Measuring Spatial Extremal Dependence

Yong Bum Cho

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2016

ABSTRACT

Measuring Spatial Extremal Dependence

Yong Bum Cho

The focus of this thesis is extremal dependence among spatial observations. In particular, this research extends the notion of the *extremogram* to the spatial process setting. Proposed by Davis and Mikosch (2009), the extremogram measures extremal dependence for a stationary time series. The versatility and flexibility of the concept made it well suited for many time series applications including from finance and environmental science.

After defining the spatial extremogram, we investigate the asymptotic properties of the empirical estimator of the spatial extremogram. To this end, two sampling scenarios are considered: 1) observations are taken on the lattice $\mathbb{Z}^d$ and 2) observations are taken on a continuous region in $\mathbb{R}^d$, in which the locations are points of a homogeneous Poisson point process. For both cases, we establish the central limit theorem for *the empirical spatial extremogram* under general mixing and dependence conditions. A high level overview is as follows. When observations are observed on a lattice, the asymptotic results generalize those obtained in Davis and Mikosch (2009). For non-lattice cases, we define a kernel estimator of the empirical spatial extremogram and establish the central limit theorem provided the bandwidth of the kernel gets smaller and the sampling region grows at proper speeds. We illustrate the performance of the empirical spatial extremogram using simulation examples, and then demonstrate the practical use of our results with a data set of rainfall in Florida and ground-level ozone data in the eastern United States.

The second part of the thesis is devoted to bootstrapping and variance estimation with a view towards constructing asymptotically correct confidence intervals. Even though the

empirical spatial extremogram is asymptotically normal, the limiting variance is intractable. We consider three approaches: for lattice data, we use the circular bootstrap adapted to spatial observations, jackknife variance estimation, and subsampling variance estimation. For data sampled according to a Poisson process, we use subsampling methods to estimate the variance of the empirical spatial extremogram. We establish the (conditional) asymptotic normality for the circular block bootstrap estimator for the spatial extremogram and show L_2 consistency of the variance estimated by jackknife and subsampling. Then, we propose a χ^2 based portmanteau style test to check the existence of extremal dependences at multiple lags. The validity of confidence intervals produced from these approaches and a χ^2 based portmanteau style test are demonstrated through simulation examples. Finally, we illustrate this methodology to two data sets. The first is the amount of rainfall over a grid of locations in northern Florida. The second is ground-level ozone in the eastern United States, which are recorded on an irregularly spaced set of stations.

Contents

List of Tables **iv**

List of Figures **v**

Acknowledgments **vii**

Chapter 1 Introduction **1**

1.1 Backgrounds 1

1.2 Regular variation 2

1.3 The extremogram 3

1.4 Modeling spatial extremes 4

1.5 Findings and Outline of Thesis 5

Chapter 2 Asymptotic properties of the empirical spatial extremogram **7**

2.1 Introduction 7

2.1.1 Definitions and notation 9

2.2 Asymptotics of the ESE 10

2.2.1 Random fields on a lattice 10

2.2.2 Random fields on $\mathbb{R}^d$ 12

2.3 Examples 15

2.3.1 Max Moving Average (MMA) 17

2.3.2 Brown-Resnick process 20

2.3.3 Simulations 26

2.4 Application 29
2.4.1 Lattice case: rainfall in a region in Florida 29
2.4.2 Non-lattice case: ground-level ozone in the eastern United States . . 32
2.5 Conclusion 36
2.6 Appendix: Proofs 36
2.6.1 Appendix A: Proof of Theorem 1 37
2.6.2 Appendix B: Proof of Theorem 3 43
2.6.3 Appendix C: Proof of Example 3 55
2.6.4 Appendix D: the condition (2.16) with delta = 1 58

Chapter 3 Resampling methods for the empirical spatial extremogram 67

3.1 Introduction 67
3.2 Bootstrapped ESE and the variance estimation of the ESE 70
3.2.1 Random fields on the lattice 70
3.2.2 Random fields on $\mathbb{R}^d$: subsampling variance estimator 77
3.3 The bias corrected confidence intervals and comparison of three approaches 80
3.3.1 The bias corrected confidence intervals 81
3.3.2 Comparison of three approaches 82
3.4 Examples 82
3.4.1 Max-Moving Average (MMA) 83
3.4.2 Brown-Resnick process 87
3.5 Portmanteau tests 91
3.6 Application 93
3.6.1 The lattice case: rainfall in a region in Florida 94
3.6.2 Non-lattice case: ground-level ozone in the eastern United States . . 97
3.7 Conclusion 99
3.8 Appendix: Proofs 100
3.8.1 Appendix A: Proof of Theorem 5 100
3.8.2 Appendix B: Proof of Theorem 6 109
3.8.3 Appendix C: Proof of Theorem 7. 115

Chapter 4 Conclusion and Future Directions **122**

Bibliography **124**

List of Tables

3.1 Estimation results from the CBB, subsampling ($s_n = 10$) and jackknife ($j_k = 10$) for the MMA (1) with $a_m = .98$ quantile 85

3.2 Estimated variance - covariance matrix for the MMA (1) using subsampling variance estimation 85

3.3 95% confidence intervals from the CBB, subsampling ($s_n = 10$) and jackknife ($j_k = 10$) for the MMA (1) with $a_m = .98$ quantile 85

3.4 Estimation results from the CBB, subsampling ($s_n = 10$) and jackknife ($j_k = 10$) for the BR process width $a_m = .98$ upper quantile 90

3.5 Estimated variance - covariance matrix for the Brown-Resnick process using subsampling variance estimation (scaled by 100). 90

3.6 95% confidence intervals from the CBB, subsampling ($s_n = 10$) and jackknife ($j_k = 10$) for the BR process with $a_m = .98$ quantile 91

3.7 The portmanteau test results for 20 hypothesis tests for the MMA(1). . . . 93

3.8 The test statistics (3.38) for $a_m = 0.70$ (top), 0.75 (middle) and 0.80 (bottom) upper quantile. 96

3.9 The cases that the null hypothesis is rejected. 96

3.10 The deviation of the ESE from subsample and the original data per different subsampling ratio. 99

List of Figures

2.1 The extremogram and the ESE for the MMA(1) measuring the right tail dependence, with $a_m = .97$ (left) and $a_m = (.90,.92,.95,.97)$ quantile (right). 18

2.2 The extremogram and the ESE for the Brown-Resnick process on lattice (left) and non-lattice (right) with two different covariance structures. 22

2.3 The distribution of the ESE for the MMA (1) and BR process on lattice (top), and the BR process on non-lattice with two bandwidth choices (bottom). . 27

2.4 The region of Florida rainfall data. 30

2.5 The locations of extremes and the ESE using the 6 year maxima of Florida rainfall data. 31

2.6 The ESE of the annual maxima of Florida rainfall from 1999 to 2004. . . . 32

2.7 The region of ozone monitoring station in the eastern United States with extreme observations defined with different thresholds of .80, .90, .95, and .97 upper quantiles. 33

2.8 QQ plot of the standard Gumbel and the empirical quantiles (dots) derived from the empirical distribution of maxima over sets with cardinality K=2. . 34

2.9 The ESE of the maximum ozone data for 1997 with different choices of bandwidths. 35

3.1 The uncertainty of the pre-asymptotic extremogram and empirical spatial extremogram. 68

3.2 Overview of the spatial CBB in $\mathbb{Z}^2$. 71

3.3 The illustration of blocks of blocks jackknife with n = 40, M =10, L =5, b =3. 76

3.4 Visual comparison of three variance estimators in d=2. 82
3.5 The 95% bootstrap and the bias corrected confidence intervals the MMA(1). 83
3.6 The 95% confidence interval from the CBB (the bias correction), jackknife, and subsampling for the MMA(1) 84
3.7 The 95% confidence interval from the CBB (the bias correction), jackknife, and subsampling for the IMMA 86
3.8 The comparison of the bootstrap and the bias-corrected confidence interval for the BR process. 88
3.9 The 95% confidence interval from the CBB (the bias correction), jackknife, and subsampling for the BR process. 89
3.10 The confidence bands for the ESE using the year of 2000 (top) and 2002 (bottom) annual maxima of Florida rainfall data with $a_m = 0.70$ (left), 0.75 (middle), 0.80 (right) upper quantiles. 94
3.11 The region of ozone monitoring station in the eastern United States and subsampling scheme illustration. 98
3.12 The ESE with the .97 upper quantile and c = 3 and 95% confidence interval from subsampling variance estimators. 99

Acknowledgments

First and foremost, I would like to express my sincere and deepest gratitude to my advisor Professor Richard A. Davis. Not only he has led me towards a better understanding of subjects and suggested better ways to deliver ideas, but also offered me many incredible opportunities such as a summer program in Technische Universität München and the Mathematics and Statistics of Quantitative Risk Management workshop at Oberwolfach. Through these opportunities, I was able to meet many renowned scholars in person and discuss many interesting problems, which broadened my knowledge. Also, I am really grateful to have opportunities to discuss many interesting topics beyond academia and experience a part of American cultures such as Thanksgiving/Christmas dinner thanks to Richard and Patti. This journey was impossible without his advice and mentoring, in and outside of the realm of academia. These experiences, guidance, and advice will never be forgotten.

Without hesitation, I would also like to thank Dr. Souvik Ghosh for his kindness and patience. I am tremendously grateful that he spent great amount of his time for helping me to formulate these problems in the thesis, guiding me to solve issues when I was unable to, and reviewing drafts for journal submissions. Through the entire process, he has encouraged me so many times in so many ways. His encouragement has made me push my boundaries of learning further. This dissertation would not have been possible without his support, extremely helpful suggestions, and kindness.

I am also deeply indebted to Professors Blanchet, Professor Lall, and Professor Ying for

agreeing to serve on my dissertation committee. Research papers by Professor Blanchet and Professor Lall, and many enlightening discussions with Professor Ying during his classes, help me to deepen my understanding in various topics related to my dissertation.

Prior to Columbia University, I was fortunate to learn from many professors at Seoul National University. I thank Professor Byeong U. Park, Jongwoo Jeon, Sangyeol Lee, and Yoon-Jae Whang for their guidance and support. They motivated me to pursue a doctoral degree at Columbia.

At Columbia University, I was very fortunate to meet many good colleagues and friends. Especially, I appreciate the advice and help from Tony Sit, Chien-Hsun Huang, Ivor Cribben, Heng Lui, Megan O'Malley, Subhankar Sadhukhan, Emilio Seijo, Gongjun Xu, Xuan Yang, Pengfei Zang, Zunyi Zhang,Christina Steinkohl, Vincenzo Ferrazzano, Vincent Dorie, Radka Pickova, Ekaterina Vinkovskaya and Phyllis Wan. The memories we have shared and the times we have spent together all along this journey will stay with me forever.

Many other friends in New York City and Seoul also have supported me in various ways. I especially thank Eunhyuk Oh from Republic of Korea Army for helping me tremendously during my first few years in New York and being a good buddy no matter what happened. I am also very grateful for Youngbok Lee ad Jaegun Jason Kim for being my mentors and big brothers. I thank for your supports when I went through a personally hard time. I also thank Wonki Hong, Joonhan Lee, and Hyungrae Yea, who always inspire me. Also, Edward Kang and Sangwon Park from Seoul National University Alumni Association are great mentors to have. Thanks to all of you, my life in New York City extremely become enjoyable and unforgettable. All of you have been very important part of my life over the years.

Most importantly, I would like to thank my family for their unconditional and endless support and encouragement. Without them, I would not have been able to achieve this goal and overcome many difficulties during the journey.

Dedicated to my family

Chapter 1

Introduction

1.1 Backgrounds

There has been increasing interest in studying the behavior of extreme observations since extreme events impact our lives in so many dimensions. Events like large swings in financial markets or extreme weather conditions such as floods and hurricanes can cause not only direct costs such as large financial/property losses and numerous casualties, but also indirect costs like increased insurance premiums, food price, and maintenance costs. Extreme events often appear to cluster and that has resulted in a growing interest in studying *extremal dependence* in many areas including finance, insurance, and atmospheric science.

One such measure of extremal dependence for a stationary time series is the *extremal index* $\theta \in (0, 1]$ proposed by Leadbetter (1983). The extremal index is a measure of clustering in the extremes corresponding to the reciprocal of the mean cluster size of extremes (i.e., the average size of neighboring excesses over a high threshold). For example, $\theta = 1$ implies that no clustering among extremes while $\theta < 1$ indicates clustering. Davis and Mikosch (2009) provided examples with different extremal dependence behaviors: the generalized autoregressive conditional heteroscedastic (GARCH) process and the heavy-tailed stochastic volatility (SV) process. Both processes posses similar *stylized features* (e.g., heavy-tailed marginal distributions, dependent but uncorrelated, volatility clustering), but a GARCH process has an extremal index $\theta < 1$, while SV process has $\theta = 1$. This suggests that

identifying extremal dependence can be a critical step in the model selection phase.

Another commonly used measure of extremal dependence for random variables is the (upper) tail dependence coefficient. For a 2-dimensional vector (X, Y) with $X \stackrel{d}{=} Y$, the tail dependence coefficient is defined as

$$\lambda(X, Y) = \lim_{x \to \infty} P(X > x | Y > x)$$

provided the limit exists. The tail dependence coefficient $\lambda \in [0, 1]$ measures the strength of dependence: $\lambda = 0$ if X and Y are independent or asymptotically independent and $\lambda = 1$ if X and Y are dependent or asymptotically dependent. The concept can be extended to a stationary time series $\{X_t\}$. In this case, the h-lag tail dependence coefficient would be $\lambda(X_0, X_h)$ that gives more information on the serial extremal dependence. This turns out to be a special case of the extremogram.

1.2 Regular variation

In this thesis, we focus on strictly stationary random fields whose finite-dimensional distributions have power law like tails. To be specific, it is assumed that the finite-dimensional distributions have regularly varying distributions with a tail index $\alpha > 0$. More formally, let $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in I\}$ be a k-dimensional strictly stationary random process where I is either $\mathbb{R}^d$ or $\mathbb{Z}^d$. For $H = \{\boldsymbol{h}_1, \ldots, \boldsymbol{h}_t\} \subset I$, $t \geq 1$, we use X_H to denote $(X_{\boldsymbol{h}_1}, \ldots, X_{\boldsymbol{h}_t})$. The random field is said to be *regularly varying* with *index* $\alpha > 0$ if for any H, the radial part $\|X_H\|$ satisfies for all $y > 0$

$$\frac{P\big(\|X_H\| > yx\big)}{P\big(\|X_H\| > x\big)} \to y^{-\alpha} \text{ as } x \to \infty, \tag{C1}$$

and the angular part $\frac{X_H}{\|X_H\|}$ is asymptotically independent of the radial part $\|X_H\|$ for large values of $\|X_H\|$, i.e., there exists a random vector $\Theta_H \in \mathbb{S}^{tk-1}$, the unit sphere in $\mathbb{R}^{tk}$ with respect to $\|\cdot\|$, such that

$$P\left(\frac{X_H}{\|X_H\|} \in \cdot \,\middle|\, \|X_H\| > x\right) \xrightarrow{w} P\left(\Theta_H \in \cdot\right) \text{ as } x \to \infty, \tag{C2}$$

where $\xrightarrow{w}$ denotes weak convergence. The distribution of $P(\Theta_H \in \cdot)$ is called the *spectral measure* of X_H.

An equivalent definition of regular variation is given as follows. There exists a sequence $a_n \to \infty, \alpha > 0$ and a family of non-null Radon measures (μ_H) on the Borel σ-field of $\bar{\mathbb{R}}^{tk} \setminus \{\mathbf{0}\}$ such that $nP(a_n^{-1} X_H \in \cdot) \xrightarrow{v} \mu_H(\cdot)$ for $t \geq 1$, where the limiting measure satisfies $\mu_H(y\cdot) = y^{-\alpha}\mu_H(\cdot)$ for $y > 0$. Here, $\xrightarrow{v}$ denotes *vague* convergence. The space of non-negative and Radon measures becomes a complete separable metric space under the vague metric. See Davis and Hsing (1995) and Section 6 of Resnick (2006) for more details.

1.3 The extremogram

Davis and Mikosch (2009) proposed the *extremogram* that is a versatile tool for assessing extremal dependence in a stationary time series. Consider $(\mathbf{X}_t)$ a strictly stationary and regularly varying sequence of $\mathbb{R}^d$-valued random vectors. For sets A and B bounded away from zero and an increasing sequence a_n, the extremogram is defined as

$$\rho_{AB}(h) = \lim_{n\to\infty} P(a_n^{-1}\mathbf{X}_h \in B | a_n^{-1}\mathbf{X}_0 \in A).$$

It can be viewed as the extreme-value analog of the autocorrelation function of a stationary time series, i.e., extremal dependence is expressed as a function of lag. In addition, it allows for measuring dependence between random variables belonging in a large variety of extremal sets. Depending on choices of sets, many of the commonly used extremal dependence measures - right (or left) tail dependence or dependence among large absolute values - can be treated as a special case of the extremogram. The flexibility coming from arbitrary choices of extreme sets has made the extremogram especially well suited for time series applications such as high-frequency foreign exchange rates (Davis and Mikosch (2009)) and cross-sectional stock indices (Davis et al. (2012) and Drees et al. (2015)).

The asymptotic normality of the empirical estimate of the extremogram is established in Davis and Mikosch (2009), while the consistency of the bootstrapped empirical estimate is discussed in Davis et al. (2012).

1.4 Modeling spatial extremes

Until recently, statistical modeling of spatial extremes has been difficult due to the lack of flexible models. The main approaches that have been proposed to overcome this problem are based on latent variables, on extreme copulas, and on spatial max-stable processes. The first approach introduces a latent process, conditional on which standard extreme models are applied (Coles and Casson (1998) and Cooley et al. (2007)). The copula based approach assumes that the marginal distributions are extreme value distributions, and then specifies a copula to model the extremal behavior of the joint distributions. Such a copula is called an extremal copula and one example is the extremal t copula (Demarta and McNeil (2005)). Lastly, max-stable modeling is first suggested by de Haan (1984) and further developed by, for example, Kabluchko et al. (2009) and Davis et al. (2013a). Max-stable modeling accounts for spatial extremal depencey that is consistent with the classical extreme value theory. Applications to rainfall data can be found in Davison et al. (2012), and to snow data in Blanchet and Davison (2011).

Regarding these approaches, Davison et al. (2012) concluded that a better spatial modeling of extremes seems to be based on appropriately-chosen copula or max-stable models since latent variable modeling fits the joint distributions of extremes poorly even though it offers a better fit to marginal distributions.

Given this discussion, we have been interested in a model identification tool for spatial extremes. The empirical extremogram in Davis and Mikosch (2009) seems to be a good starting point since it is a nonparametric estimate and it is capable to distinguishing different extremal dependence among competing models for a stationary time series (i.e., GARCH and SV processes). For example, one can compare models and empirical extremograms in the model selection phase. Moreover, empirical extremograms can be applied to residuals to check model performance: if significant extremal dependences among residuals are detected, this implies that a selected model fails to capture extremal dependence.

1.5 Findings and Outline of Thesis

This thesis generalizes the extremogram from a time series to a spatial setting. An empirical extremogram is defined and its properties are studied. In particular, the central limit theorem is established in the increasing domain setting, and we suggest a way to construct a confidence interval for the extremogram, which builds on the previous studies such as Davis and Mikosch (2009), Karr (1986), Politis and Romano (1991), Politis and Romano (1993), and Politis and Sherman (2001). Through simulation examples and real data applications, the empirical spatial extremogram appears to be a useful and practical tool to explore spatial extremal dependence.

The thesis consists of two main chapters. Chapter 2 contains the results of the paper

Cho, Y., Davis, R.A., Ghosh, S., 2016. Asymptotic Properties of the Empirical Spatial Extremogram, Scandinavian Journal of Statistics. 43. 757-773.

Chapter 2 introduces the *spatial extremogram*, and then investigates the asymptotic properties of the corresponding empirical extremogram under two sampling scenarios. For strictly stationary random fields that are regularly spaced in $\mathbb{Z}^d, d > 1$, we generalize the idea in Davis and Mikosch (2009). For irregularly spaced observations in $\mathbb{R}^d, d > 1$, we show the asymptotic normality for a kernel estimator of the spatial extremogram provided that the bandwidth of the kernel and the sampling region are coordinated in the right fashion. This approach adopts the idea of kernel based covariance estimators for irregularly spaced observations in Karr (1986) and Li et al. (2008). Chapter 2 also contains details of the examples, simulation method descriptions, and proofs of the theorems.

Chapter 3 discusses the bootstrap and variance estimation procedures for the empirical spatial extremogram. Chapter 3 is based on

Cho, Y., Davis, R.A., Ghosh, S., Resampling methods for the Empirical Spatial Extremogram (in preparation).

Estimating the variance of the empirical extremogram is critical in order to make inferences about extremal dependence. Unfortunately the limiting variance of the empirical spatial extremogram is intractable. To construct credible confidence intervals, we resort to bootstrap

procedures and other variance estimation techniques applied to the spatial setting. As in Chapter 2, the results are presented for the lattice and non-lattice cases. For the lattice case, the asymptotic (conditional) normality for a bootstrapped empirical spatial extremogram is established under the circular block bootstrap adapted to spatial observations. The L_2 consistency of the variance estimated by jackknife and subsampling are also proved. For non-lattice cases, L_2 consistency of the subsampling variance estimator is obtained. These asymptotic properties are based on those established in Politis and Romano (1991), Politis and Romano (1993), and Politis and Sherman (2001). The confidence intervals derived from these approaches are compared using simulation examples. In addition, we propose a χ^2 based portmanteau test to check the existence of the extremal dependence at multiple lags and study the performance of this test with max-moving average of order 1.

Both chapters contain application sections, where the results of each chapter are applied to two data sets. The first is rainfall data in a region of Florida and the second is ground-level ozone in the eastern United States.

Chapter 2

Asymptotic properties of the empirical spatial extremogram

2.1 Introduction

Extreme events can affect our lives in many dimensions. Events like large swings in financial markets or extreme weather conditions such as floods and hurricanes can cause not only direct cost such as large financial/property losses and numerous casualties, but also indirect cost like increased insurance premium, food price, and maintenance cost. Extreme events often appear to cluster and that has resulted in a growing interest in measuring *extremal dependence* in many areas including finance, insurance, and atmospheric science.

Extremal dependence between two random vectors X and Y can be viewed as the probability that X is extreme given Y belongs to an extreme set. One of well-known summary measure of extremal dependence is *the tail dependence coefficient* defined as

$$\lambda(X, Y) = \lim_{x \to \infty} P(X > x | Y > x).$$

Davis and Mikosch (2009) proposed the *extremogram* that is a versatile tool for assessing extremal dependence in a stationary time series. The extremogram has two main features:

- It can be viewed as the extreme-value analog of the autocorrelation function of a stationary time series, i.e., extremal dependence is expressed as a function of lag.

- It allows for measuring dependence between random variables belonging in a large variety of extremal sets. Depending on choices of sets, many of the commonly used extremal dependence measures - right (or left) tail dependence or dependence among large absolute values - can be treated as a special case of the extremogram. The flexibility coming from arbitrary choices of extreme sets has made it especially well suited for time series applications such as high-frequency foreign exchange rates (Davis and Mikosch (2009)), cross-sectional stock indices (Davis et al. (2012)), and credit default swap spreads (Cont and Kan (2011)).

In this thesis, we will define the notion of the extremogram for random fields defined on $\mathbb{R}^d$ for some $d > 1$ and investigate the asymptotic properties of its corresponding empirical estimate. Let $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^d\}$ be a stationary $\mathbb{R}^k$-valued random field. For measurable sets $A, B \subset \mathbb{R}^k$ bounded away from $\mathbf{0}$, we define the *spatial extremogram* as

$$\rho_{AB}(\boldsymbol{h}) = \lim_{x\to\infty} P(X_{\boldsymbol{h}} \in xB | X_{\mathbf{0}} \in xA), \quad \boldsymbol{h} \in \mathbb{R}^d, \tag{2.1}$$

provided the limit exists. We call (2.1) the *spatial extremogram* to emphasize that it is for a random field in $\mathbb{R}^d$. If one takes $A = B = (1, \infty)$ in the $k = 1$ case, we recover the tail dependence coefficient between $X_{\boldsymbol{h}}$ and $X_{\mathbf{0}}$. For light tailed time series, such as stationary Gaussian processes, $\rho_{AB}(\boldsymbol{h}) = 0$ for $\boldsymbol{h} \neq \mathbf{0}$ in which case there is no extremal dependence. However, for heavy tailed processes in either time or space, $\rho_{AB}(\boldsymbol{h})$ is often non-zero for many lags $\boldsymbol{h} \neq \mathbf{0}$ and for most choices of sets A and B bounded away from the origin.

We will consider estimates of $\rho_{AB}(\boldsymbol{h})$ under two different sampling scenarios. In the first, observations are taken on the lattice $\mathbb{Z}^d$. Analogous to Davis and Mikosch (2009), we define the *empirical spatial extremogram* (ESE) as

$$\hat{\rho}_{AB,m}(\boldsymbol{h}) = \frac{\sum_{\boldsymbol{s},\boldsymbol{t}\in\Lambda_n, \boldsymbol{s}-\boldsymbol{t}=\boldsymbol{h}} I_{\{a_m^{-1}X_{\boldsymbol{s}}\in A, a_m^{-1}X_{\boldsymbol{t}}\in B\}}/n(\boldsymbol{h})}{\sum_{\boldsymbol{s}\in\Lambda_n} I_{\{a_m^{-1}X_{\boldsymbol{s}}\in A\}}/\#\Lambda_n}, \tag{2.2}$$

where

- $\Lambda_n = \{1, 2, \ldots, n\}^d$ is the d-dimensional cube with side length n,
- $\boldsymbol{h} \in \mathbb{Z}^d$ are observed lags in Λ_n,

- $m = m_n$ is an increasing sequence satisfying $m \to \infty$ and $m/n \to 0$ as $n \to \infty$,
- a_m is a sequence such that $P(|X| > a_m) \sim m^{-1}$,
- $n(\boldsymbol{h})$ is the number of pairs in Λ_n with lag $\boldsymbol{h}$, and
- $\#\Lambda_n$ is the cardinality of Λ_n.

The asymptotic normality of (2.2) is established under appropriate mixing/dependence conditions, which is an extension of the results found for a stationary time series in Davis and Mikosch (2009).

In the second case, the data are assumed to come from a stationary random field $X_{\boldsymbol{s}}$, where the locations $\{\boldsymbol{s}_1, ..., \boldsymbol{s}_N\}$ are assumed to be points of a homogeneous Poisson point process on $S_n \subset \mathbb{R}^d$. Adapting ideas from Karr (1986) and Li et al. (2008), we define the empirical spatial extremogram as a kernel estimator of $\rho_{AB}(\boldsymbol{h})$, in the spirit of the estimate of autocorrelation in space. Under suitable growth conditions on S_n and restrictions on the kernel function, we show that the weighted estimator of $\rho_{AB}(\boldsymbol{h})$ is consistent and asymptotically normal.

The organization of the chapter is as follows: In Section 2.2, we present the asymptotic properties of the ESE for both cases described above. Section 2.3 provides examples illustrating the results of Section 2.2 together with a simulation study demonstrating the performance of the ESE. In Section 2.4, the spatial extremogram is applied to a spatial rainfall data set in Florida and ground-level ozone data measured in the eastern United States. The proofs of all the results are in Appendix.

2.1.1 Definitions and notation

Let $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in I\}$ be a k-dimensional strictly stationary random process where I is either $\mathbb{R}^d$ or $\mathbb{Z}^d$. For $H = \{\boldsymbol{h}_1, \ldots, \boldsymbol{h}_t\} \subset I$, we use X_H to denote $(X_{\boldsymbol{h}_1}, \ldots, X_{\boldsymbol{h}_t})$. The random field is said to be *regularly varying* with *index* $\alpha > 0$ if for any H, the radial part $\|X_H\|$ satisfies for all $y > 0$

(C1)
$$\frac{P\big(\|X_H\| > yx\big)}{P\big(\|X_H\| > x\big)} \to y^{-\alpha} \text{ as } x \to \infty,$$

and the angular part $\frac{X_H}{\|X_H\|}$ is asymptotically independent of the radial part $\|X_H\|$ for large values of $\|X_H\|$, i.e., there exists a random vector $\Theta_H \in \mathbb{S}^{tk-1}$, the unit sphere in $\mathbb{R}^{tk}$ with respect to $\|\cdot\|$, such that

$$\text{(C2)} \qquad P\left(\frac{X_H}{\|X_H\|} \in \cdot \Big| \|X_H\| > x\right) \xrightarrow{w} P\left(\Theta_H \in \cdot\right) \quad \text{as } x \to \infty,$$

where $\xrightarrow{w}$ denotes weak convergence. The distribution of $P\left(\Theta_H \in \cdot\right)$ is called the *spectral measure* of X_H.

An equivalent definition of regular variation is given as follows. There exists a sequence $a_n \to \infty, \alpha > 0$ and a family of non-null Radon measures (μ_H) on the Borel σ-field of $\bar{\mathbb{R}}^{tk} \setminus \{\mathbf{0}\}$ such that $nP(a_n^{-1} X_H \in \cdot) \xrightarrow{v} \mu_H(\cdot)$ for $t \geq 1$, where the limiting measure satisfies $\mu_H(y\cdot) = y^{-\alpha}\mu_H(\cdot)$ for $y > 0$. Here, $\xrightarrow{v}$ denotes vague convergence. Under the regularly varying assumption, one can show that (2.1) is well defined. See Section 6.1 of Resnick (2006) for more details.

2.2 Asymptotics of the ESE

2.2.1 Random fields on a lattice

Let $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{Z}^d\}$ be a strictly stationary random field and suppose we have observations $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \Lambda_n = \{1, ..., n\}^d\}$. Let $d(\cdot, \cdot)$ be a metric on $\mathbb{Z}^d$. We denote the α-mixing coefficient by

$$\alpha_{j,k}(r) = \sup\left\{\alpha\big(\sigma(X_{\boldsymbol{s}}, \boldsymbol{s} \in S), \sigma(X_{\boldsymbol{s}}, \boldsymbol{s} \in T)\big) : S, T \subset \mathbb{Z}^d, \#S \leq j, \#T \leq k, d(S,T) \geq r\right\},$$

where for any two σ-fields $\mathcal{A}$ and $\mathcal{B}$,

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup\{|P(A \cap B) - P(A)P(B)| : A \in \mathcal{A}, B \in \mathcal{B}\}$$

and for any $S, T \subset \mathbb{Z}^d$, $d(S,T) = \inf\{d(\boldsymbol{s}, \boldsymbol{t}) : \boldsymbol{s} \in S, \boldsymbol{t} \in T\}$.

In order to study asymptotic properties of (2.2), we impose regularly varying and certain mixing conditions in the random field. In particular, we use the big/small block argument, thus the side length of big blocks (m_n) and the distance between big blocks (r_n) have to be coordinated in the right fashion. To be precise, we assume the following conditions.

(**M1**) Let B_γ be the ball of radius γ centered at 0, i.e., $B_\gamma = \{\boldsymbol{s} \in \mathbb{Z}^d : d(\boldsymbol{s}, \mathbf{0}) \leq \gamma\}$, and set $c = \#B_\gamma$. For a fixed γ, assume that there exist $m_n, r_n \to \infty$ with $m_n^{2+2d}/n^d \to 0$, $r_n^d/m_n \to 0$ such that

$$\lim_{k\to\infty} \limsup_{n\to\infty} m_n \sum_{\boldsymbol{l}\in\mathbb{Z}^d, k<d(\mathbf{0},\boldsymbol{l})\leq r_n} P\left(\max_{\boldsymbol{s}\in B_\gamma} |X_{\boldsymbol{s}}| > \epsilon a_m, \max_{\boldsymbol{s}'\in B_\gamma + l} |X_{\boldsymbol{s}'}| > \epsilon a_m\right) = 0, \text{ for } \forall \epsilon > 0, \tag{2.3}$$

$$\lim_{n\to\infty} m_n \sum_{\boldsymbol{l}\in\mathbb{Z}^d, r_n<d(\mathbf{0},\boldsymbol{l})} \alpha_{c,c}(d(\mathbf{0},\boldsymbol{l})) = 0, \tag{2.4}$$

$$\sum_{\boldsymbol{l}\in\mathbb{Z}^d} \alpha_{j_1,j_2}(d(\mathbf{0},\boldsymbol{l})) < \infty \quad \text{for} \quad 2c \leq j_1 + j_2 \leq 4c, \tag{2.5}$$

$$\lim_{n\to\infty} n^{d/2} m_n^{1/2} \alpha_{c,cn^d}(m_n) = 0. \tag{2.6}$$

The interpretation of conditions in (**M1**) are as follow:

- Condition (2.3) restricts the joint distributions for exceedance as two sets of points become far apart,
- Conditions (2.4) - (2.6) impose restrictions on the decaying rate of the mixing functions together with the level of the threshold specified by m_n, and
- Conditions are adapted from Bolthausen (1982) and Davis and Mikosch (2009).

As in Davis and Mikosch (2009), $\hat{\rho}_{AB,m}(\boldsymbol{h})$ is centered by the *Pre-Asymptotic* (PA) extremogram

$$\rho_{AB,m}(\boldsymbol{h}) = \frac{\tau_{AB,m}(\boldsymbol{h})}{p_m(A)}, \tag{2.7}$$

where $\tau_{AB,m}(\boldsymbol{h}) = m_n P(X_{\mathbf{0}} \in a_m A, X_{\boldsymbol{h}} \in a_m B)$ and $p_m(A) = m_n P(X_{\mathbf{0}} \in a_m A)$. Notice that (2.7) is the ratio of the expected values of the numerator and denominator in (2.2).

Theorem 1. *Suppose a strictly stationary regularly varying random field $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{Z}^d\}$ with index $\alpha > 0$ is observed on $\Lambda_n = \{1, ..., n\}^d$. For any finite $H \subset \mathbb{Z}^d$ which does not contain 0, assume (**M1**), where $B_\gamma \supseteq H$ for some γ. Then*

$$\sqrt{\frac{n^d}{m_n}}\Big[\hat{\rho}_{AB,m}(\boldsymbol{h}) - \rho_{AB,m}(\boldsymbol{h})\Big]_{\boldsymbol{h}\in H} \xrightarrow{d} N(\mathbf{0}, \Sigma).$$

where the matrix Σ in normal distribution is specified in Appendix A.

We present the proof of Theorem 1 in Appendix A. Examples of heavy-tailed processes satisfying (**M1**) are presented in Section 2.3.

Remark 1. *Theorem 1 recovers Corollary 3.4 in Davis and Mikosch (2009) when $d = 1$.*

Remark 2. *In Theorem 1, the PA extremogram $\rho_{AB,m}(\boldsymbol{h})$ is replaced by the extremogram $\rho_{AB}(\boldsymbol{h})$ if*

$$\lim_{n\to\infty} \sqrt{\frac{n^d}{m_n}} |\rho_{AB,m}(\boldsymbol{h}) - \rho_{AB}(\boldsymbol{h})| = 0, \quad for \quad \boldsymbol{h} \in H. \tag{2.8}$$

Thus, if the bias condition (2.8) is met, the asymptotic normality of the ESE is achieved by centering at its true value. Example 2 discusses a case for which (2.8) is not satisfied.

2.2.2 Random fields on $\mathbb{R}^d$

Now consider the case of a random field defined on $\mathbb{R}^d$ and the sampling locations are given by points of a Poisson process. In this case, we adopt the ideas from Karr (1986) and Li et al. (2008) and use a kernel estimate of the extremogram. For convenience, we restrict our attention to $\mathbb{R}^2$. The extension to $\mathbb{R}^d (d > 1)$ is straightforward, but notationally more complex.

Let $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^2\}$ be a stationary regularly varying random field with index $\alpha > 0$. Suppose N is a homogeneous 2-dimensional Poisson process with intensity parameter ν and is independent of X. Define

$$N^{(2)}(d\boldsymbol{s}_1, d\boldsymbol{s}_2) = N(d\boldsymbol{s}_1)N(d\boldsymbol{s}_2)I(\boldsymbol{s}_1 \neq \boldsymbol{s}_2).$$

Now consider a sequence of compact and convex sets $S_n \subset \mathbb{R}^2$ with Lebesgue measure $|S_n| \to \infty$ as $n \to \infty$. Assume that for each $\boldsymbol{y} \in \mathbb{R}^2$

$$\lim_{n\to\infty} \frac{|S_n \cap (S_n - \boldsymbol{y})|}{|S_n|} = 1, \tag{2.9}$$

where $S_n - \boldsymbol{y} = \{\boldsymbol{x} - \boldsymbol{y} : \boldsymbol{x} \in S_n\}$ and that

$$|S_n| = O(n^2), \qquad |\partial S_n| = O(n). \tag{2.10}$$

Here, ∂S_n denotes the boundary of S_n.

The spatial extremogram in (2.1) is estimated by $\hat{\rho}_{AB,m}(\boldsymbol{h}) = \hat{\tau}_{AB,m}(\boldsymbol{h})/\hat{p}_m(A)$, where

$$\hat{p}_m(A) = \frac{m_n}{\nu|S_n|}\int_{S_n} I\left(\frac{X_{\boldsymbol{s}_1}}{a_m} \in A\right) N(d\boldsymbol{s}_1), \tag{2.11}$$

and

$$\hat{\tau}_{AB,m}(\boldsymbol{h}) = \frac{m_n}{\nu^2}\frac{1}{|S_n|}\int_{S_n}\int_{S_n} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)\, I\left(\frac{X_{\boldsymbol{s}_1}}{a_m} \in A\right) I\left(\frac{X_{\boldsymbol{s}_2}}{a_m} \in B\right) N^{(2)}(d\boldsymbol{s}_1, d\boldsymbol{s}_2). \tag{2.12}$$

Note that $w_n(\cdot) = \frac{1}{\lambda_n^2} w(\frac{\cdot}{\lambda_n})$ is a sequence of weight functions, where $w(\cdot)$ on $\mathbb{R}^2$ is a positive, bounded, isotropic probability density function and λ_n is the bandwidth satisfying $\lambda_n \to 0$ and $\lambda_n^2|S_n| \to \infty$.

To establish a central limit theorem for $\hat{\rho}_{AB,m}(\boldsymbol{h})$, we derive asymptotics of the denominator $\hat{p}_m(A)$ and numerator $\hat{\tau}_{AB,m}(\boldsymbol{h})$. In order to show consistency of $\hat{p}_m(A)$, we assume the following conditions, which are the non-lattice analogs of (2.3) and (2.4).

(**M2**) There exist an increasing sequence m_n and r_n with $m_n = o(n)$ and $r_n^2 = o(m_n)$ such that

$$\lim_{k\to\infty}\limsup_{n\to\infty}\int_{B[k,r_n]} m_n P(|X_{\boldsymbol{y}}| > \epsilon a_m, |X_{\boldsymbol{0}}| > \epsilon a_m) d\boldsymbol{y} = 0, \text{ for } \forall \epsilon > 0, \tag{2.13}$$

$$\lim_{n\to\infty}\int_{\mathbb{R}^2\setminus B[0,r_n)} m_n \alpha_{1,1}(\boldsymbol{y}) d\boldsymbol{y} = 0, \tag{2.14}$$

$$\int_{\mathbb{R}^2} \tau_{AA}(\boldsymbol{y}) d\boldsymbol{y} < \infty, \tag{2.15}$$

where $B[a,b) = \{\boldsymbol{s} : a \le d(\boldsymbol{0},\boldsymbol{s}) < b, \boldsymbol{s} \in \mathbb{R}^2\}$ and $\tau_{AA}(\boldsymbol{y}) = \lim_{n\to\infty} \tau_{AA,m}(\boldsymbol{y})$.

For a central limit theorem for $\hat{\tau}_{AB,m}(\boldsymbol{h})$, the following conditions are required.

(**M3**) Consider a cube $B_n \subset S_n$ with $|B_n| = O(n^{2\alpha})$ and $|\partial B_n| = O(n^\alpha)$ for $0 < \alpha < 1$. Assume that there exist an increasing sequence m_n with $m_n = o(n^\alpha)$ and $\lambda_n^2 m_n \to 0$ such that

$$\sup_n E\left\{\sqrt{\frac{|B_n|\lambda_n^2}{m_n}}\left|\hat{\tau}_{AB,m}(\boldsymbol{h}:B_n) - E\hat{\tau}_{AB,m}(\boldsymbol{h}:B_n)\right|^{2+\delta}\right\} \le C_\delta, \quad \delta > 0,\ C_\delta < \infty \tag{2.16}$$

where $\hat{\tau}_{AB,m}(\boldsymbol{h} : B_n)$ is the quantity (2.12) with B_n instead of S_n. Further assume that

$$\int_{\mathbb{R}^2} \tau_{AB}(\boldsymbol{y})d\boldsymbol{y} < \infty \quad \text{and} \quad \int_{\mathbb{R}^2} \alpha_{2,2}(d(\mathbf{0}, \boldsymbol{y}))d\boldsymbol{y} < \infty, \tag{2.17}$$

$$\sup_l \frac{\alpha_{l,l}(\|\boldsymbol{h}\|)}{l^2} = O(\|\boldsymbol{h}\|^{-\epsilon}) \quad for\ some\ \epsilon > 0. \tag{2.18}$$

The condition (2.16) is required to apply Lyapunov's condition. By choosing n, m_n, λ_n and a appropriately, one can show (**M3**) is satisfied. See Example 3.

Lastly, the proof requires some smoothness of the random field.

Definition 1. *A stationary regularly varying random field $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^d\}$ satisfies a local uniform negligibility condition (LUNC) if for an increasing sequence a_n satisfying $P(|X| > a_n) \sim \frac{1}{n}$ and for all $\epsilon, \delta > 0$, there exists $\delta' > 0$ such that*

$$\limsup_{n\to\infty} nP\left(\sup_{\|\boldsymbol{s}\|<\delta'} \frac{|X_{\boldsymbol{s}} - X_{\mathbf{0}}|}{a_n} > \delta\right) < \epsilon. \tag{2.19}$$

Remark 3. *If the process is regularly varying in the space of continuous functions in every compact set, then LUNC is satisfied. See Hult and Lindskog (2006), Theorem 4.4.*

Proposition 2 (Hult and Lindskog (2006))**.** *Suppose the space $\boldsymbol{M}_0$ of measures on a complete separable metric space $\boldsymbol{C} = \boldsymbol{C}([0,1] : \boldsymbol{R}^d)$ of continuous functions $[0,1] \to \boldsymbol{R}^d$ with the uniform topology given by the supremum norm and $B_{x,r} = \{y \in \boldsymbol{C} : d(x,y) \le r\}$. Let $\nu, \mu \in \boldsymbol{M}_0(\boldsymbol{C})$ be nonzero and let $\{c_n\}$ be a regularly varying sequence of positive numbers. Then $c_n\nu(n\cdot) \to \mu(\cdot)$ in $\boldsymbol{M}_0(\boldsymbol{C})$ as $n \to \infty$ if and only if for each integer $k \ge 1$ and $(t_1, ..., t_k) \in [0,1]^k$*

$$c_n\nu\pi^{-1}_{t_1,...,t_k}(n\cdot) \to \mu\pi^{-1}_{t_1,...,t_k}(\cdot)$$

in $\boldsymbol{M}_0(\boldsymbol{R}^{dk})$ as $n \to \infty$, and for each $r > 0$ and each $\epsilon > 0$

$$\sup_n c_n\nu(n[\boldsymbol{C} \setminus B_{0,r}]) < \infty,$$

$$\lim_{\delta\to 0} \sup_n c_n\nu(x : w_x(\delta) \ge n\epsilon) = 0,$$

where $w_x(\delta) = \sup_{|s-t|\le\delta} |x(s) - x(t)|$.

Theorem 3. *Let $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^2\}$ be a stationary regularly varying random field with index $\alpha > 0$ satisfying LUNC. Assume N is a homogeneous 2-dimensional Poisson process with intensity parameter ν and is independent of X. Consider a sequence of compact and convex sets $S_n \subset \mathbb{R}^2$ satisfying $|S_n| \to \infty$ as $n \to \infty$. Assume conditions (**M2**) and (**M3**). Then for any finite set of non-zero lags H in $\mathbb{R}^2$,*

$$\sqrt{\frac{|S_n|\lambda_n^2}{m_n}} \left[\hat{\rho}_{AB,m}(\boldsymbol{h}) - \rho_{AB,m}(\boldsymbol{h})\right]_{\boldsymbol{h}\in H} \to N(\mathbf{0}, \Sigma), \tag{2.20}$$

where the matrix Σ is specified in the proof of Theorem 1.

We present the proof of Theorem 3 in Appendix B. As in Remark 2, $\rho_{AB,m}(\boldsymbol{h})$ can be replaced by $\rho_{AB}(\boldsymbol{h})$ if $\rho_{AB,m}(\boldsymbol{h})$ converges fast enough.

Remark 4. *In (2.20), $\rho_{AB,m}(\boldsymbol{h})$ can be replaced by $\rho_{AB}(\boldsymbol{h})$ if*

$$\lim_{n\to\infty} \sqrt{\frac{|S_n|\lambda_n^2}{m_n}} |\rho_{AB,m}(\boldsymbol{h}) - \rho_{AB}(\boldsymbol{h})| = 0 \qquad for\ \boldsymbol{h} \in H. \tag{2.21}$$

Remark 5. *Asymptotic normality of the non-parametric estimates of the extremogram happens at some bandwidth dependent rate, which is consistent to the corresponding estimate for the auto-covariance in Karr (1986).*

2.3 Examples

As an example, two max-stable processes are provided to illustrate the results of Section 2.2.

As discussed in Davison et al. (2012), for spatial extremes, the main types of models are based on latent variables, on a spatial copula, and on spatial max-stable processes. The latent variable modeling does not fit the joint distributions of extremes well since the approach postulate independence among extremes conditional on the latent process, even though it has a better marginal distribution fit. Since max-stable modeling has the potential advantage of considering for spatial extremal dependences that is consistent with the classical extreme value theory, our examples focus on max-stable process.

A max-stable process with unit Fréchet marginals $Z(\boldsymbol{s})$ is a stochastic process satisfying that for $Z(\boldsymbol{s})^i,\ i = 1, ..., n$, that are independent copies of $Z(\boldsymbol{s})$

$$\frac{1}{n} \max_{i=1,...,n} Z(\boldsymbol{s})^i \stackrel{d}{=} Z(\boldsymbol{s}).$$

For more background on max-stable processes, see de Haan (1984) and de Haan and Ferreira (2006). In order to check conditions on strong mixing coefficients, we need the result from Dombry and Eyi-Minko (2012). We start with the concept of *the extremal coefficient function.*

Definition 2 (The extremal coefficient function, Schlather and Tawn (2003))**.** *Consdier* $\{Z_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^2\}$ *a stationary max-stable spatial process with unit Fréchet marginals:*

$$P(Z_{\boldsymbol{s}} < z) = \exp\left(-\int \max_{\boldsymbol{s} \in R^2} \frac{g(\boldsymbol{x}, \boldsymbol{s})}{z} v(d\boldsymbol{x})\right)$$

where $g(\cdot, \cdot)$ *is a non-negative function such that* $\int g(\boldsymbol{x}, \boldsymbol{s}) v(d\boldsymbol{x}) = 1$ *for* $s \in \mathbb{R}^2$*. Then,*

$$P(Z_{\boldsymbol{s}} < z, Z_{\boldsymbol{s}+\boldsymbol{h}} < z) = \exp\left(-\frac{\theta(\boldsymbol{h})}{z}\right)$$

where $\theta(\boldsymbol{h}) = \int \max\{g(\boldsymbol{x}, \boldsymbol{s}), g(\boldsymbol{x}, \boldsymbol{s_h})\} v(d\boldsymbol{x})$ *is defined as the extremal coefficient function.*

The below result from Dombry and Eyi-Minko (2012) provides upper bound for α-mixing coefficient for max-stable processes.

Proposition 2.3.1 (Dombry and Eyi-Minko (2012))**.** *Suppose* $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in S\}$ *is a max-stable random field with unit Fréchet marginals. If* S_1 *and* S_2 *are finite or countable disjoint closed subsets of* S*, and* $\mathcal{S}_1$ *and* $\mathcal{S}_2$ *are* σ*-field generated by each set, then*

$$\beta(\mathcal{S}_1, \mathcal{S}_2) \leq 4 \sum_{\boldsymbol{s}_1 \in S_1} \sum_{\boldsymbol{s}_2 \in S_2} [2 - \theta(\boldsymbol{s}_1, \boldsymbol{s}_2)] \tag{2.22}$$

where $\beta(\cdot, \cdot)$ *is* $\beta-$*mixing coefficient and* $\theta(\cdot, \cdot)$ *is the extremal coefficient function. From Lemma 2, Davis et al. (2013b), the upper bound is equivalent to*

$$4 \sum_{\boldsymbol{s}_1 \in S_1} \sum_{\boldsymbol{s}_2 \in S_2} \rho_{(1,\infty)(1,\infty)}(\|\boldsymbol{s}_1 - \boldsymbol{s}_2\|).$$

Notice that (2.22) provides the upper bound for α-mixing coefficient since $2\alpha(\mathcal{S}_1, \mathcal{S}_2) \leq \beta(\mathcal{S}_1, \mathcal{S}_2)$. See Bradley (1993).

2.3.1 Max Moving Average (MMA)

Let $\{Z_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{Z}^2\}$ be an iid sequence of unit Fréchet random variables. The max-moving average (MMA) process is defined by

$$X_t = \max_{\boldsymbol{s}\in\mathbb{Z}^2} w(\boldsymbol{s}) Z_{\boldsymbol{t}-\boldsymbol{s}}, \tag{2.23}$$

where $w(\boldsymbol{s}) > 0$ and $\sum_{\boldsymbol{s}\in\mathbb{Z}^2} w(\boldsymbol{s}) < \infty$. Note that the summability of $w(\cdot)$ implies the process is well defined. Also, notice that $a_m = O(m)$ since marginal distributions are Fréchet. Consider the Euclidean metric $d(\cdot,\cdot)$ and write $\|\boldsymbol{l}\| = d(\boldsymbol{0}, \boldsymbol{l})$ for notational convenience. With $w(\boldsymbol{s}) = I(\|\boldsymbol{s}\| \le 1)$, the process (2.23) becomes the MMA(1)

$$X_{\boldsymbol{t}} = \max_{\|\boldsymbol{t}-\boldsymbol{s}\|\le 1} Z_{\boldsymbol{s}}.$$

Using $A = B = (1, \infty)$, the extremogram for the MMA(1) is

$$\rho_{AB}(h) = \lim_{n\to\infty} P(X_{\boldsymbol{h}} > a_{m_n} | X_{\boldsymbol{0}} > a_{m_n}) = \begin{cases} 1 & \text{if} \quad \|\boldsymbol{h}\| = 0, \\ 2/5 & \text{if} \quad \|\boldsymbol{h}\| = 1, \sqrt{2}, \\ 1/5 & \text{if} \quad \|\boldsymbol{h}\| = 2, \\ 0 & \text{if} \quad \|\boldsymbol{h}\| > 2. \end{cases} \tag{2.24}$$

Because the process is *2-dependent*, conditions for Theorem 1 are easily checked.

Figure 2.1 (left) shows $\rho_{AB}(\boldsymbol{h})$ and $\hat{\rho}_{AB,m}(\boldsymbol{h})$ from a realization of MMA(1) generated by *rmaxstab* in the *SpatialExtremes* package in R. We use 1600 points ($\Lambda_n = \{1, ..., 40\}^2 \in \mathbb{Z}^2$) and set $A = B = (1, \infty)$ and $a_m = .97$ quantile of the process. In the figure, the dots and the bars correspond to $\rho_{AB}(\boldsymbol{h})$ and $\hat{\rho}_{AB,m}(\boldsymbol{h})$ for observed distances in the sample. The dashed line corresponds to 0.03 ($= 1 - 0.97$) and two horizontal lines are 95% *random permutation* confidence bands to check the existence of extremal dependence. Since random permutation breaks the spatial dependence, the ESE from such data would show no extremal dependency. See Davis et al. (2012). The bands suggest

$$\rho_{(1,\infty)(1,\infty),m}(\|\boldsymbol{h}\|) = 0 \text{ for } \|\boldsymbol{h}\| > 2,$$

which is consistent with (2.24).

