Asymptotics for survey data based on spatial processes

Leshun Xu

A thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy in Statistics

Department of Statistics
University of Auckland
2017

In memory of my academic advisor

Alastair John Scott

Abstract

This thesis is the beginning of research into indicated spatial processes for analysing complex survey data. When modelling practical problems, it is widely accepted that a spatial process is generally described by a random field. However, before survey methods can be applied, research into the asymptotic properties of estimators, particularly estimators based upon random fields, is essential.

In this thesis, we introduce a new survey sampling method which has a potentially wide application, and is described by an indicated sampling method. This sampling strategy is appropriate for setting up estimators such as Horvitz-Thompson estimator and others. We also introduce some assumptions about the spatial structures of a population, developed from examining real situations. Based on this method and these assumptions, we develop central limit theorems, functional central limit theorems, and consistent estimators of variances on non-stationary dependent random fields.

Since it is important to understand the asymptotics of a new complex survey method, for this indicated sampling method, we consider central limit theorems with the assumption of conditional independence properties. In some results, we assume the partial derivatives of functions defining estimators are bounded. This assumption gives an opportunity to apply the Mean Value Theorem to the consideration of asymptotics.

With our new insight into the factor m^{d-1} in the assumptions of the central limit theorem, Theorem 3.3.1, by Guyon in 1995, we see that this is essentially an estimation of the number of the pairs which share the same dependencies. We therefore introduce a function h to stand for the number of pairs with the same dependencies. Then, with an additional assumption on the joint-blocks spatial structures, we prove the L_2 consistency for the estimators of the variance of the population. We then generalize the results on estimating the variance by Carlstein in 1986 and Fuller's central limit theorem, Theorem 1.3.2, in his book in 2011.

It is rare to see functional central limit theorems on non-stationary dependent random fields. This is because it is hard to verify the tightness in the high dimensional random fields. By using two criteria introduced by Billingsley in 1968, one of which, Theorem 15.6, is rarely

used by other scholars, in addition to the assumptions on the nested spatial structures and the proper estimation of the fourth moment of the sample sum, we provide some original results on functional central limit theorems, where the estimation of the fourth moment develops Rio's result, Theorem 2.1, in his report in 2013.

Acknowledgements

I acknowledge the support from The Marsden Fund, The University of Auckland Doctoral Scholarship and The Faculty Awards of The Department of Statistics. I would also like to thank the open sources websites, which helped me a great deal in my studies, such as r-project.org, Google Scholar, Wikipedia and many university websites. I thank everyone and everything I experienced in the past five years, which will stay in my mind forever as beautiful memories. Especially, I appreciate Dr Rosemary K Barraclough for proofreading this thesis. I appreciate my examiners, Associate Professor Alexandra Schmidt, Professor Soumendra Lahiri and the Head of Department nominee, Dr Jesse Goodman, for their important comments, questions and suggestions on my final corrections. Especially, I thank Prof Lahiri for his crucial comments and questions on the variance estimation technique based on overlapping re-sample methods.

I gratefully appreciate my advisory committees, Professor Alan Lee, Professor Thomas Lumley, Professor Chris Wild and Emeritus Professor Alastair Scott (1939–2017) for their excellent guidance and important help on my research. I appreciate the warm welcome, great encouragement and invaluable suggestions from my co-supervisor, Professor Chris Wild, and my academic advisor, Emeritus Professor Alastair Scott.

Especially, I thank my supervisor Professor Alan Lee and co-supervisor Professor Thomas Lumley. They brought me into a wonderful research field, which is very important in both statistical theory and statistical applications. I thank Thomas for his directional guidance and crucial help on my research project. I thank Alan for his bright wisdom and great patience in our regular meetings, which have been sustained for so many years. I am deeply moved by his influence and invaluable help.

I am greatly indebted to the love from my family.

Contents

Chapter 1

Introduction

Spatial statistics and asymptotics of survey data are extensive research fields with comprehensive applications. Advanced complex survey techniques and modern data processing skills are required to tackle many practical problems. The asymptotics studied in this thesis are a crucial first step in performing complex surveys. This chapter mainly answers *what* and *why* questions on the research presented in this thesis. An indicated sampling strategy is introduced for the preparation of the discussion in the following chapters, which are mainly relative to *how* we achieve the goals of asymptotics.

1.1 *What is a survey?* and *Why asymptotics?*

Nowadays it is a quite common idea to conduct a survey to effectively use the vast information and data available, for all kinds of purposes. To keep the pace with the times, the definition of survey was updated by Fritz Scheuren from the first version in 1980, to the second version published in the American Statistical Association, 2004.

> *"Today the word 'survey' is used most often to describe a method of gathering information from a sample of individuals."*

This is the definition used in his publication [65]. Scheuren suggests six stages for conducting surveys, which are introduced by Blair et al in the book [6] and its older editions. The six stages of a survey in [65] are adapted as a five-stage survey in [6], which is described in Figure 1.1.

There are four branches of survey design in the horizontal direction throughout five stages in Figure 1.1: sample related, questionnaire related, operation plan related and analysis plan related. Our research is with respect to the first and the last branches of "Design Survey" and "Stage 5".

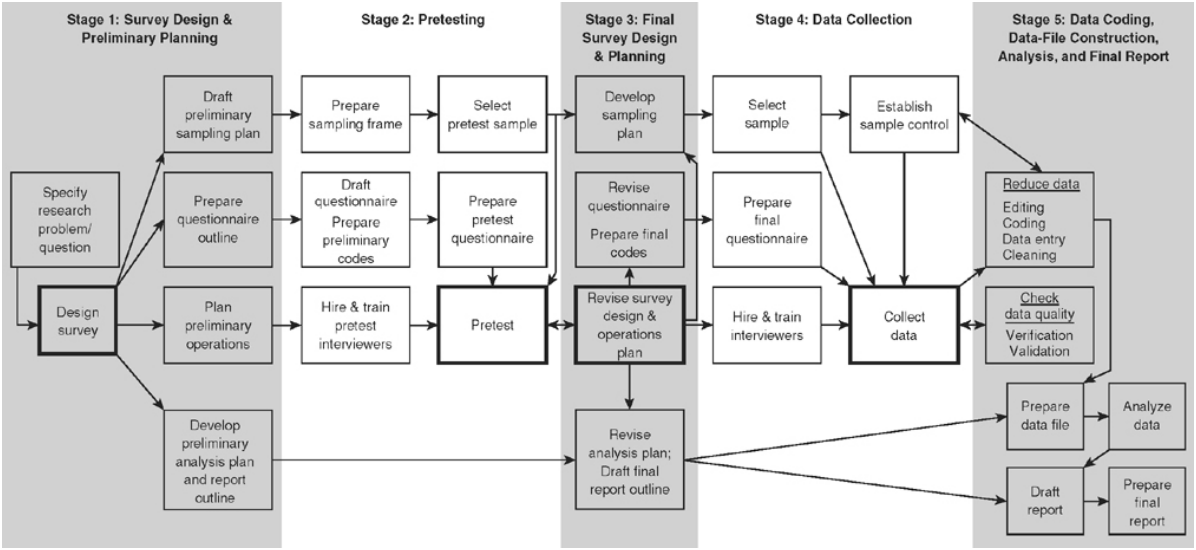


Figure 1.1: Five stages of a survey (Exhibit 3.1 in [6])

When we analysis data as per in the lowest branch and in Stage 5 (Figure 1.1), a basic setting should be taken into account, i.e. the role of randomization in surveys. This mainly involves two sampling strategies, a design-based strategy and a model-based strategy. In survey theory, both require asymptotic properties to ensure that their estimated results reflect the population. Therefore, in this case, the finite or infinite population from which the data derives is regarded as the *superpopulation*[1] in asymptotic statistics.

Generally speaking, the design-based strategy is based on a finite population for description purposes. The population is considered as fixed and all the variation is from the sampling mechanism. There is no randomization associated with the population.

The model-based strategy uses a superpopulation for the purpose of analysis. Here, the superpopulation is modelled by random variables. Its variation is impacted by the model and the sampling mechanism. The asymptotic properties of a model-based strategy are connected to its infinite population.

Probability sampling plays an important role in finite population sampling (see [54]). Although the model-based methods are based on statistical models, and the design-based methods are based on probability sampling principles[2], in most of the practical cases of complex surveys, we have to introduce a suitable sampling design to an on-going model-based frame. On the other hand, for a design-based frame, it is also possible to involve the modelled population or some auxiliary information with randomization properties.

For more details of model-based and design-based methods we suggest the book [14] by Chambers

[1] The superpopulation can be regarded as an infinite or finite population being generated theoretically from an infinite population. For more details of this definition, we recommend some books such as [14], [26] and [44].

[2] Please refer to the "Preface of Handbook 29A" by Danny Pfeffermann and C. R. Rao, on page vi, of [53].

and Skinner, and papers such as the discussion on the evaluation of model-based inference and design-based inference in [37], the comparison between model-based theory and design-based theory with some examples in [64], and the discussion on the hybrid framework in [69]. In this thesis we work with the superpopulation which is modelled by random variables. At the same time, some auxiliary information is also taken into account in the sample design.

Asymptotic properties describe approximation of statistical procedures. A. W. van der Varrt explained why asymptotic statistics in his book [70]:

> *"Why asymptotic statistics? The use of asymptotic approximations is twofold. First, they enable us to find approximate tests and confidence regions. Second, approximations can be used theoretically to study the quality (efficiency) of statistical procedures"*

There are many aspects of asymptotic statistics, such as the law of large numbers, central limit theorems, functional central limit theorems, and both the efficiency and consistency of estimators. Asymptotic normality is an extremely important property in this field. Therefore in this thesis, we mainly consider central limit theorems (CLTs), functional central limit theorems (FCLTs) and the consistency of the estimators of variance for general sampling strategies. All results are based on spatial processes.

1.2 Spatial processes in complex surveys

A spatial process is the generalization of a 1-dimensional process. Since many practical problems arise from high dimensional space, and a 1-dimensional space can be regarded as a degeneration of high dimensional one, spatial processes are widely accepted in modelling problems. As a further consideration, we endow a spatial process with dependent properties, which are more general than independent assumptions in surveys. In this section, we give two examples. One is used to show the potential dependence in practical problems, the other is a complex, multi-stage survey.

Example 1: The Centers for Disease Control and Prevention (CDC) is an important institute in the United States. It conducts many surveys each year. In 2017, the CDC produced the result of a survey for the U.S. diabetes data (see [24]). The goal was to estimate the percentage of adults at least twenty years of age with diagnosed diabetes. To map the percentages, three years of data from the Behavioral Risk Factor Surveillance System and the U.S. Census Bureau Population Estimates Program were used to improve the precision. In Figure 1.2, the gradually changing colours imply there are dependencies between those counties, and these dependences are likely related to the distance between geographical locations. This implies that if we take this kind of spatial dependent structure into account, it will be possible to improve estimates

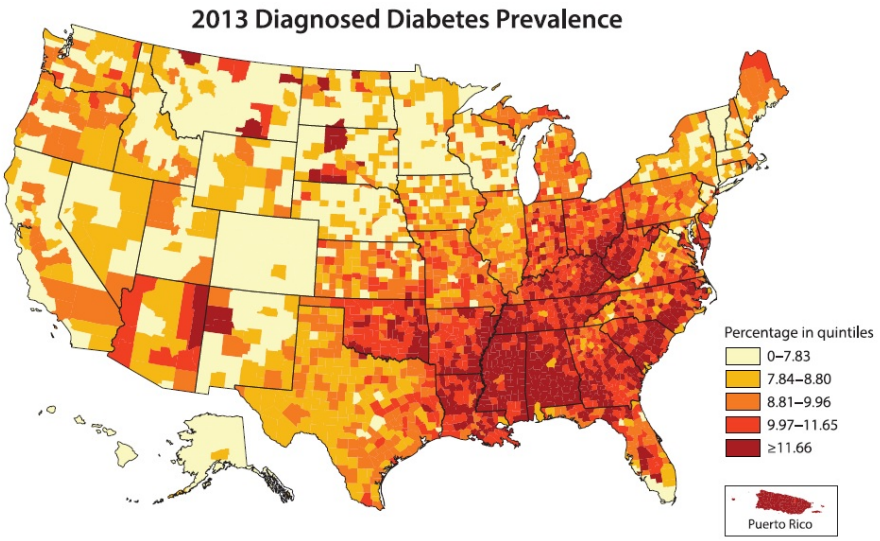


Figure 1.2: Spatial correlation pattern in diabetes data [24]

and reduce cost.

Example 2: : The example which is used to illustrate a complex survey in space is the National Health and Nutrition Examination Surveys (NHANES) conducted annually by the CDC in the United States. Four things that NHANES 1999-2010 (see [75]) hoped to achieve are:

> "To provide prevalence data on selected diseases and risk factors for the U.S. population; To monitor trends in selected diseases, behaviors, and environmental exposures; To explore emerging public health needs; To maintain a national probability sample of baseline information on health and nutritional status."

Data on health, nutritional status, and health behaviours, were collected from across the population of the United States. Three survey methods (in-person interviews, face-to-face interviews, and physical examinations) were used in the participants homes and at a mobile examination centre (MEC). Specifically, they administrated 48 items in a four part household interview, gathered 17 variables from a medical examination, and conducted 2 extra interviews at a MEC.

There were four stages in the NHANES complex, multistage probability sampling design, where the sample weights were also introduced. The first stage was the selection of primary sampling units (PSUs). About 30 counties were selected out of about 3000 counties. The second stage was the selection of segments within each PSU: All households in a relative small region were selected into the sample. The third stage was the selection of specific households in each segment. The last stage is the selection of individuals in a household. Then around 10,000 persons were selected and visited in a 2-year survey.

As it is shown in Figure A1 in Appendix, this kind of four stage sample design works through three steps in NHANES 2011-2014 [36]. To achieve a broad description, the U.S. population was arranged with domains and subdomains, i.e. for each sex-age group, there are four subdomains, non-Hispanic Black persons, non-Hispanic non-Black Asian, Hispanic, and non-Hispanic White and other, which is divided by non-low income group and low income group. To set up a clustering criteria on dividing PSUs, the measure of size (MOS) of a PSU was introduced for a self-weighting sample. The MOS of a PSU, which was indexed by h, was defined as

$$M_h = \sum_k A_k C_{hk}, \quad A_k = \sum_l r_{kl} \frac{C^*_{.kl}}{C^*_{.k.}},$$

where k was the race-Hispanic origin-income subdomain, l was the sex-age subdomain, C_{hk} was the most recent population estimate for race-Hispanic origin-income subdomain k in PSU h, r_{kl} was the sampling rate of persons in the (k, l)-th race-Hispanic origin-income-sex-age subdomain, $C^*_{.kl}$ was the most recent projection of the 2008 total population count for race-Hispanic origin-income-sex-age subdomain (k, l), $C^*_{.k.}$ was the most recent projection of the 2008 total population count for race-Hispanic origin-income subdomain k. Then the criteria of separating PSUs into certainty PSUs and noncertainty PSUs was 75% of the initial sampling interval, i.e. (see [36])

$$0.75 \frac{\sum_{h=1}^{H} M_h}{60},$$

where H was the number of PSUs in the whole sampling frame. Figure A2, in Appendix, shows how this criteria works in conducting the survey in the first stage. Here we omit the details of the rest of the first stage and the remaining three stages of the NHANES' sample design.

In NHANES, the geographical location of the population determines that a spatial process should be a well-chosen model. Furthermore, it is reasonable that we make the following three assumptions:

Firstly, the measurements of the sample are influenced by some random effects. For example, the health issues, the sample was clearly affected by M_h. M_h was with respect to C_{hk}, r_{kl}, $C^*_{.kl}$, which were all related to income. Therefore if we imagine a low-income leads a low nutrition status, it is plausible the sample is driven by incomes to some extent. Furthermore, NHANES 1999–2010 [75] has shown whether an MEC can move in could be decided by the traffic and community status. It also implies the measurements could be affected by some other random factors.

Secondly, the sample of the population is affected by some random information. For example, we suppose a person, i, is indicated by a random variable R_i, i.e. R_i decides whether i will be observed. The probability of $\mathbb{P}(R_i = 1)$ could be related to the incomes since the MOS was designed to reflect two groups of non-low income and low income.

Lastly, it is reasonable to make an assumption on dependence, which could be similar to that shown in Figure 1.2. We can endow some potential dependence between PSUs, between strata, or even between persons and the measurements. For example, the people could have the same health problems if they live closed to each other.

1.3 An indicated sampling strategy

Using a grid index

Throughout this thesis, we assume that all sample points are located in a grid space, which is a spatial process indexed by high dimensional integer points. This assumption is widely accepted in the research of random fields, and it is also technically supported by some computer software. A method of transferring a real situation into a grid map is to use cartograms. For example, in Figure 1.3, in the left graph, the dependences between these three points are described by three numbers. Based on the dependences, the graph on the right is a cartogram, where the dependences are represented by distances. Figure 1.4 is a cartogram, which is introduced on the

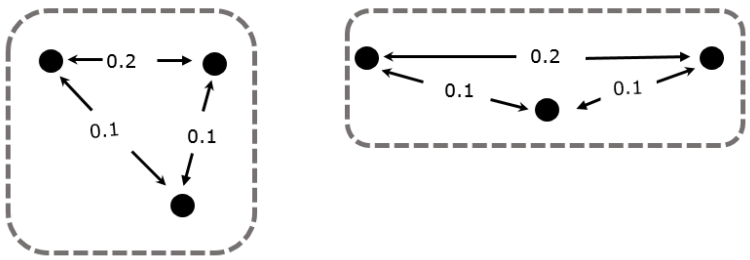


Figure 1.3: An illustration of the cartogram

website [49] and by the paper [27], where a complete rectangular grid of points is transformed as a map. It means a two dimensional rectangular grid of points can be used to describe map features by interpolating it into a cartogram.

The cartogram guarantees our basic abstraction for asymptotics is acceptable and practicable if we use grid-indexed random fields to model populations.

Basic sampling methods in surveys

In general, there are four probability sampling methods: simple random sampling, systematic sampling, cluster sampling, and stratified sampling. They are widely accepted in the first branch in Figure 1.1. Sometimes we also use multi-stage sampling if a sampling procedure is divided into more than one stage. In each stage, one of these sampling methods can be performed.

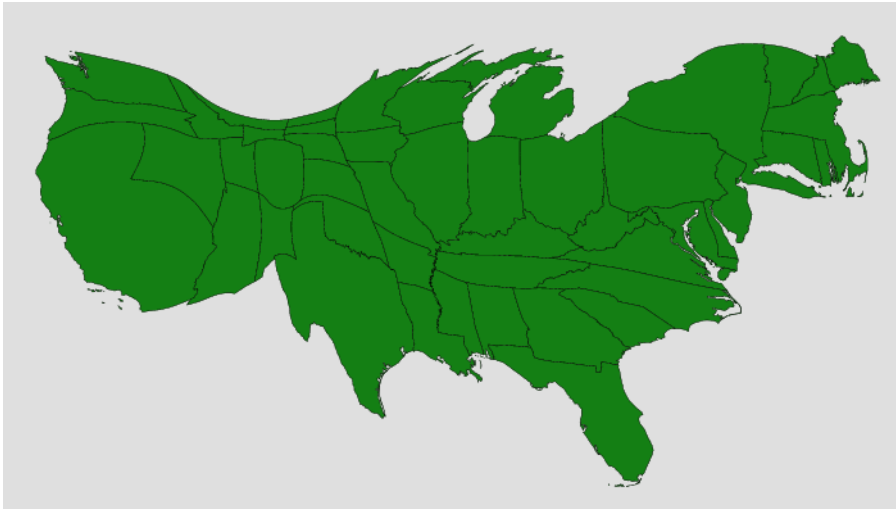
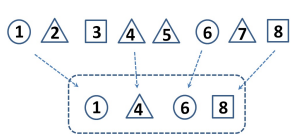
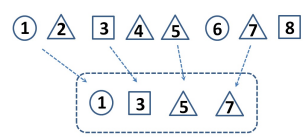


Figure 1.4: A population cartogram of the United States [49]

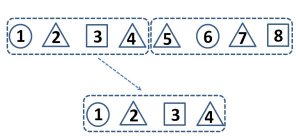
Figure 1.5 exhibits four sampling methods. For a finite population, if we select persons with equal probability, the sampling is simple random, see (a); if we chose persons equally spaced, e.g. every second, as it is shown in (b), we call it systematic sampling method; if a population is clustered, and we sample some of those clusters, then we call it cluster sampling, see (c); if a population is stratified, for example in (d), where the population is divided into three stratum, we select a suitable proportion of persons within each strata, we call this method stratified sampling.



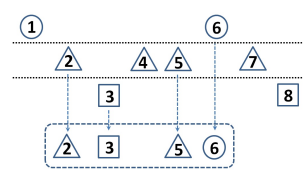
(a) Simple random sampling



(b) Systematic sampling



(c) Clustering sampling



(d) Stratified sampling

Figure 1.5: Four sampling methods for finite populations

In Figure 1.5, the size of the finite population is $N = 8$, and the sample size is $n = 4$. Let the selection probability of the person x_i be $\pi_i = \frac{1}{2}$ for all $i = 1, \cdots, 8$. If the sample is as (a), say persons 1 4 6 8 are selected as a sample, then the population mean may be estimated by

$$m_n = \frac{1}{N}(\frac{x_1}{\pi_1} + \frac{x_4}{\pi_4} + \frac{x_6}{\pi_6} + \frac{x_8}{\pi_8}) = \frac{1}{4}(x_1 + x_4 + x_6 + x_8).$$

Similarly, we have $m_n = \frac{1}{4}(x_1 + x_3 + x_5 + x_7)$, $m_n = \frac{1}{4}(x_1 + x_2 + x_3 + x_4)$, and $m_n = \frac{1}{4}(x_2 + x_3 + x_5 + x_6)$ for the other three cases, (b), (c) and (d), in Figure 1.5.

The indicated sampling method

It is obvious that we can express the estimated mean of the above four cases with one form. In fact, in most real cases, whether a sample point is observed could be affected by some auxiliary information. Therefore, we introduce an *indicated sampling strategy*, which is related to the following chapters.

Let random variables, R_i's, be indexed by $i \in \{1, \cdots, N\}$. We use R_i to indicate the sampling, i.e. $R_i = 1$ implies the person i is selected into a sample; otherwise, $R_i = 0$. Specifically, we let that R_i have Bernoulli distributions with two possible values, 0 and 1, for all $i = 1, 2, \cdots, N$. Thereafter $R_i X_i$ expresses each person, X_i, is selected into a sample with a probability, say p_i. Then we have $\mathbb{E}(R_i) = p_i$, $Var(R_i) = p_i - p_i^2$, for all $i = 1, \cdots, N$. However we do not know the correlations between R_i's, because, at this moment, there is no more information that could support the calculation of $Cov(R_i, R_j)$.

Now let's go back to Figure 1.5 with indicator R_i's. Since we assumed that the sample size is n, we have to set

$$\sum_{i=1}^{N} R_i = n. \tag{1.1}$$

Then, with a suitable setting of the design, say N is even and $n = \frac{1}{2}N$. The sample mean of the first three cases can be expressed by

$$m_n = \frac{\sum_{i=1}^{N} R_i x_i}{\sum_{i=1}^{N} R_i}. \tag{1.2}$$

For the stratified sampling method (d), the sample size is divided into several parts or strata. For example, let us imagine we have m strata, $s_1, \cdots, s_m$, within the sample, s, i,e, $s_1 \cup \cdots \cup s_m = s$. We let

$$\sum_{i \in s_1} R_i = n_i, \quad \cdots, \sum_{i \in s_m} R_i = n_m, \quad \text{where } \sum_{i=1}^{m} n_i = n. \tag{1.3}$$

Then (1.1) is satisfied, and (1.2) holds for four cases in Figure 1.5.

Furthermore, to show (1.2) is feasible, we assume that the joint distribution of $R_1, \cdots, R_N$ is given by

$$\mathbb{P}(R_1 = r_1, \cdots, R_N = r_N) = \begin{cases} \binom{N}{n}^{-1}, & \text{if } r_1 + \cdots + r_N = n, \\ 0, & \text{otherwise,} \end{cases}$$

where $r_i \in \{0, 1\}$. This means we have $\binom{N}{n}$ possible cases satisfying $r_1 + \cdots + r_N = n$. It also

means there are $\binom{N}{n}$ possible subsets T of $\{1,\cdots,N\}$, such that $R_i=1$ for all $i\in T$, where the amount of the elements of T is $|T|=n$. Then for each sample, the estimated mean m_n can be represented by $\bar{X}=\frac{1}{n}\sum_{i\in T}X_i=\frac{1}{n}(R_1X_1+\cdots+R_NX_N)$ with the probability $\binom{N}{n}^{-1}$. We note all $\binom{N}{n}$ possible values of $\bar{X}$ by $\bar{X}_1,\bar{X}_2,\cdots,\bar{X}_{\binom{N}{n}}$. Then the expectation of this estimator is

$$\mathbb{E}(\bar{X})=\binom{N}{n}^{-1}\bar{X}_1+\cdots+\binom{N}{n}^{-1}\bar{X}_{\binom{N}{n}}=\frac{1}{N}\sum_{i=1}^{N}X_i.$$

This gives the unbiased property of this indicated estimation for case (a). For case (b), we define $R_i=1$ if i is odd, $R_i=0$ if i is even, and for case (c), we define $R_i=1$ if $i=1,2,3,4$, $R_i=0$ if $i=5,6,7,8$. Then we will deduce the same result as in case (a). Similarly, for case (d), the stratified sampling, let S_t and s_t be the sub-population and sub-sample in the t-th stratum, and let $N_t=|S_t|$ stand for the size of the sub-population, $n_t=|s_t|$ be the size of the sub-sample, $t=1,\cdots,m$. Let $N=N_1+\cdots+N_m$ be the population total, $n=n_1+\cdots+n_m$ be the sample. We suppose

$$\mathbb{P}(i\in S_t)=\frac{N_t}{N},$$

and

$$\mathbb{P}\left(R_i=r_i,i\in S_t\Big|\sum_{i=1}^{N}r_i=n_t\right)=\binom{N_t}{n_t}^{-1},$$

where $t=1,\cdots,m$. We note that

$$\bar{X}=\sum_{t=1}^{m}\frac{N_t}{N}\frac{1}{n_t}\sum_{i\in T\cap S_t}X_i=\sum_{t=1}^{m}\frac{N_t}{N}\frac{1}{n_t}\sum_{i\in s_t}X_i=\sum_{t=1}^{m}\frac{N_t}{N}\frac{1}{n_t}\sum_{i\in S_t}R_iX_i,$$

where $\sum_{i\in S_t}R_i=n_t$ and the term $\sum_{i\in S_t}R_iX_i$ has $\binom{N_t}{n_t}$ possible cases, and happens with the probability $\binom{N_t}{n_t}^{-1}$. We use

$$\Big(\sum_{i\in s_t}X_i\Big)_1,\quad\Big(\sum_{i\in s_t}X_i\Big)_2,\quad\cdots,\quad\Big(\sum_{i\in s_t}X_i\Big)_{\binom{N_t}{n_t}}$$

stand for those cases.

Therefore

$$\begin{aligned}\mathbb{E}\Big(\bar{X}\Big)&=\sum_{t=1}^{m}\frac{N_t}{N}\frac{1}{n_t}\mathbb{E}\left(\sum_{i\in S_t}R_iX_i\right)\\&=\sum_{t=1}^{m}\frac{N_t}{N}\frac{1}{n_t}\left(\underbrace{\Big(\sum_{i\in s_t}X_i\Big)_1\binom{N_t}{n_t}^{-1}+\cdots+\Big(\sum_{i\in s_t}X_i\Big)_{\binom{N_t}{n_t}}\binom{N_t}{n_t}^{-1}}_{\text{There are }\binom{N_t}{n_t}\text{ terms}}\right)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{t=1}^{m} \frac{N_t}{N} \frac{1}{n_t} \left(\sum_{i \in S_t} X_i \right) \binom{N_t}{n_t}^{-1} \binom{N_t - 1}{n_t - 1} \\
&= \frac{1}{N} \sum_{i=1}^{N} X_i.
\end{aligned}
$$

The above discussion means that all of those four sampling methods can be described by using indicators.

A general sampling strategy

To generalize the above indicated sampling strategy, we assume further that a sample point, X_i, could be influenced by information Z_i and ϵ_i, where i is generalized to $\mathbb{Z}^d$ as well, i.e. $i \in \mathbb{Z}^d$. Therefore X_i, Z_i and ϵ_i are random fields.

We use the random field, which will be defined in the following chapter, to model the superpopulation, i.e. the population is described by random variables in a spatial process. A function $f(\cdot)$ is used to describe random effects from any aspects. We also introduce an indicating random field to select observations.

Specifically, we set $Y_i = R_i f(Z_i, \epsilon_i)$, where $i \in \mathbb{Z}^d$, Y_i is the sample under a sampling method, Z_i and ϵ_i stand for some random effects, which could be white noise or any effects independent of Z_i's. Here, R_i is as in (1.3), but the subscript is in $\mathbb{Z}^d$.

Let Z_i be an m–dimensional random variable, $Z_i \in \mathbb{R}^m$. The function $\phi(\cdot)$ stands for the density function of the joint distribution of Z_i's, that is the marginal density function of Z_i is $\phi(z_i)$, and similarly, $\phi(z_i, z_j)$ stands for the marginal density function of (Z_i, Z_j) and so on. In some of the results within this thesis, we assume the probability distribution of R_i is related to Z_i's. Let Z_i's effect R_i in the following way:

$$
\mathbb{P}(R_i = 1 | \{Z_t, t \in D_n\}) = g_i, \tag{1.4}
$$

where $D_n \subset \mathbb{Z}^d$ is a finite subset of $\mathbb{Z}^d$, g_i could be a function with respect to the subscript of R_i or equivalently of Z_i, therefore it is possible to force g_i being a function of Z_i, i.e. $g_i = g(Z_i, i \in D_n)$. Then we have

$$
\mathbb{P}(R_i = 1) = \int g_i \phi(z) dz,
$$

where $dz = \prod_{i \in D_n} dz_i$, z is a generalized vector of $\{z_i\}_{i \in D_n}$, i.e. $z = (z_{i, i \in D_n})$. For $r_i \in \{0, 1\}$, we note that

$$
\mathbb{P}(R_i = r_i | \{Z_t, t \in D_n\}) = g_i^{r_i} (1 - g_i)^{1 - r_i}.
$$

We assume R_i is conditionally independent given Z_t, i.e.

$$\mathbb{P}(\{R_i = r_i, i \in D_n\}|\{Z_t, t \in Dn\}) = \prod_{i \in D_n} \mathbb{P}(R_i = r_i|\{Z_t, t \in Dn\}), \tag{1.5}$$

then we have

$$\mathbb{P}(\{R_i = r_i, i \in D_n\}) = \int_{(\mathbb{R}^m)^M} \prod_{i \in D_n} g_i^{r_i}(1-g_i)^{1-r_i}\phi(z)dz,$$

where $M = |D_n|$ is the number of indices in D_n. This expression implies, with the previous condition (1.5), we can find a consistent setting for the finite dimensional distribution of $\{R_i\}_{i \in D_n}$. It satisfies two consistency conditions in the Kolmogorov Existence Theorem, see Theorem 15.1.3 in [61] or Theorem A4 in Appendix. Therefore, we can conclude that the random field R_i satisfies (1.5). For the first condition (C1) in Theorem A4, it is satisfied automatically by the setting of (1.5). For the second condition (C2), it is also satisfied, because we have

$$\begin{aligned}
\mathbb{P}(R_i = r_i, R_j \in \mathbb{R}) &= \mathbb{P}(R_i = r_i, R_j = 1) + \mathbb{P}(R_i = r_i, R_j = 0) \\
&= \iint g_i^{r_i}(1-g_i)^{1-r_i} g_j \phi(z)dz \\
&\qquad + \iint g_i^{r_i}(1-g_i)^{1-r_i}(1-g_j)\phi(z)dz \\
&= \iint g_i^{r_i}(1-g_i)^{1-r_i}\phi(z)dz \\
&= \int g_i^{r_i}(1-g_i)^{1-r_i}\phi(z)dz \\
&= \mathbb{P}(R_i = r_i),
\end{aligned}$$

and this can be easily generalized to a M–dimensional space for Z_i, $i \in D_n$. For example, we suppose Z_i is a standard multivariate normal distribution in $\mathbb{R}^m$. Let

$$g_i = g(Z_i) = \frac{e^{Z_i^\top \beta}}{1 + e^{Z_i^\top \beta}},$$

where $\beta \in \mathbb{R}^m$ is a vector with the same dimensional of Z_i's. Then we have

$$\mathbb{P}(\{R_i = r_i, i \in D_n\}) = (2\pi)^{-\frac{mM}{2}}|\Sigma|^{-\frac{1}{2}} \int_{(\mathbb{R}^m)^M} \prod_{i \in D_n} g^{r_i}(z_i)(1-g(z_i))^{1-r_i} e^{-\frac{1}{2}z^\top \Sigma z} dz,$$

where Σ is the covariance matrix of the multivariate normal distribution, $|\Sigma|$ is the determinant of Σ, and $-\frac{1}{2}z^\top \Sigma z$ is a quadratic form in the vector z.

For convenience, sometimes we call $f(Z_i, \epsilon_i)$ a *focal* random field, Z_i and ϵ_i are *extra* random fields. This focal random field is indicated by an *indicating* random field, R_i. The condition (1.5) implies that one of the extra random fields, Z_i, effects both the focal random field and the indicating random field. In following chapters, with this condition, we will study the asymptotics

of the estimator in the form:

$$e_n = \sum_{i \in D_n} R_i f(Z_i, \epsilon_i),$$

where e_n stands for an estimator. Let $a_0, a_1 \in \mathbb{R}$, and

$$c_i = \begin{cases} a_1, & \text{if } R_i = 1, \\ a_0, & \text{if } R_i = 0, \end{cases} \qquad i \in D_n.$$

Then the above estimator can be generalized by

$$e_n = \sum_{i \in D_n} c_i f(Z_i, \epsilon_i).$$

Intuitively, these two estimators will share the same asymptotic properties. We will demonstrate this in the following chapters.

In fact, there are many sampling strategies can be transformed into this indicated sampling method with the above estimators.

Example 3: For sampling methods in Figure 1.5, if we let $f(X_i) = X_i$, $a_0 = 0$ and $a_1 = \frac{1}{n}$ with the constraint of (1.1), then the estimator in (1.2) can be transformed into

$$m_n = \sum_{i=1}^{N} c_i f(X_i).$$

Example 4: Similarly, if we suppose the population is a superpopulation in NHANES, then for the first stage of the sample design, we can set indicating random variables, R_i's, to indicate a PSU or a part of divided certainty PSUs. Furthermore, by the previous discussion, we suppose $\mathbb{P}(R_i = 1) = g(Z_i)$, where Z_i could be some random effects from MOS, incomes, community status and/or traffic status. The dependence between R_i and R_j is relative to the corresponding Z_i and Z_j to some extent. All the dependences between variables, e.g. between Z_i and Z_j, within the survey will not be affected by the sample design. This is because of all the possible information that the dependences could depend on (e.g. MOS, incomes, the status of a community and/or the traffic situation) will not be influenced by the sample design. We can also imagine that the focal random field is $f(Z_i, \epsilon_i)$, where ϵ_i could be some other random effects or white noise. It means the focal random field, the health issues of the population, is driven by some random effects such as incomes, rural areas, community and traffic status, and so on. For different strata, the design restricts the sample sizes which is similar to what is shown in (1.3). Therefore, the sample mean is expressed as

$$m_n = \sum_{j=1}^{m} \frac{n_j}{n} \frac{\sum_{i \in s_j} R_i f(Z_i, \epsilon_i)}{\sum_{i \in s_j} R_i} = \sum_{i \in D_n} \frac{R_i f(Z_i, \epsilon_i)}{n},$$

where $D_n \in \mathbb{Z}^d$ is the index of the superpopulation, the subsets of D_n, $s_1, \cdots, s_m$ are used to describe m strata of PSUs in the first stage, $s_1 \cup \cdots \cup s_m = D_n$, n is the design required sample size, n_j is the required sample size within stratum j, $\sum_{i \in s_j} R_i = n_j$ and $\sum_{i \in D_n} R_i = \sum_{j=1}^m n_j = n$. If we follow the same setting of c_i in Example 3, the sample mean of the current example can be given by

$$m_n = \sum_{i \in D_n} c_i f(Z_i, \epsilon_i).$$

Example 5: Let's suppose the selection probability of $f(Z_i, \epsilon_i)$ is $\pi_i = \mathbb{P}(R_i = 1)$, which is driven by Z_i, e.g. it is defined as in (1.4), then the total of the sample can be estimated by a Horvitz-Thompson estimator,

$$T_n = \sum_{i \in D_n} \frac{R_i f(Z_i, \epsilon_i)}{\pi_i} = \sum_{i \in D_n} R_i \frac{f(Z_i, \epsilon_i)}{\mathbb{P}(R_i = 1)}.$$

Again, if we define $\tilde{f}(Z_i, \epsilon_i) = \frac{f(Z_i, \epsilon_i)}{\mathbb{P}(R_i = 1)}$, then the estimator T_n is also in the form of

$$T_n = \sum_{i \in D_n} R_i \tilde{f}(Z_i, \epsilon_i).$$

Example 6: Let $D_N \subset \mathbb{Z}^d$ be the finite population, $A = \{i \in D_N : R_i = 1\}$. Then the sample mean and the population mean are

$$\bar{x} = \frac{1}{|A|} \sum_{i \in A} x_i \quad \text{and} \quad \bar{x}_N = \frac{1}{|D_N|} \sum_{i \in D_N} x_i$$

respectively. If we further set $a_0 = -\frac{1}{|D_N|}$ and $a_1 = \frac{1}{|A|} - \frac{1}{|D_N|}$, then the estimator of the difference of these two means is in the form of

$$e_N = \bar{x} - \bar{x}_N = \sum_{i \in D_N} c_i x_i.$$

The rest of this thesis will contribute to the asymptotics of estimators under the indicated sampling strategy, where some reasonable conditions will be introduced for dependent populations, and the feasibility of the conditions will be discussed.

1.4 Assumptions on spatial structures

Many asymptotic results are based on assumptions that are convenient for the proofs. For example, the relationship between sample points, sampled clusters, or sampled strata in spatial

structures, are assumed to exhibit either an i.i.d property, the non-identical but independent distributions property, the stationary property, the Markov property, or the Martingale property, and so on. These properties are realistic for solving practical problems.

However, there are many more practical problems that cannot be described by these properties. For example, in [45], Lumley and Scott claimed that, based on the practical data set, the stroke death rates of NHANES (1996) had a complex spatial correlation structure. This correlation structure is similar to that of Example 1 in Section 1.2. In another work [46], Lumley and Scott noted that "for future research that it would be valuable to have sampling asymptotics better founded in the spatial structure of populations, not only for a better match to reality but also because it could simplify the development of Donsker-type theorems, uniform tail bounds, and other machinery of modern mathematical statistics".

Practically, we introduce three basic assumptions on spatial structures. The first assumption is on the dependence. We assume the correlations can be described by mixing coefficients. This represents a major relaxation of the assumption of independence. The second assumption is on the divided population, i.e. a random field is assumed to be divided into joint blocks. This assumption is clarified and used in Chapter 4. The third assumption is that we assume the sample regions are nested. This assumption is introduced and mainly used in Chapter 5.

1.5 Outline of this thesis

In order to have asymptotics for our general sampling strategy, we introduce preliminary definitions, properties of random fields, and properties of strong mixing dependence in Chapter 2. We prepare some lemmas on the relationship between strong mixing coefficient of the indicated sample and the coefficients of indicating random field and the extra random fields. We introduce an equivalent definition of the strong mixing coefficient in the first lemma. We also introduce a new definition on the strong mixing coefficient between two random fields. Then the last lemma implies that the strong mixing coefficient of a focal random field can be estimated by the strong mixing coefficients of the extra random fields and the strong mixing coefficient between them. Strong mixing describes the kind of dependence that often occurs in practice.

Chapter 3 is devoted to setting up conditions for the asymptotic properties of complex surveys. Two main aspects are considered in this chapter. One is the conditional independence, i.e. given limited information within one extra random field, the focal random field and the indicating random field have some independent properties. The other is that extra random fields and the indicating random field satisfy strong mixing conditions simultaneously. Moreover, some other developments are considered to prove CLTs for some specific complex surveys.

Chapter 4 gives some results on L_2–consistent estimators of the variance of the total for the

dependent survey data. We develop Carlstein's results of [13] and Fuller's Theorem 1.3.2 of [26] into dependent non-stationary random fields, with the uniform integrability assumption and the assumption on the smooth property of the function which specifies the focal random fields. These results also work for indicated focal random fields.

In Chapter 5, we provide an upper bound for the absolute covariance of two complex-valued random variables, with their quantile functions. We also generalized Rio's result, in [57], on the upper bound of the fourth moment of the sample sum. All the results are under the Skorohod topology in space D, where the tightness, uniformly integrable and asymptotic independent increments are considered for the random elements sequence whose normalized time dependent summation converges weakly to a Brownian motion. Based on these results of functional central limit theorems, the convergence of the sequence of random elements on the indicated focal random fields are shown to hold.

Materials which are used but are not quite central to the logic of this thesis are presented in the Appendix.

Chapter 2

Random fields

A random field is a generalization of a random process where the index set has dimension greater than one. Let $\{X_i\}$ be a collection of random variables. If the index satisfies $i \in \mathbb{Z}^d$, where $d > 1$, we say that the X_i's constitute a random field. Sometimes we use the notation X to stand for the random field, or use $\{X_i\}_{i\in\mathbb{Z}^d}$ and $X_{i\in\mathbb{Z}^d}$ to emphasize the index set. In this chapter, we will introduce random fields with strong mixing properties and prepare some basic results for the following chapters.

2.1 Strong mixing dependence

It is widely accepted that independence is a special case of dependence for random variables. Dependence is more comprehensive than independence in modelling real situations. There are many ways to describe the dependence between variables in a random field. Typically, we can assume the random field has some dependence properties, such as the Markov property, the martingale property and the stationary property. There are many results that follow from these assumptions on random fields. For Gaussian Markov random fields, we recommend the book [63], which provides the comprehensive account of the properties of Gaussian Markov random fields. For martingale random fields, because the spatial structure in a martingale random field has no completed order like a line, it is not very productive as a research field. However, for some specific orders, it will be very helpful to have asymptotics on martingale random fields, such as the paper [74] with the partial order and the work [17] with the lexicographic order. Sometimes a stationary property is added to martingale random fields for proving asymptotics, see [2]. For random fields with the stationary property, we suggest some books, such as [40], [41] and [42].

Mixing properties are a kind of weak dependence on random fields. There are mainly five

definitions of mixing; α-mixing, β-mixing, ϕ-mixing, ψ-mixing and ρ-mixing, as discussed by Paul Doukhan in [19]. weaker than others. This means that if the correlation can be bounded by α-mixing it can be bounded by others. Therefore, throughout this thesis, we use α–mixing to describe the dependence on random fields.

The α-mixing property is sometimes also called strong mixing. It was first introduced by Murray Rosenblatt in [60] in 1956. Based on the strong mixing dependence, there are many results on asymptotics, e.g. for CLTs, see [7], [10], [47] and [30]which have stationary assumptions, [21] for triangular array settings, and [73] considered linear processes; for FCLTs, see [20], [48], [52] and [59] which have stationary assumptions, [1] for triangular array settings,[32] for one dimensional non-stationary random processes; for the estimation of covariance, see [58]. For basic properties of strong mixing, we recommend the survey paper by Richard Bradley [8] and the updated version [11] in 2005.

Let $i, j \in \mathbb{Z}^d$, i.e. $i = (i_1, i_2, \cdots, i_d) \in \mathbb{Z}^d$ and $j = (j_1, j_2, \cdots, j_d) \in \mathbb{Z}^d$. We define the distance between i and j as

$$d(i, j) = \sum_{s=1}^{d} |i_s - j_s|.$$

Let Λ_1, Λ_2 be subsets of $\mathbb{Z}^d$. We call $X_{i\in\Lambda_1}$ and $X_{i\in\Lambda_2}$ *blocks* in the random field $X_{i\in\mathbb{Z}^d}$. Let $\sigma(X_{i\in\Lambda_1})$ and $\sigma(X_{i\in\Lambda_2})$ be sigma fields generated by $X_{i\in\Lambda_1}$ and $X_{i\in\Lambda_2}$ respectively. We define $dist(\Lambda_1, \Lambda_2)$ as the shortest distance between two sets Λ_1 and Λ_2. We use $dist(\Lambda_1, \Lambda_2)$ to stand for the distance between two blocks, $X_{i\in\Lambda_1}$ and $X_{i\in\Lambda_2}$. For the special case of $\emptyset$ being one or both of those two sets, we take the distance as zero, i.e. $dist(\emptyset, \Lambda_2) = dist(\Lambda_1, \emptyset) = dist(\emptyset, \emptyset) = 0$. For each subset, $\Lambda \subset \mathbb{Z}^d$, we define the size of Λ as the cardinality, $|\Lambda| = \#(\Lambda)$, i.e. the number of the elements in Λ. It is obvious that the union and the intersection of Λ_1 and Λ_2 satisfy $|\Lambda_1 \cup \Lambda_2| = |\Lambda_1| + |\Lambda_2| - |\Lambda_1 \cap \Lambda_2|$.

Definition 2.1. A *strong mixing coefficient* between two blocks within a random field is defined by a function with respect to the distance between the blocks, and the size of the blocks, specifically,

$$\begin{aligned} \alpha_{k,l}(m) = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|\colon A \in \sigma(X_{i\in\Lambda_1}), B \in \sigma(X_{i\in\Lambda_2}), \\ k = |\Lambda_1|, l = |\Lambda_2|, dist(\Lambda_1, \Lambda_2) \geq m\}, \end{aligned} \tag{2.1}$$

where the supremum is taken over A, B, Λ_1 and Λ_2.

In this thesis we also use $\alpha^X_{k,l}(m)$, or $\alpha_{k,l}(X; m)$, to stand for the strong mixing coefficient of the random field $X_{i\in\mathbb{Z}^d}$. We note that, for any specific index sets Λ_1 and Λ_2, $|\Lambda_1| = k$, $|\Lambda_2| = l$, we define

$$\alpha_{|\Lambda_1|,|\Lambda_2|}(m) = \sup_{A,B}\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|\colon A \in \sigma(X_{i\in\Lambda_1}), B \in \sigma(X_{i\in\Lambda_2}), dist(\Lambda_1, \Lambda_2) \geq m\}.$$

Since this supremum is taken over only A and B, we have

$$\alpha_{|\Lambda_1|,|\Lambda_2|}(m) \leq \alpha_{k,l}(m).$$

In this thesis, strong mixing coefficients are assumed to be vanishing as the distance goes to infinity. If $\alpha_{k,l}(m) \to 0$ as $m \to \infty$, then the process is called a *strong mixing process*.

If the distance is not involved, or the distance is not taken into account in some cases, for X and Y, we can also define strong mixing coefficient for them, i.e.

$$\alpha(X,Y) = \sup_{A,B}\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|\colon A \in \sigma(X), B \in \sigma(Y)\}.$$

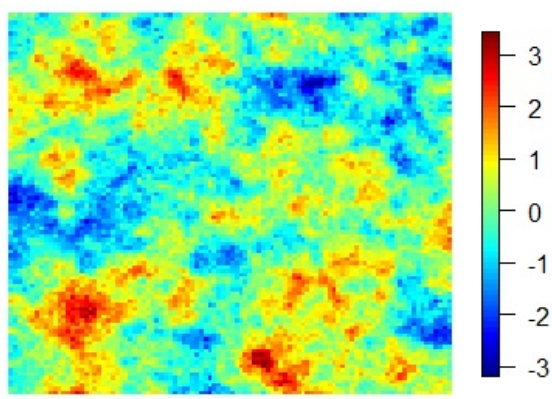


Figure 2.1: Simulation of random fields with 100×100 grids

Figure 2.1 is the simulation of a Gaussian random field with an exponential covariance function. It is an example of sim.rf function in fields package [51] on CRAN website. It illustrates that a random field is related to Figure 1.2 and Figure 1.4 in Chapter 1. If the boundary of domains is clarified, and the dependence is fitted, in practice, Figure 1.2 will be transformed into a cartogram map, which can be used to describe a random field with dependence.

Following from the fact that if X and Y are random variables, then (X,Y) is a random vector, $(X_i, Y_i)_{i\in\mathbb{Z}^d}$ is also a random field provided that X_i and Y_i are random variables.

Definition 2.2. The *strong mixing coefficient* within a combined random field can be described by

$$\begin{aligned}\alpha_{k,l}^{(X,Y)}(m) = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|\colon A \in \sigma((X_i,Y_i)_{i\in\Lambda_1}), B \in \sigma((X_i,Y_i)_{i\in\Lambda_2}),\\ |\Lambda_1| = k, |\Lambda_2| = l, dist(\Lambda_1,\Lambda_2) \geq m\}.\end{aligned} \tag{2.2}$$

Similarly, let f be a measurable function from random fields $X_{i\in\mathbb{Z}^d}$ and $Y_{i\in\mathbb{Z}^d}$ to a field $f(X_i,Y_i)_{i\in\mathbb{Z}^d}$.

Let $f(X_j, Y_j) \in \mathbb{R}^k$, B be a Borel set of $\mathcal{B}$, which is the class of Borel sets of $\mathbb{R}^k$. Then, for any event ω of the whole space Ω of (X_j, Y_j), we have

$$\{\omega : f(X_j, Y_j) \in B\} = \{\omega : (X_j, Y_j) \in f^{-1}(B)\} \in \sigma(X_j, Y_j), \tag{2.3}$$

which implies $f(X_i, Y_i)$ is also a random variable. Therefore $f(X_i, Y_i)_{i\in\mathbb{Z}^d}$ is a random field. The strong mixing coefficient of $f(X_i, Y_i)_{i\in\mathbb{Z}^d}$ can be noted by $\alpha_{k,l}^{f(X,Y)}(m)$.

