]^2) \leq M$, *then* $\int_{[0,1]^2}^R \varphi(t,s)d\psi_n(t,s) \stackrel{n\to\infty}{\longrightarrow} 0$, *for* $\varphi(\cdot) \in \mathcal{C}([0,1]^2)$.

Proof. See Appendix A, see also Högnäs (1977) and Johnson (1985) for a version of this result for a function of one variable. □

2.4.1 Extension to finite Lebesgue-Stieltjes signed measures

Definition 2.4.6. *A signed measure ν on $((0,1]^2, \mathcal{B}^2 \cap (0,1]^2)$ is said to be finite if $|\nu(A)| < \infty$, for every $A \in \mathcal{B}^2 \cap (0,1]^2$ (Elstrodt, 2005, p. 271). In the sequel $\mathcal{SM}((0,1]^2)$ denotes the set of finite Lebesgue-Stieltjes signed measures ν on $((0,1]^2, \mathcal{B}^2 \cap (0,1]^2)$. The set of functions which have bounded variation on $[0,1]^2$ in the sense of Vitali and are right continuous on $[0,1)^2$ will be denoted by $BVV_c([0,1]^2)$. It is clear that $\mathcal{SM}((0,1]^2)$ and $BVV_c([0,1]^2)$ are linear spaces.*

Remark 2.4.7. *If $\psi(\cdot)$ is in $BVV_c([0,1]^2)$, there exists a signed measure $\nu_\psi \in \mathcal{SM}((0,1]^2)$ given by*

$$\nu_\psi((t_1,t_2] \times (s_1,s_2]) := \Delta_{[t_1,t_2]\times[s_1,s_2]}\psi,\ 0 \le t_1 < t_2 \le 1,\ 0 \le s_1 < s_2 \le 1. \tag{2.4.3}$$

Let $V_+(\psi;A)$ and $V_-(\psi;A)$ be the positive and negative total variations of ψ on a set $A \subseteq [0,1]^2$, see Definition A.1.9 for the notion of positive and negative variation. By Theorem A.1.11, we further get

$$\nu_\psi((t_1,t_2] \times (s_1,s_2]) = V_+(\psi;[t_1,t_2]\times[s_1,s_2]) - V_-(\psi;[t_1,t_2]\times[s_1,s_2]). \tag{2.4.4}$$

Moreover, let ν_ψ have Jordan decomposition $\nu_\psi = \nu_\psi^+ - \nu_\psi^-$, and let $\|\nu_\psi\|_\nu := \nu_\psi^+ + \nu_\psi^-$ be the total variation of ν_ψ. Then for every $(t_1,t_2]\times(s_1,s_2] \in \mathcal{B}^2 \cap (0,1]^2$

$$\nu_\psi^+((t_1,t_2] \times (s_1,s_2]) = V_+(\psi;[t_1,t_2]\times[s_1,s_2]), \tag{2.4.5}$$

$$\nu_\psi^-((t_1,t_2] \times (s_1,s_2]) = V_-(\psi;[t_1,t_2]\times[s_1,s_2]), \tag{2.4.6}$$

$$\|\nu_\psi\|_\nu((t_1,t_2] \times (s_1,s_2]) = V_+(\psi;[t_1,t_2]\times[s_1,s_2]) + V_-(\psi;[t_1,t_2]\times[s_1,s_2]). \tag{2.4.7}$$

Corollary 2.4.8. *Let $\psi(\cdot)$ be in $BVV_c([0,1]^2)$ and let $\nu_\psi \in \mathcal{SM}((0,1]^2)$ correspond to $\psi(\cdot)$ according to (2.4.3) with the Jordan decomposition $\nu_\psi = \nu_\psi^+ - \nu_\psi^-$. Suppose* *that $\varphi(\cdot):[0,1]^2\to\mathbb{R}$ is bounded on $[0,1]^2$. Then $\varphi(\cdot)$ is Riemann-Stieltjes integrable on $[0,1]^2$ with respect to $\psi(\cdot)$, if $\varphi(\cdot)$ is continuous $\bar\nu_\psi^+$ a.e. and $\bar\nu_\psi^-$ a.e. on $(0,1]^2$, where $\bar\nu$ is the completion of ν. Moreover, we have*

$$\int_{(0,1]^2}\varphi(t,s)\bar\nu_\psi(dt,ds)=\int_{[0,1]^2}^R\varphi(t,s)d\psi(t,s). \tag{2.4.8}$$

Proof. By definition and Theorem A.1.13, we have $\psi_+(\cdot):=V_+(\psi;[0,\cdot]\times[0,\cdot])$ and $\psi_-(\cdot):=V_-(\psi;[0,\cdot]\times[0,\cdot])$ are type I non decreasing on $[0,1]^2$ and right continuous on $[0,1)^2$. Moreover, for every half-open interval $A:=(t_1,t_2]\times(s_1,s_2]$ in $((0,1]^2\cap\mathcal{B}^2)$, where $0\le t_1<t_2\le 1$ and $0\le s_1<s_2\le 1$, we obtain the following equations:

$$\begin{aligned}\Delta_{\bar A}\psi &\overset{(2.4.3)}{=} \nu_\psi(A)\overset{(2.4.4)}{=}V_+(\psi;\bar A)-V_-(\psi;\bar A)=\Delta_{\bar A}\psi_+-\Delta_{\bar A}\psi_-,\\ \nu_\psi^+(A) &\overset{(2.4.5)}{=} V_+(\psi;\bar A)=\Delta_{\bar A}\psi_+,\ \text{and}\ \nu_\psi^-(A)\overset{(2.4.6)}{=}V_-(\psi;\bar A)=\Delta_{\bar A}\psi_-,\end{aligned}$$

where $\bar A$ is the *closure* of A. Since $\varphi(\cdot)$ is continuous $\bar\nu_\psi^+$ a.e. and $\bar\nu_\psi^-$ a.e. on $(0,1]^2$, by Theorem 2.4.4, $\varphi(\cdot)$ is Riemann-Stieltjes integrable with respect to $\psi_+(\cdot)$ and $\psi_-(\cdot)$. Hence, by the linearity of the Lebesgue-Stieltjes and Riemann-Stieltjes integrals, we further get

$$\begin{aligned}\int_{(0,1]^2}\varphi(t,s)\bar\nu_\psi(dt,ds) &= \int_{(0,1]^2}\varphi(t,s)\bar\nu_\psi^+(dt,ds)-\int_{(0,1]^2}\varphi(t,s)\bar\nu_\psi^-(dt,ds)\\ &=\int_{[0,1]^2}^R\varphi(t,s)d\psi_+(t,s)-\int_{[0,1]^2}^R\varphi(t,s)d\psi_-(t,s)\\ &=\int_{[0,1]^2}^R\varphi(t,s)d\psi(t,s).\end{aligned}$$

The proof is complete. □

Corollary 2.4.9. *Let $(\psi_n)_{n\ge1}$ be a sequence of functions in $BVV_c([0,1]^2)$ such that the sequence of total variations $(V(\psi_n;[0,1]^2))_{n\ge1}$ is bounded. Let $(\nu_{\psi_n})_{n\ge1}$ be the* *sequence in* $\mathcal{SM}((0,1]^2)$ *that corresponds to the sequence* $(\psi_n)_{n\geq 1}$ *in the sense of (2.4.3). If* $\|\psi_n\|_\infty \stackrel{n\to\infty}{\longrightarrow} 0$, *then* $\int_{(0,1]^2} \varphi(t,s)\bar{\nu}_{\psi_n}(dt,ds) \stackrel{n\to\infty}{\longrightarrow} 0$, $\varphi(\cdot) \in \mathcal{C}([0,1]^2)$.

Proof. Since $\varphi(\cdot) \in \mathcal{C}([0,1]^2)$, then it is continuous $\bar{\nu}^+_{\psi_n}$ a.e. and $\bar{\nu}^-_{\psi_n}$ a.e. on $(0,1]^2$, for $n \geq 1$. By Corollary 2.4.8, we have $\int_{(0,1]^2} \varphi(t,s)\bar{\nu}_{\psi_n}(dt,ds) = \int^R_{[0,1]^2} \varphi(t,s)d\psi_n(t,s)$, for $n \geq 1$. Hence, the result follows by Proposition 2.4.5. $\square$

Chapter 3

Residual partial sums limit processes

In this chapter we derive the limit process of the least squares residual partial sums process for the linear regression models introduced in Chapter 1. Mac Neill and Jandhyala (1993) and Xie and Mac Neill (2004) showed a functional central limit theorem for such a residual partial sums process. In this chapter we propose a different and simpler method for deriving such limit processes by generalizing the approach due to Bischoff (1998), Bischoff and Miller (2000), and Bischoff (2002), from the one-dimensional to a higher-dimensional case. As a by-product we obtain the structure of the limit process: it is the projection of the standard Brownian (2) motion onto a certain subspace of the *reproducing kernel Hilbert space (RKHS)* of the standard Brownian (2) motion.

Firstly we introduce in Section 3.1 the notion of the RKHS of the standard Brownian (2) motion which is also connected directly to the partial sums operator $\mathbf{T}_n$. Some important properties of this space are also discussed from analytical aspects. Secondly, by using the notations and the results in Section 3.1, we derive in Section

3.2 (Theorem 3.2.6) the limit process of the sequence of the residual partial sums processes. Several examples of the residual partial sums limit process are also discussed.

3.1 Reproducing kernel Hilbert space of the standard Brownian (2) motion

We furnish the vector space $\mathbb{R}^{n\times n}$, i.e., the space of $n \times n$ dimensional matrices with the inner product $\langle\cdot,\cdot\rangle_{\mathbb{R}^{n\times n}}$ and the norm $\|\cdot\|_{\mathbb{R}^{n\times n}}$ defined in Chapter 1. Let us first consider the image space of the space $\mathbb{R}^{n\times n}$ under the partial sums operator $\mathbf{T}_n$ defined in Chapter 2,

$$\mathbf{T}_n(\mathbb{R}^{n\times n})(\cdot) := \left\{\mathbf{T}_n(\mathbf{A})(\cdot) \mid \mathbf{A} := (a_{\ell k})_{\ell=1,\ k=1}^{n,\ n} \in \mathbb{R}^{n\times n}\right\}.$$

We note that $\mathbf{T}_n(\mathbf{A})(\cdot) \in \mathcal{C}([0,1]^2)$, for every $\mathbf{A} \in \mathbb{R}^{n\times n}$, see Definition 2.3.1. Let us furnish $\mathbf{T}_n(\mathbb{R}^{n\times n})(\cdot)$ with the inner product

$$\langle \mathbf{T}_n(\mathbf{A})(\cdot), \mathbf{T}_n(\mathbf{B})(\cdot)\rangle_{\mathbf{T}_n(\mathbb{R}^{n\times n})} := \frac{1}{n^2}\ \langle \mathbf{A}, \mathbf{B}\rangle_{\mathbb{R}^{n\times n}} = \frac{1}{n^2}\sum_{k=1}^{n}\sum_{\ell=1}^{n} a_{\ell k} b_{\ell k}, \tag{3.1.1}$$

for every $\mathbf{A}$, $\mathbf{B} \in \mathbb{R}^{n\times n}$. Obviously, by Equation (3.1.1), $\mathbf{T}_n(\mathbb{R}^{n\times n})(\cdot)$ and $\mathbb{R}^{n\times n}$ are *isomorphic* Hilbert spaces.

Definition 3.1.1. ***(reproducing kernel Hilbert space (RKHS))***
For the standard Brownian (2) motion $B_2 = \{B_2(t,s) : (t,s) \in [0,1]^2\}$ let us define a linear subspace $\mathcal{H}_{\mathbf{B}}$, given by

$$\begin{aligned}\mathcal{H}_{\mathbf{B}} := \{h : [0,1]^2 \to \mathbb{R} : \exists \hat{h} \in L_2([0,1]^2),\\ h(z_1,z_2) = \int_{[0,z_1]\times[0,z_2]} \hat{h}(\cdot) d\lambda^2,\ (z_1,z_2) \in [0,1]^2\}\end{aligned} \tag{3.1.2}$$

We call the subspace $\mathcal{H}_{\mathbf{B}}$ the reproducing kernel Hilbert space (RKHS) of B_2. Moreover, any $\hat{h} \in L_2([0,1]^2)$ such that $h(z_1,z_2) = \int_{[0,z_1]\times[0,z_2]} \hat{h}\ d\lambda^2$, $(z_1,z_2) \in [0,1]^2$, is called a producing function of $h \in \mathcal{H}_{\mathbf{B}}$.

Remark 3.1.2. *Let $h(\cdot) \in \mathcal{H}_{\mathbf{B}}$ with a reproducing function $\hat{h}(\cdot) \in L_2([0,1]^2)$. Then $h(\cdot)$ is absolutely continuous and has bounded variation in the sense of Hardy on $[0,1]^2$. Furthermore, we have*

$$\frac{\partial^2 h(\cdot)}{\partial t \partial s} = \hat{h}(\cdot), \ \lambda^2 \ a.e. \ on \ [0,1]^2,$$

where $\partial^2 h/\partial t \partial s$ stands for the second derivative of $h(\cdot)$ with respect to t and s. Let ν_h be the finite Lebesgue-Stieltjes signed measure on $((0,1]^2, \mathcal{B}^2 \cap (0,1]^2)$ corresponding to $h(\cdot)$, given by

$$\nu_h((t_1,t_2] \times (s_1,s_2]) := \Delta_{[t_1,t_2]\times[s_1,s_2]} h, \ 0 \le t_1 < t_2 \le 1, \ 0 \le s_1 < s_2 \le 1.$$

Then ν_h is absolutely continuous with respect to Lebesgue measure λ^2, see Cohn (1993), p. 134-137 and Elstrodt (2005), p. 279-281, for the notion of absolute continuity of a signed measure. The Radon-Nikodym derivative of ν_h with respect to λ^2 is given by

$$\frac{d\nu_h}{d\lambda^2} = \hat{h}(\cdot) = \hat{h}^+(\cdot) - \hat{h}^-(\cdot), \tag{3.1.3}$$

where $\hat{h}^+(\cdot)$ and $\hat{h}^-(\cdot)$ are the positive and the negative parts of $\hat{h}(\cdot)$.

Let us furnish $\mathcal{H}_{\mathbf{B}}$ with the inner product

$$\langle h_1, h_2 \rangle_{\mathcal{H}_{\mathbf{B}}} := \int_{[0,1]^2} \hat{h}_1(t,s)\hat{h}_2(t,s)\lambda^2(dt,ds) = \langle \hat{h}_1, \hat{h}_2 \rangle_{L_2}, \tag{3.1.4}$$

for every $h_1(\cdot), \ h_2(\cdot) \in \mathcal{H}_{\mathcal{B}}$, with

$$h_i(z_1,z_2) = \int_{[0,z_1]\times[0,z_2]} \hat{h}_i(t,s)\lambda^2(dt,ds), \ (z_1,z_2) \in [0,1]^2,$$

for some $\hat{h}_i(\cdot) \in L_2([0,1]^2)$, $i = 1,2$. Hence, the norm induced by $\langle \cdot, \cdot \rangle_{\mathcal{H}_{\mathbf{B}}}$ is

$$\|h\|^2_{\mathcal{H}_{\mathbf{B}}} := \int_{[0,1]^2} \left|\hat{h}(t,s)\right|^2 \lambda^2(dt,ds) = \left\|\hat{h}\right\|^2_{L_2}. \tag{3.1.5}$$

Moreover, by (3.1.3), we have

$$\begin{aligned}
\langle h_1, h_2 \rangle_{\mathcal{H}_{\mathbf{B}}} &= \int_{[0,1]^2} \hat{h}_1(t,s)\hat{h}_2(t,s)\lambda^2(dt,ds) \\
&= \int_{(0,1]^2} \hat{h}_2(t,s)\, d\bar{\nu}_{h_1}(dt,ds) \\
&\overset{(2.4.8)}{=} \int_{[0,1]^2}^{R} \hat{h}_2(t,s)\, dh_1(t,s), \qquad (3.1.6)
\end{aligned}$$

where the last equation follows provided $\hat{h}_2(\cdot)$ is continuous $\bar{\nu}_{h_1}$ a.e. on $(0,1]^2$.

Proposition 3.1.3. *For every* $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{T}_n(\mathbf{A})(\cdot) \in \mathcal{H}_{\mathbf{B}}$. *Moreover, for any* $\mathbf{A}$ *and* $\mathbf{B}$ *in* $\mathbb{R}^{n\times n}$, *we have*

$$\langle \mathbf{T}_n(\mathbf{A})(\cdot), \mathbf{T}_n(\mathbf{B})(\cdot) \rangle_{\mathcal{H}_{\mathbf{B}}} = \langle \mathbf{A}, \mathbf{B} \rangle_{\mathbb{R}^{n\times n}}.$$

Proof. Let $\mathbf{A}$, $\mathbf{B}$ and n be fixed. Let us consider the non-overlapping finite exact cover Γ_n of $[0,1]^2$, $\Gamma_n = \{\mathbf{I}_{\ell k} := [(\ell-1)/n, \ell/n] \times [(k-1)/n, k/n] : 1 \le \ell, k \le n\}$, see Definition A.1.1 in the appendix for the notion of a non-overlapping finite exact cover of $[0,1]^2$. By Definition 2.3.1, the second partial derivative of $\mathbf{T}_n(\mathbf{A})(\cdot)$ with respect to both variables exists λ^2 a.s. on $[0,1]^2$, with

$$\frac{\partial^2 \mathbf{T}_n(\mathbf{A})(z_1,z_2)}{\partial z_1 \partial z_2} = n a_{\ell k}, \qquad (3.1.7)$$

for every (z_1, z_2) in the open rectangle $\Delta_{\ell k} := ((\ell-1)/n, \ell/n) \times ((k-1)/n, k/n)$, $1 \le \ell, k \le n$. We therefore have

$$\begin{aligned}
&\int_{[0,1]^2} \left| \frac{\partial^2 \mathbf{T}_n(\mathbf{A})(z_1,z_2)}{\partial z_1 \partial z_2} \right|^2 \lambda^2(dz_1, dz_2) \\
&= \int_{[0,1]^2} \left(\frac{\partial^2 \mathbf{T}_n(\mathbf{A})(t,s)}{\partial t \partial s} \right) \left(\frac{\partial^2 \mathbf{T}_n(\mathbf{A})(t,s)}{\partial t \partial s} \right) \lambda^2(dt,ds), \\
&= \sum_{k=1}^{n} \sum_{\ell=1}^{n} \int_{\mathbf{I}_{\ell k}} (n a_{\ell k})\, (n a_{\ell k})\, \lambda^2(dt,ds), \\
&= \sum_{k=1}^{n} \sum_{\ell=1}^{n} a_{\ell k}^2 = \|\mathbf{A}\|_{\mathbb{R}^{n\times n}}^2 < \infty. \qquad (3.1.8)
\end{aligned}$$

This leads us to the conclusion $\frac{\partial^2 \mathbf{T}_n(\mathbf{A})(z_1,z_2)}{\partial z_1 \partial z_2} \in L_2([0,1]^2)$. Since $\frac{\partial^2 \mathbf{T}_n(\mathbf{A})(\cdot)}{\partial t \partial s}$ is continuous λ^2 a.s. on $[0,1]^2$, by applying Fubini's theorem or the *fundamental theorem of calculus* for Lebesgue integrals, see Elstrodt (2005), p. 304, for each $(z_1, z_2) \in [0,1]^2$, we have

$$\int_{[0,z_1]\times[0,z_2]} \frac{\partial^2 \mathbf{T}_n(\mathbf{A})(t,s)}{\partial t \partial s} \lambda^2(dt,ds) = \mathbf{T}_n(\mathbf{A})(z_1,z_2). \tag{3.1.9}$$

We notice that (3.1.9) can also be derived by computing the integral on the left hand side directly and by recalling (3.1.7), see below.

$$\begin{aligned}
&\int_{[0,z_1]\times[0,z_2]} \frac{\partial^2 \mathbf{T}_n(\mathbf{A})(t,s)}{\partial t \partial s} \lambda^2(dt,ds) \\
&\quad = \int_{[0,[nz_1]/n]\times[0,[nz_2]/n]} \frac{\partial^2 \mathbf{T}_n(\mathbf{A})(t,s)}{\partial t \partial s} \lambda^2(dt,ds) \\
&\qquad + \int_{([nz_1]/n,nz_1/n]\times[0,[nz_2]/n]} \frac{\partial^2 \mathbf{T}_n(\mathbf{A})(t,s)}{\partial t \partial s} \lambda^2(dt,ds) \\
&\qquad + \int_{[0,[nz_1]/n]\times([nz_2]/n,nz_2/n]} \frac{\partial^2 \mathbf{T}_n(\mathbf{A})(t,s)}{\partial t \partial s} \lambda^2(dt,ds) \\
&\qquad + \int_{([nz_1]/n,nz_1/n]\times([nz_2]/n,nz_2/n]} \frac{\partial^2 \mathbf{T}_n(\mathbf{A})(t,s)}{\partial t \partial s} \lambda^2(dt,ds) \\
&\quad = \frac{1}{n}\sum_{k=1}^{[nz_2]}\sum_{\ell=1}^{[nz_1]} a_{\ell k} + \frac{1}{n}(nz_1 - [nz_1])\sum_{k=1}^{[nz_2]} a_{[nz_1]+1,k} + \frac{1}{n}(nz_2 - [nz_2])\sum_{\ell=1}^{[nz_1]} a_{\ell,[nz_2]+1} \\
&\qquad + \frac{1}{n}(nz_1 - [nz_1])(nz_2) - [nz_2]\ a_{[nz_1]+1,[nz_2]+1} \\
&\quad = \mathbf{T}_n(\mathbf{A})(z_1,z_2).
\end{aligned}$$

Thus from (3.1.8) and (3.1.9), $\mathbf{T}_n(\mathbf{A})(\cdot) \in \mathcal{H}_{\mathbf{B}}$ with the producing function given by $\frac{\partial^2 \mathbf{T}_n(\mathbf{A})(\cdot)}{\partial z_1 \partial z_2} \in L_2([0,1]^2)$. The second assertion can be similarly derived as in establishing (3.1.8), from which we get

$$\langle \mathbf{T}_n(\mathbf{A})(\cdot)\ , \mathbf{T}_n(\mathbf{B})(\cdot)\rangle_{\mathcal{H}_{\mathbf{B}}} = \langle \mathbf{A}, \mathbf{B}\rangle_{\mathbb{R}^{n\times n}}. \tag{3.1.10}$$

□

Motivated by the previous results, we investigate the properties of $\mathbf{T}_n(\mathbf{A})(\cdot) \in \mathbf{T}_n(\mathbb{R}^{n\times n})(\cdot) \subseteq \mathcal{H}_{\mathbf{B}}$, for $\mathbf{A} \in \mathbb{R}^{n\times n}$ generated by assigning any real-valued function $f(\cdot)$ to the regular lattice $\mathcal{E}_n$, $n \in \mathbb{N}$.

Definition 3.1.4. *Let $f(\cdot)$ be a real-valued function defined on $[0,1]^2$, and let*

$$f(\mathcal{E}_n) := (f(\ell/n, k/n))_{k=1,\ell=1}^{n,\ n} \in \mathbb{R}^{n\times n}, n \geq 1,$$

where $\mathcal{E}_n$ is the regular lattice given by (1.1.1). Corresponding to the sequence of matrices $\{f(\mathcal{E}_n)\}_{n\geq 1}$ in $\mathbb{R}^{n\times n}$, we define a sequence of functions $\{h_{\mathbf{f}_n}(\cdot)\}_{n\geq 1}$ in $\mathcal{H}_{\mathbf{B}}$ as follows:

$$h_{\mathbf{f}_n}(\cdot) : \begin{cases} [0,1]^2 & \to \mathbb{R} \\ (z_1, z_2) & \mapsto \frac{1}{n}\mathbf{T}_n(f(\mathcal{E}_n))(z_1, z_2). \end{cases} \tag{3.1.11}$$

Remark 3.1.5. *From Equation (3.1.9), we have*

$$h_{\mathbf{f}_n}(z_1, z_2) = \int_{[0,z_1]\times[0,z_2]} \frac{1}{n}\frac{\partial^2 \mathbf{T}_n(f(\mathcal{E}_n))(t,s)}{\partial t \partial s} \lambda^2(dt, ds),$$

for every $(z_1, z_2) \in [0,1]^2$. By definition, and by recalling (3.1.7), the second partial derivative of $h_{\mathbf{f}_n}(\cdot)$ with respect to both variables exists λ^2 a.s. on $[0,1]^2$, where

$$\frac{\partial^2 h_{\mathbf{f}_n}(z_1, z_2)}{\partial z_1 \partial z_2} = f(\ell/n, k/n), \tag{3.1.12}$$

for every $(z_1, z_2) \in \Delta_{\ell k} = ((\ell-1)/n, \ell/n) \times ((k-1)/n, k/n)$, $1 \leq \ell, k \leq n$.

For each $n \geq 1$ let us define a function $g_{\mathbf{f}_n}(\cdot)$ on $[0,1]^2$ by

$$\begin{aligned} g_{\mathbf{f}_n}(\cdot) := & \sum_{k=1}^{n}\sum_{\ell=1}^{n} f(\ell/n, k/n)\mathbf{1}_{[(\ell-1)/n,\ell/n)\times[(k-1)/n,k/n)}(\cdot) \\ & + \sum_{\ell=1}^{n} f(\ell/n, 1)\mathbf{1}_{[(\ell-1)/n,\ell/n)\times\{1\}}(\cdot) \\ & + \sum_{k=1}^{n} f(1, k/n)\mathbf{1}_{\{1\}\times[(k-1)/n,k/n)}(\cdot) \\ & + f(1,1)\mathbf{1}_{\{(1,1)\}}(\cdot) \in L_2([0,1]^2), \end{aligned} \tag{3.1.13}$$

where $\mathbf{1}_A(\cdot)$*,* $A \subseteq [0,1]^2$*, stands for the indicator function of* A*. Hence, for each* $n \geq 1$*,* $g_{\mathbf{f}_n}(\cdot)$ *is defined everywhere on* $[0,1]^2$ *and is clearly right continuous in the sense of Definition 2.4.1 on the half-open rectangle* $[0,1)^2$*. Furthermore by applying either Fubini's theorem or the fundamental theorem of calculus for Lebesgue integrals, we have*

$$h_{\mathbf{f}_n}(z_1,z_2) = \frac{1}{n}\mathbf{T}_n(f(\mathcal{E}_n))(z_1,z_2) = \int_{[0,z_1]\times[0,z_2]} g_{\mathbf{f}_n}(t,s)\lambda^2(dt,ds), \tag{3.1.14}$$

$(z_1,z_2) \in [0,1]^2$*. Note that Equation (3.1.14) can also be shown by directly applying the definition of the integral on the right hand side.*

Remark 3.1.6. *For* $n \geq 1$*,* $g_{\mathbf{f}_n}(\cdot)$ *is in* $BVV([0,1]^2)$*, see Definition A.1.5 for the definition of* $BVV([0,1]^2)$*. Moreover, for* $c = 0$ *or* 1 *and* $n \geq 1$*, the "marginal" functions* $g_{\mathbf{f}_n}(c,\cdot)$ *and* $g_{\mathbf{f}_n}(\cdot,c)$ *are in* $BV([0,1])$*. To see this, we consider the finite exact cover* $\Gamma_n = \{[(\ell-1)/n,\ell/n] \times [(k-1)/n,k/n] : 1 \leq \ell,k \leq n\}$*. Note that* $\Gamma_n = \Gamma_n^{(1)} \times \Gamma_n^{(2)}$*, where* $\Gamma_n^{(1)} := \{[(\ell-1)/n,\ell/n] : 1 \leq \ell \leq n\}$*,* $\Gamma_n^{(2)} := \{[(k-1)/n,k/n] : 1 \leq k \leq n\}$*. By (3.1.13) and by Definition A.1.4, it can be shown easily, that*

$$\begin{aligned}
&V(g_{\mathbf{f}_n};[0,1]^2) \leq v(f;\Gamma_n),\\
&V(g_{\mathbf{f}_n}(1,\cdot);[0,1]) \leq v(f(1,\cdot);\Gamma_n^{(2)}),\\
&V(g_{\mathbf{f}_n}(\cdot,1);[0,1]) \leq v(f(\cdot,1);\Gamma_n^{(1)}),\\
&V(g_{\mathbf{f}_n}(0,\cdot);[0,1]) \leq v(f(1/n,\cdot);\Gamma_n^{(2)}),\\
&V(g_{\mathbf{f}_n}(\cdot,0);[0,1]) \leq v(f(\cdot,1/n);\Gamma_n^{(1)}).
\end{aligned}$$

Here and throughout $V(\psi;A)$ *stands for the total variation of* $\psi(\cdot)$ *on* $A \subset [0,1]^2$*. Consequently, by Definition A.1.6, for* $n \geq 1$*,* $g_{\mathbf{f}_n}(\cdot)$ *is in* $BV([0,1]^2)$*.*

Lemma 3.1.7. *Let $f(\cdot)$ be any function in $L_2([0,1]^2)$. Let us define a function $h_f(\cdot) \in \mathcal{H}_{\mathbf{B}}$ as follows:*

$$h_f(\cdot) : \begin{cases} [0,1]^2 & \to \mathbb{R} \\ (z_1, z_2) & \mapsto \int_{[0,z_1]\times[0,z_2]} f(t,s)\lambda^2(dt,ds). \end{cases} \tag{3.1.15}$$

If the function $f(\cdot)$ is continuous on $[0,1]^2$, then $g_{\mathbf{f}_n}(\cdot)$ converges uniformly on $[0,1]^2$ to $\frac{\partial^2 h_f(\cdot)}{\partial z_1 \partial z_2} = f(\cdot)$, as $n \to \infty$.

Proof. By the hypothesis, given $\varepsilon > 0$, there exists a $\delta > 0$, such that

$$\sup_{\|(t,s)-(t',s')\| \le \delta} |f(t',s') - f(t,s)| \le \varepsilon.$$

We consider the non-overlapping finite exact cover Γ_n of $[0,1]^2$ given in Remark 3.1.6, with $\|\Gamma_n\| = \sqrt{2}/n$. Take $n_0 := [\sqrt{2}/\delta]+1$. Then $\|\Gamma_n\| \le \delta$ for all $n \ge n_0$. Let $(t,s) \in [0,1]^2$ be arbitrarily fixed, then there exist $\ell, k \in \mathbb{N}$, $(1 \le \ell, k \le n)$, $n_0 \le n$, such that either $(t,s) \in [(\ell-1)/n, \ell/n) \times [(k-1)/n, k/n)$, or $(t,s) \in [(\ell-1)/n, \ell/n) \times \{1\}$, or $(t,s) \in \{1\} \times [(k-1)/n, k/n)$, or $(t,s) \in \{(1,1)\}$. By (3.1.13), in each case we have $|g_{\mathbf{f}_n}(t,s) - f(t,s)| \le \varepsilon$, $n \ge n_0$. Since $(t,s) \in [0,1]^2$ is arbitrary, we get $\|g_{\mathbf{f}_n}(\cdot) - f(\cdot)\|_\infty \le \varepsilon$, for all $n \ge n_0$. This completes the proof. □

Corollary 3.1.8. *If the function $f(\cdot)$ is continuous on $[0,1]^2$, then*

$$\|h_{\mathbf{f}_n} - h_f\|_{\mathcal{H}_{\mathbf{B}}} \overset{n\to\infty}{\longrightarrow} 0 \text{ and } \|h_{\mathbf{f}_n} - h_f\|_\infty \overset{n\to\infty}{\longrightarrow} 0.$$

Proof. By definition, the producing functions of $h_{\mathbf{f}_n}(\cdot)$ and $h_f(\cdot)$ are given by $g_{\mathbf{f}_n}(\cdot)$ and $f(\cdot)$, respectively. Thus by (3.1.14), (3.1.15), and Lemma 3.1.7, we get

$$\begin{aligned} 0 \le \|h_{\mathbf{f}_n} - h_f\|^2_{\mathcal{H}_{\mathbf{B}}} &= \int_{[0,1]^2} |g_{\mathbf{f}_n}(t,s) - f(t,s)|^2 \lambda^2(dt,ds) \\ &\le \int_{[0,1]^2} (\sup_{(t,s)\in[0,1]^2} |g_{\mathbf{f}_n}(t,s) - f(t,s)|)^2 \lambda^2(dt,ds), \\ &= \|g_{\mathbf{f}_n} - f\|^2_\infty \overset{n\to\infty}{\longrightarrow} 0. \end{aligned}$$

Similarly, for any $(z_1, z_2) \in [0,1]^2$, we have

$$\begin{aligned} |h_{\mathbf{f}_n}(z_1,z_2) - h_f(z_1,z_2)| &= \left| \int_{[0,z_1]\times[0,z_2]} (g_{\mathbf{f}_n}(t,s) - f(t,s))\lambda^2(dt,ds) \right| \\ &\le \int_{[0,z_1]\times[0,z_2]} \sup_{(t,s)\in[0,1]^2} |g_{\mathbf{f}_n}(t,s) - f(t,s)|\, \lambda^2(dt,ds), \\ &\le \|g_{\mathbf{f}_n} - f\|_\infty \overset{n\to\infty}{\longrightarrow} 0. \end{aligned}$$

□

We now extent the preceding results to the regression functions $f_1(\cdot),\ldots,f_p(\cdot)$ starting with a definition of subspaces required in establishing the limit processes of the residual partial sums processes associated to $f_1(\cdot),\ldots,f_p(\cdot)$ defined on $[0,1]^2$.

Definition 3.1.9. *Suppose that $f_1(\cdot),\ldots,f_p(\cdot)$ are linearly independent as functions in $L_2([0,1]^2)$. Let us define the following linear subspaces:*

(1) $\mathbf{W} := [f_1(\cdot),\ldots,f_p(\cdot)]$, *i.e., the subspace spanned by the regression functions* $\{f_1(\cdot),\ldots,f_p(\cdot)\}$.