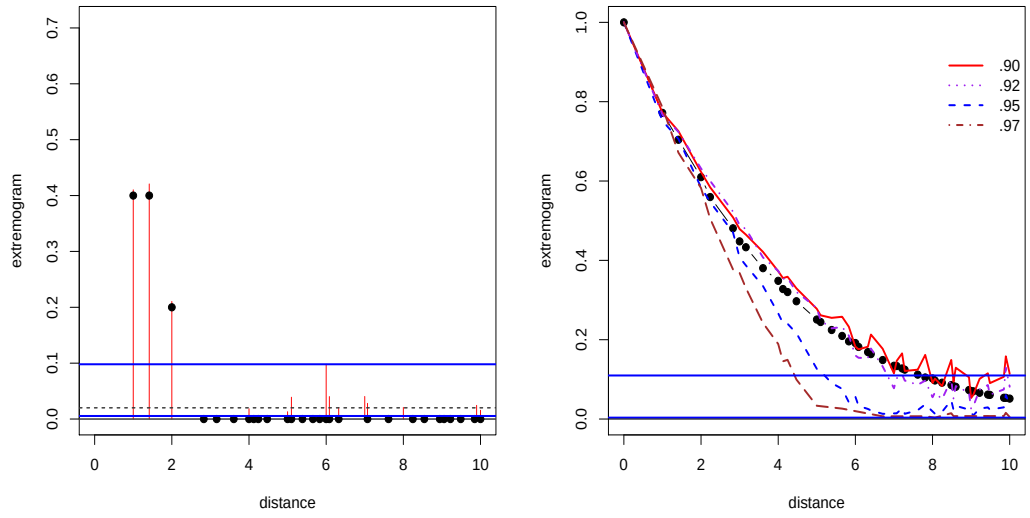

Figure 2.1: The extremogram and the ESE for the MMA(1) measuring the right tail dependence, with $a_m = .97$ (left) and $a_m = (.90,.92,.95,.97)$ quantile (right).

Now consider $w(\boldsymbol{s}) = \phi^{\|s\|}$ where $0 < \phi < 1$. Then the process (2.23) becomes

$$X_{\boldsymbol{t}} = \max_{\boldsymbol{s}\in\mathbb{Z}^2} \phi^{\|\boldsymbol{s}\|} Z_{\boldsymbol{t}-\boldsymbol{s}} \quad \text{for} \quad \sum_{\boldsymbol{l}\in\mathbb{Z}^2} \phi^{\|l\|} = \sum_{0\le\|\boldsymbol{l}\|<\infty} \phi^{\|\boldsymbol{l}\|} p(\|\boldsymbol{l}\|) < \infty, \tag{2.25}$$

where $p(\|\boldsymbol{l}\|) = \#\{\boldsymbol{s} \in \mathbb{Z}^2 : d(\boldsymbol{0}, \boldsymbol{s}) = \|\boldsymbol{l}\|\}$. Observe that the process (2.25) is istotropic and that $p(\|\boldsymbol{l}\|) = O(\|\boldsymbol{l}\|)$ from Lemma A.1 in Jenish and Prucha (2009), and

$$P(X_{\boldsymbol{t}} \le x) = \exp\Big\{-\frac{1}{x}\sum_{0\le\|\boldsymbol{l}\|<\infty} \phi^{\|\boldsymbol{l}\|} p(\|\boldsymbol{l}\|)\Big\}, \tag{2.26}$$

$$\begin{aligned} P(X_{\boldsymbol{0}} \le x, X_{\boldsymbol{h}} \le x) &= \exp\Big\{-\frac{1}{x}\sum_{\boldsymbol{s}\in\mathbb{Z}^2} \max\left(\phi^{\|\boldsymbol{s}\|}, \phi^{\|\boldsymbol{h}+\boldsymbol{s}\|}\right)\Big\} \\ &= \exp\Big\{-\frac{1}{x}\sum_{0\le\|\boldsymbol{l}\|<\infty} \phi^{\|\boldsymbol{l}\|} q_{\boldsymbol{h}}(\|\boldsymbol{l}\|)\Big\}, \end{aligned} \tag{2.27}$$

where $q_{\boldsymbol{h}}(\|\boldsymbol{l}\|) = \#\{\boldsymbol{s} \in \mathbb{Z}^2 : \min(\|\boldsymbol{s}\|, \|\boldsymbol{h} + \boldsymbol{s}\|) = \|\boldsymbol{l}\|\}$, the number of pairs with minimum distance to $\boldsymbol{0}$ or $\boldsymbol{h}$ equals $\|\boldsymbol{l}\|$. If $\|\boldsymbol{l}\| < \|\boldsymbol{h}\|/2$, then no point in $\mathbb{Z}^2$ is at the distance $\|\boldsymbol{l}\|$ from both the origin and $\boldsymbol{h}$, thus $q_{\boldsymbol{h}}(\|\boldsymbol{l}\|) = 2p(\|\boldsymbol{l}\|)$. On the other hand, $q_{\boldsymbol{h}}(\|\boldsymbol{l}\|)/p(\|\boldsymbol{l}\|) \to 1$ as $\|\boldsymbol{l}\| \to \infty$. In other words,

$$q_{\boldsymbol{h}}(\|\boldsymbol{l}\|) = 2p(\|\boldsymbol{l}\|) \text{ for } \|\boldsymbol{l}\| < \tfrac{\|\boldsymbol{h}\|}{2} \text{ and } \lim_{\|\boldsymbol{l}\|\to\infty} \frac{q_{\boldsymbol{h}}(\|\boldsymbol{l}\|)}{p(\|\boldsymbol{l}\|)} = 1.$$

Now we check the properties of (2.25). First, the max-stability can be easily checked from

$$P\left(\max_{i=1,\ldots,n} X^i_{\boldsymbol{t}}/n \le x\right) = P(X^i_{\boldsymbol{t}} \le nx)^n = \exp\left(-\frac{n}{nx}\sum_{0\le\|\boldsymbol{l}\|<\infty} \phi^{\|\boldsymbol{l}\|} p(\|\boldsymbol{l}\|)\right) = P(X_{\boldsymbol{t}} \le x).$$

For the extremogram with $A = B = (1,\infty)$, using the joint distribution in (2.27) and a Taylor series expansion,

$$\rho_{(1,\infty)(1,\infty)}(\boldsymbol{h}) = \frac{\sum_{\frac{\|\boldsymbol{h}\|}{2}\le\|\boldsymbol{l}\|<\infty} \phi^{\|\boldsymbol{l}\|}[2p(\|\boldsymbol{l}\|) - q_{\boldsymbol{h}}(\|\boldsymbol{l}\|)]}{\sum_{0\le\|\boldsymbol{l}\|<\infty} \phi^{\|\boldsymbol{l}\|} p(\|\boldsymbol{l}\|)}. \tag{2.28}$$

This is because

$$\begin{aligned} P(X_{\boldsymbol{0}} > x, X_{\boldsymbol{h}} > x) &= 1 - (P(X_{\boldsymbol{0}} \le x) + P(X_{\boldsymbol{h}} \le x) - P(X_{\boldsymbol{0}} \le x, X_{\boldsymbol{h}} \le x)) \\ &= \frac{1}{x}\sum_{0\le\|\boldsymbol{l}\|<\infty} \phi^{\|\boldsymbol{l}\|}\left[2p(\|\boldsymbol{l}\|) - q(\|\boldsymbol{l}\|)\right] + O\left(\frac{1}{x^2}\right). \end{aligned}$$

Example 1. *For the process (2.25), the conditions (2.3)-(2.6) in Theorem 1 are satisfied if* $r_n^2 = o(m_n), \log m_n = o(r_n)$ *and* $\log n = o(m_n)$.

Proof. Observe that (2.25) is isotropic. By Lemma A.1 in Jenish and Prucha (2009), $p(\|\boldsymbol{l}\|) = O(\|\boldsymbol{l}\|)$. Thus, (2.22) implies that

$$\begin{aligned} \alpha_{c,c}(k) &\le const\,(2 - \theta(0,k)) \\ &= const\left(2 - \frac{\sum_{0\le\|\boldsymbol{l}\|<\infty} \phi^{\|\boldsymbol{l}\|} q(\|\boldsymbol{l}\|)}{\sum_{0\le\|\boldsymbol{l}\|_\infty} \phi^{\|\boldsymbol{l}\|} p(\|\boldsymbol{l}\|)}\right) \\ &= const\left(\sum_{\frac{k}{2}\le\|\boldsymbol{l}\|<\infty} \phi^{\|\boldsymbol{l}\|} q(\|\boldsymbol{l}\|)\right) \\ &\le const\int_{k/2}^{\infty} j\phi^j dj \\ &= O(k\phi^k) \end{aligned}$$

for any $k > 0$. Then (2.4) is satisfied as

$$\begin{aligned} m_n \sum_{\boldsymbol{l}\in\mathbb{Z}^2, r_n\le\|\boldsymbol{l}\|} \alpha_{c,c}(\|\boldsymbol{l}\|) &= m_n \sum_{r_n\le\|\boldsymbol{l}\|} p(\|\boldsymbol{l}\|)\alpha_{c,c}(\|\boldsymbol{l}\|) \\ &= O\left(m_n\, r_n^3 \phi^{r_n}\right) \\ &\to 0 \end{aligned}$$

if $\log m_n = o(r_n)$.

Similarly, (2.5) can be shown. If $\log n = o(m_n)$, (2.6) holds since (2.22) implies

$$n^{d/2} m_n^{1/2} \alpha_{c,cn^d}(m_n) \leq const\ n^{3d/2} m_n^{1/2} m_n \phi^{m_n}.$$

Turning to (2.3), notice from (2.26) and (2.27) that

$$\begin{aligned}
&P\left(\max_{\boldsymbol{s}\in B_\gamma} |X_{\boldsymbol{s}}| > \epsilon a_m, \max_{\boldsymbol{s}'\in B_\gamma + l} |X_{\boldsymbol{s}'}| > \epsilon a_m\right) \\
&\leq \sum_{\boldsymbol{s}\in B_\gamma} \sum_{\boldsymbol{s}'\in B_\gamma + l} P\left(X_{\boldsymbol{s}} > \epsilon a_m, X_{\boldsymbol{s}'} > \epsilon a_m\right) \\
&\leq \sum_{\boldsymbol{s}\in B_\gamma} \sum_{\boldsymbol{s}'\in B_\gamma + l} \left[\frac{const}{\epsilon a_m} \sum_{\frac{\|\boldsymbol{s}-\boldsymbol{s}'\|}{2} \leq j < \infty} \phi^j j + O\left(\frac{1}{a_m^2}\right)\right] \\
&\leq const\ \frac{\phi^{\|\boldsymbol{l}\|} \|\boldsymbol{l}\|^2}{\epsilon a_m} + O\left(\frac{1}{a_m^2}\right).
\end{aligned}$$

Hence, the term in (2.3) is bounded by

$$\begin{aligned}
&\limsup_{n\to\infty} m_n \sum_{\boldsymbol{l}\in\mathbb{Z}^2, k<\|\boldsymbol{l}\|\leq r_n} \left[const\ \frac{\phi^{\|\boldsymbol{l}\|^2} \|\boldsymbol{l}\|}{\epsilon a_m} + O\left(\frac{1}{a_m^2}\right)\right] \\
&= \sum_{k<\|\boldsymbol{l}\|<\infty} const\ \phi^{\|\boldsymbol{l}\|} \|\boldsymbol{l}\|^3 + \limsup_{n\to\infty} O\left(\frac{m_n r_n^2}{a_m^2}\right),
\end{aligned}$$

where the second term is 0 as $a_m = O(m_n)$ and $r_n^2 = o(m_n)$. Letting $k \to \infty$, we obtain (2.3). □

Figure 2.1 (right) shows $\rho_{AB}(\boldsymbol{h})$ and $\hat{\rho}_{AB,m}(\boldsymbol{h})$ from a realization of the process (2.25) with $\phi = 0.5$. Here, $A = B = (1, \infty)$ and a_m = (.90,.92,.95,.97) quantiles. The dots are $\rho_{AB}(\boldsymbol{h})$ and the dashed lines are $\hat{\rho}_{AB,m}(\boldsymbol{h})$ with different a_m. The ESE with $a_m = .90$ and .92 quantiles are close to the extremogram for all observed distances while the ESE with $a_m = .95$ and .97 quantiles decay faster for the observed distances greater than 3. The two horizontal lines are 95% confidence bands based on random permutations.

2.3.2 Brown-Resnick process

We begin with the definition of the Brown-Resnick process with Fréchet marginals. Details can be found in Kabluchko et al. (2009) or Davis et al. (2013a). Consider a stationary

Gaussian process $\{Z_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^d\}$ with mean 0 and variance 1. For the correlation function $\rho(\boldsymbol{h}) = E[Z_{\boldsymbol{s}} Z_{\boldsymbol{s}+\boldsymbol{h}}]$, assume that there exist sequences $d_n \to 0$ such that

$$\log(n)\{1 - \rho(d_n \boldsymbol{h})\} \to \delta(\boldsymbol{h}) > 0, \quad \text{as} \quad n \to \infty.$$

Use $\{Z^j_{\boldsymbol{s}}, s \in \mathbb{R}^d\}, j \in 1, ..., n$, to denote independent replications of $\{Z_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^d\}$. Let $(\Gamma_i)_{i \geq 1}$ be an increasing enumeration of a unit rate Poisson process, and suppose $Y^1, Y^2, ...,$ is an iid sequence of random fields on $\mathbb{R}^d$ independent of $(\Gamma_i)_{i \geq 1}$.

Then, the random fields defined by

$$X_s(n) = \frac{1}{n} \bigvee_{i=1}^{n} -\frac{1}{\log\left(\Phi(Z^i_{\boldsymbol{s}})\right)}, \qquad \boldsymbol{s} \in \mathbb{R}^d, n \in \mathbb{N},$$

converge weakly in the space of continuous function to the stationary Brown-Resnick process

$$X_{\boldsymbol{s}} = \sup_{j \geq 1} \Gamma_j^{-1} Y^j_{\boldsymbol{s}} = \sup_{j \geq 1} \Gamma_j^{-1} \exp\{W^j_{\boldsymbol{s}} - \delta(\boldsymbol{s})\}, \qquad \boldsymbol{s} \in \mathbb{R}^d, \tag{2.29}$$

where $\{W^j_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^d\}, j \in \mathbb{N}$, are independent replications of a Gaussian random field with stationary increments, $W_{\boldsymbol{0}} = 0$ and $E[W_{\boldsymbol{s}}] = 0$ and covariance function by $\text{cov}(W_{\boldsymbol{s}_1}, W_{\boldsymbol{s}_2}) = \delta(\boldsymbol{s}_1) + \delta(\boldsymbol{s}_2) - \delta(\boldsymbol{s}_1 - \boldsymbol{s}_2)$. Here, Φ is the cumulative distribution function of $N(0, 1)$. We refer to Davis et al. (2013a).

Proposition 2.3.2 (Davis et al. (2013a))**.** *Let $\{Z_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^d\}$ be a space Gaussian process with mean 0, variance 1 and correlation funciton $\rho(\cdot)$ which is smooth around the origin. Use $Z_{\boldsymbol{s}}(i), i = 1, 2, ...$ to denote independent copies of $Z_{\boldsymbol{s}}$. Assume that there are nonnegative sequences of constants $s_n \to 0$ as $n \to \infty$ and a non-negative function δ satisfying*

$$\log n(1 - \rho(s_n(\boldsymbol{s}_1 - \boldsymbol{s}_2)) \to \delta(\boldsymbol{s}_1 - \boldsymbol{s}_2) \in (0, \infty), \qquad n \to \infty \tag{2.30}$$

for all $\boldsymbol{s}_1 \neq \boldsymbol{s}_2,\ \boldsymbol{s}_1, \boldsymbol{s}_2 \in \mathbb{R}^d$.

If there exists a metric d on $\mathbb{R}^d$ such that

$$\delta(\boldsymbol{s}_1 - \boldsymbol{s}_2) \leq (d(\boldsymbol{s}_1, \boldsymbol{s}_2))^2$$

then

$$\frac{1}{n} \bigvee_{i=1}^{n} -\frac{1}{\log\left(\Phi\left(Z_{\boldsymbol{s}_n \boldsymbol{s}}(i)\right)\right)} \tag{2.31}$$

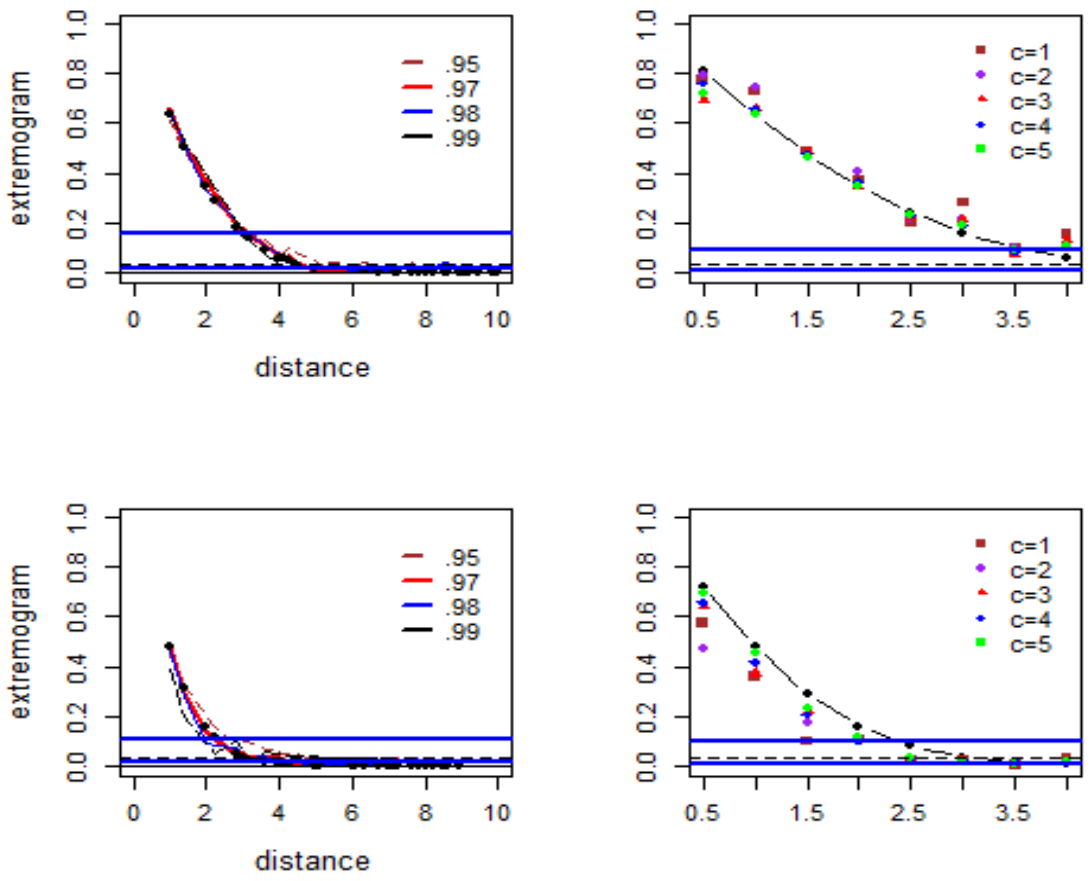

Figure 2.2: The extremogram and the ESE for the Brown-Resnick process on lattice (left) and non-lattice (right) with two different covariance structures.

converges weakly to a space max-stable process whose bivariate function is given as

$$
\begin{aligned}
&F(y_1, y_2) \\
&:= P(X_{\mathbf{0}} \le y_1, X_{\boldsymbol{h}} \le y_2) \\
&= \exp\left\{ -\frac{1}{y_1}\Phi\left(\frac{\log(y_2/y_1)}{\sqrt{2\delta(\boldsymbol{h})}} + \sqrt{\frac{\delta(\boldsymbol{h})}{2}} \right) - \frac{1}{y_2}\Phi\left(\frac{\log(y_1/y_2)}{\sqrt{2\delta(\boldsymbol{h})}} + \sqrt{\frac{\delta(\boldsymbol{h})}{2}} \right) \right\}. \quad (2.32)
\end{aligned}
$$

The extremogram for the Brown-Resnick process $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^d\}$ with $A = (c_A, \infty)$ and

$B = (c_B, \infty)$ is

$$\rho_{AB}(\boldsymbol{h}) = \bar{\Phi}_{c_A, c_B}(\delta(\boldsymbol{h})) + \frac{c_A}{c_B}\bar{\Phi}_{c_B, c_A}(\delta(\boldsymbol{h})), \tag{2.33}$$

where $\Phi_{y_1,y_2}(\delta(\boldsymbol{h})) = \Phi\left(\frac{\log(y_2/y_1)}{\sqrt{2\delta(\boldsymbol{h})}} + \sqrt{\frac{\delta(\boldsymbol{h})}{2}}\right)$. Here, Φ is the cumulative distribution function of $N(0,1)$ and $\bar{\Phi}$ represents the tail cumulative distribution function.

To see (2.33), recall (2.32) from Hüsler and Reiss (1989). As $a_m = O(m_n)$, we assume without loss of generality that $\lim_{n\to\infty} \frac{m_n}{a_m} = 1$.

$$\begin{aligned} p_m(A) = m_n\left(1 - e^{-\frac{1}{c_A a_m}}\right) &= \frac{m_n}{c_A a_m} + O(\frac{m_n}{a_m^2}) \\ &\to \frac{1}{c_A} \\ &= \mu(A) \end{aligned}$$

and

$$\begin{aligned} \tau_{AB,m}(\boldsymbol{h}) &= m_n\left[1 - e^{-\frac{1}{c_A a_m}} - e^{-\frac{1}{c_B a_m}} + F(a_m c_A, a_m c_B)\right] \\ &\to \frac{1}{c_A}\bar{\Phi}_{c_A,c_B}(\delta(\boldsymbol{h})) + \frac{1}{c_B}\bar{\Phi}_{c_B,c_A}(\delta(\boldsymbol{h})). \end{aligned} \tag{2.34}$$

Now we find the upper bound for α-mixing coefficient for the Brown-Resnick process. Similar to Lemma 2 in Davis et al. (2013b), using Proposition 2.3.1, (2.33) and the inequality $\bar{\Phi}(x) \le \frac{1}{\sqrt{2\pi}x}e^{-x^2/2}$ for $x > 0$, α-mixing coefficient of the process is bounded by

$$\begin{aligned} \alpha_{m,n}(\|\boldsymbol{h}\|) &\le 4mn \sup_{\boldsymbol{l}\ge\|\boldsymbol{h}\|} \bar{\Phi}(\sqrt{\delta(\boldsymbol{l})/2}) \\ &\le const \sup_{\boldsymbol{l}\ge\|\boldsymbol{h}\|} \frac{1}{\sqrt{\delta(\boldsymbol{l})}} e^{-\delta(\boldsymbol{l})/4}. \end{aligned} \tag{2.35}$$

In the following examples, the correlation function $\rho(\boldsymbol{h})$ of a Gaussian process $\{Z_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^d\}$ is assumed to have an expansion around zero as

$$\rho(\|\boldsymbol{h}\|) = 1 - \theta\|\boldsymbol{h}\|^\alpha + O(\|\boldsymbol{h}\|^\alpha), \quad \boldsymbol{h} \in \mathbb{R}^d, \tag{2.36}$$

where $\alpha \in (0, 2]$ and $\theta > 0$. For this choice of correlation function, we have $\delta(\boldsymbol{h}) = \theta\|\boldsymbol{h}\|^\alpha$ as in Davis et al. (2013b). Recall the condition for continuous sample path from Lindgren (2012), Section 2.2.

Proposition 2.3.3 (Lindgren (2012), Section 2.2)**.** *Any stationary process with*

$$r(t) = r(0) - const\,|t|^{\alpha} + o(|t|^{\alpha})$$

as $t \to 0$*, where* $r(\cdot)$ *is a covariance function, has continuous sample paths if* $1 < \alpha \leq 2$*. If it is a Gaussian process, it is continuous provided* $0 < \alpha \leq 2$*.*

Example 2. *Consider the Brown-Resnick process* $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{Z}^d\}$ *with* $\delta(\boldsymbol{h}) = \theta \|\boldsymbol{h}\|^{\alpha}$ *for* $0 < \alpha \leq 2$ *and* $\theta > 0$*. The conditions of Theorem 1 hold if* $\log n = o(m_n^{\alpha}), \log m_n = o(r_n^{\alpha})$ *and* $r_n^d/m_n \to 0$*. In this case, (2.8) is not satisfied for* $d > 0$*.*

Proof. From (2.35), we have $\alpha_{c,c}(\|\boldsymbol{h}\|) \leq const\ \|\boldsymbol{h}\|^{-\alpha/2} e^{-\theta\|\boldsymbol{h}\|^{\alpha}/2}$. If $\log m_n = o(r_n^{\alpha})$, (2.4) holds as

$$\begin{aligned} m_n \sum_{\boldsymbol{l} \in \mathbb{Z}^d, r_n \leq \|\boldsymbol{l}\|} \alpha_{c,c}(\|\boldsymbol{l}\|) &\leq const\ m_n \sum_{r_n \leq \|\boldsymbol{l}\|} \|\boldsymbol{l}\|^{d-1} \alpha_{c,c}(\|\boldsymbol{l}\|) \\ &\leq const\ m_n \sum_{r_n \leq \|\boldsymbol{l}\|} \|\boldsymbol{l}\|^{d-1-\alpha/2} e^{-\theta\|\boldsymbol{l}\|^{\alpha}/4} \\ &\leq const\ m_n r_n^{d-\alpha/2} e^{-\theta r_n^{\alpha}/4} \\ &\to 0. \end{aligned}$$

Similarly, (2.5) can be checked. For (2.6),

$$n^{d/2} m_n^{1/2} \alpha_{c,cn^d}(m_n) \leq const\ n^{3d/2} m_n^{(1-\alpha)/2} \exp\{-\theta m_n^{\alpha}/4\}$$

which converges to 0 if $\log n = o(m_n^{\alpha})$.

Showing (2.3) is similar to Example 1. From (2.34),

$$P\left(\max_{\boldsymbol{s} \in B_\gamma} |X_{\boldsymbol{s}}| > \epsilon a_m, \max_{\boldsymbol{s}' \in B_\gamma + \boldsymbol{l}} |X_{\boldsymbol{s}'}| > \epsilon a_m\right) \leq const\ \frac{\bar{\Phi}_{1,1}(\delta(\|\boldsymbol{l}\|))}{\epsilon a_m} + O\left(\frac{1}{a_m^2}\right).$$

Hence, the term in (2.3) is bounded by

$$\begin{aligned} &\limsup_{n \to \infty} \sum_{\boldsymbol{l} \in \mathbb{Z}^d, k < \|\boldsymbol{l}\| \leq r_n} \left[const\ m_n \frac{\bar{\Phi}_{1,1}(\delta(\|\boldsymbol{l}\|))}{\epsilon a_m} + O\left(\frac{1}{a_m^2}\right)\right] \\ \leq\ & const \sum_{k < \|\boldsymbol{l}\| < \infty} \|\boldsymbol{l}\|^{d-1} e^{-\frac{\theta\|\boldsymbol{l}\|^{\alpha}}{4}} + \limsup_{n \to \infty} O\left(\frac{r_n^d m_n}{a_m^2}\right), \end{aligned}$$

where the second term is 0 since $r_n^d = o(m_n)$. Letting $k \to \infty$, (2.3) is obtained

Lastly, we show (2.8) does not hold for any $d > 0$. From (2.34),

$$\begin{aligned} |\rho_{AB,m}(\boldsymbol{h}) - \rho_{AB}(\boldsymbol{h})| &= \frac{1+o(1)}{\mu(A)} |\tau_{AB,m}(\boldsymbol{h})\mu(A) - \tau_{AB}(\boldsymbol{h}) p_m(A)| \\ &= \frac{1+o(1)}{\mu(A)} O(m_n/a_m^2) \\ &= O(1/m_n). \end{aligned}$$

Thus, $\frac{n^d}{m_n^3} \to 0$ is required to have (2.8), which conflicts with the assumption $m_n^{2+2d} = o(n^d)$. Note that similar calculation was done in Lemma 5.4 in Davis et al. (2013b). □

In Figure 2.2 (left), we have $\rho_{AB,m}(\boldsymbol{h})$ and $\hat{\rho}_{AB,m}(\boldsymbol{h})$ from a realization of the Brown-Resnick process with $\delta(\boldsymbol{h}) = \frac{4}{9}\|\boldsymbol{h}\|^2$ (top) and $\|\boldsymbol{h}\|^2$ (bottom). We use 1600 points ($[1,40]^2 \in \mathbb{Z}^2$) to compute the extremogram with $A = B = (1,\infty)$ and $a_m = (.95, .97, .98, .99)$ upper quantiles. The extremogram is marked by dots and the ESE with different line types corresponding to various choices of a_m. From the figure, the ESE is not overly sensitive to different a_m, but $\hat{\rho}_{(1,\infty)(1,\infty),m}(\boldsymbol{h})$ with $a_m = .97$ quantile looks most robust. Also the extremal dependence seems to disappear for $\|\boldsymbol{h}\| > 4$ based on the random permutation bands (two horizontal lines).

Example 3. *Consider the Brown-Resnick process* $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^2\}$ *with* $\delta(\boldsymbol{h}) = \theta\|\boldsymbol{h}\|^\alpha$ *for* $\alpha \in (0,2]$ *and* $\theta > 0$. *Assume that* $\log m_n = o(r_n^\alpha)$ *and*

$$\sup_n \frac{\lambda_n^2 n^{2a}}{m_n} < \infty \quad \textit{and} \quad \sup_n \frac{m_n}{\lambda_n^2 n^{2a}} < \infty \quad \textit{for} \quad 0 < a < 1. \tag{2.37}$$

*Then LUNC, (**M2**), and (**M3**) can be verified (see Appendix C), so that Theorem 3 applies. Furthermore, (2.21) holds if* $\frac{|S_n|\lambda_n^2}{m_n^3} \to 0$.

Remark 6. *Using a similar change of variable technique, one can verify that condition (2.37) satisfies (2.16) with* $\delta = 1$. *See Appendix D. One of choices that satisfies condition (2.37) and* $\frac{|S_n|\lambda_n^2}{m_n^3} \to 0$ *is* $a = \frac{7}{12}, \lambda_n = n^{-1/3}$ *and* $m_n = n^{1/2}$.

To simulate the Brown-Resnick process in $\mathbb{R}^2$, we use *RPbrownresnick* in the *RandomFields* package in R for $\delta(\boldsymbol{h}) = \frac{4}{9}\|\boldsymbol{h}\|^2$ (top) and $\|\boldsymbol{h}\|^2$ (bottom). In each simulation, first we generate 1600 random locations in $\{1, ..., 40\}^2$, where the process is simulated with the

scale of $(1/\log(1600))^{1/a}$ and $\rho(\cdot) = (1 + \theta \|\cdot\|^a)^{-1}$ with $a = 2$. For the ESE computation, we use $A = B = (1, \infty)$, $a_m = .97$ upper quantile. We set $w(\cdot) = I_{[-\frac{1}{2}, \frac{1}{2}]}(\cdot)$, and distances $\|\boldsymbol{h}\| = (0.5, 1, ..., 4.5, 5)$. In Figure 2.2 (right), the extremogram and ESE from one realization are displayed. The extremogram $\rho_{AB}(\boldsymbol{h})$ corresponds to connected solid circles and $\hat{\rho}_{AB,m}(\boldsymbol{h})$ for different bandwidths λ_n are displayed in different point types. As will be seen in Section 2.3.3, smaller variances and larger biases are observed for a larger bandwidth. The two horizontal lines are the random permutation bands.

2.3.3 Simulations

We use a simulation experiment to examine performances of the ESE. Samples are generated from models with Fréchet marginals for both lattice and non-lattice cases. For lattice cases, we consider MMA(1) and the Brown-Resnick process with $\delta(\boldsymbol{h}) = \|\boldsymbol{h}\|^2$. In each simulation, $\hat{\rho}_{AB,m}(\boldsymbol{h})$ with $A = B = (1, \infty)$ and $a_m = .97$ upper quantile is calculated for observed distances less or equal to 10. This is repeated 1000 times.

For the MMA(1), the steps of simulations are

Step 1: Generate unit Fréchet $Z_{\boldsymbol{t}}, \boldsymbol{t} \in \{1, ..., 40\}^2$ using *rmaxstab* in R.
Step 2: Construct the MMA(1) using $X_{\boldsymbol{t}} = \max_{\|\boldsymbol{t}-\boldsymbol{s}\| \le 1} Z_{\boldsymbol{s}}$.
Step 3: Compute the ESE with $A = B = (1, \infty)$ and $a_m = .97$ upper quantile for $\|\boldsymbol{h}\| \le 10$.

Figure 2.3 (upper left) shows the distributions of $\hat{\rho}_{AB,m}(\boldsymbol{h})$ (box plots), $\rho_{AB}(\boldsymbol{h})$ (red solid squares) and $\rho_{AB,m}(\boldsymbol{h})$ (blue solid circles) for the MMA(1). Observe that the distributions is centered at $\rho_{AB,m}(\boldsymbol{h})$, not $\rho_{AB}(\boldsymbol{h})$. Notice that $\rho_{AB,m}(\boldsymbol{h})$ is computed by

$$\begin{aligned} P(X_{\boldsymbol{h}} > a_m | X_{\boldsymbol{0}} > a_m) &= \frac{1 - 2P(X_{\boldsymbol{0}} \le a_m) + P(X_{\boldsymbol{h}} \le a_m, X_{\boldsymbol{0}} \le a_m)}{P(X_{\boldsymbol{0}} > a_m)} \\ &= \frac{\frac{2}{m} - 1 + (1 - \frac{1}{m})^{8/5}}{1/m} \quad \text{for } \|\boldsymbol{h}\| = 1, \sqrt{2}, \\ &= \frac{\frac{2}{m} - 1 + (1 - \frac{1}{m})^{9/5}}{1/m} \quad \text{for } \|\boldsymbol{h}\| = 2, \\ &= \frac{1}{m} \quad \text{for } \|\boldsymbol{h}\| > 2, \end{aligned}$$

using $P(X > a_m) = \frac{1}{m}$ and $P(X \le x) = \exp^{-5/x}$ for $x > 0$. For example, if we use $a_m = .97$

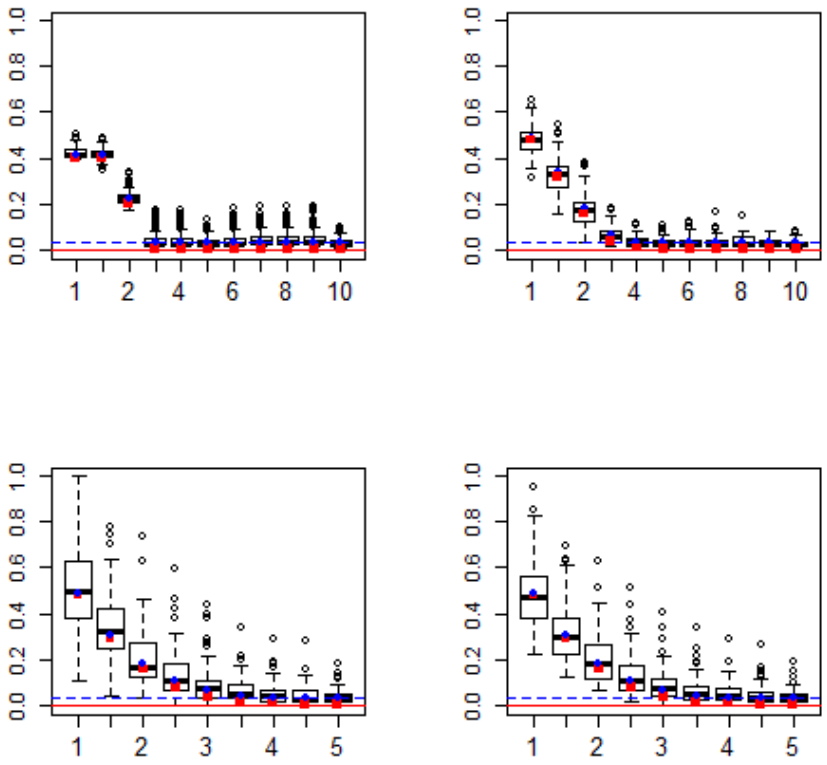

Figure 2.3: The distribution of the ESE for the MMA (1) and BR process on lattice (top), and the BR process on non-lattice with two bandwidth choices (bottom).

upper quantile, or equivalently $m = 0.03^{-1}$, $P(X_{\boldsymbol{h}} > a_m | X_{\mathbf{0}} > a_m)$ and the biases are computed as below.

$\|\boldsymbol{h}\|$	1	$\sqrt{2}$	2	2 <
PA extremogram	0.4145	0.4145	0.2216	0.03
Bias	0.0145	0.0145	0.0216	0.03

All numbers are rounded to four decimal place.

For the Brown-Resnick process on the lattice, the simulations are the same except Step 1 - Step 2 are replaced by

Step 1*: Generate the Brown-Resnick process $X_{\boldsymbol{t}}, \boldsymbol{t} \in \{1, ..., 40\}^2$ by *RPbrownresnick* in R.

The upper right panel of the figure presents the distributions of the ESE with $\rho_{AB}(\boldsymbol{h})$ (solid squares) and $\rho_{AB,m}(\boldsymbol{h})$ (solid circles) for the Brown-Resnick process on the lattice. The derivation of $\hat{\rho}_{AB,m}(\boldsymbol{h})$ is from (2.34). Again, the ESE is centered around the PA-extremogram. In the figure, the PA-extremogram is approximated by

$$P(X_{\boldsymbol{h}} > a_m | X_{\boldsymbol{0}} > a_m) \sim \frac{\frac{2\bar{\Phi}(\sqrt{\delta(\boldsymbol{h})/2})}{a_m} + \frac{2\Phi(\sqrt{\delta(\boldsymbol{h})/2})^2 - 1}{(a_m)^2}}{1 - e^{-1/a_m}}$$

since

$$\begin{aligned}
& P(X_{\boldsymbol{h}} > a_m | X_{\boldsymbol{0}} > a_m) \\
= \; & \frac{1 - 2e^{-1/a_m} + e^{-(2\Phi(\sqrt{\delta(\boldsymbol{h})/2}))/a_m}}{1 - e^{-1/a_m}} \\
= \; & \frac{1 - 2\left(1 - \frac{1}{xa_m} + \frac{1}{2}(\frac{1}{a_m})^2 + O((\frac{1}{a_m})^3)\right)}{1 - e^{-1/a_m}} \\
& + \; \frac{\left(1 - \frac{1}{a_m} 2\Phi(\sqrt{\delta(\boldsymbol{h})/2}) + \frac{1}{2}(\frac{2\Phi(\sqrt{\delta(\boldsymbol{h})/2})}{a_m})^2 + O((\frac{1}{a_m})^3)\right)}{1 - e^{-1/a_m}} \\
= \; & \frac{\frac{2\bar{\Phi}(\sqrt{\delta(\boldsymbol{h})/2})}{a_m} + \frac{2\Phi(\sqrt{\delta(\boldsymbol{h})/2})^2 - 1}{(a_m)^2} + O((\frac{1}{a_m})^3)}{1 - e^{-1/a_m}}.
\end{aligned} \tag{2.38}$$

Observe that (2.38) implies that

$$\lim_{m \to \infty} P(X_{\boldsymbol{h}} > a_m | X_{\boldsymbol{0}} > a_m) = 2\bar{\Phi}(\sqrt{\delta(\boldsymbol{h})/2}),$$

which is consistent to (2.33). From the above, we see that $\rho_{AB,m}(\boldsymbol{h}) - \rho_{AB}(\boldsymbol{h}) = \mathrm{O}\left(\frac{1}{m}\right)$ since the right-hand side of (2.38) subtracted by $2\bar{\Phi}(\sqrt{\delta(\boldsymbol{h})/2})$ equals to

$$\begin{aligned}
& \frac{\frac{2\bar{\Phi}(\sqrt{\delta(\boldsymbol{h})/2})}{a_m} + \frac{2\Phi(\sqrt{\delta(\boldsymbol{h})/2})^2 - 1}{(a_m)^2} + O((\frac{1}{a_m})^3) - \left(1 - e^{-1/a_m}\right) 2\bar{\Phi}(\sqrt{\delta(\boldsymbol{h})/2})}{1 - e^{-1/a_m}} \\
= \; & \frac{\frac{2\bar{\Phi}(\sqrt{\delta(\boldsymbol{h})/2})}{a_m} + \frac{2\Phi(\sqrt{\delta(\boldsymbol{h})/2})^2 - 1}{(a_m)^2} + O((\frac{1}{a_m})^3)}{1 - e^{-1/a_m}} \\
& - \; \frac{\left(\frac{1}{a_m} - \frac{1}{2}(\frac{1}{a_m})^2 + O((\frac{1}{a_m})^3)\right) 2\bar{\Phi}(\sqrt{\delta(\boldsymbol{h})/2}))}{1 - e^{-1/a_m}} \\
= \; & O(1/a_m^2)/O(1/a_m).
\end{aligned}$$

Turning to a non-lattice case, the simulations involve the below steps.

Step 1: Randomly generate 1600 locations using Uniform $(0, 40)\times$Uniform $(0, 40)$.

Step 2: Given locations, calculate pairwise distances, then compute $\delta(\cdot)$.

Step 3: Simulate the Brown-Resnick process using *RPbrownresnick* and compute the ESE with $A = B = (1, \infty)$ and $a_m = .97$ upper quantile.

Alternatively, one can generate spatial Gaussian process $W_{\boldsymbol{s}_i}, i = 1, ..., 1600$ from multivariate normal distribution with mean 0 and variance Σ where Σ is derived from

$$cov(W_{\boldsymbol{s}_1}, W_{\boldsymbol{s}_2}) = \delta(\boldsymbol{s}_1) + \delta(\boldsymbol{s}_2) - \delta(\boldsymbol{s}_1 - \boldsymbol{s}_2).$$

Then, repeat the above 100 times to derive the "approximated" max-stable process using (2.31). Then, the ESE is computed on the "approximated" max-stable process.

The bottom panels of Figure 2.3 is the simulation results from the Brown-Resnick process in the non-lattice case. For each simulation, 1600 points are generated from a Poisson process in $\{1, ..., 40\}^2$, then we compute $\hat{\rho}_{AB,m}(\boldsymbol{h})$ for $\|\boldsymbol{h}\| = (0.5, 1, ..., 4.5, 5)$ using the bandwidths $\lambda_n = 1/\log n$ and $5/\log n$. This is repeated 100 times. Notice that the ESE using $\lambda_n = 1/\log n$ has smaller biases but larger variances compared to the ESE using $\lambda_n = 5/\log n$ for $h \leq 2$. For longer distances, the differences is not apparent. This indicates that the ESE with wider bandwidths tends to have smaller variance but larger biases.

2.4 Application

2.4.1 Lattice case: rainfall in a region in Florida

In this section, we apply the ESE to analyze geographical dependence of heavy rainfall in a region in Florida. The source is Southwest Florida Water Management District. The raw data is total rainfall in 15 minute intervals from 1999 to 2004, measured on a 120 $\times$ 120 $(\text{km})^2$ region containing 3600 grid locations. The region of the measurements is shown in Figure 2.4.

Buhl and Klüppelberg (2016) studied the spatio-temporal process, constructed by taking the spatial maximum over a non-overlapping block and daily maximum for each location is studied and investigated the properties such as max-stability, anisotropy and etc. In

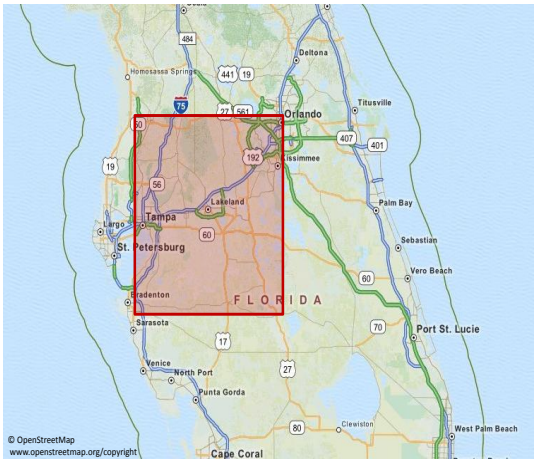

Figure 2.4: The region of Florida rainfall data.

particular, it was found that 1) a Gumbel distribution fits a block-maxima data over time for each location well, 2) there is not enough statistical evidence that the data set is not generated from a max-stable process, and 3) there is little extremal temporal dependences even for a daily frequency.

For our analysis, we focus on the spatial domain. For each fixed time, we first calculate the spatial maximum over a non-overlapping block of size 10×10 $(\text{km})^2$, which provides a 12×12 grid of spatial maxima. Then, we calculate the annual maxima from 1999 to 2004 and the 6 year maxima from the corresponding time series for each spatial maximum. The 7 spatial data sets on a 12×12 grid under consideration consist of annual maxima and 6 year maxima of spatial maxima. Following Buhl and Klüppelberg (2016), it is not unreasonable to view these 6 spatial data sets as realizations from a max-stable process. Moreover, the data are constructed as a maxima over a spatial grid of 25 locations and a temporal resolution of 15 minutes intervals.

We first look at the spatial extremal dependence for 6 year maxima rainfall. In Figure 2.5, the locations of extremes (left) and the ESE (right) are displayed, where the ESE is computed using $A = B = (1, \infty)$ and $a_m = .70$ (dotted line), .75 (dashed line) and .80 (solid line) upper quantiles. Since the number of spatial locations is small (144), we chose modest thresholds in order to ensure enough exceedances for estimation of the ESE. Such thresholds should provide good estimates of the pre-asymptotic extremogram for a max-stable process.

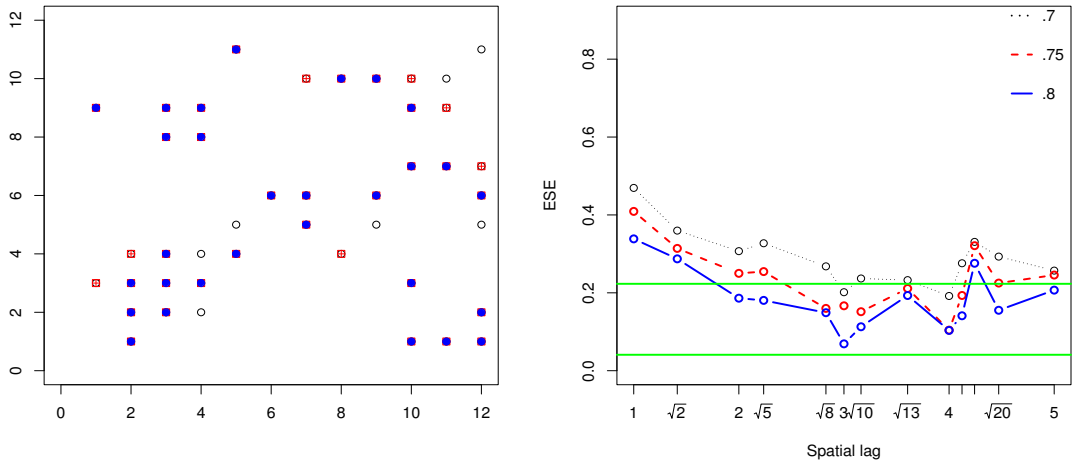

Figure 2.5: The locations of extremes and the ESE using the 6 year maxima of Florida rainfall data.

The locations of extremes are marked corresponding to choices of a_m by .70 (empty circles), .75 (empty squares) and .80 (solid circles) upper quantiles. For the ESE plot, the horizontal lines are permutation based confidence bands. For example, if extreme events are defined by any rainfall heavier than the .70 upper quantile of the maxima rainfall observed for the entire periods, there is a significant extremal dependence between two clusters at distance 2. On the other hand, using the .80 upper quantile, the extremal dependence at the same distance is no longer significant. In the case of 6 year maxima rainfall, the ESE from the .70 upper quantile indicates that no spatial extremal dependence for spatial lags larger than 3. A small spike of the ESE at spatial lags around 4 may be the result of two extremal clusters that are 4 units apart, as seen in the left panel of Figure 2.5.

By looking at the ESE of annual maxima rainfall from 1999 to 2004, we see year-over-year changes in spatial extremal dependence. Figure 2.6 presents the locations of extremes and the ESE from 1999 to 2004 (left to right, top to bottom). For example, the ESE suggests that the spatial extremal dependence for lags less than 3 in 2000 is stronger than at any other year between 1999 and 2004. Using the .80 upper quantile, there is significant extremal dependence for spatial lag $\sqrt{8}$ in 2000, but not for any other years. In 2002, the spatial extremal dependence is not significant at $\|\boldsymbol{h}\| = \sqrt{8}$ using the .80 upper quantile. Similarly, the year-to-year comparisons of the ESE with .70 and .75 upper quantiles confirm

that the spatial extremal dependence for spatial lags up to 3 is stronger in 2000 than in any other years.

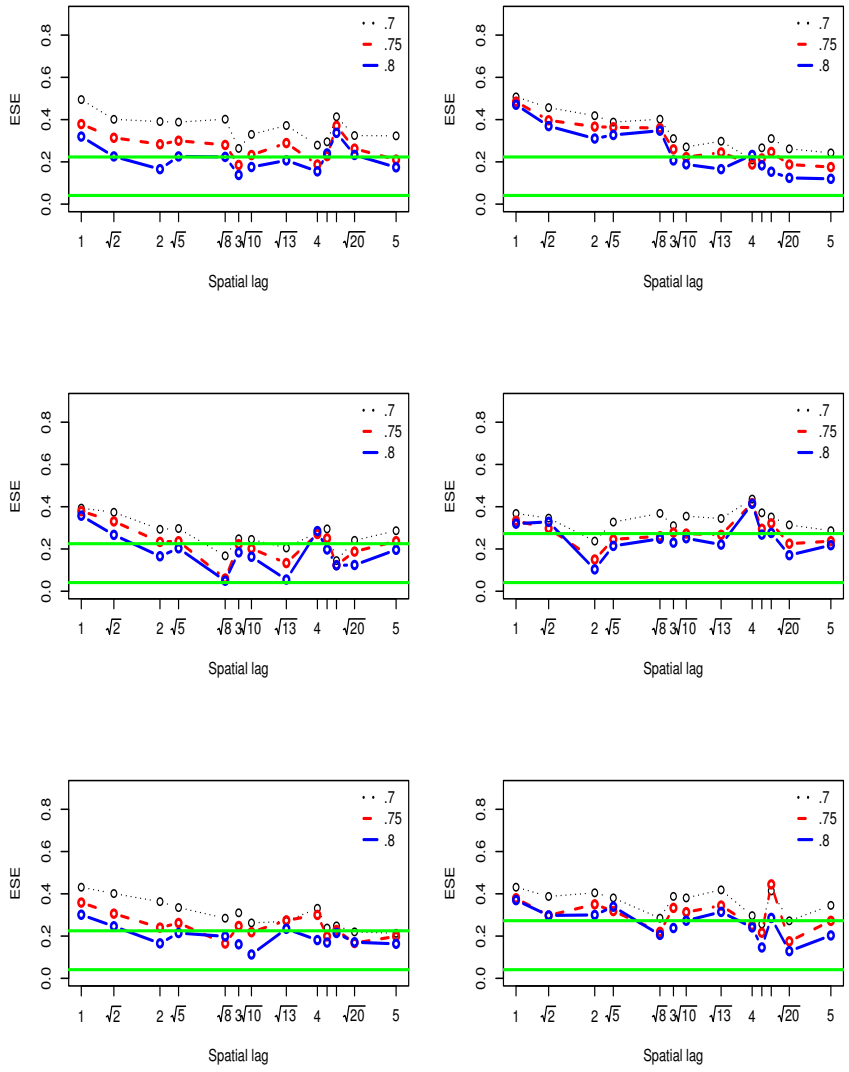

Figure 2.6: The ESE of the annual maxima of Florida rainfall from 1999 to 2004.

2.4.2 Non-lattice case: ground-level ozone in the eastern United States

Now we apply the ESE to analyze geographical extremal dependence of ground-level ozone in the eastern United States. The data is the maximum ozone reading of maximum daily 8-hour averages ozone levels in part per billion (ppb) from April to October 1997, measured

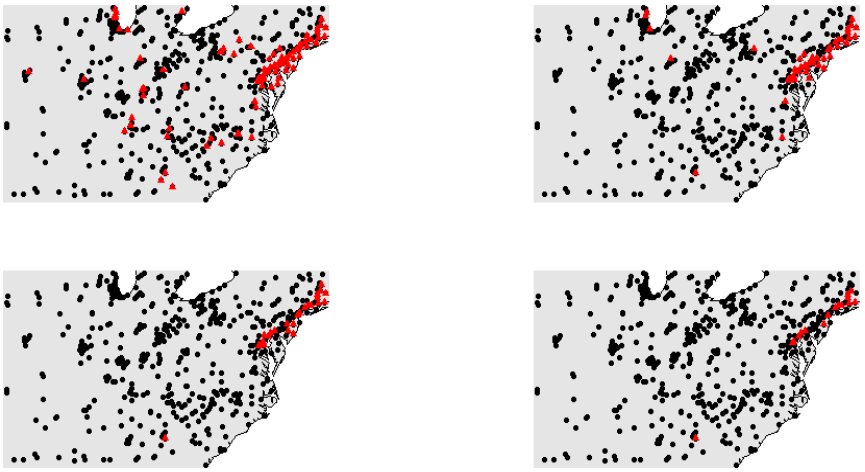

Figure 2.7: The region of ozone monitoring station in the eastern United States with extreme observations defined with different thresholds of .80, .90, .95, and .97 upper quantiles.

at 513 ozone monitoring stations. The source is *Ozone4H* in the *extRemes* package in R, but the raw data source is the U.S. Environmental Protection Agency (EPA) National Ambient Air Quality Standard (NAAQS).

