TESTING FUNCTIONAL INEQUALITIES

SOKBAE LEE, KYUNGCHUL SONG, AND YOON-JAE WHANG

ABSTRACT.
This paper develops tests for inequality constraints of nonparametric regression functions. The test statistics involve a one-sided version of L_p-type functionals of kernel estimators. Drawing on the approach of Poissonization, this paper establishes that the tests are asymptotically distribution free, admitting asymptotic normal approximation. Furthermore, the tests have nontrivial local power against a certain class of local alternatives converging to the null at the rate of $n^{-1/2}$. Some results from Monte Carlo simulations are presented.

KEY WORDS. Kernel estimation, one-sided test, local power, L_p norm, Poissonization.

JEL SUBJECT CLASSIFICATION. C12, C14.

AMS SUBJECT CLASSIFICATION. 62G10, 62G08, 62G20.

1. INTRODUCTION

Suppose that we observe $\{(Y_i', X_i')'\}_{i=1}^n$ that are i.i.d. copies from a random vector, $(Y', X')' \in \mathbf{R}^J \times \mathbf{R}^d$. Write $Y_i = (Y_{1i}, \cdot\cdot\cdot, Y_{Ji})' \in \mathbf{R}^J$ and define $m_j(x) \equiv \mathbf{E}[Y_{ji}|X_i = x]$, $j = 1, 2, \cdot\cdot\cdot, J$. The notation $\equiv$ indicates definition.

This paper focuses on the problem of testing functional inequalities:

$$\begin{aligned} &H_0 : m_j(x) \leq 0 \text{ for all } (x, j) \in \mathcal{X} \times \mathcal{J}, \text{ vs.} \\ &H_1 : m_j(x) > 0 \text{ for some } (x, j) \in \mathcal{X} \times \mathcal{J}, \end{aligned} \tag{1.1}$$

where $\mathcal{X} \subset \mathbf{R}^d$ is the domain of interest and $\mathcal{J} \equiv \{1, \ldots, J\}$. The testing problem is relevant in various applied settings. For example, a researcher may be interested in testing whether a new medical treatment brings about a better outcome than an existing treatment uniformly across different treatment environments. Or it may be of interest to see if a particular job training program leads to a better earnings prospect uniformly across all the regions in which the program has been implemented. Details of the examples and references are provided in the next section.

Date: 7 February 2011.
We would like to thank David Mason and Oliver Linton for their helpful comments on earlier versions of this paper. Lee thanks the Economic and Social Research Council for the ESRC Centre for Microdata Methods and Practice (RES-589-28-0001) and the European Research Council for the research grant (ERC-2009-StG-240910-ROMETA).

This paper proposes a one-sided L_p approach in testing nonparametric functional inequalities. While measuring the quality of an estimated nonparametric function by its L_p-distance from the true function has long received attention in the literature (see Devroye and Györfi (1985), for an elegant treatment of the L_1 norm of nonparametric density estimation), the advance of this approach for general nonparametric testing seems to have been rather slow relative to other approaches, perhaps due to its technical complexity. See Section 1.1 for details on the related literature.

Csörgő and Horváth (1988) first established a central limit theorem for the L_p-distance of a kernel density estimator from its population counterpart, and Horváth (1991) introduced a Poissonization technique into the analysis of the L_p-distance. Beirlant and Mason (1995) developed a different Poissonization technique and established a central limit theorem for the L_p-distance of kernel density estimators and regressograms from their expected values without assuming smoothness conditions for the nonparametric functions. Giné, Mason and Zaitsev (2003: GMZ, hereafter) employed this technique to prove the weak convergence of an L_1-distance process indexed by kernel functions in kernel density estimators.

There are other applications of the Poissonization method. For example, Anderson, Linton, and Whang (2009) developed methodology for kernel estimation of a polarization measure; Lee and Whang (2009) established asymptotic null distributions for the L_1-type test statistics for conditional treatment effects; and Mason (2009) established both finite sample and asymptotic moment bounds for the L_p risk for kernel density estimators. See also Mason and Polonik (2009) and Biau, Cadre, Mason, and Pelletier (2009) for asymptotic distribution theory in support estimation.

This paper builds on the contributions of Beirlant and Mason (1995) and GMZ, and develops testing procedures that have the following desirable properties:

(i) The nonparametric tests do not require usual smoothness conditions for nonparametric functions for their asymptotic validity and consistency. While this is partly expected from the results of Beirlant and Mason (1995) and GMZ, this paper demonstrates that we can construct a kernel-based nonparametric test enjoying this flexibility in a much broader context.

(ii) The nonparametric tests of this paper are distribution free on the "boundary" of the null hypothesis where $m_j(x) = 0$, for all $x \in \mathcal{X}$ and for all $j \in \mathcal{J}$ and at the same time have nontrivial power against a wide class of $n^{-1/2}$-converging Pitman sequences. This is somewhat unexpected, given that nonparametric goodness-of-fit tests that involve random vectors of a multi-dimension and have nontrivial power against $n^{-1/2}$-converging Pitman sequences are not often distribution free. Exceptions are tests that use an innovation martingale approach (see, e.g., Khmaladze

(1993), Stute, Thies and Zhu (1998), Bai (2003), and Khmaladze and Koul (2004)) or tests for a null hypothesis that has a specific functional form (see, e.g., Blum, Kiefer, and Rosenblatt (1961), Delgado and Mora (2000) and Song (2009)).

The tests that we propose are based on one-sided L_p-type functionals. For $1 \leq p < \infty$, let $\Lambda_p : \mathbf{R} \mapsto \mathbf{R}$ be such that $\Lambda_p(v) \equiv \max\{v, 0\}^p,\ v \in \mathbf{R}$. Consider the following one-sided L_p-type functionals:

$$\varphi \mapsto \Gamma_j(\varphi) \equiv \int_{\mathcal{X}} \Lambda_p(\varphi(x)) w_j(x) dx, \text{ for } j \in \mathcal{J},$$

where $w_j : \mathbf{R}^d \to [0, \infty)$ is a nonnegative weight function. Let f denote the density function of X and define $g_j(x) \equiv m_j(x) f(x)$. Using Γ_j, we reduce the testing problem to that of testing whether $\Gamma_j(g) = 0$ for all $j \in \mathcal{J}$, or $\Gamma_j(g) > 0$ for some $j \in \mathcal{J}$. To construct a test statistic, define

$$\hat{g}_{jn}(x) \equiv \frac{1}{nh^d} \sum_{i=1}^{n} Y_{ji} K\left(\frac{x - X_i}{h}\right),$$

where $K : \mathbf{R}^d \mapsto \mathbf{R}$ is a kernel function and h a bandwidth parameter satisfying $h \to 0$ as $n \to \infty$.

This paper shows that under weak assumptions, for any $t = (t_1, \cdots, t_J)' \in (0, \infty)^J$, there exist nonstochastic sequences $a_{jn} \in \mathbf{R}$, $j \in \mathcal{J}$, and $\sigma_{t,n} \in (0, \infty)$ such that as $n \to \infty$,

$$T_n \equiv \frac{1}{\sigma_{t,n}} \sum_{j=1}^{J} t_j \left\{ n^{p/2} h^{(p-1)d/2} \Gamma_j(\hat{g}_{jn}) - a_{jn} \right\} \xrightarrow{d} N(0, 1),$$

on the boundary of the null hypothesis. This is done first by deriving asymptotic results for the Poissonized version of the processes, $\{\hat{g}_{jn}(x) : x \in \mathcal{X}\}$, $j \in \mathcal{J}$, and then by translating them back into those for the original processes through the de-Poissonization lemma of Beirlant and Mason (1995). To construct a test statistic, we replace a_{jn} and $\sigma_{t,n}$ by appropriate estimators to obtain a feasible version of T_n, say, $\hat{T}_n$, and show that the limiting distribution remains the same under a stronger bandwidth condition. Hence, for each $t \in (0, \infty)^J$, we obtain a distribution free and consistent test for the nonparametric functional inequality constraints. We also discuss the choice of t in Section 3.

1.1. **Related Literature.** The literature on hypothesis testing involving nonparametric functions has a long history. Many studies have focused on testing parametric or semiparametric specifications of regression functions against nonparametric alternatives. See, e.g., Bickel and Rosenblatt (1973), Härdle and Mammen (1993), Stute (1997), Delgado and González Manteiga (2000) and Khmaladze and Koul (2004) among many others. The testing problem in this paper is different from the aforementioned papers, as the focus is on whether certain inequality (or equality) restrictions hold, rather than on whether certain parametric specifications are plausible.

When $J = 1$, our testing problem is also different from testing

$$\begin{aligned} H_0 &: m(x) = 0 \text{ for all } x \in \mathcal{X}, \text{ against} \\ H_1 &: m(x) \geq 0 \text{ for all } x \in \mathcal{X} \text{ with strict inequality for some } x \in \mathcal{X}. \end{aligned}$$

Related to this type of testing problems, see Hall, Huber, and Speckman (1997) and Koul and Schick (1997, 2003) among others. In their setup, the possibility that $m(x) < 0$ for some x is excluded, so that a consistent test can be constructed using a linear functional of $m(x)$. On the other hand, in our setup, negative values of $m(x)$ for some x are allowed under both H_0 and H_1. As a result, a linear functional of $m(x)$ would not be suitable for our purpose.

There also exist some papers that consider the testing problem in (1.1). For example, Hall and Yatchew (2005) and Andrews and Shi (2010) considered functions of the form $u \mapsto \max\{u, 0\}^p$ to develop tests for (1.1). However, their tests are not distribution free, although they achieve some local power against $n^{-1/2}$-converging sequences. See also Hall and van Keilegom (2005) for the use of the one-sided L_p-type functionals for testing for monotone increasing hazard rate. None of the aforementioned papers developed test statistics of one-sided L_p-type functionals with kernel estimators like ours. See some remarks of Ghosal, Sen, and van der Vaart (2000, p.1070) on difficulty in dealing with one-sided L_p-type functionals with kernel estimators.

In view of Bickel and Rosenblatt (1973) who considered both L_2 and sup tests, a one-sided sup test appears to be a natural alternative to the L_p-type tests studied in this paper. For example, Chernozhukov, Lee, and Rosen (2009) considered a sup norm approach in testing inequality constraints of nonparametric functions. Also, it may be of interest to develop sup tests based on a one-sided version of a bootstrap uniform confidence interval of $\hat{g}_n$, similar to Claeskens and van Keilegom (2003). The sup tests typically do not have nontrivial power against any $n^{-1/2}$-converging alternatives, but they may have better power against some "sharp peak" type alternatives (Liero, Läuter and Konakov, 1998).

1.2. **Structure of the Paper.** The remainder of the paper is as follows. In Section 2, we illustrate the usefulness of our testing framework by discussing various potential applications. Section 3 establishes conditions under which our tests have asymptotically valid size when the null hypothesis is true and also are consistent against fixed alternatives. We also obtain local power results for the leading cases when $p = 1$ and $p = 2$ and make comparison with functional equality tests. In Section 4, we report results of some Monte

Carlo simulations that show that our tests perform well in finite samples. All the proofs are contained in Section 5.

2. Examples

In this section we briefly describe several potential applications of our testing framework.

2.1. **Testing for average treatment effects conditional on covariates.** In a randomized controlled trial, a researcher observes either an outcome with treatment (W_1) or an outcome without treatment (W_0) along with observable pre-determined characteristics of the subjects (X). Let $D = 1$ if the subject belongs to the treatment group and 0 otherwise. We assume that assignment to treatment is random and independent of X and that the assignment probability $p \equiv P\{D = 1\}$, $0 < p < 1$, is fixed by the experiment design. Then the average treatment effect $\mathbf{E}(W_1 - W_0|X = x)$, conditional on $X = x$, can be written as

$$\mathbf{E}(W_1 - W_0|X = x) = \mathbf{E}\left[\frac{DW}{p} - \frac{(1-D)W}{1-p}\middle| X = x\right],$$

where $W \equiv DW_1 + (1 - D)W_0$. In this setup, it may be of interest to test whether or not $m(x) \equiv \mathbf{E}(W_1 - W_0|X = x) \leq 0$ for all x. For example, suppose that a new treatment is introduced to cancer patients and the outcome variable is cancer recurrence in a fixed time horizon. Then one interesting null hypothesis H_0 is that average cancer recurrence has decreased for all demographic groups (X). Rejecting H_0 in this setting implies that there may exist a certain demographic group for which the new treatment causes some adverse effects. Previously, Lee and Whang (2009) developed nonparametric tests for conditional treatment effects using only the L_1-type functionals when the assignment probability is unknown and may depend on x.

2.2. **Testing for a monotonic dose-response relationship.** In medical or toxicological studies, researchers are interested in confirming a monotonic dose-response relationship between the severity of disease and the dosage of medicine. Suppose that increasing doses $\{d_j, j = 0, \cdots, J\}$ are given to $J+1$ groups of study subjects and responses $\{W_j, j = 0, \ldots, J\}$ are recorded. Let $D_j = 1$ if the study subject belongs to the j-th dose level group and 0 otherwise. As in the previous example, each subject is randomly assigned to a dose group in a manner that is independent of the subject's observable predetermined state (X), and $\{p_j \equiv P\{D_j = 1\} : j = 0, \cdots, J,\ 0 < p_j < 1\}$ are fixed by the experiment design. Then the increment $W_j - W_{j-1}$ in response to the dose increase from the $j-1$-th level to the j th conditional on $X = x$ is written as

$$m_j(x) \equiv \mathbf{E}(W_j - W_{j-1}|X = x) = \mathbf{E}\left[\frac{D_jW}{p_j} - \frac{D_{j-1}W}{p_{j-1}}\middle| X = x\right],$$

where $W \equiv \sum_{j=0}^{J} D_j W_j$. Testing for a decreasing dose-response relationship for all levels of dosage and for all values of X falls within the framework of this paper.

2.3. **Testing the "realistic expectations" assumption in insurance markets.** In economic theory, primitive assumptions of economic models generate certain testable implications in the form of functional inequalities. For example, Chiappori, Jullien, Salanié, and Salanié (2006) formulated some testable restrictions in the study of insurance markets. In their set-up, economic agents face the risk of a monetary loss $L \geq 0$, and are allowed to purchase an insurance contract either with a high deductible d_1 or with a low deductible d_2 ($d_2 < d_1$). As in many actual insurance markets, insurers can set the levels of insurance premia based on observable characteristics (X) of the insured, while fixing the deductibles (d_1 and d_2) the same across the contracts. As a consequence, the premia offered to an agent with $X = x$ are deterministic functions $P_1(x)$ and $P_2(x)$ of x. If each agent can correctly assess their accident probability and loss distribution, it follows that (see equation (4) of Chiappori, Jullien, Salanié, and Salanié (2006))

$$m(x) \equiv P\{L > d_1 | X = x\} - \frac{P_2(x) - P_1(x)}{d_1 - d_2} \leq 0 \text{ for all } x \tag{2.1}$$

for all the insured who have bought the contract with d_1. Hence, testing (2.1) is viewed as testing the "realistic expectations" assumption. See a recent review by Einav, Finkelstein and Levin (2010) for related testing problems in insurance markets.

2.4. **Inference on intersection bounds.** One-sided L_p-type tests developed in this paper can be used to construct confidence regions for a parameter that is partially identified. Various partially identified models have received increasing attention in the recent literature of econometrics. See, among many others, Manski (2003), Chernozhukov, Lee, and Rosen (2009), Andrews and Shi (2010) and references therein.

Suppose that the true parameter value, say $\theta^* \in \Theta \subset \mathbf{R}$, is known to lie within the bounds $[\theta^l(x), \theta^u(x)]$ for each value $x \in \mathcal{X}$, where $\theta^l(x) \equiv \mathbf{E}[W_l | X = x]$ and $\theta^u(x) \equiv \mathbf{E}[W_u | X = x]$ are conditional expectations of W_l and W_u given $X = x$. The identification region for θ^* takes the form of intersection bounds:

$$\Theta_I = \cap_{x \in \mathcal{X}} \left[\theta^l(x), \theta^u(x)\right] = \left[\sup_{x \in \mathcal{X}} \theta^l(x), \inf_{x \in \mathcal{X}} \theta^u(x)\right]. \tag{2.2}$$

For each $\theta \in \Theta$, letting $Y_1 = W_l - \theta$ and $Y_2 = \theta - W_u$, we define our test statistic $\hat{T}_n(\theta)$ as proposed in this paper. Then the $(1-\alpha)\%$ level confidence region for θ^* is constructed as

$$\{\theta \in \Theta : \hat{T}_n(\theta) \leq z_{1-\alpha}\}.$$

The computational merit of our test being distribution free is prominent in this context, because we do not need to simulate the critical value $z_{1-\alpha}$ for each choice of θ. See Chernozhukov, Lee, and Rosen (2009) for an alternative approach based on sup tests.

3. Test Statistics and Asymptotic Properties

3.1. **Test Statistics and Asymptotic Validity.** Define $\mathcal{S}_j \equiv \{x \in \mathcal{X} : w_j(x) > 0\}$ for each $j \in \mathcal{J}$, and, given $\varepsilon > 0$, let $\mathcal{S}_j^\varepsilon$ be an ε-enlargement of $\mathcal{S}_j$, i.e., $\mathcal{S}_j^\varepsilon \equiv \{x + a : x \in \mathcal{S}_j,\ a \in [-\varepsilon, \varepsilon]^d\}$. For $1 \leq p < \infty$, let

$$r_{j,p}(x) \equiv \mathbf{E}[|Y_{ji}|^p | X_i = x] f(x). \tag{3.1}$$

We introduce the following assumptions.

Assumption 1: (i) $\min_{j \in \mathcal{J}} \inf_{x \in \mathcal{X}} r_{j,2}(x) > 0$, and $\max_{j \in \mathcal{J}} \sup_{x \in \mathcal{X}} r_{j,2p+2}(x) < \infty$.
(ii) For each $j \in \mathcal{J}$, $w_j(\cdot)$ is nonnegative on $\mathcal{X}$ and $0 < \int_{\mathcal{X}} w_j^s(x)dx < \infty$, where $s \in \{1, 2\}$.
(iii) There exists $\varepsilon > 0$ such that $\mathcal{S}_j^\varepsilon \subset \mathcal{X}$ for all $j \in \mathcal{J}$.

Assumption 2: $K(u) = \Pi_{s=1}^d K_s(u_s)$, $u = (u_1, \cdots, u_d)$, with each $K_s : \mathbf{R} \to \mathbf{R}$, $s = 1, \cdots, d$, satisfying that (a) $K_s(u_s) = 0$ for all $u_s \in \mathbf{R} \backslash [-1/2, 1/2]$, (b) K_s is of bounded variation, and (c) $||K_s||_\infty \equiv \sup_{u_s \in \mathbf{R}} |K_s(u_s)| < \infty$ and $\int K_s(u_s) du_s = 1$.

Assumption 1(iii) is introduced to avoid the boundary problem of kernel estimators by requiring that w_j have support inside an ε-shrunk subset of $\mathcal{X}$. The conditions for the kernel function in Assumption 2 are quite flexible, except that the kernel functions have bounded support.

Define for $j, k \in \mathcal{J}$ and $x \in \mathbf{R}^d$,

$$\begin{aligned}
\rho_{jk,n}(x) &\equiv \frac{1}{h^d} \mathbf{E}\left[Y_{ji} Y_{ki} K^2 \left(\frac{x - X_i}{h} \right) \right], \\
\rho_{jn}^2(x) &\equiv \frac{1}{h^d} \mathbf{E}\left[Y_{ji}^2 K^2 \left(\frac{x - X_i}{h} \right) \right], \\
\rho_{jk}(x) &\equiv \mathbf{E}\left[Y_{ji} Y_{ki} | X_i = x \right] f(x) \int K^2(u) du, \text{ and} \\
\rho_j^2(x) &\equiv \mathbf{E}[Y_{ji}^2 | X_i = x] f(x) \int K^2(u) du.
\end{aligned}$$

Let $\mathbb{Z}_1$ and $\mathbb{Z}_2$ denote mutually independent standard normal random variables. We introduce the following quantities:

$$\begin{aligned} a_{jn} &\equiv h^{-d/2} \int_{\mathcal{X}} \rho_{jn}^{p}(x) w_j(x) dx \cdot \mathbf{E}\Lambda_p(\mathbb{Z}_1) \text{ and} \\ \sigma_{jk,n} &\equiv \int_{\mathcal{X}} q_{jk,p}(x) \rho_{jn}^{p}(x) \rho_{kn}^{p}(x) w_j(x) w_k(x) dx, \end{aligned} \tag{3.2}$$

where $q_{jk,p}(x) \equiv \int_{[-1,1]^d} Cov(\Lambda_p(\sqrt{1-t_{jk}^2(x,u)}\mathbb{Z}_1 + t_{jk}(x,u)\mathbb{Z}_2),\ \Lambda_p(\mathbb{Z}_2))du$ and

$$t_{jk}(x,u) \equiv \frac{\rho_{jk}(x)}{\rho_j(x)\rho_k(x)} \cdot \frac{\int K(x)\,K(x+u)\,dx}{\int K^2(x)\,dx}.$$

Let Σ_n be a $J \times J$ matrix whose (j,k)-th entry is given by $\sigma_{jk,n}$. We also define Σ to be a $J \times J$ matrix whose (j,k)-th entry is given by σ_{jk}, where

$$\sigma_{jk} \equiv \int_{\mathcal{X}} q_{jk,p}(x) \rho_j^p(x) \rho_k^p(x) w_j(x) w_k(x) dx.$$

As for Σ, we introduce the following assumption.

Assumption 3: Σ is positive definite.

Assumption 3 excludes the case where Y_{ji} and Y_{ki} are perfectly correlated conditional on $X_i = x$ for almost all x and $\mathbf{E}[Y_{ji}^2|X_i = x]$ is a constant function for each $j \in \mathcal{J}$.

The following theorem is the main result of this paper.