(2) $\mathbf{W}_{\mathcal{H}_\mathbf{B}} := [h_{f_1}(\cdot),\ldots,h_{f_p}(\cdot)]$, *i.e., the subspace of $\mathcal{H}_\mathbf{B}$ spanned by the set of functions $\{h_{f_1}(\cdot),\ldots,h_{f_p}(\cdot)\}$, where $h_{f_i}(\cdot)$ is given by (3.1.15).*

(3) $\mathbf{W}_n := [f_1(\mathcal{E}_n),\ldots,f_p(\mathcal{E}_n)]$ *is a subspace in $\mathbb{R}^{n\times n}$, where $f_i(\mathcal{E}_n)$ is given in Definition 3.1.4.*

(4) $\mathbf{W}_{n\mathcal{H}_\mathbf{B}} := [h_{\mathbf{f}_{1,n}}(\cdot),\ldots,h_{\mathbf{f}_{p,n}}(\cdot)]$, *i.e., the subspace of $\mathcal{H}_\mathbf{B}$ spanned by the set of functions $\{h_{\mathbf{f}_{1,n}}(\cdot),\ldots,h_{\mathbf{f}_{p,n}}(\cdot)\}$, where $h_{\mathbf{f}_{i,n}}(\cdot)$ is defined in (3.1.11).*

(5) $\mathbf{L}_{2\mathbf{W}_{n\mathcal{H}_\mathbf{B}}} := [g_{\mathbf{f}_{1,n}}(\cdot),\ldots,g_{\mathbf{f}_{p,n}}(\cdot)]$, *i.e., the subspace of $L_2([0,1]^2)$ spanned by the set of functions $\{g_{\mathbf{f}_{1,n}}(\cdot),\ldots,g_{\mathbf{f}_{p,n}}(\cdot)\}$, where $g_{\mathbf{f}_{i,n}}(\cdot)$ is defined in (3.1.13).*

Remark 3.1.10. *We furnish the subspaces* $\mathbf{W}$ *and* $\mathbf{L}_{2\mathbf{W}_{n\mathcal{H}_\mathbf{B}}}$ *with the inner product* $\langle\cdot,\cdot\rangle_{L_2}$*, the subspaces* $\mathbf{W}_{\mathcal{H}_\mathbf{B}}$ *and* $\mathbf{W}_{n\mathcal{H}_\mathbf{B}}$ *with the inner product* $\langle\cdot,\cdot\rangle_{\mathcal{H}_\mathbf{B}}$*, while the subspace* $\mathbf{W}_n$ *is furnished with the inner product* $\langle\cdot,\cdot\rangle_{\mathbb{R}^{n\times n}}$*. By definition it can be shown that* $\{h_{f_1}(\cdot),\ldots,h_{f_p}(\cdot)\}$ *is a basis of* $\mathbf{W}_{\mathcal{H}_\mathbf{B}} \subseteq \mathcal{H}_\mathbf{B}$*. The matrices* $f_1(\mathcal{E}_n),\ldots,f_p(\mathcal{E}_n)$ *form a basis of* $\mathbf{W}_n \subseteq \mathbb{R}^{n\times n}$*, if and only if, the functions* $h_{\mathbf{f}_{1,n}}(\cdot),\ldots,h_{\mathbf{f}_{p,n}}(\cdot)$ *are a basis of* $\mathbf{W}_{n\mathcal{H}_\mathbf{B}} \subseteq \mathcal{H}_\mathbf{B}$*. The latter holds if and only if,* $\{g_{\mathbf{f}_{1,n}}(\cdot),\ldots,g_{\mathbf{f}_{p,n}}(\cdot)\}$ *is a basis of* $\mathbf{L}_{2\mathbf{W}_{n\mathcal{H}_\mathbf{B}}}$*. Let us define the matrix* $\mathbf{J} := (\langle f_i, f_j\rangle_{L_2})_{i=1,j=1}^{p,\ p}$*. Since the design matrix* $\mathbf{X}_n := (vec(f_1(\mathcal{E}_n)),\ldots,vec(f_p(\mathcal{E}_n)))$ *has the property* $(\frac{1}{n^2}\mathbf{X}_n^\top\mathbf{X}_n) \stackrel{n\to\infty}{\longrightarrow} \mathbf{J}$*, component-wise, then* $det(\frac{1}{n^2}\mathbf{X}_n^\top\mathbf{X}_n) \stackrel{n\to\infty}{\longrightarrow} det(\mathbf{J})$*, where* $det(\cdot)$ *is the determinant function. The matrix* $\mathbf{J}$ *is clearly invertible (*$det(\mathbf{J}) \neq 0$*), hence there exists an* $n_0 \in \mathbb{N}$ *such that* $det(\frac{1}{n^2}\mathbf{X}_n^\top\mathbf{X}_n) \neq 0$*, for* $n \geq n_0$*. Hence* $\mathbf{X}_n$ *has full-column rank* p *for every* $n \geq n_0$*, which directly implies that the vectors* $vec(f_1(\mathcal{E}_n)),\ldots,vec(f_p(\mathcal{E}_n))$ *are linearly independent in* $\mathbb{R}^{n^2}$*, for* $n \geq n_0$*. By the linearity of vec operator, we further get that* $f_1(\mathcal{E}_n),\ldots,f_p(\mathcal{E}_n)$ *are linearly independent in* $\mathbb{R}^{n\times n}$*, for every* $n \geq n_0$*. Thus, as stated above, this implies that, for every* $n \geq n_0$*,* $\{h_{\mathbf{f}_{1,n}}(\cdot),\ldots,h_{\mathbf{f}_{p,n}}(\cdot)\}$ *and* $\{g_{\mathbf{f}_{1,n}}(\cdot),\ldots,g_{\mathbf{f}_{p,n}}(\cdot)\}$ *are bases of* $\mathbf{W}_{n\mathcal{H}_\mathbf{B}}$ *and* $\mathbf{L}_{2\mathbf{W}_{n\mathcal{H}_\mathbf{B}}}$*, respectively. In the sequel we always assume that* $n \geq n_0$*.*

Remark 3.1.11. *Let* $\{\tilde{f}_1(\cdot),\ldots,\tilde{f}_p(\cdot)\}$ *be the Gram-Schmidt orthonormal basis of* $\mathbf{W}$ *associated with the basis* $\{f_1(\cdot),\ldots,f_p(\cdot)\}$*, and let* $\{\tilde{h}_{f_1}(\cdot),\ldots,\tilde{h}_{f_p}(\cdot)\}$ *be the Gram-Schmidt orthonormal basis of* $\mathbf{W}_{\mathcal{H}_\mathbf{B}}$ *associated with the basis* $\{h_{f_1}(\cdot),\ldots,h_{f_p}(\cdot)\}$*. Then by (3.1.15) it can be shown that, for every* $(z_1,z_2)\in[0,1]^2$

$$\tilde{h}_{f_i}(z_1,z_2) = \int_{[0,z_1]\times[0,z_2]} \tilde{f}_i(t,s)\lambda^2(dt,ds),\ i=1,\ldots,p. \tag{3.1.16}$$

Similarly, let $\{\tilde{h}_{\mathbf{f}_{1,n}}(\cdot),\ldots,\tilde{h}_{\mathbf{f}_{p,n}}(\cdot)\}$ *be the Gram-Schmidt orthonormal basis of* $\mathbf{W}_{n\mathcal{H}_\mathbf{B}}$ *corresponding to the basis* $\{h_{\mathbf{f}_{1,n}}(\cdot),\ldots,h_{\mathbf{f}_{p,n}}(\cdot)\}$*, and let* $\{\tilde{g}_{\mathbf{f}_{1,n}}(\cdot),\ldots,\tilde{g}_{\mathbf{f}_{p,n}}(\cdot)\}$ *be the*

Gram-Schmidt orthonormal basis of $\mathbf{L}_{2\mathbf{W}_{n\mathcal{H}_{\mathbf{B}}}}$ *that corresponds to* $\{g_{\mathbf{f}_{1,n}}(\cdot), \ldots, g_{\mathbf{f}_{p,n}}(\cdot)\}$. *Then by (3.1.14), for every* $(z_1, z_2) \in [0,1]^2$, *we have*

$$\tilde{h}_{\mathbf{f}_{i,n}}(z_1, z_2) = \int_{[0,z_1]\times[0,z_2]} \tilde{g}_{\mathbf{f}_{i,n}}(t,s)\lambda^2(dt, ds). \tag{3.1.17}$$

Corollary 3.1.12. *By definition, the functions* $\tilde{g}_{\mathbf{f}_{i,n}}(\cdot)$, $n \geq 1$ *are right continuous on* $[0,1)^2$. *Moreover, if* $f_i(\cdot)$ *is continuous on* $[0,1]^2$, *then*

$$\left\|\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i\right\|_\infty \overset{n\to\infty}{\longrightarrow} 0,\ i = 1, \ldots, p.$$

The same properties are also satisfied by the corresponding sequences of marginal functions $\left(\tilde{g}_{\mathbf{f}_{i,n}}(c,\cdot)\right)_{n\geq 1}$ *and* $\left(\tilde{g}_{\mathbf{f}_{i,n}}(\cdot,c)\right)_{n\geq 1}$, $c = 0, 1$, *i.e., they are right continuous on the half-open interval* $[0,1)$, *and we have*

$$\left\|\tilde{g}_{\mathbf{f}_{i,n}}(c,\cdot) - \tilde{f}_i(c,\cdot)\right\|_\infty \overset{n\to\infty}{\longrightarrow} 0 \text{ and } \left\|\tilde{g}_{\mathbf{f}_{i,n}}(\cdot,c) - \tilde{f}_i(\cdot,c)\right\|_\infty \overset{n\to\infty}{\longrightarrow} 0.$$

Proof. We first consider the case $i = 1$. Since $g_{\mathbf{f}_{1,n}}(\cdot) \overset{\|\cdot\|_\infty}{\longrightarrow} f_1(\cdot)$ as $n \to \infty$, then $g_{\mathbf{f}_{1,n}}(\cdot) \overset{\|\cdot\|_{L_2}}{\longrightarrow} f_1(\cdot)$, which, by the continuity of $\|\cdot\|_{L_2}$, implies $\left\|g_{\mathbf{f}_{1,n}}\right\|_{L_2} \overset{n\to\infty}{\longrightarrow} \|f_1\|_{L_2}$. Hence, for $n \to \infty$, we have

$$\tilde{g}_{\mathbf{f}_{1,n}}(\cdot) := \frac{g_{\mathbf{f}_{1,n}}(\cdot)}{\left\|g_{\mathbf{f}_{1,n}}(\cdot)\right\|_{L_2}} \overset{\|\cdot\|_\infty}{\longrightarrow} \frac{f_1(\cdot)}{\|f_1(\cdot)\|_{L_2}} =: \tilde{f}_1(\cdot).$$

Now let us consider the case $i = 2$. Since $\tilde{g}_{\mathbf{f}_{1,n}}(\cdot) \overset{\|\cdot\|_\infty}{\longrightarrow} \tilde{f}_1(\cdot)$ and $g_{\mathbf{f}_{2,n}}(\cdot) \overset{\|\cdot\|_\infty}{\longrightarrow} f_2(\cdot)$, as $n \to \infty$, then $\langle \tilde{g}_{\mathbf{f}_{1,n}}, g_{\mathbf{f}_{2,n}} \rangle_{L_2} \overset{n\to\infty}{\longrightarrow} \langle \tilde{f}_1, f_2 \rangle_{L_2}$. These yield $g_{\mathbf{f}_{2,n}}(\cdot) - \langle \tilde{g}_{\mathbf{f}_{1,n}}, g_{\mathbf{f}_{2,n}} \rangle_{L_2} \tilde{g}_{\mathbf{f}_{1,n}}(\cdot) \overset{\|\cdot\|_\infty}{\longrightarrow} f_2(\cdot) - \langle \tilde{f}_1, f_2 \rangle_{L_2} \tilde{f}_1(\cdot)$ as $n \to \infty$. Hence, by applying the same argument as before, as $n \to \infty$, $g_{\mathbf{f}_{2,n}}(\cdot) - \langle \tilde{g}_{\mathbf{f}_{1,n}}, g_{\mathbf{f}_{2,n}} \rangle_{L_2} \tilde{g}_{\mathbf{f}_{1,n}}(\cdot) \overset{\|\cdot\|_{L_2}}{\longrightarrow} f_2(\cdot) - \langle \tilde{f}_1, f_2 \rangle_{L_2} \tilde{f}_1(\cdot)$. Furthermore, since $\|\cdot\|_{L_2}$ is continuous, we further obtain

$$\left\|g_{\mathbf{f}_{2,n}}(\cdot) - \langle \tilde{g}_{\mathbf{f}_{1,n}}, g_{\mathbf{f}_{2,n}} \rangle_{L_2} \tilde{g}_{\mathbf{f}_{1,n}}(\cdot)\right\|_{L_2} \overset{n\to\infty}{\longrightarrow} \left\|f_2(\cdot) - \langle \tilde{f}_1, f_2 \rangle_{L_2} \tilde{f}_1(\cdot)\right\|_{L_2}.$$

Thus, by combining these results, we finally get

$$\tilde{g}_{\mathbf{f}_{2,n}}(\cdot) := \frac{g_{\mathbf{f}_{2,n}}(\cdot) - \langle \tilde{g}_{\mathbf{f}_{1,n}}, g_{\mathbf{f}_{2,n}} \rangle_{L_2} \tilde{g}_{\mathbf{f}_{1,n}}(\cdot)}{\left\| g_{\mathbf{f}_{2,n}}(\cdot) - \langle \tilde{g}_{\mathbf{f}_{1,n}}, g_{\mathbf{f}_{2,n}} \rangle_{L_2} \tilde{g}_{\mathbf{f}_{1,n}}(\cdot) \right\|_{L_2}} \overset{\|\cdot\|_\infty}{\longrightarrow} \frac{f_2(\cdot) - \langle \tilde{f}_1, f_2 \rangle_{L_2} \tilde{f}_1(\cdot)}{\left\| f_2(\cdot) - \langle \tilde{f}_1, f_2 \rangle_{L_2} \tilde{f}_1(\cdot) \right\|_{L_2}} =: \tilde{f}_2(\cdot).$$

The assertion for $i = 3, \ldots, p$ can be handled analogously. This proves the first assertion. The second assertion is a direct consequence of the first assertion. □

Corollary 3.1.13. *For $i = 1, \ldots, p$, if the regression function $f_i(\cdot)$ is continuous on $[0,1]^2$, then $\left\| \tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i} \right\|_{\mathcal{H}_{\mathbf{B}}} \overset{n\to\infty}{\longrightarrow} 0$ and $\left\| \tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i} \right\|_{\infty} \overset{n\to\infty}{\longrightarrow} 0$.*

Proof. This result follows directly from Corollary 3.1.12. □

Proposition 3.1.14. *Let $f_i(\cdot)$ be in $BV([0,1]^2)$, and let $c = 0$ or 1. Then for $i = 1, \ldots, p$, there exist positive real numbers M_i, M_{ic} and K_{ic}, such that for $n \geq 1$, $V(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i; [0,1]^2) \leq M_i$, $V((\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(c, \cdot); [0,1]) \leq M_{ic}$ and $V((\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(c, \cdot); [0,1]) \leq K_{ic}$. That is, the sequences $(V(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i; [0,1]^2))_{n\geq 1}$, $(V((\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(c, \cdot); [0,1]))_{n\geq 1}$ and $(V((\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(c, \cdot); [0,1]))_{n\geq 1}$ are bounded uniformly.*

Proof. Let us consider first the case $i = 1$. Since $\tilde{g}_{\mathbf{f}_{1,n}} = \frac{g_{\mathbf{f}_{1,n}}}{\left\| g_{\mathbf{f}_{1,n}} \right\|_{L_2}}$, by Remark 3.1.6 and by the *triangle* inequality, for a fixed $n \geq 1$, it follows that

$$\begin{aligned}
V(\tilde{g}_{\mathbf{f}_{1,n}} - \tilde{f}_1; [0,1]^2) &\leq \frac{V(f_1; [0,1]^2)}{\left\| g_{\mathbf{f}_{1,n}} \right\|_{L_2}} + V(\tilde{f}_1; [0,1]^2), \\
V((\tilde{g}_{\mathbf{f}_{1,n}} - \tilde{f}_1)(1, \cdot); [0,1]) &\leq \frac{V(f_1(1, \cdot); [0,1])}{\left\| g_{\mathbf{f}_{1,n}} \right\|_{L_2}} + V(\tilde{f}_1(1, \cdot); [0,1]), \\
V((\tilde{g}_{\mathbf{f}_{1,n}} - \tilde{f}_1)(\cdot, 1); [0,1]) &\leq \frac{V(f_1(\cdot, 1); [0,1])}{\left\| g_{\mathbf{f}_{1,n}} \right\|_{L_2}} + V(\tilde{f}_1(\cdot, 1); [0,1]), \\
V((\tilde{g}_{\mathbf{f}_{1,n}} - \tilde{f}_1)(0, \cdot); [0,1]) &\leq \frac{V(f_1(1/n, \cdot); [0,1])}{\left\| g_{\mathbf{f}_{1,n}} \right\|_{L_2}} + V(\tilde{f}_1(0, \cdot); [0,1]), \\
V((\tilde{g}_{\mathbf{f}_{1,n}} - \tilde{f}_1)(\cdot, 0); [0,1]) &\leq \frac{V(f_1(\cdot, 1/n); [0,1])}{\left\| g_{\mathbf{f}_{1,n}} \right\|_{L_2}} + V(\tilde{f}_1(\cdot, 0); [0,1]).
\end{aligned}$$

Since $f_1(\cdot) \in BV([0,1]^2)$, by Proposition A.1.7, $V(f_1(1/n,\cdot);[0,1]) \leq \|\Phi_1\|_\infty$ and $V(f_1(\cdot,1/n);[0,1]) \leq \|\Psi_1\|_\infty$, where $\Phi_1(\cdot)$ and $\Psi_1(\cdot)$ are the total variation functions on $[0,1]$ associated with $f_1(\cdot)$, see Appendix A. Since $\left\|g_{\mathbf{f}_{1,n}}\right\|_{L_2}$ converges to $\|f_1\|_{L_2}$ as $n \to \infty$, the sequences on the right hand side of the preceding inequalities are clearly bounded from above by positive numbers. Consequently, the sequences on the left hand side are also bounded from above.

We now consider the case $i = 2$. Since $\tilde{g}_{\mathbf{f}_{2,n}}(\cdot) = \frac{g_{\mathbf{f}_{2,n}}(\cdot)-\langle \tilde{g}_{\mathbf{f}_{1,n}}, g_{\mathbf{f}_{2,n}}\rangle_{L_2}\tilde{g}_{\mathbf{f}_{1,n}}(\cdot)}{\left\|g_{\mathbf{f}_{2,n}}(\cdot)-\langle \tilde{g}_{\mathbf{f}_{1,n}}, g_{\mathbf{f}_{2,n}}\rangle_{L_2}\tilde{g}_{\mathbf{f}_{1,n}}(\cdot)\right\|_{L_2}}$, then by the triangle inequality, Remark 3.1.6 and by using the results obtained for $i = 1$, we obtain

$$\begin{aligned}
V(\tilde{g}_{\mathbf{f}_{2,n}} - \tilde{f}_2;[0,1]^2) &\leq \frac{V(f_2;[0,1]^2)}{\left\|g^*_{\mathbf{f}_{2,n}}\right\|_{L_2}} + \frac{\left|\langle \tilde{g}_{\mathbf{f}_{1,n}}, g_{\mathbf{f}_{2,n}}\rangle_{L_2}\right| V(f_1;[0,1]^2)}{\left\|g^*_{\mathbf{f}_{2,n}}\right\|_{L_2}\left\|g_{\mathbf{f}_{1,n}}\right\|_{L_2}} \\
&\quad + V(\tilde{f}_2;[0,1]^2), \\
V((\tilde{g}_{\mathbf{f}_{2,n}} - \tilde{f}_2)(1,\cdot);[0,1]) &\leq \frac{V(f_2(1,\cdot);[0,1])}{\left\|g^*_{\mathbf{f}_{2,n}}\right\|_{L_2}} + \frac{\left|\langle \tilde{g}_{\mathbf{f}_{1,n}}, g_{\mathbf{f}_{2,n}}\rangle_{L_2}\right| V(f_1(1,\cdot);[0,1])}{\left\|g^*_{\mathbf{f}_{2,n}}\right\|_{L_2}\left\|g_{\mathbf{f}_{1,n}}\right\|_{L_2}} \\
&\quad + V(\tilde{f}_2(1,\cdot);[0,1]), \\
V((\tilde{g}_{\mathbf{f}_{2,n}} - \tilde{f}_2)(\cdot,1);[0,1]) &\leq \frac{V(f_2(\cdot,1);[0,1])}{\left\|g^*_{\mathbf{f}_{2,n}}\right\|_{L_2}} + \frac{\left|\langle \tilde{g}_{\mathbf{f}_{1,n}}, g_{\mathbf{f}_{2,n}}\rangle_{L_2}\right| V(f_1(\cdot,1);[0,1])}{\left\|g^*_{\mathbf{f}_{2,n}}\right\|_{L_2}\left\|g_{\mathbf{f}_{1,n}}\right\|_{L_2}} \\
&\quad + V(\tilde{f}_2(\cdot,1);[0,1]), \\
V((\tilde{g}_{\mathbf{f}_{2,n}} - \tilde{f}_2)(0,\cdot);[0,1]) &\leq \frac{\|\Phi_2\|_\infty}{\left\|g^*_{\mathbf{f}_{2,n}}\right\|_{L_2}} + \frac{\left|\langle \tilde{g}_{\mathbf{f}_{1,n}}, g_{\mathbf{f}_{2,n}}\rangle_{L_2}\right| \|\Phi_1\|_\infty}{\left\|g^*_{\mathbf{f}_{2,n}}\right\|_{L_2}\left\|g_{\mathbf{f}_{1,n}}\right\|_{L_2}} + V(\tilde{f}_2(0,\cdot);[0,1]), \\
V((\tilde{g}_{\mathbf{f}_{2,n}} - \tilde{f}_2)(\cdot,0);[0,1]) &\leq \frac{\|\Psi_2\|_\infty}{\left\|g^*_{\mathbf{f}_{2,n}}\right\|_{L_2}} + \frac{\left|\langle \tilde{g}_{\mathbf{f}_{1,n}}, g_{\mathbf{f}_{2,n}}\rangle_{L_2}\right| \|\Psi_1\|_\infty}{\left\|g^*_{\mathbf{f}_{2,n}}\right\|_{L_2}\left\|g_{\mathbf{f}_{1,n}}\right\|_{L_2}} + V(\tilde{f}_2(\cdot,0);[0,1]),
\end{aligned}$$

where Φ_2 and Ψ_2 are total variation functions on $[0,1]$ associated with $f_2(\cdot)$, while

$$g^*_{\mathbf{f}_{2,n}}(\cdot) := g_{\mathbf{f}_{2,n}}(\cdot) - \langle \tilde{g}_{\mathbf{f}_{1,n}}, g_{\mathbf{f}_{2,n}}\rangle_{L_2}\tilde{g}_{\mathbf{f}_{1,n}}(\cdot).$$

By Corollary 3.1.12, all sequences on the right hand side of the preceding inequalities clearly converge to positive real numbers. Hence the sequences on the left hand side are bounded from above. The assertion for $i = 3, \ldots, p$ can be derived analogously. □

By Corollary 3.1.12 and Proposition 3.1.14, the conditions of Proposition 2.4.5 are clearly satisfied for $\psi_n(\cdot) = (\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(\cdot)$ as well as for $\psi_n(c, \cdot) = (\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(c, \cdot)$ and $\psi_n(\cdot, c) = (\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(\cdot, c)$, $n \geq 1$, $c = 0$ and 1. Thus the following version of Proposition 2.4.5 is straightforward.

Corollary 3.1.15. *If the regression functions $f_1(\cdot), \ldots, f_p(\cdot)$ are continuous and have bounded variation on $[0,1]^2$ in the sense of Hardy (in $BV([0,1]^2)$), then for $u(\cdot) \in \mathcal{C}([0,1]^2)$, and $c = 0$ and 1, we have*

$$\int_{[0,1]^2}^{R} u(t,s)\ d(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(t,s) \overset{n\to\infty}{\longrightarrow} 0,$$
$$\int_{[0,1]}^{R} u(c,s)\ d(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(c,s) \overset{n\to\infty}{\longrightarrow} 0,$$
$$\int_{[0,1]}^{R} u(t,c)\ d(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(t,c) \overset{n\to\infty}{\longrightarrow} 0.$$

Corollary 3.1.16. *Suppose the regression functions $f_1(\cdot), \ldots, f_p(\cdot)$ are continuous and have bounded variation on $[0,1]^2$ in the sense of Hardy. Let $(\nu_{(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)})_{n\geq 1}$, $(\nu_{(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(c,\cdot)})_{n\geq 1}$, and $(\nu_{(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(\cdot,c)})_{n\geq 1}$ be the sequence of finite Lebesgue-Stieltjes signed measures that correspond to the sequences $((\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i))_{n\geq 1}$, $((\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(c,\cdot))_{n\geq 1}$, and $((\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(\cdot,c))_{n\geq 1}$, respectively, $c = 0, 1$. Then for $u(\cdot) \in \mathcal{C}([0,1]^2)$, we have*

$$\int_{(0,1]^2} u(\cdot)\ d\bar{\nu}_{(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)} \overset{n\to\infty}{\longrightarrow} 0, \quad \int_{(0,1]} u(c,\cdot)\ d\bar{\nu}_{(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(c,\cdot)} \overset{n\to\infty}{\longrightarrow} 0,$$
$$\int_{(0,1]} u(\cdot,c)\ d\bar{\nu}_{(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(\cdot,c)} \overset{n\to\infty}{\longrightarrow} 0.$$

Proof. The result follows by Corollary 2.4.9 and Corollary 3.1.15. □

3.2 Residual partial sums limit processes

Having defined the notion of the reproducing kernel Hilbert space for Brownian (2) motion and the partial sums operator $\mathbf{T}_n$, we now establish the limit processes of the sequence of the residual partial sums processes $\{\mathbf{T}_n(\mathbf{R}_{n\times n})(t,s) : (t,s) \in [0,1]^2\}_{n\geq 1}$ for the linear regression model defined on the unit square $[0,1]^2$ starting with Lemma 3.2.1 below.

Lemma 3.2.1. *For any* $\mathbf{A} = (a_{\ell k})_{\ell=1,k=1}^{n,\ n} \in \mathbb{R}^{n\times n}$, *we have*

$$\mathbf{T}_n(pr_{\mathbf{W}_n}(\mathbf{A}))(\cdot) = (pr_{\mathbf{T}_n(\mathbf{W}_n)}\mathbf{T}_n(\mathbf{A})(\cdot))(\cdot) = (pr_{\mathbf{W}_{n\mathcal{H}_\mathbf{B}}}\mathbf{T}_n(\mathbf{A})(\cdot))(\cdot),\ \forall\ n \geq n_0,$$

where $\mathbf{T}_n(\mathbf{W}_n) := [\mathbf{T}_n(f_1(\mathcal{E}_n))(\cdot), \ldots, \mathbf{T}_n(f_p(\mathcal{E}_n))(\cdot)]$, *and* n_0 *is the natural number defined in Remark 3.1.10.*

Proof. Let $n \geq n_0$ be arbitrarily fixed. Since an orthogonal projection does not depend on any specific choice of a basis, without loss of generality we can assume that $\{f_1(\mathcal{E}_n), \ldots, f_p(\mathcal{E}_n)\}$ is an orthonormal basis of $\mathbf{W}_n$. By (3.1.10) the orthonormal basis of $\mathbf{W}_{n\mathcal{H}_\mathbf{B}}$ that corresponds to this basis is $\{\mathbf{T}_n(f_1(\mathcal{E}_n))(\cdot), \ldots, \mathbf{T}_n(f_p(\mathcal{E}_n))(\cdot)\}$. Hence, the orthogonal projection of $\mathbf{A} \in \mathbb{R}^{n\times n}$ onto $\mathbf{W}_n$ with respect to this basis is

$$pr_{\mathbf{W}_n}(\mathbf{A}) = \sum_{i=1}^{p} \langle \mathbf{A}, f_i(\mathcal{E}_n)\rangle_{\mathbb{R}^{n\times n}} f_i(\mathcal{E}_n). \tag{3.2.1}$$

Thus, by the linearity of the partial sums operator $\mathbf{T}_n$ on $\mathbb{R}^{n\times n}$, we get

$$\begin{aligned}
\mathbf{T}_n(pr_{\mathbf{W}_n}(\mathbf{A}))(\cdot) &= \mathbf{T}_n\left(\sum_{i=1}^{p} \langle \mathbf{A}, f_i(\mathcal{E}_n)\rangle_{\mathbb{R}^{n\times n}} f_i(\mathcal{E}_n)\right)(\cdot) \\
&= \sum_{i=1}^{p} \langle \mathbf{A}, f_i(\mathcal{E}_n)\rangle_{\mathbb{R}^{n\times n}} \mathbf{T}_n(f_i(\mathcal{E}_n))(\cdot) \\
&\overset{(3.1.10)}{=} \sum_{i=1}^{p} \langle \mathbf{T}_n(\mathbf{A})(\cdot), \mathbf{T}_n(f_i(\mathcal{E}_n))(\cdot)\rangle_{\mathcal{H}_\mathbf{B}} \mathbf{T}_n(f_i(\mathcal{E}_n))(\cdot) \\
&= (pr_{\mathbf{W}_{n\mathcal{H}_\mathbf{B}}}\mathbf{T}_n(\mathbf{A})(\cdot))(\cdot).
\end{aligned}$$

□

Definition 3.2.2. *Let $\mathcal{H}_{\mathbf{B}_{BV([0,1]^2)}}$ be the set of all functions in $\mathcal{H}_{\mathbf{B}}$ whose associated producing functions are in $BV([0,1]^2)$, i.e.,*

$$\mathcal{H}_{\mathbf{B}_{BV([0,1]^2)}} := \{h(\cdot) \in \mathcal{H}_{\mathbf{B}} : h(t,s) = \int_{[0,t]\times[0,s]} \hat{h}(\cdot)d\lambda^2,\ \hat{h}(\cdot) \in BV([0,1]^2)\}.$$

The mapping $\langle \cdot, \cdot \rangle$ is defined on $\mathcal{H}_{\mathbf{B}_{BV([0,1]^2)}} \times \mathcal{C}([0,1]^2)$, by

$$\begin{aligned}\langle h, u\rangle := \Delta_{[0,1]^2}(u\hat{h}) &- \int_{[0,1]}^{R} u(t,1)d\hat{h}(t,1) + \int_{[0,1]}^{R} u(t,0)d\hat{h}(t,0)\\ &- \int_{[0,1]}^{R} u(1,s)d\hat{h}(1,s) + \int_{[0,1]}^{R} u(0,s)d\hat{h}(0,s)\\ &+ \int_{[0,1]^2}^{R} u(t,s)d\hat{h}(t,s). \end{aligned} \tag{3.2.2}$$

Remark 3.2.3. *By Remark A.1.8, Theorem A.2.4 and Theorem 1.2.18 in Stroock (1994), all Riemann-Stieltjes integrals on the right hand side of (3.2.2) exist, hence the mapping $\langle \cdot, \cdot \rangle$ is well defined on $\mathcal{H}_{\mathbf{B}_{BV([0,1]^2)}} \times \mathcal{C}([0,1]^2)$. Since the Riemann-Stieltjes integral (3.2.2) is linear in $\hat{h}(\cdot)$ as well as in $u(\cdot)$, the mapping $\langle \cdot, \cdot \rangle$ constitutes a bilinear mapping on $\mathcal{H}_{\mathbf{B}_{BV([0,1]^2)}} \times \mathcal{C}([0,1]^2)$. Moreover, by using integration by parts for the Riemann-Stieltjes integral on the closed interval $[0,1]$, see Theorem 1.2.7 in Stroock (1994), we further get*

$$\begin{aligned}\langle h, u\rangle = &-\Delta_{[0,1]^2}(u\hat{h}) + \int_{[0,1]}^{R} \hat{h}(t,1)du(t,1) - \int_{[0,1]}^{R} \hat{h}(t,0)du(t,0)\\ &+ \int_{[0,1]}^{R} \hat{h}(1,s)du(1,s) - \int_{[0,1]}^{R} \hat{h}(0,s)du(0,s)\\ &+ \int_{[0,1]^2}^{R} u(t,s)d\hat{h}(t,s). \end{aligned} \tag{3.2.3}$$

Suppose for the moment that $h(\cdot) \in \mathcal{H}_{\mathbf{B}_{BV([0,1]^2)}}$ and $u(\cdot) \in \mathcal{H}_{\mathbf{B}}$ with the producing functions $\hat{h}(\cdot) \in BV([0,1]^2)$ and $\hat{u}(\cdot)$, respectively. By using integration by parts for *the Riemann-Stieltjes integrals on* $[0,1]^2$ *(see Theorem A.2.6), we obtain*

$$\begin{aligned}\int_{[0,1]^2}^{R} u(t,s)d\hat{h}(t,s) = \Delta_{[0,1]^2}(u\hat{h}) - \int_{[0,1]}^{R} \hat{h}(t,1)du(t,1) + \int_{[0,1]}^{R} \hat{h}(t,0)du(t,0)\\ - \int_{[0,1]}^{R} \hat{h}(1,s)du(1,s) + \int_{[0,1]}^{R} \hat{h}(0,s)du(0,s)\\ + \int_{[0,1]^2}^{R} \hat{h}(t,s)du(t,s). \qquad (3.2.4)\end{aligned}$$

We note that each of the Riemann-Stieltjes integrals that contribute to the right hand side of (3.2.4) exist, see also Young (1917a). By combining (3.2.3) and (3.2.4), we finally obtain

$$\begin{aligned}\langle h,u\rangle &= \int_{[0,1]^2}^{R} \hat{h}(t,s)du(t,s) \overset{(2.4.8)}{=} \int_{(0,1]^2} \hat{h}(t,s)\bar{\nu}_u(dt,ds)\\ &\overset{(3.1.3)}{=} \int_{[0,1]^2} \hat{h}(t,s)\hat{u}(t,s)\lambda^2(dt,ds) = \langle h,u\rangle_{\mathcal{H}_{\mathbf{B}}}, \qquad (3.2.5)\end{aligned}$$

provided $\hat{h}$ *is continuous* $\bar{\nu}_u$ *a.e. on* $(0,1]^2$*. Thus, the inner product* $\langle\cdot,\cdot\rangle_{\mathcal{H}_{\mathbf{B}}}$ *given by Equation (3.1.4) coincides with the bilinear form* $\langle\cdot,\cdot\rangle$ *on* $\mathcal{H}_{\mathbf{B}_{BV([0,1]^2)}} \times \mathcal{H}_{\mathbf{B}}$.

Remark 3.2.4. *By combining (3.2.2) and (A.2.1) given in the appendix, for every* $(h,u) \in \mathcal{H}_{\mathbf{B}_{BV([0,1]^2)}} \times \mathcal{C}([0,1]^2)$*, we have*

$$\begin{aligned}|\langle h,u\rangle| \leq \|u\|_\infty \Big(4\left\|\hat{h}\right\|_\infty + V(\hat{h}(\cdot,1);[0,1]) + V(\hat{h}(\cdot,0);[0,1])\\ + V(\hat{h}(1,\cdot);[0,1]) + V(\hat{h}(0,\cdot);[0,1]) + V(\hat{h};[0,1]^2)\Big), \qquad (3.2.6)\end{aligned}$$

where $V(\psi(\cdot);A)$*,* $A \subseteq [0,1]^2$ *denotes the total variation of* ψ *on* A*.*

Let us consider again Remark 3.1.11. The remark suggests that if the regression function $f_i(\cdot)$ is in $BV([0,1]^2)$, then the associated Gram-Schmidt orthonormal basis $\{\tilde{h}_{f_1}(\cdot),\ldots,\tilde{h}_{f_p}(\cdot)\}$ of $\mathbf{W}_{\mathcal{H}_{\mathbf{B}}}$ is in $\mathcal{H}_{\mathbf{B}_{BV([0,1]^2)}}$. The producing function of $\tilde{h}_{f_i}(\cdot)$ is given by $\tilde{f}_i(\cdot)$, $i=1,\ldots,p$. Based on this fact, by means of the bilinear mapping $\langle\cdot,\cdot\rangle$ just defined, we extend the orthogonal projection of $\mathcal{H}_{\mathbf{B}}$ onto $\mathbf{W}_{\mathcal{H}_{\mathbf{B}}}$ to a "projection" of $\mathcal{C}([0,1]^2)$ onto $\mathbf{W}_{\mathcal{H}_{\mathbf{B}}}$, see Proposition 3.2.5, below.

Proposition 3.2.5. *Let the regression function $f_i(\cdot)$ be continuous on $[0,1]^2$ and be in $BV([0,1]^2)$. Let $pr^*_{\mathbf{W}_{\mathcal{H}_{\mathbf{B}}}} : \mathcal{C}([0,1]^2) \to \mathbf{W}_{\mathcal{H}_{\mathbf{B}}}$ be the mapping defined on $\mathcal{C}([0,1]^2)$, such that for each $u(\cdot) \in \mathcal{C}([0,1]^2)$, and $(t,s) \in [0,1]^2$,*

$$(pr^*_{\mathbf{W}_{\mathcal{H}_{\mathbf{B}}}} u)(t,s) := \sum_{i=1}^{p} \langle \tilde{h}_{f_i}, u\rangle \tilde{h}_{f_i}(t,s), \tag{3.2.7}$$

where for $i = 1, \ldots, p$,

$$\begin{aligned}\langle \tilde{h}_{f_i}, u\rangle = \Delta_{[0,1]^2}(u\tilde{f}_i) &- \int_{[0,1]}^{R} u(t,1)d\tilde{f}_i(t,1) + \int_{[0,1]}^{R} u(t,0)d\tilde{f}_i(t,0)\\ &- \int_{[0,1]}^{R} u(1,s)d\tilde{f}_i(1,s) + \int_{[0,1]}^{R} u(0,s)d\tilde{f}_i(0,s)\\ &+ \int_{[0,1]^2}^{R} u(t,s)d\tilde{f}_i(t,s).\end{aligned}$$

Then $pr^*_{\mathbf{W}_{\mathcal{H}_{\mathbf{B}}}}$ constitutes a projection of $u(\cdot)$ onto $\mathbf{W}_{\mathcal{H}_{\mathbf{B}}}$.

Proof. By Definition 5.15 in Rudin (1991), it suffices to show that $pr^*_{\mathbf{W}_{\mathcal{H}_{\mathbf{B}}}}$ is linear and idempotent, i.e., $pr^*_{\mathbf{W}_{\mathcal{H}_{\mathbf{B}}}} \circ pr^*_{\mathbf{W}_{\mathcal{H}_{\mathbf{B}}}} = pr^*_{\mathbf{W}_{\mathcal{H}_{\mathbf{B}}}}$, and *surjective.* The first two conditions follow easily from the linearity of $\langle \cdot, \cdot \rangle$, the orthonormality of $\{\tilde{h}_{f_1}, \ldots, \tilde{h}_{f_p}\}$ in $\mathbf{W}_{\mathcal{H}_{\mathbf{B}}} \subset \mathcal{H}_{\mathbf{B}}$ and by the continuity of $\tilde{f}_i$, for $i = 1, \ldots, p$. The last condition is obvious by the definition and by inclusion $\mathbf{W}_{\mathcal{H}_{\mathbf{B}}} \subseteq \mathcal{C}([0,1]^2)$. We notice that in case $u(\cdot) \in \mathcal{H}_{\mathbf{B}}$, $pr^*_{\mathbf{W}_{\mathcal{H}_{\mathbf{B}}}}$ is an orthogonal projection of $u(\cdot)$ onto $\mathbf{W}_{\mathcal{H}_{\mathbf{B}}}$ with respect to the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}_{\mathbf{B}}}$. □

Let us define a mapping $pr^*_{\mathbf{W}_{n\mathcal{H}_{\mathbf{B}}}} : \mathcal{C}([0,1]^2) \to \mathbf{W}_{n\mathcal{H}_{\mathbf{B}}}$ by means of

$$(pr^*_{\mathbf{W}_{n\mathcal{H}_{\mathbf{B}}}} u_n)(t,s) := \sum_{i=1}^{p} \langle \tilde{h}_{\mathbf{f}_{i,n}}, u_n\rangle \tilde{h}_{\mathbf{f}_{i,n}}(t,s),\ (t,s) \in [0,1]^2, \tag{3.2.8}$$

where

$$\langle \tilde{h}_{\mathbf{f}_{i,n}}, u_n\rangle = \Delta_{[0,1]^2}(u_n\tilde{g}_{\mathbf{f}_{i,n}}) - \int_{[0,1]}^{R} u_n(t,1)d\tilde{g}_{\mathbf{f}_{i,n}}(t,1) + \int_{[0,1]}^{R} u_n(t,0)d\tilde{g}_{\mathbf{f}_{i,n}}(t,0)$$
$$- \int_{[0,1]}^{R} u_n(1,s)d\tilde{g}_{\mathbf{f}_{i,n}}(1,s) + \int_{[0,1]}^{R} u_n(0,s)d\tilde{g}_{\mathbf{f}_{i,n}}(0,s)$$
$$+ \int_{[0,1]^2}^{R} u_n(t,s)d\tilde{g}_{\mathbf{f}_{i,n}}(t,s).$$

By Remark 3.1.6 and Equation (3.1.17), $\{\tilde{h}_{\mathbf{f}_{1,n}}(\cdot)\ldots,\tilde{h}_{\mathbf{f}_{p,n}}(\cdot)\} \subseteq \mathcal{H}_{\mathbf{B}_{BV([0,1]^2)}}$ with the producing function of $\tilde{h}_{\mathbf{f}_{i,n}}(\cdot)$ given by $\tilde{g}_{\mathbf{f}_{i,n}}(\cdot)$ being in $BV([0,1]^2)$. Hence (3.2.8) is well defined on $\mathcal{C}([0,1]^2)$. Furthermore, by applying similar argument as in Remark 3.2.3, for any $u_n(\cdot) \in \mathcal{H}_{\mathbf{B}}$ with producing function $\hat{u}_n$, we have

$$\langle \tilde{h}_{\mathbf{f}_{i,n}}, u_n\rangle = \int_{[0,1]^2}^{R} \tilde{g}_{\mathbf{f}_{i,n}}(t,s)du_n(t,s) = \int_{(0,1]^2} \tilde{g}_{\mathbf{f}_{i,n}}(t,s)d\bar{\nu}_{u_n}$$
$$= \int_{[0,1]^2} \tilde{g}_{\mathbf{f}_{i,n}}(t,s)\hat{u}_n(t,s)\lambda^2(dt,ds) = \langle \tilde{h}_{\mathbf{f}_{i,n}}, u_n\rangle_{\mathcal{H}_{\mathbf{B}}},$$

since $\tilde{g}_{\mathbf{f}_{i,n}}(\cdot)$ is continuous $\bar{\nu}_{u_n}$ a.e. on $(0,1]^2$. Hence, for $u_n(\cdot) \in \mathcal{H}_{\mathbf{B}}$, we get

$$\begin{aligned}(pr^*_{\mathbf{W}_n\mathcal{H}_{\mathbf{B}}}u_n)(\cdot) &\overset{(3.2.8)}{=} \sum_{i=1}^{p}\langle \tilde{h}_{\mathbf{f}_{i,n}}, u_n\rangle \tilde{h}_{\mathbf{f}_{i,n}}(\cdot) \overset{(3.2.5)}{=} \sum_{i=1}^{p}\langle \tilde{h}_{\mathbf{f}_{i,n}}, u_n\rangle_{\mathcal{H}_{\mathbf{B}}} \tilde{h}_{\mathbf{f}_{i,n}}(\cdot)\\ &= (pr_{\mathbf{W}_n\mathcal{H}_{\mathbf{B}}}u_n)(\cdot).\end{aligned} \tag{3.2.9}$$

This leads us to the conclusion that $pr_{\mathbf{W}_n\mathcal{H}_{\mathbf{B}}}$ is a restriction of $pr^*_{\mathbf{W}_n\mathcal{H}_{\mathbf{B}}}$ on $\mathcal{H}_{\mathbf{B}}$, i.e., $pr^*_{\mathbf{W}_n\mathcal{H}_{\mathbf{B}}}|_{\mathcal{H}_{\mathbf{B}}} = pr_{\mathbf{W}_n\mathcal{H}_{\mathbf{B}}}$. We notice that this result will be important in establishing the limit process of the sequence of the residual partial sums processes, see Theorem 3.2.6 below.