The region of the measurements in the eastern United States is presented in Figure 2.7, where the locations of monitoring stations are denoted as circles. The range of the maximum ozone reading observed from the region is from 56 ppb to 153 ppb. In the figure, the station with ground-level ozone greater than the .80 (top left), .90 (top right), .95 (bottom left), and .97 (bottom right) upper quantiles are marked with triangles. For example, in the bottom right panel, the station with ground-level ozone greater than the .97 upper quantile, corresponding to 136 ppb, is presented with triangles. One can see that the extreme cluster is located on the north eastern part of the region and that the extreme clusters do not spread until a low threshold, such as .80 quantile, is used.

Given only the maximum ozone reading is available, marginal distributional analysis over time at a fixed location is not possible. However, some assumptions can bee justified. First, we check if the data are from a max-stable process. Because the data are constructed as a maxima over a temporal resolution of 8-hour averages for 6 month periods, it is not unreasonable to view these data points as realizations from a max-stable process. How-

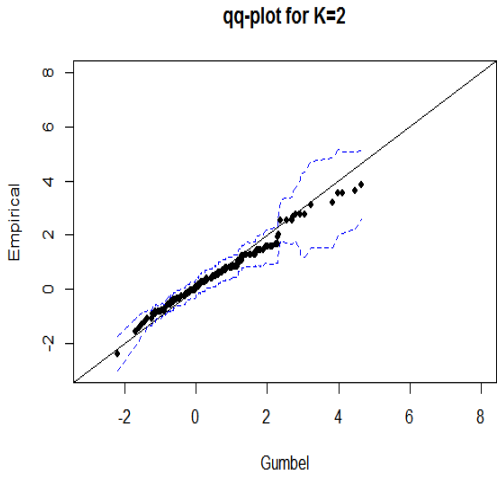

Figure 2.8: QQ plot of the standard Gumbel and the empirical quantiles (dots) derived from the empirical distribution of maxima over sets with cardinality K=2.

ever, we used the procedure outlined in Gabda et al. (2012) to check max-stability of the bivariate distributions. This test is based on the idea that under the null hypothesis of max-stability, maxima taken over sets follow the standard Gumbel distribution after location shifts, where the latter can be estimated. Then one can graphically compare the transformed sample maxima with the standard Gumbel distribution using a QQ plot. The test showed no contradictory evidence of max-stability, as seen in Figure 2.8, where the empirical distribution of maxima from bivariate distributions (dots) and 95% confidence bands (two dashed lines) are presented. As one can see, the diagonal line is between the confidence bands, indicating that there is no statistically significant evidence that the data is not max-stable. For the convenience, the graphical results use 200 points that are randomly selected from $\binom{513}{2}$ pairs. We did not pursue higher dimensional distributions.

In terms of stationarity, it is often a reasonable assumption from the modeling perspective. Spatial stationarity is a difficult condition to verify from real data. For example, data can often look stationary even when generated by stochastic volatility data. It is often difficult to discriminate between regular variation and exact max-stable marginals from a sample. So although max-stability is a simplifying assumption, it is not a bad first step in modeling.

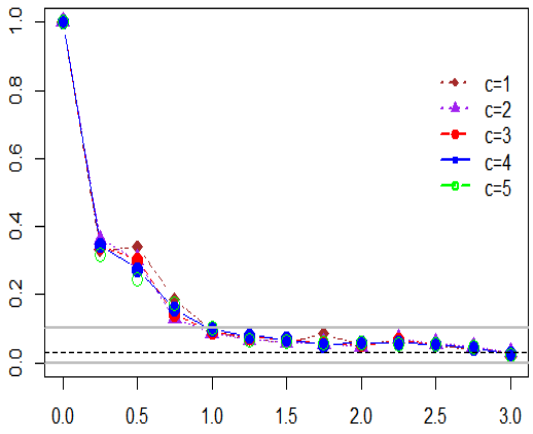

Figure 2.9: The ESE of the maximum ozone data for 1997 with different choices of bandwidths.

For the ESE, the great-circle distances are used as in Gilleland and Nychka (2005). The shortest distance over the earth's surface is calculated from longitude/latitude of the stations using the haversine formula. The unit for distance is chosen as 100 mile. Then, the ESE using $A = B = (1, \infty)$ and the .97 upper quantile with different bandwidths $\lambda_n = c/\log n$ for $c = 1, 2, 3, 4$, and 5 are calculated for every 25 miles interval, as shown in Figure 2.9. It turns out that $\frac{1}{\log n}$ in the bandwidth corresponds to 8 miles. For example, $\hat{\rho}_{AB,m}(1)$ with $\lambda_n = \frac{1}{\log n}$ is estimated by using pairs of points whose distances are between 92 and 108 miles.

From the figure, it is noted that the ESE is robust with respect to bandwidths choices. The extremal dependence of ground-level ozone is below 0.4 even for the short distance of 25 miles and it disappears for $\|\boldsymbol{h}\| > 100$ miles, indicated by random permutation bands (two horizontal lines). These observations are aligned with Gilleland and Nychka (2005) and Gilleland et al. (2006), where the authors found that the spatial dependence in the fourth-highest daily maximum 8-hour average ozone level fields is limited to a short distance less than 100 miles. The dashed line corresponds to 0.03 $(= 1 - 0.97)$.

2.5 Conclusion

In Chapter 2, we study the spatial extremogram that is defined as the extremogram for data for higher dimensions. We also examine the asymptotic properties for the empiricial spatial extremogram (the ESE) under two sampling scenarios - when the samples are from the lattice and non-lattice. For the lattice case, the asymptotic results of the ESE can be viewed as a generalization of the asymptotic results for a stationary time series in Davis and Mikosch (2009). In particular, the limiting variance and the scaling term need to be coordinated by a dimension to achieve the central limit theorem. For a non-lattice case, we consider a kernal estimator following ideas in Karr (1986) and Li et al. (2008). We find that the central limit theorem for the ESE holds when the growth rates of bandwidths and sampling regions, and the decay rates of dependence and mixing functions are coordinated. The performance of the ESE under both sampling schemes is demonstrated with the simulations, which confirms that the ESE is capable of capturing theoretical aspects of the simulated processes. Moreover, real data applications to a heavy rain fall data in a region of Florida and ozone data in the eastern United States show that the ESE provides consistent results with the existing literatures, thus it can be used as a tool to analyze the spatial extremal dependency.

2.6 Appendix: Proofs

The following proposition presented by Li et al. (2008) is used in the proof. The proposition is analogous to Theorem 17.2.1 in Ibragimov and Linnik (1971).

Proposition 2.6.1 (Lemma A.1. in Li et al. (2008))**.** *Let U and V be two closed and connected sets in $\mathbb{R}^d$ such that $\#U = \#V \leq b$ and $d(U, V) \geq r$ for some constants b and r. For a stationary process $X_{\boldsymbol{s}}$, consider ξ and η measurable random variables with respect to $\sigma(X_{\boldsymbol{s}} : \boldsymbol{s} \in U)$ and $\sigma(X_{\boldsymbol{s}} : \boldsymbol{s} \in V)$ with $|\xi| \leq C_1, |\eta| \leq C_2$. Then $|cov(\xi, \eta)| \leq 4C_1C_2\alpha_{b,b}(r)$.*

Proof. Using the properties of conditional expectation,

$$\begin{aligned} |E(\xi\eta) - E\xi E\eta| &= |E(\xi[E(\eta|A) - E(\eta)])| \\ &\leq C_1 E|E(\eta|A) - E(\eta)| \\ &= C_1 E(\tilde{\xi}(E(\eta|A) - E(\eta))) \end{aligned}$$

where $\tilde{\xi} = sgn\{E(\eta|A) - E(\eta)\}$, measurable w.r.t. A and C_1 is a constant. Thus,

$$|E(\xi\eta) - E\xi E\eta| \leq C_1 |E(\tilde{\xi}\eta) - E(\tilde{\xi})E(\eta)|.$$

Similarly, define $\tilde{\eta} = sgn\{E(\tilde{\xi}|B) - E(\tilde{\xi})\}$, to give

$$|E(\xi\eta) - E\xi E\eta| \leq C_1 C_2 |E(\tilde{\xi}\tilde{\eta}) - E(\tilde{\xi})E(\tilde{\eta})|$$

where C_2 is a constant. Now introduce the events $C = \{\tilde{\xi} = 1\} \in A$ and $D = \{\tilde{\eta} = 1\} \in B$. Then, $|E(\tilde{\xi}\tilde{\eta}) - E(\tilde{\xi})E(\tilde{\eta})|$ is bounded by

$$|P(CD) + P(C^c D^c) - P(C^c D) - P(CD^c) - P(C)P(D) - P(C^c)P(D^c) + P(C)P(D^c) + P(C^c)P(D)|$$

which is again bounded by $4\alpha_{b,b}(r)$ □

2.6.1 Appendix A: Proof of Theorem 1

Theorem 1 is derived from Theorem 4. For notation, we suppress the dependence of m on n and write m for m_n. Define a vector valued random field by

$$Y_{\boldsymbol{t}} = X_{D_{\boldsymbol{t}}}, \text{ where } D_{\boldsymbol{t}} = \boldsymbol{t} + B_\gamma = \{\boldsymbol{s} \in \mathbb{Z}^d : d(\boldsymbol{t}, \boldsymbol{s}) \leq \gamma\}.$$

Theorem 4 will prove a central limit theorem for

$$\begin{aligned} \hat{P}_m(C) &= \frac{m_n}{n^d} \sum_{\boldsymbol{t} \in \Lambda_n} I_{\{Y_{\boldsymbol{t}}/a_m \in C\}} \\ &= \frac{m_n}{n^d} \sum_{\boldsymbol{t} \in \Lambda_n^p} I_{\{Y_{\boldsymbol{t}}/a_m \in C\}} + \frac{m_n}{n^d} \sum_{\boldsymbol{t} \in \Lambda_n \setminus \Lambda_n^p} I_{\{Y_{\boldsymbol{t}}/a_m \in C\}}, \end{aligned} \tag{2.39}$$

where $\Lambda_n^p = \{\boldsymbol{t} \in \Lambda_n | d(\boldsymbol{t}, \partial\Lambda_n) \geq p\}$ and $\partial\cdot$ denotes the boundary. In fact, showing a central limit theorem for the first term in (2.39) is sufficient as the second term is negligible as $n \to \infty$.

Recall that $p_m(A) = mP(X_{\mathbf{0}} \in a_m A)$ and $\tau_{AB,m}(\boldsymbol{h}) = mP(X_{\mathbf{0}} \in a_m A, X_{\boldsymbol{h}} \in a_m B)$, where A and B are sets bounded away from the origin. Write

$$\begin{aligned}
&\mu(A) = \lim_{n\to\infty} p_m(A),\\
&\tau_{AB}(\boldsymbol{h}) = \lim_{n\to\infty} \tau_{AB,m}(\boldsymbol{h}),\\
&\mu_A(D_{\mathbf{0}}) = \lim_{x\to\infty} P\left(\frac{Y_{\boldsymbol{t}}}{\|Y_{\boldsymbol{t}}\|} \in A \Big| \|Y_{\boldsymbol{t}}\| > x\right), \text{and}\\
&\tau_{A\times B}(D_{\mathbf{0}} \times D_{\boldsymbol{l}}) = \lim_{x\to\infty} P\left(\frac{(Y_{\mathbf{0}}, Y_{\boldsymbol{l}})}{\|vec\{Y_{\mathbf{0}}, Y_{\boldsymbol{l}}\}\|} \in A\times B \Big| \|vec\{Y_{\mathbf{0}}, Y_{\boldsymbol{l}}\}\| > x\right).
\end{aligned}$$

Theorem 4. *Assume the conditions of Theorem 1. Let C be a set bounded away from zero and a continuity set with respect to $\mu(\cdot)$ and $\tau(\boldsymbol{h})$ for $\boldsymbol{h} \in \mathbb{R}^d$. Then*

$$S_n = \left(\frac{m_n}{n^d}\right)^{1/2} \sum_{\boldsymbol{s}\in\Lambda_n} \left[I\left(\frac{Y_{\boldsymbol{s}}}{a_m} \in C\right) - P\left(\frac{Y_{\boldsymbol{s}}}{a_m} \in C\right)\right] \xrightarrow{d} N(0, \sigma_Y^2(C))$$

where $\sigma_Y^2(C) = \mu_C(D_{\mathbf{0}}) + \sum_{\boldsymbol{l}\neq\mathbf{0}\in\mathbb{Z}^d} \tau_{C\times C}(D_{\mathbf{0}} \times D_{\boldsymbol{l}})$.

Proof. We use ideas from Bolthausen (1982) and Davis and Mikosch (2009) to show the central limit theorem for

$$\hat{P}_m(C) = m_n \textstyle\sum_{\boldsymbol{s}\in\Lambda_n} I_{\boldsymbol{s}}/|\Lambda_n| \text{ where } I_{\boldsymbol{s}} = I_{\{X_{\boldsymbol{s}}/a_m \in C\}}.$$

The proof for the central limit theorem of $X_{\boldsymbol{s}}$ replaced by a vector valued random field $Y_{\boldsymbol{s}}$ in indicator is analogous.

Define $H[a,b] = \{d(\boldsymbol{s},\boldsymbol{t}) : a \le d(\boldsymbol{s},\boldsymbol{t}) \le b\}$ and $\|\boldsymbol{l}\| = d(\mathbf{0},\boldsymbol{l})$ for convenience. Assume that $m_n^{2+2d} = o(n^d)$ and

$$\lim_{k\to\infty} \limsup_{n\to\infty} m_n \sum_{\boldsymbol{l}\in\mathbb{Z}^d, \|\boldsymbol{l}\|\in H(k,r_n]} P(|X_{\boldsymbol{l}}| > \varepsilon a_m, |X_{\mathbf{0}}| > \varepsilon a_m) = 0, \text{ for } \forall \epsilon > 0, \quad (2.40)$$

$$\lim_{n\to\infty} m_n \sum_{\boldsymbol{l}\in\mathbb{Z}^d, \|\boldsymbol{l}\|\in H(r_n,\infty)} \alpha_{1,1}(\|\boldsymbol{l}\|) = 0, \quad (2.41)$$

$$\sum_{\boldsymbol{l}\in\mathbb{Z}^d} \alpha_{j_1,j_2}(\|\boldsymbol{l}\|) < \infty \quad \text{for} \quad 2 \le j_1 + j_2 \le 4, \quad (2.42)$$

$$\lim_{n\to\infty} n^{d/2} m_n^{1/2} \alpha_{1,n^d}(m_n) = 0, \quad (2.43)$$

which are univariate case analog of conditions (2.3) - (2.6).

By the same arguments in Davis and Mikosch (2009),

$$E\hat{P}_m(C) \to \mu(C), \tag{2.44}$$

$$var\left(\hat{P}_m(C)\right) \sim \frac{m_n}{n^d}\left[\mu(C) + \sum_{\boldsymbol{l}\neq\boldsymbol{0}\in\mathbb{Z}^d} \tau_{CC}(\boldsymbol{l})\right] = \frac{m_n}{n^d}\,\sigma_X^2(C), \tag{2.45}$$

where (2.44) is implied by the regularly varying assumption. To see (2.45), observe that

$$\begin{aligned}\frac{n^d}{m_n} var\left(\hat{P}_m(C)\right) &= \frac{m_n}{n^d}\sum_{\boldsymbol{s}\in\Lambda_n} var(I_{\boldsymbol{s}}) + \frac{m_n}{n^d}\sum_{\boldsymbol{s},\boldsymbol{t}\in\Lambda_n,\boldsymbol{s}\neq\boldsymbol{t}} cov(I_{\boldsymbol{s}}, I_{\boldsymbol{t}}) \\ &= A_1 + A_2.\end{aligned}$$

By the regularly varying assumption, $A_1 = p_m(C) - (p_m(C))^2/m_n \to \mu(C)$.

Turning to A_2, for $k \geq 1$ fixed, we have

$$\begin{aligned}A_2 &\sim \frac{m_n}{n^d}\sum_{\boldsymbol{l}=(\boldsymbol{l}_1,\ldots,\boldsymbol{l}_d)\neq\boldsymbol{0},\|\boldsymbol{l}\|\leq\max\Lambda_n} \Pi_{i=1}^d(n-|\boldsymbol{l}_i|)cov(I_{\boldsymbol{0}}, I_{\boldsymbol{l}}) \\ &= \frac{m_n}{n^d}\left[\sum_{\boldsymbol{l}\in\mathbb{Z}^d,\|\boldsymbol{l}\|\in H(0,k]} \cdot + \sum_{\boldsymbol{l}\in\mathbb{Z}^d,\|\boldsymbol{l}\|\in H(k,r_n]} \cdot + \sum_{\boldsymbol{l}\in\mathbb{Z}^d,\|\boldsymbol{l}\|\in H(r_n,\max\Lambda_n]} \cdot\right] \\ &= A_{21} + A_{22} + A_{23},\end{aligned}$$

where $\max\Lambda_n = \{\max(d(\boldsymbol{s},\boldsymbol{t})) : \boldsymbol{s},\boldsymbol{t}\in\Lambda_n\}$. Note that $\Pi_{i=1}^d(n-|\boldsymbol{l}_i|)$ counts a number of cubes with lag $\boldsymbol{l}$ in Λ_n.

From the regularly varying assumption, $\lim_{k\to\infty}\limsup_{n\to\infty} A_{21} = \sum_{\boldsymbol{l}\neq\boldsymbol{0}\in\mathbb{Z}^d}\tau_{CC}(l)$ since

$$\begin{aligned}\limsup_{n\to\infty} A_{21} &= \sum_{\boldsymbol{l}\in\mathbb{Z}^d,\|\boldsymbol{l}\|\in H(0,k]} \limsup_{n\to\infty}\left(\tau_{CC,m}(C) - \frac{p_m(C)^2}{m_n} + \left(\frac{m_n}{n}\right)\right) \\ &= \sum_{\boldsymbol{l}\in\mathbb{Z}^d,\|\boldsymbol{l}\|\in H(0,k]} \tau_{CC}(\boldsymbol{l}).\end{aligned}$$

Thus, it is sufficient to show

$$\lim_{k\to\infty}\limsup_{n\to\infty}(|A_{22}| + |A_{23}|) = 0$$

to achieve (2.45). Recall that C is bounded away from the origin. Notice that

$$\begin{aligned}A_{22} &\leq const\ m_n \sum_{\boldsymbol{l}\in\mathbb{Z}^d,\|\boldsymbol{l}\|\in H(k,r_n]}\left[P(|X_{\boldsymbol{l}}| > \varepsilon a_m, |X_{\boldsymbol{0}}| > \varepsilon a_m) + \left(\frac{p_m(C)}{m_n}\right)^2\right], \\ A_{23} &\leq const\ \ m_n \sum_{\boldsymbol{l}\in\mathbb{Z}^d,\|\boldsymbol{l}\|\in H(r_n,\infty)} \alpha_{1,1}(\|\boldsymbol{l}\|),\end{aligned}$$

so (2.45) holds assuming (2.40), (2.41) and $r_n^d/m_n \to 0$.

Now, we prove

$$\sqrt{\frac{n^d}{m_n}}(\hat{P}_m(C) - p_m(C)) = \sqrt{\frac{m_n}{n^d}} \sum_{\boldsymbol{s} \in \Lambda_n} \bar{I}_{\boldsymbol{s}}, \xrightarrow{d} N(0, \sigma_X^2(C)) \tag{2.46}$$

where $\bar{I}_{\boldsymbol{t}} = I\left(\frac{X_{\boldsymbol{t}}}{a_m} \in C\right) - P\left(\frac{X}{a_m} \in C\right)$. First, we infer from the argument showing (2.45) that

$$\frac{m_n}{n^d} \sum_{\boldsymbol{s},\boldsymbol{t} \in \Lambda_n} |cov(\bar{I}_{\boldsymbol{s}}, \bar{I}_{\boldsymbol{t}})| = |A_1 + A_2| < \infty. \tag{2.47}$$

As the next step, define

$$\begin{aligned}
S_{\alpha,n} &= \sum_{\beta \in \Lambda_n, d(\alpha,\beta) \le m_n} \sqrt{\frac{m_n}{n^d}} \bar{I}_\beta, \\
v_n &= \sum_{\alpha \in \Lambda_n} E(\sqrt{\frac{m_n}{n^d}} \bar{I}_\alpha S_{\alpha,n}), \\
\bar{S}_n &= v_n^{-1/2} S_n, \\
\bar{S}_{\alpha,n} &= v_n^{-1/2} S_{\alpha,n}.
\end{aligned}$$

From the definition, $v_n \sim var(S_n) \to \sigma_X^2(C)$. Now, we use Stein's lemma to show (2.46) as done in Bolthausen (1982) by checking $\lim_{n\to\infty} E((i\lambda - \bar{S}_n)e^{i\lambda \bar{S}_n}) = 0$ for all $\lambda \in R$. Write

$$\begin{aligned}
(i\lambda - \bar{S}_n)e^{i\lambda \bar{S}_n} &= i\lambda e^{i\lambda \bar{S}_n}(1 - v_n^{-1} \sum_{\alpha \in \Lambda_n} \sqrt{\frac{m_n}{n^d}} \bar{I}_\alpha S_{\alpha,n}) \\
&\quad - v_n^{-1/2} e^{i\lambda \bar{S}_n} \sum_{\alpha \in \Lambda_n} \sqrt{\frac{m_n}{n^d}} \bar{I}_\alpha [1 - e^{-i\lambda \bar{S}_{\alpha,n}} - i\lambda \bar{S}_{\alpha,n}] \\
&\quad - v_n^{-1/2} \sum_{\alpha \in \Lambda_n} \sqrt{\frac{m_n}{n^d}} \bar{I}_\alpha e^{-i\lambda(\bar{S}_n - \bar{S}_{\alpha,n})} \\
&= B_1 + B_2 + B_3.
\end{aligned}$$

We will show $E|B_1|^2 = \lambda^2 v_n^{-2} \sum_{\alpha,\alpha',\beta,\beta', d(\alpha,\beta) \le m_n, d(\alpha',\beta') \le m_n} \frac{m_n^2}{n^{2d}} \; cov(\bar{I}_\alpha \bar{I}_\beta, \bar{I}_{\alpha'} \bar{I}_{\beta'}) \to 0$. From Proposition 2.6.1,

$$|cov(\bar{I}_\alpha \bar{I}_\beta, \bar{I}_{\alpha'} \bar{I}_{\beta'})| \le 4 \, \alpha_{2,2}(d(\alpha, \alpha') - 2m_n) \quad if \; d(\alpha, \alpha') \ge 3m_n.$$

When $d(\alpha,\alpha') < 3m_n$, let $j = \min\{d(\alpha,\alpha'), d(\alpha,\beta'), d(\beta,\alpha'), d(\beta,\beta')\}$. Then

$$|cov(\bar{I}_\alpha \bar{I}_\beta, \bar{I}_{\alpha'} \bar{I}_{\beta'})| \;\; \leq \;\; 4\, \alpha_{p,q}(j)$$

for $2 \leq p+q \leq 4$. To see this, for $d(\alpha,\alpha') < 3m_n$, let $\min\{d(\alpha,\alpha'), d(\alpha,\beta), d(\alpha,\beta')\} = j$ and $k = \min\{d(\beta,\beta'), d(\beta,\alpha')\}$. Then we have

$$\begin{aligned}
|cov(\bar{I}_\alpha \bar{I}_\beta, \bar{I}_{\alpha'} \bar{I}_{\beta'})| &= |E(\bar{I}_\alpha \bar{I}_\beta \bar{I}_{\alpha'} \bar{I}_{\beta'}) - E(\bar{I}_\alpha \bar{I}_\beta) E(\bar{I}_{\alpha'} \bar{I}_{\beta'})| \\
&\leq |E(\bar{I}_\alpha \bar{I}_\beta \bar{I}_{\alpha'} \bar{I}_{\beta'}) - E(\bar{I}_\alpha) E(\bar{I}_\beta \bar{I}_{\alpha'} \bar{I}_{\beta'})| \\
&\quad + |E(\bar{I}_\alpha) E(\bar{I}_\beta \bar{I}_{\alpha'} \bar{I}_{\beta'}) - E(\bar{I}_\alpha) E(\bar{I}_\beta) E(\bar{I}_{\alpha'} \bar{I}_{\beta'})| \\
&\quad + |E(\bar{I}_\alpha) E(\bar{I}_\beta) E(\bar{I}_{\alpha'} \bar{I}_{\beta'}) - E(\bar{I}_\alpha \bar{I}_\beta) E(\bar{I}_{\alpha'} \bar{I}_{\beta'})| \\
&\leq \alpha_{1,3}(j) + \alpha_{1,2}(j)\, I\,(j \leq k) + \alpha_{1,2}(k)\, I\,(j > k) + \alpha_{1,1}(j).
\end{aligned}$$

Provided $m_n^{2+2d} = o(n^d)$, we have $E|B_1|^2 \to 0$ since

$$\begin{aligned}
E|B_1|^2 &\leq \frac{\lambda^2}{v_n^2} \frac{m_n^2}{n^{2d}} \left[\sum_{\alpha \in \Lambda_n} \sum_{\alpha' \in \Lambda_n \cap \{d(\alpha,\alpha') > 3m_n\}} \sum_{\beta,\beta'} \left| cov(\bar{I}_\alpha \bar{I}_\beta, \bar{I}_{\alpha'} \bar{I}_{\beta'}) \right| \right. \\
&\qquad \left. + \sum_{\alpha \in \Lambda_n} \sum_{\alpha' \in \Lambda_n \cap \{d(\alpha,\alpha') \leq 3m_n\}} \sum_{\beta,\beta'} \left| cov(\bar{I}_\alpha \bar{I}_\beta, \bar{I}_{\alpha'} \bar{I}_{\beta'}) \right| \right] \\
&\leq \frac{\lambda^2}{v_n^2} \frac{m_n^2}{n^{2d}}\, 4 \left[\sum_{\alpha \in \Lambda_n} \sum_{\alpha' \in \Lambda_n \cap \{d(\alpha,\alpha') > 3m_n\}} \sum_{\beta,\beta'} \alpha_{2,2}(d(\alpha,\alpha') - 2m_n) \right. \\
&\qquad \left. + \sum_{\alpha \in \Lambda_n} \sum_{\alpha' \in \Lambda_n \cap \{d(\alpha,\alpha') \leq 3m_n\}} \sum_{\beta,\beta'} \alpha_{p,q}(j) \right] \\
&\leq \text{const}\, \frac{\lambda^2}{v_n^2} \frac{m_n^2}{n^{2d}}\, n^d m_n^{2d} \left[\sum_{\boldsymbol{l} \in \mathbb{Z}^d, \|\boldsymbol{l}\| \in H(3m_n, \infty)} \alpha_{2,2}(\|\boldsymbol{l}\| - 2m_n) \right. \\
&\qquad \left. + \sum_{\boldsymbol{l} \in \mathbb{Z}^d, \|\boldsymbol{l}\| \in H[0, 3m_n]} \alpha_{p,q}(\|\boldsymbol{l}\|) \right] \quad (*) \\
&= O(m_n^{2+2d}/n^d).
\end{aligned}$$

Notice that in (*), $n^d m_n^{2d}$ is from summing over α (giving n^d), β (giving $O(m_n^d)$), and β' (giving $O(m_n^d)$) for the first summation and α (giving n^d), β (giving $O(m_n^d)$) and β' (giving

$O(m_n^d)$) or α' (giving $O(m_n^d)$) depending on the location of points for the second summation. The last equation is from (2.42).

Now we show $E|B_2| \to 0$ if $m_n^{2+2d} = o(n^d)$. Recall that $|e^{ix} - 1 - ix| \leq \frac{1}{2}x^2$, then

$$\begin{aligned} E|B_2| &\leq cv_n^{-1/2} n^d \sqrt{\frac{m_n}{n^d}} E\bar{S}_{\alpha,n}^2 \\ &= cv_n^{-1/2} \sqrt{\frac{m_n}{n^d}} m_n \sum_{\beta,\beta', d(0,\beta)\leq m_n, d(0,\beta')\leq m_n} E(\bar{I}_\beta \bar{I}_{\beta'}) \\ &\leq c\sqrt{\frac{m_n}{n^d}} m_n^{d+1} \sum_{\boldsymbol{l}\in\Lambda_n} E(\bar{I}_{\mathbf{0}} \bar{I}_{\boldsymbol{l}}) \\ &= O\left(\sqrt{\frac{m_n^{1+2d}}{n^d}}\right) \end{aligned}$$

where $m_n \sum_{\boldsymbol{l}\in\Lambda_n} E(\bar{I}_{\mathbf{0}} \bar{I}_{\boldsymbol{l}}) < \infty$ is inferred from (2.47).

The condition (2.43) implies

$$\begin{aligned} |EB_3| &\leq cv_n^{-1/2} n^d \sqrt{\frac{m_n}{n^d}} \alpha_{1,n^d}(m_n) \\ &= cn^{d/2} m_n^{1/2} \alpha_{1,n^d}(m_n) \\ &\to 0. \end{aligned}$$

Thus, Stein's lemma is satisfied, which completes the proof. □

Remark 7. *$\hat{P}_m(C)$ is a consistent estimator of $\mu(C)$. If $\mu(C) = 0$, var $(\hat{P}_m(C)) = o(m_n/n^d)$.*

Remark 8. *In fact, (2.3) - (2.6) are derived from (2.40) - (2.43) by replacing univariate process $(X_{\boldsymbol{t}})$ by vectorized process $(Y_{\boldsymbol{t}})$. In order to see (2.3) is derived from (2.40), for example, consider Euclidean norm for $(Y_{\boldsymbol{t}})$ process. Then, the vectorized analog of (2.40) is*

$$\lim_{k\to\infty} \limsup_{n\to\infty} m_n \sum_{\boldsymbol{l}\in\mathbb{Z}^d, \|\boldsymbol{l}\|\in H(k,r_n]} P(\|Y_{\mathbf{0}}\| > \epsilon a_m, \|Y_{\boldsymbol{l}}\| > \epsilon a_m) = 0,$$

which holds under (2.3) by triangular inequality, i.e.,

$$\begin{aligned} P(\|Y_{\mathbf{0}}\| > \epsilon a_m, \|Y_{\boldsymbol{l}}\| > \epsilon a_m) &\leq P(\sum_{\boldsymbol{s}\in D_{\mathbf{0}}} |X_{\boldsymbol{s}}| > \epsilon a_m, \sum_{\boldsymbol{s}'\in D_{\boldsymbol{l}}} |X_{\boldsymbol{s}'}| > \epsilon a_m) \\ &\leq P\left(\max_{\boldsymbol{s}\in D_{\mathbf{0}}} |X_{\boldsymbol{s}}| > \frac{\epsilon a_m}{|D_{\mathbf{0}}|}, \max_{\boldsymbol{s}'\in D_l} |X_{\boldsymbol{s}'}| > \frac{\epsilon a_m}{|D_{\boldsymbol{l}}|}\right). \end{aligned}$$

The rest of the derivations are straightforward.

Proof of Theorem 1. Apply the Cramér-Wold device to Theorem 4 to achieve the multivariate central limit theorem, then use δ-method to obtain the central limit theorem for the ESE. To specify the limiting variance Σ, redefine

$$\mu(A) = \lim_{x\to\infty} P\left(\frac{X_{\boldsymbol{t}}}{\|Y_{\boldsymbol{t}}\|} \in A \Big| \|Y_{\boldsymbol{t}}\| > x\right).$$

Then, $\Sigma = \mu(A)^{-4} F \Pi F^t$ where

$$\begin{aligned}
\Pi_{i,i} &= \mu_{S_i}(D_{\mathbf{0}}) + \sum_{\boldsymbol{l}\neq\mathbf{0}\in\mathbb{Z}^d} \tau_{S_i\times S_i}(D_{\mathbf{0}} \times D_{\boldsymbol{l}}) \\
\Pi_{i,j} &= \mu_{S_i\cap S_j}(D_{\mathbf{0}}) + \sum_{\boldsymbol{l}\neq\mathbf{0}\in\mathbb{Z}^d} \tau_{S_i\times S_j}(D_{\mathbf{0}} \times D_{\boldsymbol{l}}) \\
F &= \begin{pmatrix}
\mu(S_{(\#H)+1}) & 0 & 0 & ... & 0 & -\mu_{S_1}(D_{\mathbf{0}}) \\
0 & \mu(S_{(\#H)+1}) & 0 & ... & 0 & -\mu_{S_2}(D_{\mathbf{0}}) \\
. & . & . & ... & . & . \\
. & . & . & ... & . & . \\
0 & 0 & 0 & ... & \mu(S_{(\#H)+1}) & -\mu_{S_{(\#H)}}(D_{\mathbf{0}})
\end{pmatrix},
\end{aligned}$$

where the sets S_i are chosen such that $\{Y_{\boldsymbol{t}} \in S_i\} = \{X_{\boldsymbol{t}} \in A, X_{\boldsymbol{s}} \in B : d(\boldsymbol{t},\boldsymbol{s}) = \boldsymbol{h}_i\}$ for $\boldsymbol{h}_i \in H$ and $i = 1, ..., (\#H)$ and $\{Y_{\boldsymbol{t}} \in S_{(\#H)+1}\} = \{X_{\boldsymbol{t}} \in A\}$. For more details, see Davis and Mikosch (2009). □

2.6.2 Appendix B: Proof of Theorem 3

Theorem 3 is derived from Proposition 2.6.3 - 2.6.5. Before proceeding to Proposition 2.6.3, we present the following result regarding LUNC.

Proposition 2.6.2. *Consider a strictly stationary regularly varying random field* $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^d\}$ *with index* $\alpha > 0$ *satisfying LUNC. For a positive integer* k *and* $\lambda_n \to 0$,

$$nP\left(\frac{X_{\mathbf{0}}}{a_n} \in A_0, \frac{X_{\boldsymbol{s}_1+\lambda_n}}{a_n} \in A_1, \cdots, \frac{X_{\boldsymbol{s}_k+\lambda_n}}{a_n} \in A_k\right) \to \tau_{A_0,A_1,\cdots,A_k}(\boldsymbol{s}_1, \cdots, \boldsymbol{s}_k)$$

provided $A_0 \times A_1 \times \cdots \times A_k$ *is a continuity set of the limit measure* $\tau_{A_0,A_1,\cdots,A_k}(\boldsymbol{s}_1, \cdots, \boldsymbol{s}_k) = \lim_{n\to\infty} nP\left(X_{\mathbf{0}}/a_n \in A_0, X_{\boldsymbol{s}_1}/a_n \in A_1, \cdots, X_{\boldsymbol{s}_k}/a_n \in A_k\right)$.

Proof. Let f be a continuous function with compact support on $\bar{\mathbb{R}}^{k+1} \setminus \{\mathbf{0}\}$. Since f has compact support, it is uniformly continuous and hence for every $\epsilon > 0$ there exists δ such that $|f(x_1, x_2, \cdots, x_{k+1}) - f(y_1, y_2, \cdots, y_{k+1})| < \epsilon$ whenever $|(x_1, x_2, \cdots, x_{k+1}) - (y_1, y_2, \cdots, y_{k+1})| < \delta$.

Let $\tilde{X}_n = (X_{\mathbf{0}}, X_{\boldsymbol{s}_1+\lambda_n}, \cdots, X_{\boldsymbol{s}_k+\lambda_n})$ and $\tilde{X} = (X_{\mathbf{0}}, X_{\boldsymbol{s}_1}, \cdots, X_{\boldsymbol{s}_k})$. Notice that

$$\begin{aligned} nE\left|f\left(\frac{\tilde{X}_n}{a_n}\right) - f\left(\frac{\tilde{X}}{a_n}\right)\right| &= nE|\cdot|I_{\{\frac{|\tilde{X}_n - \tilde{X}|}{a_n} > \delta\}} + nE|\cdot|I_{\{\frac{|\tilde{X}_n - \tilde{X}|}{a_n} \le \delta\}} \\ &= A_1 + A_2. \end{aligned}$$

Let $M = \max f\left(\frac{\mathbf{x}}{a_n}\right)$. By (2.19), there exists $\epsilon > 0$ such that

$$\begin{aligned} \limsup_n A_1 &\le \limsup_n 2Mn\left[P\left(|X_{\boldsymbol{s}_1+\lambda_n} - X_{\boldsymbol{s}_1}| > \frac{\delta a_n}{k}\right) + \cdots \right. \\ &\qquad \left. + P\left(|X_{\boldsymbol{s}_k+\lambda_n} - X_{\boldsymbol{s}_k}| > \frac{\delta a_n}{k}\right)\right] \\ &< 2M\epsilon \end{aligned}$$

since $|X_{\lambda_n} - X_{\mathbf{0}}| \le \sup_{|\boldsymbol{s}|<\delta'} |X_{\boldsymbol{s}} - X_{\mathbf{0}}|$ as $n \to \infty$ for $|\lambda_n| < \delta'$. For A_2, since the support of $f \in \{|\tilde{X}| > C\} \subset \{|X_{\mathbf{0}}| > \frac{C}{k+1}\} \cup \cdots \cup \{|X_{\boldsymbol{s}_k}| > \frac{C}{k+1}\}$

$$\begin{aligned} \limsup_n A_2 &\le \limsup_n \epsilon n\left[P\left(\frac{|\tilde{X}_n|}{a_n} > C\right) + P\left(\frac{|\tilde{X}|}{a_n} > C\right)\right] \\ &= \limsup_n \epsilon\ n\ 2(k+1)\ P\left(|X_{\mathbf{0}}| > a_n C/(k+1)\right) \\ &= \epsilon\ 2(k+1)\tau_{BB}(\mathbf{0}) \qquad \text{where } B = \{x : x > C/(k+1)\}. \end{aligned}$$

Take ϵ small by choosing appropriate δ and δ', then for a positive integer k and $\lambda_n \to 0$,

$$nEf\left(\frac{X_{\mathbf{0}}, X_{\boldsymbol{s}_1+\lambda_n}, \cdots, X_{\boldsymbol{s}_k+\lambda_n}}{a_n}\right) \to \int f(u_1, u_2, \cdots, u_k)\mu(du_1, du_2, \cdots, du_k)$$

for any continuous function with compact support f. Using Portmanteau theorem for vague convergence, we complete the proof. See Theorem 3.2 in Resnick (2006). □

We discuss asymptotics of the denominator and the numerator of the ESE in turn.

Proposition 2.6.3. *Under the setting of Theorem 3 and condition (**M2**),*

$$E(\hat{p}_m(A)) = p_m(A) \to \mu(A) \quad \text{and} \quad \frac{|S_n|}{m_n} var(\hat{p}_m(A)) \to \frac{\mu(A)}{\nu} + \int_{\mathbb{R}^2} \tau_{AA}(\boldsymbol{y})d\boldsymbol{y}.$$

Hence, $\hat{p}_m(A) \xrightarrow{p} \mu(A)$.

Proof. By the regularly varying property, $E(\hat{p}_m(A)) = p_m(A) \to \mu(A)$.

For $var(\hat{p}_m(A))$, recall that $N^{(2)}(d\boldsymbol{s}_1, d\boldsymbol{s}_2) = N(d\boldsymbol{s}_1)N(d\boldsymbol{s}_2)I(\boldsymbol{s}_1 \neq \boldsymbol{s}_2)$ and observe that

$$\begin{aligned}
&E(\hat{p}_m(A)^2) \\
&= \left(\frac{m_n}{\nu|S_n|}\right)^2 E\left[\int_{S_n} I\left(\frac{X_{\boldsymbol{s}_1}}{a_m} \in A\right) N(d\boldsymbol{s}_1)\right. \\
&\qquad \left. + \int_{S_n}\int_{S_n} I\left(\frac{X_{\boldsymbol{s}_1}}{a_m} \in A, \frac{X_{\boldsymbol{s}_2}}{a_m} \in A\right) N^{(2)}(d\boldsymbol{s}_1, d\boldsymbol{s}_2)\right] \\
&= \left(\frac{m_n}{\nu|S_n|}\right)^2 \left[\int_{S_n} \frac{p_m(A)}{m_n}\nu d\boldsymbol{s}_1\right. \\
&\qquad \left. + \int_{S_n}\int_{S_n} \left[P\left(\frac{X_{\boldsymbol{s}_1}}{a_m} \in A, \frac{X_{\boldsymbol{s}_2}}{a_m} \in A\right) - \frac{p_m(A)^2}{m_n^2}\right] \nu^2 d\boldsymbol{s}_1 d\boldsymbol{s}_2\right] + E(\hat{p}_m(A))^2 \\
&= \left(\frac{m_n}{|S_n|}\right)\left[\frac{E(\hat{p}_m(A))}{\nu} + \int_{S_n - S_n} m_n \left[\frac{\tau_{AA,m}(\boldsymbol{h})}{m_n} - \frac{p_m(A)^2}{m_n^2}\right] \frac{|S_n \cap (S_n - \boldsymbol{y})|}{|S_n|} d\boldsymbol{y}\right] \\
&\qquad\qquad + E(\hat{p}_m(A))^2
\end{aligned}$$

where the change of variables $\boldsymbol{s}_2 - \boldsymbol{s}_1 = \boldsymbol{y}$ is used in the last line. Using the above, we show

$$\begin{aligned}
\frac{|S_n|}{m_n} var(\hat{p}_m(A)) &= \frac{E(\hat{p}_m(A))}{\nu} + \int_{S_n - S_n} m_n \left[\frac{\tau_{AA,m}(\boldsymbol{h})}{m_n} - \frac{p_m(A)^2}{m_n^2}\right] \frac{|S_n \cap (S_n - \boldsymbol{y})|}{|S_n|} d\boldsymbol{y} \\
&\to \frac{\mu(A)}{\nu} + \int_{\mathbb{R}^2} \tau_{AA}(\boldsymbol{y}) d\boldsymbol{y}. \qquad (2.48)
\end{aligned}$$

To see (2.48), notice that for a fixed $k > 0$

$$\begin{aligned}
&\int_{S_n - S_n} m_n \left[\frac{\tau_{AA,m}(\boldsymbol{h})}{m_n} - \frac{p_m(A)^2}{m_n^2}\right] \frac{|S_n \cap (S_n - \boldsymbol{y})|}{|S_n|} d\boldsymbol{y} \\
&= \int_{B[0,k)} [\cdot] d\boldsymbol{y} + \int_{B[k,r_n]} [\cdot] d\boldsymbol{y} + \int_{S_n - S_n \setminus B[0,r_n]} [\cdot] d\boldsymbol{y} \\
&= A_1 + A_2 + A_3.
\end{aligned}$$

For each fixed $k > 0$, $\lim_n A_1 = \int_{B[0,k)} \tau_{AA}(\boldsymbol{y}) d\boldsymbol{y}$ by Proposition 2.6.2. Now, we show

$$\lim_{k\to\infty} \limsup_{n\to\infty} (|A_2 + A_3|) = 0.$$

Recall that set A is bounded away from the origin. Using (2.13) and $r_n^2 = o(m_n)$,

$$\begin{aligned}
|A_2| &\leq \int_{B[k,r_n]} m_n P(|X_{\boldsymbol{y}}| > \epsilon a_m, |X_{\boldsymbol{0}}| > \epsilon a_m) d\boldsymbol{y} + const\ r_n^2 \frac{p_m(A)^2}{m_n} \\
&\to 0.
\end{aligned}$$

From (2.14), $\lim_n |A_3| \le \lim_n \int_{\mathbb{R}^2 \setminus B[0,r_n)} m_n \alpha_{1,1}(\boldsymbol{y}) d\boldsymbol{y} = 0$. This completes the proof. □

Proposition 2.6.4. *Assume that a stationary regularly varying random field satisfies LUNC. Further, assume the conditions of Proposition 2.6.3, and (2.17) in (**M3**). Then*

$$
\begin{aligned}
&(i)\ E\hat{\tau}_{AB,m}(\boldsymbol{h}) \to \tau_{AB}(\boldsymbol{h}),\\
&(ii)\ \frac{|S_n|\lambda_n^2}{m_n}\mathrm{cov}\left(\hat{\tau}_{AB,m}(\boldsymbol{h}_1), \hat{\tau}_{AB,m}(\boldsymbol{h}_2)\right)\\
&\qquad\qquad \to \frac{\int_{\mathbb{R}^2} w(\boldsymbol{y})^2 d\boldsymbol{y}}{\nu^2}\left[\tau_{AB}(\boldsymbol{h}_1)\, I_{\{\boldsymbol{h}_1=\boldsymbol{h}_2\}} + \tau_{A\cap BA\cap B}(\boldsymbol{h}_1)\, I_{\{\boldsymbol{h}_1=-\boldsymbol{h}_2\}}\right],\\
&(iii)\ \left(\frac{|S_n|\lambda_n^2}{m_n}\right)\mathrm{var}\left(\hat{\tau}_{AB,m}(\boldsymbol{h})\right) \to \frac{1}{\nu^2}\left(\int_{\mathbb{R}^2} w(\boldsymbol{y})^2 d\boldsymbol{y}\right)\tau_{AB}(\boldsymbol{h}).
\end{aligned}
$$

Proof. (i) From (2.12) and stationarity of $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^2\}$,

$$E\hat{\tau}_{AB,m}(\boldsymbol{h}) = \frac{m_n}{\nu^2}\frac{1}{|S_n|}\int_{S_n}\int_{S_n} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)\, P\left(\frac{X_{\boldsymbol{0}}}{a_m} \in A, \frac{X_{\boldsymbol{s}_2-\boldsymbol{s}_1}}{a_m} \in B\right)\nu^2 d\boldsymbol{s}_1 d\boldsymbol{s}_2$$

which after making the transformation $\frac{\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2}{\lambda_n} = \boldsymbol{y}$ and $\boldsymbol{s}_2 = \boldsymbol{u}$ becomes

$$
\begin{aligned}
& \frac{1}{|S_n|}\int_{\frac{S_n-S_n+\boldsymbol{h}}{\lambda_n}}\int_{S_n\cap(S_n-\lambda_n\boldsymbol{y}+\boldsymbol{h})} w(\boldsymbol{y})\, \tau_{AB,m}(\boldsymbol{h}-\boldsymbol{y}\lambda_n) d\boldsymbol{u} d\boldsymbol{y}\\
= & \int_{\frac{S_n-S_n+\boldsymbol{h}}{\lambda_n}} w(\boldsymbol{y})\, \tau_{AB,m}(\boldsymbol{h}-\boldsymbol{y}\lambda_n)\frac{|S_n\cap(S_n-\lambda_n\boldsymbol{y}+\boldsymbol{h})|}{|S_n|} d\boldsymbol{y}\\
\to & \tau_{AB}(\boldsymbol{h}).
\end{aligned}
$$

The limit in the last line follows from the dominated convergence theorem since

$$\tau_{AB,m}(\boldsymbol{h}-\boldsymbol{y}\lambda_n)\frac{|S_n\cap(S_n-\lambda_n\boldsymbol{y}+\boldsymbol{h})|}{|S_n|} \le\ p_m(A)$$

and

$$\int_{\mathbb{R}^2} w(\boldsymbol{y})\, p_m(A) d\boldsymbol{y} < \infty.$$

(ii) For sets A and B let $\tau^*_m(\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3,\boldsymbol{s}_4) = m_n P\left(\frac{X_{\boldsymbol{s}_1}}{a_m} \in A, \frac{X_{\boldsymbol{s}_2}}{a_m} \in B, \frac{X_{\boldsymbol{s}_3}}{a_m} \in A, \frac{X_{\boldsymbol{s}_4}}{a_m} \in B\right)$.