Theorem 1: *Suppose that Assumptions 1-3 hold and that $h \to 0$ and $n^{-1/2}h^{-d} \to 0$ as $n \to \infty$. Furthermore, assume that $m_j(x) = 0$ for almost all $x \in \mathcal{X}$ and for all $j \in \mathcal{J}$. Then, for any $t = (t_1, \cdots, t_J)' \in \mathbf{R}^J \backslash \{0\}$,*

$$T_n \equiv \frac{1}{\sigma_{t,n}} \sum_{j=1}^{J} t_j \left\{ n^{p/2} h^{(p-1)d/2} \Gamma_j(\hat{g}_{jn}) - a_{jn} \right\} \xrightarrow{d} N(0,1),$$

where $\sigma_{t,n}^2 \equiv t'\Sigma_n t$.

Remark 1(a): The asymptotic theory does not require assumptions for m_j's and f beyond those in Assumption 1(i). In particular, the theory does not require continuity or differentiability of f or m_j's. This is due to our using the powerful Poissonization approach in Beirlant and Mason (1995) and GMZ, combined with our using $\rho_{jn}(x)$ instead of its limit as $n \to \infty$.

REMARK 1(B): When $J=1$ and $t=1$, $\sigma_{t,n}^2$ takes the simple form of $q_p \int_{\mathcal{X}} \rho_{1n}^{2p}(x) w_1^2(x) dx$, where

$$\begin{aligned} q_p &\equiv \int_{[-1,1]^d} Cov(\Lambda_p(\sqrt{1-t^2(u)}\mathbb{Z}_1 + t(u)\mathbb{Z}_2),\ \Lambda_p(\mathbb{Z}_2))du, \text{ and} \\ t(u) &\equiv \int K(x) K(x+u) dx / \int K^2(x) dx. \end{aligned}$$

To develop a feasible testing procedure, we construct estimators of a_{jn}'s and $\sigma_{t,n}^2$ as follows. First, define

$$\begin{aligned} \hat{\rho}_{jk,n}(x) &\equiv \frac{1}{nh^d} \sum_{i=1}^n Y_{ji} Y_{ki} K^2\left(\frac{x-X_i}{h}\right), \text{ and} \\ \hat{\rho}_{jn}^2(x) &\equiv \frac{1}{nh^d} \sum_{i=1}^n Y_{ji}^2 K^2\left(\frac{x-X_i}{h}\right). \end{aligned} \tag{3.3}$$

We estimate a_{jn} and $\sigma_{jk,n}$ by:

$$\begin{aligned} \hat{a}_{jn} &\equiv h^{-d/2} \int_{\mathcal{X}} \hat{\rho}_{jn}^p(x) w_j(x) dx \cdot \mathbf{E}\Lambda_p(\mathbb{Z}_1) \text{ and} \\ \hat{\sigma}_{jk,n} &\equiv \int_{\mathcal{X}} \hat{q}_{jk,p}(x) \hat{\rho}_{jn}^p(x) \hat{\rho}_{kn}^p(x) w_j(x) w_k(x) dx, \end{aligned}$$

where $\hat{q}_{jk,p}(x) \equiv \int_{[-1,1]^d} Cov(\Lambda_p(\sqrt{1-\hat{t}_{jk}^2(x,u)}\mathbb{Z}_1 + \hat{t}_{jk}(x,u)\mathbb{Z}_2),\ \Lambda_p(\mathbb{Z}_2))du$ and

$$\hat{t}_{jk}(x,u) \equiv \frac{\hat{\rho}_{jk,n}(x)}{\hat{\rho}_{jn}(x)\hat{\rho}_{kn}(x)} \cdot \frac{\int K(x) K(x+u) dx}{\int K^2(x) dx}.$$

Note that $\mathbf{E}\Lambda_1(\mathbb{Z}_1) = 1/\sqrt{2\pi} \approx 0.39894$ and $\mathbf{E}\Lambda_2(\mathbb{Z}_1) = 1/2$. When p is an integer, the covariance expression in $q_{jk,p}(x)$ can be computed using the moment generating function of a truncated multivariate normal distribution (Tallis, 1961). More practically, simulated draws from $\mathbb{Z}_1$ and $\mathbb{Z}_2$ can be used to compute the quantities $\mathbf{E}\Lambda_p(\mathbb{Z}_1)$ and $q_{jk,p}(x)$ for general values of p. The integrals appearing above can be evaluated using methods of numerical integration. We define $\hat{\Sigma}_n$ to be a $J \times J$ matrix whose (j,k)-th entry is given by $\hat{\sigma}_{jk,n}$.

Fix $t = (t_1, \cdot\cdot\cdot, t_J)' \in (0,\infty)^J$, and let $\hat{\sigma}_{t,n}^2 \equiv t'\hat{\Sigma}_n t$. Our test statistic is taken to be

$$\hat{T}_n \equiv \frac{1}{\hat{\sigma}_{t,n}} \sum_{j=1}^J t_j \left\{ n^{p/2} h^{(p-1)d/2} \Gamma_j(\hat{g}_{jn}) - \hat{a}_{jn} \right\}. \tag{3.4}$$

Let $z_{1-\alpha} \equiv \Phi^{-1}(1-\alpha)$, where Φ denotes the cumulative distribution function of $N(0,1)$. This paper proposes using the following test:

$$\text{Reject } H_0 \text{ if and only if } \hat{T}_n > z_{1-\alpha}. \tag{3.5}$$

The following theorem shows that the test has an asymptotically valid size.

THEOREM 2: *Suppose that Assumptions 1-3 hold and that* $h \to 0$ *and* $n^{-1/2}h^{-3d/2} \to 0$, *as* $n \to \infty$. *Furthermore, assume that the kernel function* K *in Assumption 3 is nonnegative. Then under the null hypothesis, for any* $t = (t_1, \cdots, t_J)' \in (0, \infty)^J$,

$$\lim_{n\to\infty} P\{\hat{T}_n > z_{1-\alpha}\} \leq \alpha,$$

with equality holding if $m_j(x) = 0$ *for almost all* $x \in \mathcal{X}$ *and for all* $j \in \mathcal{J}$.

REMARK 2(A): The nonparametric test does not require smoothness conditions for m_j's and f, even after replacing a_{jn}'s and $\sigma^2_{t,n}$ by their estimators. This result uses the assumption that the kernel function K is nonnegative to control the size of the test. (See the proof of Theorem 2 for details.)

REMARK 2(B): The bandwidth condition for Theorem 2 is stronger than that in Theorem 1. This is mainly due to the treatment of the estimation errors in $\hat{a}_{jn}$ and $\hat{\sigma}^2_{t,n}$. For the bandwidth parameter, it suffices to take $h = c_1 n^{-s}$ with $0 < s < 1/(3d)$ for a constant $c_1 > 0$.

In general, optimal bandwidth choice for nonparametric testing is different from that for nonparametric estimation as we need to balance the size and power of the test instead of the bias and variance of an estimator. For example, Gao and Gijbels (2008) considered testing a parametric null hypothesis against a nonparametric alternative and derived a bandwidth-selection rule by utilizing an Edgeworth expansion of the asymptotic distribution of the test statistic concerned. The methods of Gao and Gijbels (2008) are not directly applicable to our tests, and it is a challenging problem to develop a theory of optimal bandwidths for our tests. We provide some simulation evidence regarding sensitivity to the choice of h in Section 4.

REMARK 2(C): The asymptotic rejection probability under the null hypothesis achieves its maximum of α when $m_j(x) = 0$ for almost all $x \in \mathcal{X}$ and for all $j \in \mathcal{J}$. Hence we call the latter case *the least favorable case of the null hypothesis.*

REMARK 2(E): According to Theorems 1-2, each choice of $t \in (0, \infty)^J$ leads to an asymptotically valid test. The actual choice of t may reflect the relative importance of individual inequality restrictions. When it is of little practical significance to treat individual inequality restrictions differently, one may choose simply $t = (1, \cdots, 1)'$. Perhaps more naturally, to avoid undue influences of different scales across Y_{ji}'s, one may use the following

$$\hat{t} \equiv (\hat{\sigma}^{-1/2}_{11,n}, \ldots, \hat{\sigma}^{-1/2}_{JJ,n})', \text{ with } \hat{\sigma}_{jj,n} \equiv q_p \int_{\mathcal{X}} \hat{\rho}^{2p}_{jn}(x) w_j^2(x) dx, j \in \mathcal{J},$$

where $\hat{\rho}^2_{jn}(x)$ is given as in (3.3). Then $\hat{t}$ is consistent for $t \equiv (\sigma_{11,n}^{-1/2}, \ldots, \sigma_{JJ,n}^{-1/2})'$ (see the proof of Theorem 2), and just as the estimation error of $\hat{\sigma}_{t,n}$ in (3.5) leaves the limiting distribution of T_n under the null hypothesis intact, so does the estimation error of $\hat{t}$.

The following result shows the consistency of the test in (3.5) against fixed alternatives.

THEOREM 3: *Suppose that Assumptions 1-3 hold and that $h \to 0$ and $n^{-1/2}h^{-3d/2} \to 0$, as $n \to \infty$. Then, under $H_1 : \Gamma_j(g_j) > 0$ for some $j \in \mathcal{J}$, we have*

$$\lim_{n\to\infty} P\{\hat{T}_n > z_{1-\alpha}\} = 1.$$

3.2. **Local Asymptotic Power.** We determine the power of the test in (3.5) against some sequences of local alternatives. Consider the following sequences of local alternatives converging to the null hypothesis at the rate $n^{-1/2}$:

$$H_\delta : g_j(x) = n^{-1/2}\delta_j(x), \text{ for each } j \in \mathcal{J}, \tag{3.6}$$

where $\delta_j(\cdot)$'s are bounded real functions on $\mathbf{R}^d$.

The following theorem establishes a representation of the local asymptotic power functions, when $p \in \{1, 2\}$. For simplicity of notation, let us introduce the following definition: for $s \in \{1, 2\}$, $z \in \{-1, 0, 1\}$, a given weight function vector $w \equiv (w_1, \cdots, w_J)$, and the direction $\delta = (\delta_1, \cdots, \delta_J)'$, let $\eta_{s,z}(w, \delta) \equiv \sum_{j=1}^J t_j \int_{\mathcal{X}} \delta_j^s(x)\rho_j^z(x)w_j(x)dx$, and let $\sigma_t^2 \equiv t'\Sigma t$.

THEOREM 4: *Suppose that Assumptions 1-3 hold and that $h \to 0$ and $n^{-1/2}h^{-3d/2} \to 0$, as $n \to \infty$.*

(i) *If $p = 1$, then under H_δ, we have*

$$\lim_{n\to\infty} P\{\hat{T}_n > z_{1-\alpha}\} = 1 - \Phi(z_{1-\alpha} - \eta_{1,0}(w, \delta)/2\sigma_t).$$

(ii) *If $p = 2$, then under H_δ, we have*

$$\lim_{n\to\infty} P\{\hat{T}_n > z_{1-\alpha}\} = 1 - \Phi(z_{1-\alpha} - \eta_{1,1}(w, \delta)/(\sigma_t\sqrt{\pi/2})).$$

REMARK 4: This theorem gives explicit local asymptotic power functions, when $p = 1$ and $p = 2$. The local power of the test is greater than the size α, whenever $\eta_{1,0}(w, \delta)$ in the case of $p = 1$ and $\eta_{1,1}(w, \delta)$ in the case of $p = 2$ are strictly positive.

When $J = 1$, we can compute an optimal weight function that maximizes the local power against a given direction δ. See Stute (1997) for related results of optimal directional tests, and Tripathi and Kitamura (1997) for results of optimal directional and average tests based on smoothed empirical likelihoods.

We let $J = 1$ and define $\sigma_p^2(w_1) \equiv q_p \int_{\mathcal{X}} \rho_{1n}^{2p}(x) w_1^2(x) dx$. The optimal weight function (denoted by w_p^*) is taken to be a maximizer of the drift term $\eta_{1,0}(w_1, \delta_1)/\sigma_1(w_1)$ (in the case of $p = 1$) or $\eta_{1,1}(w_1, \delta_1)/\sigma_2(w_1)$ (in the case of $p = 2$) with respect to w_1 under the constraint that $w_1 \geq 0$ and $\int_{\mathcal{X}} w(x)\rho^{2p}(x)dx = 1$. The latter condition is for a scale normalization. Let $\delta_1^+ = \max\{\delta_1, 0\}$. Since ρ_1 and w_1 are nonnegative, the Cauchy-Schwarz inequality suggests that the optimal weight function is given by

$$w_p^*(x) = \begin{cases} \frac{\delta_1^+(x)\rho_1^{-2}(x)}{\sqrt{\int_{\mathcal{X}} (\delta_1^+)^2(x)\rho_1^{-2}(x)dx}}, & \text{if } p = 1, \text{ and} \\ \frac{\delta_1^+(x)\rho_1^{-3}(x)}{\sqrt{\int_{\mathcal{X}} (\delta_1^+)^2(x)\rho_1^{-2}(x)dx}}, & \text{if } p = 2. \end{cases} \tag{3.7}$$

To satisfy Assumption 1(iii), we assume that the support of δ_1 is contained in an ε-shrunk subset of $\mathcal{X}$. With this choice of an optimal weight function, the local power function becomes:

$$\begin{array}{ll} 1 - \Phi\left(z_{1-\alpha} - \frac{\sqrt{\int_{\mathcal{X}} (\delta_1^+)^2(x)\rho_1^{-2}(x)dx}}{2\sqrt{q_1}}\right), & \text{if } p = 1, \text{ and} \\ 1 - \Phi\left(z_{1-\alpha} - \frac{\sqrt{\int_{\mathcal{X}} (\delta_1^+)^2(x)\rho_1^{-2}(x)dx}}{\sqrt{q_2\pi/2}}\right), & \text{if } p = 2. \end{array}$$

3.3. **Comparison with Testing Functional Equalities.** It is straightforward to follow the proofs of Theorems 1-3 to develop a test for equality restrictions:

$$\begin{aligned} H_0 &: m_j(x) = 0 \text{ for all } (x, j) \in \mathcal{X} \times \mathcal{J}, \text{ vs.} \\ H_1 &: m_j(x) \neq 0 \text{ for some } (x, j) \in \mathcal{X} \times \mathcal{J}. \end{aligned} \tag{3.8}$$

For this test, we redefine $\Lambda_p(v) = |v|^p$ and, using this, redefine $\hat{T}_n$ in (3.4) and σ_t^2. Then under the null hypothesis,

$$\hat{T}_n \xrightarrow{d} N(0, 1).$$

Therefore, we can take a critical value in the same way as before. The asymptotic validity of this test under the null hypothesis follows under precisely the same conditions as in Theorem 2. However, the convergence rates of the inequality tests and the equality tests under local alternatives are different, as we shall see now.

Consider the local alternatives converging to the null hypothesis at the rate $n^{-1/2}h^{-d/4}$:

$$H_\delta^* : g_j(x) = n^{-1/2}h^{-d/4}\delta_j(x), \text{ for each } j \in \mathcal{J}, \tag{3.9}$$

where $\delta_j(\cdot)$'s are again bounded real functions on $\mathbf{R}^d$. The following theorem establishes the local asymptotic power functions of the test based on $\hat{T}_n$.

Theorem 5: *Suppose that Assumptions 1-3 hold and that $h \to 0$ and $n^{-1/2}h^{-3d/2} \to 0$, as $n \to \infty$.*

(i) *If* $p = 1$, *then under* H_δ^*, *we have*

$$\lim_{n\to\infty} P\{\hat{T}_n > z_{1-\alpha}\} = 1 - \Phi(z_{1-\alpha} - \eta_{2,-1}(w,\delta)/(\sqrt{2\pi}\sigma_t)).$$

(ii) *If* $p = 2$, *then under* H_δ^*, *we have*

$$\lim_{n\to\infty} P\{\hat{T}_n > z_{1-\alpha}\} = 1 - \Phi(z_{1-\alpha} - \eta_{2,0}(w,\delta)/\sigma_t).$$

Remark 5(a): Theorem 5 shows that the equality tests (on (3.8)), in contrast to the inequality tests (on (1.1)), have non-trivial local power against alternatives converging to the null at rate $n^{-1/2}h^{-d/4}$, which is slower than $n^{-1/2}$.

Remark 5(b): Since $\eta_{2,-1}(w,\delta)$ and $\eta_{2,0}(w,\delta)$ are always nonnegative, the equality tests are *locally asymptotically unbiased* against any local alternatives. In contrast, the terms $\eta_{1,0}(w,\delta)$ and $\eta_{1,1}(w,\delta)$ in the local asymptotic power functions of the inequality tests in Theorem 4 can take negative values for some local alternatives, implying that the inequality tests might be asymptotically biased against such local alternatives.

When $J = 1$, an optimal directional test under (3.9) can also be obtained by following the arguments leading up to (3.7) so that

$$w_p^*(x) = \begin{cases} \frac{\delta_1^2(x)\rho_1^{-3}(x)}{\sqrt{\int_{\mathcal{X}} \delta_1^4(x)\rho_1^{-4}(x)dx}}, & \text{if } p = 1, \text{ and} \\ \frac{\delta_1^2(x)\rho_1^{-4}(x)}{\sqrt{\int_{\mathcal{X}} \delta_1^4(x)\rho_1^{-4}(x)dx}}, & \text{if } p = 2. \end{cases}$$

Similarly as before, let the support of δ_1 be contained in an ε-shrunk subset of $\mathcal{X}$. The optimal weight function yields the following local power functions:

$$\begin{array}{ll} 1 - \Phi\left(z_{1-\alpha} - \frac{\sqrt{\int \delta_1^4(x)\rho_1^{-4}(x)dx}}{\sqrt{2\pi}\bar{q}_1}\right), & \text{if } p = 1, \text{ and} \\ 1 - \Phi\left(z_{1-\alpha} - \frac{\sqrt{\int \delta_1^4(x)\rho_1^{-4}(x)dx}}{\sqrt{\bar{q}_2}}\right), & \text{if } p = 2, \end{array}$$

where $\bar{q}_p \equiv \int_{[-1,1]^d} Cov(\Lambda_p(|\sqrt{1-t^2(u)}\mathbb{Z}_1 + t(u)\mathbb{Z}_2|^p,\ |\mathbb{Z}_2|^p)du$, for $p \in \{1,2\}$.

4. Monte Carlo Experiments

This section reports the finite-sample performance of the one-sided L_1- and L_2-type tests from a Monte Carlo study. In the experiments, n observations of a pair of random variables (Y, X) were generated from $Y = m(X) + \sigma(X)U$, where $X \sim$ Unif$[0,1]$ and $U \sim N(0,1)$ and X and U are independent. In all the experiments, we set $\mathcal{X} = [0.05, 0.95]$.

To evaluate the finite-sample size of the tests, we first set $m(x) \equiv 0$. In addition, we consider the following alternative model

$$m(x) = x(1-x) - c_m \tag{4.1}$$

where $c_m \in \{0.25, 0.15, 0.05\}$. When $c_m = 0.25$, we have $m(x) < 0$ for all $x \neq 0.5$ and $m(x) = 0$ with $x = 0.5$. Hence, this case corresponds to the "interior" of the null hypothesis. In view of asymptotic theory, we expect the empirical probability of rejecting H_0 will go to zero as n gets large. When $c_m = 0.15$ or $c_m = 0.05$, we have $m(x) > 0$ for some x. Therefore, these two cases are considered to see the the finite-sample power of our tests. Two different functions of $\sigma(x)$ are considered: $\sigma(x) \equiv 1$ (homoskedastic error) and $\sigma(x) = x$ (heteroskedastic error).

The experiments use sample sizes of $n = 50, 200, 1000$ and the nominal level of $\alpha = 0.05$. We performed 1000 Monte Carlo replications in each experiment. In implementing both L_1 and L_2-type tests, we used $K(u) = (3/2)(1-(2u)^2)I(|u| \leq 1/2)$ and $h = c_h \times \hat{s}_X \times n^{-1/5}$, where $I(A)$ is the usual indicator function that has value one if A is true and zero otherwise, c_h is a constant and $\hat{s}_X$ is the sample standard deviation of X. To check the sensitivity to the choice of the bandwidth, three different values of c_h are considered: $c_h \in \{1, 1.5, 2\}$. Finally, we considered the uniform weight function: $w(x) = 1$.

The results are shown in Tables 1 and 2. First, see Table 1 for the simulation results when H_0 is true. When H_0 is true and $m(x) \equiv 0$, the differences between the nominal and empirical rejection probabilities are small, especially when $n = 1000$. Also, the results are not very sensitive to the bandwidth choice. When H_0 is true and $m(x)$ is (4.1) with $c_m = 0.25$ (the interior case), the empirical rejection probabilities are smaller than the nominal level and become almost zero for $n = 1000$.

When H_0 is false and the correct model is (4.1) with $c_m = 0.15$ or $c_m = 0.05$, (see Table 2), the powers of both the L_1 and L_2 tests are better for the model with $c_m = 0.05$ than with $c_m = 0.15$. This finding is consistent with asymptotic theory since it is likely that our test will be more powerful when $\int_{\mathcal{X}} m(x)w(x)dx$ is larger. Note that in both cases ($c_m = 0.15$ and 0.05), the rejection probabilities increase as n gets large. This is in line with the asymptotic theory in the preceding sections, for our test is consistent for both values of c_m.

5. Proofs

The proof of Theorem 1 overall follows the structure of the proof of the finite-dimensional convergence in Theorem 1.1 of GMZ. The lemma below is an extension of Lemma 6.1 of GMZ.

Lemma A1: *Let $J(\cdot) : \mathbf{R}^d \to \mathbf{R}$ be a Lebesgue integrable function and $H : \mathbf{R}^d \to \mathbf{R}$ be a bounded function with compact support S. Then, for almost every $y \in \mathbf{R}^d$,*

$$\int_{\mathbf{R}^d} J(x)H_h(y-x)\,dx \to J(y)\int_S H(x)dx, \text{ as } h \to 0,$$

where $H_h(x) \equiv H(x/h)/h^d$.

