DSSR

Discussion Paper No. 133

LOCAL POLYNOMIAL REGRESSION FOR SPATIAL DATA ON $\mathbb{R}^d$

Daisuke Kurisu and Yasumasa Matsuda

November, 2022

Data Science and Service Research
Discussion Paper

Center for Data Science and Service Research
Graduate School of Economic and Management
Tohoku University
27-1 Kawauchi, Aobaku
Sendai 980-8576, JAPAN

LOCAL POLYNOMIAL REGRESSION FOR SPATIAL DATA ON $\mathbb{R}^d$

DAISUKE KURISU AND YASUMASA MATSUDA

Abstract. In this study, we develop a general asymptotic theory of local polynomial (LP) regression for spatial data observed at irregularly spaced locations in a sampling region $R_n \subset \mathbb{R}^d$. We adopt a stochastic sampling design that can generate irregularly spaced sampling sites in a flexible manner and include both pure increasing and mixed increasing domain framework. We first introduce a nonparametric regression model for spatial data defined on $\mathbb{R}^d$ and then establish the asymptotic normality of LP estimators with general order $p \geq 1$. We also propose methods for constructing confidence intervals and establish uniform convergence rates of LP estimators. Our dependence structure conditions on the underlying random field cover a wide class of random fields such as Lévy-driven continuous autoregressive moving average random fields. As an application of our main results, we also discuss a two-sample testing problem for mean functions and their partial derivatives.

1. Introduction

Recently, a considerable interest has been paid on statistical inference of spatial regression models for geostatistical data analysis in many economic and scientific fields such as spatial econometrics, ecology, and seismology. Particularly, nonparametric methods for spatial data have also been the focus of attention. There is fairly extensive literature on the local constant (LC) , local linear (LL), and local polynomial (LP) estimators for dependent data. For stationary time series, we refer to Hansen (2008) and Zhao and Wu (2008) for LC estimators and Masry (1996a,b), and Masry and Fan (1997) for LP estimators. For nonstationary time series, we refer to Kristensen (2009) and Vogt (2012) for LC estimators, and Zhou and Wu (2009) and Zhang and Wu (2015) for LL estimators for quantile curves and conditional mean functions, respectively. For stationary spatial data on $\mathbb{Z}^d$, we refer to El Machkouri and Stoica (2010) and Jenish (2012) for LC estimators and Hallin et al. (2004) and El Machkouri et al. (2017) for LL estimators. For spatial data on $\mathbb{R}^d$, we refer to Kurisu (2019) and Kurisu (2022) who investigate LC estimators for the stationary and locally stationary case, respectively. We also refer to Robinson (2011) for other recent contribution for possibly nonstationary spatial data. Notably, there seems no theoretical results on the statistical properties of LP estimators for (irregularly spaced) spatial data on both $\mathbb{Z}^d$ and $\mathbb{R}^d$ and even the properties of LC estimators are not known, especially under our model.

The goal of this paper is to make progress in this literature by developing a general asymptotic theory for LP estimators of any order $p \geq 1$ for nonstationary spatial data on $\mathbb{R}^d$. The contributions of this paper are as follows.

Date: First version: July 18, 2022. This version: November 1, 2022.

Key words and phrases. irregularly spaced spatial data, Lévy-driven moving average random field, local polynomial regression, two-sample test

MSC2020 subject classifications: 62M30, 62G08, 62G20.

D. Kurisu is partially supported by JSPS KAKENHI Grant Number 20K13468. The authors would like to thank Takuya Ishihara, Taisuke Otsu, Peter Robinson, Masayuki Sawada, and Yoshihiro Yajima for their helpful comments and suggestions.

First, we propose a nonparametric model for spatial data $\{Y(\boldsymbol{x}_{n,i})\}_{i=1}^n$ observed at irregularly spaced sampling sites $\{\boldsymbol{x}_{n,i}\}_{i=1}^n$ over a sampling region $R_n \subset \mathbb{R}^d$ $(d \geq 1)$. Precisely, each $Y(\boldsymbol{x}_{n,i})$ is explained by the sum of a deterministic spatial trend function (i.e. mean function), a random field on $\mathbb{R}^d$ that represents spatial dependence, and a location specific measurement error (see Section 2.1 for details). In many scientific fields, such as ecology, geology, meteorology, and seismology, sampling points are naturally irregular. In fact, measurement stations cannot be placed on a regular grid owing to physical constraints. To cope with irregularly spaced sampling sites, we adopt the stochastic sampling scheme of Lahiri (2003a), which allows the sampling sites to have a non-uniform density in the sampling region and allows the number of sampling sites n to grow at a different rate compared to the volume of the sampling region A_n. This scheme accommodates both the pure increasing domain case ($\lim_{n\to\infty} A_n/n = \kappa \in (0, \infty)$) and the mixed increasing domain case ($\lim_{n\to\infty} A_n/n = 0$). From a theoretical point of view, this scheme covers all possible asymptotic regimes, since it is well known that the sample mean is not consistent under the infill asymptotics (Lahiri (1996)). See Lahiri (2003b), Lahiri and Zhu (2006), Matsuda and Yajima (2009), Bandyopadhyay et al. (2015), Kurisu et al. (2021), and Kurisu (2022) for discussions on the stochastic spatial sampling design. We note that our model can be seen as an extension of the model considered in Müller and Watson (2021) who investigate inference on sample means of irregularly spaced spatial data under stochastic sampling design, to nonparametric regression models.

Second, we (i) establish the asymptotic normality of the LP estimators of the mean function of the proposed model, (ii) construct an estimator for their asymptotic variances, and (iii) derive uniform convergence rates of the LP estimators over compact sets. The LP estimators are broad enough to include the estimation of conditional moment, distribution, density functions, and their partial derivatives, and our theoretical results are applicable to irregularly spaced time series ($d = 1$) as well as spatial data ($d \geq 2$). The results (i) and (ii) enable us to give expressions for the bias and variance/covariance matrix (of the joint asymptotic distribution) of these LP estimators and to construct confidence intervals of the LP estimators, which is also important when performing a hypothesis test on the mean function. To establish the result (iii), we first consider general kernel estimators and derive their uniform convergence rates. Since the estimators include many kernel-based estimators such as, kernel density, LC, LL, and LP estimators for random fields on $\mathbb{R}^d$ with irregularly spaced sampling sites, the results are of independent theoretical interest. We note that the general results are also useful for evaluating both the bias and variance terms of the LP estimators. Particularly, the results on uniform convergence rates enable us to predict the values of the mean function uniformly on a spatial region that does not contain sampling sites. As an application of our main results, we discuss a two-sample test for the mean functions and their partial derivatives, which also seems a novel result for irregularly spaced spatial data. Additionally, in the literature of causal inference, regression discontinuity designs (RDDs), which are based on local polynomial fitting for the mean functions for both treatment and control groups, is known as an important tool for analyzing the (local) average treatment effect of interventions (cf. Hahn et al. (2001) and Calonico et al. (2014)). Existing methods for RDDs build on asymptotic properties of LP estimators for i.i.d. data even for spatial data (cf. Keele and Titiunik (2015) and Ehrlich and Seidel (2018)). We believe our results pave the way for a new framework of RDDs for spatially dependent data.

Third, we provide examples of random fields that can be covered by our assumptions. Specifically, we show that a broad class of Lévy-driven moving average (MA) random fields, which includes continuous autoregressive moving average (CARMA) random fields (cf. Brockwell and Matsuda (2017)), satisfies our assumptions. The CARMA random fields are known as a rich class of models for spatial data that can represent non-Gaussian random fields as well as Gaussian random fields if the driving Lévy random measures are purely non-Gaussian (cf. Brockwell and Matsuda (2017), Matsuda and Yajima (2018), and Kurisu (2022)). However, mixing properties of Lévy-driven MA random fields have not been investigated since it is often difficult to check mixing conditions as considered in Lahiri and Zhu (2006) and Bandyopadhyay et al. (2015) for general (possibly non-Gaussian) random fields on $\mathbb{R}^d$ except for a class of Gaussian processes. We show that a wide class of Lévy-driven MA random fields can be approximated by m_n-dependent random fields with $m_n \to \infty$ as $n \to \infty$. As a result, this study also contributes to the flexible modeling of nonparametric, nonstationary and possibly non-Gaussian random fields on $\mathbb{R}^d$ by addressing an open question on the dependence structure of statistical models built on Lévy-driven MA random fields.

To the best of our knowledge, our work is the first attempt to establish an asymptotic theory on local polynomial fitting for spatial models on $\mathbb{R}^d$ by (i) establishing the asymptotic normality and uniform convergence rates of the LP estimators of the mean function of the proposed model, (ii) proposing methods for constructing confidence intervals of the LP estimators, and (iii) showing the applicability of our theoretical results for a wide class of Lévy-driven MA random fields. From a theoretical point of view, the present paper builds on Lahiri (2003a) and Lahiri and Zhu (2006), but our theoretical analysis differs substantially from those references in several important points. Specifically, (i) we extend the coupling technique used in Yu (1994) to irregularly spatial data to establish uniform convergence rates of the LP estimators. This extension is non-trivial since there is no natural ordering for spatial data and the number of observations in each block constructed is random, and hence our approach to blocking construction for establishing uniform rates is quite different from those in Lahiri (2003a) and Lahiri and Zhu (2006) whose proofs essentially rely on approximating the characteristic function of the weighted sample mean by that of independent blocks. (ii) We explore concrete random fields that satisfy our assumptions in detail, while Lahiri (2003a) and Lahiri and Zhu (2006) lack a detailed discussion on random fields that satisfy their mixing conditions and other regularity conditions. Verification of our regularity conditions to Lévy-driven MA fields is indeed non-trivial and relies on several probabilistic techniques from Lévy process theory and theory of infinitely divisible random measures (cf. Bertoin (1996), Sato (1999), and Rajput and Rosinski (1989)). Further, in our framework, we cannot use common techniques for the analysis of (equidistant) time series to show the asymptotic normality or uniform convergence of estimators due to the irregularly spaced observations. Specifically, it seems not possible to construct a martingale difference sequence, as is common in the analysis of temporally dependent data.

The rest of the paper is organized as follows. In Section 2, we introduce our nonparametric regression model for spatial data with irregularly spaced sampling sites. In Section 3, we define local polynomial estimators as solutions of a multivariate weighted least squares problem. In Section 4, we establish the asymptotic normality of the LP estimators and construct estimators of their asymptotic variances. In Section 5 we provide the uniform convergence rates of a general kernel estimators and as a special case, we provide the uniform convergence rates of the LP estimators.

In Section 6, we provide examples of the random fields that satisfies our assumptions. All proofs are included in Appendix.

1.1. **Notation.** For any vector $\boldsymbol{x} = (x_1, \dots, x_q)' \in \mathbb{R}^q$, let $|\boldsymbol{x}| = \sum_{j=1}^q |x_j|$ and $\|\boldsymbol{x}\| = \sqrt{\sum_{j=1}^q x_j^2}$ denote the ℓ^1-norm and ℓ^2-norms of $\boldsymbol{x}$, respectively. For any set $A \subset \mathbb{R}^d$ and any vector $\boldsymbol{a} = (a_1, \dots, a_d)' \in (0, \infty)^d$, let $|A|$ denote the Lebesgue measure of A, let $[\![A]\!]$ denote the number of elements in A, and let $\boldsymbol{a}A = \{(a_1x_1, \dots a_dx_d) : \boldsymbol{x} = (x_1, \dots, x_d) \in A\}$. For any positive sequences a_n, b_n, we write $a_n \lesssim b_n$ if there is a constant $C > 0$ independent of n such that $a_n \leq Cb_n$ for all n, $a_n \sim b_n$ if $a_n \lesssim b_n$ and $b_n \lesssim a_n$. For a sequence of random variables $\{\boldsymbol{X}_i\}_{i\geq 1}$, let $\sigma(\{\boldsymbol{X}_i\}_{i\geq 1})$ denote the σ-field generated by $\{\boldsymbol{X}_i\}_{i\geq 1}$. Let $E_{\boldsymbol{X}}$ denote the expectation with respect to a sequence of random variables $\{\boldsymbol{X}_i\}_{i\geq 1}$ and let $P_{\cdot|\boldsymbol{X}}$ and $E_{\cdot|\boldsymbol{X}}$ denote the conditional probability and expectation given $\sigma(\{\boldsymbol{X}_i\}_{i\geq 1})$. For any real-valued random variable X and $\tau \in (0, 1)$, let $q_{1-\tau} = \inf\{x \in \mathbb{R} : P(X \leq x) \geq 1 - \tau\}$ be the $(1 - \tau)$-quantile of X. For $a \in \mathbb{R}$ and $b > 0$, we use the shorthand notation $[a \pm b] = [a - b, a + b]$.

2. Settings

In this section, we discuss the mathematical settings of our model (Section 2.1), sampling design (Section 2.2) and spatial dependence structure (Section 2.3).

2.1. **Model.** Consider the following nonparametric regression model:

$$\begin{aligned} Y(\boldsymbol{x}_{n,i}) &= m\left(\frac{\boldsymbol{x}_{n,i}}{A_n}\right) + \eta\left(\frac{\boldsymbol{x}_{n,i}}{A_n}\right) e(\boldsymbol{x}_{n,i}) + \sigma_\varepsilon\left(\frac{\boldsymbol{x}_{n,i}}{A_n}\right) \varepsilon_i, \qquad (2.1) \\ &:= m\left(\frac{\boldsymbol{x}_{n,i}}{A_n}\right) + e_{n,i} + \varepsilon_{n,i},\ \boldsymbol{x}_{n,i} = (x_{ni,1}, \dots, x_{ni,d})' \in R_n,\ i = 1, \dots, n, \end{aligned}$$

where $R_n = \prod_{j=1}^d [-A_{n,j}/2, A_{n,j}/2]^d$, $A_n = \prod_{j=1}^d A_{n,j}$, $\frac{\boldsymbol{x}_{n,i}}{A_n} = \left(\frac{x_{ni,1}}{A_{n,1}}, \dots, \frac{x_{ni,d}}{A_{n,d}}\right)'$ with $A_{n,j} \to \infty$ as $n \to \infty$, $m : \mathbb{R}^d \to \mathbb{R}$ is the mean function, $\boldsymbol{e} = \{e(\boldsymbol{x}) : \boldsymbol{x} \in \mathbb{R}^d\}$ is a random field defined on $\mathbb{R}^d$ with $E[e(\boldsymbol{x})] = 0$ and $E[e^2(\boldsymbol{x})] = 1$ for any $\boldsymbol{x} \in \mathbb{R}^d$, $\eta : \mathbb{R}^d \to (0, \infty)$ is the variance function of spatially dependent random variables $\{e_{n,i}\}$, $\{\varepsilon_i\}$ is a sequence of i.i.d. random variables such that $E[\varepsilon_i] = 0$ and $E[\varepsilon_i^2] = 1$, and $\sigma_\varepsilon : \mathbb{R}^d \to (0, \infty)$ is the variance function of random variables $\{\varepsilon_{n,i}\}$. Intuitively, the mean function m represents deterministic spatial trend, the random field $\boldsymbol{e}$ represents spatial correlation, and the random variables $\{\varepsilon_{n,i}\}$ can represent location specific measurement error. We note that our model is an extension of the model considered in Müller and Watson (2021). The above setup (2.1) is broad enough to include estimating function of the form $m_F(\boldsymbol{z}) = E[F(Y(\boldsymbol{x}))|\boldsymbol{x}/A_n = \boldsymbol{z}]$ by using the new data set $\{(F(Y(\boldsymbol{x}_{n,i})), \boldsymbol{x}_{n,i})\}_{i=1}^n$. Note that these functions include the conditional moment, conditional distribution, conditional density functions, and their partial derivatives with respect to $\boldsymbol{z}$.

We assume the following condition on the mean function m, the variance function η, and $\{\varepsilon_{n,j}\}$:

Assumption 2.1. *Let $U_{\boldsymbol{z}}$ be a neighborhood of $\boldsymbol{z} = (z_1, \dots, z_d) \in (-1/2, 1/2)^d$.*

- *(i) The mean function m is $(p + 1)$-times continuously partial differentiable on $U_{\boldsymbol{z}}$ and define $\partial_{j_1 \dots j_L} m(\boldsymbol{z}) := \partial m(\boldsymbol{z})/\partial z_{j_1} \dots z_{j_L}$, $1 \leq j_1, \dots, j_L \leq d$, $0 \leq L \leq p + 1$. When $L = 0$, we set $\partial_{j_1 \dots j_L} m(\boldsymbol{z}) = \partial_{j_0} m(\boldsymbol{z}) = m(\boldsymbol{z})$.*
- *(ii) The function η is continuous over $U_{\boldsymbol{z}}$ and $\eta(\boldsymbol{z}) > 0$.*

(iii) The random variables $\{\varepsilon_i\}_{i=1}^n$ are i.i.d. with $E[\varepsilon_1] = 0$, $E[\varepsilon_1^2] = 1$, $E[|\varepsilon_1|^{q_1}] < \infty$ for some integer $q_1 > 4$, and the function $\sigma_\varepsilon(\cdot)$ is continuous over $U_{\boldsymbol{z}}$ with $\sigma_\varepsilon(\boldsymbol{z}) > 0$.

2.2. **Sampling design.** To account for irregularly spaced data, we consider the stochastic sampling design. First, we define the sampling region R_n. For $j = 1, \dots, d$, let $\{A_{n,j}\}_{n\geq 1}$ be a sequence of positive numbers such that $A_{n,j} \to \infty$ as $n \to \infty$. We consider the following set as the sampling region.

$$R_n = \prod_{j=1}^{d} [-A_{n,j}/2, A_{n,j}/2].$$

Next, we introduce our (stochastic) sampling designs. Let $g(\boldsymbol{z}) = g(z_1, \dots, z_d)$ be a probability density function on $R_0 = [-1/2, 1/2]^d$, and let $\{\boldsymbol{X}_{n,i}\}_{i\geq 1}$ be a sequence of i.i.d. random vectors with probability density $A_n^{-d} g(\boldsymbol{x}/A_n) = A_n^{-d} g(x_1/A_{n,1}, \dots, x_d/A_{n,d})$ where $A_n = \prod_{j=1}^d A_{n,j}$. We assume that the sampling sites $\boldsymbol{x}_{n,1}, \dots, \boldsymbol{x}_{n,n}$ are obtained from the realizations of random vectors $\boldsymbol{X}_{n,1}, \dots, \boldsymbol{X}_{n,n}$. To simplify the notation, we will write $\boldsymbol{x}_{n,i}$ and $\boldsymbol{X}_{n,i}$ as $\boldsymbol{x}_i = (x_{i,1}, \dots, x_{i,d})'$ and $\boldsymbol{X}_i = (X_{i,1}, \dots, X_{i,d})'$, respectively.

We summarize conditions on the stochastic sampling design as follows:

Assumption 2.2. *Recall that $U_{\boldsymbol{z}}$ is a neighborhood of $\boldsymbol{z} \in (-1/2, 1/2)^d$. Let g be a probability density function with support $R_0 = [-1/2, 1/2]^d$.*

(i) $A_n/n \to \kappa \in [0, \infty)$ as $n \to \infty$,
(ii) $\{\boldsymbol{X}_i = (X_{i,1}, \dots, X_{i,d})'\}_{i=1}^n$ is a sequence of i.i.d. random vectors with density $A_n^{-d} g(\cdot/A_n)$ and g is continuous over $U_{\boldsymbol{z}}$ and $g(\boldsymbol{z}) > 0$.
(iii) $\{\boldsymbol{X}_i\}_{i=1}^n$, $\boldsymbol{e} = \{e(\boldsymbol{x}) : \boldsymbol{x} \in \mathbb{R}^d\}$, and $\{\varepsilon_i\}_{i=1}^n$ are mutually independent.

Condition (i) implies that our sampling design allows both the pure increasing domain case ($\lim_{n\to\infty} A_n/n = \kappa \in (0, \infty)$) and the mixed increasing domain case ($\lim_{n\to\infty} A_n/n = 0$). This implies that our study addresses the infill sampling criteria in the stochastic design case (cf. Cressie (1993) and Lahiri (2003b)), which is of interest in geostatistical and environmental monitoring applications (cf. Lahiri and Zhu (2006)). Condition (ii) implies that the sampling density can be nonuniformly distributed over the sampling region $R_n = \prod_{j=1}^d [-A_{n,j}/2, A_{n,j}/2]$. It is straightforward to extend the definition of the sampling region R_n to a more general case that includes non-standard shapes (e.g., ellipsoids, polyhedrons, and non-convex sets) as considered in Lahiri and Zhu (2006).

2.3. **Dependence structure.** We assume that random field $\boldsymbol{e}$ satisfies a mixing condition. First, we define the α- and β-mixing coefficients for the random field $\boldsymbol{e}$. Let $\mathcal{F}_{\boldsymbol{e}}(T) = \sigma(\{e(\boldsymbol{x}) : \boldsymbol{x} \in T\})$ be the σ-field generated by the variables $\{e(\boldsymbol{x}) : \boldsymbol{x} \in T\}$, $T \subset \mathbb{R}^d$. For any two subsets T_1 and T_2 of $\mathbb{R}^d$, let

$$\bar{\alpha}(T_1, T_2) = \sup\{|P(A \cap B) - P(A)P(B)| : A \in \mathcal{F}_{\boldsymbol{e}}(T_1), B \in \mathcal{F}_{\boldsymbol{e}}(T_2)\},$$
$$\bar{\beta}(T_1, T_2) = \sup \frac{1}{2} \sum_{j=1}^{J} \sum_{k=1}^{K} |P(A_j \cap B_k) - P(A_j)P(B_k)|,$$

where the supremum for $\bar{\beta}(T_1, T_2)$ is taken over all pairs of (finite) partitions $\{A_1, \dots, A_J\}$ and $\{B_1, \dots, B_K\}$ of $\mathbb{R}^d$ such that $A_j \in \mathcal{F}_{\boldsymbol{e}}(T_1)$ and $B_k \in \mathcal{F}_{\boldsymbol{e}}(T_2)$. The α- and β-mixing coefficients of the random field $\boldsymbol{e}$ is defined as

$$\alpha(a;b) = \sup\{\bar{\alpha}(T_1, T_2) : d(T_1, T_2) \geq a, T_1, T_2 \in \mathcal{R}(b)\},$$
$$\beta(a;b) = \sup\{\bar{\beta}(T_1, T_2) : d(T_1, T_2) \geq a, T_1, T_2 \in \mathcal{R}(b)\}.$$

where $a, b > 0$, $d(T_1, T_2) = \inf\{|\boldsymbol{x} - \boldsymbol{y}| : \boldsymbol{x} \in T_1, \boldsymbol{y} \in T_2\}$, and $\mathcal{R}(b)$ is the collection of all the finite disjoint unions of cubes in $\mathbb{R}^d$ with a total volume not exceeding b. Moreover, we assume that there exist a non-increasing functions α_1 and β_1 with $\alpha_1(a), \beta_1(a) \to 0$ as $a \to \infty$ and a non-decreasing functions ϖ_1 and ϖ_2 (that may be unbounded) such that

$$\alpha(a;b) \leq \alpha_1(a)\varpi_1(b),\ \beta(a;b) \leq \beta_1(a)\varpi_2(b).$$

Remark 2.1. The definitions of the α- and β-mixing coefficients are based on the argument in Bradley (1989). It is important to restrict the size of the index sets T_1 and T_2 in the definition of α- (or β-) mixing coefficients. Let us define the β-mixing coefficient of a random field $\boldsymbol{e}$ similarly to the time series as follows: For any subsets T_1 and T_2 of $\mathbb{R}^d$, the β-mixing coefficient between $\mathcal{F}_{\boldsymbol{e}}(T_1)$ and $\mathcal{F}_{\boldsymbol{e}}(T_2)$ is defined by $\widetilde{\beta}(T_1, T_2) = \sup \sum_{j=1}^{J} \sum_{k=1}^{K} |P(A_j \cap B_k) - P(A_j)P(B_k)|/2$, where the supremum is taken over all partitions $\{A_j\}_{j=1}^J \subset \mathcal{F}_{\boldsymbol{e}}(T_1)$ and $\{B_k\}_{k=1}^K \subset \mathcal{F}_{\boldsymbol{e}}(T_2)$ of $\mathbb{R}^d$. Let $\mathcal{O}_1$ and $\mathcal{O}_2$ be half-planes with boundaries L_1 and L_2, respectively. For each $a > 0$, define $\beta(a) = \sup\{\widetilde{\beta}(\mathcal{O}_1, \mathcal{O}_2) : d(\mathcal{O}_1, \mathcal{O}_2) \geq a\}$. According to Theorem 1 in Bradley (1989), if $\{e(\boldsymbol{x}) : \boldsymbol{x} \in \mathbb{R}^2\}$ is strictly stationary, then $\beta(a) = 0$ or 1 for $a > 0$. This implies that if a random field $\boldsymbol{e}$ is β-mixing ($\lim_{a\to\infty} \beta(a) = 0$), then it is automatically m dependent, that is, $\beta(a) = 0$ for some $a > m$, where m is a positive constant. To allow a certain flexibility, we restrict the size of T_1 and T_2 in the definitions of $\alpha(a;b)$ and $\beta(a;b)$. We refer to Bradley (1993) and Doukhan (1994) for more details on mixing coefficients for random fields.

For the asymptotic normality of the LP estimators, we assume the following conditions for the random field $\boldsymbol{e}$:

Assumption 2.3. *For $j = 1, \ldots, d$, let $\{A_{n1,j}\}_{n\geq 1}$ and $\{A_{n2,j}\}_{n\geq 1}$ be sequences of positive numbers such that* $\min\left\{A_{n2,j}, \frac{A_{n1,j}}{A_{n2,j}}, \frac{A_{n,j}h_j}{A_{n1,j}}\right\} \to \infty$ *as $n \to \infty$.*

(i) The random field $\mathbf{e}$ is stationary and $E[|\boldsymbol{e}(\mathbf{0})|^{q_2}] < \infty$ for some integer $q_2 > 4$.

(ii) Define $\sigma_{\boldsymbol{e}}(\boldsymbol{x}) = E[e(\mathbf{0})e(\boldsymbol{x})]$. Assume that $\sigma_{\boldsymbol{e}}(\mathbf{0}) = 1$ and $\int_{\mathbb{R}^d} |\sigma_{\boldsymbol{e}}(\boldsymbol{v})| d\boldsymbol{v} < \infty$.

(iii) The random field $\mathbf{e}$ is α-mixing with mixing coefficients $\alpha(a;b)$ such that as $n \to \infty$,

$$A_n^{(1)} \left(\alpha_1^{1-2/q}(\underline{A}_{n2}) + \sum_{k=\underline{A}_{n1}}^{\infty} k^{d-1} \alpha_1^{1-2/q}(k) \right) \varpi_1^{1-2/q}(A_n^{(1)}) \to 0,$$

where $q = \min\{q_1, q_2\}$, $A_n^{(1)} = \prod_{j=1}^d A_{n1,j}$, $\underline{A}_{n1} = \min_{1\leq j\leq d} A_{n1,j}$, and $\underline{A}_{n2} = \min_{1\leq j\leq d} A_{n2,j}$.

The sequences $\{A_{n1,j}\}$ and $\{A_{n2,j}\}$ will be used in the large-block-small-block argument, which is commonly used in proving CLTs for sums of mixing random variables. Specifically, $A_{n1,j}$ corresponds to the side length of large blocks, while $A_{n2,j}$ corresponds to the side length of small blocks. In Section 6, we provide examples of random fields that satisfy Assumptions 2.3 and 4.1 below. In particular, a wide class of Lévy-driven moving average (MA) random fields that includes continuous autoregressive and moving average (CARMA) random fields (cf. Brockwell and Matsuda (2017)) satisfies our assumptions.

3. Local polynomial regression of order p

In this section, we introduce local polynomial (LP) estimators of order $p \geq 1$ for the estimation of derivatives of the mean function m of the model (2.1).

Define

$$D = [\![\{(j_1, \ldots, j_L) : 1 \leq j_1 \leq \cdots \leq j_L \leq d, 0 \leq L \leq p\}]\!],$$
$$\bar{D} = [\![\{(j_1, \ldots, j_{p+1}) : 1 \leq j_1 \leq \cdots \leq j_{p+1} \leq d\}]\!],$$

$(s_{j_1 \ldots j_L 1}, \ldots, s_{j_1 \ldots j_L d}) \in \mathbb{Z}^d_{\geq 0}$ such that $s_{j_1 \ldots j_L k} = [\![\{j_\ell : j_\ell = k, 1 \leq \ell \leq L\}]\!]$, and define

$$\boldsymbol{s}_{j_1 \ldots j_L}! = \prod_{k=1}^{d} s_{j_1 \ldots j_L k}!.$$

When $L = 0$, we set $(j_1, \ldots, j_L) = j_0 = 0$ and $\boldsymbol{s}_{j_1 \ldots j_L}! = 1$. Note that $\sum_{k=1}^d s_{j_1 \ldots j_L k} = L$. Further, for $p \geq 1$ and $\boldsymbol{z} \in [-1/2, 1/2]^d$, define

$$\begin{aligned}
\boldsymbol{M}(\boldsymbol{z}) &:= \Bigg(m(\boldsymbol{z}), \partial_1 m(\boldsymbol{z}), \ldots, \partial_d m(\boldsymbol{z}), \frac{\partial_{11} m(\boldsymbol{z})}{2!}, \frac{\partial_{12} m(\boldsymbol{z})}{1!1!}, \ldots, \frac{\partial_{dd} m(\boldsymbol{z})}{2!}, \\
&\qquad \ldots, \frac{\partial_{1 \ldots 1} m(\boldsymbol{z})}{p!}, \frac{\partial_{1 \ldots 2} m(\boldsymbol{z})}{(p-1)!1!} \ldots, \frac{\partial_{d \ldots d} m(\boldsymbol{z})}{p!} \Bigg)' \\
&= \left(\frac{1}{\boldsymbol{s}_{j_1 \ldots j_L}!} \partial_{j_1, \ldots j_L} m(\boldsymbol{z}) \right)'_{1 \leq j_1 \leq \cdots \leq j_L \leq d, 0 \leq L \leq p} \in \mathbb{R}^D.
\end{aligned}$$

We define the local polynomial regression estimator of order p for $M(\boldsymbol{z})$ as a solution of the following problem:

$$\widehat{\boldsymbol{\beta}}(\boldsymbol{z}) := \arg\min_{\boldsymbol{\beta} \in \mathbb{R}^D} \sum_{i=1}^n \left(Y(\boldsymbol{X}_i) - \sum_{L=0}^{p} \sum_{1 \leq j_1 \leq \cdots \leq j_L \leq d} \beta_{j_1 \ldots j_L} \prod_{\ell=1}^L \left(\frac{X_{i,j_\ell} - A_{n,j_\ell} z_{j_\ell}}{A_{n,j_\ell}} \right) \right)^2 K_{Ah}\left(\boldsymbol{X}_i - A_n \boldsymbol{z}\right) \tag{3.1}$$

$$\begin{aligned}
&= (\widehat{\beta}_0(\boldsymbol{z}), \widehat{\beta}_1(\boldsymbol{z}), \ldots, \widehat{\beta}_d(\boldsymbol{z}), \widehat{\beta}_{11}(\boldsymbol{z}), \ldots, \widehat{\beta}_{dd}(\boldsymbol{z}), \ldots, \widehat{\beta}_{1\ldots 1}(\boldsymbol{z}), \ldots, \widehat{\beta}_{d \ldots d}(\boldsymbol{z}))' \\
&= (\widehat{\beta}_{j_1 \ldots j_L}(\boldsymbol{z}))'_{1 \leq j_1 \leq \cdots \leq j_L \leq d, 0 \leq L \leq p},
\end{aligned}$$

where $\boldsymbol{\beta} = (\beta_{j_1 \ldots j_L})'_{1 \leq j_1 \leq \cdots \leq j_L \leq d, 0 \leq L \leq p}$, $K : \mathbb{R}^d \to \mathbb{R}$ is a kernel function, and each h_j is a sequence of positive constants (bandwidths) such that $h_j \to 0$ as $n \to \infty$, and where

$$K_{Ah}(\boldsymbol{X}_i - A_n \boldsymbol{z}) = K\left(\frac{X_{i,1} - A_{n,1} z_1}{A_{n,1} h_1}, \ldots, \frac{X_{i,d} - A_{n,d} z_d}{A_{n,d} h_d} \right)$$

and $\sum_{1 \leq j_1 \leq \cdots \leq j_L \leq d} \beta_{j_1 \ldots j_L} \prod_{\ell=1}^L (X_{i,j_\ell} - A_{n,j_\ell} z_{j_\ell}) / A_{n,j_\ell} = \beta_0$ when $L = 0$.