Then,

$$\frac{|S_n|\lambda_n^2}{m_n} E\left(\hat{\tau}_{AB,m}(\boldsymbol{h}_1)\hat{\tau}_{AB,m}(\boldsymbol{h}_2)\right) \tag{2.49}$$
$$= \frac{m_n\lambda_n^2}{\nu^4|S_n|} \iiiint_{S_n^4} w_n(h_1+\boldsymbol{s}_1-\boldsymbol{s}_2)w_n(h_2+\boldsymbol{s}_3-\boldsymbol{s}_4)$$
$$\frac{\tau_m^*(\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3,\boldsymbol{s}_4)}{m_n} E[N^{(2)}(d\boldsymbol{s}_1,d\boldsymbol{s}_2)N^{(2)}(d\boldsymbol{s}_3,d\boldsymbol{s}_4)],$$

where $N^{(2)}(d\boldsymbol{s}_1,d\boldsymbol{s}_2) = N(d\boldsymbol{s}_1)N(d\boldsymbol{s}_2)I(\boldsymbol{s}_1 \neq \boldsymbol{s}_2)$. From Karr (1986)

$$\begin{aligned}
&E[N^{(2)}(d\boldsymbol{s}_1,d\boldsymbol{s}_2)N^{(2)}(d\boldsymbol{s}_3,d\boldsymbol{s}_4)] \\
&\quad = \nu^4 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4 \\
&\quad +\nu^3 d\boldsymbol{s}_1 d\boldsymbol{s}_2 \varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_3)d\boldsymbol{s}_4 + \nu^3 d\boldsymbol{s}_1 d\boldsymbol{s}_2 \varepsilon_{\boldsymbol{s}_2}(d\boldsymbol{s}_3)d\boldsymbol{s}_4 \\
&\quad +\nu^3 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 \varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_4) + \nu^3 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 \varepsilon_{\boldsymbol{s}_2}(d\boldsymbol{s}_4) \\
&\quad +\nu^2 d\boldsymbol{s}_1 d\boldsymbol{s}_2 \varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_3)\varepsilon_{\boldsymbol{s}_2}(d\boldsymbol{s}_4) + \nu^2 d\boldsymbol{s}_1 d\boldsymbol{s}_2 \varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_4)\varepsilon_{\boldsymbol{s}_2}(d\boldsymbol{s}_3).
\end{aligned} \tag{2.50}$$

Now, let I_i, for $i = 1, ..., 7$, be the integral in (2.49) corresponding to these seven scenarios of $E[N^{(2)}(d\boldsymbol{s}_1,d\boldsymbol{s}_2)N^{(2)}(d\boldsymbol{s}_3,d\boldsymbol{s}_4)]$. The only cases that contribute to a non-zero limit are I_1, I_6, and I_7. For example, if $\boldsymbol{h}_1 = \boldsymbol{h}_2$,

$$\begin{aligned}
I_6 &= \frac{m_n\lambda_n^2}{\nu^4|S_n|} \iiiint_{S_n^4} w_n(\boldsymbol{h}_1+\boldsymbol{s}_1-\boldsymbol{s}_2)w_n(\boldsymbol{h}_2+\boldsymbol{s}_3-\boldsymbol{s}_4) \\
&\qquad \frac{\tau_m^*(\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3,\boldsymbol{s}_4)}{m_n} \nu^2 d\boldsymbol{s}_1 d\boldsymbol{s}_2 \varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_3)\varepsilon_{\boldsymbol{s}_2}(d\boldsymbol{s}_4) \\
&= \frac{\lambda_n^2}{\nu^2|S_n|} \iint_{S_n^2} w_n(\boldsymbol{h}_1+\boldsymbol{s}_1-\boldsymbol{s}_2)w_n(\boldsymbol{h}_1+\boldsymbol{s}_1-\boldsymbol{s}_2)\tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1)d\boldsymbol{s}_1 d\boldsymbol{s}_2 \qquad (2.51) \\
&= \frac{\lambda_n^2}{\nu^2} \int_{\frac{S_n-S_n+\boldsymbol{h}_1}{\lambda_n}} \frac{1}{\lambda_n^2} w(\boldsymbol{y})^2 \tau_{AB,m}(\boldsymbol{h}_1-\lambda_n\boldsymbol{y}) \frac{|S_n \cap (S_n+\boldsymbol{h}_1-\lambda_n\boldsymbol{y})|}{|S_n|} d\boldsymbol{y} \\
&\to \frac{1}{\nu^2}\left(\int_{\mathbb{R}^2} w(\boldsymbol{y})^2 d\boldsymbol{y}\right)\tau_{AB}(\boldsymbol{h}_1)
\end{aligned}$$

by taking $\boldsymbol{y} = \frac{\boldsymbol{h}_1+\boldsymbol{s}_1-\boldsymbol{s}_2}{\lambda_n}$ and $\boldsymbol{u} = \boldsymbol{s}_2$ in the last equation. The convergence is from the dominated convergence theorem. On the other hand, if $\boldsymbol{h}_1 \neq \boldsymbol{h}_2$, I_6 has the order of

$$\begin{aligned}
&\frac{\lambda_n^2}{\nu^2} \int_{\frac{S_n-S_n+\boldsymbol{h}_1}{\lambda_n}} \frac{1}{\lambda_n^2} w(\boldsymbol{y}) w\left(\boldsymbol{y}+\frac{\boldsymbol{h}_2-\boldsymbol{h}_1}{\lambda_n}\right) \tau_{AB,m}(\boldsymbol{h}_1-\lambda_n\boldsymbol{y}) \frac{|S_n \cap (S_n+\boldsymbol{h}_1-\lambda_n\boldsymbol{y})|}{|S_n|} d\boldsymbol{y} \\
\to\ & 0.
\end{aligned}$$

Similarly,

$$I_7 \to \frac{1}{\nu^2}\left(\int_{\mathbb{R}^2} w(\boldsymbol{y})^2 d\boldsymbol{y}\right)\tau_{A\cap BA\cap B}(\boldsymbol{h}_1). \tag{2.52}$$

Turning to I_1, we claim

$$\left|I_1 - \frac{|S_n|\lambda_n^2}{m_n} E\left(\hat{\tau}_{AB,m}(\boldsymbol{h}_1)\right) E\left(\hat{\tau}_{AB,m}(\boldsymbol{h}_2)\right)\right| \to 0. \tag{2.53}$$

To see this, observe that the left-hand side in (2.53) is bounded by

$$\begin{aligned}
&\frac{m_n\lambda_n^2}{\nu^4|S_n|}\iiiint_{S_n^4} w_n(\boldsymbol{h}_1+\boldsymbol{s}_1-\boldsymbol{s}_2)w_n(\boldsymbol{h}_2+\boldsymbol{s}_3-\boldsymbol{s}_4)\\
&\left|\frac{\tau_m^*(\boldsymbol{0},\boldsymbol{s}_2-\boldsymbol{s}_1,\boldsymbol{s}_3-\boldsymbol{s}_1,\boldsymbol{s}_4-\boldsymbol{s}_1)}{m_n} - \frac{\tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1)}{m_n}\frac{\tau_{AB,m}(\boldsymbol{s}_4-\boldsymbol{s}_3)}{m_n}\right| \nu^4 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4\\
\leq\ & \lambda_n^2 m_n \iiint_{(S_n-S_n)^3} w_n(\boldsymbol{h}_1-\boldsymbol{v}_1)w_n(\boldsymbol{h}_2-(\boldsymbol{v}_3-\boldsymbol{v}_2))\\
&\qquad\left|\frac{\tau_m^*(\boldsymbol{0},\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3)}{m_n} - \frac{\tau_{AB,m}(\boldsymbol{v}_1)}{m_n}\frac{\tau_{AB,m}(\boldsymbol{v}_3-\boldsymbol{v}_2)}{m_n}\right| d\boldsymbol{v}_1 d\boldsymbol{v}_2 d\boldsymbol{v}_3,
\end{aligned}$$

where the change of variables $\boldsymbol{v}_1 = \boldsymbol{s}_2 - \boldsymbol{s}_1, \boldsymbol{v}_2 = \boldsymbol{s}_3 - \boldsymbol{s}_1,$ and $\boldsymbol{v}_3 = \boldsymbol{s}_4 - \boldsymbol{s}_1$ are used. By taking $u = \boldsymbol{v}_2, \boldsymbol{y}_1 = \frac{\boldsymbol{h}_1-\boldsymbol{v}_1}{\lambda_n}$ and $y_2 = \frac{\boldsymbol{h}_2-(\boldsymbol{v}_3-\boldsymbol{v}_2)}{\lambda_n}$, the right-hand side of the inequality is equivalent to

$$\begin{aligned}
&\lambda_n^2 m_n \int_{\frac{(S_n-S_n)-(S_n-S_n)+\boldsymbol{h}_2}{\lambda_n}} \int_{\frac{S_n-S_n+\boldsymbol{h}_1}{\lambda_n}} \int_{S_n-S_n} w(\boldsymbol{y}_1)w(\boldsymbol{y}_2)\\
&\left|\frac{\tau_m^*(0,\boldsymbol{h}_1-\boldsymbol{y}_1\lambda_n,u,u+\boldsymbol{h}_2-\boldsymbol{y}_2\lambda_n)}{m_n} - \frac{\tau_{AB,m}(\boldsymbol{h}_1-\boldsymbol{y}_1\lambda_n)}{m_n}\frac{\tau_{AB,m}(\boldsymbol{h}_2-\boldsymbol{y}_2\lambda_n)}{m_n}\right| du d\boldsymbol{y}_1 d\boldsymbol{y}_2\\
&= \lambda_n^2 m_n\, O\left(\int_{\mathbb{R}^2}\alpha_{2,2}(\|\boldsymbol{y}\|)d\boldsymbol{y}\right).
\end{aligned} \tag{2.54}$$

To see (2.54), observe that

$$\begin{aligned}
&\min d(\{\boldsymbol{0},\boldsymbol{h}_1-\boldsymbol{y}_1\lambda_n\}\{u,u+\boldsymbol{h}_2-\boldsymbol{y}_2\lambda_n\})\\
&\quad\leq \|\boldsymbol{u}\| + \|\boldsymbol{u}-\boldsymbol{h}_1+\boldsymbol{y}_1\lambda_n\| + \|\boldsymbol{u}+\boldsymbol{h}_2-\boldsymbol{y}_2\lambda_n\| + \|\boldsymbol{u}+\boldsymbol{h}_2-\boldsymbol{y}_2\lambda_n\ -\boldsymbol{h}_1+\boldsymbol{y}_1\lambda_n\|.
\end{aligned}$$

Thus, the integral in (2.54) is bounded by

$$\begin{aligned}
&\int_{\mathbb{R}^2} \alpha_{2,2}(\|\boldsymbol{u}\|)d\boldsymbol{u} \left(\int_{\mathbb{R}^2} w(\boldsymbol{y}_1)d\boldsymbol{y}_1\right)^2 \\
&+ \int_{\frac{S_n - S_n + \boldsymbol{h}_1}{\lambda_n}} \int_{S_n - S_n} w(\boldsymbol{y}_1)\alpha_{2,2}(\|\boldsymbol{u} - \boldsymbol{h}_1 + \boldsymbol{y}_1\lambda_n\|)d\boldsymbol{u}d\boldsymbol{y}_1 \int_{\mathbb{R}^2} w(\boldsymbol{y}_2)d\boldsymbol{y}_2 \\
&+ \int_{\frac{S_n - S_n + \boldsymbol{h}_2}{\lambda_n}} \int_{S_n - S_n} w(\boldsymbol{y}_2)\alpha_{2,2}(\|\boldsymbol{u} - \boldsymbol{h}_2 + \boldsymbol{y}_2\lambda_n\|)d\boldsymbol{u}d\boldsymbol{y}_2 \int_{\mathbb{R}^2} w(\boldsymbol{y}_1)d\boldsymbol{y}_1 \\
&+ \int_{\frac{S_n - S_n + \boldsymbol{h}_2}{\lambda_n}} \int_{\frac{S_n - S_n + \boldsymbol{h}_1}{\lambda_n}} \int_{S_n - S_n} w(\boldsymbol{y}_1)w(\boldsymbol{y}_2)\alpha_{2,2}(\|\boldsymbol{u} + \boldsymbol{h}_2 - \boldsymbol{h}_1 - \boldsymbol{y}_2\lambda_n + \boldsymbol{y}_1\lambda_n\|)d\boldsymbol{u}d\boldsymbol{y}_1d\boldsymbol{y}_2 \\
&= A_1 + A_2 + A_3 + A_4.
\end{aligned}$$

Notice that $A_1 = \int_{R^2} \alpha_{2,2}(\|\boldsymbol{u}\|)d\boldsymbol{u}$. Also, by taking $\boldsymbol{x} = \boldsymbol{u} - \boldsymbol{h}_1 + \boldsymbol{y}_1\lambda_n$,

$$\begin{aligned}
A_2 &\leq \int_{\frac{S_n - S_n + \boldsymbol{h}_1}{\lambda_n}} \int_{\mathbb{R}^2} w(\boldsymbol{y}_1)\alpha_{2,2}(\|\boldsymbol{x}\|)d\boldsymbol{x}d\boldsymbol{y}_1 \\
&\leq \int_{\mathbb{R}^2} \alpha_{2,2}(\|\boldsymbol{x}\|)d\boldsymbol{x} \int_{\mathbb{R}^2} w(\boldsymbol{y}_1)d\boldsymbol{y}_1 \\
&= \int_{\mathbb{R}^2} \alpha_{2,2}(\|\boldsymbol{x}\|)d\boldsymbol{x}.
\end{aligned}$$

Similarly $A_3 \leq \int_{\mathbb{R}^2} \alpha_{2,2}(\|\boldsymbol{x}\|)d\boldsymbol{x}$ can be shown. Using the similar change of variable technique,

$$\begin{aligned}
A_4 &\leq \int_{\frac{S_n - S_n + \boldsymbol{h}_2}{\lambda_n}} \int_{\frac{S_n - S_n + \boldsymbol{h}_1}{\lambda_n}} \int_{S_n - S_n + \boldsymbol{h}_2 - \boldsymbol{h}_1 - \boldsymbol{y}_2\lambda_n + \boldsymbol{y}_1\lambda_n} w(\boldsymbol{y}_1)w(\boldsymbol{y}_2)\alpha_{2,2}(\|\boldsymbol{x}\|)d\boldsymbol{x}d\boldsymbol{y}_1d\boldsymbol{y}_2 \\
&\leq \int_{\mathbb{R}^2} \alpha_{2,2}(\|\boldsymbol{x}\|)d\boldsymbol{x}.
\end{aligned}$$

Hence, (2.54) is verified and (2.53) is proved.

Lastly, using the same argument in Lemma A.4. in Li et al. (2008), we have

$$I_j \to 0, \text{ if } j = 2, 3, 4, 5.$$

In fact,

$$I_j = O(\lambda_n^2)$$

for $j = 2, 3, 4, 5$. To see this, take one example of I_2, equivalent to

$$\frac{m_n\lambda_n^2}{\nu^4|S_n|} \int_{S_n} \int_{S_n} \int_{S_n} w_n(\boldsymbol{h}_1 + \boldsymbol{s}_1 - \boldsymbol{s}_2)w_n(\boldsymbol{h}_2 + \boldsymbol{s}_1 - \boldsymbol{s}_4) \frac{\tau_{A,B,B,m}(\boldsymbol{s}_2 - \boldsymbol{s}_1, \boldsymbol{s}_4 - \boldsymbol{s}_1)}{m_n} \nu^3 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_4.$$

After considering the change of variables of $\boldsymbol{s}_2 - \boldsymbol{s}_1 = \boldsymbol{v}_1, \boldsymbol{s}_4 - \boldsymbol{s}_1 = \boldsymbol{v}_2$, then consider $\frac{\boldsymbol{h}_1 - \boldsymbol{v}_1}{\lambda_n} = \boldsymbol{y}_1, \frac{\boldsymbol{h}_2 - \boldsymbol{v}_2}{\lambda_n} = \boldsymbol{y}_2$. Then,

$$\begin{aligned} I_2 &\leq \frac{p_m(A)\lambda_n^2}{\nu|S_n|} \int_{S_n - S_n} \int_{S_n - S_n} \int_{S_n} w_n(\boldsymbol{h}_1 - \boldsymbol{v}_1) w_n(\boldsymbol{h}_2 - \boldsymbol{v}_2) d\boldsymbol{u} d\boldsymbol{v}_1 d\boldsymbol{v}_2 \\ &\leq \frac{p_m(A)\lambda_n^2}{\nu} \left(\int_{\mathbb{R}^2} w(\boldsymbol{y}_1) d\boldsymbol{y}_1 \right)^2 \\ &= O(\lambda_n^2). \end{aligned}$$

Other cases can be shown in the same way. Thus, we conclude

$$I_j \to 0, \text{ if } j = 2, 3, 4, 5.$$

Combining the result (2.51)-(2.53), (ii) is proved, which completes the proof. □

Next, we establish the asymptotic normality for $\hat{\tau}_{AB,m}(\boldsymbol{h})$.

Proposition 2.6.5. *Assume that the conditions of Proposition 2.6.4 and (**M3**) hold. Then*

$$\sqrt{\frac{|S_n|\lambda_n^2}{m_n}} \left(\hat{\tau}_{AB,m}(\boldsymbol{h}) - E\hat{\tau}_{AB,m}(\boldsymbol{h}) \right) \to N(0, \sigma^2),$$

where $\sigma^2 = \frac{1}{\nu^2} \left(\int_{\mathbb{R}^2} w(\boldsymbol{y})^2 d(\boldsymbol{y}) \right) \tau_{AB}(\boldsymbol{h})$. *Furthermore,*

$$\sqrt{\frac{|S_n|\lambda_n^2}{m_n}} \left(\hat{\tau}_{AB,m}(\boldsymbol{h}) - \tau_{AB}(\boldsymbol{h}) \right) \to N(0, \sigma^2)$$

if $E\hat{\tau}_{AB,m}(\boldsymbol{h}) - \tau_{AB}(\boldsymbol{h}) = o\left(\left(\frac{|S_n|\lambda_n^2}{m_n} \right)^{-1/2} \right)$.

Proof. We follow Li et al. (2008) with focusing our attention to $\mathbb{R}^2$ and using a classical blocking technique. Let D_n^i be non-overlapping cubes that divide S_n for $i = 1, ..., k_n$, where $k_n = |S_n|/|D_n^i|$. Within each D_n^i, B_n^i is an inner cube sharing the same center and $d(\partial D_n^i, B_n^i) \geq n^\eta$. Let $|D_n^i| = n^{2\alpha}$ and $|B_n^i| = (n^\alpha - n^\eta)^2$ where $6/(2+\epsilon) < \eta < \alpha < 1$ for some $\epsilon > \frac{2+4\alpha}{\eta}$. Let k_n' be the additional number of cubes to cover S_n. From Lemma A.3. in Li et al. (2008),

$$k_n = O(n^{2(1-\alpha)}), \quad k_n' = O(n^{1-\alpha}). \tag{2.55}$$

Now define

$$
\begin{aligned}
A_n &= \sqrt{\frac{m_n\lambda_n^2}{|S_n|}}\frac{1}{\nu^2}\iint\limits_{S_n\times S_n} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)\, I\left(\frac{X_{\boldsymbol{s}_1}}{a_m}\in A\right) I\left(\frac{X_{\boldsymbol{s}_2}}{a_m}\in B\right) N^{(2)}(d\boldsymbol{s}_1, d\boldsymbol{s}_2),\\
a_{ni} &= \sqrt{\frac{m_n\lambda_n^2}{|S_n|}}\frac{1}{\nu^2}\iint\limits_{B_n^i\times B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)\, I\left(\frac{X_{\boldsymbol{s}_1}}{a_m}\in A\right) I\left(\frac{X_{\boldsymbol{s}_2}}{a_m}\in B\right) N^{(2)}(d\boldsymbol{s}_1, d\boldsymbol{s}_2),\\
&= \frac{1}{\sqrt{k_n}}\sqrt{\frac{m_n\lambda_n^2}{|B_n^i|}}\frac{1}{\nu^2}\iint\limits_{B_n^i\times B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)\, I\left(\frac{X_{\boldsymbol{s}_1}}{a_m}\in A\right) I\left(\frac{X_{\boldsymbol{s}_2}}{a_m}\in B\right) N^{(2)}(d\boldsymbol{s}_1, d\boldsymbol{s}_2),\\
\tilde{A}_n &= A_n - EA_n, \quad \tilde{a}_{ni} = a_{ni} - Ea_{ni}, \quad a_n = \sum_{i=1}^{k_n} a_{ni}, \quad \tilde{a}_n = \sum_{i=1}^{k_n}\tilde{a}_{ni}, \quad \text{and} \quad \tilde{a}_n' = \sum_{i=1}^{k_n}\tilde{a}_{ni}',
\end{aligned}
$$

where $\tilde{a}_{ni}'$ denotes an independent copy of $\tilde{a}_{ni}$.

Step 1. Show $var(\tilde{A}_n - \tilde{a}_n) \to 0$.

We will prove Step 1 by showing:

i) $var(\tilde{A}_n) \to \frac{1}{\nu^2}\left(\int_{\mathbb{R}^2} w(\boldsymbol{y})^2 d\boldsymbol{y}\right)\tau_{AB}(\boldsymbol{h})$,

ii) $cov(\tilde{A}_n, \tilde{a}_n) \to \frac{1}{\nu^2}\left(\int_{\mathbb{R}^2} w(\boldsymbol{y})^2 d\boldsymbol{y}\right)\tau_{AB}(\boldsymbol{h})$, and

iii) $var(\tilde{a}_n) \to \frac{1}{\nu^2}\left(\int_{\mathbb{R}^2} w(\boldsymbol{y})^2 d\boldsymbol{y}\right)\tau_{AB}(\boldsymbol{h})$.

i) This follows from Proposition 2.6.4 (iii).

ii) Recall $\tau_m^*(\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4)$ defined in Proposition 2.6.4 (ii). Then

$$
\begin{aligned}
& E\left(A_n a_n\right)\\
=\ & \frac{\lambda_n^2}{\nu^4|S_n|}\sum_{i=1}^{k_n}\iiiint\limits_{S_n\times S_n\times B_n^i\times B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)\, w_n(\boldsymbol{h}+\boldsymbol{s}_3-\boldsymbol{s}_4)\\
& \qquad \tau_m^*(\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4) E[N^{(2)}(d\boldsymbol{s}_1, d\boldsymbol{s}_2) N^{(2)}(d\boldsymbol{s}_3, d\boldsymbol{s}_4)]\\
=\ & \frac{\lambda_n^2}{\nu^4|S_n|}\sum_{i=1}^{k_n}\left[\iiiint\limits_{S_n\setminus B_n^i\times S_n\setminus B_n^i\times B_n^i\times B_n^i}\cdot + \iiiint\limits_{S_n\setminus B_n^i\times B_n^i\times B_n^i\times B_n^i}\cdot + \iiiint\limits_{B_n^i\times S_n\setminus B_n^i\times B_n^i\times B_n^i}\cdot + \iiiint\limits_{(B_n^i)^4}\cdot\right]\\
=\ & D_1 + D_2 + D_3 + D_4\\
=\ & \sum_{i=1}^{4}\sum_{j=1}^{7} D_i^j,
\end{aligned}
$$

where D_i^j be the integral in D_i corresponding to seven cases of $E[N^{(2)}(d\boldsymbol{s}_1, d\boldsymbol{s}_2)N^{(2)}(d\boldsymbol{s}_3, d\boldsymbol{s}_4)]$ as in (2.50) for $i = 1, ..., 4$ and $j = 1, ..., 7$. As shown in the proof of Proposition 2.6.4 (ii), non-zero contributions only arise when $j = 1, 6$, and 7. By the similar arguments in (2.53),

$$|\sum_{i=1}^{4} D_i^j - E(A_n)E(a_n)| \to 0.$$

Since $j = 6$ and 7 only occur when $\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4 \in B_n^i$, we only consider $D_4^6 + D_4^7$ which equals to

$$\begin{aligned}
&\frac{\lambda_n^2}{\nu^4|S_n|}\sum_{i=1}^{k_n}\iint_{(B_n^i)^2}\left[w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)^2 + w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)w_n(\boldsymbol{h}+\boldsymbol{s}_2-\boldsymbol{s}_1)\right] \\
&\qquad\qquad \tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1)\nu^2 d\boldsymbol{s}_1 d\boldsymbol{s}_2 \\
&= \frac{m_n\lambda_n^2}{\nu^2|B_n^1|}\iint_{(B_n^1)^2}\left[w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)^2 + w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)w_n(\boldsymbol{h}+\boldsymbol{s}_2-\boldsymbol{s}_1)\right] \\
&\qquad\qquad \tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1) d\boldsymbol{s}_1 d\boldsymbol{s}_2 \\
&\to \frac{1}{v^2}\int_{\mathbb{R}^2} w(\boldsymbol{y})^2 d\boldsymbol{y}\ \tau_{AB}(\boldsymbol{h}).
\end{aligned}$$

The convergence is derived from arguments in (2.51) and (2.53). Thus, we conclude

$$\begin{aligned}
cov(\tilde{A}_n, \tilde{a}_n) &= \left(\sum_{i=1}^{4}\sum_{j=1}^{7} D_i^j\right) - E(A_n)E(a_n) \\
&= D_4^6 + D_4^7 + o(1) \\
&\to \frac{1}{\nu^2}\left(\int_{\mathbb{R}^2} w(\boldsymbol{y})^2 d\boldsymbol{y}\right)\tau_{AB}(\boldsymbol{h}).
\end{aligned}$$

iii) Let $var(\tilde{a}_n) = \sum_{i=1}^{k_n} var(\tilde{a}_{ni}) + \sum_{1\le i\ne j\le k_n} cov(\tilde{a}_{ni}, \tilde{a}_{nj})$. Note from Proposition 2.6.4 (iii) that

$$\sum_{i=1}^{k_n} var(\tilde{a}_{ni}) = k_n var(a_{n1}) \to \frac{1}{\nu^2}\left(\int_{\mathbb{R}^2} w(\boldsymbol{y})^2 d\boldsymbol{y}\right)\tau_{AB}(\boldsymbol{h}).$$

Also note that since $\tilde{a}_{ni}$ and $\tilde{a}_{nj}$ are integrals over disjoint sets for $i \ne j$ and X_s is independent of N, $E[\tilde{a}_{ni}|N]$ and $E[\tilde{a}_{nj}|N]$ are independent. Thus,

$$\begin{aligned}
\sum_{1\le i\ne j\le k_n} |cov(\tilde{a}_{ni}, \tilde{a}_{nj})| &= \sum_{1\le i\ne j\le k_n} |E\{cov(\tilde{a}_{ni}, \tilde{a}_{nj}|N)\} + cov\{E(\tilde{a}_{ni}|N), E(\tilde{a}_{nj}|N)\}| \\
&= \sum_{1\le i\ne j\le k_n} |E\{cov(\tilde{a}_{ni}, \tilde{a}_{nj}|N)\}|.
\end{aligned}$$

Notice from Proposition 2.6.1 and $|a_{ni}| \leq \sqrt{\frac{m_n\lambda_n^2}{|S_n|}}|B_n^i|$ that

$$\begin{aligned} E\{cov(\tilde{a}_{ni}, \tilde{a}_{nj}|N)\} &\leq const\ \frac{m_n\lambda_n^2}{|S_n|}|B_n^i||B_n^j|\ \ |E(\alpha_{M,M}(n^\eta)|N)| \\ &\leq const\ \frac{m_n\lambda_n^2}{|S_n|}|B_n^1|^2\ E(M^2)n^{-\epsilon\eta} \end{aligned}$$

where $M = \max\{N(B_n^i), N(B_n^j)\}$ and the last inequality is from (2.18). Since $k_n = |S_n|/|D_n^1|$ where $|S_n| = n^2, |D_n^1| = n^{2\alpha}, |B_n^1| = O(n^{2\alpha})$,

$$\begin{aligned} \sum_{1\leq i\neq j\leq k_n} |cov(\tilde{a}_{ni}, \tilde{a}_{nj})| &\leq const\ k_n^2\frac{m_n\lambda_n^2}{|S_n|}|B_n^1|^2\ |B_n^1|^2 n^{-\epsilon\eta} \\ &= O\left(m_n\lambda_n^2 n^{2+4\alpha-\epsilon\eta}\right) \end{aligned}$$

which converges to 0 as $m_n\lambda_n^2 \to 0$ and $\epsilon > \frac{2+4\alpha}{\eta}$.

Step 2. Show $|\phi_n(x) - \phi_n'(x)| \to 0$ where $\phi_n(x)$ and $\phi_n'(x)$ are the characteristic functions of $\tilde{a}_n$ and $\tilde{a}_n'$.

Analogously to the idea presented in (6.2) in Davis and Mikosch (2009),

$$\begin{aligned} |\phi_n(x) - \phi_n'(x)| &= \left|\sum_{l=1}^{k_n} E\Pi_{j=1}^{l-1} e^{ix\frac{\tilde{a}_{nj}}{\sqrt{k_n}}}\left(e^{ix\frac{\tilde{a}_{nl}}{\sqrt{k_n}}} - e^{ix\frac{\tilde{a}'_{nl}}{\sqrt{k_n}}}\right)\Pi_{j=l+1}^{k_n} e^{ix\frac{\tilde{a}'_{nj}}{\sqrt{k_n}}}\right| \\ &\leq \sum_{l=1}^{k_n}\left|cov(\Pi_{j=1}^{l-1}e^{ix\frac{\tilde{a}_{nj}}{\sqrt{k_n}}}, e^{ix\frac{\tilde{a}_{nl}}{\sqrt{k_n}}})\right|. \end{aligned}$$

Using the same technique in Step 1 iii),

$$\begin{aligned} \left|cov(\Pi_{j=1}^{l-1}e^{ix\frac{\tilde{a}_{nj}}{\sqrt{k_n}}}, e^{ix\frac{\tilde{a}_{nl}}{\sqrt{k_n}}})\right| &\leq |E\ cov(\Pi_{j=1}^{l-1}e^{ix\frac{\tilde{a}_{nj}}{\sqrt{k_n}}}, e^{ix\frac{\tilde{a}_{nl}}{\sqrt{k_n}}}|N)| \\ &\leq c\ E\left(\alpha_{M,M}(n^\eta)\right) \\ &\leq c\ E\left(M^2\right)n^{-\epsilon\eta} \\ &\leq c\ l^2|n^{2\alpha}|^2 n^{-\epsilon\eta}, \end{aligned}$$

where $M = N(\cup_{j=1}^{l} B_n^j)$. The second and the last inequality is from (2.18) and $|B_n^1| = O(n^{2\alpha})$ respectively. Hence, from $k_n = n^{2-2\alpha}$, we have

$$\begin{aligned} |\phi_n(x) - \phi_n'(x)| &\leq const\ \sum_{l=1}^{k_n} l^2|n^{2\alpha}|^2 n^{-\epsilon\eta} \\ &\leq O(n^{6-2\alpha-\epsilon\eta}) \end{aligned}$$

which converges to 0 by setting $6/(2+\epsilon) < \eta < \alpha < 1$.

Step 3. Show the central limit theorem holds for $\tilde{a}'_n$.

Let $I_{ni} = \int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)\; I\left(\frac{X_{\boldsymbol{s}_1}}{a_m}\in A\right) I\left(\frac{X_{\boldsymbol{s}_2}}{a_m}\in B\right) N^{(2)}(d\boldsymbol{s}_1, d\boldsymbol{s}_2)$. By (2.16), we have

$$\begin{aligned} E|\sqrt{k_n}\tilde{a}'_{ni}|^{2+\delta} &= E\left|\sqrt{\frac{m_n\lambda_n^2}{|B_n^i|}}\frac{1}{\nu^2}\left[I_{ni}-E(I_{ni})\right]\right|^{2+\delta} \\ &= E\left|\sqrt{\frac{|B_n^i|\lambda_n^2}{m_n}}\left[\hat{\tau}_{AB,m}(\boldsymbol{h}:B_n^i)-E(\hat{\tau}_{AB,m}(\boldsymbol{h}:B_n^i))\right]\right|^{2+\delta} \\ &< C_\delta. \end{aligned}$$

As $(\tilde{a}'_{ni})$ is triangular array of independent random variables with $var(\sum_{i=1}^{k_n}\tilde{a}'_{ni}) = \sigma_n^2 \to \sigma^2$,

$$\frac{\sum_{i=1}^{k_n} E|\tilde{a}'_{ni}|^{2+\delta}}{(\sigma_n)^{2+\delta}} \le \frac{k_n k_n^{-(1+\delta/2)}C_\delta}{(\sigma_n)^{2+\delta}} \to 0.$$

Thus, Lyapunov's condition is satisfied and hence the central limit theorem holds. □

From the results of Proposition 2.6.3 - 2.6.5, Theorem 3 is proved.

Proof of Theorem 3. Proposition 2.6.3 implies that

$$\hat{p}_m(A) \xrightarrow{p} \mu(A).$$

By Slutsky's theorem and Proposition 2.6.5,

$$\begin{aligned} \sqrt{\frac{|S_n|\lambda_n^2}{m_n}}\left(\frac{\hat{\tau}_{AB,m}(\boldsymbol{h})}{\hat{p}_m(A)}-\frac{\tau_{AB,m}(\boldsymbol{h})}{\hat{p}_m(A)}\right) &= \sqrt{\frac{|S_n|\lambda_n^2}{m_n}}\left(\hat{\rho}_{AB,m}(\boldsymbol{h})-\frac{\tau_{AB,m}(\boldsymbol{h})}{\hat{p}_m(A)}\right) \\ &\to N(0,\sigma^2/\mu(A)^2). \end{aligned}$$

Recall from Proposition 2.6.3 that $var(\hat{p}_m(A)) = O\left(m_n/|S_n|\right)$, thus

$$\sqrt{\frac{|S_n|\lambda_n^2}{m_n}}\left(\hat{\rho}_{AB,m}(\boldsymbol{h})-\frac{\tau_{AB,m}(\boldsymbol{h})}{\hat{p}_m(A)}\right) = \sqrt{\frac{|S_n|\lambda_n^2}{m_n}}\left(\hat{\rho}_{AB,m}(\boldsymbol{h})-\rho_{AB,m}(\boldsymbol{h})\right)+o_p(1).$$

Thus, the central limit theorem for $\sqrt{\frac{|S_n|\lambda_n^2}{m_n}}\left(\hat{\rho}_{AB,m}(\boldsymbol{h})-\rho_{AB,m}(\boldsymbol{h})\right)$ is proved. The joint normality (2.20) is established using the Cramér-Wold device. □

2.6.3 Appendix C: Proof of Example 3

First, we show that X_s satisfies LUNC in (2.19). Notice that the process has continuous sample paths a.s. since the Gaussian process $\{W_{\boldsymbol{s}} - \delta(\boldsymbol{s}), \boldsymbol{s} \in \mathbb{R}^2\}$ in (2.29) has continuous sample paths. Notice from Lindgren (2012), Section 2.2, that a Gaussian process with a continuous correlation function satisfying (2.36) has continuous sample paths.

From (2.29), let $X_{\boldsymbol{s}} = U^1_{\boldsymbol{s}} \vee U^2_{\boldsymbol{s}}$, where $U^1_{\boldsymbol{s}} = \Gamma_1^{-1} Y^1_{\boldsymbol{s}}$ and $U^2_{\boldsymbol{s}} = \sup_{j \geq 2} \Gamma_j^{-1} Y^j_{\boldsymbol{s}}$. Then

$$
\begin{aligned}
& nP\left(\sup_{\|\boldsymbol{s}\| < \delta'} \frac{|X_{\boldsymbol{s}} - X_{\mathbf{0}}|}{a_n} > \delta \right) \\
&= nP\left(\sup_{\|\boldsymbol{s}\| < \delta'} |U^1_{\boldsymbol{s}} \vee U^2_{\boldsymbol{s}} - U^1_{\mathbf{0}} \vee U^2_{\mathbf{0}}| > a_n \delta \right) \\
&\leq nP\left(\sup_{\|\boldsymbol{s}\| < \delta'} |U^1_{\boldsymbol{s}} - U^1_{\mathbf{0}}| > \frac{a_n \delta}{2} \right) + nP\left(\sup_{\|\boldsymbol{s}\| < \delta'} |U^2_{\boldsymbol{s}} - U^2_{\mathbf{0}}| > \frac{a_n \delta}{2} \right) \\
&= A_1 + A_2.
\end{aligned}
$$

Since $E|\sup_{\|\boldsymbol{s}\| < \delta'} |Y(\boldsymbol{s})|| < \infty$ (see Proposition 13 in Kabluchko et al. (2009)), we can apply the dominated convergence theorem to obtain

$$
\begin{aligned}
A_1 &= nP\left(\Gamma_1 < \frac{2 \sup_{\|\boldsymbol{s}\| < \delta'} |Y^1_{\boldsymbol{s}} - Y^1_{\mathbf{0}}|}{\delta a_n} \right) \\
&= n \int \left(1 - e^{-z/\delta a_n}\right) g(Z) dZ \\
&\to \frac{2E(\sup_{\|\boldsymbol{s}\| < \delta'} |Y_{\boldsymbol{s}} - Y_{\mathbf{0}}|)}{\delta} \\
&\to 0,
\end{aligned}
$$

where $Z = 2 \sup_{\|\boldsymbol{s}\| < \delta'} |Y_{\boldsymbol{s}} - Y_{\mathbf{0}}|$.

To show $A_2 \to 0$, we follow the arguments in Davis and Mikosch (2008).

$$\begin{aligned} A_2 &= nP\left(\sup_{\|\boldsymbol{s}\|<\delta'} \bigvee_{j\geq 2}^{\infty} \Gamma_j^{-1}|Y_{\boldsymbol{s}}^j - Y_{\boldsymbol{s}}^j| > \frac{a_n\delta}{2}\right) \\ &\leq n\sum_{j=2}^{\infty} P\left(2\sup_{\|\boldsymbol{s}\|<\delta'}|Y_{\boldsymbol{s}}| > \Gamma_j\delta a_n/2\right) \\ &= n\int\left(\sum_{j\geq 2}^{\infty} P\left(4y > \Gamma_j\delta a_n\right)\right) P\left(\sup_{\|\boldsymbol{s}\|<\delta'}|Y_{\boldsymbol{s}}| \in dy\right) \\ &= n\int_0^{\infty}\left(\frac{4y}{\delta a_n} - \left(1-e^{-\frac{4y}{\delta a_n}}\right)\right) P\left(\sup_{\|\boldsymbol{s}\|<\delta'}|Y_{\boldsymbol{s}}| \in dy\right) \end{aligned}$$

The last line is from $ET[0,\frac{4y}{\delta a_n}] = \sum_{j=1}^{\infty} P\left(\Gamma_j < \frac{4y}{\delta a_n}\right) = \frac{4y}{\delta a_n}$, where $T = \sum_{j=1}^{\infty}\epsilon_{\Gamma_j}$ is a homogeneous point process. The dominated convergence theorem applies as $f_n(y) = n\left(\frac{4y}{\delta a_n} - (1-e^{-\frac{4y}{\delta a_n}})\right) \leq cy$ for some $c > 0$, all $y > 0$ and $f_n(y) \to 0$ as $n \to 0$, and $E\sup_{\|\boldsymbol{s}\|<\delta'}|Y_{\boldsymbol{s}}| < \infty$ from Kabluchko et al. (2009).

Now we check conditions (2.13)- (2.18). Recall from (2.35) that

$$\alpha_{c,c}(\boldsymbol{h}) \leq const\ \frac{1}{\sqrt{\|\boldsymbol{h}\|^{\alpha}}} e^{-\theta\|\boldsymbol{h}\|^{\alpha}/2}$$

holds for the process. For convenience in the calculations, set $g(\boldsymbol{h}) = \frac{1}{\sqrt{\|\boldsymbol{h}\|^{\alpha}}} e^{-\theta\|\boldsymbol{h}\|^{\alpha}/2}$. We will find the sufficient conditions for (2.13)- (2.18). For (2.13),

$$\int_{\mathbb{R}^2} g(\boldsymbol{y})d\boldsymbol{y} < \infty \tag{2.56}$$

is sufficient. To see this, infer from (2.34) that

$$\begin{aligned} m_n P(X_{\boldsymbol{y}} > \epsilon a_m, X_{\boldsymbol{0}} > \epsilon a_m) &= m_n\left[1 - 2e^{-1/a_m} + e^{-2\Phi(\sqrt{\delta(\boldsymbol{h})/2})/a_m}\right] \\ &= \frac{2m_n}{a_m}\bar{\Phi}(\sqrt{\delta(\boldsymbol{h})/2}) + O\left(\frac{m_n}{a_m^2}\right). \end{aligned}$$

Thus,

$$\begin{aligned} m_n\int_{B[k,r_n]} P(X_{\boldsymbol{y}} > \epsilon a_m, X_{\boldsymbol{0}} > \epsilon a_m)d\boldsymbol{y} &= \int_{B[k,r_n]} \frac{2m_n}{a_m}\bar{\Phi}(\sqrt{\delta(\boldsymbol{y})/2})d\boldsymbol{y} + O\left(\frac{r_n^2}{m_n}\right) \\ &\leq const\int_{B[k,\infty]} g(\boldsymbol{y})d\boldsymbol{y} + o(1), \end{aligned}$$

where the last inequality is from (2.35).

Using (2.35), condition (2.14) is satisfied if

$$\int_{\mathbb{R}^2 \setminus B[0,r_n)} m_n g(\boldsymbol{y}) d\boldsymbol{y} \to 0. \tag{2.57}$$

Similarly, using (2.35), the second condition in (2.17) is implied if (2.56) holds. The condition (2.18) is checked immediately from (2.22) as

$$\sup_l \frac{\alpha_{l,l}(\|\boldsymbol{h}\|)}{l^2} \leq c \, \frac{1}{\sqrt{\|\boldsymbol{h}\|^\alpha}} e^{-\theta \|\boldsymbol{h}\|^\alpha/2} = O(\|\boldsymbol{h}\|^{-\epsilon}).$$

The condition (2.16) with $\delta = 1$ is satisfied if (2.37) is assumed. See Appendix D in Section 2.6.4. Hence, it suffices to find conditions under which (2.56) - (2.57) hold.

Proposition 2.6.6. *For Example 3, the conditions (2.56) - (2.57) hold if* $\log m_n = o(r_n^a)$.

Proof. First, (2.56) is satisfied for $a \in (0, 2]$ since

$$\int_{\mathbb{R}^2} g(\boldsymbol{y}) d\boldsymbol{y} < \Gamma\left(\frac{1}{a} + \frac{1}{2}\right) < \infty$$

which can be shown using change of variables to polar coordinates and $r^a/2 = t$:

$$\begin{aligned}
\int_{\mathbb{R}^2} g(y) dy &\leq const \int_{\mathbb{R}^2} \|\boldsymbol{y}\|^{-a/2} e^{-\|\boldsymbol{y}\|^a/2} d\boldsymbol{y} \\
&\leq const \int_0^\infty r^{1-a/2} e^{-r^a/2} dr \\
&= const \int_0^\infty t^{\frac{1}{a}-\frac{1}{2}} e^{-t} dt \\
&= const \, \Gamma\left(\frac{1}{a} + \frac{1}{2}\right) \\
&< \infty
\end{aligned}$$

for $a \in (0, 2]$.

For (2.57), notice that $\delta(r_n) \geq 1$ for sufficiently large n. Thus, $m_n g(r_n) \leq m_n e^{-\theta r_n^\alpha/2} = o(1)$ provided $\log m_n = o(r_n^a)$. This completes the proof. □

Finally, we find the condition under which (2.21) holds.

Proposition 2.6.7. *For the Brown-Resnick process, (2.21) holds if* $\frac{|S_n|\lambda_n^2}{m_n^3} \to 0$.

Proof. From (2.34),

$$\begin{aligned} |\rho_{AB,m}(\boldsymbol{h}) - \rho_{AB}(\boldsymbol{h})| &= \frac{1+o(1)}{\mu(A)} |\tau_{AB,m}(\boldsymbol{h})\mu(A) - \tau_{AB}(\boldsymbol{h}) p_m(A)| \\ &= \frac{1+o(1)}{\mu(\boldsymbol{h})} O(m_n/a_m^2) \\ &= O(1/m_n). \end{aligned}$$

Therefore, (2.21) holds if $\frac{|S_n|\lambda_n^2}{m_n^3} \to 0$. □

2.6.4 Appendix D: the condition (2.16) with delta = 1

Here, we show

$$E|\sqrt{k_n}\tilde{a}'_{ni}|^3 < \infty \tag{2.58}$$

under the following assumptions.

$$\frac{\lambda_n^2 n^{2a}}{m_n} = O(1), \quad \int_{\mathbb{R}^2} \tau_{AA,m}(\boldsymbol{y}) d\boldsymbol{y} < \infty, \quad \text{and} \quad \int_{\mathbb{R}^2} \tau_{AB,m}(\boldsymbol{y}) d\boldsymbol{y} < \infty.$$

The first condition is derived from combining the following conditions

$$\sup_n \frac{\lambda_n^2 |B_n^i|}{m_n} < \infty \text{ and } \sup_n \frac{m_n}{\lambda_n^2 |B_n^i|} < \infty,$$

where $|B_n^i| = n^{2a}$ for $0 < a < 1$.

To see (2.58), first observe that

$$
\begin{aligned}
& E|\sqrt{k_n}\tilde{a}'_{ni}|^3 \\
= & \left(\sqrt{\frac{m_n\lambda_n^2}{|B_n^i|}}\frac{1}{\nu^2}\right)^3 \\
& \quad E\left[\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)\, I\left(\frac{X_{\boldsymbol{s}_1}}{a_m}\in A\right) I\left(\frac{X_{\boldsymbol{s}_2}}{a_m}\in B\right) N^{(2)}(d\boldsymbol{s}_1, d\boldsymbol{s}_2)\right. \\
& \qquad \left. - \int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)\, \frac{\tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1)}{m_n}\nu^2 d\boldsymbol{s}_1 d\boldsymbol{s}_2\right]^3 \\
= & \left(\sqrt{\frac{m_n\lambda_n^2}{|B_n^i|}}\frac{1}{\nu^2}\right)^3 E\left[A-E(A)\right]^3 \\
= & \left(\sqrt{\frac{m_n\lambda_n^2}{|B_n^i|}}\frac{1}{\nu^2}\right)^3 \left[EA^3 - 3EA^2E(A) + 2E(A)^3\right] \\
= & A_1 - 3A_2 + 2A_3.
\end{aligned}
$$

Step 1: Show $A_3 < \infty$ if $\sup_n \dfrac{\lambda_n^2|B_n^i|}{m_n} < \infty$.

Note that

$$
\begin{aligned}
A_3 &= \left(\sqrt{\frac{m_n\lambda_n^2}{|B_n^i|}}\frac{1}{\nu^2}\right)^3 \left(\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)\frac{\tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1)}{m_n}\nu^2 d\boldsymbol{s}_1 d\boldsymbol{s}_2\right)^3 \\
&= \left(\sqrt{\frac{\lambda_n^2}{|B_n^i|m_n}}\right)^3 |B_n^i|^3 \left(\frac{1}{|B_n^i|}\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)\tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1) d\boldsymbol{s}_1 d\boldsymbol{s}_2\right)^3 \\
&= \left(\sqrt{\frac{\lambda_n^2|B_n^i|}{m_n}}^3\right) (\cdot)^3 .
\end{aligned}
$$

To show $(\cdot)$ is bounded, we use the change of variable technique. As the same techinque is used later repeatedly, we decide to describe the detailed step here.

Claim 2.6.1. *Using the change of variables with* $\frac{\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2}{\lambda_n}$*, the below holds.*

$$
\frac{1}{|B_n^i|}\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)\tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1) d\boldsymbol{s}_1 d\boldsymbol{s}_2 \leq \infty. \tag{2.59}
$$

Proof. Let $\frac{\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2}{\lambda_n}=\boldsymbol{y}$ and $\boldsymbol{s}_1=\boldsymbol{u}$. Then, the left-hand side of (2.59) is bounded since

$$\begin{aligned}\frac{1}{|B_n^i|}\int_{B_n^i}\int_{\frac{\boldsymbol{h}+B_n^i-B_n^i}{\lambda_n}} w(\boldsymbol{y})\tau_{AB,m}(\boldsymbol{h}-\boldsymbol{y}\lambda_n)d\boldsymbol{y}d\boldsymbol{u} &\leq \frac{1}{|B_n^i|}\int_{B_n^i}\int_{\frac{\boldsymbol{h}+B_n^i-B_n^i}{\lambda_n}} w(\boldsymbol{y})p_m(A)d\boldsymbol{y}d\boldsymbol{u}\\ &\leq \frac{1}{|B_n^i|}\int_{B_n^i}\int_{R^2} w(\boldsymbol{y})p_m(A)d\boldsymbol{y}d\boldsymbol{u}\\ &\leq \frac{p_m(A)}{|B_n^i|}\int_{B_n^i} d\boldsymbol{u}\\ &< \infty.\end{aligned}$$

This completes the proof. □

Using Claim 2.6.1, we conclude that A_3 is finite provided $\sup_n \frac{\lambda_n^2|B_n^i|}{m_n}<\infty$.

Step 2: Show $A_2<\infty$ if $\sup_n \frac{\lambda_n^2|B_n^i|}{m_n}<\infty$ and $\int_{R^2}\tau_{AB,m}(y)dy<\infty$.

By the similar idea in (2.53), we consider the seven cases of $E[N^{(2)}(ds_1,ds_2)N^{(2)}(ds_3,ds_4)]$ for EA^2 as in (2.50). Say $A_2^i, i=1,...7$.

Case 1) $A_2^i, i=6,7$.

Using the similar technique in Claim 2.6.1,

$$\begin{aligned}A_2^6 &= \left(\sqrt{\frac{m_n\lambda_n^2}{|B_n^i|}}\frac{1}{\nu^2}\right)^3\int_{B_n^i}\int_{B_n^i}\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)^2 w_n(h+\boldsymbol{s}_3-\boldsymbol{s}_4)\\ &\qquad \frac{\tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1)}{m_n}\frac{\tau_{AB,m}(\boldsymbol{s}_4-\boldsymbol{s}_3)}{m_n}\nu^4 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4\\ &= \left(\sqrt{\frac{\lambda_n^2}{|B_n^i|}}\right)^3\left(\frac{1}{\nu^2\sqrt{m_n}}\right)\int_{B_n^i}\int_{B_n^i}\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)^2 w_n(h+\boldsymbol{s}_3-\boldsymbol{s}_4)\\ &\qquad \tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1)\tau_{AB,m}(\boldsymbol{s}_4-\boldsymbol{s}_3)d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4\\ &\leq \text{const}\left(\sqrt{\frac{\lambda_n^2}{|B_n^i|}}\right)^3\frac{|B_n^i|^2}{\nu^2\sqrt{m_n}\lambda_n^2}.\end{aligned}$$

Now take $\frac{\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2}{\lambda_n}=\boldsymbol{y}_1, \boldsymbol{s}_1=\boldsymbol{u}_1, \frac{\boldsymbol{h}+\boldsymbol{s}_3-\boldsymbol{s}_4}{\lambda_n}=\boldsymbol{y}_2,$ and $\boldsymbol{s}_4=\boldsymbol{u}_2$, then

$$\begin{aligned}
&\int_{B_n^i}\int_{B_n^i}\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)^2 w_n(\boldsymbol{h}+\boldsymbol{s}_3-\boldsymbol{s}_4) \\
&\qquad\qquad \tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1)\tau_{AB,m}(\boldsymbol{s}_4-\boldsymbol{s}_3)d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4 \\
&= \int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)^2 \tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1)d\boldsymbol{s}_1 d\boldsymbol{s}_2 \\
&\qquad\qquad \int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_3-\boldsymbol{s}_4)\tau_{AB,m}(\boldsymbol{s}_4-\boldsymbol{s}_3)d\boldsymbol{s}_3 d\boldsymbol{s}_4
\end{aligned}$$

which is bounded by

$$\int_{B_n^i}\int_{\mathbb{R}^2} \frac{w(\boldsymbol{y}_1)^2}{\lambda_n^2} p_m(A) d\boldsymbol{y}_1 d\boldsymbol{u}_1 \ |B_n^i| \leq \frac{|B_n^i|^2}{\lambda_n^2} \int_{\mathbb{R}^2} w(\boldsymbol{y}_1)^2 d\boldsymbol{y}_1 \leq \text{const}\ \frac{|B_n^i|^2}{\lambda_n^2},$$

where the first and the second inequality are the direct application of the logic of Claim 2.6.1 to the two double integrals. Hence, the term is is finite if $\sup_n \frac{\lambda_n^2|B_n^i|}{m_n} < \infty$. The same argument is applied for A_2^7, so we skip this.

Case 2) $A_2^i, i=2,3,4,$ and 5.

We will only consider $i=2$ since $A_2^i, i=3,4,$ and 5 can be shown similarly.

$$\begin{aligned}
A_2^2 &= \left(\sqrt{\frac{m_n\lambda_n^2}{|B_n^i|}}\frac{1}{\nu^2}\right)^3 \left(\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_5-\boldsymbol{s}_6)\ \frac{\tau_{AB,m}(\boldsymbol{s}_6-\boldsymbol{s}_5)}{m_n}\nu^2 d\boldsymbol{s}_5 d\boldsymbol{s}_6\right) \\
&\qquad \int_{B_n^i}\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}_1+\boldsymbol{s}_1-\boldsymbol{s}_2)w_n(\boldsymbol{h}_2+\boldsymbol{s}_1-\boldsymbol{s}_4) \\
&\qquad\qquad \frac{\tau_{A,B,B,m}(\boldsymbol{s}_2-\boldsymbol{s}_1,\boldsymbol{s}_4-\boldsymbol{s}_1)}{m_n}\nu^3 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_4 \\
&\leq \left(\sqrt{\frac{\lambda_n^2}{|B_n^i|}}\right)^3 \left(\frac{1}{\nu\sqrt{m_n}}\right)\ |B_n^i| \\
&\qquad \int_{B_n^i}\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}_1+\boldsymbol{s}_1-\boldsymbol{s}_2)w_n(\boldsymbol{h}_2+\boldsymbol{s}_1-\boldsymbol{s}_4)\tau_{A,B,B,m}(\boldsymbol{s}_2-\boldsymbol{s}_1,\boldsymbol{s}_4-\boldsymbol{s}_1)d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_4.
\end{aligned}$$

The right-hand side is bounded by $\left(\sqrt{\frac{\lambda_n^2}{|B_n^i|}}\right)^3 \left(\frac{1}{\nu\sqrt{m_n}}\right)\ |B_n^i|^2$, where $|B_n^i|$ in the first inequality is from applying Claim 2.6.1 to the double integral. Similarly, the last inequality is from the change of variables with $\frac{\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2}{\lambda_n}=\boldsymbol{y}_1, \frac{\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_4}{\lambda_n}=\boldsymbol{y}_2,$ and $\boldsymbol{s}_1=\boldsymbol{u}$. Hence, $A_2^2<\infty$ if $\sup_n \frac{\lambda_n^2|B_n^i|}{m_n} < \infty$.

Case 3) A_2^1.

$$\begin{aligned}
A_2^1 &= \left(\sqrt{\frac{m_n\lambda_n^2}{|B_n^i|}}\frac{1}{\nu^2}\right)^3 \left(\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_5-\boldsymbol{s}_6)\,\frac{\tau_{AB,m}(\boldsymbol{s}_6-\boldsymbol{s}_5)}{m_n}\nu^2 d\boldsymbol{s}_5 d\boldsymbol{s}_6\right) \\
&\quad \int_{B_n^i}\int_{B_n^i}\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)w_n(\boldsymbol{h}+\boldsymbol{s}_3-\boldsymbol{s}_4) \\
&\qquad \frac{\tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1,\boldsymbol{s}_3-\boldsymbol{s}_1,\boldsymbol{s}_4-\boldsymbol{s}_1)}{m_n}\nu^4 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4 \\
&= \left(\sqrt{\lambda_n^2}\right)^3 \left(\frac{1}{\sqrt{|B_n^i|^3 m_n}}\right)\left(\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_5-\boldsymbol{s}_6)\,\tau_{AB,m}(\boldsymbol{s}_6-\boldsymbol{s}_5)d\boldsymbol{s}_5 d\boldsymbol{s}_6\right) \\
&\quad \int_{B_n^i}\int_{B_n^i}\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)w_n(\boldsymbol{h}+\boldsymbol{s}_3-\boldsymbol{s}_4) \\
&\qquad \tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1,\boldsymbol{s}_3-\boldsymbol{s}_1,\boldsymbol{s}_4-\boldsymbol{s}_1)d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4 \\
&\leq \left(\sqrt{\lambda_n^2}\right)^3 \left(\frac{1}{\sqrt{|B_n^i|^3 m_n}}\right)\,|B_n^i| \\
&\quad \int_{B_n^i}\int_{B_n^i}\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)w_n(h+\boldsymbol{s}_3-\boldsymbol{s}_4) \\
&\qquad \tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1,\boldsymbol{s}_3-\boldsymbol{s}_1,\boldsymbol{s}_4-\boldsymbol{s}_1)d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4 \\
&\leq \left(\sqrt{\lambda_n^2}\right)^3 \left(\frac{1}{\sqrt{|B_n^i|^3 m_n}}\right)\,|B_n^i|^2 \int_{\mathbb{R}^2}\tau_{AB,m}(\boldsymbol{y})d\boldsymbol{y} \\
&\leq \text{const}\,\sqrt{\frac{\lambda_n^6|B_n^i|}{m_n}}.
\end{aligned}$$

Notice that the technique in Claim 2.6.1 is repeatedly used for each inequality.

