The Annals of Statistics
1999, Vol. 27, No. 2, 697–731

NEAREST NEIGHBOR INVERSE REGRESSION

BY TAILEN HSING

Texas A&M University and National University of Singapore

Sliced inverse regression (SIR), formally introduced by Li, is a very general procedure for performing dimension reduction in nonparametric regression. This paper considers a version of SIR in which the "slices" are determined by nearest neighbors and the response variable takes value possibly in a multidimensional space. It is shown, under general conditions, that the "effective dimension reduction space" can be estimated with rate $n^{-1/2}$ where n is the sample size.

1. Introduction. In nonparametric regression, the presence of a large number of potential predictors renders model fitting ineffective. This, known as the curse of dimensionality, is caused in large part by the sparseness of the data scattered in a relatively high-dimensional space. As a result, it is a standard practice to consider reducing the dimension of the predictor variable at some stage. The quality of this reduction step is crucial to the success of the model fitting. To handle this important problem there are many proposals. With no intention of being complete, we mention here Friedman and Stuetzle (1981), Brieman, Friedman, Olshen and Stone (1984), Huber (1985), Hastie and Tibshirani (1986), Hall (1989), Härdle and Stoker (1989), Chen (1991), Li (1991) and Samorov (1993). In this paper we focus on the procedure sliced inverse regression, or SIR, formally introduced in Li (1991).

Throughout this paper let $(\mathbf{X}_i, \mathbf{Y}_i)$, $1 \le i \le n$, be iid random vectors where $\mathbf{X}_i \in \mathbb{R}^{d_{\mathbf{x}}}$ is the predictor variable and $\mathbf{Y}_i \in \mathbb{R}^{d_{\mathbf{y}}}$ the response variable. Both $d_{\mathbf{x}}$ and $d_{\mathbf{y}}$ may be bigger than 1. For convenience $(\mathbf{X}, \mathbf{Y})$ will be a generic variable having the same distribution as $(\mathbf{X}_1, \mathbf{Y}_1)$. We first briefly introduce the essential elements of SIR for which the details can be found in Li (1991). There are a number of basic assumptions in SIR.

(SIRa) The distribution of $\mathbf{Y}$ given $\mathbf{X}$ depends only on K linear combinations of $\mathbf{X}$, say $\boldsymbol{\beta}_1'\mathbf{X}, \ldots, \boldsymbol{\beta}_K'\mathbf{X}$.
(SIRb) For any $\mathbf{b} \in \mathbb{R}^{d_{\mathbf{x}}}$, the conditional expectation $E(\mathbf{b}'\mathbf{X} | \boldsymbol{\beta}_1'\mathbf{X}, \ldots, \boldsymbol{\beta}_K'\mathbf{X})$ is linear in $\boldsymbol{\beta}_1'\mathbf{X}, \ldots, \boldsymbol{\beta}_K'\mathbf{X}$.

Clearly, the $\boldsymbol{\beta}_k$ in (a) and (b) are not identifiable. However, the linear space $\mathscr{E}$ spanned by the $\boldsymbol{\beta}_k$ is. The space $\mathscr{E}$ is called the effective dimension reduction (e.d.r.) space. SIR achieves dimension reduction by identifying the e.d.r. space. The assumption (b) is implied by but not equivalent to spherical symmetry of the distribution of $\mathbf{X}$. Hall and Li (1993) contains an interesting justifi-

Received August 1997; revised December 1998.
AMS 1991 *subject classifications*. Primary 62G05; secondary 60F05.
Key words and phrases. Central limit theorem, dimension reduction, nonparametric regression, sliced inverse regression.

cation of the assumption. Assume also that var($\mathbf{X}$) is nonsingular to avoid complication. It is easier to work with the standardized version of $\mathbf{X}$ and $\mathscr{E}$. Let

$$\tilde{\mathbf{X}}_i = \mathbf{\Sigma}^{-1/2}(\mathbf{X}_i - \boldsymbol{\mu}) \tag{1}$$

where $\boldsymbol{\mu}$ is the mean of $\mathbf{X}$ and $\mathbf{\Sigma}^{1/2}$ is any square root of var($\mathbf{X}$). Define

$$\boldsymbol{\nu}(\mathbf{y}) = E(\tilde{\mathbf{X}}|\mathbf{Y} = \mathbf{y})$$

and

$$\mathbf{\Lambda} = \mathrm{var}(\nu(\mathbf{Y})) = E[\nu(\mathbf{Y})\nu(\mathbf{Y})'].$$

Also define the standardized e.d.r. space [cf. Li (1991)]

$$\tilde{\mathscr{E}} = \mathbf{\Sigma}^{1/2}(\mathscr{E}) = \{\mathbf{\Sigma}^{1/2}\mathbf{v}\colon \mathbf{v} \in \mathscr{E}\}.$$

Under (a) and (b), it is shown as in Theorem 3.1 and Corollary 3.1 of Li (1991) that

$$\boldsymbol{\nu}(\mathbf{y})'\mathbf{v} = 0 \quad \text{for all } \mathbf{v} \in \tilde{\mathscr{E}}^{\perp} \text{ and all } \mathbf{y}, \tag{2}$$

where $\tilde{\mathscr{E}}^{\perp}$ is the linear space orthogonal to $\tilde{\mathscr{E}}$. That is, the standardized inverse regression curve $\{\boldsymbol{\nu}(\mathbf{y})\colon \text{all } \mathbf{y}\}$ is contained in $\tilde{\mathscr{E}}$. This implies that

$$\mathbf{\Lambda}\mathbf{v} = \mathbf{0} \quad \text{for all } \mathbf{v} \in \tilde{\mathscr{E}}^{\perp},$$

from which it follows readily (using symmetry) that

$$\mathscr{C}(\mathbf{\Lambda}) \subset \tilde{\mathscr{E}}, \tag{3}$$

where $\mathscr{C}(\mathbf{A})$ denotes the column space (or range space) of the matrix $\mathbf{A}$. We will assume throughout that

(SIRc) $\tilde{\mathscr{E}} = \mathscr{C}(\mathbf{\Lambda})$.

This is of course sometimes violated and modifications would be required to make SIR work. See Li (1992) and Cook (1998).

In Li (1991), the e.d.r. space is estimated in the following manner in the setting where $\mathbf{Y}$ is one-dimensional.

1. Divide the observed $\mathbf{Y}$ values into, say, r "slices" where slice i has n_i observations. For example, the smallest n_1 of the $\mathbf{Y}$ form the first slice, the next n_2 form the second slice and so on.
2. Within the ith slice, compute the average, $\bar{\mathbf{X}}^{(i)}$, of the corresponding $\mathbf{X}$ values. Let

$$\hat{\mathbf{\Lambda}} = \sum_{i=1}^{r}(n_i/n)\hat{\mathbf{\Sigma}}_n^{-1/2}(\bar{\mathbf{X}}^{(i)} - \bar{\mathbf{X}}_n)(\bar{\mathbf{X}}^{(i)} - \bar{\mathbf{X}}_n)'(\hat{\mathbf{\Sigma}}_n^{-1/2})',$$

where $\bar{\mathbf{X}}_n$ is the overall sample mean and $\hat{\mathbf{\Sigma}}_n$ is the sample covariance matrix.

3. Estimate $\tilde{\mathscr{E}}$ with the basis comprising the eigenvectors that correspond to the K largest eigenvalues of $\hat{\mathbf{\Lambda}}$. Finally, estimate $\mathscr{E}$ by performing the transformation $\hat{\mathbf{\Sigma}}_n^{-1/2}$.

The present work is motivated by the following two issues. First it is pointed out in Li (1991) that the choice of slices is flexible. However, it is not clear how to characterize what constitutes a configuration that will lead to a good estimate. Second, when $\mathbf{Y}$ is multidimensional, to generalize the way in which the slices are defined from the one-dimensional setting is not entirely straightforward. See Cook (1995) and Li, Aragon and Thomas-Agnan (1994). In this paper, we consider a simple variation of the above procedure that is free of the problems mentioned.

Assume that $\mathbf{Y}$ has a continuous distribution to avoid ties. Unless otherwise noted, throughout this paper for each $i \in \{1, \dots, n\}$, let $i* \in \{1, \dots, n\} - \{i\}$ be the index for which

$$d(\mathbf{Y}_i, \mathbf{Y}_{i*}) = \min_{\substack{1 \le j \le n \\ j \ne i}} d(\mathbf{Y}_i, \mathbf{Y}_j),$$

where $d(\cdot, \cdot)$ is some metric. That is, $\mathbf{Y}_{i*}$ is the nearest neighbor of $\mathbf{Y}_i$. To simplify notation we will assume that this metric $d(\cdot, \cdot)$ is Euclidean, although, in sofar as proofs go, $d(\cdot, \cdot)$ only has to be a metric generated by some norm. Let $\mathbf{\Sigma}^{1/2}$ be any nonsingular matrix such that $\mathbf{\Sigma}^{1/2}(\mathbf{\Sigma}^{1/2})' = \operatorname{var}(\mathbf{X})$. Let $\hat{\mathbf{\Sigma}}_n^{-1/2}$ be a root-n consistent estimate of $\mathbf{\Sigma}^{-1/2}$ and

$$\hat{\mathbf{X}}_i = \hat{\mathbf{\Sigma}}_n^{-1/2}(\mathbf{X}_i - \bar{\mathbf{X}}_n).$$

Define

$$\mathbf{\Lambda}_n = (2n)^{-1} \sum_{i=1}^{n} \big(\hat{\mathbf{X}}_i \hat{\mathbf{X}}_{i*}' + \hat{\mathbf{X}}_{i*} \hat{\mathbf{X}}_i'\big). \tag{4}$$

Intuitions suggest that $\mathbf{\Lambda}_n$ estimates $\mathbf{\Lambda}$ and hence the e.d.r. space can be estimated following the steps in (3) above.

This paper focuses on the asymptotic properties of the above procedure. We will show that $\mathbf{\Lambda}_n - E\mathbf{\Lambda}_n = O_P(n^{-1/2})$ and that $\mathbf{\Lambda}_n$ is asymptotically normally distributed if the estimated mean and variance in $\mathbf{\Lambda}_n$ are replaced by the corresponding population versions. To address the rate of convergence of the e.d.r. space estimate, it seems that one could just conclude from the above-mentioned asymptotic results that the eigenvectors of $\mathbf{\Lambda}_n$ estimate those of $\mathbf{\Lambda}$ at rate $n^{-1/2}$. Unfortunately, while the differences between the eigenvectors of $\mathbf{\Lambda}_n$ and the matching eigenvectors of $E\mathbf{\Lambda}_n$ are $O_P(n^{-1/2})$, one cannot readily conclude that the same holds for $E\mathbf{\Lambda}_n$ and $\mathbf{\Lambda}$ for a general dimension $d_{\mathbf{y}}$. Indeed, for a sample of size n, the "typical distance" between an observation and its nearest neighbor in the d-dimensional space is $n^{-1/d}$. Hence one could expect the bias in our problem to be of rate $O(n^{-1/d_{\mathbf{y}}})$, implying that root-n consistency could not be achieved if $d_{\mathbf{y}} > 2$. Fortunately, estimating the eigenvectors of $\mathbf{\Lambda}$ is not the goal. Estimating the e.d.r. space is. To address the rate of convergence of an estimated space to a target space, it is only fitting to consider a proper distance between the spaces. Li (1991) addresses this by an R^2 statistic. Here we consider a distance which is somewhat different. We will show that the procedure outlined above is root-n consistent in this sense. The main results are formulated and stated in Section 2. The proofs are lengthy and are presented in the remaining sections.

It is interesting to mention the connection between the problem considered in this paper and certain problems in geometric probability. Recently there has been some considerable interest on limit theories for random graphs constructed according to different optimization algorithms. See Aldous and Steele (1992) and Avram and Bertsimas (1993) for some recent results and references. The way of grouping in (4) corresponds to the nearest neighbor graph. Other possibilities in that regard include the k-nearest neighbor graph, minimal spanning tree, sphere of influence graph, and so on. It is plausible to try to adapt those algorithms to the slicing step in SIR. On the other hand, the techniques developed in this paper to tackle the asymptotic theory of nearest neighbor inverse regression may be beneficial for some of the unresolved problems in the context of such random graphs.