To compute the LP estimators, we introduce some notations: $\boldsymbol{Y} := (Y(\boldsymbol{X}_1), \ldots, Y(\boldsymbol{X}_n))'$,

$$\boldsymbol{X} := (\widetilde{\boldsymbol{X}}_1, \ldots, \widetilde{\boldsymbol{X}}_n) = \begin{pmatrix} 1 & \ldots & 1 \\ \frac{(\boldsymbol{X}_1 - A_n \boldsymbol{z})_1}{A_n} & \ldots & \frac{(\boldsymbol{X}_n - A_n \boldsymbol{z})_1}{A_n} \\ \vdots & \ldots & \vdots \\ \frac{(\boldsymbol{X}_1 - A_n \boldsymbol{z})_p}{A_n} & \ldots & \frac{(\boldsymbol{X}_n - A_n \boldsymbol{z})_p}{A_n} \end{pmatrix} = \begin{pmatrix} 1 & \ldots & 1 \\ (\boldsymbol{X}_1 \check{-} A_n \boldsymbol{z}) & \ldots & (\boldsymbol{X}_n \check{-} A_n \boldsymbol{Z}) \end{pmatrix},$$

$$\boldsymbol{W} := \operatorname{diag}\left(K_{Ah}\left(\boldsymbol{X}_1 - A_n \boldsymbol{z}\right), \ldots, K_{Ah}\left(\boldsymbol{X}_n - A_n \boldsymbol{z}\right)\right),$$

where

$$\frac{(\boldsymbol{X}_i - A_n\boldsymbol{z})_L}{A_n} = \left(\prod_{\ell=1}^{L}\left(\frac{X_{i,j_\ell} - A_{n,j_\ell}z_{j_\ell}}{A_{n,j_\ell}}\right)\right)'_{1\le j_1\le\cdots\le j_L\le d}.$$

The minimization problem (3.1) can be rewritten as

$$\widehat{\boldsymbol{\beta}}(\boldsymbol{z}) = \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^D}(\boldsymbol{Y} - \boldsymbol{X}'\boldsymbol{\beta})'\boldsymbol{W}(\boldsymbol{Y} - \boldsymbol{X}'\boldsymbol{\beta}) =: \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^D} Q_n(\boldsymbol{\beta}).$$

Then the first order condition of the problem (3.1) is given by

$$\frac{\partial}{\partial\boldsymbol{\beta}}Q_n(\boldsymbol{\beta}) = -2\boldsymbol{X}\boldsymbol{W}\boldsymbol{Y} + 2\boldsymbol{X}\boldsymbol{W}\boldsymbol{X}'\boldsymbol{\beta} = 0.$$

Hence the solution of the problem (3.1) is given by

$$\begin{aligned}\widehat{\boldsymbol{\beta}}(\boldsymbol{z}) &= (\boldsymbol{X}\boldsymbol{W}\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{W}\boldsymbol{Y}\\ &= \left[\sum_{i=1}^{n} K_{Ah}\left(\boldsymbol{X}_i - A_n\boldsymbol{z}\right)\widetilde{\boldsymbol{X}}_i\widetilde{\boldsymbol{X}}_i'\right]^{-1}\sum_{i=1}^{n} K_{Ah}\left(\boldsymbol{X}_i - A_n\boldsymbol{z}\right)\widetilde{\boldsymbol{X}}_i Y(\boldsymbol{X}_i).\end{aligned}$$

We assume the following conditions on the kernel function K:

Assumption 3.1. *Let $K : \mathbb{R}^d \to \mathbb{R}$ be a kernel function such that*

(i) $\int K(\boldsymbol{z})d\boldsymbol{z} = 1$.

(ii) The kernel function K is bounded and supported on $S_K \subset [-1/2, 1/2]^d$ with $U_{\boldsymbol{z}} \subset S_K$.

(iii) Define $\kappa_0^{(r)} := \int K^r(\boldsymbol{z})d\boldsymbol{z}$, $\kappa_{j_1,\dots,j_M}^{(r)} := \int \prod_{\ell=1}^{M} z_{j_\ell}K^r(\boldsymbol{z})d\boldsymbol{z}$, and

$$\check{\boldsymbol{z}} := (1, (\boldsymbol{z})_1', \dots, (\boldsymbol{z})_p')',\ (\boldsymbol{z})_L = \left(\prod_{\ell=1}^{L} z_{j_\ell}\right)'_{1\le j_1\le\cdots\le j_L\le d},\ 1 \le L \le p.$$

The matrix $S = \int \begin{pmatrix} 1 \\ \check{\boldsymbol{z}} \end{pmatrix}(1\ \check{\boldsymbol{z}}')K(\boldsymbol{z})d\boldsymbol{z}$ *is non-singular.*

4. Main results

In this section, we discuss asymptotic properties of the LP estimators defined in Section 3. In particular, we establish the asymptotic normality of the LP estimator (Section 4.1) and estimation of the asymptotic variance of the LP estimators (Section 4.2).

4.1. **Asymptotic normality of local polynomial estimators.** We assume the following conditions for the sample size n, bandwidths h_j, constants $A_{n,j}$, $A_{n1,j}$, and $A_{n2,j}$, and mixing coefficients $\alpha(a; b)$:

Assumption 4.1. *Recall $q = \min\{q_1, q_2\}$, $A_n^{(1)} = \prod_{j=1}^{d} A_{n1,j}$, and $\underline{A}_{n1} = \min_{1\le j\le d} A_{n1,j}$. Define $\overline{A}_{n1} = \max_{1\le j\le d} A_{n1,j}$, $\overline{A}_{n2} = \max_{1\le j\le d} A_{n2,j}$, and $\overline{A_n h} = \max_{1\le j\le d} A_{n,j}h_j$. As $n \to \infty$,*

(i) $h_j \to 0$ *for* $1 \le j \le d$.

(ii) $nh_1 \dots h_d \to \infty$.

(iii) $A_n h_1 \dots h_d \times h_{j_1}^2 \dots h_{j_p}^2 \to \infty$ *for* $1 \le j_1 \le \cdots \le j_p \le d$.

(iv) $A_n h_1 \dots h_d \times h_{j_1}^2 \dots h_{j_p}^2 h_{j_{p+1}}^2 \to c_{j_1\dots j_{p+1}} \in [0, \infty)$ *for* $1 \le j_1 \le \cdots \le j_{p+1} \le d$.

(v) $\left(\frac{A_n h_1 \ldots h_d}{A_n^{(1)}}\right) \alpha_1(\underline{A}_{n2}) \varpi_1(A_n h_1 \ldots h_d) \to 0,$

$$\left(\frac{A_n^{(1)}}{A_n h_1 \ldots h_d}\right) \sum_{k=1}^{\overline{A}_{n1}} k^{2d-1} \alpha_1^{1-4/q}(k) \to 0,$$

$$\left\{ \left(\frac{\overline{A}_{n1}}{\underline{A}_{n1}}\right)^d \left(\frac{\overline{A}_{n2}}{\overline{A}_{n1}}\right) + \left(\frac{A_n^{(1)}}{\underline{A}_{n1}^d}\right) \left(\frac{(\overline{A_n h})^d}{A_n h_1 \ldots h_d}\right) \left(\frac{\overline{A}_{n1}}{\overline{\overline{A_n h}}}\right) \right\} \sum_{k=1}^{\overline{A}_{n1}} k^{d-1} \alpha_1^{1-2/q}(k) \to 0.$$

We need Condition (ii) to compute the asymptotic variances of the LP estimators. Conditions (iii) and (iv) are concerned with the rates of convergence of variance and bias terms of the LP estimators, respectively. Condition (v) is concerned with the large-block-small-block argument to show the asymptotic normality of the LP estimators. Indeed, we use the first condition to approximate a weighted sum of spatially dependent data of the form

$$\sum_{i=1}^n K_{Ah}(\boldsymbol{X}_i) H^{-1} \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_i \end{pmatrix} (e_{n,i} + \varepsilon_{n,i})$$

by a sum of independent large blocks. The second condition is used to apply Lyapunov's central limit theorem to the sum of independent blocks. The third conditions is used to show the asymptotic negligibility of a sum of small blocks. See the proof of Theorem 4.1 for detailed definitions of large and small blocks.

Define

$$H := \mathrm{diag}(1, h_1, \ldots, h_d, h_1^2, h_1 h_2, \ldots, h_d^2, \ldots, h_1^p, h_1^{p-1} h_2, \ldots, h_d^p) \in \mathbb{R}^{D \times D}.$$

Throughout Sections 4.1, 4.2, and 4.3, we set $\boldsymbol{z} = \boldsymbol{0}$ without loss of generality. Extending the results in this section to the case $\boldsymbol{z} \in (-1/2, 1/2)^d$ is straightforward.

Theorem 4.1 (Asymptotic normality of local polynomial estimators)**.** *Suppose Assumptions 2.1, 2.2, 2.3, 3.1, and 4.1 hold. Then, as $n \to \infty$, the following result holds:*

$$\sqrt{A_n h_1 \ldots h_d} \left(H\left(\widehat{\boldsymbol{\beta}}(\boldsymbol{0}) - \boldsymbol{M}(\boldsymbol{0})\right) - S^{-1} B^{(d,p)} \boldsymbol{M}_n^{(d,p)}(\boldsymbol{0}) \right)$$

$$\xrightarrow{d} N\left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \left\{ \frac{\kappa(\eta^2(\boldsymbol{0}) + \sigma_\varepsilon^2(\boldsymbol{0}))}{g(\boldsymbol{0})} + \eta^2(\boldsymbol{0}) \int \sigma_{\boldsymbol{e}}(\boldsymbol{v}) d\boldsymbol{v} \right\} S^{-1} \mathcal{K} S^{-1} \right),$$

where

$$B^{(d,p)} = \int \begin{pmatrix} 1 \\ \check{\boldsymbol{z}} \end{pmatrix} (\boldsymbol{z})'_{p+1} K(\boldsymbol{z}) d\boldsymbol{z} \in \mathbb{R}^{D \times \bar{D}},\ \mathcal{K} = \int \begin{pmatrix} 1 \\ \check{\boldsymbol{z}} \end{pmatrix} (1\ \check{\boldsymbol{z}}') K^2(\boldsymbol{z}) d\boldsymbol{z} \in \mathbb{R}^{D \times D},$$

$$\boldsymbol{M}_n^{(d,p)}(\boldsymbol{z}) = \left(\frac{\partial_{j_1 \ldots j_{p+1}} m(\boldsymbol{z})}{\boldsymbol{s}_{j_1 \ldots j_{p+1}}!} \prod_{\ell=1}^{p+1} h_{j_\ell} \right)'_{1 \le j_1 \le \cdots \le j_{p+1} \le d}$$

$$= \left(\frac{\partial_{1 \ldots 1} m(\boldsymbol{z})}{(p+1)!} h_1^{p+1}, \frac{\partial_{1 \ldots 2} m(\boldsymbol{z})}{p!} h_1^p h_2, \ldots, \frac{\partial_{d \ldots d} m(\boldsymbol{z})}{(p+1)!} h_d^p \right)' \in \mathbb{R}^{\bar{D}}.$$

Theorem 4.1 differs from the asymptotic normality of LP estimators under i.i.d. observations in several points. First, the convergence rates of the LP estimators depends not on the sample size n explicitly but on the volume of the sampling region A_n. Second, the asymptotic variance is represented as a sum of two components $\{\kappa(\eta^2(\mathbf{0})+\sigma_\varepsilon^2(\mathbf{0}))\}S^{-1}\mathcal{K}S^{-1}/g(\mathbf{0})$ and $\eta^2(\mathbf{0})\left(\int \sigma_{\boldsymbol{e}}(\boldsymbol{v})d\boldsymbol{v}\right)S^{-1}\mathcal{K}S^{-1}$. When the sampling design satisfies the mixed increasing domain asymptotics, that is, $\kappa=0$, then the asymptotic variance depends only on the second term, which represents the effect of the spatial dependence, and does not includes $\sigma_\varepsilon^2(\mathbf{0})$, the effect of the measurement error $\{\varepsilon_{n,j}\}$. This is completely different from i.i.d. case. We also note that the form of the asymptotic variance in Theorem 4.1 is different from that of Theorem 4 in Masry (1996b) who investigates asymptotic properties of LP estimators for equidistant time series. Indeed, in his result, the variance term that corresponds to the second term of the asymptotic variance in our result does not appear. When the sampling design satisfies the pure increasing domain asymptotics, that is, $\kappa \in (0,\infty)$, then the asymptotic variance depends on both first and second terms. In this case, the asymptotic variance includes the effect of the sampling design $1/g(\mathbf{0})$, which implies that the more likely the sampling sites are distributed around $\mathbf{0}$, the more accurate the estimation of $M(\mathbf{0})$. Moreover, if $\eta(\cdot)\equiv 0$, then the asymptotic variance coincides with that of i.i.d. case.

Remark 4.1 (General form of the mean squared error of $\partial_{j_1\dots j_L}\widehat{m}(\mathbf{0})$)**.** Define

$$\begin{aligned}\boldsymbol{b}_n^{(d,p)}(\boldsymbol{x}) &:= B^{(d,p)}\boldsymbol{M}_n^{(d,p)}(\boldsymbol{x})\\ &= (b_{n,0}(\boldsymbol{x}), b_{n,1}(\boldsymbol{x}),\dots,b_{n,d}(\boldsymbol{x}),\\ &\qquad b_{n,11}(\boldsymbol{x}), b_{n,12}(\boldsymbol{x}),\dots,b_{n,dd}(\boldsymbol{x}),\dots,b_{n,1\dots,1}(\boldsymbol{x}), b_{n,1\dots 2}(\boldsymbol{x}),\dots,b_{n,d\dots d}(\boldsymbol{x}))'\end{aligned}$$

and let $e_{j_1\dots j_L}=(0,\dots,0,1,0,\dots,0)'$ be a D-dimensional vector such that $e'_{j_1\dots j_L}\boldsymbol{b}_n^{(d,p)}(\boldsymbol{x}) = b_{j_1\dots j_L}(\boldsymbol{x})$. Theorem 4.1 yields that

$$b_{n,j_1,\dots,j_L}(\mathbf{0}) = \sum_{1\le j_{1,1}\le\cdots\le j_{1,p+1}\le d}\frac{\partial_{j_{1,1}\dots j_{1,p+1}}m(\mathbf{0})}{\boldsymbol{s}_{j_{1,1}\dots j_{1,p+1}}!}\prod_{\ell_1=1}^{p+1}h_{j_{1,\ell_1}}\kappa^{(1)}_{j_1\dots j_L j_{1,1}\dots j_{1,p+1}},$$

for $1\le j_1\le\cdots\le j_L\le d$, $0\le L\le p$, and the mean squared error (MSE) of the LP estimator $\partial_{j_1\dots j_L}\widehat{m}(\boldsymbol{x})$ is given as follows:

$$\begin{aligned}&\mathrm{MSE}(\partial_{j_1\dots j_L}\widehat{m}(\mathbf{0}))\\ &= E\left[(\partial_{j_1\dots j_L}m(\mathbf{0}) - \partial_{j_1\dots j_L}\widehat{m}(\mathbf{0}))^2\right]\\ &= \left\{\boldsymbol{s}_{j_1\dots j_L}!\frac{(S^{-1}e_{j_1\dots j_L})'B^{(d,p)}\boldsymbol{M}_n^{(d,p)}(\mathbf{0})}{\prod_{\ell=1}^L h_{j_\ell}}\right\}^2\\ &\quad + \left(\frac{\kappa(\eta^2(\mathbf{0})+\sigma_\varepsilon^2(\mathbf{0}))}{g(\mathbf{0})} + \eta^2(\mathbf{0})\int\sigma_{\boldsymbol{e}}(\boldsymbol{v})d\boldsymbol{v}\right)(\boldsymbol{s}_{j_1\dots j_L}!)^2\frac{e'_{j_1\dots j_L}S^{-1}\mathcal{K}S^{-1}e_{j_1\dots j_L}}{A_n h_1\dots h_d\times\left(\prod_{\ell=1}^L h_{j_\ell}\right)^2}.\end{aligned}\tag{4.1}$$

4.2. **Estimation of asymptotic variances.** An estimator of the asymptotic variance of the statistics $\widehat{\boldsymbol{\beta}}(\mathbf{0})$ can be constructed by using leave-one (or two)-out estimators. For $\boldsymbol{z}\in(-1/2,1/2)^d$, let $\widehat{m}_{-I}(\boldsymbol{z})$ be the LP estimator (of order p) of $m(\boldsymbol{z})$ computed without $\{(Y(\boldsymbol{X}_i),\boldsymbol{X}_i)\}_{i\in I}$, $I\subset\{1,\dots,n\}$.

Define

$$\widehat{g}(\mathbf{0}) = \frac{1}{nh_1\dots h_d}\sum_{i=1}^n K_{Ah}(\boldsymbol{X}_i),$$

$$\widehat{V}_{n,1}(\mathbf{0}) = \frac{1}{nh_1 \dots h_d} \sum_{i=1}^{n} K_{Ah}(\boldsymbol{X}_i) \left(Y(\boldsymbol{X}_i) - \widehat{m}_{-\{i\}}(\boldsymbol{X}_i/A_n)\right)^2,$$

$$\widehat{V}_{n,2}(\mathbf{0}) = \frac{A_n}{nh_1 \dots h_d} \sum_{i=1}^{n-1} K_{Ah}(\boldsymbol{X}_i) K_{Ah}(\boldsymbol{X}_{i+1}) \times \left(Y(\boldsymbol{X}_i) - \widehat{m}_{-\{i,i+1\}}(\boldsymbol{X}_i/A_n)\right)\left(Y(\boldsymbol{X}_{i+1}) - \widehat{m}_{-\{i,i+1\}}(\boldsymbol{X}_{i+1}/A_n)\right),$$

Note that $\widehat{m}_{-\{i\}}(\boldsymbol{z})$ and $\widehat{m}_{-\{i,i+1\}}(\boldsymbol{z})$ are leave-i-out and leave-$(i,i+1)$-out version of $\widehat{m}(\boldsymbol{z})$, respectively and then $\widehat{m}_{-\{i\}}(\boldsymbol{z})$ and $\boldsymbol{X}_i$ (or $\widehat{m}_{-\{i,i+1\}}(\boldsymbol{z})$ and $\{\boldsymbol{X}_i, \boldsymbol{X}_{i+1}\}$) are independent under Assumption 2.2.

Proposition 4.1. *Under the assumptions of Theorem 4.1, as* $n \to \infty$,

$$\widehat{V}_n(\mathbf{0}) := \frac{(A_n/n)\widehat{V}_{n,1}(\mathbf{0})}{\widehat{g}^2(\mathbf{0})} + \frac{(\kappa_0^{(2)})^{-1}\widehat{V}_{n,2}(\mathbf{0})}{\widehat{g}^2(\mathbf{0})} \xrightarrow{p} \frac{\kappa(\eta^2(\mathbf{0}) + \sigma_\varepsilon^2(\mathbf{0}))}{g(\mathbf{0})} + \eta^2(\mathbf{0}) \int \sigma_{\mathbf{e}}(\boldsymbol{v}) d\boldsymbol{v}.$$

Theorem 4.1 and Proposition 4.1 enable us to construct confidence intervals of $\partial_{j_1 \dots j_L} m(\mathbf{0})$. Consider a confidence interval of the form

$$C_{n,j_1\dots j_L}(1-\tau) = \left[\partial_{j_1\dots j_L}\widehat{m}(\mathbf{0}) \pm \sqrt{\frac{\widehat{V}_n(\mathbf{0})\,(\boldsymbol{s}_{j_1\dots j_L}!)^2 \left(e'_{j_1\dots j_L} S^{-1}\mathcal{K}S^{-1}e_{j_1\dots j_L}\right)}{A_n h_1 \dots h_d \left(\prod_{\ell=1}^L h_{j_\ell}\right)^2}} q_{1-\tau/2}\right],$$

where $q_{1-\tau}$ is the $(1-\tau)$-quantile of the standard normal random variable. Then we can show the asymptotic validity of the confidence interval as follows:

Corollary 4.1. *Let* $\tau \in (0,1)$. *Under the assumptions of Theorem 4.1 with*

$$A_n h_1 \dots h_d \left((S^{-1}e_{j_1\dots j_L})' B^{(d,p)} M_n^{(d,p)}(\mathbf{0})\right)^2 \to 0$$

as $n \to \infty$. *Then,* $\lim_{n\to\infty} P(\partial_{j_1\dots j_L} m(\mathbf{0}) \in C_{n,j_1\dots j_L}(1-\tau)) = 1-\tau$.

4.3. **Two-sample test for spatially dependent data.** In this section, we discuss two-sample tests for the derivatives of the mean function as an application of our main results.

Consider the following nonparametric regression model:

$$Y_1(\boldsymbol{x}_{1,\ell_1}) = m_1\left(\frac{\boldsymbol{x}_{1,\ell_1}}{A_n}\right) + \eta_1\left(\frac{\boldsymbol{x}_{1,\ell_1}}{A_n}\right) e_1(\boldsymbol{x}_{1,\ell_1}) + \sigma_{\varepsilon,1}\left(\frac{\boldsymbol{x}_{1,\ell_1}}{A_n}\right)\varepsilon_{1,\ell_1},\ \ell_1 = 1,\dots,n_1$$

$$Y_2(\boldsymbol{x}_{2,\ell_2}) = m_2\left(\frac{\boldsymbol{x}_{2,\ell_2}}{A_n}\right) + \eta_2\left(\frac{\boldsymbol{x}_{2,\ell_2}}{A_n}\right) e_2(\boldsymbol{x}_{2,\ell_2}) + \sigma_{\varepsilon,2}\left(\frac{\boldsymbol{x}_{2,\ell_2}}{A_n}\right)\varepsilon_{2,\ell_2},\ \ell_2 = 1,\dots,n_2,$$

where $\boldsymbol{x}_{1,\ell_1}, \boldsymbol{x}_{2,\ell_2} \in R_n$, $\mathbf{e} = \{e(\boldsymbol{x}) = (e_1(\boldsymbol{x}), e_2(\boldsymbol{x}))' : \boldsymbol{x} \in \mathbb{R}^d\}$ is a bivariate stationary random field such that $E[e_k(\mathbf{0})] = 0$, $E[e_k^2(\mathbf{0})] = 1$, and $\{\varepsilon_{k,\ell_k}\}$ is a sequence of i.i.d. random variables such that $E[\varepsilon_{k,\ell_k}] = 0$, $k = 1,2$.

Assume that $\{\boldsymbol{x}_{k,\ell_k}\}$ are realizations of a sequence of random variables $\{\boldsymbol{X}_{k,\ell_k}\}$ with density $A_n^{-1} g_k(\cdot/A_n)$ where $g_k(\cdot)$ is a probability density function with support $[-1/2,1/2]^d$, $k = 1,2$.

Assumption 4.2. *The bivariate random field* $\mathbf{e}$ *satisfies the following conditions:*

- *(i)* $E[|e_k(\mathbf{0})|^{q_2}] < \infty$, $k = 1,2$ *for some integer* $q_2 > 4$.
- *(ii) Define* $\Sigma_{\mathbf{e}}(\boldsymbol{x}) = (\sigma_{\mathbf{e},jk}(\boldsymbol{x}))_{1\le j,k\le 2}$ *where* $\sigma_{\mathbf{e},jk}(\boldsymbol{x}) = E[e_j(\mathbf{0})e_k(\boldsymbol{x})]$, $j,k = 1,2$. *Assume that* $\sigma_{\mathbf{e},kk}(\mathbf{0}) = 1$, $k = 1,2$ *and* $\int_{\mathbb{R}^d} |\sigma_{\mathbf{e},jk}(\boldsymbol{v})| d\boldsymbol{v} < \infty$, $j,k = 1,2$.

(iii) *The random field* $\boldsymbol{e}$ *is* α*-mixing with mixing coefficients* $\alpha(a;b) \le \alpha_1(a)\varpi_1(b)$ *such that as* $n \to \infty$,

$$A_n^{(1)} \left(\alpha_1^{1-2/q}(\underline{A}_{n2}) + \sum_{k=\underline{A}_{n1}}^{\infty} k^{d-1} \alpha_1^{1-2/q}(k) \right) \varpi_1^{1-2/q}(A_n^{(1)}) \to 0,$$

where $q = \min\{q_1, q_2\}$, $A_n^{(1)} = \prod_{j=1}^d A_{n1,j}$, $\underline{A}_{n1} = \min_{1\le j\le d} A_{n1,j}$, *and* $\underline{A}_{n2} = \min_{1\le j\le d} A_{n2,j}$. *Here,* $\{A_{n1,j}\}_{n\ge 1}$ *and* $\{A_{n2,j}\}_{n\ge 1}$ *are sequences of constants such that* $\min\left\{A_{n2,j}, \frac{A_{n1,j}}{A_{n2,j}}, \frac{A_{n,j}h_j}{A_{n1,j}}\right\} \to \infty$ *as* $n \to \infty$, *and* q_1 *is the integer that appear in Assumption 2.1.*

(iv) $\{\boldsymbol{X}_{1,\ell_1}\}_{\ell_1=1}^{n_1}$, $\{\boldsymbol{X}_{2,\ell_2}\}_{\ell_2=1}^{n_2}$, $\boldsymbol{e}$, $\{\varepsilon_{1,\ell_1}\}_{\ell_1=1}^{n_1}$, *and* $\{\varepsilon_{2,\ell_2}\}_{\ell_2=1}^{n_2}$ *are mutually independent.*

In Section 6, we give examples of bivariate random fields that satisfies Assumptions 4.1 and 4.2. We note that a wide class of bivariate Lévy-driven MA random fields satisfies our assumptions.

We are interested in testing the null hypothesis

$$\mathbb{H}_{0,j_1\dots j_L} : \partial_{j_1\dots j_L} m_1(\mathbf{0}) - \partial_{j_1\dots j_L} m_2(\mathbf{0}) = 0 \tag{4.2}$$

against the alternative $\mathbb{H}_{1,j_1\dots j_L} : \partial_{j_1\dots j_L} m_1(\mathbf{0}) - \partial_{j_1\dots j_L} m_2(\mathbf{0}) \neq 0$.

Define $\boldsymbol{M}_k(\mathbf{0})$ as $\boldsymbol{M}(\mathbf{0})$ with $m = m_k$ and $\overline{\boldsymbol{\beta}}_k(\mathbf{0})$ as LP estimators of order p for $\boldsymbol{M}_k(\mathbf{0})$ computed by using $\{(Y_k(\boldsymbol{x}_{k,\ell_k}), \boldsymbol{x}_{k,\ell_k})\}$, bandwidths $h_1, \dots, h_d$, and a common kernel function K, $k = 1, 2$, respectively. The next theorem is a building block of the two-sample test (4.2).