Furthermore, suppose that $\bar{J} \equiv \int |J(z)|dz > 0$. *Then for all* $0 < \varepsilon < \bar{J} \equiv \int |J(z)|dz$, *there exist* $M > 0$, $\nu > 0$ *and a Borel set* B *of finite Lebesgue measure* $m(B)$ *such that* $B \subset [-M+v, M-v]^d$, $\alpha \equiv \int_{\mathbf{R}^d \setminus [-M,M]^d} |J(z)|dz > 0$, $\int_B |J(z)|dz > \bar{J} - \varepsilon$, J *is bounded and continuous on* B, *and*

$$\sup_{y \in B} \left| \int_{\mathbf{R}^d} J(x) H_h(y-x)\, dx - J(y) \int_S H(x) dx \right| \to 0, \text{ as } h \to 0.$$

Proof : We consider a simple function $H_m(x) \equiv \sum_{i=1}^m a_i 1\{x \in A_i \cap S\}$ for some numbers $a_i \in \mathbf{R}$ and measurable sets $A_i \subset \mathbf{R}^d$ such that $|H_m(x) - H(x)| \to 0$ as $m \to \infty$. Without loss of generality, we let $A_i \cap S$ be a rectangle. Let $H_{h,m}(x) \equiv H_m(x/h)/h^d$ and note that

$$\begin{aligned} \int_{\mathbf{R}^d} J(x) H_{h,m}(y-x)\, dx &= \frac{1}{h^d} \sum_{i=1}^m a_i \int_{\mathbf{R}^d} J(x) 1\{(y-x)/h \in A_i \cap S\}\, dx \\ &= \sum_{i=1}^m \frac{a_i}{h^d} \int_{y - h(A_i \cap S)} J(x) dx, \end{aligned}$$

where $y - h(A_i \cap S) = \{y - hz : z \in A_i \cap S\}$. Hence

$$\begin{aligned} & \left| \int_{\mathbf{R}^d} J(x) H_{h,m}(y-x)\, dx - J(y) \int_S H_m(x) dx \right| \\ = & \left| \sum_{i=1}^m a_i \left\{ \frac{1}{h^d} \int_{y-h(A_i \cap S)} J(x) dx - J(y) m(A_i \cap S) \right\} \right| \\ = & \left| \sum_{i=1}^m \frac{a_i}{h^d} \int_{y-h(A_i \cap S)} \{J(x) - J(y)\}\, dx \right| \\ \le & \sum_{i=1}^m \frac{a_i}{h^d} \int_{y - hB_i} |J(x) - J(y)|\, dx, \end{aligned}$$

where B_i is a compact ball in $\mathbf{R}^d$ centered at zero containing $A_i \cap S$. For almost every $y \in \mathbf{R}^d$ (with respect to the Lebesgue measure), the last sum converges to zero as $h \to 0$ by the Lebesgue Differentiation Theorem (e.g. Theorem 11.1 of DiBenedetto (2001), p.192). By sending $m \to \infty$, we obtain the first desired result.

The second statement can be proved following the proof of Lemma 6.1 of GMZ. Since J is Lebesgue integrable, the integral $\int_{\mathbf{R}^d \setminus [-M,M]^d} |J(z)|dz$ is continuous in M and converges to zero as $M \to \infty$. We can find $M > 0$ and $v > 0$ such that

$$\int_{\mathbf{R}^d \setminus [-M,M]^d} |J(z)|dz = \varepsilon/8 \text{ and } \int_{\mathbf{R}^d \setminus [-M+v, M-v]^d} |J(z)|dz = \varepsilon/4.$$

The construction of the desired set $B \subset [-M+v, M-v]^d$ can be done using the arguments in the proof of Lemma 6.1 of GMZ. ■

The following result is a special case of Theorem 1 of Sweeting (1977) with $g(x) = \min(x,1)$ (in his notation). See also Fact 6.1 of GMZ and Fact 4 of Mason (2009) for applications of Theorem 1 of Sweeting (1977).

LEMMA A2 (SWEETING (1977)): *Let $\mathbb{Z} \in \mathbf{R}^k$ be a mean zero normal random vector with covariance matrix I and $\{W_i\}_{i=1}^n$ is a set of i.i.d. random vectors in $\mathbf{R}^k$ such that $\mathbf{E}W_i = 0$, $\mathbf{E}W_iW_i' = I$, and $\mathbf{E}||W_i||^r < \infty$, $r \geq 3$. Then for any Borel measurable function $\varphi : \mathbf{R}^k \to \mathbf{R}$ such that*

$$\sup_{x \in \mathbf{R}^k} \frac{|\varphi(x) - \varphi(0)|}{1 + ||x||^r \min(||x||, 1)} < \infty,$$

we have

$$\begin{aligned} & \left| \mathbf{E}\left[\varphi\left(\frac{1}{\sqrt{n}} \sum_{i=1}^n W_i \right) \right] - \mathbf{E}\left[\varphi(\mathbb{Z}) \right] \right| \\ \leq \; & c_1 \left(\sup_{x \in \mathbf{R}^k} \frac{|\varphi(x) - \varphi(0)|}{1 + ||x||^r \min(||x||, 1)} \right) \left\{ \frac{1}{\sqrt{n}} \mathbf{E}||W_i||^3 + \frac{1}{n^{(r-2)/2}} \mathbf{E}||W_i||^r \right\} \\ & + c_2 \mathbf{E}\left[\omega_\varphi \left(\mathbb{Z}; \frac{c_3}{\sqrt{n}} \mathbf{E}||W_i||^3 \right) \right], \end{aligned}$$

where c_1, c_2 and c_3 are positive constants that depend only on k and r and

$$\omega_\varphi(x; \varepsilon) \equiv \sup\left\{ |\varphi(x) - \varphi(y)| : y \in \mathbf{R}^k, ||x - y|| \leq \varepsilon \right\}.$$

The following algebraic inequality is used frequently throughout the proofs.

LEMMA A3: *For any $a, b \in \mathbf{R}$, let $a_+ = \max(a,0)$ and $b_+ = \max(b,0)$. Furthermore, for any real $a \geq 0$, if $a = 0$, we define $\lceil a \rceil = 1$, and if $a > 0$, we define $\lceil a \rceil$ to be the smallest integer among the real numbers greater than or equal to a. Then for any $p \geq 1$,*

$$\begin{aligned} \max\left\{ |a_+^p - b_+^p|, ||a|^p - |b|^p| \right\} & \leq 2p|a-b| \left(\sum_{k=0}^{\lceil p-1 \rceil} \frac{\lceil p-1 \rceil!}{k!} |a-b|^{\lceil p-1 \rceil - k} |b|^k \right)^{(p-1)/\lceil p-1 \rceil} \\ & \leq C \sum_{k=0}^{\lceil p-1 \rceil} |a-b|^{p - \frac{(p-1)k}{\lceil p-1 \rceil}} |b|^k, \end{aligned}$$

for some $C > 0$ that depends only on p.

PROOF : First, we show the inequality for the case where p is a positive integer. We prove first that $||a|^p - |b|^p|$ has the desired bound. Note that in this case of p being a positive integer, the bound takes the following form:

$$2\sum_{k=0}^{p-1}\frac{p!}{k!}|a-b|^{p-k}|b|^k.$$

When $p=1$, the bound is trivially obtained. Suppose now that the inequality holds for a positive integer q. First, note that using the mean-value theorem, convexity of the function $f(x)=x^q$ for $q\geq 1$, and the triangular inequality,

$$\begin{aligned}||a|^{q+1}-|b|^{q+1}| &\leq (q+1)|a-b|\sup_{\alpha\in[0,1]}\left(\alpha|a|+(1-\alpha)|b|\right)^q\\ &\leq (q+1)|a-b|\sup_{\alpha\in[0,1]}\left(\alpha|a|^q+(1-\alpha)|b|^q\right)\\ &\leq (q+1)|a-b|\left(||a|^q-|b|^q|+2|b|^q\right).\end{aligned}$$

As for $||a|^q-|b|^q|$, we apply the inequality to bound the last term by

$$\begin{aligned}&(q+1)|a-b|\left(2\sum_{k=0}^{q-1}\frac{q!}{k!}|a-b|^{q-k}|b|^k+2|b|^q\right)\\ =\;&2\sum_{k=0}^{q}\frac{(q+1)!}{k!}|a-b|^{q-k+1}|b|^k.\end{aligned}$$

Therefore, by the principle of mathematical induction, the desired bound in the case of p being a positive integer follows.

Certainly, we obtain the same bound for $|a_+^p-b_+^p|$ when $p=1$. When $p>1$, we observe that

$$|a_+^p-b_+^p|\leq p|a-b|\left(||a|^{p-1}-|b|^{p-1}|+2|b|^{p-1}\right).$$

By applying the previous inequality to $||a|^{p-1}-|b|^{p-1}|$, we obtain the desired bound for $|a_+^p-b_+^p|$ when p is a positive integer.

Now, let us consider the case where p is a real number greater than or equal to 1. Again, we first show that $||a|^p-|b|^p|$ has the desired bound. Using the mean-value theorem as before and the fact that $|a+b|\leq 2^{1-1/s}\left(|a|^s+|b|^s\right)^{1/s}$ for all $s\in[1,\infty)$ and all $a,b\in\mathbf{R}$, we find that for $u\equiv\lceil p-1\rceil$,

$$\begin{aligned}||a|^p-|b|^p| &\leq p\,|a-b|\left(|a|^{p-1}+|b|^{p-1}\right)\\ &\leq p\,|a-b|\,2^{1-(p-1)/u}\left(|a|^u+|b|^u\right)^{(p-1)/u}\\ &\leq p\,|a-b|\,2^{1-(p-1)/u}\left(||a|^u-|b|^u|+2|b|^u\right)^{(p-1)/u}.\end{aligned}$$

Since u is a positive integer, using the previous bound, we bound the right-hand side by

$$p\,|a-b|\,2^{1-(p-1)/u}\left(2\sum_{k=0}^{u-1}\frac{u!}{k!}|a-b|^{u-k}|b|^k+2|b|^u\right)^{(p-1)/u}.$$

Consolidating the sum in the parentheses, we obtain the wanted bound.

As for the second inequality, observe that

$$\begin{aligned}
&2p|a-b|\left(\sum_{k=0}^{\lceil p-1\rceil}\frac{\lceil p-1\rceil!}{k!}|a-b|^{\lceil p-1\rceil-k}|b|^k\right)^{(p-1)/\lceil p-1\rceil}\\
\leq\ & C\max_{k\in\{0,1,\cdots,\lceil p-1\rceil\}}|a-b|^{p-k\{(p-1)/\lceil p-1\rceil\}}|b|^k\leq C\sum_{k=0}^{\lceil p-1\rceil}|a-b|^{p-k\{(p-1)/\lceil p-1\rceil\}}|b|^k,
\end{aligned}$$

for some $C>0$ that depends only on p. We can obtain the same bound for $|a_+^p-b_+^p|$ by noting that $|a_+^p-b_+^p|\leq p\,|a-b|\,(|a|^{p-1}+|b|^{p-1})$ and following the same arguments afterwards as before. ■

Define for $j\in\mathcal{J}$,

$$k_{jn,r}(x)\equiv h^{-d}\mathbf{E}\left[\left|Y_{ji}K\left(\frac{x-X_i}{h}\right)\right|^r\right],\ r\geq 1. \tag{5.1}$$

LEMMA A4: *Suppose that Assumptions 1(i)(iii) and 2 hold and $h\to 0$ as $n\to\infty$. Then for $\varepsilon>0$ in Assumption 1(iii), there exist positive integer n_0 and constants $c_1,c_2>0$ such that for all $n\geq n_0$, all $r\in[1,2p+2]$, and all $j\in\mathcal{J}$,*

$$0<c_1\leq\inf_{x\in\mathcal{S}_j^{\varepsilon/2}}\rho_{jn}^2(x)\ \textit{and}$$

$$\sup_{x\in\mathcal{S}_j^{\varepsilon/2}}k_{jn,r}(x)\leq c_2<\infty.$$

PROOF: Since $h\to 0$ as $n\to\infty$, we apply change of variables to find that from large n on,

$$\begin{aligned}
\inf_{x\in\mathcal{S}_j^{\varepsilon/2}}\rho_{jn}^2(x) &= \inf_{x\in\mathcal{S}_j^{\varepsilon/2}}\frac{1}{h^d}\mathbf{E}\left[Y_{ji}^2K^2\left(\frac{x-X_i}{h}\right)\right]\\
&\geq \inf_{x\in\mathcal{S}_j^{\varepsilon}}\mathbf{E}\left[Y_{ji}^2|X=x\right]f(x)\int_{[-1/2,1/2]^d}K^2(u)\,du\\
&\geq \inf_{x\in\mathcal{X}}\mathbf{E}\left[Y_{ji}^2|X=x\right]f(x)\int_{[-1/2,1/2]^d}K^2(u)\,du>c_1,
\end{aligned}$$

for some $c_1>0$ by Assumptions 1(i) and 2. Similarly, from some large n on,

$$\sup_{x\in\mathcal{S}_j^{\varepsilon/2}}k_{jn,r}(x)\leq\sup_{x\in\mathcal{X}}\mathbf{E}\left[|Y_{ji}|^r|X=x\right]f(x)\int|K(u)|^r du<\infty,$$

by Assumptions 1(i) and 2. ■

Define for each $j \in \mathcal{J}$,

$$\hat{g}_{jN}(x) \equiv \frac{1}{nh^d} \sum_{i=1}^{N} Y_{ji} K\left(\frac{x - X_i}{h}\right), \ x \in \mathcal{X},$$

where N is a Poisson random variable that is common across $j \in \mathcal{J}$, has mean n, and is independent of $\{(Y_{ji}, X_i) : j \in \mathcal{J}\}_{i=1}^{\infty}$. Let for each $j \in \mathcal{J}$,

$$v_{jn}(x) \equiv \hat{g}_{jn}(x) - \mathbf{E}\hat{g}_{jn}(x), \text{ and } v_{jN}(x) \equiv \hat{g}_{jN}(x) - \mathbf{E}\hat{g}_{jn}(x).$$

We define, for each $j \in \mathcal{J}$,

$$\begin{aligned} \xi_{jn}(x) &\equiv \frac{\sqrt{nh^d} v_{jN}(x)}{\rho_{jn}(x)} \text{ and} \\ V_{jn}(x) &\equiv \frac{\sum_{i \geq N_1} \{Y_{ji} K((x - X_i)/h) - \mathbf{E}(Y_{ji} K((x - X_i)/h))\}}{\sqrt{\mathbf{E}[Y_{ji}^2 K^2((x - X_i)/h)]}}, \end{aligned} \tag{5.2}$$

where N_1 denotes a Poisson random variable with mean 1 that is independent of $\{(Y_{ji}, X_i) : j \in \mathcal{J}\}_{i=1}^{\infty}$. Then, $Var(V_{jn}(x)) = 1$. Let $V_{jn}^{(i)}(x), \ i = 1, \cdots, n$, be i.i.d. copies of $V_{jn}(x)$ so that

$$\xi_{jn}(x) \overset{d}{=} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} V_{jn}^{(i)}(x). \tag{5.3}$$

Lemma A5: *Suppose that Assumptions 1(i)(iii) and 2 hold and $h \to 0$ as $n \to \infty$ and*

$$\text{limsup}_{n \to \infty} n^{-r/2+1} h^{(1-r/2)d} < C,$$

for some constant $C > 0$ and for $r \in [2, 2p+2]$. Then, for $\varepsilon > 0$ in Assumption 1(iii),

$$\sup_{x \in \mathcal{S}_j^{\varepsilon/2}} \mathbf{E}\left[|V_{jn}(x)|^r\right] \leq C_1 h^{(1-r/2)d} \text{ and } \sup_{x \in \mathcal{S}_j^{\varepsilon/2}} \mathbf{E}\left[|\xi_{jn}(x)|^r\right] \leq C_2,$$

where C_1 and C_2 are constants that depend only on r.

Proof : For all $x \in \mathcal{S}_j^{\varepsilon/2}$, $\mathbf{E}[V_{jn}^2(x)] = 1$. Recall the definition of $k_{jn,r}(x)$ in (5.1). Then for some $C_1 > 0$,

$$\sup_{x \in \mathcal{S}_j^{\varepsilon/2}} \mathbf{E}\left|V_{jn}(x)\right|^r \leq \sup_{x \in \mathcal{S}_j^{\varepsilon/2}} \frac{h^d k_{jn,r}(x)}{h^{rd/2} \rho_{jn}^r(x)} \leq C_1 h^{(1-r/2)d}, \tag{5.4}$$

by Lemma A4, completing the proof of the first statement.

As for the second statement, using (5.3) and applying Rosenthal's inequality (e.g. (2.3) of GMZ), we deduce that for positive constants C_3, C_4 and C_5 that depend only on r,

$$\begin{aligned}\sup_{x\in\mathcal{S}_j^{\varepsilon/2}} \mathbf{E}\left[|\xi_{jn}(x)|^r\right] &\leq C_3 \sup_{x\in\mathcal{S}_j^{\varepsilon/2}} \max\{(\mathbf{E}V_{jn}^2(x))^{r/2}, n^{-r/2+1}\mathbf{E}|V_{jn}(x)|^r\} \\ &\leq C_4 \max\left\{1, C_5 n^{-r/2+1}h^{(1-r/2)d}\right\}\end{aligned}$$

by (5.4). By the condition that $\text{limsup}_{n\to\infty} n^{-r/2+1}h^{(1-r/2)d} < C$, the desired result follows. ■

The following lemma is adapted from Lemma 6.3 of GMZ. The result is obtained by combining Lemmas A2-A5.

LEMMA A6: *Suppose that Assumptions 1 and 2 hold and $h \to 0$ and $n^{-1/2}h^{-d} \to 0$ as $n \to \infty$. Then for any Borel set $A \subset \mathbf{R}^d$ and for any $j \in \mathcal{J}$,*

$$\begin{aligned}\int_A \left\{n^{p/2}h^{(p-1)d/2}\mathbf{E}\Lambda_p(v_{jN}(x)) - h^{-d/2}\rho_{jn}^p(x)\mathbf{E}\Lambda_p(\mathbb{Z}_1)\right\} w_j(x)dx &\to 0, \\ \int_A \left\{n^{p/2}h^{(p-1)d/2}\mathbf{E}\Lambda_p(v_{jn}(x)) - h^{-d/2}\rho_{jn}^p(x)\mathbf{E}\Lambda_p(\mathbb{Z}_1)\right\} w_j(x)dx &\to 0.\end{aligned}$$

PROOF : Recall the definition of $\xi_{jn}(x)$ in (5.2) and write

$$\begin{aligned}& n^{p/2}h^{(p-1)d/2}\mathbf{E}\Lambda_p(v_{jN}(x)) - h^{-d/2}\rho_{jn}^p(x)\mathbf{E}\Lambda_p(\mathbb{Z}_1) \\ =\; & h^{-d/2}\rho_{jn}^p(x)\left\{\mathbf{E}\Lambda_p(\xi_{jn}(x)) - \mathbf{E}\Lambda_p(\mathbb{Z}_1)\right\}.\end{aligned}$$

In view of Lemma A4 and Assumption 1(ii), we find that it suffices for the first statement of the lemma to show that

$$\sup_{x\in\mathcal{S}_j} \left|\mathbf{E}\Lambda_p(\xi_{jn}(x)) - \mathbf{E}\Lambda_p(\mathbb{Z}_1)\right| = o(h^{d/2}). \tag{5.5}$$

By Lemma A5, $\sup_{x\in\mathcal{S}_j} \mathbf{E}\left|V_{jn}(x)\right|^3 \leq Ch^{-d/2}$ for some $C > 0$. Using Lemma A2 and taking $r = \max\{p, 3\}$ and $V_{jn}^{(i)}(x) = W_i$, and $\Lambda_p(\cdot) = \varphi(\cdot)$, we deduce that

$$\begin{aligned}& \sup_{x\in\mathcal{S}_j} \left|\mathbf{E}\Lambda_p(\xi_{jn}(x)) - \mathbf{E}\Lambda_p(\mathbb{Z}_1)\right| \\ \leq\; & C_1 n^{-1/2} \sup_{x\in\mathcal{S}_j} \mathbf{E}\left|V_{jn}(x)\right|^3 + C_2 n^{-(r-2)/2} \sup_{x\in\mathcal{S}_j} \mathbf{E}\left|V_{jn}(x)\right|^r \\ & + C_3 \mathbf{E}\left[\omega_{\Lambda_p}\left(\mathbb{Z}_1; \frac{C_4}{\sqrt{n}}\mathbf{E}\left|V_{jn}(x)\right|^3\right)\right],\end{aligned} \tag{5.6}$$

for some constants $C_s > 0$, $s = 1, 2, 3$. The first two terms are $o(h^{d/2})$. As for the last expectation, observe that by Lemma A3, for all $x \in \mathcal{S}_j$,

$$\mathbf{E}\left[\omega_{\Lambda_p}\left(\mathbb{Z}_1; \frac{C_4}{\sqrt{n}}\mathbf{E}\left|V_{jn}(x)\right|^3\right)\right] \leq C\sum_{k=0}^{\lceil p-1\rceil}\left(\frac{C_4}{\sqrt{n}}\mathbf{E}\left|V_{jn}(x)\right|^3\right)^{p-\frac{(p-1)k}{\lceil p-1\rceil}}\mathbf{E}|\mathbb{Z}_1|^k = O(n^{-1/2}h^{-d/2}).$$

Since $O(n^{-1/2}h^{-d/2}) = o(h^{d/2})$, this completes the proof of (5.5).