2. Main results.

Throughout, let

$$F(\cdot) = P((\tilde{\mathbf{X}}, \mathbf{Y}) \in \cdot),\ \ \bar{F}(\cdot) = P((\tilde{\mathbf{X}}, \mathbf{Y}) \notin \cdot),\ \ G(\cdot) = P(\mathbf{Y} \in \cdot),\ \ \bar{G}(\cdot) = P(\mathbf{Y} \notin \cdot)$$

and, where applicable, let f be the joint density of $(\tilde{\mathbf{X}}, \mathbf{Y})$, g the marginal density of $\mathbf{Y}$ and $f(\cdot|\mathbf{y})$ be the conditional density of $\tilde{\mathbf{X}}$ given $\mathbf{Y} = \mathbf{y}$. For convenience of notation we denote $\mathbf{Z}_i = (\tilde{\mathbf{X}}_i, \mathbf{Y}_i)$ and $\mathbf{Z} = (\tilde{\mathbf{X}}, \mathbf{Y})$. Unless otherwise stated, vector and matrix norms will be Euclidean and denoted by $|\cdot|$. Finally, denote the sphere in $\mathbb{R}^{d_{\mathrm{y}}}$ which is centered at $\mathbf{u}$ with radius r by

$$S(\mathbf{u}; r) := \{\mathbf{v} \in \mathbb{R}^{d_{\mathrm{y}}}\colon |\mathbf{u} - \mathbf{v}| \le r\}.$$

As explained in Section 1, the standardized e.d.r. space $\tilde{\mathscr{E}}$ is estimated with the basis composed of the eigenvectors that correspond to the largest K eigenvalues of $\mathbf{\Lambda}_n$ where K is the dimension of $\tilde{\mathscr{E}}$. Hence it would be convenient to have a notation describing this operation. For a symmetric matrix $\mathbf{A}$ of dimension $p \times p$, let $\lambda_1(\mathbf{A}) \ge \cdots \ge \lambda_p(\mathbf{A})$ be the ordered eigenvalues of $\mathbf{A}$ and $\boldsymbol{\eta}_1(\mathbf{A}), \dots, \boldsymbol{\eta}_p(\mathbf{A})$ be a corresponding set of unit eigenvectors. Let $1 \le j < p$. Define

$$\mathscr{H}_j(\mathbf{A}) = \text{the linear space spanned by } \boldsymbol{\eta}_1(\mathbf{A}), \dots, \boldsymbol{\eta}_j(\mathbf{A}).$$

Note that if $\lambda_j(\mathbf{A}) > \lambda_{j+1}(\mathbf{A})$ then $\mathscr{H}_j(\mathbf{A})$ doesn't depend on the particular choice of eigenvectors.

With this, our estimates for $\tilde{\mathscr{E}}$ and $\mathscr{E}$ are $\mathscr{H}_K(\mathbf{\Lambda}_n)$ and $\hat{\mathbf{\Sigma}}_n^{-1/2}[\mathscr{H}_K(\mathbf{\Lambda}_n)]$, respectively. We wish to address the issue of the speed of convergence of these estimates to $\tilde{\mathscr{E}}$ and $\mathscr{E}$. In order to do so, we must first come up with a notion of distance between linear spaces. The following notion is reasonable although not unique. Let $\mathscr{S}_1$ and $\mathscr{S}_2$ be two Euclidean spaces. Define

$$\Delta(\mathscr{S}_1, \mathscr{S}_2) = \sup_{\mathbf{a}\in\mathscr{S}_1,\, |\mathbf{a}|=1} |\mathbf{a} - \boldsymbol{\pi}_{\mathscr{S}_2}(\mathbf{a})| + \sup_{\mathbf{b}\in\mathscr{S}_2,\, |\mathbf{b}|=1} |\mathbf{b} - \boldsymbol{\pi}_{\mathscr{S}_1}(\mathbf{b})|, \tag{5}$$

where $\boldsymbol{\pi}_{\mathscr{S}_1}$ and $\boldsymbol{\pi}_{\mathscr{S}_2}$ denote the projections on $\mathscr{S}_1$ and $\mathscr{S}_2$, respectively. It is straightforward to verify that Δ satisfies the triangle inequality. When this notion of distance is applied to the problem on hand, it is related to that of R^2 in Li (1991).

The first theorem states that our procedures give root-n rates of convergence in Δ under very general conditions.

THEOREM 1. *Assume that $E|\mathbf{X}|^{4+\varepsilon} < \infty$ for some $\varepsilon > 0$ and that*

$$\lim_{\delta\to 0} E\Big[\sup_{|\mathbf{Y}-\mathbf{y}|<\delta} |\boldsymbol{\nu}(\mathbf{Y}) - \boldsymbol{\nu}(\mathbf{y})|^2\Big] = 0. \tag{6}$$

Then both $\Delta(\mathscr{H}_K(\boldsymbol{\Lambda}_n), \tilde{\mathscr{E}})$ and $\Delta(\hat{\boldsymbol{\Sigma}}_n^{-1/2}[\mathscr{H}_K(\boldsymbol{\Lambda}_n)], \mathscr{E})$ are of rate $O_P(n^{-1/2})$.

The smoothness condition (6) is very weak indeed, which is satisfied for most situations. The proof of Theorem 1 will be given in Section 3. The proof is partially based on the idea that $\boldsymbol{\Lambda}_n$ can be approximated by the following quantity which is easier to analyze:

$$\tilde{\boldsymbol{\Lambda}}_n := (2n)^{-1} \sum_{i=1}^{n} \big(\tilde{\mathbf{X}}_i \tilde{\mathbf{X}}_{i*}' + \tilde{\mathbf{X}}_{i*} \tilde{\mathbf{X}}_i'\big),$$

where $\tilde{\mathbf{X}}_i$ is defined in (1). The following result concerning the variance of $\tilde{\boldsymbol{\Lambda}}_n$ is therefore instrumental.

THEOREM 2. *Suppose that $E|\mathbf{X}|^{4+\varepsilon} < \infty$ for some $\varepsilon > 0$. Then for any constants $\mathbf{a}_1, \mathbf{b}_1, \dots, \mathbf{a}_S, \mathbf{b}_S$ in $\mathbb{R}^{d_\mathbf{x}}$, there exists a bounded function γ such that for any Borel sets $A, A* \subset \mathbb{R}^{d_\mathbf{x}+d_\mathbf{y}}$ and any n,*

$$n^{-1} \operatorname{var}\bigg(\sum_{i=1}^{n} \sum_{s=1}^{S} \mathbf{a}_s' \tilde{\mathbf{X}}_i \tilde{\mathbf{X}}_{i*}' \mathbf{b}_s I(\mathbf{Z}_i \in A,\ \mathbf{Z}_{i*} \in A*) \bigg) \le \gamma(A, A*), \tag{7}$$

where γ has the continuity property that $\gamma(A_m, A_m*) \to 0$ for any sequence of Borel sets $\{A_m, A_m*\}$ for which $F(A_m) \wedge F(A_m*) \to 0$. In particular,

$$n \operatorname{var}\bigg(\sum_{s=1}^{S} \mathbf{a}_s' \tilde{\boldsymbol{\Lambda}}_n \mathbf{b}_s \bigg) < \gamma(\mathbb{R}^{d_\mathbf{x}+d_\mathbf{y}}, \mathbb{R}^{d_\mathbf{x}+d_\mathbf{y}}) < \infty.$$

The proof of Theorem 2 will be given in Section 4. The strength of this result is its generality. Note that the distributions are not even required to have densities for the result to be valid.

Alternatively to Theorem 2, the following central limit theorem can also be used to establish Theorem 1. While Theorem 1 does not give the precise asymptotic distribution of $\Delta(\mathscr{H}_K(\boldsymbol{\Lambda}_n), \tilde{\mathscr{E}})$, this central limit theorem offers some insight into how to go about that.

THEOREM 3. *Assume that* g *is continuous on* $(0 < g < \infty)$ *and for almost every* $\mathbf{x}$, $f(\mathbf{x}, \mathbf{y})$ *is continuous in* $\mathbf{y}$ *on* $(0 < g < \infty)$. *Also assume that* $E|\mathbf{X}|^{4+\varepsilon} < \infty$ *for some* $\varepsilon > 0$. *Then* $n^{1/2}(\tilde{\mathbf{\Lambda}}_n - E\tilde{\mathbf{\Lambda}}_n)$ *converges in distribution to some random matrix* $\mathbf{W}$ *where for any* $S = 1, 2 \ldots$ *and constants* $\mathbf{a}_1, \mathbf{b}_1, \ldots, \mathbf{a}_S, \mathbf{b}_S$ *in* $\mathbb{R}^{d_{\mathbf{x}}}$, $\sum_{s=1}^{S} \mathbf{a}_s' \mathbf{W} \mathbf{b}_s$ *is distributed as Normal* $(0, \sigma^2)$ *with* $\sigma^2 = \sum_{i=1}^{4} \kappa_i \omega_i$; *the* κ_i *are finite constants depending only on* $d_{\mathbf{y}}$ *while the* ω_i *are determined by* f *and the* $\mathbf{a}_i, \mathbf{b}_i$ *as*

$$\omega_1 = \int \psi^2(\mathbf{x}_1, \mathbf{x}_2) f(\mathbf{x}_1|\mathbf{y}) f(\mathbf{x}_2|\mathbf{y}) g(\mathbf{y})\, d\mathbf{x}_1\, d\mathbf{x}_2\, d\mathbf{y},$$

$$\omega_2 = \int \left(\int \psi(\mathbf{x}_1, \mathbf{x}_2) f(\mathbf{x}_2|\mathbf{y})\, d\mathbf{x}_2 \right)^2 f(\mathbf{x}_1|\mathbf{y}) g(\mathbf{y})\, d\mathbf{x}_1\, d\mathbf{y},$$

$$\omega_3 = \int \left(\int \psi(\mathbf{x}_1, \mathbf{x}_2) f(\mathbf{x}_1|\mathbf{y}) f(\mathbf{x}_2|\mathbf{y})\, d\mathbf{x}_1\, d\mathbf{x}_2 \right)^2 g(\mathbf{y})\, d\mathbf{y},$$

$$\omega_4 = \left(\int \psi(\mathbf{x}_1, \mathbf{x}_2) f(\mathbf{x}_1|\mathbf{y}) f(\mathbf{x}_2|\mathbf{y}) g(\mathbf{y})\, d\mathbf{x}_1\, d\mathbf{x}_2\, d\mathbf{y} \right)^2,$$

where $\psi(\mathbf{x}_1, \mathbf{x}_2) = (1/2) \sum_{s=1}^{S} \mathbf{a}_s' (\mathbf{x}_1 \mathbf{x}_2' + \mathbf{x}_2 \mathbf{x}_1') \mathbf{b}_s$.

Note that the precise expressions of the constants κ_i are delayed until Section 4 in (35)–(38) since they contribute little insight at this point. The proof of Theorem 3 is contained in Section 5.

We end this section with the remark that both Theorems 2 and 3 are easily extended to cover the partial sums of a class of functions of $(\mathbf{X}_i, \mathbf{Y}_i)$, $(\mathbf{X}_{i*}, \mathbf{Y}_{i*})$.

3. The speed of convergence of the e.d.r. space estimate. We need some lemmas first.

LEMMA 4. *Let* $\mathbf{A}, \mathbf{B}$ *be two symmetric matrices of dimension* $p \times p$, *each having* k *nonzero eigenvalues. Assume that* $\|\mathbf{A} - \mathbf{B}\| \le \delta$ *for some* δ *where* $\|\cdot\|$ *denotes the* sup *norm. Let* $\mathscr{S}_{\mathbf{A}}$ (*respectively*, $\mathscr{S}_{\mathbf{B}}$) *be the linear spaces spanned by the eigenvectors of* $\mathbf{A}$ (*respectively*, $\mathbf{B}$) *that correspond to* $\lambda_1(\mathbf{A}), \ldots, \lambda_k(\mathbf{A})$ [*respectively*, $\lambda_1(\mathbf{B}), \ldots, \lambda_k(\mathbf{B})$]. *Then*

$$\Delta(\mathscr{S}_{\mathbf{A}}, \mathscr{S}_{\mathbf{B}}) \le k^{3/2} \delta \{ |\lambda_k(\mathbf{A})|^{-1} + |\lambda_k(\mathbf{B})|^{-1} \}.$$

PROOF. Let $\mathbf{u}_1, \ldots, \mathbf{u}_p$ and $\mathbf{v}_1, \ldots, \mathbf{v}_p$ be two orthonormal bases of $\mathbb{R}^p$ for which $\mathbf{u}_1, \ldots, \mathbf{u}_k$ are the eigenvectors that correspond to $\lambda_1(\mathbf{A}), \ldots, \lambda_k(\mathbf{A})$ for $\mathbf{A}$ and $\mathbf{v}_1, \ldots, \mathbf{v}_k$ the eigenvectors that correspond to $\lambda_1(\mathbf{B}), \ldots, \lambda_k(\mathbf{B})$ for $\mathbf{B}$. First focus on $\sup_{\mathbf{a} \in \mathscr{S}_{\mathbf{A}},\, |\mathbf{a}|=1} |\mathbf{a} - \boldsymbol{\pi}_{\mathscr{S}_{\mathbf{B}}}(\mathbf{a})|$. Write

$$\mathbf{U}_1 = [\mathbf{u}_1, \ldots, \mathbf{u}_k], \qquad \mathbf{U}_2 = [\mathbf{u}_{k+1}, \ldots, \mathbf{u}_p]$$

and

$$\mathbf{V}_1 = [\mathbf{v}_1, \ldots, \mathbf{v}_k], \qquad \mathbf{V}_2 = [\mathbf{v}_{k+1}, \ldots, \mathbf{v}_p].$$