Proposition 4.2. *Suppose Assumptions 2.1, 2.2 (i), 3.1, 4.1, and 4.2 hold with* $m = m_k$, $\eta = \eta_k$, $\sigma_\varepsilon = \sigma_{\varepsilon,k}$, $\{\varepsilon_j\} = \{\varepsilon_{k,\ell_k}\}$, $g = g_k$, $k = 1, 2$. *Moreover, assume that* $n = n_1$, $n_1/n_2 \to \theta \in (0,\infty)$ *as* $n_1 \to \infty$ *and* $(\eta_1(\mathbf{0}), -\eta_2(\mathbf{0})) \left(\int \Sigma_{\boldsymbol{e}}(\boldsymbol{v})d\boldsymbol{v}\right) (\eta_1(\mathbf{0}), -\eta_2(\mathbf{0}))' \ge 0$. *Then, as* $n \to \infty$,

$$\begin{aligned} &\sqrt{A_n h_1 \dots h_d} \left\{ H\left((\overline{\boldsymbol{\beta}}_1(\mathbf{0}) - \overline{\boldsymbol{\beta}}_2(\mathbf{0})) - (\boldsymbol{M}_1(\mathbf{0}) - \boldsymbol{M}_2(\mathbf{0})) \right) - (\overline{B}_{n1}(\mathbf{0}) - \overline{B}_{n2}(\mathbf{0})) \right\} \\ &\quad \stackrel{d}{\to} N\left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \left(\overline{V}_1(\mathbf{0}) + \overline{V}_2(\mathbf{0}) - 2\overline{V}_3(\mathbf{0})\right) S^{-1}\mathcal{K}S^{-1} \right), \end{aligned}$$

where

$$\begin{aligned} \overline{B}_{n1}(\mathbf{0}) &= S^{-1}B^{(d,p)}\boldsymbol{M}_{n1}^{(d,p)}(\mathbf{0}),\ \overline{B}_{n2}(\mathbf{0}) = S^{-1}B^{(d,p)}\boldsymbol{M}_{n2}^{(d,p)}(\mathbf{0}), \\ \overline{V}_1(\mathbf{0}) &= \left(\frac{\kappa(\eta_1^2(\mathbf{0}) + \sigma_{\varepsilon,1}^2(\mathbf{0}))}{g_1(\mathbf{0})} + \eta_1^2(\mathbf{0}) \int \sigma_{\boldsymbol{e},11}(\boldsymbol{v})d\boldsymbol{v} \right), \\ \overline{V}_2(\mathbf{0}) &= \left(\frac{\theta\kappa(\eta_2^2(\mathbf{0}) + \sigma_{\varepsilon,2}^2(\mathbf{0}))}{g_2(\mathbf{0})} + \eta_2^2(\mathbf{0}) \int \sigma_{\boldsymbol{e},22}(\boldsymbol{v})d\boldsymbol{v} \right), \\ \overline{V}_3(\mathbf{0}) &= \eta_1(\mathbf{0})\eta_2(\mathbf{0}) \int \sigma_{\boldsymbol{e},12}(\boldsymbol{v})d\boldsymbol{v}, \end{aligned}$$

where $\boldsymbol{M}_{nk}^{(d,p)}(\mathbf{0})$ *are defined as* $\boldsymbol{M}_n^{(d,p)}(\mathbf{0})$ *with* $m = m_k$.

An estimator of the asymptotic variance of the statistics $\overline{\boldsymbol{\beta}}_1(\mathbf{0}) - \overline{\boldsymbol{\beta}}_2(\mathbf{0})$ can be constructed as follows. For $\boldsymbol{z} \in (-1/2, 1/2)^d$, let $\widehat{m}_{k,-I_k}(\boldsymbol{z})$ be the LP estimator (of order p) of $m_k(\boldsymbol{z})$ computed without $\{(Y_k(\boldsymbol{X}_{k,\ell_k}), \boldsymbol{X}_{k,\ell_k})\}_{\ell_k\in I_k}$, $I_k \subset \{1, \dots, n_k\}$, $k = 1, 2$.

Define

$$
\begin{aligned}
\overline{g}_{n_k}(\mathbf{0}) &= \frac{1}{n_k h_1 \dots h_d} \sum_{\ell_k=1}^{n_k} K_{Ah}(\boldsymbol{X}_{k,\ell_k}), \\
\overline{V}_{n,1k}(\mathbf{0}) &= \frac{1}{n_k h_1 \dots h_d} \sum_{\ell_k=1}^{n_k} K_{Ah}(\boldsymbol{X}_{k,\ell_k}) \left(Y_k(\boldsymbol{X}_{k,\ell_k}) - \widehat{m}_{k,-\{\ell_k\}}(\boldsymbol{X}_{k,\ell_k}/A_n)\right)^2, \ k = 1, 2, \\
\overline{V}_{n,2k}(\mathbf{0}) &= \frac{A_n}{n_k h_1 \dots h_d} \sum_{\ell_k=1}^{n_k - 1} K_{Ah}(\boldsymbol{X}_{k,\ell_k}) K_{Ah}(\boldsymbol{X}_{k,\ell_k+1}) \\
&\quad \times \left(Y_k(\boldsymbol{X}_{k,\ell_k}) - \widehat{m}_{k,-\{\ell_k,\ell_k+1\}}(\boldsymbol{X}_{k,\ell_k}/A_n)\right) \\
&\quad \times \left(Y_k(\boldsymbol{X}_{k,\ell_k+1}) - \widehat{m}_{k,-\{\ell_k,\ell_k+1\}}(\boldsymbol{X}_{k,\ell_k+1}/A_n)\right), \ k = 1, 2, \\
\overline{V}_{n,3}(\mathbf{0}) &= \frac{A_n}{n_1 n_2 h_1 \dots h_d} \sum_{\ell_1=1}^{n_1} \sum_{\ell_2=1}^{n_2} K_{Ah}(\boldsymbol{X}_{1,\ell_1}) K_{Ah}(\boldsymbol{X}_{2,\ell_2}) \\
&\quad \times \left(Y_1(\boldsymbol{X}_{1,\ell_1}) - \widehat{m}_{1,-\{\ell_1\}}(\boldsymbol{X}_{1,\ell_1}/A_n)\right) \left(Y_2(\boldsymbol{X}_{2,\ell_2}) - \widehat{m}_{2,-\{\ell_2\}}(\boldsymbol{X}_{2,\ell_2}/A_n)\right).
\end{aligned}
$$

Proposition 4.3. *Under the assumptions of Theorem 4.2, as* $n \to \infty$,

$$
\begin{aligned}
\check{V}_n(\mathbf{0}) &:= \left\{ \frac{(A_n/n_1)\overline{V}_{n,11}(\mathbf{0}) + (\widehat{V}_{n,21}(\mathbf{0})/\kappa_0^{(2)})}{\overline{g}_{n_1}^2(\mathbf{0})} \right\} + \left\{ \frac{(A_n/n_2)\overline{V}_{n,12}(\mathbf{0}) + (\widehat{V}_{n,22}(\mathbf{0})/\kappa_0^{(2)})}{\overline{g}_{n_2}^2(\mathbf{0})} \right\} \\
&\quad - 2\frac{(\overline{V}_{n,3}(\mathbf{0})/\kappa_0^{(2)})}{\overline{g}_{n_1}(\mathbf{0})\overline{g}_{n_2}(\mathbf{0})} \stackrel{p}{\to} \overline{V}_1(\mathbf{0}) + \overline{V}_2(\mathbf{0}) - 2\overline{V}_3(\mathbf{0}).
\end{aligned}
$$

Define the test statistics

$$
T_{n,j_1\dots j_L} := \frac{\sqrt{A_n h_1 \dots h_d \left(\prod_{\ell=1}^L h_{j_\ell}\right)^2} \left(\partial_{j_1\dots j_L}\widehat{m}_1(\mathbf{0}) - \partial_{j_1\dots j_L}\widehat{m}_2(\mathbf{0})\right)}{\sqrt{\overline{V}_n(\mathbf{0}) \left(\boldsymbol{s}_{j_1\dots j_L}!\right)^2 \left(e'_{j_1\dots j_L} S^{-1}\mathcal{K}S^{-1}e_{j_1\dots j_L}\right)}}.
$$

The asymptotic properties of the test statistics under both null and alternative hypotheses are given as follows:

Corollary 4.2. *Let* $\tau \in (0, 1/2)$. *Under the assumptions of Theorem 4.2 with*

$$
A_n h_1 \dots h_d \left((S^{-1}e_{j_1\dots j_L})' B^{(d,p)} M_{n1}^{(d,p)}(\mathbf{0})\right)^2 \to 0, \ n \to \infty.
$$

Then, $\lim_{n\to\infty} P(|T_{n,j_1\dots j_L}| \geq q_{1-\tau/2}) = \tau$ *under* $\mathbb{H}_{0,j_1\dots j_L}$ *and* $\lim_{n\to\infty} P(|T_{n,j_1\dots j_L}| \geq q_{1-\tau/2}) = 1$ *under* $\mathbb{H}_{1,j_1\dots j_L}$, *where* $q_{1-\tau}$ *is the* $(1-\tau)$*-quantile of the standard normal random variable.*

5. Uniform convergence of local polynomial estimators

In this section, we consider general kernel estimators and derive their uniform convergence rates (Section 5.1). Building on the results, we derive the uniform convergence rates of the LP estimators for the mean function of the model (2.1) (Section 5.2).

5.1. **Uniform convergence rates for general kernel estimators.** For $j = 1, 2, 3$, let $f_j : \mathbb{R}^d \to \mathbb{R}$ be functions such that f_j is continuous on $R_{0,\delta} := (-1/2 - \delta, 1/2 + \delta)^d$ for some $\delta > 0$. Define

$$\begin{aligned}\widehat{\Psi}_{\mathrm{I}}(\boldsymbol{z}) &= \frac{1}{n^2 A_n^{-1} h_1 \dots h_d} \sum_{i=1}^{n} K_{Ah}(\boldsymbol{X}_i - A_n \boldsymbol{z}) \\ &\quad \times f_{1,Ah}\left(\boldsymbol{X}_i - A_n \boldsymbol{z}\right) f_{2,A}\left(\boldsymbol{X}_i - A_n \boldsymbol{z}\right) f_{3,A}\left(\boldsymbol{X}_i\right) Z_{\boldsymbol{X}_i}, \qquad (5.1)\\ \widehat{\Psi}_{\mathrm{II}}(\boldsymbol{z}) &= \frac{1}{n h_1 \dots h_d} \sum_{i=1}^{n} K_{Ah}(\boldsymbol{X}_i - A_n \boldsymbol{z}) \\ &\quad \times f_{1,Ah}\left(\boldsymbol{X}_i - A_n \boldsymbol{z}\right) f_{2,A}\left(\boldsymbol{X}_i - A_n \boldsymbol{z}\right) f_{3,A}\left(\boldsymbol{X}_i\right), \qquad (5.2)\end{aligned}$$

where $f_{j,Aa}(\boldsymbol{x}) = f_j\left(\frac{x_1}{A_{n,1}a_1}, \dots, \frac{x_d}{A_{n,d}a_d}\right)$ for $\boldsymbol{a} = (a_1, \dots a_d)' \in (0,\infty)^d$ and $\{Z_{\boldsymbol{X}_i}\}_{i=1}^n$ is a sequence of real-valued random variables. Many kernel estimators, such as kernel density, Nadaraya-Watson, and LP estimators, can be represented by combining special cases of estimators (5.1) or (5.2). In this study, we use the uniform convergence rates of these estimators with

$$\begin{aligned} &f_1 \in \left\{ e'_{j_1 \dots j_L} \begin{pmatrix} 1 \\ \check{\boldsymbol{x}} \end{pmatrix}, e'_{j_{1,1} \dots j_{1,L_1}} \begin{pmatrix} 1 \\ \check{\boldsymbol{x}} \end{pmatrix} (1\ \check{\boldsymbol{x}}') e_{j_{2,1} \dots j_{2,L_2}} \right\}, \\ &f_2 \in \left\{ 1, \prod_{\ell=1}^{L} x_{j_\ell} \right\},\ f_3 \in \left\{1, \eta, \sigma_\varepsilon, \{\partial_{j_1 \dots j_{p+1}} m\}_{1 \le j_1 \le \dots \le j_{p+1} \le d}\right\},\ Z_{\boldsymbol{X}_i} \in \{e(\boldsymbol{X}_i), \varepsilon_i\}. \end{aligned}$$

We assume the following conditions for the sampling sites $\{\boldsymbol{X}_i\}_{i=1}^n$:

Assumption 5.1. *Let g be a probability density function with support $R_0 = [-1/2, 1/2]^d$.*

(i) $A_n/n \to \kappa \in [0, \infty)$ *as* $n \to \infty$,

(ii) $\{\boldsymbol{X}_i = (X_{i,1}, \dots, X_{i,d})'\}_{i=1}^n$ *is a sequence of i.i.d. random vectors with density* $A_n^{-d} g(\cdot/A_n)$ *and g is continuous and positive on R_0.*

(iii) $\{\boldsymbol{X}_i\}_{i=1}^n$ *and* $\{Z_{\boldsymbol{x}} : \boldsymbol{x} \in \mathbb{R}^d\}$ *are independent.*

We also assume the following conditions on the bandwidth h_j, the random field $\{Z_{\boldsymbol{x}} : \boldsymbol{x} \in \mathbb{R}^d\}$, and functions f_j:

Assumption 5.2. *For $j = 1, \dots, d$, let $\{A_{n1,j}\}_{n \ge 1}$, $\{A_{n2,j}\}_{n \ge 1}$ be sequence of positive numbers.*

(i) The random field $\{Z_{\boldsymbol{x}} : \boldsymbol{x} \in \mathbb{R}^d\}$ *is stationary and* $E[|Z_{\boldsymbol{0}}|^{q_2}] < \infty$ *for some integer* $q_2 > 4$.

(ii) Define $\sigma_{\boldsymbol{Z}}(\boldsymbol{x}) = E[Z_{\boldsymbol{0}} Z_{\boldsymbol{x}}]$. *Assume that* $\int_{\mathbb{R}^d} |\sigma_{\boldsymbol{Z}}(\boldsymbol{v})| d\boldsymbol{v} < \infty$.

(iii) $\min\left\{A_{n2,j}, \frac{A_{n1,j}}{A_{n2,j}}, \frac{A_{n,j}h_j}{A_{n1,j}}\right\} \to \infty$ *as* $n \to \infty$.

(iv) The random field $\{Z_{\boldsymbol{x}} : \boldsymbol{x} \in \mathbb{R}^d\}$ *is β-mixing with mixing coefficients* $\beta(a; b) \le \beta_1(a)\varpi_2(b)$ *such that as* $n \to \infty$, $h_j \to 0$, $1 \le j \le d$,

$$\sup_{\boldsymbol{v} \in R_{0,\delta}} \left| \frac{f_2(h_1 v_1, \dots, h_d v_d)}{f_2(h_1, \dots, h_d)} \right| \in (c_{f_2}, C_{f_2}) \text{ for some } 0 < c_{f_2} < C_{f_2} < \infty, \qquad (5.3)$$

$$\frac{A_n^{(1)}}{(\overline{A}_{n1})^d} \sim 1,\ \frac{A_n^{\frac{1}{2}} (h_1 \dots h_d)^{\frac{1}{2}}}{n^{1/q_2} (\overline{A}_{n1})^d (\log n)^{\frac{1}{2}+\iota}} \gtrsim 1 \text{ for some } \iota \in (0, \infty), \qquad (5.4)$$

$$\sqrt{\frac{n^2 A_n h_1 \dots h_d}{(A_n^{(1)})^2 \log n}} \beta_1(\underline{A}_{n2}) \varpi_2(A_n h_1 \dots h_d) \to 0, \qquad (5.5)$$

where

$$A_n^{(1)} = \prod_{j=1}^{d} A_{n1,j}, \ \overline{A}_{n1} = \max_{1 \le j \le d} A_{n1,j}, \ \underline{A}_{n1} = \min_{1 \le j \le d} A_{n1,j},$$
$$\overline{A}_{n2} = \max_{1 \le j \le d} A_{n2,j}, \ \underline{A}_{n2} = \min_{1 \le j \le d} A_{n2,j}.$$

(iv) $f_1 : \mathbb{R}^d \to \mathbb{R}$ *is Lipschitz continuous on* $\mathbb{R}^d$*, i.e.,* $|f_1(\boldsymbol{v}_1) - f_1(\boldsymbol{v}_2)| \le L_{f_1}|\boldsymbol{v}_1 - \boldsymbol{v}_2|$ *for some* $L_{f_1} \in (0, \infty)$ *and all* $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{R}^d$*, and* f_2 *and* f_3 *are continuous on* $R_{0,\delta}$*.*

When $Z_{\boldsymbol{X}_i} = \varepsilon_i$, we interpret $\{Z_{\boldsymbol{x}} : \boldsymbol{x} \in \mathbb{R}^d\}$ as a set of i.i.d. random variable and in this case $\sigma_{\boldsymbol{Z}}(\boldsymbol{x}) = 0$ if $\boldsymbol{x} \neq 0$. Condition (5.5) is concerned with large-block-small-block argument for β-mixing sequences. To derive uniform convergence rates of kernel estimators, we need to care about the effect of non-equidistant sampling sites when applying a maximal inequality and it requires additional work compared with the case that sampling sites are equidistant. Indeed, in place of using results for (regularly spaced) stationary sequence, which cannot be applied to the analysis of irregularly spaced nonstationary data, we construct "exactly" independent blocks of observations and apply results for independent data to the independent blocks since there is no practical guidance for introducing an order to spatial points as opposed to time series. Precisely, we first reduce the dependent data to not asymptotically but exactly independent blocks in finite sample by extending the blocking technique in Yu (1994)(Corollary 2.7), which does not require regularly spaced sampling sites. Then apply a maximal inequality for independent and possibly not identically distributed random variables to the independent blocks. In Section 6, we will show that a wide class of Lévy-driven MA random fields satisfies our β-mixing conditions.

Remark 5.1 (Discussion on β-mixing conditions)**.** Lahiri (2003b) established central limit theorems for weighted sample means of bounded spatial data under α-mixing conditions. Lahiri's proof relies essentially on approximating the characteristic function of the weighted sample mean by that of independent blocks using the Volkonskii-Rozanov inequality (cf. Proposition 2.6 in Fan and Yao (2003)) and then showing that the characteristic function corresponding to the independent blocks converges to the characteristic function of its Gaussian limit. However, characteristic functions are difficult to capture the uniform behavior of the LP estimators over compact sets so we rely on a different argument than that of Lahiri (2003b). Indeed, we use a stronger blocking argument tailored to β-mixing sequences; cf. Lemma 4.1 in Yu (1994). Further, we cannot apply other techniques for dependent data such as m-dependent approximation under a physical dependence structure (cf. El Machkouri et al. (2013)) since the technique is designed for regularly spaced random fields on $\mathbb{Z}^d$. We also note that it is not known that the results corresponding to Corollary 2.7 in Yu (1994) hold for α-mixing sequences; see Remark (ii) right after the proof of Lemma 4.1 in Yu (1994).

We assume the following conditions on the kernel function K:

Assumption 5.3. *Let* $K : \mathbb{R}^d \to \mathbb{R}$ *be a kernel function such that*

(i) $\int K(\boldsymbol{z})d\boldsymbol{z} = 1$.

(ii) The kernel function K *is bounded and supported on* $[-C_K, C_K]^d \subset [-1/2, 1/2]^d$ *for some* $C_K > 0$*. Moreover,* K *is Lipschitz continuous on* $\mathbb{R}^d$*, i.e.,* $|K(\boldsymbol{v}_1) - K(\boldsymbol{v}_2)| \le L_K|\boldsymbol{v}_1 - \boldsymbol{v}_2|$ *for some* $L_K \in (0, \infty)$ *and all* $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{R}^d$*.*

(iii) Define $\kappa_0^{(r)} := \int K^r(\boldsymbol{z})d\boldsymbol{z}$, $\kappa_{j_1,\dots,j_M}^{(r)} := \int \prod_{\ell=1}^M z_{j_\ell} K^r(\boldsymbol{z})d\boldsymbol{z}$, *and*

$$\check{\boldsymbol{z}} := (1, (\boldsymbol{z})_1', \dots, (\boldsymbol{z})_p')',\ (\boldsymbol{z})_L = \left(\prod_{\ell=1}^L z_{j_\ell}\right)'_{1\le j_1\le\cdots\le j_L\le d},\ 1 \le L \le p.$$

The matrix $S = \int \begin{pmatrix} 1 \\ \check{\boldsymbol{z}} \end{pmatrix} (1\ \check{\boldsymbol{z}}')K(\boldsymbol{z})d\boldsymbol{z}$ *is non-singular.*

The next result provides uniform convergence rates of $\widehat{\Psi}_{\mathrm{I}}$ and $\widehat{\Psi}_{\mathrm{II}}$.

Proposition 5.1. *Suppose that Assumptions 5.1, 5.2, and 5.3 hold. Then as* $n \to \infty$, *we have*

$$\sup_{\boldsymbol{z}\in[-1/2,1/2]^d} \left|\widehat{\Psi}_{\mathrm{I}}(\boldsymbol{z}) - E[\widehat{\Psi}_{\mathrm{I}}(\boldsymbol{z})]\right| = O_p\left(|f_2(h_1,\dots,h_d)|\sqrt{\frac{\log n}{n^2 A_n^{-1} h_1\dots h_d}}\right), \tag{5.6}$$

$$\sup_{\boldsymbol{z}\in[-1/2,1/2]^d} \left|\widehat{\Psi}_{\mathrm{II}}(\boldsymbol{z}) - E[\widehat{\Psi}_{\mathrm{II}}(\boldsymbol{z})]\right| = O_p\left(|f_2(h_1,\dots,h_d)|\sqrt{\frac{\log n}{n h_1\dots h_d}}\right). \tag{5.7}$$

5.2. **Uniform estimation of the derivatives of the mean function.** In this section, we provide uniform convergence rates of the LP estimators. We assume the following condition on the mean function m, the variance function η, and $\{\varepsilon_{n,j}\}$:

Assumption 5.4. *Recall* $R_0 = [-1/2, 1/2]^d$.

(i) The mean function m *is* $(p+1)$*-times continuously partial differentiable on* R_0 *and define* $\partial_{j_1\dots j_L} m(\boldsymbol{z}) := \partial m(\boldsymbol{z})/\partial z_{j_1}\dots z_{j_L}$, $1 \le j_1,\dots,j_L \le d$, $0 \le L \le p+1$. *When* $L = 0$, *we set* $\partial_{j_1\dots j_L} m(\boldsymbol{z}) = \partial_{j_0} m(\boldsymbol{z}) = m(\boldsymbol{z})$.

(ii) The function η *is continuous over* R_0 *and* $\inf_{\boldsymbol{z}\in R_0}\eta(\boldsymbol{z}) > 0$.

(iii) The sequence of random variables $\{\varepsilon_j\}_{j=1}^n$ *are i.i.d. with* $E[\varepsilon_1] = 0$, $E[\varepsilon_1^2] = 1$, $E[|\varepsilon_1|^{q_1}] < \infty$ *for some integer* $q_1 > 4$, *and the function* $\sigma_\varepsilon(\cdot)$ *is continuous over* R_0 *and* $\inf_{\boldsymbol{z}\in R_0}\sigma_\varepsilon(\boldsymbol{z}) > 0$.

The next result provides uniform convergence rates of LP estimators $\partial_{j_1\dots j_L}\widehat{m}(\boldsymbol{z})$.

Theorem 5.1. *Define* $\mathrm{T}_n = \prod_{j=1}^d[-1/2 + C_K h_j, 1/2 - C_K h_j]$. *Suppose that Assumptions 5.1, 5.2 (i), (ii), (iii), (5.4), (5.5), 5.3, and 5.4 hold with (with* $Z_{\boldsymbol{x}} = e(\boldsymbol{x})$ *and* $q_1 \ge q_2$*). Moreover, assume that* $\frac{\log n}{nh_1\dots h_d} \to 0$ *as* $n \to \infty$. *Then for* $1 \le j_1 \le \cdots \le j_L \le d$, $0 \le L \le p$, *as* $n \to \infty$, *we have*

$$\sup_{\boldsymbol{z}\in\mathrm{T}_n} |\partial_{j_1\dots j_L}\widehat{m}(\boldsymbol{z}) - \partial_{j_1\dots j_L}\widehat{m}(\boldsymbol{z})|$$
$$= O_p\left(\frac{\sum_{1\le j_1\le\cdots\le j_{p+1}\le d}\prod_{\ell=1}^{p+1} h_{j_\ell}}{\prod_{\ell=1}^L h_{j_\ell}} + \sqrt{\frac{\log n}{A_n h_1\dots h_d\left(\prod_{\ell=1}^L h_{j_\ell}\right)^2}}\right).$$

6. Examples

In this section, we discuss examples of random fields to which our theoretical results can be applied. To this end, we consider Lévy-driven moving average (MA) random fields and discuss their dependence structure. Lévy-driven MA random fields include many Gaussian and non-Gaussian random fields and constitute a flexible class of models for spatial data. We refer to Bertoin (1996) and Sato (1999) for standard references on Lévy processes, and Rajput and Rosinski (1989) and Kurisu (2022) for details on the theory of infinitely divisible measures and fields. In particular, we show that a broad class of Lévy-driven MA random fields, which includes continuous autoregressive and moving average (CARMA) random fields as special cases (cf. Brockwell and Matsuda (2017)), satisfies our assumptions.

Let $\boldsymbol{L} = \{\boldsymbol{L}(A) = (L_1(A), L_2(A))' : A \in \mathcal{B}(\mathbb{R}^d)\}$ be an $\mathbb{R}^2$-valued random measure on the Borel subsets $\mathcal{B}(\mathbb{R}^d)$ that satisfies the following conditions:

1. For each sequence $\{A_m\}_{m\geq 1}$ of disjoint sets in $\mathbb{R}^d$,
   (a) $\boldsymbol{L}(\cup_{m\geq 1} A_m) = \sum_{m\geq 1} \boldsymbol{L}(A_m)$ a.s. whenever $\cup_{m\geq 1} A_m \in \mathcal{B}(\mathbb{R}^d)$,
   (b) $\{\boldsymbol{L}(A_m)\}_{m\geq 1}$ is a sequence of independent random variables.
2. For every Borel subset A of $\mathbb{R}^d$ with finite Lebesgue measure $|A|$, $\boldsymbol{L}(A)$ has an infinitely divisible distribution, that is,
$$E[\exp(\mathrm{i}\boldsymbol{\theta}'\boldsymbol{L}(A))] = \exp(|A|\psi(\boldsymbol{\theta})),\ \boldsymbol{\theta} \in \mathbb{R}^2, \tag{6.1}$$
   where $\mathrm{i} = \sqrt{-1}$ and ψ is the logarithm of the characteristic function of an $\mathbb{R}^2$-valued infinitely divisible distribution, which is given by
$$\psi(\boldsymbol{\theta}) = \mathrm{i}\boldsymbol{\theta}'\boldsymbol{\gamma}_0 - \frac{1}{2}\boldsymbol{\theta}'\Sigma_0\boldsymbol{\theta} + \int_{\mathbb{R}^2} \left\{e^{\mathrm{i}\boldsymbol{\theta}'\boldsymbol{x}} - 1 - \mathrm{i}\boldsymbol{\theta}'\boldsymbol{x}1_{\{\|\boldsymbol{x}\|\leq 1\}}\right\} \nu_0(d\boldsymbol{x}),$$
   where $\boldsymbol{\gamma}_0 = (\gamma_{0,1}, \gamma_{0,2})' \in \mathbb{R}^2$, $\Sigma_0 = (\sigma_{0,jk})_{1\leq j,k\leq 2}$ is a 2×2 positive semi-definite matrix, and ν_0 is a Lévy measure with $\int_{\mathbb{R}^2} \min\{1, \|\boldsymbol{x}\|^2\}\nu_0(d\boldsymbol{x}) < \infty$. If $\nu_0(d\boldsymbol{x})$ has a Lebesgue density, i.e., $\nu_0(d\boldsymbol{x}) = \nu_0(\boldsymbol{x})d\boldsymbol{x}$, we call $\nu_0(\boldsymbol{x})$ as the Lévy density. The triplet $(\boldsymbol{\gamma}_0, \Sigma_0, \nu_0)$ is called the Lévy characteristic of $\boldsymbol{L}$ and uniquely determines the distribution of $\boldsymbol{L}$.

By equation (6.1), the first and second moments of the random measure L are determined by
$$E[L_j(A)] = \mu_j^{(\boldsymbol{L})}|A|,\ \mathrm{Cov}(L_j(A), L_k(A)) = \sigma_{j,k}^{(\boldsymbol{L})}|A|,$$
where $\mu_j^{(\boldsymbol{L})} = -\mathrm{i}\frac{\partial\psi(\boldsymbol{0})}{\partial\theta_j}$ and $\sigma_{j,k}^{(\boldsymbol{L})} = -\frac{\partial^2\psi(\boldsymbol{0})}{\partial\theta_j\partial\theta_k}$.

The following are a couple of examples of Lévy random measures.

- If $\psi(\boldsymbol{\theta}) = -\boldsymbol{\theta}'\Sigma_0^2\boldsymbol{\theta}/2$ with a 2×2 positive semi-definite matrix Σ_0, then $\boldsymbol{L}$ is a Gaussian random measure.
- If $\psi(\boldsymbol{\theta}) = \lambda\int_{\mathbb{R}^2}(\exp(\mathrm{i}\boldsymbol{\theta}'\boldsymbol{x}) - 1)F(d\boldsymbol{x})$, where $\lambda > 0$ and F is a probability distribution function with no jump at the origin, then $\boldsymbol{L}$ is a compound Poisson random measure with intensity λ and jump size distribution F. More specifically,
$$\boldsymbol{L}(A) = \sum_{i=1}^{\infty} \boldsymbol{J}_i 1_{\{\boldsymbol{s}_i\}}(A),\ A \in \mathcal{B}(\mathbb{R}^d),$$
  where $\boldsymbol{s}_i$ denotes the location of the ith unit point mass of a Poisson random measure on $\mathbb{R}^d$ with intensity $\lambda > 0$ and $\{\boldsymbol{J}_i\}$ is a sequence of i.i.d. random vectors in $\mathbb{R}^2$ with distribution function F independent of $\{\boldsymbol{s}_i\}$.

Let $\boldsymbol{\phi} = (\phi_{j,k})_{1\leq j,k\leq 2}$ be a measurable function on $\mathbb{R}^d$ with $\phi_{j,k} \in L^1(\mathbb{R}^d)\cap L^\infty(\mathbb{R}^d)$. A bivariate Lévy-driven MA random field with kernel $\boldsymbol{\phi}$ driven by a Lévy random measure $\boldsymbol{L}$ is defined by
$$\boldsymbol{e}(\boldsymbol{x}) = \int_{\mathbb{R}^d} \boldsymbol{\phi}(\boldsymbol{x}-\boldsymbol{u})\boldsymbol{L}(d\boldsymbol{u}),\ \boldsymbol{x} \in \mathbb{R}^d. \tag{6.2}$$

Define $\boldsymbol{\mu}_{\boldsymbol{L}} = (\mu_1^{(\boldsymbol{L})}, \mu_2^{(\boldsymbol{L})})'$ and $\Sigma_{\boldsymbol{L}} = (\sigma_{j,k}^{(\boldsymbol{L})})_{1\leq j,k\leq 2}$. The first and second moments of $\boldsymbol{e}(\boldsymbol{x})$ satisfy
$$E[\boldsymbol{e}(\boldsymbol{0})] = \boldsymbol{\mu}_{(L)}\int_{\mathbb{R}^d} \boldsymbol{\phi}(\boldsymbol{u})d\boldsymbol{u},\ \mathrm{Cov}(\boldsymbol{e}(\boldsymbol{0}), \boldsymbol{e}(\boldsymbol{x})) = \int_{\mathbb{R}^d} \boldsymbol{\phi}(\boldsymbol{x}-\boldsymbol{u})\Sigma_{\boldsymbol{L}}\boldsymbol{\phi}(\boldsymbol{u})d\boldsymbol{u}.$$

We refer to Brockwell and Matsuda (2017) for more details on the computation of moments of Lévy-driven MA processes.