Step 3: Show $A_1 < \infty$.

This step involves $E[N^{(2)}(d\boldsymbol{s}_1, d\boldsymbol{s}_2)N^{(2)}(d\boldsymbol{s}_3, d\boldsymbol{s}_4)N^{(2)}(d\boldsymbol{s}_5, d\boldsymbol{s}_6)]$. As done in Step 2, we consider each scenarios. We need to consider the following 5 representative cases:

- $\nu^6 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4 d\boldsymbol{s}_5 d\boldsymbol{s}_6$,
- $\nu^5 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4 d\boldsymbol{s}_5 \varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_6)$ or any other combinations involving ν^5,
- $\nu^4 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4 \varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_5)\varepsilon_{\boldsymbol{s}_2}(d\boldsymbol{s}_6)$ or any other combinations involving ν^4,
- $\nu^3 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 \varepsilon_{\boldsymbol{s}_2}(d\boldsymbol{s}_4)\varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_5)\varepsilon_{\boldsymbol{s}_2}(d\boldsymbol{s}_6)$ or any other combinations involving ν^3, and
- $\nu^2 d\boldsymbol{s}_1 d\boldsymbol{s}_2 \varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_3)\varepsilon_{\boldsymbol{s}_2}(d\boldsymbol{s}_4)\varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_5)\varepsilon_{\boldsymbol{s}_2}(d\boldsymbol{s}_6)$ or any other combinations involving ν^2.

Case 1) $\nu^2 d\boldsymbol{s}_1 d\boldsymbol{s}_2 \varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_3)\varepsilon_{\boldsymbol{s}_2}(d\boldsymbol{s}_4)\varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_5)\varepsilon_{\boldsymbol{s}_2}(d\boldsymbol{s}_6)$ or other combinations involving ν^2. In this case, the corresponding integral becomes

$$\begin{aligned}
& \left(\sqrt{\frac{m_n\lambda_n^2}{|B_n^i|}}\frac{1}{\nu^2}\right)^3 \int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)^3 \, \frac{\tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1)}{m_n}\nu^2 d\boldsymbol{s}_1 d\boldsymbol{s}_2 \\
= & \left(\sqrt{\frac{\lambda_n^2}{|B_n^i|}}\right)^3 \left(\frac{\sqrt{m_n}}{\nu^4}\right) \int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)^3 \, \tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1) d\boldsymbol{s}_1 d\boldsymbol{s}_2 \\
\leq & \text{ const } \left(\sqrt{\frac{\lambda_n^6 m_n}{|B_n^i|^3}}\right) \frac{1}{\lambda_n^4}|B_n^i| \\
\leq & \text{ const } \left(\sqrt{\frac{m_n}{|B_n^i|\lambda_n^2}}\right),
\end{aligned}$$

where the change of variable of $\frac{\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2}{\lambda_n} = \boldsymbol{t}, \boldsymbol{s}_2 = \boldsymbol{u}$ are used for the first inequality as Claim 2.6.1. Therefore, the right-hand side is finite if $\sup_n \frac{m_n}{\lambda_n^2|B_n^i|} < \infty$.

Case 2) $\nu^3 d\boldsymbol{s}_1 d\boldsymbol{s}_2 \varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_3) d\boldsymbol{s}_4 \varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_5)\varepsilon_{\boldsymbol{s}_2}(d\boldsymbol{s}_6)$ or other combinations involving ν^3. In this case, the corresponding integral becomes

$$\begin{aligned}
& \left(\sqrt{\frac{m_n\lambda_n^2}{|B_n^i|}}\frac{1}{\nu^2}\right)^3 \int_{B_n^i}\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}_1+\boldsymbol{s}_1-\boldsymbol{s}_2)^2 w_n(\boldsymbol{h}_2+\boldsymbol{s}_1-\boldsymbol{s}_4) \\
& \qquad \frac{\tau_{A,B,B,m}(\boldsymbol{s}_2-\boldsymbol{s}_1,\boldsymbol{s}_4-\boldsymbol{s}_1)}{m_n}\nu^3 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_4 \\
= & \left(\sqrt{\frac{m_n\lambda_n^6}{|B_n^i|^3\nu^6}}\right) \int_{B_n^i}\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}_1+\boldsymbol{s}_1-\boldsymbol{s}_2)^2 w_n(\boldsymbol{h}_2+\boldsymbol{s}_1-\boldsymbol{s}_4) \\
& \qquad \tau_{A,B,B,m}(\boldsymbol{s}_2-\boldsymbol{s}_1,\boldsymbol{s}_4-\boldsymbol{s}_1) d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_4 \\
\leq & \text{ const } \left(\sqrt{\frac{m_n\lambda_n^6}{|B_n^i|^3\nu^6}}\right) \frac{|B_n^i|}{\lambda_n^2},
\end{aligned}$$

where the last inequality is from Claim 2.6.1. The right-hand side is finite if $\sup_n \frac{m_n\lambda_n^2}{|B_n^i|} < \infty$. Other cases of $card\{\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3,\boldsymbol{s}_4,\boldsymbol{s}_5,\boldsymbol{s}_6\} = 3$ can be considered similarly.

Case 3) $\nu^4 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4 \varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_5)\varepsilon_{\boldsymbol{s}_2}(d\boldsymbol{s}_6)$ or other combinations involving ν^4. Let $\tau^* = \tau_{AB,m}(\boldsymbol{s}_2-\boldsymbol{s}_1,\boldsymbol{s}_3-\boldsymbol{s}_1,\boldsymbol{s}_4-\boldsymbol{s}_1)$. The corresponding integral is

$$\left(\sqrt{\frac{m_n\lambda_n^2}{|B_n^i|}}\frac{1}{\nu^2}\right)^3 \int_{B_n^i}\int_{B_n^i}\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)^2$$
$$w_n(\boldsymbol{h}+\boldsymbol{s}_3-\boldsymbol{s}_4)\,\frac{\tau^*}{m_n}\nu^4 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4$$
$$= \left(\sqrt{\frac{\lambda_n^2}{|B_n^i|}}\right)^3 \left(\frac{\sqrt{m_n}}{\nu^2}\right) \int_{B_n^i}\int_{B_n^i}\int_{B_n^i}\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2)^2$$
$$w_n(\boldsymbol{h}+\boldsymbol{s}_3-\boldsymbol{s}_4)\,\tau^* d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4$$

after the change of variables which is bounded by

$$\left(\sqrt{\frac{\lambda_n^2}{|B_n^i|}}\right)^3 \left(\frac{\sqrt{m_n}}{\nu^4}\right) \frac{|B_n^i|}{\lambda_n^2} \int_{\mathbb{R}^2} w(\boldsymbol{y}_1)^2 d\boldsymbol{y}_1 \int_{\mathbb{R}^2} w(\boldsymbol{y}_2)^2 d\boldsymbol{y}_2 \int_{\mathbb{R}^2} \tau_{AB,m}(\boldsymbol{y}_3) d\boldsymbol{y}_3.$$

Thus, the corresponding integral is finite if $\sup_n \frac{m_n\lambda_n^2}{|B_n^i|} < \infty$ and $\int_{R^2} \tau_{AB,m}(\boldsymbol{y})d\boldsymbol{y} < \infty$. Other cases involving ν^4 are considered similarly.

Case 4) $\nu^5 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4 d\boldsymbol{s}_5 \varepsilon_{\boldsymbol{s}_1}(d\boldsymbol{s}_6)$ or other combinations involving ν^5.
This case has smaller order than the case with $\nu^6 d\boldsymbol{s}_1 d\boldsymbol{s}_2 d\boldsymbol{s}_3 d\boldsymbol{s}_4 d\boldsymbol{s}_5 d\boldsymbol{s}_6$. So we skip this part.

Case 5) If $card\{\boldsymbol{s}_1, \boldsymbol{s}_2, \boldsymbol{s}_3, \boldsymbol{s}_4, \boldsymbol{s}_5, \boldsymbol{s}_6\} = 6$. For the convenience, let $\tau^* = \tau_{AB,m}(\boldsymbol{s}_2 - \boldsymbol{s}_1, \boldsymbol{s}_3 - \boldsymbol{s}_1, \boldsymbol{s}_4 - \boldsymbol{s}_1, \boldsymbol{s}_5 - \boldsymbol{s}_1, \boldsymbol{s}_6 - \boldsymbol{s}_1)$. Then the corresponding integral is equivalent to

$$\left(\sqrt{\frac{m_n\lambda_n^2}{|B_n^i|}}\frac{1}{\nu^2}\right)^3 \int_{B_n^i}\cdots\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2) w_n(\boldsymbol{h}+\boldsymbol{s}_3-\boldsymbol{s}_4)$$
$$w_n(\boldsymbol{h}+\boldsymbol{s}_5-\boldsymbol{s}_6)\,\frac{\tau^*}{m_n}\nu^6 d\boldsymbol{s}_1\cdots d\boldsymbol{s}_6$$
$$= \left(\sqrt{\frac{\lambda_n^2}{|B_n^i|}}\right)^3 \sqrt{m_n} \int_{B_n^i}\cdots\int_{B_n^i} w_n(\boldsymbol{h}+\boldsymbol{s}_1-\boldsymbol{s}_2) w_n(\boldsymbol{h}+\boldsymbol{s}_3-\boldsymbol{s}_4)$$
$$w_n(\boldsymbol{h}+\boldsymbol{s}_5-\boldsymbol{s}_6)\,\tau^* d\boldsymbol{s}_1\cdots d\boldsymbol{s}_6.$$

Consider the change of variable with $\boldsymbol{v}_1 = \boldsymbol{s}_2 - \boldsymbol{s}_1, \boldsymbol{v}_2 = \boldsymbol{s}_3 - \boldsymbol{s}_1, \boldsymbol{v}_3 = \boldsymbol{s}_4 - \boldsymbol{s}_1, \boldsymbol{v}_4 = \boldsymbol{s}_5 - \boldsymbol{s}_1, \boldsymbol{v}_5 = \boldsymbol{s}_6 - \boldsymbol{s}_1$, and $\boldsymbol{s}_1 = \boldsymbol{u}_1$ and let $\tau^{**} = \tau_{AB,m}(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4, \boldsymbol{v}_5)$. Then the

right-hand side of the above equation becomes

$$
\begin{aligned}
&= \left(\sqrt{\frac{\lambda_n^2}{|B_n^i|}}\right)^3 \sqrt{m_n} \int_{B_n^i(B_n^i - B_n^i)^2(B_n^i - B_n^i)^3} \iint \iiint w_n(\boldsymbol{h}-\boldsymbol{v}_1) w_n(\boldsymbol{h}-(\boldsymbol{v}_3-\boldsymbol{v}_2)) \\
&\qquad\qquad w_n(\boldsymbol{h}-(\boldsymbol{v}_5-\boldsymbol{v}_4))\, \tau^{**} d\boldsymbol{v}_1 \cdots d\boldsymbol{v}_5 d\boldsymbol{u}_1 \\
&= \left(\sqrt{\frac{\lambda_n^2}{|B_n^i|}}\right)^3 \sqrt{m_n} \int_{B_n^i(B_n^i - B_n^i)^2\left(\frac{\boldsymbol{h}+B_n^i - B_n^i}{\lambda_n}\right)^3} \iint \iiint w(\boldsymbol{y}_1) w(\boldsymbol{y}_2) w(\boldsymbol{y}_3)\, \tau^{***} d\boldsymbol{y}_1 d\boldsymbol{y}_2 d\boldsymbol{y}_3 d\boldsymbol{u}_2 d\boldsymbol{u}_3 d\boldsymbol{u}_1,
\end{aligned}
$$

where the line is by the change of variables $\frac{\boldsymbol{h}-\boldsymbol{v}_1}{\lambda_n} = \boldsymbol{y}_1$, $\frac{\boldsymbol{h}-(\boldsymbol{v}_3-\boldsymbol{v}_2)}{\lambda_n} = \boldsymbol{y}_2$, $\frac{\boldsymbol{h}-(\boldsymbol{v}_5-\boldsymbol{v}_4)}{\lambda_n} = \boldsymbol{y}_3$, $\boldsymbol{v}_2 = \boldsymbol{u}_2$, and $\boldsymbol{v}_4 = \boldsymbol{u}_3$ and $\tau^{***} = \tau_{AB,m}(\boldsymbol{h} - \boldsymbol{y}_1\lambda_n, \boldsymbol{u}_2, \boldsymbol{h} - \boldsymbol{y}_2\lambda_n + \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{h} - \boldsymbol{y}_3\lambda_n + \boldsymbol{u}_3)$. From $\int_{\mathbb{R}^2} w(\boldsymbol{y}) d\boldsymbol{y} = 1$, $\int_{\mathbb{R}^2} \tau_{AB,m}(\boldsymbol{y}) d\boldsymbol{y} < \infty$ and $|B_n^i - B_n^i| = O(|B_n^i|)$, the right-hand side is bounded

$$
\begin{aligned}
&\left(\sqrt{\frac{\lambda_n^2}{|B_n^i|}}\right)^3 \sqrt{m_n} \int_{B_n^i(B_n^i - B_n^i)^2\left(\frac{\boldsymbol{h}+B_n^i - B_n^i}{\lambda_n}\right)^3} \iint \iiint w(\boldsymbol{y}_1) w(\boldsymbol{y}_2) w(\boldsymbol{y}_3) \\
&\qquad\qquad \tau_{AB,m}(\boldsymbol{u}_2, \boldsymbol{u}_3) d\boldsymbol{y}_1 d\boldsymbol{y}_2 d\boldsymbol{y}_3 d\boldsymbol{u}_2 d\boldsymbol{u}_3 d\boldsymbol{u}_1 \\
\leq\ &\left(\sqrt{\frac{\lambda_n^2}{|B_n^i|}}\right)^3 \sqrt{m_n} \left(\int_{\mathbb{R}^2} w(\boldsymbol{y}) d\boldsymbol{y}\right)^3 \int_{B_n^i\,(B_n^i - B_n^i)^2} \iint \tau_{AB,m}(\boldsymbol{u}_2, \boldsymbol{u}_3) d\boldsymbol{u}_2 d\boldsymbol{u}_3 d\boldsymbol{u}_1 \\
\leq\ &\left(\sqrt{\frac{\lambda_n^2}{|B_n^i|}}\right)^3 \sqrt{m_n} \left(\int_{\mathbb{R}^2} w(\boldsymbol{y}) d\boldsymbol{y}\right)^3 \left(\int_{\mathbb{R}^2} \tau_{AB,m}(\boldsymbol{y}) d\boldsymbol{y}\right) |B_n^i - B_n^i||B_n^i| \\
\leq\ &\text{const}\ \lambda_n^6 m_n |B_n^i|.
\end{aligned}
$$

Thus, $\lambda_n^6 m_n |B_n^i| \to 0$ is required.

In conclusion, $E|\sqrt{k_n}\tilde{a}'_{ni}|^3 < \infty$ if

$$
\begin{aligned}
&\sup_n \frac{\lambda_n^2 n^{2a}}{m_n} < \infty,\ \sup_n \frac{m_n}{\lambda_n^2 n^{2a}} < \infty,\ \sup_n \frac{m_n \lambda_n^2}{n^{2a}} < \infty, \\
&\lambda_n^6 m_n n^{2a} \to 0,\ \text{and} \int_{R^2} \tau_{AB,m}(\boldsymbol{y}) d\boldsymbol{y} < \infty.
\end{aligned}
$$

Notice that the third condition is implied by the second condition since $\lambda_n \to 0$ as $n \to 0$ and the fourth condition is inferred by the first two conditions and $\lambda_n^2 m_n \to 0$ in condition

(**M3**) as $\lambda_n^6 m_n n^{2a} \sim \lambda_n^4 m_n^2 \to 0$. Hence, $E|\sqrt{k_n}\tilde{a}'_{ni}|^3 < \infty$ if

$$\sup_n \frac{\lambda_n^2 n^{2a}}{m_n} < \infty, \ \sup_n \frac{m_n}{\lambda_n^2 n^{2a}} < \infty \text{ and } \int_{\mathbb{R}^2} \tau_{AB,m}(\boldsymbol{y}) d\boldsymbol{y} < \infty.$$

Acknowledgments

We thank Christina Steinkohl for discussions on the proofs and suggestions. We also would like to thank Chin Man (Bill) Mok for providing the Florida rainfall data and acknowledge that the data is provided by the Southwest Florida Water Management District (SWFWMD).

Chapter 3

Resampling methods for the empirical spatial extremogram

3.1 Introduction

In Chapter 2, we introduce the concept of *the spatial extremogram* and propose its empirical estimator under two different sampling schemes: the lattice and non-lattice case. The asymptotic normality for the empirical spatial extremogram (ESE) is well established for both cases, but unfortunately the limiting variance is intractable since it is a function of an infinite sum of unknown quantities.

Estimating the variance of the empirical spatial extremogram is critical in order to make inferences about spatial extremal dependence. To construct credible confidence intervals, one can resort to bootstrap procedures or other variance estimation techniques applied to the spatial setting. We establish consistency properties of a bootstrapped and a subsampling variance estimator to facilitate the use of the ESE in practice. This is the main motivation of this chapter.

Figure 3.1 illustrates the variability of the PA-extremogram and the ESE. Using 100 simulations of max-stable process described in Smith (1990), we plot the extremogram (dark black solid line), PA-extremogram (darker gray lines), and ESE (lighter gray lines) using $A = B = (1, \infty)$ and $a_m = .95$ upper quantile. The uncertainty in the PA-extremogram

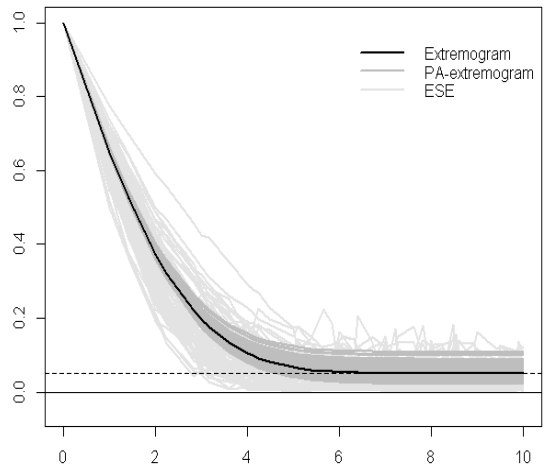

Figure 3.1: The uncertainty of the pre-asymptotic extremogram and empirical spatial extremogram.

and ESE suggests that it is more reasonable to consider an interval estimation rather than a point estimate.

By now, there are many methods for applying the bootstrap to stationary time series. The non-overlapping block bootstrap by Carlstein (1986), the moving block bootstrap by Künsch (1989), the circular block bootstrap by Politis and Romano (1991), and the stationary bootstrap by Politis and Romano (1994) are just a few options. These bootstrap methods can be categorized by overlapping or non-overlapping of data, using fixed or random bootstrap block size, and wrapping or non-wrapping the sample data. To begin with, we review bootstrap methods for stationary time series data.

- The moving block bootstrap (MBB) by Künsch (1989) uses the fixed block sizes, where blocks can overlap.

- The non-overlapping block bootstrap (NBB) inspired by Carlstein (1986) uses the fixed block sizes, where blocks do not overlap.

- The circular block bootstrap (CBB) by Politis and Romano (1991) uses the fixed block

sizes in the wrapped data, where blocks can overlap.

- The stationary bootstrap (SB) by Politis and Romano (1994) uses random block sizes by geometric distribution in the wrapped data, where blocks can overlap.

Regarding the properties of these methods, the MBB and CBB estimators have the smallest variances of bias and have the smallest mean square error, as mentioned in Lahiri (1999). On the other hand, Politis and Romano (1994) finds that pseudo-time series from the SB is stationary conditioned on the original data, but random block size increases the variances of bias. For the theoretical aspects of these methods, we refer to Lahiri (1999). For the bootstrapped ESE for a stationary time series, Davis et al. (2012) shows the (conditional) central limit theorem under the stationary bootstrap.

Other variance estimation schemes can be used to construct confidence intervals for the ESE. Politis and Romano (1993) proposes the blocks of blocks resampling scheme, such as jackknife, for random fields on the lattice. For irregularly spaced observations, Politis and Sherman (2001) shows L_2 consistency of the subsampling estimators of the moments of general statistics. These results are directly applicable to the ESE under appropriate assumptions.

For a variance estimation for the ESE, two sampling schemes are considered as in the previous chapter. When samples are from the lattice, three methods are considered:

- the circular block bootstrap (or "CBB") by Politis and Romano (1991),
- 'blocks of blocks' jackknife variance estimator by Politis and Romano (1993), and
- subsampling variance estimator by Politis and Sherman (2001).

We choose the CBB since it is easy to implement in space. In particular, the spatially adapted CBB creates pseudo-space data that allows unbiased estimator of the sample mean. Under the same setting, we estimate variance of the ESE by the blocks of blocks jackknife and subsamplig variance estimator.

For the non-lattice case, the data $\{X_{s_1}, ..., X_{s_N}\}$ is assumed to be from a stationary random field, where the locations $\{s_1, ..., s_N\}$ are points of a homogeneous Poisson point

process on R^d. In this case, we devise the subsampling to estimate the variance in order to provide confidence intervals for the ESE following Politis and Sherman (2001). This will allow us to construct confidence intervals for the ESE. The CBB may be possible, but it is more complex in the non-lattice setting, thus it is difficult to implement in practice and establish asymptotic results.

The organization of the chapter is as follows. In Section 3.2.1, we present the asymptotic results for the bootstrapped ESE and the variance estimated by jackknife and subsampling when the underlying data is observed on a d-dimensional lattice. We also consider in Section 3.2.2 subsampling for variance estimation when the data is observed at Poisson points in $\mathbb{R}^d$. The latter resampling method can be used to provide a portmanteau style test for extremal dependence at multiple lags. In Section 3.4, these results are applied and compared through simulation examples. Section 3.5 explores portmanteau test ideas to check the existence of extremal dependences at multiple lags. In Section 3.6, these methods are applied to two data sets: the first is rainfall in Florida and the second is ground-level ozone in the eastern United States. The proofs of all results are given in Section 3.8.

3.2 Bootstrapped ESE and the variance estimation of the ESE

3.2.1 Random fields on the lattice

In this section, we assume that observations are from the lattice. Consider an increasing sequence $m := m_n$ and a_m such that $m_n/n \to 0$ as $n \to \infty$ and $P(|X| > a_m) \sim \frac{1}{m_n}$, which will be assumed throughout the rest of the thesis. Following Chapter 2, define the ESE from a set $E \subset \mathbb{Z}^d$ by

$$\hat{\rho}_{AB,m}(\boldsymbol{h}, E) = \frac{\sum_{\boldsymbol{s},\boldsymbol{t}\in E, \boldsymbol{s}-\boldsymbol{t}=\boldsymbol{h}} I_{\{a_m^{-1}X_{\boldsymbol{s}}\in A, a_m^{-1}X_{\boldsymbol{t}}\in B\}}/n(\boldsymbol{h})}{\sum_{\boldsymbol{s}\in E} I_{\{a_m^{-1}X_{\boldsymbol{s}}\in A\}}/\#E}, \tag{3.1}$$

given the numerator is non-zero. In (3.1), $\boldsymbol{h} \in \mathbb{Z}^d$ are observed lags in E, $n(\boldsymbol{h})$ is the number of pairs in E with lag $\boldsymbol{h}$, and $\#E$ is the cardinality of E. When the numerator in (3.1) is zero, $\hat{\rho}_{AB,m}(\boldsymbol{h}, E)$ is defined to be 0, so is the PA-extremogram (2.7). Since the

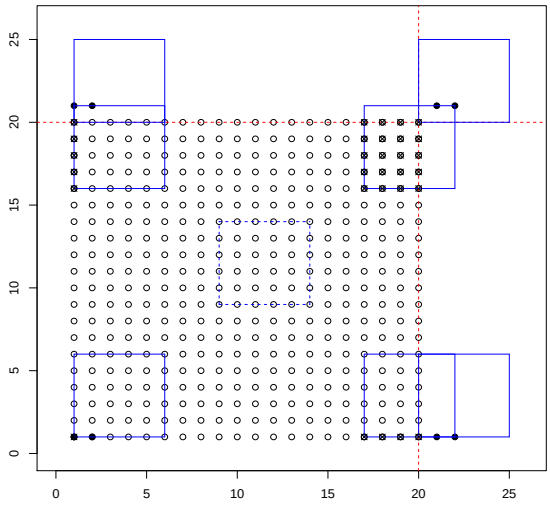

Figure 3.2: Overview of the spatial CBB in $\mathbb{Z}^2$.

generalization would be straightforward, we focus on the case of d=2.

3.2.1.1 The bootstrapped ESE using the circular block bootstrap (CBB)

For a stationary time series, Davis et al. (2012) establishes the central limit theorem for the bootstrapped extremogram under proper dependence and mixing conditions. In this section, we establish the central limit theorem for the bootstrapped ESE under the space-adapted CBB, which can be viewed as the extension of Davis et al. (2012).

For the ESE variance estimation, we choose the CBB since it creates pseudo-space data which allows unbiased estimator of sample mean. To see this, note that the CBB wraps the data before blocking them, which enables each point to have the equal probability of being selected. Also, for a stationary time series, it produces the smallest variance of bias and the smallest mean square error, which may remain true for spatial data, which is inferred from Lahiri (1999). Lastly, it is easy to implement. For example, the implementation of the SB in space is challenging since reconstructing the original sample dimension with random block sizes is difficult. On the other hand, the CBB uses a fixed block size, thus it can be easily applied to a higher dimension.

Now we illustrates the CBB. Suppose the data are observed on $\Lambda_n = \{1, ..., n\}^2 \in \mathbb{Z}^2$ and b_n be a bootstrap block length. The space-adapted CBB consists of the following steps. First, the original data is *wrapped* to get the extended region $\{1, ..., n + b_n\}^2$. For example, the data on $\{1, ..., n\} \times \{1, ..., b_n\}$ is copied and pasted to $\{1, ..., n\} \times \{n+1, ..., n+b_n\}$. The same procedure is repeated for $\{1, ..., b_n\} \times \{1, ..., n\}$ and $\{1, ..., b_n\} \times \{1, ..., b_n\}$ to achieve the extended region. Then, a sub block of dimensions $b_n \times b_n$ is scooped randomly from the extended region and filled into the re-sample space. This is repeated until the re-sample space becomes $\{1, ..., n\}^2$. Notice that the direction of reconstructing the pseudo sample does not matter since each bootstrap block is independent conditioned on the data. Figure 3.2 has as an illustration with $n = 20$ and $b_n = 5$. To be specific, the space-adapted CBB is implemented as

- $\{1, ..., 5\} \times \{1, ..., 20\}$ is wrapped to $\{21, ..., 20 + 5\} \times \{1, ..., 20\}$,
- $\{1, ..., 20\} \times \{1, ..., 5\}$ is wrapped to $\{1, ..., 20\} \times \{21, ..., 20 + 5\}$, and
- $\{1, ..., 5\} \times \{1, ..., 5\}$ is wrapped to $\{21, ..., 20 + 5\} \times \{21, ..., 20 + 5\}$.

For example, $\{(1, 1), (2, 1)\}$ (solid black dots on the left bottom corner) are wrapped to three locations of $\{(21, 1), (22, 1)\}, \{(1, 21), (2, 21)\}$ and $\{(21, 21), (22, 21)\}$. If $\{(1, 1), (2, 1)\}$ is randomly chosen, the bootstrap block (blue box in the lower left) is constructed from the chosen point as an anchoring point and copied to the pseudo data. This is repeated until the pseudo data is filled.

The central limit theorem is established for the bootstrapped ESE conditioned on the sample under the spatially adapted CBB. In the theorem, we use P^* and $\hat{\rho}^*_{AB,m}(\boldsymbol{h}, \Lambda_n)$ to denote the probability measure generated by the CBB and the ESE under P^*, respectively.

Theorem 5. *Consider a strictly stationary regularly varying random field* $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{Z}^2\}$ *with index* $\alpha > 0$ *is observed on* $\Lambda_n = \{1, ..., n\}^2$. *Define* $B_\gamma = \{\boldsymbol{s} \in \mathbb{Z}^2 : d(\boldsymbol{0}, \boldsymbol{s}) \leq \gamma\}$, *where* γ *is an observed distance in a bootstrapped block* E *with the block length* b_n. *Suppose there exists an increasing sequence* m_n *with* $m_n = o(n)$ *and* $b_n = o(n)$. *For any finite set of*

non-zero lags $H = \{\boldsymbol{h}_1, ..., \boldsymbol{h}_t\}$ *in* $\mathbb{Z}^2$*, where* $H \subset B_\gamma$*, assume*

$$\sqrt{\frac{b_n^2}{m_n}}\Big[\hat{\rho}_{AB,m_n}(\boldsymbol{h}, E) - \rho_{AB,m_n}(\boldsymbol{h})\Big]_{\boldsymbol{h}\in H} \xrightarrow{d} N(\mathbf{0}, \Sigma) \tag{3.2}$$

holds, where the asymptotic variance Σ *is specified in Theorem 2.1, Chapter 2. Further, assume that*

$$m_n b_n^4 = o(n^2), \tag{3.3}$$

$$m_n^7 = o(b_n^2), \quad \textit{and} \quad \lim_{n\to\infty} m_n^2 b_n^2 \sum_{\boldsymbol{l}\in\mathbb{Z}^2\setminus B_{m_n}} \alpha_{i,j}(d(\mathbf{0}, \boldsymbol{l})) = 0 \tag{3.4}$$

for $2c \le i + j \le 4c$*, where* $c = \#B_\gamma$*. Then*

$$P^*(|\hat{\rho}^*_{AB,m}(\boldsymbol{h}, \Lambda_n) - \rho_{AB,m}(\boldsymbol{h}, \Lambda_n)| > \delta) \xrightarrow{P} 0, \qquad \delta > 0, \tag{3.5}$$

and the central limit theorem holds

$$P^*((n^2/m_n)^{1/2}(\hat{\rho}^*_{AB,m}(\boldsymbol{h}, \Lambda_n) - \hat{\rho}_{AB,m}(\boldsymbol{h}, \Lambda_n))_{\boldsymbol{h}\in H} \in C) \xrightarrow{P} \Phi_{\mathbf{0},\sum}(C), \tag{3.6}$$

where C *is any continuity set of* $\Phi_{\mathbf{0},\sum}$*, the normal distribution with mean zero and covariance* $\sum$*.*

Proof. See Appendix A. □

Remark 9. *Theorem 5 implies that the bootstrapped ESE,* $\hat{\rho}^*_{AB,m}(\boldsymbol{h})$*, is asymptotically correct to estimate* $\hat{\rho}_{AB,m}(\boldsymbol{h})$*. Moreover, Theorem 5 can be extended to dimension* $d > 2$*, as mentioned in Corollary 3.8.2.*

Remark 10. *Condition (3.2) indicates that Theorem 2.1 in Chapter 2 is applicable to each bootstrapped block with block length* b_n*. Conditions (3.3)-(3.4) impose restrictions on the decaying rate of the mixing functions together with the level of the threshold specified by* m_n *and the bootstrap block length* b_n*.*

3.2.1.2 Blocks of blocks jackknife variance estimator

In this section, we use the blocks of blocks jackknife variance estimator proposed by Politis and Romano (1993) to construct a confidence interval for the ESE. The method in Politis

and Romano (1993) extends the idea of Künsch (1989) to have asymptotically valid confidence intervals for parametric in the whole distribution under suitable mixing and moment conditions. We borrow notation from Politis and Romano (1993), but assume rectangular blocking for the convenience. The extension to non-rectangular case is mentioned in Appendix B.

Consider observations on $\Lambda_n = \{1, ..., n\}^2 \in \mathbb{Z}^2$. Let $M \in \mathbb{N}$ be a side length of block and $L \in \mathbb{N}$ be a window lag between blocks, thus overlapping between adjacent blocks is $(M - L)$. For $Q = [(n - M)/L + 1]$, there are Q^2 blocks with the side length of M, which we index using column and row of $\{1, ..., Q\}^2$

$$\Lambda_Q = \{\boldsymbol{i} = (i_1, i_2) : 1 \leq i_k \leq Q \text{ for } k = 1 \text{ and } 2\}.$$

Then the set of indices inside of the block indexed by $\boldsymbol{i} \in \Lambda_Q$ is defined as

$$E_{\boldsymbol{i}} = \{(s_1, s_2) : (i_k - 1)L + 1 \leq s_k \leq (i_k - 1)L + M \text{ for } k = 1 \text{ and } 2\}.$$

Now *blocks of blocks* are defined by stacking neighboring blocks indexed by $\boldsymbol{i} \in \Lambda_Q$. Let b be a number of blocks to be stacked both vertically and horizontally and w be a window lag between blocks of blocks. For $q = [(Q - b)/w + 1]$, we index blocks of blocks using the column and row of $\{1, ..., q\}^2$

$$\Lambda_q^b = \{\boldsymbol{j} = (j_1, j_2) : 1 \leq j_k \leq q \text{ for } k = 1 \text{ and } 2\}.$$

Then the set of the block indices belongs to the blocks of blocks indexed by $\boldsymbol{j}$ is defined by

$$E_{\boldsymbol{j}}^b = \{\boldsymbol{i} = (i_1, i_2) : (j_k - 1)w + 1 \leq i_k \leq (j_k - 1)w + b \text{ for } k = 1 \text{ and } 2\}.$$

Figure 3.3 illustrates blocks of blocks jackknife with n = 40, M =10, L =5, and b =3. Each block is 10×10 square and indexed by (i_1, i_2) for $1 \leq i_1, i_2 \leq 7$. With the choice of b=3, a block of blocks consists of 9 blocks. For example, the block of blocks indexed by $(1, 3)$ consists of indices from E_{i_1,i_2} for $1 \leq i_1 \leq 3$ and $5 \leq i_2 \leq 7$, total 9 blocks, as presented in the upper left corner of the figure.

Now consider the ESE computed from observations from $E_{\boldsymbol{i}}$

$$T_{\boldsymbol{i}}(\boldsymbol{h}) = \sqrt{\frac{M^2}{m_M}} \left(\hat{\rho}_{AB,m_M}(\boldsymbol{h}, E_{\boldsymbol{i}}) - \rho_{AB,m}(\boldsymbol{h})\right), \tag{3.7}$$

where m_M satisfies that $m_M \to \infty$ and $m_M/M \to 0$ as $n \to \infty$. Note that the central limit theorem for (3.7) is established in Chapter 2, Theorem 2.1. Then the jackknife variance estimator for the ESE can be defined as

$$\hat{V}(\sqrt{Q^2}\bar{T}(\boldsymbol{h})) = \frac{b^2}{q^2} \sum_{\boldsymbol{j} \in \Lambda_q^b} (J_{\boldsymbol{j}} - \bar{T}(\boldsymbol{h}))^2 \tag{3.8}$$

where $J_{\boldsymbol{j}} = \frac{1}{b^2} \sum_{\boldsymbol{i} \in E_{\boldsymbol{j}}^b} T_{\boldsymbol{i}}(\boldsymbol{h})$ and $\bar{T} = \frac{1}{Q^2} \sum_{\boldsymbol{i} \in \Lambda_Q} T_{\boldsymbol{i}}(\boldsymbol{h})$. The L_2 consistency of (3.8) follows from Theorem 1 in Politis and Romano (1993), which will be discussed in the following theorem. In the theorem, we use α^T to denote α-mixing associated with $T_{\boldsymbol{i}}(\boldsymbol{h})$.

Theorem 6. *Suppose that a strictly stationary regularly varying random field* $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{Z}^2\}$ *with* $\alpha > 0$ *is observed on* $\Lambda_n = \{1, ..., n\}^2$. *Use* $B_\gamma = \{\boldsymbol{s} \in \mathbb{Z}^2 : d(\boldsymbol{0}, \boldsymbol{s}) \leq \gamma\}$ *to denote the ball with the radius* γ. *For a non-zero lag* $\boldsymbol{h} \in B_\gamma$ *and increasing sequences* M *and* m_M *satisfying* $m_M/M \to 0$ *and* $M/n \to 0$, *assume*

$$T_{\boldsymbol{i}}(\boldsymbol{h}) \xrightarrow{d} N(0, \sigma^2). \tag{3.9}$$

For some $p > 2$ *and* $\delta, k > 0$ *and* $k_1, k_2 > 0$ *with* $k_1 + k_2 \leq 12$, *and* $c = \#B_\gamma$, *further assume that*

$$\sum_{\boldsymbol{l} \in \mathbb{Z}^2} \alpha_{k^2,k^2}(\|\boldsymbol{l}\|)^{\frac{p-2}{p}} < \infty, \tag{3.10}$$

$$m_M^3/M \to 0, \sum_{\boldsymbol{l} \in \mathbb{Z}^2} \alpha_{c,c}(\|\boldsymbol{l}\|) < \infty, m_M^6 M^{10} \sum_{\boldsymbol{l} \in \mathbb{Z}^2 \setminus B_{m_M}} \alpha_{k_1 c, k_2 c}(\|\boldsymbol{l}\|) \to 0, \tag{3.11}$$

$$\left(\frac{M^2}{m_M^2}\right) \sum_{\boldsymbol{l} \in \mathbb{Z}^2} \alpha_{(Mb)^2,(Mb)^2}(\|\boldsymbol{l}\|)^{\frac{p-2}{p}} < \infty \text{ and } \left(\frac{M^2}{m_M^2}\right) \sum_{\boldsymbol{l} \in \mathbb{Z}^2 \setminus B_L} \alpha_{(Mb)^2,(Mb)^2}(\|\boldsymbol{l}\|)^{\frac{p-2}{p}} \to 0, \tag{3.12}$$

$$\sum_{\boldsymbol{l} \in \mathbb{Z}^2} \|\boldsymbol{l}\| \ \alpha_{k_1,k_2}^T(\|\boldsymbol{l}\|) < \infty, \ \alpha_{1,\infty}^T(\|\boldsymbol{l}\|) = o(\|\boldsymbol{l}\|^{-2}), \text{ and } \sum_{\boldsymbol{l} \in \mathbb{Z}^2} \|\boldsymbol{l}\| \ \alpha_{1,1}^T(\|\boldsymbol{l}\|)^{\delta/(2+\delta)} < \infty, \tag{3.13}$$

$$n^2 m_M/M^4 \to 0, \tag{3.14}$$

$$L/M \to a_1 \in (0, 1], \ w/b \to a_2 > 0, \ b \to \infty \text{ and } b = o(Q). \tag{3.15}$$

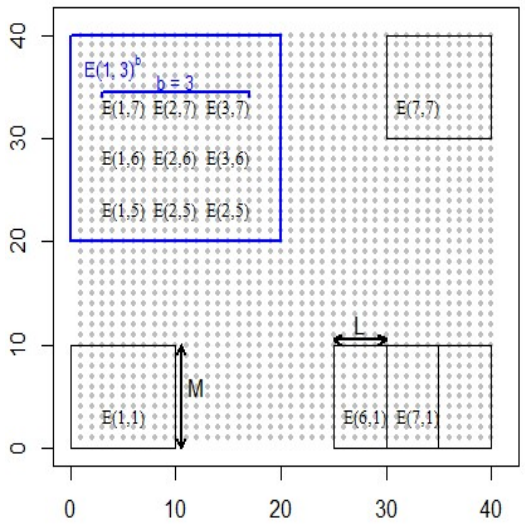

Figure 3.3: The illustration of blocks of blocks jackknife with n = 40, M =10, L =5, b =3.

Then

$$E\hat{V}(\sqrt{Q^2}\bar{T}(\boldsymbol{h})) = var\left(\frac{1}{\sqrt{b^2}}\sum_{\boldsymbol{i}\in E_j^b} T_{\boldsymbol{i}}(\boldsymbol{h})\right) + O(b^2/Q^2) \text{ and } var(\hat{V}(\sqrt{Q^2}\bar{T}(\boldsymbol{h}))) = O(b^2/Q^2). \tag{3.16}$$

Also it is followed that

$$E(\hat{V}(\sqrt{Q^2}\bar{T}(\boldsymbol{h})) - \sigma_\infty^2)^2 \to 0,$$

where $\sigma_\infty^2 = \lim_{n\to\infty} var(\sqrt{Q^2}\bar{T}(\boldsymbol{h})) > 0$.

Proof. All necessary conditions of Theorem 1 in Politis and Romano (1993) are satisfied by (3.9)-(3.15). Thus, the L_2 consistency is proved. See Appendix B for the details. □

Remark 11. *Note that blocks of blocks jackknife estimator (3.8) has two scaling factors: the block size* $1/\sqrt{b}$ *and* $1/\sqrt{q}$ *used in triangular arrays. In the theorem, the conditions (3.10) and (3.11) are required to apply Theorem 1 in Politis and Romano (1993).*

Remark 12. *Theorem 1 in Politis and Romano (1993) assumes*

$$\sqrt{Q^2}(\bar{T}(\boldsymbol{h}) - E\bar{T}(\boldsymbol{h})) \xrightarrow{d} N(0, \sigma_\infty^2),$$

which is verified under (3.12) and (3.13). See Lemma 3.8.3 in Appendix B. Using Theorem 6, the $100(1-\alpha)\%$ *confidence intervals for the PA-extremogram is derived as*

$$\rho_{AB,m}(\boldsymbol{h}) \in \left(\frac{1}{Q^2} \sum_{\boldsymbol{i} \in \Lambda_Q} \hat{\rho}_{AB,m_M}(\boldsymbol{h}, E_{\boldsymbol{i}}) \pm \sqrt{Q^{-2}}\, \hat{\sigma}_\infty\, \Phi^{-1}_{1-\alpha/2} \right),$$

where $\Phi^{-1}_{1-\alpha/2}$ *indicates the inverse cumulative distribution function of the standard normal and*

$$\hat{\sigma}^2_\infty = \left(\frac{b^2}{q^2} \sum_{\boldsymbol{j} \in \Lambda_q^b} \left(\frac{1}{b^2} \sum_{\boldsymbol{i} \in E_{\boldsymbol{j}}^b} \hat{\rho}_{AB,m_M}(\boldsymbol{h}, E_{\boldsymbol{i}}) - \frac{1}{Q^2} \sum_{\boldsymbol{i} \in \Lambda_Q} \hat{\rho}_{AB,m_M}(\boldsymbol{h}, E_{\boldsymbol{i}}) \right)^2 \right).$$

3.2.1.3 Random fields on $\mathbb{Z}^d$: subsampling the variance of the ESE

Subsampling variance estimator for the lattice case is presented in the next section since its derivation is analogous to Theorem 7.

3.2.2 Random fields on $\mathbb{R}^d$: subsampling variance estimator

Now consider observations are from $\mathbb{R}^d$. We borrow the settings of Karr (1986) and restrict our attention to $\mathbb{R}^2$. Let $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^2\}$ be a stationary regularly varying random field with index $\alpha > 0$. Suppose N is a homogeneous 2-dimensional Poisson process with intensity parameter ν and is independent of X. Define

$$N^{(2)}(d\boldsymbol{s}^1, d\boldsymbol{s}^2) = N(d\boldsymbol{s}^1) N(d\boldsymbol{s}^2) I(\boldsymbol{s}^1 \neq \boldsymbol{s}^2).$$

Consider a sequence of compact and convex sets $\Lambda_n \subset \mathbb{R}^2$ with Lebesgue measure $|\Lambda_n| = O(n^2)$ and $|\partial\Lambda_n| = O(n)$, where $\partial\cdot$ denotes the boundary. Define $\Lambda_n - \boldsymbol{y} = \{\boldsymbol{x} - \boldsymbol{y} : \boldsymbol{x} \in \Lambda_n\}$ and suppose that for each $\boldsymbol{y} \in \mathbb{R}^2$

$$\lim_{n \to \infty} \frac{|\Lambda_n \cap (\Lambda_n - \boldsymbol{y})|}{|\Lambda_n|} = 1.$$

Under this setting, the central limit theorem for $\hat{\rho}_{AB,m}(\boldsymbol{h}, \Lambda_n) = \hat{\tau}_{AB,m}(\boldsymbol{h}, \Lambda_n)/\hat{p}_m(A)$, where

$$\hat{p}_m(A) = \frac{m_n}{\nu|\Lambda_n|} \int_{\Lambda_n} I\left(\frac{X_{\boldsymbol{s}^1}}{a_m} \in A\right) N(d\boldsymbol{s}^1), \tag{3.17}$$

$$\begin{aligned}&\hat{\tau}_{AB,m}(\boldsymbol{h}, \Lambda_n)\\ &= \frac{m_n}{\nu^2} \frac{1}{|\Lambda_n|} \int_{\Lambda_n} \int_{\Lambda_n} w_n(h + \boldsymbol{s}^1 - \boldsymbol{s}^2)\, I\left(\frac{X_{\boldsymbol{s}^1}}{a_m} \in A\right) I\left(\frac{X_{\boldsymbol{s}^2}}{a_m} \in B\right) N^{(2)}(d\boldsymbol{s}^1, d\boldsymbol{s}^2),\end{aligned} \tag{3.18}$$

is established in Theorem 2.3 in Chapter 2. Recall that m_n is an increasing sequence such that $m_n = o(n)$ and $w_n(\cdot) = \frac{1}{\lambda_n^2} w(\frac{\cdot}{\lambda_n})$ is a sequence of weight functions, where $w(\cdot)$ on $\mathbb{R}^2$ is a positive, bounded, isotropic probability density function and λ_n is the bandwidth satisfying $\lambda_n \to 0$ and $\lambda_n^2|\Lambda_n| \to \infty$.

For variance estimator, we follow Politis and Sherman (2001). Consider a subsampling region $B_n = (0, cn]^2$, where $c = c_n := \sqrt{|B_n|/|\Lambda_n|} \in (0, 1)$ is a scaling factor. Use $B_n + \boldsymbol{y} = \{\boldsymbol{s} + \boldsymbol{y} : \boldsymbol{s} \in B_n\}$ and $\Lambda_n^{1-c} = \{\boldsymbol{s} : B_n + \boldsymbol{s} \in \Lambda_n\}$ to denote a $\boldsymbol{y}-$shifted subsampling region and a set of anchoring points in Λ_n. The subsampling variance and covariance estimator for the ESE for non-zero lag $\boldsymbol{h}, \boldsymbol{h}^1$ and $\boldsymbol{h}^2$, are defined as

$$\hat{\Sigma}_n(\boldsymbol{h}) = \left(\frac{|B_n|\lambda_n^2}{m_n|\Lambda_n^{1-c}|}\right) \int_{\Lambda_n^{1-c}} \left(\hat{\rho}_{AB,m}(\boldsymbol{h}, B_n + \boldsymbol{y}) - \bar{\rho}_{AB,m}(\boldsymbol{h}, B_n)\right)^2 d\boldsymbol{y}, \tag{3.19}$$

$$\begin{aligned}&\hat{\Sigma}_n(\boldsymbol{h}^1, \boldsymbol{h}^2)\\ = \;& \left(\frac{|B_n|\lambda_n^2}{m_n|\Lambda_n^{1-c}|}\right) \int_{\Lambda_n^{1-c}} (\hat{\rho}_{AB,m}(\boldsymbol{h}^1, B_n + \boldsymbol{y}) - \bar{\rho}_{AB,m}(\boldsymbol{h}^1, B_n))\\ &\qquad (\hat{\rho}_{AB,m}(\boldsymbol{h}^2, B_n + \boldsymbol{y}) - \bar{\rho}_{AB,m}(\boldsymbol{h}^2, B_n)) d\boldsymbol{y},\end{aligned} \tag{3.20}$$

where $\bar{\rho}_{AB}(\boldsymbol{h}, B_n) = \frac{1}{|\Lambda_n^{1-c}|} \int_{\Lambda_n^{1-c}} \hat{\rho}_{AB,m}(\boldsymbol{h}, B_n + \boldsymbol{y}) d\boldsymbol{y}$. It will be shown that (3.19) and (3.20) are L_2 consistent by Theorem 2 in Politis and Sherman (2001).

Theorem 7. *Consider a strictly stationary regularly varying random field* $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^d\}$ *with* $\alpha > 0$. *Suppose the locations are generated by a Poisson process* N *and observed in the compact set* $\Lambda_n = (0, n]^2 \in \mathbb{R}^d$. *Assume that there exist increasing sequences* m_n *and* n,

a decreasing sequence $c = c_n$*, and* $B_n = (0, cn]^2$ *such that*

$$g(\boldsymbol{h}, \Lambda_n) = \sqrt{\frac{|\Lambda_n|\lambda_n^2}{m_n}}[\hat{\rho}_{AB,m}(\boldsymbol{h}, \Lambda_n) - \rho_{AB,m}(\boldsymbol{h})] \to N(0, \sigma^2) \quad \textit{and}$$

$$g(\boldsymbol{h}, B_n + \boldsymbol{y}) = \sqrt{\frac{|B_n|\lambda_n^2}{m_n}}[\hat{\rho}_{AB,m}(\boldsymbol{h}, B_n + \boldsymbol{y}) - \rho_{AB,m}(\boldsymbol{h})] \to N(0, \sigma^2) \tag{3.21}$$

as $n \to \infty$*. Further assume that*

$$cn \to \infty \textit{ and } m_n/cn \to 0, \tag{3.22}$$

$$\alpha_{k^2,k^2}(k) \to 0 \textit{ as } k \to \infty, \tag{3.23}$$

$$\lambda_n^8 n^6 m_n^4 = O(1), \textit{ and} \tag{3.24}$$

$$\int_{\mathbb{R}^2} \alpha_{1,k}(\boldsymbol{y}) d\boldsymbol{y} < \infty \textit{ for } k = 1, ..., 5. \tag{3.25}$$

Then

$$\hat{\Sigma}_n(\boldsymbol{h}) \xrightarrow{L_2} \lim_{n\to\infty} var(g(\boldsymbol{h}, \Lambda_n)) \qquad \textit{and} \qquad \hat{\Sigma}_n(\boldsymbol{h}^1, \boldsymbol{h}^2) \xrightarrow{L_2} \lim_{n\to\infty} cov\left(g(\boldsymbol{h}^1, \Lambda_n), g(\boldsymbol{h}^2, \Lambda_n)\right). \tag{3.26}$$

Proof. The proof follows Theorem 2 of Politis and Sherman (2001). See Appendix C. □

Remark 13. *Theorem 2 in Politis and Sherman (2001) is applicable to the ESE given additional conditions on the bandwidth, (3.24), and mixing, (3.25).*

Remark 14. *Using Theorem 7, the confidence interval for* $\rho_{AB,m}(\boldsymbol{h})$ *is constructed. From Theorem 2.3 in Chapter 2, we have*

$$\sqrt{\frac{|\Lambda_n|\lambda_n^2}{m_n}}(\hat{\rho}_{AB,m}(\boldsymbol{h}, \Lambda_n) - \rho_{AB,m}(\boldsymbol{h})) \xrightarrow{d} N(0, \sigma^2).$$

Since $\hat{\Sigma}_n(\boldsymbol{h}) \xrightarrow{L_2} \sigma^2$*, it follows that*

$$\rho_{AB,m}(\boldsymbol{h}) \in \left[\hat{\rho}_{AB,m}(\boldsymbol{h}, \Lambda_n) \pm \left(\frac{m_n}{|\Lambda_n|\lambda_n^2}\, \hat{\Sigma}_n(\boldsymbol{h})\right)^{1/2} \Phi^{-1}_{1-\alpha/2}\right] \tag{3.27}$$

is asymptotic $100(1-\alpha)\%$ *confidence interval for* $\rho_{AB,m}(\boldsymbol{h})$*.*

Now, the lattice case analog of Theorem 7 is discussed.