Since

$$\boldsymbol{\pi}_{\mathscr{S}_{\mathbf{B}}}(\mathbf{a}) = \mathbf{V}_1\mathbf{V}_1'\mathbf{a},$$

we obtain

$$\begin{aligned}\sup_{\mathbf{a}\in\mathscr{S}_{\mathbf{A}},\,|\mathbf{a}|=1} |\mathbf{a} - \boldsymbol{\pi}_{\mathscr{S}_{\mathbf{B}}}(\mathbf{a})| &= \sup_{\mathbf{a}\in\mathscr{S}_{\mathbf{A}},\,|\mathbf{a}|=1} |(\mathbf{I}_p - \mathbf{V}_1\mathbf{V}_1')\mathbf{a}| \\ &= \Big(\sup_{\mathbf{a}\in\mathscr{S}_{\mathbf{A}},\,|\mathbf{a}|=1} \mathbf{a}'(\mathbf{I}_p - \mathbf{V}_1\mathbf{V}_1')\mathbf{a}\Big)^{1/2},\end{aligned}$$

where $\mathbf{I}_p$ is the identity matrix of dimension p. Using the fact that

$$\mathbf{V}_1\mathbf{V}_1' + \mathbf{V}_2\mathbf{V}_2' = \mathbf{I}_p,$$

we get

$$\sup_{\mathbf{a}\in\mathscr{S}_{\mathbf{A}},\,|\mathbf{a}|=1} |\mathbf{a} - \boldsymbol{\pi}_{\mathscr{S}_{\mathbf{B}}}(\mathbf{a})| = \Big(\sup_{\mathbf{a}\in\mathscr{S}_{\mathbf{A}},\,|\mathbf{a}|=1} \mathbf{a}'\mathbf{V}_2\mathbf{V}_2'\mathbf{a}\Big)^{1/2}$$

By the fact that every unit vector in $\mathscr{S}_{\mathbf{A}}$ can be written as $\mathbf{U}_1\mathbf{x}$ for some unit vector $\mathbf{x} \in \mathbb{R}^k$, we conclude that

$$\begin{aligned}\sup\nolimits_{\mathbf{a}\in\mathscr{S}_{\mathbf{A}},\,|\mathbf{a}|=1} |\mathbf{a} - \boldsymbol{\pi}_{\mathscr{S}_{\mathbf{B}}}(\mathbf{a})| &= \Big(\sup_{\mathbf{x}\in\mathbb{R}^k,|\mathbf{x}|=1} \mathbf{x}'\mathbf{U}_1'\mathbf{V}_2\mathbf{V}_2'\mathbf{U}_1\mathbf{x}\Big)^{1/2} \\ &= \big(\text{largest eigenvalue of } \mathbf{U}_1'\mathbf{V}_2\mathbf{V}_2'\mathbf{U}_1\big)^{1/2}.\end{aligned} \tag{8}$$

Express $\mathbf{A}$ and $\mathbf{B}$ by the spectral decompositions

$$\mathbf{A} = \sum_{i=1}^{k} \lambda_i(\mathbf{A})\mathbf{u}_i\mathbf{u}_i' \quad \text{and} \quad \mathbf{B} = \sum_{i=1}^{k} \lambda_i(\mathbf{B})\mathbf{v}_i\mathbf{v}_i'.$$

For $1 \le i \le k$ and $k+1 \le j \le p$, it is clear that

$$|\mathbf{v}_j'\mathbf{u}_i| = |\mathbf{v}_j'(\mathbf{A} - \mathbf{B})\mathbf{u}_i|/|\lambda_i(\mathbf{A})| \le \delta/|\lambda_i(A)|.$$

Hence

$$|\mathbf{U}_1'\mathbf{V}_2\mathbf{V}_2'\mathbf{U}_1| \le k^3\delta^2/\lambda_k^2(\mathbf{A}).$$

Consequently, for each nonzero eigenvalue ν of $\mathbf{U}_1'\mathbf{V}_2\mathbf{V}_2'\mathbf{U}_1$ and a corresponding unit eigenvector $\mathbf{w}$,

$$|\nu| = |\mathbf{U}_1'\mathbf{V}_2\mathbf{V}_2'\mathbf{U}_1\mathbf{w}| \le k^3\delta^2/\lambda_k^2(\mathbf{A}).$$

In view of (8), we have shown that

$$\sup_{\mathbf{a}\in\mathscr{S}_{\mathbf{A}},\,|\mathbf{a}|=1} |\mathbf{a} - \boldsymbol{\pi}_{\mathscr{S}_{\mathbf{B}}}(\mathbf{a})| \le k^{3/2}\delta/|\lambda_k(\mathbf{A})|.$$

By symmetry, we also have

$$\sup_{\mathbf{b}\in\mathscr{S}_{\mathbf{B}},\,|\mathbf{b}|=1} |\mathbf{b} - \boldsymbol{\pi}_{\mathscr{S}_{\mathbf{A}}}(\mathbf{b})| \le k^{3/2}\delta/|\lambda_k(\mathbf{B})|. \qquad \square$$

LEMMA 5. *Under the assumptions of Theorem* 1, $E\tilde{\boldsymbol{\Lambda}}_n \to \boldsymbol{\Lambda}$ *as* $n \to \infty$.

PROOF. Assume for convenience that $\mathbf{Y}$ has a density g. It is clear that the pdf of $\mathbf{Y}_*$ is

$$g(\mathbf{y}_*) \int_{\mathbf{y}} (n-1)\bar{G}^{n-2}(S(\mathbf{y}; |\mathbf{y}-\mathbf{y}_*|))\, dG(\mathbf{y}),$$

which is bounded by

$$g(\mathbf{y}_*) \int_{\mathbf{y}} (n-1)\exp(-(n-2)G(S(\mathbf{y}; |\mathbf{y}-\mathbf{y}_*|)))\, dG(\mathbf{y}).$$

By Theorem A.1 of Bickel and Breiman (1983),

$$P(G(S(\mathbf{Y}; |\mathbf{Y}-\mathbf{y}|)) \le u) \le bu \tag{9}$$

for any $\mathbf{y} \in \mathbb{R}^{d_{\mathbf{y}}}$ and $u \in [0,1]$ where b is a constant depending only on $d_{\mathbf{y}}$. Hence for any $\mathbf{y}_*$,

$$\begin{aligned}
&\int_{\mathbf{y}} \exp(-(n-2)G(S(\mathbf{y}; |\mathbf{y}-\mathbf{y}_*|)))\, dG(\mathbf{y}) \\
&\qquad = \int_{u=0}^{1} P(\exp(-(n-2)G(S(\mathbf{Y}; |\mathbf{Y}-\mathbf{y}_*|))) > u)\, du \\
&\qquad \le \int_{u=0}^{1} P(G(S(\mathbf{Y}; |\mathbf{Y}-\mathbf{y}_*|)) < -(\log u)/(n-2))\, du \\
&\qquad \le b/(n-2) \int_{u=0}^{1} (-\log u)\, du = b/(n-2).
\end{aligned}$$

As a result, the pdf of $\mathbf{Y}_*$ is bounded by a constant multiple C of that of $\mathbf{Y}$. Hence

$$E[|\boldsymbol{\nu}(\mathbf{Y}_*)|^4] < CE[|\boldsymbol{\nu}(\mathbf{Y})|^4] < \infty. \tag{10}$$

Note that we assumed the existence of a density for convenience and that this argument will go through in general if we replace $g(\mathbf{y})$ by $G(d\mathbf{y})$. Now

$$\begin{aligned}
E\tilde{\boldsymbol{\Lambda}}_n - \boldsymbol{\Lambda} = {} & (E[\boldsymbol{\nu}(\mathbf{Y})\boldsymbol{\nu}(\mathbf{Y}_*)'] - E[\boldsymbol{\nu}(\mathbf{Y})\boldsymbol{\nu}(\mathbf{Y})'])/2 \\
& + (E[\boldsymbol{\nu}(\mathbf{Y}_*)\boldsymbol{\nu}(\mathbf{Y})'] - E[\boldsymbol{\nu}(\mathbf{Y})\boldsymbol{\nu}(\mathbf{Y})'])/2.
\end{aligned}$$

It suffices to deal with the first term. By the Cauchy–Schwarz inequality,

$$|E[\boldsymbol{\nu}(\mathbf{Y})\boldsymbol{\nu}(\mathbf{Y}_*)'] - E[\boldsymbol{\nu}(\mathbf{Y})\boldsymbol{\nu}(\mathbf{Y})']| \le E^{1/2}[|\boldsymbol{\nu}(\mathbf{Y})|^2]E^{1/2}[|\boldsymbol{\nu}(\mathbf{Y}_*) - \boldsymbol{\nu}(\mathbf{Y})|^2].$$

Now write

$$E[|\boldsymbol{\nu}(\mathbf{Y}_*) - \boldsymbol{\nu}(\mathbf{Y})|^2] = A_n(\delta) + B_n(\delta),$$

where $\delta > 0$,

$$A_n(\delta) = E[|\boldsymbol{\nu}(\mathbf{Y}_*) - \boldsymbol{\nu}(\mathbf{Y})|^2 I(|\mathbf{Y}_* - \mathbf{Y}| > \delta)]$$

and

$$B_n(\delta) = E[|\boldsymbol{\nu}(\mathbf{Y}_*) - \boldsymbol{\nu}(\mathbf{Y})|^2 I(|\mathbf{Y}_* - \mathbf{Y}| \le \delta)].$$

It is easy to show that $A_n(\delta) \to 0$ for each fixed δ by (10) and the Cauchy–Schwarz inequality and that

$$\lim_{\delta \to 0} \sup_n B_n(\delta) = 0$$

by (6). This completes the proof. □

PROOF OF THEOREM 1. We focus on $\Delta(\mathscr{H}_K(\mathbf{\Lambda}_n), \mathscr{E})$, the other one being similar. By the triangle inequality,

$$\Delta(\mathscr{H}_K(\mathbf{\Lambda}_n), \mathscr{E}) \le \Delta(\mathscr{H}_K(\mathbf{\Lambda}_n), \mathscr{H}_K(\tilde{\mathbf{\Lambda}}_n)) + \Delta(\mathscr{H}_K(\tilde{\mathbf{\Lambda}}_n), \mathscr{E}).$$

Our task is therefore to show that both terms on the right are $O_P(n^{-1/2})$. By the root-n rate of convergence of $\bar{\mathbf{X}}_n$ and $\hat{\mathbf{\Sigma}}_n^{-1/2}$, we conclude that

$$\mathbf{\Lambda}_n - \tilde{\mathbf{\Lambda}}_n = O_P(n^{-1/2}). \tag{11}$$

Clearly, $\mathbf{\Lambda}_n$ and $\tilde{\mathbf{\Lambda}}_n$ both converge in probability to $\mathbf{\Lambda}$ and hence the convergence of the eigenvalues follows. As a result, for any $\varepsilon \in (0, \lambda_K(\mathbf{\Lambda}))$, we have

$$P(\lambda_K(\mathbf{\Lambda}_n) > \varepsilon > \lambda_{K+1}(\mathbf{\Lambda}_n)) \to 1 \tag{12}$$

and the same can be said for $\tilde{\mathbf{\Lambda}}_n$. It follows simply [cf. Lemma 3.1 of Bai, Miao and Radhakrishna (1991)] from (11) and (12) that

$$\sum_{i=1}^{K} \lambda_i(\mathbf{\Lambda}_n)\boldsymbol{\eta}_i(\mathbf{\Lambda}_n)\boldsymbol{\eta}_i(\mathbf{\Lambda}_n)' - \sum_{i=1}^{K} \lambda_i(\tilde{\mathbf{\Lambda}}_n)\boldsymbol{\eta}_i(\tilde{\mathbf{\Lambda}}_n)\boldsymbol{\eta}_i(\tilde{\mathbf{\Lambda}}_n)' = O_P(n^{-1/2}),$$

where $\boldsymbol{\eta}_i$ was defined in the beginning of Section 2. Hence (12) and Lemma 4 imply that

$$\Delta(\mathscr{H}_K(\mathbf{\Lambda}_n), \mathscr{H}_K(\tilde{\mathbf{\Lambda}}_n)) = O_P(n^{-1/2}).$$

It follows from (2) that

$$E(\tilde{\mathbf{\Lambda}}_n)\mathbf{v} = \mathbf{0} \quad \text{for all } \mathbf{v} \in \mathscr{E}^{\perp}.$$