Before discussing theoretical results, we look at some examples of univariate random fields defined by (6.2). Let $a_*(z) = z^{p_0} + a_1 z^{p_0-1} + \cdots + a_{p_0} = \prod_{i=1}^{p_0}(z - \lambda_i)$ be a polynomial of degree p_0 with real coefficients and distinct negative zeros $\lambda_1, \dots, \lambda_{p_0}$, and let $b_*(z) = b_0 + b_1 z + \cdots + b_{q_0} z^{q_0} = \prod_{i=1}^{q_0}(z - \xi_i)$ be a polynomial of degree q_0 with real coefficients and real zeros $\xi_1, \dots, \xi_{q_0}$ such that $b_{q_0} = 1$ and $0 \le q_0 < p_0$ and $\lambda_i^2 \neq \xi_j^2$ for all i and j. Define $a(z) = \prod_{i=1}^{p_0}(z^2 - \lambda_i^2)$ and $b(z) = \prod_{i=1}^{q_0}(z^2 - \xi_i^2)$. Then, the Lévy-driven MA random field driven by an infinitely divisible random measure L with

$$\phi(\boldsymbol{x}) = \sum_{i=1}^{p_0} \frac{b(\lambda_i)}{a'(\lambda_i)} e^{\lambda_i \|\boldsymbol{x}\|},$$

where a' denotes the derivative of the polynomial a, is called a univariate (isotropic) CARMA(p_0, q_0) random field. For example, if the Lévy random measure of a CARMA random field is compound Poisson, then the resulting random field is called a compound Poisson-driven CARMA random field. In particular, when

$$\phi(\boldsymbol{x}) = (1 - \varsigma)\exp(\lambda_1 \|\boldsymbol{x}\|) + \varsigma \exp(\lambda_2 \|\boldsymbol{x}\|),$$

where ς is a parameter that satisfies

$$-\frac{\lambda_2^2 - \xi^2 \lambda_1}{\lambda_1^2 - \xi^2 \lambda_2} = \frac{\varsigma}{1 - \varsigma}, \ \lambda_1 < \lambda_2 < 0, \ \xi \le 0,$$

then the random field (6.2) is called a CARMA$(2, 1)$ random field. This random field includes normalized CAR(1) (when $\varsigma = 0$) and CAR(2) (when $\varsigma = -\lambda_1/(\lambda_2 - \lambda_1)$) as special cases. See Brockwell and Matsuda (2017) for more details. We note that although we focus on isotropic case, it is possible to extend the results in this section to anisotropic Lévy-driven MA random fields.

Remark 6.1 (Connections to Matérn covariance functions)**.** In spatial statistics, Gaussian random fields with the following Matérn covariance functions play an important role (cf. Matérn, 1986; Stein, 1999; Guttorp and Gneiting, 2006):

$$M(\boldsymbol{x}; \nu, a, \sigma) = \sigma^2 \|a\boldsymbol{x}\|^\nu K_\nu(\|a\boldsymbol{x}\|), \ \nu > 0, a > 0, \sigma > 0,$$

where K_ν denotes the modified Bessel function of the second kind of order ν (we call ν the index of Matérn covariance function). Brockwell and Matsuda (2017) showed that in the univariate case, when the kernel function is $\phi(\boldsymbol{x}) = \|a\boldsymbol{x}\|^\nu K_\nu(\|a\boldsymbol{x}\|)$, which they call a Matérn kernel with index ν, then the Levy-driven MA random field has a Matérn covariance function with index $d/2 + \nu$. For example, a normalized CAR(1) random field has a Matérn covariance function since its kernel function is given by $\phi(\boldsymbol{x}) = \exp(-\|\lambda_1 \boldsymbol{x}\|) = \sqrt{(2/\pi)}\|\lambda_1 \boldsymbol{x}\|^{1/2} K_{1/2}(\|\lambda_1 \boldsymbol{x}\|)$ for some $\lambda_1 < 0$.

In general, if $\boldsymbol{\phi}$ depends only on $\|\boldsymbol{x}\|$, i.e., $\boldsymbol{\phi}(\boldsymbol{x}) = \boldsymbol{\phi}(\|\boldsymbol{x}\|)$, then $\boldsymbol{e}$ is a strictly stationary isotropic random field and the second moment of $\boldsymbol{e}(\boldsymbol{x})$ satisfies

$$\mathrm{Cov}(\boldsymbol{e}(\boldsymbol{0}), \boldsymbol{e}(\boldsymbol{x})) = \int_{\mathbb{R}^d} \boldsymbol{\phi}(\|\boldsymbol{x} - \boldsymbol{u}\|) \Sigma_{\boldsymbol{L}} \boldsymbol{\phi}(\|\boldsymbol{u}\|) d\boldsymbol{u}.$$

Consider the following decomposition:

$$\boldsymbol{e}(\boldsymbol{x}) = \int_{\mathbb{R}^d} \boldsymbol{\phi}(\boldsymbol{x} - \boldsymbol{u}) \psi_0 \left(\|\boldsymbol{x} - \boldsymbol{u}\| : m_n\right) \boldsymbol{L}(d\boldsymbol{u}) + \int_{\mathbb{R}^d} \boldsymbol{\phi}(\boldsymbol{x} - \boldsymbol{u}) \left(1 - \psi_0 \left(\|\boldsymbol{x} - \boldsymbol{u}\| : m_n\right)\right) \boldsymbol{L}(d\boldsymbol{u})$$

$$=: \boldsymbol{e}_{1,m_n}(\boldsymbol{x}) + \boldsymbol{e}_{2,m_n}(\boldsymbol{x}),$$

where m_n is a sequence of positive constants with $m_n \to \infty$ as $n \to \infty$ and $\psi_0(\cdot : c) : \mathbb{R} \to [0,1]$ is a truncation function defined by

$$\psi_0(x : c) = \begin{cases} 1 & \text{if } |x| \leq c/4, \\ -\frac{4}{c}\left(x - \frac{c}{2}\right) & \text{if } c/4 < |x| \leq c/2, \\ 0 & \text{if } x > c/2. \end{cases}$$

The random field $\boldsymbol{e}_{1,m_n} = \{\boldsymbol{e}_{1,m_n}(\boldsymbol{x}) = (e_{11,m_n}(\boldsymbol{x}), e_{12,m_n}(\boldsymbol{x}))' : \boldsymbol{x} \in \mathbb{R}^d\}$ is m_n-dependent (with respect to the ℓ^2-norm), i.e., $\boldsymbol{e}_{1,m_n}(\boldsymbol{x}_1)$ and $\boldsymbol{e}_{1,m_n}(\boldsymbol{x}_2)$ are independent if $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \geq m_n$. Also, if the tail of the kernel function $\boldsymbol{\phi}(\cdot)$ decays sufficiently fast, then the random field $\boldsymbol{e}_{2,m_n} = \{\boldsymbol{e}_{2,m_n}(\boldsymbol{x}) = (e_{21,m_n}(\boldsymbol{x}), e_{22,m_n}(\boldsymbol{x}))' : \boldsymbol{x} \in \mathbb{R}^d\}$ is asymptotically negligible. In such cases, we can approximate $\boldsymbol{e}$ by the m_n-dependent process $\boldsymbol{e}_{1,m_n}$ and verify conditions on mixing coefficients in Assymptions 2.3, 4.1, and 4.2 as shown in the following proposition.

Proposition 6.1. *Consider a Lévy-driven MA random field* $\boldsymbol{e}$ *defined by (6.2). Assume that* $\phi_{j,k}(\boldsymbol{x}) = r_{0,jk}e^{-r_{1,jk}\|\boldsymbol{x}\|}$ *where* $|r_{0,jk}| > 0$ *and* $r_{1,jk} > 0$, $j,k = 1,2$. *Additionally, assume that*

- *(a) the random measure* $\boldsymbol{L}(\cdot)$ *is Gaussian with triplet* $(0, \Sigma_0, 0)$ *or*
- *(b) the random measure* $\boldsymbol{L}(\cdot)$ *is non-Gaussian with triplet* $(\boldsymbol{\gamma}_0, 0, \nu_0)$, $\boldsymbol{\mu}_{(L)} = (0,0)'$, *and the marginal Lévy density* $\nu_{0,j}(x)$ *of* $L_j(\cdot)$ *is given by*

$$\nu_{0,j}(x) = \frac{1}{|x|^{1+\beta_{0,j}}}\left(C_{0,j}e^{-c_{0,j}|x|^{\alpha_{0,j}}} + \frac{C_{1,j}}{1+|x|^{\beta_{1,j}}}\right)1_{\mathbb{R}\backslash\{0\}}(x), \tag{6.3}$$

where $\alpha_{0,j} > 0$, $\beta_{0,j} \geq -1$, $\beta_{1,j} > 0$, $\beta_{0,j} + \beta_{1,j} > 6$, $c_{0,j} > 0$, $C_{0,j} \geq 0$, $C_{1,j} \geq 0$, *and* $C_{0,j} + C_{1,j} > 0$, $j = 1,2$.

Then $\boldsymbol{e}_{2,m_n}$ *is asymptotically negligible, that is, we can replace* $\boldsymbol{e}$ *with* $\boldsymbol{e}_{1,m_n}$ *in the results in Section 4. Further,* $\boldsymbol{e}_{1,m_n}$ *satisfies Assumptions 2.3, 4.1, and 4.2 with* $A_{n,j} \sim n^{\zeta_0/d}$, $A_{n1,j} = A_{n,j}^{\zeta_1}$, $A_{n2,j} = A_{n1,j}^{\zeta_2}$, $m_n = \underline{A}_{n2}^{1/2}$, *and* $h_j \sim n^{-\zeta_3/d}$ *where* ζ_0, ζ_1, ζ_2, *and* ζ_3 *are positive constants such that*

$$\begin{aligned}
&\zeta_0 \in \left(0, \frac{2p+2}{d}\right),\ \zeta_1 \in \left(\frac{\zeta_0 d}{d+2p+2}, \frac{2p+2}{d+2p+2}\right), \\
&\zeta_2 \in \left(0, \min\left\{\frac{2}{2+d\max\{1,\zeta_0\}}, 1 - \frac{\zeta_0 d}{\zeta_1(d+2p+2)}, \frac{2p+2}{\zeta_1(d+2p+2)} - 1\right\}\right), \\
&\zeta_3 \in \left(\frac{d\zeta_0}{2p+d+2}, \min\left\{\frac{d\zeta_0}{2p+d}, \zeta_0\left(1-\zeta_1(1+\zeta_2)\right), \zeta_1\left(1 - \left(1 + \frac{d}{2}\zeta_0\right)\zeta_2\right)\right\}\right).
\end{aligned}$$

Remark 6.2. When $d = 2$ and $p \geq 1$, the conditions on $\{\zeta_j\}_{j=0}^3$ are typically satisfied when $\zeta_0 = 1$, $\zeta_1 = \frac{3}{2p+4}$, $\zeta_2 \in \left(0, \frac{1}{6}\right)$. The Lévy density of the form (6.3) corresponds to a compound Poisson random measure if $\beta_{0,j} \in [-1, 0)$, a Variance Gamma random measure if $\alpha_{0,j} = 1$, $\beta_{0,j} = 0$, $C_{1,j} = 0$, and a tempered stable random measure if $\beta_{0,j} \in (0,1)$, $C_{1,j} = 0$ (cf. Section 5 in Kato and Kurisu (2020)). It is straight forward to extend Proposition 6.1 to the case that ϕ is a finite sum of kernel functions with exponential decay. Therefore, our results in Section 4 can be applied to a wide class of CARMA(p_0, q_0) random fields and extending the results to anisotropic CARMA random fields (cf. Brockwell and Matsuda (2017)) is straightforward.

The next result provides examples of Lévy-driven MA random fields that satisfies assumptions in Theorem 5.1.

Proposition 6.2. *Consider a univariate Lévy-driven MA random field* $\boldsymbol{e}$ *defined by (6.2). Assume that* $\phi(\boldsymbol{x}) = r_0 e^{-r_1\|\boldsymbol{x}\|}$ *where* $|r_0| > 0$ *and* $r_1 > 0$*. Additionally, assume Conditions (a) or (b) in Proposition 6.1. Then* $\boldsymbol{e}_{2,m_n}$ *is asymptotically negligible, that is, we can replace* $\boldsymbol{e}$ *with* $\boldsymbol{e}_{1,m_n}$ *in the results Theorem 5.1. Further,* $\boldsymbol{e}_{1,m_n}$ *satisfies Assumption 5.2 with* $A_{n,j} \sim n^{\zeta_0/d}$*,* $A_{n1,j} = A_{n,j}^{\zeta_1}$*,* $A_{n2,j} = A_{n1,j}^{\zeta_2}$*,* $m_n = \underline{A}_{n2}^{1/2}$*, and* $h_j \sim n^{-\zeta_3/d}$ *where* ζ_0*,* ζ_1*,* ζ_2*, and* ζ_3 *are positive constants such that* $\zeta_0 > \frac{2}{q_2}$*,* $\zeta_1 \in \left(0, \frac{1}{2} - \frac{1}{\zeta_0 q_2}\right)$*,* $\zeta_2 \in (0,1)$*, and* $\zeta_3 \in \left(0, \min\{1, \zeta_0(1-2\zeta_1) - \frac{2}{q_2}\}\right)$*.*

7. Conclusion

In this paper, we have advanced statistical theory of nonparametric regression for irregularly spaced spatial data. For this, we introduced a nonparametric regression model defined on a sampling region $R_n \subset \mathbb{R}^d$ and derived asymptotic normality and uniform convergence rates of the local polynomial estimators of order $p \geq 1$ for the mean function of the model under a stochastic sampling design. As an application of our main results, we discussed a two-sample test for the mean functions and their derivatives. We also provided examples of random fields that satisfy our assumptions. In particular, our assumptions hold for a wide class of random fields that includes Lévy-driven moving average random fields and popular Gaussian random fields as special cases.

Appendix A. Proofs for Section 4

A.1. **Proof of Theorem 4.1.**

Proof. Define $\boldsymbol{h} := (h_1, \dots, h_d)'$ and for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$, let $\boldsymbol{x} \circ \boldsymbol{y} = (x_1 y_1, \dots, x_d y_d)'$ be the Hadamard product. Considering Taylor's expansion of $m(\boldsymbol{z})$ around $\boldsymbol{z}$,

$$m(\boldsymbol{X}_i/A_n) = (1, \check{\boldsymbol{X}}_i')M(\boldsymbol{z}) + \frac{1}{(p+1)!} \sum_{1 \le j_1 \le \dots \le j_{p+1} \le d} \frac{(p+1)!}{\boldsymbol{s}_{j_1 \dots j_{p+1}}!} \partial_{j_1, \dots, j_{p+1}} m(\dot{\boldsymbol{X}}_i/A_n) \prod_{\ell=1}^{p+1} \frac{X_{i,j_\ell}}{A_{n,j_\ell}},$$

where $\dot{\boldsymbol{X}}_i = \boldsymbol{z} + \theta_i \boldsymbol{X}_i$ for some $\theta_i \in [0,1)$. Then we have

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}(\boldsymbol{0}) - \boldsymbol{M}(\boldsymbol{0}) &= (\boldsymbol{X}\boldsymbol{W}\boldsymbol{X}')^{-1}\boldsymbol{X}\boldsymbol{W}(\boldsymbol{Y} - \boldsymbol{X}'\boldsymbol{M}(\boldsymbol{0})) \\
&= \left[\sum_{i=1}^n K_{Ah}(\boldsymbol{X}_i) \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_i \end{pmatrix} (1 \ \check{\boldsymbol{X}}_i')\right]^{-1} \sum_{i=1}^n K_{Ah}(\boldsymbol{X}_i) \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_i \end{pmatrix} \\
&\quad \times \left(e_{n,i} + \varepsilon_{n,i} + \sum_{1 \le j_1 \le \dots \le j_{p+1} \le d} \frac{1}{\boldsymbol{s}_{j_1 \dots j_{p+1}}!} \partial_{j_1, \dots, j_{p+1}} m(\dot{\boldsymbol{X}}_i/A_n) \prod_{\ell=1}^{p+1} \frac{X_{i,j_\ell}}{A_{n,j_\ell}}\right).
\end{aligned}$$

This yields

$$\sqrt{A_n h_1 \dots h_d} H(\widehat{\boldsymbol{\beta}}(\boldsymbol{0}) - \boldsymbol{M}(\boldsymbol{0})) = S_n^{-1}(\boldsymbol{0})(V_n(\boldsymbol{0}) + B_n(\boldsymbol{0})),$$

where

$$\begin{aligned}
S_n(\boldsymbol{0}) &= \frac{1}{n h_1 \dots h_d} \sum_{i=1}^n K_{Ah}(\boldsymbol{X}_i) H^{-1} \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_i \end{pmatrix} (1 \ \check{\boldsymbol{X}}_i') H^{-1}, \\
V_n(\boldsymbol{0}) &= \frac{\sqrt{A_n h_1 \dots h_d}}{n h_1 \dots h_d} \sum_{i=1}^n K_{Ah}(\boldsymbol{X}_i) H^{-1} \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_i \end{pmatrix} (e_{n,i} + \varepsilon_{n,i}) \\
&=: (V_{n, j_1 \dots j_L}(\boldsymbol{0}))'_{1 \le j_1 \le \dots \le j_L \le d, 0 \le L \le p}, \\
B_n(\boldsymbol{0}) &= \frac{\sqrt{A_n h_1 \dots h_d}}{n h_1 \dots h_d} \sum_{i=1}^n K_{Ah}(\boldsymbol{X}_i) H^{-1} \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_i \end{pmatrix} \\
&\quad \times \sum_{1 \le j_1 \le \dots \le j_{p+1} \le d} \frac{1}{\boldsymbol{s}_{j_1 \dots j_{p+1}}!} \partial_{j_1, \dots, j_{p+1}} m(\dot{\boldsymbol{X}}_i/A_n) \prod_{\ell=1}^{p+1} \frac{X_{i,j_\ell}}{A_{n,j_\ell}} \\
&=: (B_{n, j_1 \dots j_L}(\dot{\boldsymbol{X}}))'_{1 \le j_1 \le \dots \le j_L \le d, 0 \le L \le p}.
\end{aligned}$$

(Step 1) Now we evaluate $S_n(\boldsymbol{0})$. By a change of variables and the dominated convergence theorem, we have

$$\begin{aligned}
E[S_n(\boldsymbol{0})] &= \frac{A_n^{-1}}{h_1 \dots h_d} \int K_{Ah}(\boldsymbol{x}) H^{-1} \begin{pmatrix} 1 \\ (\boldsymbol{x}\check{/}A_n) \end{pmatrix} (1 \ (\boldsymbol{x}\check{/}A_n)') H^{-1} g(\boldsymbol{x}/A_n) d\boldsymbol{x} \\
&= \frac{A_n^{-1}}{h_1 \dots h_d} A_n h_1 \dots h_d \int K(\boldsymbol{w}) \begin{pmatrix} 1 \\ \check{\boldsymbol{w}} \end{pmatrix} (1 \ \check{\boldsymbol{w}}') g(\boldsymbol{w} \circ \boldsymbol{h}) d\boldsymbol{w} \\
&= \left(g(\boldsymbol{0}) \int K(\boldsymbol{w}) \begin{pmatrix} 1 \\ \check{\boldsymbol{w}} \end{pmatrix} (1 \ \check{\boldsymbol{w}}') d\boldsymbol{w}\right)(1 + o(1)).
\end{aligned}$$

For $1 \le j_{1,1} \le \cdots \le j_{1,L_1} \le d$, $1 \le j_{2,1} \le \cdots \le j_{2,L_2} \le d, 0 \le L_1, L_2 \le p$, we define

$$I_{n,j_{1,1}\dots j_{1,L_1},j_{2,1}\dots j_{2,L_2}} := \frac{1}{nh_1 \dots h_d} \sum_{i=1}^{n} K_{Ah}(\boldsymbol{X}_i) \prod_{\ell_1=1}^{L_1} \left(\frac{X_{i,j_{1,\ell_1}}}{A_{n,j_{1,\ell_1}} h_{j_{1,\ell_1}}} \right) \prod_{\ell_2=1}^{L_2} \left(\frac{X_{i,j_{2,\ell_2}}}{A_{n,j_{2,\ell_2}} h_{j_{2,\ell_2}}} \right).$$

Then, by a change of variables and the dominated convergence theorem, we have

$$\begin{aligned}
&\mathrm{Var}(I_{n,j_{1,1}\dots j_{1,L_1},j_{2,1}\dots j_{2,L_2}}) \\
&= \frac{1}{n(h_1 \dots h_d)^2} \mathrm{Var}\left(K_{Ah}(\boldsymbol{X}_1) \prod_{\ell_1=1}^{L_1} \left(\frac{X_{i,j_{1,\ell_1}}}{A_{n,j_{1,\ell_1}} h_{j_{1,\ell_1}}} \right) \prod_{\ell_2=1}^{L_2} \left(\frac{X_{i,j_{2,\ell_2}}}{A_{n,j_{2,\ell_2}} h_{j_{2,\ell_2}}} \right) \right) \\
&= \frac{1}{nh_1 \dots h_d} \left\{ \int \prod_{\ell_1=1}^{L_1} z_{j_{1,\ell_1}}^2 \prod_{\ell_2=1}^{L_2} z_{j_{2,\ell_2}}^2 K^2(\boldsymbol{z}) g(\boldsymbol{z} \circ \boldsymbol{h}) d\boldsymbol{z} \right. \\
&\quad \left. - h_1 \dots h_d \left(\int \prod_{\ell_1=1}^{L_1} z_{j_{1,\ell_1}} \prod_{\ell_2=1}^{L_2} z_{j_{2,\ell_2}} K(\boldsymbol{z}) g(\boldsymbol{z} \circ \boldsymbol{h}) d\boldsymbol{z} \right)^2 \right\} \\
&= \frac{1}{nh_1 \dots h_d} \left(g(\boldsymbol{0}) \kappa^{(2)}_{j_{1,1}\dots j_{1,L_1} j_{2,1}\dots j_{2,L_2} j_{1,1}\dots j_{1,L_1} j_{2,1}\dots j_{2,L_2}} + o(1) \right) \\
&\quad - \frac{1}{n} (g(\boldsymbol{0}) \kappa^{(1)}_{j_{1,1}\dots j_{1,L_1} j_{2,1}\dots j_{2,L_2}} + o(1))^2 \\
&= \frac{g(\boldsymbol{0}) \kappa^{(2)}_{j_{1,1}\dots j_{1,L_1} j_{2,1}\dots j_{2,L_2} j_{1,1}\dots j_{1,L_1} j_{2,1}\dots j_{2,L_2}}}{nh_1 \dots h_d} + o\left(\frac{1}{nh_1 \dots h_d} \right).
\end{aligned}$$

Then for any $\rho > 0$,

$$\begin{aligned}
&P\left(|I_{n,j_{1,1}\dots j_{1,L_1},j_{2,1}\dots j_{2,L_2}} - g(\boldsymbol{0}) \kappa^{(1)}_{j_{1,1}\dots j_{1,L_1} j_{2,1}\dots j_{2,L_2}}| > \rho \right) \\
&\quad \le \rho^{-1} \left\{ \mathrm{Var}(I_{n,j_{1,1}\dots j_{1,L_1},j_{2,1}\dots j_{2,L_2}}) + \left(E[I_{n,j_{1,1}\dots j_{1,L_1},j_{2,1}\dots j_{2,L_2}}] - g(\boldsymbol{0}) \kappa^{(1)}_{j_{1,1}\dots j_{1,L_1} j_{2,1}\dots j_{2,L_2}} \right)^2 \right\} \\
&= O\left(\frac{1}{nh_1 \dots h_d} \right) + o(1) = o(1).
\end{aligned}$$

This yields $I_{n,j_{1,1}\dots j_{1,L_1},j_{2,1}\dots j_{2,L_2}} \stackrel{p}{\to} g(\boldsymbol{0}) \kappa^{(1)}_{j_{1,1}\dots j_{1,L_1} j_{2,1}\dots j_{2,L_2}}$. Hence we have

$$S_n(\boldsymbol{0}) \stackrel{p}{\to} g(\boldsymbol{0}) S.$$

(Step 2) Now we evaluate $V_n(\boldsymbol{0})$. For any $\boldsymbol{t} = (t_0, t_1, \dots, t_d, t_{11}, \dots, t_{dd}, \dots, t_{1\dots1}, \dots, t_{d\dots d})' \in \mathbb{R}^D$, we define

$$\widetilde{V}_n(\boldsymbol{0}) := \frac{nh_1 \dots h_d}{\sqrt{A_n h_1 \dots h_d}} \boldsymbol{t}' V_n(\boldsymbol{0}) = \sum_{i=1}^{n} K_{Ah}(\boldsymbol{X}_i) \left[\boldsymbol{t}' H^{-1} \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_i \end{pmatrix} \right] (e_{n,i} + \varepsilon_{n,i}).$$

In this step, we will show that

$$\begin{aligned}
&\boldsymbol{t}' V_n(\boldsymbol{0}) \\
&\stackrel{d}{\to} N\left(\boldsymbol{0}, g(\boldsymbol{0}) \left\{ \kappa(\eta^2(\boldsymbol{0}) + \sigma_\varepsilon^2(\boldsymbol{0})) + \eta^2(\boldsymbol{0}) g(\boldsymbol{0}) \int \sigma_{\boldsymbol{e}}(\boldsymbol{v}) d\boldsymbol{v} \right\} \int K^2(\boldsymbol{z}) \left[\boldsymbol{t}' \begin{pmatrix} 1 \\ \check{\boldsymbol{z}} \end{pmatrix} \right]^2 d\boldsymbol{z} \right). \quad \text{(A.1)}
\end{aligned}$$

Before we show (A.1), we introduce some notations. For $\boldsymbol{z}_0 = (z_{0,1}, \dots z_{0,d})' \in \mathbb{R}^d$ and $\boldsymbol{\ell} = (\ell_1, \dots, \ell_d)' \in \mathbb{Z}^d$, let

$$\Gamma_{n,\boldsymbol{z}_0}(\boldsymbol{\ell}; \mathbf{0}) = \prod_{j=1}^{d} (A_{n,j} z_{0,j} + (\ell_j - 1/2) A_{n3,j}, A_{n,j} z_{0,j} + (\ell_j + 1/2) A_{n3,j}]$$

with $A_{n3,j} = A_{n1,j} + A_{n2,j}$, and define the following hypercubes,

$$\Gamma_{n,\boldsymbol{z}_0}(\boldsymbol{\ell}; \boldsymbol{\Delta}) = \prod_{j=1}^{d} I_{j,\boldsymbol{z}_0}(\Delta_j), \ \boldsymbol{\Delta} = (\Delta_1, \dots, \Delta_d)' \in \{1, 2\}^d,$$

where

$$I_{j,\boldsymbol{z}_0}(\Delta_j) = \begin{cases} (A_{n,j} z_{0,j} + (\ell_j - 1/2) A_{n3,j}, A_{n,j} z_{0,j} + (\ell_j - 1/2) A_{n3,j} + A_{n1,j}] & \text{if } \Delta_j = 1, \\ (A_{n,j} z_{0,j} + (\ell_j - 1/2) A_{n3,j} + A_{n1,j}, A_{n,j} z_{0,j} + (\ell_j + 1/2) A_{n3,j}] & \text{if } \Delta_j = 2. \end{cases}$$

Let $\boldsymbol{\Delta}_0 = (1, \dots, 1)'$. The partitions $\Gamma_{n,\boldsymbol{z}_0}(\boldsymbol{\ell}; \boldsymbol{\Delta}_0)$ correspond to "large blocks" and the partitions $\Gamma_{n,\boldsymbol{z}_0}(\boldsymbol{\ell}; \boldsymbol{\Delta})$ for $\boldsymbol{\Delta} \neq \boldsymbol{\Delta}_0$ correspond to "small blocks". Let $L_{n1}(\boldsymbol{z}_0) = \{\boldsymbol{\ell} \in \mathbb{Z}^d : \Gamma_{n,\boldsymbol{z}_0}(\boldsymbol{\ell}; \mathbf{0}) \subset R_n \cap (\boldsymbol{h} R_n + A_n \boldsymbol{z}_0)\}$ denote the index set of all hypercubes $\Gamma_{n,\boldsymbol{z}_0}(\boldsymbol{\ell}; \mathbf{0})$ that are contained in $R_n \cap (\boldsymbol{h} R_n + A_n \boldsymbol{z}_0)$, and let $L_{n2}(\boldsymbol{z}_0) = \{\boldsymbol{\ell} \in \mathbb{Z}^d : \Gamma_{n,\boldsymbol{z}_0}(\boldsymbol{\ell}; \mathbf{0}) \cap R_n \cap (\boldsymbol{h} R_n + A_n \boldsymbol{z}_0) \neq 0, \Gamma_n(\boldsymbol{\ell}; \mathbf{0}) \cap (R_n \cap (\boldsymbol{h} R_n + A_n \boldsymbol{z}_0))^c \neq \emptyset\}$ be the index set of boundary hypercubes. Define $\Gamma_n(\boldsymbol{\ell}; \boldsymbol{\Delta}) = \Gamma_{n,\mathbf{0}}(\boldsymbol{\ell}; \boldsymbol{\Delta})$, $L_{n1} = L_{n1}(\mathbf{0})$, $L_{n2} = L_{n2}(\mathbf{0})$, and

$$\widetilde{V}_n(\boldsymbol{\ell}; \boldsymbol{\Delta}) = \sum_{i: \boldsymbol{X}_i \in \Gamma_n(\boldsymbol{\ell}; \boldsymbol{\Delta}) \cap \boldsymbol{h} R_n} K_{Ah}(\boldsymbol{X}_i) \left[\boldsymbol{t}' H^{-1} \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_i \end{pmatrix} \right] (e_{n,i} + \varepsilon_{n,i}).$$

Note that by our summation convention, $V_n(\boldsymbol{\ell}; \boldsymbol{\Delta}) = 0$ if the set $\{i : \boldsymbol{X}_i \in \Gamma_n(\boldsymbol{\ell}; \boldsymbol{\Delta}) \cap \boldsymbol{h} R_n\}$ is empty for some $\boldsymbol{\ell}$. Then we have

$$\begin{aligned} \widetilde{V}_n(\mathbf{0}) &= \sum_{\boldsymbol{\ell} \in L_{n1}} \widetilde{V}_n(\boldsymbol{\ell}; \boldsymbol{\Delta}_0) + \sum_{\boldsymbol{\Delta} \neq \boldsymbol{\Delta}_0} \sum_{\boldsymbol{\ell} \in L_{n1}} \widetilde{V}_n(\boldsymbol{\ell}; \boldsymbol{\Delta}) + \sum_{\boldsymbol{\Delta} \in \{1,2\}^d} \sum_{\boldsymbol{\ell} \in L_{n2}} \widetilde{V}_n(\boldsymbol{\ell}; \boldsymbol{\Delta}) \\ &=: \widetilde{V}_{n1} + \widetilde{V}_{n2} + \widetilde{V}_{n3}. \end{aligned}$$

Note that for $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2 \in L_{n1}$,

$$d\left(\Gamma_n(\boldsymbol{\ell}_1; \boldsymbol{\Delta}_0), \Gamma_n(\boldsymbol{\ell}_2; \boldsymbol{\Delta}_0)\right) \geq \min\{|\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2|, 0\} \underline{A}_{n3} + \underline{A}_{n2}, \tag{A.2}$$

where $\underline{A}_{n3} = \min_{1 \leq j \leq d} A_{n3,j}$ and $\underline{A}_{n2} = \min_{1 \leq j \leq d} A_{n2,j}$.