We consider the second statement. Let $\bar{V}_{jn}^{(i)}(x),\ i = 1, \cdots, n$, be i.i.d. copies of

$$\frac{Y_{ji}K\left(\frac{x-X_i}{h}\right) - \mathbf{E}\left(Y_{ji}K\left(\frac{x-X_i}{h}\right)\right)}{\sqrt{\mathbf{E}\left[Y_{ji}^2K^2\left(\frac{x-X_i}{h}\right)\right] - \left(\mathbf{E}\left[Y_{ji}K\left(\frac{x-X_i}{h}\right)\right]\right)^2}}$$

so that $Var(\bar{V}_{jn}^{(i)}(x)) = 1$. Observe that for some constant $C > 0$,

$$\sup_{x\in\mathcal{S}_j}\mathbf{E}\left|\bar{V}_{jn}^{(i)}(x)\right|^3 \leq Ch^{-d/2}\sup_{x\in\mathcal{S}_j}\frac{k_{jn,3}(x)}{\left(\rho_{jn}^2(x) - h^d b_{jn}^2(x)\right)^{3/2}}, \tag{5.7}$$

where $b_{jn}(x) \equiv h^{-d}\mathbf{E}\left[Y_{ji}K\left((x - X_i)/h\right)\right]$. Again, the last supremum is bounded by $Ch^{-d/2}$ by Lemma A4 for some $C > 0$. Define

$$\bar{\xi}_{jn}(x) \equiv \frac{\sqrt{nh^d}v_{jn}(x)}{\tilde{\rho}_{jn}(x)},$$

where $\tilde{\rho}_{jn}^2(x) \equiv nh^dVar(v_{jn}(x))$. Then $\bar{\xi}_{jn}(x) \overset{d}{=} \frac{1}{\sqrt{n}}\sum_{i=1}^n \bar{V}_{jn}^{(i)}(x)$. Using Lemma A2 and following the arguments in (5.6) analogously, we deduce that

$$\sup_{x\in\mathcal{S}_j}\left|\mathbf{E}\Lambda_p\left(\bar{\xi}_{jn}(x)\right) - \mathbf{E}\Lambda_p(\mathbb{Z}_1)\right| = o(h^{d/2}).$$

This leads us to conclude that

$$\int_A \left\{n^{p/2}h^{(p-1)d/2}\mathbf{E}\Lambda_p(v_{jn}(x)) - h^{-d/2}\tilde{\rho}_{jn}^p(x)\mathbf{E}\Lambda_p(\mathbb{Z}_1)\right\} w_j(x)dx = o(1).$$

Now, observe that

$$\begin{aligned}
h^{-d/2}\left|\tilde{\rho}_{jn}^p(x) - \rho_{jn}^p(x)\right| &= h^{-d/2}\left|(\rho_{jn}^2(x) - h^d b_{jn}^2(x))^{p/2} - (\rho_{jn}^2(x))^{p/2}\right| \\
&\leq ph^{d/2}b_{jn}^2(x)\left(\rho_{jn}^2(x) + h^d b_{jn}^2(x)\right)^{p/2-1} \\
&= O(h^{d/2}) = o(1),\ \text{uniformly over } x \in \mathcal{S}_j,
\end{aligned}$$

where the second to the last equality follows by Lemma A4. This completes the proof. ■

Recall the definition: $\rho_j^2(x) \equiv \mathbf{E}[Y_{ji}^2|X_i = x]f(x)\int K^2(u)du$. Let

$$
\begin{aligned}
(5.8)\ \sigma_{jk,n}(A) &\equiv n^p h^{(p-1)d} \int_A \int_A Cov\left(\Lambda_p(v_{jN}(x)), \Lambda_p(v_{kN}(z))\right) w_j(x) w_k(z) dx dz, \text{ and} \\
\sigma_{jk}(A) &\equiv \int_A q_{jk,p}(x)\rho_j^p(x)\rho_k^p(x) w_j(x) w_k(x) dx,
\end{aligned}
$$

where we recall the definition:

$$q_{jk,p}(x) \equiv \int_{[-1,1]^d} Cov\left(\Lambda_p(\sqrt{1 - t_{jk}^2(x,u)}\mathbb{Z}_1 + t_{jk}(x,u)\mathbb{Z}_2), \Lambda_p(\mathbb{Z}_2)\right) du.$$

Now, let $(Z_{1n}(x), Z_{2n}(z)) \in \mathbf{R}^2$ be a jointly normal centered random vector whose covariance matrix is the same as that of $(\xi_{jn}(x), \xi_{kn}(z))$ for all $x, z \in \mathbf{R}^d$. We define

$$\tau_{jk,n}(A) \equiv \int_A \int_{[-1,1]^d} g_{jk,n}(x,u)\lambda_{jk,n}(x, x+uh) du dx,$$

where

$$
\begin{aligned}
\lambda_{jk,n}(x,z) &\equiv \rho_{jn}^p(x)\rho_{kn}^p(z) w_j(x) w_k(z) 1_A(x) 1_A(z), \text{ and} \\
g_{jk,n}(x,u) &\equiv Cov\left(\Lambda_p(Z_{1n}(x)), \Lambda_p(Z_{2n}(x+uh))\right).
\end{aligned}
$$

The following result generalizes Lemma 6.5 of GMZ from a univariate X to a multivariate X. The truncation arguments in their proof on pages 752 and 753 do not apply in the case of multivariate X. The proof of the following lemma employs a different approach for this part.

LEMMA A7: *Suppose that Assumptions 1 and 2 hold and let $h \to 0$ as $n \to \infty$ satisfying* $\limsup_{n\to\infty} n^{-r/2+1}h^{(1-r/2)d} < C$ *for any $r \in [2, 2p+2]$ for some $C > 0$.*
(i) *Suppose that $A \subset \mathcal{S}_j \cap \mathcal{S}_k$ is any Borel set. Then*

$$\sigma_{jk,n}(A) = \tau_{jk,n}(A) + o(1).$$

(ii) *Suppose further that A has a finite Lebesgue measure, $\rho_j(\cdot)\rho_k(\cdot)$ and $w_j(\cdot)w_k(\cdot)$ are continuous and bounded on A, and*

$$(5.9) \qquad \sup_{x\in A}\left|\rho_{l,n}(x) - \rho_l(x)\right| \to 0, \text{ as } n \to \infty, \text{ for } l \in \{j,k\}.$$

Then, as $n \to \infty$, $\tau_{jk,n}(A) = \sigma_{jk}(A) + o(1)$, and hence from (i),

$$\sigma_{jk,n}(A) \to \sigma_{jk}(A).$$

PROOF: (i) By change of variables, we write $\sigma_{jk,n}(A) = \tilde{\tau}_{jk,n}(A)$, where

$$\tilde{\tau}_{jk,n}(A) \equiv \int_A \int_{[-1,1]^d} Cov\left(\Lambda_p(\xi_{jn}(x)), \Lambda_p(\xi_{kn}(x+uh))\right) \lambda_{jk,n}(x, x+uh) du dx.$$

Fix $\varepsilon_1 \in (0,1]$ and let $c(\varepsilon_1) = (1+\varepsilon_1)^2 - 1$. Let η_1 and η_2 be two independent random variables that are independent of $(\{Y_{ji}, X_i : j \in \mathcal{J}\}_{i=1}^{\infty}, N)$, each having a two-point distribution that gives two points, $\{\sqrt{c(\varepsilon_1)}\}$ and $\{-\sqrt{c(\varepsilon_1)}\}$, the equal mass of $1/2$, so that $\mathbf{E}\eta_1 = \mathbf{E}\eta_2 = 0$ and $Var(\eta_1) = Var(\eta_2) = c(\varepsilon_1)$. Furthermore, observe that for any $r \geq 1$,

$$\mathbf{E}|\eta_1|^r = \frac{1}{2}|c(\varepsilon_1)|^{r/2} + \frac{1}{2}|c(\varepsilon_1)|^{r/2} \leq C\varepsilon_1^{r/2}, \tag{5.10}$$

for some constant $C > 0$ that depends only on r. Define

$$\xi_{jn,1}^{\eta}(x) \equiv \frac{\xi_{jn}(x) + \eta_1}{1+\varepsilon_1} \text{ and } \xi_{kn,2}^{\eta}(x+uh) \equiv \frac{\xi_{kn}(x+uh) + \eta_2}{1+\varepsilon_1}.$$

Note that $Var(\xi_{jn,1}^{\eta}(x)) = Var(\xi_{kn,2}^{\eta}(x+uh)) = 1$. Let $(Z_{1n}^{\eta}(x), Z_{2n}^{\eta}(x+uh))$ be a jointly normal centered random vector whose covariance matrix is the same as that of $(\xi_{jn,1}^{\eta}(x), \xi_{kn,2}^{\eta}(x+uh))$ for all $(x,u) \in \mathbf{R}^d \times [-1,1]^d$. Define

$$\begin{aligned}
\tilde{\tau}_{jk,n}^{\eta}(A) &\equiv \int_A \int_{[-1,1]^d} Cov\left(\Lambda_p(\xi_{jn,1}^{\eta}(x)), \Lambda_p(\xi_{kn,2}^{\eta}(x+uh))\right) \lambda_{jk,n}(x, x+uh) du dx, \\
\tau_{jk,n}^{\eta}(A) &\equiv \int_A \int_{[-1,1]^d} Cov\left(\Lambda_p(Z_{1n}^{\eta}(x)), \Lambda_p(Z_{2n}^{\eta}(x+uh))\right) \lambda_{jk,n}(x, x+uh) du dx.
\end{aligned}$$

Then first observe that

$$\begin{aligned}
\left|\tilde{\tau}_{jk,n}(A) - \tilde{\tau}_{jk,n}^{\eta}(A)\right| &\leq \int_A \int_{[-1,1]^d} |\Delta_{jk,n,1}^{\eta}(x,u)| \lambda_{jk,n}(x, x+uh) du dx \\
&\quad + \int_A \int_{[-1,1]^d} |\Delta_{jk,n,2}^{\eta}(x,u)| \lambda_{jk,n}(x, x+uh) du dx,
\end{aligned}$$

where

$$\begin{aligned}
\Delta_{jk,n,1}^{\eta}(x,u) &\equiv \mathbf{E}\Lambda_p(\xi_{jn}(x))\mathbf{E}\Lambda_p(\xi_{kn}(x+uh)) \\
&\quad - \mathbf{E}\Lambda_p(\xi_{jn,1}^{\eta}(x))\mathbf{E}\Lambda_p(\xi_{kn,2}^{\eta}(x+uh)) \text{ and} \\
\Delta_{jk,n,2}^{\eta}(x,u) &\equiv \mathbf{E}\Lambda_p(\xi_{jn}(x))\Lambda_p(\xi_{kn}(x+uh)) \\
&\quad - \mathbf{E}\Lambda_p(\xi_{jn,1}^{\eta}(x))\Lambda_p(\xi_{kn,2}^{\eta}(x+uh)).
\end{aligned}$$

For any $a, b \in \mathbf{R}$ and $a_+ = \max(a,0)$ and $b_+ = \max(b,0)$,

$$\begin{aligned}
|a_+^p - b_+^p| &\leq p|a_+ - b_+|\left(|a|^{p-1} + |b|^{p-1}\right) \\
&\leq p|a - b|\left(|a|^{p-1} + |b|^{p-1}\right).
\end{aligned} \tag{5.11}$$

We bound $|\Delta^{\eta}_{jk,n,2}(x,u)|$ by

$$\begin{aligned} & p\mathbf{E}\left[|\xi_{jn}(x)-\xi^{\eta}_{jn,1}(x)|\left(|\xi_{jn}(x)|^{p-1}+|\xi^{\eta}_{jn,1}(x)|^{p-1}\right)|\xi_{kn}(x+uh)|^{p}\right] \\ & +p\mathbf{E}\left[|\xi_{kn}(x+uh)-\xi^{\eta}_{kn,2}(x+uh)|\left(|\xi_{kn}(x)|^{p-1}+|\xi^{\eta}_{kn,2}(x)|^{p-1}\right)|\xi^{\eta}_{jn,1}(x)|^{p}\right] \\ \equiv \; & A_{1n}(x,u)+A_{2n}(x,u), \text{ say.} \end{aligned}$$

As for $A_{1n}(x,u)$,

$$\begin{aligned} A_{1n}(x,u) \leq \; & p\left(\mathbf{E}\left[|\xi_{jn}(x)-\xi^{\eta}_{jn,1}(x)|^{2}(|\xi_{jn}(x)|^{p-1}+|\xi^{\eta}_{jn,1}(x)|^{p-1})^{2}\right]\right)^{1/2} \\ & \times\left(\mathbf{E}\left[|\xi_{kn}(x+uh)|^{2p}\right]\right)^{1/2}. \end{aligned}$$

Define

$$s \equiv \begin{cases} (p+1)/(p-1) \text{ if } p>1 \\ 2 \text{ if } p=1, \end{cases}$$

and $q \equiv (1-1/s)^{-1}$. Note that

$$\begin{aligned} & \mathbf{E}\left[|\xi_{jn}(x)-\xi^{\eta}_{jn,1}(x)|^{2}(|\xi_{jn}(x)|^{p-1}+|\xi^{\eta}_{jn,1}(x)|^{p-1})^{2}\right] \\ \leq \; & \left(\mathbf{E}\left[|\xi_{jn}(x)-\xi^{\eta}_{jn,1}(x)|^{2q}\right]\right)^{1/q}\left(\mathbf{E}\left[(|\xi_{jn}(x)|^{p-1}+|\xi^{\eta}_{jn,1}(x)|^{p-1})^{2s}\right]\right)^{1/s}. \end{aligned}$$

Now,

$$\begin{aligned} & \mathbf{E}\left[|\xi_{jn}(x)-\xi^{\eta}_{jn,1}(x)|^{2q}\right]=(1+\varepsilon_1)^{-2q}\mathbf{E}\left[|\varepsilon_1\xi_{jn}(x)-\eta_1|^{2q}\right] \\ \leq \; & 2^{2q-1}(1+\varepsilon_1)^{-2q}\left\{\varepsilon_1^{2q}\mathbf{E}\left[|\xi_{jn}(x)|^{2q}\right]+\mathbf{E}\left[|\eta_1|^{2q}\right]\right\}. \end{aligned}$$

Applying Lemma A5 and (5.10) to the last bound, we conclude that

$$\sup_{x\in\mathcal{S}_j}\mathbf{E}\left[|\xi_{jn}(x)-\xi^{\eta}_{jn,1}(x)|^{2q}\right] \leq \frac{C_1(\varepsilon_1^{2q}+\varepsilon_1^{q})}{(1+\varepsilon_1)^{2q}} \leq C_2\varepsilon_1^{q},$$

for some constants $C_1, C_2 > 0$. Using Lemma A5, we can also see that for some constants $C_3, C_4 > 0$,

$$\sup_{x\in\mathcal{S}_j}\mathbf{E}\left[(|\xi_{jn}(x)|^{p-1}+|\xi^{\eta}_{jn,1}(x)|^{p-1})^{2s}\right] \leq C_3$$

and from some large n on,

$$\sup_{u\in[-1,1]^d}\sup_{x\in\mathcal{S}_k}\mathbf{E}\left[|\xi_{kn}(x+uh)|^{2p}\right] \leq \sup_{x\in\mathcal{S}_k^{\varepsilon/2}}\mathbf{E}\left[|\xi_{kn}(x)|^{2p}\right] \leq C_4,$$

for $\varepsilon > 0$ in Assumption 1(iii). Therefore, for some constant $C > 0$,

$$\sup_{u\in[-1,1]^d}\sup_{x\in\mathcal{S}_j\cap\mathcal{S}_k} A_{1n}(x,u) \leq C\sqrt{\varepsilon_1}.$$

Using similar arguments for $A_{2n}(x,u)$, we deduce that for some constant $C>0$,

$$\sup_{u\in[-1,1]^d}\sup_{x\in\mathcal{S}_j\cap\mathcal{S}_k}|\Delta^\eta_{jk,n,2}(x,u)|\leq C\sqrt{\varepsilon_1}. \tag{5.12}$$

Let us turn to $\Delta^\eta_{jk,n,1}(x,u)$. We bound $|\Delta^\eta_{jk,n,1}(x,u)|$ by

$$\begin{aligned}&p\mathbf{E}\left[|\xi_{jn}(x)-\xi^\eta_{jn,1}(x)|\left(|\xi_{jn}(x)|^{p-1}+|\xi^\eta_{jn,1}(x)|^{p-1}\right)\right]\mathbf{E}\left[|\xi_{kn}(x+uh)|^p\right]\\&+p\mathbf{E}\left[|\xi_{kn}(x+uh)-\xi^\eta_{kn,2}(x+uh)|\left(|\xi_{jn}(x)|^{p-1}+|\xi^\eta_{jn,1}(x)|^{p-1}\right)\right]\mathbf{E}[|\xi^\eta_{jn,1}(x)|^p].\end{aligned}$$

Using similar arguments for $\Delta^\eta_{jk,n,2}(x,u)$, we find that for some constant $C>0$,

$$\sup_{u\in[-1,1]^d}\sup_{x\in\mathcal{S}_j\cap\mathcal{S}_k}|\Delta^\eta_{jk,n,1}(x,u)|\leq C\sqrt{\varepsilon_1}. \tag{5.13}$$

By Lemma A4 and Assumption 1(ii), there exist $n_0>0$ and $C_1,C_2>0$ such that for all $n\geq n_0$,

$$\begin{aligned}&\int_A\int_{[-1,1]^d}\lambda_{jk,n}(x,x+uh)dudx\\ \leq\ & C_1\int_A\int_{[-1,1]^d}w_j(x)w_k(x+uh)dudx\\ \leq\ & C_2\sqrt{\int_A w_j^2(x)dx}\sqrt{\int_A\int_{[-1,1]^d}w_k^2(x+uh)dudx}<\infty.\end{aligned} \tag{5.14}$$

Hence

$$\left|\tilde{\tau}_{jk,n}(A)-\tilde{\tau}^\eta_{jk,n}(A)\right|\leq C_5\sqrt{\varepsilon_1}\int_A\int_{[-1,1]^d}\lambda_{jk,n}(x,x+uh)dudx\leq C_6\sqrt{\varepsilon_1},$$

for some constants $C_5>0$ and $C_6>0$.

Since the choice of $\varepsilon_1>0$ was arbitrary, it remains for the proof of Lemma A7(i) to prove that

$$\left|\tilde{\tau}^\eta_{jk,n}(A)-\tau_{jk,n}(A)\right|=o(1), \tag{5.15}$$

as $n\to\infty$ and then $\varepsilon_1\to 0$. For any $x\in\mathcal{S}_j\cap\mathcal{S}_k$,

$$\left(\xi^\eta_{jn,1}(x),\xi^\eta_{kn,2}(x+uh)\right)'\overset{d}{=}\frac{1}{\sqrt{n}}\sum_{i=1}^n W_n^{(i)}(x,u),$$

where $W_n^{(i)}(x,u)$'s are i.i.d. copies of $W_n(x,u)\equiv(q_{jn}(x),q_{kn}(x+uh))'$ with

$$q_{jn}(x)\equiv\frac{\sum_{i\leq N_1}Y_{ji}K\left(\frac{x-X_i}{h}\right)-\mathbf{E}\left[Y_{ji}K\left(\frac{x-X_i}{h}\right)\right]+\eta_1}{h^{d/2}\rho_{jn}(x)(1+\varepsilon_1)}.$$

Using the same arguments as in the proof of Lemma A5, we find that for $j \in \mathcal{J}$,

$$\sup_{x \in \mathcal{S}_j^{\varepsilon}} \mathbf{E}\left[|q_{jn}(x)|^3\right] \leq Ch^{-d/2}, \text{ for some } C > 0. \tag{5.16}$$

Let Σ_{1n} be the 2×2 covariance matrix of $(\xi^{\eta}_{jn,1}(x), \xi^{\eta}_{kn,2}(x+uh))'$. Define

$$\tilde{\Lambda}_{n,p}(v) \equiv \Lambda_p([\Sigma_{1n}^{1/2} v]_1)\Lambda_p([\Sigma_{1n}^{1/2} v]_2),\ v \in \mathbf{R}^2,$$

where $[a]_j$ of a vector $a \in \mathbf{R}^2$ indicates its j-th entry. There exists some $C > 0$ such that for all n,

$$\begin{aligned} \sup_{v \in \mathbf{R}^2} \frac{\left|\tilde{\Lambda}_{n,p}(v) - \tilde{\Lambda}_{n,p}(0)\right|}{1 + ||v||^{2p+2} \min\{||v||, 1\}} &\leq C \text{ and} \\ \int \sup_{u \in \mathbf{R}^2 : ||z-u|| \leq \delta} \left|\tilde{\Lambda}_{n,p}(z) - \tilde{\Lambda}_{n,p}(u)\right| d\Phi(z) &\leq C\delta \text{ for all } \delta \in (0,1]. \end{aligned} \tag{5.17}$$

The correlation between $\xi^{\eta}_{jn,1}(x)$ and $\xi^{\eta}_{kn,2}(x+uh)$ is equal to

$$\mathbf{E}\left[\xi^{\eta}_{jn,1}(x)\xi^{\eta}_{kn,2}(x+uh)\right] = \frac{\mathbf{E}[\xi_{jn}(x)\xi_{kn}(x+uh)]}{(1+\varepsilon_1)^2} \in \left[-(1+\varepsilon_1)^{-2}, (1+\varepsilon_1)^{-2}\right].$$

Hence, as for $\tilde{W}_n^{(i)}(x,u) \equiv \Sigma_{1n}^{-1/2} W_n^{(i)}(x,u)$, by (5.16),

$$\begin{aligned} &\sup_{x \in \mathcal{S}_j \cap \mathcal{S}_k} \mathbf{E}||\tilde{W}_n^{(i)}(x,u)||^3 \\ &\leq C_1(1 - (\mathbf{E}[\xi^{\eta}_{jn,1}(x)\xi^{\eta}_{kn,2}(x+uh)])^2)^{-3/2}\left\{\sup\nolimits_{x \in \mathcal{S}_j^{\varepsilon}} \mathbf{E}[|q_{jn}(x)|^3] + \sup\nolimits_{x \in \mathcal{S}_k^{\varepsilon}} \mathbf{E}[|q_{kn}(x)|^3]\right\} \\ &\leq C_1(1 - (1+\varepsilon_1)^{-4})^{-3/2}\left\{\sup\nolimits_{x \in \mathcal{S}_j^{\varepsilon}} \mathbf{E}[|q_{jn}(x)|^3] + \sup\nolimits_{x \in \mathcal{S}_k^{\varepsilon}} \mathbf{E}[|q_{kn}(x)|^3]\right\} \\ &\leq C_2(1 - (1+\varepsilon_1)^{-4})^{-3/2} h^{-d/2}, \text{ for some } C_1, C_2 > 0, \end{aligned} \tag{5.18}$$

so that $n^{-1/2} \sup_{x \in \mathcal{S}_j \cap \mathcal{S}_k} \mathbf{E}||\tilde{W}_n^{(i)}(x,u)||^3 = O(n^{-1/2}h^{-d/2})$. By Lemma A2 and following the arguments in (5.6) analogously,

$$\begin{aligned} &\sup_{x \in \mathcal{S}_j \cap \mathcal{S}_k} \left|\mathbf{E}\tilde{\Lambda}_{n,p}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n \tilde{W}_n^{(i)}(x,u)\right) - \mathbf{E}\tilde{\Lambda}_{n,p}\left(\tilde{Z}_n^{\eta}(x,u)\right)\right| \\ &= O\left(n^{-1/2}h^{-d/2}\right) = o(1), \end{aligned}$$

where $\tilde{Z}_n^{\eta}(x,u) \equiv \Sigma_{1n}^{-1/2}(Z_{1n}^{\eta}(x), Z_{2n}^{\eta}(x+uh))'$. Certainly by (5.10) and Lemma A5,

$$\begin{aligned} &Cov(\Lambda_p(Z_{1n}^{\eta}(x)), \Lambda_p(Z_{2n}^{\eta}(x+uh))) \\ &\leq \sqrt{\mathbf{E}|Z_{1n}^{\eta}(x)|^{2p}}\sqrt{\mathbf{E}|Z_{2n}^{\eta}(x+uh)|^{2p}} < C, \end{aligned}$$

for some $C > 0$ that does not depend on ε_1. Using (5.14), we apply the dominated convergence theorem to obtain that

$$\left|\tau^{\eta}_{jk,n}(A) - \tilde{\tau}^{\eta}_{jk,n}(A)\right| = o(1) \tag{5.19}$$

as $n \to \infty$ for each $\varepsilon_1 > 0$.