Conclude as in (3) using symmetry that

$$\mathscr{C}(E(\tilde{\mathbf{\Lambda}}_n)) \subset \mathscr{E}. \tag{13}$$

The question is whether the left-hand side of (13) can be a strict subset of $\mathscr{E}$. In view of (SIRc), this can happen only if

$$E(\tilde{\mathbf{\Lambda}}_n)\mathbf{v} = \mathbf{0} \tag{14}$$

for some eigenvector $\mathbf{v}$ corresponding to a nonzero eigenvalue of $\mathbf{\Lambda}$. Note that

$$E(\tilde{\mathbf{\Lambda}}_n)\mathbf{v} = \mathbf{\Lambda}\mathbf{v} + \mathbf{R}_n\mathbf{v},$$

where $\mathbf{R}_n \to \mathbf{0}$ by Lemma 5. Hence the scenario in (14) cannot happen for large n. Consequently, we conclude that for all large n,

$$\mathscr{C}(E(\tilde{\mathbf{\Lambda}}_n)) = \mathscr{H}_K(E(\tilde{\mathbf{\Lambda}}_n)) = \mathscr{E}$$

and as a result,

$$\Delta(\mathscr{H}_K(\tilde{\mathbf{\Lambda}}_n), \mathscr{E}) = \Delta(\mathscr{H}_K(\tilde{\mathbf{\Lambda}}_n), \mathscr{H}_K(E(\tilde{\mathbf{\Lambda}}_n))).$$

That the last quantity is $O_P(n^{-1/2})$ now follows straightforwardly from Theorem 2. This concludes the proof. □

4. Derivation of variance. This section contains the proof of Theorem 2. Throughout we will assume without loss of generality that the $\mathbf{X}_i$ are already standardized to have mean $\mathbf{0}$ and variance equal to the identity matrix. Below let γ denote a generic function with the continuity property described in the theorem but which may take various forms in different inequalities. In the end, the maximum of these γ functions will be the one in (7). Also for convenience of notation, write

$$\omega(\mathbf{z}_1, \mathbf{z}_2) := \omega(\mathbf{z}_1, \mathbf{z}_2; A, A*) := \sum_{s=1}^{S} \mathbf{a}_s' \mathbf{x}_1 \mathbf{x}_2' \mathbf{b}_s \; I(\mathbf{z}_1 \in A, \mathbf{z}_2 \in A*)$$

for $\mathbf{z}_1 = (\mathbf{x}_1, \mathbf{y}_1), \mathbf{z}_2 = (\mathbf{x}_2, \mathbf{y}_2) \in \mathbb{R}^{d_\mathbf{x}+d_\mathbf{y}}$ and Borel sets A, $A*$ of $\mathbb{R}^{d_\mathbf{x}+d_\mathbf{y}}$. Thus,

$$\sum_{i=1}^{n} \sum_{s=1}^{S} \mathbf{a}_s' \mathbf{X}_i \mathbf{X}_{i*}' \mathbf{b}_s I((\mathbf{X}_i, \mathbf{Y}_i) \in A, (\mathbf{X}_{i*}, \mathbf{Y}_{i*}) \in A*) = \sum_{i=1}^{n} \omega(\mathbf{Z}_i, \mathbf{Z}_{i*}).$$

By symmetry,

$$\begin{aligned} E\biggl(\sum_{i=1}^{n} \omega(\mathbf{Z}_i, \mathbf{Z}_{i*})\biggr)^2 &= \sum_{i=1}^{n} \sum_{j=1}^{n} E[\omega(\mathbf{Z}_i, \mathbf{Z}_{i*})\omega(\mathbf{Z}_j, \mathbf{Z}_{j*})] \\ &= nE[\omega^2(\mathbf{Z}_1, \mathbf{Z}_{1*})] + n(n-1)E[\omega(\mathbf{Z}_1, \mathbf{Z}_{1*})\omega(\mathbf{Z}_2, \mathbf{Z}_{2*})] \end{aligned}$$

and so

$$\begin{aligned} \operatorname{var}\biggl(\sum_{i=1}^{n} \omega(\mathbf{Z}_i, \mathbf{Z}_{i*})\biggr) &= M_1^{(2)}(n) + M_2^{(2)}(n) + M_3^{(2)}(n) + M_4^{(2)}(n) + M_5^{(2)}(n) \\ &\quad + \biggl(M_6^{(2)}(n) - E^2\biggl(\sum_{i=1}^{n} \omega(\mathbf{Z}_i, \mathbf{Z}_{i*})\biggr)\biggr), \end{aligned}$$

where

$$\begin{aligned} M_1^{(2)}(n) &= n(n-1)E[\omega^2(\mathbf{Z}_1, \mathbf{Z}_2)I(1* = 2)], \\ M_2^{(2)}(n) &= n(n-1)E[\omega^2(\mathbf{Z}_1, \mathbf{Z}_2)I(1* = 2, 2* = 1)], \\ M_3^{(2)}(n) &= n(n-1)(n-2)E[\omega(\mathbf{Z}_1, \mathbf{Z}_3)\omega(\mathbf{Z}_2, \mathbf{Z}_3)I(1* = 3,\ 2* = 3)], \\ M_4^{(2)}(n) &= n(n-1)(n-2)E[\omega(\mathbf{Z}_1, \mathbf{Z}_2)\omega(\mathbf{Z}_2, \mathbf{Z}_3)I(1* = 2,\ 2* = 3)], \end{aligned}$$

$$M_5^{(2)}(n) = n(n-1)(n-2)E[\omega(\mathbf{Z}_1, \mathbf{Z}_3)\omega(\mathbf{Z}_2, \mathbf{Z}_1)I(1* = 3,\ 2* = 1)],$$
$$M_6^{(2)}(n) = n(n-1)(n-2)(n-3)E[\omega(\mathbf{Z}_1, \mathbf{Z}_2)\omega(\mathbf{Z}_3, \mathbf{Z}_4)I(1* = 2,\ 3* = 4)].$$

It follows from Cauchy–Schwarz and the triangle inequalities that

$$\sum_{s=1}^{S} \mathbf{a}_s' \mathbf{x}_1 \mathbf{x}_2' \mathbf{b}_s \le \sum_{s=1}^{S} |\mathbf{a}_s||\mathbf{b}_s||\mathbf{x}_1||\mathbf{x}_2| \le C(|\mathbf{x}_1|^2 + |\mathbf{x}_2|^2) \tag{15}$$

for some constant $C < \infty$. Below C will denote a generic constant whose value changes from line to line. By (15), the arguments that led to (10) and Hölder's inequality,

$$\begin{aligned} M_1^{(2)}(n) &\le CnE[(|\mathbf{X}_1|^4 + |\mathbf{X}_{1*}|^4)I(\mathbf{Z}_1 \in A, \mathbf{Z}_{1*} \in A*)] \\ &\le CnE^{4/(4+\varepsilon)}[|\mathbf{X}_1|^{4+\varepsilon}]P^{\varepsilon/(4+\varepsilon)}(\mathbf{Z}_1 \in A, \mathbf{Z}_{1*} \in A*). \end{aligned}$$

Consequently,

$$n^{-1}M_1^{(2)}(n) \le \gamma(A, A*). \tag{16}$$

and, since $M_2^{(2)}(n) \le M_1^{(2)}(n)$,

$$n^{-1}M_2^{(2)}(n) \le \gamma(A, A*). \tag{17}$$

Now consider $M_2^{(3)}(n)$. Again, by (15),

$$\begin{aligned} M_3^{(2)}(n) &\le Cn(n-1)(n-2) \\ &\qquad \times E\big[(|\mathbf{X}_1|^2 + |\mathbf{X}_3|^2)(|\mathbf{X}_2|^2 + |\mathbf{X}_3|^2) \\ &\qquad\qquad \times I(1* = 3,\ 2* = 3,\ \mathbf{Z}_1, \mathbf{Z}_2 \in A,\ \mathbf{Z}_3 \in A*)\big] \\ &= Cn(n-1)(n-2) \\ &\qquad \times \big(E[|\mathbf{X}_1|^2|\mathbf{X}_2|^2 I(1* = 3,\ 2* = 3,\ \mathbf{Z}_1, \mathbf{Z}_2 \in A,\ \mathbf{Z}_3 \in A*)] \\ &\qquad\quad + 2E[|\mathbf{X}_1|^2|\mathbf{X}_3|^2 I(1* = 3,\ 2* = 3,\ \mathbf{Z}_1, \mathbf{Z}_2 \in A,\ \mathbf{Z}_3 \in A*)] \\ &\qquad\qquad + E[|\mathbf{X}_3|^4 I(1* = 3,\ 2* = 3,\ \mathbf{Z}_1, \mathbf{Z}_2 \in A,\ \mathbf{Z}_3 \in A*)]\big). \end{aligned} \tag{18}$$

First consider the leading term on the right of (18). By the Cauchy–Schwarz inequality and symmetry,

$$\begin{aligned} &E[|\mathbf{X}_1|^2|\mathbf{X}_2|^2 I(1* = 3,\ 2* = 3,\ \mathbf{Z}_1, \mathbf{Z}_2 \in A,\ \mathbf{Z}_3 \in A*)] \\ &\quad \le E[|\mathbf{X}_1|^4 I(1* = 3,\ 2* = 3,\ \mathbf{Z}_1 \in A,\ \mathbf{Z}_3 \in A*)] \\ &\quad \le E^{4/(4+\varepsilon)}[|\mathbf{X}_1|^{4+\varepsilon} I(1* = 3,\ 2* = 3,\ \mathbf{Z}_1 \in A)] \\ &\qquad \times P^{\varepsilon/(4+\varepsilon)}(1* = 3,\ 2* = 3,\ \mathbf{Z}_3 \in A*). \end{aligned} \tag{19}$$

The two terms in the product on the right-hand side are bounded slightly differently, depending on the order in which the variables are to be averaged. By the inequalities

$$1 - x \le e^{-x}, \ x > 0 \quad \text{and} \quad P(A \cup B) \ge (1/2)(P(A) + P(B)), \tag{20}$$

$$\begin{aligned}
&E[|\mathbf{X}_1|^{4+\varepsilon} I(1* = 3,\ 2* = 3,\ \mathbf{Z}_1 \in A)] \\
&\quad = \int_{\mathbf{z}_1 \in A} \int_{\mathbf{y}_3} \int_{\mathbf{y}_2} |\mathbf{x}_1|^{4+\varepsilon} \bar{G}^{n-3}\big(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_3|) \cup S(\mathbf{y}_2; |\mathbf{y}_2 - \mathbf{y}_3|)\big) \\
&\qquad\qquad \times dG(\mathbf{y}_2)\, dG(\mathbf{y}_3)\, dF(\mathbf{z}_1) \\
&\quad \le \int_{\mathbf{z}_1 \in A} \int_{\mathbf{y}_3} \int_{\mathbf{y}_2} |\mathbf{x}_1|^{4+\varepsilon} \exp\big(-nG(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_3|) \cup S(\mathbf{y}_2; |\mathbf{y}_2 - \mathbf{y}_3|))\big) \\
&\qquad\qquad \times dG(\mathbf{y}_2)\, dG(\mathbf{y}_3)\, dF(\mathbf{z}_1) \\
&\quad \le \int_{\mathbf{z}_1 \in A} \int_{\mathbf{y}_3} \int_{\mathbf{y}_2} |\mathbf{x}_1|^{4+\varepsilon} \exp(-(n/2)G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_3|))) \\
&\qquad\qquad \times \exp(-(n/2)G(S(\mathbf{y}_2; |\mathbf{y}_2 - \mathbf{y}_3|)))\, dG(\mathbf{y}_2)\, dG(\mathbf{y}_3)\, dF(\mathbf{z}_1).
\end{aligned}$$

By (9) and the derivations that immediately follow,

$$\int_{\mathbf{y}_2} \exp(-(n/2)G(S(\mathbf{y}_2; |\mathbf{y}_2 - \mathbf{y}_3|)))\, dG(\mathbf{y}_2) \le (2b/n).$$

Consequently,

$$\begin{aligned}
&E[|\mathbf{X}_1|^{4+\varepsilon} I(1* = 3,\ 2* = 3, \mathbf{Z}_1 \in A)] \\
&\quad \le (2b/n) \int_{\mathbf{z}_1 \in A} \int_{\mathbf{y}_3} |\mathbf{x}_1|^{4+\varepsilon} \exp(-(n/2)G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_3|)))\, dG(\mathbf{y}_3)\, dF(\mathbf{z}_1).
\end{aligned}$$

Clearly,

$$\begin{aligned}
&\int_{\mathbf{y}_3} \exp(-(n/2)G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_3|)))\, dG(\mathbf{y}_3) \\
&\quad = \int_{r=0}^{\infty} \exp(-(n/2)G(S(\mathbf{y}_1; r)))\, d_r G(S(\mathbf{y}_1; r)) \\
&\quad = \int_{u=0}^{1} \exp(-(n/2)u)\, du = 2/n.
\end{aligned}$$

Thus,

$$E[|\mathbf{X}_1|^{4+\varepsilon} I(1* = 3,\ 2* = 3,\ \mathbf{Z}_1 \in A)] \le (4b/n^2) E[|\mathbf{X}_1|^{4+\varepsilon} I(\mathbf{Z}_1 \in A)]. \tag{21}$$