Hence, by the Volkonskii-Rozanov inequality (cf. Proposition 2.6 in Fan and Yao (2003)), we have

$$\left| E[\exp(\mathrm{i} u \widetilde{V}_{n1})] - \prod_{\ell \in L_{n1}} E[\exp(\mathrm{i} u \widetilde{V}_n(\boldsymbol{\ell}; \boldsymbol{\Delta}_0))] \right| \lesssim \left(\frac{A_n h_1 \dots h_d}{A_n^{(1)}} \right) \alpha(\underline{A}_{n2}; A_n h_1 \dots h_d). \tag{A.3}$$

From Lyapounov's CLT, it is sufficient to verify the following conditions to show (A.1): As $n \to \infty$,

$$\begin{aligned} E[\widetilde{V}_n^2(\mathbf{0})] \to g(\mathbf{0}) &\left\{ \kappa(\eta^2(\mathbf{0}) + \sigma_\varepsilon^2(\mathbf{0})) + \eta^2(\mathbf{0}) g(\mathbf{0}) \int \sigma_{\boldsymbol{e}}(\boldsymbol{v}) d\boldsymbol{v} \right\} \\ &\times \int K^2(\boldsymbol{z}) \left[\boldsymbol{t}' \begin{pmatrix} 1 \\ \check{\boldsymbol{z}} \end{pmatrix} \right]^2 d\boldsymbol{z}, \end{aligned} \tag{A.4}$$

$$\sum_{\boldsymbol{\ell}\in L_{n1}} E[\widetilde{V}_n^2(\boldsymbol{\ell};\boldsymbol{\Delta}_0)] - E[\widetilde{V}_n^2(\mathbf{0})] = o\left(n^2A_n^{-1}h_1\dots h_d\right), \tag{A.5}$$

$$\sum_{\boldsymbol{\ell}\in L_{n1}} E[\widetilde{V}_n^4(\boldsymbol{\ell};\boldsymbol{\Delta}_0)] = o\left(\left(n^2A_n^{-1}h_1\dots h_d\right)^2\right), \tag{A.6}$$

$$\mathrm{Var}(\widetilde{V}_{n2}) = o\left(n^2A_n^{-1}h_1\dots h_d\right), \tag{A.7}$$

$$\mathrm{Var}(\widetilde{V}_{n3}) = o\left(n^2A_n^{-1}h_1\dots h_d\right). \tag{A.8}$$

In the following steps, we show (A.4) (Step 2-1), (A.6) (Step 2-2), (A.7) and (A.8) (Step 2-3), and (A.5) (Step 2-4).

(Step 2-1) Now we show (A.4). Let δ_{ij} be a function such that $\delta_{ij}=1$ if $i=j$ and $\delta_{ij}=0$ if $i\neq j$. Observe that

$$\begin{aligned}\sigma_n^2(\mathbf{0}) &:= E_{\cdot|\boldsymbol{X}}\left(\widetilde{V}_n^2(\mathbf{0})\right)\\ &= \sum_{i,j=1}^n \boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_i\end{pmatrix}\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_j\end{pmatrix}K_{Ah}(\boldsymbol{X}_i)K_{Ah}(\boldsymbol{X}_j)\\ &\quad\times\left\{\eta(\boldsymbol{X}_i/A_n)\eta(\boldsymbol{X}_j/A_n)\sigma_{\boldsymbol{e}}(\boldsymbol{X}_i-\boldsymbol{X}_j)+\sigma_\varepsilon^2(\boldsymbol{X}_i/A_n)\delta_{ij}\right\}.\end{aligned}$$

Thus we have

$$\begin{aligned}&E_{\boldsymbol{X}}\left[\sigma_n^2(\mathbf{0})\right]\\ &= nA_n^{-1}\int\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ (\boldsymbol{x}/A_n)^{\vee}\end{pmatrix}\right]^2K_{Ah}^2(\boldsymbol{x})\left\{\eta^2(\boldsymbol{x}/A_n)+\sigma_\varepsilon^2(\boldsymbol{x}/A_n)\right\}g(\boldsymbol{x}/A_n)d\boldsymbol{x}\\ &\quad+n(n-1)A_n^{-2}\int\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ (\boldsymbol{x}_1/A_n)^{\vee}\end{pmatrix}\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ (\boldsymbol{x}_2/A_n)^{\vee}\end{pmatrix}K_{Ah}(\boldsymbol{x}_1)K_{Ah}(\boldsymbol{x}_2)\\ &\quad\times\eta(\boldsymbol{x}_1/A_n)\eta(\boldsymbol{x}_2/A_n)\sigma_{\boldsymbol{e}}(\boldsymbol{x}_1-\boldsymbol{x}_2)g(\boldsymbol{x}_1/A_n)g(\boldsymbol{x}_2/A_n)d\boldsymbol{x}_1d\boldsymbol{x}_2\\ &=: \sigma_{n,1}^2+\sigma_{n,2}^2.\end{aligned}$$

For $\sigma_{n,1}^2$, we have

$$\begin{aligned}\sigma_{n,1}^2 &= nh_1\dots h_d\int\left[\boldsymbol{t}'\begin{pmatrix}1\\ \check{\boldsymbol{z}}\end{pmatrix}\right]^2K^2(\boldsymbol{z})\left\{\eta^2(\boldsymbol{z}\circ\boldsymbol{h})+\sigma_\varepsilon^2(\boldsymbol{z}\circ\boldsymbol{h})\right\}g(\boldsymbol{z}\circ\boldsymbol{h})d\boldsymbol{z}\\ &= nh_1\dots h_d(\eta^2(\mathbf{0})+\sigma_\varepsilon^2(\mathbf{0}))g(\mathbf{0})\left(\int K^2(\boldsymbol{z})\left[\boldsymbol{t}'\begin{pmatrix}1\\ \check{\boldsymbol{z}}\end{pmatrix}\right]^2d\boldsymbol{z}\right)(1+o(1)).\end{aligned} \tag{A.9}$$

For $\sigma_{n,2}^2$, we have

$$\begin{aligned}\sigma_{n,2}^2 &= n(n-1)\int_{R_0^2}\sigma_{\boldsymbol{e}}(A_n(\boldsymbol{y}_1-\boldsymbol{y}_2))\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{y}}_1\end{pmatrix}\right]\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{y}}_2\end{pmatrix}\right]\\ &\quad\times K_h(\boldsymbol{y}_1)K_h(\boldsymbol{y}_2)\eta(\boldsymbol{y}_1)\eta(\boldsymbol{y}_2)g(\boldsymbol{y}_1)g(\boldsymbol{y}_2)d\boldsymbol{y}_1d\boldsymbol{y}_2\\ &= n(n-1)(h_1\dots h_d)^2\int_{\boldsymbol{h}^{-1}R_0^2}\sigma_{\boldsymbol{e}}(A_n(\boldsymbol{z}_1-\boldsymbol{z}_2)\circ\boldsymbol{h})\left[\boldsymbol{t}'\begin{pmatrix}1\\ \check{\boldsymbol{z}}_1\end{pmatrix}\right]\left[\boldsymbol{t}'\begin{pmatrix}1\\ \check{\boldsymbol{z}}_2\end{pmatrix}\right]\\ &\quad\times K(\boldsymbol{z}_1)K(\boldsymbol{z}_2)\eta(\boldsymbol{z}_1\circ\boldsymbol{h})\eta(\boldsymbol{z}_2\circ\boldsymbol{h})g(\boldsymbol{z}_1\circ\boldsymbol{h})g(\boldsymbol{z}_2\circ\boldsymbol{h})d\boldsymbol{z}_1d\boldsymbol{z}_2\end{aligned}$$

$$= n(n-1)(h_1 \ldots h_d)^2 \int_{R'_{\boldsymbol{h},0}} \sigma_{\boldsymbol{e}}(A_n \boldsymbol{w} \circ \boldsymbol{h}) \left(\int_{R_{\boldsymbol{h},0}(\boldsymbol{w})} \left[\boldsymbol{t}' \begin{pmatrix} 1 \\ (\check{\boldsymbol{z}_2 + \boldsymbol{w}}) \end{pmatrix} \right] \left[\boldsymbol{t}' \begin{pmatrix} 1 \\ \check{\boldsymbol{z}}_2 \end{pmatrix} \right] \right.$$

$$\left. \times K(\boldsymbol{z}_2 + \boldsymbol{w}) K(\boldsymbol{z}_2) \eta((\boldsymbol{z}_2 + \boldsymbol{w}) \circ \boldsymbol{h}) \eta(\boldsymbol{z}_2 \circ \boldsymbol{h}) g((\boldsymbol{z}_2 + \boldsymbol{w}) \circ \boldsymbol{h}) g(\boldsymbol{z}_2 \circ \boldsymbol{h}) d\boldsymbol{z}_2 \right) d\boldsymbol{w}$$

$$= n(n-1) h_1 \ldots h_d \int_{\boldsymbol{h}R'_{\boldsymbol{h},0}} \sigma_{\boldsymbol{e}}(A_n \boldsymbol{u}) \left(\int_{R_{\boldsymbol{h},0}(\boldsymbol{u}/\boldsymbol{h})} \left[\boldsymbol{t}' \begin{pmatrix} 1 \\ (\check{\boldsymbol{z}_2 + \boldsymbol{u} \circ \boldsymbol{h}^{-1}}) \end{pmatrix} \right] \left[\boldsymbol{t}' \begin{pmatrix} 1 \\ \check{\boldsymbol{z}}_2 \end{pmatrix} \right] \right.$$

$$\left. \times K(\boldsymbol{z}_2 + \boldsymbol{u} \circ \boldsymbol{h}^{-1}) K(\boldsymbol{z}_2) \eta(\boldsymbol{z}_2 \circ \boldsymbol{h} + \boldsymbol{u}) \eta(\boldsymbol{z}_2 \circ \boldsymbol{h}) g((\boldsymbol{z}_2 \circ \boldsymbol{h} + \boldsymbol{u}) g(\boldsymbol{z}_2 \circ \boldsymbol{h}) d\boldsymbol{z}_2 \right) d\boldsymbol{u}$$

$$= n(n-1) A_n^{-1} h_1 \ldots h_d \int_{A_n \boldsymbol{h}R'_{\boldsymbol{h},0}} \sigma_{\boldsymbol{e}}(\boldsymbol{v}) \left(\int_{R_{\boldsymbol{h},0}((\boldsymbol{v} \circ \boldsymbol{h}^{-1})/A_n)} \left[\boldsymbol{t}' \begin{pmatrix} 1 \\ \left(\check{\boldsymbol{z}_2 + \frac{\boldsymbol{v} \circ \boldsymbol{h}^{-1}}{A_n}} \right) \end{pmatrix} \right] \left[\boldsymbol{t}' \begin{pmatrix} 1 \\ \check{\boldsymbol{z}}_2 \end{pmatrix} \right] \right.$$

$$\left. \times K\left(\boldsymbol{z}_2 + \frac{\boldsymbol{v} \circ \boldsymbol{h}^{-1}}{A_n} \right) K(\boldsymbol{z}_2) \eta\left(\boldsymbol{z}_2 \circ \boldsymbol{h} + \frac{\boldsymbol{v}}{A_n} \right) \eta(\boldsymbol{z}_2 \circ \boldsymbol{h}) g\left(\boldsymbol{z}_2 \circ \boldsymbol{h} + \frac{\boldsymbol{v}}{A_n} \right) g(\boldsymbol{z}_2 \circ \boldsymbol{h}) d\boldsymbol{z}_2 \right) d\boldsymbol{v}$$

where

$$R'_{\boldsymbol{h},0} = \{\boldsymbol{w} = \boldsymbol{z}_1 - \boldsymbol{z}_2 : \boldsymbol{z}_1, \boldsymbol{z}_2 \in \boldsymbol{h}^{-1} R_0\}, \ R_{\boldsymbol{h},0}(\boldsymbol{w}) = \{\boldsymbol{z}_2 : \boldsymbol{z}_2 \in \boldsymbol{h}^{-1} R_0 \cap (\boldsymbol{h}^{-1} R_0 + \boldsymbol{w})\},$$

$$A_n \boldsymbol{h} R'_{\boldsymbol{h},0} = \{(A_{n,1} x_1, \ldots, A_{n,d} x_d) : \boldsymbol{x} = (x_1, \ldots, x_d)' \in \boldsymbol{h} R'_{\boldsymbol{h},0}\}.$$

We divide the integral $\int_{A_n \boldsymbol{h} R'_{\boldsymbol{h},0}}$ into two parts $\int_{A_n \boldsymbol{h} R'_{\boldsymbol{h},0} \cap \{|\boldsymbol{v}| \leq M\}}$ and $\int_{A_n \boldsymbol{h} R'_{\boldsymbol{h},0} \cap \{|\boldsymbol{v}| > M\}}$ for some $M > 0$ and define these as $\sigma^2_{n,21}$ and $\sigma^2_{n,22}$, respectively. Observe that as $n \to \infty$

$$|\sigma^2_{n,22}| \lesssim \int_{\{|\boldsymbol{v}| > M\}} |\sigma_{\boldsymbol{e}}(\boldsymbol{v})| d\boldsymbol{v}$$

which can be made arbitrary small by choosing a large M. Further, observe that as $n \to \infty$

$$1\{A_n \boldsymbol{h} R'_{\boldsymbol{h},0} \cap \{|\boldsymbol{v}| \leq M\}\} \int_{R_{\boldsymbol{h},0}(\boldsymbol{v}/(A_n \boldsymbol{h}))} \left[\boldsymbol{t}' \begin{pmatrix} 1 \\ \left(\check{\boldsymbol{z}_2 + \frac{\boldsymbol{v} \circ \boldsymbol{h}^{-1}}{A_n}} \right) \end{pmatrix} \right] \left[\boldsymbol{t}' \begin{pmatrix} 1 \\ \check{\boldsymbol{z}}_2 \end{pmatrix} \right]$$

$$\times K\left(\boldsymbol{z}_2 + \frac{\boldsymbol{v} \circ \boldsymbol{h}^{-1}}{A_n} \right) K(\boldsymbol{z}_2) \eta\left(\boldsymbol{z}_2 \circ \boldsymbol{h} + \frac{\boldsymbol{v}}{A_n} \right) \eta(\boldsymbol{z}_2 \circ \boldsymbol{h}) g\left(\boldsymbol{z}_2 \circ \boldsymbol{h} + \frac{\boldsymbol{v}}{A_n} \right) g(\boldsymbol{z}_2 \circ \boldsymbol{h}) d\boldsymbol{z}_2$$

$$= 1\{|\boldsymbol{v}| \leq M\} \eta^2(\boldsymbol{0}) g^2(\boldsymbol{0}) \left(\int K^2(\boldsymbol{z}_2) \left[\boldsymbol{t}' \begin{pmatrix} 1 \\ \check{\boldsymbol{z}}_2 \end{pmatrix} \right]^2 d\boldsymbol{z}_2 \right) (1 + o(1)).$$

Then as $n \to \infty$, we have

$$\sigma^2_{n,21} = \eta^2(\boldsymbol{0}) g^2(\boldsymbol{0}) \left(\int_{\{|\boldsymbol{v}| \leq M\}} \sigma_{\boldsymbol{e}}(\boldsymbol{v}) d\boldsymbol{v} \right) \left(\int K^2(\boldsymbol{z}_2) \left[\boldsymbol{t}' \begin{pmatrix} 1 \\ \check{\boldsymbol{z}}_2 \end{pmatrix} \right]^2 d\boldsymbol{z}_2 \right) (1 + o(1)).$$

Therefore, we have

$$\sigma^2_{n,2} = n^2 A_n^{-1} h_1 \ldots h_d \eta^2(\boldsymbol{0}) g^2(\boldsymbol{0}) \left(\int \sigma_{\boldsymbol{e}}(\boldsymbol{v}) d\boldsymbol{v} \right) \left(\int K^2(\boldsymbol{z}) \left[\boldsymbol{t}' \begin{pmatrix} 1 \\ \check{\boldsymbol{z}} \end{pmatrix} \right]^2 d\boldsymbol{z} \right) (1 + o(1)). \quad \text{(A.10)}$$

By (A.9) and (A.10), we have

$$\mathrm{Var}(\boldsymbol{t}' V_n(\boldsymbol{0}))$$

$$= g(\boldsymbol{0}) \left\{ \kappa(\eta^2(\boldsymbol{0}) + \sigma^2_\varepsilon(\boldsymbol{0})) + \eta^2(\boldsymbol{0}) g(\boldsymbol{0}) \int \sigma_{\boldsymbol{e}}(\boldsymbol{v}) d\boldsymbol{v} \right\} \left(\int K^2(\boldsymbol{z}) \left[\boldsymbol{t}' \begin{pmatrix} 1 \\ \check{\boldsymbol{z}} \end{pmatrix} \right]^2 d\boldsymbol{z} \right) (1 + o(1)).$$

(Step 2-2) Now we show (A.6). Define $I_n(\boldsymbol{\ell}) = \{\boldsymbol{i} \in \mathbb{Z}^d : \boldsymbol{i} + (-1/2, 1/2]^d \subset \Gamma_n(\boldsymbol{\ell}; \boldsymbol{\Delta}_0)\}$ for $\boldsymbol{\ell} \in L_{n1}$ and

$$\widetilde{V}_n(\boldsymbol{i}) = \sum_{i=1}^{n} K_{Ah}(\boldsymbol{X}_i) \left[\boldsymbol{t}' H^{-1} \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_i \end{pmatrix}\right] (e_{n,i} + \varepsilon_{n,i}) 1\{\boldsymbol{X}_i \in [\boldsymbol{i} + (-1/2, 1/2]^d] \cap R_n\}.$$

Observe that

$$\begin{aligned}
&E[\widetilde{V}_n^4(\boldsymbol{\ell}; \boldsymbol{\Delta}_0)] \\
&= E\left[\left(\sum_{\boldsymbol{i} \in I_n(\boldsymbol{\ell})} \widetilde{V}_n(\boldsymbol{i})\right)^4\right] \\
&= \sum_{\boldsymbol{i} \in I_n(\boldsymbol{\ell})} E\left[\widetilde{V}_n^4(\boldsymbol{i})\right] + \sum_{\boldsymbol{i},\boldsymbol{j} \in I_n(\boldsymbol{\ell}), \boldsymbol{i} \neq \boldsymbol{j}} E\left[\widetilde{V}_n^3(\boldsymbol{i}) \widetilde{V}_n(\boldsymbol{j})\right] + \sum_{\boldsymbol{i},\boldsymbol{j} \in I_n(\boldsymbol{\ell}), \boldsymbol{i} \neq \boldsymbol{j}} E\left[\widetilde{V}_n^2(\boldsymbol{i}) \widetilde{V}_n^2(\boldsymbol{j})\right] \\
&\quad + \sum_{\boldsymbol{i},\boldsymbol{j},\boldsymbol{k} \in I_n(\boldsymbol{\ell}), \boldsymbol{i} \neq \boldsymbol{j} \neq \boldsymbol{k}} E\left[\widetilde{V}_n^2(\boldsymbol{i}) \widetilde{V}_n(\boldsymbol{j}) \widetilde{V}_n(\boldsymbol{k})\right] + \sum_{\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p} \in I_n(\boldsymbol{\ell}), \boldsymbol{i} \neq \boldsymbol{j} \neq \boldsymbol{k} \neq \boldsymbol{p}} E\left[\widetilde{V}_n(\boldsymbol{i}) \widetilde{V}_n(\boldsymbol{j}) \widetilde{V}_n(\boldsymbol{k}) \widetilde{V}_n(\boldsymbol{p})\right] \\
&=: Q_{n1} + Q_{n2} + Q_{n3} + Q_{n4} + Q_{n5}.
\end{aligned}$$

For Q_{n1}, we have

$$\begin{aligned}
&E[\widetilde{V}_n^4(\boldsymbol{i})] \\
&= E_{\boldsymbol{X}}[E_{\cdot|\boldsymbol{X}}[\widetilde{V}_n^4(\boldsymbol{i})]] \\
&= \sum_{j_1,j_2,j_3,j_4=1}^{n} E\left[\prod_{k=1}^{4} K_{Ah}(\boldsymbol{X}_{j_k}) \left[\boldsymbol{t}' H^{-1} \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_{j_k} \end{pmatrix}\right] 1\{\boldsymbol{X}_{j_k} \in [\boldsymbol{i} + (-1/2, 1/2]^d] \cap R_n\}\right. \\
&\qquad \left. \times E_{\cdot|\boldsymbol{X}}[e_{n,j_k} + \varepsilon_{n,j_k}]\right] \\
&\lesssim \sum_{j_1,j_2,j_3,j_4=1}^{n} E\left[\prod_{k=1}^{4} \left|K_{Ah}(\boldsymbol{X}_{j_k}) \left[\boldsymbol{t}' H^{-1} \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_{j_k} \end{pmatrix}\right]\right| 1\{\boldsymbol{X}_{j_k} \in [\boldsymbol{i} + (-1/2, 1/2]^d] \cap R_n\} \eta(\boldsymbol{X}_{j_k}/A_n)\right] \\
&\quad + \sum_{j_1,j_2,j_3,j_4=1}^{n} E\left[\prod_{k=1}^{4} \left|K_{Ah}(\boldsymbol{X}_{j_k}) \left[\boldsymbol{t}' H^{-1} \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_{j_k} \end{pmatrix}\right]\right| 1\{\boldsymbol{X}_{j_k} \in [\boldsymbol{i} + (-1/2, 1/2]^d] \cap R_n\} \sigma_{\varepsilon}(\boldsymbol{X}_{j_k}/A_n)\right] \\
&=: Q_{n11} + Q_{n12}.
\end{aligned}$$

For Q_{n11}, we have

$$\begin{aligned}
&Q_{n11} \\
&\lesssim nE\left[\left|K_{Ah}(\boldsymbol{X}_1) \left[\boldsymbol{t}' H^{-1} \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_1 \end{pmatrix}\right]\right|^4 1\{\boldsymbol{X}_1 \in [\boldsymbol{i} + (-1/2, 1/2]^d] \cap R_n\} \eta^4(\boldsymbol{X}_1/A_n)\right] \\
&\quad + n^2 E\left[\left|K_{Ah}(\boldsymbol{X}_1) \left[\boldsymbol{t}' H^{-1} \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_1 \end{pmatrix}\right]\right|^3 1\{\boldsymbol{X}_1 \in [\boldsymbol{i} + (-1/2, 1/2]^d] \cap R_n\}\right. \\
&\quad \left. \times \left|K_{Ah}(\boldsymbol{X}_2) \left[\boldsymbol{t}' H^{-1} \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_2 \end{pmatrix}\right]\right| 1\{\boldsymbol{X}_2 \in [\boldsymbol{i} + (-1/2, 1/2]^d] \cap R_n\} \eta^3(\boldsymbol{X}_1/A_n) \eta(\boldsymbol{X}_2/A_n)\right] \\
&\quad + n^2 E\left[\left|K_{Ah}(\boldsymbol{X}_1) \left[\boldsymbol{t}' H^{-1} \begin{pmatrix} 1 \\ \check{\boldsymbol{X}}_1 \end{pmatrix}\right]\right|^2 1\{\boldsymbol{X}_1 \in [\boldsymbol{i} + (-1/2, 1/2]^d] \cap R_n\}\right.
\end{aligned}$$

$$\times \left|K_{Ah}\left(\boldsymbol{X}_2\right)\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_2\end{pmatrix}\right]\right|^2 1\{\boldsymbol{X}_2\in[\boldsymbol{i}+(-1/2,1/2]^d]\cap R_n\}$$
$$\times\eta^2(\boldsymbol{X}_1/A_n)\eta^2(\boldsymbol{X}_2/A_n)]$$
$$+n^3E\left[\left|K_{Ah}\left(\boldsymbol{X}_1\right)\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_1\end{pmatrix}\right]\right|^2 1\{\boldsymbol{X}_1\in[\boldsymbol{i}+(-1/2,1/2]^d]\cap R_n\}\right.$$
$$\times \left|K_{Ah}\left(\boldsymbol{X}_2\right)\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_2\end{pmatrix}\right]\right| 1\{\boldsymbol{X}_2\in[\boldsymbol{i}+(-1/2,1/2]^d]\cap R_n\}$$
$$\times \left|K_{Ah}\left(\boldsymbol{X}_3\right)\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_3\end{pmatrix}\right]\right| 1\{\boldsymbol{X}_3\in[\boldsymbol{i}+(-1/2,1/2]^d]\cap R_n\}$$
$$\times\eta^2(\boldsymbol{X}_1/A_n)\eta(\boldsymbol{X}_2/A_n)\eta(\boldsymbol{X}_3/A_n)]$$
$$+n^4E\left[\left|K_{Ah}\left(\boldsymbol{X}_1\right)\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_1\end{pmatrix}\right]\right| 1\{\boldsymbol{X}_1\in[\boldsymbol{i}+(-1/2,1/2]^d]\cap R_n\}\right.$$
$$\times \left|K_{Ah}\left(\boldsymbol{X}_2\right)\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_2\end{pmatrix}\right]\right| 1\{\boldsymbol{X}_2\in[\boldsymbol{i}+(-1/2,1/2]^d]\cap R_n\}$$
$$\times \left|K_{Ah}\left(\boldsymbol{X}_3\right)\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_3\end{pmatrix}\right]\right| 1\{\boldsymbol{X}_3\in[\boldsymbol{i}+(-1/2,1/2]^d]\cap R_n\}$$
$$\times \left|K_{Ah}\left(\boldsymbol{X}_4\right)\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_4\end{pmatrix}\right]\right| 1\{\boldsymbol{X}_4\in[\boldsymbol{i}+(-1/2,1/2]^d]\cap R_n\}$$
$$\times\eta^2(\boldsymbol{X}_1/A_n)\eta(\boldsymbol{X}_2/A_n)\eta(\boldsymbol{X}_3/A_n)\eta(\boldsymbol{X}_4/A_n)]$$
$$=: Q_{n111}+Q_{n112}+Q_{n113}+Q_{n114}.$$

For Q_{n111}, we have

$$\begin{aligned}
Q_{n111} &= nA_n^{-1}\int\left|K_{Ah}(\boldsymbol{x})\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ (\check{\boldsymbol{x}/A_n})\end{pmatrix}\right]\right|^4 1\{\boldsymbol{x}\in[\boldsymbol{i}+(-1/2,1/2]^d]\cap R_n\}\\
&\quad\times\eta^4(\boldsymbol{x}/A_n)g(\boldsymbol{x}/A_n)d\boldsymbol{x}\\
&= nA_n^{-1}A_nh_1\dots h_d\int\left|K(\boldsymbol{z})\left[\boldsymbol{t}'\begin{pmatrix}1\\ \check{\boldsymbol{z}}\end{pmatrix}\right]\right|^4 1\{\boldsymbol{z}\circ\boldsymbol{h}\in[\boldsymbol{i}+(-1/2,1/2]^d]/A_n\cap[-1/2,1/2]^d\}\\
&\quad\times\eta^4(\boldsymbol{z}\circ\boldsymbol{h})g(\boldsymbol{z}\circ\boldsymbol{h})d\boldsymbol{z}\\
&= O\left(nA_n^{-1}\right).
\end{aligned}$$

Likewise, $Q_{n112}=O(n^2A_n^{-2})$, $Q_{n113}=O(n^3A_n^{-3})$, and $Q_{n114}=O(n^4A_n^{-4})$. Then we have $Q_{n11}=O(n^4A_n^{-4})$. We can also show that $Q_{n12}=O(n^4A_n^{-4})$. Therefore, we have

$$Q_{n1}\lesssim[\![I_n(\boldsymbol{\ell})]\!]n^4A_n^{-4}\lesssim A_n^{(1)}(nA_n^{-1})^4. \tag{A.11}$$

For Q_{n2}, by the α-mixing property of $\boldsymbol{e}$ and Proposition 2.5 in Fan and Yao (2003), we have

$$Q_{n2}\lesssim\sum_{k=1}^{\overline{A}_{n1}}\sum_{\boldsymbol{i},\boldsymbol{j}\in I_n(\boldsymbol{\ell}),|\boldsymbol{i}-\boldsymbol{j}|=k}\alpha^{1-4/q}(\min\{k-d,0\};1)E[|\widetilde{V}_n(\boldsymbol{i})|^q]^{3/q}E[|\widetilde{V}_n(\boldsymbol{j})|^q]^{1/q}$$

$$\lesssim A_n^{(1)}(nA_n^{-1})^4\left(1+\sum_{k=1}^{\overline{A}_{n1}}k^{d-1}\alpha_1^{1-4/q}(k)\right). \tag{A.12}$$

where $\overline{A}_{n1}=\max_{1\le j\le d}A_{n1,j}$. Likewise,

$$Q_{n3}\lesssim A_n^{(1)}(nA_n^{-1})^4\left(1+\sum_{k=1}^{\overline{A}_{n1}}k^{d-1}\alpha_1^{1-4/q}(k)\right). \tag{A.13}$$