The following theorem provides an equivalent definition of strong mixing coefficients.

Theorem 2.1. *Let $\mathcal{H}[0,1]$ be a class of measurable functions taking values in $[0,1]$. Let $X_{i\in\mathbb{Z}^d}$ be a random field, $\Lambda_1, \Lambda_2 \subseteq \mathbb{Z}^d$, h_1 and h_2 be two measurable functions of $X_{i\in\Lambda_1}$ and $X_{j\in\Lambda_2}$ respectively. Then we have*

$$\sup_{h_1,h_2\in\mathcal{H}[0,1]} |Cov(h_1(X_{i\in\Lambda_1}), h_2(X_{i\in\Lambda_2}))| = \sup_{A_1\in\sigma(X_{i\in\Lambda_1}), A_2\in\sigma(X_{i\in\Lambda_2})} |\mathbb{P}(A_1\cap A_2) - \mathbb{P}(A_1)\mathbb{P}(A_2)|. \tag{2.4}$$

Proof. We note that $A_1 \in \sigma(X_{i\in\Lambda_1})$ is equivalent to saying there exists a Borel set B_1 such that $A_1 = \{X_{i\in\Lambda_1} \in B_1\}$. Similarly, we have B_2 such that $A_2 = \{X_{j\in\Lambda_2} \in B_2\} \in \sigma(X_{i\in\Lambda_2})$.

On one hand, let $\mathbb{1}_A$ be an indicator function of an event A, i.e.

$$\mathbb{1}_A = \begin{cases} 1, & \text{if } A \text{ happened,} \\ 0, & \text{otherwise.} \end{cases}$$

We have

$$\begin{aligned} \sup_{A_1,A_2} |\mathbb{P}(A_1\cap A_2) - \mathbb{P}(A_1)\mathbb{P}(A_2)| &= \sup_{A_1,A_2} |Cov(\mathbb{1}_{A_1}, \mathbb{1}_{A_2})| \\ &\leq \sup_{0\leq h_1,h_2\leq 1} |Cov(h_1(X_{i\in\Lambda_1}), h_2(X_{j\in\Lambda_2}))|, \end{aligned}$$

where h_1 and h_2 are measurable functions.

On the other hand, using the idea in the Proof of Theorem 1.1 in [57], we set

$$Y_1 = h_1(X_{i\in\Lambda_1}), \quad Y_2 = h_2(X_{j\in\Lambda_2}).$$

We note that, for any Borel set B of $\mathbb{R}$, we have $\{\omega : h_1(X_{i\in\Lambda_1}) \in B\} \in \sigma(Y_1)$ and

$$\{\omega : h_1(X_{i\in\Lambda_1}) \in B\} \subseteq \{\omega : X_i \in h_1^{-1}(B)\} \in \sigma(X_{i\in\Lambda_1}),$$

i.e. $\sigma(Y_1) \subseteq \sigma(X_{i\in\Lambda_1})$. Similarly we have $\sigma(Y_2) \subseteq \sigma(X_{j\in\Lambda_2})$. We also note that, $0 \leq h_1, h_2 \leq 1$

implies $0 \leq Y_1, Y_2 \leq 1$, and

$$Y_1 = \int_0^1 \mathbb{1}_{\{Y_1 > y_1\}} dy_1, \quad Y_2 = \int_0^1 \mathbb{1}_{\{Y_2 > y_2\}} dy_2.$$

Using Fubini's Theorem, we have

$$Cov(Y_1, Y_2) = \int_0^1 \int_0^1 Cov(\mathbb{1}_{Y_1 > y_1}, \mathbb{1}_{Y_2 > y_2}) dy_1 dy_2.$$

Then we have

$$\begin{aligned}
|Cov(Y_1, Y_2)| &= \left| \int_0^1 \int_0^1 Cov(\mathbb{1}_{Y_1 > y_1}, \mathbb{1}_{Y_2 > y_2}) dy_1 dy_2 \right| \\
&\leq \int_0^1 \int_0^1 |Cov(\mathbb{1}_{Y_1 > y_1}, \mathbb{1}_{Y_2 > y_2})| \, dy_1 dy_2 \\
&\leq \int_0^1 \int_0^1 \sup_{A_1 \in \sigma(Y_1), A_2 \in \sigma(Y_2)} |Cov(\mathbb{1}_{A_1}, \mathbb{1}_{A_2})| \, dy_1 dy_2 \\
&= \sup_{A_1 \in \sigma(Y_1), A_2 \in \sigma(Y_2)} |Cov(\mathbb{1}_{A_1}, \mathbb{1}_{A_2})| \\
&\leq \sup_{A_1 \in \sigma(X_{i \in \Lambda_1}), A_2 \in \sigma(X_{i \in \Lambda_2})} |Cov(\mathbb{1}_{A_1}, \mathbb{1}_{A_2})| .
\end{aligned}$$

This completes the proof. □

In the paper [56], Prakasa Rao introduced a definition of conditional strong mixing, Definition 4, where the coefficient was generalized to random variables. In another paper [55] by the same author, the mixing stably concept, Definition 2.6 in [55], was introduced for ϕ –mixing. Inspired by these definitions, we introduce the following definition.

Definition 2.3. Let $Z_{i \in \mathbb{Z}^{d_1}}$ be a random field. A *conditional strong mixing coefficient* of $X_{i \in \mathbb{Z}^d}$, which is conditional on Z, is defined by

$$\begin{aligned}
\alpha_{k,l}(X|Z; m) = \sup\{ &|\mathbb{P}(A \cap B | C_1 \cap C_2) - \mathbb{P}(A|C_1)\mathbb{P}(B|C_2)| \colon |\Lambda_1| = k, |\Lambda_2| = l, \\
&A \in \sigma(X_{i \in \Lambda_1}), B \in \sigma(X_{j \in \Lambda_2}), C_1, C_2 \in \sigma(Z_{i \in \Lambda_1 \cup \Lambda_2}), \\
&\mathbb{P}(C_1 \cap C_2) > 0, dist(\Lambda_1, \Lambda_2) \geq m\}.
\end{aligned} \tag{2.5}$$

In the above definition, in order to measure the dependence within the random field X, a conditional random field Z is introduced to X.

2.2 Properties of the strong mixing coefficient

For a random field X, the strong mixing coefficient between two blocks, $X_{i\in\Lambda_1}$ and $X_{i\in\Lambda_2}$, being zero is equivalent to X_i and X_j being independent for all $i \in \Lambda_1$ and $j \in \Lambda_2$. To estimate the upper bound of strong mixing coefficients or $|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|$, a direct calculation provides $|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \leq \frac{1}{4}$. Alternatively, by using the indicator function, the discussion in the proof of Theorem 2.1 and Cauchy–Schwarz inequality, we can also have the same bound. Therefore, those definitions of strong mixing coefficients which have the form of $|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|$ can be bounded by $\frac{1}{4}$. If we denote $\alpha_{k,l}(m)$ and $\alpha_{k,l}^{(X,Y)}(m)$ by one symbol α, then we have

$$0 \leq \alpha \leq \frac{1}{4}. \tag{2.6}$$

For the above inequality we refer to the result on page 8 in Rio's book [57] and the remark on page 4 in Doukhan's book [19]. Since $\sigma(X_{i\in\Lambda}) \subseteq \sigma(X_{i\in\tilde{\Lambda}})$ for all $\Lambda \subseteq \tilde{\Lambda} \subseteq \mathbb{Z}^d$, the strong mixing coefficient in Definition 2.1 implies that for all $k_1 \leq k_2$, $k_1, k_2, l \in \mathbb{Z}^+$, $m \geq 0$, we have

$$\alpha_{k_1,l}(m) \leq \alpha_{k_2,l}(m), \quad \alpha_{l,k_1}(m) \leq \alpha_{l,k_2}(m). \tag{2.7}$$

Similarly, for other strong mixing coefficients in Definition 2.2 and Definition 2.3, we have the same results as in (2.7). Also in those definitions, the restriction provided by the distance between two blocks being no less than m implies those strong mixing coefficients are non-increasing functions with respect to m.

We define

$$\|X\|_p = \sup_{i\in\mathbb{Z}^d} \|X_i\|_p, \quad \|X_i\|_p = (\mathbb{E}|X_i|^p)^{\frac{1}{p}}, \quad 1 \leq p < \infty,$$

$$\|X\|_\infty = \sup_{i\in\mathbb{Z}^d} \|X_i\|_\infty, \quad \|X_i\|_\infty = \text{ess-sup}|X_i| = \inf\{x \in \mathbb{R} : \mathbb{P}(|X_i| \leq x) = 1\}.$$

Theorem 2.2. *Let $X_{i\in\mathbb{Z}^d}$ be a random field, $\Lambda_1, \Lambda_2 \subseteq \mathbb{Z}^d$, $|\Lambda_1| = k$, $|\Lambda_2| = l$. Then for each $\delta > 0$, we have*

$$|Cov(X_i, X_j)| \leq 8\alpha_{k,l}^{\frac{\delta}{2+\delta}}(m)\|X\|_{2+\delta}^2, \quad \forall i \in \Lambda_1, \quad \forall j \in \Lambda_2, \tag{2.8}$$

where $m = dist(\Lambda_1, \Lambda_2)$.

The proof of this theorem can be found in [19] and [34], or see the Appendix in this thesis.

Lemma 2.3. *Let $X_{t\in\mathbb{Z}^d}$ be a random field,*

$$E_i \in \sigma(X_{t\in\Lambda_i}), \quad \Lambda_i \subset \mathbb{Z}^d, \quad i = 1, \cdots, r,$$

$$m = \min\{dist(\Lambda_i, \Lambda_j), i \neq j, i, j = 1, \cdots, r\},$$

$$k = \sum_{i=1}^{r} |\Lambda_i|.$$

Then we have

$$\left| \mathbb{P}\left(\bigcap_{i=1}^{r} E_i \right) - \prod_{i=1}^{r} \mathbb{P}(E_i) \right| \leq (r-1)\alpha_{k,k}(m). \tag{2.9}$$

Proof. We use the induction method. For $r = 1$, (2.9) holds. We suppose it holds for $r = p$, i.e. let

$$E = \bigcap_{i=1}^{p} E_i, \quad P = \prod_{i=1}^{p} \mathbb{P}(E_i),$$

$$m^* = \min\{dist(\Lambda_i, \Lambda_j), i \neq j, i, j = 1, \cdots, p\},$$

$$k^* = \sum_{i=1}^{p} |\Lambda_i|,$$

then we suppose

$$|\mathbb{P}(E) - P| \leq (p-1)\alpha_{k^*,k^*}(m^*).$$

Following this, we need to prove it is still true for $r = p + 1$.

Now let

$$E_{p+1} \in \sigma(X_{i \in \Lambda_{p+1}}), \quad |\Lambda_{p+1}| = k_*,$$

$$m = \min\{dist(\Lambda_i, \Lambda_j), i \neq j, i, j = 1, \cdots, p+1\},$$

$$k = \sum_{i=1}^{p+1} |\Lambda_i|,$$

then, by using the monotonicity properties of strong mixing coefficients, we have

$$\alpha_{k^*,k^*}(m^*) \leq \alpha_{k,k}(m), \quad \alpha_{k^*,k_*}(m) \leq \alpha_{k,k}(m),$$

which implies

$$|\mathbb{P}(E) - P| \leq (p-1)\alpha_{k,k}(m)$$

and

$$|\mathbb{P}(E \cap E_{k+1}) - \mathbb{P}(E)\mathbb{P}(E_{p+1})| \leq \alpha_{k,k}(m).$$

Therefore, we have

$$\begin{aligned}
\left| \mathbb{P}\left(\bigcap_{i=1}^{p+1} E_i \right) - \prod_{i=1}^{p+1} \mathbb{P}(E_i) \right| &= |\mathbb{P}(E \cap E_{p+1}) - P\mathbb{P}(E_{p+1})| \\
&\leq |\mathbb{P}(E \cap E_{k+1}) - \mathbb{P}(E)\mathbb{P}(E_{p+1})| + |\mathbb{P}(E)\mathbb{P}(E_{p+1}) - P\mathbb{P}(E_{p+1})| \\
&\leq \alpha_{k,k}(m) + \mathbb{P}(E_{k+1})(p-1)\alpha_{k,k}(m)
\end{aligned}$$

$$\leq \quad p\alpha_{k,k}(m).$$

This completes the proof. □

Let $R_{i\in\mathbb{Z}^d}$ be an indicating random field, so that

$$R_i \text{ takes two possible values, 0 and 1, for all } i \in \mathbb{Z}^d. \tag{2.10}$$

Then we have the following lemma.

Lemma 2.4. *Let $R_{i\in\mathbb{Z}^d}$ be defined by (2.10). Then for any random field $X_{i\in\mathbb{Z}^d}$, we have*

$$\alpha_{k,l}^{RX}(m) \leq \alpha_{k,l}^{(R,X)}(m), \tag{2.11}$$

where m, k, l are the same as in Definition 2.1 and Definition 2.2.

Proof. It is enough to prove that

$$\sigma(\{R_tX_t\}; t \in \Lambda) \subset \sigma(\{R_t\}, \{X_t\}; t \in \Lambda), \quad \Lambda \subseteq \mathbb{Z}^d.$$

Here we use Λ stand for any possible subset of $\mathbb{Z}^d$, say Λ_1 or Λ_2, in the definitions. Because, for any size of Λ, the discussions in the proof are the same, we suppose there are only two indices in Λ, i.e. $\Lambda = \{t_1, t_2\}$. Then we are going to prove

$$\sigma(R_{t_1}X_{t_1}, R_{t_2}X_{t_2}) \subset \sigma(R_{t_1}, X_{t_1}, R_{t_2}, X_{t_2}).$$

Let A be a Borel set. We write an event E, which is in $\sigma(R_{t_1}X_{t_1}, R_{t_2}X_{t_2})$, as

$$\begin{aligned}
E &= \Big\{\omega : \Big(R_{t_1}(\omega)X_{t_1}(\omega),\ R_{t_2}(\omega)X_{t_2}(\omega)\Big) \in A\Big\} \\
&= \Big(\Big\{\omega : \Big(X_{t_1}, X_{t_2}\Big) \in A\Big\} \cap \Big\{\omega : \Big(R_{t_1}, R_{t_2}\Big) = (1,1)\Big\}\Big) \\
&\quad \cup \Big(\Big\{\omega : \Big(X_{t_1}, 0\Big) \in A\Big\} \cap \Big\{\omega : \Big(R_{t_1}, R_{t_2}\Big) = (1,0)\Big\}\Big) \\
&\quad \cup \Big(\Big\{\omega : \Big(0, X_{t_2}\Big) \in A\Big\} \cap \Big\{\omega : \Big(R_{t_1}, R_{t_2}\Big) = (0,1)\Big\}\Big) \\
&\quad \cup \Big(\Big\{\omega : \Big(0, 0\Big) \in A\Big\} \cap \Big\{\omega : \Big(R_{t_1}, R_{t_2}\Big) = (0,0)\Big\}\Big).
\end{aligned}$$

Since $\{\omega : (X_{t_1}, X_{t_2}) \in A\}$, $\{\omega : (X_{t_1}, 0) \in A\}$, $\{\omega : (0, X_{t_2}) \in A\}$, $\{\omega : (0,0) \in A\}$ are in $\sigma(X_{t_1}, X_{t_2})$; and $\{\omega : (R_{t_1}, R_{t_2}) = (1,1)\}$, $\{\omega : (R_{t_1}, R_{t_2}) = (1,0)\}$, $\{\omega : (R_{t_1}, R_{t_2}) = (0,1)\}$, $\{\omega : (R_{t_1}, R_{t_2}) = (0,0)\}$ are in $\sigma(R_{t_1}, R_{t_2})$, then we have $E \in \sigma(R_{t_1}, X_{t_1}, R_{t_2}, X_{t_2})$. This completes the proof. □

Let $Z_{i\in\mathbb{Z}^d}$ be another random field. For any two subsets, $\Lambda_1, \Lambda_2 \in \mathbb{Z}^d$, if for all $A_1 \in \sigma(X_{i\in\Lambda_1})$,

$A_2 \in \sigma(X_{i\in\Lambda_2})$, $T_1 \in \sigma(Z_{i\in\Lambda_1})$ and $T_2 \in \sigma(Z_{i\in\Lambda_2})$, we suppose

$$\mathbb{P}(A_1A_2|T_1T_2) = \mathbb{P}(A_1|T_1)\mathbb{P}(A_2|T_2). \tag{2.12}$$

Then we say the random field $(X_i)_{i\in\mathbb{Z}^d}$ is *conditionally independent* on the *conditional random field*, $Z_{i\in\mathbb{Z}^d}$.

We use $\phi(\cdot)$ stands for the joint density function of specific random vectors, e.g.,

$$\phi(z_{t\in\Lambda}), \quad \Lambda \subseteq \mathbb{Z}^d, \tag{2.13}$$

means the density of a joint distribution of $Z_{t\in\Lambda}$, and $\phi(z_t)$ is the density of Z_t, $t \in \mathbb{Z}^d$. Let $|\Lambda_1|= k$ and $|\Lambda_2|= l$. The above condition, (2.12), implies

$$\begin{aligned}
\alpha^X_{k,l} &= \sup_{A_1,A_2} |\mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1)\mathbb{P}(A_2)| \\
&= \sup_{A_1,A_2} \Bigg| \int\int \mathbb{P}(A_1 \cap A_2|\{Z_t = z_t, t \in \Lambda_1\} \cap \{Z_s = z_s, s \in \Lambda_2\}) \\
&\qquad\qquad \phi(z_{t\in\Lambda_1}, z_{s\in\Lambda_2})dz_{t\in\Lambda_1}dz_{s\in\Lambda_2} \\
&\qquad - \int \mathbb{P}(A_1|\{Z_t = z_t\}_{t\in\Lambda_1})\phi(z_{t\in\Lambda_1})dz_{t\in\Lambda_1} \\
&\qquad\qquad \int \mathbb{P}(A_2|\{Z_s = z_s\}_{s\in\Lambda_2})\phi(z_{s\in\Lambda_2})dz_{s\in\Lambda_2}\Bigg| \\
&= \sup_{A_1,A_2} \Bigg| \int\int \mathbb{P}(A_1|\{Z_t = z_t\}_{t\in\Lambda_1})\mathbb{P}(A_2|\{Z_s = z_s\}_{s\in\Lambda_2}) \\
&\qquad \times [\phi(z_{t\in\Lambda_1}, z_{s\in\Lambda_2}) - \phi(z_{t\in\Lambda_1})\phi(z_{s\in\Lambda_2})]dz_{t\in\Lambda_1}dz_{s\in\Lambda_2}\Bigg|.
\end{aligned}$$

We set

$$h_1(z_t) = \mathbb{P}(A_1|\{Z_t = z_t\}_{t\in\Lambda_1}), \quad h_2(z_s) = \mathbb{P}(A_2|\{Z_s = z_s\}_{s\in\Lambda_2}).$$

Then, by using Theorem 2.1, we have

$$\alpha^X_{k,l}(m) \le \sup_{0\le h_1,h_2\le 1} |Cov(h_1(Z_{t\in\Lambda_1}), h_2(Z_{s\in\Lambda_2}))| = \alpha^Z_{k,l}(m).$$

It gives a lemma:

Lemma 2.5. *For any random field $X_{i\in\mathbb{Z}^d}$, if it is conditionally independent on another random field $Z_{i\in\mathbb{Z}^d}$, i.e. (2.12) is satisfied, then $\alpha^X_{k,l}(m) \le \alpha^Z_{k,l}(m)$.*

This lemma works for any random fields, e.g. if a combined random field (R, X) satisfies (2.12), then we have $\alpha^{(R,X)}_{k,l}(m) \le \alpha^Z_{k,l}(m)$. Obviously, let $Z_{i\in\mathbb{Z}^d}$ and $\epsilon_{i\in\mathbb{Z}^d}$ be two extra random fields, $X_i = f(Z_i, \epsilon_i)$ be the focal random field, $R_{i\in\mathbb{Z}^d}$ be the indicating random field. If (R, X) are conditionally independent given Z which satisfies (2.12), then we also have $\alpha^{(R,X)}_{k,l}(m) \le \alpha^Z_{k,l}(m)$.

Furthermore, if we set up a conditional independent property for the indicating random field, we have the following lemma.

Lemma 2.6. *Let $R_{i\in\mathbb{Z}^d}$ be the indicating random field defined by (2.10), the focal random field $X_{i\in\mathbb{Z}^d}$ be a function with respect to an extra random field $Z_{i\in\mathbb{Z}^d}$, i.e. $X_i = f(Z_i)$. We assume that for any $\Lambda_1, \Lambda_2 \subseteq \mathbb{Z}^d$, and any $A_1 \in \sigma(R_{i\in\Lambda_1})$, $A_2 \in \sigma(R_{j\in\Lambda_2})$, $T_1 \in \sigma(Z_{i\in\Lambda_1})$ and $T_2 \in \sigma(Z_{j\in\Lambda_2})$,*

$$\mathbb{P}(A_1A_2|T_1T_2) = \mathbb{P}(A_1|T_1)\mathbb{P}(A_2|T_2). \tag{2.14}$$

Then we have

$$\alpha^{RX}_{k,l}(m) \le \alpha^{Z}_{k,l}(m),$$

where k, l, m are these parameters used in the Definition 2.1.

Proof. By using Theorem 2.1, we consider the equivalent expression of the strong mixing coefficient of RX. Let h_1 and h_2 be any functions which satisfy $0 \le h_1, h_2 \le 1$. Then we have

$$\alpha^{RX}_{k,l}(m) = \sup_{0\le h_1,h_2\le 1} \Big| Cov\Big(h_1(\{R_iX_i\}_{i\in\Lambda_1}), h_2(\{R_jX_j\}_{j\in\Lambda_2})\Big)\Big|.$$

Let $\phi(\cdot)$ be the density function defined in (2.13). We note that

$$\begin{aligned}
&\mathbb{E}\Big(h_1(\{R_iX_i\}_{i\in\Lambda_1})h_2(\{R_jX_j\}_{j\in\Lambda_1})\Big)\\
=\ &\iint \mathbb{E}\Big(h_1(\{R_iX_i\}_{i\in\Lambda_1})h_2(\{R_jX_j\}_{j\in\Lambda_2})\Big|\{Z_i = z_i\}_{i\in\Lambda_1}, \{Z_j = z_j\}_{j\in\Lambda_2}\Big)\\
&\qquad\qquad \phi(z_{i\in\Lambda_1\cup\Lambda_2})dz_{i\in\Lambda_1}dz_{i\in\Lambda_2},
\end{aligned}$$

and, for $r_i \in \{0,1\}$, using assumption (H0), we have

$$\begin{aligned}
&\mathbb{E}\Big(h_1(\{R_iX_i\}_{i\in\Lambda_1})h_2(\{R_jX_j\}_{j\in\Lambda_2})\Big|\{Z_i = z_i\}_{i\in\Lambda_1}, \{Z_j = z_j\}_{j\in\Lambda_2}\Big)\\
=\ &\sum_{r_{i\in\Lambda_1}\in\{0,1\}^k}\ \sum_{r_{j\in\Lambda_2}\in\{0,1\}^l} h_1(\{r_if(z_i)\}_{i\in\Lambda_1})h_2(\{r_jf(z_j)\}_{j\in\Lambda_1}) \times\\
&\qquad \mathbb{P}\Big(\{R_i = r_i\}_{i\in\Lambda_1}, \{R_j = r_j\}_{j\in\Lambda_2}\Big|\{Z_i = z_i\}_{i\in\Lambda_1}, \{Z_j = z_j\}_{j\in\Lambda_2}\Big)\\
=\ &\sum_{r_{i\in\Lambda_1}\in\{0,1\}^k}\ \sum_{r_{j\in\Lambda_2}\in\{0,1\}^l} h_1(\{R_if(z_i)\}_{i\in\Lambda_1})h_2(\{R_jf(z_j)\}_{j\in\Lambda_1}) \times\\
&\qquad \mathbb{P}\Big(\{R_i = r_i\}_{i\in\Lambda_1}\Big|\{Z_i = z_i\}_{i\in\Lambda_1}\Big)\mathbb{P}\Big(\{R_j = r_j\}_{j\in\Lambda_2}\Big|\{Z_j = z_j\}_{j\in\Lambda_2}\Big)\\
=\ &\sum_{r_{i\in\Lambda_1}\in\{0,1\}^k} h_1(\{R_if(z_i)\}_{i\in\Lambda_1})\mathbb{P}\Big(\{R_i = r_i\}_{i\in\Lambda_1}\Big|\{Z_i = z_i\}_{i\in\Lambda_1}\Big) \times\\
&\qquad \sum_{r_{j\in\Lambda_2}\in\{0,1\}^l} h_2(\{R_jf(z_j)\}_{j\in\Lambda_1}\mathbb{P}\Big(\{R_j = r_j\}_{j\in\Lambda_2}\Big|\{Z_j = z_j\}_{j\in\Lambda_2}\Big).
\end{aligned}$$

Since

$$\sum_{r_{i\in\Lambda_1}\in\{0,1\}^k} \mathbb{P}\left(\{R_i=r_i\}_{i\in\Lambda_1}\,\middle|\,\{Z_i=z_i\}_{i\in\Lambda_1}\right)=1$$

and

$$\sum_{r_{j\in\Lambda_2}\in\{0,1\}^l} \mathbb{P}\left(\{R_j=r_j\}_{j\in\Lambda_2}\,\middle|\,\{Z_j=z_j\}_{j\in\Lambda_2}\right)=1,$$

we may define

$$\tilde{h}_1(z_{i\in\Lambda_1})=\sum_{r_{i\in\Lambda_1}\in\{0,1\}^k} h_1(\{R_if(z_i)\}_{i\in\Lambda_1})\mathbb{P}\left(\{R_i=r_i\}_{i\in\Lambda_1}\,\middle|\,\{Z_i=z_i\}_{i\in\Lambda_1}\right)$$

and

$$\tilde{h}_2(z_{j\in\Lambda_2})=\sum_{r_{j\in\Lambda_2}\in\{0,1\}^l} h_2(\{R_jf(z_j)\}_{j\in\Lambda_1})\mathbb{P}\left(\{R_j=r_j\}_{j\in\Lambda_2}\,\middle|\,\{Z_j=z_j\}_{j\in\Lambda_2}\right).$$

Therefore we have

$$\mathbb{E}\left(h_1(\{R_iX_i\}_{i\in\Lambda_1})h_2(\{R_jX_j\}_{j\in\Lambda_1})\right)=\iint \tilde{h}_1(z_{i\in\Lambda_1})\tilde{h}_2(z_{j\in\Lambda_2})\phi(z_{i\in\Lambda_1\cup\Lambda_2})dz_{i\in\Lambda_1}dz_{i\in\Lambda_2},$$

and, similarly, we have

$$\mathbb{E}\left(h_1(\{R_iX_i\}_{i\in\Lambda_1})\right)=\int \tilde{h}_1(z_{i\in\Lambda_1})\phi(z_{i\in\Lambda_1})dz_{i\in\Lambda_1}$$

and

$$\mathbb{E}\left(h_2(\{R_jX_j\}_{j\in\Lambda_1})\right)=\int \tilde{h}_2(z_{j\in\Lambda_2})\phi(z_{i\in\Lambda_2})dz_{i\in\Lambda_2}.$$

Now we have

$$\alpha^{RX}_{k,l}(m)\le\sup_{0\le\tilde{h}_1,\tilde{h}_2\le 1}\left|Cov\left(\tilde{h}_1(z_{i\in\Lambda_1}),\tilde{h}_2(z_{j\in\Lambda_2}))\right)\right|=\alpha^{Z}_{k,l}(m).$$

This completes the proof. □

Remark: The conditional independence of (1.5) may be regarded as an example of the assumption (H0) in Section 3.2, e.g. we may assume $\mathbb{P}(R_i=1|Z_i)=g(Z_i)$ and $\mathbb{P}(R_i=1,R_j=1|Z_i,Z_j)=g(Z_i)g(Z_j)$.

Let Λ_1 and Λ_2 be subsets in the definition of $\alpha^{R}_{k,l}(m)$, $\alpha^{RX}_{k,l}(m)$, $\alpha^{X}_{k,l}(m)$ and $\alpha_{k,l}(X|R;m)$, $A_1\in\sigma(\{R_iX_i\}_{i\in\Lambda_1})$ and $A_2\in\sigma(\{R_iX_i\}_{i\in\Lambda_2})$. For the indicating random field, R, we define the event

$$R^{\Lambda}_T=\{R_i=1,i\in T\subseteq\Lambda,R_j=0,j\in\Lambda-T\},\quad \Lambda\subset\mathbb{Z}^d,$$

$$A_{1,T_1}=A_1\cap R^{\Lambda_1}_T \text{ and } A_{2,T_2}=A_2\cap R^{\Lambda_2}_T.$$

Because $R_T^\Lambda \cap R_{T'}^\Lambda = \emptyset$ for any $T \neq T'$, $T, T' \in \Lambda_1$ or $T, T' \in \Lambda_2$, it means we have partitions of the whole space Ω, i.e.

$$\bigcup_{T \subseteq \Lambda_1} R_T^{\Lambda_1} = \Omega = \bigcup_{T \subseteq \Lambda_2} R_T^{\Lambda_2}.$$

To simplify the notations, we use R_{T_1} and R_{T_2} to stand for $R_T^{\Lambda_1}$ and $R_T^{\Lambda_2}$ respectively. Then we have

$$A_1 = A_1 \cap \Omega = A_1 \cap \left(\bigcup_{T_1 \subseteq \Lambda_1} R_{T_1} \right) = \bigcup_{T_1 \subseteq \Lambda_1} (A_1 \cap R_{T_1}) = \bigcup_{T_1 \subseteq \Lambda_1} A_{1,T_1},$$

and similarly

$$A_2 = \bigcup_{T_2 \subseteq \Lambda_2} A_{2,T_2}.$$

Therefore

$$\mathbb{P}(A_1) = \mathbb{P}(\bigcup_{T_1 \subseteq \Lambda_1} A_{1,T_1}) = \sum_{T_1 \subseteq \Lambda_1} \mathbb{P}(A_{1,T_1}),$$

$$\mathbb{P}(A_2) = \mathbb{P}(\bigcup_{T_2 \subseteq \Lambda_2} A_{2,T_2}) = \sum_{T_2 \subseteq \Lambda_2} \mathbb{P}(A_{2,T_2}),$$

$$\mathbb{P}(A_1)\mathbb{P}(A_2) = \sum_{T_1 \subseteq \Lambda_1} \sum_{T_2 \subseteq \Lambda_2} \mathbb{P}(A_{1,T_1})\mathbb{P}(A_{2,T_2}),$$

and

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}\left((\bigcup_{T_1 \subseteq \Lambda_1} A_{1,T_1}) \cap (\bigcup_{T_2 \subseteq \Lambda_2} A_{2,T_2}) \right) = \sum_{T_1 \subseteq \Lambda_1} \sum_{T_2 \subseteq \Lambda_2} \mathbb{P}(A_{1,T_1} \cap A_{2,T_2}).$$

If we have dependent information between random fields R and X, say $\alpha_{k,l}(X|R;m)$, then we have

$$\begin{aligned}
|\mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1)\mathbb{P}(A_2)| &\leq \sum_{T_1 \subseteq \Lambda_1} \sum_{T_2 \subseteq \Lambda_2} |\mathbb{P}(A_{1,T_1} \cap A_{2,T_2}) - \mathbb{P}(A_{1,T_1})\mathbb{P}(A_{2,T_2})| \\
&= \sum_{T_1 \subseteq \Lambda_1} \sum_{T_2 \subseteq \Lambda_2} |\mathbb{P}(A_1 \cap A_2 | R_{T_1} \cap R_{T_2})\mathbb{P}(R_{T_1} \cap R_{T_2}) \\
&\qquad - \mathbb{P}(A_1|R_{T_1})\mathbb{P}(R_{T_1})\mathbb{P}(A_2|R_{T_2})\mathbb{P}(R_{T_2})| \\
&\leq \sum_{T_1 \subseteq \Lambda_1} \sum_{T_2 \subseteq \Lambda_2} \mathbb{P}(A_1 \cap A_2 | R_{T_1} \cap R_{T_2}) \\
&\qquad \times |\mathbb{P}(R_{T_1} \cap R_{T_2}) - \mathbb{P}(R_{T_1})\mathbb{P}(R_{T_2})| \\
&\quad + \sum_{T_1 \subseteq \Lambda_1} \sum_{T_2 \subseteq \Lambda_2} \mathbb{P}(R_{T_1})\mathbb{P}(R_{T_2}) \\
&\qquad \times |\mathbb{P}(A_1 \cap A_2 | R_{T_1} \cap R_{T_2}) - \mathbb{P}(A_1|R_{T_1})\mathbb{P}(A_2|R_{T_2})| \\
&\leq \sum_{T_1 \subseteq \Lambda_1} \sum_{T_2 \subseteq \Lambda_2} [\alpha_{k,l}^R(m) + \alpha_{k,l}(X|R;m)]
\end{aligned}$$

$$= \quad 2^{k+l}[\alpha^R_{k,l}(m) + \alpha_{k,l}(X|R;m)].$$

Then we have

Lemma 2.7. *Let $R_{i\in\mathbb{Z}^d}$ be a random field, R_i be defined by (2.10), $\alpha^R_{k,l}(m)$, $\alpha^{RX}_{k,l}(m)$ and $\alpha^X_{k,l}(m)$ be strong mixing coefficients of random fields $R_{i\in\mathbb{Z}^d}$, $\{R_iX_i\}_{i\in\mathbb{Z}^d}$ and $X_{i\in\mathbb{Z}^d}$ respectively. If the conditional strong mixing coefficient of X under the condition R is defined by $\alpha_{k,l}(R|X;m)$, then we have*

$$\alpha^{RX}_{k,l}(m) \le 2^{k+l}\left(\alpha_{k,l}(R|X;m) + \alpha^X_{k,l}(m)\right). \tag{2.15}$$

Lemma 2.8. *Let a random field $Z_{i\in\mathbb{Z}^d}$ be independent of another random field $\epsilon_{i\in\mathbb{Z}^d}$, f be a Borel measurable function and $X_i = f(Z_i, \epsilon_i)$. Then $X_{i\in\mathbb{Z}^d}$ is a random field and*

$$\alpha^X_{k,l}(m) \le \alpha^Z_{k,l}(m) + \alpha^\epsilon_{k,l}(m), \tag{2.16}$$

where k, l, m are these parameters used in the Definition 2.1.

Proof. Since Z_i and ϵ_i are random variables, the same discussion of (2.3) concludes that $X_{i\in\mathbb{Z}^d}$ is a random field.

Let $\Lambda_1, \Lambda_2 \subseteq \mathbb{Z}^d$, $k = |\Lambda_1|$, $l = |\Lambda_2|$, using Theorem 2.1 and the independent property, we have

$$\begin{aligned}
& \alpha^X_{k,l}(m) = \alpha^{f(Z,\epsilon)}_{k,l}(m) \\
= & \sup_{0\le h_1,h_2\le 1} |Cov(h_1[f(Z_{t\in\Lambda_1}, \epsilon_{t\in\Lambda_1})], h_2[f(Z_{s\in\Lambda_2}, \epsilon_{s\in\Lambda_2})])| \\
= & \sup_{0\le v_1,v_2\le 1} |Cov(v_1(Z_{t\in\Lambda_1}, \epsilon_{t\in\Lambda_1}), v_2(Z_{s\in\Lambda_2}, \epsilon_{s\in\Lambda_2}))| \\
= & \sup_{0\le v_1,v_2\le 1} \Bigg| \iiiint v_1v_2\phi(z_{t\in\Lambda_1}, z_{s\in\Lambda_2})\phi(\epsilon_{t\in\Lambda_1}, \epsilon_{s\in\Lambda_2}) dz_{t\in\Lambda_1} dz_{s\in\Lambda_2} d\epsilon_{t\in\Lambda_1} d\epsilon_{s\in\Lambda_2} \\
& \qquad - \iiiint v_1v_2\phi(z_{t\in\Lambda_1})\phi(z_{s\in\Lambda_2})\phi(\epsilon_{t\in\Lambda_1})\phi(\epsilon_{s\in\Lambda_2}) dz_{t\in\Lambda_1} dz_{s\in\Lambda_2} d\epsilon_{t\in\Lambda_1} d\epsilon_{s\in\Lambda_2} \Bigg| \\
\le & \sup_{0\le v_1,v_2\le 1} \iint \left| \iint v_1v_2\phi(z_t, z_s) dz_{t\in\Lambda_1} dz_s - \iint v_1v_2\phi(z_t)\phi(z_s) dz_t dz_s \right| \phi(\epsilon_t, \epsilon_s) d\epsilon_t d\epsilon_s \\
& \quad + \sup_{0\le v_1,v_2\le 1} \iint \left| \iint v_1v_2\phi(\epsilon_t, \epsilon_s) d\epsilon_t d\epsilon_s - \iint v_1v_2\phi(\epsilon_t)\phi(\epsilon_s) d\epsilon_t d\epsilon_s \right| \phi(z_t)\phi(z_s) dz_t dz_s \\
= & \alpha^Z_{k,l}(m) + \alpha^\epsilon_{k,l}(m),
\end{aligned}$$

where h_1 and h_2 are the same functions in Theorem 2.1, ϕ is defined in (2.13), v_1 and v_2 are introduced new functions which have the same property as h_1 and h_2. On the right hand side of the above inequality, we use z_t to stand for $z_{t\in\Lambda_1}$ and similarly for $z_{s\in\Lambda_2}$, $\epsilon_{t\in\Lambda_1}$, $\epsilon_{s\in\Lambda_2}$. Then this lemma is proved. □

Remark: If we put $f(X_i, R_i) = R_iX_i$, then f is a measurable function. Furthermore, the

independence property between X and R, by using Lemma 2.8, implies

$$\alpha_{k,l}^{RX}(m) \leq \alpha_{k,l}^{X}(m) + \alpha_{k,l}^{R}(m),$$

which provides a lower upper bound compared to Lemma **??**. In the case of ϵ being a white noise random field in Lemma 2.8, we have $\alpha_{k,l}^{\epsilon}(m) = 0$. Hence $\alpha_{k,l}^{X}(m) = \alpha_{k,l}^{f(Z,\epsilon)}(m) \leq \alpha_{k,l}^{Z}(m)$. Therefore, if R is independent of X, we have

$$\alpha_{k,l}^{RX}(m) = \alpha_{k,l}^{Rf(Z,\epsilon)}(m) \leq \alpha_{k,l}^{Z}(m) + \alpha_{k,l}^{R}(m).$$

By the similar discussion in the above lemma, for conditional dependent random fields, we have the following lemma.

Lemma 2.9. *For any two random fields, X and Y, we have*

$$\alpha_{k,l}^{X}(m) \leq \alpha_{k,l}(X|Y;m) + \alpha_{k,l}^{Y}(m). \tag{2.17}$$

Proof. Let Λ_1 and Λ_2 be index sets in the definition of $\alpha_{k,l}^{X}(m)$, $\alpha_{k,l}(X|Y;m)$ and $\alpha_{k,l}^{Y}(m)$, $A \in \sigma(\{X_i\}_{i\in\Lambda_1})$ and $B \in \sigma(\{X_i\}_{i\in\Lambda_2})$. Let $\phi(y_{i\in\Lambda})$ be the joint distribution of $Y_{i\in\Lambda}$, $\Lambda \in \mathbb{Z}^d$. Then we have

$$\begin{aligned}
\alpha_{k,l}^{X}(m) &= \sup_{A,B} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)| \\
&= \sup_{A,B} \Bigg| \iint \mathbb{P}(A\cap B|\{Y_i = y_i\}_{i\in\Lambda_1}, \{Y_j = y_j\}_{j\in\Lambda_2})\phi(y_{i\in\Lambda_1\cup\Lambda_2})dy_{i\in\Lambda_1}dy_{i\in\Lambda_2} \\
&\qquad - \int \mathbb{P}(A|\{Y_i = y_i\}_{i\in\Lambda_1})\phi(y_{i\in\Lambda_1})dy_{i\in\Lambda_1} \int \mathbb{P}(B\{Y_j = y_j\}_{j\in\Lambda_2})\phi(y_{j\in\Lambda_2})dy_{j\in\Lambda_2} \Bigg| \\
&= \sup_{A,B} \Bigg| \iint \Big[\mathbb{P}(A\cap B|\{Y_i = y_i\}_{i\in\Lambda_1}, \{Y_j = y_j\}_{j\in\Lambda_2}) \\
&\qquad - \mathbb{P}(A|\{Y_i = y_i\}_{i\in\Lambda_1})\mathbb{P}(B\{Y_j = y_j\}_{j\in\Lambda_2}) \Big] \phi(y_{i\in\Lambda_1\cup\Lambda_2})dy_{i\in\Lambda_1}dy_{i\in\Lambda_2} \\
&\qquad + \iint \mathbb{P}(A|\{Y_i = y_i\}_{i\in\Lambda_1})\mathbb{P}(B\{Y_j = y_j\}_{j\in\Lambda_2}) \\
&\qquad\qquad \Big[\phi(y_{i\in\Lambda_1\cup\Lambda_2}) - \phi(y_{i\in\Lambda_1})\phi(y_{i\in\Lambda_2}) \Big] dy_{i\in\Lambda_1}dy_{i\in\Lambda_2} \Bigg| \\
&\leq \alpha_{k,l}(X|Y;m) + \alpha_{k,l}^{Y}(m),
\end{aligned}$$

where $h_1(y_{i\in\Lambda_1}) = \mathbb{P}(A|\{Y_i = y_i\}_{i\in\Lambda_1})$ and $h_2(y_{j\in\Lambda_2}) = \mathbb{P}(B|\{Y_j = y_j\}_{j\in\Lambda_2})$. This completes the proof. □

Chapter 3

Central limit theorems

Central limit theorems play a crucial role in statistical theory and applications. Whenever we set up a new sampling strategy or a new estimator, CLTs are basic tools to ensure the analysis is reliable and the precision of estimators is measurable with confidence intervals. In this chapter, we provide CLTs for the strategy in Chapter 1 on strong mixing random fields.

The CLT with strong mixing coefficients was first started by Rosenblatt in his paper [60] in 1956. His work was intuitively inspired by Bernstein's work [3] and Hopf's ergotic theory in [33]. It also inspired many followed works on asymptotic properties with strong mixing dependences. For example, in [71] and [72], the authors considered limit theorems for random additive functions with the strong mixing condition and a further assumption on the variances of quantities. In [38], for stationary random processes, the relationship between the maximal correlation coefficient and the strong mixing coefficient was discussed, as well as the necessary and sufficient condition of a stationary Gaussian process that possesses the property of strong mixing. In [62], the author introduced a CLT under the condition of the restricted strong mixing coefficient.

In 1982, by using Stein's Lemma of [68], in the paper [7], the author, Bolthausen, proved a CLT for high dimensional random fields with the stationary assumption. There are some other papers using stationary assumptions such as [10], [47] and [30]. The author of the paper [21], Ekström, considered CLTs for non-stationary strong mixing sequences but with triangular array settings. In 1995, along the same line of the proof of Bolthausen's work [7], Guyon proved a CLT, Theorem 1.3 in his book [28], for non-stationary high dimensional random fields. This theorem is our main tool to set up asymptotics for the indicated sampling method.

3.1 CLTs on random fields

Let $X = X_{i\in\mathbb{Z}^d}$ be a random field, $\mathbb{E}(X_i) = 0$,

$$S_n = \sum_{i\in D_n} X_i, \quad \sigma_n^2 = Var(S_n), \tag{3.1}$$

where $n \in \mathbb{Z}^+$, $D_n \in \mathbb{Z}^d$. $|D_n|$ increases strictly, and $|D_n|\to \infty$ as $n \to \infty$. Let $\alpha_{k,l}^X(m)$ be strong mixing coefficient of random field X, which was introduced in Definition 2.1.

The following theorem, Theorem 3.1, is an important tool in proving our results. This theorem is a generalization from stationary random fields to non-stationary random fields of the Theorem in [7]. In these proofs, Stein's Lemma and the truncation technique are used. We provide some details on Stein's Lemma and the truncation technique in the Appendix.

Theorem 3.1. (Theorem 3.3.1 in [28]) *If X satisfies*

$$\sum_{m\geq 1} m^{d-1}\alpha_{k,l}^X(m) < \infty, \quad k + l \leq 4, \tag{H1}$$

$$\alpha_{1,\infty}^X(m) = o(m^{-d}), \tag{H2}$$

$$\exists\delta > 0 \quad s.t. \quad \|X\|_{2+\delta}< \infty, \quad \sum_{m\geq 1} m^{d-1}\left[\alpha_{1,1}^X(m)\right]^{\frac{\delta}{2+\delta}} < \infty, \tag{H3}$$

then

$$\limsup_n \frac{1}{|D_n|}\sum_{i,j\in D_n} |Cov(X_i, X_j)|< \infty.$$

If we assume additionally that

$$\liminf_n \frac{\sigma_n^2}{|D_n|} > 0, \tag{H4}$$

then we have

$$\frac{S_n}{\sigma_n} \xrightarrow{\mathcal{D}} N(0, 1).$$

Remark: The proof of this theorem can be found in [28], which is along the same lines as the proof in [7]. For details, see the Appendix. For the assumption (H3), the coefficient m^{d-1} is derived from the estimation of the number of the pairs, X_i and X_j, where $i \in \mathbb{Z}^d$ is fixed and $|i - j|= m$. If we define this estimation in the same way as we do in Chapter 4, i.e. define $h(m) = \#\{j : |i - j|= m, i \in \Lambda_1, j \in \Lambda_2\}$, then we will have the same result with m^{d-1} been replaced by $h(m)$.

We introduce an indicating random field, R, to the random field X, in Theorem 3.1. Then we have the following result.

Theorem 3.2. *Let $R_{i\in\mathbb{Z}^d}$ be introduced as in (2.10), R be independent of $X_{i\in\mathbb{Z}^d}$. Assume R*

satisfies (H1) and (H2). Let X satisfy (H1)–(H4), and for the δ in (H3) of Theorem 3.1, assume that R also satisfies (H3), i.e.

$$\sum_{m\geq 1} m^{d-1}\left[\alpha^{R}_{k,l}(m)\right]^{\frac{\delta}{2+\delta}} < \infty. \tag{3.2}$$

We set

$$S_n = \sum_{i\in D_n} R_t X_t, \quad \sigma_n^2 = Var(S_n).$$

If

$$\liminf_n \frac{\sigma_n^2}{|D_n|} > 0, \tag{3.3}$$

then we have

$$\frac{S_n}{\sigma_n} \xrightarrow{\mathcal{D}} N(0,1). \tag{3.4}$$

Proof. The proof follows using Theorem 3.1. We set $f(R_i, X_i) = R_i X_i$ in Lemma 2.8, then for any k and l, the strong mixing coefficient of RX is bounded by that of R and X, i.e.

$$\alpha^{RX}_{k,l}(m) \leq \alpha^{X}_{k,l}(m) + \alpha^{R}_{k,l}(m),$$

which ensures (H1)–(H4) are satisfied for RX. This completes the proof. □

Similarly, let $a_0, a_1 \in \mathbb{R}$, $a_1 \neq 0$ and assume that

$$c_i \text{ have Bernoulli distributions with two possible values, } a_0 \text{ and } a_1, \text{ for all } i \in \mathbb{Z}^d. \tag{3.5}$$

i.e. c_i is a generalization of R_i. We set $f(c_i, X_i) = c_i X_i$. Because $f(c_i, X_i)$ is a continuous function, it is a Borel measurable function. Therefore, if $c_{i\in\mathbb{Z}^d}$ is independent to $X_{i\in\mathbb{Z}^d}$, Lemma 2.8 gives:

$$\alpha^{cX}_{k,l}(m) \leq \alpha^{c}_{k,l}(m) + \alpha^{X}_{k,l}(m). \tag{3.6}$$

Then with a similar discussion in the previous proofs, we have the following theorem, which is a general version of Theorem 4.39 for superpopulations.

Theorem 3.3. *Let $X_{i\in\mathbb{Z}^d}$ be independent to $c_{i\in\mathbb{Z}^d}$ and assume $c_{i\in\mathbb{Z}^d}$ satisfies (H1) and (H2). Let $X_{i\in\mathbb{Z}^d}$ satisfy (H1)–(H3). For the δ in (H3) of X, assume that*

$$\sum_{m\geq 1} m^{d-1}\left[\alpha^{c}_{k,l}(m)\right]^{\frac{\delta}{2+\delta}} < \infty. \tag{3.7}$$

We set

$$S_n = \sum_{i\in D_n} c_t X_t, \quad \sigma_n^2 = Var(S_n).$$

If

$$\liminf_n \frac{\sigma_n^2}{|D_n|} > 0, \tag{3.8}$$

then we have

$$\frac{S_n}{\sigma_n} \xrightarrow{\mathcal{D}} N(0,1).$$

3.2 CLTs on conditionally independent random fields

To study the asymptotics of the new strategy, which is mentioned in Chapter 1, we introduce four random fields $\{X_t\}$, $\{Z_t\}$, $\{\epsilon_t\}$ and $\{R_t\}$, $t \in \mathbb{Z}^d$. Here R_t is defined as it is in (2.10). Let $X_{t\in\mathbb{Z}^d}$ be a focal random field driven by two extra random fields, $Z_{t\in\mathbb{Z}^d}$ and $\epsilon_{t\in\mathbb{Z}^d}$, i.e.

$$X_t = f(Z_t, \epsilon_t). \tag{3.9}$$

The assumption (2.14) in Lemma 2.6 is the basic assumption for our general sampling method in Chapter 1. This assumption is adopted by Theorem 3.4. We set $\Lambda_1, \Lambda_2 \subseteq \mathbb{Z}^d$, $A_1 \in \sigma(R_{i\in\Lambda_1})$, $A_2 \in \sigma(R_{j\in\Lambda_2})$, $T_1 \in \sigma(Z_{i\in\Lambda_1})$ and $T_2 \in \sigma(Z_{j\in\Lambda_2})$. Then, to integrate the labels of our key assumptions in Theorem 3.1, the before assumption (2.14) in Lemma 2.6 and the expression (2.12) is described by

$$\mathbb{P}(A_1A_2|T_1T_2) = \mathbb{P}(A_1|T_1)\mathbb{P}(A_2|T_2). \tag{H0}$$

Theorem 3.4. *Let (H0) be satisfied, and $X_i = f(Z_i, \epsilon_i)$. We assume there exists $\delta > 0$, such that the strong mixing coefficient, $\alpha^Z_{k,l}(m)$, of $Z_{t\in\mathbb{Z}^d}$ satisfies (H1)–(H3). Furthermore, for the same $\delta > 0$, we assume $\|\epsilon\|_{2+\delta} < \infty$, and there exist constants $K_0, K_1, K_2 > 0$ such that*

$$\sup_{i\in\mathbb{Z}^d} \|f(Z_i, \epsilon_i)\|_{2+\delta} \leq K_0 + K_1\|Z\|_{2+\delta} + K_2\|\epsilon\|_{2+\delta}. \tag{3.10}$$

We set

$$S_n = \sum_{i\in D_n} R_t X_t, \quad \sigma_n^2 = Var(S_n). \tag{3.11}$$

Let

$$\liminf_n \frac{\sigma_n^2}{|D_n|} > 0. \tag{3.12}$$

Then we have

$$\frac{S_n}{\sigma_n} \xrightarrow{\mathcal{D}} N(0,1).$$

Proof. To apply Theorem 3.1, we check the conditions of that theorem.