Theorem 3.2.6. ***(residual partial sums limit processes)***

Suppose that $f_1(\cdot),\ldots,f_p(\cdot)$ are linearly independent. If $f_i(\cdot)$, $i=1,\ldots,p$, is continuous and has bounded variation on $[0,1]^2$ in the sense of Hardy, than, as $n\to\infty$,

$$\frac{1}{\sigma}\mathbf{T}_n(\mathbf{R}_{n\times n})(\cdot) \xrightarrow{\mathcal{D}} B_2(\cdot) - (pr^*_{\mathbf{W}_{\mathcal{H}_{\mathbf{B}}}}B_2)(\cdot), \text{ in } \mathcal{C}([0,1]^2),$$

where $\mathbf{R}_{n\times n}$ *is the matrix of residuals, and* $B_2(\cdot)$ *is standard Brownian (2) motion.*

Proof. Since $\mathbf{R}_{n\times n} = \mathbf{E}_{n\times n} - pr_{\mathbf{W}_n}\mathbf{E}_{n\times n}$, by Lemma 3.2.1, we have

$$\begin{aligned} \frac{1}{\sigma}\mathbf{T}_n(\mathbf{R}_{n\times n})(\cdot) &= \frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(\cdot) - \frac{1}{\sigma}\mathbf{T}_n(pr_{\mathbf{W}_n}\mathbf{E}_{n\times n})(\cdot) \\ &= \frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(\cdot) - (pr_{\mathbf{W}_{n\mathcal{H}_\mathbf{B}}}\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(\cdot))(\cdot) \qquad (3.2.10) \end{aligned}$$

We use the results of the preceding section to show that whenever $(u_n(\cdot))_{n\geq 1}$ is a sequence in $\mathcal{C}([0,1]^2)$ such that $u_n(\cdot) \stackrel{\|\cdot\|_\infty}{\longrightarrow} u(\cdot)$, then $(pr^*_{\mathbf{W}_{n\mathcal{H}_\mathbf{B}}}u_n)(\cdot) \stackrel{\|\cdot\|_\infty}{\longrightarrow} (pr^*_{\mathbf{W}_{\mathcal{H}_\mathbf{B}}}u)(\cdot)$, as $n\to\infty$. By the triangle inequality and by the linearity of the Riemann-Stieltjes integral, we have

$$\begin{aligned} &\sup_{(t,s)\in[0,1]^2}\left|(pr^*_{\mathbf{W}_{n\mathcal{H}_\mathbf{B}}}u_n)(t,s) - (pr^*_{\mathbf{W}_{\mathcal{H}_\mathbf{B}}}u)(t,s)\right| \\ &= \sup_{(t,s)\in[0,1]^2}\left|\sum_{i=1}^p\langle\tilde{h}_{\mathbf{f}_{i,n}},u_n\rangle\tilde{h}_{\mathbf{f}_{i,n}}(t,s) - \sum_{i=1}^p\langle\tilde{h}_{f_i},u\rangle\tilde{h}_{f_i}(t,s)\right| \\ &\leq \sup_{(t,s)\in[0,1]^2}\sum_{i=1}^p\left|\langle\tilde{h}_{\mathbf{f}_{i,n}},u_n\rangle\tilde{h}_{\mathbf{f}_{i,n}}(t,s) - \langle\tilde{h}_{f_i},u\rangle\tilde{h}_{f_i}(t,s)\right| \\ &= \sup_{(t,s)\in[0,1]^2}\sum_{i=1}^p\Big|\langle\tilde{h}_{\mathbf{f}_{i,n}},u_n\rangle\tilde{h}_{\mathbf{f}_{i,n}}(t,s) - \langle\tilde{h}_{\mathbf{f}_{i,n}},u_n\rangle\tilde{h}_{f_i}(t,s) + \langle\tilde{h}_{\mathbf{f}_{i,n}},u_n\rangle\tilde{h}_{f_i}(t,s) \\ &\quad -\langle\tilde{h}_{f_i},u_n\rangle\tilde{h}_{f_i}(t,s) + \langle\tilde{h}_{f_i},u_n\rangle\tilde{h}_{f_i}(t,s) - \langle\tilde{h}_{f_i},u\rangle\tilde{h}_{f_i}(t,s)\Big| \\ &\leq \sum_{i=1}^p\left|\langle(\tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i}) + \tilde{h}_{f_i},(u_n-u)+u\rangle\right|\left\|\tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i}\right\|_\infty \\ &\quad + \sum_{i=1}^p\left|\langle\tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i},(u_n-u)+u\rangle\right|\left\|\tilde{h}_{f_i}\right\|_\infty + \sum_{i=1}^p\left|\langle\tilde{h}_{f_i},u_n-u\rangle\right|\left\|\tilde{h}_{f_i}\right\|_\infty \\ &\leq \sum_{i=1}^p\left|\langle\tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i},u_n-u\rangle\right|\left(\left\|\tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i}\right\|_\infty + \left\|\tilde{h}_{f_i}\right\|_\infty\right) \\ &\quad + \sum_{i=1}^p\left|\langle\tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i},u\rangle\right|\left(\left\|\tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i}\right\|_\infty + \left\|\tilde{h}_{f_i}\right\|_\infty\right) \\ &\quad + \sum_{i=1}^p\left|\langle\tilde{h}_{f_i},u_n-u\rangle\right|\left(\left\|\tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i}\right\|_\infty + \left\|\tilde{h}_{f_i}\right\|_\infty\right) + \sum_{i=1}^p\left|\langle\tilde{h}_{f_i},u\rangle\right|\left\|\tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i}\right\|_\infty. \end{aligned}$$

We now define sequences $(\tilde{K}_{in}^{(1)})_{n\geq 1}$, $(\tilde{K}_{in}^{(2)})_{n\geq 1}$, and positive numbers $\tilde{K}_i$ by putting

$$\begin{aligned}
\tilde{K}_{in}^{(1)} &:= V((\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(\cdot,1);[0,1]) + V((\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(\cdot,0);[0,1]) \\
&\quad + V((\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(1,\cdot);[0,1]) + V((\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(0,\cdot);[0,1]) + V(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i;[0,1]^2), \\
\tilde{K}_{in}^{(2)} &:= \left|\int_{[0,1]}^{R} u(t,1)d(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(t,1)\right| + \left|\int_{[0,1]}^{R} u(t,0)d(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(t,0)\right| \\
&\quad + \left|\int_{[0,1]}^{R} u(1,s)d(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(1,s)\right| + \left|\int_{[0,1]}^{R} u(0,s)d(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(0,s)\right| \\
&\quad + \left|\int_{[0,1]^2}^{R} u(t,s)d(\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i)(t,s)\right|, \\
\tilde{K}_i &:= V(\tilde{f}_i(\cdot,1);[0,1]) + V(\tilde{f}_i(\cdot,0);[0,1]) \\
&\quad + V(\tilde{f}_i(1,\cdot);[0,1]) + V(\tilde{f}_i(0,\cdot);[0,1]) + V(\tilde{f}_i;[0,1]^2),
\end{aligned}$$

for $i = 1, \ldots, p$. By applying Inequality (A.2.1) to the absolute value of the bilinear forms involved in the last inequality, we further obtain

$$\begin{aligned}
&\sup_{(t,s)\in[0,1]^2} \left|(pr^*_{\mathbf{W}_{n\mathcal{H}_\mathbf{B}}} u_n)(t,s) - (pr^*_{\mathbf{W}_{\mathcal{H}_\mathbf{B}}} u)(t,s)\right| \\
&\leq \sum_{i=1}^{p} \|u_n - u\|_\infty \left(4\left\|\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i\right\|_\infty + \tilde{K}_{in}^{(1)}\right)\left(\left\|\tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i}\right\|_\infty + \left\|\tilde{h}_{f_i}\right\|_\infty\right) \\
&\quad + \sum_{i=1}^{p} \left(4\|u\|_\infty \left\|\tilde{g}_{\mathbf{f}_{i,n}} - \tilde{f}_i\right\|_\infty + \tilde{K}_{in}^{(2)}\right)\left(\left\|\tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i}\right\|_\infty + \left\|\tilde{h}_{f_i}\right\|_\infty\right) \\
&\quad + \sum_{i=1}^{p} \|u_n - u\|_\infty \left(4\left\|\tilde{f}_i\right\|_\infty + \tilde{K}_i\right)\left(\left\|\tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i}\right\|_\infty + \left\|\tilde{h}_{f_i}\right\|_\infty\right) \\
&\quad + \sum_{i=1}^{p} \|u\|_\infty \left(4\left\|\tilde{f}_i\right\|_\infty + \tilde{K}_i\right)\left\|\tilde{h}_{\mathbf{f}_{i,n}} - \tilde{h}_{f_i}\right\|_\infty. \qquad (3.2.11)
\end{aligned}$$

Since $\tilde{K}_{in}^{(1)}$ is bounded (see Proposition 3.1.14), $\tilde{K}_{in}^{(2)} \overset{n\to\infty}{\longrightarrow} 0$ (see Corollary 3.1.15) and $\tilde{K}_i$ is bounded (since $\tilde{f}_i(\cdot) \in BV([0,1]^2)$), $i = 1, \ldots, p$, then by Corollary 3.1.12, Corollary 3.1.13 and by the hypothesis, the right hand side of (3.2.11) converges to zero as $n \to \infty$. Therefore, $(pr^*_{\mathbf{W}_{n\mathcal{H}_\mathbf{B}}} u_n)(\cdot)$ converges to $(pr^*_{\mathbf{W}_{\mathcal{H}_\mathbf{B}}} u)(\cdot)$ uniformly on

$[0,1]^2$, whenever $(u_n(\cdot))_{n\geq 1}$ converges uniformly to $u(\cdot)$ in $\mathcal{C}([0,1]^2)$. Furthermore, since $\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(\cdot) \xrightarrow{D} B_2(\cdot)$ (see Theorem 2.3.4), by (3.2.10) and Theorem 2.3.7, for $n \to \infty$ we have

$$(pr_{\mathbf{W}_{n\mathcal{H}_\mathbf{B}}}\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(\cdot))(\cdot) = (pr^*_{\mathbf{W}_{n\mathcal{H}_\mathbf{B}}}\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(\cdot))(\cdot) \xrightarrow{D} (pr^*_{\mathbf{W}_{\mathcal{H}_\mathbf{B}}}B_2)(\cdot).$$

This completes the proof. □

Remark 3.2.7. *By Theorem 2.4.4, the sequence* $(\tilde{K}^{(2)}_{in})_{n\geq 1}$ *in (3.2.11) can be replaced by the sequence* $(\tilde{K}^{(2)'}_{in})_{n\geq 1}$*, given by*

$$\begin{aligned}\tilde{K}^{(2)'}_{in} := &\left|\int_{(0,1]^2} u(\cdot)d\bar{\nu}_{(\tilde{g}_{\mathbf{f}_{i,n}}-\tilde{f}_i)}\right| + \left|\int_{(0,1]} u(1,\cdot)d\bar{\nu}_{(\tilde{g}_{\mathbf{f}_{i,n}}-\tilde{f}_i)(1,\cdot)}\right| \\ &+ \left|\int_{(0,1]} u(\cdot,1)d\bar{\nu}_{(\tilde{g}_{\mathbf{f}_{i,n}}-\tilde{f}_i)}(\cdot,1)\right| + \left|\int_{(0,1]} u(0,\cdot)d\bar{\nu}_{(\tilde{g}_{\mathbf{f}_{i,n}}-\tilde{f}_i)}(0,\cdot)\right| \\ &+ \left|\int_{(0,1]} u(\cdot,0)d\bar{\nu}_{(\tilde{g}_{\mathbf{f}_{i,n}}-\tilde{f}_i)}(\cdot,0)\right|,\end{aligned}$$

which by Corollary 3.1.16 converges to zero as $n \to \infty$*, for* $i = 1,\ldots,p$*.*

Remark 3.2.8. *Unless otherwise stated, we abbreviate the limit process* $\{B_2(t,s) - (pr^*_{\mathbf{W}_{\mathcal{H}_\mathbf{B}}}B_2)(t,s) : (t,s) \in [0,1]^2\}$ *by* $B_\mathbf{f}$*, where the index* $\mathbf{f}$ *stands for the vector of regression functions* $\mathbf{f} = (f_1(\cdot),\ldots,f_p(\cdot))^\top$*. The function* $K_\mathbf{f}(\cdot) : [0,1]^2 \times [0,1]^2 \to \mathbb{R}$ *given by* $K_\mathbf{f}((t,s),(t',s')) := Cov(B_\mathbf{f}(t,s), B_\mathbf{f}(t',s'))$*, for* $(t,s),(t',s') \in [0,1]^2 \times [0,1]^2$ *is called the covariance function of* $B_\mathbf{f}$*. For any* $(t,s) \in [0,1]^2$*, it is obvious that* $B_\mathbf{f}(t,s)$ *is a zero mean Gaussian process with variance* $K_\mathbf{f}((t,s),(t,s)) > 0$*.*

Remark 3.2.9. *By Slutsky's theorem, in case the variance* σ^2 *is unknown, we can replace* σ *in (3.2.10) by* $\hat{\sigma}_n := \sqrt{\hat{\sigma}^2_n}$ *given by (1.2.3) without altering the convergence in distribution in* $\mathcal{C}([0,1]^2)$*.*

3.3 Examples

In this subsection we discuss the limit process of the residual partial sums process associated with constant, first-order and second-order models. For each considered model we identify the distribution of $B_{\mathbf{f}}(t,s)$ for each $(t,s) \in [0,1]^2$.

Example 3.3.1. *As a simple case, we consider a constant model* $\mathbb{E}(Y) = \beta$, *where* β *is an unknown parameter. For this model we have* $f(t,s) = 1$, $(t,s) \in [0,1]^2$, *and* $\mathbf{W}_{\mathcal{H}_\mathbf{B}} = [\tilde{h}_f(\cdot)]$, *where* $\tilde{h}_f(t,s) = ts$, $(t,s) \in [0,1]^2$. *Since* $B_2(t,s) = 0$ *a.s. if* $t = 0$ *or* $s = 0$, *the projection of* $B_2(\cdot)$ *onto* $\mathbf{W}_{\mathcal{H}_\mathbf{B}}$ *is*

$$(pr^*_{\mathbf{W}_{\mathcal{H}_\mathbf{B}}} B_2)(t,s) = ts\langle \tilde{h}_f, B_2 \rangle = tsB_2(1,1),\ (t,s) \in [0,1]^2.$$

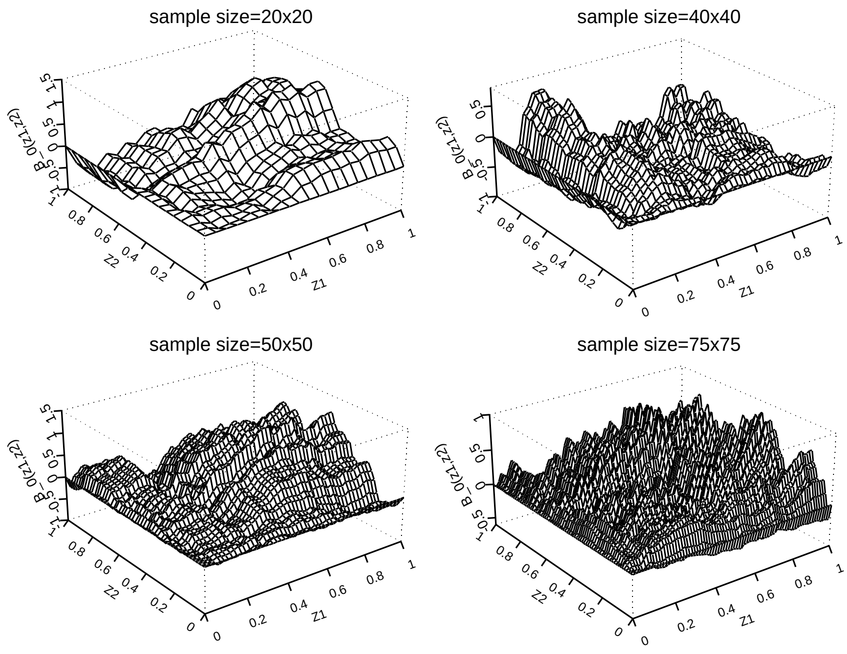

Figure 5. *Approximation to sample paths of the standard Brownian (2) bridge.*

Thus, the residual partial sums limit process is

$$B_{\mathbf{f}}(t,s) = B_2(t,s) - tsB_2(1,1),\ (t,s) \in [0,1]^2,$$

which is the standard Brownian (2) bridge defined in Section 2.2. An example of sample paths of the standard Brownian (2) bridge can be seen in Figure 5 for which the paths are approximated by residual partial sums processes $\frac{1}{\sigma}\mathbf{T}_n(\mathbf{R}_{n\times n})(t,s)$ corresponding to the constant model.

Example 3.3.2. *If the model is a full first-order model (plane), i.e.,* $\mathbb{E}(Y) = \beta_0 + \beta_1 t + \beta_2 s$, $(t,s) \in [0,1]^2$, *where* β_0, β_1 *and* β_2 *are unknown parameters, with regression functions* $f_1(t,s) = 1$, $f_2(t,s) = t$, $f_3(t,s) = s$ *that are clearly continuous on* $[0,1]^2$ *and linearly independent as functions in* $L_2([0,1]^2)$. *The Gram-Schmidt orthonormal basis in* $\mathbf{W}_{\mathcal{H}_\mathbf{B}}$ *associated with this model is*

$$\begin{aligned}
\tilde{h}_{f_1}(t,s) &= ts,\\
\tilde{h}_{f_2}(t,s) &= \sqrt{3}ts(t-1),\\
\tilde{h}_{f_3}(t,s) &= \sqrt{3}ts(s-1),\ (t,s)\in[0,1]^2.
\end{aligned}$$

Let us denote the associated limit process by $B_{\mathbf{f}^{(1)}}(t,s)$. *Since* $B_2(t,s) = 0$ *a.s. if* $t = 0$ *or* $s = 0$, *by (3.2.7) we get*

$$\begin{aligned}
B_{\mathbf{f}^{(1)}}(t,s) = B_2^0(t,s) &- 3ts(t+s-2)B_2(1,1)\\
&+ 6ts(t-1)\int_{[0,1]} B_2(t,1)dt + 6ts(s-1)\int_{[0,1]} B_2(1,s)ds,
\end{aligned}$$

where B_2^0 *is the standard Brownian (2) bridge. Furthermore, by a little computation the covariance function* $K_{\mathbf{f}^{(1)}}(\cdot,\cdot)$ *of* $B_{\mathbf{f}^{(1)}}(\cdot)$ *turns out to be*

$$\begin{aligned}
K_{\mathbf{f}^{(1)}}((t,s),(t',s')) =& (t\wedge t')(s\wedge s') - tst's' + 3t's'(t'-1)s + 3t's'(s'-1)t\\
&- 3t's'(t'-1)ts - 3t's'(s'-1)ts + 3ts(t-1)s' - 3ts(t-1)t's'\\
&+ 3ts(s-1)t' - 3ts(s-1)t's'.
\end{aligned}$$

Example 3.3.3. *We consider a full second-order model, i.e.,*

$$\mathbb{E}(Y) = \beta_0 + t\beta_1 + s\beta_2 + t^2\beta_{11} + ts\beta_{12} + s^2\beta_{22},\ (t,s) \in [0,1]^2,$$

where $\beta_0, \beta_1, \beta_2, \beta_{11}, \beta_{22}$*, and* β_{12} *are unknown parameters. It is clear that the regression functions* $f_1(t,s) = 1$*,* $f_2(t,s) = t$*,* $f_3(t,s) = s$*,* $f_4(t,s) = t^2$*,* $f_5(t,s) = ts$ *and* $f_6(t,s) = s^2$ *are continuous on* $[0,1]^2$ *and linearly independent in* $L_2([0,1]^2)$*. The Gram-Schmidt orthonormal basis of* $\mathbf{W}_{\mathcal{H}_\mathbf{B}}$ *associated with this model is*

$$\begin{aligned}
\tilde{h}_{f_1}(t,s) &= ts, \\
\tilde{h}_{f_2}(t,s) &= \sqrt{3}ts(t-1), \\
\tilde{h}_{f_3}(t,s) &= \sqrt{3}ts(s-1), \\
\tilde{h}_{f_4}(t,s) &= \sqrt{5}(2t^3s - 3t^2s + ts), \\
\tilde{h}_{f_5}(t,s) &= \frac{1}{3}(t^2s^2 - t^2s - ts^2 + ts), \\
\tilde{h}_{f_6}(t,s) &= \sqrt{5}(2ts^3 - 3ts^2 + ts).
\end{aligned}$$

We denote the associated residual partial sums limit process by $B_{\mathbf{f}^{(2)}}(t,s)$*. Then, after simplifying the Riemann-Stieltjes integrals involved, we get*

$$\begin{aligned}
B_{\mathbf{f}^{(2)}}(t,s) = {} & B_{\mathbf{f}^{(1)}}(t,s) \\
& - (10t^3s + t^2s^2/9 - 136t^2s/9 - 136ts^2/9 + 10ts^3 + 91ts/9)B_2(1,1) \\
& + (120t^3s - 180t^2s + 60ts)\int_{[0,1]} B_2(t,1)tdt \\
& + (2t^2s^2/9 - 60t^3s + 808t^2s/9 - 2ts^2/9 - 268ts/9)\int_{[0,1]} B_2(t,1)dt \\
& + (2t^2s^2/9 - 2t^2s/9 + 808ts^2/9 - 60ts^3 - 268ts/9)\int_{[0,1]} B_2(1,s)ds \\
& + (120ts^3 - 180ts^2 + 60ts)\int_{[0,1]} B_2(1,s)sds
\end{aligned}$$

$$- 4/9(t^2s^2 - t^2s - ts^2 + ts) \int_{[0,1]^2} B_2(t,s)dtds.$$

3.4 Extension to an $n \times m$ regular lattice on $[0,1]^2$

The extension of the preceding results to an $n \times m$ regular lattice, denoted by $\mathcal{E}_{nm}$ is obvious and can be directly derived. However, it is important to give a little description how we should work in such a situation. If model (1.1.2) is extended to the $n \times m$ regular lattice, the sample space is $\mathbb{R}^{m\times n}$, the matrix of observations is $\mathbf{Y}_{m\times n} := (Y_{\ell k})_{k=1,\ell=1}^{m,\ n} \in \mathbb{R}^{m\times n}$, which is assumed to satisfy the equation $\mathbf{Y}_{m\times n} = \mathbf{M}_{m\times n} + \mathbf{E}_{m\times n}$, for some $\mathbf{M}_{m\times n} \in V_{mn} \subset \mathbb{R}^{m\times n}$, with $V_{mn} = [f_1(\mathcal{E}_{nm}), \ldots, f_p(\mathcal{E}_{nm})]$, where $\mathbf{E}_{m\times n} = (\varepsilon_{\ell,k})_{k=1,\ell=1}^{m,\ n}$ is an $m \times n$ dimensional random matrix with components that are i.i.d $(0,\sigma^2)$, $\sigma^2 \in (0,\infty)$, and $f_1(\cdot), \ldots, f_p(\cdot)$ are the regression functions defined on $[0,1]^2$. The matrix of least squares residuals is

$$\mathbf{R}_{m\times n} = pr_{\mathbf{V}_{mn}^{\perp}} \mathbf{Y}_{m\times n} = pr_{\mathbf{V}_{mn}^{\perp}} \mathbf{E}_{m\times n}.$$

By using the *vec* operator the model can also be presented in the form

$$vec(\mathbf{Y}_{m\times n}) = \mathbf{X}_{mn}\beta + vec(\mathbf{E}_{m\times n}),$$

where the design matrix $X_{mn} := (vec(f_1(\mathcal{E}_{nm})), \ldots, vec(f_p(\mathcal{E}_{nm})))$ is in $\mathbb{R}^{nm\times p}$. Suppose that $(f_1(\mathcal{E}_{nm}), \ldots, f_p(\mathcal{E}_{nm})) \subset \mathbb{R}^{m\times n}$ is a basis matrix of $\mathbf{V}_{mn}$. Then the statistic

$$\begin{aligned}\hat{\sigma}^2_{nm} &:= \frac{\left\| pr_{\mathbf{V}_{mn}^{\perp}} \mathbf{Y}_{m\times n} \right\|^2_{\mathbb{R}^{m\times n}}}{nm - p} = \frac{\left\| pr_{\mathbf{X}_{mn}^{\perp}} vec(\mathbf{Y}_{m\times n}) \right\|^2_{\mathbb{R}^{nm}}}{nm - p} \\ &= \frac{vec^{\top}(\mathbf{Y}_{m\times n})(\mathbf{I}_{mn\times mn} - \mathbf{X}_{mn}(\mathbf{X}_{mn}^{\top}\mathbf{X}_{mn})^{-1}\mathbf{X}_{mn}^{\top})vec(\mathbf{Y}_{m\times n})}{nm - p} \end{aligned} \tag{3.4.1}$$

is a consistent estimator of σ^2 in the sense that $\hat{\sigma}^2_{nm}$ converges in probability to σ^2, as $n, m \to \infty$ simultaneously.

Let us define a linear operator $\mathbf{T}_{nm} : \mathbb{R}^{m\times n} \to \mathcal{C}([0,1]^2)$, such that for every $\mathbf{A} := (a_{\ell k})_{\ell=1,k=1}^{n,\ m}$ and $(z_1, z_2) \in [0,1]^2$,

$$\begin{aligned}\mathbf{T}_{nm}(\mathbf{A})(z_1,z_2) = &\sum_{k=1}^{[mz_2]}\sum_{\ell=1}^{[nz_1]} a_{\ell k} + (nz_1 - [nz_1])\sum_{k=1}^{[mz_2]} a_{[nz_1]+1,k}\\ &+ (mz_2 - [mz_2])\sum_{\ell=1}^{[nz_1]} a_{\ell,[mz_2]+1}\\ &+ (nz_1 - [nz_1])(mz_2 - [mz_2])a_{[nz_1]+1,[mz_2]+1},\end{aligned}$$

where $\sum_{k=1}^{j}\sum_{\ell=1}^{i} a_{\ell k} = 0$, for $j = 0$ or $i = 0$. Analogous to the partial sums operator $\mathbf{T}_n$ defined in Chapter 2, the operator $\mathbf{T}_{nm}$ embeds the random matrix $\mathbf{E}_{m\times n}$ and the matrix of least squares residuals $\mathbf{R}_{m\times n}$ into stochastic processes $\{\mathbf{T}_{nm}(\mathbf{E}_{m\times n})(t,s) : (t,s) \in [0,1]^2\}$ and $\{\mathbf{T}_{nm}(\mathbf{R}_{m\times n})(t,s) : (t,s) \in [0,1]^2\}$, respectively, whose sample paths are in $\mathcal{C}([0,1]^2)$. Quite analogous to Theorem 2.3.4, for $n,m \to \infty$, we get $\frac{1}{\sigma\sqrt{nm}}\mathbf{T}_{nm}(\mathbf{E}_{m\times n})(\cdot) \xrightarrow{\mathcal{D}} B_2(\cdot)$, in $\mathcal{C}([0,1]^2)$, where $B_2(\cdot)$ is the standard Brownian (2) motion. Finally, if the regression functions $f_1(\cdot),\ldots,f_p(\cdot)$ are continuous and have bounded variation on $[0,1]^2$, then for $n,m \to \infty$, we have $\frac{1}{\sigma\sqrt{nm}}\mathbf{T}_{nm}(\mathbf{R}_{m\times n})(\cdot) \xrightarrow{\mathcal{D}} B_{\mathbf{f}}(\cdot)$ in $\mathcal{C}([0,1]^2)$, where $B_{\mathbf{f}}(\cdot)$ is the residual partial sums limit process defined in Theorem 3.2.6. Thus the residual partial sums limit processes associated to both models are the same.

Chapter 4

Tests based on residual partial sums processes

In this chapter we apply the weak convergence theory elaborated in Chapter 3 to the theory of linear models. In particular we consider *model-checks* for spatial data. Based on a set of data from an experiment conducted on the regular lattice $\mathcal{E}_n \subseteq [0,1]^2$ or more generally on $\mathcal{E}_{nm} \subseteq [0,1]^2$, we test the hypothesis that a true but unknown regression function belongs to a certain subspace generated by finitely many known regression functions. More exactly, we decide the problem by a test based on the residual partial sums process associated to the linear regression model formulated in Chapter 1.

We start this chapter with the formulation of the hypotheses, then define asymptotic tests, finally study the asymptotic behavior of these procedures by applying the results derived in Chapter 3. We shall propose three test statistics which are familiar in nonparametric statistical theory: a *Kolmogorov* type statistic, a *Kolmogorov-Smirnov* type statistic and a *Cramér-von Mises* type statistic. These will be defined in Section 4.2, Section 4.3 and Section 4.4, respectively. To construct the asymptotic critical region of the tests, we calculate the quantiles of the three test statistics by simulation.

Finally, in Section 4.5 we observe the *consistency* and the asymptotic power of the tests by establishing the asymptotic distribution of each statistic under *alternatives*. At the end of Section 4.5 we conduct Monte Carlo simulations to approximate the limiting power of the tests. Throughout this chapter, we use the definitions and notations introduced in the preceding chapters.

4.1 Formulation of the Hypotheses

We consider the regression model

$$Y(t,s) = g(t,s) + \varepsilon(t,s),\ \ (t,s) \in [0,1]^2, \tag{4.1.1}$$

where $g : [0,1]^2 \to \mathbb{R}$ is an unknown real-valued function. We assume that $\varepsilon(t,s)$ is a real-valued random variable with mean 0 and variance $\sigma^2 \in (0,\infty)$ and $g(\cdot)$ is in $BVV([0,1]^2)$. We are interested in testing the hypothesis

$$H_0 : g \in \mathbf{W} \text{ versus } K : g \notin \mathbf{W}, \tag{4.1.2}$$

where $\mathbf{W} = [f_1(\cdot),\dots,f_p(\cdot)]$ is the linear subspace generated by known regression functions $f_1(\cdot),\dots,f_p(\cdot)$. We assume that the conditions of Theorem 3.2.6 are fulfilled by the vector of regression functions $\mathbf{f} = (f_1(\cdot),\dots,f_p(\cdot))^\top$, i.e., they are linearly independent, continuous and have bounded variation in the sense of Hardy on the closed square $[0,1]^2$.

Suppose that $\mathbf{Y}_{n\times n} = (Y_{\ell k})_{k=1,\ell=1}^{n,\ n}$ is an $n \times n$ dimensional matrix of independent observations associated with model (4.1.1), taken from the regular lattice $\mathcal{E}_n$. By using matrix notation this may be rewritten as

$$\mathbf{Y}_{n\times n} = \mathbf{G}_{n\times n} + \mathbf{E}_{n\times n}, \tag{4.1.3}$$

where $\mathbf{E}_{n\times n} = (\varepsilon_{\ell k})_{k=1,\ell=1}^{n,\ n}$ is the matrix of random errors with $\mathbb{E}(\varepsilon_{\ell k}) = 0$ and $Var(\varepsilon_{\ell k}) = \sigma^2 \in (0,\infty)$, and $\mathbf{G}_{n\times n} := (g(\ell/n, k/n))_{k=1,\ell=1}^{n,\ n} \in \mathbb{R}^{n\times n}$. Thus based on these observations, the problem of testing (4.1.2) can be implemented by performing a test of the hypotheses

$$H_0 : \mathbf{G}_{n\times n} \in \mathbf{W}_n \text{ versus } K : \mathbf{G}_{n\times n} \notin \mathbf{W}_n, \tag{4.1.4}$$

for each $n \in \mathbb{N}$, where $\mathbf{W}_n = [f_1(\mathcal{E}_n), \ldots, f_p(\mathcal{E}_n)]$, see Definition 3.1.9. Firstly, we fix the *asymptotic critical region* of the tests by establishing the asymptotic distribution of each test statistic under H_0. Later, we consider the asymptotic distribution under K, which is important in determining the power of the tests. For our test problems, the sample space is $(\mathbb{R}^{n\times n}, \mathcal{B}(\mathbb{R}^{n\times n}), \{P_g : g \in BVV([0,1]^2)\})$, $n \in \mathbb{N}$, where $\{P_g : g \in BVV([0,1]^2)\}$ is a family of unknown probability measures that depend on $g(\cdot) \in BVV([0,1]^2$ induced by the matrix of observations $\mathbf{Y}_{n\times n}$ which is assumed to be defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Under the operator $\mathbf{T}_n$, the subspace $pr_{\mathbf{W}_n^\perp}\mathbb{R}^{n\times n}$ is isomorphic to $\mathbf{T}_n(pr_{\mathbf{W}_n^\perp}\mathbb{R}^{n\times n})(\cdot) \subseteq \mathcal{H}_{\mathbf{B}} \subseteq \mathcal{C}([0,1]^2)$, see (3.1.1).

Definition 4.1.1. *Let $\gamma_n : \mathcal{C}([0,1]^2) \to \{0,1\}$, $n \geq 1$, be a sequence of nonrandomized tests based on residual partial sums processes $\{\mathbf{T}_n(\mathbf{R}_{n\times n})(t,s) : (t,s) \in [0,1]^2\}_{n\geq 1}$ for testing (4.1.2) or (4.1.4). Let $\alpha \in (0,1)$ and for any $g(\cdot) \in BVV([0,1]^2)$, let $\mathbb{E}_g(\cdot)$ be the expectation operator with respect to the probability measure P_g.*

- *We say that $(\gamma_n)_{n\geq 1}$ is a sequence of pointwise asymptotically level α tests, if and only if*

$$\limsup_{n\to\infty} \mathbb{E}_g(\gamma_n) \leq \alpha \text{ for each } g(\cdot) \in \mathbf{W}.$$

- *The sequence $(\gamma_n)_{n\geq 1}$ is said to be a sequence of uniformly asymptotically level α tests, if and only if*

$$\limsup_{n\to\infty} \sup_{g\in\mathbf{W}} \mathbb{E}_g(\gamma_n) \leq \alpha.$$

- *We say that* $(\gamma_n)_{n\geq 1}$ *is a sequence of asymptotically size* α *tests, if and only if*

$$\lim_{n\to\infty} \sup_{g\in\mathbf{W}} \mathbb{E}_g(\gamma_n) = \alpha.$$

- *The sequence* $(\gamma_n)_{n\geq 1}$ *is said to be pointwise consistent in power, if and only if*

$$\lim_{n\to\infty} \mathbb{E}_g(\gamma_n) = 1 \text{ for each } g(\cdot) \notin \mathbf{W}.$$

We refer the reader to Lehmann and Romano (2005), p. 422-423, for these Definitions. Note that every sequence of asymptotically size α *tests is a sequence of uniformly asymptotically level* α *tests.*

4.2 Kolmogorov type test

A test of Kolmogorov type for testing the hypotheses (4.1.2) or (4.1.4) is defined by means of the following functional of the residual partial sums process:

$$K_{n,\mathbf{f}} := \max_{0\leq k,\ell\leq n} \frac{1}{n}\sum_{j=0}^{k}\sum_{i=0}^{\ell} r_{ij} = \max_{0\leq k,\ell\leq n} \mathbf{T}_n(\mathbf{R}_{n\times n})(\ell/n, k/n).$$

Here and throughout, $r_{ij} := 0$ for $i = 0$ or $j = 0$, and $\mathbf{R}_{n\times n}$ is the matrix of residuals. Due to (2.3.1) it is obvious that

$$K_{n,\mathbf{f}} = \sup_{(t,s)\in[0,1]^2} \mathbf{T}_n(\mathbf{R}_{n\times n})(t,s). \tag{4.2.1}$$

Proposition 4.2.1. *Suppose* σ^2 *is known. For a fixed* $\alpha \in (0,1)$*, an asymptotically size* α *test for testing (4.1.2) or (4.1.4) based on* $K_{n,\mathbf{f}}$ *is given by*

$$\textit{reject } H_0, \textit{ if and only if } K_{n,\mathbf{f}}/\sigma \geq \tilde{c}_{1-\alpha},$$

where $\tilde{c}_{1-\alpha}$ *is the* $(1-\alpha)$*-quantile of* $\sup_{(t,s)\in[0,1]^2} B_{\mathbf{f}}(t,s)$.