Now we evaluate Q_{n4} and Q_{n5}. For distinct indices $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p}\in I_n(\boldsymbol{\ell})$, let

$$\begin{aligned}
d_1(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})&=\max\{d(\{\boldsymbol{i}\},\{\boldsymbol{j},\boldsymbol{k}\}),d(\{\boldsymbol{k}\},\{\boldsymbol{i},\boldsymbol{j}\})\},\\
d_2(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p})&=\max\{d(J,\{\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p}\}):J\subset\{\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p}\},[\![J]\!]=1\},\\
d_3(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p})&=\max\{d(J,\{\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p}\}):J\subset\{\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p}\},[\![J]\!]=2\}.
\end{aligned}$$

Here, d_1 denotes the maximal gap in the set of integer-indices $\{\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}\}$ from either $\boldsymbol{j}$ or $\boldsymbol{k}$ which corresponds to $E\left[\widetilde{V}_n^2(\boldsymbol{i})\widetilde{V}_n(\boldsymbol{j})\widetilde{V}_n(\boldsymbol{k})\right]$. Similarly, d_2 and d_3 are the maximal gap in the index set $\{\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p}\}$ from any of its single index-subsets or two-index subsets, respectively. Applying the argument in the proof of Lemma 4.1 of Lahiri (1999), for any given values $1\le d_{01},d_{02},d_{03}<[\![I_n(\boldsymbol{\ell})]\!]$, we have

$$[\![\{(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})\in I_n^3(\boldsymbol{\ell}):\boldsymbol{i}\neq\boldsymbol{j}\neq\boldsymbol{k}\text{ and }d_1(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k})=d_{01}\}]\!]\lesssim d_{01}^{2d-1}[\![I_n(\boldsymbol{\ell})]\!], \tag{A.14}$$

$$\begin{aligned}
&[\![\{(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p})\in I_n^4(\boldsymbol{\ell}):\boldsymbol{i}\neq\boldsymbol{j}\neq\boldsymbol{k}\neq\boldsymbol{p},\ d_2(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p})=d_{02},\text{ and }d_3(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p})=d_{03}\}]\!]\\
&\quad\lesssim(d_{02}+d_{03})^{3d-1}[\![I_n(\boldsymbol{\ell})]\!].
\end{aligned} \tag{A.15}$$

For Q_{n4}, by (A.14) and applying the same argument to show (A.12), we have

$$\begin{aligned}
Q_{n4}&\lesssim A_n^{(1)}\sum_{k=1}^{\overline{A}_{n1}}k^{2d-1}\alpha^{1-4/q}(\min\{k-d,0\};2)E[|\widetilde{V}_n(\boldsymbol{i})|^q]^{2/q}E[|\widetilde{V}_n(\boldsymbol{j})|^q]^{1/q}E[|\widetilde{V}_n(\boldsymbol{k})|^q]^{1/q}\\
&\lesssim A_n^{(1)}(nA_n^{-1})^4\left(1+\sum_{k=1}^{\overline{A}_{n1}}k^{2d-1}\alpha_1^{1-4/q}(k)\right).
\end{aligned} \tag{A.16}$$

Define

$$\begin{aligned}
I_{n1}(\boldsymbol{\ell})&=\{(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p})\in I_n^4(\boldsymbol{\ell}):\boldsymbol{i}\neq\boldsymbol{j}\neq\boldsymbol{k}\neq\boldsymbol{p},\ d_2(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p})\ge d_3(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p})\},\\
I_{n2}(\boldsymbol{\ell})&=\{(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p})\in I_n^4(\boldsymbol{\ell}):\boldsymbol{i}\neq\boldsymbol{j}\neq\boldsymbol{k}\neq\boldsymbol{p},\ d_2(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p})< d_3(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p})\}.
\end{aligned}$$

For Q_{n5}, by (A.15) and applying the same argument to show (A.12), we have

$$\begin{aligned}
Q_{n5}&=\sum_{(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p})\in I_{n1}(\boldsymbol{\ell})}E\left[\widetilde{V}_n(\boldsymbol{i})\widetilde{V}_n(\boldsymbol{j})\widetilde{V}_n(\boldsymbol{k})\widetilde{V}_n(\boldsymbol{p})\right]+\sum_{(\boldsymbol{i},\boldsymbol{j},\boldsymbol{k},\boldsymbol{p})\in I_{n2}(\boldsymbol{\ell})}E\left[\widetilde{V}_n(\boldsymbol{i})\widetilde{V}_n(\boldsymbol{j})\widetilde{V}_n(\boldsymbol{k})\widetilde{V}_n(\boldsymbol{p})\right]\\
&\lesssim A_n^{(1)}\sum_{k=1}^{\overline{A}_{n1}}k^{3d-1}\alpha^{1-4/q}(\min\{k-d,0\};3)\\
&\quad\times E[|\widetilde{V}_n(\boldsymbol{i})|^q]^{1/q}E[|\widetilde{V}_n(\boldsymbol{j})|^q]^{1/q}E[|\widetilde{V}_n(\boldsymbol{k})|^q]^{1/q}E[|\widetilde{V}_n(\boldsymbol{p})|^q]^{1/q}\\
&\quad+\left(\sum_{\boldsymbol{i},\boldsymbol{j}\in I_n(\boldsymbol{\ell}),\boldsymbol{i}\neq\boldsymbol{j}}\left|E[\widetilde{V}_n(\boldsymbol{i})\widetilde{V}_n(\boldsymbol{j})]\right|\right)^2
\end{aligned}$$

$$+ A_n^{(1)} \sum_{k=1}^{\overline{A}_{n1}} k^{3d-1} \alpha^{1-4/q}(\min\{k-d,0\};2)$$
$$\times E[|\widetilde{V}_n(\boldsymbol{i})|^q]^{1/q} E[|\widetilde{V}_n(\boldsymbol{j})|^q]^{1/q} E[|\widetilde{V}_n(\boldsymbol{k})|^q]^{1/q} E[|\widetilde{V}_n(\boldsymbol{p})|^q]^{1/q}$$
$$\lesssim (A_n^{(1)})^2 (nA_n^{-1})^4 \left(1 + \sum_{k=1}^{\overline{A}_{n1}} k^{2d-1} \alpha_1^{1-4/q}(k)\right). \tag{A.17}$$

Combining (A.11)-(A.17), we have

$$\begin{aligned}
\sum_{\boldsymbol{\ell} \in L_{n1}} E[\widetilde{V}_n^4(\boldsymbol{\ell};\boldsymbol{\Delta}_0)] &= \sum_{\boldsymbol{\ell} \in L_{n1}} E\left[\left(\sum_{\boldsymbol{i} \in I_n(\boldsymbol{\ell})} \widetilde{V}_n(\boldsymbol{i})\right)^4\right] \\
&\lesssim [\![L_{n1}]\!] (A_n^{(1)})^2 (nA_n^{-1})^4 \left(1 + \sum_{k=1}^{\overline{A}_{n1}} k^{2d-1} \alpha_1^{1-4/q}(k)\right) \\
&\lesssim \left(\frac{A_n h_1 \dots h_d}{A_n^{(1)}}\right) (A_n^{(1)})^2 (nA_n^{-1})^4 \left(1 + \sum_{k=1}^{\overline{A}_{n1}} k^{2d-1} \alpha_1^{1-4/q}(k)\right) \\
&= o\left((n^2 A_n^{-1} h_1 \dots h_d)^2\right).
\end{aligned}$$

(Step 2-3) Now we show (A.7) and (A.8). Define

$$\begin{aligned}
J_n &= \{\boldsymbol{i} \in \mathbb{Z}^d : (\boldsymbol{i} + (-1/2, 1/2]^d) \cap \boldsymbol{h} R_n \neq \emptyset\}, \\
J_{n1} &= \cup_{\boldsymbol{\ell} \in L_{n1}} I_n(\boldsymbol{\ell}), \\
J_{n2} &= \{\boldsymbol{i} \in J_n : \boldsymbol{i} + (-1/2, 1/2]^d \subset \Gamma_n(\boldsymbol{\ell};\boldsymbol{\Delta}) \text{ for some } \boldsymbol{\ell} \in L_{n1}, \boldsymbol{\Delta} \neq \boldsymbol{\Delta}_0\}, \\
J_{n3} &= J_n \backslash (J_{n1} \cup J_{n2}).
\end{aligned}$$

Note that $[\![J_{n2}]\!] \lesssim (\overline{A}_{n1})^{d-1} \overline{A}_{n2} \left(\frac{A_n h_1 \dots h_d}{A_n^{(1)}}\right)$ and $[\![J_{n3}]\!] \lesssim A_n^{(1)} \left(\frac{\overline{A_n h}}{\underline{A}_{n1}}\right)^{d-1}$. Then, applying the same argument to show (A.12), we have

$$\begin{aligned}
\mathrm{Var}(\widetilde{V}_{n2}) &\lesssim [\![J_{n2}]\!] (nA_n^{-1})^2 \left(1 + \sum_{k=1}^{\overline{A}_{n1}} k^{d-1} \alpha_1^{1-2/q}(k)\right) \\
&\lesssim \left(\frac{\overline{A}_{n1}}{\underline{A}_{n1}}\right)^d \left(\frac{\overline{A}_{n2}}{\overline{A}_{n1}}\right) A_n h_1 \dots h_d (nA_n^{-1})^2 \left(1 + \sum_{k=1}^{\overline{A}_{n1}} k^{d-1} \alpha_1^{1-2/q}(k)\right) \\
&= o\left(n^2 A_n^{-1} h_1 \dots h_d\right). \\
\mathrm{Var}(\widetilde{V}_{n3}) &\lesssim [\![J_{n3}]\!] (nA_n^{-1})^2 \left(1 + \sum_{k=1}^{\overline{A}_{n1}} k^{d-1} \alpha_1^{1-2/q}(k)\right) \\
&\lesssim \left(\frac{A_n^{(1)}}{\underline{A}_{n1}^d}\right) \left(\frac{(\overline{A_n h})^d}{A_n h_1 \dots h_d}\right) \left(\frac{\underline{A}_{n1}}{\underline{A_n h}}\right) A_n h_1 \dots h_d (nA_n^{-1})^2 \left(1 + \sum_{k=1}^{\overline{A}_{n1}} k^{d-1} \alpha_1^{1-2/q}(k)\right) \\
&= o\left(n^2 A_n^{-1} h_1 \dots h_d\right).
\end{aligned}$$

(Step 2-4) Now we show (A.5). By (A.7) and (A.8), we have for sufficiently large n,

$$E[\widetilde{V}_{n1}^2] = E[(\widetilde{V}_n(\mathbf{0}) - (\widetilde{V}_{n2} + \widetilde{V}_{n3}))^2] \leq 2\left(E[(\widetilde{V}_n(\mathbf{0}))^2] + E[(\widetilde{V}_{n2} + \widetilde{V}_{n3})^2]\right) \leq 4E[\widetilde{V}_n^2(\mathbf{0})].$$

Thus, by (A.2), (A.7), and (A.8), we have

$$\begin{aligned}
&\left|\sum_{\boldsymbol{\ell}\in L_{n1}} E[\widetilde{V}_n^2(\boldsymbol{\ell};\boldsymbol{\Delta}_0)] - E[\widetilde{V}_n^2(\mathbf{0})]\right| \\
&\leq \left|\sum_{\boldsymbol{\ell}\in L_{n1}} E[\widetilde{V}_n^2(\boldsymbol{\ell};\boldsymbol{\Delta}_0)] - E[\widetilde{V}_{n1}^2]\right| + 2E[(\widetilde{V}_{n2} + \widetilde{V}_{n3})^2]^{1/2}E[\widetilde{V}_{n1}^2]^{1/2} + E[(\widetilde{V}_{n2} + \widetilde{V}_{n3})^2] \\
&\lesssim \left(A_n^{(1)} n A_n^{-1}\right)^2 \sum_{\boldsymbol{\ell}_1\neq\boldsymbol{\ell}_2} \alpha^{1-2/q}(\min\{|\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2| - d, 0\}\underline{A}_{n3} + \underline{A}_{n2}; A_n^{(1)}) \\
&\quad + o\left(n^2 A_n^{-1} h_1 \dots h_d\right) \\
&\lesssim \left(A_n^{(1)} n A_n^{-1}\right)^2 \left(\frac{A_n h_1 \dots h_d}{A_n^{(1)}}\right) \\
&\quad \times \left(\alpha^{1-2/q}(\underline{A}_{n2}; A_n^{(1)}) + \sum_{k=1}^{\overline{A}_n/\underline{A}_{n1}} k^{d-1}\alpha^{1-2/q}(\min\{|\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2| - d, 0\}\underline{A}_{n3} + \underline{A}_{n2}; A_n^{(1)})\right) \\
&\quad + o\left(n^2 A_n^{-1} h_1 \dots h_d\right) \\
&= o\left(n^2 A_n^{-1} h_1 \dots h_d\right),
\end{aligned}$$

where $\overline{A}_n = \max_{1\leq j\leq d} A_{n,j}$.
(Step 3) Now we evaluate $B_n(\mathbf{0})$. Decompose

$$\begin{aligned}
B_{n,j_1\dots j_L}(\dot{\boldsymbol{X}}) &= \left\{B_{n,j_1\dots j_L}(\dot{\boldsymbol{X}}) - B_{n,j_1\dots j_L}(\mathbf{0}) - E\left[B_{n,j_1\dots j_L}(\dot{\boldsymbol{X}}) - B_{n,j_1\dots j_L}(\mathbf{0})\right]\right\} \\
&\quad + E\left[B_{n,j_1\dots j_L}(\dot{\boldsymbol{X}}) - B_{n,j_1\dots j_L}(\mathbf{0})\right] \\
&\quad + \{B_{n,j_1\dots j_L}(\mathbf{0}) - E\left[B_{n,j_1\dots j_L}(\mathbf{0})\right]\} \\
&\quad + E\left[B_{n,j_1\dots j_L}(\mathbf{0})\right] \\
&=: \sum_{\ell=1}^{4} B_{n,j_1\dots j_L \ell}.
\end{aligned}$$

Define $N_{\boldsymbol{x}}(h) := \prod_{j=1}^d [x_j - h_j, x_j + h_j]$ and $\boldsymbol{x} = (x_1, \dots, x_d) \in (-1/2, 1/2)^d$. For $B_{n,j_1\dots j_L 1}$, by a change of variables and the dominated convergence theorem, we have

$$\begin{aligned}
&\mathrm{Var}(B_{n,j_1\dots j_L 1}) \\
&\leq \frac{A_n}{\{(p+1)!\}^2 n h_1 \dots h_d} E\left[K_{Ah}^2(\boldsymbol{X}_i) \prod_{\ell=1}^{L}\left(\frac{X_{i,j_\ell}}{A_{n,j_\ell} h_{j_\ell}}\right)^2\right. \\
&\quad \times \sum_{1\leq j_{1,1}\leq\dots\leq j_{1,p+1}\leq d, 1\leq j_{2,1}\leq\dots\leq j_{2,p+1}\leq d} \frac{1}{\boldsymbol{s}_{j_{1,1}\dots j_{1,p+1}}!}\frac{1}{\boldsymbol{s}_{j_{2,1}\dots j_{2,p+1}}!} \\
&\quad \times (\partial_{j_{1,1}\dots j_{1,p+1}} m(\dot{\boldsymbol{X}}_i/A_n) - \partial_{j_{1,1}\dots j_{1,p+1}} m(\mathbf{0}))(\partial_{j_{2,1}\dots j_{2,p+1}} m(\dot{\boldsymbol{X}}_i/A_n) - \partial_{j_{2,1}\dots j_{2,p+1}} m(\mathbf{0}))
\end{aligned}$$

$$\times \prod_{\ell_1=1}^{p+1} \frac{X_{i,j_{1,\ell_1}}}{A_{n,j_{\ell_1}}} \prod_{\ell_2=1}^{p+1} \frac{X_{i,j_{2,\ell_2}}}{A_{n,j_{\ell_2}}} \Bigg]$$

$$\le \frac{A_n}{\{(p+1)!\}^2 n} \max_{1\le j_1\le\cdots\le j_{p+1}\le d} \sup_{\boldsymbol{y}\in N_{\mathbf{0}}(h)} |\partial_{j_1\ldots j_{p+1}} m(\boldsymbol{y}) - \partial_{j_1\ldots j_{p+1}} m(\mathbf{0})|^2$$

$$\times \sum_{1\le j_{1,1}\le\cdots\le j_{1,p+1}\le d, 1\le j_{2,1}\le\cdots\le j_{2,p+1}\le d} \prod_{\ell_1=1}^{p+1} h_{j_{1,\ell_1}} \prod_{\ell_2=1}^{p+1} h_{j_{2,\ell_2}}$$

$$\times \int \left(\prod_{\ell=1}^{L} z_{j_\ell}^2 \prod_{\ell_1=1}^{p+1} |z_{j_{1,\ell_1}}| \prod_{\ell_2=1}^{p+1} |z_{j_{2,\ell_2}}| \right) K^2(\boldsymbol{z}) g(\boldsymbol{z}\circ\boldsymbol{h}) d\boldsymbol{z}$$

$$= o\left(\frac{A_n}{n} \sum_{1\le j_{1,1}\le\cdots\le j_{1,p+1}\le d, 1\le j_{2,1}\le\cdots\le j_{2,p+1}\le d} \prod_{\ell_1=1}^{p+1} h_{j_{1,\ell_1}} \prod_{\ell_2=1}^{p+1} h_{j_{2,\ell_2}} \right)$$

$$= o(1). \tag{A.18}$$

Then we have $B_{n,j_1\ldots j_L 1} = o_p(1)$.

For $B_{n,j_1\ldots j_L 2}$,

$$|B_{n,j_1\ldots j_L 2}|$$

$$\le \frac{1}{(p+1)!} \max_{1\le j_1,\ldots,j_{p+1}\le d} \sup_{\boldsymbol{y}\in N_{\mathbf{0}}(h)} |\partial_{j_1\ldots j_{p+1}} m(\boldsymbol{y}) - \partial_{j_1\ldots j_{p+1}} m(\mathbf{0})|$$

$$\times \sqrt{A_n h_1\ldots h_d} \sum_{1\le j_{1,1}\le\cdots\le j_{1,p+1}\le d} \prod_{\ell_1=1}^{p+1} h_{j_{1,\ell_1}} \int \left(\prod_{\ell=1}^{L} |z_{j_\ell}| \prod_{\ell_1=1}^{p+1} |z_{j_{1,\ell_1}}| \right) |K(\boldsymbol{z})| g(\boldsymbol{z}\circ\boldsymbol{h}) d\boldsymbol{z}$$

$$= o(1). \tag{A.19}$$

For $B_{n,j_1\ldots j_L 3}$,

$$\mathrm{Var}(B_{n,j_1\ldots j_L 3})$$

$$\le \frac{A_n h_1\ldots h_d}{\{(p+1)!\}^2 n h_1\ldots h_d} \sum_{1\le j_{1,1}\le\cdots\le j_{1,p+1}\le d, 1\le j_{2,1}\le\cdots\le j_{2,p+1}\le d} \partial_{j_{1,1}\ldots j_{1,p+1}} m(\mathbf{0}) \partial_{j_{2,1}\ldots j_{2,p+1}} m(\mathbf{0})$$

$$\times \prod_{\ell_1=1}^{p+1} h_{j_{1,\ell_1}} \prod_{\ell_2=1}^{p+1} h_{j_{2,\ell_2}} \int \left(\prod_{\ell=1}^{L} z_{j_\ell}^2 \prod_{\ell_1}^{p+1} |z_{j_{1,\ell_1}}| \prod_{\ell_2=1}^{p+1} |z_{j_{2,\ell_2}}| \right) K^2(\boldsymbol{z}) g(\boldsymbol{z}\circ\boldsymbol{h}) d\boldsymbol{z}$$

$$= O\left(\frac{A_n}{n} \sum_{1\le j_{1,1}\le\cdots\le j_{1,p+1}\le d, 1\le j_{2,1}\le\cdots\le j_{2,p+1}\le d} \prod_{\ell_1=1}^{p+1} h_{j_{1,\ell_1}} \prod_{\ell_2=1}^{p+1} h_{j_{2,\ell_2}} \right). \tag{A.20}$$

Then we have $B_{n,j_1\ldots j_L 3} = o_p(1)$.

For $B_{n,j_1\ldots j_L 4}$,

$$B_{n,j_1\ldots j_L 4} = \sqrt{A_n h_1\ldots h_d} \sum_{1\le j_{1,1}\le\cdots\le j_{1,p+1}\le d} \frac{\partial_{j_{1,1}\ldots j_{1,p+1}} m(\mathbf{0})}{\boldsymbol{s}_{j_{1,1}\ldots j_{1,p+1}}!}$$

$$\times \prod_{\ell_1=1}^{p+1} h_{j_{1,\ell_1}} \int \left(\prod_{\ell=1}^{L} z_{j_\ell} \prod_{\ell_1=1}^{p+1} z_{j_{1,\ell_1}} \right) K(\boldsymbol{z}) g(\boldsymbol{z} \circ \boldsymbol{h}) d\boldsymbol{z}$$

$$= g(\mathbf{0}) \sqrt{A_n h_1 \dots h_d} \sum_{1 \le j_{1,1} \le \dots \le j_{1,p+1} \le d} \frac{\partial_{j_{1,1} \dots j_{1,p+1}} m(\mathbf{0})}{\boldsymbol{s}_{j_{1,1} \dots j_{1,p+1}}!} \prod_{\ell_1=1}^{p+1} h_{j_{1,\ell_1}} \kappa^{(1)}_{j_1 \dots j_L j_{1,1} \dots j_{1,p+1}} + o(1). \tag{A.21}$$

Combining (A.18)-(A.21),

$$B_{n,j_1 \dots j_L}(\dot{\boldsymbol{X}}) = g(\mathbf{0}) \sqrt{A_n h_1 \dots h_d} \sum_{1 \le j_{1,1} \le \dots \le j_{1,p+1} \le d} \frac{\partial_{j_{1,1} \dots j_{1,p+1}} m(\mathbf{0})}{\boldsymbol{s}_{j_{1,1} \dots j_{1,p+1}}!}$$

$$\times \prod_{\ell_1=1}^{p+1} h_{j_{1,\ell_1}} \kappa^{(1)}_{j_1 \dots j_L j_{1,1} \dots j_{1,p+1}} + o_p(1)$$

$$= g(\mathbf{0}) \sqrt{A_n h_1 \dots h_d} (B^{(d,p)} \boldsymbol{M}_n^{(d,p)}(\mathbf{0}))_{j_1 \dots j_L} + o_p(1).$$

(Step 4) Combining the results in Steps 2 and 3, we have

$$A_n(\mathbf{0}) := V_n(\mathbf{0}) + \left(B_n(\mathbf{0}) - g(\mathbf{0}) \sqrt{A_n h_1 \dots h_d} B^{(d,p)} M_n^{(d,p)}(\mathbf{0}) \right)$$

$$\xrightarrow{d} N \left(\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, g(\mathbf{0}) \left\{ \kappa(\eta^2(\mathbf{0}) + \sigma_\varepsilon^2(\mathbf{0})) + \eta^2(\mathbf{0}) g(\mathbf{0}) \int \sigma_{\boldsymbol{e}}(\boldsymbol{v}) d\boldsymbol{v} \right\} \mathcal{K} \right).$$

This and the result in Step 1 yield the desired result. □

A.2. **Proof of Proposition 4.1.**

Proof. It is easy to see that $\widehat{g}(\mathbf{0}) \xrightarrow{p} g(\mathbf{0})$ as $n \to \infty$. For $\widehat{V}_{n,1}(\mathbf{0})$, observe that

$$E\left[\left(Y(\boldsymbol{X}_i) - \widehat{m}_{-\{i\}}(\boldsymbol{X}_i/A_n) \right)^2 \Big| \boldsymbol{X}_i \right] = E\left[\left(m(\boldsymbol{X}_i/A_n) - \widehat{m}_{-\{i\}}(\boldsymbol{X}_i/A_n) \right)^2 \Big| \boldsymbol{X}_i \right]$$

$$+ 2E\left[\left(m(\boldsymbol{X}_i/A_n) - \widehat{m}_{-\{i\}}(\boldsymbol{X}_i/A_n) \right) (e_{n,i} + \varepsilon_{n,i}) \Big| \boldsymbol{X}_i \right]$$

$$+ E\left[(e_{n,i} + \varepsilon_{n,i})^2 \Big| \boldsymbol{X}_i \right]$$

$$=: \widehat{V}_{n1,i} + \widehat{V}_{n2,i} + E\left[(e_{n,i} + \varepsilon_{n,i})^2 \Big| \boldsymbol{X}_i \right].$$

The representation of the MSE of $\partial_{j_1 \dots j_L} \widehat{m}(\mathbf{0})$ (4.1) implies that for $\boldsymbol{z} \in (-1/2, 1/2)^d$,

$$\text{MSE}(\widehat{m}(\boldsymbol{z})) = \left\{ (S^{-1} e_0)' B^{(d,p)} M_n^{(d,p)}(\boldsymbol{z}) \right\}^2$$

$$+ \left(\frac{\kappa(\eta^2(\boldsymbol{z}) + \sigma_\varepsilon^2(\boldsymbol{z}))}{g(\boldsymbol{z})} + \eta^2(\boldsymbol{z}) \int \sigma_{\boldsymbol{e}}(\boldsymbol{v}) d\boldsymbol{v} \right) \frac{e_0' S^{-1} \mathcal{K} S^{-1} e_0}{A_n h_1 \dots h_d}$$

$$= O\left(\left(\sum_{1 \le j_1 \le \dots \le j_{p+1} \le d} \prod_{\ell=1}^{p+1} h_{j_\ell} \right)^2 + \frac{1}{A_n h_1 \dots h_d} \right),$$

where $e_0 = (1, 0, \dots, 0) \in \mathbb{R}^D$. Then by Cauchy-Schwarz inequality, we have

$$\max_{1 \le i \le n} K_{Ah}(\boldsymbol{X}_i) \left(\widehat{V}_{n1,i} + \widehat{V}_{n2,i} \right)$$

$$= O\left(\left(\sum_{1\le j_1\le\cdots\le j_{p+1}\le d}\prod_{\ell=1}^{p+1}h_{j_\ell}\right)+\sqrt{\frac{1}{A_nh_1\dots h_d}}\right)=o(1)\ a.s.,\ n\to\infty.$$

Applying a similar argument in the proof of Theorem 4.1, this implies that

$$\begin{aligned}\widehat{V}_{n,1}(\mathbf{0}) &= \frac{1}{nh_1\dots h_d}\sum_{i=1}^{n}K_{Ah}(\boldsymbol{X}_i)(e_{n,i}+\varepsilon_{n,i})^2+o_p(1)\\ &= \frac{1}{nh_1\dots h_d}\sum_{i=1}^{n}K_{Ah}(\boldsymbol{X}_i)\left(\eta^2(\boldsymbol{X}_i/A_n)e^2(\boldsymbol{X}_i)+\sigma_\varepsilon^2(\boldsymbol{X}_i/A_n)\varepsilon_i^2\right)+o_p(1)\\ &= \left(\eta^2(\mathbf{0})+\sigma_\varepsilon^2(\mathbf{0})\right)g(\mathbf{0})+o_p(1),\ n\to\infty.\end{aligned}$$

Likewise,

$$\begin{aligned}\widehat{V}_{n,2}(\mathbf{0}) &= \frac{A_n}{nh_1\dots h_d}\sum_{i=1}^{n-1}K_{Ah}(\boldsymbol{X}_i)K_{Ah}(\boldsymbol{X}_{i+1})(e_{n,i}+\varepsilon_{n,i})(e_{n,i+1}+\varepsilon_{n,i+1})+o_p(1)\\ &= \frac{A_n}{nh_1\dots h_d}\sum_{i=1}^{n-1}K_{Ah}(\boldsymbol{X}_i)K_{Ah}(\boldsymbol{X}_{i+1})\eta(\boldsymbol{X}_i/A_n)\eta(\boldsymbol{X}_{i+1}/A_n)e(\boldsymbol{X}_i)e(\boldsymbol{X}_{i+1})+o_p(1)\\ &= \kappa_0^{(2)}\eta^2(\mathbf{0})g^2(\mathbf{0})\int\sigma_{\boldsymbol{e}}(\boldsymbol{v})d\boldsymbol{v}+o_p(1),\ n\to\infty.\end{aligned}$$

Therefore, as $n\to\infty$,

$$\widehat{V}_n(\mathbf{0}):=\frac{(A_n/n)\widehat{V}_{n,1}(\mathbf{0})+(\widehat{V}_{n,2}(\mathbf{0})/\kappa_0^{(2)})}{\widehat{g}^2(\mathbf{0})}\xrightarrow{p}\frac{\kappa(\eta^2(\mathbf{0})+\sigma_\varepsilon^2(\mathbf{0}))}{g(\mathbf{0})}+\eta^2(\mathbf{0})\int\sigma_{\boldsymbol{e}}(\boldsymbol{v})d\boldsymbol{v}.$$

□

A.3. **Proof of Corollary 4.1.**

Proof. Corollary 4.1 follows immediately from Theorem 4.1 and Proposition 4.1. □

A.4. **Proof of Proposition 4.2.**

Proof. For any $\boldsymbol{t}=(t_0,t_1,\dots,t_d,t_{11},\dots,t_{dd},\dots,t_{1\dots1},\dots,t_{d\dots d})'\in\mathbb{R}^D$, we define

$$\overline{W}_{n1}(\mathbf{0}):=\sum_{\ell_1=1}^{n_1}K_{Ah}\left(\boldsymbol{X}_{1,\ell_1}\right)\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_{1,\ell_1}\end{pmatrix}\right]\underbrace{\left(\eta_1\left(\frac{\boldsymbol{X}_{1,\ell_1}}{A_n}\right)e_1(\boldsymbol{X}_{1,\ell_1})+\sigma_{\varepsilon,1}\left(\frac{\boldsymbol{X}_{1,\ell_1}}{A_n}\right)\varepsilon_{1,\ell_1}\right)}_{=:\overline{e}_{n1,\ell_1}+\overline{\varepsilon}_{n1,\ell_1}},$$

$$\overline{W}_{n2}(\mathbf{0}):=\sum_{\ell_2=1}^{n_2}K_{Ah}\left(\boldsymbol{X}_{2,\ell_2}\right)\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_{2,\ell_2}\end{pmatrix}\right]\underbrace{\left(\eta_2\left(\frac{\boldsymbol{X}_{2,\ell_2}}{A_n}\right)e_2(\boldsymbol{X}_{2,\ell_2})+\sigma_{\varepsilon,2}\left(\frac{\boldsymbol{X}_{2,\ell_2}}{A_n}\right)\varepsilon_{2,\ell_2}\right)}_{=:\overline{e}_{n2,\ell_2}+\overline{\varepsilon}_{n2,\ell_2}}.$$