Finally, note from (5.12) and (5.13) that, for all $x \in A$ and all $u \in [-1,1]^d$,

$$\begin{aligned} & Cov(\Lambda_p(Z^{\eta}_{1n}(x)), \Lambda_p(Z^{\eta}_{2n}(x+uh))) \\ = \; & Cov(\Lambda_p(Z_{1n}(x)), \Lambda_p(Z_{2n}(x+uh))) + o(1), \end{aligned}$$

where the $o(1)$ term is one that converges to zero as $n \to \infty$ and then $\varepsilon_1 \to 0$. Therefore, by the dominated convergence theorem,

$$\left|\tau^{\eta}_{jk,n}(A) - \tau_{jk,n}(A)\right| = o(1),$$

as $n \to \infty$ and then $\varepsilon_1 \to 0$. In view of (5.19), this completes the proof of (5.15) and, as a consequence, that of (i).

(ii) Define $t_{jk,n}(x,u) \equiv \mathbf{E}(\xi_{jn}(x)\xi_{kn}(x+uh))$,

$$\begin{aligned} e_{jk,n}(x,u) &\equiv \frac{1}{h^d}\mathbf{E}\left[Y_{ji}Y_{ki}K\left(\frac{x-X_i}{h}\right)K\left(\frac{x-X_i}{h}+u\right)\right] \text{ and} \\ e_{jk}(x,u) &\equiv \rho_{jk}(x)\frac{\int K(z)K(z+u)dz}{\int K^2(u)du}. \end{aligned}$$

By Assumption 1(i), and Lemma A4, for almost every $x \in A$ and for each $u \in [-1,1]^d$,

$$\begin{aligned} t_{jk,n}(x,u) &= \frac{1}{\rho_{jn}(x)\rho_{kn}(x+uh)}\frac{1}{h^d}\mathbf{E}\left[Y_{ji}Y_{ki}K\left(\frac{x-X_i}{h}\right)K\left(\frac{x-X_i}{h}+u\right)\right] \\ &= \frac{e_{jk,n}(x,u)}{\rho_{jn}(x)\rho_{kn}(x+uh)} = \frac{e_{jk}(x,u)}{\rho_j(x)\rho_k(x+uh)} + o(1) = t_{jk}(x,u) + o(1), \end{aligned} \tag{5.20}$$

where we recall that $t_{jk}(x,u) = e_{jk}(x,u)/(\rho_j(x)\rho_k(x))$ by the definition of $t_{jk}(\cdot,\cdot)$.

By (5.9),

$$\tau_{jk,n}(A) = \int_A \int_{[-1,1]^d} g_{jk,n}(x,u)\lambda_{jk}(x,x+uh)dudx + o(1),$$

where $\lambda_{jk}(x,z) \equiv \rho^p_j(x)\rho^p_k(z)w_j(x)w_k(z)1_A(x)1_A(z)$. By (5.20), for almost every $x \in A$ and for each $u \in [-1,1]^d$,

$$g_{jk,n}(x,u) \to g_{jk}(x,u), \text{ as } n \to \infty,$$

where $g_{jk}(x,u) \equiv Cov(\Lambda_p(\sqrt{1-t^2_{jk}(x,u)}\mathbb{Z}_1 + t_{jk}(x,u)\mathbb{Z}_2), \Lambda_p(\mathbb{Z}_2))$. Furthermore, since $\rho_j(\cdot)\rho_k(\cdot)$ and $w_j(\cdot)w_k(\cdot)$ are continuous on A and A has a finite Lebesgue measure, we follow the proof of Lemma 6.4 of GMZ to find that $g_{jk,n}(x,u)\lambda_{jk}(x,x+uh)$ converges in measure to

$g_{jk}(x,u)\lambda_{jk}(x,x)$ on $A \times [-1,1]^d$, as $n \to \infty$. Using the bounded convergence theorem, we deduce the desired result. ■

The following lemma is a generalization of Lemma 6.2 of GMZ from $p = 1$ to $p \geq 1$. The proof of GMZ does not carry over to this general case because the majorization inequality of Pinelis (1994) used in GMZ does not apply here. (Note that (4) in Pinelis (1994) does not apply when $p > 1$.)

LEMMA A8: *Suppose that Assumptions 1 and 2 hold. Furthermore, assume that as $n \to \infty$, $h \to 0$, $n^{-1/2}h^{-d} \to 0$. Then there exists a constant $C > 0$ such that for any Borel set $A \subset \mathbf{R}^d$ and for all $j \in \mathcal{J}$,*

$$\begin{aligned}
&\text{limsup}_{n\to\infty}\mathbf{E}\left[\left|n^{p/2}h^{(p-1)d/2}\int_A \{\Lambda_p(v_{jn}(x)) - \mathbf{E}[\Lambda_p(v_{jn}(x))]\}\, w_j(x)dx\right|\right] \\
\leq\ & C\int_A w_j(x)dx + C\sqrt{\int_A w_j^2(x)dx}.
\end{aligned}$$

PROOF : It suffices to show that there exists $C > 0$ such that for any Borel set $A \subset \mathbf{R}^d$,

STEP 1: $\mathbf{E}\left[\left|n^{p/2}h^{(p-1)d/2}\int_A \left(\Lambda_p(v_{jn}(x)) - \Lambda_p(v_{jN}(x))\right) w_j(x)dx\right|\right] \leq C\int_A w_j(x)dx$,

STEP 2: $\mathbf{E}\left[\left|n^{p/2}h^{(p-1)d/2}\int_A \left(\Lambda_p(v_{jN}(x)) - \mathbf{E}[\Lambda_p(v_{jN}(x))]\right) w_j(x)dx\right|\right] \leq C\sqrt{\int_A w_j^2(x)dx}$, and

STEP 3: $n^{p/2}h^{(p-1)d/2}\left|\int_A \left(\mathbf{E}\Lambda_p(v_{jN}(x)) - \mathbf{E}[\Lambda_p(v_{jn}(x))]\right) w_j(x)dx\right| \to 0$ as $n \to \infty$.

Indeed, by chaining Steps 1, 2 and 3, we obtain the desired result.

PROOF OF STEP 1: For simplicity, let

$$\begin{aligned}
u_{jn}^2(x) &\equiv \mathbf{E}\left[Y_{ji}^2K^2\left(\frac{x-X_i}{h}\right)\right] - \left(\mathbf{E}\left[Y_{ji}K\left(\frac{x-X_i}{h}\right)\right]\right)^2 \text{ and} \\
\bar{V}_{n,ji}(x) &\equiv \frac{1}{u_{jn}(x)}\left\{Y_{ji}K\left(\frac{x-X_i}{h}\right) - \mathbf{E}\left[Y_{ji}K\left(\frac{x-X_i}{h}\right)\right]\right\}.
\end{aligned}$$

We write, if $N = n$, $\sum_{i=N+1}^n = 0$, and if $N > n$, $\sum_{i=N+1}^n = -\sum_{i=n+1}^N$. Using this notation, write

$$v_{jn}(x) = \frac{1}{nh^d}\sum_{i=1}^N \bar{V}_{n,ji}(x)u_{jn}(x) + \frac{1}{nh^d}\sum_{i=N+1}^n \bar{V}_{n,ji}(x)u_{jn}(x).$$

Now, observe that

$$\begin{aligned}
\frac{1}{\sqrt{nh^d}}\sum_{i=1}^{N}\bar{V}_{n,ji}(x)u_{jn}(x) &= \frac{1}{\sqrt{nh^d}}\sum_{i=1}^{N}\left\{Y_{ji}K\left(\frac{x-X_i}{h}\right)-\mathbf{E}\left[Y_{ji}K\left(\frac{x-X_i}{h}\right)\right]\right\} \\
&= \sqrt{nh^d}\left\{\hat{g}_{jN}(x)-\frac{1}{h^d}\mathbf{E}\left[Y_{ji}K\left(\frac{x-X_i}{h}\right)\right]\right\} \\
&\quad +\sqrt{nh^d}\left(\frac{n-N}{n}\right)\cdot\frac{1}{h^d}\cdot\mathbf{E}\left[Y_{ji}K\left(\frac{x-X_i}{h}\right)\right] \\
&= \sqrt{nh^d}v_{jN}(x)+\sqrt{nh^d}\left(\frac{n-N}{n}\right)\cdot\frac{1}{h^d}\cdot\mathbf{E}\left[Y_{ji}K\left(\frac{x-X_i}{h}\right)\right].
\end{aligned}$$

Letting

$$\begin{aligned}
\eta_{jn}(x) &\equiv \sqrt{n}\left(\frac{n-N}{n}\right)\cdot\frac{1}{h^d}\cdot\mathbf{E}\left[Y_{ji}K\left(\frac{x-X_i}{h}\right)\right] \text{ and} \\
s_{jn}(x) &\equiv \frac{1}{\sqrt{nh^d}}\sum_{i=N+1}^{n}\bar{V}_{n,ji}(x)u_{jn}(x),
\end{aligned}$$

we can write

$$\sqrt{nh^d}v_{jn}(x)=\sqrt{nh^d}v_{jN}(x)+(\sqrt{h^d}\eta_{jn}(x)+s_{jn}(x)).$$

First, note that for some constant $C>0$,

$$\sup_{x\in\mathcal{S}_j}u_{jn}^2(x)\leq Ch^d, \tag{5.21}$$

from some large n on, by Lemma A4. Recall the definition of $\tilde{\rho}_{jn}(x):\tilde{\rho}_{jn}(x)\equiv\sqrt{nh^dVar(v_{jn}(x))}$. As in the proof of Lemma A5, there exist n_0, $C_1>0$ and $C_2>0$ such that for all $n\geq n_0$,

$$\begin{aligned}
C_1 &> \sup_{x\in\mathcal{S}_j}\sqrt{\rho_{jn}^2(x)-h^db_{jn}^2(x)}=\sup_{x\in\mathcal{S}_j}\tilde{\rho}_{jn}(x) \\
&\geq \inf_{x\in\mathcal{S}_j}\tilde{\rho}_{jn}(x)\geq\inf_{x\in\mathcal{S}_j}\sqrt{\rho_{jn}^2(x)-h^db_{jn}^2(x)}>C_2.
\end{aligned} \tag{5.22}$$

Using (5.11), (5.21) and (5.22), we deduce that for some $C > 0$,

$$\begin{aligned}
&\left|n^{p/2}h^{(p-1)d/2}\int_A \left(\Lambda_p(v_{jn}(x)) - \Lambda_p(v_{jN}(x))\right) w_j(x)dx\right| \\
&= h^{-d/2}\left|\int_A \left(\Lambda_p\left(\sqrt{nh^d}v_{jn}(x)\right) - \Lambda_p\left(\sqrt{nh^d}v_{jN}(x)\right)\right) w_j(x)dx\right| \\
&\le C\int_A |\eta_{jn}(x)|\left(\left|\sqrt{nh^d}\frac{v_{jn}(x)}{\tilde{\rho}_{jn}(x)}\right|^{p-1} + \left|\sqrt{nh^d}\frac{v_{jN}(x)}{\tilde{\rho}_{jn}(x)}\right|^{p-1}\right) w_j(x)dx \\
&\quad + C\int_A \left|\frac{s_{jn}(x)}{u_{jn}(x)}\right|\left(\left|\sqrt{nh^d}\frac{v_{jn}(x)}{\tilde{\rho}_{jn}(x)}\right|^{p-1} + \left|\sqrt{nh^d}\frac{v_{jN}(x)}{\tilde{\rho}_{jn}(x)}\right|^{p-1}\right) w_j(x)dx \\
&= A_{1n} + A_{2n}, \text{ say.}
\end{aligned}$$

To deal with A_{1n} and A_{2n}, we first show the following:

Claim 1: $\sup_{x\in\mathcal{S}_j}\mathbf{E}[\eta_{jn}^2(x)] = O(1)$.

Claim 2: $\sup_{x\in\mathcal{S}_j}\mathbf{E}[|s_{jn}(x)/u_{jn}(x)|^2] = o(1)$.

Claim 3: $\sup_{x\in\mathcal{S}_j}\mathbf{E}[|\sqrt{nh^d}v_{jN}(x)/\tilde{\rho}_{jn}(x)|^{2p-2}] = O(1)$.

Proof of Claim 1: By Lemma A4 and the fact that $\mathbf{E}|n^{-1/2}(n-N)|^2 = O(1)$,

$$\sup_{x\in\mathcal{S}_j}\mathbf{E}\left[\eta_{jn}^2(x)\right] \le \mathbf{E}\left|\sqrt{n}\left(\frac{n-N}{n}\right)\right|^2 \cdot \sup_{x\in\mathcal{S}_j}\left|\frac{1}{h^d}\cdot\mathbf{E}\left[Y_{ji}K\left(\frac{x-X_i}{h}\right)\right]\right|^2 = O(1).$$

Proof of Claim 2: Note that

$$\left|\sqrt{nh^d}\frac{s_{jn}(x)}{u_{jn}(x)}\right| = \left|\sum_{i=n+1}^{N}\bar{V}_{n,ji}(x)\right|. \tag{5.23}$$

Certainly $Var(\bar{V}_{n,ji}(x)) = 1$. Hence by Lemma 1(i) of Horváth (1991), for some $C > 0$,

$$\begin{aligned}
\mathbf{E}\left(\sqrt{nh^d}\frac{s_{jn}(x)}{u_{jn}(x)}\right)^2 &\le \mathbf{E}|N-n|\mathbf{E}|\mathbb{Z}_1|^2 \\
&\quad + C\left\{\mathbf{E}\left|N-n\right|^{1/2}\mathbf{E}\left|\bar{V}_{n,ji}(x)\right|^3 + \mathbf{E}\left|\bar{V}_{n,ji}(x)\right|^4\right\}.
\end{aligned}$$

As seen in (5.7), $\sup_{x\in\mathcal{S}_j}\mathbf{E}\left|\bar{V}_{n,ji}(x)\right|^3 \le Ch^{-d/2}$ for some $C > 0$. Similarly,

$$\sup_{x\in\mathcal{S}_j}\mathbf{E}\left|\bar{V}_{n,ji}(x)\right|^4 \le \frac{h^d k_{jn,4}(x)}{h^{2d}\left(\rho_{jn}^2(x) - h^d b_{jn}^2(x)\right)^2} \le Ch^{-d},$$

for some $C > 0$. Furthermore, $\mathbf{E}|N-n| = O(n^{1/2})$ and $\mathbf{E}\,|N-n|^{1/2} = O(n^{1/4})$ (e.g. (2.21) and (2.22) of Horváth (1991)). Therefore, there exists $C > 0$ such that

$$\sup_{x\in\mathcal{S}_j} \mathbf{E}\left(\sqrt{nh^d}\frac{s_{jn}(x)}{u_{jn}(x)}\right)^2 \leq C\left\{n^{1/2} + n^{1/4}h^{-d/2} + h^{-d}\right\}.$$

Since $n^{-1/2}h^{-d} \to 0$, $\sup_{x\in\mathcal{S}_j}\mathbf{E}[(s_{jn}(x)/u_{jn}(x))^2] = o(1)$.

Proof of Claim 3: By (5.4), Lemmas A3-A4, and (5.22), we have

$$\sup_{x\in\mathcal{S}_j}\mathbf{E}\left[\left(\sqrt{nh^d}\frac{v_{jN}(x)}{\tilde{\rho}_{jn}(x)}\right)^{2p-2}\right] = \sup_{x\in\mathcal{S}_j}\left|\frac{\rho_{jn}(x)}{\tilde{\rho}_{jn}(x)}\right|^{2p-2}\mathbf{E}\left(\left|\frac{\sqrt{nh^d}v_{jN}(x)}{\rho_{jn}(x)}\right|^{2p-2}\right) \leq C,$$

for some $C > 0$. This completes the proof of Claim 3.

Now, using Claims 1-3, we prove Step 1. Let $\mu_j(A) \equiv \int_A w_j(x)dx$. Since $h^{(p-1)d/2} = O(1)$ when $p = 1$, and $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for any $a \geq 0$ and $b \geq 0$,

$$\begin{aligned}
\mathbf{E}[A_{1n}] &\leq C\int_A \mathbf{E}\left[|\eta_{jn}(x)|\left(\left|\sqrt{nh^d}\frac{v_{jn}(x)}{\tilde{\rho}_{jn}(x)}\right|^{p-1} + \left|\sqrt{nh^d}\frac{v_{jN}(x)}{\tilde{\rho}_{jn}(x)}\right|^{p-1}\right)\right]w_j(x)dx \\
&\leq C\mu_j(A)\sup_{x\in\mathcal{S}_j}\mathbf{E}\left[|\eta_{jn}(x)|\left(\left|\sqrt{nh^d}\frac{v_{jn}(x)}{\tilde{\rho}_{jn}(x)}\right|^{p-1} + \left|\sqrt{nh^d}\frac{v_{jN}(x)}{\tilde{\rho}_{jn}(x)}\right|^{p-1}\right)\right] \\
&\leq C\mu_j(A) \times \left(\sup_{x\in\mathcal{S}_j}\mathbf{E}\left[\eta_{jn}^2(x)\right]\right)^{1/2} \\
&\quad \times \left(\left(\sup_{x\in\mathcal{S}_j}\mathbf{E}\left[\left|\sqrt{nh^d}\frac{v_{jn}(x)}{\tilde{\rho}_{jn}(x)}\right|^{2p-2}\right]\right)^{1/2} + \left(\sup_{x\in\mathcal{S}_j}\mathbf{E}\left[\left|\sqrt{nh^d}\frac{v_{jN}(x)}{\tilde{\rho}_{jn}(x)}\right|^{2p-2}\right]\right)^{1/2}\right).
\end{aligned}$$

Certainly, as in the proof of Lemma A5,

$$\sup_{x\in\mathcal{S}_j}\mathbf{E}\left[\left|\sqrt{nh^d}\frac{v_{jn}(x)}{\tilde{\rho}_{jn}(x)}\right|^{2p-2}\right] \leq C, \tag{5.24}$$

for some constant $C > 0$. Hence using Claims 1 and 3, we conclude that $\mathbf{E}[A_{1n}] \leq C\mu_j(A)$. As for A_{2n}, similarly, we obtain that

$$\begin{aligned}
\mathbf{E}[A_{2n}] &\leq C\int_A \mathbf{E}\left[\left|\frac{s_{jn}(x)}{u_{jn}(x)}\right|\left(\left|\sqrt{nh^d}\frac{v_{jn}(x)}{\tilde{\rho}_{jn}(x)}\right|^{p-1} + \left|\sqrt{nh^d}\frac{v_{jN}(x)}{\tilde{\rho}_{jn}(x)}\right|^{p-1}\right)\right] w_j(x)dx \\
&\leq C\mu_j(A) \times \left(\sup_{x\in\mathcal{S}_j}\mathbf{E}\left[\left|\frac{s_{jn}(x)}{u_{jn}(x)}\right|^2\right]\right)^{1/2} \\
&\quad \times \left(\left(\sup_{x\in\mathcal{S}_j}\mathbf{E}\left[\left|\sqrt{nh^d}\frac{v_{jn}(x)}{\tilde{\rho}_{jn}(x)}\right|^{2p-2}\right]\right)^{1/2} + \left(\sup_{x\in\mathcal{S}_j}\mathbf{E}\left[\left|\sqrt{nh^d}\frac{v_{jN}(x)}{\tilde{\rho}_{jn}(x)}\right|^{2p-2}\right]\right)^{1/2}\right).
\end{aligned}$$

Therefore, by Claims 2 and 3 and (5.24), $\mathbf{E}[A_{2n}] = o(1)$.