Theorem 8. *Consider a strictly stationary regularly varying random field* $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{Z}^2\}$ *with index* $\alpha > 0$ *observed on* $\Lambda_n = \{1, ..., n\}^2$*. Assume that there exist increasing sequences* m_n *and* n*, a decreasing sequence* $c = c_n$*, and* $B_n = (0, cn]^2$*. For any finite set of non-zero lags* $H = (\boldsymbol{h}_1, ..., \boldsymbol{h}_t)$ *in* $\mathbb{Z}^2$*, where* $H \subset B_\gamma := \{\boldsymbol{s} \in \mathbb{Z}^2 : d(\boldsymbol{0}, \boldsymbol{s}) \leq \gamma\}$ *and* γ *is an observed distance in a subsampling block* B_n*, assume*

$$[g(\boldsymbol{h}, \Lambda_n)]_{\boldsymbol{h} \in H} \to N(0, \Sigma) \quad \textit{and} \quad [g(\boldsymbol{h}, B_n + \boldsymbol{y})]_{\boldsymbol{h} \in H} \to N(0, \Sigma)$$

as $n \to \infty$*, where* $\hat{\rho}_{AB,m}(\boldsymbol{h}, \cdot)$ *as defined in (3.1) and*

$$g(\boldsymbol{h}, \Lambda_n) = \sqrt{\frac{|\Lambda_n|}{m_n}}[\hat{\rho}_{AB,m}(\boldsymbol{h}, \Lambda_n) - \rho_{AB,m}(\boldsymbol{h})],$$
$$g(\boldsymbol{h}, B_n + \boldsymbol{y}) = \sqrt{\frac{|B_n|}{m_n}}[\hat{\rho}_{AB,m}(\boldsymbol{h}, B_n + \boldsymbol{y}) - \rho_{AB,m}(\boldsymbol{h})].$$

Further assume that (3.22), (3.23), $m_n^3 = o(n)$*, and* $\lim_{n\to\infty} m_n^4 n^6 \sum_{\boldsymbol{l} \in \mathbb{Z}^2 \setminus B_{mn}} \alpha_{i,j}(d(\boldsymbol{0}, \boldsymbol{l})) = 0$ *for* $2c \leq i + j \leq 8c$*, where* $c = \#B_\gamma$*. Then (3.26) holds, where*

$$\hat{\Sigma}_n(\boldsymbol{h}) = \left(\frac{|B_n|}{m_n|\Lambda_n^{1-c}|}\right) \sum_{\boldsymbol{y} \in \Lambda_n^{1-c}} \left(\hat{\rho}_{AB,m}(\boldsymbol{h}, B_n + \boldsymbol{y}) - \bar{\rho}_{AB,m}(\boldsymbol{h}, B_n)\right)^2,$$

$$\begin{aligned} &\hat{\Sigma}_n(\boldsymbol{h}^1, \boldsymbol{h}^2) \\ = \; &\left(\frac{|B_n|}{m_n|\Lambda_n^{1-c}|}\right) \sum_{\boldsymbol{y} \in \Lambda_n^{1-c}} (\hat{\rho}_{AB,m}(\boldsymbol{h}^1, B_n + \boldsymbol{y}) - \bar{\rho}_{AB,m}(\boldsymbol{h}^1, B_n)) \\ &\qquad (\hat{\rho}_{AB,m}(\boldsymbol{h}^2, B_n + \boldsymbol{y}) - \bar{\rho}_{AB,m}(\boldsymbol{h}^2, B_n)). \end{aligned}$$

Proof. The proof follows Theorem 7 and the techniques in Remark 16. See Appendix C. □

3.3 The bias corrected confidence intervals and comparison of three approaches

Before discussing examples and applications, we introduce bias corrected confidence intervals proposed by Efron (1981) and compare three variance methods.

3.3.1 The bias corrected confidence intervals

Often a bias is observed when a ratio estimator is estimated through bootstraps. For example, skewness of sampling distribution is observed when a sample correlation is estimated through bootstrap. Since the ESE is a ratio estimator as well, we suspect the existence of bias when constructing confidence intervals.

Without bias correction, one assumes the distribution of $\hat{\rho}^*_{AB,m}(\boldsymbol{h}) - \hat{\rho}_{AB,m}(\boldsymbol{h})$ is similar to that of $\hat{\rho}_{AB,m}(\boldsymbol{h}) - \rho_{AB,m}(\boldsymbol{h})$, thus the $100(1-\alpha)\%$ bootstrap confidence interval for the PA-extremogram is

$$[2\hat{\rho}_{AB,m}(\boldsymbol{h}) - \hat{\rho}^*_{AB,m}(\boldsymbol{h})_{1-\alpha/2}, 2\hat{\rho}_{AB,m}(\boldsymbol{h}) - \hat{\rho}^*_{AB,m}(\boldsymbol{h})_{\alpha/2}].$$

Efron (1981) proposed the bias corrected confidence intervals to resolve the issue of bias. For example, the bias corrected $100(1-\alpha)\%$ confidence intervals for the ESE is

$$[\widehat{CDF}^{-1}(\Phi(2z_0 - z_{\alpha/2})), \widehat{CDF}^{-1}(\Phi(2z_0 + z_{\alpha/2}))], \tag{3.28}$$

where $\Phi(z)$ is the cumulative distribution function of $N(0,1)$, $\widehat{CDF}$ is the empirical cumulative distribution function of $\hat{\rho}^*_{AB,m}(\boldsymbol{h})$, and $z_0 = \Phi^{-1}\left[\widehat{CDF}(\hat{\rho}_{AB,m}(\boldsymbol{h}))\right]$. The main idea of (3.28) is that there is a monotone increasing transformation that allows the transformed random variable to follow a normal distribution and that the bias can be carried through the transformation.

To see (3.28), assume that there is a monotonic increasing function $g(\cdot)$ such that

$$\phi = g(\rho_{AB,m}(\boldsymbol{h})),\ \hat{\phi} = g(\hat{\rho}_{AB,m}(\boldsymbol{h})),\ \text{and}\ \hat{\phi}^* = g(\hat{\rho}^*_{AB,m}(\boldsymbol{h})).$$

Then

$$\hat{\phi} - \phi \sim N(-z_0\sigma, \sigma^2) \text{ and } \hat{\phi}^* - \hat{\phi} \sim^* N(-z_0\sigma, \sigma^2).$$

Let $\widehat{CDG}$ be the empirical CDF of $\hat{\phi}^*$. By the monotonicity of g,

$$\widehat{CDG}(\hat{\phi}) = \widehat{CDF}(\hat{\rho}_{AB,m}(\boldsymbol{h})) = \Phi(z_0) \text{ and}$$
$$\widehat{CDG}(\hat{\phi} + z_0\sigma \pm z_{\alpha/2}\sigma) = P(\hat{\phi} \le \phi + z_0\sigma \pm z_{\alpha/2}\sigma) = \Phi(2z_0 \pm z_{\alpha/2}),$$

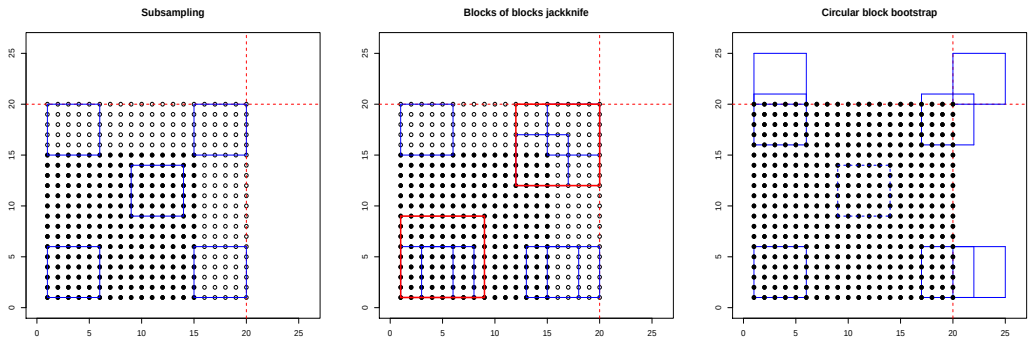

Figure 3.4: Visual comparison of three variance estimators in d=2.

which implies that

$$g^{-1}(\hat{\phi} + z_0\sigma \pm z_{\alpha/2}\sigma) = g^{-1}\widehat{CDG}^{-1}[\Phi(2z_0 \pm z_{\alpha/2})] = \widehat{CDF}^{-1}[\Phi(2z_0 \pm z_{\alpha/2})].$$

Since the $100(1-\alpha)\%$ confidence interval for ϕ is $[\hat{\phi} + z_0\sigma - z_{\alpha/2}\sigma, \hat{\phi} + z_0\sigma + z_{\alpha/2}\sigma]$, the bias corrected confidence interval (3.28) is derived. See Efron (1981) for more.

3.3.2 Comparison of three approaches

Here, we provide graphical comparison the CBB, blocks of blocks jackknife, and subsampling variance estimators to highlight their differences. See Figure 3.4. Subsampling variance estimator is a sample variance of $\hat{\rho}_{AB,m}(\boldsymbol{h})$ from all subsample boxes (solid boxes). On the other hand, blocks of blocks jackknife variance estimator is a sample variance of $\hat{\rho}_{AB,m}(\boldsymbol{h})$ from blocks of subsample boxes (bigger boxes). Lastly, the CBB estimates a variance from re-samples constructed from the wrapped original data.

All three methods use a fixed block size, but the difference is that the CBB wraps the original data. This allows the anchoring points to be selected from the entire original data, while subsampling and blocks of blocks jackknife do not.

3.4 Examples

We revisit examples discussed in Chapter 2.

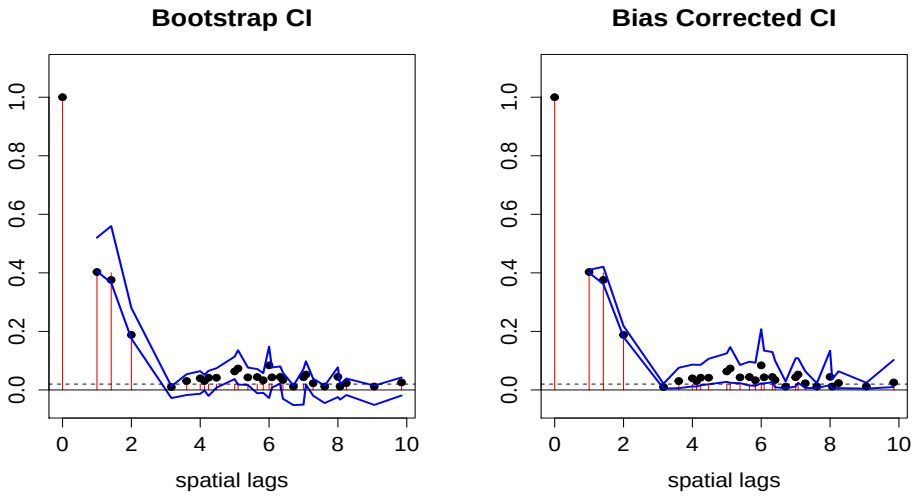

Figure 3.5: The 95% bootstrap and the bias corrected confidence intervals the MMA(1).

3.4.1 Max-Moving Average (MMA)

Recall that the max-moving average (MMA) process is defined as

$$X_{\boldsymbol{t}} = \max_{\boldsymbol{s}\in\mathbb{Z}^2} w(\boldsymbol{s}) Z_{\boldsymbol{t}-\boldsymbol{s}}, \tag{3.29}$$

where $\{Z_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{Z}^2\}$ is an iid sequence of unit Fréchet random variables, $w(\boldsymbol{s}) > 0$ and $\sum_{\boldsymbol{s}\in\mathbb{Z}^2} w(\boldsymbol{s}) < \infty$. With $w(\boldsymbol{s}) = I(\|\boldsymbol{s}\| \leq 1)$ the process (3.29) becomes the MMA(1).

$$\rho_{(1,\infty)(1,\infty)}(\boldsymbol{h}) = \lim_{n\to\infty} P(X_{\boldsymbol{h}} > a_{m_n} | X_{\boldsymbol{0}} > a_{m_n}) = \begin{cases} 1, & \text{if} \quad \|\boldsymbol{h}\| = 0, \\ 2/5, & \text{if} \quad \|\boldsymbol{h}\| = 1, \sqrt{2}, \\ 1/5, & \text{if} \quad \|\boldsymbol{h}\| = 2, \\ 0, & \text{if} \quad \|\boldsymbol{h}\| > 2. \end{cases} \tag{3.30}$$

Since the process is *2-dependent*, conditions for Theorem 5, 6, and 7 are easily checked.

We use *rmaxstab* in *SpatialExtremes* package in R to generate the MMA(1) for $\Lambda_n = \{1, ..., 40\}^2 \in \mathbb{Z}^2$. For the ESE, we use $A = B = (1, \infty)$ and $a_m = .98$ upper quantile. Figure 3.5 shows the confidence interval from the CBB (left) and the bias corrected confidence interval using Efron (1981) (right). In the figure, the bars and dots correspond to $\rho_{AB}(\boldsymbol{h})$ and $\hat{\rho}_{AB,m}(\boldsymbol{h})$, respectively. The solid lines around the dots are 95% confidence intervals.

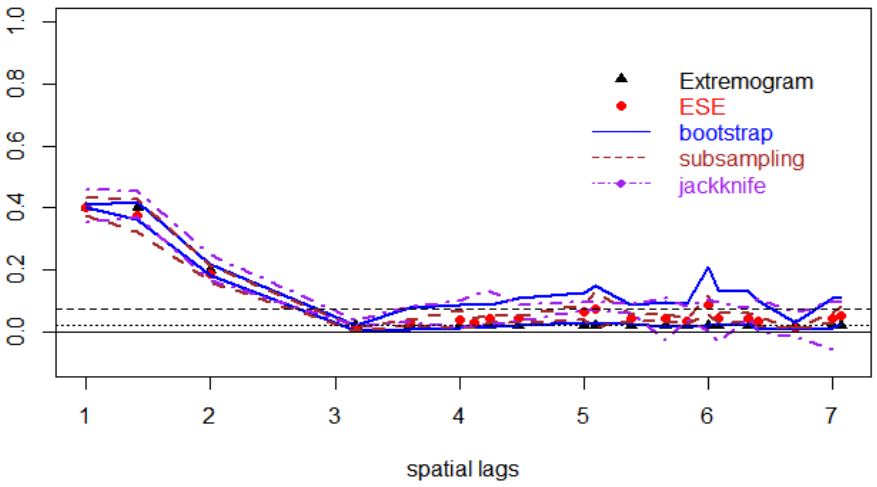

Figure 3.6: The 95% confidence interval from the CBB (the bias correction), jackknife, and subsampling for the MMA(1)

Observe that the bias corrected confidence bands are more symmetric and do not include negative value for all distance. The dotted line corresponds to $P(X_{\mathbf{0}} > a_m) = 0.02$.

Figure 3.6 compares confidence intervals from bootstrap and two variance estimators. The PA-extremogram (triangles), ESE (circles), and confidence intervals from the CBB (solid line), blocks of blocks jackknife (dotted line), and subsampling (dashed line with diamonds) are presented. For the CBB, we use a block size of 10 and 1000 simulations, then compute the bias corrected confidence interval. For jackknife, 36 blocks are created by choosing the block length of 35 and window length of 1. Then, 9 blocks of blocks are formed by stacking 4 blocks horizontally and vertically. For subsampling, a subsampling block size of 30 is used. Three methods give similar confidence band widths for $\|\boldsymbol{h}\| \leq 2$.

Table 3.1 - Table 3.3 present statistics and the 95% confidence interval estimated from the three methods. For example, to get the confidence interval of $(0.385, 0.421)$ for $\|\boldsymbol{h}\| = 1$ based on subsampling, recall (3.27). Then

$$\begin{aligned} & (0.40293 - 1.96 \times \sqrt{|B_n|/|\Lambda_n|} \times \sqrt{0.00141}, 0.40293 + 1.96 \times \sqrt{|B_n|/|\Lambda_n|} \times \sqrt{0.00141}) \\ = \; & (0.40293 - 1.96 \times 0.25 \times \sqrt{0.00141}, 0.40293 + 1.96 \times 0.25 \times \sqrt{0.00141}). \end{aligned}$$

$\|\boldsymbol{h}\|$	$\rho(\|\boldsymbol{h}\|)$	$\hat{\rho}(\|\boldsymbol{h}\|)$	CBB	SS	JK	$\|\boldsymbol{h}\|$	SS (var)	JK (var)
1	0.40	0.40	0.36	0.38	0.39	1	0.00141	0.00080
$\sqrt{2}$	0.40	0.38	0.30	0.34	0.36	$\sqrt{2}$	0.00567	0.00264
2	0.20	0.19	0.15	0.17	0.18	2	0.00117	0.00066
$\sqrt{10}$	0.02	0.01	0.02	0.03	0.03	$\sqrt{10}$	0.00007	0.00038
$\sqrt{13}$	0.02	0.03	0.03	0.09	0.09	$\sqrt{13}$	0.00040	0.00337
4	0.02	0.04	0.05	0.12	0.12	4	0.00088	0.00604

Table 3.1: Estimation results from the CBB, subsampling ($s_n = 10$) and jackknife ($j_k = 10$) for the MMA (1) with $a_m = .98$ quantile

The rest of confidence bands are derived in the same fashion. Using the subsampling variance estimation, the variance-covariance matrix can be estimated as shone in Table 3.2. From the permutation bands in Figure 3.6, one can infer that the extremal dependence for spatial lags greater than 2 is not significant. Thus, only the first 3×3 of the variance covariance matrix are considered.

$\|\boldsymbol{h}\|$	**1**	$\sqrt{2}$	**2**
1	0.00141	0.00202	0.00013
$\sqrt{2}$	0.00202	0.00567	0.00077
2	0.00013	0.00077	0.00117

Table 3.2: Estimated variance - covariance matrix for the MMA (1) using subsampling variance estimation

$\|\boldsymbol{h}\|$	CBB (lower)	CBB (upper)	SS (l)	SS (u)	JK (l)	JK (u)
1	0.401	0.410	0.385	0.421	0.389	0.417
$\sqrt{2}$	0.362	0.421	0.339	0.413	0.351	0.401
2	0.181	0.218	0.171	0.205	0.175	0.201

Table 3.3: 95% confidence intervals from the CBB, subsampling ($s_n = 10$) and jackknife ($j_k = 10$) for the MMA (1) with $a_m = .98$ quantile

Now we consider another example, the process (3.29) with $w(\boldsymbol{s}) = \phi^{\|\boldsymbol{s}\|}$, where $0 < \phi < 1$. Then

$$X_{\boldsymbol{t}} = \max_{\boldsymbol{s} \in \mathbb{Z}^2} \phi^{\|\boldsymbol{s}\|} Z_{\boldsymbol{t}-\boldsymbol{s}} \quad \text{for} \quad \sum_{\boldsymbol{l} \in \mathbb{Z}^2} \phi^{\|\boldsymbol{l}\|} = \sum_{0 \le \|\boldsymbol{l}\| < \infty} \phi^{\|\boldsymbol{l}\|} p(\|\boldsymbol{l}\|) < \infty, \tag{3.31}$$

where $p(\|\boldsymbol{l}\|) = \#\{\boldsymbol{s} \in \mathbb{Z}^2 : d(\boldsymbol{0}, \boldsymbol{s}) = \|\boldsymbol{l}\|\}$. The extremogram with $A = B = (1, \infty)$ is

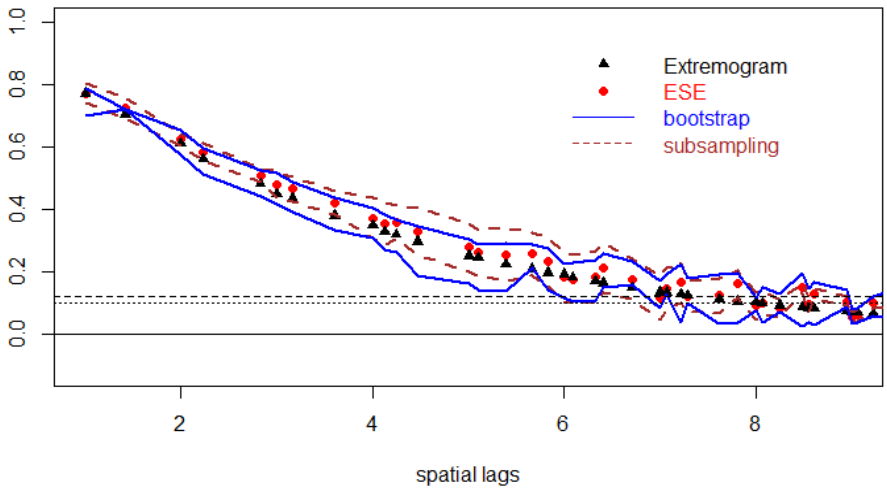

Figure 3.7: The 95% confidence interval from the CBB (the bias correction), jackknife, and subsampling for the IMMA

derived as

$$\rho_{(1,\infty)(1,\infty)}(\boldsymbol{h}) = \frac{\sum_{\frac{\|\boldsymbol{h}\|}{2} \le \|\boldsymbol{l}\| < \infty} \phi^{\|\boldsymbol{l}\|} [2p(\|\boldsymbol{l}\|) - q_{\boldsymbol{h}}(\|\boldsymbol{l}\|)]}{\sum_{0 \le \|\boldsymbol{l}\| < \infty} \phi^{\|\boldsymbol{l}\|} p(\|\boldsymbol{l}\|)}, \tag{3.32}$$

where $q_{\boldsymbol{h}}(\|\boldsymbol{l}\|) = \#\{\boldsymbol{s} \in \mathbb{Z}^2 : \min(\|\boldsymbol{s}\|, \|\boldsymbol{h} + \boldsymbol{s}\|) = \|\boldsymbol{l}\|\}$, the number of pairs with minimum distance to $\mathbf{0}$ or $\boldsymbol{h}$ equals $\|\boldsymbol{l}\|$. For the derivation of (3.32), we refer to Chapter 2.

Figure 3.7 presents confidence intervals computed from bootstrap and subsampling variance estimators for the process (3.31). The setting for the resampling methods and representation in the figure are same as Figure 3.6 except that $a_m = .90$ upper quantile is used. Two confidence intervals suggest that there is no extremal dependence for $\|\boldsymbol{h}\| > 7$.

Example 4. *The process (3.31) satisfies conditions in Theorem 5 if* $b_n = n^{2/5}$ *and* $m_n = n^{1/10}$.

Proof. From Theorem 2.1 in Chapter 2 (3.2) is checked since the choice of b_n, m_n, and the selection of $r_n = n^{1/25}$ satisfy necessary conditions of the theorem. Other conditions in Theorem 5, such as $b_n = o(n)$, (3.3), and $m_n^7/b_n^2 \to 0$ in (3.4), are easily verified. Lastly,

the mixing condition in (3.4) can be checked by Corollary 2.2 in Dombry and Eyi-Minko (2012) since

$$m_n^2 b_n^2 \sum_{\boldsymbol{l} \in \mathbb{Z}^2 \setminus B_{m_n}} \alpha_{i,j}(d(\mathbf{0}, \boldsymbol{l})) = O\left(m_n^4 b_n^2 \phi^{m_n}\right).$$

The right-hand side converges to 0 as $n \to \infty$ since $0 < \phi < 1$. See Example 3.2 in Chapter 2. □

Example 5. *The process (3.31) satisfies conditions in Theorem 6 if* $M = n^{9/10}$ *and* $m_M = n^{1/5}$.

Proof. The proof is analogous to Example 4. One can check the selected M, m_M and the choice of $r_n = n^{1/25}$ satisfy the conditions of Theorem 2.1 in Chapter 2, thus (3.9) follows. Growth rates conditions on m_M, M and n in (3.11) and (3.14) are easily verified. To show (3.10), recall (3.32), $p(\|\boldsymbol{l}\|) = O(\|\boldsymbol{l}\|)$, and $|\phi| < 1$. Then by Corollary 2.2 in Dombry and Eyi-Minko (2012),

$$\sum_{\boldsymbol{l} \in \mathbb{Z}^2} \alpha_{k^2,k^2}(d(\mathbf{0}, \boldsymbol{l}))^{\frac{p-2}{p}} \leq \text{ const} \sum_{\boldsymbol{l} \in \mathbb{Z}^2} \alpha_{1,1}(d(\mathbf{0}, \boldsymbol{l}))^{\frac{p-2}{p}} \leq \text{ const} \sum_{\boldsymbol{l} \in \mathbb{Z}^2} \|\boldsymbol{l}\|^{\frac{p-2}{p}} \ \phi^{\|\boldsymbol{l}\| \frac{p-2}{p}} < \infty.$$

Other mixing conditions (3.11)-(3.13) are shown in a similar fashion. In particular, the second condition in (3.13) requires Corollary 2.2 in Dombry and Eyi-Minko (2012), or Lemma 2 in Davis et al. (2013b),

$$\alpha_{1,\infty}(\|\boldsymbol{l}\|) \leq \sum_{k \geq \|\boldsymbol{l}\|} \rho_{(1,\infty)(1,\infty)}(k) = O\left(\phi^{\|\boldsymbol{l}\|/2}\right) \tag{3.33}$$

and $\alpha_{1,\infty}^T(k) = \alpha_{c,\infty}(k)$ for sufficiently large k, where $c = \#E_{\boldsymbol{i}}$. □

3.4.2 Brown-Resnick process

For the definition of the Brown-Resnick process, we refer to Chapter 2 or Kabluchko et al. (2009). As done in Chapter 2, the extremogram for the Brown-Resnick process $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^d\}$ with $A = (c_A, \infty)$ and $B = (c_B, \infty)$ is

$$\rho_{AB}(\boldsymbol{h}) = \bar{\Phi}_{c_A, c_B}\left(\delta(\boldsymbol{h})\right) + \frac{c_A}{c_B} \bar{\Phi}_{c_B, c_A}\left(\delta(\boldsymbol{h})\right), \tag{3.34}$$

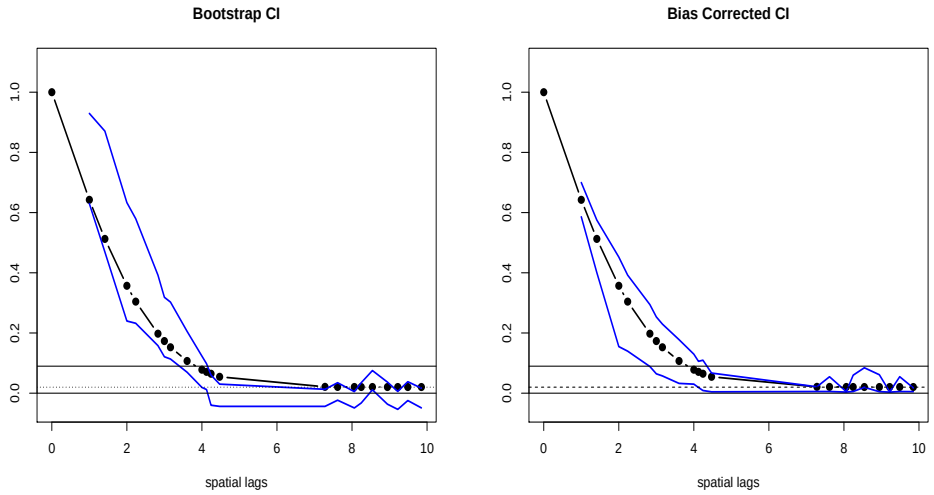

Figure 3.8: The comparison of the bootstrap and the bias-corrected confidence interval for the BR process.

where $\delta(\boldsymbol{h})$ is associated with the covariance of a underlying Gaussian random fields for the Brown-Resnick process and $\Phi_{\boldsymbol{y}^1,\boldsymbol{y}^2}(\delta(\boldsymbol{h})) = \Phi\left(\frac{\log(\boldsymbol{y}^2/\boldsymbol{y}^1)}{\sqrt{2\delta(\boldsymbol{h})}} + \sqrt{\frac{\delta(\boldsymbol{h})}{2}}\right)$.

Figure 3.8 compares the bootstrap confidence interval (left) and the bias-corrected confidence interval (right) for $\rho_{(1,\infty),(1,\infty),.98}(\|\boldsymbol{h}\|)$ for the Brown-Resnick process. The solid circles correspond to the ESE and two horizontal solid lines are permutation based confidence bands which confirms the extremal dependence is not significant when spatial lags are greater than 4. The bias-corrected confidence intervals are less asymmetric and all non-negative. Sample region is $(0, 40]^2$, the bootstrap block size 10, and 1000 simulations.

Figure 3.9 shows the 95% confidence interval for $\hat{\rho}_{(1,\infty)(1,\infty),m}(\boldsymbol{h})$ from the Brown-Resnick process with $\delta(\boldsymbol{h}) = \sqrt{4/9}\,\|\boldsymbol{h}\|$ and $\Lambda_n = \{1, ..., 40\}^2 \in \mathbb{Z}^2$. For the ESE, we use $a_m = .98$ upper quantile. In the figure, the PA-extremogram (triangles), ESE (circles), and confidence intervals from the CBB with the bias correction (solid line), blocks of blocks jackknife (dotted line), and subsampling (dashed line with diamonds) are shown, where the setting are the same as the MMA(1) in Figure 3.6. Table 3.4 - Table 3.6 present inference results regarding confidence intervals.

Example 6. *Consider the Brown-Resnick process* $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{Z}^2\}$ *with* $\delta(\boldsymbol{h}) = \theta\,\|\boldsymbol{h}\|^{\alpha}$ *for* $\alpha \in$

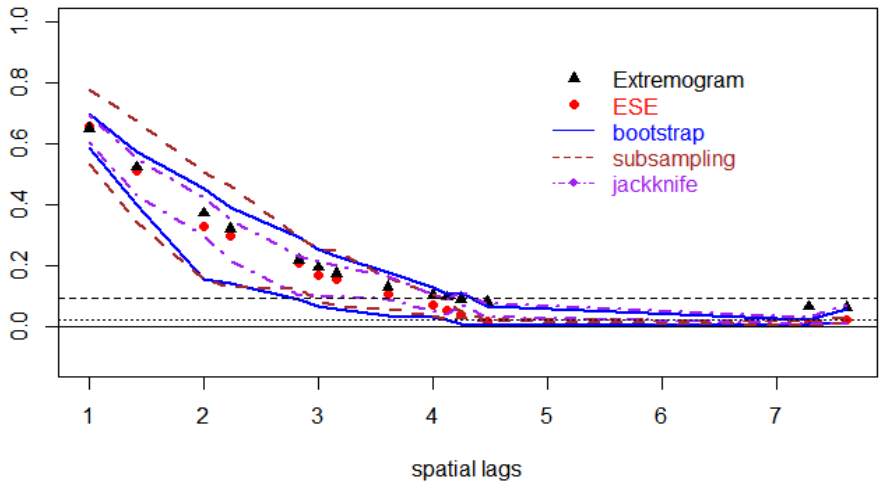

Figure 3.9: The 95% confidence interval from the CBB (the bias correction), jackknife, and subsampling for the BR process.

$(0, 2]$ *and* $\theta > 0$. *The process satisfies conditions in Theorem 5 if* $b_n = n^{2/5}$ *and* $m_n = n^{1/10}$.

Proof. The proof is identical to Example 4 except that the mixing condition in (3.4) satisfies

$$m_n^2 b_n^2 \sum_{\boldsymbol{l} \in \mathbb{Z}^2 \setminus B_{mn}} \alpha_{i,j}(d(\mathbf{0}, \boldsymbol{l})) = O\left(m_n^2 b_n^2 \sum_{\boldsymbol{l} \in \mathbb{Z}^2 \setminus B_{mn}} \|\boldsymbol{l}\|^{1-\alpha/2} e^{-\theta \|\boldsymbol{l}\|^{\alpha}/4} \right),$$

where $\alpha_{i,j}(d(\mathbf{0}, \boldsymbol{l})) \leq \text{const } ij \, \|\boldsymbol{l}\|^{-\alpha/2} e^{-\theta \|\boldsymbol{l}\|^{\alpha}/4}$ is from Corollary 2.2 in Dombry and Eyi-Minko (2012). The right-hand side converge to 0 as $n \to \infty$ under the imposed conditions. □

Example 7. *Consider the Brown-Resnick process* $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{Z}^2\}$ *with* $\delta(\boldsymbol{h}) = \theta \, \|\boldsymbol{h}\|^{\alpha}$ *for* $\alpha \in (0, 2]$ *and* $\theta > 0$. *The process satisfies conditions in Theorem 6 if* $M = n^{9/10}$ *and* $m_M = n^{1/5}$.

Proof. The proof is identical to Example 5 except that α-mixing coefficient has different upper bound, i.e.,

$$\sum_{\boldsymbol{l} \in \mathbb{Z}^2} \alpha_{k^2,k^2}(d(\mathbf{0}, \boldsymbol{l}))^{\frac{p-2}{p}} \leq \text{const} \sum_{\boldsymbol{l} \in \mathbb{Z}^2} \left(\|\boldsymbol{l}\|^{-\alpha/2} e^{-\theta \|\boldsymbol{l}\|^{\alpha}/4} \right)^{\frac{p-2}{p}} < \infty.$$

$\|\boldsymbol{h}\|$	$\rho(\|\boldsymbol{h}\|)$	$\hat{\rho}(\|\boldsymbol{h}\|)$	SS (var)	JK (var)
1.000	0.649	0.657	0.05013	0.05070
1.414	0.522	0.510	0.05016	0.01866
2.000	0.370	0.329	0.02429	0.01610
2.236	0.319	0.295	0.02716	0.00204
2.828	0.216	0.208	0.01159	0.00141
3.000	0.193	0.169	0.00695	0.00225
3.162	0.173	0.156	0.00773	0.00178
3.606	0.129	0.107	0.00374	0.00034
4.000	0.101	0.069	0.00131	0.00045
4.123	0.094	0.053	0.00079	0.00062
4.243	0.088	0.037	0.00052	0.00016
4.472	0.078	0.018	0.00018	0.00187

Table 3.4: Estimation results from the CBB, subsampling ($s_n = 10$) and jackknife ($j_k = 10$) for the BR process width $a_m = .98$ upper quantile

$\|\boldsymbol{h}\|$	1	$\sqrt{2}$	2	$\sqrt{5}$	$\sqrt{8}$	3	$\sqrt{10}$	$\sqrt{13}$	4
1	5.013	-0.188	0.159	-0.112	0.025	0.063	0.179	-0.079	-0.027
$\sqrt{2}$	-0.188	5.016	0.241	-0.062	0.050	0.105	0.043	-0.035	-0.007
2	0.159	0.241	2.429	1.186	0.004	0.319	0.214	-0.064	0.051
$\sqrt{5}$	-0.112	-0.062	1.186	2.716	0.184	0.656	0.682	-0.017	0.049
$\sqrt{8}$	0.025	0.050	0.004	0.184	1.159	-0.057	0.171	0.364	-0.044
3	0.063	0.105	0.319	0.656	-0.057	0.695	0.385	-0.067	0.072
$\sqrt{10}$	0.179	0.043	0.214	0.682	0.171	0.385	0.773	0.031	0.085
$\sqrt{13}$	-0.079	-0.035	-0.064	-0.017	0.364	-0.067	0.031	0.374	-0.054
4	-0.027	-0.007	0.051	0.049	-0.044	0.072	0.085	-0.054	0.131

Table 3.5: Estimated variance - covariance matrix for the Brown-Resnick process using subsampling variance estimation (scaled by 100).

□

Example 8. *Consider the Brown-Resnick process* $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{R}^2\}$ *with* $\delta(\boldsymbol{h}) = \theta \|\boldsymbol{h}\|^{\alpha}$ *for* $\alpha \in (0, 2]$ *and* $\theta > 0$. *Theorem 7 holds if* $c = c_n = n^{-1/35}, m_n = n^{1/9},$ *and* $\lambda_n = n^{-8/9}$ *are chosen.*

Proof. Following Example 3.4 in Chapter 2, (3.21) hold if

$$\log m_n = o(r_n^{\alpha}), \quad \sup_n \frac{\lambda_n^2 (cn)^{2a}}{m_n} < \infty \quad \text{and} \quad \sup_n \frac{m_n}{\lambda_n^2 (cn)^{2a}} < \infty \quad \text{for} \quad 0 < a < 1,$$

which are satisfied with $a = 35/36$. Also (3.22) and (3.24) are verified by the proposed rates.

$\|\boldsymbol{h}\|$	CBB (lower)	CBB (upper)	SS (l)	SS (u)	JK (l)	JK (u)
1.000	0.586	0.700	0.547	0.767	0.551	0.763
1.414	0.401	0.575	0.400	0.619	0.399	0.620
2.000	0.155	0.452	0.253	0.405	0.262	0.396
2.236	0.140	0.392	0.214	0.376	0.233	0.357
2.828	0.089	0.294	0.155	0.261	0.186	0.230
3.000	0.064	0.253	0.128	0.210	0.151	0.187
3.162	0.057	0.230	0.113	0.199	0.133	0.179
3.606	0.033	0.178	0.077	0.137	0.086	0.127
4.000	0.030	0.130	0.052	0.087	0.060	0.079
4.123	0.019	0.106	0.040	0.067	0.043	0.064
4.243	0.009	0.110	0.025	0.048	0.024	0.049
4.472	0.004	0.067	0.012	0.025	0.012	0.024

Table 3.6: 95% confidence intervals from the CBB, subsampling ($s_n = 10$) and jackknife ($j_k = 10$) for the BR process with $a_m = .98$ quantile

Other condition (3.23) and (3.25) are checked by Corollary 2.2 in Dombry and Eyi-Minko (2012). For example,

$$\alpha_{k^2,k^2}(k) = O\left(k^{4-\alpha/2}e^{-\theta k^{\alpha/4}}\right) = o(1).$$

□

3.5 Portmanteau tests

Often, one is interested in lack of spatial dependences. Since the extremal dependence tend to decay as the distance between two points are further apart, one may be interested in portmanteau tests lead-in from time series analysis,

$$H_0 : \rho_{AB,m}(\boldsymbol{h}^i) \leq 1/m \quad \text{for } i = 1, ..., k \quad \text{vs.} \quad H_1 :\neq H_0. \tag{3.35}$$

Using the L_2 consistency of the variance/covariance estimators in Theorem 7, we suggest χ^2 based test.

Proposition 9. *Assume the condition of Theorem 7 and*

$$\sqrt{\frac{n^2\lambda_n^2}{m_n}}[\hat{\rho}_{AB,m}(\boldsymbol{h}^i) - \rho_{AB,m}(\boldsymbol{h}^i)]_{i=1,...,k} \xrightarrow{d} N(\mathbf{0}, \Sigma), \tag{3.36}$$

where Σ indicates the variance-covariance matrix. Under H_0 in (3.35),

$$\sum_{i=1}^{k}\left(\hat{L}^{-1}\frac{[\hat{\rho}_{AB,m}(\boldsymbol{h}^i)-1/m]}{\sqrt{\hat{D}_n^i}}\right)^2 \sim \chi^2(k), \tag{3.37}$$

where $\hat{L}$ and $\hat{D}_n^i$ are from Cholesky decomposition of $\frac{m_n}{n^2\lambda_n^2}\Sigma$.

Proof. By Cholesky decomposition, we have

$$\Sigma = LDL^t \text{ and } \tfrac{m_n}{n^2\lambda_n^2}\Sigma = LD_nL^t,$$

where both D and $D_n = \frac{m_n}{n^2\lambda_n^2}D$ are diagonal matrices. Use D_n^i to denote i-th diagonal element of D_n. Under H_0 in (3.35),

$$\sqrt{\tfrac{n^2\lambda_n^2}{m_n}}L^{-1}[\hat{\rho}_{AB,m}(\boldsymbol{h}^i)-\tfrac{1}{m}]_{i=1,...,k} \xrightarrow{d} N(\mathbf{0},D),$$

thus it follows that

$$\sum_{i=1}^{k}\left(L^{-1}\frac{[\hat{\rho}_{AB,m}(\boldsymbol{h}^i)-\frac{1}{m}]}{\sqrt{D_n^i}}\right)^2 \sim \chi^2(k).$$

Recall from Theorem 7 that the subsampling variance/covariance estimators (3.19) and (3.20) are consistent in L_2. Write $\hat{\Sigma}$ to denote the matrix consists of (3.19) and (3.20) for diagonal and off-diagonal entries, and $\hat{L}$ and $\hat{D}$ to denote terms from Cholesky decomposition of $\hat{\Sigma}$. Then, by replacing L and D by $\hat{L}$ and $\hat{D}$, we conclude (3.37). □

The performance of portmanteau test in Proposition 9 is tested with the MMA(1) generated from $\{1,...,40\} \in \mathbb{Z}^2$. Since the MMA(1) has no spatial dependence for $\|\boldsymbol{h}\| > 2$, we focus on cases $\|\boldsymbol{h}\| \leq 5$. In the simulated data it turns out that the observed distances less or equal to 5 are $\{1, \sqrt{2}, 2, \sqrt{10}, \sqrt{13}, 4, \sqrt{17}, \sqrt{18}, \sqrt{20}, 5\}$. In Table 3.7, 20 hypothesis tests with the structure of (3.35) using a_m=.98 upper quantile are listed. The first four columns provide information on (3.35). For example, Test 1 is the simple hypothesis test of

$$H_0 : \rho_{AB,m}(1) \leq 0.02 \quad \text{vs.} \quad H_1 :\neq H_0.$$

On the other hand, Test 11 is the multiple hypothesis test of

$$H_0 : \rho_{AB,m}(\|\boldsymbol{h}^i\|) \leq 0.02 \quad \text{for } i = \{1,...,10\} \quad \text{vs.} \quad H_1 :\neq H_0.$$

Test	$\|\boldsymbol{h}^1\|$	$\|\boldsymbol{h}^k\|$	k	χ^2	$\alpha = 0.05$	$\alpha = 0.01$	theory	df	0.05	0.01
1	1	1	1	103.98	R	R	R	1	3.84	6.64
2	$\sqrt{2}$	$\sqrt{2}$	1	22.31	R	R	R	2	5.99	9.21
3	2	2	1	24.02	R	R	R	3	7.82	11.35
4	$\sqrt{10}$	$\sqrt{10}$	1	1.49	A	A	A	4	9.49	13.28
5	$\sqrt{13}$	$\sqrt{13}$	1	0.27	A	A	A	5	11.07	15.09
6	4	4	1	0.44	A	A	A	6	12.59	16.81
7	$\sqrt{17}$	$\sqrt{17}$	1	0.60	A	A	A	7	14.07	18.48
8	$\sqrt{18}$	$\sqrt{18}$	1	1.72	A	A	A	8	15.51	20.09
9	$\sqrt{20}$	$\sqrt{20}$	1	5.35	R	A	A	9	16.92	21.67
10	5	5	1	1.11	A	A	A	10	18.31	23.21
11	1	5	10	190.41	R	R	R	11	19.68	24.73
12	$\sqrt{2}$	5	9	52.17	R	R	R	12	21.03	26.22
13	2	5	8	36.27	R	R	R	13	22.36	27.69
14	$\sqrt{10}$	5	7	12.29	A	A	A	14	23.69	29.14
15	$\sqrt{13}$	5	6	9.88	A	A	A	15	25	30.58
16	4	5	5	9.49	A	A	A	16	26.3	32
17	$\sqrt{17}$	5	4	8.23	A	A	A	17	27.59	33.41
18	$\sqrt{18}$	5	3	8.23	A	A	A	18	28.87	34.81
19	$\sqrt{20}$	5	2	6.38	A	A	A	19	30.14	36.19
20	1	2	3	146.67	R	R	R	20	31.41	37.57

Table 3.7: The portmanteau test results for 20 hypothesis tests for the MMA(1).

In Table 3.7, the fifth column shows χ^2 statistics proposed in (3.37). Using the degrees of freedom and χ^2 distribution table (in the last three columns), H_0 is accepted or rejected as indicated in the sixth (seventh) columns for $\alpha = 0.05$ (0.01). The column labeled as "theory" is the expectation based on $\rho_{(1,\infty)(1,\infty)}(\|\boldsymbol{h}\|) = 0$ for $\|\boldsymbol{h}\| > 2$. For example, the valid test should reject H_0 in Test 13, but accept it for Test 14 since the MMA(1) has extremal dependence at $\|\boldsymbol{h}\| = 2$. As seen in the table, the expectation is consistent with the χ^2 based tests for both $\alpha = 0.01$ and 0.05.

Remark 15. *Proposition 9 assumes that the diagonal entry of Σ are non-zeros and that no bias exists. Hence, the χ^2 based test may fail if any of these assumptions is violated.*

3.6 Application

In this section, we revisit the applications considered in Chapter 2.

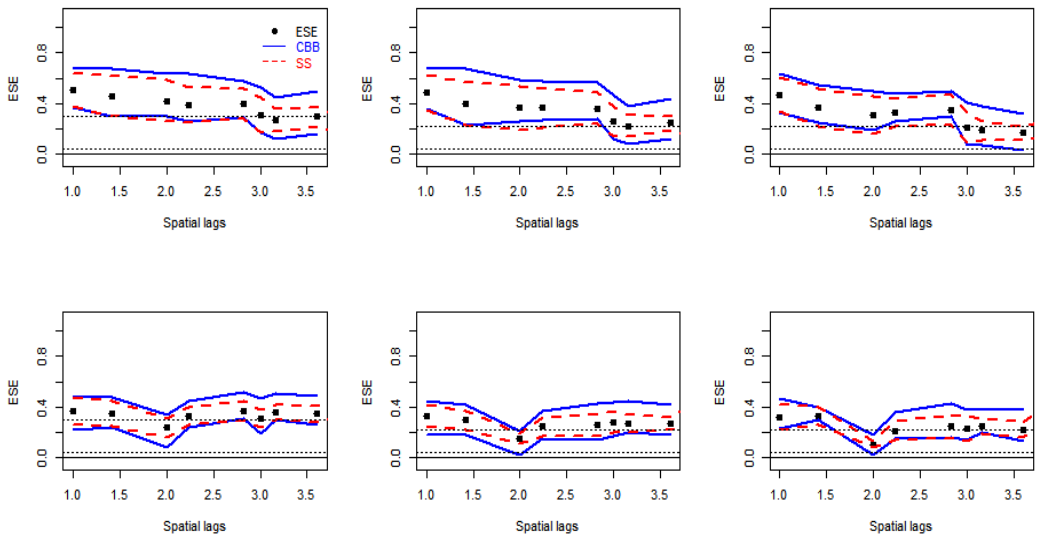

Figure 3.10: The confidence bands for the ESE using the year of 2000 (top) and 2002 (bottom) annual maxima of Florida rainfall data with $a_m = 0.70$ (left), 0.75 (middle), 0.80 (right) upper quantiles.

3.6.1 The lattice case: rainfall in a region in Florida

In this section, the bootstrapped ESE and the subsampling variance estimator for the ESE are applied to a heavy rainfall in a region in Florida. Recall from Chapter 2 that the raw data is total rainfall in 15 minute intervals from 1999 to 2004, measured on a 120×120 $(\text{km})^2$ region containing 3600 grid locations and that the 6 spatial data sets are created on a 12×12 grid under consideration consist of annual maxima of spatial maxima. Most assumptions such as max-stability are already discussed in Chapter 2.

We compute the ESE using $A = B = (1, \infty)$ and $a_m = 0.70, 0.75$, and 0.80 upper quantile. The modest thresholds were chosen to ensure enough exceedances for the ESE estimation due to the small sample size. In order to measure uncertainty, we apply the CBB and subsampling variance estimators with the block size of 4. The choice of the block size is based on the preliminary analysis showing that the ESE becomes insignificant beyond a spatial lag 3.

Figure 3.10 presents the confidence intervals for the ESE using 2000 annual maxima

(top), and 2002 annual maxima rainfall (bottom), where $a_m = 0.70$ (left), 0.75 (middle), and 0.80 (right) upper quantiles, respectively. In each panel, the ESE (dots) and the confidence intervals from the CBB (solid lines) and subsampling (dashed lines) are plotted. Based on these confidence intervals, one can check whether the extremal dependence at a certain lag is significant or not. For example, using the 0.80 upper quantile from 2002 data, one can conclude that there is a significant extremal dependence at lag 1 since the confidence interval at that lag is above 0.02 ($= P(|X| > a_m)$). On the other hand, there is no statistical evidence that extremal dependence is significant at lag 2 since the confidence intervals includes 0.2.

Another application of the variance estimation is a hypothesis testing. Observe from Figure 3.10 that the extremal dependence in 2000 is higher than 2002 overall. Thus, one may be interested to test whether the ESE from two different years are statistically the same or not. To be specific, consider

$$H_0 : \rho_{AB,m}(\boldsymbol{h})^{y_1} = \rho_{AB,m}(\boldsymbol{h})^{y_2} \qquad \text{vs.} \qquad H_1 : \rho_{AB,m}(\boldsymbol{h})^{y_1} \neq \rho_{AB,m}(\boldsymbol{h})^{y_2},$$

where $\rho_{AB,m}(\boldsymbol{h})^{\mathrm{y}}$ represents the PA-extremogram for the year y. We assume the independence of events in different years, which is not unreasonable given the data has little extremal temporal dependence even for daily frequency, as discussed in Buhl and Klüppelberg (2016). Then, the hypothesis can be tested with the test statistic

$$\frac{\hat{\rho}_{AB,m}(\boldsymbol{h})^{y_1} - \hat{\rho}_{AB,m}(\boldsymbol{h})^{y_2}}{\sqrt{\hat{\sigma}^2(\boldsymbol{h})^{y_1} + \hat{\sigma}^2(\boldsymbol{h})^{y_2}}} \sim N(0,1), \tag{3.38}$$

where $\hat{\sigma}^2(\boldsymbol{h})^{y_1}$ and $\hat{\sigma}^2(\boldsymbol{h})^{y_2}$ are the estimated sample variances of the ESE using the CBB.

Table 3.8 presents the test statistics (3.38) computed with $a_m = 0.70$ (top), 0.75 (middle) and 0.80 (bottom) upper quantile. For example, the second column of the table represents the statistics (3.38), where $y_1 = 1999$ and $y_2 = 2000$. The rest of the table can be read in a similar fashion.

From Table 3.8 and 3.9, one can observe that test statistics with $a_m = 0.70$ quantile are within $(\Phi^{-1}_{0.025}, \Phi^{-1}_{0.975})$ except 4 cases, which can be explained by 5% significance level. On the other hand, H_0 is rejected 7 and 8 times out of 90 tests when $a_m = 0.75$ and 0.80

is used for thee ESE calculation, respectively. This seems higher than $\alpha = 5\%$, but these cases are mostly associated with $\|\boldsymbol{h}\| = \sqrt{8}$ in 2001. In fact, $\hat{\rho}_{AB,m}(\sqrt{8})^{2001} = 0.06$ (and 0.05) for $a_m = 0.75$ (and 0.80), which is significantly lower than 0.25 (and 0.2). No other years shows such low level of extremal dependency. Hence, we draw a conclusion that the extremal spatial dependence in 1999-2004 are not different with 5% significance level for the spatial lag of $1, ..., 3$ with the exception of $\|\boldsymbol{h}\| = \sqrt{8}$ in 2001.

$\|\boldsymbol{h}\|$	99-00	99-01	99-02	99-03	99-04	00-01	00-02	00-03	00-04	01-02	01-03	01-04
1	-0.11	1.06	1.43	0.65	0.79	0.92	1.18	0.61	0.68	0.25	-0.35	-0.40
$\sqrt{2}$	-0.46	0.27	0.63	0.00	0.14	0.63	0.91	0.42	0.53	0.27	-0.24	-0.12
2	-0.25	0.98	1.71	0.28	-0.15	1.10	1.71	0.49	0.13	0.60	-0.69	-1.19
$\sqrt{5}$	0.00	1.01	0.78	0.54	0.09	0.71	0.50	0.39	0.06	-0.36	-0.36	-0.88
$\sqrt{8}$	0.00	3.11	0.40	1.02	1.07	2.27	0.31	0.87	0.90	-3.12	-1.15	-1.23
3	-0.38	0.15	-0.46	-0.39	-1.24	0.53	0.00	0.00	-0.70	-0.66	-0.55	-1.50
lag	99-00	99-01	99-02	99-03	99-04	00-01	00-02	00-03	00-04	01-02	01-03	01-04
1	-1.00	0.00	0.49	0.20	0.00	0.98	1.39	1.07	0.99	0.48	0.19	0.00
$\sqrt{2}$	-0.64	-0.15	0.17	0.07	0.15	0.52	0.83	0.69	0.77	0.34	0.22	0.31
2	-0.81	0.51	1.71	0.47	-0.73	1.17	2.21	1.15	0.15	0.91	-0.06	-1.12
$\sqrt{5}$	-0.61	0.67	0.62	0.39	-0.19	1.24	1.23	0.96	0.43	-0.11	-0.26	-0.86
$\sqrt{8}$	-0.75	2.94	0.19	1.25	0.56	3.93	0.94	2.09	1.29	-2.69	-1.93	-2.06
3	-0.65	-0.38	-0.96	-0.61	-1.48	0.32	-0.16	0.10	-0.63	-0.56	-0.25	-1.09
lag	99-00	99-01	99-02	99-03	99-04	00-01	00-02	00-03	00-04	01-02	01-03	01-04
1	-1.45	-0.37	0.00	0.21	-0.78	0.91	1.30	1.50	1.05	0.33	0.51	-0.14
$\sqrt{2}$	-1.33	-0.38	-1.62	-0.23	-0.87	0.79	0.42	1.06	0.65	-0.63	0.18	-0.28
2	-1.52	0.00	0.85	0.00	-1.91	1.50	2.47	1.43	0.13	0.84	0.00	-1.87
$\sqrt{5}$	-1.30	0.29	0.17	0.13	-1.47	1.40	1.44	1.20	-0.13	-0.14	-0.12	-1.55
$\sqrt{8}$	-1.47	2.77	-0.26	0.27	0.25	5.21	1.07	1.66	2.12	-2.70	-2.13	-4.26
3	-0.71	-0.55	-1.10	-0.26	-1.25	0.22	-0.22	0.42	-0.30	-0.50	0.24	-0.61

Table 3.8: The test statistics (3.38) for $a_m = 0.70$ (top), 0.75 (middle) and 0.80 (bottom) upper quantile.

a_m	y_1	y_2	$\|\boldsymbol{h}\|$	a_m	y_1	y_2	$\|\boldsymbol{h}\|$	a_m	y_1	y_2	$\|\boldsymbol{h}\|$
0.7	1999	2001	$\sqrt{8}$	0.75	1999	2001	$\sqrt{8}$	0.8	1999	2001	$\sqrt{8}$
0.7	2000	2001	$\sqrt{8}$	0.75	2000	2001	$\sqrt{8}$	0.8	2000	2001	$\sqrt{8}$
0.7	2001	2002	$\sqrt{8}$	0.75	2000	2002	2	0.8	2000	2002	2
0.7	2002	2004	2	0.75	2000	2003	$\sqrt{8}$	0.8	2000	2004	$\sqrt{8}$
				0.75	2001	2002	$\sqrt{8}$	0.8	2001	2002	$\sqrt{8}$
				0.75	2001	2004	$\sqrt{8}$	0.8	2001	2003	$\sqrt{8}$
				0.75	2002	2004	2	0.8	2001	2004	$\sqrt{8}$
								0.8	2002	2004	2

Table 3.9: The cases that the null hypothesis is rejected.