Lemma 2.6 implies $\alpha_{k,l}^{RX}(m) \leq \alpha_{k,l}^{Z}(m)$. Therefore, we have

$$\sum_{m\geq 1} m^{d-1}\alpha_{k,l}^{RX}(m) \leq \sum_{m\geq 1} m^{d-1}\alpha_{k,l}^{Z}(m) < \infty, \quad k+l \leq 4,$$

and $\alpha_{k,\infty}^{RX}(m) = o(m^{-d})$. Therefore $(H1)$ and $(H2)$ are satisfied. The assumption (3.10) and the bounded ϵ imply

$$\|RX\|_{2+\delta} \quad \leq \quad \|X\|_{2+\delta} < \infty.$$

Since $\alpha_{k,l}^{RX}(m) \leq \alpha_{k,l}^{Z}(m)$, condition (H3) of Z also implies

$$\sum_{m\geq 1} m^{d-1}[\alpha_{1,1}^{RX}(m)]^{\frac{\delta}{2+\delta}} \leq \sum_{m\geq 1} m^{d-1}[\alpha_{1,1}^{Z}(m)]^{\frac{\delta}{2+\delta}} < \infty,$$

which means random field RX also satisfies (H3). Equation (3.12) directly satisfies (H4). This completes the proof. □

Remark: The assumption (3.10) is in the general case of f satisfying some smoothness conditions, such as the following result.

Proposition 3.5. *Let $Y_{i\in\mathbb{Z}^d}$ and $Z_{i\in\mathbb{Z}^d}$ be random fields, $X_i = f(Y_i, Z_i)$, $i \in \mathbb{Z}^d$. If f has bounded partial derivatives with respect to Y and Z, then there exist constants K_0, K_1 and K_2 such that for all $r \geq 1$, we have $\|X\|_r \leq K_0 + K_1\|Y\|_r + K_2\|Z\|_r$.*

Proof. Using the Taylor's Theorem, for any pair of (Y_t, Z_t), there exist Y_t^ξ and Z_t^ξ such that

$$X_t = f(0,0) + Y_t\frac{\partial f(Y_t^\xi, Z_t^\xi)}{\partial Y_t} + Z_t\frac{\partial f(Y_t^\xi, Z_t^\xi)}{\partial Z_t}.$$

The Minkowski's inequality of Theorem A3 implies

$$\begin{aligned}\|X\|_r &= \sup_{t\in\mathbb{Z}^d}\left\|f(0,0) + Y_t\frac{\partial f(Y_t^\xi, Z_t^\xi)}{\partial Y_t} + Z_t\frac{\partial f(Y_t^\xi, Z_t^\xi)}{\partial Z_t}\right\|_r \\ &\leq \sup_{t\in\mathbb{Z}^d}\left(|f(0,0)| + \left\|Y_t\frac{\partial f(Y_t^\xi, Z_t^\xi)}{\partial Y_t}\right\|_r + \left\|Z_t\frac{\partial f(Y_t^\xi, Z_t^\xi)}{\partial Z_t}\right\|_r\right)\end{aligned}$$

The bounded partial derivatives implies $|f(0,0)| < \infty$. We set $f(0,0) = K_0$, $|\partial f/\partial Y| \leq K_1$ and $|\partial f/\partial Z| \leq K_2$. Then we have

$$\left\|Y_t\frac{\partial f(Y_t^\xi, Z_t^\xi)}{\partial Y_t}\right\|_r \leq \|Y_t\|_r\left|\frac{\partial f(Y_t^\xi, Z_t^\xi)}{\partial Y_t}\right| \leq \|Y_t\|_r K_1$$

and

$$\left\| Z_t \frac{\partial f(Y_t^\xi, Z_t^\xi)}{\partial Z_t} \right\|_r \leq \|Z_t\|_r \left| \frac{\partial f(Y_t^\xi, Z_t^\xi)}{\partial Z_t} \right| \leq \|Z_t\|_r K_2.$$

This completes the proof. □

3.3 CLTs on dependent random fields

If random fields R and X have the conditional dependence defined in (2.5), then we have the following theorem.

Theorem 3.6. *Let $X_{t\in\mathbb{Z}^d}$ be a random field, and $R_{t\in\mathbb{Z}^d}$ be an indicating random field defined in (2.10). Let $\alpha_{k,l}^R(m)$ and $\alpha_{k,l}(X|R;m)$ satisfy (H1), and there exists a $\delta > 0$ such that $\alpha_{k,l}^R(m)$ satisfies (H3), and $\alpha_{k,l}(X|R;m)$ satisfies*

$$\sum_{m\geq 1} m^{d-1} \left[\alpha_{k,l}(X|R;m)\right]^{\frac{\delta}{2+\delta}} < \infty \quad \forall k+l \leq 4.$$

If $\{R_t X_t\}_{t\in\mathbb{Z}^d}$ also satisfies (3.12), S_n and σ_n^2 are defined as in (3.11). Then we have

$$\frac{S_n}{\sigma_n} \xrightarrow{\mathcal{D}} N(0,1).$$

Proof. Lemma 2.9 gives $\alpha_{k,l}^{RX}(m) \leq \alpha_{k,l}^X(m) + \alpha_{k,l}(X|R;m)$. Then it is straightforward result by checking that the random field, RX, satisfies (H1)–(H4) of Theorem 3.1. This completes the proof. □

Chapter 4

Estimation of the variance

The estimation of variances provides not only confidence intervals but also an immediate estimation of the variations of random elements. In this chapter, under the L_2-consistancy, we mainly develop Carlstein's results in [13] and Fuller's Theorem 1.3.2 in the book [26]. Carlstein's results are mainly on one dimensional stationary random processes. Fuller's result is a central limit theorem for independent identical distributed finite populations. After generalizing this CLT to dependent random fields, we set up the variance estimation for dependent finite populations. We generalize these results, with strong mixing conditions, to non-stationary dependent random processes in Section 4.2 and random fields in Section 4.3. Section 4.2 and Section 4.3 have a parallel structure, i.e. the first part of each section presents results with uniform integrability conditions, and the second part considers the results with smooth conditions.

To set up consitent variance estimation on dependent random fields, we introduce a well-divided sub-sampling technique, which is different from the commonly used description of sub-sampled region. For example, in some works such as [66] by Sherman, [22] by Ekström and Sjöstedt de-Luna, [50] by Nordman & Lahiri, and the book [39] by Lahiri, increasing rectangle blocks are used to describe increasing sub-sampled regions, and to avoid strange regions, a further assumption on the length of the boundary of the region has to be introduced. The well-divided sub-sampling technique introduced in this chapter avoided to use these two assumptions. At present, our research focuses on the case of non-overlapping regions. For further consideration on overlapping cases, Lahiri's book [39] is definitely a crucial reference for us, especially Chapter 7 and Chapter 12 in his book.

4.1 Preliminaries

Let $X_{i\in\mathbb{Z}^d}$ be a random field with constant mean, say $\mathbb{E}(X_i) = \mu$, and let $D_n \subseteq \mathbb{Z}^d$ be the index set of the sample. To get a CLT which provides a confidence interval, we need results in the form of

$$\frac{\sum_{i\in D_n}(X_i - \mu)}{\hat{\sigma}} \xrightarrow{D} N(0,1),$$

where $\hat{\sigma}^2$ is an estimator of

$$Var\left(\sum_{i\in D_n}(x_i - \mu)\right).$$

For the case of non-identical mean, i.e. $\mathbb{E}(X_i) = \mu_i$, this kind of CLT also works, i.e.

$$\frac{\sum_{i\in D_n}(X_i - \mu_i)}{\hat{\sigma}} \xrightarrow{D} N(0,1),$$

where $\hat{\sigma}^2$ is an estimator of

$$Var\left(\sum_{i\in D_n}(x_i - \mu_i)\right),$$

if we assume

$$\frac{1}{|D_n|}\sum_{i\in D_n}\mathbb{E}(X_i) \to \mu.$$

Furthermore, we note that the CLT may also in the form of

$$\frac{\frac{1}{\sqrt{|D_n|}}\sum_{i\in D_n}(X_i - \mu_i)}{\sqrt{\frac{Var(S_n)}{|D_n|}}} \to N(0,1).$$

Then for non-identical variances of X_i's, i.e. $Var(X_i) = \sigma_i^2$, it is reasonable to assume

$$\frac{Var(S_n)}{|D_n|} \to \sigma^2,$$

where σ^2 is a constant, $S_n = \sum_{i\in D_n} X_i$. This assumption is introduced in some theorems in Section 4.2 and Section 4.3. It can be relaxed by

$$\liminf_{n\to\infty}\frac{Var(S_n)}{|D_n|} > 0,$$

which implies $\frac{Var(S_n)}{|D_n|}$ might not have a limit. This assumption will be made in the following two sections. In this section, we provide some preliminary results which are used in the following sections.

Lemma 4.1. *Let $\{t_n\}_{n\in\mathbb{Z}^+}$ be a sequence of random variables, θ is a real number. If $\lim_{n\to\infty}\mathbb{E}(t_n) =$*

θ and $\lim_{n\to\infty} Var(t_n) = 0$, then $t_n \xrightarrow{L_2} \theta$.

Proof. This directly follows the definition of convergence in L_2 and the fact of $\mathbb{E}|t_n - \theta|^2 = Var(t_n) + [\mathbb{E}(t_n) - \theta]^2$. □

Definition 4.1. Let T be an arbitrary index set, and let $\{X_i\}_{i\in T}$ be a family of random variables. $\{X_i\}_{i\in T}$ is said to be *uniformly integrable* (abbreviated *u.i.*) iff

$$\lim_{A\to\infty} \sup_i \mathbb{E}|^A X_i| = 0,$$

where ${}^A X_i = X_i \mathbb{1}_{\{|X_i|\geq A\}}$.

In the following results, we will introduce two index sets for a family of random variables, say $\{f_m^l\}_{l\in L, m\in M}$. If we say $\{f_m^l\}_{l\in L, m\in M}$ is u.i., it means $\lim_{A\to\infty} \sup_{l,m} \mathbb{E}|^A f_m^l| = 0$, which is a special case of the above definition.

Theorem 4.2. (Theorem 4.5.3 in [16]) *Let T be an arbitrary index set. The family $\{X_i\}_{i\in T}$ is u.i. iff the following two conditions are satisfied:*

(a) $\mathbb{E}|X_i|$ is bounded in $i \in T$;

(b) For every $\epsilon > 0$, there exists $\lambda(\epsilon) > 0$ s.t. for any $E \in \sigma(X_i)$:

$$\mathbb{P}(E) < \lambda(\epsilon) \quad \Rightarrow \int_E |X_i| d\mathbb{P} < \epsilon \quad \text{for every } i \in T.$$

Theorem 4.3. (Theorem 4.5.4 in [16]) *Let $\{t_n\}_{n\in\mathbb{Z}^+}$ be a random process, $0 < p < \infty$, $t_n \in L_p$ and $t_n \xrightarrow{P} t$. Then the following three propositions are equivalent:*

(i) $\{|t_n|^p\}$ is u.i.;

(ii) $t_n \xrightarrow{L_p} t$;

(iii) $\mathbb{E}|t_n|^p \to \mathbb{E}|t|^p$.

Because $\mathbb{P}(|t_n - t| \geq \epsilon) \leq \epsilon^{-p}\mathbb{E}|t_n - t|^p$, the second proposition in Theorem 4.3 implies $t_n \xrightarrow{P} t$. Then we have the following lemma.

Lemma 4.4. *If $t_n \xrightarrow{L_2} t$, then $t_n \xrightarrow{P} t$.*

Lemma 4.5. (Lemma 1 in [13]) *Let $X_{i\in\mathbb{Z}^d}$, $\Lambda_1, \Lambda_2 \subset \mathbb{Z}^d$, $|\Lambda_1| = k$, $|\Lambda_2| = l$, $m = dist(\Lambda_1, \Lambda_2)$, $\alpha_{k,l}(m)$ be the strong mixing coefficient of X_i and X_j, where $i \in \Lambda_1$, $j \in \Lambda_2$. We assume that $\max_{i\in\Lambda_1\cup\Lambda_2}\{\mathbb{E}X_i^2\} \leq C < \infty$, where C is a constant. Then, for any $A > 0$, we have*

$$|Cov(X_i, X_j)| \leq 4A^2\alpha_{k,l}(m) + 3\sqrt{C}(\sqrt{\mathbb{E}|^A X_i|^2} + \sqrt{\mathbb{E}|^A X_j|^2}).$$

Lemma 4.6. *Let T be an arbitrary index set, $r \geq 1$. If the family of random variables $\{X_i^r\}_{i\in T}$ is u.i., then $\{X_i\}_{i\in T}$ is u.i.*

Proof. This is a straightforward result using the fact that $\mathbb{E}|^A X_i| \leq (\mathbb{E}|^A X_i^r|)^{\frac{1}{r}}$. □

Lemma 4.7. *Let T be an arbitrary index set, $\{X_i\}_{i\in T}$ and $\{Y_i\}_{i\in T}$ be two families of random variables. If $\{Y_i\}_{i\in T}$ is u.i. and $|X_i| \leq |Y_i|$ for all $i \in T$, then $\{X_i\}_{i\in T}$ is u.i..*

Proof. This is because $\mathbb{E}|^A X_i| \leq \mathbb{E}|^A Y_i|$ for any $A \geq 0$. □

Lemma 4.8. *Let $k \in \mathbb{Z}^+$, and T be an arbitrary index set. If $\{X_{1,n}^k\}_{n\in T}$, $\{X_{2,n}^k\}_{n\in T}, \cdots, \{X_{k,n}^k\}_{n\in T}$ are u.i., then $\{X_{1,n}X_{2,n}\cdots X_{k,n}\}_{n\in T}$ is u.i..*

Proof. Let $I = \{1, 2, \cdots, k\}$. Theorem 4.2 implies that there exists a constant C such that $\|X_{i,n}\|_k \leq C$ for all $i \in I$. By using Hölder's inequality k times, we have

$$\|X_{1,n}X_{2,n}\cdots X_{k,n}\|_1 \leq \prod_{i\in I} \|X_{i,n}\|_k \leq C^k.$$

Then, the rest of this proof is to check condition (b) of Theorem 4.2.

We note that, for each $j \in I$, Hölder's inequality also implies

$$\|T_j\|_{\frac{k}{k-1}} = \left\| \prod_{i\in I\setminus\{j\}} X_{i,n} \right\|_{\frac{k}{k-1}} \leq \prod_{i\in I\setminus\{j\}} \|X_{i,n}\|_k \leq C^{k-1}.$$

Since for all $A > 0$,

$$\{|X_{1,n}X_{2,n}\cdots X_{k,n}| > A\} \subseteq \bigcup_{i=1}^{k} \{|X_{i,n}| > A^{\frac{1}{k}}\},$$

then we have

$$\begin{aligned} \sup_n \mathbb{E}^A \Big| \prod_{i\in I} X_{i,n} \Big| &\leq \sup_n \left(\sum_{i\in I} \left(\mathbb{E}^A |X_{i,n}|^k \right)^{\frac{1}{k}} \|T_i\|_{\frac{k}{k-1}} \right) \\ &\leq C^{k-1} \sum_{i\in I} \left(\sup_n \mathbb{E}^{A^{\frac{1}{k}}} |X_{i,n}|^k \right)^{\frac{1}{k}} \\ &\to 0 \quad \text{as} \quad A \to \infty. \end{aligned}$$

This completes the proof. □

Lemma 4.9. *For any two random variables X and Y, we have the following inequalities:*

$$\mathbb{E}^A(|X|+|Y|) \leq 2\mathbb{E}^{\frac{A}{2}}|X| + 2\mathbb{E}^{\frac{A}{2}}|Y|,$$

$$\mathbb{E}^A(|X|+|Y|) \leq \mathbb{E}^{\frac{A}{2}}|X|+\mathbb{E}^{\frac{A}{2}}|Y|+\mathbb{E}|Y|.$$

Proof. We note that $\{|X|+|Y|\geq A\} \subseteq \{|X|> \frac{A}{2}\} \cup \{|Y|\geq \frac{A}{2}\}$.

$$\begin{aligned}
\mathbb{E}^A(|X|+|Y|) &= \int_{|X|+|Y|\geq A}(|X|+|Y|)d\mathbb{P} \\
&\leq \int_{\{|X|\geq \frac{A}{2}\}\cup\{|Y|\geq \frac{A}{2}\}}(|X|+|Y|)d\mathbb{P} \\
&= \int_{\{|X|\geq \frac{A}{2}\}\cup\{|Y|\geq \frac{A}{2}\}}|X|d\mathbb{P} + \int_{\{|X|\geq \frac{A}{2}\}\cup\{|Y|\geq \frac{A}{2}\}}|Y|d\mathbb{P} \\
&= \text{(I)}+\text{(II)}.
\end{aligned}$$

We note that

$$\begin{aligned}
\text{(I)} &\leq \int_{\{|X|\geq \frac{A}{2}\}}|X|d\mathbb{P} + \int_{\{|X|< \frac{A}{2}\}\cap\{|Y|\geq \frac{A}{2}\}\}}|X|d\mathbb{P} \\
&= \mathbb{E}|^{\frac{A}{2}}|X||+\mathbb{E}|^{\frac{A}{2}}|Y||,
\end{aligned}$$

and, with the same argument as (I), we have (II)$\leq \mathbb{E}|^{\frac{A}{2}}|Y||+\mathbb{E}|^{\frac{A}{2}}|X||$. Therefore the first inequality is proved.

If we note (II) $\leq \mathbb{E}|Y|$ and

$$\begin{aligned}
\text{(I)} &= \int_{\{|X|\geq \frac{A}{2}\}}|X|d\mathbb{P} + \int_{\{|Y|\geq \frac{A}{2}\}\cap\{|X|< \frac{A}{2}\}}|X|d\mathbb{P} \\
&\leq \int_{\{|X|\geq \frac{A}{2}\}}|X|d\mathbb{P} + \int_{\{|Y|\geq \frac{A}{2}\}}|Y|d\mathbb{P}.
\end{aligned}$$

Then the second inequality holds. □

Corollary 4.10. *Let $k \in \mathbb{Z}^+$, and T be an arbitrary index set. If $\{X_{1,n}\}_{n\in T}$, $\{X_{2,n}\}_{n\in T}, \cdots, \{X_{k,n}\}_{n\in T}$ are u.i., then $\{X_{1,n} + X_{2,n} + \cdots + X_{k,n}\}_{n\in T}$ is u.i..*

Proof. We inductively use the first inequality in Lemma 4.9 $k-1$ times. Then we have

$$\mathbb{E}^A\Big|\sum_{i=1}^{k} X_{i,n}\Big| \leq \sum_{i=1}^{k-2} 2^i\mathbb{E}^{\frac{A}{2^i}}\Big|X_{i,n}\Big| + 2^{k-1}\mathbb{E}^{\frac{A}{2^{k-1}}}\Big|X_{k-1,n}\Big| + 2^{k-1}\mathbb{E}^{\frac{A}{2^{k-1}}}\Big|X_{k,n}\Big|.$$

This completes the proof. □

Lemma 4.11. *Let $S_{i\in\mathbb{Z}}$ be a random process. If there exists $\delta > 0$ s.t.*

$$\|S_i\|_{1+\delta}\leq M < \infty,$$

then S_i are u.i.

Proof. We only need to check condition (a) and (b) in Theorem 4.2.

For (a), $\mathbb{E}|S_i| \leq \|S\|_{1+\delta} \leq M < \infty$.

For (b), for all $\epsilon > 0$, there exists $\lambda(\epsilon) = (\epsilon/M)^{\frac{1+\delta}{\delta}}$ s.t. for any $E \in \sigma(S_i)$, if $\mathbb{P}(E) < \lambda(\epsilon)$, then

$$\begin{aligned}
\int_E |S_i| d\mathbb{P} &\leq \left(\int_E 1^{\frac{1+\delta}{\delta}}\right)^{\frac{\delta}{1+\delta}} \left(\int_E |S_i|^{1+\delta}\right)^{\frac{1}{1+\delta}} \\
&\leq (\mathbb{P}(E))^{\frac{\delta}{1+\delta}} M \\
&< \frac{\epsilon}{M} M = \epsilon.
\end{aligned}$$

This completes the proof. □

Lemma 4.12. *Let $\alpha(1) \geq \alpha(2) \geq \cdots$ be a decreasing sequence. For any $k, m \in \mathbb{Z}^+$, $k \geq 2$, we have*

$$\sum_{l=0}^{k-2} \sum_{j=l+1}^{k-1} \alpha(jm - lm) \leq \frac{k}{m} \sum_{i=1}^{(k-1)m} \alpha(i),$$

and

$$\sum_{i=1}^{k-1} \alpha(im) \leq \frac{1}{m} \sum_{j=1}^{(k-1)m} \alpha(j).$$

Proof. We note that

$$\sum_{l=0}^{k-2} \sum_{j=l+1}^{k-1} \alpha(jm - lm) = \sum_{i=1}^{k-1} (k-i)\alpha(im) \leq k \sum_{i=1}^{k-1} \alpha(im),$$

and

$$m \sum_{i=1}^{k-1} \alpha(im) = \sum_{i=1}^{k-1} m\alpha(im) \leq \sum_{i=1}^{k-1} \left(\sum_{j=(i-1)m+1}^{im} \alpha(j) \right) = \sum_{j=1}^{(k-1)m} \alpha(j).$$

These two facts complete the proof. □

4.2 Estimators on 1-dimension random processes

In this section, Theorem 4.13 and Theorem 4.14 generalize Carlstein's Theorem 2 and Theorem 3 in [13] respectively. The first subsection carries on the uniform integrability conditions which Carlstein used. In the second subsection, we drop the uniform integrability conditions, and introduce smoothness conditions on estimators.

4.2.1 Estimators with uniformly integrability conditions

Theorem 4.13. (To generalize Carlstein's Theorem 2 in [13] to non–stationary processes) *Let $X_{i\in\mathbb{Z}}$ be a strong mixing process, $\alpha_{\cdot,\cdot}(\cdot)$ be the strong mixing coefficient, $f_m(\cdot)$ be a measurable function. We define*

$$f_m^l = f_m(X_l, X_{l+1}, \cdots, X_{l+m+1}),$$

for all $m > 0$, $m, l, \in \mathbb{Z}$. Let n be the sample size, $n \in \mathbb{Z}^+$, $\{m_n\}_{n\in\mathbb{Z}^+}$ and $k_n = [\frac{n}{m_n}]$ be such that $m_n \to \infty$ and $k_n \to \infty$ as $n \to \infty$. Suppose that $\alpha_{m_n,m_n}(m_n) \to 0$ as $n \to \infty$, and that $\{(f_{m_n}^{lm_n})^2\}$ are u.i. with respect to l and m_n. Let

$$\bar{f}_n = \frac{1}{k_n}\sum_{l=0}^{k_n-1} f_{m_n}^{lm_n}$$

be an estimator. If $\lim_{n\to\infty}\mathbb{E}(\bar{f}_n) = \varphi$ exists, then we have $\bar{f}_n \xrightarrow{L_2} \varphi$.

Proof. We note that

$$\begin{aligned} Var(\bar{f}_n) &= Var\left(\frac{1}{k_n}\sum_{l=0}^{k_n-1} f_{m_n}^{lm_n}\right) \\ &\leq \frac{1}{k_n^2}\left[\sum_{l=0}^{k_n-1}|Cov(f_{m_n}^{lm_n}, f_{m_n}^{lm_n})| + 2\sum_{l=0}^{k_n-2}|Cov(f_{m_n}^{lm_n}, f_{m_n}^{(l+1)m_n})| \right. \\ &\qquad \left. + 2\sum_{l=0}^{k_n-3}\sum_{j=l+2}^{k_n-1}|Cov(f_{m_n}^{lm_n}, f_{m_n}^{jm_n})|\right]. \end{aligned}$$

Since $\{(f_{m_n}^{lm_n})^2\}$ are u.i., Theorem 4.2 implies $\mathbb{E}(f_{m_n}^{lm_n})^2$ are bounded for all l and m_n. Therefore, there exists a constant C such that

$$|Cov(f_{m_n}^{lm_n}, f_{m_n}^{lm_n})| = Var(f_{m_n}^{lm_n}) \leq C$$

for all l and m_n. Then we have

$$\sum_{l=0}^{k_n-1}|Cov(f_{m_n}^{lm_n}, f_{m_n}^{lm_n})| \leq k_n C,$$

and

$$\sum_{l=0}^{k_n-2}|Cov(f_{m_n}^{lm_n}, f_{m_n}^{(l+1)m_n})| \leq \sum_{l=0}^{k_n-2}\sqrt{Var(F_{m_n}^{lm_n})}\sqrt{Var(F_{m_n}^{(l+1)m_n})} \leq (k_n-1)C.$$

For any $A > 0$, by using Lemma 4.5, we have

$$\begin{aligned}
|Cov(f_{m_n}^{lm_n}, f_{m_n}^{jm_n})| &\leq 4A^2\alpha_{m_n,m_n}(m^*) + 3\sqrt{C}\left(\sqrt{\mathbb{E}|^A f_{m_n}^{lm_n}|^2} + \sqrt{\mathbb{E}|^A f_{m_n}^{jm_n}|^2}\right) \\
&\leq 4A^2\alpha_{m_n,m_n}(m_n) + 3\sqrt{C}\left(\sqrt{\mathbb{E}|^A f_{m_n}^{lm_n}|^2} + \sqrt{\mathbb{E}|^A f_{m_n}^{jm_n}|^2}\right) \\
&= B(n, A),
\end{aligned}$$

where m^* is the distance between $f_{m_n}^{lm_n}$ and $f_{m_n}^{jm_n}$, $m^* \geq m_n$ for all $j \geq l+2$. By using the u.i. condition and the assumption on strong mixing coefficient, given $\epsilon > 0$, we choose sufficiently large A so that

$$3\sqrt{C}\left(\sqrt{\mathbb{E}|^A f_{m_n}^{lm_n}|^2} + \sqrt{\mathbb{E}|^A f_{m_n}^{jm_n}|^2}\right) < \epsilon.$$

Then we have

$$B(n, A) < 4A^2\alpha_{m_n,m_n}(m_n) + \epsilon,$$

and

$$\limsup_n B(n, A) = 0.$$

Therefore, we have $|Cov(f_{m_n}^{lm_n}, f_{m_n}^{jm_n})| \to 0$ as $n \to \infty$. Hence

$$\begin{aligned}
Var(\bar{f}_n) &\leq \frac{1}{k_n^2}[k_nC + 2(k_n-1)C + 2k_n^2B(n, A)] \\
&\leq \frac{3C}{k_n} + 2B(n, A) \to 0 \quad \text{as} \quad n \to \infty.
\end{aligned}$$

By using Lemma 4.1, if $\lim_{n\to\infty}\mathbb{E}(\bar{f}_n) = \varphi$ exists, then $\bar{f}_n \xrightarrow{L_2} \varphi$. This completes the proof. □

Example: *(Example of $(f_m^l)^2$ being u.i.)*

Let $X_{i\in\mathbb{Z}}$ be a zero-mean process. We assume there exists a constant $\delta > 0$ such that $\|X\|_{4+2\delta} < \infty$. Let

$$f_m^l = \frac{1}{m}\sum_{i=0}^{m-1} X_{l+i}.$$

Then $(f_m^l)^2$ are u.i.

Proof. Since $i \in \mathbb{Z}$ are one-dimensional integers, the distance j between $\{i_1 + l, i_2 + l\}$ and $\{i_3 + l, i_4 + l\}$ is no more than m, for all $i_1, i_2, i_3, i_4 \in \{0, 1, \cdots, m-1\}$, i.e. all possible values of the distance j are $0, 1, \cdots, m-1$. Therefore, in this case,

$$\frac{1}{m}\sum_{j=0}^{m-1}\alpha_{2,2}^{\frac{\delta}{2+\delta}}(j) \leq \frac{1}{m}m\frac{1}{4} < 1 \quad \text{and} \quad \frac{1}{m}\sum_{j=0}^{m-1}\alpha_{1,1}^{\frac{\delta}{2+\delta}}(j) < 1.$$

By using Lemma 4.11, to prove this example, it is sufficient to show $\|(f_m^l)^2\|_2$ is bounded. We write

$$\begin{aligned}
\mathbb{E}\left|(f_m^l)^2\right|^2 &= \frac{1}{m^4}\sum_{i_1=0}^{m-1}\sum_{i_2=0}^{m-1}\sum_{i_3=0}^{m-1}\sum_{i_4=0}^{m-1} Cov(X_{i_1+l}X_{i_2+l}, X_{i_3+l}X_{i_4+l}) \\
&\qquad + \left(\frac{1}{m^2}\sum_{i_1=0}^{m-1}\sum_{i_2=0}^{m-1} Cov(X_{i_1+l}, X_{i_2+l})\right)^2 \\
&= \text{(I)}+\text{(II)}^2.
\end{aligned}$$

For (II), since

$$\#\Big\{(i_1, i_2) : dist(\{i_1\}, \{i_2\}) = j\Big\} = O(m),$$

Theorem 2.2 implies, for a constant C,

$$\begin{aligned}
\text{(II)} &\leq \frac{1}{m^2}O(m)8\sum_{j\geq 0}\alpha_{1,1}^{\frac{\delta}{2+\delta}}(j)\|X_{i_1+l}\|_{2+\delta}\|X_{i_2+l}\|_{2+\delta} \\
&\leq \frac{C}{m}\sum_{j=0}^{m-1}\alpha_{1,1}^{\frac{\delta}{2+\delta}}(j) \\
&= O(1).
\end{aligned}$$

For (I), let C be a constant, we have $\|X_{i_1+l}X_{i_2+l}\|_{2+\delta}\leq C$, $\|X_{i_3+l}X_{i_4+l}\|_{2+\delta}\leq C$, and

$$\#\Big\{(i_1, i_2, i_3, i_4) : dist(\{i_1, i_2\}, \{i_3, i_4\}) = j\Big\} = O(m^3).$$

Then Theorem 2.2 implies

$$\begin{aligned}
\text{(I)} &\leq \frac{1}{m^4}O(m^3)8\sum_{j\geq 0}\alpha_{2,2}^{\frac{\delta}{2+\delta}}(j)\|X_{i_1+l}X_{i_2+l}\|_{2+\delta}\|X_{i_3+l}X_{i_4+l}\|_{2+\delta} \\
&\leq \frac{C}{m}\sum_{j=0}^{m-1}\alpha_{2,2}^{\frac{\delta}{2+\delta}}(j) \\
&= O(1),
\end{aligned}$$

where C is a new constant. □

Theorem 4.14. (To generalize Carlstein's Theorem 3 in [13] to non–stationary processes) *Let $X_{i\in\mathbb{Z}}$ be a strong mixing process, $S_m(\cdot)$ be a measurable function. We define*

$$S_m^l = S_m^l(X_l, X_{l+1}, \cdots, X_{l+m-1}),$$

$$t_m^l = \sqrt{m}(S_m^l - \mathbb{E}(S_m^l)),$$

for all $m > 0$, $m, l \in \mathbb{Z}$. Let m_n, k_n and n be defined as in Theorem 4.13,

$$\Sigma_n = \frac{1}{k_n} \sum_{l=0}^{k_n-1} (t_{m_n}^{lm_n})^2,$$

$$\bar{S}_n = \frac{1}{k_n} \sum_{l=0}^{k_n-1} S_{m_n}^{lm_n},$$

$$\hat{\sigma}_n^2 = \frac{m_n}{k_n} \sum_{l=0}^{k_n-1} (S_{m_n}^{lm_n} - \bar{S}_n)^2.$$

We assume

1) $\{(t_{m_n}^{lm_n})^4\}$ are u.i. with respect to l and m_n;

2) $\alpha_{m_n,m_n}(m_n) \to 0$ as $n \to \infty$;

3) $\lim_{n\to\infty} \mathbb{E}(\Sigma_n) = \phi$;

4)

$$E_n \equiv \frac{m_n}{k_n} \sum_{l=0}^{k_n-1} \left(\mathbb{E}S_{m_n}^{lm_n} - \mathbb{E}\bar{S}_n\right)^2 \to 0 \quad as \quad n \to \infty.$$

Then we have $\hat{\sigma}_n^2 \xrightarrow{L_2} \phi$.

Proof. We set

$$\bar{t}_n = \frac{1}{k_n} \sum_{l=0}^{k_n-1} t_{m_n}^{lm_n}.$$

Then we have

$$\begin{aligned}
\hat{\sigma}_n^2 &= \frac{m_n}{k_n} \sum_{l=0}^{k_n-1} \left(S_{m_n}^{lm_n} - \mathbb{E}S_{m_n}^{lm_n} + \mathbb{E}S_{m_n}^{lm_n} - \bar{S}_n\right)^2. \\
&= \frac{m_n}{k_n} \sum_{l=0}^{k_n-1} \Bigg[\left(S_{m_n}^{lm_n} - \mathbb{E}S_{m_n}^{lm_n}\right)^2 \\
&\qquad +2\left(S_{m_n}^{lm_n} - \mathbb{E}S_{m_n}^{lm_n}\right)\left(\mathbb{E}S_{m_n}^{lm_n} - \bar{S}_n\right) + \left(\mathbb{E}S_{m_n}^{lm_n} - \bar{S}_n\right)^2\Bigg] \\
&= \frac{m_n}{k_n} \sum_{l=0}^{k_n-1} \left(S_{m_n}^{lm_n} - \mathbb{E}S_{m_n}^{lm_n}\right)^2 \\
&\qquad +2\frac{m_n}{k_n} \sum_{l=0}^{k_n-1} \left(S_{m_n}^{lm_n} - \mathbb{E}S_{m_n}^{lm_n}\right)\left(\mathbb{E}S_{m_n}^{lm_n} - \bar{S}_n\right) + \frac{m_n}{k_n} \sum_{l=0}^{k_n-1} \left(\mathbb{E}S_{m_n}^{lm_n} - \bar{S}_n\right)^2 \\
&= \Sigma_n + 2A_n + B_n,
\end{aligned}$$

where

$$A_n = \frac{m_n}{k_n} \sum_{l=0}^{k_n-1} \left(S_{m_n}^{lm_n} - \mathbb{E}S_{m_n}^{lm_n} \right) \left(\mathbb{E}S_{m_n}^{lm_n} - \bar{S}_n \right),$$

and

$$B_n = \frac{m_n}{k_n} \sum_{l=0}^{k_n-1} \left(\mathbb{E}S_{m_n}^{lm_n} - \bar{S}_n \right)^2.$$

The rest of this proof is to show $\Sigma_n \xrightarrow{L_2} \phi$, $A_n \xrightarrow{L_2} 0$ and $B_n \xrightarrow{L_2} 0$.

First, consider Σ_n.

Because the distance between $S_{m_n}^{lm_n}$ and $S_{m_n}^{jm_n}$ can be regarded as the distance between $t_{m_n}^{lm_n}$ and $t_{m_n}^{jm_n}$, Theorem 4.13 implies

$$\Sigma_n \xrightarrow{L_2} \phi, \quad \text{as} \quad n \to \infty. \tag{4.1}$$

We note the fact that $\mathbb{E}(t_{m_n}^{lm_n}) = 0$ for all l and m_n. Therefore $\mathbb{E}(\bar{t}_n) = 0$ for all $n \in \mathbb{Z}^+$. Then we certainly have $\lim_{n\to\infty} \mathbb{E}(\bar{t}_n) = 0$. Furthermore, we note

$$(\bar{t}_n)^2 = \left(\frac{1}{k_n} \sum_{l=0}^{k_n-1} t_{m_n}^{lm_n} \right)^2 \leq \frac{1}{k_n} \sum_{l=0}^{k_n-1} (t_{m_n}^{lm_n})^2 = \Sigma_n,$$

that is $(\bar{t}_n)^4 \leq \Sigma_n^2$. By using Theorem 4.3, equation (4.1) implies Σ_n^2 are u.i. Then Lemma 4.7 implies $(\bar{t}_n)^4$ are u.i. Thereafter, by using Lemma 4.6, the condition of $(t_{m_n}^{lm_n})^4$ being u.i. implies $(t_{m_n}^{lm_n})^2$ being u.i. Then Theorem 4.13 implies

$$\mathbb{E}|\bar{t}_n - 0|^2 \to 0 \quad \text{as} \quad n \to \infty. \tag{4.2}$$

Therefore $\bar{t}_n \xrightarrow{P} 0$. Since $(\bar{t}_n)^4$ are u.i., Theorem 4.3 implies $\bar{t}_n \xrightarrow{L_4} 0$, which is equivalent to $(\bar{t}_n)^2 \xrightarrow{L_2} 0$.

To proof $B_n \xrightarrow{L_2} 0$, we note

$$\begin{aligned} B_n &= \frac{m_n}{k_n} \sum_{l=0}^{k_n-1} \left(\mathbb{E}S_{m_n}^{lm_n} - \mathbb{E}\bar{S}_n + \mathbb{E}\bar{S}_n - \bar{S}_n \right)^2 \\ &\leq E_n + (\bar{t}_n)^2. \end{aligned}$$

Then, by using Lemma 4.7, assumption 4 and $(\bar{t}_n)^2 \xrightarrow{L_2} 0$, we have $B_n \xrightarrow{L_2} 0$.

To proof $A_n \xrightarrow{L_2} 0$, we note, by using Hölder's inequality,

$$|A_n| \leq \frac{m_n}{k_n} \left[\sum_{l=0}^{k_n-1} \left(S_{m_n}^{lm_n} - \mathbb{E}S_{m_n}^{lm_n} \right)^2 \sum_{l=0}^{k_n-1} \left(\mathbb{E}S_{m_n}^{lm_n} - \bar{S}_n \right)^2 \right]^{\frac{1}{2}}$$

$$= \quad [\Sigma_n B_n]^{\frac{1}{2}}.$$

Since $\Sigma_n \xrightarrow{L_2} \sigma^2$ and $B_n \xrightarrow{L_2} 0$, we have $A_n \xrightarrow{L_2} 0$. This completes the proof. □

Theorem 4.15. *Let $X_{i\in\mathbb{Z}}$ be a zero-mean strong mixing random process. We assume that there exists $\delta > 0$ such that $\|X\|_{2+\delta}= C < \infty$, C is a constant. Let $\alpha_{1,1}(j)$ be the strong mixing coefficient. We suppose*

$$\sum_{j=0}^{\infty} \alpha_{1,1}^{\frac{\delta}{2+\delta}}(j) < \infty.$$

We set

$$S_{m_n}^l = \sum_{i=0}^{m_n-1} X_{l+i}, \quad S_n = \sum_{l=0}^{k_n-1} S_{m_n}^{lm_n},$$

$$\nu_{m_n}^l = \frac{S_{m_n}^l}{\sqrt{m_n}}, \quad \bar{\nu}_n = \frac{1}{k_n}\sum_{l=0}^{k_n-1} \nu_{m_n}^{lm_n},$$

$$\hat{\phi} = \frac{1}{k_n}\sum_{l=0}^{k_n-1} (\nu_{m_n}^{lm_n} - \bar{\nu}_n)^2,$$

where k_n, m_n and n are defined as in Theorem 4.13. If $(\nu_{m_n}^{lm_n})^4$ are u.i. and

$$\lim_{n\to\infty} \frac{Var(S_n)}{n} = \phi, \tag{4.3}$$

then

$$\hat{\phi} \xrightarrow{L_2} \phi.$$

Proof. Let

$$\Sigma_n = \frac{1}{k_n}\sum_{l=0}^{k_n-1} (\nu_{m_n}^{lm_n})^2,$$

then we have

$$\hat{\phi} = \Sigma_n - (\bar{\nu})^2.$$

By the same argument in Theorem 4.14, since $(\nu_{m_n}^{lm_n})^4$ are u.i., and $\mathbb{E}(\bar{\nu}_n) = 0$, we have

$$\mathbb{E}|\bar{\nu}_n|^2 \to 0 \quad \text{as} \quad n \to \infty.$$

For the rest of the proof, we only need to show $\mathbb{E}\Sigma_n \to \phi$, i.e.

$$\lim_{n\to\infty} \frac{1}{k_n}\sum_{l=0}^{k_n-1} \mathbb{E}(\nu_{m_n}^{lm_n})^2 = \phi.$$

We note that

$$\begin{aligned}
\frac{Var(S_n)}{n} &= \frac{1}{m_n k_n}\sum_{l=0}^{k_n-1}\sum_{j=0}^{k_n-1} Cov(S_{m_n}^{lm_n}, S_{m_n}^{jm_n}) \\
&= \frac{1}{m_n k_n}\sum_{l=0}^{k_n-1} Var(S_{m_n}^{lm_n}) + \frac{1}{m_n k_n}\sum_{l\neq j} Cov(S_{m_n}^{lm_n}, S_{m_n}^{jm_n}) \\
&= \text{(I)+(II)}.
\end{aligned}$$

Because (I)$= \frac{1}{k_n}\sum_{l=0}^{k_n-1}\mathbb{E}(\nu_{m_n}^{lm_n})^2$, we only need to show (II) converges to zero. We note that

$$\begin{aligned}
\text{(II)} &\leq \frac{2}{m_n k_n}\sum_{l=0}^{k_n-1}\sum_{j=l+1}^{k_n-1} |Cov(S_{m_n}^{lm_n}, S_{m_n}^{lm_n})| \\
&\leq \frac{2}{m_n k_n}\sum_{l=0}^{k_n-1}\sum_{j=l+1}^{k_n-1}\sum_{p=0}^{m_n-1}\sum_{q=0}^{m_n-1} |Cov(X_{lm_n+p}, X_{jm_n+q})| \\
&\leq \frac{2}{m_n k_n}\sum_{l=0}^{k_n-1}\sum_{j=l+1}^{k_n-1}\sum_{p=0}^{m_n-1}\sum_{q=0}^{m_n-1} 8\cdot\alpha_{1,1}^{\frac{\delta}{2+\delta}}(|jm_n+q-lm_n-p|)\|X\|_{2+\delta}^2 \\
&\leq \frac{16C}{m_n k_n}\sum_{l=0}^{k_n-1}\sum_{j=l+1}^{k_n-1}\sum_{p=0}^{m_n-1}\sum_{q=0}^{m_n-1} \alpha_{1,1}^{\frac{\delta}{2+\delta}}(|(j-l)m_n+q-p|).
\end{aligned}$$

Since the distance of any two random variables is directly driven by their 1–dimensional integer index, we have

$$\begin{aligned}
&\sum_{j=l+1}^{k_n-1}\sum_{p=0}^{m_n-1}\sum_{q=0}^{m_n-1}\alpha(|(j-l)m_n+q-p|) \\
=\ &\sum_{i=1}^{k_n-l-1}\sum_{p=0}^{m_n-1}\sum_{q=0}^{m_n-1}\alpha(|im_n+q-p|) \\
=\ &\alpha(1)+2\alpha(2)+\cdots+m_n\alpha(m_n) \\
&\qquad +(m_n-1)\alpha(m_n+1)+\cdots+2\alpha(2m_n-2)+\alpha(2m_n-1) \\
&+\alpha(m_n+1)+2\alpha(m_n+2)+\cdots+m_n\alpha(2m_n) \\
&\qquad +(m_n-1)\alpha(2m_n+1)+\cdots+2\alpha(3m_n-2)+\alpha(3m_n-1) \\
&+\cdots \\
&+\alpha((k_n-l-2)m_n+1)+2\alpha((k_n-l-2)m_n+2) \\
&\quad +\cdots+m_n\alpha((k_n-l-2)m_n+m_n) \\
&\qquad +(m_n-1)\alpha((k_n-l-1)m_n+1) \\
&\qquad\quad +\cdots+2\alpha((k_n-l)m_n-2)+\alpha((k_n-l)m_n-1)
\end{aligned}$$

$$\leq \quad \alpha(1) + \cdots + m_n\alpha(m_n) + m_n \sum_{i=m_n+1}^{k_n m_n} \alpha(i),$$

where we use $\alpha(\cdot)$ as the abbreviation of $\alpha_{1,1}^{\frac{\delta}{2+\delta}}(\cdot)$.

Therefore,

$$\text{(II)} \leq 16C\frac{1}{m_n}\left[\alpha(1) + \cdots + m_n\alpha(m_n)\right] + 16C \sum_{i=m_n+1}^{k_n m_n} \alpha_i.$$

By using the condition $\sum_{j=0}^{\infty} \alpha_{1,1}^{\frac{\delta}{2+\delta}}(j) < \infty$, the second term on the right hand side converges to zero. The first term on the right side is also converges to zero by Kronecker's Lemma (see [67], page 390).

Now we have $\lim_{n\to} \mathbb{E}\hat{\phi} = \phi$, which completes the proof. □

Theorem 4.16. *Let $X_{i\in\mathbb{Z}}$ be a strong mixing random process. We assume*

1) There exists a constant M and $\delta > 0$ such that

$$\left\| \frac{(S_{m_n}^{lm_n})^2}{m_n} \right\|_{2+\delta} \leq M;$$

2)

$$\sum_{j\geq 0} \alpha_{m,m}^{\frac{\delta}{2+\delta}}(j) = o(m) \qquad \text{as } m \to \infty;$$

3)

$$\liminf_{n\to\infty} \frac{Var(S_n)}{n} > 0;$$

4) $\mathbb{E}(S_{m_n}^{lm_n}) \equiv 0$ for all l and m_n.

Let

$$\hat{\sigma}_n^2 = \sum_{\substack{l=0,j=0 \\ |l-j|\leq 1}}^{k_n-1} S_{m_n}^{lm_n} S_{m_n}^{jm_n}.$$

Then we have

$$\frac{\hat{\sigma}_n^2}{Var(S_n)} \xrightarrow{L_2} 1 \qquad \text{as} \qquad n \to \infty,$$

where $S_{m_n}^{lm_n}$, m_n, k_n and n are defined as in Theorem 4.14, and

$$S_n = \sum_{l=0}^{k_n-1} S_{m_n}^{lm_n}.$$

Proof. We write

$$\mathbb{E}\left|\frac{\hat{\sigma}_n^2}{Var(S_n)}-1\right|^2=\frac{Var(\hat{\sigma}_n^2)}{[Var(S_n)]^2}+\frac{\left[\mathbb{E}(\hat{\sigma}_n^2)-Var(S_n)\right]^2}{[Var(S_n)]^2}.$$

Then we will show

(i)

$$\frac{Var(\hat{\sigma}_n^2)}{[Var(S_n)]^2}\to 0 \quad \text{as } n\to\infty,$$

and

(ii)

$$\frac{\mathbb{E}(\hat{\sigma}_n^2)}{Var(S_n)}\to 1 \quad \text{as } n\to\infty.$$

For (i), by assumption 3, for sufficiently large n, we have a constant C such that

$$[Var(S_n)]^2\geq n^2C.$$

Therefore it is enough to prove

$$\frac{Var(\hat{\sigma}_n^2)}{n^2}\to 0 \quad \text{as } n\to\infty,$$

or, equivalently, by $n=m_nk_n$,

$$\frac{Var\left[\sum_{l=0,j=0,|l-j|\leq 1}^{k_n-1}\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\right)\left(\frac{S_{m_n}^{jm_n}}{\sqrt{m_n}}\right)\right]}{k_n^2}\to 0 \quad \text{as } n\to\infty.$$

We note that

$$\begin{aligned}\frac{Var\left[\sum_{l=0}^{k_n-1}\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\right)^2\right]}{k_n^2} &\leq \frac{1}{k_n^2}\sum_{l=0}^{k_n-1}Var\left[\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\right)^2\right]\\ &\quad+\frac{2}{k_n^2}\sum_{l=0}^{k_n-2}\left|Cov(\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\right)^2,\left(\frac{S_{m_n}^{(l+1)m_n}}{\sqrt{m_n}}\right)^2)\right|\\ &\quad+\frac{1}{k_n^2}\sum_{|l-j|\geq 2}\left|Cov(\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\right)^2,\left(\frac{S_{m_n}^{jm_n}}{\sqrt{m_n}}\right)^2)\right|\\ &= \text{(I)+(II)+(III)}.\end{aligned}$$

Consider term (I). By assumption 1, we have

$$Var\left[\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\right)^2\right] \leq \mathbb{E}\left[\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\right)^4\right] \leq \left\|\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\right)^2\right\|_{2+\delta}^2 \leq M^2.$$

Therefore, the first term (I)$\leq \frac{1}{k_n}M^2 \to 0$ as $n \to \infty$.

Consider term (II).

$$\left|Cov\left(\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\right)^2, \left(\frac{S_{m_n}^{jm_n}}{\sqrt{m_n}}\right)^2\right)\right| \leq \sqrt{Var\left[\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\right)^2\right]}\sqrt{Var\left[\left(\frac{S_{m_n}^{jm_n}}{\sqrt{m_n}}\right)^2\right]} \leq M^2$$

implies (II) converges to zero.