Proof of Step 2: We can follow the proof of Lemma A7(i) to show that

$$\mathbf{E}\left[n^{p/2}h^{(p-1)d/2}\int_A \left(|v_{jN}(x)|^p - \mathbf{E}\left[|v_{jN}(x)|^p\right]\right) w_j(x)dx\right]^2 = \kappa_{jn}(A) + o(1),$$

where $\kappa_{jn}(A) \equiv \int_A \int_{[-1,1]^d} r_{jn}(x,u)\lambda_{jn}(x, x+uh)dudx$,

$$\begin{aligned}
\lambda_{jn}(x,z) &\equiv \rho_{jn}^p(x)\rho_{jn}^p(z)w_j(x)w_j(z)1_{A\cap\mathcal{S}_j}(x)1_{A\cap\mathcal{S}_j}(z) \text{ and} \\
r_{jn}(x,u) &\equiv Cov\left(|Z_{jn,A}(x)|^p, |Z_{jn,B}(x+uh)|^p\right),
\end{aligned}$$

with $(Z_{jn,A}(x), Z_{jn,B}(x+uh))' \in \mathbf{R}^2$ denoting a centered normal random vector whose covariance matrix is equal to that of $(\xi_{jn}(x), \xi_{jn}(x+uh))'$. By Cauchy-Schwarz inequality and Lemma A5,

$$\sup_{x\in\mathcal{S}_j} r_{jn}(x,u) \leq \sup_{x\in\mathcal{S}_j}\sqrt{\mathbf{E}\left|Z_{jn,A}(x)\right|^{2p}\mathbf{E}\left|Z_{jn,B}(x+uh)\right|^{2p}} < \infty.$$

Furthermore, for each $u \in [-1,1]^d$,

$$\int_A \lambda_{jn}(x, x+uh)dx \leq \sqrt{\int_A w_j^2(x)dx}\sqrt{\int_{A+uh} w_j^2(x)dx}.$$

Since $\int_{\mathcal{S}_j^\varepsilon} w_j^2(x)dx < \infty$ for some $\varepsilon > 0$ (Assumption 1(ii)), we find that as $h \to 0$, the last term converges to $\int_A w_j^2(x)dx$. We obtain the desired result of Step 2.

Proof of Step 3: The convergence above follows from the proof of Lemma A6. ■

Let $\mathcal{C} \subset \mathbf{R}^d$ be a bounded Borel set such that

$$\alpha \equiv P\left\{X \in \mathbf{R}^d \backslash \mathcal{C}\right\} > 0.$$

For any Borel set $A \subset \mathcal{C}$ and $t \in \mathbf{R}^J$, let

$$\begin{aligned}\zeta_{t,n}(A) &\equiv \sum_{j=1}^{J} t_j \int_A \Lambda_p(v_{jn}(x)) w_j(x) dx \text{ and} \\ \zeta_{t,N}(A) &\equiv \sum_{j=1}^{J} t_j \int_A \Lambda_p(v_{jN}(x)) w_j(x) dx.\end{aligned}$$

We also let $\sigma_{t,n}^2(A) \equiv \sum_{j=1}^J \sum_{k=1}^J t_j t_k \sigma_{jk,n}(A)$, and $\sigma_t^2(A) \equiv \sum_{j=1}^J \sum_{k=1}^J t_j t_k \sigma_{jk}(A)$. We define

$$S_{t,n}(A) \equiv \frac{n^{p/2} h^{(p-1)d/2} \{\zeta_{t,N}(A) - \mathbf{E}\zeta_{t,N}(A)\}}{\sigma_{t,n}(A)},$$

where

$$\begin{aligned}U_n &\equiv \frac{1}{\sqrt{n}} \left\{ \sum_{i=1}^{N} 1\{X_i \in \mathcal{C}\} - nP\{X \in \mathcal{C}\} \right\}, \text{ and} \\ V_n &\equiv \frac{1}{\sqrt{n}} \left\{ \sum_{i=1}^{N} 1\{X_i \in \mathbf{R}^d \backslash \mathcal{C}\} - nP\{X \in \mathbf{R}^d \backslash \mathcal{C}\} \right\}.\end{aligned}$$

LEMMA A9: *Suppose that Assumptions 1 and 2 hold. Furthermore, assume that as* $n \to \infty$, $h \to 0$, *and* $n^{-1/2} h^{-d} \to 0$. *Let* $t \in \mathbf{R}^J \backslash \{0\}$ *and* $A \subset \mathcal{C}$ *be such that* $\sigma_t^2(A) > 0$, $\alpha \equiv P\{X \in \mathbf{R}^d \backslash \mathcal{C}\} > 0$, $\rho_j(\cdot)$*'s and* $w_j(\cdot)$*'s are continuous and bounded on* A, *and condition in* (5.9) *is satisfied for all* $l = 1, \cdots, J$. *Then,*

$$(S_{t,n}(A), U_n) \xrightarrow{d} (\mathbb{Z}_1, \sqrt{1-\alpha}\mathbb{Z}_2).$$

PROOF : First, we show that

$$Cov\,(S_{t,n}(A), U_n) \to 0. \tag{5.25}$$

Write

$$Cov\,(S_{t,n}(A), U_n) = \frac{n^{p/2} h^{(p-1)d/2}}{\sigma_{t,n}(A)} \sum_{j=1}^{J} t_j \int_A Cov\,(\Lambda_p(v_{jN}(x)), U_n)\, w_j(x) dx.$$

It suffices for (5.25) to show that

$$Cov\left(n^{p/2} h^{pd/2} \{\zeta_{t,N}(A) - \mathbf{E}\zeta_{t,N}(A)\}, U_n\right) = o(h^{d/2}), \tag{5.26}$$

since $\sigma_{t,n}^2(A) \to \sigma_t^2(A) \equiv \sum_{j=1}^J \sum_{k=1}^J t_j t_k \sigma_{jk}(A) > 0$ by Lemma A7. For any $x \in \mathcal{S}_j$,

$$\left(\frac{\sqrt{nh^d} v_{jN}(x)}{\rho_{jn}(x)}, \frac{U_n}{\sqrt{P\{X \in \mathcal{C}\}}} \right) \stackrel{d}{=} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Q_n^{(i)}(x), \frac{1}{\sqrt{n}} \sum_{i=1}^{n} U^{(i)} \right),$$

where $(Q_n^{(i)}(x), U^{(i)})$'s are i.i.d. copies of $(Q_n(x), U)$ with

$$\begin{aligned} Q_n(x) &\equiv \frac{1}{h^{d/2}\rho_{jn}(x)}\left\{\sum_{i\le N_1} Y_{ji}K\left(\frac{x-X_i}{h}\right) - \mathbf{E}\left[Y_{ji}K\left(\frac{x-X_i}{h}\right)\right]\right\} \text{ and} \\ U &\equiv \frac{\sum_{i\le N_1} 1\{X_i\in\mathcal{C}\} - P\{X\in\mathcal{C}\}}{\sqrt{P\{X\in\mathcal{C}\}}}. \end{aligned}$$

Uniformly over $x\in\mathcal{S}_j$,

$$r_n(x) \equiv \mathbf{E}[Q_n(x)U] = O(h^{d/2}) = o(1), \tag{5.27}$$

by Lemma A4. Let $(Z_{1n}, Z_{2n})'$ be a centered normal random vector with the same covariance matrix as that of $(Q_n(x), U)'$. Let the 2 by 2 covariance matrix be $\Sigma_{n,2}$.

Recall the definition of $\xi_{jn}(x)$ in (5.2) and write

$$\begin{aligned} & Cov\left(\Lambda_p\left(\xi_{jn}(x)\right), n^{-1/2}\Sigma_{i=1}^n U^{(i)}\right) - Cov\left(\Lambda_p\left(Z_{1n}\right), Z_{2n}\right) \\ = \;& \mathbf{E}\left[\Lambda_p\left(\xi_{jn}(x)\right) n^{-1/2}\Sigma_{i=1}^n U^{(i)}\right] - \mathbf{E}\left[\Lambda_p\left(Z_{1n}\right) Z_{2n}\right] \equiv A_{1n}(x), \text{ say.} \end{aligned}$$

Define $\bar{\Lambda}_{n,p}(v) \equiv \Lambda_p([\Sigma_{n,2}^{1/2}v]_1)[\Sigma_{n,2}^{1/2}v]_2$, $v\in\mathbf{R}^2$. There exists some $C>0$ such that for all $n\ge 1$,

$$\begin{aligned} \sup_{v\in\mathbf{R}^2} \frac{\left|\bar{\Lambda}_{n,p}(v) - \bar{\Lambda}_{n,p}(0)\right|}{1+||v||^{p+1}\min\{||v||,1\}} &\le C \text{ and} \\ \int \sup_{u\in\mathbf{R}^2:||z-u||\le\delta} \left|\bar{\Lambda}_{n,p}(z) - \bar{\Lambda}_{n,p}(u)\right| d\Phi(z) &\le C\delta, \text{ for all } \delta\in(0,1]. \end{aligned}$$

Letting $W_n^{(i)}(x) \equiv \Sigma_{n,2}^{-1/2}(Q_n^{(i)}(x), U^{(i)})'$, observe that using (5.27) and following the arguments in (5.18), from some large n on, for some $C>0$,

$$\begin{aligned} \mathbf{E}||W_n^{(i)}(x)||^3 &= \mathbf{E}||\Sigma_{n,2}^{-1/2}(Q_n^{(i)}(x), U^{(i)})'||^3 \\ &= \mathbf{E}[\{tr(\Sigma_{n,2}^{-1/2}(Q_n^{(i)}(x), U^{(i)})'(Q_n^{(i)}(x), U^{(i)})\Sigma_{n,2}^{-1/2})\}^{3/2}] \\ &\le C(1-r_n^2(x))^{-3/2}\mathbf{E}\left[|Q_n(x)|^3 + |U|^3\right] \le Ch^{-d/2}. \end{aligned}$$

Hence, by Lemma A2,

$$\begin{aligned} & \sup_{x\in\mathcal{S}_j}|A_{1n}(x)| = \sup_{x\in\mathcal{S}_j}\left|\mathbf{E}\bar{\Lambda}_{n,p}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n W_n^{(i)}(x)\right) - \mathbf{E}\bar{\Lambda}_{n,p}\left(\tilde{Z}_n\right)\right| \\ = \;& O\left(n^{-1/2}h^{-d/2}\right) = o(h^{d/2}), \end{aligned}$$

where $\tilde{Z}_n \equiv \Sigma_{n,2}^{-1/2}(Z_{1n}, Z_{2n})'$. This completes the proof of (5.26) and hence that of (5.25).

Now, define

$$\Delta_{t,n}(x) \equiv n^{p/2}h^{(p-1)d/2}\sum_{j=1}^{J} t_j\{\Lambda_p(v_{jN}(x)) - \mathbf{E}[\Lambda_p(v_{jN}(x))]\}w_j(x).$$

Let $\{R_{n,i} : i = 1, \cdots, L_n\}$ be the collection of rectangles in $\mathbf{R}^d$ such that all the rectangles $R_{n,i}$ are of the form $R_{n,i} \equiv \Pi_{s=1}^d (a_s, b_s]$, where $h \leq b_s - a_s \leq 2h$. Let $B_{n,i} \equiv R_{n,i} \cap \mathcal{C}$ and $\mathcal{I}_n \equiv \{i : \mathbf{R} \times B_{n,i} \neq \varnothing\}$. Then, $B_{n,i}$ has Lebesgue measure $m(B_{n,i})$ bounded by $C_1 h^d$ and the cardinality of the set $\mathcal{I}_n$ is bounded by $C_2 h^{-d}$ for some positive constants C_1 and C_2. Define

$$\begin{aligned}
\alpha_{i,n} &\equiv \frac{1}{\sigma_{t,n}(A)}\int_{B_{n,i}\cap A}\Delta_{t,n}(x)dx \text{ and} \\
u_{i,n} &\equiv \frac{1}{\sqrt{n}}\left\{\sum_{j=1}^{N} 1\{X_j \in B_{n,i}\} - nP\{X_j \in B_{n,i}\}\right\}.
\end{aligned}$$

Then, we can write

$$S_{t,n}(A) = \sum_{i\in\mathcal{I}_n}\alpha_{i,n} \text{ and } U_n = \sum_{i\in\mathcal{I}_n} u_{i,n}.$$

Certainly $Var(S_{t,n}(A)) = 1$ and it is easy to check that $Var(U_n) = 1 - \alpha$. Take $\mu_1, \mu_2 \in \mathbf{R}$ and let

$$y_{i,n} \equiv \mu_1\alpha_{i,n} + \mu_2 u_{i,n}.$$

From (5.25),

$$Var\left(\sum_{i\in\mathcal{I}_n} y_{i,n}\right) \to \mu_1^2 + \mu_2^2(1-\alpha) \text{ as } n \to \infty.$$

Since $\sigma_{t,n}^r(A) = \sigma_t^r(A) + o(1),\ r > 0$, by Lemma A7 and $m(B_{n,i}) \leq Ch^d$ for a constant $C > 0$, we take $r \in (2, (2p+2)/p]$ and bound

$$\begin{aligned}
&\sigma_{t,n}^r(A)\sum_{i\in\mathcal{I}_n}\mathbf{E}|y_{i,n}|^r \\
&\leq C\sup_{x\in A}\mathbf{E}\left|\Delta_{t,n}(x)\right|^r\sum_{i\in\mathcal{I}_n}\left(\int_A\int_A\int_A 1_{B_{n,i}}(u,v,s)dudvds\right)^{r/3},
\end{aligned}$$

where $1_B(u,v,s) \equiv 1\{u \in B\}1\{v \in B\}1\{s \in B\}$. Using Jensen's inequality, we have

$$\begin{aligned}
\sup_{x\in A}\mathbf{E}\left|\Delta_{t,n}(x)\right|^r &\leq C_1 n^{rp/2}h^{r(p-1)d/2}\sup_{x\in A}\sum_{j=1}^{J} t_j^{rp}\mathbf{E}\left|v_{jN}(x)\right|^{rp} w_j^r(x) \\
&\leq C_2 n^{rp/2}h^{r(p-1)d/2}\max_{1\leq j\leq J}\sup_{x\in A\cap\mathcal{S}_j}\mathbf{E}\left|v_{jN}(x)\right|^{rp}
\end{aligned}$$

for some $C_1, C_2 > 0$. As for the last term, we apply Rosenthal's inequality (see. e.g. Lemma 2.3. of GMZ): for some constant $C > 0$,

$$\begin{aligned}
& n^{rp/2} h^{r(p-1)d/2} \sup_{x \in A \cap \mathcal{S}_j} \mathbf{E} \left| v_{jN}(x) \right|^{rp} \\
\leq\ & C h^{r(p-1)d/2} \sup_{x \in A \cap \mathcal{S}_j} \left(\frac{1}{h^{2d}} \mathbf{E} \left[Y_{ji}^2 K^2 \left(\frac{x - X_i}{h} \right) \right] \right)^{rp/2} \\
& + C h^{r(p-1)d/2} \sup_{x \in A \cap \mathcal{S}_j} \left(\frac{n}{n^{rp/2} h^{rpd}} \mathbf{E} \left[\left| Y_{ji} K \left(\frac{x - X_i}{h} \right) \right|^{rp} \right] \right).
\end{aligned}$$

By Lemma A4, the first term is $O(h^{-rd/2})$ and the last term is $O(n^{1-rp/2} h^{-rdp/2-rd/2+d})$. Hence we find that

$$\begin{aligned}
\sum_{i \in \mathcal{I}_n} \mathbf{E}|y_{i,n}|^r &= \text{Cardinality of } \mathcal{I}_n \times O\left(m(B_{n,i})^r h^{-rd/2} \{1 + n^{1-rp/2} h^{-rdp/2+d}\} \right) \\
&= O\left(h^{rd/2-d} \{1 + n^{1-rp/2} h^{-rdp/2+d}\} \right) = o(1)
\end{aligned}$$

for any $r \in (2, (2p+2)/p]$, because $n^{-1/2} h^{-d} \to 0$. Therefore, as $n \to \infty$,

$$\sum_{i \in \mathcal{I}_n} \mathbf{E}|y_{i,n}|^r \to 0 \text{ for any } r \in (2, (2p+2)/p].$$

The sequence $\{y_{i,n}\}_{i=1}^n$ is a one-dependent triangular array because X_i's are common across different j's. The desired result follows by Corollary 2 of Shergin (1979). ■

LEMMA A10: *Suppose that the conditions of Lemma A9 are satisfied, and let $A \subset \mathbf{R}^d$ be a Borel set in Lemma A9. Then,*

$$\frac{n^{p/2} h^{(p-1)d/2} \left\{ \zeta_{t,n}(A) - \mathbf{E}\zeta_{t,n}(A) \right\}}{\sigma_{t,n}(A)} \xrightarrow{d} N(0,1), \text{ as } n \to \infty.$$

PROOF: The conditional distribution of $S_{t,n}(A)$ given $N = n$ is equal to that of

$$\frac{n^{p/2} h^{(p-1)d/2}}{\sigma_{t,n}(A)} \sum_{j=1}^{J} t_j \int_A \left\{ \Lambda_p(v_{jn}(x)) - \mathbf{E}\Lambda_p(v_{jN}(x)) \right\} w_j(x) dx.$$

Using Lemma A9 and the de-Poissonization argument of Beirlant and Mason (1995) (see also Lemma 2.4 of GMZ), this conditional distribution converges to $N(0,1)$. Now by Lemma A6, it follows that

$$n^{p/2} h^{(p-1)d/2} \sum_{j=1}^{J} t_j \int_A \left\{ \mathbf{E}\Lambda_p(v_{jN}(x)) - \mathbf{E}\Lambda_p(v_{jn}(x)) \right\} w_j(x) dx \to 0,$$

as $n \to \infty$. This completes the proof. ■

Proof of Theorem 1 : Fix $\varepsilon > 0$ as in Assumption 1(iii), and take $n_0 > 0$ such that for all $n \geq n_0$,

$$\{x - uh : x \in \mathcal{S}_j, u \in [-1/2, 1/2]^d\} \subset \mathcal{S}_j^{\varepsilon} \subset \mathcal{X} \text{ for all } j \in \mathcal{J}.$$

Since we are considering the least favorable case of the null hypothesis,

$$\mathbf{E}[Y_{ji}K((x - X_i)/h)]/h^d = \int_{[-1/2,1/2]^d} m_j(x - uh)K(u)du = 0, \text{ for almost all } x \in \mathcal{S}_j,$$

for all $n \geq n_0$ and for all $j \in \mathcal{J}$. Therefore, $\hat{g}_{jn}(x) = v_{jn}(x)$ for almost all $x \in \mathcal{S}_j$, $j \in \mathcal{J}$, and for all $n \geq n_0$. From here on, we consider only $n \geq n_0$.

We fix $0 < \varepsilon_l \to 0$ as $l \to \infty$ and take a compact set $\mathcal{W}_l \subset \mathcal{S}_j$ such that for each $j \in \mathcal{J}$, w_j is bounded and continuous on $\mathcal{W}_l$ and for $s \in \{1, 2\}$,

$$\int_{\mathcal{X} \backslash \mathcal{W}_l} w_j^s(x)dx \to 0 \text{ as } l \to \infty. \tag{5.28}$$

We can choose such $\mathcal{W}_l$ following the arguments in the proof of Lemma 6.1 of GMZ because w_j^s is integrable by Assumption 1(ii). Take $M_{l,j}, v_{l,j} > 0$, $j = 1, 2, \cdots, J$, such that for $\mathcal{C}_{l,j} \equiv [-M_{l,j} + v_{l,j}, M_{l,j} - v_{l,j}]^d$,

$$P\left\{X_i \in \mathbf{R}^d \backslash \mathcal{C}_{l,j}\right\} > 0,$$

and for some Borel $A_{l,j} \subset \mathcal{C}_{l,j} \cap \mathcal{W}_l$, $\rho_j(\cdot)$ is bounded and continuous on $A_{l,j}$,

$$\begin{aligned} \sup_{x \in A_{l,j}} \left|\rho_{jn}(x) - \rho_j(x)\right| &\to 0, \text{ as } n \to \infty, \text{ and} \\ \int_{\mathcal{W}_l \backslash A_{l,j}} \rho_j(x) w_j^s(x)dx &\to 0, \text{ as } l \to \infty, \text{ for } s \in \{1, 2\}. \end{aligned} \tag{5.29}$$

The existence of $M_{l,j}, v_{l,j}$ and ε_l and the sets $A_{l,j}$ are ensured by Lemma A1. By Assumption 1(i), we find that the second convergence in (5.29) implies that $\int_{\mathcal{W}_l \backslash A_{l,j}} w_j^s(x)dx \to 0$ as $l \to \infty$, for $s \in \{1, 2\}$. Now, take $A_l \equiv \cap_{j=1}^J A_{l,j}$ and $\mathcal{C}_l \equiv \cap_{j=1}^J \mathcal{C}_{l,j}$, and observe that for $s \in \{1, 2\}$,

$$\int_{\mathcal{W}_l \backslash A_l} w_j^s(x)dx \leq \sum_{j=1}^J \int_{\mathcal{W}_l \backslash A_{l,j}} w_j^s(x)dx \to 0, \tag{5.30}$$

as $l \to \infty$ for all $j \in \mathcal{J}$.