By inspecting the proof of Theorem 4.1, to show Proposition 4.2, it is sufficient to verify

$$\begin{aligned}&E\left[\left(\overline{W}_{n1}(\mathbf{0})-\overline{W}_{n1}(\mathbf{0})\right)^2\right]/(h_1\dots h_d)\\ &=\Bigg(n_1\left\{(\eta_1^2(\mathbf{0})+\sigma_{\varepsilon,1}^2(\mathbf{0}))g_1(\mathbf{0})+n_1A_n^{-1}\eta_1^2(\mathbf{0})g_1^2(\mathbf{0})\int\sigma_{\boldsymbol{e},11}(\boldsymbol{v})d\boldsymbol{v}\right\}\\ &\quad+n_2\left\{(\eta_2^2(\mathbf{0})+\sigma_{\varepsilon,2}^2(\mathbf{0}))g_2(\mathbf{0})+n_2A_n^{-1}\eta_2^2(\mathbf{0})g_2^2(\mathbf{0})\int\sigma_{\boldsymbol{e},22}(\boldsymbol{v})d\boldsymbol{v}\right\}\end{aligned}$$

$$-2n_1n_2A_n^{-1}\eta_1(\mathbf{0})\eta_2(\mathbf{0})g_1(\mathbf{0})g_2(\mathbf{0})\int\sigma_{\boldsymbol{e},12}(\boldsymbol{v})d\boldsymbol{v}\Bigg)$$
$$\times\left(\int K^2(\boldsymbol{z})\left[\boldsymbol{t}'\begin{pmatrix}1\\ \check{\boldsymbol{z}}\end{pmatrix}\right]^2 d\boldsymbol{z}\right)(1+o(1)),\ n\to\infty.$$

Let $E_{\boldsymbol{X}_{12}}$ denote the expectation with respect to $\{\boldsymbol{X}_{1,\ell_1}\}$ and $\{\boldsymbol{X}_{2,\ell_2}\}$ and let $E_{\cdot|\boldsymbol{X}_{12}}$ denote the conditional expectation given $\sigma(\{\boldsymbol{X}_{1,\ell_1}\}\cup\{\boldsymbol{X}_{2,\ell_2}\})$. Observe that

$$\begin{aligned}
&E_{\cdot|\boldsymbol{X}_{12}}\left[\left(\overline{W}_{n1}(\mathbf{0})-\overline{W}_{n2}(\mathbf{0})\right)^2\right]\\
&=\sum_{\ell_{11},\ell_{12}=1}^{n_1}E_{\cdot|\boldsymbol{X}_{12}}\left[K_{Ah}\left(\boldsymbol{X}_{1,\ell_{11}}\right)K_{Ah}\left(\boldsymbol{X}_{1,\ell_{12}}\right)\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_{1,\ell_{11}}\end{pmatrix}\right]\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_{1,\ell_{12}}\end{pmatrix}\right]\right.\\
&\quad\left.\times(\overline{e}_{n1,\ell_{11}}+\overline{\varepsilon}_{n1,\ell_{11}})(\overline{e}_{n1,\ell_{12}}+\overline{\varepsilon}_{n1,\ell_{12}})\right]\\
&\quad+\sum_{\ell_{21},\ell_{22}=1}^{n_2}E_{\cdot|\boldsymbol{X}_{12}}\left[K_{Ah}\left(\boldsymbol{X}_{2,\ell_{21}}\right)K_{Ah}\left(\boldsymbol{X}_{2,\ell_{22}}\right)\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_{2,\ell_{21}}\end{pmatrix}\right]\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_{2,\ell_{22}}\end{pmatrix}\right]\right.\\
&\quad\left.\times(\overline{e}_{n2,\ell_{21}}+\overline{\varepsilon}_{n2,\ell_{21}})(\overline{e}_{n2,\ell_{22}}+\overline{\varepsilon}_{n2,\ell_{22}})\right]\\
&\quad-2\sum_{\ell_1=1}^{n_1}\sum_{\ell_2=1}^{n_2}E_{\cdot|\boldsymbol{X}_{12}}\left[K_{Ah}\left(\boldsymbol{X}_{1,\ell_1}\right)K_{Ah}\left(\boldsymbol{X}_{2,\ell_2}\right)\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_{1,\ell_1}\end{pmatrix}\right]\left[\boldsymbol{t}'H^{-1}\begin{pmatrix}1\\ \check{\boldsymbol{X}}_{2,\ell_2}\end{pmatrix}\right]\right.\\
&\quad\left.\times(\overline{e}_{n1,\ell_1}+\overline{\varepsilon}_{n1,\ell_1})(\overline{e}_{n2,\ell_2}+\overline{\varepsilon}_{n2,\ell_2})\right]\\
&=:\overline{W}_{n11}+\overline{W}_{n12}-2\overline{W}_{n13}.
\end{aligned}$$

Applying the same argument in Step 2 of the proof of Theorem 4.1, we have

$$\begin{aligned}
E_{\boldsymbol{X}_{12}}[\overline{W}_{n1\ell}]&=n_\ell h_1\dots h_d g_\ell(\mathbf{0})\left\{(\eta_\ell^2(\mathbf{0})+\sigma_{\varepsilon,\ell}^2(\mathbf{0}))+n_\ell A_n^{-1}\eta_\ell^2(\mathbf{0})g_\ell(\mathbf{0})\int\sigma_{\boldsymbol{e},\ell\ell}(\boldsymbol{v})d\boldsymbol{v}\right\}\\
&\quad\times\left(\int K^2(\boldsymbol{z})\left[\boldsymbol{t}'\begin{pmatrix}1\\ \check{\boldsymbol{z}}\end{pmatrix}\right]^2 d\boldsymbol{z}\right)(1+o(1)),\ \ell=1,2,\\
E_{\boldsymbol{X}_{12}}[\overline{W}_{n13}]&=n_1n_2A_n^{-1}h_1\dots h_d\left(\eta_1(\mathbf{0})\eta_2(\mathbf{0})g_1(\mathbf{0})g_2(\mathbf{0})\int\sigma_{\boldsymbol{e},12}(\boldsymbol{v})d\boldsymbol{v}\right)(1+o(1))
\end{aligned}$$

as $n\to\infty$. Therefore, we obtain the desired result. □

A.5. **Proof of Proposition 4.3.**

Proof. Applying the same argument in the proof of Proposition 4.1, we have that as $n\to\infty$,

$$\begin{aligned}
\overline{g}_{n_k}(\mathbf{0})&=g_k(\mathbf{0})+o_p(1),\ k=1,2,\\
\overline{V}_{n,1k}(\mathbf{0})&=(\eta_k^2(\mathbf{0})+\sigma_{\varepsilon,k}^2(\mathbf{0}))g_k(\mathbf{0})+o_p(1),\ k=1,2,\\
\overline{V}_{n,2k}(\mathbf{0})&=\kappa_0^{(2)}\eta_k^2(\mathbf{0})g_k^2(\mathbf{0})\int\sigma_{\boldsymbol{e},kk}(\boldsymbol{v})d\boldsymbol{v}+o_p(1),\ k=1,2,\\
\overline{V}_{n,3}(\mathbf{0})&=\kappa_0^{(2)}\eta_1(\mathbf{0})\eta_2(\mathbf{0})g_1(\mathbf{0})g_2(\mathbf{0})\int\sigma_{\boldsymbol{e},12}(\boldsymbol{v})d\boldsymbol{v}+o_p(1).
\end{aligned}$$

Therefore, $\check{V}_n(\mathbf{0})\stackrel{p}{\to}\overline{V}_1(\mathbf{0})+\overline{V}_2(\mathbf{0})-2\overline{V}_3(\mathbf{0})$ as $n\to\infty$. □

A.6. **Proof of Corollary 4.2.**

Proof. Corollary 4.2 follows immediately from Propositions 4.2 and 4.3. □

Appendix B. Proofs for Section 5

B.1. **Proof of Proposition 5.1.**

Proof. We only provide the proof of (5.6) since the proof of (5.7) is almost the same. Let $a_n = \sqrt{\frac{\log n}{n^2 A_n^{-1} h_1 \ldots h_d}}$ and $\tau_n = \rho_n n^{1/q_2}$ with $\rho_n = (\log n)^\iota$ for some $\iota > 0$. Define

$$\widehat{\Psi}_1(\boldsymbol{z}) = \frac{|f_2^{-1}(h_1,\ldots,h_d)|}{n^2 A_n^{-1} h_1 \ldots h_d} \sum_{i=1}^n K_{Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}) \times f_{1,Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}) f_{2,A}(\boldsymbol{X}_i - A_n\boldsymbol{z}) f_{3,A}(\boldsymbol{X}_i) Z_{\boldsymbol{X}_i} 1\{|Z_{\boldsymbol{X}_i}| \leq \tau_n\},$$

$$\widehat{\Psi}_2(\boldsymbol{z}) = \frac{|f_2^{-1}(h_1,\ldots,h_d)|}{n^2 A_n^{-1} h_1 \ldots h_d} \sum_{i=1}^n K_{Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}) \times f_{1,Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}) f_{2,A}(\boldsymbol{X}_i - A_n\boldsymbol{z}) f_{3,A}(\boldsymbol{X}_i) Z_{\boldsymbol{X}_i} 1\{|Z_{\boldsymbol{X}_i}| > \tau_n\}.$$

Note that

$$\widehat{\Psi}(\boldsymbol{z}) - E[\widehat{\Psi}(\boldsymbol{z})] = \widehat{\Psi}_1(\boldsymbol{z}) - E[\widehat{\Psi}_1(\boldsymbol{z})] + \widehat{\Psi}_2(\boldsymbol{z}) - E[\widehat{\Psi}_2(\boldsymbol{z})].$$

(Step 1) First we consider the term $\widehat{\Psi}_2(\boldsymbol{z}) - E[\widehat{\Psi}_2(\boldsymbol{z})]$. Observe that

$$\begin{aligned} P\left(\sup_{\boldsymbol{z}\in R_0} |\widehat{\Psi}_2(\boldsymbol{z})| > a_n\right) &\leq P\left(|Z_{\boldsymbol{X}_i}| > \tau_n \text{ for some } i = 1,\ldots,n\right) \\ &\leq \tau_n^{-q_2} \sum_{i=1}^n E\left[E_{\cdot|\boldsymbol{X}}[|Z_{\boldsymbol{X}_i}|^{q_2}]\right] \leq n\tau_n^{-q_2} = \rho_n^{-q_2} \to 0. \end{aligned}$$

Further, for $\boldsymbol{z} \in [-1/2, 1/2]^d$,

$$\begin{aligned} &E\left[\left|\widehat{\Psi}_2(\boldsymbol{z})\right|\right] \\ &\leq \frac{|f_2^{-1}(h_1,\ldots,h_d)|}{n^2 A_n^{-1} h_1 \ldots h_d} \sum_{i=1}^n E\left[|K_{Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z})|\right. \\ &\quad \left.\times |f_{1,Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}) f_{2,A}(\boldsymbol{X}_i - A_n\boldsymbol{z})| f_{3,A}(\boldsymbol{X}_i) E_{\cdot|\boldsymbol{X}}[|Z_{\boldsymbol{X}_i}| 1\{|Z_{\boldsymbol{X}_i}| > \tau_n\}]\right] \\ &\lesssim \frac{nA_n^{-1}|f_2^{-1}(h_1,\ldots,h_d)|}{n^2 A_n^{-1} h_1 \ldots h_d \tau_n^{q_2-1}} \int_{R_n} |K_{Ah}(\boldsymbol{x} - A_n\boldsymbol{z})| |f_{1,Ah}(\boldsymbol{x} - A_n\boldsymbol{z}) f_{2,A}(\boldsymbol{X}_i - A_n\boldsymbol{z})| \\ &\quad \times f_{3,A}(\boldsymbol{x}) g(\boldsymbol{x}/A_n) d\boldsymbol{x} \\ &= \frac{|f_2^{-1}(h_1,\ldots,h_d)|}{nA_n^{-1}\tau_n^{q_2-1}} \int_{\boldsymbol{h}^{-1}(R_0-\boldsymbol{z})} |K(\boldsymbol{v})| |f_1(\boldsymbol{v}) f_2(\boldsymbol{v}\circ\boldsymbol{h})| f_3(\boldsymbol{z}+\boldsymbol{v}\circ\boldsymbol{h}) g(\boldsymbol{z}+\boldsymbol{v}\circ\boldsymbol{h}) d\boldsymbol{v} \\ &\lesssim \frac{1}{nA_n^{-1}\tau_n^{q_2-1}} \lesssim \frac{1}{\tau_n^{q_2-1}} \lesssim a_n. \end{aligned}$$

Then we have

$$\sup_{\boldsymbol{z}\in R_0} \left|\widehat{\Psi}(\boldsymbol{z}) - E[\widehat{\Psi}(\boldsymbol{z})]\right| = O_p(a_n).$$

(Step 2) Now we consider the term $\widehat{\Psi}_1(\boldsymbol{z}) - E[\widehat{\Psi}_1(\boldsymbol{z})]$.

Define

$$\Psi_{1,\boldsymbol{X}_i}(\boldsymbol{z}) = K_{Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z})f_{1,Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}) f_{2,A}(\boldsymbol{X}_i - A_n\boldsymbol{z}) f_{3,A}(\boldsymbol{X}_i) Z_{\boldsymbol{X}_i} 1\{|Z_{\boldsymbol{X}_i}| \le \tau_n\}$$
$$- E\left[K_{Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z})f_{1,Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}) f_{2,A}(\boldsymbol{X}_i - A_n\boldsymbol{z}) f_{3,A}(\boldsymbol{X}_i) Z_{\boldsymbol{X}_i} 1\{|Z_{\boldsymbol{X}_i}| \le \tau_n\}\right].$$

Observe that

$$\sum_{i=1}^{n} \Psi_{1,\boldsymbol{X}_i}(\boldsymbol{z}) = \sum_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z})} \Psi_1^{(\boldsymbol{\ell};\boldsymbol{\Delta}_0)}(\boldsymbol{z}) + \sum_{\boldsymbol{\Delta}\neq\boldsymbol{\Delta}_0} \sum_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z})} \Psi_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}) + \sum_{\boldsymbol{\Delta}\in\{1,2\}^d} \sum_{\boldsymbol{\ell}\in L_{n2}(\boldsymbol{z})} \Psi_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}),$$

where

$$\Psi_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}) = \sum_{i=1}^{n} \Psi_{1,\boldsymbol{X}_i}(\boldsymbol{z}) 1\{\boldsymbol{X}_i \in \Gamma_{n,\boldsymbol{z}}(\boldsymbol{\ell};\boldsymbol{\Delta}) \cap R_n \cap (\boldsymbol{h}R_n + A_n\boldsymbol{z})\}.$$

For $\boldsymbol{\Delta} \in \{1,2\}^d$, let $\{\widetilde{\Psi}_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z})\}_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z})\cup L_{n2}(\boldsymbol{z})}$ be independent random variables such that $\Psi_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}) \overset{d}{=} \widetilde{\Psi}_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z})$. Applying Lemma D.2 below with $M_h = 1$, $m \sim \left(\frac{A_n h_1 \dots h_d}{A_n^{(1)}}\right)$ and $\tau \sim \beta(\underline{A}_{n2}; A_n h_1 \dots h_d)$, we have that for $\boldsymbol{\Delta} \in \{1,2\}^d$,

$$\sup_{t>0}\left| P\left(\left|\sum_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z})} \Psi_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z})\right| > t\right) - P\left(\left|\sum_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z})} \widetilde{\Psi}_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z})\right| > t\right)\right|$$
$$\lesssim \left(\frac{A_n h_1 \dots h_d}{A_n^{(1)}}\right) \beta(\underline{A}_{n2}; A_n h_1 \dots h_d), \tag{B.1}$$

$$\sup_{t>0}\left| P\left(\left|\sum_{\boldsymbol{\ell}\in L_{n2}(\boldsymbol{z})} \Psi_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z})\right| > t\right) - P\left(\left|\sum_{\boldsymbol{\ell}\in L_{n2}(\boldsymbol{z})} \widetilde{\Psi}_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z})\right| > t\right)\right|$$
$$\lesssim \left(\frac{A_n h_1 \dots h_d}{A_n^{(1)}}\right) \beta(\underline{A}_{n2}; A_n h_1 \dots h_d). \tag{B.2}$$

Since $\left(\frac{A_n h_1 \dots h_d}{A_n^{(1)}}\right) \beta(\underline{A}_{n2}; A_n h_1 \dots h_d) \to 0$ as $n \to \infty$, these results imply that

$$\sum_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z})} \Psi_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}) = O_p\left(\sum_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z})} \widetilde{\Psi}_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z})\right),$$
$$\sum_{\boldsymbol{\ell}\in L_{n2}(\boldsymbol{z})} \Psi_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}) = O_p\left(\sum_{\boldsymbol{\ell}\in L_{n2}(\boldsymbol{z})} \widetilde{\Psi}_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z})\right).$$

Now we show $\sup_{\boldsymbol{z}\in R_0}\left|\widehat{\Psi}_1(\boldsymbol{z}) - E[\widehat{\Psi}_1(\boldsymbol{z})]\right| = O_p(a_n)$. Cover the region R_0 with $N \le (h_1 \dots h_d)^{-1} a_n^{-d}$ balls $B_k = \{\boldsymbol{z} \in \mathbb{R}^d : |z_j - z_{k,j}| \le a_n h_j\}$ and use $\boldsymbol{z}_k = (z_{k,1}, \dots, z_{k,d})$ to denote the mid point of B_k, $k = 1, \dots, N$. In addition, let $K^*(\boldsymbol{v}) = C^* \prod_{j=1}^{d} I(|v_j| \le 2C_K)$ for $\boldsymbol{v} \in \mathbb{R}^d$ and sufficiently large $C^* > 0$. Note that for $\boldsymbol{z} \in B_k$ and sufficiently large n,

$$|K_{Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}) f_{1,Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}) - K_{Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}_k) f_{1,Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}_k)|$$
$$\le a_n K^*_{Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}_k).$$

For $\boldsymbol{\ell} \in L_{n1}(\boldsymbol{z}) \cup L_{n2}(\boldsymbol{z})$ and $\boldsymbol{\Delta} \in \{1,2\}^d$, define

$$\Psi_2^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}) = \sum_{i=1}^{n} \Psi_{2,\boldsymbol{X}_i}(\boldsymbol{z}) 1\{\boldsymbol{X}_i \in \Gamma_{n,\boldsymbol{z}}(\boldsymbol{\ell};\boldsymbol{\Delta}) \cap R_n \cap (\boldsymbol{h}R_n + A_n\boldsymbol{z})\},$$

where

$$\begin{aligned}\Psi_{2,\boldsymbol{X}_i}(\boldsymbol{z}) &= K^*_{Ah}\left(\boldsymbol{X}_i - A_n\boldsymbol{z}_n\right) f_{2,A}\left(\boldsymbol{X}_i - A_n\boldsymbol{z}\right) f_{3,A}\left(\boldsymbol{X}_i\right) Z_{\boldsymbol{X}_i} 1\{|Z_{\boldsymbol{X}_i}| \leq \tau_n\} \\ &\quad - E\left[K^*_{Ah}\left(\boldsymbol{X}_i - A_n\boldsymbol{z}_n\right) f_{2,A}\left(\boldsymbol{X}_i - A_n\boldsymbol{z}\right) f_{3,A}\left(\boldsymbol{X}_i\right) Z_{\boldsymbol{X}_i} 1\{|Z_{\boldsymbol{X}_i}| \leq \tau_n\}\right].\end{aligned}$$

Moreover, define

$$\bar{\Psi}_1(\boldsymbol{z}) = \frac{|f_2^{-1}(h_1,\ldots,h_d)|}{n^2 A_n^{-1} h_1 \ldots h_d} \sum_{i=1}^{n} K^*_{Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}) f_{2,A}\left(\boldsymbol{X}_i - A_n\boldsymbol{z}\right) f_{3,A}\left(\boldsymbol{X}_i\right) Z_{\boldsymbol{X}_i} 1\{|Z_{\boldsymbol{X}_i}| \leq \tau_n\}.$$

Observe that for $\boldsymbol{z} \in R_0$,

$$\begin{aligned}E\left[|\bar{\Psi}_1(\boldsymbol{z})|\right] &\lesssim \frac{A_n^{-1}|f_2^{-1}(h_1,\ldots,h_d)|}{nA_n^{-1}h_1\ldots h_d} \int_{R_n} |K^*_{Ah}(\boldsymbol{x} - A_n\boldsymbol{z}) f_{2,A}(\boldsymbol{x} - A_n\boldsymbol{z}) f_{3,A}(\boldsymbol{x})| g(\boldsymbol{x}/A_n) d\boldsymbol{x} \\ &= \frac{|f_2^{-1}(h_1,\ldots,h_d)|}{nA_n^{-1}} \int_{\boldsymbol{h}^{-1}(R_0-\boldsymbol{z})} |K^*(\boldsymbol{v})||f_2(\boldsymbol{v}\circ\boldsymbol{h})||f_3(\boldsymbol{z}+\boldsymbol{v}\circ\boldsymbol{h}) g(\boldsymbol{z}+\boldsymbol{v}\circ\boldsymbol{h})| d\boldsymbol{v} \\ &\lesssim \frac{1}{nA_n^{-1}} \leq M.\end{aligned}$$

for sufficiently large $M > 0$. Then we have

$$\begin{aligned}&\sup_{\boldsymbol{z}\in B_k} \left|\widehat{\Psi}_1(\boldsymbol{z}) - E[\widehat{\Psi}_1(\boldsymbol{z})]\right| \\ &\leq \left|\widehat{\Psi}_1(\boldsymbol{z}_k) - E[\widehat{\Psi}_1(\boldsymbol{z}_k)]\right| + a_n\left(\left|\bar{\Psi}_1(\boldsymbol{z}_k)\right| + E\left[\left|\bar{\Psi}_1(\boldsymbol{z}_k)\right|\right]\right) \\ &\leq \left|\widehat{\Psi}_1(\boldsymbol{z}_k) - E[\widehat{\Psi}_1(\boldsymbol{z}_k)]\right| + \left|\bar{\Psi}_1(\boldsymbol{z}_k) - E[\bar{\Psi}_1(\boldsymbol{z}_k)]\right| + 2Ma_n \\ &\leq \frac{|f_2^{-1}(h_1,\ldots,h_d)|}{n^2A_n^{-1}h_1\ldots h_d}\left(\left|\sum_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z}_k)} \Psi_1^{(\boldsymbol{\ell};\boldsymbol{\Delta}_0)}(\boldsymbol{z}_k)\right| + \sum_{\boldsymbol{\Delta}\neq\boldsymbol{\Delta}_0}\left|\sum_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z}_k)} \Psi_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}_k)\right| + \sum_{\boldsymbol{\Delta}\in\{1,2\}^d}\left|\sum_{\boldsymbol{\ell}\in L_{n2}(\boldsymbol{z}_k)} \Psi_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}_k)\right|\right) \\ &+ \frac{|f_2^{-1}(h_1,\ldots,h_d)|}{n^2A_n^{-1}h_1\ldots h_d}\left(\left|\sum_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z}_k)} \Psi_2^{(\boldsymbol{\ell};\boldsymbol{\Delta}_0)}(\boldsymbol{z}_k)\right| + \sum_{\boldsymbol{\Delta}\neq\boldsymbol{\Delta}_0}\left|\sum_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z}_k)} \Psi_2^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}_k)\right| + \sum_{\boldsymbol{\Delta}\in\{1,2\}^d}\left|\sum_{\boldsymbol{\ell}\in L_{n2}(\boldsymbol{z}_k)} \Psi_2^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}_k)\right|\right) \\ &+ 2Ma_n.\end{aligned}$$

For $\boldsymbol{\Delta} \in \{1,2\}^d$, let $\{\widetilde{\Psi}_2^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z})\}_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z})\cup L_{n2}(\boldsymbol{z})}$ be independent random variables such that $\Psi_2^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}) \stackrel{d}{=} \widetilde{\Psi}_2^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z})$. From (B.1) and (B.2), and applying Lemma D.2 below to $\{\widetilde{\Psi}_2^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z})\}_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z})\cup L_{n2}(\boldsymbol{z})}$, we have

$$\begin{aligned}&P\left(\sup_{\boldsymbol{z}\in R_0}\left|\widehat{\Psi}_1(\boldsymbol{z}) - E[\widehat{\Psi}_1(\boldsymbol{z})]\right| > 2^{d+3}Ma_n\right) \\ &\leq N \max_{1\leq k\leq N} P\left(\sup_{\boldsymbol{z}\in B_k}\left|\widehat{\Psi}_1(\boldsymbol{z}) - E[\widehat{\Psi}_1(\boldsymbol{z})]\right| > 2^{d+3}Ma_n\right) \\ &\leq \sum_{\boldsymbol{\Delta}\in\{1,2\}^d} \widehat{Q}_{n1}(\boldsymbol{\Delta}) + \sum_{\boldsymbol{\Delta}\in\{1,2\}^d} \widehat{Q}_{n2}(\boldsymbol{\Delta}) + \sum_{\boldsymbol{\epsilon}\in\{1,2\}^d} \bar{Q}_{n1}(\boldsymbol{\Delta}) + \sum_{\boldsymbol{\epsilon}\in\{1,2\}^d} \bar{Q}_{n2}(\boldsymbol{\Delta})\end{aligned}$$

$$+ 2^{d+2}N\left(\frac{A_n h_1 \dots h_d}{A_n^{(1)}}\right)\beta(\underline{A}_{n2}; A_n h_1 \dots h_d),$$

where

$$\widehat{Q}_{nj}(\boldsymbol{\Delta}) = N \max_{1\le k\le N} P\left(\left|\sum_{\boldsymbol{\ell}\in L_{nj}(\boldsymbol{z}_k)} \widetilde{\Psi}_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}_k)\right| > M a_n \frac{n^2 A_n^{-1} h_1 \dots h_d}{|f_2^{-1}(h_1,\dots,h_d)|}\right),\ j = 1,2,$$

$$\bar{Q}_{nj}(\boldsymbol{\Delta}) = N \max_{1\le k\le N} P\left(\left|\sum_{\boldsymbol{\ell}\in L_{nj}(\boldsymbol{z}_k)} \widetilde{\Psi}_2^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}_k)\right| > M a_n \frac{n^2 A_n^{-1} h_1 \dots h_d}{|f_2^{-1}(h_1,\dots,h_d)|}\right),\ j = 1,2.$$

Now we restrict our attention to $\widehat{Q}_{n1}(\boldsymbol{\Delta})$, $\boldsymbol{\Delta} \neq \boldsymbol{\Delta}_0$. The proofs for other cases are similar. Note that

$$\begin{aligned} &P\left(\left|\sum_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z}_k)} \widetilde{\Psi}_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}_k)\right| > M a_n \frac{n^2 A_n^{-1} h_1 \dots h_d}{|f_2^{-1}(h_1,\dots,h_d)|}\right) \\ &\le 2P\left(\sum_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z}_k)} \widetilde{\Psi}_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}_k) > M a_n \frac{n^2 A_n^{-1} h_1 \dots h_d}{|f_2^{-1}(h_1,\dots,h_d)|}\right). \end{aligned}$$

Observe that $\widetilde{\Psi}_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}_k)$ are zero-mean independent random variables and

$$\left|\widetilde{\Psi}_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}_k)\right| \le C_{\widetilde{\Psi}_1}(\overline{A}_{n1})^{d-1}\overline{A}_{n2} n A_n^{-1}|f_2(h_1,\dots,h_d)|\tau_n,\ \textit{a.s.}\ \text{(from Lemma D.1)}$$

$$E\left[\left(\widetilde{\Psi}_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}_k)\right)^2\right] \le C_{\widetilde{\Psi}_1}(\overline{A}_{n1})^{d-1}\overline{A}_{n2} n^2 A_n^{-2} f_2^2(h_1,\dots,h_d), \tag{B.3}$$

for some $C_{\widetilde{\Psi}_1} > 0$, where (B.3) can be shown by applying the same argument in (Step 2-1) in the proof of Theorem 4.1. Then Lemma D.3 yields that

$$P\left(\sum_{\boldsymbol{\ell}\in L_{n1}(\boldsymbol{z}_k)} \widetilde{\Psi}_1^{(\boldsymbol{\ell};\boldsymbol{\Delta})}(\boldsymbol{z}_k) > M a_n \frac{n^2 A_n^{-1} h_1 \dots h_d}{|f_2^{-1}(h_1,\dots,h_d)|}\right) \le \exp\left(-\frac{\frac{M^2 n^2 A_n^{-1} h_1 \dots h_d \log n}{2|f_2^{-1}(h_1,\dots,h_d)|^2}}{E_{n1}+E_{n2}}\right),$$

where

$$E_{n1} = C_{\widetilde{\Psi}_1}\left(\frac{A_n h_1 \dots h_d}{A_n^{(1)}}\right)(\overline{A}_{n1})^{d-1}\overline{A}_{n2} n^2 A_n^{-2} f_2^2(h_1,\dots,h_d),$$

$$E_{n2} = \frac{M C_{\widetilde{\Psi}_1} n^2 A_n^{-3/2}(h_1 \dots h_d)^{1/2}(\log n)^{1/2}(\overline{A}_{n1})^{d-1}\overline{A}_{n2}\tau_n}{3|f_2^{-1}(h_1,\dots,h_d)|^2}.$$