Consider term (III). We note that

$$\left|Cov\left(\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\right)^2, \left(\frac{S_{m_n}^{jm_n}}{\sqrt{m_n}}\right)^2\right)\right| \leq 8\alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}((|l-j|-1)m_n)M^2.$$

Because $|l-j| \geq 2$, by using Lemma 4.12 and assumption 2, we have

$$\begin{aligned}
\text{(III)} &\leq 8M^2\frac{1}{k_n^2}\sum_{|l-j|\geq 2}\alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}((|l-j|-1)m_n) \\
&\leq 16M^2\frac{1}{k_n m_n}\sum_{j\geq 1}\alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}(j) \\
&\to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

Therefore

$$\frac{Var\left[\sum_{l=0}^{k_n-1}\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\right)^2\right]}{k_n^2} \to 0, \quad \text{as} \quad n \to \infty. \tag{4.4}$$

Similarly, we note, for any $|l-j| \leq 2$,

$$\left|Cov\left(\frac{S_{m_n}^{lm_n}S_{m_n}^{(l+1)m_n}}{m_n}, \frac{S_{m_n}^{jm_n}S_{m_n}^{(j+1)m_n}}{m_n}\right)\right| \leq M^2.$$

and for any $|l-j| \geq 3$,

$$\left|Cov\left(\frac{S_{m_n}^{lm_n}S_{m_n}^{(l+1)m_n}}{m_n}, \frac{S_{m_n}^{jm_n}S_{m_n}^{(j+1)m_n}}{m_n}\right)\right| \leq 8\alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}((|l-j|-2)m_n)M^2.$$

Then we have

$$\frac{Var\left[\sum_{l=0}^{k_n-2}\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\frac{S_{m_n}^{(l+1)m_n}}{\sqrt{m_n}}\right)\right]}{k_n^2} \to 0, \quad \text{as} \quad n\to\infty, \tag{4.5}$$

and

$$\frac{Cov\left[\sum_{l=0}^{k_n-1}\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\right)^2, \sum_{l=0}^{k_n-2}\left(\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\frac{S_{m_n}^{(l+1)m_n}}{\sqrt{m_n}}\right)\right]}{k_n^2} \to 0, \quad \text{as} \quad n\to\infty. \tag{4.6}$$

Now (4.4), (4.5) and (4.6) prove (i).

Lastly, we prove (ii). We write

$$\begin{aligned} Var(S_n) &= Var\left(\sum_{l=0}^{k_n-1} S_{m_n}^{lm_n}\right) \\ &= \sum_{\substack{l=0,j=0\\|l-j|\le 1}}^{k_n-1} Cov(S_{m_n}^{lm_n}, S_{m_n}^{jm_n}) + \sum_{\substack{l=0,j=0\\|l-j|\ge 2}}^{k_n-1} Cov(S_{m_n}^{lm_n}, S_{m_n}^{jm_n}) \\ &= \mathbb{E}(\hat{\sigma}_n^2) + (\mathrm{I}). \end{aligned}$$

Since $Var(S_{mn}^{lm_n}) \le m_n M$, for (I), by using Lemma 4.12, we have

$$\begin{aligned} (\mathrm{I}) &\le 2m_n \sum_{\substack{l=0,j=0\\|l-j|\ge 2}}^{k_n-1} 8\alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}((|l-j|-1)m_n)M^2 \\ &\le 16M^2 k_n \sum_{j\ge 1} \alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}(j). \end{aligned}$$

Assumption 3 implies that we have a constant C such that $Var(S_n) \ge nC$. Then by using assumption 2, we have

$$\frac{(\mathrm{I})}{Var(S_n)} \le \frac{(\mathrm{I})}{m_n k_n C} \le \frac{16M^2}{C}\frac{1}{m_n}\sum_{j\ge 0}\alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}(j) \to 0,$$

as $n\to\infty$. Now we have

$$\frac{\mathbb{E}(\hat{\sigma}_n^2)}{Var(S_n)} = 1 - \frac{(\mathrm{I})}{Var(S_n)} \to 1 \quad \text{as } n\to\infty.$$

Therefore (ii) is proved. This completes the poof of this theorem. □

Example: *(Example of $\|m^{-1}(S_m^l)^2\|_{2+\delta}$ being bounded)*

For any given l, let $\delta = 1$, X_i be a Gaussian process and $S_m^1 = X_1 + X_2 + \cdots + X_m$. We note that $\|m^{-1}(S_m^l)^2\|_3 < \infty$ is equivalent to $\mathbb{E}[m^{-1}(S_m^l)^2]^3 < \infty$. By using the Isserlis Theorem,

Theorem A6 in the Appendix, we have

$$
\begin{aligned}
\mathbb{E}(S_m^l)^6 &= \sum_{i_1=1}^{m}\sum_{i_2=1}^{m}\sum_{i_3=1}^{m}\sum_{i_4=1}^{m}\sum_{i_5=1}^{m}\sum_{i_6=1}^{m}\mathbb{E}(X_{i_1}X_{i_2}X_{i_3}X_{i_4}X_{i_5}X_{i_6}) \\
&= \sum_{i_1=1}^{m}\sum_{i_2=1}^{m}\sum_{i_3=1}^{m}\sum_{i_4=1}^{m}\sum_{i_5=1}^{m}\sum_{i_6=1}^{m}\left(\sum_{(i,j)\in\{i_1i_2i_3i_4i_5i_6\}}\prod\mathbb{E}(X_iX_j)\right) \\
&= \sum_{i_1=1}^{m}\sum_{i_2=1}^{m}\sum_{i_3=1}^{m}\sum_{i_4=1}^{m}\sum_{i_5=1}^{m}\sum_{i_6=1}^{m} \\
&\qquad \Big(\mathbb{E}(X_{i_1}X_{i_2})\mathbb{E}(X_{i_3}X_{i_4})\mathbb{E}(X_{i_5}X_{i_6})+\mathbb{E}(X_{i_1}X_{i_2})\mathbb{E}(X_{i_3}X_{i_5})\mathbb{E}(X_{i_4}X_{i_6})+\cdots \\
&\qquad +\mathbb{E}(X_{i_1}X_{i_3})\mathbb{E}(X_{i_2}X_{i_4})\mathbb{E}(X_{i_5}X_{i_6})+\mathbb{E}(X_{i_1}X_{i_3})\mathbb{E}(X_{i_2}X_{i_5})\mathbb{E}(X_{i_4}X_{i_6})+\cdots \\
&\qquad +\mathbb{E}(X_{i_1}X_{i_4})\mathbb{E}(X_{i_2}X_{i_3})\mathbb{E}(X_{i_5}X_{i_6})+\mathbb{E}(X_{i_1}X_{i_4})\mathbb{E}(X_{i_2}X_{i_5})\mathbb{E}(X_{i_3}X_{i_6})+\cdots \\
&\qquad +\mathbb{E}(X_{i_1}X_{i_5})\mathbb{E}(X_{i_2}X_{i_3})\mathbb{E}(X_{i_4}X_{i_6})+\mathbb{E}(X_{i_1}X_{i_5})\mathbb{E}(X_{i_2}X_{i_4})\mathbb{E}(X_{i_3}X_{i_6})+\cdots \\
&\qquad +\mathbb{E}(X_{i_1}X_{i_6})\mathbb{E}(X_{i_2}X_{i_3})\mathbb{E}(X_{i_4}X_{i_5})+\mathbb{E}(X_{i_1}X_{i_6})\mathbb{E}(X_{i_2}X_{i_4})\mathbb{E}(X_{i_3}X_{i_5})+\cdots\Big) \\
&= 15\left(\sum_{i=1}^{m}\sum_{j=1}^{m}\mathbb{E}(X_iX_j)\right)^3.
\end{aligned}
$$

By using the fact of $\mathbb{E}(X_i,X_j)=Cov(X_i,X_j)$ and Theorem 2.2, we have

$$
\begin{aligned}
\mathbb{E}(S_m^1)^6 &\le 15\left(\sum_{i=1}^{m}\sum_{j=1}^{m}Cov(X_i,X_j)\right)^3 \\
&\le 7680\left(\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_{1,1}^{\frac{\delta}{2+\delta}}(|i-j|)\|X\|_{2+\delta}^2\right)^3 \\
&\le 7680\left(\sum_{i=1}^{m}\sum_{r\ge1}2\alpha_{1,1}^{\frac{\delta}{2+\delta}}(r)\|X\|_{2+\delta}^2\right)^3 \\
&= 61440m^3\left(\sum_{r\ge1}\alpha_{1,1}^{\frac{\delta}{2+\delta}}(r)\|X\|_{2+\delta}^2\right)^3.
\end{aligned}
$$

Therefore, if we have the assumption such as (H3) for $d=1$, i.e. $\|X\|_{2+\delta}$ and $\sum_{r\ge1}\alpha_{1,1}^{\frac{\delta}{2+\delta}}(r)$ are bounded, then we have $\mathbb{E}[m^{-1}(S_m^1)^2]^3<\infty$.

In the following theorem, Theorem 4.17, we relax the assumption on $\mathbb{E}S_{m_n}^{lm_n}\equiv 0$ in Theorem 4.16. We also introduce a different estimator to approach the variance in L_2.

Theorem 4.17. *Let $X_{i\in\mathbb{Z}}$ be a strong mixing random process. We assume*

1) There exists a constant M and $\delta > 0$ such that

$$\left\| \frac{(S_{m_n}^{lm_n})^2}{m_n} \right\|_{2+\delta} \leq M;$$

2)

$$\sum_{j \geq 0} \alpha_{\infty,\infty}^{\frac{\delta}{2+\delta}}(j) < \infty;$$

3)

$$\liminf_{n \to \infty} \frac{Var(S_n)}{n} > 0;$$

4)

$$\frac{1}{m_n k_n} \sum_{i=0}^{k_n - 1} \left(\mathbb{E} S_{m_n}^{(i)} - \mathbb{E} \bar{S}_{m_n} \right)^2 \to 0.$$

Let

$$\hat{\sigma}_n^2 = \sum_{\substack{l=0, j=0 \\ |l-j| \leq 1}}^{k_n - 1} \left(S_{m_n}^{lm_n} - \bar{S}_{m_n} \right) \left(S_{m_n}^{jm_n} - \bar{S}_{m_n} \right).$$

Then we have

$$\frac{\hat{\sigma}_n^2}{Var(S_n)} \xrightarrow{L_2} 1 \quad \text{as} \quad n \to \infty,$$

where $S_{m_n}^{lm_n}$, m_n, k_n and n are defined as in Theorem 4.14, and

$$S_n = \sum_{l=0}^{k_n - 1} S_{m_n}^{lm_n}, \quad \bar{S}_{m_n} = \frac{1}{k_n} S_n.$$

Proof. We use the same idea which is used in Theorem 4.16. With the similar argument of (i) in the proof of Theorem 4.16, we have

$$\frac{Var(\hat{\sigma}_n^2)}{[Var(S_n)]^2} \to 0 \quad \text{as } n \to \infty.$$

Therefore, the rest of this proof is to show

$$\frac{\mathbb{E}(\hat{\sigma}_n^2)}{Var(S_n)} \to 1 \quad \text{as } n \to \infty.$$

Let $E_i = \mathbb{E}S_{m_n}^{im_n}$, $\bar{E} = \mathbb{E}\bar{S}$ and

$$t_i = \frac{1}{\sqrt{m_n}}(S_{m_n}^{im_n} - E_i), \quad i = 0, 1, 2, \cdots, k_n.$$

Then we have

$$\begin{aligned}
\hat{\sigma}_n^2 &= \sum_{|l-j|\le 1} m_n t_l t_j - \sum_{|l-j|\le 1} m_n(t_l\bar{t} + t_j\bar{t} + \bar{t}^2) \\
&\quad + \sum_{|l-j|\le 1} \sqrt{m_n}(t_l - \bar{t})(E_i - \bar{E}) + \sum_{|l-j|\le 1} \sqrt{m_n}(t_j - \bar{t})(E_i - \bar{E}) \\
&\quad + \sum_{|l-j|\le 1} (E_l - \bar{E})(E_j - \bar{E}) \\
&= \sum_{|l-j|\le 1} m_n t_l t_j - (\mathrm{I}) + (\mathrm{II}) + (\mathrm{III}) + (\mathrm{IV}).
\end{aligned}$$

By using the fact that $\mathbb{E}t_i = 0$ for all $i = 0, 1, \cdots, k_n$, we have $\mathbb{E}(\mathrm{II}) = 0$ and $\mathbb{E}(\mathrm{III}) = 0$. By using assumption 1 and 2, we have

$$\frac{\mathbb{E}(\mathrm{I})}{n} \to 0.$$

Assumption 4 implies

$$\frac{\mathbb{E}(\mathrm{IV})}{n} \to 0.$$

Therefore we have

$$\frac{\mathbb{E}\hat{\sigma}_n^2}{n} = A_n + o(1),$$

where

$$A_n = \frac{1}{n} \sum_{|l-j|\le 1} \mathbb{E}m_n t_l t_j.$$

We write

$$\begin{aligned}
Var(S_n) &= \sum_{|l-j|\le 1} m_n \mathbb{E}(t_l t_j) + m_n \sum_{|l-j|\ge 2} Cov(t_l, t_j) \\
&= \sum_{|l-j|\le 1} m_n \mathbb{E}(t_l t_j) + (\mathrm{V}).
\end{aligned}$$

Since assumption 2 implies

$$\frac{\mathbb{E}(\mathrm{V})}{n} \to 0,$$

we have

$$\frac{Var(S_n)}{n} = A_n + o(1).$$

Therefore, assumption 3 implies

$$\left|\frac{\mathbb{E}\hat{\sigma}_n^2}{Var(S_n)} - 1\right| = \left|\frac{\mathbb{E}\hat{\sigma}_n^2/n - Var(S_n)/n}{Var(S_n)/n}\right| = \left|\frac{o(1)}{Var(S_n)/n}\right| \to 0.$$

This completes the proof. □

4.2.2 Estimators with smoothness conditions

In this subsection and the subsection 4.3.2, we assume the estimators have bounded partial derivatives, which we call the *smoothness condition*, i.e. the second assumption in Theorem 4.18, Theorem 4.19, Theorem 4.29, Theorem 4.30, Theorem 4.32 and Theorem 4.38. We also assume that the population is divided into blocks. For example, when the population size is $n = 48$, we may divide it into $k_{48} = 16$ blocks, and put $m_{48} = 3$ points in each block, see Figure 4.1. Therefore, if the population has asymptotically independent properties, the distance between any

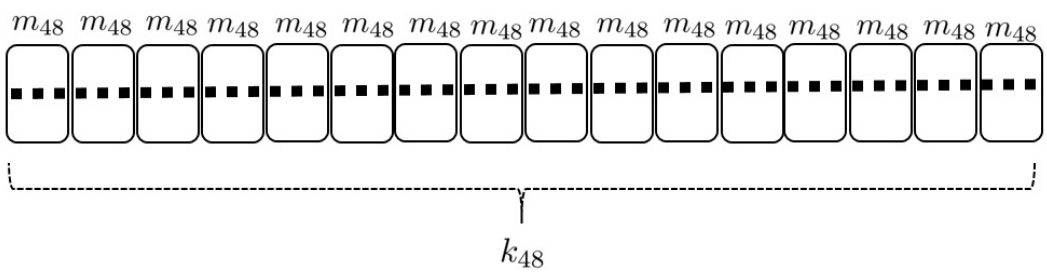


Figure 4.1: The divided population with 16 blocks

one pair of blocks will contribute to the estimation of the variance between them. Furthermore, the assumed relationship between k_n and m_n will also contribute to the estimation of the variance of the estimators.

Theorem 4.18. *Let $X_{i\in\mathbb{Z}}$ be a strong mixing random process, $f_m(\cdot)$ be a function. We define*

$$f_m^l = f_m(X_l, X_{l+1}, \cdots, X_{l+m+1}),$$

for all $m > 0$, $m, l, \in \mathbb{Z}$. Let n be the sample size, $n \in \mathbb{Z}^+$, $\{m_n\}_{n\in\mathbb{Z}^+}$ and $k_n = [\frac{n}{m_n}]$ be such that $m_n \to \infty$ and $k_n \to \infty$ as $n \to \infty$. Let

$$\bar{f}_n = \frac{1}{k_n}\sum_{l=0}^{k_n-1} f_{m_n}^{lm_n}.$$

We assume that

1) $m_n^2 = o(k_n)$ as $n \to \infty$;

2) $\left\|\frac{\partial f}{\partial X}\right\|_\infty = \sup_{m,i}\left|\frac{\partial f_m}{\partial X_i}\right| < \infty$, and $\sup_m |f_m(0, 0, \cdots, 0)| < \infty$;

3) There exists $\delta > 0$ s.t. $\|X\|_{2+\delta} < \infty$;

4)

$$\sum_{j=0}^{\infty} \alpha_{m,m}^{\frac{\delta}{2+\delta}}(j) = O(m) \quad as \quad m \to \infty.$$

If $\lim_{n\to\infty} \mathbb{E}(\bar{f}_n) = \varphi$ exists, then we have $\bar{f}_n \xrightarrow{L_2} \varphi$.

Proof. We only need to show $Var(\bar{f}_n) \to 0$ as $n \to \infty$.

By using the Mean Value Theorem for $f_{m_n}^{lm_n}$, there exists $(X_l^{(\xi)}, X_{l+1}^{(\xi)}, \cdots, X_{l+m_n-1}^{(\xi)})$ which is in the segment joining $(0, 0, \cdots, 0)$ and $(X_l, X_{l+1}, \cdots, X_{l+m_n-1})$ such that

$$f_{m_n}^{lm_n} = f_{m_n}(0, 0, \cdots, 0) + \sum_{i=l}^{l+m_n-1} \frac{\partial f_{m_n}(X_l^{(\xi)}, X_{l+1}^{(\xi)}, \cdots, X_{l+m_n-1}^{(\xi)})}{\partial X_i} X_i.$$

Then Minkovski inequality implies

$$\|f_{m_n}^{lm_n}\|_{2+\delta} \leq \|f_{m_n}(0, 0, \cdots, 0)\|_{2+\delta} + \sum_{i=l}^{l+m_n-1} \left\| \frac{\partial f}{\partial X} \right\|_{\infty} \|X\|_{2+\delta}.$$

Assumptions 2 and 3 imply $\|f_{m_n}^{lm_n}\|_{2+\delta} \leq (m_n + 1)C$, where C is a constant.

We note that

$$Var(\bar{f}_n) \leq \frac{1}{k_n^2} \left[\sum_{\substack{l=0,j=0 \\ |l-j|\leq 1}}^{k_n-1} |Cov(f_{m_n}^{lm_n}, f_{m_n}^{jm_n})| + \sum_{\substack{l=0,j=0 \\ |l-j|\geq 2}}^{k_n-1} |Cov(f_{m_n}^{lm_n}, f_{m_n}^{jm_n})| \right],$$

and

$$\sum_{\substack{l=0,j=0 \\ |l-j|\leq 1}}^{k_n-1} |Cov(f_{m_n}^{lm_n}, f_{m_n}^{jm_n})| \leq 3k_n \|f_{m_n}^{lm_n}\|_{2+\delta}^2 \leq 3k_n(m_n+1)^2 C^2.$$

By using Lemma 4.12, we have

$$\begin{aligned} \sum_{\substack{l=0,j=0 \\ |l-j|\geq 2}}^{k_n-1} |Cov(f_{m_n}^{lm_n}, f_{m_n}^{jm_n})| &\leq 8 \sum_{\substack{l=0,j=0 \\ |l-j|\geq 2}}^{k_n-1} \alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}(|j-l|m_n - m_n) \|f_{m_n}^{lm_n}\|_{2+\delta} \|f_{m_n}^{jm_n}\|_{2+\delta} \\ &\leq 16 \frac{(m_n+1)^2 C^2 k_n}{m_n} \sum_{j=1}^{\infty} \alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}(j). \end{aligned}$$

Therefore, by using assumption 1 and 4, we have

$$\begin{aligned} Var(\bar{f}_n) &\leq 3\frac{(m_n+1)^2C^2}{k_n} + 16\frac{(m_n+1)^2C^2}{k_n m_n}\sum_{j=0}^{\infty}\alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}(j) \\ &\to 0 \quad \text{as } n\to\infty. \end{aligned}$$

This completes the proof. □

Remark: In the above theorem, for an example of assumption 1, take $m_n = [n^{(1-r)}]$ and $k_n = [n^r]$, where $\frac{2}{3} < r < 1$. The assumption 4 implies, comparing with the assumption (H1) and the assumption (H3), a stronger dependence is taken into account in this theorem, i.e. $\sum_{j=0}^{\infty}\alpha_{m,m}^{\frac{\delta}{2+\delta}}(j)$ could be divergent, rather than a constant, as $m\to\infty$.

Theorem 4.19. *Let $X_{i\in\mathbb{Z}}$ be a strong mixing process, $S_m(\cdot)$ be a function. We define*

$$S_m^l = S_m^l(X_l, X_{l+1}, \cdots, X_{l+m-1}),$$

$$t_m^l = \sqrt{m}(S_m^l - \mathbb{E}(S_m^l)),$$

for all $m>0$, $m,l\in\mathbb{Z}$. Let m_n, k_n and n be defined as in Theorem 4.13 and set

$$\bar{t}_n = \frac{1}{k_n}\sum_{l=0}^{k_n-1} t_{m_n}^{lm_n},$$

$$\Sigma_n = \frac{1}{k_n}\sum_{l=0}^{k_n-1} (t_{m_n}^{lm_n})^2,$$

$$\bar{S}_n = \frac{1}{k_n}\sum_{l=0}^{k_n-1} S_{m_n}^{lm_n},$$

$$\hat{\sigma}_n^2 = \frac{m_n}{k_n}\sum_{l=0}^{k_n-1} (S_{m_n}^{lm_n} - \bar{S}_n)^2.$$

We assume

1) $m_n^6 = o(k_n)$ as $n\to\infty$;

2) $\left\|\frac{\partial S}{\partial X}\right\|_\infty = \sup_{m,i}\left|\frac{\partial S_m}{\partial X_i}\right| = C_1 < \infty$, and $\sup_m |S_m(0,0,\cdots,0)| = C_2 < \infty$, C_1 and C_2 are constants;

3) There exists $\delta>0$ s.t. $\|X\|_{6+3\delta} = C_3 < \infty$, C_3 is a constant;

4)

$$\sum_{j=0}^{\infty}\alpha_{m,m}^{\frac{\delta}{2+\delta}}(j) = O(m) \quad \text{as} \quad m\to\infty;$$

5)

$$E_n \equiv \frac{m_n}{k_n} \sum_{l=0}^{k_n-1} \left(\mathbb{E} S_{m_n}^{lm_n} - \mathbb{E}\bar{S}_n \right)^2 \to 0 \quad as \quad n \to \infty;$$

6) $\mathbb{E}S_m^l \leq O(m)$, *for all* $l = 0, 1, \cdots$.

If $\lim_{n\to\infty} \mathbb{E}(\Sigma_n) = \sigma^2$ *exists, then we have* $\hat{\sigma}_n^2 \xrightarrow{L_2} \sigma^2$ *as* $n \to \infty$.

Proof. Following the idea of the proof of Theorem 4.14, since

$$\hat{\sigma}_n^2 = \Sigma_n + 2A_n + B_n, \quad |A_n| \leq |\Sigma_n B_n|^{\frac{1}{2}}, \quad B_n \leq 2E_n + 2(\bar{t}_n)^2,$$

where A_n and B_n are introduced in Theorem 4.14, we only need to show

(i) $\Sigma_n \xrightarrow{L_2} \sigma^2$, i.e. $Var(\Sigma_n) \to 0$ as $n \to \infty$

and

(ii) $\bar{t}_n \xrightarrow{L_2} 0$, as $n \to \infty$.

With the same argument as in Theorem 4.18, there exists $(X_l^{(\xi)}, X_{l+1}^{(\xi)}, \cdots, X_{l+m_n-1}^{(\xi)})$ such that

$$S_m^l = S_m^l(0, 0, \cdots, 0) + \sum_{i=l}^{l+m-1} \frac{\partial S_m^l(X_l^{(\xi)}, X_{l+1}^{(\xi)}, \cdots, X_{l+m-1}^{(\xi)})}{\partial X_i} X_i. \tag{4.7}$$

Then, by using Minkovski inequality and Hölder inequality, assumptions 2 and 3 imply $\|(S_m^l)^2\|_{2+\delta} \leq (m+1)^2 C$, C is a constant. Therefore, assumption 6 implies $\|(t_m^{lm})^2\|_{2+\delta} \leq (m+1)^2 m C_4$, where C_4 is a constant.

For (i), we note that, by using Theorem 2.2 and Lemma 4.12,

$$\begin{aligned} Var(\Sigma_n) &\leq \frac{1}{k_n^2} \left[\sum_{\substack{l=0,j=0 \\ |l-j|\leq 1}}^{k_n-1} |Cov\left((t_{m_n}^{lm_n})^2, (t_{m_n}^{jm_n})^2 \right)| + \sum_{\substack{l=0,j=0 \\ |l-j|\geq 2}}^{k_n-1} |Cov\left((t_{m_n}^{lm_n})^2, (t_{m_n}^{jm_n})^2 \right)| \right] \\ &\leq \frac{1}{k_n^2} \left[3k_n(m_n+1)^4 m_n^2 C_3 + 16 C_4 (m_n+1)^4 m_n^2 k_n \frac{1}{m_n} \sum_{j\geq 1} \alpha_{m_n.m_n}^{\frac{\delta}{2+\delta}}(j) \right]. \end{aligned}$$

Then assumptions 1 and 4 imply $Var(\Sigma_n) \to 0$ as $n \to \infty$.

For (ii), by using Theorem 4.3, we are going to show $\mathbb{E}|\bar{t}_n|^4 \to 0$. We write

$$\mathbb{E}(\bar{t}_n)^4 = \frac{1}{k_n^4}\left[\sum_{l_1=0}^{k_n-1}\sum_{l_2=0}^{k_n-1}\sum_{l_3=0}^{k_n-1}\sum_{l_4=0}^{k_n-1} Cov\left(t_{m_n}^{l_1 m_n}, t_{m_n}^{l_2 m_n} t_{m_n}^{l_3 m_n} t_{m_n}^{l_4 m_n}\right)\right].$$

Theorem 2.2 implies

$$\begin{aligned}
\mathbb{E}(\bar{t}_n)^4 &\leq \frac{1}{k_n^4}\left[\sum_{l_1=0}^{k_n-1}\sum_{l_2=0}^{k_n-1}\sum_{l_3=0}^{k_n-1}\sum_{l_4=0}^{k_n-1}\left|Cov\left(t_{m_n}^{l_1 m_n}, t_{m_n}^{l_2 m_n} t_{m_n}^{l_3 m_n} t_{m_n}^{l_4 m_n}\right)\right|\right] \\
&\leq \frac{1}{k_n^4}\left[\sum_{l_1=0}^{k_n-1}\sum_{l_2=0}^{k_n-1}\sum_{l_3=0}^{k_n-1}\sum_{l_4=0}^{k_n-1}\right. \\
&\qquad \left. 8\alpha_{m_n,3m_n}^{\frac{\delta}{2+\delta}}\Big(dist(\Lambda_{l_1}, \Lambda_{l_2 l_3 l_4})\Big)\|t_{m_n}^{l_1 m_n}\|_{2+\delta}\|t_{m_n}^{l_2 m_n} t_{m_n}^{l_3 m_n} t_{m_n}^{l_4 m_n}\|_{2+\delta}\right],
\end{aligned}$$

where, we use d to denote the distance between $t_{m_n}^{l_1 m_n}$ and $t_{m_n}^{l_2 m_n} t_{m_n}^{l_3 m_n} t_{m_n}^{l_4 m_n}$, i.e.

$$d := dist(\Lambda_{l_1}, \Lambda_{l_2 l_3 l_4}),$$

Λ_{l_1} and $\Lambda_{l_2 l_3 l_4}$ are the index sets of X_i's which are involved in $t_{m_n}^{l_1 m_n}$ and $t_{m_n}^{l_2 m_n} t_{m_n}^{l_3 m_n} t_{m_n}^{l_4 m_n}$ respectively. It means d may take values:

$$d = 0, \quad d = m_n, \quad d = 2m_n, \quad d = 3m_n, \quad \cdots,$$

i.e. $d = \lambda m_n$, where $\lambda = 0, 1, 2, 3, \cdots$. By using the symmetry of l_2, l_3 and l_4, we use $l_1 - l_2$ to describe λ, we have the following estimation

$$\mathbb{E}(\bar{t}_n)^4 \leq \frac{1}{k_n^4}\left[k_n^2 \sum_{l_1=0}^{k_n-1}\sum_{l_2=0}^{k_n-1} 8\alpha_{m_n,3m_n}^{\frac{\delta}{2+\delta}}\Big(|l_1 - l_2|)m_n\Big)\|t_{m_n}^{l_1 m_n}\|_{2+\delta}\|t_{m_n}^{l_2 m_n} t_{m_n}^{l_3 m_n} t_{m_n}^{l_4 m_n}\|_{2+\delta}\right],$$

We note that

$$\|t_{m_n}^{l_1 m_n}\|_{2+\delta} \leq \sqrt{m_n}[(m_n+1)\|X\|_{2+\delta} + O(m_n)], \tag{4.8}$$

and

$$\|t_{m_n}^{l_1 m_n} t_{m_n}^{l_2 m_n} t_{m_n}^{l_3 m_n}\|_{2+\delta} \leq \Big(\sqrt{m_n}[(m_n+1)\|X\|_{2+\delta} + O(m_n)]\Big)^3. \tag{4.9}$$

Therefore,

$$\mathbb{E}(\bar{t}_n)^4 \leq \frac{1}{k_n^4} k_n^2 m_n^2 (m_n+1)^4 C\left[\sum_{l_1=0}^{k_n-1}\sum_{l_2=0}^{k_n-1} \alpha_{m_n,3m_n}^{\frac{\delta}{2+\delta}}\Big(|l_1 - l_2| m_n\Big)\right]$$

$$\leq \quad \frac{1}{k_n^2} m_n^2 (m_n+1)^4 C \left[k_n + \frac{k_n - 1}{m_n} \sum_{\lambda=1}^{(k_n-1)m_n} \alpha_{m_n,3m_n}^{\frac{\delta}{2+\delta}} \Big(\lambda\Big) \right].$$

Then assumption 1 and 4 imply $\mathbb{E}(\bar{t}_n)^4 \to 0$. This complete the proof. □

Theorem 4.20. *Let $X_{i\in\mathbb{Z}}$ be a zero-mean strong mixing random process. We set*

$$S_{m_n}^l = \sum_{i=0}^{m_n-1} X_{l+i}, \quad S_n = \sum_{l=0}^{k_n-1} S_{m_n}^{lm_n},$$

$$\nu_{m_n}^l = \frac{S_{m_n}^l}{\sqrt{m_n}}, \quad \bar{\nu}_n = \frac{1}{k_n} \sum_{l=0}^{k_n-1} \nu_{m_n}^{lm_n},$$

$$\hat{\sigma}_n^2 = \frac{1}{k_n} \sum_{l=0}^{k_n-1} (\nu_{m_n}^{lm_n} - \bar{\nu}_n)^2, \quad \Sigma_n = \frac{1}{k_n} \sum_{l=0}^{k_n-1} (\nu_{m_n}^{lm_n})^2,$$

where k_n, m_n and n are defined as in Theorem 4.13. We assume

1) $m_n^6 = o(k_n)$ as $n \to \infty$;

2) $\exists \delta > 0$ s.t. $\|X\|_{4+2\delta} < \infty$;

3)

$$\sum_{j=0}^{\infty} \alpha_{m,m}^{\frac{\delta}{2+\delta}}(j) = O(m) \quad \text{as} \quad m \to \infty.$$

4)

$$\sum_{j=0}^{\infty} \alpha_{1,1}^{\frac{\delta}{2+\delta}}(j) \leq \infty;$$

5) $\lim_{n\to\infty} \frac{Var(S_n)}{n} = \sigma^2$.

Then we have $\hat{\sigma}_n^2 \xrightarrow{L_2} \sigma^2$ as $n \to \infty$.

Proof. Theorem 4.19 implies that if $\lim_{n\to\infty} \mathbb{E}(\Sigma_n) = \sigma^2$ exists, then $\hat{\sigma}_n^2 \xrightarrow{L_2} \sigma^2$.

We note that, by using the property of zero-mean,

$$\begin{aligned} \frac{Var(S_n)}{n} &= \frac{1}{k_n m_n} Var\left(\sum_{l=0}^{k_n-1} S_{m_n}^{lm_n} \right) \\ &= \frac{1}{k_n m_n} \mathbb{E}\left(\sum_{l=0}^{k_n-1} S_{m_n}^{lm_n} \right)^2 \end{aligned}$$

$$= \frac{1}{k_n}\mathbb{E}\left(\sum_{l=0}^{k_n-1}\frac{S_{m_n}^{lm_n}}{\sqrt{m_n}}\right)^2 = \mathbb{E}\Sigma_n.$$

This completes this proof. □

In Theorem 4.18, Theorem 4.19 and Theorem 4.20, if we introduce a stronger assumption on strong mixing coefficients, then the "little-o" assumption may become weaker. For example, Theorem 4.20 still holds if we replace the assumptions 1, 3 and 4 with the following two assumptions:

$$m_n^5 = o(k_n), \quad \text{as} \quad n \to \infty$$

and

$$\sum_{j=0}^{\infty} \alpha_{\infty,\infty}^{\frac{\delta}{2+\delta}}(j) < \infty.$$

Since m_n and k_n can be practically controlled by the sampling design, we adopt the weaker condition on strong mixing coefficients and the stronger condition on the "little-o" assumptions.

Proposition 4.21. *Let $X_{i\in\mathbb{Z}}$ be a random process, $S_m(\cdot)$ be a measurable function, $S_m^l = S_m(X_l, X_{l+1}, \cdots, X_{l+m-1})$. We assume*

1) There exists $\delta > 0$ s.t. $\|X\|_{1+\delta} < \infty$;

2) $\left\|\frac{\partial S}{\partial X}\right\|_\infty = \sup_{m,i}\left|\frac{\partial S_m}{\partial X_i}\right| < \infty$, and $\sup_m S_m(0, 0, \cdots, 0) < \infty$.

Then $m^{-1}S_m^l$ are u.i., for the indices l and m.

Proof. By using the Mean Value Theorem and the Minkovski's inequality, assumption 1 and 2 imply $\|S_m^l\|_{1+\delta} \leq (m+1)C$, C is a constant. Therefore $\|m^{-1}S_m^l\|_{1+\delta}$ is bounded. Lemma 4.11 completes this proof. □

Remark: If we define $S_m^l = S_m(X_{i\in D_m})$, where $D_m \subset \mathbb{Z}^d$ and $|D_m| = m$, then this proposition still holds for a random field $X_{\in\mathbb{Z}^d}$.

4.3 Estimators on random fields

In this section, we are going to generalize the results of Section 4.2 to random fields. Let D_n stand for the sample of the random field $X_{i\in\mathbb{Z}^d}$, which is divided into blocks. We assume $D_{m_n}^{(i)}$

are blocks of D_n, where $i \in K_n \subseteq \mathbb{Z}^d$ is the index of $D_{m_n}^{(i)}$, and

$$D_n = \bigcup_{i \in K_n} D_{m_n}^{(i)}.$$

For all $i \in K_n$, we set

$$D_{m_n}^{(i)} \cap D_{m_n}^{(j)} = \emptyset \quad \text{if} \quad i \neq j,$$

$$|D_{m_n}^{(i)}| = m_n, \quad |K_n| = k_n, \quad m_n \to \infty, \quad k_n \to \infty, \quad \text{as} \quad n \to \infty,$$

and therefore $|D_n| = m_n k_n$. We introduce J_i to stand for the index set of the neighbours of the block $D_{m_n}^{(i)}$, i.e.

$$J_i = \{j : dist(D_{m_n}^{(i)}, D_{m_n}^{(j)}) < d_{m_n}\},$$

where d_{m_n} is a real number with respect to m_n. We use $|J_i|$ to stand for the number of elements in J_i. We assume we may control $D_{m_n}^{(i)}$ through sampling strategies. Therefore we introduce the following definition:

Definition 4.2. We say D_n is *well-divided* by $\{D_{m_n}^{(i)}\}_{i \in K_n}$ in the random field $X_{i \in \mathbb{Z}^d}$ if and only if there exist positive constants b_1, b_2 and b such that, for all m_n, we have: (1) $d_{m_n} = b_1$ and $|J_i| \leq b$ for all $i \in K_n$. (2) $dist(D_{m_n}^{(i)}, D_{m_n}^{(j)}) \geq m_n^{b_2}$ when $j \notin J_i$.

This definition means by using a suitable well-divided sampling strategy, say the blocks are regularly growing with $|D_n|$, the number $|J_i|$ is not related to the size of the blocks. For example,

(a) Well-divided by squares (b) Well-divided by different shapes

Figure 4.2: Well-divided sample in 2-dimension random fields

in Figure 4.2, all sides of these blocks are growing with the same ratio which is with respect to the sample size, i.e. this sampling strategy ensures these growing blocks keep similar shapes. Therefore, for Figure 4.2(a), say the centred block is $D_{m_n}^{(i)}$, for any real number $d_{m_n} \in [2, m_n^{\frac{1}{2}}]$, we have $|J_i| \leq 8$. Similarly, for Figure 4.2(b), if we translate this sampling strategy into well-divided blocks as in (a), we conclude that for any $|J_i|$ of (b) will not greater than that of (a), which means $|J_i| < 8l$.

On the other hand, we may have some sampling strategies which are not well-divided. For example in Figure 4.3, say the block at the bottom is $D_{m_n}^{(i)}$, if one side of these rectangles is 1,

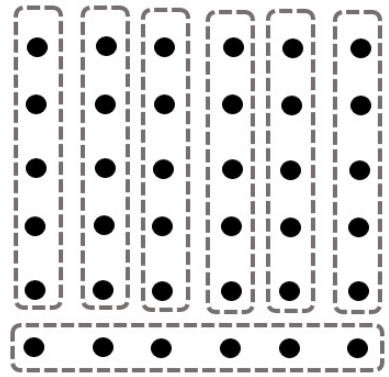


Figure 4.3: Not well-divided sample

and the other side is m_n, then the sampling strategy in Figure 4.3 implies

$$|J_i| = \#\{j : dist(D_{m_n}^{(i)}, D_{m_n}^{(j)}) \leq 1\} = m_n,$$

which means $|J_i| \to \infty$ as $n \to \infty$. This infinite number of pairs within the finite distance brings difficulties on estimating the moments of some estimators, therefore most of the results within this subsection are based on a well-divided sampling strategy.

The following two theorems are applications of a well-divided sampling strategy. This sampling strategy contributes to the estimation of fourth moments and variances of some estimators.

Theorem 4.22. *Let $X_{i\in\mathbb{Z}^d}$ be a random field, $D_n \subseteq \mathbb{Z}^d$, $K_n \subseteq \mathbb{Z}^d$,*

$$\bigcup_{i\in K_n} D_{m_n}^{(i)} = D_n, \quad D_{m_n}^{(i)} \cap D_{m_n}^{(j)} = \emptyset \quad \text{if} \quad i \neq j,$$

$$|D_{m_n}^{(i)}| = m_n, \quad |K_n| = k_n, \quad |D_n| = m_n k_n,$$

$$k_n \to \infty, \quad m_n \to \infty \quad \text{as} \quad n \to \infty.$$

Let D_n be well-divided by $\{D_{m_n}^{(i)}\}_{i\in K_n}$, i.e. there exist constants $b_1, b_2, b > 0$, such that, for all $i \in K_n$, $|J_i| \leq b$, where $J_i = \{j : dist(D_{m_n}^{(i)}, D_{m_n}^{(j)}) \leq b_1\}$, and $dist(D_{m_n}^{(i)}, D_{m_n}^{(j)}) \geq m_n^{b_2}$ when $j \notin J_i$. Let $f_{m_n}^{(i)}$ be a measurable function,

$$f_{m_n}^{(i)} = f_{m_n}(X_{j\in D_{m_n}^{(i)}}),$$

$$\bar{f}_{D_n} = \frac{1}{k_n} \sum_{i\in K_n} f_{m_n}^{(i)}.$$

We assume

1) $(f_{m_n}^{(i)})^{2(k-1)}$ are u.i., where $k \geq 2$, $k \in \mathbb{Z}^+$ is a constant;

2) $\mathbb{E}f_{m_n}^{(i)} \equiv 0$*;*

3) $\alpha_{\infty,\infty}(m^{b_2}) \to 0$ *as* $m \to \infty$*.*

Then we have $\mathbb{E}|\bar{f}_{D_n}|^k \to 0$ *as* $n \to \infty$*.*

Proof. We only prove this theorem for $k = 4$, similar arguments apply in other cases. We write

$$\mathbb{E}(\bar{f}_{D_n})^4 = \frac{1}{k_n^4} \sum_{i\in K_n} \sum_{j\in K_n} \sum_{q\in K_n} \sum_{l\in K_n} Cov(f_{m_n}^{(i)}, f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)}).$$

Since assumption 1 implies $Var(f_{m_n}^{(i)})$ and $Var(f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)})$ are bounded, then we have a constant C such that

$$\begin{aligned}
\text{(I)} &= \frac{1}{k_n^4} \sum_{i\in K_n} \sum_{j\in J_i} \sum_{q\in K_n} \sum_{l\in K_n} Cov(f_{m_n}^{(i)}, f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)}) \\
&\leq \frac{1}{k_n^4} \sum_{i\in K_n} \sum_{j\in J_i} \sum_{q\in K_n} \sum_{l\in K_n} \left| Cov(f_{m_n}^{(i)}, f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)}) \right| \\
&\leq \frac{1}{k_n^4} \sum_{i\in K_n} \sum_{j\in J_i} \sum_{q\in K_n} \sum_{l\in K_n} C \\
&= \frac{1}{k_n^4} \sum_{i\in K_n} |J_i| \sum_{q\in K_n} \sum_{l\in K_n} C \\
&\leq \frac{1}{k_n^4} k_n b k_n^2 C \to 0.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\text{(II)} &= \frac{1}{k_n^4} \sum_{i\in K_n} \sum_{j\in J_i} \sum_{q\in J_i} \sum_{l\in K_n} Cov(f_{m_n}^{(i)}, f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)}) \\
&\leq \frac{1}{k_n^4} \sum_{i\in K_n} \sum_{j\in J_i} \sum_{q\in J_i} \sum_{l\in K_n} \left| Cov(f_{m_n}^{(i)}, f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)}) \right| \\
&\leq \frac{1}{k_n^4} k_n b^2 k_n C \to 0
\end{aligned}$$

and

$$\begin{aligned}
\text{(III)} &= \frac{1}{k_n^4} \sum_{i\in K_n} \sum_{j\in J_i} \sum_{q\in J_i} \sum_{l\in J_i} Cov(f_{m_n}^{(i)}, f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)}) \\
&\leq \frac{1}{k_n^4} \sum_{i\in K_n} \sum_{j\in J_i} \sum_{q\in J_i} \sum_{l\in J_i} \left| Cov(f_{m_n}^{(i)}, f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)}) \right| \\
&\leq \frac{1}{k_n^4} k_n b^3 C \to 0.
\end{aligned}$$

If $j, q, l \notin J_i$, then we have

$$dist\Big(f_{m_n}^{(i)}, f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)}\Big) > m_n^{b_2}.$$

Assumption 1 and Lemma 4.8 imply there exists a constant C such that

$$\mathbb{E}(f_{m_n}^{(i)})^2 \leq C \quad \text{and} \quad \mathbb{E}(f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)})^2 \leq C.$$

Then Lemma 4.5 implies, for any $A > 0$,

$$\begin{aligned}
\text{(IV)} &= \frac{1}{k_n^4} \sum_{i \in K_n} \sum_{j \notin J_i} \sum_{q \notin J_i} \sum_{l \notin J_i} Cov(f_{m_n}^{(i)}, f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)}) \\
&\leq \frac{1}{k_n^4} \sum_{i \in K_n} \sum_{j \notin J_i} \sum_{q \notin J_i} \sum_{l \notin J_i} \Big| Cov(f_{m_n}^{(i)}, f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)}) \Big| \\
&\leq \frac{1}{k_n^4} k_n^4 \left[4A^2 \alpha_{m_n, 3m_n}(m_n^{b_2}) + 3\sqrt{C} \sup_{i,j,q,l} \left(\sqrt{\mathbb{E}|^A f_{m_n}^{(i)}|^2} + \sqrt{\mathbb{E}|^A \Big(f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)} \Big)|^2} \right) \right] \\
&= 4A^2 \alpha_{m_n, 3m_n}(m_n^{b_2}) + 3\sqrt{C} \sup_{i,j,q,l} \left(\sqrt{\mathbb{E}|^A f_{m_n}^{(i)}|^2} + \sqrt{\mathbb{E}|^A \Big(f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)} \Big)|^2} \right).
\end{aligned}$$

By using assumption 1, for all $\epsilon > 0$, there exists a sufficiently large A, such that

$$3\sqrt{C} \sup_{i,j,q,l} \left(\sqrt{\mathbb{E}|^A f_{m_n}^{(i)}|^2} + \sqrt{\mathbb{E}|^A \Big(f_{m_n}^{(j)} f_{m_n}^{(q)} f_{m_n}^{(l)} \Big)|^2} \right) \leq \epsilon,$$

and

$$\text{(IV)} \leq 4A^2 \alpha_{m_n, 3m_n}(m_n^{b_2}) + \epsilon.$$

Hence we have (IV) $\to 0$ as $n \to \infty$. Then, by using the symmetry of $f_{m_n}^{(j)}, f_{m_n}^{(k)}, f_{m_n}^{(l)}$, we have

$$\mathbb{E}(\bar{f}_{D_n})^4 = 3(\text{I}) + 3(\text{II}) + (\text{III}) + (\text{IV}) \to 0.$$

This completes the proof. □

Theorem 4.23. *Let $X_{i \in \mathbb{Z}^d}$ be a random field, D_n be well-divided, $D_{m_n}^{(i)}$, K_n, J_i, b_1, b_2, b, m_n and k_n be defined as in Theorem 4.22. Let $f_{m_n}^{(i)}$ be a measurable function,*

$$f_{m_n}^{(i)} = f_{m_n}(X_{j \in D_{m_n}^{(i)}}),$$

$$\bar{f}_{D_n} = \frac{1}{k_n} \sum_{i \in K_n} f_{m_n}^{(i)}.$$

We assume

1) $(f_{m_n}^{(i)})^{2k}$ are u.i., where $k \in \mathbb{Z}^+$ is a constant;

2) $\alpha_{\infty,\infty}(m^{b_2}) \to 0$ as $m \to \infty$.

Then we have $Var(\bar{f}^k_{D_n}) \to 0$ *as* $n \to \infty$.

Proof. We write

$$Var(\bar{f}^k_{D_n}) \;=\; \frac{1}{(k_n)^{2k}} \sum_{i_1 \in K_n} \cdots \sum_{i_k \in K_n} \sum_{j_1 \in K_n} \cdots \sum_{j_k \in K_n} Cov\Big(f^{(i_1)}_{m_n} \cdots f^{(i_k)}_{m_n}, f^{(j_1)}_{m_n} \cdots f^{(j_k)}_{m_n} \Big).$$

Since the idea used in this proof works for any positive integer k, we only prove this theorem for $k = 2$, and omit other cases.

The proof is similar to that of Theorem 4.22.

Assumption 1 implies $Var(f^{(i_1)}_{m_n} f^{(i_2)}_{m_n})$ and $Var(f^{(j_1)}_{m_n} f^{(j_2)}_{m_n})$ are bounded, then we have a constant C such that

$$\text{(I)} \;=\; \frac{1}{k_n^4} \sum_{i_1 \in K_n} \sum_{i_2 \in K_n} \sum_{\substack{j_1 \in J_{i_1} \cup J_{i_2} \\ \text{or} \\ j_2 \in J_{i_1} \cup J_{i_2}}} Cov(f^{(i_1)}_{m_n} f^{(i_2)}_{m_n}, f^{(j_1)}_{m_n} f^{(j_2)}_{m_n}) \le \frac{1}{k_n^4} k_n b k_n^2 C \to 0.$$

If $j_1 \notin J_{i_1} \cup J_{i_2}$ and $j_2 \notin J_{i_1} \cup J_{i_2}$, then the well-divided D_n implies

$$dist\Big(f^{(i_1)}_{m_n} f^{(i_2)}_{m_n}, f^{(j_1)}_{m_n} f^{(j_2)}_{m_n} \Big) > m_n^{b_2}.$$

Assumption 1 and Lemma 4.8 imply there exists a constant C such that

$$\mathbb{E}(f^{(i_1)}_{m_n} f^{(i_2)}_{m_n})^2 \le C \quad \text{and} \quad \mathbb{E}(f^{(j_1)}_{m_n} f^{(j_2)}_{m_n})^2 \le C.$$

Then Lemma 4.5 implies, for any $A > 0$,

$$\begin{aligned} \text{(II)} \;&=\; \frac{1}{k_n^4} \sum_{i_1 \in K_n} \sum_{i_2 \in K_n} \sum_{\substack{j_1 \notin J_{i_1} \cup J_{i_2} \\ \text{and} \\ j_2 \notin J_{i_1} \cup J_{i_2}}} Cov(f^{(i_1)}_{m_n} f^{(i_2)}_{m_n}, f^{(j_1)}_{m_n} f^{(j_2)}_{m_n}) \\ &\le\; 4A^2 \alpha_{2m_n, 2m_n}(m_n^{b_2}) + 3\sqrt{C} \sup_{i_1, i_2, j_1, j_2} \left(\sqrt{\mathbb{E}|^A \Big(f^{(i_1)}_{m_n} f^{(i_2)}_{m_n} \Big)|^2} + \sqrt{\mathbb{E}|^A \Big(f^{(j_1)}_{m_n} f^{(j_2)}_{m_n} \Big)|^2} \right). \end{aligned}$$

By using assumption 1, for all $\epsilon > 0$, there exists a sufficiently large A, such that

$$3\sqrt{C} \sup_{i_1, i_2, j_1, j_2} \left(\sqrt{\mathbb{E}|^A \Big(f^{(i_1)}_{m_n} f^{(i_2)}_{m_n} \Big)|^2} + \sqrt{\mathbb{E}|^A \Big(f^{(j_1)}_{m_n} f^{(j_2)}_{m_n} \Big)|^2} \right) \le \epsilon,$$

and

$$\text{(II)} \le 4A^2 \alpha_{2m_n, 2m_n}(m_n^{b_2}) + \epsilon.$$

Thus as $n \to \infty$, (II) converges to aero. Then we have

$$Var(\bar{f}_{D_n}) = (\mathrm{I}) + (\mathrm{II}) \to 0.$$

This completes the proof. □

4.3.1 Estimators with u.i. conditions

In this subsection we generalize the results in Section 4.2.1 to strong mixing non-stationary random fields.