Proof. We define a sequence of *nonrandomized* tests $(\delta_n)_{n\geq 1}: \mathcal{C}([0,1]^2) \longrightarrow \{0,1\}$, by

$$\delta_n(\mathbf{T}_n(\mathbf{R}_{n\times n})(\cdot)) := \begin{cases} 1, & \text{if } K_{n,\mathbf{f}}/\sigma \geq \tilde{c}, \\ 0, & \text{otherwise,} \end{cases} \tag{4.2.2}$$

where the constant $\tilde{c}$ is the *critical value* of the test. Let $g(\cdot) \in \mathbf{W}$ be fixed . From the continuity of the supremum function and Theorem 3.2.6 we obtain for $\tilde{c} = \tilde{c}_{1-\alpha}$

$$\begin{aligned} \lim_{n\to\infty} \mathbb{E}_g(\delta_n) &= \lim_{n\to\infty} \mathbb{P}\left\{K_{n,\mathbf{f}}/\sigma \geq \tilde{c}_{1-\alpha}\right\} \\ &= \lim_{n\to\infty} \mathbb{P}\left\{\sup_{(t,s)\in[0,1]^2} \frac{1}{\sigma}\mathbf{T}_n(pr_{\mathbf{W}_n^\perp}\mathbf{Y}_{n\times n})(t,s) \geq \tilde{c}_{1-\alpha}\right\} \\ &\overset{(1.2.1)}{=} \lim_{n\to\infty} \mathbb{P}\left\{\sup_{(t,s)\in[0,1]^2} \frac{1}{\sigma}\mathbf{T}_n(pr_{\mathbf{W}_n^\perp}\mathbf{E}_{n\times n})(t,s) \geq \tilde{c}_{1-\alpha}\right\} \\ &= \mathbb{P}\left\{\sup_{(t,s)\in[0,1]^2} B_{\mathbf{f}}(t,s) \geq \tilde{c}_{1-\alpha}\right\} = \alpha. \end{aligned} \tag{4.2.3}$$

□

Remark 4.2.2. *If σ^2 is unknown, we can replace it by any consistent estimator. By Slutsky's theorem and the preceding result, for $n \to \infty$ we have*

$$K_{n,\mathbf{f}}/\sqrt{\hat{\sigma}_n^2} \xrightarrow{\mathcal{D}} \sup_{(t,s)\in[0,1]^2} B_{\mathbf{f}}(t,s),$$

where $\hat{\sigma}_n^2$ is given by (1.2.2), see also Arnold (1981), p. 142-148.

The problem in realizing the asymptotically size α test formulated above is to find the $(1-\alpha)$-quantile of $\sup_{(t,s)\in[0,1]^2} B_{\mathbf{f}}(t,s)$ analytically or approximately. In case the stochastic process $B_{\mathbf{f}}(\cdot)$ is standard Brownian motion with parameter space $[0,1]$, it is well known by the *reflection principle* of the standard Brownian motion that $\mathbb{P}\{\sup_{t\in[0,1]} B(t) \geq \tilde{c}\} = 2\mathbb{P}\{Z > \tilde{c}\}$, $\tilde{c} > 0$, where Z is a standard normal random variable, see also Shorack and Wellner (1986), p. 33-37 and Billingsley (1999), p.

91-93. For our case, if $B_{\mathbf{f}}(\cdot)$ is the standard Brownian (2) motion B_2, Zimmerman (1972) derived the inequality $\mathbb{P}\left\{\sup_{(t,s)\in[0,1]^2} B_2(t,s) \geq \tilde{c}\right\} \leq 4\,\mathbb{P}\{Z \geq \tilde{c}\}$. This yields an upper bound of $\mathbb{P}\{\sup_{(t,s)\in[0,1]^2} B_2(t,s) \geq \tilde{c}\}$ for each preassigned value of $\tilde{c}$, so that we can use this inequality for checking our simulation results, see below.

4.2.1 Approximation of the quantiles of $\sup_{(t,s)\in[0,1]^2} B_{\mathbf{f}}(t,s)$

As an alternative solution to the problem described above, we estimate the $(1-\alpha)$-quantile of $\sup_{(t,s)\in[0,1]^2} B_{\mathbf{f}}(t,s)$ by applying *Monte Carlo* simulations generated according to Algorithm 1 below. The simulations are constructed only for several polynomial models under H_0: null, constant, first order and second order models.

Numerically, the matrix of least squares residuals is computed by the equation $\mathbf{R}_{n\times n} = \mathbf{Y}_{n\times n} - \sum_{i=1}^{p} \hat{\beta}_{ni} f_i(\mathcal{E}_n)$, or equivalently $vec(\mathbf{R}_{n\times n}) = vec(\mathbf{Y}_{n\times n}) - \mathbf{X}_n\hat{\beta}_n$, where $\hat{\beta}_n := (\hat{\beta}_{n1}, \ldots, \hat{\beta}_{np})^\top$ is the solution of the *normal* equation $(\mathbf{X}_n^\top \mathbf{X}_n)\beta = \mathbf{X}_n^\top vec(\mathbf{Y}_{n\times n})$ for β. This system of linear equations can be solved numerically by applying either *Gaussian elimination*, *Cholesky factorization* or *QR factorization*, see Gentle (1998), p. 87-112. The statistical software package S-PLUS and R provide a *macro* for solving such a normal equation, i.e., by applying the command: **solve**$(\mathbf{X}_n^\top \mathbf{X}_n, \mathbf{X}_n^\top vec(\mathbf{Y}_{n\times n}))$.

`Begin Algorithm 1`

step 1: Fix $\tilde{n} \in \mathbb{N}$.

step 2: Generate M i.i.d. pseudo random matrices $\mathbf{E}^{(j)}_{\tilde{n}\times\tilde{n}} := (\varepsilon_{\ell k j})_{k=1,\ell=1}^{\tilde{n},\,\tilde{n}}$, with components $\varepsilon_{\ell k j}$ generated from i.i.d. $\mathcal{N}(0,1)$ random variables, $j = 1, \ldots, M$.

step 3: Calculate $\hat{\beta}^{(j)}_{\tilde{n}}$ by solving the equation $(\mathbf{X}_{\tilde{n}}^\top \mathbf{X}_{\tilde{n}})\beta = \mathbf{X}_{\tilde{n}}^\top vec(\mathbf{Y}^{(j)}_{\tilde{n}\times\tilde{n}})$.

step 4: Calculate the matrix of residuals $\mathbf{R}^{(j)}_{\tilde{n}\times\tilde{n}} := \mathbf{Y}^{(j)}_{\tilde{n}\times\tilde{n}} - \sum_{i=1}^{p} \hat{\beta}^{(j)}_{\tilde{n}i} f_i(\mathcal{E}_n)$.

step 5: Calculate the statistic $K_{\tilde{n},\mathbf{f}}^{(j)} := \max_{0\le k,\ell\le\tilde{n}} \mathbf{T}_{\tilde{n}}(\mathbf{R}_{\tilde{n}\times\tilde{n}}^{(j)})(\ell/\tilde{n}, k/\tilde{n})$.

step 6: Calculate the simulated $(1-\alpha)$-quantiles of $\sup_{(t,s)\in[0,1]^2} B_{\mathbf{f}}(t,s)$: Sort all M values of $K_{\tilde{n},\mathbf{f}}^{(j)}$ in ascending order. Let $K_{\tilde{n},\mathbf{f}}^{(M:j)}$ be the j'th smallest observation, i.e., $K_{\tilde{n},\mathbf{f}}^{(M:1)} \le \ldots \le K_{\tilde{n},\mathbf{f}}^{(M:j)} \le K_{\tilde{n},\mathbf{f}}^{(M:j+1)} \le \ldots \le K_{\tilde{n},\mathbf{f}}^{(M:M)}$, the simulated $(1-\alpha)$-quantile is

$$\tilde{c}_{1-\alpha} = \begin{cases} K_{\tilde{n},\mathbf{f}}^{(M:M(1-\alpha))}, & \text{if } M(1-\alpha) \in \mathbb{N}, \\ \\ K_{\tilde{n},\mathbf{f}}^{(M:[M(1-\alpha)]+1)}, & \text{otherwise,} \end{cases}$$

where $[M(1-\alpha)] = \max\{k \in \mathbb{N} : k \le M(1-\alpha)\}$.

`End Algorithm 1`

$\mathbb{P}\{\sup_{(t,s)\in[0,1]^2} B_{\mathbf{f}}(t,s) \ge \tilde{c}_{1-\alpha}\} = \alpha$					
Models	$\tilde{c}_{0.5000}$	$\tilde{c}_{0.6500}$	$\tilde{c}_{0.7500}$	$\tilde{c}_{0.8000}$	$\tilde{c}_{0.8500}$
Zero	0.8648	1.1002	1.2942	1.4129	1.5564
Constant	0.8101	0.9416	1.0452	1.1067	1.1794
First order	0.7232	0.7917	0.8470	0.8810	0.9211
Second order	0.6392	0.6889	0.7293	0.7540	0.7838
Models	$\tilde{c}_{0.9000}$	$\tilde{c}_{0.9500}$	$\tilde{c}_{0.9750}$	$\tilde{c}_{0.9900}$	$\tilde{c}_{0.9950}$
Zero	1.7443	2.0348	2.3001	2.6126	2.8353
Constant	1.2728	1.4152	1.5408	1.6906	1.7965
First order	0.9739	1.0570	1.1329	1.2264	1.2930
Second order	0.8231	0.8849	0.9413	1.0098	1.0590

Table 1. The simulated $(1-\alpha)$-quantiles of $\sup_{(t,s)\in[0,1]^2} B_{\mathbf{f}}(t,s)$.

The simulation results obtained by using the statistical software package R 2.0.1 are presented in Table 1, for $\alpha = 0.0050, 0.0100, 0.0250, 0.0500, 0.1000, 0.1500, 0.2000,$

0.2500, 0.3600 and 0.5000. For our results we chose sample size $\tilde{n} \times \tilde{n} = 30 \times 30$ and the number of replications is $M = 10^6$. It can be seen therein that for each α, the higher the order of the model is, the smaller the simulated quantile $\tilde{c}_{1-\alpha}$ is. In case the residual partial sums limit process is standard Brownian (2) motion B_2 we can use Zimmerman's inequality for checking our simulation results. As is shown in Table 2, for each value of $\tilde{c}_{1-\alpha}$ associated to the zero model, the simulated value of $\mathbb{P}\{\sup_{(t,s)\in[0,1]^2} B_2(t,s) \geq \tilde{c}_{1-\alpha}\}$ is smaller than the corresponding value of $4\mathbb{P}\{Z \geq \tilde{c}_{1-\alpha}\}$, where $Z \sim \mathcal{N}(0,1)$.

$\tilde{c}_{1-\alpha}$	0.8648	1.1002	1.2942	1.4129	1.5564
$\mathbb{P}\{\sup B_2(t,s) \geq \tilde{c}_{1-\alpha}\}$	0.5000	0.3500	0.2500	0.2000	0.1500
$4\mathbb{P}\{Z \geq \tilde{c}_{1-\alpha}\}$	0.7743	0.5425	0.3912	0.3154	0.2392
$\tilde{c}_{1-\alpha}$	1.7443	2.0348	2.3001	2.6126	2.8353
$\mathbb{P}\{\sup B_2(t,s) \geq \tilde{c}_{1-\alpha}\}$	0.1000	0.0500	0.0250	0.0100	0.0050
$4\mathbb{P}\{Z \geq \tilde{c}_{1-\alpha}\}$	0.1622	0.0837	0.0429	0.0180	0.0092

Table 2. Comparison of $\mathbb{P}\{\sup_{(t,s)\in[0,1]^2} B_2(t,s) \geq \tilde{c}_{1-\alpha}\}$ and $4\mathbb{P}\{Z \geq \tilde{c}_{1-\alpha}\}$.

Remark 4.2.3. *For the regular lattice $\mathcal{E}_{nm}$ defined in Subsection 3.2.2, the Kolmogorov type statistic is defined by the following functional of the residual partial sums process:*

$$\begin{aligned} K_{nm,\mathbf{f}} &:= \max_{0\leq k\leq m;\ 0\leq \ell\leq n} \frac{1}{\sqrt{nm}} \sum_{j=0}^{k}\sum_{i=0}^{\ell} r_{ij} \\ &= \max_{0\leq k\leq m;\ 0\leq \ell\leq n} \frac{1}{\sqrt{nm}} \mathbf{T}_{nm}(\mathbf{R}_{m\times n})(\ell/n, k/m) \\ &= \sup_{0\leq t\leq 1;\ 0\leq s\leq 1} \frac{1}{\sqrt{nm}} \mathbf{T}_{nm}(\mathbf{R}_{m\times n})(t,s), \end{aligned} \tag{4.2.4}$$

where $\mathbf{T}_{nm}$ is the partial sums operator defined in Subsection 3.2.2. Analogous to Proposition 4.2.1, for a fixed $\alpha \in (0,1)$ and known σ, the asymptotically size α test *for testing (4.1.2) or (4.1.4) based on the statistic* $K_{nm,\mathbf{f}}$ *is given by*

$$\textit{reject } H_0, \textit{ if and only if } K_{nm,\mathbf{f}}/\sigma \geq \tilde{c}_{1-\alpha},$$

where $\tilde{c}_{1-\alpha}$ *is the* $(1-\alpha)$*-quantile of* $\sup_{(t,s)\in[0,1]^2} B_{\mathbf{f}}(t,s)$. *If* σ^2 *is unknown, it may be replaced by* $\sqrt{\hat{\sigma}^2_{nm}}$ *given by (3.4.1).*

4.3 Kolmogorov-Smirnov type test

A Kolmogorov-Smirnov type procedure for testing the hypotheses (4.1.2) or (4.1.4) is defined by means of the following functional of the residual partial sums process:

$$KS_{n,\mathbf{f}} := \max_{0\leq k,\ell\leq n} \frac{1}{n}\left|\sum_{j=0}^{k}\sum_{i=0}^{\ell} r_{ij}\right| = \max_{0\leq k,\ell\leq n} \left|\mathbf{T}_n(\mathbf{R}_{n\times n})(\ell/n, k/n)\right|.$$

By analogy with the Kolmogorov statistic, it is obvious that

$$KS_{n,\mathbf{f}} = \sup_{(t,s)\in[0,1]^2} \left|\mathbf{T}_n(\mathbf{R}_{n\times n})(t,s)\right|. \tag{4.3.1}$$

Proposition 4.3.1. *Let* σ^2 *be known. For a fixed* $\alpha \in (0,1)$*, an asymptotically size* α *test for testing hypotheses (4.1.2) or (4.1.4) based on the statistic* $KS_{n,\mathbf{f}}$ *is given by*

$$\textit{reject } H_0, \textit{ if and only if } KS_{n,\mathbf{f}}/\sigma \geq \tilde{q}_{1-\alpha},$$

where $\tilde{q}_{1-\alpha}$ *is the* $(1-\alpha)$*-quantile of* $\sup_{(t,s)\in[0,1]^2} |B_{\mathbf{f}}(t,s)|$. *If* σ^2 *is unknown, then the test is given by*

$$\textit{reject } H_0, \textit{ if and only if } KS_{n,\mathbf{f}}/\sqrt{\hat{\sigma}^2_n} \geq \tilde{q}_{1-\alpha},$$

where $\hat{\sigma}^2_n$ *is a consistent estimator of* σ^2.

Proof. We first assume that σ^2 is known and define a sequence of non randomized tests $(\lambda_n)_{n\geq 1} : \mathcal{C}([0,1]^2) \longrightarrow \{0,1\}$ by

$$\lambda_n(\mathbf{T}_n(\mathbf{R}_{n\times n})(\cdot)) := \begin{cases} 1, & \text{if } KS_{n,\mathbf{f}}/\sigma \geq \tilde{q}, \\ 0, & \text{otherwise,} \end{cases} \tag{4.3.2}$$

where $\tilde{q}$ is a constant. Analogous to the Kolmogorov type test, by Theorem 3.2.6 and the continuous mapping theorem, for $g(\cdot) \in \mathbf{W}$, we get for $\tilde{q} = \tilde{q}_{1-\alpha}$,

$$\begin{aligned} \lim_{n\to\infty} \mathbb{E}_g(\lambda_n) &= \lim_{n\to\infty} \mathbb{P}\{KS_{n,\mathbf{f}}/\sigma \geq \tilde{q}_{1-\alpha}\} \\ &= \lim_{n\to\infty} \mathbb{P}\left\{ \sup_{(t,s)\in[0,1]^2} \left| \frac{1}{\sigma}\mathbf{T}_n(pr_{\mathbf{W}_n^{\perp}}\mathbf{Y}_{n\times n})(t,s) \right| \geq \tilde{q}_{1-\alpha} \right\} \\ &\overset{(1.2.1)}{=} \lim_{n\to\infty} \mathbb{P}\left\{ \sup_{(t,s)\in[0,1]^2} \left| \frac{1}{\sigma}\mathbf{T}_n(pr_{\mathbf{W}_n^{\perp}}\mathbf{E}_{n\times n})(t,s) \right| \geq \tilde{q}_{1-\alpha} \right\} \\ &= \mathbb{P}\left\{ \sup_{(t,s)\in[0,1]^2} |B_{\mathbf{f}}(t,s)| \geq \tilde{q}_{1-\alpha} \right\} = \alpha. \end{aligned} \tag{4.3.3}$$

Since $g(\cdot)$ is arbitrary in $\mathbf{W}$, the test is asymptotically size α. By using the same argument as that in the test of Kolmogorov type, in case σ^2 is unknown, we can replace σ by the square root of the consistent estimator $\hat{\sigma}_n^2$ given by (1.2.2) without altering (4.3.3). Thus, instead of $KS_{n,\mathbf{f}}/\sigma$ we evaluate $KS_{n,\mathbf{f}}/\sqrt{\hat{\sigma}_n^2}$. □

4.3.1 Approximation of the quantiles of $\sup_{(t,s)\in[0,1]^2} |B_{\mathbf{f}}(t,s)|$

Similar to the Kolmogorov type test, we conduct Monte Carlo simulation to approximate the $(1-\alpha)$-quantiles of $\sup_{(t,s)\in[0,1]^2} |B_{\mathbf{f}}(t,s)|$ by applying Algorithm 1 with the only modification that $K_{n,\mathbf{f}}$ in *step 5* and *step 6* of Algorithm 1 is replaced by $KS_{n,\mathbf{f}}$ and the quantity $\tilde{c}_{1-\alpha}$ is replaced by $\tilde{q}_{1-\alpha}$. The simulations are constructed for several polynomial models under H_0: null, constant, first order and second order model, each of which is calculated for sample size $\tilde{n} \times \tilde{n} = 30 \times 30$ and the number of replication is

$M = 10^6$. Table 3 presents the simulation results corresponding to the zero, constant, first order and second order model executed by using the statistical software package R 2.0.1, for $\alpha = 0.0050$, 0.0100, 0.0250, 0.0500, 0.1000, 0.1500, 0.2000, 0.2500, 0.3500 and 0.5000.

$\mathbb{P}\{\sup_{(t,s)\in[0,1]^2} \lvert B_{\mathbf{f}}(t,s)\rvert \geq \tilde{q}_{1-\alpha}\} = \alpha$					
Models	$\tilde{q}_{0.5000}$	$\tilde{q}_{0.6500}$	$\tilde{q}_{0.7500}$	$\tilde{q}_{0.8000}$	$\tilde{q}_{0.8500}$
Zero	1.2922	1.4804	1.6421	1.7432	1.8676
Constant	1.0326	1.1352	1.2194	1.2710	1.3330
First order	0.8212	0.8845	0.9356	0.9670	1.0040
Second order	0.7010	0.7493	0.7885	0.8123	0.8408
Models	$\tilde{q}_{0.9000}$	$\tilde{q}_{0.9500}$	$\tilde{q}_{0.9750}$	$\tilde{q}_{0.9900}$	$\tilde{q}_{0.9950}$
Zero	2.0335	2.2973	2.5370	2.8335	3.0380
Constant	1.4146	1.54036	1.6562	1.7967	1.8958
First order	1.0536	1.1306	1.2029	1.2920	1.3527
Second order	0.9920	1.0577	1.1050	1.0577	1.1048

Table 3: The simulated $(1-\alpha)$ quantiles of $\sup_{(t,s)\in[0,1]^2} |B_{\mathbf{f}}(t,s)|$.

Remark 4.3.2. *The Kolmogorov-Smirnov type statistic for the residual partial sums process associated to the linear regression model defined on the regular lattice $\mathcal{E}_{nm}$ is*

$$\begin{aligned} KS_{nm,\mathbf{f}} &:= \max_{0\leq k\leq m;\ 0\leq \ell\leq n} \frac{1}{\sqrt{nm}} \left| \sum_{j=0}^{k} \sum_{i=0}^{\ell} r_{ij} \right| \\ &= \max_{0\leq k\leq m;\ 0\leq \ell\leq n} \frac{1}{\sqrt{nm}} \left| \mathbf{T}_{nm}(\mathbf{R}_{m\times n})(\ell/n, k/m) \right| \\ &= \sup_{0\leq t\leq 1;\ 0\leq s\leq 1} \frac{1}{\sqrt{nm}} \left| \mathbf{T}_{nm}(\mathbf{R}_{m\times n})(t,s) \right|. \end{aligned} \tag{4.3.4}$$

For a fixed $\alpha \in (0,1)$, if σ is known, the asymptotically size α test for testing (4.1.2) or (4.1.4) based on $KS_{nm,\mathbf{f}}$ is given by

$$\textit{reject } H_0, \textit{ if and only if } KS_{nm,\mathbf{f}}/\sigma \geq \tilde{q}_{1-\alpha},$$

where $\tilde{q}_{1-\alpha}$ is the $(1-\alpha)$ quantile of $\sup_{(t,s)\in[0,1]^2} |B_{\mathbf{f}}(t,s)|$. *If σ^2 is unknown, we can replace σ by $\sqrt{\hat{\sigma}^2_{nm}}$ given by (3.4.1).*

4.4 Cramér-von Mises type test

In this section we present a Cramér-von Mises type procedure for testing hypothesis (4.1.2) or (4.1.4) based on the residual partial sums processes.

Proposition 4.4.1. *Let us consider the following functional of the residual partial sums processes associated with model (4.1.3),*

$$C_{n,\mathbf{f}} := \frac{1}{n^4}\sum_{k=0}^{n}\sum_{\ell=0}^{n}\left(\sum_{j=0}^{k}\sum_{i=0}^{\ell} r_{ij}\right)^2 = \frac{1}{n^2}\sum_{k=0}^{n}\sum_{\ell=0}^{n}(\mathbf{T}_n(\mathbf{R}_{n\times n})(\ell/n,k/n))^2. \qquad (4.4.1)$$

Suppose σ^2 is known. For a fixed $\alpha\in(0,1)$, the asymptotically size α test for testing (4.1.2) or (4.1.4) based on $C_{n,\mathbf{f}}$ is given by

$$\text{reject } H_0, \text{ if and only if } C_{n,\mathbf{f}}/\sigma^2 \geq \tilde{t}_{1-\alpha},$$

where $\tilde{t}_{1-\alpha}$ is the $(1-\alpha)$-quantile of $\int_{[0,1]^2} B^2_{\mathbf{f}}(t,s)\lambda^2(dt,ds)$. If σ^2 is unknown, then the test is given by

$$\text{reject } H_0, \text{ if and only if } C_{n,\mathbf{f}}/\hat{\sigma}^2_n \geq \tilde{t}_{1-\alpha},$$

where $\hat{\sigma}^2_n$ is a consistent estimator of σ^2.

Proof. By definition, we have

$$C_{n,\mathbf{f}} = \int_{[0,1]^2} (\mathbf{T}_n(\mathbf{R}_{n\times n})(\ell/n,k/n))^2\,\lambda^2(dt,ds),$$

where $\mathbf{R}_{n\times n}$ is the matrix of least squares residuals. Consider a sequence of nonrandomized tests $(\psi_n)_{n\geq 1} : \mathcal{C}([0,1]^2) \longrightarrow \{0,1\}$, given by

$$\psi_n(\mathbf{T}_n(\mathbf{R}_{n\times n})(\cdot)) := \begin{cases} 1, & \text{if } C_{n,\mathbf{f}}/\sigma^2 \geq \tilde{t}, \\ 0, & \text{otherwise,} \end{cases} \tag{4.4.2}$$

where $\tilde{t}$ is a constant. By Theorem 3.2.6 and the continuous mapping theorem, for a fixed $\alpha \in (0,1)$ and an arbitrary $g(\cdot) \in \mathbf{W}$, we obtain for $\tilde{t} = \tilde{t}_{1-\alpha}$

$$\begin{aligned} \lim_{n\to\infty} \mathbb{E}_g(\psi_n) &= \lim_{n\to\infty} \mathbb{P}\left\{ \int_{[0,1]^2} \left(\frac{1}{\sigma} \mathbf{T}_n(pr_{\mathbf{W}_n^\perp} \mathbf{Y}_{n\times n})(\frac{\ell}{n}, \frac{k}{n}) \right)^2 \lambda^2(dt, ds) \geq \tilde{t}_{1-\alpha} \right\} \\ &\overset{(1.2.1)}{=} \lim_{n\to\infty} \mathbb{P}\left\{ \int_{[0,1]^2} \left(\frac{1}{\sigma} \mathbf{T}_n(pr_{\mathbf{W}_n^\perp} \mathbf{E}_{n\times n})(\frac{\ell}{n}, \frac{k}{n}) \right)^2 \lambda^2(dt, ds) \geq \tilde{t}_{1-\alpha} \right\} \\ &= \mathbb{P}\left\{ \int_{[0,1]^2} B_{\mathbf{f}}^2(t,s) \lambda^2(dt, ds) \geq \tilde{t}_{1-\alpha} \right\} = \alpha, \end{aligned} \tag{4.4.3}$$

where $\int_{[0,1]^2} B_{\mathbf{f}}^2(t,s)\lambda^2(dt,ds)$ can be understood as a *pathwise* Lebesgue or Riemann integral on $[0,1]^2$. Then the test is asymptotically of size α. The second assertion is trivial. □

4.4.1 Approximation of the quantiles of $\int_{[0,1]^2} B_{\mathbf{f}}^2(\cdot)\ d\lambda^2$

Though its very intensive use in testing problems based on the residual partial sums process, for instance, MacNeill and Jandhyala (1993) and Xie and MacNeill (2004) used the statistic $\int_{[0,1]^2} B_{\mathbf{f}}^2(t,s)\lambda^2(dt,ds)$ in change-point problems for spatial data, but exact as well as approximation methods for calculating the quantiles of the limiting statistic have not yet been derived. In this subsection we conduct Monte Carlo simulations for approximating the $(1-\alpha)$-quantiles of this statistic by applying a similar algorithm as Algorithm 1 with the modification that $K_{n,\mathbf{f}}$ in *step 5* and *step 6* of Algorithm 1 is replaced by $C_{n,\mathbf{f}}$ and $\tilde{c}_{1-\alpha}$ is replaced by $\tilde{t}_{1-\alpha}$. The simulations are constructed for the null, the constant, the first order, and the second order model, each of which is calculated for the sample size $\tilde{n} \times \tilde{n} = 30 \times 30$ and the number of replications $M = 10^6$. Table 4 presents our simulation results which are executed by using software package R 2.0.1, for $\alpha = 0.0050, 0.0100, 0.0250, 0.0500, 0.1000, 0.1500, 0.2000, 0.2500, 0.3600$ and 0.5000.

$\mathbb{P}\{\int_{[0,1]^2} B_{\mathbf{f}}^2(t,s)\lambda^2(dt,ds) \geq \tilde{t}_{1-\alpha}\} = \alpha$					
Models	$\tilde{t}_{0.5000}$	$\tilde{t}_{0.6500}$	$\tilde{t}_{0.7500}$	$\tilde{t}_{0.8000}$	$\tilde{t}_{0.8500}$
Zero	0.1803	0.2521	0.3293	0.3850	0.4590
Constant	0.1156	0.1463	0.1761	0.1966	0.2240
First order	0.0683	0.0794	0.0890	0.0951	0.1027
Second order	0.0450	0.0506	0.0553	0.0582	0.0618
Models	$\tilde{t}_{0.9000}$	$\tilde{t}_{0.9500}$	$\tilde{t}_{0.9750}$	$\tilde{t}_{0.9900}$	$\tilde{t}_{0.9950}$
Zero	0.5699	0.7694	0.9769	1.2606	1.4802
Constant	0.2643	0.3370	0.4127	0.5160	0.5973
First order	0.1131	0.1305	0.1477	0.1695	0.1860
Second order	0.0668	0.0750	0.0828	0.0928	0.1001

Table 4. The approximated $(1-\alpha)$ quantiles of $\int_{[0,1]^2} B_{\mathbf{f}}^2(t,s)\lambda^2(dt,ds)$.

Remark 4.4.2. *The Cramér-von Mises type statistic for the residual partial sums process associated to the linear regression model defined on the regular lattice $\mathcal{E}_{nm}$ is*

$$\begin{aligned} C_{nm,\mathbf{f}} &:= \frac{1}{nm}\sum_{k=0}^{m}\sum_{\ell=0}^{n}\left(\frac{1}{\sqrt{nm}}\sum_{j=0}^{k}\sum_{i=0}^{\ell} r_{ij}\right)^2 \\ &= \frac{1}{nm}\sum_{k=0}^{m}\sum_{\ell=0}^{n}\left(\frac{1}{\sqrt{nm}}\mathbf{T}_{nm}(\mathbf{R}_{m\times n})(\ell/n,k/m)\right)^2. \end{aligned} \tag{4.4.4}$$

For a fixed $\alpha \in (0,1)$, if σ is known, the asymptotically size α test for testing (4.1.2) or (4.1.4) based on the statistic $C_{nm,\mathbf{f}}$ is given by

$$\text{reject } H_0, \text{ if and only if } C_{nm,\mathbf{f}}/\sigma^2 \geq \tilde{t}_{1-\alpha},$$

where $\tilde{t}_{1-\alpha}$ *is the* $(1-\alpha)$*-quantile of* $\int_{[0,1]^2} B_{\mathbf{f}}^2(t,s)\lambda^2(dt,ds)$. *If* σ^2 *is unknown, it can be replaced by* $\hat{\sigma}^2_{nm}$ *given by (3.4.1).*

4.5 Consistency and power of the tests

In this section we investigate the *consistency* and the *asymptotic power* under alternatives of the sequence of asymptotically size α tests $(\delta_n)_{n\geq 1}$, $(\lambda_n)_{n\geq 1}$ and $(\psi_n)_{n\geq 1}$ defined in Section 4.2, Section 4.3 and Section 4.4, respectively.

Let us consider a *localized version* of the linear model (4.1.3) with *localizing rate* $\frac{1}{n}$, denoted by $\mathbf{Y}^{loc}_{n\times n} = \mathbf{G}^{loc}_{n\times n} + \mathbf{E}_{n\times n}$, where $\mathbf{G}^{loc}_{n\times n} := (g(\ell/n, k/n)/n)_{k=1,\ell=1}^{n,\ n}$, see also Bischoff and Miller (2000), and Bader (2001) for the notion of a localized linear model on a closed interval. We consider now the problem of testing the hypotheses $H_0 : \mathbf{G}^{loc}_{n\times n} \in \mathbf{W}_n$ versus $K : \mathbf{G}^{loc}_{n\times n} \notin \mathbf{W}_n$ observing the residual partial sums process $\mathbf{T}_n(pr_{\mathbf{W}_n^{\perp}}\mathbf{Y}^{loc}_{n\times n})(\cdot)$ and the hypothesis $H_0 : \mathbf{G}_{n\times n} \in \mathbf{W}_n$ versus $K : \mathbf{G}_{n\times n} \notin \mathbf{W}_n$ observing the residual partial sums process $\mathbf{T}_n(pr_{\mathbf{W}_n^{\perp}}\mathbf{Y}_{n\times n})(\cdot)$. Since under H_0 we have the equality $pr_{\mathbf{W}_n^{\perp}}\mathbf{Y}^{loc}_{n\times n} = pr_{\mathbf{W}_n^{\perp}}\mathbf{Y}_{n\times n}$, which in turn results the same residual partial sums limit process, then based on the statistics $K_{n,\mathbf{f}}$, $KS_{n,\mathbf{f}}$ and $C_{n,\mathbf{f}}$, both testing problems will produce the same asymptotically size α tests. We are interested in comparing the behavior of the tests under alternatives.

Since the partial sums operator $\mathbf{T}_n$ is linear on $\mathbb{R}^{n\times n}$, then by applying Donsker's theorem, for $g(\cdot) \in BVV_c([0,1]^2)$, i.e., the space of functions which have bounded variation on $[0,1]^2$ in the sense of Vitali and are right continuous on $[0,1)^2$, we get

$$\frac{1}{\sigma}\mathbf{T}_n(\mathbf{G}^{loc}_{n\times n} + \mathbf{E}_{n\times n})(\cdot) \xrightarrow{\mathcal{D}} \frac{1}{\sigma}h_g(\cdot) + B_2(\cdot),\ \text{in } \mathcal{C}([0,1]^2),\ n\to\infty, \tag{4.5.1}$$

where

$$h_g(z_1,z_2) := \int_{[0,z_1]\times[0,z_2]} g(t,s)\lambda^2(dt,ds),\ (z_1,z_2)\in[0,1]^2,$$

and $B_2(\cdot)$ is standard Brownian (2) motion. Clearly, $\frac{\partial^2 h_g(\cdot)}{\partial t \partial s} = g(\cdot)$, λ^2 a.s. on $[0,1]^2$, and $g(\cdot) \in BVV_c([0,1]^2) \subset L_2([0,1]^2)$. Hence $h_g(\cdot) \in \mathcal{H}_{\mathbf{B}}$.

Corollary 4.5.1. *Suppose that the regression functions $f_1(\cdot), \dots, f_p(\cdot)$ are continuous, have bounded variation in the sense of Hardy on $[0,1]^2$ and are linearly independent. Consider the localized linear model $\mathbf{Y}^{loc}_{n\times n} = \mathbf{G}^{loc}_{n\times n} + \mathbf{E}_{n\times n}$. If the unknown regression function $g(\cdot) \in BVV_c([0,1]^2)$, then for $n \to \infty$*

$$\frac{1}{\sigma}\mathbf{T}_n(pr_{\mathbf{W}_n^\perp}\mathbf{Y}^{loc}_{n\times n})(\cdot) \xrightarrow{\mathcal{D}} \frac{1}{\sigma}\,\varphi_g(\cdot) + B_{\mathbf{f}}(\cdot), \;\; in\ \mathcal{C}([0,1]^2), \tag{4.5.2}$$

where

$$\varphi_g(\cdot) := h_g(\cdot) - (pr^*_{\mathbf{W}_{\mathcal{H}_\mathbf{B}}} h_g)(\cdot),$$
$$B_{\mathbf{f}}(\cdot) = B_2(\cdot) - (pr^*_{\mathbf{W}_{\mathcal{H}_\mathbf{B}}} B_2)(\cdot).$$

Proof. Without loss of generality we assume $\sigma = 1$. By the linearity of the partial sum operator $\mathbf{T}_n$ and Lemma 3.2.1, we get

$$\begin{aligned}\mathbf{T}_n(pr_{\mathbf{W}_n^\perp}\mathbf{Y}^{loc}_{n\times n})(\cdot) &= \mathbf{T}_n(\mathbf{Y}^{loc}_{n\times n})(\cdot) - (pr_{\mathbf{W}_{n\mathcal{H}_\mathbf{B}}}\mathbf{T}_n(\mathbf{Y}^{loc}_{n\times n})(\cdot))(\cdot)\\ &= \mathbf{T}_n(\mathbf{G}^{loc}_{n\times n} + \mathbf{E}_{n\times n})(\cdot) - (pr_{\mathbf{W}_{n\mathcal{H}_\mathbf{B}}}\mathbf{T}_n(\mathbf{G}^{loc}_{n\times n} + \mathbf{E}_{n\times n})(\cdot))(\cdot).\end{aligned}$$

Hence, the assertion follows from (4.5.1) and Theorem 3.2.6. □

By Corollary 4.5.1, under the *localized* alternative $K : \mathbf{G}^{loc}_{n\times n} \notin \mathbf{W}_n$, the limiting distribution of $\mathbf{T}_n(pr_{\mathbf{W}_n^\perp}\mathbf{Y}^{loc}_{n\times n})(\cdot)$ is a *signal-plus-noise model* with deterministic *signal* $\varphi_g(\cdot)$ and the residual partial sums limit process $B_{\mathbf{f}}(\cdot)$ as *noise*. Such a limit process for the localized linear regression model on the closed interval $[a,b]$ was studied deeply in Bischoff and Miller (2000) and Bischoff and et al. (2003a). Hence, the hypotheses that correspond to (4.1.1) or (4.1.2) are

$$H_0 : \varphi_g(t,s) \le 0,\ (t,s) \in [0,1]^2 \text{ vs. } K : \exists (t,s) \in [0,1]^2,\ \varphi_g(t,s) > 0 \tag{4.5.3}$$

for a one-sided alternative, and

$$H_0 : \varphi_g(t,s) = 0,\ (t,s) \in [0,1]^2 \text{ vs. } K : \exists (t,s) \in [0,1]^2,\ \varphi_g(t,s) \neq 0 \tag{4.5.4}$$

for a two-sided alternative. Moreover, because of Corollary 4.5.1, the limiting distributions of the statistics $K_{n,\mathbf{f}}/\sigma$, $KS_{n,\mathbf{f}}/\sigma$ and $C_{n,\mathbf{f}}/\sigma^2$ under the localized alternative exist by the continuous mapping theorem.