By (9) and (20),

$$
\begin{aligned}
&P(1* = 3,\ 2* = 3,\ \mathbf{Z}_3 \in A*)\\
&\quad \le \int_{\mathbf{z}_3 \in A*} \int_{\mathbf{y}_1} \int_{\mathbf{y}_2} \exp(-(n/2)G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_3|)))\\
&\qquad\qquad \times \exp(-(n/2)G(S(\mathbf{y}_2; |\mathbf{y}_2 - \mathbf{y}_3|)))\, dG(\mathbf{y}_2)\, dG(\mathbf{y}_1)\, dF(\mathbf{z}_3)\\
&\quad = \int_{\mathbf{z}_3 \in A*} \left(\int_{\mathbf{y}} \exp(-(n/2)G(S(\mathbf{y}; |\mathbf{y} - \mathbf{y}_3|)))\, dG(\mathbf{y}) \right)^2 dF(\mathbf{z}_3)\\
&\quad \le (2b/n)^2 F(A*).
\end{aligned} \tag{22}
$$

The first term on the right of (18) is taken care of by (19), (21) and (22). It is clear that the other two terms on the right of (18) can be dealt with similarly, giving

$$
n^{-1} M_3^{(2)}(n) \le \gamma(A, A*). \tag{23}
$$

Next we consider $M_4^{(2)}(n)$ and $M_5^{(2)}(n)$, which are the same, and we use the notation of the former. By (15),

$$
\begin{aligned}
M_4^{(2)}(n) \le{}& Cn(n-1)(n-2)E[(|\mathbf{X}_1|^2 + |\mathbf{X}_2|^2)(|\mathbf{X}_2|^2 + |\mathbf{X}_3|^2)\\
&\times I(1* = 2,\ 2* = 3,\ \mathbf{Z}_1 \in A,\ \mathbf{Z}_2 \in A \cap A*,\ \mathbf{Z}_3 \in A*)].
\end{aligned}
$$

Repeated applications of the techniques used in the previous step give

$$
n^{-1}(M_4^{(2)}(n) + M_5^{(2)}(n)) \le \gamma(A, A*). \tag{24}
$$

Finally we consider the interplay between $M_6^{(2)}(n)$ and $E^2[\sum_{i=1}^n \omega(\mathbf{Z}_i, \mathbf{Z}_{i*})]$, which is the most crucial part of the computation of the variance. Define

$$
\xi(\mathbf{y}_1, \mathbf{y}_2) = E[\omega(\mathbf{Z}_1, \mathbf{Z}_2) | (\mathbf{Y}_1, \mathbf{Y}_2) = (\mathbf{y}_1, \mathbf{y}_2)].
$$

Then

$$
\begin{aligned}
M_6^{(2)}(n) ={}& n(n-1)(n-2)(n-3)\\
&\times \int \xi(\mathbf{y}_1, \mathbf{y}_2)\xi(\mathbf{y}_3, \mathbf{y}_4)\\
&\quad \times I(\mathbf{y}_1, \mathbf{y}_2 \notin S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|)) I(\mathbf{y}_3, \mathbf{y}_4 \notin S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\\
&\quad \times \bar{G}^{n-4}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|) \cup S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|))\\
&\quad \times dG(\mathbf{y}_1)\, dG(\mathbf{y}_2)\, dG(\mathbf{y}_3)\, dG(\mathbf{y}_4)\\
\le{}& n^2(n-1)^2\\
&\times \int \xi(\mathbf{y}_1, \mathbf{y}_2)\xi(\mathbf{y}_3, \mathbf{y}_4)\\
&\quad \times \bar{G}^{n-4}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|) \cup S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|))\\
&\quad \times dG(\mathbf{y}_1)\, dG(\mathbf{y}_2)\, dG(\mathbf{y}_3)\, dG(\mathbf{y}_4).
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
E^2\bigg(\sum_{i=1}^{n} \omega(\mathbf{Z}_i, \mathbf{Z}_{i*})\bigg)
&= n^2(n-1)^2 \int \xi(\mathbf{y}_1, \mathbf{y}_2)\xi(\mathbf{y}_3, \mathbf{y}_4) \\
&\qquad \times \bar{G}^{n-2}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\bar{G}^{n-2}(S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|)) \\
&\qquad \times dG(\mathbf{y}_1)\, dG(\mathbf{y}_2)\, dG(\mathbf{y}_3)\, dG(\mathbf{y}_4).
\end{aligned}
$$

Hence

$$
\begin{aligned}
M_6^{(2)}(n) - E^2\bigg(\sum_{i=1}^{n} \omega(\mathbf{Z}_i, \mathbf{Z}_{i*})\bigg)
&\le n^2(n-1)^2 \\
&\quad \times \int \xi(\mathbf{y}_1, \mathbf{y}_2)\xi(\mathbf{y}_3, \mathbf{y}_4) \\
&\qquad \times \big(\bar{G}^{n-4}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|) \cup S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|)) \\
&\qquad\quad - \bar{G}^{n-2}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\bar{G}^{n-2}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\big) \\
&\qquad \times dG(\mathbf{y}_1)\, dG(\mathbf{y}_2)\, dG(\mathbf{y}_3)\, dG(\mathbf{y}_4).
\end{aligned}
\tag{25}
$$

If $\bar{G}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) + \bar{G}(S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|)) < 1$ then clearly,

$$
\begin{aligned}
&\bar{G}^{n-2}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\bar{G}^{n-2}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) \\
&\qquad \ge \big(1 - G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) - G(S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|))\big)^{n-2},
\end{aligned}
$$

while the integral in (25) has an exponential rate, say $e^{-cn}\gamma(A, A*)$ for some $c > 0$ and some γ, when restricted to the set $\bar{G}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) + \bar{G}(S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|)) \ge 1$. Hence,

$$
\begin{aligned}
&M_6^{(2)}(n) - E^2\bigg(\sum_{i=1}^{n} \omega(\mathbf{Z}_i, \mathbf{Z}_{i*})\bigg) \\
&\quad \le e^{-cn}\gamma(A, A*) + n^2(n-1)^2 \\
&\qquad \times \int \xi(\mathbf{y}_1, \mathbf{y}_2)\xi(\mathbf{y}_3, \mathbf{y}_4) \\
&\qquad\quad \times I\big(G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) + \bar{G}(S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|)) < 1\big) \\
&\qquad\quad \times \big(\bar{G}^{n-4}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|) \cup S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|)) \\
&\qquad\qquad - (1 - G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) - G(S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|)))^{n-2}\big) \\
&\qquad\quad \times dG(\mathbf{y}_1)\, dG(\mathbf{y}_2)\, dG(\mathbf{y}_3)\, dG(\mathbf{y}_4).
\end{aligned}
$$

Since we also have

$$
\begin{aligned}
&\bar{G}^{n-4}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|) \cup S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|)) \\
&\qquad \ge \big(1 - G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) - G(S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|))\big)^{n-2},
\end{aligned}
$$

upon applying the Cauchy–Schwarz inequality and symmetry,

$$\begin{aligned} M_6^{(2)}(n) - E^2\biggl(\sum_{i=1}^{n} \omega(\mathbf{Z}_i, \mathbf{Z}_{i*})\biggr) \\ \le e^{-cn}\gamma(A, A*) + n^2(n-1)^2[A_{n,1} + (A_{n,2} - B_n)], \end{aligned} \tag{26}$$

where

$$\begin{aligned} A_{n,1} = \int & \xi^2(\mathbf{y}_1, \mathbf{y}_2) I(\mathbf{y}_4 \in S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) \\ & \times \bar{G}^{n-4}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|) \cup S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|)) \\ & \times dG(\mathbf{y}_1)\, dG(\mathbf{y}_2)\, dG(\mathbf{y}_3)\, dG(\mathbf{y}_4), \\ A_{n,2} = \int & \xi^2(\mathbf{y}_1, \mathbf{y}_2) I(\mathbf{y}_4 \notin S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) \\ & \times \bar{G}^{n-4}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|) \cup S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|)) \\ & \times dG(\mathbf{y}_1)\, dG(\mathbf{y}_2)\, dG(\mathbf{y}_3)\, dG(\mathbf{y}_4) \end{aligned}$$

and

$$\begin{aligned} B_n = \int & \xi^2(\mathbf{y}_1, \mathbf{y}_2) I\bigl(G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) + \bar{G}(S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|)) < 1\bigr) \\ & \times \bigl(1 - G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) - G(S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|))\bigr)^{n-2} \\ & \times dG(\mathbf{y}_1)\, dG(\mathbf{y}_2)\, dG(\mathbf{y}_3)\, dG(\mathbf{y}_4). \end{aligned}$$

Using (9) and (20) as before, we obtain

$$\begin{aligned} A_{n,1} \le & \int_{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_4} \xi^2(\mathbf{y}_1, \mathbf{y}_2) I(\mathbf{y}_4 \in S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) \\ & \times \exp\biggl(-\frac{n-4}{2} G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\biggr) \\ & \times \biggl(\int_{\mathbf{y}_3} \exp\biggl(-\frac{n-4}{2} G(S(\mathbf{y}_3; |\mathbf{y}_3 - \mathbf{y}_4|))\biggr) dG(\mathbf{y}_3)\biggr) \\ & \times dG(\mathbf{y}_1)\, dG(\mathbf{y}_2)\, dG(\mathbf{y}_4) \\ \le & \frac{2b}{n-4} \int_{\mathbf{y}_1, \mathbf{y}_2} \xi^2(\mathbf{y}_1, \mathbf{y}_2) G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) \\ & \qquad \times \exp\biggl(-\frac{n-4}{2} G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\biggr) dG(\mathbf{y}_1)\, dG(\mathbf{y}_2) \qquad (27) \\ \le & \frac{2b}{n-4} \int_{\mathbf{y}_1, \mathbf{y}_2} \eta_1(\mathbf{y}_1, A)\eta_2(\mathbf{y}_2, A*) G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) \\ & \qquad \times \exp\biggl(-\frac{n-4}{2} G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\biggr) dG(\mathbf{y}_1)\, dG(\mathbf{y}_2) \\ & + \frac{2b}{n-4} \int_{\mathbf{y}_1, \mathbf{y}_2} \eta_2(\mathbf{y}_1, A)\eta_1(\mathbf{y}_2, A*) G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) \\ & \qquad \times \exp\biggl(-\frac{n-4}{2} G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\biggr) dG(\mathbf{y}_1)\, dG(\mathbf{y}_2), \end{aligned}$$

where

$$\eta_1(\mathbf{y}, B) = E[|\mathbf{X}|^4 I(\mathbf{Z} \in B)|\mathbf{Y} = \mathbf{y}], \qquad \eta_2(\mathbf{y}, B) = P(\mathbf{Z} \in B|\mathbf{Y} = \mathbf{y}).$$

For any $\varepsilon > 0$, by the Cauchy–Schwarz inequality

$$\begin{aligned}
&\int_{\mathbf{y}_1, \mathbf{y}_2} \eta_1(\mathbf{y}_1, A)\eta_2(\mathbf{y}_2, A*)G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) \\
&\quad \times \exp\left(-\frac{n-4}{2}G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\right) dG(\mathbf{y}_1)\, dG(\mathbf{y}_2) \\
&\qquad \le \left(\int_{\mathbf{y}_1, \mathbf{y}_2} \eta_1^{1+\varepsilon}(\mathbf{y}_1, A)G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) \right. \\
&\qquad\qquad \left. \times \exp\left(-\frac{n-4}{2}G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\right) dG(\mathbf{y}_1)\, dG(\mathbf{y}_2)\right)^{1/(1+\varepsilon)} \\
&\qquad\quad \times \left(\int_{\mathbf{y}_1, \mathbf{y}_2} \eta_2^{(1+\varepsilon)/\varepsilon}(\mathbf{y}_2, A*)G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) \right. \\
&\qquad\qquad \left. \times \exp\left(-\frac{n-4}{2}G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\right) dG(\mathbf{y}_1)\, dG(\mathbf{y}_2)\right)^{\varepsilon/(1+\varepsilon)}.
\end{aligned}$$

Clearly,

$$\begin{aligned}
&\int_{\mathbf{y}_1, \mathbf{y}_2} \eta_1^{1+\varepsilon}(\mathbf{y}_1, A)G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) \\
&\quad \times \exp\left(-\frac{n-4}{2}G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\right) dG(\mathbf{y}_1)\, dG(\mathbf{y}_2) \\
&\qquad = \int_{\mathbf{y}_1} \eta_1^{1+\varepsilon}(\mathbf{y}_1, A)\left[\int_{u=0}^{1} u \exp\left(-\frac{n-4}{2}u\right) du\right] dG(\mathbf{y}_1) \\
&\qquad = O(n^{-2})E[\eta_1^{1+\varepsilon}(\mathbf{Y}, A)].
\end{aligned}$$

Similarly, for any $p, q > 0,\ p + q = 1$,

$$\begin{aligned}
&\int_{\mathbf{y}_1, \mathbf{y}_2} \eta_2^{p(1+\varepsilon)/\varepsilon}(\mathbf{y}_2, A*)G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) \\
&\quad \times \exp\left(-\frac{n-4}{2}G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\right) dG(\mathbf{y}_1)\, dG(\mathbf{y}_2) \\
&\qquad \le \left(\int_{\mathbf{y}_1, \mathbf{y}_2} \eta_2^{p(1+\varepsilon)/\varepsilon}(\mathbf{y}_2, A*) \right. \\
&\qquad\qquad \left. \times \exp\left(-\frac{n-4}{2}G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\right) dG(\mathbf{y}_1)\, dG(\mathbf{y}_2)\right)^{1/p} \\
&\qquad\quad \times \left(\int_{\mathbf{y}_1, \mathbf{y}_2} G^q(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) \right. \\
&\qquad\qquad \left. \times \exp\left(-\frac{n-4}{2}G(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\right) dG(\mathbf{y}_1)\, dG(\mathbf{y}_2)\right)^{1/q} \\
&\qquad = O(n^{-2})E^{1/p}[\eta_2^{p(1+\varepsilon)/\varepsilon}(\mathbf{Y}, A*)].
\end{aligned}$$