Theorem 4.24. (To generalize Carlstein's Theorem 2 in [13] to random fields)

Let $X_{i\in\mathbb{Z}^d}$ be a random field, D_n be well-divided, $D^{(i)}_{m_n}$, K_n, J_i, b_1, b_2, b, m_n and k_n be defined as in Theorem 4.22. Let $f^{(i)}_{m_n}$ be a measurable function,

$$f^{(i)}_{m_n} = f_{m_n}(X_{j\in D^{(i)}_{m_n}}),$$

$$\bar{f}_{D_n} = \frac{1}{k_n}\sum_{i\in K_n} f^{(i)}_{m_n}.$$

We assume

1) $(f^{(i)}_{m_n})^2$ are u.i.;

2) $\alpha_{\infty,\infty}(m^{b_2}) \to 0$ as $m \to \infty$.

If $\mathbb{E}(\bar{f}_{D_n}) \to \varphi$ as $n \to \infty$, Then we have $\bar{f}_{D_n} \xrightarrow{L_2} \varphi$ as $n \to \infty$.

Proof. By Lemma 4.1, we only need to show $Var(\bar{f}_{D_n}) \to 0$, which follows directly using Theorem 4.23 for the case of $k = 1$. This completes the proof. □

Theorem 4.25. *Let $X_{i\in\mathbb{Z}^d}$ be a random field, D_n be well-divided, $D^{(i)}_{m_n}$, K_n, J_i, b_1, b_2, b, m_n and k_n be defined as in Theorem 4.22. Let $S^{(i)}_{m_n}$ be a measurable function,*

$$S^{(i)}_{m_n} = S_{m_n}(X_{j\in D^{(i)}_{m_n}}),$$

$$t^{(i)}_{m_n} = \sqrt{m_n}(S^{(i)}_{m_n} - \mathbb{E}S^{(i)}_{m_n}),$$

$$\Sigma_n = \frac{1}{k_n}\sum_{i\in K_n}\sum_{j\in J_i} t^{(i)}_{m_n} t^{(j)}_{m_n},$$

$$\bar{S}_n = \frac{1}{k_n}\sum_{i\in K_n} S^{(i)}_{m_n},$$

$$\hat{\sigma}_n^2 = \frac{m_n}{k_n} \sum_{i \in K_n} \sum_{j \in J_i} (S_{m_n}^{(i)} - \bar{S}_n)(S_{m_n}^{(j)} - \bar{S}_n).$$

We assume

1) $\lim_{n\to\infty} \mathbb{E}(\Sigma_n) = \sigma^2$*;*

2) $(t_{m_n}^{(i)})^4$ *are u.i.;*

3)

$$E_n \equiv \frac{m_n}{k_n} \sum_{i \in K_n} \left(\mathbb{E} S_{m_n}^{(i)} - \mathbb{E}\bar{S}_n \right)^2 \to 0 \quad as \quad n \to \infty;$$

4) $\alpha_{\infty,\infty}(m^{b_2}) \to 0$ *as* $m \to \infty$*.*

Then we have $\hat{\sigma}_n^2 \xrightarrow{L_2} \sigma^2$ *as* $n \to \infty$*.*

Proof. To simplify notations, we set

$$t_i = t_{m_n}^{(i)}, \quad \bar{t} = \frac{1}{k_n} \sum_{i \in K_n} t_i, \quad E_i = \mathbb{E} S_{m_n}^{(i)}, \quad \bar{E} = \mathbb{E}\bar{S}_{m_n}.$$

We write

$$\hat{\sigma}_n^2 \quad = \quad \Sigma_n - \text{(I)} + \text{(II)} + \text{(III)} + \text{(IV)} + \text{(V)},$$

where

$$\text{(I)} = \frac{1}{k_n} \sum_{i \in K_n} \sum_{j \in J_i} (t_i + t_j)\bar{t},$$

$$\text{(II)} = \frac{1}{k_n} \sum_{i \in K_n} \sum_{j \in J_i} \bar{t}^2,$$

$$\text{(III)} = \frac{\sqrt{m_n}}{k_n} \sum_{i \in K_n} \sum_{j \in J_i} (t_i - \bar{t})(E_j - \bar{E}),$$

$$\text{(IV)} = \frac{\sqrt{m_n}}{k_n} \sum_{i \in K_n} \sum_{j \in J_i} (t_j - \bar{t})(E_i - \bar{E}),$$

and

$$\text{(V)} = \frac{m_n}{k_n} \sum_{i \in K_n} \sum_{j \in J_i} (E_i - \bar{E})(E_j - \bar{E}).$$

Then for the rest of this proof, it is sufficient to show (I)–(V) converge to zero in L_2, and $Var(\Sigma_n) \to 0$.

By using Theorem 4.23, since assumption 2 and Lemma 4.8 imply $(\sum_{j \in J_i} t_i t_j)^2$ is u.i., then we have $Var(\Sigma_n) \to 0$.

For (II), Theorem 4.22 implies $\mathbb{E}(\bar{t}^2) \to 0$. We write

$$\bar{t}^2 = \frac{1}{k_n^2} \sum_{i \in K_n} \sum_{j \in K_n} t_i t_j.$$

Since $(t_i t_j)^2$ is u.i., By re-indexing the pair (i, j), then we have, by using Theorem 4.23, $Var(\bar{t}^2) \to 0$. Therefore, by using Lemma 4.1, we have $\bar{t}^2 \xrightarrow{L_2} 0$. Because D_n is well-divided, we have

$$\text{(II)} \leq \sup_i |J_i| \bar{t}^2 \leq b\bar{t}^2,$$

which implies, by using Lemma 4.7, (II)$\xrightarrow{L_2} 0$.

For (I), since $\bar{t}^2 \xrightarrow{L_2} 0$, Theorem 4.3 implies $(\bar{t})^4$ is u.i.. Assumption 2 and Lemma 4.8 also imply $[\sum_{j \in J_i}(t_i + t_j)]^4$ is u.i.. Therefore, Hölder's inequality ensures $\mathbb{E}(\text{I}) \to 0$, and Theorem 4.23 implies $Var((\text{I})) \to 0$. Then we have (I)$\xrightarrow{L_2} 0$.

For (III), we have $\mathbb{E}(\text{III}) = 0$ and

$$\begin{aligned}
(\text{III})^2 &= \frac{m_n}{k_n^2}\Big(\sum_{i \in K_n} \sum_{j \in J_i} (t_i - \bar{t})(E_j - \bar{E})\Big)^2 \\
&\leq \Big(\frac{1}{k_n} \sum_{i \in K_n} \sum_{j \in J_i} (t_i - \bar{t})^2\Big)\Big(\frac{m_n}{k_n} \sum_{i \in K_n} \sum_{j \in J_i} (E_i - \bar{E})^2\Big) \\
&= A_n B_n.
\end{aligned}$$

Assumption 2, Lemma 4.8 and Theorem 4.2 imply $\mathtt{E}|t_i t_j|$ is bounded. Therefore $\mathbb{E}A_n$ is bounded. Furthermore, by using assumption 3, we have

$$B_n \leq \frac{m_n}{k_n} \sum_{i \in K_n} \sum_{j \in J_i} |J_i|(E_i - \bar{E})^2 \to 0.$$

Hence (III)$\xrightarrow{L_2} 0$.

For (IV), with the same argument in (III), we have (IV)$\xrightarrow{L_2} 0$.

For (V), we note

$$\sum_{i \in K_n} \sum_{j \in J_i} (E_i - \bar{E})(E_j - \bar{E}) \quad \leq \quad \sum_{i \in K_n} |J_i|(E_i - \bar{E})^2.$$

Therefore, assumption 3 implies

$$(\text{V}) \leq \frac{m_n}{k_n} \sum_{i \in K_n} |J_i|(E_i - \bar{E})^2 \to 0.$$

This completes the proof. □

Theorem 4.26. *Let $X_{i\in\mathbb{Z}^d}$ be a random field, D_n be well-divided, $D_{m_n}^{(i)}$, K_n, J_i, b_1, b_2, b, m_n and k_n be defined as in Theorem 4.22. Let*

$$S_{m_n}^{(i)} = \sum_{j\in D_{m_n}^{(i)}} X_j, \quad S_n = \sum_{i\in K_n} S_{m_n}^{(i)}, \quad \bar{S}_{m_n} = \frac{1}{k_n} S_n,$$

$$\nu_{m_n}^{(i)} = \frac{S_{m_n}^{(i)}}{\sqrt{m_n}}, \quad \bar{\nu}_{m_n} = \frac{\bar{S}_{m_n}}{\sqrt{m_n}}, \quad \hat{\phi}_n = \frac{1}{k_n} \sum_{i\in K_n} \sum_{j\in J_i} \Big(\nu_{m_n}^{(i)} - \bar{\nu}_{m_n}\Big)\Big(\nu_{m_n}^{(j)} - \bar{\nu}_{m_n}\Big).$$

We assume

1) $(\nu_{m_n}^{(i)})^4$ are u.i.,

2)

$$\lim_{n\to\infty} \frac{Var(S_n)}{|D_n|} = \sigma^2;$$

3)

$$\frac{1}{m_n k_n} \sum_{i\in K_n} \Big(\mathbb{E}S_{m_n}^{(i)} - \mathbb{E}\bar{S}_{m_n}\Big)^2 \to 0;$$

4) $\alpha_{\infty,\infty}(m^{b_2}) \to 0$ as $m \to \infty$.

Then $\hat{\phi}_n \xrightarrow{L_2} \sigma^2$ as $n \to \infty$.

Proof. Let $t_i = \nu_{m_n}^{(i)} - \mathbb{E}\nu_{m_n}^{(i)}$ and $\bar{t} = \bar{\nu}_{m_n} - \mathbb{E}\bar{\nu}_{m_n}$. Then we have $\mathbb{E}t_i = \mathbb{E}\bar{t} = 0$. Assumption 1 implies t_i^4 is u.i.. Let

$$\Sigma_n = \frac{1}{k_n} \sum_{i\in K_n} \sum_{j\in J_i} t_i t_j, \quad E_i = \mathbb{E}S_{m_n}^{(i)}, \quad \bar{E} = \mathbb{E}\bar{S}_{m_n}.$$

Then we have

$$\hat{\phi}_n \quad = \quad \Sigma_n - (\mathrm{I}) + (\mathrm{II}) + (\mathrm{III}) + (\mathrm{IV}) + (\mathrm{V}),$$

where

$$(\mathrm{I}) = \frac{1}{k_n} \sum_{i\in K_n} \sum_{j\in J_i} (t_i + t_j)\bar{t},$$

$$(\mathrm{II}) = \frac{1}{k_n} \sum_{i\in K_n} \sum_{j\in J_i} \bar{t}^2,$$

$$(\mathrm{III}) = \frac{1}{\sqrt{m_n} k_n} \sum_{i\in K_n} \sum_{j\in J_i} (t_i - \bar{t})(E_j - \bar{E}),$$

$$(\text{IV}) = \frac{1}{\sqrt{m_n k_n}} \sum_{i \in K_n} \sum_{j \in J_i} (t_j - \bar{t})(E_i - \bar{E}),$$

and

$$(\text{V}) = \frac{1}{m_n k_n} \sum_{i \in K_n} \sum_{j \in J_i} (E_i - \bar{E})(E_j - \bar{E}).$$

By using the properties of zero means and u.i. of t_i, with the similar discussion in the proof of Theorem 4.25, we have that (I)–(IV) converge to zero in L_2, (V)$\to 0$ and $Var(\Sigma_n) \to 0$. We also have $\hat{\phi}_n - \Sigma_n \xrightarrow{L_2} 0$. Therefore the rest of this proof is to show $\mathbb{E}\Sigma_n \to \sigma^2$. We note that

$$\begin{aligned} \frac{Var(S_n)}{|D_n|} &= \frac{1}{k_n} \sum_{i \in K_n} \sum_{j \in K_n} Cov(t_i, t_j) \\ &= \mathbb{E}\Sigma_n + \frac{1}{k_n} \sum_{i \in K_n} \sum_{j \notin J_i} Cov(t_i, t_j). \end{aligned}$$

The well-divided D_n and assumption 1 imply the above second term converges to zero. Therefore assumption 2 gives $\mathbb{E}\Sigma_n \to \sigma^2$. This completes the proof. □

Theorem 4.27. *Let $X_{i \in \mathbb{Z}^d}$ be a random field, D_n be well-divided, $D^{(i)}_{m_n}$, K_n, J_i, b_1, b_2, b, m_n and k_n be defined as in Theorem 4.22. Let S_m be a measurable function,*

$$S^{(i)}_{m_n} = S_{m_n}(X_{j \in D^{(i)}_{m_N}}), \quad S_n = \sum_{i \in K_n} S^{(i)}_{m_n}, \quad \bar{S}_{m_n} = \frac{1}{k_n} S_n.$$

Let

$$\hat{\phi}_n = \sum_{i \in K_n} \sum_{j \in J_i} \left(S^{(i)}_{m_n} - \bar{S}_{m_n} \right) \left(S^{(j)}_{m_n} - \bar{S}_{m_n} \right).$$

We assume

1) There exists a constant M and δ, $\delta > 0$, such that

$$\sup_{i, m_n} \left\| \frac{(S^{(i)}_{m_n})^2}{m_n} \right\|_{2+\delta} \leq M;$$

2)

$$\liminf_{n \to \infty} \frac{Var(S_n)}{|D_n|} > 0;$$

3)

$$\frac{1}{m_n k_n} \sum_{i \in K_n} \left(\mathbb{E} S^{(i)}_{m_n} - \mathbb{E} \bar{S}_{m_n} \right)^2 \to 0 \quad as \quad n \to \infty.$$

4) $\alpha_{\infty,\infty}(m^{b_2}) \to 0$ as $m \to \infty$.

Then we have

$$\frac{\hat{\phi}_n}{Var(S_n)} \xrightarrow{L_2} 1 \quad as \quad n \to \infty.$$

Proof. We follow the idea in the proofs of Theorem 4.17 and Theorem 4.26. We need to show

(i)

$$\frac{Var(\hat{\phi}_n)}{[Var(S_n)]^2} \to 0 \quad \text{as } n \to \infty,$$

and

(ii)

$$\frac{\mathbb{E}(\hat{\phi}_n)}{Var(S_n)} \to 1 \quad \text{as } n \to \infty.$$

Let

$$\nu_{m_n}^{(i)} = \frac{S_{m_n}^{(i)}}{\sqrt{m_n}}, \quad t_i = \nu_{m_n}^{(i)} - \mathbb{E}\nu_{m_n}^{(i)}, \quad \bar{t} = \bar{\nu}_{m_n} - \mathbb{E}\bar{\nu}_{m_n}.$$

Then we have $\mathbb{E}t_i = \mathbb{E}\bar{t} = 0$. Lemma 4.11 and assumption 1 imply t_i^4 is u.i.. Let

$$\Sigma_n = \frac{1}{k_n} \sum_{i \in K_n} \sum_{j \in J_i} t_i t_j, \quad E_i = \mathbb{E}S_{m_n}^{(i)}, \quad \bar{E} = \mathbb{E}\bar{S}_{m_n}.$$

We write

$$\hat{\phi}_n = m_n k_n \Big[\Sigma_n - \text{(I)}+\text{(II)}+\text{(III)}+\text{(IV)}+\text{(V)}\Big],$$

where (I)–(V) are defined in the proof of Theorem 4.26.

For (i), assumption 2 implies it is sufficient to show

$$\frac{Var(\hat{\phi}_n)}{m_n^2 k_n^2} \to 0.$$

With the same argument as in the proof of Theorem 4.26, (I)–(IV) converge to zero in L_2, (V)$\to 0$ and $Var(\Sigma_n) \to 0$. Then we have

$$Var\Big(\frac{\hat{\phi}_n}{m_n k_n}\Big) \to 0,$$

which proves (i).

For (ii), we write

$$Var(S_n) = m_n k_n \Big[\mathbb{E}\Sigma_n + \mathbb{E}A_n\Big],$$

where

$$A_n = \frac{1}{k_n} \sum_{i \in K_n} \sum_{j \notin J_i} t_i, t_j.$$

By using Lemma 4.5, with the same argument in the proof of Theorem 4.22, assumption 4 implies $\mathbb{E}A_n \to 0$. Therefore, by using assumption 2, we have

$$\liminf_{n\to\infty} \mathbb{E}\Sigma_n > 0. \tag{4.10}$$

Since (I)–(IV) converging to zero in L_2 implies their expectation converging to zero, then

$$\frac{\mathbb{E}(\hat{\phi}_n)}{Var(S_n)} = \frac{\frac{\mathbb{E}\hat{\phi}_n}{m_n k_n}}{\frac{Var(S_n)}{|D_n|}} = \frac{\mathbb{E}\Sigma_n + \Big[\mathbb{E}(\mathrm{I}) + \mathbb{E}(\mathrm{II}) + \mathbb{E}(\mathrm{III}) + \mathbb{E}(\mathrm{IV}) + (\mathrm{V})\Big]}{\mathbb{E}\Sigma_n + \mathbb{E}A_n} \to 0,$$

which is ensured by (4.10). This completes the proof. □

4.3.2 Estimators with smooth conditions

This subsection generalizes those results in Section 4.2.2. We introduce some general assumptions and some estimators which are different to those in 1-dimension random processes.

To avoid estimating the number of these paired blocks, in this subsection, we also introduce assumptions about the following function, a function of the number of the paired blocks with respect to their distance, i.e.

$$h_{m_n}(m) = \#\left\{(i,j) : dist(D^{(i)}_{m_n}, D^{(j)}_{m_n}) = m, i, j \in K_n\right\}.$$

For example in the 2-dimensional random field in Figure 4.2(a), if we use the definition of the distance introduced in Section 2.1, we have

$$h_{m_n}(m) = \begin{cases} 9, & \text{if } m = 0, \\ 0, & \text{if } 0 < m < 1, \\ 24, & \text{if } m = 1, \\ 0, & \text{if } 1 < m < 2, \\ 16, & \text{if } 2. \end{cases}$$

We have two properties of h_{m_n} function:

(P1) Let $\alpha_{k,l}(m)$ be the strong mixing coefficient. For any $p \geq 0$, we have

$$\sum_{i\in K_n}\sum_{j\in K_n} \alpha^p_{m_n,m_n}\Big(dist(D^{(i)}_{m_n}, D^{(j)}_{m_n})\Big) = \sum_{m=0}^{\infty} h_{m_n}(m)\alpha^p_{m_n,m_n}(m).$$

(P2) We note that, for any $q, k \in \mathbb{Z}^+$ and $i_1, i_2, \cdots, i_q, j_1, j_2, \cdots, j_k \in K_n$, we have

$$\begin{aligned} & \#\left\{(i_1, i_2, \cdots, i_q, j_1, j_2, \cdots, j_k) : dist\Big(\bigcup_{t=1}^{q} D_{m_n}^{(i_t)}, \bigcup_{t=1}^{k} D_{m_n}^{(j_t)}\Big) = m\right\} \\ \leq \quad & k_n^{q+k-2} \#\left\{(i, j) : dist\Big(D_{m_n}^{(i)}, D_{m_n}^{(j)}\Big) = m\right\}, \end{aligned}$$

where $i \in \{i_1, i_2, \cdots, i_q\}$ and $j \in \{j_1, j_2, \cdots, j_k\}$. Therefore we have

$$\begin{aligned} & \sum_{i_1 \in K_n} \cdots \sum_{i_q \in K_n} \sum_{j_1 \in K_n} \cdots \sum_{j_k \in K_n} \alpha_{qm_n, km_n}^p \Big(dist\Big(\bigcup_{t=1}^{q} D_{m_n}^{(i_t)}, \bigcup_{t=1}^{k} D_{m_n}^{(j_t)}\Big)\Big) \\ = \quad & \sum_{m=0}^{\infty} \#\left\{(i_1, i_2, \cdots, i_q, j_1, j_2, \cdots, j_k) : dist\Big(\bigcup_{t=1}^{q} D_{m_n}^{(i_t)}, \bigcup_{t=1}^{k} D_{m_n}^{(j_t)}\Big) = m\right\} \alpha_{qm_n, km_n}^p \Big(m\Big) \\ \leq \quad & k_n^{q+k-2} \sum_{m=0}^{\infty} h_{m_n}(m) \alpha_{qm_n, km_n}^p(m). \end{aligned}$$

By using the above two properties, for a special case of well-divided D_n, we have the following result.

Theorem 4.28. *Let $b > 0$, $p > 0$ be constants, D_n, $D_{m_n}^{(i)}$ and K_n be defined as before, i.e.*

$$D_n = \bigcup_{i \in K_n} D_{m_n}^{(i)}.$$

We assume

1) For all $l \in \mathbb{Z}^+$ and fixed $i \in K_n$, we have

$$\#\left\{j : (l-1)m_n^b \leq dist\Big(D_{m_n}^{(i)}, D_{m_n}^{(j)}\Big) < l m_n^b\right\} \leq l^{d-1};$$

2)

$$\sum_{l=1}^{\infty} l^{d-1} \alpha_{m_n, m_n}^p(l) < \infty.$$

Then we have

$$\sum_{m=0}^{\infty} h_{m_n}(m) \alpha_{m_n, m_n}^p(m) = O(k_n) \quad as \quad n \to \infty.$$

Proof. By using the above property (P1), we have

$$\sum_{m=0}^{\infty} h_{m_n}(m) \alpha_{m_n, m_n}^p(m)$$

$$
\begin{aligned}
&= \sum_{i\in K_n}\sum_{j\in K_n} \alpha^p_{m_n,m_n}\Big(dist(D^{(i)}_{m_n}, D^{(j)}_{m_n})\Big) \\
&\le \sum_{i\in K_n}\sum_{l=1}^{\infty} \#\left\{j : (l-1)m_n^b \le dist\Big(D^{(i)}_{m_n}, D^{(j)}_{m_n}\Big) < lm_n^b\right\} \alpha^p_{m_n,m_n}\Big((l-1)m_n^b\Big) \\
&\le \sum_{i\in K_n}\sum_{l=1}^{\infty} l^{d-1}\alpha^p_{m_n,m_n}\Big((l-1)m_n^b\Big) \\
&\le \sum_{i\in K_n}\left(\alpha^p_{m_n,m_n}(0) + \sum_{l=1}^{\infty}(lm_n^b)^{d-1}\alpha^p_{m_n,m_n}\Big(lm_n^b\Big)\right) \\
&\le \sum_{i\in K_n}\left(\alpha^p_{m_n,m_n}(0) + \sum_{m=lm_n^b}^{\infty} m^{d-1}\alpha^p_{m_n,m_n}(m)\right).
\end{aligned}
$$

Because assumption 2 implies

$$
\sum_{m=lm_n^b}^{\infty} m^{d-1}\alpha^p_{m_n,m_n}(m) \to 0,
$$

we have

$$
\frac{1}{k_n}\sum_{m=0}^{\infty} h_{m_n}(m)\alpha^p_{m_n,m_n}(m) \le \alpha^p_{\infty,\infty}(0) + o(1).
$$

This completes the proof. □

Similarly, if we use assumption 1 and 2 for union blocks as in property (P2), we will have the following estimation:

$$
\begin{aligned}
&\sum_{i_1\in K_n}\cdots\sum_{i_q\in K_n}\sum_{j_1\in K_n}\cdots\sum_{j_k\in K_n} \alpha^p_{qm_n,km_n}\Big(dist\Big(\bigcup_{t=1}^{q} D^{(i_t)}_{m_n}, \bigcup_{t=1}^{k} D^{(j_t)}_{m_n}\Big)\Big) \\
\le\ & k_n^{q+k-2}\sum_{m=0}^{\infty} h_{m_n}(m)\alpha^p_{qm_n,km_n}(m) = k_n^{q+k-2}O(k_n) = O(k_n^{q+k-1}).
\end{aligned}
$$

Therefore, in following theorems, we state assumptions in terms of the h_{m_n} function. Furthermore, if D_n is well-divided, it is possible for us to introduce weaker assumptions. For example, in Theorem 4.30, Theorem 4.31 and Theorem 4.32, we assume the tail part of the sum, rather than the total part of the sum, shares the same power with k_n, i.e.

$$
\sum_{m=m_n^{b_2}}^{\infty} h_{m_n}(m)\alpha^{\frac{\delta}{2+\delta}}_{4m_n,4m_n}(m) = O(k_n) \quad \text{as} \quad n\to\infty,
$$

where $b_2 > 0$ is the constant introduced in these theorems.

Theorem 4.29. *Let $X_{i\in\mathbb{Z}^d}$ be a strong mixing random field, $D_n \subseteq \mathbb{Z}^d$, $K_n \subseteq \mathbb{Z}^d$,*

$$\bigcup_{i\in K_n} D_{m_n}^{(i)} = D_n, \quad D_{m_n}^{(i)} \cap D_{m_n}^{(j)} = \emptyset \quad \text{if} \quad i \neq j,$$

$$|D_{m_n}^{(i)}| = m_n, \quad |K_n| = k_n, \quad |D_n| = m_n k_n,$$

$$k_n \to \infty, \quad m_n \to \infty \quad \text{as} \quad n \to \infty.$$

$$h_{m_n}(m) = \#\left\{(i,j) : dist(D_{m_n}^{(i)}, D_{m_n}^{(j)}) = m, i, j \in K_n\right\}.$$

Let $f_m(\cdot)$ be a function,

$$f_m^{(i)} = f_m(X_{j\in D_{m_n}^{(i)}}), \quad \bar{f}_n = \frac{1}{k_n}\sum_{i=1}^{k_n} f_{m_n}^{(i)}.$$

We assume that

1) $m_n^2 = o(k_n)$ as $n \to \infty$;

2) $\left\|\frac{\partial f}{\partial X}\right\|_\infty = \sup_{m,i}\left|\frac{\partial f_m}{\partial X_i}\right| < \infty$, and $\sup_m f_m(0,0,\cdots,0) < \infty$;

3) There exists $\delta > 0$ s.t. $\|X\|_{2+\delta} < \infty$;

4)

$$\sum_{m=0}^{\infty} h_{m_n}(m)\alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}(m) = O(k_n) \quad \text{as} \quad n \to \infty.$$

If $\lim_{n\to\infty}\mathbb{E}(\bar{f}_n) = \varphi$ exists, then $\bar{f}_n \xrightarrow{L_2} \varphi$.

Proof. The idea of the proof is similar to that of the previous theorems. We simply need to show $Var(\bar{f}_n) \to 0$.

By using assumption 2, assumption 3, Mean Value Theorem and Minkovski inequality, we have

$$\begin{aligned}
\|f_{m_n}^{(i)}\|_{2+\delta} &\leq \|f_{m_n}(0,0,\cdots,0)\|_{2+\delta} + \sum_{i\in D_{m_n}^{(i)}} \left\|\frac{\partial f}{\partial X}\right\|_\infty \|X\|_{2+\delta} \\
&\leq C_0 + |D_{m_n}^{(i)}| C_1 C_2 \\
&\leq m_n C,
\end{aligned}$$

where C_0, C_1, C_2 and C are constants.

By using assumption 1 and 4, we have

$$\begin{aligned}
Var(\bar{f}_n) &\leq \frac{1}{k_n^2}\sum_{i=1}^{k_n}\sum_{j=1}^{k_n}|Cov(f_{m_n}^{(i)}, f_{m_n}^{(j)})| \\
&\leq \frac{1}{k_n^2}\sum_{i=1}^{k_n}\sum_{j=1}^{k_n} 8\alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}\left(dist(f_{m_n}^{(i)}, f_{m_n}^{(j)})\right)\|f_{m_n}^{(i)}\|_{2+\delta}\|f_{m_n}^{(j)}\|_{2+\delta} \\
&= \frac{1}{k_n^2}\sum_{dist(f_{m_n}^{(i)},f_{m_n}^{(j)})=0}^{\max\{dist(f_{m_n}^{(i)},f_{m_n}^{(j)})\}} 8h_{m_n}\left(dist(f_{m_n}^{(i)}, f_{m_n}^{(j)})\right)\alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}\left(dist(f_{m_n}^{(i)}, f_{m_n}^{(j)})\right) \\
&\qquad \times \|f_{m_n}^{(i)}\|_{2+\delta}\|f_{m_n}^{(j)}\|_{2+\delta} \\
&\leq \frac{8m_n^2C^2}{k_n^2}\sum_{m=0}^{\infty} h_{m_n}(m)\alpha_{m_n,m_n}^{\frac{\delta}{2=\delta}}(m) \\
&\to 0.
\end{aligned}$$

This completes the proof. □

Theorem 4.30. *Let $X_{i\in\mathbb{Z}^d}$ be a strong mixing random field, D_n, K_n, $D_{m_n}^{(i)}$, m_n, k_n and $h_n(m)$ be defined as in Theorem 4.29. Let $S_{m_n}^{(i)}$, $t_{m_n}^{(i)}$, Σ_n, $\bar{S}_n$ and $\hat{\sigma}_n^2$ be defined as in Theorem 4.25. Let D_n be well-divided, b_1, b_2 and b be defined as in Theorem 4.22. We assume*

1) $m_n^6 = o(k_n)$ as $n \to \infty$;

2) $\left\|\frac{\partial S}{\partial X}\right\|_\infty = \sup_{m,i}\left|\frac{\partial S_m}{\partial X_i}\right| < \infty$, and $\sup_m S_m(0,0,\cdots,0) < \infty$;

3) There exists $\delta > 0$ s.t. $\|X\|_{4+2\delta} < \infty$;

4)

$$\sum_{m=m_n^{b_2}}^{\infty} h_{m_n}(m)\alpha_{4m_n,4m_n}^{\frac{\delta}{2+\delta}}(m) = O(k_n) \quad as \quad n \to \infty;$$

5)

$$E_n \equiv \frac{m_n}{k_n}\sum_{i\in K_n}\left(\mathbb{E}S_{m_n}^{(i)} - \mathbb{E}\bar{S}_n\right)^2 \to 0 \quad as \quad n \to \infty.$$

If $\lim_{n\to\infty}\mathbb{E}(\Sigma_n) = \sigma^2$ exists, then we have $\hat{\sigma}_n^2 \xrightarrow{L_2} \sigma^2$ as $n \to \infty$.

Proof. Using the same idea and notations as in the proof of Theorem 4.25, since assumption 4 implies (V)$\to 0$, we are going to show (I)–(IV) converge to zero in L_2 without the u.i. assumption. It is sufficient to show $Var(\Sigma_n) \to 0$ and $\mathbb{E}(\bar{t}_n^4) \to 0$ as $n \to \infty$.

Assumptions 2 and 3 imply that, for all $i, j \in K_n$, we have $\|t_i t_j\|_{2+\delta} \leq m_n^3 C$, where C is a constant. Since D_n is well-divided, it implies $\sup|J_i| \leq b$, by using property (P2) of the h_{m_n} function, we have

$$\begin{aligned}
Var(\Sigma_n) &= \frac{1}{k_n^2} \sum_{i_1 \in K_n} \sum_{j_1 \in J_{i_1}} \sum_{i_2 \in K_n} \sum_{j_2 \in J_{i_2}} Cov\Big(t_{i_1} t_{j_1}, t_{i_2} t_{j_2}\Big) \\
&\leq \frac{b^2}{k_n^2} \sum_{i_1 \in K_n} \sum_{i_2 \in K_n} \Big|Cov\Big(t_{i_1} t_{j_1}, t_{i_2} t_{j_2}\Big)\Big| \\
&\leq \frac{b^2}{k_n^2} \sum_{i_1 \in K_n} \left(\sum_{i_2 \in J_{i_1}} \Big|Cov\Big(t_{i_1} t_{j_1}, t_{i_2} t_{j_2}\Big)\Big| + \sum_{i_2 \notin J_{i_1}} \Big|Cov\Big(t_{i_1} t_{j_1}, t_{i_2} t_{j_2}\Big)\Big| \right) \\
&\leq \frac{b^2}{k_n^2} \sum_{i_1 \in K_n} \left(b m_n^6 C^2 + \sum_{i_2 \notin J_{i_1}} 8 \alpha_{2m_n, 2m_n}^{\frac{\delta}{2+\delta}} \Big(d_{i_1, i_2}\Big) m_n^6 C^2 \right) \\
&\leq \frac{b^3 m_n^6 C^2}{k_n} + \frac{b^2}{k_n^2} \sum_{d_{i_1, i_2} = m_n^{b2}}^{\infty} h_{m_n}(d_{i_1, i_2}) 8 \alpha_{2m_n, 2m_n}^{\frac{\delta}{2+\delta}} \Big(d_{i_1, i_2}\Big) m_n^6 C^2,
\end{aligned}$$

where

$$d_{i_1, i_2} = dist\Big(t_{i_1} t_{j_1}, t_{i_2} t_{j_2}\Big).$$

Then assumptions 1 and 4 imply $Var(\Sigma_n) \to 0$.

We write

$$\begin{aligned}
\mathbb{E}(\bar{t}^4) &= \frac{1}{k_n^4} \sum_{i \in K_n} \sum_{j \in K_n} \sum_{q \in K_n} \sum_{l \in K_n} Cov\Big(t_i t_j, t_q t_l\Big) \\
&\quad + \frac{1}{k_n^4} \sum_{i \in K_n} \sum_{j \in K_n} \sum_{q \in K_n} \sum_{l \in K_n} Cov\Big(t_i, t_j\Big) Cov\Big(t_q, t_l\Big) \\
&= \text{(I)+(II)}.
\end{aligned}$$

For (I), with the similar argument to $Var(\Sigma_n)$, since $\|t_i t_j\|_{2+\delta} \leq m_n^3 C$, and D_n is well-divided, then, by using property (P2) of h_{m_n} function, we have

$$\begin{aligned}
\text{(I)} &\leq \frac{1}{k_n^4} \sum_{i \in K_n} \sum_{j \in K_n} \sum_{q \in K_n} \sum_{l \in K_n} \Big|Cov\Big(t_i t_j, t_q t_l\Big)\Big| \\
&= \frac{1}{k_n^4} \left(\sum_{i \in K_n} \sum_{j \in K_n} \sum_{q \text{ or } l \in J_i \cup J_j} \Big|Cov\Big(t_i t_j, t_q t_l\Big)\Big| \right. \\
&\qquad \left. + \sum_{i \in K_n} \sum_{j \in K_n} \sum_{q \text{ and } l \notin J_i \cup J_j} \Big|Cov\Big(t_i t_j, t_q t_l\Big)\Big| \right)
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{k_n^4}\Big(\sum_{i\in K_n}\sum_{j\in K_n} k_n m_n^6 C^2(\sup_i|J_i|)^2 + k_n^2 8 \sum_{d_{ij,ql}=m_n^{b_2}}^{\infty} h_{m_n}(d_{ij,ql})\alpha_{m_n,3m_n}^{\frac{\delta}{2+\delta}}(d_{ij,ql}) m_n^6 C^2\Big) \\
&\leq \frac{b^2 m_n^6 C^2}{k_n} + \frac{8}{k_n^2}\sum_{m=m_n^{b_2}}^{\infty} h_{m_n}(m) 8\alpha_{m_n,3m_n}^{\frac{\delta}{2+\delta}}(m) m_n^6 C^2 \to 0,
\end{aligned}
$$

where

$$
d_{ij,ql} = dist\Big(t_i t_j, t_q t_l\Big).
$$

For (II), we write

$$
\sqrt{(\mathrm{II})} = \frac{1}{k_n^2}\sum_{i\in K_n}\sum_{j\in K_n} Cov\Big(t_i, t_j\Big) = \mathbb{E}\bar{t}^2.
$$

Then Theorem 4.22 implies $\sqrt{(\mathrm{II})} \to 0$. This completes the proof. □

Theorem 4.31. *Let $X_{i\in\mathbb{Z}^d}$ be a strong mixing random field, D_n, K_n, $D_{m_n}^{(i)}$, m_n, k_n and $h_n(m)$ be defined as in Theorem 4.29. Let S_n, $S_{m_n}^{(i)}$, $\bar{S}_{m_n}$,*

$$
\hat{\phi}_n = \frac{1}{m_n k_n}\sum_{i\in K_n}\sum_{j\in J_i}\Big(S_{m_n}^{(i)} - \bar{S}_{m_n}\Big)\Big(S_{m_n}^{(j)} - \bar{S}_{m_n}\Big)
$$

be defined as in Theorem 4.26. Let D_n be well-divided, b_1, b_2 and b be defined as in Theorem 4.22. We assume

1) $m_n^2 = o(k_n)$;

2) There exists $\delta > 0$ s.t. $\|X\|_{4+2\delta} < \infty$;

3)

$$
\lim_{n\to\infty}\frac{Var(S_n)}{|D_n|} = \sigma^2;
$$

4)

$$
\sum_{m=m_n^{b_2}}^{\infty} h_n(m)\alpha_{4m_n,4m_n}^{\frac{\delta}{2+\delta}}(m) = O(k_n) \quad \text{as } n\to\infty.
$$

5)

$$
\frac{1}{m_n k_n}\sum_{i\in K_n}\Big(\mathbb{E}S_{m_n}^{(i)} - \mathbb{E}\bar{S}_{m_n}\Big)^2 \to 0;
$$

Then we have $\hat{\phi}_n \xrightarrow{L_2} \sigma^2$ as $n\to\infty$.

Proof. Let Σ_n, t_i, $\bar{t}$, (I)–(V) be defined as in Theorem 4.26. We write

$$\hat{\phi} \;=\; \Sigma_n + \text{(I)}+\text{(II)}+\text{(III)}+\text{(IV)}+\text{(V)}.$$

Since assumption 5 implies (V)$\to$ 0, for the rest of this proof, using the same idea as in the proof of Theorem 4.26 but without the u.i. assumption, we need to show (I)–(IV) converge to zero in L_2, and $Var(\Sigma_n) \to 0$. It is sufficient to show $\mathbb{E}(\bar{t}^4) \to 0$.

With the similar argument on $\mathbb{E}(\bar{t}^4)$ as in the proof of Theorem 4.30, we write $\mathbb{E}\bar{t}^4$=(I)+(II), where (I) and (II) are defined as in Theorem 4.30. Then we have, by using Theorem 4.22, $\sqrt{(\text{II})} \to 0$. Assumption 2 implies $\|t_i t_j\|_{2+\delta} \leq m_n C$, where C is a constant. Therefore the argument involving (I) in the proof of Theorem 4.30 implies

$$\begin{aligned}
\text{(I)} \;&\leq\; \frac{1}{k_n^4}\Big(\sum_{i\in K_n}\sum_{j\in K_n} k_n m_n^2 C^2 (\sup_i |J_i|)^2 + k_n^2 8 \sum_{d_{ij,ql}=m_n^{b_2}}^{\infty} h_{m_n}(d_{ij,ql})\alpha_{m_n,3m_n}^{\frac{\delta}{2+\delta}}(d_{ij,ql}) m_n^2 C^2\Big) \\
&\leq\; \frac{b^2 m_n^2 C^2}{k_n} + \frac{8}{k_n^2}\sum_{m=m_n^{b_2}}^{\infty} h_{m_n}(m) 8\alpha_{m_n,3m_n}^{\frac{\delta}{2+\delta}}(m) m_n^2 C^2 \to 0,
\end{aligned}$$

This completes the proof. □

Theorem 4.32. *Let $X_{i\in\mathbb{Z}^d}$ be a strong mixing random field, D_n, K_n, $D_{m_n}^{(i)}$, m_n, k_n and $h_n(m)$ be defined as in Theorem 4.29. Let S_n, $S_{m_n}^{(i)}$, $\bar{S}_{m_n}$ and*

$$\hat{\phi}_n = \sum_{i\in K_n}\sum_{j\in J_i}\Big(S_{m_n}^{(i)} - \bar{S}_{m_n}\Big)\Big(S_{m_n}^{(j)} - \bar{S}_{m_n}\Big)$$

be defined as in Theorem 4.27 Let D_n be well-divided, b_1, b_2 and b be defined as in Theorem 4.22.

We assume

1) $m_n^2 = o(k_n)$ as $n \to \infty$;

2) $\left\|\frac{\partial S}{\partial X}\right\|_\infty = \sup_{m,i}\left|\frac{\partial S_m}{\partial X_i}\right| < \infty$, and $\sup_m S_m(0,0,\cdots,0) < \infty$;

3) there exists $\delta > 0$ s.t. $\|X\|_{4+2\delta} < \infty$;

4) $\liminf_{n\to\infty} \frac{Var(S_n)}{|D_n|} > 0$;

5)

$$\sum_{m=m_n^{b_2}}^{\infty} h_{m_n}(m)\alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}(m) = O(k_n) \qquad \text{as } n \to \infty;$$

6)

$$\frac{1}{m_n k_n}\sum_{i\in K_n}\Big(\mathbb{E}S_{m_n}^{(i)} - \mathbb{E}\bar{S}_{m_n}\Big)^2 \to 0 \quad \text{as} \quad n \to \infty.$$

Then we have

$$\frac{\hat{\phi}_n}{Var(S_n)} \xrightarrow{L_2} 1 \quad as \quad n \to \infty.$$

Proof. We use the same ideas as in the proofs of Theorem 4.27 and Theorem 4.30. Let Σ_n, t_i, $\bar{t}$, (I)–(V) be defined as in the proof of Theorem 4.27. Therefore, similar to the arguments in the proofs of Theorem 4.30 and Theorem 4.31, it is sufficient to show $\mathbb{E}(\bar{t}^4) \to 0$.

Assumption 2 and 3 imply $\|t_i t_j\|_{2+\delta} \leq m_n C$, where C is a constant. Now, with the same argument on $\mathbb{E}(\bar{t}^4)$ in the proof of Theorem 4.31, we have $\mathbb{E}(\bar{t}^4) \to 0$. This completes the proof. □

4.3.3 CLTs with estimated variances

This subsection provides CLTs with those estimated variances from previous theorems. We firstly introduce two preliminary lemmas.

Lemma 4.33. *Let X_n be random sequence and a_n be a real-value sequence. If $\mathbb{E}|X_n - a_n|^2 \to 0$ and $\liminf_n a_n > 0$, then $X_n/a_n \xrightarrow{P} 1$.*

Proof. Because $\liminf_n a_n > 0$ implies $1/a_n < C$, where C is a constant, we have

$$\mathbb{P}\left(\left|\frac{X_n}{a_n} - 1\right| > \epsilon\right) \leq \frac{\mathbb{E}|X_n - a_n|^2}{\epsilon^2 a_n^2} = \frac{1}{a_n^2}\frac{\mathbb{E}|X_n - a_n|^2}{\epsilon^2} \to 0 \quad \text{as} \quad n \to \infty.$$

□

Lemma 4.34. *Let X_n and Y_n be two random sequences. If*

1) $X_n/\sqrt{Var(X_n)} \xrightarrow{D} N(0,1)$;

2) $\mathbb{E}|Y_n - \sqrt{Var(X_n)}|^2 \to 0$;

3) $\liminf_n Var(X_n) > 0$.

Then $X_n/Y_n \xrightarrow{D} N(0,1)$.

Proof. By using Lemma 4.33, assumption 2 and 3 imply $Y_n/\sqrt{Var(X_n)} \xrightarrow{P} 1$. Then Slutsky's Theorem implies

$$\frac{X_n}{Y_n} = \frac{X_n}{\sqrt{Var(X_n)}}\frac{\sqrt{Var(X_n)}}{Y_n} \xrightarrow{D} N(0,1).$$

□

For a random field $X_{i\in\mathbb{Z}^d}$, $D_n \subseteq \mathbb{Z}^d$, Lemma 4.34 can be easily generalized by using $|D_n|$ to take the place of n with the condition of $|D_n| \to \infty$ as $n \to \infty$. It means Lemma 4.34 also holds for random fields. The following theorems are directly combined by previous L_2 consistency estimators, CLTs and Lemma 4.34.

Theorem 4.35. *Let $X_{i\in\mathbb{Z}^d}$ be a zero-mean random field. The strong mixing coefficient of X satisfy (H1)–(H3) of Theorem 3.1. Let D_n be well-divided, $D_{m_n}^{(i)}$, K_n, J_i, b_1, b_2, b, m_n and k_n be defined as in Theorem 4.22. Let S_n, $S_{m_n}^{(i)}$, $\bar{S}_{m_n}$,*

$$\hat{\phi}_n = \frac{1}{m_n k_n} \sum_{i\in K_n} \sum_{j\in J_i} \left(S_{m_n}^{(i)} - \bar{S}_{m_n}\right)\left(S_{m_n}^{(j)} - \bar{S}_{m_n}\right)$$

be defined as in Theorem 4.26. We assume

1) $(\frac{S_{m_n}^{(i)}}{\sqrt{m_n}})^4$ *are u.i.;*

2)

$$\lim_{n\to\infty} \frac{Var(S_n)}{|D_n|} = \sigma^2 > 0;$$

3)

$$\frac{1}{m_n k_n} \sum_{i\in K_n} \left(\mathbb{E}S_{m_n}^{(i)} - \mathbb{E}\bar{S}_{m_n}\right)^2 \to 0;$$

4) $\alpha_{\infty,\infty}(m^{b_2}) \to 0$ *as* $m \to \infty$.

Then $S_n/\sqrt{\hat{\phi}_n} \xrightarrow{D} N(0,1)$ *as* $n \to \infty$.

Proof. Because assumption 2 implies (H4) of Theorem 3.1, this proof is completed by using Theorem 3.1, Theorem 4.26 and Lemma 4.34. In using Lemma 4.34, we set $X_n = S_n$ and $Y_n = \sqrt{\hat{\phi}_n}$. □

Theorem 4.36. *Let $X_{i\in\mathbb{Z}^d}$ be a zero-mean random field, and suppose that the strong mixing coefficient of X satisfies (H1)–(H4) of Theorem 3.1. Let D_n be well-divided, $D_{m_n}^{(i)}$, K_n, J_i, b_1, b_2, b, m_n and k_n be defined as in Theorem 4.22. Let*

$$S_{m_n}^{(i)} = \sum_{j\in D_{m_n}^{(i)}} X_j, \quad S_n = \sum_{i\in K_n} S_{m_n}^{(i)}, \quad \bar{S}_{m_n} = \frac{1}{k_n} S_n,$$

Let

$$\hat{\phi}_n = \sum_{i\in K_n} \sum_{j\in J_i} \left(S_{m_n}^{(i)} - \bar{S}_{m_n}\right)\left(S_{m_n}^{(j)} - \bar{S}_{m_n}\right).$$

We assume

1) For the δ in (H3),

$$\sup_{i,m_n}\left\|\frac{(S_{m_n}^{(i)})^2}{m_n}\right\|_{2+\delta}\leq M;$$

2)

$$\frac{1}{m_nk_n}\sum_{i\in K_n}\left(\mathbb{E}S_{m_n}^{(i)}-\mathbb{E}\bar{S}_{m_n}\right)^2\to 0\quad as\quad n\to\infty.$$

Then $S_n/\sqrt{\hat{\phi}_n}\xrightarrow{D}N(0,1)$ as $n\to\infty$.

Proof. This proof is completed by using Theorem 3.1, Theorem 4.27 and Lemma 4.34. □

Theorem 4.37. *Let $X_{i\in\mathbb{Z}^d}$ be a zero-mean random field, and suppose that the strong mixing coefficient of X satisfies (H1)–(H3) of Theorem 3.1. Let $X_{i\in\mathbb{Z}^d}$ be a strong mixing random field, D_n, K_n, $D_{m_n}^{(i)}$, m_n, k_n and $h_n(m)$ be defined as in Theorem 4.29. Let S_n, $S_{m_n}^{(i)}$, $\bar{S}_{m_n}$,*

$$\hat{\phi}_n=\frac{1}{m_nk_n}\sum_{i\in K_n}\sum_{j\in J_i}\left(S_{m_n}^{(i)}-\bar{S}_{m_n}\right)\left(S_{m_n}^{(j)}-\bar{S}_{m_n}\right)$$

be defined as in Theorem 4.26. Let D_n be well-divided, b_1,b_2 and b be defined as in Theorem 4.22. We assume

1) $m_n^4=o(k_n)$;

2) There exists $\delta>0$ s.t. $\|X\|_{4+2\delta}<\infty$ (implies $\|X\|_{2+\delta}<\infty$ in (H3));

3)

$$\lim_{n\to\infty}\frac{Var(S_n)}{|D_n|}=\sigma^2>0\quad (implies\ (H4));$$

4)

$$\sum_{m=m_n^{b_2}}^{\infty}h_n(m)\alpha_{4m_n,4m_n}^{\frac{\delta}{2+\delta}}(m)=O(k_n)\qquad as\ n\to\infty.$$

5)

$$\frac{1}{m_nk_n}\sum_{i\in K_n}\left(\mathbb{E}S_{m_n}^{(i)}-\mathbb{E}\bar{S}_{m_n}\right)^2\to 0;$$

Then $S_n/\hat{\sigma}_n\xrightarrow{D}N(0,1)$ as $n\to\infty$.