4.5.1 Test of Kolmogorov type

We consider the hypotheses (4.5.3) and assume for the moment that σ^2 is known. The *power functions* $(\Psi_{\delta_n}(\cdot))_{n\geq 1} : BVV_c([0,1]^2) \longrightarrow (0,1)$ of the sequence of asymptotically size α tests $(\delta_n)_{n\geq 1}$, are given by

$$\Psi_{\delta_n}(g) :=\ \mathbb{P}\left\{\frac{1}{\sigma}K_{n,\mathbf{f}} \geq \tilde{c}_{1-\alpha}\right\},\ g(\cdot) \in BVV_c([0,1]^2),\ \text{with } \mathbf{G}_{n\times n} \notin \mathbf{W}_n,\ n \geq 1,$$

where $\tilde{c}_{1-\alpha}$ is the $(1-\alpha)$-quantile of $\sup_{0\leq t,s\leq 1} B_{\mathbf{f}}(t,s)$. Hence,

$$\begin{aligned}
\Psi_{\delta_n}(g) &= \mathbb{P}\left\{\sup_{0\leq t,s\leq 1}\left(\frac{1}{\sigma}\mathbf{T}_n\left(pr_{\mathbf{W}_n^{\perp}}(\mathbf{G}_{n\times n} + \mathbf{E}_{n\times n})\right)(t,s)\right) \geq \tilde{c}_{1-\alpha}\right\}\\
&= \mathbb{P}\left\{\sup_{0\leq t,s\leq 1}\left(\frac{1}{\sigma}\mathbf{T}_n\left(pr_{\mathbf{W}_n^{\perp}}\mathbf{G}^{loc}_{n\times n} + \frac{1}{n}pr_{\mathbf{W}_n^{\perp}}\mathbf{E}_{n\times n}\right)(t,s)\right) \geq \frac{\tilde{c}_{1-\alpha}}{n}\right\}\\
&\overset{n\to\infty}{\longrightarrow}\ \mathbb{P}\left\{\sup_{0\leq t,s\leq 1}\frac{1}{\sigma}\varphi_g(t,s) \geq 0\right\},
\end{aligned}$$

pointwise on $BVV_c([0,1]^2)$. Under the associated one-sided alternative $K : \exists (t,s) \in [0,1]^2,\ \varphi_g(t,s) > 0$, the last limiting probability is equal to 1. Hence, the test is asymptotically pointwise consistent in power. Under the localized alternative $K : \mathbf{G}^{loc}_{n\times n} \notin \mathbf{W}_n$, $g(\cdot) \in BVV_c([0,1]^2)$, Corollary 4.5.1 yields

$$\begin{aligned}
\Psi_{\delta_n}(g) &= \mathbb{P}\left\{\sup_{0\leq t,s\leq 1}\left(\frac{1}{\sigma}\mathbf{T}_n\left(pr_{\mathbf{W}_n^{\perp}}(\mathbf{G}^{loc}_{n\times n} + \mathbf{E}_{n\times n})\right)(t,s)\right) \geq \tilde{c}_{1-\alpha}\right\}\\
&\overset{n\to\infty}{\longrightarrow}\ \mathbb{P}\left\{\sup_{0\leq t,s\leq 1}\left(\frac{1}{\sigma}\varphi_g(t,s) + B_{\mathbf{f}}(t,s)\right) \geq \tilde{c}_{1-\alpha}\right\} =: \Psi_\delta(g),
\end{aligned} \tag{4.5.5}$$

pointwise on $BVV_c([0,1]^2)$. Thus the limiting power function of sequence of asymptotically size α test $(\delta_n)_{n\geq 1}$ based on the statistic $K_{n,\mathbf{f}}$ is given by the *boundary crossing* probability

$$\mathbb{P}\left\{\exists (t,s) \in [0,1]^2 : \frac{1}{\sigma}\varphi_g(t,s) + B_{\mathbf{f}}(t,s) \geq \tilde{c}_{1-\alpha}\right\}. \tag{4.5.6}$$

4.5.2 Test of Kolmogorov-Smirnow type

The sequence of power functions $(\Psi_{\lambda_n}(\cdot))_{n\geq 1} : BVV_c([0,1]^2) \longrightarrow (0,1)$ of the sequence of asymptotically size α tests $(\lambda_n)_{n\geq 1}$ based on the statistic $KS_{n,\mathbf{f}}$ is given by

$$\Psi_{\lambda_n}(g) := \ \mathbb{P}\left\{\frac{1}{\sigma}KS_{n,\mathbf{f}} \geq \tilde{q}_{1-\alpha}\right\},\ g(\cdot) \in BVV_c([0,1]^2),\ \text{with } \mathbf{G}_{n\times n} \notin \mathbf{W}_n,\ n \geq 1,$$

where $\tilde{q}_{1-\alpha}$ is the $(1-\alpha)$-quantile of $\sup_{0\leq t,s\leq 1} |B_{\mathbf{f}}(t,s)|$. Under the two-sided alternative $K : \exists (t,s) \in [0,1]^2,\ \varphi_g(t,s) \neq 0$, we get

$$\begin{aligned}\Psi_{\delta_n}(g) &= \mathbb{P}\left\{\sup_{0\leq t,s\leq 1}\left|\frac{1}{\sigma}\mathbf{T}_n\left(pr_{\mathbf{W}_n^\perp}(\mathbf{G}_{n\times n} + \mathbf{E}_{n\times n})\right)(t,s)\right| \geq \tilde{q}_{1-\alpha}\right\}\\ &\overset{n\to\infty}{\longrightarrow}\ \mathbb{P}\left\{\sup_{0\leq t,s\leq 1}\left|\frac{1}{\sigma}\varphi_g(t,s)\right| \geq 0\right\} = 1,\end{aligned}$$

pointwise on $BVV_c([0,1]^2)$. Hence, the test is asymptotically pointwise consistent in power.

By analogy with the test of Kolmogorov type, under the localized alternative $K : \mathbf{G}_{n\times n}^{loc} \notin \mathbf{W}_n$, we obtain

$$\begin{aligned}\Psi_{\lambda_n}(g) &= \mathbb{P}\left\{\sup_{0\leq t,s\leq 1}\left|\frac{1}{\sigma}\mathbf{T}_n\left(pr_{\mathbf{W}_n^\perp}(\mathbf{G}_{n\times n}^{loc} + \mathbf{E}_{n\times n})\right)(t,s)\right| \geq \tilde{q}_{1-\alpha}\right\}\\ &\overset{n\to\infty}{\longrightarrow}\ \mathbb{P}\left\{\sup_{0\leq t,s\leq 1}\left|\frac{1}{\sigma}\varphi_g(t,s) + B_{\mathbf{f}}(t,s)\right| \geq \tilde{q}_{1-\alpha}\right\} =: \Psi_\lambda(g),\end{aligned} \tag{4.5.7}$$

pointwise on $BVV_c([0,1]^2)$. Thus the limiting power function of $(\lambda_n)_{n\geq 1}$ is given by the boundary crossing probability

$$\mathbb{P}\left\{\exists (t,s) \in [0,1]^2 : \left|\frac{1}{\sigma}\varphi_g(t,s) + B_{\mathbf{f}}(t,s)\right| \geq \tilde{q}_{1-\alpha}\right\}. \tag{4.5.8}$$

4.5.3 Test of Cramér-von Mises type

The power functions $(\Psi_{\psi_n}(\cdot))_{n\geq 1} : BVV_c([0,1]^2) \longrightarrow (0,1)$ of the sequence of asymptotically size α tests $(\psi_n)_{n\geq 1}$ are given by

$$\Psi_{\psi_n}(g) := \ \mathbb{P}\left\{\frac{1}{\sigma^2}C_{n,\mathbf{f}} \geq \tilde{t}_{1-\alpha}\right\}, \ g(\cdot) \in BVV_c([0,1]^2), \text{ with } \mathbf{G}_{n\times n} \notin \mathbf{W}_n, \ n \geq 1,$$

where $\tilde{t}_{1-\alpha}$ is the $(1-\alpha)$-quantile of $\int_{[0,1]^2} B_{\mathbf{f}}^2(t,s)\lambda^2(dt,ds)$.

Under the two-sided alternative $K : \exists (t,s) \in [0,1]^2, \ \varphi_g(t,s) \neq 0$, we get

$$\begin{aligned}\Psi_{\psi_n}(g) &= \mathbb{P}\left\{\int_{[0,1]^2}\left(\frac{1}{\sigma}\mathbf{T}_n\left(pr_{\mathbf{W}_n^{\perp}}(\mathbf{G}_{n\times n} + \mathbf{E}_{n\times n})\right)(\frac{\ell}{n},\frac{k}{n})\right)^2 \lambda^2(dt,ds) \geq \tilde{t}_{1-\alpha}\right\} \\ &\overset{n\to\infty}{\longrightarrow} \mathbb{P}\left\{\int_{[0,1]^2}\left(\frac{1}{\sigma}\varphi_g(t,s)\right)^2 \lambda^2(dt,ds) \geq 0\right\} = 1,\end{aligned}$$

pointwise on $BVV_c([0,1]^2)$. Thus the test based on the Cramér-von Mises statistic is also asymptotically pointwise consistent. Furthermore, by Corollary 4.5.1, the power function under the localized alternative is

$$\begin{aligned}\Psi_{\psi_n}(g) = \ &\mathbb{P}\left\{\int_{[0,1]^2}\left(\frac{1}{\sigma}\mathbf{T}_n\left(pr_{\mathbf{W}_n^{\perp}}(\mathbf{G}_{n\times n}^{loc} + \mathbf{E}_{n\times n})\right)(\frac{\ell}{n},\frac{k}{n})\right)^2 \lambda^2(dt,ds) \geq \tilde{t}_{1-\alpha}\right\} \\ \overset{n\to\infty}{\longrightarrow} \ &\mathbb{P}\left\{\int_{[0,1]^2}\left(\frac{1}{\sigma}\varphi_g(t,s) + B_{\mathbf{f}}(t,s)\right)^2 \lambda^2(dt,ds) \geq \tilde{t}_{1-\alpha}\right\} =: \Psi_{\psi}(g),\end{aligned} \tag{4.5.9}$$

pointwise on $BVV_c([0,1]^2)$.

Remark 4.5.2. *In case σ^2 is unknown, by Slutsky's theorem, we can replace σ^2 by any consistent estimator, for instance by the estimator $\hat{\sigma}_n^2$ given by (1.2.2).*

4.5.4 Approximation of the localized power

In this subsection we present Monte Carlo simulations to approximate the localized power $\Psi_\delta(g)$, $\Psi_\lambda(g)$ and $\Psi_\psi(g)$ of the three tests, for $g(\cdot)$ varies in $BVV_c([0,1]^2)$,

$g(\cdot) \notin \mathbf{W}$. By (4.5.5), (4.5.7) and (4.5.9), for any $g(\cdot) \in BVV_c([0,1]^2)$ with $g(\cdot) \notin \mathbf{W}$ such that $0 \leq \varphi_g(t,s)$, $\forall (t,s) \in [0,1]^2$, the localized powers of the tests evaluated at $g(\cdot)$ are in the interval $[\alpha, 1)$.

Let us again consider the regression functions $f_1(t,s) = 1$, $f_2(t,s) = t$ and $f_3(t,s) = s$, for $(t,s) \in [0,1]^2$ studied in Example 3.3.2. Suppose that under H_0 we have a first order model, i.e., $\mathbf{W} = [f_1(\cdot), f_2(\cdot), f_3(\cdot)]$. We shall approximate the localized power of the tests at $\rho g(\cdot)$, for ρ varies in $(0,\infty)$ and a function $g(t,s) := 1 - 2t + s - 2t^2 \notin \mathbf{W}$. Since for $(t,s) \in [0,1]^2$, we have $h_g(t,s) = ts - t^2 s + t^2 s/2 - 2t^3 s/3 \in \mathcal{H}_\mathbf{B}$, then by a little computation we get

$$\begin{aligned}
\varphi_g(t,s) &= h_g(t,s) - (pr^*_{\mathbf{W}_{\mathcal{H}_\mathbf{B}}} h_g)(t,s) \\
&= h_g(t,s) - \sum_{i=1}^{3} \langle \tilde{f}_i, g \rangle_{L_2} \tilde{h}_{f_i}(t,s) \\
&= 17ts/3 - 7t^2 s/2 - 3ts^2/2 - 2t^3 s/3, \ (t,s) \in [0,1]^2.
\end{aligned}$$

Clearly, $\varphi_g(\cdot)$ satisfies the condition specified under the one-sided alternative $K : \varphi_g(t,s) > 0$ for some $(t,s) \in [0,1]^2$. It also fulfills the condition specified under the two-sided alternative $K : \varphi_g(t,s) \neq 0$ for some $(t,s) \in [0,1]^2$. We notice that by Corollary 4.5.1 $\varphi_g(\cdot)$ is approximated by $\mathbf{T}_n(pr_{\mathbf{W}_n^\perp} \mathbf{G}^{loc}_{n\times n})(\cdot)$ uniformly on $[0,1]^2$, where

$$\begin{aligned}
\mathbf{G}^{loc}_{n\times n} &= \left(\frac{1}{n} g(\ell/n, k/n) \right)_{k=1,\ \ell=1}^{n,\ n} \\
&= \left(\frac{1}{n} (1 - 2l/n + k/n - 2(l/n)^2) \right)_{k=1, \ell=1}^{n,\ n} \notin \mathbf{W}_n,
\end{aligned}$$

for $\mathbf{W}_n = [f_1(\mathcal{E}_n), f_2(\mathcal{E}_n), f_3(\mathcal{E}_n)]$.

We now present an algorithm for approximating the localized power of the Kolmogorov type test. Algorithms for approximating the localized power of the Kolmogorov-Smirnov and Cramér-von Mises type tests are completely similar.

`Begin Algorithm 2`

step 1: Fix $\rho \in (0, \infty)$ and $\tilde{n} \in \mathbb{N}$.

step 2: Generate M i.i.d. pseudo random matrices $\mathbf{E}^{(j)}_{\tilde{n}\times\tilde{n}} := (\varepsilon_{\ell kj})^{\tilde{n},\ \tilde{n}}_{k=1,\ell=1}$, with components $\varepsilon_{\ell kj}$ generated from i.i.d. $\mathcal{N}(0,1)$ random variables, $j = 1, \ldots, M$.

step 3: Generate M i.i.d. matrix of observations $\mathbf{Y}^{(j)}_{\tilde{n}\times\tilde{n}} := \rho\mathbf{G}^{loc}_{\tilde{n}\times\tilde{n}} + \mathbf{E}^{(j)}_{\tilde{n}\times\tilde{n}}$.

step 4: Calculate $\hat{\beta}^{(j)}_{\tilde{n}}$ by solving the equation $(\mathbf{X}^\top_{\tilde{n}}\mathbf{X}_{\tilde{n}})\beta = \mathbf{X}^\top_{\tilde{n}} vec(\mathbf{Y}^{(j)}_{\tilde{n}\times\tilde{n}})$.

step 5: Calculate the matrix of residuals $\mathbf{R}^{(j)}_{\tilde{n}\times\tilde{n}} := \mathbf{Y}^{(j)}_{\tilde{n}\times\tilde{n}} - \sum_{i=1}^{p} \hat{\beta}^{(j)}_{\tilde{n}i} f_i(\mathcal{E}_n)$.

step 6: Calculate the statistic $K^{(j)}_{\tilde{n},\mathbf{f}} := \max_{0\le k,\ell\le\tilde{n}} \mathbf{T}_{\tilde{n}}\left(\mathbf{R}^{(j)}_{\tilde{n}\times\tilde{n}}\right)(\ell/\tilde{n}, k/\tilde{n})$.

step 7: Calculate the power $\Psi_\delta(g) \approx \frac{1}{M}\sum_{j=1}^{M} \mathbf{1}\{K^{(j)}_{\tilde{n},\mathbf{f}} \ge \tilde{c}_{1-\alpha}\}$.

`End Algorithm 2`

Table 5-Table 7 present the approximated localized power of the three sequences of asymptotically size α tests computed according to Algorithm 2, executed by using the software package R 2.0.1, for $\rho = 0.5$, 1.0, 1.5, 2.0, 2.5, 3.0, and 5.0 and several values of α. For all cases we chose the sample size $\tilde{n}\times\tilde{n} = 30\times 30$, and the number of replications is $M = 10^6$. The graphs of the localized power of the tests versus ρ for $\alpha = 0.0100$, 0.0500, 0.1000, 0.1500, 0.2000 and 0.2500 are given in Figure 6. The simulation results show that for such a function $g(\cdot) \in BVV_c([0,1]^2)$ and ρ, all tests are *powerful* in the sense that the localized power of the tests are larger than α. Of the three type tests, the Kolmogorov type test seems to be the most powerful.

If we consider the simulated power of the Kolmogorov-Smirnov and Cramér-von Mises type tests for testing against a two-sided alternative, the second test is in general more powerful than the first one. It happens only for some values of ρ and α that the first test is more powerful than the second.

First-order model	Kolmogorov type test						
sample, rep.	$\tilde{n} \times \tilde{n} = 30 \times 30,\ M = 10^6,\ \varepsilon_{\ell kj} \stackrel{iid}{\sim} \mathcal{N}(0,1)$						
α	0.3500	0.2500	0.2000	0.1500	0.1000	0.0500	0.0100
$\tilde{c}_{1-\alpha}$	0.7917	0.8470	0.8805	0.9211	0.9739	1.0570	1.2264
$\Psi_{\delta}^{loc}(g/2)$	0.3553	0.2507	0.2058	0.1548	0.1044	0.0529	0.0110
$\Psi_{\delta}^{loc}(2g/2)$	0.3620	0.2568	0.2117	0.1599	0.1086	0.0555	0.0118
$\Psi_{\delta}^{loc}(3g/2)$	0.3695	0.2639	0.2181	0.1657	0.1133	0.0586	0.0127
$\Psi_{\delta}^{loc}(4g/2)$	0.3775	0.2714	0.2252	0.1717	0.1183	0.0617	0.0138
$\Psi_{\delta}^{loc}(5g/2)$	0.3863	0.2798	0.2330	0.1785	0.1239	0.0653	0.0149
$\Psi_{\delta}^{loc}(6g/2)$	0.3955	0.2885	0.2410	0.1856	0.1298	0.0691	0.0162
$\Psi_{\delta}^{loc}(10g/2)$	0.4392	0.3293	0.2794	0.2197	0.1583	0.0881	0.0224

Table 5. The approximated localized power of the Kolmogorov type test.

First-order model	Kolmogorov-Smirnov type test						
sample, rep.	$\tilde{n} \times \tilde{n} = 30 \times 30,\ M = 10^6,\ \varepsilon_{\ell kj} \stackrel{iid}{\sim} \mathcal{N}(0,1)$						
α	0.3500	0.2500	0.2000	0.1500	0.1000	0.0500	0.0100
$\tilde{q}_{1-\alpha}$	0.8845	0.9356	0.9666	1.0040	1.0536	1.1306	1.2916
$\Psi_{\delta}^{loc}(g/2)$	0.3500	0.2501	0.1999	0.1503	0.0997	0.0501	0.0099
$\Psi_{\lambda}^{loc}(2g/2)$	0.3507	0.2510	0.2010	0.1514	0.1006	0.0504	0.0102
$\Psi_{\delta}^{loc}(3g/2)$	0.3529	0.2535	0.2031	0.1530	0.1019	0.0514	0.0102
$\Psi_{\delta}^{loc}(4g/2)$	0.3563	0.2564	0.2055	0.1552	0.1035	0.0525	0.0106
$\Psi_{\delta}^{loc}(5g/2)$	0.3604	0.2598	0.2089	0.1582	0.1059	0.0539	0.0110
$\Psi_{\delta}^{loc}(6g/2)$	0.3654	0.2642	0.2129	0.1621	0.1086	0.0556	0.0115
$\Psi_{\delta}^{loc}(10g/2)$	0.3940	0.2903	0.2370	0.1828	0.1254	0.0664	0.0149

Table 6. The approximated localized power of the Kolmogorov-Smirnow type test.

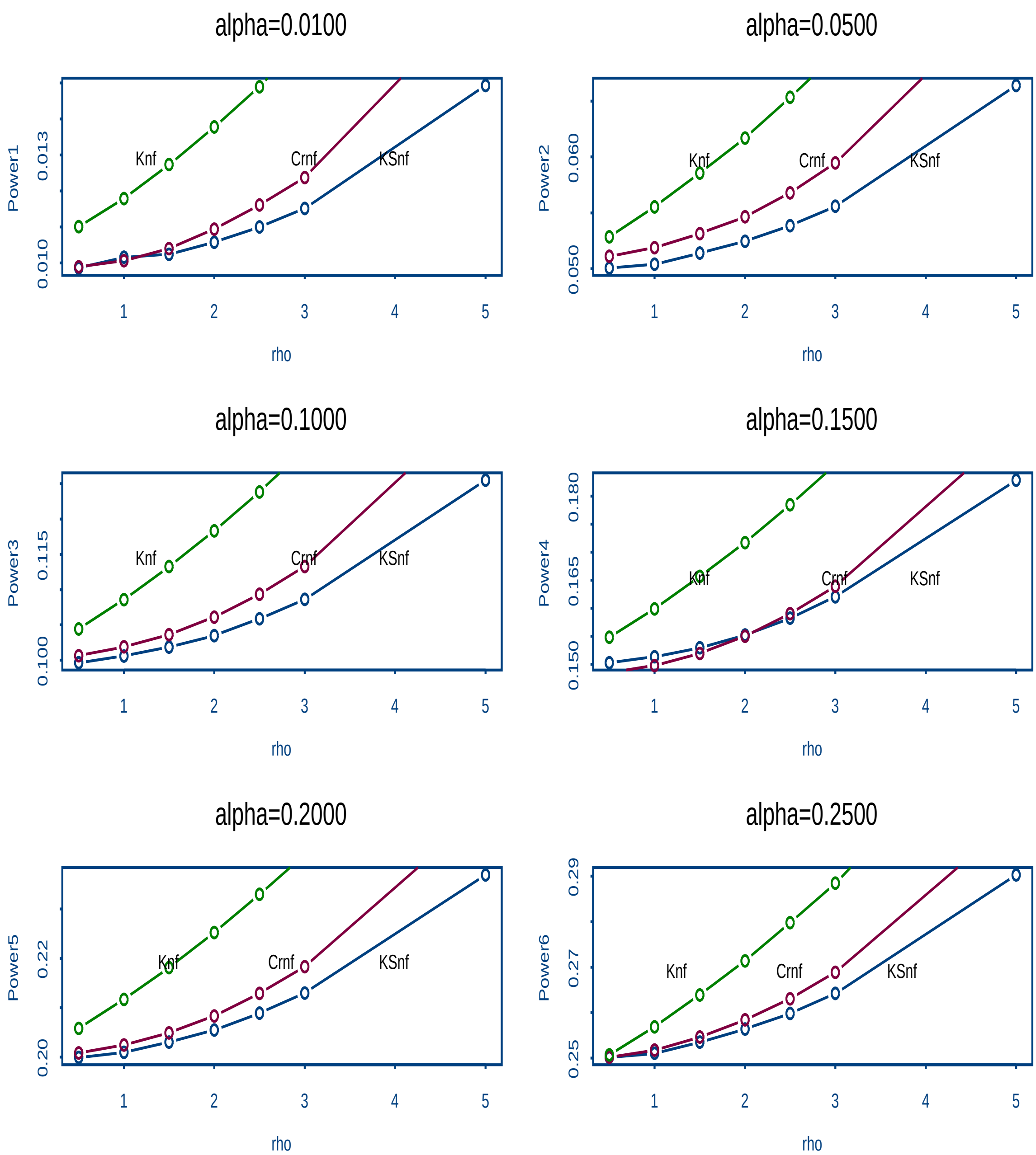

Figure 6. Graphs of $\Psi_\delta(\rho g)$, $\Psi_\lambda(\rho g)$ and $\Psi_\psi(\rho g)$. K_{nf}, Cr_{nf} and KS_{nf} denote the power of the Kolmogorov, Kolmogorov-Smirnov and Cramér-von Mises type tests.

First-order model	Cramér-von Mises type test						
sample, rep.	$\tilde{n} \times \tilde{n} = 30 \times 30,\ M = 10^6,\ \varepsilon_{\ell kj} \overset{iid}{\sim} \mathcal{N}(0,1)$						
α	0.3500	0.2500	0.2000	0.1500	0.1000	0.0500	0.0100
$\tilde{t}_{1-\alpha}$	0.0794	0.0890	0.0951	0.1027	0.1131	0.1305	0.1695
$\Psi_{\delta}^{loc}(g/2)$	0.3548	0.2502	0.2008	0.1485	0.1007	0.0511	0.0099
$\Psi_{\psi}^{loc}(2g/2)$	0.3565	0.2517	0.2024	0.1498	0.1019	0.0519	0.0101
$\Psi_{\delta}^{loc}(3g/2)$	0.3594	0.2546	0.2049	0.1520	0.1036	0.0532	0.0104
$\Psi_{\delta}^{loc}(4g/2)$	0.3635	0.2584	0.2083	0.1550	0.1061	0.0546	0.0109
$\Psi_{\delta}^{loc}(5g/2)$	0.3687	0.2631	0.2129	0.1590	0.1093	0.0568	0.0116
$\Psi_{\delta}^{loc}(6g/2)$	0.3749	0.2689	0.2183	0.1639	0.1133	0.0595	0.0124
$\Psi_{\delta}^{loc}(10g/2)$	0.4108	0.3029	0.2503	0.1923	0.1370	0.0752	0.0176

Table 7. The approximated localized power of the Cramér-von Mises type test.

4.6 Weighted tests

In this section we make a generalization to the foregoing tests by introducing a weight function $w(\cdot) : [0,1]^2 \to [0,\infty)$, $w(\cdot) \in \mathcal{C}([0,1]^2)$. For a fixed $n \in \mathbb{N}$ and $w(\cdot) \in \mathcal{C}([0,1]^2)$, let $\Gamma_n := \{\mathbf{I}_{\ell k} := [(\ell-1)/n, \ell/n] \times [(k-1)/n, k/n] :\ 1 \le \ell, k \le n\}$ and

$$
\begin{aligned}
w_n(t,s) := {} & (\ell - nt)(k - ns)w((\ell-1)/n, (k-1)/n) \\
& + (nt - (\ell-1))(k - ns)w(\ell/n, (k-1)/n) \\
& + (\ell - nt)(ns - (k-1))w((\ell-1)/n, k/n) \\
& + (nt - (\ell-1))(ns - (k-1))w(\ell/n, k/n),\ \forall (t,s) \in \mathbf{I}_{\ell k}.
\end{aligned}
$$

The sequence $(w_n(\cdot))_{n \ge 1}$ has the following characteristics:

- For $n \to \infty$ $w_n(\cdot) \overset{\|\cdot\|_\infty}{\longrightarrow} w(\cdot)$ in $\mathcal{C}([0,1]^2)$.

- For any $n \in \mathbb{N}$, $\max_{0\le k,\ell\le n} w(\ell/n, k/n) = \sup_{0\le t,s\le 1} w_n(t,s)$.
- For any $n \in \mathbb{N}$,

$$\max_{0\le k,\ell\le n} w(\ell/n, k/n)\mathbf{T}_n(\mathbf{R}_{n\times n})(\ell/n, k/n) = \sup_{0\le t,s\le 1} w_n(t,s)\mathbf{T}_n(\mathbf{R}_{n\times n})(t,s).$$

Based on these characteristics of $w_n(\cdot)$ and the weak convergence of the sequence of the residual partial sums processes, the following result can be directly verified.

Proposition 4.6.1. *Suppose σ^2 is known. For a fixed $\alpha \in (0,1)$, an asymptotically size α weighted Kolmogorov type test for testing (4.1.2) or (4.1.4), i.e., a test based on the statistic $K_{w_n,\mathbf{f}} := \max_{0\le k,\ell\le n} \frac{1}{n} w(\ell/n, k/n) \sum_{j=0}^{k}\sum_{i=0}^{\ell} r_{ij}$ is*

$$\textit{reject } H_0, \textit{ if and only if } K_{w_n,\mathbf{f}}/\sigma \ge \tilde{c}_{w,1-\alpha},$$

where $\tilde{c}_{w,1-\alpha}$ is the $(1-\alpha)$-quantile of $\sup_{(t,s)\in[0,1]^2} w(t,s)B_{\mathbf{f}}(t,s)$. The test is consistent. Moreover, the limiting localized power function of this test is given by the boundary crossing probability

$$\Psi_{w,\delta}(g) := \mathbb{P}\left\{\sup_{0\le t,s\le 1}\left(w(t,s)\left(\frac{1}{\sigma}\varphi_g(t,s) + B_{\mathbf{f}}(t,s)\right)\right) \ge \tilde{c}_{w,1-\alpha}\right\},$$

for $g(\cdot) \in BVV_c([0,1]^2)$ such that $\mathbf{G}^{loc}_{n\times n} \notin \mathbf{W}_n$. If σ^2 is unknown, then the test is given by

$$\textit{reject } H_0, \textit{ if and only if } K_{w_n,\mathbf{f}}/\hat{\sigma}^2_n \ge \tilde{c}_{w,1-\alpha}.$$

Here and throughout $\hat{\sigma}^2_n$ is a consistent estimator of σ^2.

As an additional information, a nice investigation to the asymptotic of such boundary crossing probability for $B_{\mathbf{f}}(\cdot)$ is a Brownian bridge with parameter space $[0,1]$ has been observed in Bischoff and et al. (2003b) in which they gave typical assumptions for the weight function and the trend.

Proposition 4.6.2. *Suppose σ^2 is known. For a fixed $\alpha \in (0,1)$, an asymptotically size α weighted Kolmogorov-Smirnov type test for testing (4.1.2) or (4.1.4), i.e., a test based on the statistic $KS_{w_n,\mathbf{f}} := \max_{0\le k,\ell\le n}\left|\frac{1}{n}w(\ell/n,k/n)\sum_{j=0}^{k}\sum_{i=0}^{\ell} r_{ij}\right|$, is given by*

$$\text{reject } H_0, \text{ if and only if } KS_{w_n,\mathbf{f}}/\sigma \ge \tilde{q}_{w,1-\alpha},$$

where $\tilde{q}_{w,1-\alpha}$ is the $(1-\alpha)$-quantile of $\sup_{(t,s)\in[0,1]^2}|w(t,s)B_{\mathbf{f}}(t,s)|$. The test is consistent. Moreover, the following boundary crossing probability give the limiting localized power function of this test:

$$\Psi_{w,\lambda}(g) := \mathbb{P}\left\{\sup_{0\le t,s\le 1}\left|w(t,s)\left(\frac{1}{\sigma}\varphi_g(t,s)+B_{\mathbf{f}}(t,s)\right)\right| \ge \tilde{q}_{w,1-\alpha}\right\},$$

for $g(\cdot) \in BVV_c([0,1]^2)$ such that $\mathbf{G}^{loc}_{n\times n} \notin \mathbf{W}_n$. In case σ^2 is unknown, the test is given by

$$\text{reject } H_0, \text{ if and only if } KS_{w_n,\mathbf{f}}/\hat{\sigma}_n^2 \ge \tilde{q}_{w,1-\alpha}.$$

The weighted Cramér-von Mises type test is based on the statistic

$$\begin{aligned} C_{w_n,\mathbf{f}} &:= \frac{1}{n^4}\sum_{k=0}^{n}\sum_{\ell=0}^{n}\left(w(\ell/n,k/n)\sum_{j=0}^{k}\sum_{i=0}^{\ell} r_{ij}\right)^2 \\ &= \int_{[0,1]^2}\left(w_n(\ell/n,k/n)\mathbf{T}_n(\mathbf{R}_{n\times n})(\ell/n,k/n)\right)^2\lambda^2(dt,ds). \end{aligned}$$

Proposition 4.6.3. *Suppose σ^2 is known and $w(\cdot)$ is right continuous on $[0,1)^2$. For a fixed $\alpha \in (0,1)$, the asymptotically size α weighted Cramér-von Mises type test for testing (4.1.2) or (4.1.4) is*

$$\text{reject } H_0, \text{ if and only if } C_{w_n,\mathbf{f}}/\sigma^2 \ge \tilde{t}_{w,1-\alpha},$$

where $\tilde{t}_{w,1-\alpha}$ is the $(1-\alpha)$-quantile of $\int_{[0,1]^2}(w(t,s)B_{\mathbf{f}}(t,s))^2\lambda^2(dt,ds)$. If σ^2 is unknown, then the test is given by

$$\text{reject } H_0, \text{ if and only if } C_{w_n,\mathbf{f}}/\hat{\sigma}_n^2 \ge \tilde{t}_{w,1-\alpha}.$$

The test is consistent. The limiting localized power function of this test is given by the following boundary crossing probability:

$$\Psi_{w,\psi}(g) := \mathbb{P}\left\{\int_{[0,1]^2} \left(w(t,s)\left(1/\sigma\varphi_g(t,s) + B_{\mathbf{f}}(t,s)\right)\right)^2 \lambda^2(dt,ds) \geq \tilde{t}_{w,1-\alpha}\right\},$$

for $g(\cdot) \in BVV_c([0,1]^2)$ *such that* $\mathbf{G}^{loc}_{n\times n} \notin \mathbf{W}_n$.

Clearly, the tests developed in Section 4.2 - Section 4.5 are only special cases of weighted tests for the hypothesis introduced in Section 4.1, for which we put the weight $w(t,s) = 1$ for every $(t,s) \in [0,1]^2$. The Monte Carlo simulations for approximating the quantiles and the power of the weighted tests can be carried out analogously as for the unweighted tests in the foregoing sections. Therefore we omit such simulations for the weighted tests.

4.7 Applications

Our aim in this subsection is to present an example of the application of the foregoing asymptotic test theory in spatial data analysis. We consider the wheat-yield data (Mercer and Hall's data) presented and discussed in Cressie (1993), p. 454-455, and Xie and MacNeill (2004). The data are yields of grains (in pounds) observed over a 25×20 lattice of plots with 20 rows running east to west and 25 columns of plots running north to south. The experiment consists of giving the 500 plots the same treatment (presumably fertilizer, water, etc.), from which we identify the data as a realization of 500 independent random variables. The exact size of the plots from the original data seems to be unknown, but as was informed in Cressie (1993), p. 454-455, we assume that the plots are equally spaced, with the dimension of each plot being 10.82 ft by 8.05 ft.

Figure 7 presents the perspective plot of the data. Other quantities of the data such as the median, mean, standard deviation, first and third quartiles are summarized in Table 8. Visually (one can generate with S PLUS or R), the shape of the histogram of the data is identified as the familiar bell-shaped curve, indicating the nearly normal distribution of the the 500 wheat yield measurements.

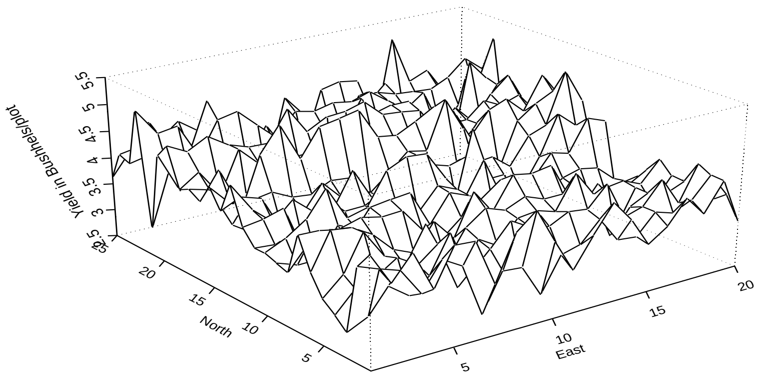

Figure 7. The Perspective plot of Mercer and Hall's data.

1st Qu.	Med.	Mean	Mode	3rd Qu.	St.Dev.	Skew.	Kurt.
3.630	3.970	3.940	3.944	4.270	0.455	0.036	-0.254

Table 8. Summary of the Mercer and Hall's data.

Observing Figure 7 we postulate under H_0 a full first-order model, i.e., by using the test statistics $K_{nm,\mathbf{f}}/\hat{\sigma}_{nm}$, $KS_{nm,\mathbf{f}}/\hat{\sigma}_{nm}$ and $C_{nm,\mathbf{f}}/\hat{\sigma}^2_{nm}$, we are interested in testing the hypotheses

$$H_0 : g(\cdot) \in [f_1(\cdot), f_2(\cdot), f_3(\cdot)] \text{ vs. } K : g(\cdot) \notin [f_1(\cdot), f_2(\cdot), f_3(\cdot)],$$

where $f_1(t,s) = 1,\ f_2(t,s) = t$, and $f_3(t,s) = s,\ (t,s) \in [0,1]^2$. Actually we perform weighted tests defined in Section 4.6, with weight function $w(t,s) = 1$, for $(t,s) \in [0,1]^2$.

Since the model variance σ^2 is unknown, we use a consistent estimator $\widehat{\sigma}^2_{nm}$. Calculated under H_0, the data give $\widehat{\sigma}^2_{nm} = 0.1898$. The following is an algorithm for calculating the approximated p-value of the Kolmogorov type test for observations taken from $\mathcal{E}_{nm}$. Algorithm for approximating the p-values of the Kolmogorov-Smirnov and Cramér-von Mises type tests are completely similar.