This together with the previous estimate give a bound of the first term in (27). The second term there is bounded by the same principle. Hence we obtain

$$n^3 A_{n,1} \le \gamma(A, A*). \tag{28}$$

Next, straightforward computations show that

$$B_n = \frac{1}{n-1}\int_{\mathbf{y}_1,\mathbf{y}_2} \xi^2(\mathbf{y}_1,\mathbf{y}_2)\bar{G}^{n-1}(S(\mathbf{y}_1;|\mathbf{y}_1-\mathbf{y}_2|))\,dG(\mathbf{y}_1)\,dG(\mathbf{y}_2). \tag{29}$$

Next note that

$$\begin{aligned}
A_{n,2} &= \int_{\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3} \xi^2(\mathbf{y}_1,\mathbf{y}_2)\\
&\qquad \times \biggl(\int_{\mathbf{y}_4} I(\mathbf{y}_4 \notin S(\mathbf{y}_1;|\mathbf{y}_1-\mathbf{y}_2|))\\
&\qquad\qquad \times \bar{G}^{n-4}(S(\mathbf{y}_1;|\mathbf{y}_1-\mathbf{y}_2|)\cup S(\mathbf{y}_3;|\mathbf{y}_3-\mathbf{y}_4|))\,dG(\mathbf{y}_4)\biggr)\\
&\qquad \times dG(\mathbf{y}_1)\,dG(\mathbf{y}_2)\,dG(\mathbf{y}_3)\\
&= \int_{\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3} \xi^2(\mathbf{y}_1,\mathbf{y}_2)\biggl(\int_{r=0}^{\infty} \bar{G}^{n-4}(S(\mathbf{y}_1;|\mathbf{y}_1-\mathbf{y}_2|)\cup S(\mathbf{y}_3;r))\\
&\qquad\qquad \times d_r G(S(\mathbf{y}_1;|\mathbf{y}_1-\mathbf{y}_2|)\cup S(\mathbf{y}_3;r))\biggr)\\
&\qquad \times dG(\mathbf{y}_1)\,dG(\mathbf{y}_2)\,dG(\mathbf{y}_3)\\
&= \int_{\mathbf{y}_1,\mathbf{y}_2,\mathbf{y}_3} \xi^2(\mathbf{y}_1,\mathbf{y}_2)\biggl(\int_{u=0}^{\bar{G}(S(\mathbf{y}_1;|\mathbf{y}_1-\mathbf{y}_2|))} u^{n-4}\,du\biggr)\\
&\qquad \times dG(\mathbf{y}_1)\,dG(\mathbf{y}_2)\,dG(\mathbf{y}_3)\\
&= \frac{1}{n-3}\int_{\mathbf{y}_1,\mathbf{y}_2} \xi^2(\mathbf{y}_1,\mathbf{y}_2)\bar{G}^{n-3}(S(\mathbf{y}_1;|\mathbf{y}_1-\mathbf{y}_2|))\,dG(\mathbf{y}_1)\,dG(\mathbf{y}_2).
\end{aligned} \tag{30}$$

Now by (29), (30) and the techniques leading to (28),

$$n^3(A_{n,2} - B_n) \le \gamma(A, A*). \tag{31}$$

By (26), (28) and (31), we obtain

$$n^{-1}\biggl(M_6^{(2)}(n) - E^2\biggl(\sum_{i=1}^{n} \omega(\mathbf{Z}_i,\mathbf{Z}_{i*})\biggr)\biggr) \le \gamma(A, A*). \tag{32}$$

Hence (7) follows from (16), (17), (23), (24) and (32). This concludes the proof of Theorem 2. □

5. Central limit theorem. The rest of this paper is devoted to proving Theorem 3. This section gives the outline of the proof and the remaining sections contain the various technical details. The complete proof is rather lengthy and is best broken into a number of components as is done below. To facilitate the proof it is crucial to have a set of clear and flexible notation. First it is obvious that we can assume without loss of generality that $\mathbf{X}$ is already standardized. Thus, $\mathbf{Z}_i = (\mathbf{X}_i, \mathbf{Y}_i)$, $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ and F and G are the probability measures corresponding to $\mathbf{Z}$ and $\mathbf{Y}$, respectively. Let $\mathbf{z}_i = (\mathbf{x}_i, \mathbf{y}_i)$, $1 \le i \le m$, be vectors where $\mathbf{x}_i$ and $\mathbf{y}_i$ are of dimensions $d_{\mathbf{x}}$ and $d_{\mathbf{y}}$, respectively, and the distances $|\mathbf{y}_i - \mathbf{y}_j|$, $1 \le i \ne j \le m$, are all distinct. From now on, fix a set of constants $\mathbf{a}_1, \mathbf{b}_1, \dots, \mathbf{a}_S, \mathbf{b}_S$ in $\mathbb{R}^{d_{\mathbf{x}}}$. For measurable sets $A, A* \subset \mathbb{R}^{d_{\mathbf{x}}+d_{\mathbf{y}}}$, define

$$\mathscr{K}(A, A*; \{\mathbf{z}_i, 1 \le i \le m\}) = \frac{1}{2}\sum_{i=1}^{m}\sum_{s=1}^{S} \mathbf{a}_s'(\mathbf{x}_i\mathbf{x}_{i*}' + \mathbf{x}_{i*}\mathbf{x}_i')\mathbf{b}_s I(\mathbf{z}_i \in A,\ \mathbf{z}_{i*} \in A*),$$

where $i*$ is the index for which $\mathbf{y}_{i*}$ is the nearest neighbor of $\mathbf{y}_i$, that is, for which

$$|\mathbf{y}_i - \mathbf{y}_{i*}| = \min_{\substack{1\le j\le m\\ j\ne i}} |\mathbf{y}_i - \mathbf{y}_j|.$$

Also define

$$\mathscr{I}_n = \{\mathbf{Z}_i, 1 \le i \le n\}$$

so that we can write

$$\sum_{s=1}^{S} \mathbf{a}_s'\tilde{\mathbf{\Lambda}}_n\mathbf{b}_s = \frac{1}{n}\mathscr{K}(\mathbb{R}^{d_{\mathbf{x}}+d_{\mathbf{y}}}, \mathbb{R}^{d_{\mathbf{x}}+d_{\mathbf{y}}}; \mathscr{I}_n).$$

As the notation will become rather complicated, let us make a minor simplification by assuming that the density f of $\mathbf{Z}$ is bounded. It will be clear that if this is not the case, we can do a truncation by attaching an indicator to every relevant expectation and then the proof will go through in more or less the same way as for the case where f is bounded. Since

$$1 = G(\{\mathbf{y}\colon g(\mathbf{y}) > 0\}) = \lim_{\varepsilon\to 0} G(\{\mathbf{y}\colon \varepsilon < g(\mathbf{y}) < \varepsilon^{-1}\}),$$

it follows that for each $\delta > 0$ there exists some $\varepsilon \in (0, 1)$ such that

$$\bar{G}(\{\mathbf{y}\colon \varepsilon < g(\mathbf{y}) < \varepsilon^{-1}\}) < \delta/4.$$

Since $\{\mathbf{y}\colon \varepsilon < g(\mathbf{y}) < \varepsilon^{-1}\}$ is open, it can be written as a countable union of bounded open rectangles, sets of the form $(e_1, f_1) \times \cdots \times (e_{d_{\mathbf{y}}}, f_{d_{\mathbf{y}}})$ for finite e_i, f_i. Then for any $\delta > 0$ it can be selected from these a finite set of bounded open rectangles whose union we denote by C such that

$$0 \le \bar{G}(C) - \bar{G}(\{\mathbf{y}\colon \varepsilon < g(\mathbf{y}) < \varepsilon^{-1}\}) < \delta/4.$$

As a consequence, for every $\delta > 0$ there exists a set C which is the union of a finite number of bounded rectangles in $\mathbb{R}^{d_{\mathbf{y}}}$ such that

$$\bar{G}(C) < \delta/2 \quad \text{and} \quad 0 < \inf_{\mathbf{y}\in C} g(\mathbf{y}) \le \sup_{\mathbf{y}\in C} g(\mathbf{y}) < \infty.$$

Now take a bounded Borel set $B \in \mathbb{R}^{d_{\mathbf{x}}}$ so that $P(\mathbf{X} \notin B) < \delta/2$ and write $A = B \times C$. Clearly,

$$\bar{F}(A) = P(\mathbf{Z} \notin B \times C) \le P(\mathbf{X} \notin B) + P(\mathbf{Y} \notin C) < \delta. \tag{33}$$

Write

$$n^{1/2} \sum_{s=1}^{S} \mathbf{a}_s' \tilde{\mathbf{\Lambda}}_n \mathbf{b}_s = \frac{1}{n^{1/2}} \Big[\mathscr{K}(A, A; \mathscr{I}_n) + \mathscr{K}(A^c, A; \mathscr{I}_n) + \mathscr{K}(A, A^c; \mathscr{I}_n) + \mathscr{K}(A^c, A^c; \mathscr{I}_n) \Big].$$

By Theorem 2 and (33), the variance of

$$n^{-1/2} \Big[\mathscr{K}(A^c, A; \mathscr{I}_n) + \mathscr{K}(A, A^c; \mathscr{I}_n) + \mathscr{K}(A^c, A^c; \mathscr{I}_n) \Big]$$

can be made as small as desired by choosing a small enough δ. Thus, the central limit theorem follows if we show

$$n^{-1/2}\big(\mathscr{K}(A, A; \mathscr{I}_n) - E\mathscr{K}(A, A; \mathscr{I}_n)\big) \longrightarrow_d \text{Normal } (0, \sigma^2), \tag{34}$$

where

$$\sigma^2 = \sum_{i=1}^{4} \kappa_i \omega_i(A)$$

and where, with

$$S(\mathbf{y}_1, \mathbf{y}_2; r_1, r_2) := S(\mathbf{y}_1; r_1) \cup S(\mathbf{y}_2; r_2), \qquad \mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^{d_{\mathbf{y}}},\ r_1, r_2 > 0,$$
$$\|S\| := \textstyle\int_S d\mathbf{y}, \qquad S \subset \mathbb{R}^{d_{\mathbf{y}}}$$

and ψ as defined in Theorem 3, the κ_i and $\omega_i(A)$ are given by

$$\kappa_1 = 1 + \int \exp(-\|S(0, \boldsymbol{\varepsilon}; |\boldsymbol{\varepsilon}|, |\boldsymbol{\varepsilon}|)\|)\, d\boldsymbol{\varepsilon}, \tag{35}$$

$$\begin{aligned} \kappa_2 = {} & \int I(|\boldsymbol{\varepsilon}_{13} - \boldsymbol{\varepsilon}_{23}| > |\boldsymbol{\varepsilon}_{13}| \vee |\boldsymbol{\varepsilon}_{23}|) \\ & \quad \times \exp(-\|S(\boldsymbol{\varepsilon}_{13}, \boldsymbol{\varepsilon}_{23}; |\boldsymbol{\varepsilon}_{13}|, |\boldsymbol{\varepsilon}_{23}|)\|)\, d\boldsymbol{\varepsilon}_{13}\, d\boldsymbol{\varepsilon}_{23} \\ & + 2\int I(|\boldsymbol{\varepsilon}_{12} - \boldsymbol{\varepsilon}_{23}| > |\boldsymbol{\varepsilon}_{12}| > |\boldsymbol{\varepsilon}_{23}|) \\ & \quad \times \exp(-\|S(0, \boldsymbol{\varepsilon}_{12}; |\boldsymbol{\varepsilon}_{12}|, |\boldsymbol{\varepsilon}_{23}|)\|)\, d\boldsymbol{\varepsilon}_{12}\, d\boldsymbol{\varepsilon}_{23}, \end{aligned} \tag{36}$$