First, choose $t \in \mathbf{R}^J \backslash \{0\}$ so that $\sigma_t^2 \equiv \sum_{j=1}^J \sum_{k=1}^J t_j t_k \sigma_{jk} > 0$ by Assumption 3. We write

$$\begin{aligned} (5.31) \qquad & \frac{\sum_{j=1}^J t_j \left\{ n^{p/2} h^{(p-1)d/2} \Gamma_j(\hat{g}_{jn}) - a_{jn} \right\}}{\sigma_{t,n}} \\ = \; & \frac{n^{p/2} h^{(p-1)d/2}}{\sigma_{t,n}} \left\{ \zeta_{t,n}(\mathcal{X}\backslash\mathcal{W}_l) - \mathbf{E}\zeta_{t,n}(\mathcal{X}\backslash\mathcal{W}_l) \right\} \\ & + \frac{n^{p/2} h^{(p-1)d/2}}{\sigma_{t,n}} \left\{ \zeta_{t,n}(\mathcal{W}_l\backslash A_l) - \mathbf{E}\zeta_{t,n}(\mathcal{W}_l\backslash A_l) \right\} \\ & + \frac{n^{p/2} h^{(p-1)d/2}}{\sigma_{t,n}} \left\{ \zeta_{t,n}(A_l) - \mathbf{E}\zeta_{t,n}(A_l) \right\}. \end{aligned}$$

Since $\mathcal{X}\backslash A_l = (\mathcal{X}\backslash\mathcal{W}_l) \cup (\mathcal{W}_l\backslash A_l)$, by Lemma A8, (5.28), and (5.30),

$$(5.32) \qquad n^{p/2} h^{(p-1)d/2} \left\{ \zeta_{t,n}(\mathcal{X}\backslash A_l) - \mathbf{E}\zeta_{t,n}(\mathcal{X}\backslash A_l) \right\} \xrightarrow{p} 0, \text{ as } n \to \infty, \text{ and } l \to \infty.$$

Furthermore, we write $\left|\sigma_{t,n}^2 - \sigma_{t,n}^2(A_l)\right|$ as

$$\begin{aligned} & \sum_{j=1}^J \sum_{k=1}^J t_j t_k \int_{\mathcal{X}} q_{jk,p}(x) \left(1 - 1_{A_l}(x)\right) \rho_{jn}^p(x) \rho_{kn}^p(x) w_j(x) w_k(x) dx \\ \leq \; & \sum_{j=1}^J \sum_{k=1}^J \sup_{x \in \mathcal{S}_j \cap \mathcal{S}_k} \left| q_{jk,p}(x) \rho_{jn}^p(x) \rho_{kn}^p(x) \right| \int_{\mathcal{X}} \left(1 - 1_{A_l}(x)\right) w_j(x) w_k(x) dx \\ = \; & \sum_{j=1}^J \sum_{k=1}^J \sup_{x \in \mathcal{S}_j \cap \mathcal{S}_k} \left| q_{jk,p}(x) \rho_{jn}^p(x) \rho_{kn}^p(x) \right| \int_{\mathcal{X}\backslash A_l} w_j(x) w_k(x) dx. \end{aligned}$$

Observe that as $l \to \infty$,

$$\left| \int_{\mathcal{X}\backslash A_l} w_j(x) w_k(x) dx \right|^2 \leq \left(\int_{\mathcal{X}\backslash A_l} w_j^2(x) dx \right) \left(\int_{\mathcal{X}\backslash A_l} w_k^2(x) dx \right) \to 0.$$

From Lemma A4, it follows that

$$(5.33) \qquad \lim_{l\to\infty} \text{limsup}_{n\to\infty} \left| \sigma_{t,n}^2 - \sigma_{t,n}^2(A_l) \right| = 0.$$

Furthermore, since $\sigma_{t,n}^2(A_l) \to \sigma_t^2(A_l)$ as $n \to \infty$ for each l by Lemma A7, and $\sigma_t^2(A_l) \to \sigma_t^2 > 0$ as $l \to \infty$, by Assumption 1, it follows that for any $\varepsilon_1 > 0$,

$$\begin{aligned} (5.34) \qquad 0 \; & < \; \sigma_t^2 - \varepsilon_1 \leq \text{liminf}_{n\to\infty} \sigma_{t,n}^2 \\ & \leq \; \text{limsup}_{n\to\infty} \sigma_{t,n}^2 \leq \sigma_t^2 + \varepsilon_1 < \infty. \end{aligned}$$

Combining this with (5.32), we find that as $n \to \infty$ and $l \to \infty$,

$$\frac{n^{p/2} h^{(p-1)d/2}}{\sigma_{t,n}} \left\{ \zeta_{t,n}(\mathcal{X}\backslash A_l) - \mathbf{E}\zeta_{t,n}(\mathcal{X}\backslash A_l) \right\} = o_P(1).$$

As for the last term in (5.31), by (5.34) and Lemma A10, as $n \to \infty$ and $l \to \infty$,

$$n^{p/2}h^{(p-1)d/2}\left|\zeta_{t,n}(A_l) - \mathbf{E}\zeta_{t,n}(A_l)\right| = O_P(1).$$

Therefore, by (5.33),

$$\begin{aligned} & \frac{n^{p/2}h^{(p-1)d/2}}{\sigma_{t,n}}\left\{\zeta_{t,n}(A_l) - \mathbf{E}\zeta_{t,n}(A_l)\right\} \\ = & \frac{n^{p/2}h^{(p-1)d/2}}{\sigma_{t,n}(A_l)}\left\{\zeta_{t,n}(A_l) - \mathbf{E}\zeta_{t,n}(A_l)\right\} + o_P(1), \end{aligned}$$

where $o_P(1)$ is a term that vanishes in probability as $n \to \infty$ and $l \to \infty$. For each $l \geq 1$, the last term converges in distribution to $N(0,1)$ by Lemma A10. Since $\sigma_{t,n}^2(A_l) \to \sigma_t^2$ as $n \to \infty$ and $l \to \infty$, we conclude that

$$\sum_{j=1}^{J} t_j \left\{n^{p/2}h^{(p-1)d/2}\Gamma_j(\hat{g}_{jn}) - a_{jn}\right\} \xrightarrow{d} N\left(0, \sigma_t^2\right).$$

■

Proof of Theorem 2 : We first show that for each $j \in \mathcal{J}$,

(5.35)
$$\begin{aligned} \hat{a}_{jn} &= a_{jn} + O_P(n^{-1/2}h^{-3d/2}) \text{ and} \\ \hat{\sigma}_{t,n}^2 &= \sigma_{t,n}^2 + O_P(n^{-1/2}h^{-3d/2}). \end{aligned}$$

For this, we show that for all $j,k = 1, \cdots, J$,

(5.36)
$$\sup_{x \in \mathcal{S}_j \cap \mathcal{S}_k} |\hat{\rho}_{jk,n}(x) - \rho_{jk,n}(x)| = O_P\left(n^{-1/2}h^{-d}\right).$$

Write $\sup_{x \in \mathcal{S}_j \cap \mathcal{S}_k} |\hat{\rho}_{jk,n}(x) - \rho_{jk,n}(x)|$ as

$$\sup_{x \in \mathcal{S}_j \cap \mathcal{S}_k}\left|\frac{1}{nh^d}\sum_{i=1}^{n}\left\{Y_{ji}Y_{ki}K^2\left(\frac{x-X_i}{h}\right) - \mathbf{E}\left[Y_{ji}Y_{ki}K^2\left(\frac{x-X_i}{h}\right)\right]\right\}\right|.$$

Let $\varphi_{n,x}(y_1,y_2,z) \equiv y_1y_2K^2((x-z)/h)$ and $\mathcal{K}_n \equiv \{\varphi_{n,x}(\cdot,\cdot,\cdot) : x \in \mathcal{S}_j \cap \mathcal{S}_k\}$. We define $N(\varepsilon, \mathcal{K}_n, L_2(Q))$ to be a covering number of $\mathcal{K}_n$ with respect to $L_2(Q)$, i.e., the smallest number of maps φ_j, $j = 1, \cdots, N_1$, such that for all $\varphi \in \mathcal{K}_n$, there exists φ_j such that $\int(\varphi_j - \varphi)^2 dQ \leq \varepsilon^2$. By Assumption 2(b), Lemma 2.6.16 of van der Vaart and Wellner (1996), and Lemma A.1 of Ghosal, Sen and van der Vaart (2000), we find that for some $C > 0$,

$$\sup_Q \log N(\varepsilon, \mathcal{K}_n, L_2(Q)) \leq C \log \varepsilon,$$

where the supremum is over all discrete probability measures. We take $\bar{\varphi}_n(y_1,y_2,z) \equiv y_1y_2||K||_\infty^2$ to be the envelope of $\mathcal{K}_n$. By Theorem 2.14.1 of van der Vaart and Wellner

(1996), we deduce that

$$n^{1/2}h^d\mathbf{E}\left[\sup_{x\in\mathcal{S}_j\cap\mathcal{S}_k}|\hat{\rho}_{jk,n}(x)-\rho_{jk,n}(x)|\right]\leq C,$$

for some positive constant C. This yields (5.36). In view of the definitions of $\hat{a}_{jn}$ and $\hat{\sigma}^2_{t,n}$, and Lemma A4, this completes the proof of (5.35).

Since $g_j(x)\leq 0$ for all $x\in\mathcal{X}$ under the null hypothesis and K is nonnegative,

$$\begin{aligned}\sup_{x\in\mathcal{S}_j}\mathbf{E}\hat{g}_{jn}(x) &= \sup_{x\in\mathcal{S}_j}\int g_j(x-uh)K(u)\,du\leq\int\sup_{x\in\mathcal{S}_j}g_j(x-uh)K(u)\,du\\ &\leq \int\sup_{x\in\mathcal{X}}g_j(x)K(u)\,du=\sup_{x\in\mathcal{X}}g_j(x)\leq 0,\end{aligned}$$

from some large n on. The second inequality follows from Assumption 1(iii). Therefore,

$$\int_{\mathcal{X}}\Lambda_p(\hat{g}_{jn}(x))w_j(x)dx\leq\int_{\mathcal{X}}\Lambda_p(\hat{g}_{jn}(x)-\mathbf{E}\hat{g}_{jn}(x))w_j(x)dx.$$

Hence by using this and (5.35), we bound $P\{\hat{T}_n>z_{1-\alpha}\}$ by

$$P\left\{\frac{1}{\sigma_{t,n}}\sum_{j=1}^{J}t_j\left\{n^{p/2}h^{(p-1)d/2}\int_{\mathcal{X}}\Lambda_p(\hat{g}_{jn}(x)-\mathbf{E}\hat{g}_{jn}(x))w_j(x)dx-a_{jn}\right\}>z_{1-\alpha}\right\}+o(1).$$

By Theorem 1, the leading probability converges to α as $n\to\infty$, delivering the desired result. ■

Proof of Theorem 3: Fix j such that $\Gamma_j(g_j)>0$. We focus on the case with $p>1$. The proof in the case with $p=1$ is simpler and hence omitted. Using the triangular inequality, we bound $|\Gamma_j(\hat{g}_{jn})-\Gamma_j(g_j)|$ by

$$\begin{aligned}&\left|\int_{\mathcal{X}}\left\{\Lambda_p(\hat{g}_{jn}(x))-\Lambda_p(\mathbf{E}\hat{g}_{jn}(x))\right\}w_j(x)dx\right|\\ &+\left|\int_{\mathcal{X}}\left\{\Lambda_p(\mathbf{E}\hat{g}_{jn}(x))-\Lambda_p(g_j(x))\right\}w_j(x)dx\right|.\end{aligned}$$

There exists n_0 such that for all $n\geq n_0$, $\sup_{x\in\mathcal{S}_j}|\mathbf{E}\hat{g}_{jn}(x)|<\infty$ by Lemma A4. Also, note that $\sup_{x\in\mathcal{S}_j}|g_j(x)|<\infty$ by Assumption 1(i). Hence, applying Lemma A3, from some large n on, for some $C_1,C_2>0$,

$$\begin{aligned}|\Gamma_j(\hat{g}_{jn})-\Gamma_j(g_j)| &\leq C_1\sum_{k=0}^{\lceil p-1\rceil}\int_{\mathcal{X}}|\hat{g}_{jn}(x)-\mathbf{E}\hat{g}_{jn}(x)|^{p-kz}w_j(x)dx\\ &\quad +C_2\sum_{k=0}^{\lceil p-1\rceil}\int_{\mathcal{X}}|\mathbf{E}\hat{g}_{jn}(x))-g_j(x)|^{p-kz}w_j(x)dx,\end{aligned}$$

where $z = (p-1)/\lceil p-1 \rceil$. Observe that $0 \leq z \leq 1$.

As for the second integral, take $\varepsilon > 0$ and a compact set $D \subset \mathbf{R}^d$ such that $\int_{\mathcal{X} \backslash D} w_j(x)dx < \varepsilon$ and g_j is continuous on D. Such a set D exists by Lemma A1. Since D is compact, g_j is in fact uniformly continuous on D. By change of variables,

$$\begin{aligned}
\mathbf{E}\hat{g}_{jn}(x) - g_j(x) &= \int_{[-1/2,1/2]^d} \{g_j(x-uh)K(u) - g_j(x)\}\, du \\
&= \int_{[-1/2,1/2]^d} \{g_j(x-uh) - g_j(x)\}\, K(u)du
\end{aligned}$$

and obtain that for $k = 0, 1, \cdots, p-1$,

$$\begin{aligned}
&\int_{\mathcal{X}} \left|\mathbf{E}\hat{g}_{jn}(x) - g_j(x)\right|^{p-kz} w_j(x)dx \\
&= \int_D \left|\mathbf{E}\hat{g}_{jn}(x) - g_j(x)\right|^{p-kz} w_j(x)dx + \int_{\mathcal{X}\backslash D} \left|\mathbf{E}\hat{g}_{jn}(x) - g_j(x)\right|^{p-kz} w_j(x)dx \\
&\leq C_3 \sup_{u \in [-1/2,1/2]^d} \sup_{x \in D \cap \mathcal{S}_j} \left|g_j(x-uh) - g_j(x)\right|^{p-kz} \\
&\quad + C_4 \int_{\mathcal{X}\backslash D} \int_{[-1/2,1/2]^d} \left|g_j(x-uh) - g_j(x)\right|^{p-kz} w_j(x)dudx,
\end{aligned}$$

for some positive constants C_3 and C_4. Note that the constant C_4 involves $||K||_\infty$. The first term is $o(1)$ as $h \to 0$, because g_j is uniformly continuous on D. By Assumption 1(i), the last term is bounded by

$$C_5 \int_{\mathcal{X}\backslash D} w_j(x)dx < C_6 \varepsilon, \text{ for some } C_5, C_6 > 0,$$

for some large n on. Since the choice of ε was arbitrary, we conclude that as $n \to \infty$,

$$|\Gamma_j(\hat{g}_{jn}) - \Gamma_j(g_j)| \leq C_1 \int_{\mathcal{X}} \left|\hat{g}_{jn}(x) - \mathbf{E}\hat{g}_{jn}(x)\right|^{p-kz} w_j(x)dx + o(1).$$

As for the leading integral, from the result of Theorem 1 (replacing $\Lambda_p(\cdot)$ there by $|\cdot|^{p-kz}$), we find that

$$\int_{\mathcal{X}} \left|\hat{g}_{jn}(x) - \mathbf{E}\hat{g}_{jn}(x)\right|^{p-kz} w_j(x)dx = O_P(n^{-(p-kz)/2} h^{-(p-kz-1)d/2-d/2}).$$

Since $n^{-1/2}h^{-d/2} \to 0$ by the condition of the theorem, we conclude that $\Gamma_j(\hat{g}_{jn}) \xrightarrow{p} \Gamma_j(g_j)$. Using the similar argument, we can also show that

$$\hat{\sigma}_{t,n}^2 \xrightarrow{p} \sigma_t^2 \text{ and } \hat{a}_{jn} = O_P(h^{-d/2}) \text{ for all } j \in \mathcal{J},$$

where $\sigma_t^2 = t'\Sigma t > 0$. Hence

$$\hat{\sigma}_{t,n}^{-1}\{\Gamma_j(\hat{g}_{jn}) - n^{-p/2}h^{-pd/2}h^{d/2}\hat{a}_{jn}\} \xrightarrow{p} \sigma_t^{-1}\Gamma_j(g_j) > 0.$$

Therefore,

$$P\{\hat{T}_n > z_{1-\alpha}\} \geq P\left\{\sigma_t^{-1}\Gamma_j(g_j) > 0\right\} + o(1) \to 1,$$

where the inequality holds by the fact that $n^{-1/2}h^{-d/2} \to 0$ and $\hat{a}_{jn} = O_P(h^{-d/2})$. ■

LEMMA A11: *Suppose that Assumptions 1-3 hold, $n^{-1/2}h^{-d} \to 0$, and that $\sqrt{n}g_j(\cdot) = \delta_j(\cdot)$, $j \in \mathcal{J}$, for real bounded functions δ_j, $j \in \mathcal{J}$, for each n. Then,*

$$\frac{1}{\sigma_{t,n}}\sum_{j=1}^{J} t_j\left\{n^{p/2}h^{(p-1)d/2}\Gamma_{j,\delta}(\hat{g}_{jn}) - \tilde{a}_{jn}\right\} \xrightarrow{d} N(0,1),$$

where $\tilde{a}_{jn} \equiv \int \mathbf{E}\Lambda_p(h^{-d/(2p)}\rho_{jn}(x)\mathbb{Z}_1 + h^{d(p-1)/(2p)}\delta_{jn}(x))w_j(x)dx$ and $\delta_{jn}(x) \equiv \int \delta_j(x - uh)K(u)du$.

PROOF: By change of variables,

$$\sqrt{n}\mathbf{E}\hat{g}_{jn}(x) = \sqrt{n}\int g_j(x-uh)K(u)du = \int \delta_j(x-uh)K(u)du.$$

Since δ_j is bounded, $\sup_{x\in\mathcal{S}_j}\sqrt{n}\,|\mathbf{E}\hat{g}_{jn}(x)| = O(1)$. Hence

$$\frac{\sqrt{nh^d}\hat{g}_{jN}(x)}{\rho_{jn}(x)} = \xi_{jn}(x) + \frac{\sqrt{nh^d}\mathbf{E}\hat{g}_{jn}(x)}{\rho_{jn}(x)} = \xi_{jn}(x) + O(h^{d/2}),$$

under the local alternatives. Using this and following the proof of Lemma A7, we find that under the local alternatives, $\sigma_{jk,n} \to \sigma_{jk}$. Also, as in the proof of Theorem 1, we deduce that

$$\frac{1}{\sigma_{t,n}}\sum_{j=1}^{J} t_j n^{p/2}h^{(p-1)d/2}\left\{\Gamma_j(\hat{g}_{jn}) - \mathbf{E}\Gamma_j(\hat{g}_{jn})\right\} \xrightarrow{d} N(0,1). \tag{5.37}$$

Now, as for $n^{p/2}h^{(p-1)d/2}\sigma_{t,n}^{-1}\mathbf{E}\Gamma_j(\hat{g}_{jn})$, we first note that

$$\begin{aligned} n^{p/2}h^{(p-1)d/2}\Gamma_j(\hat{g}_{jn}) &= h^{-d/2}\Gamma_j(n^{1/2}h^{d/2}\{\hat{g}_{jn} - \mathbf{E}\hat{g}_{jn}\} + n^{1/2}h^{d/2}\mathbf{E}\hat{g}_{jn}) \\ &= \Gamma_j(h^{-d/(2p)}\rho_{jn}(x)\xi_{jn}(x) + h^{(p-1)d/(2p)}\delta_{jn}(x)). \end{aligned}$$

We follow the proof of Lemma A4 and Lemma A6 (using (5.11) and applying Lemma A2 with $\Lambda_p(v)$ in Lemma A6 replaced by $\Lambda_p(v + h^{d(p-1)/(2p)}\delta_{jn}(x)/\rho_{jn}(x)))$ to deduce that

$$\int\left\{n^{p/2}h^{(p-1)d/2}\mathbf{E}\Lambda_p(\hat{g}_{jn}(x)) - \mathbf{E}\Lambda_p(\bar{Z}_{jn}(x))\right\}w_j(x)dx \to 0,$$

where $\bar{Z}_{jn}(x) \equiv h^{-d/(2p)}\rho_{jn}(x)\mathbb{Z}_1 + h^{d(p-1)/(2p)}\delta_{jn}(x)$. ■

PROOF OF THEOREM 4: Under the local alternatives, by (5.35) and (5.37),

$$
\begin{aligned}
(5.38) \qquad & P\{\hat{T}_n > z_{1-\alpha}\} \\
&= P\{\hat{\sigma}_{t,n}^{-1}\Sigma_{j=1}^{J} t_j\{n^{p/2}h^{(p-1)d/2}\Gamma_j(\hat{g}_{jn}) - \hat{a}_{jn}\} > z_{1-\alpha}\} \\
&= P\{\sigma_t^{-1}\Sigma_{j=1}^{J} t_j\{n^{p/2}h^{(p-1)d/2}\Gamma_j(\hat{g}_{jn}) - \tilde{a}_{jn} + \tilde{a}_{jn} - \hat{a}_{jn}\} > z_{1-\alpha}\} + o(1) \\
&= P\{\mathbb{Z}_1 + \sigma_t^{-1}\Sigma_{j=1}^{J} t_j\{\tilde{a}_{jn} - a_{jn}\}) > z_{1-\alpha}\} + o(1).
\end{aligned}
$$

Fix $\varepsilon > 0$ and take a compact set $A_\varepsilon \subset \mathcal{S}_j$ such that $\int_{\mathcal{S}_j \backslash A_\varepsilon} w_j(x)dx < \varepsilon$. Furthermore, without loss of generality, let A_ε be a set on which $\delta_j(\cdot)$ and $\rho_j(\cdot)$ are uniformly continuous. Then for any $\varepsilon_1 > 0$. Then there exists $\lambda > 0$ such that $\sup_{z \in \mathbf{R}^d : ||x-z|| < \lambda} |\delta_j(z) - \delta_j(x)| \leq \varepsilon_1$ uniformly over $x \in A_\varepsilon$. Hence from some large n on,

$$
\sup_{x \in A_\varepsilon} |\delta_{jn}(x) - \delta_j(x)| \leq \int_{[-1/2,1/2]^d} \sup_{x \in A_\varepsilon} |\delta_j(x - uh) - \delta_j(x)| \, K(u)du \leq \varepsilon_1.
$$

Since the choice of ε_1 was arbitrary, we conclude that $|\delta_{jn}(x) - \delta_j(x)| \to 0$ uniformly over $x \in A_\varepsilon$. Similarly, we also conclude that $|\rho_{jn}(x) - \rho_j(x)| \to 0$ uniformly over $x \in A_\varepsilon$. Using these facts, we analyze $\sigma_t^{-1}\Sigma_{j=1}^{J} t_j\{\tilde{a}_{jn} - a_{jn}\}$ for each case of $p \in \{1, 2\}$.