Since

$$\frac{M^2 n^2 A_n^{-1} h_1 \dots h_d \log n}{2|f_2^{-1}(h_1,\dots,h_d)|^2 E_{n1}} = \frac{M^2}{2C_{\widetilde{\Psi}_1}}\left(\frac{A_n^{(1)}}{(\overline{A}_{n1})^{d-1}\overline{A}_{n2}}\right)\log n,$$

$$\frac{M^2 n^2 A_n^{-1} h_1 \dots h_d \log n}{2|f_2^{-1}(h_1,\dots,h_d)|^2 E_{n2}} = \frac{3M}{2C_{\widetilde{\Psi}_1}} \frac{A_n^{1/2}(h_1 \dots h_d)^{1/2}}{n^{1/q_2}(\overline{A}_{n1})^{d-1}\overline{A}_{n2}(\log n)^{-1/2+\iota}},$$

by taking $M > 0$ sufficiently large, we obtain the desired result. □

B.2. **Proof of Theorem 5.1.**

Proof. Define

$$S_n(\boldsymbol{z}) = \frac{1}{nh_1\dots h_d}\sum_{i=1}^{n} K_{Ah}\left(\boldsymbol{X}_i - A_n\boldsymbol{z}\right)H^{-1}\begin{pmatrix} 1 \\ (\boldsymbol{X}_i \check{-} A_n\boldsymbol{z}) \end{pmatrix}(1\ (\boldsymbol{X}_i \check{-} A_n\boldsymbol{z})')H^{-1},$$

$$V_n(\boldsymbol{z}) = \frac{1}{nh_1\dots h_d}\sum_{i=1}^{n} K_{Ah}\left(\boldsymbol{X}_i - A_n\boldsymbol{z}\right)H^{-1}\begin{pmatrix} 1 \\ (\boldsymbol{X}_i \check{-} A_n\boldsymbol{z}) \end{pmatrix}(e_{n,i}+\varepsilon_{n,i}),$$

$$B_n(\boldsymbol{z}) = \frac{1}{nh_1\dots h_d}\sum_{i=1}^{n} K_{Ah}\left(\boldsymbol{X}_i - A_n\boldsymbol{z}\right)H^{-1}\begin{pmatrix} 1 \\ (\boldsymbol{X}_i \check{-} A_n\boldsymbol{z}) \end{pmatrix}$$
$$\times \sum_{1\le j_1\le\cdots\le j_{p+1}\le d}\frac{1}{\boldsymbol{s}_{j_1\dots j_{p+1}}!}\partial_{j_1,\dots,j_{p+1}}m(\dot{\boldsymbol{X}}_i/A_n)\prod_{\ell=1}^{p+1}\left(\frac{X_{i,j_\ell}}{A_{n,j_\ell}} - z_{j_\ell}\right).$$

Note that

$$H(\widehat{\boldsymbol{\beta}}(\boldsymbol{z}) - \boldsymbol{M}(\boldsymbol{z})) = S_n^{-1}(\boldsymbol{z})(V_n(\boldsymbol{z}) + B_n(\boldsymbol{z})).$$

Applying Proposition 5.1 (5.7) to $e'_{j_{1,1}\dots j_{1,L_1}}S_n(\boldsymbol{z})e_{j_{2,1}\dots j_{2,L_2}}$ with

$$f_1(\boldsymbol{x}) = e'_{j_{1,1}\dots j_{1,L_1}}\begin{pmatrix} 1 \\ \check{\boldsymbol{x}} \end{pmatrix}(1\ \check{\boldsymbol{x}}')e_{j_{2,1}\dots j_{2,L_2}},\ f_2(\boldsymbol{x}) = 1,\ f_3(\boldsymbol{x}) = 1,$$

we have that

$$\begin{aligned}&\sup_{\boldsymbol{z}\in \mathrm{T}_n}|e'_{j_{1,1}\dots j_{1,L_1}}(S_n(\boldsymbol{z}) - g(\boldsymbol{z})S)e_{j_{2,1}\dots j_{2,L_2}}|\\ &\le \sup_{\boldsymbol{z}\in R_0}|e'_{j_{1,1}\dots j_{1,L_1}}(S_n(\boldsymbol{z}) - E[S_n(\boldsymbol{z})])e_{j_{2,1}\dots j_{2,L_2}}| + \sup_{\boldsymbol{z}\in \mathrm{T}_n}|e'_{j_{1,1}\dots j_{1,L_1}}(E[S_n(\boldsymbol{z})] - g(\boldsymbol{z})S)e_{j_{2,1}\dots j_{2,L_2}}|\\ &= O_p\left(\sqrt{\frac{\log n}{nh_1\dots h_d}}\right) + o(1) = o_p(1). \end{aligned}\tag{B.4}$$

Applying Proposition 5.1 (5.6) to $A_n n^{-1}e'_{j_1\dots j_L}V_n(\boldsymbol{z})$ with

$$f_1(\boldsymbol{x}) = e'_{j_1\dots j_L}\begin{pmatrix} 1 \\ \check{\boldsymbol{x}} \end{pmatrix},\ f_2(\boldsymbol{x}) = 1, (f_3(\boldsymbol{x}), Z_{\boldsymbol{X}_i}) \in \left\{(\eta(\boldsymbol{x}), e(\boldsymbol{X}_i)), (\sigma_\varepsilon(\boldsymbol{x}), \varepsilon_i)\right\},$$

we have that

$$\frac{n}{A_n}\sup_{\boldsymbol{z}\in \mathrm{T}_n}\left|\frac{A_n}{n}e'_{j_1\dots j_L}(V_n(\boldsymbol{z}) - E[V_n(\boldsymbol{z})])\right| \le \frac{n}{A_n}\sup_{\boldsymbol{z}\in R_0}\left|\frac{A_n}{n}e'_{j_1\dots j_L}V_n(\boldsymbol{z})\right| = O_p\left(\sqrt{\frac{\log n}{A_n h_1\dots h_d}}\right).\tag{B.5}$$

Applying Proposition 5.1 (5.7) to $e'_{j_1\dots j_L}B_n(\boldsymbol{z})$ with

$$f_1(\boldsymbol{x}) = e'_{j_1\dots j_L}\begin{pmatrix} 1 \\ \check{\boldsymbol{x}} \end{pmatrix},\ f_2(\boldsymbol{x}) = \prod_{\ell=1}^{L}x_{j_\ell},\ f_3(\boldsymbol{x}) = \sum_{1\le j_1\le\cdots\le j_{p+1}\le d}\frac{1}{\boldsymbol{s}_{j_1\dots j_{p+1}}!}\partial_{j_1,\dots,j_{p+1}}m(\boldsymbol{x}),$$

we have that

$$\sup_{\boldsymbol{z}\in \mathrm{T}_n}\left|e'_{j_1\dots j_L}B_n(\boldsymbol{z})\right| \le \sup_{\boldsymbol{z}\in R_0}\left|e'_{j_1\dots j_L}(B_n(\boldsymbol{z}) - E[B_n(\boldsymbol{z})])\right| + \sup_{\boldsymbol{z}\in \mathrm{T}_n}\left|e'_{j_1\dots j_L}E[B_n(\boldsymbol{z})]\right|$$

$$= O_p\left(\prod_{\ell=1}^{L} h_{j_\ell}\sqrt{\frac{\log n}{nh_1\dots h_d}}\right) + O\left(\sum_{1\le j_1\le\cdots\le j_{p+1}\le d}\prod_{\ell=1}^{p+1} h_{j_\ell}\right) \tag{B.6}$$

Combining (B.4)-(B.6), we have that

$$\begin{aligned}
&\sup_{\boldsymbol{z}\in \mathrm{T}_n} |\partial_{j_1\dots j_L}\widehat{m}(\boldsymbol{z}) - \partial_{j_1\dots j_L} m(\boldsymbol{z})| \\
&\le \left(\prod_{\ell=1}^{L} h_{j_\ell}\right)^{-1} \frac{1}{\inf_{\boldsymbol{z}\in R_0} g(\boldsymbol{z})}\sup_{\boldsymbol{z}\in \mathrm{T}_n}\left|e_{j_1\dots j_L}S^{-1}(V_n(\boldsymbol{z})+B_n(\boldsymbol{z}))\right| \\
&\quad + \left(\prod_{\ell=1}^{L} h_{j_\ell}\right)^{-1} \sup_{\boldsymbol{z}\in \mathrm{T}_n}\left|e_{j_1\dots j_L}(S_n^{-1}(\boldsymbol{z}) - g^{-1}(\boldsymbol{z})S^{-1})(V_n(\boldsymbol{z})+B_n(\boldsymbol{z}))\right| \\
&\lesssim \left(\prod_{\ell=1}^{L} h_{j_\ell}\right)^{-1}\left(\sup_{\boldsymbol{z}\in \mathrm{T}_n}|V_n(\boldsymbol{z})| + \sup_{\boldsymbol{z}\in \mathrm{T}_n}|B_n(\boldsymbol{z})|\right) \\
&= O_p\left(\frac{\sum_{1\le j_1\le\cdots\le j_{p+1}\le d}\prod_{\ell=1}^{p+1} h_{j_\ell}}{\prod_{\ell=1}^{L} h_{j_\ell}} + \sqrt{\frac{\log n}{A_n h_1\dots h_d\left(\prod_{\ell=1}^{L} h_{j_\ell}\right)^2}}\right).
\end{aligned}$$

□

Appendix C. Proofs for Section 6

C.1. Proof of Proposition 6.1.

Proof. Define $r_1 = \min_{1\le j,k\le 2} r_{1,jk}$. We first check the asymptotic negligibility of the random field $\boldsymbol{e}_{2,m_n}$, that is,

$$\max_{1\le i\le n} e_{2j,m_n}(\boldsymbol{X}_i) = O_p\left(\exp\left(-\frac{r_1 n^{\zeta_0\zeta_1\zeta_2}}{2}\right)\right),\ n\to\infty, \tag{C.1}$$

Note that under Condition (a), we have $E[|e_j(\mathbf{0})|^6]<\infty$ since $\boldsymbol{e}$ is Gaussian. Under Condition (b), we also have $E[|L_j([0,1]^d)|^6]<\infty$ since $\int_{|x|>1}|x|^6\nu_{0,j}(x)dx<\infty$ (cf. Theorem 25.3 in Sato (1999)). Define $\sigma_{\boldsymbol{e}_{1,m_n}}^{(j,k)}(\boldsymbol{x}) = E[e_{1j,m_n}(\mathbf{0})e_{1k,m_n}(\boldsymbol{x})]$, $j,k=1,2$. Then we have that

$$\begin{aligned}
E[|e_{1j,m_n}(\mathbf{0})|^6] &\le E[|e_j(\mathbf{0})|^6] \lesssim \int e^{-6r_1\|\boldsymbol{u}\|}d\boldsymbol{u} < \infty, \\
|\sigma_{\boldsymbol{e}_{1,m_n}}^{(j,k)}(\boldsymbol{x})| &\lesssim |E[e_j(\mathbf{0})e_k(\boldsymbol{x})]| \lesssim \int e^{-r_1\|\boldsymbol{u}\|}e^{-r_1\|\boldsymbol{x}-\boldsymbol{u}\|}d\boldsymbol{u} \\
&\le \int e^{-r_1\|\boldsymbol{u}\|}e^{-\frac{r_1}{2}(\|\boldsymbol{x}\|-\|\boldsymbol{u}\|)}d\boldsymbol{u} \lesssim e^{-\frac{r_1}{2}\|\boldsymbol{x}\|}.
\end{aligned}$$

The latter implies that $\int|\sigma_{\boldsymbol{e}_{1,m_n}}^{(j,k)}(\boldsymbol{v})|d\boldsymbol{v}<\infty$, $j,k=1,2$. Likewise,

$$\begin{aligned}
E[(e_{2j,m_n}(\mathbf{0}))^4] &\lesssim \int_{\mathbb{R}^d} e^{-4r_1\|\boldsymbol{u}\|}\left(1-\psi_0\left(\|\boldsymbol{u}\| : m_n\right)\right)^4 d\boldsymbol{u} \\
&\lesssim \int_{\|\boldsymbol{u}\|\ge m_n/4} e^{-4r_1\|\boldsymbol{u}\|}\left|1+\frac{4}{m_n}\left(\|\boldsymbol{u}\|-\frac{m_n}{2}\right)\right|^4 d\boldsymbol{u}
\end{aligned}$$

$$
\begin{aligned}
&\lesssim \int_{\|\boldsymbol{u}\| \geq m_n/4} e^{-4r_1\|\boldsymbol{u}\|} \left|1 + \frac{4\|\boldsymbol{u}\|}{m_n}\right|^4 d\boldsymbol{u} \\
&\leq 2^{q-1} \int_{\|\boldsymbol{u}\| \geq m_n/4} e^{-4r_1\|\boldsymbol{u}\|} \left(1 + \frac{4^4\|\boldsymbol{u}\|^4}{m_n^4}\right) d\boldsymbol{u} \\
&\lesssim \int_{m_n/4}^{\infty} e^{-4r_1 t} \left(1 + \frac{4^4 t^4}{m_n^4}\right) t^{d-1} dt \\
&\lesssim m_n^{d-1} e^{-r_1 m_n}.
\end{aligned}
$$

By Markov's inequality and Lemma 2.2.2 in van der Vaart and Wellner (1996), we have

$$
\begin{aligned}
P_{\cdot|\boldsymbol{X}}\left(\left|\max_{1\leq i\leq n} e_{2j,m_n}(\boldsymbol{X}_i)\right| > \varrho\right) &\leq \varrho^{-1} E_{\cdot|\boldsymbol{X}}\left[\max_{1\leq i\leq n} |e_{2j,m_n}(\boldsymbol{X}_i)|\right] \\
&\leq \varrho^{-1} n^{1/4} \max_{1\leq i\leq n} \left(E_{\cdot|\boldsymbol{X}}\left[|e_{2j,m_n}(\mathbf{0})|^4\right]\right)^{1/4} \\
&\lesssim \varrho^{-1} n^{1/4} m_n^{(d-1)/4} e^{-r_1 m_n/4}.
\end{aligned}
$$

Therefore, under the assumptions of Proposition 6.1, we have (C.1), which implies that $\boldsymbol{e}_{2,m_n}$ is asymptotically negligible. Hence we can replace $\boldsymbol{e}$ with $\boldsymbol{e}_{1,m_n}$ in the results in Section 4.

Next we check the mixing conditions on $\boldsymbol{e}_{1,m_n}$. Let $\alpha_{\boldsymbol{e}_1}(a;b)$ be the α-mixing coefficients of $\boldsymbol{e}_{1,m_n}$. Note that $\alpha_{\boldsymbol{e}_1}(a;b) \leq \alpha(a;b)$. Since $\boldsymbol{e}_{1,m_n}$ is m_n-dependent, under the assumptions of Proposition 6.1, we have $\alpha_1(\underline{A}_{n2}) = 0$, which yields

$$
\begin{aligned}
&\left(\frac{A_n h_1 \dots h_d}{A_n^{(1)}}\right) \alpha_1(\underline{A}_{n2}) \varpi_1(A_n h_1 \dots h_d) = 0, \\
&A_n^{(1)} \left(\alpha_1^{1-2/q}(\underline{A}_{n2}) + \sum_{k=\underline{A}_{n1}}^{\infty} k^{d-1} \alpha_1^{1-2/q}(k)\right) \varpi_1^{1-2/q}(A_n^{(1)}) = 0.
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
\left(\frac{A_n^{(1)}}{A_n h_1 \dots h_d}\right) \sum_{k=1}^{\overline{A}_{n1}} k^{2d-1} \alpha_1^{1-4/q}(k) &\lesssim \left(\frac{A_n^{(1)}}{A_n h_1 \dots h_d}\right) \sum_{k=1}^{m_n} k^{2d-1} \\
&\leq \left(\frac{A_n^{(1)}}{A_n h_1 \dots h_d}\right) m_n^{2d} \\
&\lesssim n^{-\zeta_0\{1-\zeta_1(1+\zeta_2)\}+\zeta_3} = o(1).
\end{aligned}
$$

$$
\begin{aligned}
&\left\{\left(\frac{\overline{A}_{n1}}{\underline{A}_{n1}}\right)^d \left(\frac{\overline{A}_{n2}}{\overline{A}_{n1}}\right) + \left(\frac{A_n^{(1)}}{\underline{A}_{n1}^d}\right) \left(\frac{(\overline{A_n h})^d}{A_n h_1 \dots h_d}\right) \left(\frac{\overline{A}_{n1}}{\overline{\overline{A_n h}}}\right)\right\} \sum_{k=1}^{\overline{A}_{n1}} k^{d-1} \alpha_1^{1-2/q}(k) \\
&\lesssim \left\{\left(\frac{\overline{A}_{n2}}{\overline{A}_{n1}}\right) + \left(\frac{\overline{A}_{n1}}{\overline{\overline{A_n h}}}\right)\right\} m_n^d \lesssim \left(n^{\frac{\zeta_0\zeta_1\zeta_2}{d} - \frac{\zeta_0\zeta_1}{d}} + n^{-\frac{\zeta_0\zeta_1}{d} - \frac{\zeta_0}{d} + \frac{\zeta_3}{d}}\right) n^{\frac{\zeta_0\zeta_1\zeta_2}{2}} \\
&= n^{\zeta_0\zeta_1\left\{\left(\frac{d+2}{2d}\right)\zeta_2 - \frac{1}{d}\right\}} + n^{-\zeta_1\left\{1-\left(1+\frac{d}{2}\zeta_0\right)\zeta_2\right\}+\zeta_3} = o(1).
\end{aligned}
$$

We can also check that $A_{n,j}h_j/A_{n1,j} \to \infty$ as $n \to \infty$ and that Assumptions 4.1 (ii), (iii), and (iv) are satisfied. Therefore, we obtain the desired result. □

C.2. **Proof of Proposition 6.2.**

Proof. Define

$$\Psi_{1,\boldsymbol{e}_2}(\boldsymbol{z}) = \frac{1}{n^2 A_n^{-1} h_1 \dots h_d} \sum_{i=1}^{n} K_{Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}) H^{-1} \begin{pmatrix} 1 \\ (\boldsymbol{X}_i \check{-} A_n\boldsymbol{z}) \end{pmatrix} \eta\left(\frac{\boldsymbol{X}_i}{A_n}\right) e_{2,m_n}(\boldsymbol{X}_i).$$

By the same argument in the proof of Proposition 6.1, we can show that

$$\max_{1 \le i \le n} |e_{2,m_n}(\boldsymbol{X}_i)| = O_p\left(\exp\left(-\frac{r_1 n^{\zeta_0\zeta_1\zeta_2}}{2}\right)\right),\ n \to \infty. \tag{C.2}$$

Then we have

$$|\Psi_{1,\boldsymbol{e}_2}(\boldsymbol{z})| = O_p\left(\frac{\exp\left(-\frac{r_1 n^{\zeta_0\zeta_1\zeta_2}}{2}\right)}{n^2 A_n^{-1} h_1 \dots h_d}\right) \left|\sum_{i=1}^{n} K_{Ah}(\boldsymbol{X}_i - A_n\boldsymbol{z}) H^{-1} \begin{pmatrix} 1 \\ (\boldsymbol{X}_i \check{-} A_n\boldsymbol{z}) \end{pmatrix} \eta\left(\frac{\boldsymbol{X}_i}{A_n}\right)\right|.$$

Applying Proposition 5.1 (5.7) with

$$f_1(\boldsymbol{x}) = e'_{j_1 \dots j_L} \begin{pmatrix} 1 \\ \check{\boldsymbol{x}} \end{pmatrix},\ f_2(\boldsymbol{x}) = 1,\ f_3(\boldsymbol{x}) = \eta(\boldsymbol{z}),$$

we have that

$$\begin{aligned}
\sup_{\boldsymbol{z} \in \mathrm{T}_n} |\Psi_{1,\boldsymbol{e}_2}(\boldsymbol{z})| &\le O_p\left(\frac{A_n}{n}\exp\left(-\frac{r_1 n^{\zeta_0\zeta_1\zeta_2}}{2}\right)\right)\left(\sup_{\boldsymbol{z} \in \mathrm{T}_n}\left|\widehat{\Psi}_{\mathrm{II}}(\boldsymbol{z}) - E[\widehat{\Psi}_{\mathrm{II}}(\boldsymbol{z})]\right| + \sup_{\boldsymbol{z} \in \mathrm{T}_n}\left|E[\widehat{\Psi}_{\mathrm{II}}(\boldsymbol{z})]\right|\right) \\
&= O_p\left(\frac{A_n}{n}\exp\left(-\frac{r_1 n^{\zeta_0\zeta_1\zeta_2}}{2}\right)\right)\left(O_p\left(\sqrt{\frac{\log n}{n h_1 \dots h_d}}\right) + O(1)\right) \\
&= O_p\left(\exp\left(-\frac{r_1 n^{\zeta_0\zeta_1\zeta_2}}{2}\right)\right)
\end{aligned}$$

and this implies that $\boldsymbol{e}_{2,m_n}$ is asymptotically negligible. Further, under the assumptions in Proposition 6.2 we have that $\beta_1(\underline{A}_{n2}) = 0$,

$$\frac{A_{n,j} h_j}{A_{n1,j}} \sim n^{\frac{\zeta_0(1-\zeta_1)-\zeta_3}{d}} \gg 1,\ \left(\frac{A_n^{(1)}}{(\overline{A}_{n1})^d}\right) \sim 1,\ \frac{A_n^{\frac{1}{2}}(h_1 \dots h_d)^{\frac{1}{2}}}{n^{\frac{1}{q_2}}(\overline{A}_{n1})^d} \sim n^{\frac{\zeta_0(1-2\zeta_1)-\zeta_3}{2} - \frac{1}{q_2}} \gg (\log n)^{\frac{1}{2}+\iota}.$$

Therefore, we can replace $\boldsymbol{e}$ with $\boldsymbol{e}_{1,m_n}$ in Theorem 5.1. □

Appendix D. Technical tools

We refer to the following lemmas without those proofs.

Lemma D.1 ((5.19) in Lahiri (2003b))**.** *Under Assumption 2.2, we have*

$$P\left(\sum_{i=1}^{n} 1\{\boldsymbol{X}_i \in \Gamma_{n,\boldsymbol{z}}(\boldsymbol{\ell};\boldsymbol{\Delta})\} > C|\Gamma_{n,\boldsymbol{z}}(\boldsymbol{\ell};\boldsymbol{\Delta})| n A_n^{-1} \text{ for some } \boldsymbol{\ell} \in L_{n1}(\boldsymbol{z}),\ i.o.\right) = 0$$

for any $\boldsymbol{\Delta} \in \{1,2\}^d$*, where* $C > 0$ *is a sufficiently large constant.*

Remark D.1. Lemma D.1 implies that each $\Gamma_{n,\boldsymbol{z}}(\boldsymbol{\ell};\boldsymbol{\Delta})$ contains at most $C|\Gamma_{n,\boldsymbol{z}}(\boldsymbol{\ell};\boldsymbol{\Delta})| n A_n^{-1}$ samples almost surely.

Lemma D.2 (Corollary 2.7 in Yu (1994)). *Let $m \in \mathbb{N}$ and let Q be a probability measure on a product space $(\prod_{i=1}^m \Omega_i, \prod_{i=1}^m \Sigma_i)$ with marginal measures Q_i on (Ω_i, Σ_i). Suppose that h is a bounded measurable function on the product probability space such that $|h| \leq M_h < \infty$. For $1 \leq a \leq b \leq m$, let Q_a^b be the marginal measure on $(\prod_{i=a}^b \Omega_i, \prod_{i=a}^b \Sigma_i)$. For a given $\tau > 0$, suppose that, for all $1 \leq k \leq m-1$,*

$$\|Q - Q_1^k \times Q_{k+1}^m\|_{TV} \leq 2\tau, \tag{D.1}$$

where $Q_1^k \times Q_{k+1}^m$ is a product measure and $\|\cdot\|_{TV}$ is the total variation. Then

$$|Qh - Ph| \leq 2M_h(m-1)\tau.$$

where $P = \prod_{i=1}^m Q_i$, $Qh = \int hdQ$, and $Ph = \int hdP$.

Lemma D.3 (Bernstein's inequality). *Let $X_1, \ldots, X_n$ be independent zero-mean random variables. Suppose that $\max_{1 \leq i \leq n} |X_i| \leq M < \infty$ a.s. Then, for all $t > 0$,*

$$P\left(\sum_{i=1}^n X_i \geq t\right) \leq \exp\left(-\frac{\frac{t^2}{2}}{\sum_{i=1}^n E[X_i^2] + \frac{Mt}{3}}\right).$$

References

Bandyopadhyay, S., Lahiri, S. N., and Nordman, D. J. (2015). A frequency domain empirical likelihood method for irregularly spaced spatial data. *Annals of Statistics*, 43:519–545.

Bertoin, J. (1996). *Lévy Processes.* Cambridge University Press.

Bradley, R. C. (1989). A caution on mixing conditions for random fields. *Statistics & Probability Letters*, 8:489–491.

Bradley, R. C. (1993). Some examples of mixing random fields. *Rocky Mountain Journal of Mathematics*, 23:495–519.

Brockwell, P. J. and Matsuda, Y. (2017). Continuous auto-regressive moving average random fields on $\mathbb{R}^n$. *Journal of the Royal Statistical Society Series B*, 79:833–857.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Cressie, N. (1993). *Statistics for Spatial Data.* Wiley, New York, NY.

Doukhan, P. (1994). *Mixing: Properties and Examples.* Springer.

Ehrlich, M. v. and Seidel, T. (2018). The persistent effects of place-based policy: Evidence from the west-german zonenrandgebiet. *American Economic Journal: Economic Policy*, 10(4):344–74.

El Machkouri, M., Es-Sebaiy, K., and Ouassou, I. (2017). On local linear regression for strongly mixing random fields. *Journal of Multivariate Analysis*, 156:103–115.

El Machkouri, M. and Stoica, R. (2010). Asymptotic normality of kernel estimates in a regression model for random fields. *Journal of Nonparametric Statistics*, 22(8):955–971.

El Machkouri, M., Volný, D., and Wu, W. B. (2013). A central limit theorem for stationary random fields. *Stochastic Processes and their Applications*, 123:1–14.

Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods.* Springer.

Guttorp, P. and Gneiting, T. (2006). Studies in the history of probability and statistics xlix on the matérn correlation family. *Biometrika*, 93:989–995.

Hahn, J., Todd, P., and Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209.

Hallin, M., Lu, Z., and Tran, L. T. (2004). Local linear spatial regression. *Annals of Statistics*, 32:2469–2500.

Hansen, B. E. (2008). Uniform convergence rates for kernel estimation with dependent data. *Econometric Theory*, 24(3):726–748.

Jenish, N. (2012). Nonparametric spatial regression under near-epoch dependence. *Journal of Econometrics*, 167:224–239.

Kato, K. and Kurisu, D. (2020). Bootstrap confidence bands for spectral estimation of lévy densities under high-frequency observations. *Stochastic Processes and their Applications*, 130:1159–1205.

Keele, L. J. and Titiunik, R. (2015). Geographic boundaries as regression discontinuities. *Political Analysis*, 23(1):127–155.

Kristensen, D. (2009). Uniform convergence rates of kernel estimators with heterogeneous dependent data. *Econometric Theory*, 25(5):1433–1445.

Kurisu, D. (2019). On nonparametric inference for spatial regression models under domain expanding and infill asymptotics. *Statistics & Probability Letters*, 154:108543.

Kurisu, D. (2022). Nonparametric regression for locally stationary random fields under stochastic sampling design. *Bernoulli*, 28:1250–1275.

Kurisu, D., Kato, K., and Shao, X. (2021). Gaussian approximation and spatially dependent wild bootstrap for high-dimensional spatial data. *arXiv:2103.10720*.

Lahiri, S. N. (1996). On inconsistency of estimators under infill asymptotics for spatial data. *Sankhya Ser. A*, 58:403–417.

Lahiri, S. N. (1999). Asymptotic distribution of the empirical spatial cumulative distribution function predictor and prediction bands based on a subsampling method. *Probability Theory and Related Fields*, 114(1):55–84.

Lahiri, S. N. (2003a). Central limit theorems for weighted sum of a spatial process under a class of stochastic and fixed design. *Sankhya*, 65:356–388.

Lahiri, S. N. (2003b). *Resampling Methods for Dependent Data.* Springer.

Lahiri, S. N. and Zhu, J. (2006). Resampling methods for spatial regression models under a class of stochastic designs. *Annals of Statistics*, 34:1774–1813.

Masry, E. (1996a). Multivariate local polynomial regression for time series: uniform strong consistency and rates. *Journal of Time Series Analysis*, 17(6):571–599.

Masry, E. (1996b). Multivariate regression estimation local polynomial fitting for time series. *Stochastic Processes and their Applications*, 65(1):81–101.

Masry, E. and Fan, J. (1997). Local polynomial estimation of regression functions for mixing processes. *Scandinavian Journal of Statistics*, 24(2):165–179.

Matérn, B. (1986). *Spatial Variation (2nd ed.)*. Springer.

Matsuda, Y. and Yajima, Y. (2009). Fourier analysis of irregularly spaced data on $\mathbb{R}^d$. *Journal of the Royal Statistical Society Series B*, 71:191–217.

Matsuda, Y. and Yajima, Y. (2018). Locally stationary spatio-temporal processes. *Japan Journal of Statistics and Data Science*, 1:41–57.

Müller, U. K. and Watson, M. W. (2021). Spatial correlation robust inference. *forthcoming in Econometrica*.

Rajput, B. S. and Rosinski, J. (1989). Spectral representations of infinitely divisible processes. *Probability Theory and Related Fields*, 82:451–487.

Robinson, P. M. (2011). Asymptotic theory for nonparametric regression with spatial data. *Journal of Econometrics*, 165:5–19.

Sato, K. (1999). *Lévy processes and infinitely divisible distributions.* Cambridge University Press.

Stein, M. L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging*. Springer.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes with Applications to Statistics.* Springer.

Vogt, M. (2012). Nonparametric regression for locally stationary time series. *The Annals of Statistics*, 40(5):2601–2633.

Yu, B. (1994). Rates of convergence for empirical processes of stationary mixing sequences. *Annals of Probability*, 22:94–116.

Zhang, T. and Wu, W. B. (2015). Time-varying nonlinear regression models: nonparametric estimation and model selection. *The Annals of Statistics*, 43(2):741–768.

Zhao, Z. and Wu, W. B. (2008). Confidence bands in nonparametric time series regression. *The Annals of Statistics*, 36(4):1854–1878.

Zhou, Z. and Wu, W. B. (2009). Local linear quantile estimation for nonstationary time series. *The Annals of Statistics*, 37(5B):2696–2729.

(D. Kurisu) Graduate School of International Social Sciences, Yokohama National University, 79-4, Tokiwadai, Hodogaya-ku, Yokohama 240-8501, Japan.

Email address: kurisu-daisuke-jr@ynu.ac.jp

(Y. Matsuda) Graduate School of Economics and Management, Tohoku University, Sendai 980-8576, Japan.

Email address: matsuda@econ.tohoku.ac.jp