$$
\begin{aligned}
\kappa_3 = -\int_{\boldsymbol{\varepsilon}_{12}}\int_{\boldsymbol{\varepsilon}_{13}}\int_{\boldsymbol{\varepsilon}_{24}} & I\big(|\boldsymbol{\varepsilon}_{13}| > |\boldsymbol{\varepsilon}_{12}| \wedge |\boldsymbol{\varepsilon}_{12} - \boldsymbol{\varepsilon}_{24}| \\
&\qquad \text{or } |\boldsymbol{\varepsilon}_{24}| > |\boldsymbol{\varepsilon}_{12}| \wedge |\boldsymbol{\varepsilon}_{12} - \boldsymbol{\varepsilon}_{13}| \\
&\qquad\qquad \text{or } |\boldsymbol{\varepsilon}_{12}| < |\boldsymbol{\varepsilon}_{13}| + |\boldsymbol{\varepsilon}_{24}|\big) \\
&\qquad \times \exp(-\|S(0; |\boldsymbol{\varepsilon}_{13}|)\| - \|S(0; |\boldsymbol{\varepsilon}_{24}|)\|)\, d\boldsymbol{\varepsilon}_{12}\, d\boldsymbol{\varepsilon}_{13}\, d\boldsymbol{\varepsilon}_{24} \\
+ \int_{\boldsymbol{\varepsilon}_{12}}\int_{\boldsymbol{\varepsilon}_{13}}\int_{\boldsymbol{\varepsilon}_{24}} & I\big(|\boldsymbol{\varepsilon}_{13}| < |\boldsymbol{\varepsilon}_{12}| \wedge |\boldsymbol{\varepsilon}_{12} - \boldsymbol{\varepsilon}_{24}|, \\
&\qquad |\boldsymbol{\varepsilon}_{24}| < |\boldsymbol{\varepsilon}_{12}| \wedge |\boldsymbol{\varepsilon}_{12} - \boldsymbol{\varepsilon}_{13}|, \\
&\qquad\qquad |\boldsymbol{\varepsilon}_{12}| \le |\boldsymbol{\varepsilon}_{13}| + |\boldsymbol{\varepsilon}_{24}|\big) \\
&\qquad \times \exp(-\|S(0, \boldsymbol{\varepsilon}_{12}; |\boldsymbol{\varepsilon}_{13}|, |\boldsymbol{\varepsilon}_{24}|)\|)\, d\boldsymbol{\varepsilon}_{12}\, d\boldsymbol{\varepsilon}_{13}\, d\boldsymbol{\varepsilon}_{24},
\end{aligned} \tag{37}
$$

$$
\kappa_4 = -1 \tag{38}
$$

and

$$
\begin{aligned}
\omega_1(A) &= \int \psi^2(\mathbf{x}_1, \mathbf{x}_2) f(\mathbf{x}_1|\mathbf{y}) f(\mathbf{x}_2|\mathbf{y}) \\
&\qquad \times I(\mathbf{x}_1, \mathbf{x}_2 \in B,\ \mathbf{y} \in C) g(\mathbf{y})\, d\mathbf{x}_1\, d\mathbf{x}_2\, d\mathbf{y}, \\
\omega_2(A) &= \int \psi(\mathbf{x}_1, \mathbf{x}_2)\psi(\mathbf{x}_1, \mathbf{x}_3) I(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \in B,\ \mathbf{y} \in C) \\
&\qquad \times f(\mathbf{x}_1|\mathbf{y}) f(\mathbf{x}_2|\mathbf{y}) f(\mathbf{x}_3|\mathbf{y}) g(\mathbf{y})\, d\mathbf{x}_1\, d\mathbf{x}_2\, d\mathbf{x}_3\, d\mathbf{y}, \\
\omega_3(A) &= \int \bigg(\int \psi(\mathbf{x}_1, \mathbf{x}_2) I(\mathbf{x}_1, \mathbf{x}_2 \in B,\ \mathbf{y} \in C) \\
&\qquad\qquad \times f(\mathbf{x}_1|\mathbf{y}) f(\mathbf{x}_2|\mathbf{y})\, d\mathbf{x}_1\, d\mathbf{x}_2 \bigg)^2 g(\mathbf{y})\, d\mathbf{y}, \\
\omega_4(A) &= \bigg(\int \psi(\mathbf{x}_1, \mathbf{x}_2) I(\mathbf{x}_1, \mathbf{x}_2 \in B,\ \mathbf{y} \in C) \\
&\qquad \times f(\mathbf{x}_1|\mathbf{y}) f(\mathbf{x}_2|\mathbf{y}) g(\mathbf{y})\, d\mathbf{x}_1\, d\mathbf{x}_2\, d\mathbf{y} \bigg)^2.
\end{aligned}
$$

Here the $\omega_i(A)$ are truncated versions of the ω_i in Theorem 3. Also note that in the κ_i, the purpose for the particular subscripts for the dummy variables $\boldsymbol{\varepsilon}$'s is to keep track of how the various κ_i arise in the proof.

The proof of (34) goes as follows. One of the novelties here is a coupling argument which simplifies the proof. For each $n \ge 1$, let P_n be a Poisson random variable with mean n and independent of all the $\mathbf{Z}_i$. Define

$$
\mathscr{P}_n = \mathscr{I}_{P_n} = \{\mathbf{Z}_1, \dots, \mathbf{Z}_{P_n}\}.
$$

Thus, $\mathscr{P}_n$ is a Poisson process with intensity measure $n\int_{\bullet} f(\mathbf{z})\, d\mathbf{z}$. Also let

$$
\tau(m) = E\mathscr{K}(A, A; \mathscr{I}_m).
$$

Define

$$\sigma_1^2 = \kappa_1\omega_1(A) + \kappa_2\omega_2(A) + (\kappa_3 - 1)\omega_3(A),$$
$$\sigma_2^2 = \omega_3(A) \text{ and } \sigma_3^2 = \omega_4(A). \tag{39}$$

Note that

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 - \sigma_3^2.$$

Write

$$\begin{aligned} &\frac{\mathscr{K}(A, A; \mathscr{I}_n) - E\mathscr{K}(A, A; \mathscr{I}_n)}{\sqrt{n}} + \frac{\tau(P_n) - \tau(n)}{\sqrt{n}} \\ &\quad = \frac{\mathscr{K}(A, A; \mathscr{P}_n) - E\mathscr{K}(A, A; \mathscr{P}_n)}{\sqrt{n}} \\ &\qquad - \frac{\mathscr{K}(A, A; \mathscr{P}_n) - \mathscr{K}(A, A; \mathscr{I}_n) - (\tau(P_n) - \tau(n))}{\sqrt{n}} \\ &\qquad + \frac{E\mathscr{K}(A, A; \mathscr{P}_n) - E\mathscr{K}(A, A; \mathscr{I}_n)}{\sqrt{n}}. \end{aligned} \tag{40}$$

The first term on the right-hand side of (40) converges in distribution to Normal $(0, \sigma_1^2 + \sigma_2^2)$ by Proposition 6. The second term on the right converges to 0 in probability by Proposition 9 and the fact that $P_n - n = O_p(n^{1/2})$. The third term on the right converges to 0 by a straightforward computation (cf. Lemma 8). Note that the l.h.s. of (40) is the sum of two independent random variables by the definition of P_n. Since Proposition 10 shows that $[\tau(P_n) - \tau(n)]/\sqrt{n}$ converges in distribution to Normal $(0, \sigma_3^2)$, the proof of (34) follows.

6. Blocking. The main purpose of this section is to prove a central limit theorem for the first term on the right of (40). Therefore we continue to work in that setting and use the notation defined there.

PROPOSITION 6. *Suppose that* $A = B \times C \in \mathbb{R}^{d_{\mathbf{x}}+d_{\mathbf{y}}}$ *where* B *is a bounded Borel set in* $\mathbb{R}^{d_{\mathbf{x}}}$ *and* $C \in \mathbb{R}^{d_{\mathbf{y}}}$ *is a finite union of bounded rectangles. Assume that* f *and* g *satisfy the assumptions of Theorem* 3 *and also* $0 < \inf_{\mathbf{y}\in C} g(\mathbf{y}) \le \sup_{\mathbf{y}\in C} g(\mathbf{y}) < \infty$ *and* $\sup_{\mathbf{z}\in A} f(\mathbf{z}) < \infty$. *Then*

$$n^{-1/2}\big(\mathscr{K}(A, A; \mathscr{P}_n) - E\mathscr{K}(A, A; \mathscr{P}_n)\big) \longrightarrow_d \text{Normal } (0, \sigma_1^2 + \sigma_2^2) \tag{41}$$

where σ_1^2 *and* σ_2^2 *are given by* (39).

PROOF. A blocking method is created for the purpose of proving (41) and is described as follows. For each n, partition C into disjoint equal-sized cubes $C_1, \ldots, C_p$. A cube is a rectangle of the form $(e_1, e_1 + \delta) \times \cdots \times (e_{d_{\mathbf{y}}}, e_{d_{\mathbf{y}}} + \delta)$ for some $\delta > 0$ and $e_1, \ldots, e_{d_{\mathbf{y}}} \in \mathbb{R}$. Call these C_i "blocks." Such a partition is possible if we choose the rectangles that form C in such a way that the ratio of the lengths of any pair of sides is a rational number. This can clearly be

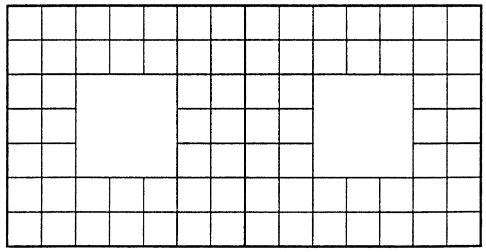

FIG. 1. *Two adjacent blocks for* $d_{\mathbf{y}} = 2$.

done without sacrificing generality. Within each C_i, place a "big" cube $\hat{C}_i$ in the center and fill the remaining space with l layers of equal-sized "small" cubes where $d_{\mathbf{y}}^{1/2} < l \le d_{\mathbf{y}}^{1/2} + 1$. Assume that the sizes of the big (small) cubes in all the C_i are the same. This way, every pair of big cubes $\hat{C}_i$ and $\hat{C}_j$ are buffered by at least $2l$ layers of small cubes. Figure 1 illustrates two adjacent blocks in dimension 2 with $l = 2$. Denote by $\tilde{C}_{i,1}, \ldots, \tilde{C}_{i,q}$ the small cubes that touch $\hat{C}_i$ (i.e., the first layer of small cubes outside of $\hat{C}_i$). The motivation for this scheme of blocking is the following. Clearly, if $\mathbf{y} \in \hat{C}_i$, it follows from the choice of l that

$$\begin{aligned}(42)\qquad &\inf(|\mathbf{y} - \mathbf{y}'|: \mathbf{y}' \notin C_i) - \min_{1\le k\le q} \sup(|\mathbf{y} - \mathbf{y}'|: \mathbf{y}' \in \tilde{C}_{i,k})\\ &\qquad \ge (l - d_{\mathbf{y}}^{1/2}) \times \text{length of the side of a small cube} > 0\end{aligned}$$

Suppose that $\mathbf{Y}_j \in \hat{C}_i$ and there is at least a $\mathbf{Y}_m$ in each of the $\tilde{C}_{i,k}$, then (42) shows that the nearest neighbor of $\mathbf{Y}_j$ must be in C_i. We will make use of this shortly. First, control the sizes of the big and small cubes by choosing

$$(43)\qquad \int_{\hat{C}_1} d\mathbf{y} \sim n^{-\beta} \quad \text{and} \quad \int_{\tilde{C}_{1,1}} d\mathbf{y} \sim n^{-\beta'} \quad \text{where } 1/2 < \beta < \beta' < 1.$$

It is then easily verified that the small cubes are asymptotically negligible and in fact

$$(44)\qquad \int_{C_1} d\mathbf{y} \sim n^{-\beta}, \ \ p = O(n^{\beta}) \quad \text{and} \quad q = O(n^{(\beta'-\beta)(d_{\mathbf{y}}-1)/d_{\mathbf{y}}}).$$

Note that we suppressed n at various places to streamline notation and furthermore write

$$\mathscr{M}(\,\cdot\,;\,\cdot\,) = \mathscr{K}(B \times \cdot\,, A;\ \cdot).$$

Since the big and small cubes are all disjoint,

$$\mathscr{M}(C; \mathscr{P}_n) = \mathscr{M}\left(\bigcup_{i=1}^{p} \hat{C}_i; \mathscr{P}_n\right) + \mathscr{M}\left(\bigcup_{i=1}^{p} (C_i - \hat{C}_i); \mathscr{P}_n\right).$$

By (43), (44) and the fact that the $\mathbf{X}_i$ are finitely truncated, it is easy to check that

$$n^{-1/2}\mathscr{M}\left(\bigcup_{i=1}^{p}(C_i - \hat{C}_i); \mathscr{P}_n\right) \to 0 \text{ uniformly almost surely}$$

and hence

$$n^{-1/2}\left(\mathscr{M}\left(\bigcup_{i=1}^{p}(C_i - \hat{C}_i); \mathscr{P}_n\right) - E\mathscr{M}\left(\bigcup_{i=1}^{p}(C_i - \hat{C}_i); \mathscr{P}_n\right)\right) \longrightarrow 0 \text{ a.s.}$$

So (41) will follow if we prove

$$\text{(45)} \quad n^{-1/2}\left(\mathscr{M}\left(\bigcup_{i=1}^{p}\hat{C}_i; \mathscr{P}_n\right) - E\mathscr{M}\left(\bigcup_{i=1}^{p}\hat{C}_i; \mathscr{P}_n\right)\right) \longrightarrow_d \text{Normal } (0, \sigma_1^2 + \sigma_2^2).$$

Let

$$\mathscr{P}_n(E) = \{(\mathbf{X}, \mathbf{Y}) \in \mathscr{P}_n \colon \mathbf{Y} \in E\} \quad \text{and} \quad N_n(E) = \#(\mathscr{P}_n(E)),\ E \subset \mathbb{R}^{d_{\mathbf{y}}}.$$

Define the following events:

$$E_i^{(1)} = \bigcap_{k=1}^{q_1}\{N_n(\tilde{C}_{i,k}) \neq 0\},$$
$$E_i^{(2)} = \{|N_n(C_i)/EN_n(C_i) - 1| \leq \zeta_n\}, \qquad 1 \leq i \leq p,$$

where ζ_n tends to 0 slowly enough so that

(46) $P[(E_i^{(1)})^c]$ and $P[(E_i^{(2)})^c]$ tend to 0 exponentially as $n \to \infty$ uniformly in i

and hence

$$\text{(47)} \qquad \sum_{i=1}^{p}\left(P[(E_i^{(1)})^c] + P[(E_i^{(2)})^c]\right) = 0.$$

This is possible by (43), (44) and the assumption $\inf_{\mathbf{y}\in C} g(\mathbf{y}) > 0$.