(i) Suppose $p = 1$. For $\gamma > 0$ and $\mu \in \mathbf{R}$,

$$
\begin{aligned}
\mathbf{E}\max\{\gamma\mathbb{Z}_1 + \mu, 0\} &= \mathbf{E}[\gamma\mathbb{Z}_1 + \mu | \gamma\mathbb{Z}_1 + \mu > 0] P\{\gamma\mathbb{Z}_1 + \mu > 0\} \\
&= \{\mu + \gamma\phi(-\mu/\gamma)/(1 - \Phi(-\mu/\gamma))\}(1 - \Phi(-\mu/\gamma)) \\
&= \mu(1 - \Phi(-\mu/\gamma)) + \gamma\phi(-\mu/\gamma) \\
&= \mu\Phi(\mu/\gamma) + \gamma\phi(\mu/\gamma).
\end{aligned}
$$

Taking $\gamma_{jn} \equiv h^{-d/2}\rho_{jn}(x)$, we have

$$
\begin{aligned}
& \mathbf{E}\max\{\gamma_{jn}\mathbb{Z}_1 + \delta_{jn}(x), 0\} - \mathbf{E}\max\{\gamma_{jn}\mathbb{Z}_1, 0\} \\
&= \delta_{jn}(x)\Phi(\delta_{jn}(x)/\gamma_{jn}) + \gamma_{jn}\phi(\delta_{jn}(x)/\gamma_{jn}) - \gamma_{jn}\phi(0) \\
&= \delta_{jn}(x)\Phi(0) + O(h^{d/2}),
\end{aligned}
$$

uniformly in $x \in \mathcal{S}_j$. Therefore, we can write $\lim_{n\to\infty}\{\tilde{a}_{jn} - a_{jn}\}$ as

$$
\begin{aligned}
& \lim_{n\to\infty} \int_{\mathcal{X}} \mathbf{E}[\Lambda_1(h^{-d/2}\rho_{jn}(x)\mathbb{Z}_1 + \delta_{jn}(x)) - \Lambda_1(h^{-d/2}\rho_{jn}(x)\mathbb{Z}_1)] w_j(x)dx \\
&= \frac{1}{2}\int_{A_\varepsilon} \delta_j(x)w_j(x)dx + \frac{1}{2}\lim_{n\to\infty}\int_{\mathcal{X}\backslash A_\varepsilon} \delta_{jn}(x)w_j(x)dx.
\end{aligned}
$$

Since δ_{jn} is uniformly bounded, there exists $C > 0$ such that the last integral is bounded by $C\varepsilon$. Since the choice of $\varepsilon > 0$ was arbitrary, in view of (5.38), this gives the desired result.

(ii) Suppose $p = 2$. For $\gamma > 0$ and $\mu \in \mathbf{R}$,

$$\begin{aligned}\mathbf{E}\max\{\gamma\mathbb{Z}_1 + \mu, 0\}^2 &= \mathbf{E}[(\gamma\mathbb{Z}_1 + \mu)^2 \,|\gamma\mathbb{Z}_1 + \mu > 0] P\{\gamma\mathbb{Z}_1 + \mu > 0\} \\ &= \left(\mu^2 + \gamma^2\right)\Phi(\mu/\gamma) + \mu\gamma\phi(\mu/\gamma).\end{aligned}$$

Taking $\gamma_{jn} \equiv h^{-d/4}\rho_{jn}(x)$ and $\mu_{jn} = h^{d/4}\delta_{jn}(x)$, we have

$$\begin{aligned}&\mathbf{E}\max\{\gamma_{jn}\mathbb{Z}_1 + \mu_{jn}, 0\}^2 - \mathbf{E}\max\{\gamma_{jn}\mathbb{Z}_1, 0\}^2 \\ =\ & \gamma_{jn}^2\{\Phi(\mu_{jn}/\gamma_{jn}) - \Phi(0)\} + \mu_{jn}^2\Phi(\mu_{jn}/\gamma_{jn}) + \mu_{jn}\gamma_{jn}\phi(\mu_{jn}/\gamma_{jn}) \\ =\ & \{\mu_{jn}\gamma_{jn}\phi(0) + O(h^{d/2})\} + O(h^{d/2}) + \{\mu_{jn}\gamma_{jn}\phi(0) + O(h^d)\} \\ =\ & 2\phi(0)\delta_{jn}(x)\rho_{jn}(x) + O(h^{d/2}),\ \text{uniformly in } x \in \mathcal{S}_j.\end{aligned}$$

Hence we write $\lim_{n\to\infty}\{\tilde{a}_{jn} - a_{jn}\}$ as

$$\begin{aligned}&\lim_{n\to\infty}\int_{\mathcal{S}_j}\mathbf{E}[\Lambda_2(h^{-d/4}\rho_{jn}(x)\mathbb{Z}_1 + h^{d/4}\delta_{jn}(x)) - \Lambda_2(h^{-d/4}\rho_{jn}(x)\mathbb{Z}_1)]w_j(x)dx \\ =\ & 2\phi(0)\lim_{n\to\infty}\int_{\mathcal{S}_j}\delta_{jn}(x)\rho_{jn}(x)w_j(x)dx + O(h^{d/2}) \\ =\ & \sqrt{\frac{2}{\pi}}\lim_{n\to\infty}\int_{A_\varepsilon}\delta_{jn}(x)\rho_{jn}(x)w_j(x)dx + \sqrt{\frac{2}{\pi}}\lim_{n\to\infty}\int_{\mathcal{S}_j\setminus A_\varepsilon}\delta_{jn}(x)\rho_{jn}(x)w_j(x)dx + O(h^{d/2}).\end{aligned}$$

The second term is bounded by $C\varepsilon$ for some $C > 0$, because $\delta_{jn}\rho_{jn}$ is bounded. Since the choice of $\varepsilon > 0$ was arbitrary and

$$\int_{A_\varepsilon}\delta_{jn}(x)\rho_{jn}(x)w_j(x)dx \to \int_{A_\varepsilon}\delta_j(x)\rho_j(x)w_j(x)dx,\ \text{as } n \to \infty,$$

in view of (5.38), this gives the desired result. ■

Proof of Theorem 5: Similarly as before, we fix $\varepsilon > 0$ and take a compact set $A_\varepsilon \subset \mathcal{S}_j$ such that $\int_{\mathcal{S}_j\setminus A_\varepsilon} w_j(x)dx < \varepsilon$ and $\delta_j(\cdot)$ and $\delta_j(\cdot)\rho_j^{-1}(\cdot)$ are uniformly continuous on A_ε. By change of variables and uniform continuity,

$$\begin{aligned}\sup_{x\in A_\varepsilon}|\delta_{jn}(x)\rho_{jn}^{-1}(x) - \delta_j(x)\rho_j^{-1}(x)| &\to 0 \text{ and} \\ \sup_{x\in A_\varepsilon}|\delta_{jn}(x) - \delta_j(x)| &\to 0.\end{aligned}$$

(i) Suppose $p = 1$. For $\gamma > 0$ and $\mu \in \mathbf{R}$,

$$\mathbf{E}\,|\gamma\mathbb{Z}_1 + \mu| = 2\gamma\phi(\mu/\gamma) + 2\mu\left[\Phi(\mu/\gamma) - 1/2\right].$$

With $\gamma_{jn} \equiv h^{-d/2}\rho_{jn}(x)$ and $\mu_{jn} = h^{-d/4}\delta_{jn}(x)$, we find that uniformly over $x \in \mathcal{S}_j$,

$$\begin{aligned}
& \mathbf{E}|\gamma_{jn}\mathbb{Z}_1 + \mu_{jn}| - \mathbf{E}|\gamma_{jn}\mathbb{Z}_1| \\
= \; & 2\gamma_{jn}[\phi(\mu_{jn}/\gamma_{jn}) - \phi(0)] + 2\mu_{jn}[\Phi(\mu_{jn}/\gamma_{jn}) - 1/2] \\
= \; & [\phi''(0) + 2\phi(0)]\, \delta_{jn}^2(x)\rho_{jn}^{-1}(x) + O(h^{d/4}).
\end{aligned}$$

Therefore, we write $\lim_{n\to\infty}\{\tilde{a}_{jn} - a_{jn}\}$ as

$$\begin{aligned}
& \lim_{n\to\infty}\int_{\mathcal{S}_j} \mathbf{E}[\Lambda_1(h^{-d/2}\rho_{jn}(x)\mathbb{Z}_1 + n^{-d/4}\delta_{jn}(x)) - \Lambda_1(h^{-d/2}\rho_{jn}(x)\mathbb{Z}_1)]w_j(x)dx \\
= \; & \frac{1}{\sqrt{2\pi}}\lim_{n\to\infty}\int_{\mathcal{S}_j} \delta_{jn}^2(x)\rho_{jn}^{-1}(x)w_j(x)dx + O(h^{d/4}) \\
= \; & \frac{1}{\sqrt{2\pi}}\int_{A_\varepsilon} \delta_j^2(x)\rho_j^{-1}(x)w_j(x)dx + \frac{1}{\sqrt{2\pi}}\lim_{n\to\infty}\int_{\mathcal{S}_j\backslash A_\varepsilon} \delta_{jn}^2(x)\rho_{jn}^{-1}(x)w_j(x)dx + o(1).
\end{aligned}$$

By Assumption 4 and Lemma A4, $\delta_{jn}^2(x)\rho_{jn}^{-1}(x)$ is bounded uniformly over $x \in \mathcal{S}_j$, enabling us to bound the second integral by $C\varepsilon$ for some $C > 0$. Since ε is arbitrarily chosen, in view of (5.38), this gives the desired result.

(ii) Suppose $p = 2$. We have, for each $x \in \mathcal{S}_j$,

$$\mathbf{E}\{h^{-d/4}\rho_{jn}(x)\mathbb{Z}_1 + \delta_{jn}(x)\}^2 - \mathbf{E}\{h^{-d/4}\rho_{jn}(x)\mathbb{Z}_1\}^2 = \delta_{jn}^2(x).$$

Therefore, we write $\lim_{n\to\infty}\{\tilde{a}_{jn} - a_{jn}\}$ as

$$\begin{aligned}
& \lim_{n\to\infty}\int_{\mathcal{X}} \mathbf{E}[\Lambda_2(h^{-d/4}\rho_{jn}(x)\mathbb{Z}_1 + \delta_{jn}(x)) - \Lambda_2(h^{-d/4}\rho_{jn}(x)\mathbb{Z}_1)]w_j(x)dx \\
= \; & \int_{A_\varepsilon} \delta_j^2(x)w_j(x)dx + \lim_{n\to\infty}\int_{\mathcal{S}_j\backslash A_\varepsilon} \delta_{jn}(x)w_j(x)dx + o(1)
\end{aligned}$$

The second integral is bounded by $C\varepsilon$ for some $C > 0$, and in view of (5.38), this gives the desired result. ■

Table 1. Empirical probability that H_0 is rejected when H_0 is true

Conditional Variance	Sample Size	Bandwidth	$m(x) \equiv 0$ L_1 test	$m(x) \equiv 0$ L_2 test	$m(x) \equiv x(1-x) - 0.25$ L_1 test	$m(x) \equiv x(1-x) - 0.25$ L_2 test
$\sigma(x) \equiv 1$	50	$c_h = 1.0$	0.067	0.043	0.035	0.022
		$c_h = 1.5$	0.070	0.049	0.033	0.025
		$c_h = 2.0$	0.073	0.050	0.033	0.029
	200	$c_h = 1.0$	0.070	0.059	0.010	0.018
		$c_h = 1.5$	0.064	0.064	0.012	0.021
		$c_h = 2.0$	0.066	0.071	0.012	0.022
	1000	$c_h = 1.0$	0.055	0.064	0.000	0.007
		$c_h = 1.5$	0.054	0.069	0.000	0.005
		$c_h = 2.0$	0.057	0.065	0.000	0.005
$\sigma(x) = x$	50	$c_h = 1.0$	0.065	0.046	0.021	0.017
		$c_h = 1.5$	0.061	0.050	0.019	0.019
		$c_h = 2.0$	0.060	0.052	0.021	0.019
	200	$c_h = 1.0$	0.063	0.043	0.004	0.008
		$c_h = 1.5$	0.068	0.052	0.006	0.009
		$c_h = 2.0$	0.068	0.054	0.006	0.010
	1000	$c_h = 1.0$	0.055	0.051	0.000	0.002
		$c_h = 1.5$	0.057	0.049	0.000	0.002
		$c_h = 2.0$	0.056	0.045	0.000	0.002

Notes: The nominal level for each test is $\alpha = 0.05$. There are 1000 Monte Carlo replications in each experiment.

TABLE 2. Empirical probability that H_0 is rejected when H_0 is false

Conditional Variance	Sample Size	Bandwidth	$m(x) \equiv x(1-x) - 0.15$ L_1 test	L_2 test	$m(x) \equiv x(1-x) - 0.05$ L_1 test	L_2 test
$\sigma(x) \equiv 1$	50	$c_h = 1.0$	0.104	0.063	0.273	0.156
		$c_h = 1.5$	0.100	0.069	0.271	0.189
		$c_h = 2.0$	0.110	0.074	0.272	0.208
	200	$c_h = 1.0$	0.156	0.142	0.601	0.515
		$c_h = 1.5$	0.158	0.142	0.603	0.556
		$c_h = 2.0$	0.164	0.150	0.612	0.571
	1000	$c_h = 1.0$	0.351	0.332	0.993	0.985
		$c_h = 1.5$	0.378	0.384	0.993	0.988
		$c_h = 2.0$	0.401	0.413	0.994	0.989
$\sigma(x) = x$	50	$c_h = 1.0$	0.152	0.075	0.535	0.271
		$c_h = 1.5$	0.160	0.090	0.549	0.315
		$c_h = 2.0$	0.161	0.098	0.548	0.349
	200	$c_h = 1.0$	0.292	0.161	0.973	0.821
		$c_h = 1.5$	0.306	0.183	0.978	0.857
		$c_h = 2.0$	0.322	0.206	0.979	0.889
	1000	$c_h = 1.0$	0.843	0.653	1.000	1.000
		$c_h = 1.5$	0.867	0.711	1.000	1.000
		$c_h = 2.0$	0.884	0.764	1.000	1.000

Notes: The nominal level for each test is $\alpha = 0.05$. There are 1000 Monte Carlo replications in each experiment.

References

[1] Anderson, G., O. Linton, and Y.-J. Whang (2009): "Nonparametric estimation of a polarization measure," *Cemmap Working Papers*, CWP 14/09, available at `http://www.cemmap.ac.uk/wps/cwp1409.pdf`.

[2] Andrews, D. W. K. and X. Shi (2010): "Inference based on conditional moment inequalities," *Cowles Foundation Discussion Paper*, No. 1761, available at `http://cowles.econ.yale.edu/P/cd/d17b/d1761.pdf`.

[3] Bai, J. (2003): "Testing parametric conditional distributions of dynamic models," *Review of Economics and Statistics* 85, 531-549.

[4] Beirlant, J., and D. M. Mason (1995): "On the asymptotic normality of L_p-norms of empirical functionals," *Mathematical Methods of Statistics* 4, 1-19.

[5] Biau, G., B. Cadre, D. M. Mason, and B. Pelletier (2009): "Asymptotic normality in density support estimation," *Electronic Journal of Probability* 14, 2617-2635.

[6] Bickel, P. J. and M. Rosenblatt (1973): "On some global measures of the deviations of density function estimates,"*Annals of Statistics* 1, 1071-1095.

[7] Blum, J. R., J. Kiefer, and M. Rosenblatt (1961): "Distribution free tests of independence based on the sample distribution function,"*Annals of Mathematical Statistics* 32, 485-498.

[8] Chernozhukov, V., S. Lee, and A. Rosen (2009): "Intersection bounds: estimation and inference,"*Cemmap Working Papers*, CWP 19/09, available at `http://www.cemmap.ac.uk/wps/cwp1909.pdf`.

[9] Chiappori, P.-A., B. Jullien, B. Salanié, and F. Salanié (2006): "Asymmetric information in insurance: general testable implications,"*Rand Journal of Economics* 37, 783-798.

[10] Claeskens, G. and I. van Keilegom (2003): "Bootstrap confidence bands for regression curves and their derivatives,"*Annals of Statistics* 31, 1852-1884.

[11] Csörgő, M. and L. Horváth (1988): "Central limit theorems for L_p-norms of density estimators," *Zeitschrits für Wahrscheinlichkeitheorie und verwandte Gebiete* 80, 269-291.

[12] Delgado, M. A. and W. González Manteiga (2001): "Significance testing in nonparametric regression based on the bootstrap," *Annals of Statistics* 29, 1469-1507.

[13] Delgado, M. A. and J. Mora (2000): "A nonparametric test for serial independence of regression errors," *Biometrika* 87, 228-234.

[14] Devroye, L. and L. Györfi (1985): *Nonparametric Density Estimation: The L1 View*, Wiley, New York.

[15] DiBenedetto, E. (2001): *Real Analysis*, Birkhäuser, New York.

[16] Liran Einav, L., A. Finkelstein, and J. Levin (2010): "Beyond testing: empirical models of insurance markets,"*Annual Review of Economics* 2, 311-336.

[17] Gao, J. and I. Gijbels (2008). "Bandwidth selection in nonparametric kernel testing," *Journal of the American Statistical Association* 103(484), 1584-1594.

[18] Ghosal, S., A. Sen and A. W. van der Vaart (2000): "Testing monotonicity of regression," *Annals of Statistics* 28, 1054-1082.

[19] Giné, E., D. M. Mason, and A. Y. Zaitsev (2003): "The L_1-norm density estimator process," *Annals of Probability* 31, 719-768.

[20] Hall, P., Huber,C., and Speckman,P.L. (1997): "Covariate-matched one-sided tests for the difference between functional means, " *J. Amer. Statist. Assoc.* 92, 1074-1083.

[21] Hall, P. and I. Van Keilegom (2005): "Testing for monotone increasing hazard rate, "*Annals of Statistics* 33, 1109-1137.

[22] Hall, P. and Yatchew, A. (2005): "Unified approach to testing functional hypotheses in semiparametric contexts,"*Journal of Econometrics* 127(2), 225-252.

[23] Härdle, W. and E. Mammen (1993): "Comparing nonparametric versus parametric regression fits," *Annals of Statistics* 21, 1926-1947.

[24] Horváth, L. (1991): "On L_p-norms of multivariate density estimators," *Annals of Statistics* 19, 1933-1949.

[25] Khmaladze, E. V. (1993): "Goodness of fit problem and scanning innovation martingales," *Annals of Statistics* 21, 798-829.

[26] Khmaladze, E. V. and H. Koul (2004): "Martingale transforms goodness-of-fit tests in regression models," *Annals of Statistics* 32, 995-1034.

[27] Koul, H.L. and Schick,A. (1997): "Testing for the equality of two nonparametric regression curves, " *J. Statist. Plann. Inference* 65, 293-314.

[28] Koul, H.L. and Schick,A. (2003): "Testing for superiority among two regression curves, " *J. Statist. Plann. Inference* 117, 15-33.

[29] Lee, S. and Y.-J. Whang (2009): "Nonparametric tests of conditional treatment effects," *Cemmap Working Papers*, CWP 36/09, available at `http://www.cemmap.ac.uk/wps/cwp3609.pdf`.

[30] Liero, H., H. Läuter, and V. Konakov (1998): "Nonparametric versus parametric goodness of fit,"*Statistics* 31, 115-149.

[31] Manski, C. F. (2003): *Partial Identification of Probability Distributions*, Springer-Verlag, New York.

[32] Mason, D. M. (2009): "Risk bounds for kernel density estimators,"*Journal of Mathematical Sciences* 163, 238-261.

[33] Mason, D. M. and W. Polonik (2009): "Asymptotic normality of plug-in level set estimates," *Annals of Applied Probability* 19, 1108-1142.

[34] Pinelis, I. F. (1994): "On a majorization inequality for sums of independent random variables," *Statistics and Probability Letters* 19, 97-99.

[35] Shergin, V. V. (1979): "On the convergence rate in the central limit theorem for m-dependent random variables," *Theory of Probability and Its Applications*, 24, 782-796 (translated from Russian).

[36] Song, K. (2009): "Testing conditional independence via Rosenblatt transforms," *Annals of Statistics* 37, 4011-4045.

[37] Stute, W. (1997): "Nonparametric model checks for regression," *Annals of Statistics* 25, 613-642.

[38] Stute, W., S. Thies, and L. Zhu (1998): "Model checks for regression: an innovation process approach," *Annals of Statistics* 26, 1916-1934.

[39] Sweeting, T. J.(1977): "Speeds of convergence in the multidimensional central limit theorem," *Annals of Probability* 5 28-41.

[40] Tallis, G. M.(1961): "The moment generating function of the truncated multi-normal distribution," *Journal of the Royal Statistical Society. Series B (Methodological)* 23, 223-229.

[41] Tripathi, G. and Y. Kitamura (2003): "Testing conditional moment restrictions," *Annals of Statistics* 31, 2059-2095.

[42] van der Vaart, A. W. and J. A. Wellner (1996): *Weak Convergence and Empirical Processes*, New York, NY, Springer-Verlag.

Department of Economics, Seoul National University, 599 Gwanak-ro, Gwanak-gu, Seoul, 151-742, Republic of Korea, and Centre for Microdata Methods and Practice, Institute for Fiscal Studies, 7 Ridgmount Street, London, WC1E 7AE, UK.

E-mail address: sokbae@gmail.com

Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19096, USA.

E-mail address: kysong@sas.upenn.edu

Department of Economics, Seoul National University, 599 Gwanak-ro, Gwanak-gu, Seoul, 151-742, Republic of Korea.

E-mail address: whang@snu.ac.kr