`Begin Algorithm` 3

step 1: Generate M i.i.d. pseudo random matrices $\mathbf{E}^{(j)}_{m\times n} = (\varepsilon_{\ell kj})^{m,\ n}_{k=1,\ell=1}$, with components $\varepsilon_{\ell kj}$ generated from i.i.d $\mathcal{N}(0,1)$ random variables, $j = 1, \ldots, M$.

step 2: Generate M i.i.d. matrix of observations $\mathbf{Y}^{(j)}_{m\times n}$.

step 3: Calculate $\widehat{\beta}^{(j)}$ by solving the equation

$$\left(\mathbf{X}^\top_{mn\times 3}\mathbf{X}_{mn\times 3}\right)\beta = \mathbf{X}^\top_{mn\times 3} vec(\mathbf{Y}^{(j)}_{m\times n}),$$

where $\mathbf{X}_{mn\times 3} = (vec(f_1(\mathcal{E}_{n\times m})), vec(f_2(\mathcal{E}_{n\times m})), vec(f_3(\mathcal{E}_{n\times m}))) \in \mathbb{R}^{mn\times 3}$.

step 4: Calculate the matrix of residuals $\mathbf{R}^{(j)}_{m\times n} := \mathbf{Y}^{(j)}_{m\times n} - \mathbf{X}_{mn\times 3}\widehat{\beta}^{(j)}$.

step 5: Calculate the Kolmogorov statistic

$$K^{(j)} := \max_{0\le \ell\le n;\ 0\le k\le m} \mathbf{T}_{nm}(\mathbf{R}^{(j)}_{m\times n})(\ell/n, k/m).$$

step 6: Based on the data, calculate the critical value $\widehat{K} := K_{n,m,\mathbf{f}}/\sqrt{\widehat{\sigma}^2_{nm}}$.

step 7: Calculate the approximated p-value $\widehat{p}_K := \frac{1}{M}\sum_{j=1}^{M} \mathbf{1}\{K^{(j)} \ge \widehat{K}\}$.

`End Algorithm` 3

The critical values and the corresponding approximated p-values of the tests for such data are presented in Table 9. The simulations were conducted by using the software package R 2.0.1 according to Algorithm 3, in which we chose sample size $n \times m = 25 \times 20$ and number of replications is $M = 10^6$. It can be seen that, based on the statistic $K_{nm,\mathbf{f}}/\hat{\sigma}_{nm}$, H_0 is rejected for $\alpha \geq 0.0002$, see also Table 9. Based on the statistics $KS_{nm,\mathbf{f}}$, H_0 is rejected for $\alpha \geq 0.0001$, whereas based on $C_{nm,\mathbf{f}}$, H_0 is rejected for $\alpha \geq 0.0010$. Thus, H_0 is rejected for all values of α used in applications.

statistics	**critical values**	**p-values**
$K_{nm,\mathbf{f}}/\hat{\sigma}_{nm}$	1.5919	0.0002
$KS_{nm,\mathbf{f}}/\hat{\sigma}_{nm}$	1.5919	0.0001
$C_{nm,\mathbf{f}}/\hat{\sigma}^2_{nm}$	0.2283	0.0010

Table 9: The critical and p-values of $K_{nm,\mathbf{f}}$, $KS_{nm,\mathbf{f}}$, and $C_{nm,\mathbf{f}}$.

Chapter 5

Lower and upper bounds for the power of the Kolmogorov type test

In this chapter we investigate the localized power of the Kolmogorov type test developed in Chapter 4. More exactly, we derive bounds for the boundary crossing probability $\mathbb{P}\{\exists(t,s)\in[0,1]^2 : \rho\varphi(t,s)+B_{\mathbf{f}}(t,s)\geq u(t,s)\}$, for $\rho>0$, a known trend $\varphi(\cdot):[0,1]^2\to\mathbb{R}$, and a general known *boundary* $u(\cdot):[0,1]^2\to\mathbb{R}$. We shall consider two cases, i.e., the case in which $B_{\mathbf{f}}(\cdot)$ is the standard Brownian (2) motion and $B_{\mathbf{f}}(\cdot)$ is the standard Brownian (2) bridge. These will be studied in Section 5.2 and Section 5.3, respectively. In Section 5.1 we study a general method for deriving the *kernel* of the residual partial sums limit processes, following an approach due to Lifshits (1996), p. 88-107. Furthermore, we denote the measure $\mathbb{P}\circ B_{\mathbf{f}}^{-1}$ defined on the infinite dimensional measurable space $(\mathcal{C}([0,1]^2),\mathcal{B}_{\mathcal{C}})$ by $\mathcal{P}_{\mathbf{f}}$.

5.1 Modelling the covariance function of $B_{\mathbf{f}}$

We say that the family $\{m_{(t,s)} : (t,s)\in[0,1]^2\}\subseteq L_2([0,1]^2)$ of functions on $[0,1]^2$ is a *model* of the covariance function $K_{\mathbf{f}}(\cdot,\cdot)$ of the residual partial sums limit process

$B_{\mathbf{f}}$, if for each $(t,s),(t',s') \in [0,1]^2$, we have

$$\begin{aligned} Cov(B_{\mathbf{f}}(t,s), B_{\mathbf{f}}(t',s')) &:= K_{\mathbf{f}}((t,s),(t',s')) = \langle m_{(t,s)}, m_{(t',s')} \rangle_{L_2} \\ &:= \int_{[0,1]^2} m_{(t,s)}(x,y) m_{(t',s')}(x,y) \lambda^2(dx,dy), \end{aligned} \tag{5.1.1}$$

see Lifshits (1996), p. 41-51.

5.1.1 Model for the standard Brownian (2) motion

By Definition 2.1.1, the covariance function of the standard Brownian (2) motion is

$$K_{\mathbf{f}}((t,s),(t',s')) = \min\{t,t'\} \min\{s,s'\},\ (t,s),(t',s') \in [0,1]^2.$$

Let us define the family of functions

$$\{m^{(2)}_{(t,s)} := \mathbf{1}_{[0,t]\times[0,s]} :\ (t,s) \in [0,1]^2\} \subseteq L_2([0,1]^2). \tag{5.1.2}$$

For any $(t,s),(t',s') \in [0,1]^2$, we then have

$$\int_{[0,1]^2} \mathbf{1}_{[0,t]\times[0,s]}(x,y) \mathbf{1}_{[0,t']\times[0,s']}(x,y) \lambda^2(dx,dy) = \min\{t,t'\} \min\{s,s'\}.$$

Thus the family of indicator functions $\{\mathbf{1}_{[0,t]\times[0,s]} : (t,s) \in [0,1]^2\}$ may be taken as a model for the covariance function of the standard Brownian (2) motion B_2.

5.1.2 Model for the standard Brownian (2) bridge

The standard Brownian (2) bridge is the residual partial sums limit process associated with a constant model having covariance function

$$K_{\mathbf{f}}((t,s),(t',s')) = \min\{t,t'\} \min\{s,s'\} - tst's',\ (t,s),(t',s') \in [0,1]^2.$$

Define the family of functions

$$\{m^{\circ}_{(t,s)} := \mathbf{1}_{[0,t]\times[0,s]} - ts\mathbf{1}_{[0,1]^2} : (t,s) \in [0,1]^2\} \subseteq L_2([0,1]^2). \tag{5.1.3}$$

For any $(t,s),(t',s') \in [0,1]^2$, we then get

$$\int_{[0,1]^2} (\mathbf{1}_{[0,t]\times[0,s]}(x,y) - ts\mathbf{1}_{[0,1]^2}(x,y))(\mathbf{1}_{[0,t']\times[0,s']}(x,y) - t's'\mathbf{1}_{[0,1]^2}(x,y))\lambda^2(dx,dy)$$
$$= \min\{t,t'\}\min\{s,s'\} - tst's'.$$

Hence by definition, the family $\{m^{\circ}_{(t,s)} : (t,s) \in [0,1]^2\}$ may be taken as a model for the covariance function of the standard Brownian (2) bridge B_2^0.

Definition 5.1.1. *(Lifshits (1996), p. 87)*

(a) *A function $h(\cdot) \in \mathcal{C}([0,1]^2)$ is called an admissible shift for the Gaussian measure $\mathcal{P}_{\mathbf{f}}$, if the measure $\mathcal{P}_{\mathbf{f}}^h$ defined on $(\mathcal{C}([0,1]^2), \mathcal{B}_{\mathcal{C}})$, given by $\mathcal{P}_{\mathbf{f}}^h\{A\} := \mathcal{P}_{\mathbf{f}}\{A-h\}$, for every $A \in \mathcal{B}_{\mathcal{C}}$, is absolutely continuous with respect to $\mathcal{P}_{\mathbf{f}}$. Here the set $\{A-h\}$ is defined as $\{x-h : x \in A\}$.*

(b) *A function $h(\cdot) \in \mathcal{C}([0,1]^2)$ is said to assign an admissible direction for $\mathcal{P}_{\mathbf{f}}$, if each vector of the family $\{c\,h(\cdot) : c \in \mathbb{R}\}$ is an admissible shift for $\mathcal{P}_{\mathbf{f}}$.*

5.2 Lower and upper bounds for the boundary crossing probabilities of B_2 with trend

In this section we derive a lower bound for the probability $\mathbb{P}\{\exists(t,s) \in [0,1]^2 : \rho\varphi(t,s) + B_2(t,s) \geq u(t,s)\}$, $\rho > 0$, where B_2 is standard Brownian (2) motion. Let us consider the model $\{m^{(2)}_{(t,s)}(\cdot) : (t,s) \in [0,1]^2\}$ given by (5.1.2) and the reproducing kernel Hilbert space (RKHS) $\mathcal{H}_{\mathbf{B}}$ of $B_2(\cdot)$ defined in Section 3.1. If $h(\cdot) \in \mathcal{H}_{\mathbf{B}}$ and $\ell(\cdot) \in L_2([0,1]^2)$ such that $h(t,s) = \left\langle \ell, m^{(2)}_{(t,s)} \right\rangle_{L_2}$, then $\ell(\cdot) = \frac{\partial^2 h(\cdot)}{\partial t \partial s}$, where $\frac{\partial^2 h(\cdot)}{\partial t \partial s}$ is the almost everywhere existing second derivative of $h(\cdot)$ with respect both variables on $[0,1]^2$.

Proposition 5.2.1. *Let $\mathcal{P}_{B_2}$ be the distribution of the standard Brownian (2) motion on $(\mathcal{C}([0,1]^2), \mathcal{B}_{\mathcal{C}})$. For any $h(\cdot) \in \mathcal{H}_{\mathbf{B}_{BV([0,1]^2)}}$, the density of the shifted measure $\mathcal{P}^h_{B_2}$ with respect to $\mathcal{P}_{B_2}$ is*

$$\frac{d\mathcal{P}^h_{B_2}}{d\mathcal{P}_{B_2}}(x) = \exp\left\{ \int_{[0,1]^2}^{R} \frac{\partial^2 h(t,s)}{\partial t \partial s}\, dx(t,s) - \frac{1}{2}\|h\|^2_{\mathcal{H}_{\mathbf{B}}} \right\}, \tag{5.2.1}$$

where $\mathcal{H}_{\mathbf{B}_{BV([0,1]^2)}} = \{h(\cdot) \in \mathcal{H}_{\mathbf{B}} : \frac{\partial^2 h(\cdot)}{\partial t \partial s} \in BV([0,1]^2)\}$, and $\|\cdot\|_{\mathcal{H}_{\mathbf{B}}}$ is the norm defined on $\mathcal{H}_{\mathbf{B}}$. See Appendix A for the definition of $BV([0,1]^2)$.

Proof. We refer the reader to Theorem 3 in Lifshits (1996), p. 88. We notice that (5.2.1) is a special case of (6) in Lifshits (1996), p. 88. A one-dimensional version of (5.2.1) is also presented there (see, Formula (13) in Lifshits (1996), p. 107). □

Remark 5.2.2. *Equation (5.2.1) is frequently called the Cameron-Martin-Girsanov formula for the standard Brownian (2) motion, see also Lifshits (1996), p. 107, and Bischoff and et al. (2005).*

Theorem 5.2.3. *Suppose that the boundary $u(\cdot)$ is continuous on $[0,1]^2$ and the trend $\varphi(\cdot) \in \mathcal{H}_{\mathbf{B}_{BV([0,1]^2)}}$ has a second derivative $\varphi''(\cdot) := \frac{\partial^2 \varphi(\cdot)}{\partial t \partial s} \in BV([0,1]^2)$ which is type I non decreasing on $[0,1]^2$. If the marginal functions $\varphi''(t,1) := \frac{\partial^2 \varphi(t,s)}{\partial t \partial s}|_{s=1}$, and $\varphi''(1,s) := \frac{\partial^2 \varphi(t,s)}{\partial t \partial s}|_{t=1}$ are non increasing on $[0,1]$, then*

$$\begin{aligned}&\mathbb{P}\left\{\forall (t,s) \in [0,1]^2 :\ \rho\varphi(t,s) + B_2(t,s) < u(t,s)\right\}\\ &\quad \le k^* \mathbb{P}\left\{\forall (t,s) \in [0,1]^2 :\ B_2(t,s) < u(t,s)\right\},\end{aligned} \tag{5.2.2}$$

where

$$\begin{aligned}k^* := \exp\Bigg\{&\rho\varphi''(1,1)u(1,1) + \rho \int_{[0,1]}^{R} u(t,1)d(-\varphi''(t,1)) + \rho \int_{[0,1]}^{R} u(1,s)d(-\varphi''(1,s))\\ &+\rho \int_{[0,1]^2}^{R} u(t,s)d\varphi''(t,s) - \frac{1}{2}\rho^2 \|\varphi\|^2_{\mathcal{H}_{\mathbf{B}}}\Bigg\},\ \rho > 0.\end{aligned}$$

Proof. By transformation of variable and the Cameron-Martin-Girsanov formula, we obtain

$$\begin{aligned}
&\mathbb{P}\left\{\omega \in \Omega : \ \rho\varphi(t,s) + B_2(\omega)(t,s) < u(t,s), \ \forall \ (t,s) \in [0,1]^2\right\} \\
&= \int_\Omega \mathbf{1}\{\omega \in \Omega : \ \rho\varphi(t,s) + B_2(\omega)(t,s) < u(t,s), \ \forall \ (t,s) \in [0,1]^2\}\mathbb{P}(d\omega) \\
&= \int_{\mathcal{C}([0,1]^2)} \mathbf{1}\{y \in \mathcal{C}([0,1]^2) : \ y(t,s) < u(t,s), \ \forall \ (t,s) \in [0,1]^2\}\mathcal{P}_{B_2}^{\rho\varphi}(dy) \\
&= \int_{\mathcal{C}([0,1]^2)} \mathbf{1}\{y \in \mathcal{C}([0,1]^2) : \ y(t,s) < u(t,s), \ \forall \ (t,s) \in [0,1]^2\} \\
&\qquad \times \exp\left\{\int_{[0,1]^2}^{R} \rho\varphi''(t,s) \ dy(t,s) - \frac{1}{2}\left\|\rho\varphi\right\|_{\mathcal{H}_\mathbf{B}}^2\right\}\mathcal{P}_{B_2}(dy) \\
&= \int_\Omega \mathbf{1}\{\omega \in \Omega : \ B_2(\omega)(t,s) < u(t,s), \ \forall \ (t,s) \in [0,1]^2\} \\
&\qquad \times \exp\left\{\int_{[0,1]^2}^{R} \rho\varphi''(t,s) \ dB_2(\omega)(t,s) - \frac{1}{2}\rho^2\left\|\varphi\right\|_{\mathcal{H}_\mathbf{B}}^2\right\}\mathbb{P}(d\omega).
\end{aligned}$$

Since $B_2(t,0) = 0$ a.s. for $t \in [0,1]$ and $B_2(0,s) = 0$ a.s. for $s \in [0,1]$, then $\Delta_{[0,1]^2}\varphi'' B_2(\cdot) = \varphi''(1,1)B_2(1,1)$ almost surely. The result follows immediately from integration by parts and the assumption that $\varphi''(t,s)$ is type I non decreasing on $[0,1]^2$ with $-\varphi''(t,1)$ and $-\varphi''(1,s)$ are non increasing on the closed interval $[0,1]$. □

Corollary 5.2.4. *Since the event* $\{\exists(t,s) \in [0,1]^2 : \ \rho\varphi(t,s)+B_2(t,s) \geq u(t,s)\}$ *is the complement of the event* $\{\forall(t,s) \in [0,1]^2 : \ \rho\varphi(t,s) + B_2(t,s) < u(t,s)\}$, *then under the conditions of Theorem 5.2.3, we get the following lower bound for the boundary crossing probability under consideration:*

$$\begin{aligned}
&\mathbb{P}\{\exists(t,s) \in [0,1]^2 : \ \rho\varphi(t,s) + B_2(t,s) \geq u(t,s)\} \\
&\geq 1 - k^*\mathbb{P}\{\forall(t,s) \in [0,1]^2 : \ B_2(t,s) < u(t,s)\} \\
&= 1 - k^* + k^*\mathbb{P}\{\exists(t,s) \in [0,1]^2 : \ B_2(t,s) \geq u(t,s)\},
\end{aligned}$$

where k^* is the constant defined above. In particular, if $u(t,s) = \tilde{t}_{1-\alpha}$, for $(t,s) \in [0,1]^2$, where $\tilde{t}_{1-\alpha}$ is the $(1-\alpha)$ quantile of $\sup_{0\le t,s\le 1} B_2(t,s)$, we get the lower bound for the localized power of the asymptotically size α Kolmogorov type test defined in Section 4.2. That is, we have

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{0\le t,s\le 1}(\rho\varphi(t,s)+B_2(t,s))\ge \tilde{t}_{1-\alpha}\right\}\\
&\ge 1-k_\alpha^*\mathbb{P}\{\sup_{0\le t,s\le 1} B_2(t,s)<\tilde{t}_{1-\alpha}\} = 1-k_\alpha^*(1-\alpha),\ \ \rho>0, \qquad (5.2.3)
\end{aligned}$$

where

$$\begin{aligned}
k_\alpha^* := &\exp\{\rho\varphi''(1,1)\tilde{t}_{1-\alpha} - \rho\Delta_{[0,1]}\varphi''(\cdot,1)\tilde{t}_{1-\alpha} - \rho\Delta_{[0,1]}\varphi''(1,\cdot)\tilde{t}_{1-\alpha}\\
&+\rho\Delta_{[0,1]^2}\varphi''(\cdot)\tilde{t}_{1-\alpha} - \frac{1}{2}\rho^2\left\|\varphi\right\|^2_{\mathcal{H}_\mathbf{B}}\}.
\end{aligned}$$

Remark 5.2.5. *The conditions of Theorem 5.2.3 are satisfied for instance by the function $\varphi(t,s) := c(ts)\mathbf{1}_{[0,1]^2} \in \mathcal{H}_{\mathbf{B}_{BV([0,1]^2)}}$, for $(t,s)\in[0,1]^2$ and $c\in\mathbb{R}$. In this case we obtain*

$$\begin{aligned}
&\mathbb{P}\left\{\forall(t,s)\in[0,1]^2:\ \rho ts + B_2(t,s) < u(t,s)\right\}\\
&\le\ \exp\{\rho c u(1,1) - \rho^2\left|c\right|^2/2\}\ \mathbb{P}\left\{\forall(t,s)\in[0,1]^2:\ B_2(t,s)<u(t,s)\right\}.
\end{aligned}$$

By Corollary 5.2.4 we further get

$$\begin{aligned}
&\mathbb{P}\{\exists(t,s)\in[0,1]^2:\ \rho\varphi(t,s)+B_2(t,s)\ge u(t,s)\}\\
&\ge 1-\exp\{\rho c u(1,1)-\rho^2\left|c\right|^2/2\}\mathbb{P}\{\forall(t,s)\in[0,1]^2:\ B_2(t,s)<u(t,s)\}\\
&=\ \exp\{\rho c u(1,1)-\rho^2\left|c\right|^2/2\}\mathbb{P}\{\exists(t,s)\in[0,1]^2:\ B_2(t,s)\ge u(t,s)\}\\
&\quad +1-\exp\{\rho c u(1,1)-\rho^2\left|c\right|^2/2\}. \qquad (5.2.4)
\end{aligned}$$

Putting $u(\cdot)=\tilde{t}_{1-\alpha}$, the right side of (5.2.3) becomes $1-\exp\{\rho c\tilde{t}_{1-\alpha}-\rho^2\left|c\right|^2/2\}(1-\alpha)$.

Corollary 5.2.6. *Under the same conditions on $\varphi(\cdot)$ and $u(\cdot)$ as in Theorem 5.2.3, we have*

$$\begin{aligned}
&\mathbb{P}\{\exists(t,s)\in[0,1]^2 : \rho\varphi(t,s)+B_2(t,s)\le u(t,s)\}\\
&\ge 1-k_1^*\mathbb{P}\{\forall(t,s)\in[0,1]^2 :\ B_2(t,s)<-u(t,s)\},
\end{aligned}\tag{5.2.5}$$

where

$$\begin{aligned}
k_1^* := \exp\Bigg\{&-\rho\varphi''(1,1)u(1,1)+\rho\int_{[0,1]}^{R}-u(t,1)d(-\varphi''(t,1))+\rho\int_{[0,1]}^{R}-u(1,s)d(-\varphi''(1,s))\\
&+\rho\int_{[0,1]^2}^{R}-u(t,s)d\varphi''(t,s)-\frac{1}{2}\rho^2\left\|\varphi\right\|^2_{\mathcal{H}_{\mathbf{B}}}\Bigg\}.
\end{aligned}$$

Proof. Since B_2 is a Brownian (2) motion, then $-B_2$ is also a Brownian (2) motion, and we have

$$\begin{aligned}
&\mathbb{P}\{\exists(t,s)\in[0,1]^2 : \rho\varphi(t,s)+B_2(t,s)\le u(t,s)\}\\
&=1-\mathbb{P}\{\forall(t,s)\in[0,1]^2 :\ -\rho\varphi(t,s)+B_2(t,s)<-u(t,s)\}.
\end{aligned}\tag{5.2.6}$$

Hence, the result follows by applying the Cameron-Martin-Girsanov formula and integration by parts. □

Remark 5.2.7. *We notice that for the function $\varphi(t,s)=c(ts)$, for $(t,s)\in[0,1]^2$ and $c\in\mathbb{R}$, we have $k_1^*=\exp\{-\rho c u(1,1)-\rho^2\left|c\right|^2/2\}$.*

Corollary 5.2.8. *For $\rho>0$, let*

$$\begin{aligned}
k^*_{\rho B_2} :=\ &\rho\varphi''(1,1)B_2(1,1)+\rho\int_{[0,1]}^{R}B_2(t,1)d(-\varphi''(t,1))\\
&+\rho\int_{[0,1]}^{R}B_2(1,s)d(-\varphi''(1,s))+\rho\int_{[0,1]^2}^{R}B_2(t,s)d\varphi''(t,s).
\end{aligned}$$

Then by the Cameron-Martin-Girsanov formula, integration by parts and Jensen's inequality, we get

$$\begin{aligned}
&\mathbb{P}\left\{\forall(t,s)\in[0,1]^2:\rho\varphi(t,s)+B_2(t,s)<u(t,s)\right\}\\
&=\exp\left\{-\frac{1}{2}\rho^2\left\|\varphi\right\|^2_{\mathcal{H}_{\mathbf{B}}}\right\}\int_\Omega \mathbf{1}\{\forall(t,s)\in[0,1]^2:B_2(t,s)<u(t,s)\}\exp\{k^*_{\rho B_2}\}d\mathbb{P}\\
&\geq\exp\left\{-\frac{1}{2}\rho^2\left\|\varphi\right\|^2_{\mathcal{H}_{\mathbf{B}}}\right\}\exp\left\{\mathbb{E}_{\mathbb{P}}\left(k^*_{\rho B_2}\mathbf{1}\{\forall(t,s)\in[0,1]^2:B_2(t,s)<u(t,s)\}\right)\right\}.
\end{aligned}$$

Hence, the upper bound of $\mathbb{P}\left\{\exists(t,s)\in[0,1]^2:\rho\varphi(t,s)+B_2(t,s)\geq u(t,s)\right\}$ *is*

$$1-\exp\left\{-\frac{1}{2}\rho^2\left\|\varphi\right\|^2_{\mathcal{H}_{\mathbf{B}}}\right\}\exp\left\{\mathbb{E}_{\mathbb{P}}\left(k^*_{\rho B_2}\mathbf{1}\{\forall(t,s)\in[0,1]^2:B_2(t,s)<u(t,s)\}\right)\right\}.$$

For the function $\varphi(\cdot)$ defined above, we have $k^*_{\rho B_2}=\rho c B_2(1,1)$. Hence (5.2.5) becomes

$$\begin{aligned}
&\mathbb{P}\{\exists(t,s)\in[0,1]^2:\rho c(ts)+B_2(t,s)\leq u(t,s)\}\\
&\geq 1-\exp\{-\rho c u(1,1)-\rho^2\left|c\right|^2/2\}\mathbb{P}\{\forall(t,s)\in[0,1]^2:\ B_2(t,s)<-u(t,s)\}. \quad (5.2.7)
\end{aligned}$$

5.3 Lower and upper bounds for the boundary crossing probabilities of B_2^0 with trend

In this section we consider the boundary crossing probability

$$\mathbb{P}\{\exists(t,s)\in[0,1]^2:\ \rho\varphi(t,s)+B_2^0(t,s)\geq u(t,s)\},\ \rho>0,$$

where B_2^0 is the standard Brownian (2) Bridge. By Theorem 4 in Lifshits (1996), p. 90, the kernel of B_2^0 can be derived analogously to that of standard Brownian (2) motion, i.e.,

$$\mathcal{H}_{B_2^0}:=\left\{h(\cdot):\exists\ell(\cdot)\in L_2([0,1]^2),h(t,s)=\langle\ell,m^\circ_{(t,s)}\rangle_{L_2},(t,s)\in[0,1]^2\right\},$$

where $m^{\circ}_{(t,s)}(\cdot) \in L_2([0,1]^2)$ is the model of the covariance function of the standard Brownian (2) bridge. For $\ell(\cdot) \in L_2([0,1]^2)$, let $\mathbb{E}_{\lambda^2}(\ell(\cdot))$ be the expectation of $\ell(\cdot)$ with respect to λ^2. Then, $\mathcal{H}_{B_2^0}$ can be equivalently written by

$$\mathcal{H}_{B_2^0} = \left\{ h(\cdot) : \exists \ell(\cdot) \in L_2([0,1]^2),\ h(t,s) = \int_{[0,t]\times[0,s]} \ell(\cdot) d\lambda^2 - ts\mathbb{E}_{\lambda^2}(\ell(\cdot)) \right\}.$$

Thus, the typical characteristic of any $h(\cdot) \in \mathcal{H}_{B_2^0}$ that distinguishes it from the elements of $\mathcal{H}_{\mathbf{B}}$ is that $h(1,1) = 0$, or more precisely $\mathcal{H}_{B_2^0} = \{h(\cdot) \in \mathcal{H}_{\mathbf{B}} : h(1,1) = 0\}$.

Analogous to the RKHS $\mathcal{H}_{\mathbf{B}}$, we furnish the kernel $\mathcal{H}_{B_2^0}$ with an inner product and an associated norm given by

$$\langle h_1, h_2 \rangle_{\mathcal{H}_{B_2^0}} := \left\langle \hat{h}_1, \hat{h}_2 \right\rangle_{L_2}, \text{ and } \|h\|_{\mathcal{H}_{B_2^0}} := \left\| \hat{h} \right\|_{L_2},$$

where

$$h_i(t,s) = \int_{[0,t]\times[0,s]} \hat{h}_i(\cdot) d\lambda^2,\ h_i(1,1) = 0,\ \hat{h}_i(\cdot) \in L_2([0,1]^2),\ i = 1,2.$$

Then with respect to these inner product and norm, $\mathcal{H}_{B_2^0}$ is a Hilbert space. Because of this reason we called $\mathcal{H}_{B_2^0}$ the reproducing kernel Hilbert space of the standard Brownian (2) bridge.

Proposition 5.3.1. *Let* $\mathcal{H}_{B^0_{2BV([0,1]^2)}} := \left\{ h(\cdot) \in \mathcal{H}_{B_2^0} : \frac{\partial^2 h(\cdot)}{\partial t \partial s} \in BV([0,1]^2) \right\}$. *For any* $h(\cdot) \in \mathcal{H}_{B^0_{2BV([0,1]^2)}}$, *the density of the shifted measure* $\mathcal{P}^h_{B_2^0}$ *with respect to* $\mathcal{P}_{B_2^0}$ *is*

$$\frac{d\mathcal{P}^h_{B_2^0}}{d\mathcal{P}_{B_2^0}}(x) = \exp\left\{ \int_{[0,1]^2}^{R} \frac{\partial^2 h(t,s)}{\partial t \partial s}\, dx(t,s) - \frac{1}{2} \|h\|^2_{\mathcal{H}_{B_2^0}} \right\}.$$

Proof. This result can also be proved similarly to the proof of Proposition 5.2.1. We also refer the reader to Theorem 3 in Lifshits (1996), p. 88. □

Remark 5.3.2. *By the definition of the standard Brownian (2) bridge (see Definition 2.2.1) and the characteristic of a function* $h(\cdot) \in \mathcal{H}_{B_2^0}$, *it can be shown that*

$$Var_{B_2^0}\left(\int_{[0,1]^2}^{R} \frac{\partial^2 h(t,s)}{\partial t \partial s}\, dx(t,s) \right) = \left\| \frac{\partial^2 h(\cdot)}{\partial t \partial s} \right\|^2_{L_2} = \|h\|^2_{\mathcal{H}_{B_2^0}},$$

where $Var_{B_2^0}(\cdot)$ *is the variance operator with respect to* B_2^0.

Theorem 5.3.3. *Suppose that the boundary* $u(\cdot)$ *is continuous on* $[0,1]^2$ *and the trend* $\varphi(\cdot) \in \mathcal{H}_{B^0_{2BV([0,1]^2)}}$ *is such that the existing second derivative* $\varphi'' := \frac{\partial^2 \varphi(t,s)}{\partial t \partial s}$ *is type I non decreasing on* $[0,1]^2$*. If the marginal functions* $\varphi''(t,1) := \frac{\partial^2 \varphi(t,s)}{\partial t \partial s}|_{s=1}$ *and* $\varphi''(1,s) := \frac{\partial^2 \varphi(t,s)}{\partial t \partial s}|_{t=1}$ *are non increasing on the closed interval* $[0,1]$*, we get*

$$\mathbb{P}\left\{\forall (t,s) \in [0,1]^2 : \ \rho\varphi(t,s) + B^0_2(t,s) < u(t,s)\right\}$$
$$\leq m^* \mathbb{P}\left\{\forall (t,s) \in [0,1]^2 : \ B^0_2(t,s) < u(t,s)\right\}, \ \rho > 0,$$

where

$$m^* := \exp\left\{\rho\varphi''(1,1)u(1,1) + \rho\int_{[0,1]} u(t,1) d(-\varphi''(t,1)) + \rho\int_{[0,1]} u(1,s) d(-\varphi''(1,s)) + \rho\int_{[0,1]^2} u(t,s) d\varphi''(t,s) - \frac{1}{2}\rho^2 \|\varphi\|^2_{\mathcal{H}_{B^0_2}}\right\}.$$

Proof. The result follows directly from Proposition 5.3.1 and integration by parts. □

Corollary 5.3.4. *If* $u(\cdot)$ *and* $\varphi(\cdot)$ *satisfy the conditions of Theorem 5.3.3, we get*

$$\mathbb{P}\{\exists (t,s) \in [0,1]^2 : \ \rho\varphi(t,s) + B^0_2(t,s) \geq u(t,s)\}$$
$$\geq 1 - m^* \mathbb{P}\{\forall (t,s) \in [0,1]^2 : \ B^0_2(t,s) < u(t,s)\}$$
$$= 1 - m^* + m^* \mathbb{P}\{\exists (t,s) \in [0,1]^2 : \ B^0_2(t,s) \geq u(t,s)\}.$$

Let $\tilde{t}^0_{1-\alpha}$ *is the* $(1-\alpha)$ *quantile of* $\sup_{0 \leq t,s \leq 1} B^0_2(t,s)$*. The lower bound for the localized power of the asymptotically size* α *Kolmogorov type test is*

$$\mathbb{P}\{\sup_{0 \leq t,s \leq 1} \left(\rho\varphi(t,s) + B^0_2(t,s)\right) \geq \tilde{t}^0_{1-\alpha}\}$$
$$\geq 1 - m^*_\alpha \mathbb{P}\{\sup_{0 \leq t,s \leq 1} B^0_2(t,s) < \tilde{t}^0_{1-\alpha}\} = 1 - m^*_\alpha(1-\alpha), \tag{5.3.1}$$

where

$$m^*_\alpha := \exp\{\rho\varphi''(1,1)\tilde{t}^0_{1-\alpha} - \rho\Delta_{[0,1]}\varphi''(\cdot,1)\tilde{t}^0_{1-\alpha} - \rho\Delta_{[0,1]}\varphi''(1,\cdot)\tilde{t}^0_{1-\alpha} + \rho\Delta_{[0,1]^2}\varphi''(\cdot)\tilde{t}^0_{1-\alpha} - \frac{1}{2}\rho^2 \|\varphi\|^2_{\mathcal{H}_{B^0_2}}\}.$$

Corollary 5.3.5. *Under the same conditions on $\varphi(\cdot)$ and $u(\cdot)$ as in Theorem 5.3.3, we obtain*

$$\begin{aligned}&\mathbb{P}\{\exists(t,s)\in[0,1]^2 : \rho\varphi(t,s)+B_2^0(t,s)\le u(t,s)\}\\&\ge 1-m_1^*\mathbb{P}\{\forall(t,s)\in[0,1]^2 :\ B_2^0(t,s)<-u(t,s)\},\end{aligned}\tag{5.3.2}$$

where

$$\begin{aligned}m_1^* := \exp\Big\{&-\rho\varphi''(1,1)u(1,1)+\rho\int_{[0,1]}-u(t,1)d(-\varphi''(t,1))\\&+\rho\int_{[0,1]}-u(1,s)d(-\varphi''(1,s))+\rho\int_{[0,1]^2}-u(t,s)d\varphi''(t,s)-\frac{1}{2}\rho^2\|\varphi\|^2_{\mathcal{H}_{B_2^0}}\Big\}.\end{aligned}$$

Corollary 5.3.6. *For $\rho>0$, let*

$$\begin{aligned}m^*_{\rho B_2^0} := &\int^R_{[0,1]}\rho B_2^0(t,1)d(-\varphi''(t,1))+\int^R_{[0,1]}\rho B_2^0(1,s)d(-\varphi''(1,s))\\&+\int^R_{[0,1]^2}\rho B_2^0(t,s)d\varphi''(t,s).\end{aligned}$$

Then by Proposition 5.3.1, integration by parts and Jensen's inequality, we get

$$\begin{aligned}&\mathbb{P}\left\{\forall(t,s)\in[0,1]^2 : \rho\varphi(t,s)+B_2^0(t,s)<u(t,s)\right\}\\&\ge\exp\left\{-\frac{1}{2}\rho^2\|\varphi\|^2_{B_2^0}\right\}\exp\left\{\mathbb{E}_{\mathcal{P}_{B_2^0}}\left(m^*_{\rho B_2^0}\mathbf{1}\{\forall(t,s)\in[0,1]^2 : B_2^0(t,s)<u(t,s)\}\right)\right\}.\end{aligned}$$

It follows that

$$\begin{aligned}&\mathbb{P}\left\{\exists(t,s)\in[0,1]^2 : \rho\varphi(t,s)+B_2^0(t,s)\ge u(t,s)\right\}\\&\le 1-\exp\left\{-\frac{1}{2}\rho^2\|\varphi\|^2_{B_2^0}\right\}\exp\left\{\mathbb{E}_{\mathcal{P}_{B_2^0}}\left(m^*_{\rho B_2^0}\mathbf{1}\{\forall(t,s)\in[0,1]^2 : B_2^0(t,s)<u(t,s)\}\right)\right\}.\end{aligned}$$

Remark 5.3.7. *We consider the model $\{m^\circ_{(t,s)}(\cdot) := \mathbf{1}_{[0,t]\times[0,s]}(\cdot)-ts\mathbf{1}_{[0,1]\times[0,1]}(\cdot) : (t,s)\in[0,1]^2\}$ of the covariance function of the standard Brownian (2) bridge. It can be easily shown that for every $(t,s)\in[0,1]^2$,*

- $m^{\circ}_{(t,s)}(\cdot)$ *is type I non decreasing on* $[0,1]^2$,

- $m^{\circ}_{(t,s)}(1,\cdot)$ *and* $m^{\circ}_{(t,s)}(\cdot,1)$ *are non increasing on* $[0,1]$.

Furthermore, we define a family $\{\varphi_{ts}(\cdot) : (t,s) \in [0,1]^2\}$ *of functions on* $[0,1]^2$ *given by* $\varphi_{ts}(x,y) := \int_{[0,x]\times[0,y]} m^{\circ}_{(t,s)}(\cdot)d\lambda^2 = \lambda^2([0,x]\times[0,y]\cap[0,t]\times[0,s]) - tsxy$, *for* $(x,y) \in [0,1]^2$. *Since for every* $(t,s) \in [0,1]^2$, $m^{\circ}_{(t,s)}(\cdot) \in BV([0,1]^2)$, *and*

$$\varphi_{ts}(1,1) = \int_{[0,1]\times[0,1]} \left(\mathbf{1}_{[0,t]\times[0,s]}(u,\nu) - ts\mathbf{1}_{[0,1]\times[0,1]}(u,\nu)\right)\lambda^2(du,d\nu) = 0,$$

then $\{\varphi_{ts}(\cdot) : (t,s) \in [0,1]^2\} \subset \mathcal{H}_{B^0_{2BV([0,1]^2)}}$. *Thus the conditions of Theorem 5.3.3 are satisfied by the family* $\{\varphi_{ts}(\cdot) : (t,s) \in [0,1]^2\}$.

Chapter 6

Discussions and conclusions

By virtue of being a new approach, model-check methods based on the residual partial sums process applied to spatial data analysis present open problems, both from a pure mathematical viewpoint as well as from the perspective of applications. Throughout this work, even under the simplest model, several difficulties and challenging mathematical problems are encountered. Since only few results concerning this subject are available, we preserve these problems as challenging open problems and directions for future research.

Throughout the thesis, wherever appropriate, there have been discussions and suggestions for modification and improvement of the model. Here we highlight only the major open questions and suggestions addressed throughout this work.