Proof. This proof is completed by using Theorem 3.1, Theorem 4.31 and Lemma 4.34. □

Theorem 4.38. *Let $X_{i\in\mathbb{Z}^d}$ be a zero-mean random field, and suppose that the strong mixing coefficient of X satisfies (H1)–(H4) of Theorem 3.1. Let $X_{i\in\mathbb{Z}^d}$ be a strong mixing random field,*

D_n, K_n, $D_{m_n}^{(i)}$, m_n, k_n and $h_n(m)$ be defined as in Theorem 4.29. Let S_n, $S_{m_n}^{(i)}$, $\bar{S}_{m_n}$ and

$$\hat{\phi}_n = \sum_{i\in K_n}\sum_{j\in J_i}\left(S_{m_n}^{(i)} - \bar{S}_{m_n}\right)\left(S_{m_n}^{(j)} - \bar{S}_{m_n}\right)$$

be defined as in Theorem 4.27 Let D_n be well-divided, b_1, b_2 and b be defined as in Theorem 4.22.

We assume

1) $m_n^4 = o(k_n)$ as $n \to \infty$;

2) $\left\|\frac{\partial S}{\partial X}\right\|_\infty = \sup_{m,i}\left|\frac{\partial S_m}{\partial X_i}\right| < \infty$, and $\sup_m S_m(0,0,\cdots,0) < \infty$;

4)

$$\sum_{m=m_n^{b_2}}^{\infty} h_{m_n}(m)\alpha_{m_n,m_n}^{\frac{\delta}{2+\delta}}(m) = O(k_n) \quad \text{as } n \to \infty;$$

5)

$$\frac{1}{m_n k_n}\sum_{i\in K_n}\left(\mathbb{E}S_{m_n}^{(i)} - \mathbb{E}\bar{S}_{m_n}\right)^2 \to 0 \quad \text{as} \quad n \to \infty.$$

Then we have

$$\frac{S_n}{\sqrt{\hat{\phi}_n}} \xrightarrow{D} N(0,1) \quad \text{as} \quad n \to \infty.$$

Proof. This proof is completed by using Theorem 3.1, Theorem 4.32, Lemma 4.4 and Lemma 4.34. □

4.3.4 Estimators for finite populations

In this subsection, we develop Fuller's central limit theorem, Theorem 1.3.2 in [26], into dependent non-stationary random fields.

Let a finite population, $x_{t\in D_N}$, be from a random field $X_{t\in\mathbb{Z}^d}$, i.e. $D_N \subseteq \mathbb{Z}^d$. Let $R_{t\in\mathbb{Z}^d}$ be defined by (2.10). We define

$$\begin{aligned}
A &= \{t \in D_N : R_t = 1\},\\
\bar{x} &= \frac{1}{|A|}\sum_{t\in A} x_t,\\
\bar{x}_N &= \frac{1}{|D_N|}\sum_{t\in D_N} x_t,\\
c_t &= \begin{cases} \frac{1}{|A|} - \frac{1}{|D_N|}, & R_t = 1,\\ -\frac{1}{|D_N|}, & R_t = 0,\end{cases}
\end{aligned}$$

Theorem 4.39. *Let $X_{i\in\mathbb{Z}^d}$ be independent to $R_{i\in\mathbb{Z}^d}$. $R_{i\in\mathbb{Z}^d}$ satisfies (H1) and (H2). Let $X_{i\in\mathbb{Z}^d}$ satisfy (H1)–(H3), and for the δ in (H3) of Theorem 3.1, assume that R also satisfies (H3). For the finite subset, D_N, of $\mathbb{Z}^d$, let $x_{t\in D_N}$ be a realization of $X_{t\in D_N}$, and $A \subset D_N$. We set*

$$S_N = \sum_{i\in D_N} c_t x_t, \quad \sigma_N^2 = Var(S_N).$$

If

$$\liminf_N \frac{\sigma_N^2}{|D_N|} > 0, \tag{4.11}$$

then we have $\bar{x} - \bar{x}_N = S_N$ and

$$\frac{S_N}{\sigma_N} \xrightarrow{\mathcal{D}} N(0,1), \qquad \text{as } N \to \infty. \tag{4.12}$$

Proof. We note that $\bar{x} - \bar{x}_N = \sum_{t\in D_N} c_t x_t$, which proved $\bar{x} - \bar{x}_N = S_N$.

Let two possible values of c_t be $a_1 = \frac{1}{|A|} - \frac{1}{|D|}$ and $a_0 = -\frac{1}{|D|}$. It is obvious that $a_0 \neq 0$, $a_1 \neq 0$ and $a_0 \neq a_1$. Then we have

$$R_t = \frac{c_t - a_1}{a_0 - a_1}, \quad t \in \mathbb{Z}^d.$$

Therefore $\sigma(c_{t\in\Lambda}) = \sigma(R_{t\in\Lambda})$ for all subsets $\Lambda \subseteq \mathbb{Z}^d$. It also means we have $\alpha^c_{k,l}(m) = \alpha^R_{k,l}(m)$. By using Lemma 2.8, it is easy to check that $\{c_t X_t\}_{t\in\mathbb{Z}^d}$ satisfies (H1)–(H3). Equation (4.11) implies that $\{c_t x_t\}_{t\in\mathbb{Z}^d}$ satisfies (H4) of Theorem 3.1. Then (4.12) is a straightforward consequence of Theorem 3.1. □

Theorem 4.39 implies that if there exists an estimator, $\hat{\sigma}_N^2$, satisfying the assumptions in Theorem 4.35, Theorem 4.36, Theorem 4.37 or Theorem 4.38, then we have

$$\frac{S_N}{\hat{\sigma}_N} \xrightarrow{\mathcal{D}} N(0,1) \quad \text{as } N \to \infty. \tag{4.13}$$

Furthermore, we suppose this finite population, D_N, is in H strata, and h stands for the h-th stratum, $h \in \{1, \cdots, H\}$. Then D_h is the set of subscripts of the population within the h-th stratum. Let $R_{t\in\mathbb{Z}^d}$ is defined by (2.10) and

$$\begin{aligned}
D_N &= \bigcup_{h=1}^{H} D_h, \\
A_h &= \{t \in D_h : R_t = 1\}, \\
A_N &= \bigcup_{h=1}^{H} A_h, \\
\bar{x}_h &= \frac{1}{|A_h|} \sum_{t\in A_h} x_t,
\end{aligned}$$

$$
\begin{aligned}
\bar{x}_N &= \frac{1}{|D_N|}\sum_{h=1}^{H}\sum_{t\in D_h} x_t = \frac{1}{|D_N|}\sum_{t\in D_N} x_t, \\
\bar{x} &= \sum_{h=1}^{H}\frac{|D_h|}{|D_N|}\bar{x}_h, \\
c_t &= \begin{cases} \frac{1}{|A_N|} - \frac{1}{|D_N|}, & R_t = 1, \\ -\frac{1}{|D_N|}, & R_t = 0. \end{cases}
\end{aligned}
$$

Then Theorem 4.39 still holds.

We note that D_h in here is a kind of partition of D_N based on strata, however $D_{m_n}^{(i)}$ in previous theorems is another partition of $\mathbb{Z}^d$ based on the settings in the assumptions. There is no confusion between them.

Since the dependence between strata can also be described by the dependence between persons, in this case, the estimator $\hat{\sigma}_N^2$ defined as in Theorem 4.26 still works for the result as per (4.13) with suitable assumptions in Theorem 4.35, Theorem 4.36, Theorem 4.37 or Theorem 4.38.

For indicated random fields, $\{R_i X_i\}_{i\in\mathbb{Z}^d}$, $d \geq 1$, central limit theorems in Chapter 3 imply that, with their required assumptions, added by the assumptions used in this section, all the theorems, such as Theorem 4.35, Theorem 4.36, Theorem 4.37 and Theorem 4.38, hold for the indicated sampling method which is introduced in Chapter 1.

Chapter 5

Functional central limit theorems

The functional central limit theorem is a kind of development of the central limit theorem on an infinite dimensional space. It considers the asymptotic property of a series of random processes which are indexed by time, t. It is also called a Donsker theorem or invariance principle. To consider the convergence of these random processes, we need to set up a suitable space and a topology with a measurement of distance between two elements in this space. Space $C = C[0,1]$, which is the space of continuous functions, is one such space. However, it cannot be used to describe random processes with jumps. Therefore, we introduce another space, space D, endowed with the Skorohod topology. The preliminaries of this chapter are mainly adapted from [4] by Billingsley and [57] by Rio. We prove FCLTs for nested sampled non-stationary dependent random fields in Section 5.2. Section 5.3 directly applies the results in Section 5.2 to the indicated sampling strategy, which was introduced in Chapter 1.

5.1 Space D and Skorohod topology

Let $D = D[0,1]$ be the space of *cadlag* functions, i.e. $x : [0,1] \mapsto \mathbb{R}$, where $x \in D$ is a right-continuous function, and the left-limits exist for any points of x. Let x and y be two elements in the space D. We introduce two metrics, which are defined in [4] by Billingsley,

$$d(x,y) = \inf\{\epsilon : \sup_t |\lambda(t) - t| \leq \epsilon, \sup_t |x(t) - y(\lambda(t))| \leq \epsilon\},$$

and

$$d_0(x,y) = \inf\{\epsilon : \|\lambda\| \leq \epsilon, \sup_t |x(t) - y(\lambda(t))| \leq \epsilon\},$$

where $\lambda(t)$ is in the class of strictly increasing, continuous mappings from $[0,1]$ onto itself,

$$\|\lambda\| = \sup_{s \neq t} \left| \log \frac{\lambda(t) - \lambda(s)}{t - s} \right|,$$

and $\lambda(t)$ satisfies $\|\lambda\| < \infty$. The Skorohod topology is defined by the metric $d(\cdot,\cdot)$. In Chapter 14 of [4], Theorem 14.1 proves that d and d_0 are equivalent metrics, Theorem 14.2 and Theorem 14.3 proved that, with the metric $d_0(\cdot,\cdot)$, the space D is complete and separable. Thereupon we can consider the weak convergence and the tightness of a sequence of random elements in the space D.

Definition 5.1. ([4]) Let $W_n(t)$ be random elements of D. We say $W_n(t)$ has *asymptotically independent increments* if

$$0 \leq s_1 \leq t_1 < s_2 \leq t_2 \cdots < s_r \leq t_r \leq 1$$

implies, for all Borel sets $B_1, \cdots, B_r$ of $\mathbb{R}$, that

$$\left| \mathbb{P}\left\{ W_n(t_i) - W_n(s_i) \in B_i, i = 1, \cdots, r \right\} - \prod_{i=1}^{r} \mathbb{P}(W_n(t_i) - W_n(s_i) \in B_i) \right| \to 0, \text{ as } n \to \infty.$$

Similarly, we say two random sequences $X_{n\in\mathbb{N}}$ and $Y_{n\in\mathbb{N}}$ are *asymptotically independent* if and only if for all Borel sets B_1, B_2 of $\mathbb{R}$, that

$$\left| \mathbb{P}\left\{ X_n \in B_1, Y_n \in B_2 \right\} - \mathbb{P}\left\{ X_n \in B_1 \right\} \mathbb{P}\left\{ Y_n \in B_2 \right\} \right| \to 0, \text{ as } n \to \infty.$$

Lemma 5.1.
a) Let two random sequences X_n and Y_n be asymptotically independent. Let X and Y be independent random variables. If $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{D} Y$, then we have $X_n \pm Y_n \xrightarrow{D} X \pm Y$.

b) Let X_n and Y_n be two random sequences. Let $\{X_n \pm Y_n\}_{n\in\mathbb{N}}$ and Y_n be asymptotically independent, $X_n \xrightarrow{D} X$, $Y_n \xrightarrow{D} Y$. If $X \pm Y$ is independent to Y, and Y has the non-zero characteristic function, then $X_n \pm Y_n \xrightarrow{D} X \pm Y$.

Proof. a) We only need to prove the addition case. Since two random sequences X_n and Y_n are asymptotically independent, then we have $|\mathbb{E}e^{it(X_n+Y_n)} - \mathbb{E}e^{itX_n}\mathbb{E}e^{itY_n})| \to 0$. We note that $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{D} Y$ imply

$$|\mathbb{E}e^{itX_n} - \mathbb{E}e^{itX}| \to 0, \quad |\mathbb{E}e^{itY_n} - \mathbb{E}e^{itY})| \to 0.$$

Therefore

$$\left| \mathbb{E}e^{it(X_n+Y_n)} - \mathbb{E}e^{it(X+Y)} \right| = \left| \mathbb{E}e^{it(X_n+Y_n)} - \mathbb{E}e^{itX}\mathbb{E}e^{itY} \right|$$

$$
\begin{aligned}
&\leq \quad |\mathbb{E}e^{it(X_n+Y_n)} - \mathbb{E}e^{itX_n}\mathbb{E}e^{itY_n})| \\
&\quad\quad +|\mathbb{E}e^{itX_n} - \mathbb{E}e^{itX}|+|\mathbb{E}e^{itY_n} - \mathbb{E}e^{itY})| \\
&\to \quad 0.
\end{aligned}
$$

This implies $X_n + Y_n \xrightarrow{D} X + Y$.

b) We note that the non-negative characteristic function of Y implies the non-negative characteristic function of $-Y$. Therefore we only prove the substitution case.

$$
\begin{aligned}
\left|\mathbb{E}e^{it(X_n-Y_n)} - \mathbb{E}e^{it(X-Y)}\right| &= \left|\frac{1}{\mathbb{E}e^{itY}}\left(\mathbb{E}e^{it(X_n-Y_n)}\mathbb{E}e^{itY} - \mathbb{E}e^{it(X-Y)}\mathbb{E}e^{itY}\right)\right| \\
&\leq \left|\frac{1}{\mathbb{E}e^{itY}}\right| \times \\
&\quad \Big(\left|\mathbb{E}e^{it(X_n-Y_n)}\left(\mathbb{E}e^{itY} - \mathbb{E}e^{itY_n}\right)\right| \\
&\quad\quad +\left|\mathbb{E}e^{it(X_n-Y_n)}\mathbb{E}e^{itY_n} - \mathbb{E}e^{it(X_n-Y_n)+itY_n}\right| \\
&\quad\quad\quad +\left|\mathbb{E}e^{itX_n} - \mathbb{E}e^{itX}\right|\Big) \\
&= \left|\frac{1}{\mathbb{E}e^{itY}}\right| \times \Big((\mathrm{I})+(\mathrm{II})+(\mathrm{III})\Big).
\end{aligned}
$$

Since (I) (II) and (III) all go to zero by the assumptions, this completes the proof. □

Theorem 5.2. (Theorem 19.2 in [4])*Suppose that $W_n(t)$ satisfies the following conditions:*

1. *$W_n(t)$ has asymptotically independent increments;*

2. *$\{W_n^2(t)\}_{n\geq 1}$ is uniformly integrable for each t;*

3. *$\mathbb{E}(W_n(t)) \to 0$ and $\mathbb{E}(W_n^2(t)) \to t$ as $n \to \infty$;*

4. *For each positive ϵ and η, there exists a positive δ such that*

$$
\mathbb{P}\left(\sup_{|s-t|<\delta} |W_n(s) - W_n(t)| \geq \epsilon\right) \leq \eta
$$

for sufficiently large n. Then we have

$$
W_n(t) \xrightarrow{\mathcal{D}} W(t).
$$

The above theorem is a criterion for convergence introduced in [4]. It has been used in some work such as [12] by Bradley and Peligrad for one dimensional strictly stationary processes with strong mixing conditions and a polynomial mixing rate, [31] by Herrndorf for non-stationary

processes with ρ-mixing conditions, and been developed by Deo for a high dimensional time index and stationary random fields with ϕ-mixing conditions in [18]. However another criteria of convergence in Billingsley's book [4], Theorem 15.6, is also used in this section.

Theorem 5.3. (Theorem 15.6 of [4]) *Let $F(t)$ be a continuous non-decreasing function on $[0,1]$. We suppose that $W_n(t)$ and $W(t)$ are elements in space D, and the following conditions are satisfied:*

1). $(W_n(t_1), W_n(t_2), \cdots, W_n(t_k)) \xrightarrow{\mathcal{D}} (W(t_1), W(t_2), \cdots, W(t_k))$ *whenever* $t_1, t_2, \cdots, t_k \in [0,1]$*;*

2). $\mathbb{P}(W(1) \neq W(1-)) = 0$*;*

3).

$$\mathbb{P}(|W_n(t) - W_n(t_1)| \geq \lambda, |W_n(t) - W_n(t_2)| \geq \lambda) \leq \frac{1}{\lambda^{2\gamma}}[F(t_2) - F(t_1)]^{2\alpha}, \tag{5.1}$$

for $t_1 \leq t \leq t_2$ *and* $n \geq 1$*, where* $\gamma \geq 0$*,* $\alpha > \frac{1}{2}$*.*

Then we have

$$W_n(t) \xrightarrow{\mathcal{D}} W(t).$$

5.2 FCLTs on non-stationary random fields

Generally speaking, the stationarity property will benefit the estimation of the variance of the sample sum. Let X_i, $i \in \mathbb{Z}$, be a one dimensional zero mean random process. Then $Var \sum_{i=1}^n X_i = \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}X_iX_j$. It means, although X_i are well bounded, the variance still has order n^2. However, if this sequence is endowed with the stationarity property, we have $Cov(X_i, X_j) = Cov(X_0, X_{j-i})$. Therefore, Theorem 1.2, in the book [42] by Lindgren et al, implies $|Cov(X_i, X_j)| \leq Cov(X_0, X_0)$. Now we have

$$Var \sum_{i=1}^n X_i = \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}X_iX_j = nCov(X_0, X_0)\left[1 + 2\sum_{k=1}^n \frac{n-k}{n}\frac{Cov(X_0, X_k)}{Cov(X_0, X_0)}\right].$$

Since $\frac{n-k}{n} < 1$ and $\frac{Cov(X_0,X_k)}{Cov(X_0,X_0)} \leq 1$, it implies the possibility of an estimation of the variance with a lower order compared to n^2.

There are many works of FCLTs on stationary random processes and/or fields. For example, in Theorem 20.1 of Billingsley's book [4], φ-mixing conditions are introduced for the one dimensional stationary processes. In the paper [48] by Merlevède and Peligrad, the strong mixing conditions and quantile conditions are introduced to one dimensional strictly stationary processes. In [57] by Rio, the strong mixing conditions are introduced to one dimensional strictly stationary processes. Fazekas [23] introduced continuous parameter fields and infill-increasing assumptions to strictly stationary random fields.

In [1] by Andrew and Pollard, bracketing conditions are introduced to FCLTs on non-stationary strong mixing random process. In this section, we will introduce a FCLT which works for non-stationary random fields with strong mixing dependent properties, quantile conditions and the nested sampling strategy. Another FCLT involves φ-mixing conditions in order to relax the previous quantile conditions.

Following Rio's definition of a "tail" function in [57], we let the quantile function of $|X_i|$ be the inverse of the non increasing and left continuous tail function of $|X_i|$, which is noted by $Q_{|X_i|}$ or Q_i. Specifically,

$$Q_{|X_i|}(u) = \inf\{x \geq 0 : \mathbb{P}(|X_i| > x) \leq u\}.$$

Some properties of the quantile function are provided in [57] such as Theorem 1.1 and Lemma 2.1 which will be our main tools in the following proofs. We also need some preliminaries on complex-valued random variables such as the following lemma, Lemma 5.5, which is a relaxed version of exercise 7 at the end of Chapter 1 in Rio's book [57].

Comparing our Definition 2.1 of strong mixing coefficient with Rio's definition in Chapter 1 of [57], we use m to specify the distance between two blocks, Λ_1 and Λ_2. We also use the subscript of $\alpha_{k,l}$ to emphasize the sizes of Λ_1 and Λ_2, where $|\Lambda_1| = k$ and $|\Lambda_2| = l$. Rio introduced a factor, 2, in front of our $\alpha_{k,l}(m)$. This factor brings convenience to the proof of the following theorem, Theorem 1.1 in the book [57]. Rio's definition is the same as the form, $2\alpha_{k,l}(m)$ in our case on random fields. In the following lemmas, we assume X is one of $X_{i\in\Lambda_1}$, and Y is one of $X_{i\in\Lambda_2}$, $dist(\Lambda_1, \Lambda_2) \geq m$, in Definition 2.1. Furthermore, for convenience, we let $\alpha = \alpha_{|\Lambda_1|,|\Lambda_2|}(m)$. Then we have the following two results, Lemma 5.4 and Lemma 5.6, by Rio.

Lemma 5.4. (Theorem 1.1 (a) in [57]) *Let X and Y be integrable real-valued random variables. Assume that XY is integrable. Then*

$$|Cov(X,Y)| \leq 2\int_0^{2\alpha} Q_{|X|}(u)Q_{|Y|}(u)du \leq 4\int_0^{\alpha} Q_{|X|}(u)Q_{|Y|}(u)du.$$

Lemma 5.5. *Let X and Y be complex-valued random variables. Then we have*

$$|Cov(X,Y)| \leq 8\int_0^{2\alpha} Q_{|X|}(u)Q_{|Y|}(u)du \leq 16\int_0^{\alpha} Q_{|X|}(u)Q_{|Y|}(u)du,$$

where α is the corresponding strong mixing coefficient with respect to X and Y, and $|X|$ is the modulus of a complex number X.

Proof. Let $X = X_1 + iX_2$, $Q_{|X_1|}(u) = x_1$, $Q_{|X_2|}(u) = x_2$ and $Q_{|X|}(u) = x^*$ for some fixed u. Since $|X| \geq |X_1|$, we have

$$\mathbb{P}(|X_1| > x^*) \leq \mathbb{P}(|X| > x^*) \leq u.$$

Therefore we have $x^* \geq x_1$, i.e.

$$Q_{|X_1|}(u) \leq Q_{|X|}(u).$$

Similarly we have $Q_{|X_2|}(u) \leq Q_{|X|}(u)$. If we let $Y = Y_1 + iY_2$, then we have

$$Q_{|Y_1|}(u) \leq Q_{|Y|}(u), \quad Q_{|Y_2|}(u) \leq Q_{|Y|}(u).$$

Using the above inequalities, and the first inequality of Lemma 5.4, we have

$$\begin{aligned} |Cov(X,Y)| &\leq |Cov(X_1,Y_1)|+|Cov(X_2,Y_1)|+|Cov(X_1,Y_2)|+|Cov(X_2,Y_2)| \\ &\leq 2\int_0^{2\alpha} \big[Q_{|X_1|}(u)Q_{|Y_1|}(u) + Q_{|X_2|}(u)Q_{|Y_1|}(u) \\ &\qquad\qquad +Q_{|X_2|}(u)Q_{|Y_1|}(u) + Q_{|X_2|}(u)Q_{|Y_2|}(u)\big]\, du \\ &\leq 8\int_0^{2\alpha} Q_{|X|}(u)Q_{|Y|}(u)du. \end{aligned}$$

Similarly, the second inequality of Lemma 5.4 implies the second inequality of this lemma. This completes the proof. □

Since $Q_{|X|}(u)$ is non-negative for the random variable X, and $0 \leq \alpha \leq 1$, then the inequalities in Lemma 5.4 and Lemma 5.5 can also be bounded by a higher upper limitation of the integration, i.e.

$$\int_0^{\alpha} Q_{|X|}(u)Q_{|Y|}(u)du \leq \int_0^{1} Q_{|X|}(u)Q_{|Y|}(u)du. \tag{5.2}$$

Now the following lemma can be directly applied to the case of the multiplication of random variables if it occurs in Lemma 5.4 and Lemma 5.5.

Lemma 5.6. (Lemma 2.1 (b) in [57]) *Let $X_1, \cdots, X_p$ be random variables. Then*

$$\int_0^1 Q_{|X_1X_2|}(u)Q_{|X_3|}(u)\cdots Q_{|X_p|}(u)du \leq \int_0^1 Q_{|X_1|}(u)Q_{|X_2|}(u)Q_{|X_3|}(u)\cdots Q_{|X_p|}(u)du.$$

Lemma 5.7. *For any $\alpha \in [0, \frac{1}{4}]$ and any random variables, $X_1, X_2, \cdots, X_p$, we have*

$$\int_0^{\alpha} Q_{|X_1X_2|}(u)Q_{|X_3|}(u)\cdots Q_{|X_p|}(u)du \leq \int_0^{\alpha} Q_{|X_1|}(u)Q_{|X_2|}(u)Q_{|X_3|}(u)\cdots Q_{|X_p|}(u)du.$$

Proof. For any α, $\alpha \in [0, \frac{1}{4}]$, we define a random variable,

$$Z = \begin{cases} 0, & \text{with the probability } 1-\alpha, \\ 1, & \text{with the probability } \alpha. \end{cases}$$

Then the tail function of Z is

$$Q_{|Z|}(u) = \inf\{z \geq 0 : \mathbb{P}(|Z| > z) \leq u\} = \begin{cases} 1, & \text{if } u \leq \alpha, \\ 0, & \text{if } u > \alpha. \end{cases}$$

Now, for any random variables, X_1 and X_2, we have

$$\begin{aligned} & \int_0^\alpha Q_{|X_1X_2|}(u)Q_{|X_3|}(u)\cdots Q_{|X_p|}(u)du \\ = & \int_0^1 Q_{|Z|}(u)Q_{|X_1X_2|}(u)Q_{|X_3|}(u)\cdots Q_{|X_p|}(u)du \\ \leq & \int_0^1 Q_{|Z|}(u)Q_{|X_1|}(u)Q_{|X_2|}(u)Q_{|X_3|}(u)\cdots Q_{|X_p|}(u)du \quad \text{(By using Lemma 5.6)} \\ = & \int_0^\alpha Q_{|X_1|}(u)Q_{|X_2|}(u)Q_{|X_3|}(u)\cdots Q_{|X_p|}(u)du. \end{aligned}$$

This completes the proof. □

For the strong mixing coefficient in Definition 2.1, $\alpha_{k,l}(m)$, we define the inverse function of $\alpha_{k,l}(m)$ as the supremum of the distance m, i.e. $\alpha_{k,l}^{-1}(u) = \sup\{m \in \mathbb{N} : \alpha_{k,l}(m) > u\}$, which is the same as

$$\alpha_{k,l}^{-1}(u) = \inf\{m \in \mathbb{N} : \alpha_{k,l}(m) \leq u\}.$$

We note that

$$\sum_{m=0}^{\infty}(m+1)\mathbb{1}_{u\in[\alpha_{k,l}(m+1),\alpha_{k,l}(m))} = \alpha_{k,l}^{-1}(u). \tag{5.3}$$

Therefore, for any constant $K > 0$, we have

$$\sum_{m=0}^{\infty}(m+1)^K\mathbb{1}_{u\in[\alpha_{k,l}(m+1),\alpha_{k,l}(m))} = \left[\alpha_{k,l}^{-1}(u)\right]^K. \tag{5.4}$$

Theorem 5.8. *Let $X_{i\in\mathbb{Z}^d}$ be a zero-mean random field with strong mixing coefficient $\alpha_{1,1}(m)$, which is defined by Definition 2.1. For any subset of $\mathbb{Z}^d$, $D \subseteq \mathbb{Z}^d$, we assume there exists a constant $K_1 > 0$ such that*

$$|D| \leq K_1(m+1)^d, \tag{5.5}$$

where $m = \sup\{d(i,j) : i,j \in D\}$. Then there exists a constant C such that

$$\mathbb{E}\left(\sum_{i\in D} X_i\right)^2 \leq C\sum_{i\in D}\int_0^{\frac{1}{4}}\left[\alpha_{1,1}^{-1}(u)\right]^d Q_{X_i}^2(u)du.$$

Proof. We note that (5.9) implies

$$\#\{j : d(i,j) \leq m\} \leq K_1(m+1)^d.$$

Therefore, by using Lemma 5.4, (5.3) and (5.4),

$$\begin{aligned}
\mathbb{E}\left(\sum_{i\in D} X_i\right)^2 &\leq \sum_{i,j\in D} |Cov(X_i, X_j)| \\
&\leq 2\sum_{i,j\in D}\int_0^{\alpha_{1,1}(d(i,j))} \left(Q^2_{X_i}(u) + Q^2_{X_j}(u)\right) du \\
&= 4\sum_{i\in D}\sum_{m=0}^{\infty}\int_0^{\alpha_{1,1}(m)} \#\{j : d(i,j) = m\} Q^2_{X_i}(u) du \\
&\leq 4\sum_{i\in D}\sum_{m=0}^{\infty}\int_{\alpha_{1,1}(m+1)}^{\alpha_{1,1}(m)} \#\{j : d(i,j) \leq m\} Q^2_{X_i}(u) du \\
&\leq 4K_1\sum_{i\in D}\sum_{m=0}^{\infty}\int_{\alpha_{1,1}(m+1)}^{\alpha_{1,1}(m)} (m+1)^d Q^2_{X_i}(u) du \\
&= 4K_1\sum_{i\in D}\int_0^{\frac{1}{4}}\sum_{m=0}^{\infty} \mathbb{1}_{u\in[\alpha_{1,1}(m+1),\alpha_{1,1}(m))}(m+1)^d Q^2_{X_i}(u) du \\
&= 4K_1\sum_{i\in D}\int_0^{\frac{1}{4}} \left[\alpha_{1,1}^{-1}(u)\right]^d Q^2_{X_i}(u) du.
\end{aligned}$$

This completes the proof. □

Let $\{S_1, S_2\}$ be a partition of distinct $\{i, j, k, l\}$, $i, j, k, l \in \mathbb{Z}^d$. Then S_1 and S_2 are two non-empty sets. We define

$$M_{ijkl} = \max\Big\{dist(S_1, S_2) : \{S_1, S_2\} \in \mathcal{P}_{ijkl}\Big\}, \tag{5.6}$$

where $\mathcal{P}_{ijkl}$ stands for the set of all partitions of $\{i, j, k, l\}$, i.e.

$$\begin{aligned}
\mathcal{P}_{ijkl} = \Big\{&\{\{i,j\},\{k,l\}\}, \quad \{\{i,k\},\{j,l\}\}, \quad \{\{i,l\},\{k,j\}\}, \\
&\{\{i,\},\{j,k,l\}\}, \quad \{\{j,\},\{i,k,l\}\}, \quad \{\{k,\},\{j,i,l\}\}, \quad \{\{l,\},\{j,k,i\}\}\Big\}
\end{aligned}$$

Lemma 5.9. *Let M_{ijkl} be defined as in (5.6), but with $\{i, j, k, l\}$ not necessary distinct. Then j, k, l must be within $3M_{ijkl}$ of i, i.e.*

$$dist(i, j) \leq 3M_{ijkl}, \quad dist(i, k) \leq 3M_{ijkl}, \quad dist(i, l) \leq 3M_{ijkl}. \tag{5.7}$$

Proof. We suppose the contrary of (5.7).

1) Firstly, we suppose $\{i, j, k, l\}$ are distinct, then there are several cases to consider.

1.1) All of j, k, l are more than $3M_{ijkl}$ from i. But this is impossible as the split $S_1 = \{i\}$ and $S_2 = \{j, k, l\}$ would have a separation greater than M_{ijkl}.

1.2) Two of j,k,l are more than $3M_{ijkl}$ from i, say k,l. Then the separation between i and k,l is more than M_{ijkl}. Therefore we must have $dist(i,j) \leq M_{ijkl}$, or else the separation $S_1 = \{i\}$ and $S_2 = \{j,k,l\}$ would be more than M_{ijkl}. Hence

$$3M_{ijkl} \leq d(i,k) \leq d(i,j) + d(j,k) \leq M_{ijkl} + dist(j,k),$$

i.e. $d(j,k) \geq 2M_{ijkl}$. By the same argument, we have $d(j,l) \geq 2M_{ijkl}$. Thus, the seperation between $S_1 = \{i,j\}$ and $S_2 = \{k,l\}$ is more than M_{ijkl}. This contradicts the definition of M_{ijkl}.

1.3) One of j,k,l is more than $3M_{ijkl}$ from i, say l. For example, if $d(i,l) > 3M_{ijkl}$, $d(i,j) < M_{ijkl}$ and $d(i,k) < M_{ijkl}$, then by the same argument, we have $d(i,l) > 2M_{ijkl}$ and $d(k,l) > 2M_{ijkl}$. This contradicts to the definition of M_{ijkl}.

Thus all these alternatives are impossible, so that j,k,l are within $3M_{ijkl}$ of i.

2) Secondly, we suppose $\{i,j,k,l\}$ are distinct. Then there are several cases to consider.

2.1) If only three of i,j,k,l are distinct, say $i = j$, then in (5.6), we can interpret S_1 and S_2 as the partition of $\{i,k,l\}$. In this case, with a similar argument, we must have the distance between i and the other two distinct points is less than or equal to $2M_{ijkl}$.

2.2) If they are equal in pairs, then we must have the distance between the two distinct points equal to M_{ijkl}.

2.3) If they are all equal, then it becomes trivial.

So, in every case, we must have the conclusion of the lemma. □

To simplify the expression in the following theorem, we combine $\alpha_{2,2}^{-1}(u)$ and $\alpha_{1,3}^{-1}(u)$ by defining

$$\alpha_4^{-1}(u) = \max\{\alpha_{2,2}^{-1}(u), \alpha_{1,3}^{-1}(u)\}. \tag{5.8}$$

Theorem 5.10. *Let $X_{i\in\mathbb{Z}^d}$ be a zero-mean random field with strong mixing coefficient $\alpha_{k,l}(m)$, which is defined by Definition 2.1, $\alpha_4^{-1}(u)$ be defined as in (5.8), and $D \subseteq \mathbb{Z}^d$. We assume there exists a constant $K_1 > 0$ such that*

$$|D| \leq K_1(m+1)^d, \tag{5.9}$$

where $m = \sup\{dist(i,j) : i,j \in D\}$. Then there exist a constant $K > 0$ such that

$$\mathbb{E}\left(\sum_{i\in D} X_i\right)^4 \leq 3\left(\sum_{i,j\in D} |\mathbb{E}(X_iX_j)|\right)^2 + K\sum_{i\in D}\int_0^{\frac{1}{4}} \left[\alpha_4^{-1}(u)\right]^{3d} Q_{X_i}^4(u)du.$$

Proof. We write

$$\mathbb{E}\left(\sum_{i\in D} X_i\right)^4 = \sum_{i,j,k,l\in D} \mathbb{E}(X_iX_jX_kX_l).$$

Consider partitioning $\{i,j,k,l\}$ into two non-empty subsets, S_1 and S_2. Let M_{ijkl} be defined as in (5.6).

Case 1: Suppose the maximum distance M_{ijkl} occurs at the partition $\{S_1, S_2\}$, where S_1 contains two elements, and S_2 contains the other two. Without loss of generality, we set $S_1 = \{i,j\}$ and $S_2 = \{k,l\}$. Then, by using the fact of $abcd \le \frac{1}{4}(a^4+b^4+c^4+d^4)$, where $a,b,c,d \in \mathbb{R}$, Lemma 5.4 and Lemma 5.7 imply

$$\begin{aligned}
\mathbb{E}(X_iX_jX_kX_l) &\le |\mathbb{E}(X_iX_j)||\mathbb{E}(X_kX_l)|+|Cov(X_iX_j, X_kX_l)| \\
&\le |\mathbb{E}(X_iX_j)||\mathbb{E}(X_kX_l)| \\
&\qquad + \int_0^{\alpha_{2,2}(M_{ij,kl})} \left(Q^4_{X_i}(u) + Q^4_{X_j}(u) + Q^4_{X_k}(u) + Q^4_{X_l}(u)\right) du.
\end{aligned}$$

Case 2: If M_{ijkl} occurs at the partition, where S_1 contains one element, and S_2 contains the other three. Without losing of generality, we set $S_1 = \{i\}$ and $S_2 = \{j,k,l\}$. Then, similarly, we have

$$\begin{aligned}
\mathbb{E}(X_iX_jX_kX_l) &= |Cov(X_i, X_jX_kX_l)| \\
&\le \int_0^{\alpha_{1,3}(M_{i,jkl})} \left(Q^4_{X_i}(u) + Q^4_{X_j}(u) + Q^4_{X_k}(u) + Q^4_{X_l}(u)\right) du.
\end{aligned}$$

The above two estimations of $\mathbb{E}(X_iX_jX_kX_l)$ also work for the cases where i,j,k,l are not distinct.

We note that

$$\begin{aligned}
\sum_{i,j,k,l\in D} \int_0^{\alpha_{2,2}(M_{il,jk})} Q^4_{X_i}(u)du &= \sum_{i\in D}\sum_{j,k,l\in D} \int_0^{\alpha_{2,2}(M_{il,jk})} Q^4_{X_i}(u)du \\
&= \sum_{i\in D}\sum_{m=0}^{\infty} \int_0^{\alpha_{2,2}(m)} \#\{(j,k,l) : M_{ijkl} = m\} Q^4_{X_i}(u)du \\
&\le \sum_{i\in D}\sum_{m=0}^{\infty} \int_{\alpha_{2,2}(m+1)}^{\alpha_{2,2}(m)} \#\{(j,k,l) : M_{ijkl} \le m\} Q^4_{X_i}(u)du.
\end{aligned}$$

Lemma 5.9 and the assumption (5.9) imply $\#\{(j,k,l) : M_{ijkl} \le m\} \le K(m+1)^{3d}$, where K is a constant. Then by using (5.3) and (5.4),

$$\sum_{i,j,k,l\in D} \int_0^{\alpha_{2,2}(M_{il,jk})} Q^4_{X_i}(u)du \le \sum_{i\in D}\sum_{m=0}^{\infty} \int_{\alpha_{2,2}(m+1)}^{\alpha_{2,2}(m)} K(m+1)^{3d} Q^4_{X_i}(u)du$$

$$
\begin{aligned}
&= \sum_{i\in D}\int_0^{\frac{1}{4}}\sum_{m=0}^{\infty}\mathbb{1}_{u\in[\alpha_{1,1}(m+1),\alpha_{1,1}(m))}K(m+1)^{3d}Q^4_{X_i}(u)du \\
&= \sum_{i\in D}\int_0^{\frac{1}{4}}\left[\alpha_{2,2}^{-1}(u)\right]^{3d}Q^4_{X_i}(u)du \\
&\leq \sum_{i\in D}\int_0^{\frac{1}{4}}\left[\alpha_4^{-1}(u)\right]^{3d}Q^4_{X_i}(u)du.
\end{aligned}
$$

Similarly, we have

$$
\sum_{i,j,k,l\in D}\int_0^{\alpha_{1,3}(M_{i,ljk})}Q^4_{X_i}(u)du \leq \sum_{i\in D}\int_0^{\frac{1}{4}}\left[\alpha_4^{-1}(u)\right]^{3d}Q^4_{X_i}(u)du.
$$

Therefore the above discussion implies the upper bound of this fourth moment is not greater than the sum of these two possible cases. It leads to

$$
\begin{aligned}
\sum_{i,j,k,l\in D}\mathbb{E}(X_iX_jX_kX_l) &\leq 3\left(\sum_{i,j\in D}|\mathbb{E}(X_iX_j)|\right)^2+3\sum_{i,j,k,l\in D}|Cov(X_iX_j,X_kX_l)| \\
&\quad +4\sum_{i,j,k,l\in D}|Cov(X_i,X_jX_kX_l)| \\
&\leq 3\left(\sum_{i,j\in D}|\mathbb{E}(X_iX_j)|\right)^2 \\
&\quad +3\sum_{i,j,k,l\in D}\int_0^{\alpha_{2,2}(M_{ij,kl})}\left(Q^4_{X_i}(u)+Q^4_{X_j}(u)+Q^4_{X_k}(u)+Q^4_{X_l}(u)\right)du \\
&\quad +4\sum_{i,j,k,l\in D}\int_0^{\alpha_{1,3}(M_{i,jkl})}\left(Q^4_{X_i}(u)+Q^4_{X_j}(u)+Q^4_{X_k}(u)+Q^4_{X_l}(u)\right)du \\
&\leq 3\left(\sum_{i,j\in D}|\mathbb{E}(X_iX_j)|\right)^2+28K\sum_{i\in D}\int_0^{\frac{1}{4}}[\alpha_4^{-1}(u)]^{3d}Q^4_{X_i}(u)du.
\end{aligned}
$$

This completes the proof. □

In Theorem 5.11 and Theorem 5.13, we introduce an assumption on the nested spatial structure, which is illustrated by Figure 5.1. We set D_0 is the index set of the original sample. Let $D_0 \subseteq D_1 \subseteq \cdots \subseteq D_{n-1} \subseteq D_n \subseteq \cdots$.

For example, in Figure 5.1, we assume n_1 and n_2 are the subscripts of the index set of the sample, $n_1 < n_2$. The index set D_{n_1} is circled by the black dash line, and D_{n_2} is circled by the black solid line, $D_{n_1} \subseteq D_{n_2}$. We also use $D^c_{n_2}$ to stand for the outside of the region D_{n_2}, i.e.

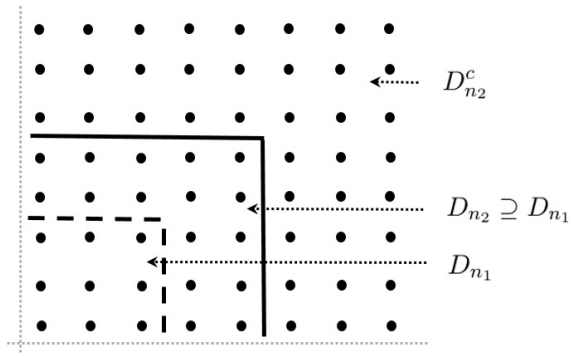


Figure 5.1: The nested spatial structure.

the indices which are not in D_{n_2}. Then $dist(D_{n_2}^c, D_{n_1})$ is used to size the distance of the gap between the dash line and the solid line, i.e. the distance between the region of $D_{n_2}^c$ and the region of D_{n_1}.

Theorem 5.11. *Let the zero-mean random field $X_{i\in\mathbb{Z}^d}$ satisfies (H1), (H2), (H3) and (H4) in Theorem 3.1. We assume*

1) $D_n \subset \mathbb{Z}^d$ is with a nested spatial structure, i.e. $D_0 \subseteq D_1 \subseteq \cdots \subseteq D_{n-1} \subseteq D_n \subseteq \cdots$. There exists constants $K_0 > 0$ and $K_1 > 0$, such that for any $n_1 \leq n_2$,

$$dist(D_{n_2}^c, D_{n_1}) \geq K_1(n_2 - n_1)^{K_0};$$

2) There exists a constant $K_3 > 0$ such that (5.9) being satisfied, i.e. for any $D \subseteq \mathbb{Z}^d$,

$$|D| \leq K_3(m+1)^d,$$

where $m = \sup\{dist(i,j) : i, j \in D\}$,

$$\sup_{i\in\mathbb{Z}^d} \int_0^{\frac{1}{4}} \left([\alpha_{1,1}^X]^{-1}(u)\right)^{3d} Q_{X_i}^4(u)du < \infty,$$

and

$$\sup_{i\in\mathbb{Z}^d} \int_0^{\frac{1}{4}} \left([\alpha_4^X]^{-1}(u)\right)^{3d} Q_{X_i}^4(u)du < \infty,$$

where $[\alpha_4^X]^{-1}(u)$ is defined as in (5.8);

3) There exists a non-decreasing continuous function $F(t)$, $t \in [0,1]$, and a constant $K > 0$, such that

$$\frac{|D_{[nt_2]} - D_{[nt_1]}|}{|D_n|} \leq K[F(t_2) - F(t_1)], \quad 0 \leq t_1 \leq t_2 \leq 1.$$

Let K_2 be a constant and

$$W_n(t) = \frac{S_{[nt]}}{\sigma_n \sqrt{K_2}}. \tag{5.10}$$

Then we have

$$W_n(t) \xrightarrow{\mathcal{D}} W(t),$$

where $W(t)$ is a one dimensional Brownian motion on $[0,1]$.

Proof. We need to check the three conditions in Theorem 5.3.

Firstly, for finite $k \in \mathbb{Z}^+$, $0 = t_0 < t_1 < \cdots < t_{k-1} < t_k = 1$, and for sufficiently large n, we set

$$\epsilon := n^{-\frac{1}{2}} < \min_{i=1\cdots,k} \{t_i - t_{i-1}\}. \tag{5.11}$$

We are going to prove

$$(W_n(t_i) - W_n(t_{i-1}))_{1\le i\le k} \xrightarrow{D} (W(t_i) - W(t_{i-1}))_{1\le i\le k}. \tag{5.12}$$

As in the proof of Theorem 4.3 in [57], we consider characteristic functions. Let $s = (s_1, s_2, \cdots, s_k) \in \mathbb{R}^k$, and

$$\varphi_n(s) = \mathbb{E}\exp\left(i\sum_{j=1}^{k} s_j(W_n(t_j) - W_n(t_{j-1}))\right),$$

$$\varphi_{n,\epsilon}(s) = \mathbb{E}\exp\left(i\sum_{j=1}^{k} s_j(W_n(t_j - \epsilon) - W_n(t_{j-1}))\right).$$

Since $W(t)$ is a Brownian motion, $(W(t_i) - W(t_{i-1}))_{1\le i\le k}$ are independent. Therefore, to prove (5.12), it is sufficient to prove

$$\left|\varphi_n(s) - \prod_{j=1}^{k} \mathbb{E}\exp\Big(is_j(W(t_j) - W(t_{j-1}))\Big)\right| \to 0.$$

We write

$$\begin{aligned}
&\left|\varphi_n(s) - \prod_{j=1}^{k} \mathbb{E}\exp\Big(is_j(W(t_j) - W(t_{j-1}))\Big)\right| \\
\le\ & \left|\varphi_n(s) - \varphi_{n,\epsilon}(s)\right| + \left|\varphi_{n,\epsilon}(s) - \prod_{j=1}^{k} \mathbb{E}\exp\Big(is_j(W_n(t_j - \epsilon) - W_n(t_{j-1}))\Big)\right| \\
+\ & \left|\prod_{j=1}^{k} \mathbb{E}\exp\Big(is_j(W_n(t_j - \epsilon) - W_n(t_{j-1}))\Big) - \prod_{j=1}^{k} \mathbb{E}\exp\Big(is_j(W(t_j - \epsilon) - W(t_{j-1}))\Big)\right|
\end{aligned}$$

$$
\begin{aligned}
&+ \quad \left| \prod_{j=1}^{k} \mathbb{E} \exp\Big(is_j (W(t_j - \epsilon) - W(t_{j-1})) \Big) - \prod_{j=1}^{k} \mathbb{E} \exp\Big(is_j (W(t_j) - W(t_{j-1})) \Big) \right| \\
&= \quad \text{(I)+(II)+(III)+(IV)}.
\end{aligned}
$$

For term (IV), by using the continuity property of the characteristic function of the normal distribution, we have

$$
\text{(IV)} \to 0 \quad \text{as} \quad n \to \infty.
$$

For term (I), by using the 1-Lipschitz property of the function of the real valued y, $y \mapsto e^{iy}$, we have, for all real valued x and y, $|e^{ix} - e^{iy}| \leq |x - y|$. We note that

$$
\begin{aligned}
& \left| \sum_{j=1}^{k} s_j \Big(W_n(t_j) - W_n(t_{j-1}) \Big) - \sum_{j=1}^{k} s_j \Big(W_n(t_j - \epsilon) - W_n(t_{j-1}) \Big) \right| \\
= \quad & \left| \sum_{j=1}^{k} s_j \Big(W_n(t_j) - W_n(t_j - \epsilon) \Big) \right| \\
\leq \quad & \|s\| \left(\sum_{j=1}^{k} \Big(W_n(t_j) - W_n(t_j - \epsilon) \Big)^2 \right)^{\frac{1}{2}}.
\end{aligned}
$$

Then we have

$$
\begin{aligned}
& \Big| \varphi_n(s) - \varphi_{n,\epsilon}(s) \Big| \\
\leq \quad & \|s\| \left(\sum_{j=1}^{k} \mathbb{E} \Big(W_n(t_j) - W_n(t_j - \epsilon) \Big)^2 \right)^{\frac{1}{2}}.
\end{aligned}
$$

Because (H4) implies $\sigma_n^2 \geq |D_n| C$, where C is a constant, we have

$$
\begin{aligned}
\mathbb{E} \Big(W_n(t_j) - W_n(t_j - \epsilon) \Big)^2 \quad &\leq \quad \frac{1}{\sigma_n^2 K_2} \sum_{i,j \in D_{[nt_j]} - D_{[n(t_j - \epsilon)]}} |Cov(X_i, X_j)| \\
&\leq \quad \frac{1}{|D_n| C K_2} |D_{[nt_j]} - D_{n[t_j - \epsilon]}| \sum_{m=1}^{\infty} m^{d-1} \alpha_{1,1}^{\frac{\delta}{2+\delta}}(m) \|X\|_{2+\delta}^2,
\end{aligned}
$$

and by using (H3) and assumption 3, we have

$$
\mathbb{E} \Big(W_n(t_j) - W_n(t_j - \epsilon) \Big)^2 \quad \leq \quad C \Big(F(t_j) - F(t_j - \epsilon) \Big).
$$

Therefore, for all s, we have $|\varphi_n(s) - \varphi_{n,\epsilon}(s)| \to 0$ as $n \to \infty$.

For (II), we note that the distance between $W_n(t_{j+1} - \epsilon) - W_n(t_j)$ and $W_n(t_j - \epsilon) - W_n(t_{j-1})$ is the same as the distance between $D_{[n(t_{j+1} - \epsilon)]} - D_{[nt_j]}$ and $D_{[n(t_j - \epsilon)]} - D_{[nt_{j-1}]}$. In Figure 5.2, the gap

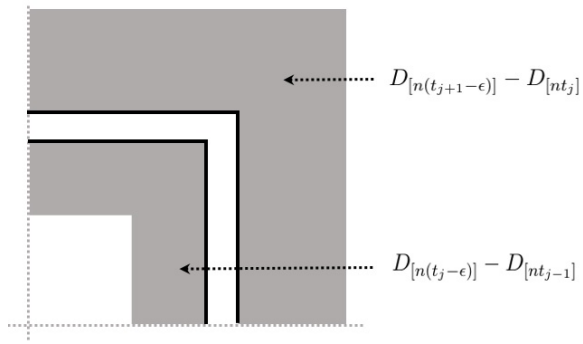


Figure 5.2: The gap between $D_{[n(t_{j+1}-\epsilon)]} - D_{[nt_j]}$ and $D_{[n(t_j-\epsilon)]} - D_{[nt_{j-1}]}$.

between two solid lines implies the distance between two grey regions. By using the definition of $D^c_{[nt_j]}$, this distance can be described by $dist(D^c_{[nt_j]}, D_{[n(t_j-\epsilon)]})$. Then the assumption 1 implies

$$\begin{aligned} dist\Big(D_{[n(t_{j+1}-\epsilon)]} - D_{[nt_j]}, \quad D_{[n(t_j-\epsilon)]} - D_{[nt_{j-1}]}\Big) &= dist\Big(D^c_{[nt_j]}, D_{[n(t_j-\epsilon)]}\Big) \\ &\geq K_1\Big([nt_j] - [n(t_j-\epsilon)]\Big)^{K_0} \\ &\geq C_2\Big([n\epsilon]\Big)^{K_0}, \end{aligned}$$

where C_2 is a constant.