3.6.2 Non-lattice case: ground-level ozone in the eastern United States

In this section, we revisit ground-level ozone in the eastern United States using the ESE. Recall that the data is the maximum ozone reading of maximum daily 8-hour averages ozone levels in part per billion (ppb) from April to October 1997, measured at 513 ozone monitoring stations, as appeared in Figure 3.11 (left). In the figure, the locations of monitoring stations are marked with circles. Triangles indicate the stations with ground-level ozone greater than the .97 upper quantile, corresponding to 136 ppb. One can observe that the north eastern part of the region has the extreme cluster. The range of the maximum ozone reading observed from the region is (56,153) ppb. Also, note that we have discussed max-stability and stationarity in Chapter 2.

For subsampling variance estimator for non-lattice data, we follow Politis et al. (1998), as illustrated in the right panel of Figure 3.11. First we consider the shape of the region that includes all monitoring stations (outer line, rectangle trapezoid). Then the subsampling region anchoring points are chosen in a way that the subsampling region (smaller rectangle trapezoid) starting from one of these points does not exceed the original shape (bigger rectangle trapezoid). For our analysis, we pick the subsampling region ratio of 0.7. In other words, the ratio of the long base of the subsamling region to that of the region is 0.6, which gives 92 subsampling regions. It will be shown later that the inference conclusion is indifferent to the subsampling ratio of $0.6, 0.65, 0.7, 0.75$, and 0.8.

For the ESE computation, the great-circle distances are used as in Gilleland and Nychka (2005). Using the haversine formula, the shortest distance over the earth's surface is calculated from longitude/latitude of the stations. Recall from Chapter 2 that we choose 100 mile as a unit distance and calculate the ESE with $A = B = (1, \infty)$, $a_m = .97$ upper quantile and that the ESE is robust with respect to bandwidths choices of $\lambda_n = c/\log n$ for $c = 1, 2, 3, 4$, and 5 . In the estimation, $\frac{1}{\log n}$ in the bandwidth corresponds to 6 miles.

For the choice of the subsample size, Politis and Sherman (2001) suggested the algorithm to minimize

$$\frac{1}{N_m} \sum_{i=1}^{N_m} \left(\hat{\rho}_{AB,m}(\|\boldsymbol{h}\|)_i^c - \hat{\rho}_{AB,m}(\|\boldsymbol{h}\|) \right)^2, \tag{3.39}$$

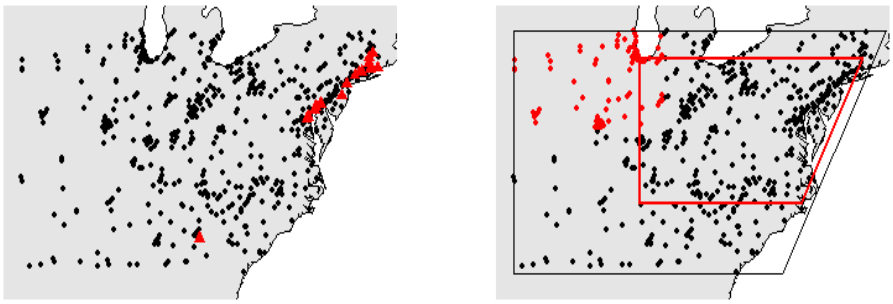

Figure 3.11: The region of ozone monitoring station in the eastern United States and subsampling scheme illustration.

where N_m is the number of subsamples, $\hat{\rho}_{AB,m}(\|\boldsymbol{h}\|)_i^c$ is the ESE computed from each subsample and $\rho_{AB,m}(\|\boldsymbol{h}\|)$ is the ESE from the original data. Table 3.10 shows (3.39) for $c = 0.6, 0.65, 0.7, 0.75$, and 0.8. The case that $c > 0.8$ is not considered since it gives too few subsampling regions. As one can see from the table, the subsample ratio choice of $c = 0.7$ minimizes the deviation from the ESE across different lags.

For the following analysis, we pick bandwidths $\lambda_n = 3/\log n$ for every 25 miles interval and calculate the confidence intervals from subsampling variance estimators (dashed lines) as presented in Fig 3.12. Two dashed lines correspond to $c = 0.6$ (red) and $c = 0.7$ (blue). The dotted line corresponds to 0.03 $(= 1 - 0.97)$. Notice that both random permutation bands (two horizontal lines) and the confidence intervals suggest that the extremal dependence disappears for $\|\boldsymbol{h}\| > 100$ miles. This finding is aligned with Gilleland and Nychka (2005) and Gilleland et al. (2006), where the authors concluded that the spatial dependence in the fourth-highest daily maximum 8-hour average ozone level fields is limited to distance less than 100 miles.

$\|\boldsymbol{h}\|$	0.6	0.65	0.7	0.75	0.8
0.25	0.048	0.034	0.032	0.045	0.044
0.5	0.045	0.034	0.032	0.029	0.034
0.75	0.008	0.008	0.007	0.007	0.008
1	0.004	0.004	0.003	0.003	0.002
1.25	0.002	0.002	0.002	0.002	0.001
N_m	92	60	21	12	8

Table 3.10: The deviation of the ESE from subsample and the original data per different subsampling ratio.

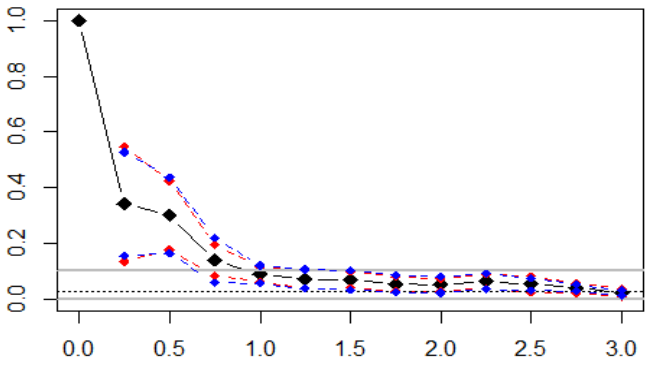

Figure 3.12: The ESE with the .97 upper quantile and c = 3 and 95% confidence interval from subsampling variance estimators.

3.7 Conclusion

In Chapter 3, we explore bootstrapped and variance estimator for the ESE. We study the asymptotic properties for the spatially adapted circular block bootstrapped estimator, and jackknife and subsampling variance estimators. When the samples are from the lattice, all three methods are investigated while only subsampling method is studied for a non-lattice case. Then, we propose χ^2 based statistics that can be used for testing the lack of extremal dependence for multiple lags. The performance of the multiple test is examined through the MMA (1) simulations, which confirms that the test is robust. In addition, real data applications to a heavy rain fall data and ozone data suggest that these methodologies

measuring an uncertainty in the ESE estimation allow us to make more interesting inference regarding the extremal spatial dependence. Thus, together with the ESE, the bootstrapped ESE and variance estimators for the ESE can be used as a informative summary tool for the dependence structures among spatial extremes.

3.8 Appendix: Proofs

3.8.1 Appendix A: Proof of Theorem 5

For the proof of Theorem 5, we introduce notation. Recall that $\Lambda_n = \{1, ..., n\}^2$ and b_n a bootstrap block size. Define $I_{\boldsymbol{s}} = I(X_{\boldsymbol{s}}/a_m \in C)$, $\bar{I} = \frac{1}{n^2}\sum_{\boldsymbol{s}\in\Lambda_n} I_{\boldsymbol{s}}$, and

$$\hat{P}_m(C) = \frac{m_n}{n^2}\sum_{\boldsymbol{s}\in\Lambda_n} I_{\boldsymbol{s}},$$

and recall that the limit exist for the numerator and denominator in the PA-extremogram, i.e.,

$$\begin{aligned} p_m(A) &= mP(X_{\mathbf{0}} \in a_m A) \to \mu(A), \\ \tau_{AB,m}(\boldsymbol{h}) &= mP(X_{\mathbf{0}} \in a_m A, X_{\boldsymbol{h}} \in a_m B) \to \tau_{AB}(\boldsymbol{h}) \end{aligned}$$

by the regularly varying assumption, where A and B are sets bounded away from the origin.

Use P^*, E^* and var* to denote the probability measure generated by the bootstrap scheme, the expected value and variance from P^*, respectively, and let $(I^*_{\boldsymbol{s}})$ be a bootstrap sequence generated from the sample $(I_{\boldsymbol{s}}, \boldsymbol{s} \in \Lambda_n)$ by the CBB. For the bootstrap block, write $B(\boldsymbol{t}, b_n) = \{(u_1, u_2) : \boldsymbol{t} = (t_1, t_2)$ and $t_i \le u_i \le t_i + b_n$ for $i = 1, 2\}$ a bootstrap block with the anchoring point $\boldsymbol{t}$. Notice that a re-sampled pseudo space consists of $B(\boldsymbol{t}^i, b_n)$, where $\boldsymbol{t}^i \in \mathbb{Z}^2$ for $i = \{1, ..., n^2/b_n^2\}$, is the i-th randomly selected anchoring point by the bootstrap.

The proof of Theorem 5 consists of three steps:

1. Establish the conditional central limit theorem of

$$\hat{P}^*_m(C) = \frac{b_n^2}{n^2}\sum_{i=1,...,n^2/b_n^2} \frac{m_n}{b_n^2}\sum_{s\in B(\boldsymbol{t}^i, b_n)} I^*_{\boldsymbol{s}} = \frac{m_n}{n^2}\sum_{\boldsymbol{s}\in\Lambda_n} I^*_s, \tag{3.40}$$

   where $I^*_{\boldsymbol{s}} = I(X^*_{\boldsymbol{s}}/a_m \in C)$.

2. Extend the result in Step 1 to the vectorized processes.

3. Apply the Cramér-Wold device to achieve the multivariate central limit theorem, then use δ-method to obtain the central limit theorem for the bootstrapped ESE.

Step 1: Show the conditional central limit theorem of $\hat{P}_m^*(C)$

Theorem 10 implies that the bootstrap estimator $\hat{P}_m^*(C)$ of $\hat{P}_m(C)$ is asymptotically correct. In the proof, we assume n^2/b_n^2 to be an integer and write $\hat{P}_m = \hat{P}_m(C)$ and $\hat{P}_m^* = \hat{P}_m^*(C)$ for the convenience.

Theorem 10. *Suppose a strictly stationary regularly varying random field* $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{Z}^2\}$ *with* $\alpha > 0$ *is observed on* $\Lambda_n = \{1, ..., n\}^2$. *Assume (3.2) and (3.3). Further assume that*

$$\sup_n E\left|\sqrt{\frac{m_n}{b_n^2}} \sum_{\boldsymbol{s}\in B(t,b_n)} (I_{\boldsymbol{s}} - p_{\mathbf{0}})\right|^4 < \infty, \tag{3.41}$$

where $p_{\mathbf{0}} = E(I_{\boldsymbol{s}})$. *Then*

$$E^*(\hat{P}_m^*) \xrightarrow{P} \mu(C), \quad \textit{and} \tag{3.42}$$

$$\textit{var}^*\left(\sqrt{\frac{n^2}{m_n}}\hat{P}_m^*\right) \xrightarrow{P} \sigma^2(C), \tag{3.43}$$

which implies $P^*(|\hat{P}_m^* - \mu(C)| > \delta) \xrightarrow{P} 0$, $\delta > 0$. *Also the central limit theorem holds*

$$\sup_x \left|P^*\left(\sqrt{\frac{n^2}{m_n s_n^2}}(\hat{P}_m^* - \hat{P}_m) \le x\right) - \Phi(x)\right| \xrightarrow{P} 0, \tag{3.44}$$

where $s_n^2 = \textit{var}^*\left(\sqrt{\frac{n^2}{m_n}}\hat{P}_m^*\right)$.

Proof. For (3.42), we have $E(\hat{P}_m) \to \mu(C)$ and $\text{var}(\hat{P}_m) \to 0$ from (5.6) and (5.7) in Chapter 2, thus $\hat{P}_m \xrightarrow{P} \mu(C)$. As the CBB allows each point to be selected equally,

$$E^*(\hat{P}_m^*) = \frac{m_n}{n^2} \sum_{\boldsymbol{s}\in\Lambda_n} E^*(I_{\boldsymbol{s}}^*) = \frac{m_n}{n^2} \sum_{\boldsymbol{s}\in\Lambda_n} I_{\boldsymbol{s}} = \hat{P}_m \xrightarrow{P} \mu(C).$$

Turning to (3.43), recall that bootstrap blocks are conditionally independent. From (3.40),

$$\text{var}^*\left(\sqrt{\frac{n^2}{m_n}}\hat{P}_m^*\right) = \frac{m_n}{b_n^2}\ \text{var}^*\left(\sum_{\boldsymbol{s}\in B(\boldsymbol{t},b_n)} I_{\boldsymbol{s}}^*\right) = \frac{m_n}{b_n^2}\left[\frac{1}{n^2}\sum_{\boldsymbol{t}\in\Lambda_n}\left(\sum_{\boldsymbol{s}\in B(\boldsymbol{t},b_n)}(I_{\boldsymbol{s}} - \bar{I})\right)^2\right].$$

Thus,

$$\begin{aligned} E\left[\text{var}^*\left(\sqrt{\frac{n^2}{m_n}}\hat{P}_m^*\right)\right] &= \frac{m_n}{b_n^2}E\left[\left(\sum_{\boldsymbol{s}\in B(\boldsymbol{t},b_n)}(I_{\boldsymbol{s}}-p_{\boldsymbol{0}})-b_n^2(\bar{I}-p_{\boldsymbol{0}})\right)^2\right] \\ &= E\left[\left(\sqrt{\frac{m_n}{b_n^2}}\sum_{\boldsymbol{s}\in B(\boldsymbol{t},b_n)}(I_{\boldsymbol{s}}-p_{\boldsymbol{0}})-\sqrt{\frac{b_n^2}{n^2}\frac{m_n}{n^2}}\sum_{\boldsymbol{s}\in\Lambda_n}(I_{\boldsymbol{s}}-p_{\boldsymbol{0}})\right)^2\right] \\ &= E[(J_1-J_2)^2]. \end{aligned} \tag{3.45}$$

We have $J_2 = o(1)$ since

$$\sqrt{\frac{m_n}{n^2}}\sum_{\boldsymbol{s}\in\Lambda_n}(I_{\boldsymbol{s}}-p_{\boldsymbol{0}}) \xrightarrow{d} N(0,\sigma^2(C))$$

by Theorem 5.2 in Chapter 2 and $b_n^2/n^2 \to 0$.

For J_1, the condition (3.2) implies $J_1 \xrightarrow{d} N(0,\sigma^2(C))$, thus $E(J_1)^2 \to \sigma^2(C)$. Moreover, J_1^2 is uniformly integrable by (3.41). Thus,

$$\begin{aligned} & E\left(J_1^2-\sigma^2(C)\right)^2 \\ = \;& E\left[\left(J_1^2-\sigma^2(C)\right)^2 I\left(|J_1^2-\sigma^2(C)|\le m\right)\right]+E\left[\left(J_1^2-\sigma^2(C)\right)^2 I\left(|J_1^2-\sigma^2(C)|> m\right)\right] \\ \le \;& mE\left[\left(J_1^2-\sigma^2(C)\right)\right]+E\left[\left(J_1^2-\sigma^2(C)\right)^2 I\left(|J_1^2-\sigma^2(C)|> m\right)\right] \\ \to \;& 0. \end{aligned} \tag{3.46}$$

To show (3.44), it suffices to check conditional versions of the Lyapunov condition by Theorem 8 of Rao (2009) since $\hat{P}_m^*$ is the sum of n^2/b_n^2 independent blocks conditional on a sample $(X_{\boldsymbol{s}})$. Write

$$\sqrt{\frac{n^2}{m_n}}(\hat{P}_m^*-\hat{P}_m) = \sqrt{\frac{m_n}{n^2}}\sum_{k=1,\ldots,\left(\frac{n}{b_n}\right)^2}\sum_{\boldsymbol{s}\in B(\boldsymbol{t}^k,b_n)}(I_{\boldsymbol{s}}^*-\bar{I}) = \sum_{k=1,\ldots,\left(\frac{n}{b_n}\right)^2} Y_{nk},$$

where $Y_{nk} = \sqrt{\frac{m_n}{n^2}}\sum_{\boldsymbol{s}\in B(\boldsymbol{t}^k,b_n)}(I_{\boldsymbol{s}}^*-\bar{I})$ are independent for different $k \in \{1,...,(n/b_n)^2\}$ condition on a sample. Observe that

$$E^*[Y_{nk}] = \sqrt{\frac{m_n}{n^2}}\sum_{\boldsymbol{s}\in B(\boldsymbol{t}^k,b_n)}(E^*(I_{\boldsymbol{s}}^*)-\bar{I}) = 0.$$

For the second moment we will verify

$$\frac{n^2}{b_n^2}\, E^*\left[Y_{nk}^2\right] \xrightarrow{P} \sigma^2(C). \tag{3.47}$$

To see (3.47) notice that $\frac{n^2}{b_n^2}\, E^*\left[Y_{nk}^2\right]$ becomes

$$\begin{aligned}\frac{m_n}{b_n^2} E^*\left[\sum_{\boldsymbol{s}\in B(\boldsymbol{t}^k,b_n)} (I_{\boldsymbol{s}}^* - \bar{I})\right]^2 &= \frac{m_n}{b_n^2}\,\frac{1}{n^2}\sum_{\boldsymbol{t}^k\in\Lambda_n}\left[\sum_{\boldsymbol{s}\in B(\boldsymbol{t}^k,b_n)} (I_{\boldsymbol{s}} - p_{\boldsymbol{0}}) - b_n^2(\bar{I} - p_{\boldsymbol{0}})\right]^2 \\ &= \frac{1}{n^2}\sum_{\boldsymbol{t}^k\in\Lambda_n}[(J_1^k)^2 - 2J_1^k J_2 + J_2^2], \end{aligned}\tag{3.48}$$

where $J_1^k = \sqrt{\frac{m_n}{b_n^2}}\sum_{\boldsymbol{s}\in B(\boldsymbol{t}^k,b_n)}(I_{\boldsymbol{s}} - p_{\boldsymbol{0}})$ and $J_2 = \sqrt{m_n b_n^2}(\bar{I} - p_{\boldsymbol{0}})$ that is defined in (3.45). Since J_2 is ignorable and $E(J_1^k)^2 \xrightarrow{P} \sigma^2(C)$ for all k, the independence between different bootstrap blocks (conditionally) implies that the right-hand side of (3.48) converges to $\sigma^2(C)$ in probability. This verifies (3.47).

Hence, $s_n^2 = \sum_{\boldsymbol{t}^k,k=1}^{n^2/b_n^2} E^*[|Y_{nk}|^2] \xrightarrow{P} \sigma^2(C)$. From $|Y_{nk}| \le \sqrt{\frac{m_n}{n^2}}b_n^2$ and (3.3),

$$\sum_{k=1}^{n^2/b_n^2}\frac{1}{s_n^3}E^*[|Y_{nk}|^3] \le \frac{\sqrt{\frac{m_n}{n^2}}b_n^2}{s_n} = \sqrt{\frac{m_n b_n^4}{n^2}}\frac{1}{s_n} \to 0.$$

This shows Lyapunov condition (conditional on a sample) is satisfied, which completes the proof. □

Step 2: Asymptotics of $\hat{P}_m^*$ based on a vectorized process in space

Define the vectorized process $Y_{\boldsymbol{t}} = X_{D_{\boldsymbol{t}}}$ with $D_{\boldsymbol{t}} = \{\boldsymbol{s}\in\mathbb{Z}^d : d(\boldsymbol{t},\boldsymbol{s}) \le p\}$ for a fixed p. Write $\mu_A(D_{\boldsymbol{0}}) = \lim_x P\left(\frac{Y_{\boldsymbol{t}}}{\|Y_{\boldsymbol{t}}\|}\in A \middle| \|Y_{\boldsymbol{t}}\| > x\right)$ and

$$\tau_{A\times B}(D_{\boldsymbol{0}}\times D_{\boldsymbol{l}}) = \lim_x P\left(\frac{(Y_{\boldsymbol{0}}, Y_{\boldsymbol{l}})}{\|vec\{Y_{\boldsymbol{0}}, Y_{\boldsymbol{l}}\}\|}\in A\times B \middle| \|vec\{Y_{\boldsymbol{0}}, Y_{\boldsymbol{l}}\}\| > x\right).$$

Then Theorem 10 can be extended to the vectorized process by replacing $X_{\boldsymbol{s}}$ by $Y_{\boldsymbol{s}}$. For example,

$$\hat{P}_m^*(C) = \frac{m_n}{n^2}\sum_{\boldsymbol{s}\in\Lambda_n} I^*(Y_{\boldsymbol{s}}/a_m \in C). \tag{3.49}$$

Some of vectorized process $Y_{\boldsymbol{s}}$ may stack (conditionally) independent data together, which can be categorized into two cases: 1) anchoring points are near the boundary of Λ_n or 2)

$Y_{\boldsymbol{s}}$ is constructed by random fields from the different blocks. These cases are ignorable as $n \to \infty$ for a fixed p. Thus, $\{Y_{\boldsymbol{s}}\}$ can be treated as $\{X_{\boldsymbol{s}}\}$ in Theorem 10. In other words, if the conditions of Theorem 10 hold for a strictly stationary regularly varying random vectors $(Y_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{Z}^2)$, then (3.42) - (3.44) hold for $\hat{P}_m^*(C)$ defined in (3.49).

Step 3: Apply the Cramér-Wold device

The following Corollary 3.8.1 shows that the application of the Cramér-Wold device on Theorem 10 to random vectors $(Y_{\boldsymbol{s}})$. To this end, we redefine

$$\mu(A) = \lim_{x\to\infty} P\left(\frac{X_{\boldsymbol{t}}}{\|Y_{\boldsymbol{t}}\|} \in A \Big| \|Y_{\boldsymbol{t}}\| > x\right).$$

The corollary implies that the multivariate central limit theorem holds. The central limit theorem for the bootstrapped ESE follows by δ-method, which is analogous to the proof of Theorem 1 in Chapter 2. For the convenience, we consider $d = 2$ for Corollary 3.8.1, but the extension to general cases is straightforward.

Corollary 3.8.1. *Assume that the conditions of Theorem 10 hold for a strictly stationary regularly varying random vectors $\{Y_s, s \in \mathbb{Z}^2\}$ with $\alpha > 0$. For sets A_1, A_2 bounded away from $\boldsymbol{0}$ with $\mu(A_2) > 0$, assume that*

$$\sup_n E\left|\frac{m}{b_n^2}\frac{1}{n^2}\sum_{\boldsymbol{t}\in\Lambda_n}\sum_{\boldsymbol{s}\in B(\boldsymbol{t},b_n)}(I_{\boldsymbol{s}}(A_1) - p_{\boldsymbol{0}}(A_1))(I_{\boldsymbol{s}}(A_2) - p_{\boldsymbol{0}}(A_2))\right|^2 < \infty, \quad \textit{and}$$

$$\sup_n E\left|\frac{m}{b_n^2}\frac{1}{n^2}\sum_{t\in\Lambda_n}\sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\in B(t,b_n)}(I_{\boldsymbol{s}^1}(A_1) - p_{\boldsymbol{0}}(A_1))(I_{\boldsymbol{s}^2}(A_2) - p_{\boldsymbol{0}}(A_2))\right|^2 < \infty, \tag{3.50}$$

where $I_{\boldsymbol{s}}(A) = I\left(Y_{\boldsymbol{s}}/a_m \in A\right)$ and $p_{\boldsymbol{0}}(A) = P(Y_{\boldsymbol{s}}/a_m \in A)$. Then

$$S_n = \sqrt{\frac{n^2}{m_n}}[\hat{P}_m^*(A_i) - \hat{P}_m(A_i)]_{i=1,2} \xrightarrow{d} N(0, \Sigma),$$

where Σ is the asymptotic covariance matrix with

$$\Sigma_{ii} = \mu_{A_i}(D_{\boldsymbol{0}}) + \sum_{\boldsymbol{l}\neq\boldsymbol{0}\in\mathbb{Z}^2}\tau_{A_i\times A_i}(D_{\boldsymbol{0}}\times D_{\boldsymbol{l}}) \quad \textit{for} \quad i = 1, 2, \textit{ and}$$

$$\Sigma_{ij} = \mu_{A_i\cap A_j}(D_{\boldsymbol{0}}) + \sum_{\boldsymbol{l}\neq\boldsymbol{0}\in\mathbb{Z}^2}\tau_{A_i\times A_j}(D_{\boldsymbol{0}}\times D_{\boldsymbol{l}}) \quad \textit{for } 1 \le i \neq j \le 2.$$

Proof. By the Cramér-Wold device, it suffices to show that

$$\boldsymbol{z}'S_n \xrightarrow{d} N(\mathbf{0}, \boldsymbol{z}'\Sigma\boldsymbol{z}),\ \boldsymbol{z} \in \mathbb{R}^2.$$

One can easily check $E^*(\boldsymbol{z}'S_n) = \mathbf{0}$. With the definition of $\tilde{P}^*_m(\cdot) = \hat{P}^*_m(\cdot) - \hat{P}_m(\cdot)$, we will verify

$$\begin{aligned}
&\mathrm{var}^*(\boldsymbol{z}'S_n) \\
&= (n^2/m)\left[z_1^2\ \mathrm{var}^*(\tilde{P}^*_m(A_1)) + z_2^2\ \mathrm{var}^*(\tilde{P}^*_m(A_2)) + 2z_1z_2\ \mathrm{cov}^*(\tilde{P}^*_m(A_1), \tilde{P}^*_m(A_2))\right] \\
&\xrightarrow{P} \boldsymbol{z}'\Sigma\boldsymbol{z}.
\end{aligned}$$

Notice from Theorem 10 that $\mathrm{var}^*\left(\sqrt{\frac{n^2}{m}}\tilde{P}^*_m(A_i)\right) \xrightarrow{P} \Sigma_{ii}$ for $i = 1, 2$. For the cross-product term, $\mathrm{cov}^*\left(\sqrt{\frac{n^2}{m}}\tilde{P}^*_m(A_1), \sqrt{\frac{n^2}{m}}\tilde{P}^*_m(A_2)\right)$ equals

$$\begin{aligned}
&E^*\left(\sqrt{\frac{n^2}{m}}\tilde{P}^*_m(A_1)\sqrt{\frac{n^2}{m}}\tilde{P}^*_m(A_2)\right) \\
= \ &\frac{m}{n^2}E^*\left(\left[\sum_{i=1}^{(\frac{n}{b_n})^2}\sum_{\boldsymbol{s}\in B(\boldsymbol{t}^i, b_n)}(I^*_{\boldsymbol{s}}(A_1) - \bar{I}(A_1))\right]\left[\sum_{i=1}^{(\frac{n}{b_n})^2}\sum_{\boldsymbol{s}\in B(\boldsymbol{t}^i, b_n)}(I^*_{\boldsymbol{s}}(A_2) - \bar{I}(A_2))\right]\right).
\end{aligned}$$

Due to the conditional independence among bootstrap blocks, the right-hand side is equivalent to

$$\frac{m}{b_n^2}E^*\left(\left[\sum_{\boldsymbol{s}\in B(\boldsymbol{t}^i, b_n)}(I^*_{\boldsymbol{s}}(A_1) - \bar{I}(A_1))\right]\left[\sum_{\boldsymbol{s}\in B(\boldsymbol{t}^i, b_n)}(I^*_{\boldsymbol{s}}(A_2) - \bar{I}(A_2))\right]\right) = E_1 + E_2,$$

where

$$\begin{aligned}
E_1 &= \frac{m}{b_n^2}\frac{1}{n^2}\sum_{\boldsymbol{t}\in\Lambda_n}\sum_{\boldsymbol{s}\in B(\boldsymbol{t}, b_n)}(I_{\boldsymbol{s}}(A_1) - \bar{I}(A_1))(I_{\boldsymbol{s}}(A_2) - \bar{I}(A_2)) \text{ and} \\
E_2 &= \frac{m}{b_n^2}\frac{1}{n^2}\sum_{\boldsymbol{t}\in\Lambda_n}\sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\in B(\boldsymbol{t}, b_n)}(I_{\boldsymbol{s}^1}(A_1) - \bar{I}(A_1))(I_{\boldsymbol{s}^2}(A_2) - \bar{I}(A_2)).
\end{aligned}$$

We will show

$$E_1 \xrightarrow{P} \mu_{A_1\cap A_2}(D_{\mathbf{0}}) \quad \text{and} \quad E_2 \xrightarrow{P} \textstyle\sum_{\boldsymbol{l}\neq\mathbf{0}\in\mathbb{Z}^2}\tau_{A_1\times A_2}(D_{\mathbf{0}}\times D_{\boldsymbol{l}}).$$

For E_1, write

$$\begin{aligned} E_1 &= \frac{m}{b_n^2}\frac{1}{n^2}\sum_{\boldsymbol{t}\in\Lambda_n}\sum_{\boldsymbol{s}\in B(\boldsymbol{t},b_n)}(I_{\boldsymbol{s}}(A_1)-\bar{I}(A_1))(I_{\boldsymbol{s}}(A_2)-\bar{I}(A_2)) \\ &= \frac{m}{b_n^2}\frac{1}{n^2}\sum_{\boldsymbol{t}\in\Lambda_n}\sum_{\boldsymbol{s}\in B(\boldsymbol{t},b_n)}[(I_{\boldsymbol{s}}(A_1)-p_{\boldsymbol{0}}(A_1))+(p_{\boldsymbol{0}}(A_1)-\bar{I}(A_1))] \\ &\qquad\qquad [(I_{\boldsymbol{s}}(A_2)-p_{\boldsymbol{0}}(A_2))+(p_{\boldsymbol{0}}(A_2)-\bar{I}(A_2))] \\ &\sim \frac{m}{b_n^2}\frac{1}{n^2}\sum_{\boldsymbol{t}\in\Lambda_n}\sum_{\boldsymbol{s}\in B(\boldsymbol{t},b_n)}[(I_{\boldsymbol{s}}(A_1)-p_{\boldsymbol{0}}(A_1))(I_{\boldsymbol{s}}(A_2)-p_{\boldsymbol{0}}(A_2))]. \end{aligned}$$

Hence, $E(E_1)\sim mP(Y_{\boldsymbol{0}}/a_m\in A_1\cap A_2)\to\mu_{A_1\cap A_2}(D_{\boldsymbol{0}})$. For the second moment, note from (3.50) that E_1 is uniformly integrable, thus by the same argument in (3.46)

$$E\left(E_1-\mu_{A_1\cap A_2}(D_{\boldsymbol{0}})\right)^2\to 0.$$

Similarly, $E_2\to\sum_{\boldsymbol{l}\neq\boldsymbol{0}\in\mathbb{Z}^2}\tau_{A_1\times A_2}(D_{\boldsymbol{0}}\times D_{\boldsymbol{l}})$ and $E\left(E_2-\sum_{\boldsymbol{l}\neq\boldsymbol{0}\in\mathbb{Z}^2}\tau_{A_1\times A_2}(D_{\boldsymbol{0}}\times D_{\boldsymbol{l}})\right)^2\to 0$ can be shown. Hence, we verify

$$\mathrm{cov}^*\left(\sqrt{\tfrac{n^2}{m}}\hat{P}_m^*(A_1),\sqrt{\tfrac{n^2}{m}}\hat{P}_m^*(A_2)\right)\xrightarrow{P}\Sigma_{12}.$$

Lastly, we show the asymptotic normality. Observe that $\boldsymbol{z}'S_n$ can be viewed as the sum of $(n/b_n)^2$ conditionally independent random variables. Thus, it suffices to apply the Lyapunov condition with the order 3, conditional on $(I_{\boldsymbol{s}})$. Write

$$\begin{aligned} \boldsymbol{z}'S_n &= \sqrt{\frac{m_n}{n^2}}\sum_{k=1,\ldots,\left(\frac{n}{b_n}\right)^2}\sum_{\boldsymbol{s}\in B(\boldsymbol{t}^k,b_n)}\left[z_1(I_{\boldsymbol{s}}^*(D_1)-\bar{I}(D_1))+z_2(I_{\boldsymbol{s}}^*(D_2)-\bar{I}(D_2))\right] \\ &= \sum_{k=1,\ldots,\left(\frac{n}{b_n}\right)^2}Y_{nk}, \end{aligned}$$

where $Y_{nk}=\sqrt{\frac{m_n}{n^2}}\sum_{\boldsymbol{s}\in B(\boldsymbol{t}^k,b_n)}\left[z_1(I_{\boldsymbol{s}}^*(D_1)-\bar{I}(D_1))+z_2(I_{\boldsymbol{s}}^*(D_2)-\bar{I}(D_2))\right]$. Recall that $\{Y_{nk}\}$ are conditionally independent and satisfy

$$|Y_{nk}|\le 2\max(z_1,z_2)\sqrt{m_nb_n^4/n^2}.$$

Since $s_n^2=\sum_{\boldsymbol{t}^k,k=1}^{n^2/b_n^2}E^*(Y_{nk})^2\xrightarrow{P}\boldsymbol{z}'\Sigma\boldsymbol{z}$ and (3.3),

$$\sum_{k=1}^{n^2/b_n^2}\frac{1}{s_n^3}E^*[|Y_{nk}|^3]\le\frac{const\ \sqrt{m_nb_n^4/n^2}}{s_n}\to 0,$$

which shows Lyapunov condition (conditional on $\{I_{\boldsymbol{s}}\}$) is satisfied. This completes the proof. □

Remark 16. *One can show the condition (3.4) implies (3.41) by using the argument in Lemma 18.5.2 in Ibragimov and Linnik (1971), where the authors calculated* $E\left(\sum_{i=1}^{n} X_j\right)^4 = o(n^3)$ *for stationary time series* $\{X_j, j \in \mathbb{N}\}$.

To see this, let $\tilde{I}_{\boldsymbol{s}} = I(Y_{\boldsymbol{s}}/a_m \in C) - P(Y_{\boldsymbol{s}}/a_m \in C)$. *Then*

$$
\begin{aligned}
& E\left|\sqrt{\frac{m_n}{b_n^2}} \sum_{\boldsymbol{s}\in B(\boldsymbol{t},b_n)} \tilde{I}_{\boldsymbol{s}}\right|^4 \\
&= \frac{m_n^2}{b_n^4}\, E\left(\sum_{\boldsymbol{s}\in B(\boldsymbol{t},b_n)} (\tilde{I}_{\boldsymbol{s}})^4 + \sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\in B(\boldsymbol{t},b_n)} (\tilde{I}_{\boldsymbol{s}^1})^3 \tilde{I}_{\boldsymbol{s}^2} + \sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\in B(\boldsymbol{t},b_n)} (\tilde{I}_{\boldsymbol{s}^1})^2 (\tilde{I}_{\boldsymbol{s}^2})^2 \right. \\
&\qquad \left. + \sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\neq\boldsymbol{s}^3\in B(\boldsymbol{t},b_n)} (\tilde{I}_{\boldsymbol{s}^1})^2 \tilde{I}_{\boldsymbol{s}^2} \tilde{I}_{\boldsymbol{s}^3} + \sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\neq\boldsymbol{s}^3\neq\boldsymbol{s}^4\in B(\boldsymbol{t},b_n)} \tilde{I}_{\boldsymbol{s}^1} \tilde{I}_{\boldsymbol{s}^2} \tilde{I}_{\boldsymbol{s}^3} \tilde{I}_{\boldsymbol{s}^4} \right) \\
&= A_1 + A_2 + A_3 + A_4 + A_5.
\end{aligned}
$$

As stated in Ibragimov and Linnik (1971), it suffices to estimate the last sum since the number of terms in other sums are smaller order. However, we will show that $A_1, ..., A_5 \to 0$ *as* $n \to \infty$ *for the completeness.*

For A_1, *note that* $|\tilde{I}_{\boldsymbol{s}}| \leq 2$ *and* $\#|B(\boldsymbol{t}, b_n)| = b_n^2$. *Thus,* $A_1 = O\left(m_n^2 b_n^2 / b_n^4\right) = O\left(m_n^2/b_n^2\right)$.

For the sums with two indices, for example for A_2,

$$
\begin{aligned}
\frac{m_n^2}{b_n^4} \sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\in B(\boldsymbol{t},b_n)} (\tilde{I}_{\boldsymbol{s}^1})^3 \tilde{I}_{\boldsymbol{s}^2} &= O\left(\frac{m_n^2}{b_n^4} \sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\in B(\boldsymbol{t},b_n)} E\left|(\tilde{I}_{\boldsymbol{s}^1})^3 \tilde{I}_{\boldsymbol{s}^2}\right| \right) \\
&= O\left(\frac{m_n^2}{b_n^4} \sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\in B(\boldsymbol{t},b_n)} \alpha_{c,c}(|\boldsymbol{s}^2 - \boldsymbol{s}^1|) \right)
\end{aligned}
$$

which becomes $O\left(\frac{m_n^2}{b_n^4}\right)$. *The same argument holds for* A_3.

For A_4, let $j = \min\{d(\boldsymbol{s}^1, \boldsymbol{s}^2), d(\boldsymbol{s}^1, \boldsymbol{s}^3), d(\boldsymbol{s}^2, \boldsymbol{s}^3)\}$. Then, the quantity becomes

$$\begin{aligned}
&\frac{m_n^2}{b_n^4} \sum_{\boldsymbol{s}^1 \neq \boldsymbol{s}^2 \neq \boldsymbol{s}^3 \in B(\boldsymbol{t}, b_n)} E|(\tilde{I}_{\boldsymbol{s}^1})^2 \tilde{I}_{\boldsymbol{s}^2} \tilde{I}_{\boldsymbol{s}^3}| \\
&= \frac{m_n^2}{b_n^4} \left[\sum_{j \leq m_n, \boldsymbol{s}^1 \neq \boldsymbol{s}^2 \neq \boldsymbol{s}^3 \in B(\boldsymbol{t}, b_n)} E|\cdot| + \sum_{j > m_n, \boldsymbol{s}^1 \neq \boldsymbol{s}^2 \neq \boldsymbol{s}^3 \in B(\boldsymbol{t}, b_n)} E|\cdot| \right] \\
&= O\left(\frac{m_n^5}{b_n^2}\right) + O\left(m_n^2 \sum_{j > m_n} \alpha_{c,2c}(j)\right).
\end{aligned}$$

To see the last equation, observe that $\boldsymbol{s}^1, \boldsymbol{s}^2$ and $\boldsymbol{s}^3$ are within m_n distance in the first term. Then fixing $\boldsymbol{s}^1$ (giving b_n^2), then $\boldsymbol{s}^2$ and $\boldsymbol{s}^3$ (giving m_n^4), using $m_n p_{\mathbf{0}} = O(1)$ produces the first term. For the second term, notice that $\#|B(\boldsymbol{t}, b_n) \cap \{\boldsymbol{s} : d(\boldsymbol{s}, \boldsymbol{t}) > m_n\}| = O(b_n^2)$ since $m_n/b_n \to 0$. Also note that

$$E|(\tilde{I}_{\boldsymbol{s}^1})^2 \tilde{I}_{\boldsymbol{s}^2} \tilde{I}_{\boldsymbol{s}^3}| \leq const\,(\alpha_{c,2c}(j) + p_{\mathbf{0}}\, \alpha_{c,c}(j)).$$

The same logic can be applied to derive $A_5 = O\left(m_n^7/b_n^2\right) + O\left(m_n^2 b_n^2 \sum_{j>m_n} \alpha_{c,3c}(j)\right)$, which converges to 0 under (3.4).

The next corollary generalizes Theorem 10 to dimension $d \geq 2$. We skip the proof as it is analogous.

Corollary 3.8.2. *Suppose a strictly stationary regularly varying random vector $\{Y_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{Z}^d\}$ with index $\alpha > 0$ is observed on $\Lambda_n = \{1, ..., n\}^d$. For any finite $H \in \mathbb{Z}^d$ which does not include $\mathbf{0}$ and sets A_1, A_2 bounded away from $\mathbf{0}$ with $\mu(A_2) > 0$, assume that*

$$\sqrt{\frac{b_n^d}{m_n}} \left[\hat{\rho}_{AB,m_n}(\boldsymbol{h}, E) - \rho_{AB,m_n}(\boldsymbol{h})\right]_{\boldsymbol{h} \in H} \xrightarrow{d} N(\mathbf{0}, \Sigma).$$

Further assume that $\frac{m_n b_n^{2d}}{n^d}, \frac{m_n^{2+3d-1}}{b_n^d} \to 0$, and $\lim_n m_n^2 b_n^d \sum_{\boldsymbol{l} \in \mathbb{Z}^d \setminus B_{m_n}} \alpha_{i,j}(d(\mathbf{0}, \boldsymbol{l})) = 0$ for $2c \leq i + j \leq 4c$, where $c = \#B_\gamma$, the ball of radius γ in $\mathbb{Z}^d$. Then

$$\sup_x \left| P^*\left(\sqrt{\frac{n^d}{m_n s_n^2}}(\hat{P}_m^* - \hat{P}_m) \leq x\right) - \Phi(x) \right| \xrightarrow{P} 0,$$

where $\hat{P}_m^* = \frac{m_n}{n^d} \sum_{\boldsymbol{s} \in \Lambda_n} I_{\boldsymbol{s}}^*$ and $s_n^2 = var^*\left(\sqrt{n^d/m_n}\hat{P}_m^*\right)$. Also,

$$P^*(|\hat{\rho}^*_{AB,m}(\boldsymbol{h}) - \hat{\rho}_{AB,m}(\boldsymbol{h})| > \delta) \xrightarrow{P} 0,\ \delta > 0,\ \textit{and}$$
$$P^*((n^d/m_n)^{1/2}(\hat{\rho}^*_{AB,m}(\boldsymbol{h}) - \hat{\rho}_{AB,m}(\boldsymbol{h}))_{\boldsymbol{h}\in H} \in C) \xrightarrow{P} \Phi_{(\mathbf{0},\Sigma)}(C)$$

hold, where C is any continuity set of the normal distribution $\Phi_{(\mathbf{0},\Sigma)}$.

3.8.2 Appendix B: Proof of Theorem 6

The proof of Theorem 6 uses the lemmas, under which the conditions of Theorem 1 of Politis and Romano (1993) hold. In particular, Theorem 1 of Politis and Romano (1993) assumes (3.10),(3.15),

$$E\left|T_{\boldsymbol{i}}(\boldsymbol{h})\right|^{2p+\delta} < \infty,\ \text{where } p \text{ is an integer such that } p > 2,\ \text{and } 0 < \delta \leq 2, \quad (3.52)$$
$$ET_{\boldsymbol{i}}(\boldsymbol{h}) = o(Q^{-1}), \textit{and} \quad (3.53)$$
$$\sqrt{Q^2}(\bar{T}(\boldsymbol{h}) - E\bar{T}(\boldsymbol{h})) \xrightarrow{d} N(0, \sigma^2_\infty),\ \text{where } \sigma^2_\infty = \lim_{n\to\infty} \text{var}(\sqrt{Q^2}\bar{T}(\boldsymbol{h})) > 0. \quad (3.54)$$

The following lemmas (Lemma 3.8.1 - 3.8.3) provide sufficient conditions for (3.52) - (3.54).

Lemma 3.8.1. *For a vectorized process $Y_{\boldsymbol{s}}$, let $\tilde{I}_{\boldsymbol{s}} = I(Y_{\boldsymbol{s}}/a_m \in C) - P(Y_{\boldsymbol{s}}/a_m \in C)$. If (3.11) holds, we have*

$$E\left|\sqrt{\frac{m_M}{M^2}}\sum_{\boldsymbol{s}\in E_{\boldsymbol{i}}}\tilde{I}_{\boldsymbol{s}}\right|^{12} < \infty \quad (3.55)$$

and it follows that

$$E\left|T_{\boldsymbol{i}}\right|^7 = E\left|\sqrt{\frac{M^2}{m_M}}\ \left(\hat{\rho}_{AB,m_M}(\boldsymbol{h}, E_{\boldsymbol{i}}) - \rho_{AB,m}(\boldsymbol{h})\right)\right|^7 < \infty. \quad (3.56)$$

Thus, (3.52) with $p = 3$ and $\delta = 1$ is satisfied.

Proof. To show (3.55), we adapt the technique from Lemma 18.5.2 in Ibragimov and Linnik

(1971), as in Remark 16. Observe that

$$
\begin{aligned}
&E\left(\sqrt{\frac{m_M}{M^2}}\sum_{\boldsymbol{s}\in E_{\boldsymbol{i}}}\tilde{I}_{\boldsymbol{s}}\right)^{12}\\
&=\frac{m_M^6}{M^{12}}\left[E\left(\sum_{\boldsymbol{s}\in E_{\boldsymbol{i}}}(\tilde{I}_{\boldsymbol{s}})^{12}+\sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\in E_{\boldsymbol{i}}}(\tilde{I}_{\boldsymbol{s}^1})^{11}\tilde{I}_{\boldsymbol{s}^2}+\sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\in E_{\boldsymbol{i}}}(\tilde{I}_{\boldsymbol{s}^1})^{10}(\tilde{I}_{\boldsymbol{s}^2})^2+\cdots\right.\right.\\
&+\sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\neq\boldsymbol{s}^3\in E_{\boldsymbol{i}}}(\tilde{I}_{\boldsymbol{s}^1})^{10}\tilde{I}_{\boldsymbol{s}^2}\tilde{I}_{\boldsymbol{s}^3}+\sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\neq\boldsymbol{s}^3\in E_{\boldsymbol{i}}}(\tilde{I}_{\boldsymbol{s}^1})^{9}(\tilde{I}_{\boldsymbol{s}^2})^2\tilde{I}_{\boldsymbol{s}^3}+\cdots\\
&+\sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\neq\boldsymbol{s}^3\neq\boldsymbol{s}^4\in E_{\boldsymbol{i}}}(\tilde{I}_{\boldsymbol{s}^1})^{9}\tilde{I}_{\boldsymbol{s}^2}\tilde{I}_{\boldsymbol{s}^3}\tilde{I}_{\boldsymbol{s}^4}+\cdots\\
&\left.\left.+\sum_{\boldsymbol{s}^i\neq\boldsymbol{s}^j\in E_{\boldsymbol{i}},1\leq i,j\leq 11}\left(\tilde{I}_{\boldsymbol{s}^1}\right)^2\prod_{j=2}^{11}\left(\tilde{I}_{\boldsymbol{s}^j}\right)+\sum_{\boldsymbol{s}^j\in E_{\boldsymbol{i}}}\prod_{j=1}^{12}\left(\tilde{I}_{\boldsymbol{s}^j}\right)\right)\right]
\end{aligned}
$$

From Remark 16, it suffices to estimate the last sum, but we show two other terms as illustrations.

For the first sum, note that $|\tilde{I}_s|\leq 2$ and $\#|E_{\mathbf{i}}|=M^2$. Thus, it is bounded by

$$2^{12}\left(m_M^6/M^{12}\right)M^2=O\left(m_M^6/M^{10}\right).$$

For the sums with four indices, let $j=\min_{1\leq i\neq j\leq 4}\{d(\boldsymbol{s}^i,\boldsymbol{s}^j)\}$. Then, for example,

$$
\begin{aligned}
&\frac{m_M^6}{M^{12}}\sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\neq\boldsymbol{s}^3\neq\boldsymbol{s}^4\in E_{\boldsymbol{i}}}E|(\tilde{I}_{\boldsymbol{s}^1})^9\tilde{I}_{\boldsymbol{s}^2}\tilde{I}_{\boldsymbol{s}^3}\tilde{I}_{\boldsymbol{s}^4}|\\
&=\frac{m_M^6}{M^{12}}\left[\sum_{j\leq m_M,\boldsymbol{s}^1\neq\boldsymbol{s}^2\neq\boldsymbol{s}^3\neq\boldsymbol{s}^4\in E_{\boldsymbol{i}}}E|\cdot|+\sum_{j>m_M,\boldsymbol{s}^1\neq\boldsymbol{s}^2\neq\boldsymbol{s}^3\neq\boldsymbol{s}^4\in E_{\boldsymbol{i}}}E|\cdot|\right]\\
&=O\left(\frac{m_M^{12}}{M^{10}}\right)+O\left(\frac{m_M^6}{M^6}\sum_{j>m_M}\alpha_{k_1c,k_2c}(j)\right).
\end{aligned}
$$

For the first term, notice that $\boldsymbol{s}^1-\boldsymbol{s}^4$ are within m_M distance. The second term is from

$$E|\tilde{I}_{\boldsymbol{s}^1}\tilde{I}_{\boldsymbol{s}^2}\tilde{I}_{\boldsymbol{s}^3}\tilde{I}_{\boldsymbol{s}^4}|=O\left(\alpha_{k_1c,k_2c}(j)\right)$$

for $1\leq k_1,k_2\leq 3$ with $k_1+k_2=4$ and $c=\#B_\gamma$.

The same logic can be applied to other terms, and the last sum becomes

$$O\left(\tfrac{m_M^{28}}{M^{10}}\right)+O\left(m_M^6M^{10}\textstyle\sum_{j>m_M}\alpha_{k_1c,k_2c}(j)\right),$$

for $1 \leq k_1, k_2 \leq 11$ with $k_1 + k_2 = 12$, which completes the proof for (3.55).