6.1 Open problems and further plans of research

- As mentioned in Section 4.2 and Section 4.3, we face difficulties in deriving analytical as well as approximation methods for computing the quantiles of $\sup_{0\le t,s\le 1} B_{\mathbf{f}}(t,s)$ and $\sup_{0\le t,s\le 1} |B_{\mathbf{f}}(t,s)|$. In this work we only propose Monte carlo simulations to approximate these quantiles. In the case of the standard Brownian (2) motion, an approximation method has been proposed, see e.g., Zimmerman (1972) and Lifshits (1996), p. 139-155. By directly extending this approach to the standard Brownian (2) bridge or more general processes $B_{\mathbf{f}}(\cdot)$ it does not seem to be satisfactory since the processes do not have independent increments. In the one-dimensional case, by applying the *Markov* property inherent to the Brownian bridge, Bischoff et al. (2003b) proposed an asymptotic method for computing such quantiles. Unfortunately, both the Markov property and the *reflection principle* are not available for standard Brownian (2) motion as well as for the standard Brownian (2) bridge, so that the last approach fails. Further investigation of the properties of $B_{\mathbf{f}}$ is therefore needed, so that these challenging mathematical problems can be solved.

- As before, we encounter in Section 4.4, a mathematical difficulty in computing the quantiles of $\int_{[0,1]^2} B_{\mathbf{f}}^2(\cdot)d\lambda^2$ analytically as well as approximately. The first idea is to extend directly either the classical Anderson and Darling's or Imhof or Slepian method, see Shorack and Wellner (1986), p. 212, or the Imhof-Eastwood's method, see Imhof (1961) and Eastwood (1993). To this end, based on Mercer's theorem, we need to derive the *principle component decomposition* (Karhunen-Loève expansion) of $B_{\mathbf{f}}$ which consists in finding the solution of a complicated integral equation with a kernel being the covariance function $K_{\mathbf{f}}(\cdot,\cdot)$ of $B_{\mathbf{f}}$. Since this typical mathematical problem appears in many tests of hypothesis based on Cramér-von Mises statistics, further alternative approaches should be developed.

- Similar complicated problems are found as we intend to compute analytically or approximately the power of the asymptotically size α tests: $\Psi_\delta(g)$, $\Psi_\delta(g)$ and $\Psi_\psi(g)$, for $g(\cdot) \in BV([0,1]^2)$. Here we handle these problems by Monte Carlo simulation, however analytical as well as approximation methods are important for the sake of comparing our results. Following an approach due to Bischoff et al. (2003b) does not seem to be satisfactory, since the Markov property does not hold for $B_{\mathbf{f}}$. To handle this problem we therefore need further development concerning the properties of Brownian (2) motion and Brownian (2) bridge.

- In our effort to establish upper and lower bounds for the power of the Kolmogorov-type test (Chapter 5) following the approach due to Lifshits (1996), p. 88-107, difficulties are encountered in deriving the model of $B_{\mathbf{f}}$ for the processes associated with first-order and second-order linear regression models. Consequently, we can not derive the kernel of the corresponding processes, and in turn the admissible shift can not be defined explicitly. This is mainly due to the structure of the covariance function $K_{\mathbf{f}}(\cdot,\cdot)$ of such processes as a complicated function of four variables in $[0,1]^2 \times [0,1]^2$. Further development of these method for deriving upper and lower bounds for the power of the Kolmogorov-Smirnov and Cramér-von Mises type tests is important and presents interesting and challenging mathematical problems.

6.2 Some remarks

- From the viewpoint of applications, regular lattice (equidistance experimental design) for spatial data analysis as well as in response surface methodology is easy to conduct on one hand, but on the other hand, it is clearly cost expensive. Therefore for the sake of establishing optimal tests based on residual partial sums processes addressed to the hypotheses formulated in Chapter 4 and efficiency in cost, it is important to further develop functional central limit theorems for residual partial sums processes in more general settings incorporating experimental design theory. Every *exact design* (for $n \times m$ observations): $\{(t_{ni}, s_{mj}) \in [0,1]^2 : i = 1, \ldots, n,\ j = 1, \ldots, m\}$ uniquely corresponds to a discrete probability measure P_{nm} on $[0,1]^2 \cap \mathcal{B}^2$, given by

$$P_{nm} := \frac{1}{nm} \sum_{j=1}^{m} \sum_{i=1}^{n} P_{\{(t_{ni}, s_{mj})\}},$$

where $P_{\{(t,s)\}}$ denotes the Dirac measure in (t,s). Conversely, given a probability measure P_0 on $[0,1]^2 \cap \mathcal{B}^2$, we can choose an exact design according to P_0: choose $(t_{ni}, s_{mj}) := Q_0(\frac{ij}{nm} - z_0)$ with $z_0 \in [0, \frac{1}{nm}]$ arbitrarily fixed, $1 \leq i \leq m$, $1 \leq j \leq n$, where Q_0 is the quantile function of P_0, see Bischoff (1998) and Bischoff and Miller (2000). In the case of a regular lattice with $m \times n$ observations, we get

$$\mathcal{L}_{nm} := \frac{1}{nm} \sum_{k=1}^{m} \sum_{\ell=1}^{n} P_{\{(\ell/n, k/m)\}} \Rightarrow_{n,m} \lambda_{[0,1]^2},$$

i.e., $\mathcal{L}_{nm}$ converges weakly to $\lambda_{[0,1]^2}$ in the sense $\mathbb{E}_{\mathcal{L}_{nm}}(f) \overset{n,m\to\infty}{\longrightarrow} \mathbb{E}_{\lambda^2}(f)$, for every bounded continuous function $f(\cdot)$ on $[0,1]^2$. Based on this fact, given a sequence $(P_{nm})_{n\geq 1, m\geq 1}$ of exact designs on $[0,1]^2 \cap \mathcal{B}^2$, under the assumptions

1. P_{nm} converges to an exact design P_0 in the sense

$$\sup_{(t,s)\in[0,1]^2} |F_{nm}(t,s) - F_0(t,s)| \overset{n,m\to\infty}{\longrightarrow} 0,$$

   where F_{nm} and F_0 are the distribution functions of P_{nm} and P_0, respectively.

2. the regression functions $f_1(\cdot), \ldots, f_p(\cdot)$ are linearly independent, continuous and have bounded variation on $[0,1]^2$,

we can derive, by applying the approach due to Bischoff (1998) and Bischoff and Miller (2000), the limit process of the sequence of the residual partial sums processes corresponding to $(P_{nm})_{n\geq 1, m\geq 1}$. A similar result as in our approach in Chater 3 will be obtained in which the limit process does not depend on the sequence of designs, but it depends only on the regression functions $\{f_1(\cdot), \ldots, f_p(\cdot)\}$.

- In applications, the assumption that $Cov(vec(\mathbf{E}_{n\times n})) = \sigma^2 (\mathbf{I}_{n^2\times n^2})$ given to (1.1.4) is sometimes reasonable, but there are certainly many occasions when it is unrealistic, i.e., the errors must be correlated. In spatial data analysis we frequently find a situation in which $Cov(vec(\mathbf{E}_{n\times n})) = \sigma^2 \mathbf{\Sigma}$, where $\sigma^2 \in (0, \infty)$ is unknown and $\mathbf{\Sigma}$ is a known $n^2 \times n^2$ non-singular matrix. An example is $Cov(\varepsilon_{uv}, \varepsilon_{ij}) = \sigma^2 \exp\{-\|(t_{nu}, s_{nv}) - (t_{ni}, s_{nj})\|\}$, for $1 \leq u, v, i, j \leq n$. Hence, in the analysis, particulary in performing tests based on residual partial sums process, the *spatial structure* inherent in the random errors should be incorporated. This problem can be well handled by applying the *generalized least squares* method, see e.g., Stapleton (1995), p. 163-165, or Schabenberger and Gotway (2005), p. 320-321. Coming back to the model described in Chapter 1, let $\mathbf{B}$ be a non-singular $n^2 \times n^2$-dimensional matrix such that $\mathbf{B}\mathbf{B}^\top = \mathbf{\Sigma}$. We transform the original observation $vec(\mathbf{Y}_{n\times n})$ to $\mathbf{Z_n} := \mathbf{B}^{-1} vec(\mathbf{Y}_{n\times n})$. Let $\eta_n := \mathbf{B}^{-1} vec(\mathbf{E}_{n\times n})$, for $i = 1, \ldots, p$, let $\mathbf{U}_i := \mathbf{B}^{-1} vec(\mathbf{M}_i)$ and let $\mathbf{U}_n := (\mathbf{U}_1, \ldots, \mathbf{U}_p) \in \mathbb{R}^{n^2\times n^2}$. Then $\eta_n \sim (\mathbf{O}, \sigma^2 \mathbf{I}_{n^2\times n^2})$ and the *generalized least squares residual* vector is given by $\hat{\eta}_n := \eta_n - pr_{\mathbf{U}} \eta_n$. We consider again the hypothesis formulated in Chapter 4. Testing the hypotheses $H_0 : \mathbf{G}_{n\times n} \in \mathbf{W}_n$ versus $K : \mathbf{G}_{n\times n} \notin \mathbf{W}_n$ based on the residual partial sums process associated with the original model is similar to testing the hypotheses $H_0 : \mathbf{B}^{-1} vec(\mathbf{G}_{n\times n}) \in \mathbf{U}_n$ versus $K : \mathbf{B}^{-1} vec(\mathbf{G}_{n\times n}) \notin \mathbf{U}_n$ based on the residual partial sums process associated with $\hat{\eta}_n$, whose limiting process can be derived similarly as that $B_{\mathbf{f}}$. All test procedures derived in Chapter 4 can be performed using the functional of residual partial sums process associated with $\hat{\eta}_n$.

- A generalization of our approach in Chapter 3 from $n \times n$ to $\underbrace{n \times \cdots \times n}_{k}$ regular lattice and $\underbrace{n_1 \times \cdots \times n_k}_{k}$ regular lattice, where $n_1, \ldots, n_k$ are not all equal, is straightforward. For this purpose we only need formula of integration by parts for Riemann-Stieltjes integral on the k-dimensional unit cube $[0,1]^k$ which is already available in Young (1917b) or Yeh (1963). Moreover we also need a k-dimensional version of Theorem 2.4.4. But this can be derived similarly as in the two-dimensional case.

6.3 Conclusions

This thesis presents an extension of some existing functional central limit theorems for least squares residual partial sums processes of the linear regression model in one dimension to higher dimension. A simple and advantageous approach for deriving the limit process was proposed and proved.

From the asymptotic, theoretical result presented, a test of hypotheses was proposed in Chapter 4 for checking the adequacy of the model. Three types of tests were proposed: Kolmogorov, Kolmogorov-Smirnov and Cramér-von Mises type tests. Monte carlo simulations were conducted for approximating the quantiles of the limiting statistics. The simulation results were very satisfactory compared to existing approaches and results. For the three type tests proposed, the consistency and the power of the tests were investigated. By conducting Monte Carlo simulations, the tests were shown to be powerful in the sense that the power is larger then the preassigned level of significance. Lower and upper bounds of the limiting power functions of Kolmogorov type test were also investigated.

Appendix A

Functions of bounded variation and the Riemann-Stieltjes integral

A.1 Definitions and Terminology

Definition A.1.1. **(1)** *A rectangle in* $[0,1]^2$ *is a subset* $\mathbf{I}$ *of* $[0,1]^2$ *which can be written as the Cartesian product* $[a_1, a_2] \times [b_1, b_2]$*, where* $0 \leq a_1 \leq a_2 \leq 1$*, and* $0 \leq b_1 \leq b_2 \leq 1$*.*

(2) *If* $\mathbf{I}$ *is a rectangle, the diameter of* $\mathbf{I}$ *is given by*

$$diam(\mathbf{I}) := \sqrt{(a_2 - a_1)^2 + (b_2 - b_1)^2}.$$

As a convention we put $diam(\emptyset) = 0$*.*

(3) *Let* Γ *be a collection of rectangles in* $[0,1]^2$*. We say that* Γ *is a non-overlapping, finite exact cover of* $[0,1]^2$ *if the following conditions hold:*

1. *Different elements of* Γ *have disjoint interiors, i.e, for any* $\mathbf{I}_1, \mathbf{I}_2 \in \Gamma$*, if* $\dot{\mathbf{I}}_1 \cap \dot{\mathbf{I}}_2 \neq \emptyset$*, then* $\mathbf{I}_1 = \mathbf{I}_2$*, where* $\dot{\mathbf{I}}$ *denotes the interior of* $\mathbf{I}$.
2. Γ *has only finitely many elements.*
3. $\bigcup_{\mathbf{I}\in\Gamma} \mathbf{I} = [0,1]^2$.

The set of all non-overlapping, finite exact covers of $[0,1]^2$ *is denoted by* $\prod[0,1]^2$.

(4) *For every* $\Gamma \in \prod[0,1]^2$*, we define a quantity* $\|\Gamma\|$ *by putting*

$$\|\Gamma\| := \max_{\mathbf{I}\in\Gamma} diam(\mathbf{I}).$$

(5) *Given a collection of rectangles* Γ*, the set* $\Xi(\Gamma)$ *is defined as follows:*

$$\Xi(\Gamma) := \{\xi : \Gamma \mapsto \bigcup_{\mathbf{I}\in\Gamma} \mathbf{I} \mid \xi(\mathbf{I}) \in \mathbf{I},\ \forall\ \mathbf{I} \in \Gamma\}.$$

(6) *Let* Γ_1 *and* Γ_2 *be in* $\prod[0,1]^2$*. We say that* Γ_2 *is more refined than* Γ_1 *and write* $\Gamma_1 \leq \Gamma_2$*, if for every* $\mathbf{I}_2 \in \Gamma_2$ *there is an* $\mathbf{I}_1 \in \Gamma_1$ *such that* $\mathbf{I}_2 \subseteq \mathbf{I}_1$.

(7) *For every pair* Γ_1 *and* $\Gamma_2 \in \prod[0,1]^2$*, the common refinement of* Γ_1 *and* Γ_2 *is given by*

$$\Gamma_1 \vee \Gamma_2 := \{\mathbf{I}_1 \cap \mathbf{I}_2 \mid \mathbf{I}_1 \in \Gamma_1,\ \mathbf{I}_2 \in \Gamma_2\}.$$

Definition A.1.2. *A real-valued function* ψ *defined on* $[0,1]^2$ *is said to be type I non decreasing on* $[0,1]^2$ *if for any* $(x,y) \in [0,1]^2$ *and any positive real numbers* h, u *such that* $x + h \leq 1$ *and* $y + u \leq 1$*, the following condition holds:*

$$\psi(x+h, y+u) - \psi(x, y+u) - \psi(x+h, y) + \psi(x,y) \geq 0.$$

Likewise, ψ *is said to be type I non increasing on* $[0,1]^2$ *if* $-\psi$ *is type I non decreasing on* $[0,1]^2$.

Definition A.1.3. *We say that ψ is type II non decreasing on $[0,1]^2$ if for any $(x,y) \in [0,1]^2$ and any positive real numbers h, u such that $x+h \leq 1$ and $y+u \leq 1$, and any $c \in [0,1]$, the following conditions hold:*

1. $\psi(x+h,c) - \psi(x,c) \geq 0$, *and* $\psi(c,y+u) - \psi(c,y) \geq 0$,

2. $\psi(x+h,y+u) - \psi(x,y+u) - \psi(x+h,y) + \psi(x,y) \geq 0$.

Likewise, ψ is type II non increasing on $[0,1]^2$ if $-\psi$ is type II non decreasing on $[0,1]^2$.

Definition A.1.4. *Let Γ be a non-overlapping, finite exact cover of $[0,1]^2$, and let ψ be a real-valued function defined on $[0,1]^2$. The variation of ψ over Γ, denoted by $v(\psi;\Gamma)$, is given by*

$$v(\psi;\Gamma) := \sum_{\mathbf{I}\in\Gamma} \left|\Delta_{\mathbf{I}}\psi\right|,$$

where for any rectangle $\mathbf{I} = [a_1,a_2] \times [b_1,b_2] \in \Gamma$,

$$\Delta_{\mathbf{I}}\psi := \psi(a_2,b_2) - \psi(a_1,b_2) - \psi(a_2,b_1) + \psi(a_1,b_1).$$

The total variation of ψ over $[0,1]^2$, denoted by $V(\psi;[0,1]^2)$, is given by

$$V(\psi;[0,1]^2) := \sup_{\Gamma\in\prod[0,1]^2} v(\psi;\Gamma).$$

Definition A.1.5. *The function ψ is said to have bounded variation on $[0,1]^2$ in the sense of Vitali, if there exists a positive real number M such that $V(\psi;[0,1]^2) \leq M$, see Clarkson and Adams (1933). We denote the class of these functions by $BVV([0,1]^2)$. The definition is analogous to the definition of a function which has bounded variation in one dimension.*

Definition A.1.6. *The function ψ is said to have bounded variation on $[0,1]^2$ in the sense of Hardy, if the following conditions hold, see Clarkson and Adams (1933):*

- *There exists a positive real number M such that $V(\psi;[0,1]^2) \leq M$.*
- *There exist $\bar{x}, \bar{y} \in [0,1]$ such that the functions $\psi(\bar{x},\cdot)$ and $\psi(\cdot,\bar{y})$ have bounded variation on $[0,1]$.*

As a convention, we denote this class of functions by $BV([0,1]^2)$. It is clear that, $BV([0,1]^2) \subseteq BVV([0,1]^2)$.

Proposition A.1.7. *Let Φ and Ψ be functions on $[0,1]^2$, $\Phi(x) := V(\psi(x,\cdot);[0,1])$ and $\Psi(y) := V(\psi(\cdot,y);[0,1])$, $x, y \in [0,1]$. If ψ has bounded variation on $[0,1]^2$ in the sense of Hardy, then Φ and Ψ have bounded variation on $[0,1]$. We shall call Φ and Ψ the total variation functions. As a convention, we denote the class of functions which have bounded variation on $[0,1]$ by $BV([0,1])$.*

Proof. See Theorem 1 in Clarkson and Adams (1933). □

Remark A.1.8. *If $\psi \in BV([0,1]^2)$, then for arbitrarily fixed $\bar{x}$ and $\bar{y}$ in $[0,1]$, the functions $\psi(\bar{x},\cdot)$ and $\psi(\cdot,\bar{y})$ defined on $[0,1]$ are in $BV([0,1])$. We refer the reader to Stroock (1994), p. 12-18, for further discussion of the notion of the class $BV([0,1])$.*

Definition A.1.9. *For ψ in $BVV([0,1]^2)$, the positive and negative variations of ψ associated with $\Gamma \in \prod[0,1]^2$ are given by*

$$v_+(\psi;\Gamma) := \sum_{\mathbf{I}\in\Gamma} (\Delta_{\mathbf{I}}\psi)^+ \text{ and } v_-(\psi;\Gamma) := \sum_{\mathbf{I}\in\Gamma} (\Delta_{\mathbf{I}}\psi)^-,$$

where $\alpha^+ := \max\{\alpha, 0\}$ and $\alpha^- := -\min\{\alpha, 0\}$. Likewise, the positive and negative total variations of ψ in $[0,1]^2$ are defined by

$$V_+(\psi;[0,1]^2) := \sup_{\Gamma\in\prod[0,1]^2} v_+(\psi;\Gamma) \text{ and } V_-(\psi;[0,1]^2) := \sup_{\Gamma\in\prod[0,1]^2} v_-(\psi;\Gamma).$$

Corollary A.1.10. *For each* $(x, y) \in [0,1]^2$ *and* ψ *in* $BVV([0,1]^2)$ *let us consider a closed rectangle* $[0,x] \times [0,y] \subset [0,1]^2$. *For any* $\Gamma_{xy} \in \prod([0,x] \times [0,y])$ *we define the quantities* $v_+(\psi; \Gamma_{xy})$, $v_-(\psi; \Gamma_{xy})$, $V_+(\psi; [0,x] \times [0,y])$, *and* $V_-(\psi; [0,x] \times [0,y])$ *similarly as in Definition A.1.9. Then for any* $(x, y) \in [0,1]^2$, *we have*

$$\left.\begin{aligned} &(1)\ V_+(\psi; [0,x] \times [0,y]) - V_-(\psi; [0,x] \times [0,y]) = \Delta_{[0,x]\times[0,y]}\psi \\ &(2)\ V_+(\psi; [0,x] \times [0,y]) + V_-(\psi; [0,x] \times [0,y]) = V(\psi; [0,x] \times [0,y]) \end{aligned}\right\} \quad \text{(A.1.1)}$$

Theorem A.1.11. *A function* ψ *is in* $BVV([0,1]^2)$ *if and only if, there exist* ψ_1 *and* ψ_2 *which are type I non decreasing on* $[0,1]^2$ *such that* $\psi = \psi_1 - \psi_2$.

Proof. See Theorem 5 in Adams and Clarkson (1934). See also Theorem 1.2.18 and Exercise 1.2.29 in Stroock (1994) for a version of this result for function of one variable. To show the necessary condition we put

$$\begin{aligned} \psi_1(x,y) :=& V_+(\psi; [0,x] \times [0,y]) + 1/2\left(\psi(0,y) + \psi(x,0) - \psi(0,0)\right), \\ \psi_2(x,y) :=& V_-(\psi; [0,x] \times [0,y]) - 1/2\left(\psi(0,y) + \psi(x,0) - \psi(0,0)\right). \end{aligned}$$

□

Theorem A.1.12. *A necessary and sufficient condition that* ψ *be in* $BV([0,1]^2)$ *is that it be expressible as the difference of two functions* ψ_1 *and* ψ_2 *which are type II non decreasing on* $[0,1]^2$.

Proof. See Theorem 6 in Adams and Clarkson (1934). We notice that the necessary condition can be shown by defining

$$\begin{aligned} \psi_1(x,y) :=& V_+(\psi; [0,x] \times [0,y]) + V_+(\psi(\cdot,0); [0,x]) + V_+(\psi(0,\cdot); [0,y]) + 1/2\psi(0,0), \\ \psi_2(x,y) :=& V_-(\psi; [0,x] \times [0,y]) + V_-(\psi(\cdot,0); [0,x]) + V_-(\psi(0,\cdot); [0,y]) + 1/2\psi(0,0). \end{aligned}$$

□

Theorem A.1.13. *Let ψ be in $BVV([0,1]^2)$. If ψ is right continuous on $[0,1)^2$ in the sense of Definition 2.4.1, then ψ_1 and ψ_2 defined in the proof of Theorem A.1.11 are also right continuous on $[0,1)^2$.*

Proof. We extend the proof of Theorem 5.5.2 in Douglass (1996) to the higher-dimensional case. We first show that ψ_1 is right continuous on $[0,1)^2$. Let $\varepsilon > 0$ and (x, y) be arbitrarily fixed in $[0,1)^2$ with $0 \le x < t$ and $0 \le y < s$, for some $(t,s) \in [0,1]^2$. Suppose that

$$\Gamma_{ts} := \{[t_{\ell-1}, t_\ell] \times [s_{k-1}, s_k] : 1 \le \ell \le m, 1 \le k \le q\},$$
$$0 = t_0 < t_1 < \cdots < t_m = t,\ 0 = s_0 < s_1 < \cdots < s_q = s,$$

such that $V_+(\psi; [0,t] \times [0,s]) - v_+(\psi; \Gamma) \le \varepsilon/2$, for all $\Gamma \in \prod([0,t] \times [0,s])$ with $\Gamma_{ts} \le \Gamma$. Let $((x_n, y_n))_{n \ge 1}$ be a sequence in $[0,1]^2$ which converges to (x,y) from above. Then there exists an $n_0 := n_0(\varepsilon) \in \mathbb{N}$ and an index $(u, \nu) \in \{(\ell, k) : 0 \le \ell \le (m-1), 0 \le k \le (q-1)\}$ such that $t_u \le x \le x_n \le t_{u+1}$ and $s_\nu \le y \le y_n \le s_{\nu+1}$, for $n \ge n_0$. Since ψ is right continuous on $[0,1)^2$, there exists an $n_0' := n_0'(\varepsilon)$ such that $|\psi(x_n, y_n) - \psi(x,y)| \le \frac{\varepsilon}{4(\nu+u+3)}$, for $n \ge n_0'$. For a fixed $n \ge n_0'' := \max\{n_0, n_0'\}$, let us define a finite exact cover Γ_{ts}' of the interval $[0,t] \times [0,s]$ by putting

$$\Gamma_{ts}' := \{[t_0, t_1], \ldots, [t_u, x], [x, x_n], [x_n, t_{u+1}], \ldots, [t_{m-1}, t_m]\}$$
$$\times \{[s_0, s_1], \ldots, [s_\nu, y], [y, y_n], [y_n, s_{\nu+1}], \ldots, [s_{q-1}, s_q]\}.$$

Thus, Γ_{ts}' is a refinement of Γ_{ts}. Moreover, we obtain

$$\begin{aligned}\varepsilon/2 &\ge V_+(\psi; [0,t] \times [0,s]) - v_+(\psi; \Gamma_{ts}') \\ &= \big[V_+(\psi; [0,t] \times [0,s]) - V_+(\psi; [0,x_n] \times [0,y_n]) - (v_+(\psi; \Gamma_{ts}') - v_+(\psi; \Gamma_{x_n y_n}'))\big]\end{aligned}$$

$$
\begin{aligned}
&+ \left[V_+(\psi;[0,x_n]\times[0,y_n]) - V_+(\psi;[0,x]\times[0,y]) - (v_+(\psi;\Gamma'_{x_ny_n}) - v_+(\psi;\Gamma'_{xy}))\right] \\
&+ \left[V_+(\psi;[0,x]\times[0,y]) - v_+(\psi;\Gamma'_{xy})\right] \\
\geq\ & V_+(\psi;[0,x_n]\times[0,y_n]) - V_+(\psi;[0,x]\times[0,y]) - (v_+(\psi;\Gamma'_{x_ny_n}) - v_+(\psi;\Gamma'_{xy})) \geq 0.
\end{aligned}
$$

It can easily be shown that $v_+(\psi;\Gamma'_{x_ny_n}) - v_+(\psi;\Gamma'_{xy}) \leq \varepsilon/2$, for $n \geq n''_0$. Hence, the last inequality gives the result $0 \leq V_+(\psi;[0,x_n]\times[0,y_n]) - V_+(\psi;[0,x]\times[0,y]) \leq \varepsilon$, for $n \geq n''_0$. This leads us to the conclusion that the function $V_+(\psi;[0,\cdot]\times[0,\cdot])$ is right continuous on $[0,1)^2$. The result follows since the marginal function $\psi(0,\cdot)$ and $\psi(\cdot,0)$ are obviously right continuous on $[0,1)$. The right continuity of ψ_2 on $[0,1)^2$ can be shown analogously. □

Theorem A.1.14. *Let ψ be in $BV([0,1]^2)$. If ψ is right continuous on $[0,1)^2$ in the sense of Definition 2.4.1, then ψ_1 and ψ_2 defined in the proof of Theorem A.1.12 are also right continuous on $[0,1)^2$.*

Proof. As before, we first prove the assertion for ψ_1 by showing that $V_+(\psi;[0,\cdot]\times[0,\cdot])$ is right continuous on $[0,1)^2$. But this was already shown in Theorem A.1.13. We refer the reader to Exercise 1.2.29 in Stroock (1994) for the right continuity of $V_+(\psi(\cdot,0);[0,\cdot])$ and $V_+(\psi(0,\cdot);[0,\cdot])$ on $[0,1)$. Hence the proof for ψ_1 is complete. The assertion for ψ_2 can be handled similarly. □

A.2 Riemann-Stieltjes integral on $[0,1]^2$

Definition A.2.1. *Let $\Gamma \in \prod[0,1]^2$, and let φ and ψ be real-valued functions defined on $[0,1]^2$. Then the Riemann-Stieltjes sum (abbreviated as RS-sum) of φ over Γ with* *respect to ψ relative to a mapping $\xi \in \Xi(\Gamma)$ is defined by*

$$\mathcal{RS}(\varphi \mid \psi, \Gamma, \xi) := \sum_{\mathbf{I} \in \Gamma} \varphi(\xi(\mathbf{I}))\Delta_{\mathbf{I}}\psi.$$

The function φ is said to be Riemann-Stieltjes integrable with respect to ψ, or more simply, ψ-RS-integrable on $[0,1]^2$, if there exists a real number A with the property that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for every $\Gamma \in \prod[0,1]^2$ with $\|\Gamma\| \leq \delta$, we have

$$\sup_{\xi \in \Xi(\Gamma)} |\mathcal{RS}(\varphi \mid \psi, \Gamma, \xi) - A| \leq \varepsilon.$$

If such a real number A exists, we call A the RS-integral of φ with respect to ψ, and denote it by

$$A = \int_{[0,1]^2}^{R} \varphi(t,s)\ d\psi(t,s).$$

Theorem A.2.2. *Let φ_i be ψ-RS-integrable on $[0,1]^2$, and let c_i be any real number, $i = 1, \ldots, k$. Then $\sum_{i=1}^{k} c_i\varphi_i$ is ψ-RS- integrable on $[0,1]^2$, with*

$$\int_{[0,1]^2}^{R} \sum_{i=1}^{k} c_i\varphi_i(t,s)d\psi(t,s) = \sum_{i=1}^{k} c_i \int_{[0,1]^2}^{R} \varphi_i(t,s)d\psi(t,s).$$

Theorem A.2.3. *Let φ be ψ_i-RS-integrable on $[0,1]^2$, and let c_i be any real number, $i = 1, \ldots, k$. Then φ is $\sum_{i=1}^{k} c_i\psi_i$-RS- integrable on $[0,1]^2$, with*

$$\int_{[0,1]^2}^{R} \varphi(t,s)\ d\left(\sum_{i=1}^{k} c_i\psi_i(t,s)\right) = \sum_{i=1}^{k} c_i \int_{[0,1]^2}^{R} \varphi(t,s)d\psi_i(t,s).$$

Theorem A.2.4. *If $\psi \in BVV([0,1]^2)$, every $\varphi \in \mathcal{C}([0,1]^2)$ is ψ-RS-integrable on $[0,1]^2$. Moreover, we have*

$$\left| \int_{[0,1]^2}^{R} \varphi(t,s)d\psi(t,s) \right| \leq \|\varphi\|_\infty V(\psi;[0,1]^2). \tag{A.2.1}$$

Proof. The existence of the integral $\int_{[0,1]^2}^{R} \varphi(t,s)d\psi(t,s)$ is due to Young (1917a). We refer the reader to Clarkson (1932) and Smirnov (1969), p. 56-60. The second assertion follows directly from the definition of the Riemann-Stieltjes integral. □

Theorem A.2.5. ***(Integration by parts)***

Let φ be ψ-RS-integrable on $[0,1]^2$, let $\varphi(c,\cdot)$ be $\psi(c,\cdot)$-RS-integrable on $[0,1]$ and let $\varphi(\cdot,c)$ be $\psi(\cdot,c)$-RS-integrable on $[0,1]$, for $c=0$ and 1. Then ψ is φ-RS-integrable on $[0,1]^2$ and we have

$$\int_{[0,1]^2}^{R} \psi(t,s)d\varphi(t,s) = \Delta_{[0,1]^2}(\varphi\psi) + \int_{[0,1]^2}^{R} \varphi(t,s)d\psi(t,s) + \int_{[0,1]}^{R} \varphi(0,s)d\psi(0,s) + \int_{[0,1]}^{R} \varphi(t,0)d\psi(t,0) - \int_{[0,1]}^{R} \varphi(t,1)d\psi(t,1) - \int_{[0,1]}^{R} \varphi(1,s)d\psi(1,s). \quad \text{(A.2.2)}$$

Proof. We consult the reader to Young (1917a) for the complete proof this Theorem. See also Móricz (2006) for Formula (A.2.2) for the case $\varphi(\cdot) \in \mathcal{C}([0,1]^2)$ and $\psi \in BV([0,1]^2)$. □

The following three Theorems (Theorem A.2.6, Theorem A.2.7 and Theorem A.2.8) are two-dimensional versions of the results in Bartle (1976), p.241-243. We refer the reader to Luxemberg (1971) for consulting the results.

Theorem A.2.6. *Let ψ be non decreasing on $[0,1]^2$ and let φ_n, $n \geq 1$ be ψ-RS-integrable on $[0,1]^2$. If $\varphi_n \xrightarrow{\|\cdot\|_\infty} \varphi$, then φ is ψ-RS-integrable on $[0,1]^2$ and*

$$\int_{[0,1]^2}^{R} \varphi(t,s)d\psi(t,s) = \lim_{n\to\infty} \int_{[0,1]^2}^{R} \varphi_n(t,s)d\psi(t,s).$$

Theorem A.2.7. ***(Bounded convergence theorem)***

Let ψ be non decreasing on $[0,1]^2$ and let φ_n, $n \geq 1$ be ψ-RS-integrable on $[0,1]^2$. Suppose that there exists an $M>0$ such that $|\varphi_n(x)| \leq M$, for $n \geq 1$ and $x \in [0,1]^2$.

If there exists a ψ-RS-integrable function φ on $[0,1]^2$ such that $\varphi_n \overset{n\to\infty}{\longrightarrow} \varphi$ point wise, then

$$\int_{[0,1]^2}^{R} \varphi(t,s)d\psi(t,s) = \lim_{n\to\infty} \int_{[0,1]^2}^{R} \varphi_n(t,s)d\psi(t,s).$$

Theorem A.2.8. ***(Monotone convergence theorem)***
Let ψ be non decreasing on $[0,1]^2$ and let φ, φ_n, $n \geq 1$ be ψ-RS-integrable on $[0,1]^2$ satisfying $\varphi_1 \leq \varphi_2 \leq \cdots \leq \varphi$. Suppose that $\lim_{n\to\infty} \varphi_n = \varphi$ point wise on $[0,1]^2$, then

$$\int_{[0,1]^2}^{R} \varphi(t,s)d\psi(t,s) = \lim_{n\to\infty} \int_{[0,1]^2}^{R} \varphi_n(t,s)d\psi(t,s).$$

A.3 Proof of Theorem 2.4.4

Proof. Let $(\Gamma_n)_{n\geq 1}$ be the following sequence of partitions of $(0,1]^2$, where

$$\Gamma_n := \{\mathbf{I}_{n;\ell k} := ((\ell-1)/2^n, \ell/2^n] \times ((k-1)/2^n, k/2^n] : \ 1 \leq k, \ell \leq 2^n, \ n \in \mathbb{N}\},$$

with $\|\Gamma_n\| = \frac{\sqrt{2}}{2^n} \overset{n\to\infty}{\longrightarrow} 0$. Associated with the sequence $(\Gamma_n)_{n\geq 1}$ and the function φ, let us define a sequence of step functions $(g_n(\varphi))_{n\geq 1}$, $(h_n(\varphi))_{n\geq 1} : (0,1]^2 \to \mathbb{R}$, by putting

$$g_n(\varphi) := \sum_{k=1}^{2^n} \sum_{\ell=1}^{2^n} (\inf_{(t,s)\in \bar{\mathbf{I}}_{n;\ell k}} \varphi(t,s)) \chi_{\mathbf{I}_{n;\ell k}}$$
$$h_n(\varphi) := \sum_{k=1}^{2^n} \sum_{\ell=1}^{2^n} (\sup_{(t,s)\in \bar{\mathbf{I}}_{n;\ell k}} \varphi(t,s)) \chi_{\mathbf{I}_{n;\ell k}},$$

where $\bar{\mathbf{I}}_{n;\ell k}$ is the *closure* of $\mathbf{I}_{n;\ell k}$, and $\chi_{\mathbf{I}_{n;\ell k}}$ stands for the indicator function of $\mathbf{I}_{n;\ell k}$ on $[0,1]^2$. Then for $n \geq 1$, $g_n(\varphi) \leq \varphi \leq h_n(\varphi)$ everywhere on $(0,1]^2$. It is clear that $g_n(\varphi)$ and $h_n(\varphi)$ are Borel-measurable and bounded, hence ν_ψ-integrable on $(0,1]^2$, where ν_ψ is the Lebesgue-Stieltjes measure on $((0,1]^2, \mathcal{B}^2 \cap (0,1]^2)$ associated with a Type I non decreasing function $\psi \in \mathcal{R}_c([0,1]^2)$. Moreover,

$$\int_{(0,1]^2} g_n(\varphi) d\nu_\psi = \sum_{k=1}^{2^n} \sum_{\ell=1}^{2^n} (\inf_{(t,s)\in\bar{\mathbf{I}}_{n;\ell k}} \varphi(t,s)) \Delta_{\bar{\mathbf{I}}_{n;\ell k}} \psi$$
$$\int_{(0,1]^2} h_n(\varphi) d\nu_\psi = \sum_{k=1}^{2^n} \sum_{\ell=1}^{2^n} (\sup_{(t,s)\in\bar{\mathbf{I}}_{n;\ell k}} \varphi(t,s)) \Delta_{\bar{\mathbf{I}}_{n;\ell k}} \psi.$$

Furthermore, $g := \lim_{n\to\infty} g_n(\varphi)$ and $h := \lim_{n\to\infty} h_n(\varphi)$ are clearly Borel-measurable and bounded, hence ν_ψ-integrable on $(0,1]^2$. Since $(|g_n(\varphi)|)_{n\geq 1}$ and $(|h_n(\varphi)|)_{n\geq 1}$ are dominated by positive constants, by *Lebesgue's dominated convergence theorem* and the Riemann-Stieltjes integrability of φ with respect to ψ on $[0,1]^2$, we get

$$\int_{(0,1]^2} g d\nu_\psi = \int_{[0,1]^2}^{R} \varphi(t,s) d\psi(t,s) = \int_{(0,1]^2} h d\nu_\psi.$$

This implies $g = h$, ν_ψ a.e. on $(0,1]^2$, hence $g = \varphi$, ν_ψ a.e. on $(0,1]^2$. By Exercise 4.5 in Elstrodt (2005), p. 143, φ is $\bar{\nu}_\psi$-integrable on $(0,1]^2$, and we obtain

$$\int_{(0,1]^2} \varphi d\bar{\nu}_\psi = \int_{(0,1]^2} g d\nu_\psi = \int_{[0,1]^2}^{R} \varphi(t,s) d\psi(t,s),$$

where $\bar{\nu}_\psi$ is the completion of ν_ψ. Furthermore, let $\mathbf{D}$ be the set of discontinuity points of φ, and let $\mathbf{R} := \cup_{n=1}^{\infty} \cup_{k=1}^{2^n} \cup_{\ell=1}^{2^n} \partial \mathbf{I}_{n,\ell k}$, where $\partial \mathbf{I}_{n,\ell k}$ is the *boundary* of $\mathbf{I}_{n,\ell k}$. Then $\mathbf{D} \subset \mathbf{R} \cup \{g < h\}$. From the preceding result, we get $\nu_\psi(\mathbf{R} \cup \{g < h\}) = 0$. Then we have $\bar{\nu}_\psi(\mathbf{D}) = 0$. This leads us to the conclusion that φ is continuous $\bar{\nu}_\psi$ a.e. on $(0,1]^2$.