We note that, if $Y_j = \exp iT_j$, where T_j, $j = 1, 2, \cdots, k$ are real valued random variables, then we have

$$|\mathbb{E}(Y_1Y_2\cdots Y_k) - \mathbb{E}Y_1\mathbb{E}Y_2\cdots\mathbb{E}Y_k| \leq |Cov(Y_1, Y_2\cdots Y_k)| + |Cov(Y_2, Y_3\cdots Y_k)| + \cdots |Cov(Y_{k-1}, Y_k)|,$$

$$|Y_c| = 1,$$

and

$$\int_0^\alpha Q_{|Y_c|}(u)du = \alpha,$$

where Y_c is the multiplication of a combination of any one subset of $\{Y_1, Y_2, \cdots, Y_k\}$.

Now, let

$$Y_j = is_j(W_n(t_j-\epsilon) - W_n(t_{j-1})),$$
$$d_j = dist\Big(\{Y_j\}, \{Y_{j+1}, Y_{j+2}, \cdots, Y_k\}\Big).$$

By using Lemma 5.5, repeatedly $(k-1)$ times,

$$\left|\varphi_{n,\epsilon}(s) - \prod_{j=1}^{k} \mathbb{E}\exp\left(is_j(W_n(t_j-\epsilon) - W_n(t_{j-1}))\right)\right|$$

$$
\begin{aligned}
&\leq 16\sum_{j=1}^{k-1}\int_0^{\alpha_{\infty,\infty}(d_j)} Q_{|Y_j|}(u)Q_{|Y_{j+1}Y_{j+2}\cdots Y_k|}(u)du \\
&= 16\sum_{j=1}^{k-1}\alpha_{\infty,\infty}(d_j).
\end{aligned}
$$

By using the non-increasing property of strong mixing coefficients with respect to the distance, since assumption 1 implies

$$\min\{d_j, j=1,2,\cdots,k\} \geq K_1([n\epsilon])^{K_0},$$

then we have

$$\left|\varphi_{n,\epsilon}(s) - \prod_{j=1}^{k}\mathbb{E}\exp\left(is_j(W_n(t_j-\epsilon)-W_n(t_{j-1}))\right)\right| \leq 16(k-1)\alpha_{\infty,\infty}\Big(K_1([n\epsilon])^{K_0}\Big),$$

which vanishes as $n\to\infty$.

For term (III), it is sufficient to prove

$$W_n(t_2)-W_n(t_1) \xrightarrow{D} W(t_2)-W(t_1).$$

Since Theorem 3.1 implies $W_n(t)\to W(t)$ for any fixed $t\in[0,1]$, by referring the proof of Lemma 5.1 b, to prove (III), we only need to show that $W_n(t_2)-W_n(t_1)$ and $W_n(t_1)$ are asymptotically independent, i.e. that

$$|\varphi_1(s)-\varphi_2(s)\varphi_3(s)|\to 0,$$

where $\varphi_1(s)$, $\varphi_2(s)$ and $\varphi_3(s)$ are characteristic functions of $W_n(t_2)$, $W_n(t_2)-W_n(t_1)$ and $W_n(t_1)$ respectively. Let $\varphi_*(s)$ be the characteristic function of $W_n(t_2)-W_n(t_1)+W_n(t_1-\epsilon)$, and $\varphi_\epsilon(s)$ be the characteristic function of $W_n(t_1-\epsilon)$. Then we have

$$
\begin{aligned}
|\varphi_1(s)-\varphi_2(s)\varphi_3(s)| &\leq |\varphi_1(s)-\varphi_*(s)|+|\varphi_*(s)-\varphi_2(s)\varphi_\epsilon(s)|+|\varphi_2(s)\varphi_\epsilon(s))-\varphi_2(s)\varphi_3(s)| \\
&= \text{(i)}+\text{(ii)}+\text{(iii)}.
\end{aligned}
$$

For (i) and (iii), we follow the similar arguments in (I). Therefore, we have

$$\text{(i)} \leq \mathbb{E}s|W_n(t_1)-W_n(t_1-\epsilon)|\to 0$$

and

$$\text{(iii)} \leq |\varphi_\epsilon(s))-\varphi_3(s)|\to 0.$$

For (ii), since the distance between $e^{is(W_n(t_2)-W_n(t_1))}$ and $e^{is(W_n(t_1-\epsilon))}$ is just the distance between

$W_n(t_2) - W_n(t_1)$ and $W_n(t_1 - \epsilon)$, we have

$$\begin{aligned}
\text{(ii)} &= Cov\Big(e^{is(W_n(t_2)-W_n(t_1))}, e^{is(W_n(t_1-\epsilon))}\Big) \\
&\leq 8\alpha_{\infty,\infty}^{\frac{\delta}{2+\delta}}\Big(K_1([n\epsilon])^{K_0}\Big) \to 0 \quad \text{as} \quad n \to \infty.
\end{aligned}$$

Hence, the first condition of Theorem 5.3 is satisfied.

Secondly, since $W(t)$ is a Brownian motion, the second condition of Theorem 5.3 is satisfied.

Thirdly, we check the restricted version of the third condition of Theorem 5.3 which is provided on page 128 of Billingsley's book [4]. We set $\gamma = 2$, $\alpha = 1$, then we have

$$\begin{aligned}
& \mathbb{E}(|W_n(t) - W_n(t_1)|^2 |W_n(t_2) - W_n(t)|^2) \\
= \; & \frac{1}{\sigma^4 K_2} \mathbb{E}\left[\left(\sum_{i \in D_{[nt]} - D_{[nt_1]}} X_i\right)^2 \left(\sum_{i \in D_{[nt_2]} - D_{[nt]}} X_i\right)^2\right] \\
\leq \; & \frac{1}{\sigma^4 K_2} \left[\mathbb{E}\left(\sum_{i \in D_{[nt]} - D_{[nt_1]}} X_i\right)^4\right]^{\frac{1}{2}} \left[\mathbb{E}\left(\sum_{i \in D_{[nt_2]} - D_{[nt]}} X_i\right)^4\right]^{\frac{1}{2}}
\end{aligned}$$

Theorem 5.10, the assumption 1, the assumption 2, (H3) and the assumption 3 imply that a constant C exists such that

$$\begin{aligned}
& \mathbb{E}(|W_n(t) - W_n(t_1)|^2 |W_n(t_2) - W_n(t)|^2) \\
\leq \; & \frac{C}{|D_n|^2} |D_{[nt]} - D_{[nt_1]}| |D_{[nt_2]} - D_{[nt]}| \\
\leq \; & [F(t_2) - F(t_1)]^2.
\end{aligned}$$

Therefore the third condition of Theorem 5.3 is satisfied. This completes the proof. □

Remark: For the assumption 2 in Theorem 5.11, the application of Corollary 3.1 in [57] has shown that, with some proper conditions, it holds for a strictly stationary and strong mixing sequence. For the assumption 3, if we define a nested sampling method which satisfies $|D_n| = n^2$ for a 2-dimensional random field, then $|D_{[nt_2]}| - |D_{[nt_1]}| = [nt_2]^2 - [nt_1]^2$. Now we only need to show

$$\frac{|D_{[nt_2]} - D_{[nt]}||D_{[nt]} - D_{[nt_1]}|}{|D_n|^2} \leq 4(t_2^2 - t_1^2)^2.$$

As in the proof of Theorem 16.1 in Billingsley's book [4] page 138, we prove this with two cases. Case 1: If $nt_2 - nt_1 \leq 1$, then one of $|D_{[nt_2]} - D_{[nt]}|$ and $|D_{[nt]} - D_{[nt_1]}|$ must be zero since $t_1 \leq t \leq t_2$. The above inequality holds. Case 2: If $nt_2 - nt_1 > 1$, we note that

$[nt_2] - [nt_1] \leq 2(nt_2 - nt_1)$ and

$$[nt_2]^2 - [nt_1]^2 = ([nt_2] - [nt_1])([nt_2] + [nt_1]) \leq 2(nt_2)^2 - 2(nt_1)^2.$$

Therefore the above inequality holds. It means we can also replace the assumption 3 with a weaker assumption such as

$$\frac{|D_{[nt_2]} - D_{[nt]}||D_{[nt]} - D_{[nt_1]}|}{|D_n|^2} \leq K[F(t_2) - F(t_1)]^2, \quad 0 \leq t_1 \leq t \leq t_2 \leq 1.$$

Furthermore, we note that

$$\frac{|D_{[nt_2]} - D_{[nt]}||D_{[nt]} - D_{[nt_1]}|}{|D_n|^2} \leq \frac{|D_{[nt_2]} - D_{[nt_1]}|^2}{|D_n|^2}.$$

This is a stronger condition which we used in Theorem 5.11, i.e.

$$\frac{|D_{[nt_2]} - D_{[nt_1]}|^2}{|D_n|^2} \leq K[F(t_2) - F(t_1)]^2,$$

where $F(t)$ can be $F(t) = t^2$ for the present example of the 2-dimensional random field. If the sample size can be controlled by its index, say $|D_n| = |D_0| + n$, then according to the discussion in the proof of Theorem 16.1 in [4] by Billingsley, $F(t) = t$ satisfies the weaker assumption we mentioned above.

In Theorem 5.13, we are going to relax the assumption 2 of Theorem 5.11. As a trade-off for this relaxing, we have to introduce the following mixing coefficient.

Definition 5.2. An *uniform mixing coefficient* between two blocks within a random field, $X_{i\in\mathbb{Z}^d}$, is defined by a function with respect to the distance between the blocks, and the size of the blocks, specifically,

$$\begin{aligned}\varphi_{k,l}(m) \quad = \quad \sup_{A,B} \Big\{ &|\mathbb{P}(B|A) - \mathbb{P}(B)|\colon A \in \sigma(X_{i\in\Lambda_1}), B \in \sigma(X_{i\in\Lambda_2}), \\ &\mathbb{P}(A) \neq 0, k = |\Lambda_1|, l = |\Lambda_2|, dist(\Lambda_1, \Lambda_2) \geq m \Big\}.\end{aligned}$$

Analogously to the strong mixing coefficient, sometimes, we use $\varphi_{|\Lambda_1|,|\Lambda_2|}(m)$, $\varphi^X_{k,l}(m)$, or $\varphi_{k,l}(X;m)$, to stand for the uniform mixing coefficient of the random field $X_{i\in\mathbb{Z}^d}$.

To check the tightness of the random fields in the proof of Theorem 5.13, we introduce the following theorem, which is similar to Theorem 8.4 in [4] but which works for non-stationary random fields.

Theorem 5.12. *Let D_n be with a nested spatial structure, $|D_n|$ be a non-decreasing function*

with respect to n,

$$W_n(t) = \frac{S_{[nt]}}{\sigma_n \sqrt{k}}$$

be defined as in (5.10). The sequence $\{W_n\}$ is tight if (H4) is satisfied and for each $\epsilon > 0$ there exists a $\lambda > 1$, and an integer $n_0 \in \mathbb{Z}^+$ such that for all $n \geq n_0$

$$\sup_m \mathbb{P}\left\{\max_{i \leq n} |S_{m+i} - S_m| \geq \lambda \sqrt{|D_n|}\sqrt{k}\right\} \leq \frac{\epsilon}{\lambda^2}. \tag{5.13}$$

Proof. Since (H4) implies there exists a constant $C \geq 1$ such that

$$\frac{\sqrt{|D_n|}}{\sigma_n} \leq C,$$

then for any random variable X, we have

$$\mathbb{P}\{X \geq \sigma_n\} \leq \mathbb{P}\{X \geq \frac{\sqrt{|D_n|}}{C}\}.$$

Therefore, according to Theorem 8.3 in [4], to prove W_n is tight, it is sufficient to show that, for any $0 < \epsilon < 1$ and $0 < \eta < 1$, there exits a δ, $0 < \delta < 1$, and an integer $n_0 > 0$, such that for all $n \geq n_0$ we have

$$\sup_m \mathbb{P}\left\{\max_{i \leq n} |S_{m+i} - S_m| \geq \epsilon \sigma_n \sqrt{k}\right\} \leq \eta\delta,$$

which is implied by

$$\sup_m \mathbb{P}\left\{\max_{i \leq n} |S_{m+i} - S_m| \geq \epsilon \frac{\sqrt{|D_n|}}{C}\sqrt{k}\right\} \leq \eta\delta.$$

Along the line of the proof of Theorem 8.4 in [4], with $\eta\epsilon^2$ taking the place of ϵ of (5.13), there exits λ and n_1 such that

$$\sup_m \mathbb{P}\left\{\max_{i \leq n} |S_{m+i} - S_m| \geq \lambda \frac{\sqrt{|D_n|}}{C}\sqrt{k}\right\} \leq \frac{\eta\epsilon^2}{\lambda^2},$$

for all $n \geq n_1$ and $k \geq 1$. Put $\delta = \frac{\epsilon^2}{\lambda^2}$, since $\lambda > 1 > \epsilon$, we have $0 < \delta < 1$. Let $n_0 \geq \frac{n_1}{\delta} + 1$. Then for sufficiently large n, $n \geq n_0$ implies $[n\delta] \geq n_1$, which means $\frac{\sqrt{|D_{[n\delta]}|}}{C}\sqrt{k} \leq \frac{\sqrt{|D_{n_1}|}}{C}\sqrt{k}$. This is followed by the assumption on the nested spatial structure and $|D_n|$ sharing the monotonicity with its index n. Now the rest of this proof is the same as the corresponding part of the proof of Theorem 8.4 in [4]. □

Remark: If σ_n is a non-decreasing function with respect to its index n, then Theorem 5.12 still

holds provided by the assumption of (5.13) being replaced by

$$\sup_m \mathbb{P}\left\{\max_{i\le n}|S_{m+i}-S_m|\ge \lambda\sigma_n\sqrt{k}\right\}\le \frac{\epsilon}{\lambda^2}.$$

The constant $\sqrt{k}$ can be ignored if we do not need to specify the W_n, which is defined in (5.10).

In the following theorem, we use both strong mixing and uniform mixing in the assumption.

Theorem 5.13. *Let the random field $X_{i\in\mathbb{Z}^d}$ satisfies (H1), (H2), (H3) and (H4) in Theorem 3.1. We assume*

1) $D_n\subset\mathbb{Z}^d$ is from a nested spatial structure, i.e. $D_0\subseteq D_1\subseteq\cdots\subseteq D_{n-1}\subseteq D_n\subseteq\cdots$; $n\le|D_n|$; and there exists constants $K_0>0$ and $K_1>0$, such that for any $n_1\le n_2$,

$$dist(D^c_{n_2},D_{n_1})\ge K_1(n_2-n_1)^{K_0};$$

2) There exists a constant $k>0$ such that, for all $t\in[0,1]$,

$$\lim_n\frac{\sigma^2_{[nt]}}{\sigma^2_n}=kt; \tag{5.14}$$

3) There exists a constant $K_2>0$ such that (5.9) is satisfied, i.e.

$$|D|\le K_2(m+1)^d,\quad \forall D\subseteq\mathbb{Z}^d,$$

where $m=\sup\{dist(i,j):i,j\in D\}$, and

$$\frac{1}{|D_n|^2}\sum_{i\in D_n}\int_0^{\frac14}\left([\alpha^X_4]^{-1}(u)\right)^{3d}Q^4_{X_i}(u)du<\infty;$$

4) Let $\phi(m)=\sup_{\Lambda_1,\Lambda_2}\left\{\varphi^X_{|\Lambda_1|,|\Lambda_2|}(m):\Lambda_1=D_{n_1}-D_{n_2},\Lambda_2=D_{n_3}-D_{n_4},n_1>n_2,n_3>n_4\right\}$, and $\phi(m)\to0$ as $m\to\infty$;

Then we have

$$W_n(t)=\frac{S_{[nt]}}{\sigma_n\sqrt{k}}\xrightarrow{\mathcal{D}}W(t),$$

where $W_n(t)$ is defined by (5.10), $W(t)$ is a one dimensional Brownian motion on $[0,1]$.

Proof. This proof is analogous to the proof of Theorem 20.1 in the book [5] by Billingsley. We are going to check four conditions in Theorem 5.2.

Firstly, we check the asymptotically independent increments. Let

$$0 \le s_1 \le t_1 < s_2 \le t_2 < \cdots < s_r \le t_r \le 1, \quad \delta = \min\{s_i - t_{i-1} | i = 2, \cdots, r\}.$$

Then we have $\delta > 0$, $[ns_i] - [nt_{i-1}] + 1 \ge [n\delta]$ and

$$D_{[ns_1]} \subseteq D_{[nt_1]} \subset D_{[ns_2]} \subseteq D_{[nt_2]} \subset \cdots \subset D_{[ns_r]} \subseteq D_{[nt_r]},$$

for all n and $i = 2, \cdots, r$. We set

$$T_i = D_{[nt_i]} - D_{[ns_i]}, \quad \mathscr{M}_i = \sigma(T_i), \quad i = 1, \cdots, r. \tag{5.15}$$

Then the assumption 1 implies that, for all $i = 1, \cdots, r-1$,

$$dist(T_{i+1}, T_i) = dist(D^c_{[ns_{i+1}]}, D_{[nt_i]}) \ge K_1([n\delta] - 1)^{K_0}.$$

Because

$$W_n(t_i) - W_n(s_i) = \frac{1}{\sigma_n \sqrt{k}} \sum_{j \in T_i} X_i,$$

if we set

$$E_i = \{W_n(t_i) - W_n(s_i) \in H_i\},$$

where $H_i \in \mathscr{B}^d$, d-dimensional Borel sets, we have $E_i \in \mathscr{M}_i$. Lemma 2.3 guarantees the following estimation,

$$|\mathbb{P}(E_1 \cap \cdots \cap E_r) - \mathbb{P}(E_1) \cdots \mathbb{P}(E_r)| \le (r-1)\alpha^X_{l,l}\Big(K_1([n\delta] - 1)^{K_0}\Big), \quad l = \sum_{i=1}^r |T_i|.$$

The definition of strong mixing coefficient and the assumption 4 imply that the right side goes zero as $n \to \infty$. Hence the first condition is satisfied.

Secondly, we prove that $W_n^2(t)$ is uniformly integrable for each t. We note that

$$\begin{aligned}
\mathbb{E}^A|W_n^2(t)| &\le \frac{1}{A}\mathbb{E}(W_n^4(t)) \\
&= \frac{1}{A}\mathbb{E}\left(\frac{1}{\sigma_n^4 k^2}\left(\sum_{i \in D_{[nt]}} X_i\right)^4\right) \\
&= \frac{1}{A}\frac{1}{\sigma_n^4 k^2}\mathbb{E}\left(\left(\sum_{i \in D_{[nt]}} X_i\right)^4\right).
\end{aligned}$$

Because $\frac{1}{k^2}$, $\frac{|D_n|}{\sigma_n^2}$ and $\frac{|D_{[nt]}|}{|D_n|}$ are bounded, Theorem 5.10 and the assumption 3 imply the right

hand side of the above inequality convergences to zero as $A \to \infty$. Then we have

$$\lim_{A\to\infty} \sup_n \mathbb{E}_A \left(W_n^2(t)\right) = 0. \tag{5.16}$$

Hence $\{W_n^2(t)\}$ is uniformly integrable for each t.

Thirdly, we have

$$\mathbb{E}\left(W_n(t)\right) = \mathbb{E}\left(\frac{1}{\sigma_n\sqrt{k}} \sum_{i\in D_{[nt]}} X_i\right) = 0$$

and

$$\begin{aligned}
\mathbb{E}\left(W_n^2(t)\right) &= \mathbb{E}\left(\frac{1}{\sigma_n^2 k}\left(\sum_{i\in D_{[nt]}} X_i\right)^2\right) \\
&= \frac{\sigma_{[nt]}^2}{\sigma_n^2 k}\mathbb{E}\left(\frac{1}{\sigma_{[nt]}} \sum_{i\in D_{[nt]}} X_i\right)^2
\end{aligned}$$

Theorem 3.1 gives $\mathbb{E}\left(\frac{1}{\sigma_{[nt]}}\sum_{i\in D_{[nt]}} X_i\right)^2 \to 1$. Therefore the assumption 2 implies

$$\mathbb{E}\left(W_n^2(t)\right) \to t.$$

Lastly, we check the tightness. Let

$$Y_j = S_j - S_m, \quad j = m, m+1, \cdots, n. \tag{5.17}$$

By using Theorem 5.12, we only need to prove that, for any $m \in \mathbb{Z}^+$ and for all $\epsilon > 0$, there exists a $\lambda > 1$, such that for sufficiently large n we have

$$\mathbb{P}\left\{\max_{m\le j\le n} |Y_j| \ge 3\lambda\sqrt{|D_n|k}\right\} \le \frac{\epsilon}{\lambda^2}.$$

We set

$$S_n^* = \sum_{i\in D_n} |X_i|,$$

$$Y_j^* = S_j^* - S_m^*$$

and

$$p(n) = p(n;m) = \sup\left\{\tau \in \mathbb{Z}^+ : m \le \tau \le n-1, |D_\tau - D_m| \le |D_n|^{\frac{\delta}{4(2+\delta)}}\right\}.$$

Then we have, for any $m \leq n-1$, $m \in \mathbb{Z}^+$, $p(n) \to \infty$ as $n \to \infty$, and

$$\begin{aligned}
\mathbb{P}\left\{Y^*_{p(n)} \geq \lambda\sqrt{|D_n|k}\right\} &\leq \frac{\mathbb{E}|Y^*_{p(n)}|^{2+\delta}}{\left(\lambda\sqrt{|D_n|k}\right)^{2+\delta}} = \frac{\|Y^*_{p(n)}\|^{2+\delta}_{2+\delta}}{(\lambda\sqrt{k})^{2+\delta}|D_n||D_n|^{\frac{\delta}{2}}} \\
&\leq \frac{\left(\sum_{i\in D_{p(n)}-D_m}\|X_i\|_{2+\delta}\right)^{2+\delta}}{(\lambda\sqrt{k})^{2+\delta}|D_n||D_n|^{\frac{\delta}{2}}} \\
&\leq \frac{\left(C|D_n|^{\frac{\delta}{4(2+\delta)}}\right)^{2+\delta}}{(\lambda\sqrt{k})^{2+\delta}|D_n||D_n|^{\frac{\delta}{2}}} = \frac{K}{|D_n||D_n|^{\frac{\delta}{4}}},
\end{aligned}$$

where C and K are constants. Therefore we have

$$\lim_n |D_n|\mathbb{P}\left\{Y^*_{p(n)} \geq \lambda\sqrt{|D_n|k}\right\} = 0. \tag{5.18}$$

We note that $|S_n|/\sqrt{|D_n|k} \geq \lambda$ implies

$$\left|\mathbb{1}_{\{|S_n|/\sqrt{|D_n|k}\geq\lambda\}}\frac{|S_n|}{\sqrt{|D_n|k}}\right| \geq \lambda,$$

we have

$$\begin{aligned}
\mathbb{P}\{|S_n|\geq \lambda\sqrt{|D_n|k}\} &\leq \mathbb{P}\left\{\left|\mathbb{1}_{\{|S_n|/\sqrt{|D_n|k}\geq\lambda\}}\frac{|S_n|}{\sqrt{|D_n|k}}\right| \geq \lambda\right\} \\
&\leq \frac{1}{\lambda^2}\mathbb{E}_\lambda\left(\frac{S_n^2}{|D_n|k}\right) \\
&= \frac{1}{\lambda^2}\frac{\sigma_n^2}{|D_n|}\mathbb{E}_\lambda\left(W_n^2(1)\right).
\end{aligned}$$

By using Theorem 3.1, (H1)–(H3) implies

$$\limsup_n \frac{1}{|D_n|}\sum_{i,j\in D_n}|Cov(X_i,X_j)| < \infty,$$

which means $\limsup_n \frac{\sigma_n^2}{|D_n|} < \infty$. Then, additionally since (5.16) is satisfied, we have that, for any $\epsilon > 0$, we can find a sufficiently large λ such that $\mathbb{E}_\lambda\left(W_n^2(1)\right) < \epsilon$, i.e. we have

$$\mathbb{P}\{|S_n|\geq \lambda\sqrt{|D_n|k}\} \leq \frac{\epsilon}{\lambda^2}. \tag{5.19}$$

Since for any $m \leq j \leq n$,

$$Y_j^2 = (S_j - S_m)^2 \leq 2(S_j^2 + S_m^2),$$

as in the discussion of $\mathbb{P}\{|S_n|\geq \lambda\sqrt{|D_n|k}\}$, we have

$$\begin{aligned}\mathbb{P}\{|Y_j|\geq \lambda\sqrt{|D_n|k}\} &\leq \frac{2}{\lambda^2}\mathbb{E}_\lambda\left(\frac{S_j^2+S_m^2}{|D_n|k}\right)\\ &= \frac{2}{\lambda^2}\frac{\sigma_n^2}{|D_n|}\mathbb{E}_\lambda\Big(W_j^2(t)+W_m^2(t)\Big).\end{aligned}$$

Again, by using the property of (5.16) and the fact of $\limsup_n \frac{\sigma_n^2}{|D_n|}<\infty$, it means for any $\epsilon>0$, there exist sufficiently large λ such that $\mathbb{E}_\lambda\left(W_j^2(t)+W_m^2(t)\right)<\epsilon$ for all $m\leq n$. Hence we have

$$\mathbb{P}\{|Y_j|\geq \lambda\sqrt{|D_n|k}\}\leq \frac{\epsilon}{\lambda^2}. \tag{5.20}$$

Now, we define

$$E_i=\left\{\max_{j<i}|Y_j|< 3\lambda\sqrt{|D_n|k}\leq |Y_i|\right\}\quad i=m,m+1,\cdots,n.$$

Then we have

$$\begin{aligned}&\mathbb{P}\left\{\max_{m\leq j\leq n}|S_j-S_m|\geq 3\lambda\sqrt{|D_n|k}\right\}\\ =\;&\mathbb{P}\left\{\max_{m\leq j\leq n}|Y_j|\geq 3\lambda\sqrt{|D_n|k}\right\}\\ \leq\;&\mathbb{P}(|Y_n|\geq \lambda\sqrt{|D_n|k})+\sum_{j=m}^{n-1}\mathbb{P}(E_j\cap\{|Y_n-Y_j|\geq 2\lambda\sqrt{|D_n|k}\})\\ \leq\;&\mathbb{P}(|Y_n|\geq \lambda\sqrt{|D_n|k})+\sum_{j=m}^{n-p(n)-1}\mathbb{P}(|Y_j-Y_{j+p(n)}|\geq \lambda\sqrt{|D_n|k})\\ &+\sum_{j=m}^{n-p(n)-1}\mathbb{P}(E_j\cap\{|Y_n-Y_{j+p(n)}|\geq \lambda\sqrt{|D_n|k}\})+\sum_{j=n-p(n)}^{n-1}\mathbb{P}(|Y_n-Y_j|\geq \lambda\sqrt{|D_n|k})\\ =\;&\text{(I)}+\text{(II)}+\text{(III)}+\text{(IV)}.\end{aligned}$$

For (II) and (IV), each term can be bounded by $\mathbb{P}\left\{Y^*_{p(n)}\geq \lambda\sqrt{|D_n|k}\right\}$. Since the assumption 1 gives $n\leq |D_n|$, by using (5.18), we have

$$\text{(II)}+\text{(IV)}\leq (n-m)\mathbb{P}\left\{Y^*_{p(n)}\geq \lambda\sqrt{|D_n|k}\right\}\leq |D_n|\mathbb{P}\left\{Y^*_{p(n)}\geq \lambda\sqrt{|D_n|k}\right\}\leq \epsilon.$$

For (I) and (III), we note that (5.20) yields

$$\text{(I)}\leq \frac{\epsilon}{\lambda^2},$$

and, for all $m + p(n) \leq q \leq n - 1$,

$$\mathbb{P}(|Y_n - Y_q| \geq \lambda\sqrt{|D_n|k}) \leq \frac{\epsilon}{\lambda^2}.$$

By using the definition of the uniform mixing coefficient, the assumption 4 implies

$$\begin{aligned}
\text{(III)} &\leq \sum_{j=m}^{n-p(n)-1} \mathbb{P}(E_j)\left[\mathbb{P}(|Y_n - Y_{j+p(n)}| \geq \lambda\sqrt{|D_n|k}) + \varphi^X_{|D_{j+1}-D_m|,|D_n-D_{j+p(n)}|}(K_1[p(n)]^{K0})\right] \\
&\leq \max\left\{\mathbb{P}(|Y_n - Y_{j+p(n)}| \geq \lambda\sqrt{|D_n|k}) : m \leq j \leq n - p(n) - 1\right\} + \phi(K_1[p(n)]^{K0}) \\
&= \mathbb{P}(|Y_n - Y_q| \geq \lambda\sqrt{|D_n|k}) + \phi(K_1[p(n)]^{K0}) \\
&\leq \frac{\epsilon}{\lambda^2} + \epsilon,
\end{aligned}$$

where q is the subscript which makes $\mathbb{P}(|Y_n - Y_q| \geq \lambda\sqrt{|D_n|k})$ reach the maximum. Therefore,

$$\text{(I)+(II)+(III)+(IV)} \leq \frac{\epsilon}{\lambda^2} + \epsilon + \frac{\epsilon}{\lambda^2} + \epsilon.$$

Theorem 5.12 completes the proof. □

Remark: In the assumption 4 of Theorem 5.13, $\phi(m)$ can be bounded by $\varphi^X_{\infty,\infty}(m)$, i.e.

$$\phi(m) \leq \varphi^X_{\infty,\infty}(m).$$

In Bradley's paper [9], if the random field is strictly stationary, $\varphi^X_{\infty,\infty}(m) = 0$ implies an "m-dependence" random field. However, we do not have the similar result for non-stationary random fields. Therefore, this assumption is adopted in Theorem 5.13.

Proposition 1 in [19] by Doukhan gives that, for any fixed k, l and m, the uniform mixing coefficient is stronger than the strong mixing coefficient, i.e.

$$2\alpha_{k,l}(m) \leq \varphi_{k,l}(m). \tag{5.21}$$

Therefore, if we define

$$\varphi^{-1}_{|\Lambda_1|,|\Lambda_2|}(u) = \sup\{m \in \mathbb{N} : \varphi_{|\Lambda_1|,|\Lambda_2|}(m) > u\},$$

since the set $\{m \in \mathbb{N} : 2\alpha_{|\Lambda_1|,|\Lambda_2|}(m) > u\}$ implies $\{m \in \mathbb{N} : \varphi_{|\Lambda_1|,|\Lambda_2|}(m) > u\}$, we have

$$\varphi^{-1}_{|\Lambda_1|,|\Lambda_2|}(u) \geq \alpha^{-1}_{|\Lambda_1|,|\Lambda_2|}\left(\frac{u}{2}\right). \tag{5.22}$$

It is obvious that the above inequality also implies, for the constant $d > 0$,

$$\int_0^{\frac{1}{4}} \left[\alpha^{-1}_{|\Lambda_1|,|\Lambda_2|}(u)\right]^{3d} Q^4_{X_i}(u)du \leq \int_0^{\frac{1}{4}} \left[\varphi^{-1}_{|\Lambda_1|,|\Lambda_2|}(2u)\right]^{3d} Q^4_{X_i}(u)du,$$

which connects the application of Theorem 5.10 to uniform mixing coefficients, and can be used in the proof of Theorem 5.13. Therefore the assumption on the tail functions of the assumption 3 in Theorem 5.13 may be replaced with

$$\frac{1}{|D_n|^2} \sum_{i \in D_n} \int_0^{\frac{1}{4}} \left([\varphi^X_{l_1,l_2}]^{-1}(2u)\right)^{3d} Q^4_{X_i}(u)du < \infty, \quad \forall l_1 + l_2 = 4. \tag{5.23}$$

By using the inequality (5.21), we can directly develop a theorem like Theorem 3.1 but with uniform mixing coefficients.

Theorem 5.14. *If X satisfies*

$$\sum_{m \geq 1} m^{d-1} \varphi^X_{k,l}(m) < \infty, \quad k + l \leq 4, \tag{H1'}$$

$$\varphi^X_{1,\infty}(m) = o(m^{-d}), \tag{H2'}$$

$$\exists \delta > 0 \quad s.t. \quad \|X\|_{2+\delta} < \infty, \quad \sum_{m \geq 1} m^{d-1} \left[\varphi^X_{1,1}(m)\right]^{\frac{\delta}{2+\delta}} < \infty, \tag{H3'}$$

then

$$\limsup_n \frac{1}{|D_n|} \sum_{i,j \in D_n} |Cov(X_i, X_j)| < \infty.$$

If we assume additionally that

$$\liminf_n \frac{\sigma_n^2}{|D_n|} > 0, \tag{H4}$$

then we have

$$\frac{S_n}{\sigma_n} \xrightarrow{\mathcal{D}} N(0,1).$$

Based on Theorem 5.14, if (H1), (H2) and (H3) are replaced by (H1'), (H2') and (H3') respectively in Theorem 5.13, and the third assumption is replaced by (5.23), then Theorem 5.13 still holds, but only uniform mixing coefficients are involved.

5.3 FCLTs on indicated random fields

Let $\{R_iX_i\}_{i\in\mathbb{Z}^d}$ be the indicated random field which we introduced before. Since $|R_iX_i|\leq |X_i|$, for all $i\in\mathbb{Z}^d$, implies

$$\{x\geq 0:\mathbb{P}(|R_iX_i|> x)\leq u\}\supseteq\{x\geq 0:\mathbb{P}(|X_i|> x)\leq u\},\quad \forall u\in[0,1],$$

we have

$$Q_{R_iX_i}(u)\leq Q_{X_i}(u),\quad \forall i\in\mathbb{Z}^d,\forall u\in[0,1]. \tag{5.24}$$

For the inverse function of strong mixing coefficient, if $\alpha^X_{k,l}(m)\leq\alpha^Z_{k,l}(m)$, then we have

$$\{m\in\mathbb{N}:\alpha^X_{k,l}(m)>u\}\subseteq\{m\in\mathbb{N}:\alpha^Z_{k,l}(m)>u\}\quad \forall u\in[0,\frac{1}{4}],$$

which implies

$$[\alpha^X_{k,l}]^{-1}(u)\leq[\alpha^Z_{k,l}]^{-1}(u),\quad \forall u\in[0,\frac{1}{4}]. \tag{5.25}$$

In the proofs of Theorem 5.13, Theorem 3.1 is used for checking the third condition of Theorem 5.2. In the proof of Theorem 5.11, Theorem 3.1 is used to check the first condition of Theorem 5.3. It means the results on CLTs in Chapter 3 can be introduced to set up FCLTs with some suitable conditions on indicated random fields. To avoid tedious repetition, we omit the details of description and proofs of similar theorems except the following example, where Theorem 3.4 is involved.

Theorem 5.15. *Let random fields $X_{t\in\mathbb{Z}^d}$, $Z_{t\in\mathbb{Z}^d}$, and $R_{t\in\mathbb{Z}^d}$ satisfy (H0), which is introduced in Theorem 3.4. $\epsilon_{i\in\mathbb{Z}^d}$ be another random field, f satisfy (3.9). We suppose there exists $\delta>0$, such that $\|\epsilon\|_{2+\delta}<\infty$, the strong mixing coefficient, $\alpha^Z_{k,l}(m)$, of $Z_{t\in\mathbb{Z}^d}$ satisfies (H1)–(H3), and there exist constants $K_0,K_1,K_2>0$ such that*

$$\sup_{i\in\mathbb{Z}^d}\|f(Z_i,\epsilon_i)\|_{2+\delta}\leq K_0+K_1\|Z\|_{2+\delta}+K_2\|\epsilon\|_{2+\delta}.$$

Let $\{R_tX_t\}_{t\in\mathbb{Z}^d}$ satisfy

$$0<\liminf_n\frac{1}{|D_n|}Var\left(\sum_{t\in D_n}R_tX_t\right),$$

and

$$S_n=\sum_{i\in D_n}R_tX_t,\quad \sigma_n^2=Var(S_n).$$

We also assume

1) $D_n\subset\mathbb{Z}^d$ is with a nested spatial structure, i.e. $D_0\subseteq D_1\subseteq\cdots\subseteq D_{n-1}\subseteq D_n\subseteq\cdots$. There

exists constants $K_0 > 0$ and $K_1 > 0$, such that for any $n_1 \leq n_2$,

$$dist(D_{n_2}^c, D_{n_1}) \geq K_1(n_2 - n_1)^{K_0};$$

2) There exists a constant $K_3 > 0$ such that (5.9) being satisfied, i.e. for any $D \subseteq \mathbb{Z}^d$,

$$|D| \leq K_3(m+1)^d,$$

where $m = \sup\{dist(i,j) : i, j \in D\}$,

$$\sup_{i \in \mathbb{Z}^d} \int_0^{\frac{1}{4}} \left([\alpha_{1,1}^Z]^{-1}(u)\right)^{3d} Q_{Z_i}^4(u)du < \infty,$$

and

$$\sup_{i \in \mathbb{Z}^d} \int_0^{\frac{1}{4}} \left([\alpha_4^Z]^{-1}(u)\right)^{3d} Q_{Z_i}^4(u)du < \infty,$$

where $[\alpha_4^Z]^{-1}(u)$ is defined as in (5.8);

3) There exists a non-decreasing continuous function $F(t)$, $t \in [0,1]$, and a constant $K > 0$, such that

$$\frac{|D_{[nt_2]} - D_{[nt_1]}|}{|D_n|} \leq K[F(t_2) - F(t_1)], \quad 0 \leq t_1 \leq t_2 \leq 1.$$

Then we have

$$W_n(t) = \frac{S_{[nt]}}{\sigma_n \sqrt{K}} \xrightarrow{\mathcal{D}} W(t), \tag{5.26}$$

where $W(t)$ is a one dimensional Brownian motion on $[0,1]$.

Proof. We use Theorem 5.11. The proof of Theorem 3.4 implies $\{R_i X_i\}_{i \in \mathbb{Z}^d}$ satisfies (H1)–(H4) and

$$\limsup_n \frac{1}{|D_n|} Var\left(\sum_{t \in D_n} R_t X_t\right) < \infty.$$

For the rest of this proof, we only need to check assumptions of Theorem 5.11. In fact, only the assumption 2 of Theorem 5.11 need to be checked. It is directly from the fact that $[\alpha_{k,l}^{RX}(m)]^{-1} \leq [\alpha_{k,l}^Z(m)]^{-1}$. This completes the proof. □

Chapter 6

Future research

The current research in this thesis focuses on three parts: central limit theorems, estimation of variances and functional central limit theorems.

6.1 On the sampling method

These asymptotic properties are original from the indicated sampling method, where each individual of the super-population has equal or unequal selection probability. Based on real situations, we further assume that each individual and its selection probability are driven by the auxiliary information. We also introduce some assumptions about the spatial structures of the population. This sampling method is appropriate for setting up estimators such as Horvitz-Thompson estimator and others.

We have set up some asymptotic results for this indicated sampling method in this thesis. The future work for this part will focus on applications. Practically, we need to specify functions f and g for real problems. This is going to lead some researches on curve fitting methods and the goodness of fit, from the auxiliary information to interested quantities and selection probabilities.

6.2 On CLTs

It is important to understand the asymptotics of a new complex survey method. For this indicated sampling method, we prove central limit theorems with the assumption of conditional independence properties. In this thesis, we also generalized Fuller's central limit theorem, Theorem 1.3.2 in his book in 2011, to dependent random fields.

For the future work on CLTs, the more general assumptions allow for wider applications. We will consider relaxing the independence assumptions and introduce more assumptions on dependences to enable more general practical problems to be tackled.

6.3 On variance estimations

In Chapter 4, with the assumption on the joint-blocks spatial structures, we proved the L_2 consistency for the estimators of the variance of the population. We then generalized the results on estimating the variance by Carlstein in 1986. Compared with the existed method, we provided a new way to describe re-sample domains. This method allows the domain have complicated geometric features such as an infinite boundary, which is technically avoided as an assumption in many works.

The future work for this part is to develop our method to over-lapping re-sample domains. We also need a further research on comparing our method to the existed method. This is because of, intuitively, our method implies the existed method. It has a massive application background such as sampling from high dimension continuous spaces.

6.4 On FCLTs

In Chapter 5, by using Billingsley's Theorem 15.6, in his book in 1968, with the assumptions on the nested spatial structures and the proper estimation of the fourth moment of the sample sum, we prove functional central limit theorems, where the estimation of the fourth moment develops Rio's Theorem 2.1 in his report in 2013.

The further work of this part lies on polishing or refining some assumptions in our results. It also includes examining whether Bradley's result in 1989 works for non-stationary random fields.

Appendix

A1: Figures on NHANES 2011-2014 (in support of Example 2, Chapter 1)

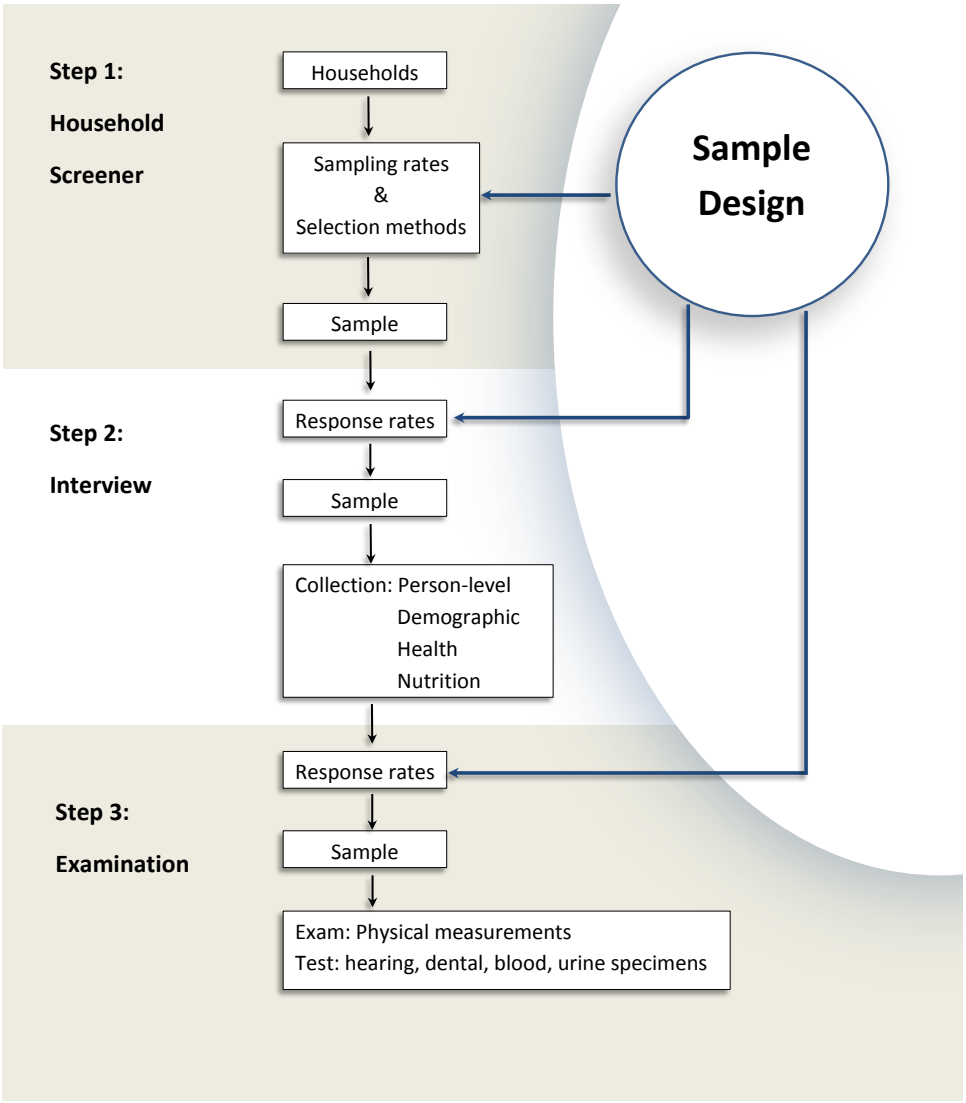


Figure A1: Three steps of NHANES 2011-2014

Here we provide two figures to exhibit three steps in the NHANES 2011–2014 survey and the first stage of this sample design. We use arrows in these figures to stand for the direction of information flow. In Figure A1, the sample design is depicted in each step. The sample is affected by response rates, sampling rates and selection methods of the sample design.

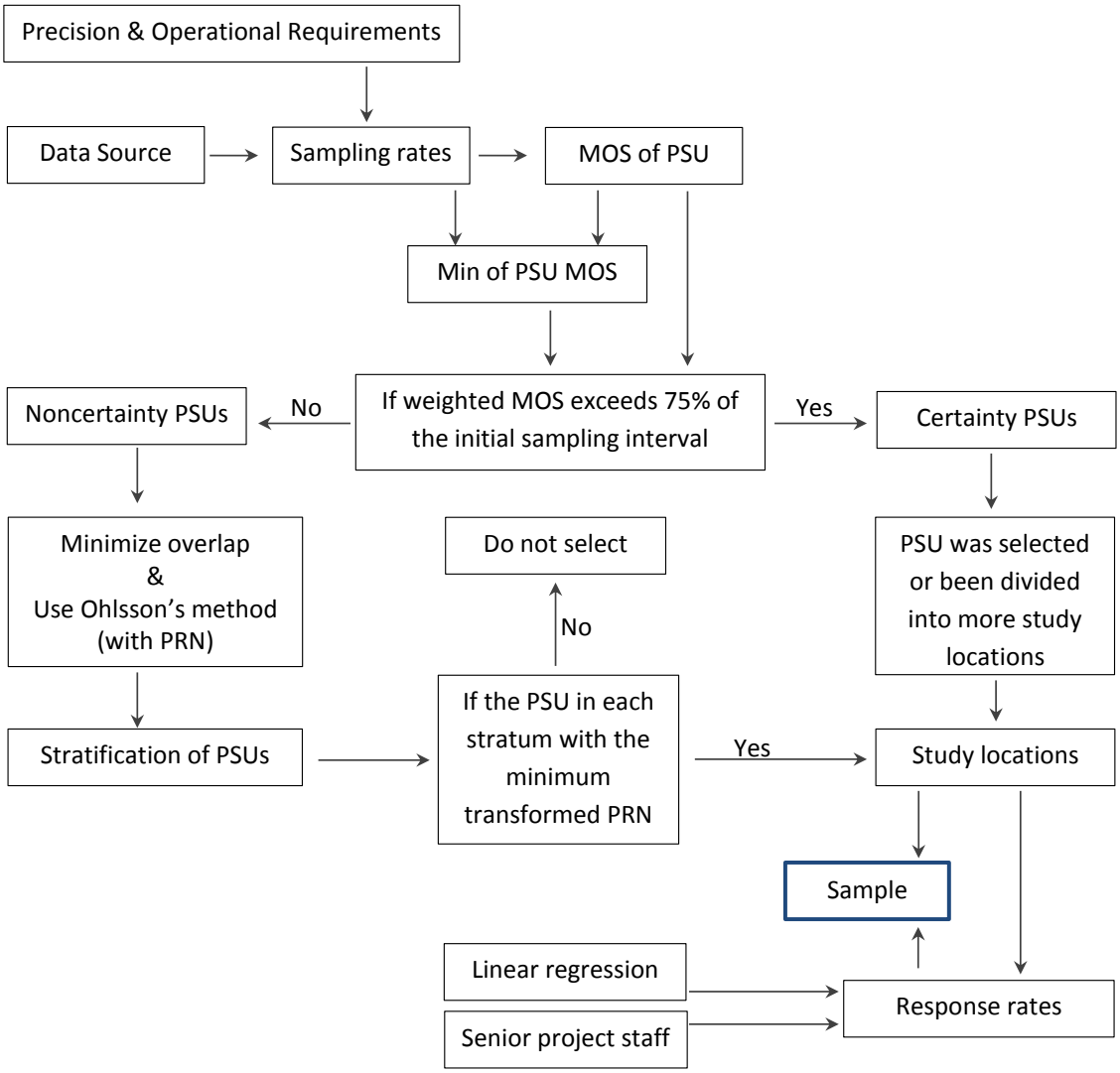


Figure A2: The first stage of sample design in NHANES 2011-2014

Figure A2 gives some details of one out of four stages in NHANES. In the processes of sampling, there are many implicit random effects, which are abstracted as $\epsilon_{i\in\mathbb{Z}^d}$ in our sampling strategy. Furthermore, the dependence is also implicit in the sampling process, e.g. the dependence within PSUs, the dependence between strata, and/or the dependence between persons. This information is taken into account in our asymptotic results with some abstracted conditions in Chapter 3, Chapter 4 and Chapter 5.