Turning to (3.56), recall $p_m(A) = m_M P(X_{\mathbf{0}} \in a_m A), \tau_{AB,m}(\boldsymbol{h}) = m_M P(X_{\mathbf{0}} \in a_m A, X_{\boldsymbol{h}} \in a_m B), \hat{p}_m(A) = m_M \sum_{\boldsymbol{s} \in E_{\boldsymbol{i}}} I_{\{a_m^{-1} X_{\boldsymbol{s}} \in A\}} / M^2$, and

$$\hat{\tau}_{AB,m}(\boldsymbol{h}) = m_M \sum_{\boldsymbol{s},\boldsymbol{t} \in E_{\boldsymbol{i}}, \boldsymbol{s}-\boldsymbol{t}=\boldsymbol{h}} I_{\{a_m^{-1} X_{\boldsymbol{s}} \in A, a_m^{-1} X_{\boldsymbol{t}} \in B\}} / M(\boldsymbol{h}),$$

where $M(\boldsymbol{h})$ denotes the number of $\boldsymbol{h}$ lag pairs in $E_{\boldsymbol{i}}$, thus $M(\boldsymbol{h}) = O(M^2)$. For the convenience, we compress $m, \boldsymbol{h}$ and A. Now, observe that

$$\begin{aligned} \hat{\rho}_{AB,m_M}(\boldsymbol{h}, E_{\boldsymbol{i}}) &= \frac{\sum_{\boldsymbol{s},\boldsymbol{t} \in E_{\boldsymbol{i}}, \boldsymbol{s}-\boldsymbol{t}=\boldsymbol{h}} I_{\{X_{\boldsymbol{s}}/a_m \in A, X_{\boldsymbol{t}}/a_m \in B\}} / M(\boldsymbol{h})}{\sum_{\boldsymbol{s} \in E_{\boldsymbol{i}}} I_{\{X_{\boldsymbol{s}}/a_m \in A\}} / M^2} \\ &\leq \frac{\sum_{\boldsymbol{s} \in E_{\boldsymbol{i}}} I_{\{X_{\boldsymbol{s}}/a_m \in A\}}\ p(\boldsymbol{h}) / M(\boldsymbol{h})}{\sum_{\boldsymbol{s} \in E_{\boldsymbol{i}}} I_{\{X_{\boldsymbol{s}}/a_m \in A\}} / M^2} \\ &= O(1), \end{aligned}$$

where $p(\boldsymbol{h})$ indicates the number of p distance pairs from the origin. Then

$$\begin{aligned} E\left|T_{\boldsymbol{i}}\right|^7 &= E\left| \sqrt{\frac{M^2}{m_M}} \left(\frac{\hat{\tau}}{\hat{p}} - \frac{\tau}{p} \right) \right|^7 \\ &= E\left| \sqrt{\frac{M^2}{m_M}} \left(\frac{\hat{\tau}(p-\hat{p})}{\hat{p}p} - \frac{(\tau - \hat{\tau})}{p} \right) \right|^7 \\ &= O\left(\sum_{0 \leq k \leq 7} \left| E Z_1^k Z_2^{7-k} \right| \right), \end{aligned}$$

where $Z_1 = \sqrt{\frac{M^2}{m_M}}\ \frac{\hat{\tau}(p-\hat{p})}{\hat{p}p}$ and $Z_2 = \sqrt{\frac{M^2}{m_M}}\ \frac{(\tau - \hat{\tau})}{p}$. Observe that each summand is bounded by (3.55),

$$E Z_1^7 = E\left(\sqrt{\tfrac{M^2}{m_M}}\ \left(\tfrac{\hat{\rho}_{AB}(p-\hat{p})}{p} \right) \right)^7 = O\left(E\left(\sqrt{\tfrac{M^2}{m_M}}\ \left(\tfrac{(p-\hat{p})}{p} \right) \right)^7 \right) < \infty.$$

Similarly, $E Z_2^7 < \infty$ can be verified by setting $Y_{\boldsymbol{s}} = \mathrm{vec}(X_{\mathbf{0}}, X_{\boldsymbol{h}})$. The rest of terms are bounded by Cauchy-Schwarz inequality and (3.55). □

Lemma 3.8.2. *For a strictly stationary regularly varying random fields $\{X_{\boldsymbol{s}}, \boldsymbol{s} \in \mathbb{Z}^2\}$, observed in $\Lambda_n = \{1, .., n\}^2$ in the lattice, assume (3.9) holds. For $\hat{\rho}_{AB,m}(\boldsymbol{h}) = \hat{\tau}_{AB,m}(\boldsymbol{h}) / \hat{p}(A)$,*

$$E(\hat{\rho}_{AB,m}(\boldsymbol{h})) - \rho_{AB,m}(\boldsymbol{h}) = O(m_n / M^2). \tag{3.57}$$

Furthermore, if $n^2 m_M/M^4 \to 0$, then (3.53) holds.

Proof. Recall that by definition in (3.1) both $\hat{\rho}_{AB,m}(\boldsymbol{h})$ and $\rho_{AB,m}(\boldsymbol{h})$ are zero when $\hat{p}(A) = 0$. Thus, we focus on the case when $\hat{p}(A) > 0$.

Consider $g(U,V) = V/U$ and $\mu = (\mu_U, \mu_V)$. By Taylor expansion,

$$\begin{aligned} g(U,V) &= g(\mu) + (U-\mu_U)\frac{\partial g}{\partial u}(\mu) + (V-\mu_V)\frac{\partial g}{\partial v}(\mu) \\ &\quad + \frac{1}{2}(U-\mu_U)^2\frac{\partial^2 g}{\partial u^2}(\mu) + (U-\mu_U)(V-\mu_V)\frac{\partial^2 g}{\partial u \partial v}(\mu) + R, \end{aligned} \tag{3.58}$$

where $R = \frac{1}{3!}\left(\frac{6}{(U^*)^3}(U-\mu_U)^2(V-\mu_V) - \frac{6V^*}{(U^*)^4}(U-\mu_U)^3\right)$. U^* and V^* are values between U and μ_U, and V and μ_V, respectively. Notice that $\frac{\partial^k g}{\partial v^k}(\mu) = 0$ and $\frac{\partial^{k+1} g}{\partial u \partial v^k}(\mu) = 0$ for $k > 1$.

Taking the expectation in (3.58) gives us

$$\begin{aligned} E(\hat{\rho}_{AB,m}(\boldsymbol{h})) &= \rho_{AB,m}(\boldsymbol{h}) + \left(\frac{\tau_{AB,m}(\boldsymbol{h})}{m_M^2\, p_{\mathbf{0}}(A)^3}\right) \mathrm{var}\,(U) \\ &\quad - \left(\frac{1}{m_M^2\, p_{\mathbf{0}}(A)^2}\right) \mathrm{cov}\,(U,V) + E(R), \end{aligned} \tag{3.59}$$

where

- $U = U_n := m_M \sum_{\boldsymbol{s} \in E_{\boldsymbol{i}}} I_{\{X_{\boldsymbol{s}}/a_m \in A\}}/M^2$,
- $V = V_n := m_M \sum_{\boldsymbol{s},\boldsymbol{t} \in E_{\boldsymbol{i}}, \boldsymbol{s}-\boldsymbol{t}=\boldsymbol{h}} I_{\{X_{\boldsymbol{s}}/a_m \in A, X_{\boldsymbol{t}}/a_m \in B\}}/M(\boldsymbol{h})$,
- $\mu_U := m_M p_{\mathbf{0}}(A) = m_M P(X_{\mathbf{0}}/a_m \in A)$, and
- $\mu_V := m_M \tau_{AB,m}(h) = m_M P(X_{\mathbf{0}}/a_m \in A, X_{\boldsymbol{h}}/a_m \in B)$.

To see (3.59), recall U is assumed to be positive. Thus, U^* is positive as well.

For the second term in (3.59), the variance calculation in Theorem 2.1 in Chapter 2 implies that

$$\left(\frac{m_M}{M^2}\right) \frac{\rho_{AB,m}(\boldsymbol{h})}{m_M^2\, p_{\mathbf{0}}(A)^2} \left[\sum_{\boldsymbol{l} \in \Lambda_n - \Lambda_n} m_M \tau_{AB,m}(\boldsymbol{l})\right] - \rho_{AB,m}(\boldsymbol{h}) = O\left(m_M/M^2\right) - \rho_{AB,m}(\boldsymbol{h}). \tag{3.60}$$

Similarly, the third term is equivalent to

$$
\begin{aligned}
&-\left(\frac{m_M}{M(\boldsymbol{h})}\right)\left[\frac{\sum_{\boldsymbol{l}\in\Lambda_n} m_M\, P(\frac{X_{\mathbf{0}}}{a_m}\in A, \frac{X_{\boldsymbol{l}}}{a_m}\in B, \frac{X_{\boldsymbol{l}+\boldsymbol{h}}}{a_m}\in A)}{m_M^2\, p_{\mathbf{0}}(A)^2}\right] + \rho_{AB,m}(\boldsymbol{h}) \\
&= O\left(m_M/M^2\right) + \rho_{AB,m}(\boldsymbol{h}). \qquad (3.61)
\end{aligned}
$$

since $M(\boldsymbol{h}) = O(M^2)$ and

$$
\sum_{\boldsymbol{l}\in\Lambda_n} m_M\, P\left(\frac{X_{\mathbf{0}}}{a_m}\in A, \frac{X_{\boldsymbol{l}}}{a_m}\in B, \frac{X_{\boldsymbol{l}+\boldsymbol{h}}}{a_m}\in A\right) = O\left(\sum_{\boldsymbol{l}\in\Lambda_n} m_M\, \tau_{AB,m}(\boldsymbol{l})\right).
$$

Lastly, $R = O_p\left(\left(\frac{m_M}{M^2}\right)^{3/2}\right)$. To see this, note that $\mu_U = O(1)$ and $\mu_V = O(1)$ by the regularly varying assumption. Also

$$
U \xrightarrow{p} \lim_{n\to\infty}\mu_U \text{ and } \sqrt{\tfrac{M^2}{m_M}}\,(U-\mu_U) = O_p(1)
$$

by Theorem 2.1, Chapter 2, and Prohorov's theorem. Similarly

$$
V \xrightarrow{p} \lim_{n\to\infty}\mu_V \text{ and } \sqrt{\tfrac{M^2}{m_M}}\,(V-\mu_V) = O_p(1).
$$

Thus, it follows that $U^* \xrightarrow{p} \lim_{n\to\infty}\mu_U$ and $V^* \xrightarrow{p} \lim_{n\to\infty}\mu_V$, and once can conclude that $R = O_p\left(\left(\frac{m_M}{M^2}\right)^{3/2}\right)$. Combined with (3.60) and (3.61), (3.59) gives us

$$
\begin{aligned}
\left(\frac{M^2}{m_M}\right)\left(E(\hat{\rho}_{AB,m}(\boldsymbol{h})) - \rho_{AB,m}(\boldsymbol{h})\right) &= \left(\frac{M^2}{m_M}\right)\left[O\left(m_M/M^2\right) + E(R)\right] \\
&= O(1) + \sqrt{\frac{m_M}{M^2}}E(O_p(1)) = O(1),
\end{aligned}
$$

which verifies (3.57). Hence, (3.53) holds if $n^2 m_M/M^4 \to 0$ since

$$
Q\left[ET_{\boldsymbol{i}}(\boldsymbol{h})\right] = QO\left(\sqrt{\frac{m_M}{M^2}}\right) = O\left(\sqrt{\frac{n^2 m_M}{M^4}}\right).
$$

□

Lemma 3.8.3. *Assume the conditions of Lemma 3.8.1 hold. Assume that there exists an increasing sequence n, M, and m_M such that $M = o(n)$ and $m_M = o(M)$ and that satisfy (3.12) and (3.13). Then, (3.54) holds, i.e.,*

$$
\sqrt{Q^2}(\bar{T}(\boldsymbol{h}) - E\bar{T}(\boldsymbol{h})) \xrightarrow{d} N(0, \sigma_\infty^2),
$$

where $\sigma_\infty^2 = \lim_{n\to\infty} var(\sqrt{Q^2}\bar{T}(\boldsymbol{h})) > 0$.

Proof. For the second moment, we will show

$$\mathrm{var}\left(\sqrt{Q^2}(\bar{T}(\boldsymbol{h}) - E\bar{T}(\boldsymbol{h}))\right) \to \sigma_\infty^2 = \sigma^2 + \sum_{h<B_C} \mathrm{cov}\left(T_{\mathbf{0}}, T_{\boldsymbol{h}}\right) < \infty, \tag{3.62}$$

where B_C is the ball of radius C. Here, $\lim_{n\to\infty}(M/L+1) = C$.

To see (3.62), observe that the left-hand side of it is

$$\begin{aligned}
& \frac{1}{Q^2}\sum_{\boldsymbol{i}\in\Lambda_Q} \mathrm{var}(T_{\boldsymbol{i}}) + \frac{1}{Q^2}\sum_{\boldsymbol{i},\boldsymbol{j}\in\Lambda_Q} \mathrm{cov}(T_{\boldsymbol{i}}, T_{\boldsymbol{j}}) \\
= \;& \sigma^2 + \sum_{\boldsymbol{h}\in\Lambda_Q} \mathrm{cov}(T_{\mathbf{0}}, T_{\boldsymbol{h}}) \\
= \;& \sigma^2 + \sum_{d(\mathbf{0},\boldsymbol{h})<\frac{M}{L}+1, \boldsymbol{h}\in\Lambda_Q} \mathrm{cov}(T_{\mathbf{0}}, T_{\boldsymbol{h}}) + \sum_{d(\mathbf{0},\boldsymbol{h})\geq\frac{M}{L}+1, \boldsymbol{h}\in\Lambda_Q} \mathrm{cov}(T_{\mathbf{0}}, T_{\boldsymbol{h}}) \\
\leq \;& \sigma^2 + \mathrm{const}\,\left(\frac{M}{L}+1\right)^2 + \mathrm{const}\,\frac{M^2}{m_M} \sum_{\boldsymbol{h}\in\mathbb{Z}^2, d(\mathbf{0},\boldsymbol{h})\geq L} \alpha_{(Mb)^2,(Mb)^2}(d(\mathbf{0},\boldsymbol{h}))^{\frac{p-2}{p}},
\end{aligned}$$

for $p > 2$. To see the inequality, for the first term in the inequality, note from (3.56) that $|\mathrm{cov}(T_{\mathbf{0}}, T_{\boldsymbol{h}})| < \infty$. For the second term, Theorem 3 in Doukhan (1994) gives

$$|\mathrm{cov}(T_{\mathbf{0}}, T_{\boldsymbol{h}})| \leq \mathrm{const}\,\left[\alpha^T(\|\boldsymbol{h}\|)\right]^{(p-2)/p} \|T_{\mathbf{0}}\|_p \|T_{\boldsymbol{h}}\|_p \quad \text{for} \quad p > 2, \tag{3.63}$$

where α^T be α-mixing coefficient is associated with $T_{\boldsymbol{i}}$. From (31) and Lemma 1 in Politis and Romano (1993),

$$\alpha^T(s) \leq \alpha_{(Mb)^2,(Mb)^2}(sL - M) \quad \text{for} \quad s \geq M/L + 1 \tag{3.64}$$

and $\|T_{\mathbf{0}}\|_p \leq \sqrt{M^2/m_M}$. Given (3.12), the dominated convergence theorem implies (3.62). The central limit theorem follows from Theorem in Bolthausen (1982) as (3.13) is assumed.

□

Proof of Theorem 6. Theorem 1 in Politis and Romano (1993) is directly applicable since all conditions required are satisfied by Lemma 3.8.1 - 3.8.3.

Remark 17. *Theorem 6 can be extended to non-rectangular cases if a growth rate of blocks has the same order for all directions, as mentioned in conditions (i) and (ii) of Theorem 1 in Politis and Romano (1993).*

3.8.3 Appendix C: Proof of Theorem 7.

For Theorem 7, we introduce lemma to provide sufficient conditions for

$$E[|g(\boldsymbol{h}, B_n)|^{4+\delta}] < \infty, \tag{3.65}$$

for some $\delta > 0$. The condition (3.65), together with (3.21), (3.22), and (3.23), implies that $\text{var}(g(\boldsymbol{h}, B_n)) \to \sigma^2$, which is assumed by Theorem 2 in Politis and Sherman (2001). Similarly, L_2 consistency for the covariance estimator requires

$$E[g(\boldsymbol{h}^1, B_n)^{2+\delta} g(\boldsymbol{h}^2, B_n)^{2+\delta}] < \infty \quad \text{for} \quad \boldsymbol{h}^1 \neq \boldsymbol{h}^2. \tag{3.66}$$

Notice that (3.66) is implied by (3.65).

Lemma 3.8.4. *Assume* $\lambda_n^8 n^6 m_n^4 = O(1), m_n = O(n^{1/3})$ *and* $\int_{\mathbb{R}^2} \alpha_{1,k}(\boldsymbol{y}) d\boldsymbol{y} < \infty$ *for* $k = 1, ..., 5$. *Then*

$$\sup_n E \left| \sqrt{\frac{n^2\lambda_n^2}{m_n}} \ (\hat{p}_m(A) - p_m(A)) \right|^8 < \infty \quad \textit{and} \quad \sup_n E \left| \sqrt{\frac{n^2\lambda_n^2}{m_n}} \ \hat{\tau}_{AB,m}(\boldsymbol{h}, \Lambda_n) \right|^8 < \infty. \tag{3.67}$$

Proof. Write $p_m = p_m(A)$ for our convenience. Note that $E \left| \sqrt{n^2\lambda_n^2/m_n} \ (\hat{p}_m - p_m) \right|^8$ is equivalent to

$$\begin{aligned} \left(\frac{n^8\lambda_n^8}{m_n^4} \right) O\left(\sum_{i=0}^{8} E(\hat{p}_m)^i p_m^{8-i} \right) &= \left(\frac{n^8\lambda_n^8}{m_n^4} \right) O\left(\sum_{i=1}^{8} C_i \ \left[E(\hat{p}_m)^i p_m^{8-i} - E(\hat{p}_m)^{i-1} p_m^{9-i} \right] \right) \\ &= O\left(\sum_{i=1}^{8} C_i \ A_i \right) \end{aligned} \tag{3.68}$$

for some constant C_i. To show the finiteness, the similar techniques presented in Lemma 16 following the argument in Lemma 18.5.2 in Ibragimov and Linnik (1971) are used. In the proof, we focus on $A_i, i = 1, ..., 8$.

First, observe that $A_1 = 0$ since, from $E(\hat{p}_m) = p_m = O(1)$,

$$E(\hat{p}_m) p_m^7 - p_m^8 = p_m^7 (E(\hat{p}_m) - p_m) = 0.$$

One can derive $A_2 = O\left(n^6 \lambda_n^8 / m_n^3\right) = o(1)$ since (5.13) in Chapter 2 implies

$$E(\hat{p}_m)^2 - E(\hat{p}_m)\, p_m = \text{var}\,(\hat{p}_m) = O\left(m_n/n^2\right).$$

For A_3, we have

$$A_3 = O\left(n^4\lambda_n^8/m_n^2\right) + O\left(n^4\lambda_n^8/m_n \int_{\mathbb{R}^2} \alpha_{1,1}(\boldsymbol{y})d\boldsymbol{y}\right) + O\left(n^6\lambda_n^8/m_n \int_{\mathbb{R}^2} \alpha_{1,2}(\boldsymbol{y})d\boldsymbol{y}\right). \tag{3.69}$$

To see the above, notice that

$$\begin{aligned}
& E(\hat{p}_m)^3 \\
= \; & \left(\frac{m_n}{\nu n^2}\right)^3 E\left[\int_{\Lambda_n} I\left(\frac{X_{\boldsymbol{s}^1}}{a_m} \in A\right) N(d\boldsymbol{s}^1)\right. \\
& \quad + \int_{\Lambda_n}\int_{\Lambda_n} I\left(\frac{X_{\boldsymbol{s}^1}}{a_m} \in A\right) I\left(\frac{X_{\boldsymbol{s}^2}}{a_m} \in A\right) N(d\boldsymbol{s}^1)N(d\boldsymbol{s}^2) \\
& \quad \left. + \int_{\Lambda_n}\int_{\Lambda_n}\int_{\Lambda_n} I\left(\frac{X_{\boldsymbol{s}^1}}{a_m} \in A\right) I\left(\frac{X_{\boldsymbol{s}^2}}{a_m} \in A\right) I\left(\frac{X_{\boldsymbol{s}^3}}{a_m} \in A\right) N(d\boldsymbol{s}^1)N(d\boldsymbol{s}^2)N(d\boldsymbol{s}^3)\right] \\
= \; & \left(\frac{m_n}{\nu n^2}\right)^3 \left[\left(\frac{p_m}{m_n}\right)\nu n^2 + \int_{\Lambda_n}\int_{\Lambda_n} P\left(\frac{X_{\boldsymbol{s}^1}}{a_m} \in A, \frac{X_{\boldsymbol{s}^2}}{a_m} \in A\right)\nu^2 d\boldsymbol{s}^1 d\boldsymbol{s}^2\right. \\
& \quad \left. + \int_{\Lambda_n}\int_{\Lambda_n}\int_{\Lambda_n} P\left(\frac{X_{\boldsymbol{s}^1}}{a_m} \in A, \frac{X_{\boldsymbol{s}^2}}{a_m} \in A, \frac{X_{\boldsymbol{s}^3}}{a_m} \in A\right)\nu^3 d\boldsymbol{s}^1 d\boldsymbol{s}^2 d\boldsymbol{s}^3\right], \quad \text{and} \\
& E(\hat{p}_m)^2\, p_m = \left(\frac{m_n}{\nu n^2}\right)^2 \left[\left(\frac{p_m}{m_n}\right)\nu n^2 + \int_{\Lambda_n}\int_{\Lambda_n} P\left(\frac{X_{\boldsymbol{s}^1}}{a_m} \in A, \frac{X_{\boldsymbol{s}^2}}{a_m} \in A\right)\nu^2 d\boldsymbol{s}^1 d\boldsymbol{s}^2\right]\, p_m.
\end{aligned}$$

Thus, $E(\hat{p}_m)^3 - E(\hat{p}_m)^2\, p_m$ becomes

$$\begin{aligned}
& \frac{m_n^2}{\nu^2 n^4} p_m + \frac{m_n^3}{\nu n^6} \int_{\Lambda_n}\int_{\Lambda_n}\left[P\left(\frac{X_{\boldsymbol{s}^1}}{a_m} \in A, \frac{X_{\boldsymbol{s}^2}}{a_m} \in A\right) - \left(\frac{p_m}{m_n}\right)^2\right] d\boldsymbol{s}^1 d\boldsymbol{s}^2 \\
& + \frac{m_n^3}{n^6} \int_{\Lambda_n}\int_{\Lambda_n}\int_{\Lambda_n}\left[P\left(\frac{X_{\boldsymbol{s}^1}}{a_m} \in A, \frac{X_{\boldsymbol{s}^2}}{a_m} \in A, \frac{X_{\boldsymbol{s}^3}}{a_m} \in A\right)\right. \\
& \qquad \left. - P\left(\frac{X_{\boldsymbol{s}^1}}{a_m} \in A, \frac{X_{\boldsymbol{s}^2}}{a_m} \in A\right)\left(\frac{p_m}{m_n}\right)\right] d\boldsymbol{s}^1 d\boldsymbol{s}^2 d\boldsymbol{s}^3,
\end{aligned}$$

where the second and the third term are bounded as below.

$$\begin{aligned}
& \int_{\Lambda_n}\int_{\Lambda_n}\left|P\left(\frac{X_{\boldsymbol{s}^1}}{a_m} \in A, \frac{X_{\boldsymbol{s}^2}}{a_m} \in A\right) - \left(\frac{p_m}{m_n}\right)^2\right| d\boldsymbol{s}^1 d\boldsymbol{s}^2 \le n^2 \int_{\mathbb{R}^2} \alpha_{1,1}(\boldsymbol{y})d\boldsymbol{y}, \quad \text{and} \\
& \int_{\Lambda_n}\int_{\Lambda_n}\int_{\Lambda_n}\left|P\left(\frac{X_{\boldsymbol{s}^1}}{a_m} \in A, \frac{X_{\boldsymbol{s}^2}}{a_m} \in A, \frac{X_{\boldsymbol{s}^3}}{a_m} \in A\right)\right. \\
& \qquad \left. - P\left(\frac{X_{\boldsymbol{s}^1}}{a_m} \in A, \frac{X_{\boldsymbol{s}^2}}{a_m} \in A\right)\left(\frac{p_m}{m_n}\right)\right| d\boldsymbol{s}^1 d\boldsymbol{s}^2 d\boldsymbol{s}^3 \le n^4 \int_{\mathbb{R}^2} \alpha_{1,2}(\boldsymbol{y})d\boldsymbol{y}.
\end{aligned}$$

Hence, we have (3.69). The similar techniques are used for $A_4 - A_8$.

$$
\begin{aligned}
A_4 &= O\left(\frac{\lambda_n^6 n^2}{m_n}\right) + O\left(\lambda_n^6 n^2 \int_{\mathbb{R}^2} \alpha_{1,1}(\boldsymbol{y})d\boldsymbol{y}\right) \\
&\quad + O\left(\lambda_n^8 n^4 \int_{\mathbb{R}^2} \alpha_{1,2}(\boldsymbol{y})d\boldsymbol{y}\right) + O\left(\lambda_n^8 n^6 \int_{\mathbb{R}^2} \alpha_{1,3}(\boldsymbol{y})d\boldsymbol{y}\right), \\
A_5 &= O\left(\lambda_n^8\right) + O\left(\lambda_n^8 m_n \int_{\mathbb{R}^2} \alpha_{1,1}(\boldsymbol{y})d\boldsymbol{y}\right) + O\left(\lambda_n^8 m_n n^2 \int_{\mathbb{R}^2} \alpha_{1,2}(\boldsymbol{y})d\boldsymbol{y}\right) \\
&\quad + O\left(\lambda_n^8 m_n n^4 \int_{\mathbb{R}^2} \alpha_{1,3}(\boldsymbol{y})d\boldsymbol{y}\right) + O\left(\lambda_n^8 m_n n^6 \int_{\mathbb{R}^2} \alpha_{1,4}(\boldsymbol{y})d\boldsymbol{y}\right), \\
A_6 &= O\left(\frac{\lambda_n^8 m_n}{n^2}\right) + O\left(\frac{\lambda_n^8 m_n^2}{n^2} \int_{\mathbb{R}^2} \alpha_{1,1}(\boldsymbol{y})d\boldsymbol{y}\right) + O\left(\lambda_n^8 m_n^2 \int_{\mathbb{R}^2} \alpha_{1,2}(\boldsymbol{y})d\boldsymbol{y}\right) \\
&\quad + O\left(\lambda_n^8 m_n^2 n^2 \int_{\mathbb{R}^2} \alpha_{1,3}(\boldsymbol{y})d\boldsymbol{y}\right) + O\left(\lambda_n^8 m_n^2 n^4 \int_{\mathbb{R}^2} \alpha_{1,4}(\boldsymbol{y})d\boldsymbol{y}\right) + O\left(\lambda_n^8 m_n^2 n^6 \int_{\mathbb{R}^2} \alpha_{1,5}(\boldsymbol{y})d\boldsymbol{y}\right), \\
A_7 &= O\left(\frac{\lambda_n^8 m_n^2}{n^4}\right) + O\left(\frac{\lambda_n^8 m_n^3}{n^4} \int_{\mathbb{R}^2} \alpha_{1,1}(\boldsymbol{y})d\boldsymbol{y}\right) + O\left(\frac{\lambda_n^8 m_n^3}{n^2} \int_{\mathbb{R}^2} \alpha_{1,2}(\boldsymbol{y})d\boldsymbol{y}\right) \\
&\quad + O\left(\lambda_n^8 m_n^3 \int_{\mathbb{R}^2} \alpha_{1,3}(\boldsymbol{y})d\boldsymbol{y}\right) + O\left(\lambda_n^8 m_n^3 n^2 \int_{\mathbb{R}^2} \alpha_{1,4}(\boldsymbol{y})d\boldsymbol{y}\right) \\
&\quad + O\left(\lambda_n^8 m_n^3 n^4 \int_{\mathbb{R}^2} \alpha_{1,5}(\boldsymbol{y})d\boldsymbol{y}\right) + O\left(\lambda_n^8 m_n^3 n^6 \int_{\mathbb{R}^2} \alpha_{1,6}(\boldsymbol{y})d\boldsymbol{y}\right), \quad \text{and} \\
A_8 &= O\left(\frac{\lambda_n^8 m_n^3}{n^6}\right) + O\left(\frac{\lambda_n^8 m_n^4}{n^6} \int_{\mathbb{R}^2} \alpha_{1,1}(\boldsymbol{y})d\boldsymbol{y}\right) + O\left(\frac{\lambda_n^8 m_n^4}{n^4} \int_{\mathbb{R}^2} \alpha_{1,2}(\boldsymbol{y})d\boldsymbol{y}\right) \\
&\quad + O\left(\frac{\lambda_n^8 m_n^4}{n^2} \int_{\mathbb{R}^2} \alpha_{1,3}(\boldsymbol{y})d\boldsymbol{y}\right) + O\left(\lambda_n^8 m_n^4 \int_{\mathbb{R}^2} \alpha_{1,4}(\boldsymbol{y})d\boldsymbol{y}\right) + O\left(\lambda_n^8 m_n^4 n^2 \int_{\mathbb{R}^2} \alpha_{1,5}(\boldsymbol{y})d\boldsymbol{y}\right) \\
&\quad + O\left(\lambda_n^8 m_n^4 n^4 \int_{\mathbb{R}^2} \alpha_{1,6}(\boldsymbol{y})d\boldsymbol{y}\right) + O\left(\lambda_n^8 m_n^4 n^6 \int_{\mathbb{R}^2} \alpha_{1,7}(\boldsymbol{y})d\boldsymbol{y}\right).
\end{aligned}
$$

These can be derived from

$$
\begin{aligned}
& E(\hat{p}_m)^4 - E(\hat{p}_m)^3\, p_m \\
= \;& \frac{m_n^3}{\nu^3 n^6} p_m + \frac{m_n^4}{\nu^2 n^8} \int\int_{\Lambda_n} \left[P_{\boldsymbol{s}^1,\boldsymbol{s}^2} - \left(\frac{p_m}{m_n}\right)^2 \right] d\boldsymbol{s}^1 d\boldsymbol{s}^2 \\
& + \frac{m_n^4}{\nu n^8} \int\int\int_{\Lambda_n} \left[P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3} - P_{\boldsymbol{s}^1,\boldsymbol{s}^2} \left(\frac{p_m}{m_n}\right) \right] d\boldsymbol{s}^1 d\boldsymbol{s}^2 d\boldsymbol{s}^3 \\
& + \frac{m_n^4}{n^8} \int\int\int\int_{\Lambda_n} \left[P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3,\boldsymbol{s}^4} - P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3} \left(\frac{p_m}{m_n}\right) \right] d\boldsymbol{s}^1 d\boldsymbol{s}^2 d\boldsymbol{s}^3 d\boldsymbol{s}^4, \\
& E(\hat{p}_m)^5 - E(\hat{p}_m)^4\, p_m \\
= \;& \frac{m_n^4}{\nu^4 n^8} p_m + \frac{m_n^5}{\nu^3 n^{10}} \int\int_{\Lambda_n} \left[P_{\boldsymbol{s}^1,\boldsymbol{s}^2} - \left(\frac{p_m}{m_n}\right)^2 \right] d\boldsymbol{s}^1 d\boldsymbol{s}^2 \\
& + \frac{m_n^5}{\nu^2 n^{10}} \int\int\int_{\Lambda_n} \left[P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3} - P_{\boldsymbol{s}^1,\boldsymbol{s}^2} \left(\frac{p_m}{m_n}\right) \right] d\boldsymbol{s}^1 d\boldsymbol{s}^2 d\boldsymbol{s}^3 \\
& + \frac{m_n^5}{\nu n^{10}} \int\int\int\int_{\Lambda_n} \left[P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3,\boldsymbol{s}^4} - P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3} \left(\frac{p_m}{m_n}\right) \right] d\boldsymbol{s}^1 d\boldsymbol{s}^2 d\boldsymbol{s}^3 d\boldsymbol{s}^4 \\
& + \frac{m_n^5}{n^{10}} \int\int\int\int\int_{\Lambda_n} \left[P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3,\boldsymbol{s}^4,\boldsymbol{s}^5} - P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3,\boldsymbol{s}^4} \left(\frac{p_m}{m_n}\right) \right] d\boldsymbol{s}^1 d\boldsymbol{s}^2 d\boldsymbol{s}^3 d\boldsymbol{s}^4 d\boldsymbol{s}^5, \text{ and}
\end{aligned}
$$

$$
\begin{aligned}
& E(\hat{p}_m)^6 - E(\hat{p}_m)^5\, p_m \\
= \;& \frac{m_n^5}{\nu^5 n^{10}} p_m + \frac{m_n^6}{\nu^4 n^{12}} \int\int_{\Lambda_n} \left[P_{\boldsymbol{s}^1,\boldsymbol{s}^2} - \left(\frac{p_m}{m_n}\right)^2 \right] d\boldsymbol{s}^1 d\boldsymbol{s}^2 \\
& + \frac{m_n^6}{\nu^3 n^{12}} \int\int\int_{\Lambda_n} \left[P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3} - P_{\boldsymbol{s}^1,\boldsymbol{s}^2} \left(\frac{p_m}{m_n}\right) \right] d\boldsymbol{s}^1 d\boldsymbol{s}^2 d\boldsymbol{s}^3 \\
& + \frac{m_n^6}{\nu^2 n^{12}} \int\int\int\int_{\Lambda_n} \left[P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3,\boldsymbol{s}^4} - P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3} \left(\frac{p_m}{m_n}\right) \right] d\boldsymbol{s}^1 d\boldsymbol{s}^2 d\boldsymbol{s}^3 d\boldsymbol{s}^4 \\
& + \frac{m_n^6}{\nu n^{12}} \int\int\int\int\int_{\Lambda_n} \left[P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3,\boldsymbol{s}^4,\boldsymbol{s}^5} - P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3,\boldsymbol{s}^4} \left(\frac{p_m}{m_n}\right) \right] d\boldsymbol{s}^1 d\boldsymbol{s}^2 d\boldsymbol{s}^3 d\boldsymbol{s}^4 d\boldsymbol{s}^5 \\
& + \frac{m_n^6}{n^{12}} \int\int\int\int\int\int_{\Lambda_n} \left[P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3,\boldsymbol{s}^4,\boldsymbol{s}^5,\boldsymbol{s}^6} - P_{\boldsymbol{s}^1,\boldsymbol{s}^2,\boldsymbol{s}^3,\boldsymbol{s}^4,\boldsymbol{s}^5} \left(\frac{p_m}{m_n}\right) \right] d\boldsymbol{s}^1 d\boldsymbol{s}^2 d\boldsymbol{s}^3 d\boldsymbol{s}^4 d\boldsymbol{s}^d \boldsymbol{s}^6,
\end{aligned}
$$

where $P_{\boldsymbol{s}^1,\ldots,\boldsymbol{s}^k} = P\left(X_{\boldsymbol{s}^1}/a_m \in A, \cdots, X_{\boldsymbol{s}^k}/a_m \in A\right)$. Hence, (3.68) is shown.

Notice that $E\left|\sqrt{\frac{n^2\lambda_n^2}{m_n}}\,\hat{p}_m(A)\right|^8 < \infty$ implies $E\left|\sqrt{\frac{n^2\lambda_n^2}{m_n}}\,\hat{\tau}_{AB,m}(\boldsymbol{h},\Lambda_n)\right|^8 < \infty$. This is

because

$$
\begin{aligned}
& |\hat{\tau}_{AB,m}(\boldsymbol{h},\Lambda_n)| \\
\leq\ & \frac{m_n}{\nu^2}\frac{1}{|\Lambda_n|}\left|\int_{\Lambda_n}\int_{\Lambda_n} w_n(\boldsymbol{h}+\boldsymbol{s}^1-\boldsymbol{s}^2)\, I\left(\frac{X_{\boldsymbol{s}^1}}{a_m}\in A\right) I\left(\frac{X_{\boldsymbol{s}^2}}{a_m}\in B\right) N^{(2)}(d\boldsymbol{s}^1,d\boldsymbol{s}^2)\right| \\
=\ & \frac{m_n}{\nu^2}\frac{1}{|\Lambda_n|}\left|\int_{\Lambda_n}\int_{\frac{\boldsymbol{h}+\Lambda_n-\Lambda_n}{\lambda_n}} w(\boldsymbol{t})\, I\left(\frac{X_{\boldsymbol{u}}}{a_m}\in A\right) I\left(\frac{X_{\boldsymbol{h}+\boldsymbol{u}-\boldsymbol{t}\lambda_n}}{a_m}\in B\right) N(d\boldsymbol{u})N(d\boldsymbol{t})\right| \\
\leq\ & \frac{m_n}{\nu^2}\frac{1}{|\Lambda_n|}\left|\int_{\Lambda_n} I\left(\frac{X_{\boldsymbol{u}}}{a_m}\in A\right) N(d\boldsymbol{u})\int_{\mathbb{R}^2} w(\boldsymbol{t})N(d\boldsymbol{t})\right| \\
=\ & O\left(\hat{p}_m(A)\right),
\end{aligned}
$$

where the change of variable of $\frac{\boldsymbol{h}+\boldsymbol{s}^1-\boldsymbol{s}^2}{\lambda_n} = \boldsymbol{t}$ and $\boldsymbol{s}^1 = \boldsymbol{u}$ are used in the first equality. Hence, (3.67) is shown, which completes the proof. □

Remark 18. *Notice that it is sufficient to check the second condition in (3.67) to show*

$$\sup_n E\left|\sqrt{\frac{n^2\lambda_n^2}{m_n}}\ \left(\hat{\tau}_{AB,m}(\boldsymbol{h},\Lambda_n)-\tau_{AB,m}(\boldsymbol{h},\Lambda_n)\right)\right|^8 < \infty$$

because $|\tau_{AB,m}(\boldsymbol{h},\Lambda_n)| \leq 1$ *and* $\sqrt{\frac{n^2\lambda_n^2}{m_n}} = O(1)$. *The latter is implied by* $\lambda_n^2 n^2 \to \infty, \lambda_n^8 n^6 m_n^4 = O(1)$ *and* $m_n = O(n^{1/3})$. *The same logic applies to the first condition in (3.67) and*

$$\sup_n E\left|\sqrt{\frac{n^2\lambda_n^2}{m_n}}\ \hat{p}_m(A)\right|^8 < \infty.$$

This explains the last argument in the proof of Lemma 3.8.4.

Proof of Theorem 7. From Lemma 3.8.4, (3.65) and (3.66) with $\delta = 1$ can be shown using the argument in Remark 16 and the Cauchy-Schwartz inequality. Thus, Theorem 2 in Politis and Sherman (2001) is directly applicable since all conditions required are satisfied.

Proof of Theorem 8.

Proof. The proof is analogous to that of Theorem 7 and the techniques in Remark 16.

Recall from Theorem 8 that n^2/m_n is the normalizing term for $\hat{\rho}_{AB,m}(\boldsymbol{h},\Lambda_n)$. Thus, the lattice analog of (3.67) becomes

$$\sup_n E\left|\sqrt{\frac{n^2}{m_n}}\ \left(\hat{p}_m(A)-p_m(A)\right)\right|^8 < \infty \quad \text{and} \quad \sup_n E\left|\sqrt{\frac{n^2}{m_n}}\ \hat{\tau}_{AB,m}(\boldsymbol{h},\Lambda_n)\right|^8 < \infty. \quad (3.70)$$

Now we apply the techniques presented in Remark 16, following Ibragimov and Linnik (1971). Let $\tilde{I}_{\boldsymbol{s}} = I(Y_{\boldsymbol{s}}/a_m \in A) - P(Y_{\boldsymbol{s}}/a_m \in A)$. Then the first term in (3.70) is equivalent to

$$\begin{aligned}
& E\left|\sqrt{\frac{m_n}{n^2}} \sum_{\boldsymbol{s}\in B(\boldsymbol{t},n)} \tilde{I}_{\boldsymbol{s}}\right|^8 \\
= \; & \frac{m_n^4}{n^8}\, E\left(\sum_{\boldsymbol{s}\in B(\boldsymbol{t},n)} (\tilde{I}_{\boldsymbol{s}})^8 + \sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\in B(\boldsymbol{t},n)} (\tilde{I}_{\boldsymbol{s}^1})^7 \tilde{I}_{\boldsymbol{s}^2} + \sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\in B(\boldsymbol{t},n)} (\tilde{I}_{\boldsymbol{s}^1})^6 (\tilde{I}_{\boldsymbol{s}^2})^2 \right. \\
& \left. + \cdots + \sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\neq\cdots\neq\boldsymbol{s}^7\neq\boldsymbol{s}^8\in B(\boldsymbol{t},n)} \tilde{I}_{\boldsymbol{s}^1}\cdots\tilde{I}_{\boldsymbol{s}^8}\right) \\
= \; & A_1 + A_2 + A_3 + A_4 + A_5 + A_6 + A_7 + A_8.
\end{aligned}$$

As discussed earlier, it suffices to check the last term, but we will select four terms and verify necessary conditions.

For A_1, note that $|\tilde{I}_{\boldsymbol{s}}| \le 2$ and $\#|B(\boldsymbol{t},n)| = n^2$. Thus, $A_1 = O\left(m_n^4 n^2/n^8\right) = O\left(m_n^4/n^6\right)$.

For the sums with two indices, like A_2, $\frac{m_n^4}{n^8}\, E\left(\sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\in B(\boldsymbol{t},n)} (\tilde{I}_{\boldsymbol{s}^1})^3 \tilde{I}_{\boldsymbol{s}^2}\right)$ equals

$$O\left(\frac{m_n^4}{n^8} \sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\in B(\boldsymbol{t},n)} E\left|(\tilde{I}_{\boldsymbol{s}^1})^7 \tilde{I}_{\boldsymbol{s}^2}\right|\right) = O\left(\frac{m_n^4}{n^6} \sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\in B(\boldsymbol{t},n)} \alpha_{c,c}(|\boldsymbol{s}^2 - \boldsymbol{s}^1|)\right) = O\left(\frac{m_n^4}{n^6}\right).$$

The same argument holds for A_3.

For A_4, let $j = \min\{d(\boldsymbol{s}^1,\boldsymbol{s}^2), d(\boldsymbol{s}^1,\boldsymbol{s}^3), d(\boldsymbol{s}^2,\boldsymbol{s}^3)\}$. Then, the term becomes

$$\begin{aligned}
& \frac{m_n^4}{n^8} \sum_{\boldsymbol{s}^1\neq\boldsymbol{s}^2\neq\boldsymbol{s}^3\in B(\boldsymbol{t},n)} E|(\tilde{I}_{\boldsymbol{s}^1})^2 \tilde{I}_{\boldsymbol{s}^2} \tilde{I}_{\boldsymbol{s}^3}| \\
= \; & \frac{m_n^4}{n^8}\left[\sum_{j\le m_n, \boldsymbol{s}^1\neq\boldsymbol{s}^2\neq\boldsymbol{s}^3\in B(\boldsymbol{t},n)} E|\cdot| + \sum_{j> m_n, \boldsymbol{s}^1\neq\boldsymbol{s}^2\neq\boldsymbol{s}^3\in B(\boldsymbol{t},n)} E|\cdot|\right] \\
= \; & O\left(\frac{m_n^7}{n^6}\right) + O\left(\frac{m_n^4}{n^4} \sum_{j> m_n} \alpha_{c,2c}(j)\right).
\end{aligned}$$

To see the last equation, recall that $\boldsymbol{s}^1, \boldsymbol{s}^2$ and $\boldsymbol{s}^3$ are within m_n distance in the first term. Thus, fixing $\boldsymbol{s}^1$ (giving n^2), then $\boldsymbol{s}^2$ and $\boldsymbol{s}^3$ (giving m_n^4), using $m_n p_{\boldsymbol{0}} = O(1)$ produces the first term. For the second term, note that $\#|B(\boldsymbol{t},n) \cap \{\boldsymbol{s} : d(\boldsymbol{s},\boldsymbol{t}) > m_n\}| = O(n^2)$ since $m_n/n \to 0$ and that $E|(\tilde{I}_{\boldsymbol{s}^1})^2 \tilde{I}_{\boldsymbol{s}^2} \tilde{I}_{\boldsymbol{s}^3}| \le \text{const}\,(\alpha_{c,2c}(j) + p_{\boldsymbol{0}}\, \alpha_{c,c}(j))$.

The same logic can be applied to find the upper bound for $A_5, ..., A_8$. In particular, we have $A_8 = O\left(m_n^{17}/n^2\right) + O\left(m_n^4 n^6 \sum_{j>m_n} \alpha_{kc,lc}(j)\right)$ for $1 \leq k, l \leq 7$ and $k + l \leq 8$, which converges to 0 under the mixing conditions of Theorem 8. □

Chapter 4

Conclusion and Future Directions

In this thesis, the notion of the extremogram that is originally defined for stationary time series has been extended to spatial observations. This includes introducing the empirical spatial extremogram (ESE) that is defined to reflect different sampling schemes and studying the asymptotic properties of it. For irregularly spaced data, for example, we define the kernel estimate of the extremogram.

Chapter 2 examines the asymptotic normality of the ESE under two sampling scenarios: the lattice and non-lattice. For the lattice case, the asymptotic results of the ESE can be viewed as a generalization of the asymptotic results for a stationary time series in Davis and Mikosch (2009). In particular, the limiting variance and the scaling term need to be coordinated by a dimension to achieve the central limit theorem. For non-lattice cases, a kernal estimator following ideas in Karr (1986) and Li et al. (2008) is considered. When the growth rates of sampling regions, and the decay rates of bandwidths and mixing coefficients are coordinated, the central limit theorem for the ESE holds.

Chapter 3 explores resampling methods to construct asymptotically correct confidence intervals. When the samples are from the lattice, the validity of the circular block bootstrapped ESE is shown under suitable assumptions on the rates of the size of the bootstrap block and the decaying rates of mixing functions. In the same setting, L_2 consistency is proved for the variance estimated by jackknife and subsampling methods. For a non-lattice observation, it is shown that subsampling the variance of the ESE is consistent in L_2. Based

on these resampling methods, a χ^2 based portmanteau style test is proposed to check the lack of extremal dependence for multiple lags.

The performance of these results are investigated through simulation examples such as max-moving average of order 1, the infinite max moving average, and the Brown-Resnick process. This study shows that the ESE is capable of capturing theoretical aspects for at least these processes. Moreover, real data applications consisting of a rainfall in a region of Florida and ozone data in the eastern United States show that the ESE provides consistent results with the existing literature. Also, the bootstrapped ESE and the variance of the ESE estimated by jackknife or subsampling allow one to construct credible confidence intervals, which facilitates the use of the ESE in practice.

In the future, there are a number of directions in which to expand this current work.

- **Extension to space-time:** Extension to spatial-temporal setting is necessary since many real data have such structure. In fact, this extension has been discussed by Davis et al. (2013b), Davis et al. (2013a), and Buhl and Klüppelberg (2016). One should consider not only an increasing spatial-temporal setting, but also a fixed space and increasing temporal setting since many applications have the fixed spatial locations with multiple observations over time.

- **Consider wider classes of random fields:** The distributional assumption in the thesis is that random field is regularly varying. So, one may investigate the possibility to weaken such distributional assumptions. This can be done in some cases such as light tailed distributions, asymptotically independent cases, and distributions posses hidden regular variation.

- **Bias analysis:** Since the ESE is the ratio estimator, it is susceptible to bias. This also explains why the bias corrected confidence interval proposed by Efron (1981) works well. Hence, it would be interesting to explore the "bias corrected" ESE and see how much enhancement can be made.

- **More data applications:** The thesis considers two data applications: rainfall in Florida and ground-level ozone in the eastern United States. Even though these

applications show the usage of the ESE as a tool to discover the spatial extremal dependence, it does not cast light on how the ESE can be used in selecting competing models or checking the goodness of the fit of models. It will be interesting to find some applications where the ESE plays a important role as a tool to select/validate different models.

Bibliography

J. Blanchet and A.C. Davison. Spatial modeling of extreme snow depth. *Annals of Applied Statistics*, 5:1699–1725, 2011.

E. Bolthausen. On the central limit theorem for stationary mixing random fields. *Annals of Probability*, 10:1047–1050, 1982.

R. C. Bradley. Equivalent mixing conditions for random fields. *Annals of Probability*, 21: 1921–1926, 1993.

S. Buhl and C. Klüppelberg. Anisotropic Brown-Resnick space-time processes: estimation and model assessment. *Extremes (to appear)*, 2016. URL `http://link.springer.com/article/10.1007/s10687-016-0257-1`.

E. Carlstein. The use of subseries methods for estimating the variance of a general statistic from a stationary time series. *Annals of Statistics*, 14:1171–1179, 1986.

S. Coles and E. Casson. Extreme value modelling of hurricane wind speeds. *Structural Safety*, 20:283–296, 1998.

R. Cont and Y.H. Kan. Statistical modeling of credit default swap portfolios. 2011. URL `http://papers.ssrn.com/sol3/papers.cfm?abstract`$_id = 1771862$.

D. Cooley, D. Nychka, and P. Naveau. Bayesian spatial modeling of extremeprecipitation return levels. *Journal of the American Statistical Association*, 102:824–840, 2007.

R. A. Davis and T. Hsing. Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Annals of Probability*, 23:879–917, 1995.

R. A. Davis and T. Mikosch. Extreme value theory for space-time processes with heavy-tailed distributions. *Stochastic Processes and Their Applications*, 118:560–584, 2008.

R. A. Davis and T. Mikosch. The extremogram: A correlogram for extreme events. *Bernoulli*, 15:977–1009, 2009.

R. A. Davis, T. Mikosch, and I. Cribben. Towards estimating extremal serial dependence via the bootstrapped extremogram. *Journal of Econometrics*, 170:142–152, 2012.

R. A. Davis, C. Klüppelberg, and C. Steinkohl. Max-stable processes for modeling extremes observed in space and time. *Journal of the Korean Statistical Society*, 42:399–414, 2013a.

R. A. Davis, C. Klüppelberg, and C. Steinkohl. Statistical inference for max-stable processes in space and time. *Journal of the Royal Statistical Society: Series B*, 75:791–819, 2013b.

A.C. Davison, S.A. Padoan, and M. Ribatet. Statistical modeling of spatial extremes. *Statistical Science*, 27:161–186, 2012.

L. de Haan. A spectral representation for max-stable processes. *Annals of Probability*, 12: 1194–1204, 1984.

L. de Haan and A. Ferreira. *Extreme Value Theory: An introduction*. Springer, 2006.

S. Demarta and A.J. McNeil. The t copula and related copulas. *Int. Stat. Rev.*, 73:111–129, 2005.

C. Dombry and F. Eyi-Minko. Strong mixing properties of max-infinitely divisible random fields. *Stochastic Processes and their Applications*, 122:3790–3811, 2012.

P. Doukhan. *Mixing: Properties and Examples*. Springer-Verlag, 1994.

H. Drees, J. Segers, and M. Warchol. Statistics for tail processes of markov chains. *Extremes*, 18:369–402, 2015.

B. Efron. Nonparametric standard errors and confidence intervals. *The Canadian Journal of Statistics*, 9, 1981.

D. Gabda, R. Towe, J. Wadsworth, and J. Tawn. Discussion of "statistical modeling of spatial extremes" by a.c. davison, s.a. padoan and m. ribatet. *Statistical Science*, 27: 189–192, 2012.

E. Gilleland and D. Nychka. Statistical models for monitoring and regulating ground-level ozone. *Environmetrics*, 16:535546, 2005.

E. Gilleland, D. Nychka, and U. Schneider. Spatial models for the distribution of extremes. In J.S. Clark and A. Gelfand, editors, *Applications of Computational Statistics in the Environmental Sciences: Hierarchical Bayes and MCMC Methods*. Oxford University Press, 2006.

H. Hult and F. Lindskog. Regular variation for measures on metric spaces. *Publications de l'Institut Mathématique, Nouvelle Série*, 80:121–140, 2006.

J. Hüsler and R.-D. Reiss. Maxima of normal random vectors: between independence and complete dependence. *Statistics and Probability Letters*, 7:283–286, 1989.

I.A. Ibragimov and Y.V. Linnik. *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, 1971.

N. Jenish and I.R. Prucha. Central limit theorems and uniform laws of large numbers for arrays of random fields. *Journal of Econometrics*, 150:86–98, 2009.

Z. Kabluchko, M. Schlather, and L. de Haan. Stationary max-stable fields associated to negative definite functions. *Annals of Probability*, 37:2042–2065, 2009.

A.F. Karr. Inference for stationary random fields given poisson samples. *Advances in Applied Probability*, 18:406–422, 1986.

H.R. Künsch. The jackknife and the bootstrap for general stationary observations. *Annals of Statistics*, 17:1217–1241, 1989.

S.N. Lahiri. Theoretical comparisons of block bootstrap methods. *Annals of Statistics*, 27, 1999.

M. R. Leadbetter. Extremes and local dependence in stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 65:291306, 1983.

B. Li, M.G. Genton, and M. Sherman. On the asymptotic joint distribution of sample space-time covariance estimators. *Bernoulli*, 14:228–248, 2008.

G. Lindgren. *Stationary Stochastic Processes: Theory and Applications*. Chapman & Hall, 2012.

D.N. Politis and J.P. Romano. A circular block-resampling procedure for stationary data. Technical report, 1991.

D.N. Politis and J.P. Romano. Nonparametric resampling for homogeneous strong mixing random fields. *Journal of Multivariate Analysis*, 47:301–328, 1993.

D.N. Politis and J.P. Romano. The stationary bootstrap. *Journal of the American Statistical Association*, 89, 1994.

D.N. Politis and M. Sherman. Moment estimation for statistics from marked point processes. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 63:261–275, 2001.

D.N. Politis, E. Paparoditis, and J.P. Romano. Large sample inference for irregularly spaced dependent observation based on subsampling. *Sankhya: The Indian Journal of Statistics, Series A*, 60:274–292, 1998.

B. L. S. Prakasa Rao. Conditional independence, conditional mixing and conditional association. *Annals of the Institute of Statistical Mathematics*, 61:441460, 2009.

S.I. Resnick. *Heavy Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, 2006.

M. Schlather and J.A. Tawn. A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika*, 90:139–156, 2003.

R. L. Smith. Max-stable processes and spatial extremes. Unpublished, 1990. URL `http://www.stat.unc.edu/postscript/rs/spatex.pdf`.