Conversely, suppose that $\bar{\nu}_\psi(\mathbf{D}) = 0$, then $g = h$, ν_ψ a.e. on $(0,1]^2$. By Theorem 4.2.c in Elstrodt (2005), we have $\int_{(0,1]^2} g d\nu_\psi = \int_{(0,1]^2} h d\nu_\psi$, since g and h are ν_ψ-integrable on $(0,1]^2$. By the dominated convergence theorem, we obtain

$$\lim_{n\to\infty} \sum_{k=1}^{2^n} \sum_{\ell=1}^{2^n} (\inf_{(t,s)\in\bar{\mathbf{I}}_{n;\ell k}} \varphi(t,s)) \Delta_{\bar{\mathbf{I}}_{n;\ell k}} \psi = \lim_{n\to\infty} \sum_{k=1}^{2^n} \sum_{\ell=1}^{2^n} (\inf_{(t,s)\in\bar{\mathbf{I}}_{n;\ell k}} \varphi(t,s)) \nu_\psi(\mathbf{I}_{n;\ell k})$$

$$= \int_{(0,1]^2} g d\nu_\psi = \int_{(0,1]^2} h d\nu_\psi = \lim_{n\to\infty} \sum_{k=1}^{2^n} \sum_{\ell=1}^{2^n} (\sup_{(t,s)\in \bar{\mathbf{I}}_{n;\ell k}} \varphi(t,s)) \nu_\psi(\mathbf{I}_{n;\ell k})$$
$$= \lim_{n\to\infty} \sum_{k=1}^{2^n} \sum_{\ell=1}^{2^n} (\sup_{(t,s)\in \bar{\mathbf{I}}_{n;\ell k}} \varphi(t,s)) \Delta_{\bar{\mathbf{I}}_{n;\ell k}} \psi.$$

Therefore φ is Riemann-Stieltjes integrable with respect to ψ on $[0,1]^2$. □

A.4 Proof of Proposition 2.4.5

Proof. Let $\varphi \in \mathcal{C}([0,1]^2)$ and let $\varepsilon > 0$. There exists a $\delta > 0$ such that

$$\sup_{\|(t_1,s_1)-(t_2,s_2)\|\le\delta} |\varphi(t_1,s_1) - \varphi(t_2,s_2)| \le \frac{\varepsilon}{M + \|\varphi\|_\infty}.$$

For such a $\delta > 0$, there exists a partition Γ of $[0,1]^2$,

$$\Gamma := \{[x_0,x_1],(x_1,x_2],\ldots,(x_{p-1},t_p]\} \times \{[y_0,y_1],(y_1,y_2],\ldots,(y_{m-1},y_m]\},$$

where

$$0 = x_0 < x_1 < \ldots < x_{\ell-1} < x_\ell < \ldots < x_p = 1$$
$$0 = y_0 < y_1 < \ldots < y_{k-1} < y_k < \ldots < y_m = 1,$$

such that $\max_{1\le\ell\le p, 1\le k\le m} \|(x_{\ell-1}, y_{k-1}) - (x_\ell, y_k)\| \le \delta$. Let us denote the elements of Γ by $A_{\ell k}$, $1 \le \ell \le p$, $1 \le k \le m$. Note that for $1 \le \ell \le p, 1 \le k \le m$, we get

$$|\Delta_{\mathbf{I}_{\ell k}} \psi_n| \le 4 \|\psi_n\|_\infty, n \ge 1.$$

Since $\|\psi_n\|_\infty \overset{n\to\infty}{\longrightarrow} 0$, there exists an $n_0 = n_0(\varepsilon) \in \mathbb{N}$ such that

$$\|\psi_n\|_\infty \le \frac{\varepsilon}{4mp(M + \|\varphi\|_\infty)}, \text{ for all } n \ge n_0.$$

Let us define a step function $g := \sum_{k=1}^{m} \sum_{\ell=1}^{p} (\min_{(t,s)\in \bar{A}_{\ell k}} \varphi(t,s))\chi_{A_{\ell k}}$, where $\bar{A}_{\ell k}$ stands for the closure of $A_{\ell k}$ and χ_A stands for the indicator function of $A \subseteq [0,1]^2$. Then $\|\varphi - g\|_\infty \le \frac{\varepsilon}{M+\|\varphi\|_\infty}$, and $\|g\|_\infty \le \|\varphi\|_\infty$. Furthermore, by the definition of g, we obtain

$$\left| \int_{[0,1]^2}^{R} g(t,s) d\psi_n(t,s) \right| \le\ 4mp \|g\|_\infty \|\psi_n\|_\infty \le 4mp \|\varphi\|_\infty \|\psi_n\|_\infty \le \frac{\varepsilon \|\varphi\|_\infty}{M + \|\varphi\|_\infty}.$$

Hence, by the triangle inequality and by inequality (A.2.1), we finally get

$$\left| \int_{[0,1]^2}^{R} \varphi(t,s) d\psi_n(t,s) \right| \le\ \|\varphi - g\|_\infty V(\psi_n; [0,1]^2) + \frac{\varepsilon \|\varphi\|_\infty}{M + \|\varphi\|_\infty} \le \varepsilon,\ n \ge n_0.$$

This leads us to the conclusion $\int_{[0,1]^2}^{R} \varphi(t,s) d\psi_n(t,s) \stackrel{n\to\infty}{\longrightarrow} 0.$ □

Appendix B

Weak convergence on $\mathcal{C}([0,1]^2)$

B.1 Etemadi's Inequality

Proposition B.1.1. *Let $(X_{\ell k})_{\ell=1,k=1}^{n,m}$ be an $n \times m$ dimensional random matrix whose components are independent random variables with finite mean. Let for $1 \leq u \leq n$, $1 \leq v \leq m$, $S_{uv} := \sum_{\ell=1}^{u} \sum_{k=1}^{v} X_{\ell k}$. Then for $\alpha \in \mathbb{R}$,*

$$\mathbb{P}\left\{ \max_{\substack{1 \leq v \leq m \\ 1 \leq u \leq n}} |S_{uv}| \geq 3\alpha \right\} \leq 3 \max_{\substack{1 \leq v \leq m \\ 1 \leq u \leq n}} \mathbb{P}\left\{ |S_{uv}| \geq \alpha \right\}.$$

Proof. See Theorem 22.5 in Billingsley (1995). □

B.2 Tightness and compactness in $\mathcal{C}([0,1]^2)$

Definition B.2.1. *Let $\mathcal{S}$ be a metric space and $\mathcal{B}_{\mathcal{S}}$ be the Borel σ-algebra over $\mathcal{S}$. A sequence of probability measures $\{P_n\}_{n\geq 1}$ on a measurable space $(\mathcal{S}, \mathcal{B}_{\mathcal{S}})$ is said to be tight, if and only if $\forall \eta > 0$ $\exists$ a compact $K \subset \mathcal{S}$, such that $P_n(K) > 1 - \eta$, $\forall n \geq 1$. Correspondingly, the sequence $\{X_n\}_{n\geq 1}$ of random elements on $(\mathcal{S}, \mathcal{B}_{\mathcal{S}})$ is tight, if and* *only if the sequence of their distributions is tight.*

Theorem B.2.2. *Let* P_n *and* P *be probability measures on* $(\mathcal{C}([0,1]^2), \mathcal{B}_{\mathcal{C}})$*. If the finite-dimensional distribution of* P_n *converges weakly to those of* P *and if* P_n *is tight, than* P_n *converges weakly to* P *for* $n \to \infty$*, denoted by* $P_n \Rightarrow_n P$*.*

Proof. See Park (1971). □

Definition B.2.3. *The modulus of continuity of any* $x \in \mathcal{C}([0,1]^2)$ *is defined by*

$$W_x(\delta) := W(x, \delta) := \sup_{\|(t_1,s_1)-(t_2,s_2)\| \le \delta} |x(t_1, s_1) - x(t_2, s_2)|\,, \ \delta \in (0,1), \qquad \text{(B.2.1)}$$

where $\|\cdot\|$ *denotes Euclidean norm.*

Theorem B.2.4. *(Arzelà-Ascoli theorem)*

A set $A \subseteq \mathcal{C}([0,1]^2)$ *is relatively compact (i.e., its closure is compact), if and only if* (1) $\sup_{x \in A} |x(0,0)| < \infty$ *and* (2) $\lim_{\delta \to 0} \sup_{x \in A} W_x(\delta) = 0$.

Theorem B.2.5. *Let* $\{P_n\}_{n \ge 1}$ *be a sequence of probability measures on* $(\mathcal{C}([0,1]^2), \mathcal{B}_{\mathcal{C}})$*. The sequence* $\{P_n\}_{n \ge 1}$ *is tight, if and only if the following two conditions are fulfilled.*

(1) For every $\eta > 0$*, there exist an* $a > 0$ *and an* $n_0 \in \mathbb{N}$ *such that*

$$P_n\{x : |x(0,0)| > a\} \le \eta, \ \forall n \ge n_0.$$

(2) For every $\eta > 0$ *and every* $\varepsilon > 0$*, there exist a* $\delta \in (0,1)$ *and an* $n_0 \in \mathbb{N}$ *such that*

$$P_n\{x : \ W_x(\delta) > \varepsilon\} \le \eta, \ \forall n \ge n_0. \qquad \text{(B.2.2)}$$

The second condition is equivalent to the following one : for every $\varepsilon > 0$

$$\lim_{\delta \to 0} \limsup_{n \to \infty} P_n\{x : \ W_x(\delta) > \varepsilon\} = 0.$$

Proof. Analogous to the proof of Theorem 7.3 in Billingsley (1999). □

Theorem B.2.6. *Let* X *and* X_n, $n \geq 1$ *be random functions on* $(\mathcal{C}([0,1]^2), \mathcal{B}_{\mathcal{C}})$ *defined on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$. *If for any finitely many distinct points* $(t_1, s_1), \ldots, (t_p, s_q) \in [0,1]^2$, $(X_n(t_1, s_1), \ldots, X_n(t_p, s_q)) \Rightarrow_n (X(t_1, s_1), \ldots, X(t_p, s_q))$, *and* $\lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}\{W(X_n, \delta) \geq \varepsilon\} = 0, \forall \varepsilon > 0$, *then* $X_n \Rightarrow_n X$.

Proof. First proof :

Let P_n and P be the probability measures corresponding to the random functions X_n and X, respectively, $n \geq 1$ and π be the projection defined in Section 2.1. By the hypothesis, $P_n \circ \pi^{-1}_{(0,0)} \Rightarrow_n P \circ \pi^{-1}_{(0,0)}$ which implies $\{P_n \circ \pi^{-1}_{(0,0)}\}_{n \geq 1}$ is tight on $(\mathbb{R}, \mathcal{B}^1)$. Consequently condition (1) in Theorem B.2.5 is satisfied by $\{P_n\}_{n\geq 1}$. In addition, since the condition $\lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}\{W(X_n, \delta) \geq \varepsilon\} = 0, \forall \varepsilon > 0$ is the same with condition (2) in Theorem B.2.5, together we have $\{X_n\}_{n \geq 1}$ is tight. The proof is complete since the finite dimensional distributions of $\{X_n\}_{n \geq 1}$ converges weakly to that of X.

Second proof :

For $u \in \mathbb{N}$, let $\Gamma_{uu} := \{\mathbf{I}_{\ell k} := [(\ell-1)/u, \ell/u] \times [(k-1)/u, k/u] : \ 1 \leq \ell, k \leq u\}$. Let us define a mapping $\mathcal{C}([0,1]^2) \ni x \to M_u x \in \mathcal{C}([0,1]^2)$, given by

$$\begin{aligned}(M_u x)(t,s) := {} & (\ell - ut)(k - us)x((\ell-1)/u, (k-1)/u) \\ & + (ut - (\ell-1))(k - us)x(\ell/u, (k-1)/u) \\ & + (\ell - ut)(us - (k-1))x((\ell-1)/u, k/u) \\ & + (ut - (\ell-1))(us - (k-1))x(\ell/u, k/u), \ \forall (t,s) \in \mathbf{I}_{\ell k},\end{aligned}$$

and a mapping $\mathcal{C}([0,1]^2) \ni x \to \pi_{\Gamma_{uu}} x \in \mathbb{R}^{(u+1)\times(u+1)}$, $\pi_{\Gamma_{uu}} x := (x(\ell/u, k/u))^{u,\ u}_{k=0,\ell=0}$.

Obviously, $\|M_u x - x\|_\infty \le W_x(\sqrt{2}/u)$. Let $L_u : \mathbb{R}^{(u+1)\times(u+1)} \to \mathcal{C}([0,1]^2)$, given by

$$(L_u\mathbf{B})(t,s) := (\ell - ut)(k - us)b_{\ell-1,k-1} + (ut - (\ell-1))(k - us)b_{\ell,k-1} + (\ell - ut)(us - (k-1))b_{\ell-1,k} + (ut - (\ell-1))(us - (k-1))b_{\ell k},$$

for $(t,s) \in \mathbf{I}_{\ell k}$ and $\mathbf{B} = (b_{\ell k})_{\ell=0,k=0}^{u,\ u} \in \mathbb{R}^{(u+1)\times(u+1)}$. For any $\mathbf{B}, \mathbf{D} \in \mathbb{R}^{(u+1)\times(u+1)}$, we have $\|L_u\mathbf{B} - L_u\mathbf{D}\|_\infty \le \max_{0\le i,j\le u} |b_{ij} - d_{ij}| =: \|B - D\|_{\mathbb{R}^{(u+1)\times(u+1)}}$, hence L_u is continuous on $\mathbb{R}^{(u+1)\times(u+1)}$. Moreover, since $M_u x = L_u(\pi_{\Gamma_{uu}} \circ x)$, for any $u \in \mathbb{N}$ and any $x \in \mathcal{C}([0,1]^2)$, by the first hypothesis and the continuous mapping theorem, we get $M_u X_n = L_u(\pi_{\Gamma_{uu}} \circ X_n) \Rightarrow_n L_u(\pi_{\Gamma_{uu}} \circ X) = M_u X$. Furthermore, since $\|M_u X - X\|_\infty \le W(X, \sqrt{2}/u) \xrightarrow{u\to\infty} 0$, we further have $M_u X \Rightarrow_u X$, see Theorem 3.1 in Billingsley (1999). Thus, together we get $M_u X_n \Rightarrow_n M_u X \Rightarrow_u X$. In addition, the inequality $\|M_u X_n - X_n\|_\infty \le W(X_n, \sqrt{2}/u)$ together with the second hypothesis lead us to the following result

$$\lim_{u\to\infty} \limsup_{n\to\infty} \mathbb{P}\{\|M_u X_n - X_n\|_\infty \ge \varepsilon\} \le \lim_{u\to\infty} \limsup_{n\to\infty} \mathbb{P}\{W(X_n, \sqrt{2}/u) \ge \varepsilon\} = 0, \ \forall \varepsilon > 0.$$

Thus, the assertion follows by applying Theorem 3.2 in Billingsley (1999). □

Proposition B.2.7. *Let* $\{\mathbf{I}_{\ell k} := [t_{\ell-1}, t_\ell] \times [s_{k-1}, s_k] : 1 \le \ell \le p, 1 \le k \le q\}$, *with* $0 = t_0 < t_1 < \ldots < t_p = 1$, $0 = s_0 < s_1 < \ldots < s_q = 1$, *and* $\min_{1<\ell<p}(t_\ell - t_{\ell-1}) \ge \delta$, $\min_{1<k<q}(s_k - s_{k-1}) \ge \delta$, $\delta \in (0,1)$. *Then for every* $x \in \mathcal{C}([0,1]^2)$, *we have*

$$W_x(\delta\sqrt{2}) \le 3 \max_{\substack{1\le\ell\le p\\ 1\le k\le q}} \sup_{(t,s)\in\mathbf{I}_{\ell k}} |x(t,s) - x(t_{\ell-1}, s_{k-1})| . \tag{B.2.3}$$

For any probability measure P *on* $(\mathcal{C}([0,1]^2, \mathcal{B}_\mathcal{C}))$ *and any* $\varepsilon > 0$, *it holds*

$$P\left\{W_x(\delta\sqrt{2}) \ge 3\varepsilon\right\} \le \sum_{k=1}^{q}\sum_{\ell=1}^{p} P\left\{\sup_{(t,s)\in\mathbf{I}_{\ell k}} |x(t,s) - x(t_{\ell-1}, s_{k-1})| \ge \varepsilon\right\}. \tag{B.2.4}$$

B.3 Proof of Theorem 2.3.4

Proof. We shall prove this theorem by showing that the sequence $\{\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(\cdot)\}_{n\geq 1}$ satisfies the sufficient conditions of Theorem B.2.6 : for any choice of distinct points $(t_1,s_1),\ldots,(t_p,s_q)\in[0,1]^2$ and any $p,q\in\mathbb{N}$, $\mu_n\circ\pi^{-1}_{(t_1,s_1),\ldots,(t_p,s_q)}\Rightarrow_n \mathbb{W}\circ\pi^{-1}_{(t_1,s_1),\ldots,(t_p,s_q)}$, i.e.,

$$(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(t_1,s_1),\ldots,\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(t_p,s_q))\Rightarrow_n (B_2(t_1,s_1),\ldots,B_2(t_p,s_q)), \quad \text{(B.3.1)}$$

and for every $\varepsilon>0$,

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\mathbb{P}\{W(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n}),\delta)\geq\varepsilon\}=0. \quad \text{(B.3.2)}$$

Consider first a single point $(t_1,s_1)\in[0,1]^2$, by (3) of Corollary 2.2.2, we have

$$\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(t_1,s_1)\Rightarrow_n B_2(t_1,s_1).$$

Let us consider any two distinct points (z_1,z_2) and $(z_1',z_2')\in[0,1]^2$, and suppose firstly that $z_1<z_1'$, and $z_2<z_2'$. For each vector $(\alpha_1,\alpha_2)\in\mathbb{R}^2$, we get

$$\begin{aligned}\frac{\alpha_1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2)+&\frac{\alpha_2}{\sigma}(\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1',z_2')-\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2))\\ &\Rightarrow_n \alpha_1\mathcal{N}(0,z_1z_2)+\alpha_2(\mathcal{N}(0,z_1'z_2'-z_1z_2)).\end{aligned}$$

Hence by the *Cramér-Wald technique*, see Billingsley (1995), p. 382, we have

$$\begin{aligned}(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2),&\frac{1}{\sigma}(\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1',z_2')-\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2)))\\ &\Rightarrow_n (\mathcal{N}(0,z_1z_2),\mathcal{N}(0,z_1'z_2'-z_1z_2)).\end{aligned}$$

Furthermore, since

$$\begin{aligned}(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2),\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1',z_2'))=(&\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2),\frac{1}{\sigma}(\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1',z_2')\\ &-\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2)+\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2))),\end{aligned}$$

then by the preceding result, we get

$$(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2),\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1',z_2'))\Rightarrow_n(\mathcal{N}(0,z_1z_2),\mathcal{N}(0,z_1'z_2')),$$

where the right side of this result is the distribution of $(B_2(z_1,z_2),B_2(z_1',z_2'))$.

Now let us consider the case $z_1'<z_1$ and $z_2<z_2'$. Then we get

$$\begin{aligned}&(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1',z_2'),\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2))\\ =&((\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1',z_2')-\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1',z_2))+\frac{1}{\sigma}\mathbf{T}_n(\varepsilon_n)(z_1',z_2),\\ &(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1,z_2)-\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1',z_2))+\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(z_1',z_2))\end{aligned}$$

which by using the analogous argument as before, converges in distribution to the vector $(\mathcal{N}(0,z_1'z_2'),\mathcal{N}(0,z_1z_2))$. The other cases can be handled analogously. Thus the assertion for $k=2$ follows. A set of three or more points can be handled in the same way, this leads us to the conclusion that (B.3.1) is satisfied.

To prove (B.3.2) we apply Proposition B.2.7. From (B.2.4), for any $\varepsilon>0$ we get

$$\begin{aligned}&\mathbb{P}\left\{W(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n}),\delta\sqrt{2})\geq 3\varepsilon\right\}\\ &\leq\sum_{k=1}^{q}\sum_{\ell=1}^{p}\mathbb{P}\left\{\sup_{(t,s)\in\mathbf{I}_{\ell k}}\left|\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(t,s)-\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n})(t_{\ell-1},s_{k-1})\right|\geq\varepsilon\right\},\end{aligned}$$

whenever $\min_{1<\ell<p}(t_\ell-t_{\ell-1})\geq\delta$, $\min_{1<k<q}(s_k-s_{k-1})\geq\delta$, $\delta\in(0,1)$. For $0\leq\ell\leq p$, and $0\leq k\leq q$, let us chose $t_\ell=m_\ell/n$, and $s_k=m_k'/n$, where m_ℓ and m_k' are integers that satisfy the condition

$$\begin{aligned}&0=m_0<m_1<\ldots<m_{\ell-1}<m_\ell<\ldots<m_p=n,\\ &0=m_0'<m_1'<\ldots<m_{\ell-1}'<m_\ell'<\ldots<m_q'=n.\end{aligned}$$

Then by the definition of $\mathbf{T}_n(\mathbf{E}_{n\times n})(\cdot)$, we get

$$\mathbb{P}\left\{W(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n}),\delta\sqrt{2})\geq 3\varepsilon\right\}\leq\sum_{k=1}^{q}\sum_{\ell=1}^{p}\mathbb{P}\left\{\max_{\substack{m_{\ell-1}\leq i_1\leq m_\ell\\ m'_{\ell-1}\leq i_2\leq m'_\ell}}\left|S_{i_1i_2}-S_{m_{\ell-1}m'_{k-1}}\right|\geq\varepsilon\sigma n\right\},$$

whenever

$$\frac{m_\ell}{n}-\frac{m_{\ell-1}}{n}\geq\delta,\text{ and }\frac{m'_\ell}{n}-\frac{m'_{\ell-1}}{n}\geq\delta,\text{ for }1<\ell<p,1<k<q,$$

where $S_{\ell k}:=\sum_{j=1}^{k}\sum_{i=1}^{\ell}\varepsilon_{ij}$. Since $\{\varepsilon_{ij}:1\leq i,j\leq n\}$ are i.i.d, the right side of the preceding inequality is the same with the following one

$$\sum_{k=1}^{q}\sum_{\ell=1}^{p}\mathbb{P}\left\{\max_{\substack{0\leq i_1\leq(m_\ell-m_{\ell-1})\\ 0\leq i_2\leq(m'_\ell-m'_{\ell-1})}}|S_{i_1i_2}|\geq\varepsilon\sigma n\right\}.$$

Hence, we have

$$\mathbb{P}\left\{W(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n}),\delta\sqrt{2})\geq 3\varepsilon\right\}\leq\sum_{k=1}^{q}\sum_{\ell=1}^{p}\mathbb{P}\left\{\max_{\substack{0\leq i_1\leq(m_\ell-m_{\ell-1})\\ 0\leq i_2\leq(m'_\ell-m'_{\ell-1})}}|S_{i_1i_2}|\geq\varepsilon\sigma n\right\}.$$

For further simplification we chose $m_\ell=\ell m$ and $m'_k=km'$, for some integers m and m' that satisfy $m_\ell-m_{\ell-1}=m\geq n\delta$ and $m'_k-m'_{k-1}=m'\geq n\delta$, for $0\leq\ell<p$ and $0\leq k<q$. Since the indexes p and q must satisfy $(p-1)m<n\leq pm$ and $(q-1)m'<n\leq pm'$, we chose $p=\lceil n/m\rceil\overset{n\to\infty}{\longrightarrow}1/\delta<2/\delta$ and $q=\lceil n/m'\rceil\overset{n\to\infty}{\longrightarrow}1/\delta<2/\delta$. Moreover, $n/m\overset{n\to\infty}{\longrightarrow}1/\delta>1/2\delta$ and $n/m'\overset{n\to\infty}{\longrightarrow}1/\delta>1/2\delta$. Hence, for large n and for every $\varepsilon>0$, we have

$$\begin{aligned}\mathbb{P}\left\{W(\frac{1}{\sigma}\mathbf{T}_n(\mathbf{E}_{n\times n}),\delta\sqrt{2})\geq 3\varepsilon\right\}\leq&\frac{4}{\delta^2}\,\mathbb{P}\left\{\max_{\substack{0\leq\ell\leq m\\0\leq k\leq m'}}|S_{\ell k}|\geq\frac{\varepsilon\sigma\sqrt{mm'}}{2\delta}\right\}\\=&\frac{16\lambda^2}{\varepsilon^2}\,\mathbb{P}\left\{\max_{\substack{0\leq\ell\leq m\\0\leq k\leq m'}}|S_{\ell k}|\geq\lambda\sigma\sqrt{mm'}\right\}\\\leq&\frac{48\lambda^2}{\varepsilon^2}\max_{\substack{0\leq\ell\leq m\\0\leq k\leq m'}}\mathbb{P}\left\{|S_{\ell k}|\geq\frac{\lambda\sigma\sqrt{mm'}}{3}\right\}\qquad\text{(B.3.3)}\end{aligned}$$

where $\lambda := \varepsilon/2\delta$. The last inequality follows by applying Etemadi's inequality. Thus, (B.3.2) will follow if we can show that

$$\lim_{\lambda\to\infty}\limsup_{m,m'\to\infty}\frac{48\lambda^2}{\varepsilon^2}\max_{\substack{0\le\ell\le m\\0\le k\le m'}}\mathbb{P}\left\{|S_{\ell k}|\ge\frac{\lambda\sigma\sqrt{mm'}}{3}\right\}=0,$$

which is the same with the condition

$$\lim_{\lambda\to\infty}\limsup_{n\to\infty}\frac{48\lambda^2}{\varepsilon^2}\max_{\substack{0\le\ell\le n\\0\le k\le n}}\mathbb{P}\left\{|S_{\ell k}|\ge\frac{\lambda\sigma n}{3}\right\}=0.$$

By the central limit theorem, if ℓ_λ and k_λ are large enough such that $\ell_\lambda\le\ell\le n$ and $k_\lambda\le k\le n$, then

$$\mathbb{P}\left\{|S_{\ell k}|\ge\frac{\lambda\sigma n}{3}\right\}\le\mathbb{P}\left\{|S_{\ell k}|\ge\frac{\lambda\sigma\sqrt{\ell k}}{3}\right\}=\mathbb{P}\left\{\left|\frac{S_{\ell k}}{\sigma\sqrt{\ell k}}\right|\ge\frac{\lambda}{3}\right\}\le\frac{243}{\lambda^4}.$$

For $\ell\le\ell_\lambda\le n$ and $k\le k_\lambda\le n$, we can apply Chebyshev's inequality to get

$$\mathbb{P}\left\{|S_{\ell k}|\ge\frac{\lambda\sigma n}{3}\right\}\le\frac{9\ell k}{\lambda^2n^2}\le\frac{9\ell_\lambda k_\lambda}{\lambda^2n^2}.$$

As a result, the maximum on the right side of (B.3.3) is dominated by $\max\left\{\frac{243}{\lambda^4},\frac{9\ell_\lambda k_\lambda}{\lambda^2n^2}\right\}$. Consequently, we have

$$\lim_{\lambda\to\infty}\limsup_{n\to\infty}\frac{48\lambda^2}{\varepsilon^2}\max_{\substack{0\le\ell\le n\\0\le k\le n}}\mathbb{P}\left\{|S_{\ell k}|\ge\frac{\lambda\sigma n}{3}\right\}\le\frac{48}{\varepsilon^2}\lim_{\lambda\to\infty}\limsup_{n\to\infty}\max\left\{\frac{243}{\lambda^2},\frac{9\ell_\lambda k_\lambda}{n^2}\right\}=0.$$

This complete the proof of the theorem. □

Symbols

Bibliography

[1] Adams, C,R. and Clarkson, J.A. (1934). Properties of function $f(x, y)$ of bounded variation. *Transactions of the American Mathematical Society*, **36**, no. 4, 711-730.

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

[3] Bader, G. (1997). *Konvergenzarten endlicher Borelmaße basierend auf schwacher Konvergenz* (Diplomarbeit). Fakultät für Mathematik der Universität Karlsruhe.

[4] Bader, G. (2001). *Asymptotik von Regressionsmodellen* (Dissertation). Fakultät für Mathematik der Universität Karlsruhe.

[5] Bartle, R.G. (1976). *The Elements of Real Analysis* (2nd. edition). John Wiley & Sons, Inc., New York, London, Sydney, Toronto.

[6] Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley & Sons, Inc., New York, Chichester, Weinheim, Brisbane, Singapore, Toronto.

[7] Billingsley, P. (1995). *Probability and Measure* (third edition). John Wiley & Sons, Inc., New York, Chichester, Weinheim, Brisbane, Singapore, Toronto.

[8] Billingsley, P. (1999). *Convergence of Probability Measures* (2nd. edition). John Wiley & Sons, Inc., New York, Chichester, Weinheim, Brisbane, Singapore, Toronto.

[9] Bischoff, W. (1998). A functional central limit theorem for regression models. *Ann. Stat.* **26**, no. 4, 1398-1410.

[10] Bischoff, W. and Miller, F. (2000). Asymptotically optimal tests and optimal designs for testing the mean in regression models with applications to change-point problems. *Ann. Inst. Statist. Math.* **52**, 658-679.

[11] Bischoff, W. (2002). The structure of residual partial sums limit processes of linear regression models. *Theory of Stochastic Processes.* **2**(**24**), no. 1-2, 23-28.

[12] Bischoff, W., Hashorva, E., Hüsler, J. and Miller, F. (2003a). Exact asymptotics for boundary crossings of the Brownian bridge with trend with application to the Kolmogorov test. *Ann. Inst. Statist. Math.* **55**, no. 4, 849-864.

[13] Bischoff, W., Hashorva, E., Hüsler, J. and Miller, F. (2003b). Asymptotic of a boundary crossing probability of a Brownian bridge with general trend. *Methodology and Computing in Applied Probability.* **5**, 271-287.

[14] Bischoff, W. and Hashorva, E. (2005). A lower bound for boundary crossing probabilities of Brownian bridge/motion with trend. *Statistics and Probability Letters*, **74**, issue 3, 265-271.

[15] Bisgaard, S. and Ankenman, B. (1996). Standard errors for the eigenvalues in second-order response surface models. *Technometrics*, **38**, no. 3, 238-246.

[16] Christensen, R. (1991). *Linear Models for Multivariate, Time Series, and Spatial Data.* Springer-Verlag, New York, Berlin, Heidelberg, London, Paris, Tokyo.

[17] Clarkson, J.A. (1932). On double Riemann-Stieltjes integrals. *Bull. Am. Math. Soc.* **39**, 929-936.

[18] Clarkson, J.A. and Adams, C,R. (1933). On definition of Bounded Variation for functions of two variables. *Transactions of the American Mathematical Society*, **35**, no. 4, 824-854.

[19] Cohn, D.L. (1993). *Measure Theory.* Birkhäuser. Boston, Basel, Berlin.

[20] Cressie, Noel A.C. (1993). *Statistics for Spatial Data* (revised edition). John Wiley & Sons, Inc., New York, Chichester, Toronto, Brisbane, Singapore.

[21] Douglass, S.A. (1996). *Introduction to Mathematical Analysis.* Addison-Wesley Publishing Company, Inc., Califonia, New York, Amsterdam, Bonn, Sydney.

[22] Elstrodt, J. (2005). *Maß-und Integrationstheorie* (vierte, korrigierte Auflage). Springer-Verlag, Berlin, Heidelberg.

[23] Eastwood, V.R. (1993). Some nonparametric methods for changepoint problems. *The Canadian Journal of Statistics*, **21**, no. 2, 209-222.

[24] Gentle, J.E. (1998). *Numerical Linear Algebra for Applications in Statistics.* Springer-Verlag New York Inc., Berlin, Heidelberg, London, Paris, Tokyo.

[25] Harville, D.A. (1997). *Matrix Algebra from a Statistician's Perspective.* Springer-Verlag New York Inc., Berlin, Heidelberg, London, Paris, Tokyo.

[26] Högnäs, G. (1977). Characterization of weak convergence of signed measures on $[0, 1]$. *Math. Scand*, **41**, 175-184.

[27] Imhof, J.P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, **48**, no. 3 and 4, 419-426.

[28] Jandhala, V.K. and Minogue, C.D. (1993). Distributions of Bayes-type change-point statistics under polynomial regression. *Journal of Statistical Planing and Inference*, **37**, 271-290.

[29] Johnson, J. (1985). An elementary characterization of weak convergence of measures. *The American Mathematical Monthly*, **92**, no. 2, 136-137.

[30] Kuelbs, J. (1968). The invariance principle for a lattice of random variables. *The Ann. of Math. Stat.*, **39**, no. 2, 382-389.

[31] Lehmann, E.L. and Romano, J.P. (2005). *Testing Statitical Hypotheses* (3th. edition). Springer Science+Business Media, Inc., New York.

[32] Lifshits, M.A. (1996). *Gaussian Random Functions*. Kluwer Academic Publishers, Dordrecht, Boston, London.

[33] Luxemburg, W.A.J. (1971). Arzelà's domonated convergence theorem for the Riemann inetegral. *Amer. Math. Monthly*, **78**, 970-979.

[34] MacNeill, I.B. (1978a). Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times. *Ann. Statist.*, **6**, 422-433.

[35] MacNeill, I.B. (1978b). Limit processes for secuences of partial sums of regression residuals. *Ann. Prob.*, **6**, 695-698.

[36] 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., 298-306.

[37] Móricz, F. (2006). Pointwise behavior of double Fourier series of functions of bounded variation. *Monatsh. Math.* **148**, 51-59.

[38] Muirhead, R.J. (1982). *Aspect of Multivariate Statistical Theory*. John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore.

[39] Xie, L. and MacNeill, I.B. (2004). Spatial residual processes and boundary detection (Preprint). Department of Statistics and Actuarial Sciences, University of Western Ontario, Canada.

[40] Park, W.J. (1970). A multi-parameter gaussian process. *The Ann. of Math. Stat.*, **41**, no. 5, 1582-1595.

[41] Park, W.J. (1971). Weak convergence of probability measures on the function space $\mathcal{C}([0,1]^2$. *J. of Multivariate Analysis*, **1**, 433-444.

[42] Rudin, W. (1991). *Functional Analysis* (2nd. edition). McGraw-Hill, Inc., New York.

[43] Schabenberger, O., Gotway, C.A. (2005). *Statistical Methods for Spatial Data Analysis*. Chapman & Hall/CRC, Boca-Raton, London, New York.

[44] Shorack, G.R., Wellner, J.A. (1986). *Empirical Processes with Applications to Statistics*. John Wiley & Sons Inc., New York, Chichester, Brisbane, Toronto, Singapore.

[45] Smirnov, W.I. (1969). *Lehrgang der höheren Mathematik* (teil V). Veb. Deutscher Verlag der Wissenschaften, Berlin.

[46] Stapleton, J. H. (1995). *Linear Statistical Models*. John Wiley & Sons Inc., New York, Chichester, Brisbane, Toronto, Singapore.

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

[48] Werner, D. (2005). *Funktionalanalysis* (fünfte erweiterte Auflage). Springer-Verlag, Berlin, Heidelberg.

[49] Yeh, J. (1960). Wiener measure in a space of functions of two variables. *Trans. Amer. Math. Soc.*, **95**, 433-450.

[50] Yeh, J. (1963). Cameron-Martin translation theorems in the Wiener space of functions of two variables. *Trans. Amer. Math. Soc.*, **107**, no. 3, 409-420.

[51] Yeh, J. (1973). *Stochastic Processes and the Wiener Integral*. Marcel Dekker, Inc., New York.

[52] Young, W.H. (1914). On integration with respect to a function of bounded variation. *Proc. London Math. Soc.*, **2**, no. 13, 109-150.

[53] Young, W.H. (1917a). On multiple integral. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, **93**, no. 647, 28-41.

[54] Young, W.H. (1917b). On multiple integration by parts and the second theorem of mean. *The Ann. of Math. Stat.*, **43**, no. 4, 1235-1246.

[55] Zimmerman, G.J. (1972). Some sample function properties of the two-parameter gaussian process. *The Ann. of Math. Stat.*, **43**, no. 4, 1235-1246.

Lebenslauf

Persönliche Daten

Name :	Wayan Somayasa
Geburtsdatum, ort :	01.12.1971, Jinengdalem, Buleleng
Zivilstand :	Verheiratet

Schulausbildung

1978 - 1984	Grundschule (SDN 3 Jinengdalem, Buleleng)
1984 - 1987	Mittelschule (SMPN 5 Singaraja, Bali)
1987 - 1990	Mittelhochschule (SMAN 1 Bau-Bau, Buton)

Studium

1991 - 1992	Lehramt Physik an der Univ. Haluoleo Kendari
1992 - 1998	Bachelor in Mathematik an der Univ. Gadjah Mada
1998 - 2001	Master in Mathematik an der Univ. Gadjah Mada
April 2003 - März 2007	DAAD Stipendiat
Seit Sept. 2003	Doktorand am Inst. für Stochastik, Univ. Karlsruhe

Berufstätigkeit

Juli 1990 - Juni 1991	Praktikum bei der PT. S.P., Buton (Indonesien)
Juli 1997 - Okt. 1997	Sozialdienst in Kulonprogo-Java (Indonesien)
März 1999 - März 2003	Wissenschaftlicher Mitarbeiter an der Univ. Haluoleo
Sept. 2003 - Feb. 2005	Wissenschaftlicher Hilfskraft an der Univ. Karlsruhe
April 2007 - Sept.2007	Wissenschaftlicher Mitarbeiter an der Univ. Karlsruhe