A2: Some basic inequalities

Theorem A1. (Hölder's inequality, equation (21.15) on page 276 in [5])

Let $p > 1$, $q > 1$ and $\frac{1}{q} + \frac{1}{p} = 1$. Then for any two random variables, X_1 and X_2, we have

$$\|X_1 X_2\|_1 \leq \|X_1\|_q \|X_2\|_p. \tag{1}$$

Theorem A2. (Hölder's inequality, a special case of Theorem 188 of [29]) *Let $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$, $p, q, r > 0$. Then for any two random variables, X_1 and X_2, we have*

$$\|X_1 X_2\|_r \leq \|X_1\|_p \|X_2\|_q. \tag{2}$$

Proof. It is a straightforward extension of Theorem A1. We note that $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ implies

$$\frac{r}{p} + \frac{r}{q} = \frac{1}{\frac{p}{r}} + \frac{1}{\frac{q}{r}} = 1, \quad \frac{p}{r} > 1, \quad \text{and} \quad \frac{q}{r} > 1.$$

Hence Theorem A1 completes the proof. □

Theorem A3. (Minkowski's inequality, a special case of equation (31) on page 194 in [43])

If $r \geq 1$, then for any random variables, X_1 and X_2, we have

$$\|X_1 + X_2\|_r \leq \|X_1\|_r + \|X_2\|_r. \tag{3}$$

A3: Kolmogorov Existence Theorem

Let $\{X_t\}_{t \in T}$ be a random process, T be any non-empty index set, $k \in \mathbb{N}$. Given two consistency conditions:

(C1) If $(s(1), s(2), \cdots, s(k))$ is any permutation of $(1, 2, \cdots, k)$, then for distinct $t_1, \cdots, t_k \in T$, and any Borel $H_1, \cdots, H_k \subseteq \mathbb{R}$, we have

$$\mu_{t_1 \cdots t_k}(H_1 \times \cdots \times H_k) = \mu_{t_{s(1)} \cdots t_{s(k)}}(H_{s(1)} \times \cdots \times H_{s(k)})$$

(C2) For distinct $t_1, \cdots, t_k \in T$, and any Borel $H_1 \cdots, H_{k-1} \subseteq \mathbb{R}$, we have

$$\mu_{t_1 \cdots t_k}(H_1 \times \cdots \times H_{k-1} \times \mathbb{R}) = \mu_{t_1 \cdots t_{k-1}}(H_1 \times \cdots \times H_{k-1})$$

Theorem A4. (Theorem 15.1.3 in [61]) *A family of Borel probability measures* $\{\mu_{t_1,\cdots,t_k}; k \in \mathbb{N}, t_i \in T \text{distinct}\}$*, with* $\mu_{t_1,\cdots,t_k}$ *a measure on* $\mathbb{R}^k$*, satisfies the consistency conditions (C1) and (C2) above iff there exists a probability triple* $(\mathbb{R}^T, \mathcal{F}^T, \mathbb{P})$*, and random variables* $\{X_t\}_{t\in T}$ *defined on this triple, such that for all* $k \in \mathbb{N}$*, distinct* $t_1, \cdots, t_k \in T$*, and Borel* $H \subseteq \mathbb{R}^k$*, we have*

$$\mathbb{P}\Big[(X_{t_1}, \cdots, X_{t_k}) \in H\Big] = \mu_{t_1\cdots t_k}(H).$$

A4: Truncation technique

Let X be a random variable,

$$\overline{X}(N) = \begin{cases} X, & |X| \leq N, \\ N, & X > N, \\ -N, & X < -N, \end{cases} \qquad \underline{X}(N) = X - \overline{X}(N), \qquad N \in \mathbb{R}.$$

We note the t-th absolute moment satisfies

$$\mathbb{E}|X|^r = \int |x|^r dF = \lim_{N\to\infty} \int_{|x|\leq N} |x|^r dF,$$

where $r > 0$, F is the distribution of X. Then we have the following lemma.

Lemma A1. *If* $\mathbb{E}|X|^r < \infty$*, then* $\mathbb{E}|\underline{X}(N)|^r \to 0$ *as* $N \to \infty$*.*

By using the definition of almost sure convergence, we also have

Lemma A2. $\underline{X}(N) \xrightarrow{a.s.} 0$ *as* $N \to \infty$*.*

If the higher order moment is bounded, we have the following lemmas.

Lemma A3. *If* $\mathbb{E}|X|^{2r} < \infty$*, then* $\mathbb{E}|\overline{X}\underline{X}|^r \to 0$ *as* $N \to \infty$*.*

Lemma A4. *Let* $\sigma^2 = Var(X)$*,* $\sigma_N^2 = Var(\overline{X})$*. If* $\mathbb{E}|X|^2 < \infty$*, then*

$$\frac{\sigma_N^2}{\sigma^2} \to 1, \quad \text{as } N \to \infty.$$

We note that

$$\frac{X}{\overline{X}} \xrightarrow{P} 1 \quad \text{as } N \to \infty.$$

Then Slutsky's Theorem implies: If

$$\frac{\overline{X}}{\sigma_N} \xrightarrow{D} N(0,1),$$

then we have

$$\frac{X}{\sigma} = \frac{\overline{X}}{\sigma_N}\frac{X}{\overline{\overline{X}}}\frac{\sigma_N}{\sigma} \xrightarrow{D} N(0,1).$$

A5: Proof of Theorem 2.2

The proof in this section refers to [19] and [34]. We prove a general case for Theorem 2.2, i.e. let $p, q > 1$, then for any two random variables X and Y with two sigma fields $\sigma_1 = \sigma(X)$ and $\sigma_2 = \sigma(Y)$ generated by them respectively, we have

$$|Cov(X,Y)| \leq 8\alpha^{1-\frac{1}{p}-\frac{1}{q}}\|X\|_q\|Y\|_p, \tag{4}$$

where $\alpha = \alpha_{k,l}(m)$, because the result is for any k, l and m, we omit them for shorthand, that is

$$\alpha = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|\colon A \in \sigma_1, B \in \sigma_2\}.$$

To prove this, we need the following two lemmas.

Lemma A5. $|Cov(X,Y)| \leq 4\alpha\|X\|_\infty\|Y\|_\infty$.

Proof. Let $u = sign[\mathbb{E}(X|\sigma_2) - \mathbb{E}(X)]$ and $v = sign[\mathbb{E}(Y|\sigma_1) - \mathbb{E}(Y)]$, where $sign(\cdot)$ is the sign function of real numbers. Then we have

$$\begin{aligned}
|Cov(X,Y)| &= |\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)| \\
&= |\mathbb{E}(X\mathbb{E}(Y|\sigma_1)) - \mathbb{E}(X)\mathbb{E}(Y)| \\
&= |\mathbb{E}[X(\mathbb{E}(Y|\sigma_1) - \mathbb{E}(Y))]| \\
&\leq \mathbb{E}[|X||\mathbb{E}(Y|\sigma_1) - \mathbb{E}(Y)|] \\
&\leq \|X\|_\infty\mathbb{E}|\mathbb{E}(Y|\sigma_1) - \mathbb{E}(Y)| \\
&= \|X\|_\infty\mathbb{E}[v\mathbb{E}(Y|\sigma_1) - \mathbb{E}(Y)] \\
&= \|X\|_\infty Cov(v,Y) \\
&\leq \|X\|_\infty|Cov(v,Y)|.
\end{aligned}$$

Following the same idea, we have

$$|Cov(v,Y)| = |Cov(Y,v)| \leq \|Y\|_\infty|Cov(u,v)|,$$

and thereafter

$$|Cov(X,Y)| \leq \|X\|_\infty\|Y\|_\infty|Cov(u,v)|.$$

We set up four events $u^+ = \{u = 1\}$, $u^- = \{u = -1\}$, $v^+ = \{v = 1\}$ and $v^- = \{v = -1\}$, then

$$u^+, u^- \in \sigma_1, \quad v^+, v^- \in \sigma_2, \quad u^+ \cap v^+ \cap u^- \cap v^- = \emptyset,$$

and the expectations can be calculated in the following ways.

$$\begin{aligned}
\mathbb{E}(uv) &= 1 \times \mathbb{P}[(u^+ \cap v^+) \cup (u^- \cap v^-)] - 1 \times \mathbb{P}[(u^+ \cap v^-) \cup (u^- \cap v^+)] \\
&= \mathbb{P}(u^+ \cap v^+) + \mathbb{P}(u^- \cap v^-) - \mathbb{P}(u^+ \cap v^+ \cap u^- \cap v^-) \\
&\quad -\mathbb{P}(u^+ \cap v^-) - \mathbb{P}(u^- \cap v^+) + \mathbb{P}(u^+ \cap v^- \cap u^- \cap v^+) \\
&= \mathbb{P}(u^+ \cap v^+) + \mathbb{P}(u^- \cap v^-) - \mathbb{P}(u^+ \cap v^-) - \mathbb{P}(u^- \cap v^+), \\
\mathbb{E}(u) &= 1 \times \mathbb{P}(u^+) - 1 \times \mathbb{P}(u^-) = \mathbb{P}(u^+) - \mathbb{P}(u^-), \\
\mathbb{E}(v) &= \mathbb{P}(v^+) - \mathbb{P}(v^-).
\end{aligned}$$

Hence

$$\begin{aligned}
|Cov(u, v)| &= |\mathbb{E}(u, v) - \mathbb{E}(u)\mathbb{E}(v)| \\
&= |\mathbb{P}(u^+ \cap v^+) + \mathbb{P}(u^- \cap v^-) - \mathbb{P}(u^+ \cap v^-) - \mathbb{P}(u^- \cap v^+) \\
&\quad -\mathbb{P}(u^+)\mathbb{P}(v^+) - \mathbb{P}(u^-)\mathbb{P}(v^-) + \mathbb{P}(u^+)\mathbb{P}(v^-) + \mathbb{P}(u^-)\mathbb{P}(v^+)| \\
&\leq 4 \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|\colon A \in \sigma_1 . B \in \sigma_2\} \\
&= 4\alpha.
\end{aligned}$$

This completes the proof. □

Lemma A6. *For all* $1 \leq p \leq \infty$,

$$|Cov(X, Y)| \leq 6\alpha^{1-\frac{1}{p}} \|X\|_p \|Y\|_\infty \tag{5}$$

Proof. For any real number $a > 0$, we define

$$\mathbb{1}_{\{|X| \leq a\}} = \begin{cases} 1, & |X| \leq a, \\ 0, & \text{otherwise}, \end{cases} \qquad \mathbb{1}_{\{|X| > a\}} = \begin{cases} 1, & |X| > a, \\ 0, & \text{otherwise}, \end{cases}$$

and

$$\overline{X} = X\mathbb{1}_{\{|X| \leq a\}} = \begin{cases} X, & |X| \leq a, \\ 0, & \text{otherwise}, \end{cases} \qquad \underline{X} = X\mathbb{1}_{\{|X| > a\}} = \begin{cases} X, & |X| > a, \\ 0, & \text{otherwise}. \end{cases}$$

Then we have $X = \overline{X} + \underline{X}$ and for any $p, q \neq 0$,

$$\mathbb{1}^p_{\{|X| > a\}} = \mathbb{1}^q_{\{|X| > a\}} = \mathbb{1}_{\{|X| > a\}} = \mathbb{1}_{\{|X|^p > a^p\}} = \mathbb{1}_{\{|X|^q > a^q\}}. \tag{6}$$

We note that

$$\begin{aligned} |Cov(X,Y)| &= |Cov(\overline{X}+\underline{X},Y)| = |Cov(\overline{X},Y)+Cov(\underline{X},Y)| \\ &\leq |Cov(\overline{X},Y)| + |Cov(\underline{X},Y)|. \end{aligned} \tag{7}$$

Lemma A5 implies

$$|Cov(\overline{X},Y)| \leq 4\alpha\|\overline{X}\|_\infty\|Y\|_\infty = 4\alpha a\|Y\|_\infty.$$

We have

$$\begin{aligned} |Cov(\underline{X},Y)| &= |\mathbb{E}(\underline{X}Y)-\mathbb{E}(\underline{X})\mathbb{E}(Y)| \\ &\leq |\mathbb{E}(\underline{X}Y)| + |\mathbb{E}(\underline{X})\mathbb{E}(Y)| \\ &\leq \|Y\|_\infty|\mathbb{E}(\underline{X})| + \|Y\|_\infty|\mathbb{E}(\underline{X})| = 2\|Y\|_\infty|\mathbb{E}(\underline{X})|, \end{aligned}$$

and by using (6) and Hölder's inequality for any $p,q>1$, $\frac{1}{p}+\frac{1}{q}=1$,

$$\begin{aligned} \mathbb{E}|\underline{X}| &= \mathbb{E}|X\mathbb{1}_{\{|X|>a\}}| \\ &\leq (\mathbb{E}|X|^p)^{\frac{1}{p}}\left(\mathbb{E}|\mathbb{1}_{\{|X|>a\}}|^q\right)^{\frac{1}{q}} = (\mathbb{E}|X|^p)^{\frac{1}{p}}\left(\mathbb{E}\mathbb{1}^q_{\{|X|>a\}}\right)^{\frac{1}{q}} \\ &= (\mathbb{E}|X|^p)^{\frac{1}{p}}\left(\mathbb{E}\mathbb{1}_{\{|X|>a\}}\right)^{\frac{1}{q}} = (\mathbb{E}|X|^p)^{\frac{1}{p}}\left(\mathbb{E}\mathbb{1}_{\{|X|^p>a^p\}}\right)^{\frac{1}{q}} \\ &= (\mathbb{E}|X|^p)^{\frac{1}{p}}\left(1\times\mathbb{P}(|X|^p>a^p)+0\times\mathbb{P}(|X|^p\leq a^p)\right)^{\frac{1}{q}} \\ &= (\mathbb{E}|X|^p)^{\frac{1}{p}}\left(\mathbb{P}(|X|^p>a^p)\right)^{\frac{1}{q}} \\ &\leq (\mathbb{E}|X|^p)^{\frac{1}{p}}\left(\frac{\mathbb{E}|X|^p}{a^p}\right)^{\frac{1}{q}} \\ &= (\mathbb{E}|X|^p)^{\frac{1}{p}+\frac{1}{q}}\,a^{-\frac{p}{q}} = a\frac{\mathbb{E}|X|^p}{a^p}. \end{aligned}$$

Since $a>0$ is arbitrary, it is acceptable if we set

$$\frac{\mathbb{E}|X|^p}{a^p}=\alpha, \quad \text{i.e. } a=(\mathbb{E}|X|^p)^{\frac{1}{p}}\,\alpha^{-\frac{1}{p}} = \|X\|_p\alpha^{-\frac{1}{p}}.$$

Then we have

$$|Cov(\underline{X},Y)| \leq 2\alpha a\|Y\|_\infty,$$

and (7) shows

$$|Cov(\underline{X},Y)| \leq 4\alpha a\|Y\|_\infty + 2\alpha a\|Y\|_\infty = 6\alpha^{1-\frac{1}{p}}\|X\|_p\|Y\|_\infty.$$

This completes the proof. □

We define

$$\overline{Y}=Y\mathbb{1}_{\{|Y|\leq b\}}, \quad \underline{Y}=Y\mathbb{1}_{\{|Y|>b\}}, \quad \forall b>0.$$

Then we have

$$|Cov(Y,X)| \leq |Cov(\overline{Y},X)| + |Cov(\underline{Y},X)|. \tag{8}$$

For the first term on the right side of (8), Lemma A6 implies

$$|Cov(\overline{Y}, X) \leq 6\alpha^{1-\frac{1}{p}}\|X\|_p\|\overline{Y}\|_\infty = 6\alpha^{1-\frac{1}{p}}\|X\|_p b.$$

For all $m, p > 1$ and $\frac{1}{m} + \frac{1}{p} = 1$, we estimate the second term on the right side of (8),

$$\begin{aligned} |Cov(\underline{Y}, X)| &\leq |\mathbb{E}(\underline{Y}X)| + |\mathbb{E}(\underline{Y})\mathbb{E}(X)| \\ &\leq (\mathbb{E}|\underline{Y}|^m)^{\frac{1}{m}}(\mathbb{E}|X|^p)^{\frac{1}{p}} + \mathbb{E}|\underline{Y}|\mathbb{E}|X| \\ &\leq 2\|X\|_p\|\underline{Y}\|_m \end{aligned}$$

Furthermore, for all $\tau, t > 1$ and $\frac{1}{\tau} + \frac{1}{t} = 1$, (6) implies

$$\begin{aligned} \mathbb{E}|\underline{Y}|^m &\leq (\mathbb{E}|\underline{Y}^m|^\tau)^{\frac{1}{\tau}}(\mathbb{E}\mathbb{1}_{\{|Y|>b\}})^{\frac{1}{t}} \\ &= (\mathbb{E}|\underline{Y}|^{m\tau})^{\frac{1}{\tau}}(\mathbb{E}\mathbb{1}_{\{|Y|^{m\tau}>b^{m\tau}\}})^{\frac{1}{t}} \\ &\leq (\mathbb{E}|\underline{Y}|^{m\tau})^{\frac{1}{\tau}}\left(\frac{\mathbb{E}|Y|^{m\tau}}{b^{m\tau}}\right)^{\frac{1}{t}} \\ &= \mathbb{E}|Y|^{m\tau}b^{m(1-\tau)} \end{aligned}$$

Now we have

$$\begin{aligned} |Cov(\underline{Y}, X)| &\leq 2\|X\|_p(\mathbb{E}|Y|^{m\tau}b^{m(1-\tau)})^{\frac{1}{m}} \\ &= 2\|X\|_p b\left(\frac{\mathbb{E}|Y|^{m\tau}}{b^{m\tau}}\right)^{\frac{1}{m}}. \end{aligned}$$

Similarly, we use the idea of setting a in Lemma A6. We set

$$\alpha^{1-\frac{1}{p}} = \left(\frac{\mathbb{E}|Y|^{m\tau}}{b^{m\tau}}\right)^{\frac{1}{m}}, \quad \text{i.e.} \quad \alpha^{\frac{1}{\tau}-\frac{1}{p\tau}} = \frac{\|Y\|_{m\tau}}{b}.$$

Then we have

$$b = \|Y\|_{m\tau}\alpha^{-\frac{1}{\tau}+\frac{1}{p\tau}},$$

and

$$\begin{aligned} |Cov(X, Y)| &\leq 6\alpha^{1-\frac{1}{p}}b\|X\|_p + 2\alpha^{1-\frac{1}{p}}b\|X\|_p \\ &= 8\alpha^{1=\frac{1}{p}}\alpha^{-\frac{1}{\tau}+\frac{1}{p\tau}}\|Y\|_{m\tau}\|X\|_p. \end{aligned}$$

Now we set $q = m\tau$. Then $\frac{1}{m} + \frac{1}{p} = 1$ implies (4). If we go further by setting $p = q = 2 + \delta$ in (4), Theorem 2.2 is proved.

A6: Stein's lemma

Stein's lemma was first introduced in [68]. Bolthausen [7] and Guyon [28] used it as a tool in proving CLTs. To give a more detailed proof of Stein's lemma, we provide some preliminaries at the beginning of this section.

Lemma A7. *Let $\nu_{n\in\mathbb{N}}$ be a sequence of probability measures. If*

$$\sup_n \int x^2 \nu_n(dx) < \infty,$$

then there exists a subsequence ν_{n_k} and a probability measure ν such that $\nu_{n_k} \xrightarrow{w} \nu$.

Proof. By using the definition of tightness on page 276 of [15], the assumption of this lemma directly implies the tightness of ν_n. Then Theorem 25.10 of [5] implies the existence of the subsequence. □

Lemma A8. *Let $\nu_{n\in\mathbb{N}}$ be a sequence of probability measure, $g(x)$ be a function, which is continuous in $\mathbb{R}$. If*

$$\sup_n \int x^2 \nu_n(dx) < \infty, \tag{9}$$

and

$$\frac{|g(x)|}{x^2} \to 0, \quad \text{as } x \to \infty, \tag{10}$$

then for the subsequence ν_{n_k} and ν in Lemma A7, we have

$$\int g(x)\nu_{n_k}(dx) \to \int g(x)\nu(dx).$$

Proof. From Lemma A7, we have $\nu_{n_k} \xrightarrow{w} \nu$. The remainder of the proof is to show that

$$\left| \int g(x)\nu_{n_k}(dx) - \int g(x)\nu(dx) \right| \to 0.$$

Condition (10) implies, for all $\epsilon > 0$, there exists $N_\epsilon > 0$, such that

$$\frac{|g(x)|}{x^2} < \epsilon, \quad x > N_\epsilon.$$

We set

$$\bar{g}_N(x) = \begin{cases} g(-N_\epsilon), & x < -N_\epsilon \\ g(x), & |x| \leq N_\epsilon \\ g(N_\epsilon), & x > N_\epsilon. \end{cases}$$

Since $g(x)$ is continuous, $\bar{g}_N(x)$ is bounded. Hence, by the definition of weak convergence, we

have

$$\left|\int \bar{g}_N(x)\nu_{n_k}(dx) - \int \bar{g}_N(x)\nu(dx)\right| \to 0. \quad \text{as } k \to \infty.$$

Furthermore, condition (9) ensures

$$\left|\int g(x)\nu_{n_k}(dx) - \int \bar{g}_N(x)\nu_{n_k}(dx)\right| \to 0. \quad \text{as } N_\epsilon \to \infty.$$

Similarly, ν is a limit of ν_{n_k} implies

$$\int x^2\nu(dx) \leq \sup_n \int x^2\nu_n(dx) < \infty.$$

Then we have

$$\left|\int g(x)\nu(dx) - \int \bar{g}_N(x)\nu(dx)\right| \to 0. \quad \text{as } N_\epsilon \to \infty.$$

We note

$$\begin{aligned} & \left|\int g(x)\nu_{n_k}(dx) - \int g(x)\nu(dx)\right| \\ \leq & \left|\int g(x)\nu_{n_k}(dx) - \int \bar{g}_N(x)\nu_{n_k}(dx)\right| + \left|\int \bar{g}_N(x)\nu_{n_k}(dx) - \int \bar{g}_N(x)\nu(dx)\right| \\ & \qquad + \left|\int g(x)\nu(dx) - \int \bar{g}_N(x)\nu(dx)\right| \\ = & \ (\mathrm{I}) + (\mathrm{II}) + (\mathrm{III}). \end{aligned}$$

For (I) and (III), since $\sup_n \int x^2\nu_n(dx) < \infty$, there exists a constant K, such that

$$\sup_{N_\epsilon}\{(\mathrm{I})\} \leq \epsilon K, \quad \sup_{N_\epsilon}\{(\mathrm{III})\} \leq \epsilon K.$$

For the second term, (II)$\to 0$ as $k \to \infty$. Therefore we have

$$\left|\int g(x)\nu_{n_k}(dx) - \int g(x)\nu(dx)\right| \to 0 \quad \text{as} \quad k \to \infty.$$

This completes the proof. □

Theorem A5. (Stein's Lemma, see [68], [7] and [28]) *Let $\nu_{n\in\mathbb{N}}$ be a sequence of probability measure on $\mathbb{R}$. If (9) is satisfied and*

$$\lim_{n\to\infty} \int (i\lambda - x)e^{i\lambda x}\nu_n(dx) = 0, \quad \forall \lambda \in \mathbb{R}. \tag{11}$$

Then

$$\nu_n \xrightarrow{\mathcal{D}} N(0,1).$$

Proof. Let $g(x) = (i\lambda - x)e^{i\lambda x}$. Then Lemma A8 implies there exists a probability measure ν,

which satisfies

$$\int g(x)\nu_{n_k}(dx) \to \int g(x)\nu(dx).$$

The uniqueness of the limit means $\int g(x)\nu(dx) = 0$, with the condition (11), i.e. we have

$$\int (i\lambda - x)e^{i\lambda x}\nu(dx) = 0.$$

Let $\phi(\lambda)$ be the characteristic function of $\nu(x)$, i.e.

$$\phi(\lambda) = \int e^{i\lambda x}\nu(dx).$$

We note

$$\left|\int \frac{\partial}{\partial\lambda}(e^{i\lambda x})\nu(dx)\right| \leq \left(\int x^2\nu(dx)\right)^{\frac{1}{2}} < \infty.$$

Hence we can calculate the derivative as the following:

$$\frac{d\phi(\lambda)}{d\lambda} = \int \frac{\partial}{\partial\lambda}(e^{i\lambda x})\nu(dx) = i\int xe^{i\lambda x}\nu(dx).$$

Then we have

$$-\lambda\phi(\lambda) = \phi'(\lambda).$$

Because $\nu(x)$ is a probability measure, we also have $\phi(0) = 1$. By solving this differential equation with the initial condition, we get

$$\phi(\lambda) = e^{-\frac{\lambda^2}{2}},$$

which means $\nu(x)$ is a normal distribution. □

A7: Proof of Theorem 3.1

By using the definition of $dist(\Lambda_1, \Lambda_2)$ in Chapter 2, we have, for fixed i_0, and for all $m \geq 1$,

$$\sum_{j:|i_0-j|=m} \alpha_{1,1}(i_0, j) \leq C(2m+1)^{d-1}\alpha_{1,1}(m),$$

where C is a constant. This inequality and the property of the strong mixing coefficient, (2.6), support

$$\sum_{i,j\in D_n} \alpha(i,j) \leq |D_n|\left[\frac{1}{4} + C\sum_{m\geq 1} m^{d-1}\alpha_{1,1}(m)\right], \tag{12}$$

where C is a new constant. Now the condition (H1) gives

$$\frac{1}{|D_n|}\sum_{i,j\in D_n}\alpha(i,j)<\infty.$$

By Theorem 2.2 and condition (H3), we have

$$\limsup_n \frac{1}{|D_n|}\sum_{i,j\in D_n}|Cov(X_i,X_j)|\leq C\|X\|_{2+\delta}^2\cdot\limsup_n\sum_{m\geq 1}m^{d-1}\alpha_{1,1}^{\frac{\delta}{2+\delta}}(m)<\infty,$$

where C is a new constant. Then the first part of the result in Theorem 3.1 is proved.

For the asymptotic normality as the second part of the result, the two conditions of Stein's Lemma, i.e. Theorem A5, need to be checked.

Firstly, we check condition (9). For truncated $X_i(N)$'s, which is introduced in Appendix A4, $\mathbb{E}(X_i)=0$ means

$$\frac{S_n}{\sigma_n}=\frac{1}{\sigma_n}\sum_{i\in D_n}\left[\overline{X}_i-\mathbb{E}(\overline{X}_i)\right]+\frac{1}{\sigma_n}\sum_{i\in D_n}\left[\underline{X_i}-\mathbb{E}(\underline{X_i})\right].$$

Theorem 2.2 and the estimation of (12) imply the estimation of the variance of the second term, in the above inequality, as

$$\mathbb{E}\left(\frac{1}{\sigma_n}\sum_{i\in D_n}\left[\underline{X_i}-\mathbb{E}(\underline{X_i})\right]\right)^2\leq\frac{|D_n|}{\sigma_n^2}C\|\underline{X}\|_{2+\delta}^2\sum_{m\geq 1}m^{d-1}\alpha_{1,1}^{\frac{\delta}{2+\delta}}(m),$$

where, by Lemma A2, $\|\underline{X}\|_{2+\delta}\to 0$ as $N\to\infty$ for all $i\in D_n$. Condition (H4) implies $|D_n|/\sigma_n^2$ is bounded. Therefore,

$$\mathbb{E}\left(\frac{1}{\sigma_n}\sum_{i\in D_n}\left[\underline{X_i}-\mathbb{E}(\underline{X_i})\right]\right)^2\to 0,\quad \text{as } N\to\infty.$$

We conclude that if $\frac{S_n}{\sigma_n}$ converges, then $\frac{1}{\sigma_n}\sum_{i\in D_n}\left[\overline{X_i}-\mathbb{E}(\overline{X_i})\right]$ converge to a same distribution. Since Lemma A2 also implies $\mathbb{E}(\overline{X}_i)\to\mathbb{E}(X)=0$ as $N\to\infty$, it is sufficient to show $\frac{1}{\sigma_n}\sum_{i\in D_n}\overline{X_i}$ converges to a normal distribution as $n\to\infty$. In the rest of this proof, we use S_n to denote the sum of the truncated X_i's.

Condition (H1) implies $\alpha_{k,l}(m)=o(m^{-d})$ for all $k+l\leq 4$. Guyon [28] and Bolthausen [7] claimed the existence of a sub-sequence m_n. Then we have

$$\alpha_{k,l}(m_n)|D_n|^{\frac{1}{2}}\to 0,\qquad \alpha_{1,\infty}(m_n)|D_n|^{\frac{1}{2}}\to 0,\qquad \frac{|D_n|^{\frac{1}{2}}}{m_n^d}\to\infty,\qquad \text{as } n\to\infty. \tag{13}$$

Now S_n can be separated into two parts by the conditions of $d(i,j) \le m_n$ or $d(i,j) > m_n$. For all $i \in \mathbb{Z}^d$, we define

$$S_{i,n} = \sum_{\substack{j \in D_n \\ d(i,j) \le m_n}} X_j, \qquad S^*_{i,n} = S_n - S_{i,n} = \sum_{\substack{j \in D_n \\ d(i,j) > m_n}} X_j.$$

$$a_n = \sum_{i \in D_n} \mathbb{E}(X_i S_{i,n}), \qquad \overline{S}_n = a_n^{-\frac{1}{2}} S_n, \qquad \overline{S}_{i,n} = a_n^{-\frac{1}{2}} S_{i,n}.$$

Then we have

$$\sigma_n^2 = a_n + \sum_{i \in D_n} \mathbb{E}(X_i S^*_{i,n}).$$

Since condition (H3) implies

$$\left| \sum_{i \in D_n} \mathbb{E}(X_i S^*_{i,n}) \right| = o(|D_n|), \quad \text{as } n \to \infty,$$

added by condition (H4), we have

$$a_n = \sigma_n^2 (1 - o(1)).$$

Thus the proof is reduced to showing that $\overline{S}_n$ is asymptotically normal.

Now the first result in this proof and condition (H4) gives

$$\sup_n \mathbb{E}(\overline{S}_n^2) = \sup_n \frac{1}{a_n} \sum_{i,j \in D_n} Cov(X_i, X_j) < \infty,$$

which satisfies the first condition, (9), of Theorem A5.

Secondly, we check condition (11) of Theorem A5. Let

$$A_1 = i\lambda e^{i\lambda \overline{S}_n} \left(1 - \frac{1}{a_n} \sum_{j \in D_n} X_j S_{j,n} \right),$$

$$A_2 = \frac{1}{\sqrt{a_n}} e^{i\lambda \overline{S}_n} \sum_{j \in D_n} X_j \left(1 - i\lambda \overline{S}_{j,n} - e^{-i\lambda \overline{S}_{j,n}} \right),$$

$$A_3 = \frac{1}{\sqrt{a_n}} \sum_{j \in D_n} X_j e^{i\lambda(\overline{S}_n - \overline{S}_{j,n})}.$$

We have

$$(i\lambda - \overline{S}_n) e^{i\lambda \overline{S}_n} = A_1 - A_2 - A_3.$$

Because

$$\mathbb{E}\left((i\lambda - \overline{S}_n)e^{i\lambda\overline{S}_n}\right) \leq |\mathbb{E}(A_1)|+|\mathbb{E}(A_2)|+|\mathbb{E}(A_3)|\leq \mathbb{E}|A_1|+\mathbb{E}|A_2|+\mathbb{E}|A_3|.$$

The remainder of this proof is to show three asymptotic behaviour of A_1, A_2 and A_3.

For A_2, Taylor's Theorem gives

$$|1 - i\lambda\overline{S}_{j,n} - e^{i\lambda\overline{S}_{j,n}}|C \leq \lambda^2\overline{S}_{j,n}^2,$$

where C is a constant. Then, by using the facts of $a_n = \sigma_n^2(1-o(1))$ and the bounded $|D_n|/\sigma_n^2$, we have

$$\begin{aligned}
\mathbb{E}|A_2| &\leq \frac{\|X\|_\infty}{\sqrt{a_n}}\mathbb{E}\left|\sum_{j\in D_n} C\lambda^2\bar{S}_{j,n}^2\right| \\
&\leq \frac{C_1}{(\sqrt{a_n}^3)}\sum_{j\in D_n}\mathbb{E}\left|S_{j,n}^2\right| \\
&\leq \frac{C_1}{(\sqrt{a_n}^3)}\sum_{j\in D_n}\sum_{\substack{i,i'\in D_n \\ d(i,j)\leq m_n \\ d(i',j)\leq m_n}}|Cov(X_i,X_j)| \\
&\leq \frac{C_2|D_n|m_n^d}{(\sqrt{a_n})^3}\sum_{\tilde{m}\geq 0}\tilde{m}^{d-1}\alpha_{1,1}(\tilde{m}) \\
&= O\left(\frac{|D_n|}{a_n}\frac{m_n^d}{\sqrt{a_n}}\right) \\
&= O\left(\frac{|D_n|^{\frac{1}{2}}}{\sigma_n}\frac{m_n^d}{|D_n|^{\frac{1}{2}}}\right) \to 0,
\end{aligned}$$

where C_1 and C_2 are constants, the last equality uses the second limitation in (13).

For A_3, the truncated setting, property of (13) and Lemma A5 yield

$$\begin{aligned}
|\mathbb{E}A_3| &\leq \frac{1}{\sqrt{a_n}}\sum_{i,j\in D_n}\left|Cov\left(X_j, e^{i\lambda(\bar{S}_n-\bar{S}_{i,n})}\right)\right| \\
&\leq \frac{1}{\sqrt{a_n}}\sum_{i\in D_n}\alpha_{1,\infty}(m_n)\|X\|_\infty\|e^{i\lambda(\bar{S}_n-\bar{S}_{i,n})}\|_\infty \\
&\leq O\left(\frac{|D_n|}{\sqrt{a_n}}\alpha_{1,\infty}(m_n)\right) \\
&= O\left(\frac{|D_n|^{\frac{1}{2}}}{\sqrt{a_n}}|D_n|^{\frac{1}{2}}\alpha_{1,\infty}(m_n)\right) \\
&= O\left(|D_n|^{\frac{1}{2}}\alpha_{1,\infty}(m_n)\right) \to 0.
\end{aligned}$$

For A_1, we note that

$$\mathbb{E}|A_1|^2 = \frac{\lambda^2}{a_n} \sum_{\substack{j,j',l,l' \in D_n \\ d(j,l) \le m_n \\ d(j',l') \le m_n}} Cov(X_j X_l, X_{j'} X_{l'})$$

and

$$|Cov(X_j X_l, X_{j'} X_{l'})| \le |\mathbb{E}(X_j X_l X_{j'} X_{l'})| + |\mathbb{E}(X_j X_l)||\mathbb{E}(X_{j'} X_{l'})|.$$

Now for fixed $j \in D_n$, we estimate the number of $|Cov(X_j X_l, X_{j'} X_{l'})|$. Let $dist(\{j\}, \{j'\}) = k$.

Caes 1: $k < 3m_n$. Let $\tilde{m} = dist(\{j\}, \{j', l, l'\})$. Then there exists a constant C such that the number of $|Cov(X_j X_l, X_{j'} X_{l'})|$ no more than $C m_n^{2d} \tilde{m}^{d-1}$. Therefore

$$\begin{aligned}
\mathbb{E}|A_1|^2 &\le \frac{\lambda^2}{a_n} \sum_{j \in D_n} \sum_{\substack{j',l,l' \in D_n \\ d(j,l) \le m_n \\ d(j',l') \le m_n}} |Cov(X_j X_l, X_{j'} X_{l'})| \\
&\le \frac{\lambda^2}{a_n} \sum_{j \in D_n} \sum_{\substack{j',l,l' \in D_n \\ d(j,l) \le m_n \\ d(j',l') \le m_n}} \Big(|Cov(X_j, X_l X_{j'} X_{l'})| + |Cov(X_j, X_l)| \|X\|_\infty^2 \Big) \\
&\le \frac{C}{a_n} |D_n| m_n^{2d} \sum_{\tilde{m} \ge 0} \tilde{m}^{d-1} \alpha_{1,3}(\tilde{m}) \\
&= O\left(\frac{|D_n|^2}{a_n^2} \frac{m_n^{2d}}{|D_n|} \right) \to 0,
\end{aligned}$$

where C is a new constant.

Case 2: $k \ge 3m_n$. Then we have

$$\min \Big\{ dist(\{j\}, \{j'\}), dist(\{j\}, \{l'\}), dist(\{l\}, \{j'\}), dist(\{l\}, \{l'\}) \Big\} \ge k - 2m_n$$

and there exists a new constant C, such that the number of $|Cov(X_j X_l, X_{j'} X_{l'})|$ is no more than $C m_n^{2d} k^{d-1}$. Let $p = k - 2m_n$, then we have

$$\begin{aligned}
\mathbb{E}|A_1|^2 &\le \frac{\lambda^2}{a_n} \sum_{j \in D_n} \sum_{\substack{j',l,l' \in D_n \\ d(j,l) \le m_n \\ d(j',l') \le m_n}} |Cov(X_j X_l, X_{j'} X_{l'})| \\
&\le \frac{C}{a_n} |D_n| m_n^{2d} \sum_{k \ge 0} k^{d-1} \alpha_{2,2}(k - 2m_n) \\
&\le \frac{C}{a_n} |D_n| m_n^{2d} \sum_{p \ge m_n} (3p)^{d-1} \alpha_{2,2}(p)
\end{aligned}$$

$$= \quad O\left(\frac{|D_n|^2}{a_n^2}\frac{m_n^{2d}}{|D_n|}\right) \to 0,$$

where C is a new constant.

Hence, $\mathbb{E}\left((i\lambda - \overline{S}_n)e^{i\lambda\overline{S}_n}\right) \to 0$. Theorem A5 completes this proof.

A8: Isserlis' Theorem

Isserlis' Theorem is first introduced by Leon Isserlis in [35], 1918. This theorem provides a formula for calculating the product–moment of Gaussian random variables.

Theorem A6. (Refer to page 44 of [25] and equation (6) in [35]) *Let $n, m, k \in \mathbb{Z}^+$, $X_1, \cdots, X_n$ be centred Guassian random variables, then we have*

$$\mathbb{E}(X_{i_1}X_{i_2}\cdots X_{i_n}) = \begin{cases} 0, & \text{if } n = 2m+1, \\ \sum\prod\limits_{(i,j)} \mathbb{E}(X_i, X_j), & \text{if } n = 2m, \end{cases}$$

where $i_k \in \{1, 2, \cdots, n\}$ and $1 \le k \le n$, $\sum\prod_{(i,j)}$ means the summation of products of all (possible) partitions of $\{i_1, i_2, \cdots, i_{2m}\}$ into pairs.

Let $n = 2m$ be an even number. For the second equality in this theorem, there are m factors in each term, with the form $\mathbb{E}(X_i, X_j)$, and $\frac{1}{m}\binom{n}{2}\binom{n-2}{2}\cdots\binom{4}{2}\binom{2}{2}$ terms in total.

Bibliography

[1] Donald W. K. Andrews and David Pollard. A functional central limit theorem for strong mixing stochastic processes. Technical report, Cowles Foundation for Research in Economics, Yale University, 1990.

[2] A. K. Basu and C. C. Y. Dorea. On functional central limit theorem for stationary martingale random fields. *Acta Mathematica Hungarica*, 33(3-4):307–316, 1979.

[3] Serge Bernstein. Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes. *Mathematische Annalen*, 97(1):1–59, 1927.

[4] Patrick Billingsley. *Convergence of probability measures.* John Wiley & Sons, 1968.

[5] Patrick Billingsley. *Probability and Measure.* John Wiley & Sons, New York, 2012.

[6] Johnny Blair, Ronald F. Czaja, and Edward A. Blair. *Designing Surveys: A Guide to Decisions and Procedures.* 3rd edition. Sage Publications, 2014.

[7] Erwin Bolthausen. On the central limit theorem for stationary mixing random fields. *The Annals of Probability*, 10(4):1047–1050, 1982.

[8] Richard C Bradley. Basic properties of strong mixing conditions. Technical report, North Carolina Univ. at Chapel Hill Center for Stochastic Processes, 1985.

[9] Richard C Bradley. A caution on mixing conditions for random fields. *Statistics & Probability Letters*, 8(5):489–491, 1989.

[10] Richard C Bradley. On quantiles and the central limit question for strongly mixing sequences. *Journal of Theoretical Probability*, 10(2):507–555, 1997.

[11] Richard C Bradley. Basic properties of strong mixing conditions. a survey and some open questions. *Probability surveys*, 2:107–144, 2005.

[12] Richard C Bradley and Magda Peligrad. Invariance principles under a two-part mixing assumption. *Stochastic processes and their applications*, 22(2):271–289, 1986.

[13] Edward Carlstein. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Annals of Statistics*, 14(3):1171–1179, 1986.

[14] Ray L. Chambers and Chris J. Skinner. *Analysis of survey data*. John Wiley & Sons, 2003.

[15] Yuan Shih Chow and Henry Teicher. *Probability theory: independence, interchangeability, martingales*. Springer Science & Business Media, 2012.

[16] Kai Lai Chung. *A course in probability theory*. Academic Press, 2001.

[17] Guy Cohen. A CLT for multi-dimensional martingale differences in a lexicographic order. *Stochastic Processes and their Applications*, 126(5):1503–1510, 2016.

[18] Chandrakant M. Deo. A functional central limit theorem for stationary random fields. *The Annals of Probability*, 3(4):708–715, 1975.

[19] Paul Doukhan. *Mixing: properties and examples*, volume 85 of Lecture Notes in Statistics. Springer–Verlag, 1994.

[20] Paul Doukhan, Pascal Massart, and Emmanuel Rio. The functional central limit theorem for strongly mixing processes. *Annales de L'IHP Probabilités et Statistiques*, 30(1):63–82, 1994.

[21] Magnus Ekström. A general central limit theorem for strong mixing sequences. *Statistics & Probability Letters*, 94:236–238, 2014.

[22] Magnus Ekström and Sara Sjöstedt de Luna. Subsampling methods to estimate the variance of sample means based on nonstationary spatial data with varying expected values. *Journal of the American Statistical Association*, 99(465):82–95, 2004.

[23] István Fazekas. Limit theorems for the empirical distribution function in the spatial case. *Statistics & Probability Letters*, 62(3):251–262, 2003.

[24] Centers for Disease Control and Prevention (CDC). National diabetes statistics report, 2017. *Atlanta, GA: Centers for Disease Control and Prevention, US Dept of Health and Human Services*, 2017.

[25] Uriel Frisch. *Turbulence: the legacy of AN Kolmogorov*. Cambridge University Press, 1995.

[26] Wayne A. Fuller. *Sampling statistics*. John Wiley & Sons, 2011.

[27] Michael T. Gastner and Mark E. J. Newman. Diffusion-based method for producing density-equalizing maps. *Proceedings of the National Academy of Sciences of the United States of America*, 101(20):7499–7504, 2004.

[28] Xavier Guyon. *Random fields on a network: modeling, statistics, and applications*. Springer–Verlag, 1995.

[29] Godfrey Harold Hardy, John Edensor Littlewood, and George Pólya. *Inequalities*. Cambridge University Press, 1952.

[30] Norbert Herrndorf. Stationary strongly mixing sequences not satisfying the central limit theorem. *The Annals of Probability*, 11(3):809–813, 1983.

[31] Norbert Herrndorf. A functional central limit theorem for ρ-mixing sequences. *Journal of multivariate analysis*, 15(1):141–146, 1984.

[32] Norbert Herrndorf. A functional central limit theorem for strongly mixing sequences of random variables. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete (Probability Theory and Related Fields)*, 69(4):541–550, 1985.

[33] Eberhard Hopf. *Ergodentheorie*. Springer, 1937.

[34] I. A. Ibragimov and Yu. V. Linnik. *Independent and stationary sequences of random variables*. Wolters-Noordhoff Publishing Groningen, The Netherlands, 1971.

[35] Leon Isserlis. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, 12(1/2):134–139, 1918.

[36] Cliffoed L. Johnson, Sylvia M. Dohrmann, Vicki L. Burt, and Leyla K. Mohadjer. *National Health and Nutrition Examination Survey: Sample Design, 2011-2014*. Hyattsville, Maryland, DHHS Publication No. 2014–1362, March 2014.

[37] Gary G. Koch and Dennis B. Gillings. Inference, Design-Based vs. Model-Based. In *Encyclopedia of statistical sciences* (S. Kotz, C. B. Read, N. Balakrishnan and B. Vidakovic, eds.), 2nd ed. Wiley, Hoboken, NJ, 2005.

[38] Andrei Nikolaevich Kolmogorov and Yu A Rozanov. On strong mixing conditions for stationary gaussian processes. *Theory of Probability & Its Applications*, 5(2):204–208, 1960.

[39] Soumendra Nath Lahiri. *Resampling methods for dependent data*. Springer Science & Business Media, 2013.

[40] Georg Lindgren. Lectures on stationary stochastic processes, *Lecture Notes, Lund University*. `http://www.maths.lth.se/matstat/staff/georg/Publications/lecture2006.pdf`, 2006.

[41] Georg Lindgren. *Stationary stochastic processes: theory and applications*. CRC Press, Taylor & Francis Group, 2013.

[42] Georg Lindgren, Holger Rootzén, and Maria Sandsten. *Stationary stochastic processes for scientists and engineers*. CRC press, Taylor & Francis Group, 2014.

[43] Michel Loeve. *Probability Theory I*, volume 46 of Graduate Texts in Mathematics. 4rd edition. Springer–Verlag, 1977.

[44] Sharon Lohr. *Sampling: design and analysis.* Nelson Education, 2009.

[45] Thomas Lumley and Alastair Scott. Partial likelihood ratio tests for the cox model under complex sampling. *Statistics in Medicine*, 32(1):110–123, 2013.

[46] Thomas Lumley and Alastair Scott. Fitting regression models to survey data. *Statistical Science*, 32(2):265–278, 2017.

[47] Mohamed El Machkouri. Asymptotic normality of the parzen–rosenblatt density estimator for strongly mixing random fields. *Statistical Inference for Stochastic Processes*, 14(1):73–84, 2011.

[48] Florence Merlevàde and Magda Peligrad. The functional central limit theorem under the strong mixing condition. *The Annals of Probability*, 28(3):1336–1352, 2000.

[49] Mark Newman. Cart: Computer software for making cartograms. `http://www-personal.umich.edu/~mejn/cart/doc/`, 2006. [Online: accessed 16-March-2016].

[50] Daniel J Nordman and Soumendra N Lahiri. On optimal spatial subsample size for variance estimation. *The Annals of Statistics*, 32(5):1981–2027, 2004.

[51] Douglas Nychka, Reinhard Furrer, John Paige, and Stephan Sain. Package "fields". `https://cran.r-project.org/web/packages/fields/fields.pdf`, 2017. [Online: accessed 25-September-2017].

[52] Hiroshi Oodaira and Ken-ichi Yoshihara. Functional central limit theorems for strictly stationary processes satisfyinc the strong mixing condition. In *Kodai Mathematical Seminar Reports*, volume 24, pages 259–269, 1972.

[53] D. Pfefferman and C.R. Rao. (Edited) *Sample Surveys: Design, Methods and Applications,* Handbook of statistics, vol. 29A. Amsterdam: Elsevier, 2009.

[54] Zuzana Prášková and Pranab Kumar Sen. *Asymptotics in finite population sampling* Handbook of Statistics, 29B, pages 489–522. Elsevier, 2009.

[55] B. L. S. Prakasa Rao. On mixing for flows of σ-algebras. *Sankhyā: The Indian Journal of Statistics, Series A*, 52(1):1–15, 1990.

[56] B. L. S. Prakasa Rao. Conditional independence, conditional mixing and conditional association. *Annals of the Institute of Statistical Mathematics*, 61(2):441–460, 2009.

[57] Emmanuel Rio. Inequalities and limit theorems for weakly dependent sequences. 2013, pp.170. <cel-00867106>.

[58] Emmanuel Rio. Covariance inequalities for strongly mixing processes. *Annales de L'IHP Probabilités et Statistiques*, 29(4):587–597, 1993.

[59] Emmanuel Rio. The functional law of the iterated logarithm for stationary strongly mixing sequences. *The Annals of Probability*, 23(3):1188–1203, 1995.

[60] Murray Rosenblatt. A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences*, 42(1):43–47, 1956.

[61] Jeffrey Seth Rosenthal. *A first look at rigorous probability theory.* 2nd ed., World Scientific, 2006.

[62] Yu A Rozanov. A central limit theorem for additive random functions. *Theory of Probability & Its Applications*, 5(2):221–223, 1960.

[63] Havard Rue and Leonhard Held. *Gaussian Markov random fields: theory and applications.* Chapman and Hall/CRC, 2005.

[64] Carl-Erik Särndal, Ib Thomsen, Jan M Hoem, DV Lindley, O Barndorff-Nielsen, and Tore Dalenius. Design-based and model-based inference in survey sampling [with discussion and reply]. *Scandinavian Journal of Statistics*, 5(1):27–52, 1978.

[65] Fritz Scheuren. What is a survey? `http://isites.harvard.edu/fs/docs/icb.topic549650.files/What_is_a_Survey.pdf`, 2004. [Online: accessed 13-July-2017].

[66] Michael Sherman. Variance estimation for statistics computed from spatial lattice data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 509–523, 1996.

[67] Albert N. Shiryaev. *Probability*, volume 95 of Graduate Texts in Mathematics. Springer-Verlag, New York, 1996.

[68] Charles Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory*, pages 583–602, 1972.

[69] Sonya K. Sterba. Alternative model-based and design-based frameworks for inference from samples to populations: From polarization to integration. *Multivariate behavioral research*, 44(6):711–740, 2009.

[70] A.W. Van der Vaart. *Asymptotic statistics.* New York: Cambridge University Press, 1998.

[71] VA Volkonskii and Yu A Rozanov. Some limit theorems for random functions. i. *Theory of Probability & Its Applications*, 4(2):178–197, 1959.

[72] VA Volkonskii and Yu A Rozanov. Some limit theorems for random functions. ii. *Theory of Probability & Its Applications*, 6(2):186–198, 1961.

[73] C.S. Withers. Conditions for linear processes to be strong-mixing. *Probability Theory and Related Fields*, 57(4):477–480, 1981.

[74] Eugene Wong and Moshe Zakai. Martingales and stochastic integrals for processes with a multi-dimensional parameter. *Probability Theory and Related Fields*, 29(2):109–122, 1974.

[75] George Zipf, Michele Chiappa, Kathryn S. Porter, Yechiam Ostchega, Brenda G. Lewis, and Jennifer Dostal. *National Health and Nutrition Examination Survey: Plan and Operations, 1999-2010*. Hyattsville, Maryland, DHHS Publication No. 2013–1332, August 2013.

獻給我的父母

(To my parents)