Consider the characteristic function

$$\phi_n(t) = E\exp\left(itn^{-1/2}\left[\mathscr{M}\left(\bigcup_{i=1}^{p}\hat{C}_i; \mathscr{P}_n\right) - E\mathscr{M}\left(\bigcup_{i=1}^{p}\hat{C}_i; \mathscr{P}_n\right)\right]\right), \qquad t \in \mathbb{R}.$$

Write

$$\phi_n(t) = \phi_{n,1}(t) + \phi_{n,2}(t),$$

where

$$\phi_{n,1}(t) = E\left[I_{\bigcap_{i=1}^{p}(E_i^{(1)} \cap E_i^{(2)})}\exp\left(itn^{-1/2}\left[\mathscr{M}\left(\bigcup_{i=1}^{p}\hat{C}_i; \mathscr{P}_n\right) - E\mathscr{M}\left(\bigcup_{i=1}^{p}\hat{C}_i; \mathscr{P}_n\right)\right]\right)\right],$$

$$\phi_{n,2}(t) = E\left[I_{\bigcup_{i=1}^{p}[(E_i^{(1)})^c \cup (E_i^{(2)})^c]}\exp\left(itn^{-1/2}\left[\mathscr{M}\left(\bigcup_{i=1}^{p}\hat{C}_i; \mathscr{P}_n\right) - E\mathscr{M}\left(\bigcup_{i=1}^{p}\hat{C}_i; \mathscr{P}_n\right)\right]\right)\right].$$

By (47), $\phi_{n,2}(t) \to 0$ and we only have to deal with $\phi_{n,1}(t)$. Take any $i = 1, \dots, p$. If $E_i^{(1)}$ holds, then by (42) and the explanation there, $\mathscr{M}(\hat{C}_i; \mathscr{P}_n)$ is completely determined by the set $\mathscr{P}_n(C_i)$, namely those points in $\mathscr{P}_n$ whose $\mathbf{y}$-coordinates are in C_i. As a result,

$$\mathscr{M}(\hat{C}_i; \mathscr{P}_n) = \mathscr{M}(\hat{C}_i; \mathscr{P}_n(C_i)) \quad \text{on } E_i^{(1)}$$

and as a consequence,

$$\phi_{n,1}(t) = E\Big\{ I_{\bigcap_{i=1}^p (E_i^{(1)} \cap E_i^{(2)})} \exp\big(itn^{-1/2}[\mathscr{M}(\hat{C}_i; \mathscr{P}_n(C_i)) - E\mathscr{M}(\hat{C}_i; \mathscr{P}_n)]\big)\Big\}.$$

Now write

$$\phi_{n,1}(t) = E\Bigg\{ I_{\bigcap_{i=1}^p E_i^{(2)}} \Bigg(\exp\Bigg[itn^{-1/2} \sum_{i=1}^p V_{n,i}\Bigg]\Bigg)\Bigg(\prod_{i=1}^p I_{E_i^{(1)}} \exp(itn^{-1/2} U_{n,i})\Bigg) \times \exp\Bigg[itn^{-1/2} \sum_{i=1}^p W_{n,i}\Bigg]\Bigg\},$$

where

$$\begin{aligned} U_{n,i} &= \mathscr{M}(\hat{C}_i; \mathscr{P}_n(C_i)) - E[\mathscr{M}(\hat{C}_i; \mathscr{P}_n(C_i)) | N_n(C_i)], \\ V_{n,i} &= E[\mathscr{M}(\hat{C}_i; \mathscr{P}_n(C_i)) | N_n(C_i)] - E[\mathscr{M}(\hat{C}_i; \mathscr{P}_n(C_i))], \\ W_{n,i} &= E[\mathscr{M}(\hat{C}_i; \mathscr{P}_n(C_i))] - E[\mathscr{M}(\hat{C}_i; \mathscr{P}_n)]. \end{aligned}$$

By (46), $\sum_{i=1}^p W_{n,i} \to 0$ exponentially fast and so we focus on the remaining terms. Clearly the random quantities $I_{E_i^{(1)}} \exp(itn^{-1/2} U_{n,i})$, $1 \le i \le p$, are conditionally independent given the $N_n(C_i)$ and $I_{\bigcap_{i=1}^p E_i^{(2)}} \exp(itn^{-1/2} \sum_{i=1}^p V_{n,i})$ is measurable with respect to $N_n(C_i)$, $1 \le i \le p$. Consequently, we have

$$\phi_{n,1}(t) = E\Bigg\{ I_{\bigcap_{i=1}^p E_i^{(2)}} \exp\Bigg[itn^{-1/2} \sum_{i=1}^p V_{n,i}\Bigg] \prod_{i=1}^p \gamma_{n,i}(t)\Bigg\} + o(1), \tag{48}$$

where

$$\gamma_{n,i}(t) = E\big(I_{E_i^{(1)}} \exp(itn^{-1/2} U_{n,i}) | N_n(C_i)\big).$$

Similar to the derivation of (47), it can be shown that on the event $\bigcap_{i=1}^p E_i^{(2)}$,

$$\lim_{n\to\infty} \sum_{i=1}^p P[(E_i^{(1)})^c | N_n(C_i)] = 0 \quad \text{uniformly}.$$

Hence, to obtain the limit of $\prod_{i=1}^p \gamma_{n,i}(t)$, we can focus on

$$\prod_{i=1}^p E\big(\exp(itn^{-1/2} U_{n,i}) | N_n(C_i)\big),$$

which is the characteristic function of the sum of p independent random variables. On the event $\bigcap_{i=1}^p E_i^{(2)}$, these random variables have zero means and the sum of the variances converges to a constant σ_1^2 uniformly by Lemma 7. Since we have restricted the $\mathbf{X}_i$ to be in a bounded set B and $n^{-1/2} \times n \times \int_{C_1} d\mathbf{y} \to 0$ by (44), it is easy to see that on the event $\bigcap_{i=1}^p E_i^{(2)}$, $n^{-1/2} U_{n,i}$ tends to 0 uniformly in i and ω. Then it follows from the elementary inequality,

$$\left| Ee^{itX} - \sum_{m=0}^{2} E\frac{(itX)^m}{m!} \right| \le \left| E\frac{(itX)^3}{3!} \right| \le \frac{t^3 \sup|X|}{6} EX^2 \quad \text{for any bounded r.v. } X$$

that

$$\lim_{n\to\infty} \prod_{i=1}^{p} E\big(\exp(itn^{-1/2} U_{n,i}) | N_n(C_i)\big) = \exp(-\sigma_1^2 t^2/2) \quad \text{uniformly on } \bigcap_{i=1}^{p} E_i^{(2)}.$$

It then follows from (48) and (46) that

$$\begin{aligned} \phi_{n,1}(t) &= \exp(-\sigma_1^2 t^2/2) E\left\{ I_{\bigcap_{i=1}^p E_i^{(2)}} \exp\left[itn^{-1/2} \sum_{i=1}^{p} V_{n,i} \right] \right\} + o(1) \\ &= \exp(-\sigma_1^2 t^2/2) E\left\{ \exp\left[itn^{-1/2} \sum_{i=1}^{p} V_{n,i} \right] \right\} + o(1). \end{aligned}$$

The same approach as before works as $V_{n,1}, \dots, V_{n,p}$ are independent (by the independent increment property of the Poisson process). It follows from Lemma 8 that the variance of $\sum_{i=1}^p V_{n,i}$ converges to σ_2^2. In view of (46), it is an easy exercise to verify the Lindeberg condition for the $V_{n,i}$ and hence (45) follows from these steps. □

We continue to use the notation defined in Proposition 6. Write

$$f_i(\mathbf{z}) = \frac{f(\mathbf{z}) I[\mathbf{y} \in C_i]}{G(C_i)}$$

and let $\mathbf{Z}_{i,j} = (\mathbf{X}_{i,j}, \mathbf{Y}_{i,j})$, $j \ge 1$, be iid random variables with distribution

$$P(\mathbf{Z}_{i,1} \in \cdot) = \int_{\mathbf{z}\in\cdot} f_i(\mathbf{z})\, d\mathbf{z}.$$

Accordingly, define $g_i(\mathbf{y})$, G_i and $f_i(\mathbf{x}|\mathbf{y})$.

LEMMA 7. *Under the conditions of Proposition* 6,

$$n^{-1} \sum_{i=1}^{p} \operatorname{var}(U_{n,i} | N_n(C_i)) \to \sigma_1^2 \tag{49}$$

$$\textit{as } n \to \infty \textit{ uniformly on the event } \bigcap_{i=1}^{p} E_i^{(2)},$$

where σ_1^2 *is defined by* (39).

PROOF. First fix $1 \le i \le p$ and $m \ge 4$ and let $\mathbf{Y}_{i,\,j*}$ be the nearest neighbor of $\mathbf{Y}_{i,\,j}$ for $1 \le j \le m$. For convenience, in this proof and that of Lemma 8 we will write

$$\zeta(\mathbf{x}_1, \mathbf{x}_2) = \tfrac{1}{2} \sum_{s=1}^{S} \mathbf{a}_s'(\mathbf{x}_1\mathbf{x}_2' + \mathbf{x}_2\mathbf{x}_1')\mathbf{b}_s I(\mathbf{x}_1, \mathbf{x}_2 \in B).$$

Then, by symmetry,

$$\begin{aligned}
&\operatorname{var}(U_{n,\,i}|N_n(C_i) = m)\\
&\quad = E[\mathscr{M}^2(\hat{C}_i; \{\mathbf{Z}_{i,\,j},\ 1 \le j \le m\})] - E^2[\mathscr{M}(\hat{C}_i; \{\mathbf{Z}_{i,\,j},\ 1 \le j \le m\})]\\
&\quad = \sum_{j=1}^{m}\sum_{k=1}^{m} E[\zeta(\mathbf{X}_{i,\,j}, \mathbf{X}_{i,\,j*})\zeta(\mathbf{X}_{i,\,k}, \mathbf{X}_{i,\,k*}) I(\mathbf{Y}_{i,\,j}, \mathbf{Y}_{i,\,k} \in \hat{C}_i)]\\
&\qquad - \big(mE[\zeta(\mathbf{X}_{i,\,1}, \mathbf{X}_{i,\,1*})\ I(\mathbf{Y}_{i,\,1} \in \hat{C}_i)]\big)^2\\
&\quad = mE[\zeta^2(\mathbf{X}_{i,\,1}, \mathbf{X}_{i,\,1*}) I(\mathbf{Y}_{i,\,1} \in \hat{C}_i)]\\
&\qquad + m(m-1)E[\zeta(\mathbf{X}_{i,\,1}, \mathbf{X}_{i,\,1*})\zeta(\mathbf{X}_{i,\,2}, \mathbf{X}_{i,\,2*}) I(\mathbf{Y}_{i,\,1}, \mathbf{Y}_{i,\,2} \in \hat{C}_i)]\\
&\qquad - \big(mE[\zeta(\mathbf{X}_{i,\,1}, \mathbf{X}_{i,\,1*})\ I(\mathbf{Y}_{i,\,1} \in \hat{C}_i)]\big)^2\\
&\quad =: M_{i,1}^{(2)}(m) + M_{i,2}^{(2)}(m) + M_{i,3}^{(2)}(m) + M_{i,4}^{(2)}(m)\\
&\qquad + M_{i,5}^{(2)}(m) + M_{i,6}^{(2)}(m) - M_i^{(1)}(m),
\end{aligned} \tag{50}$$

where

$$\begin{aligned}
M_{i,1}^{(2)}(m) &= mE[\zeta^2(\mathbf{X}_{i,\,1}, \mathbf{X}_{i,\,1*}) I(\mathbf{Y}_{i,\,1} \in \hat{C}_i)]\\
&= m(m-1)E[\zeta^2(\mathbf{X}_{i,\,1}, \mathbf{X}_{i,\,2}) I(1* = 2,\ \mathbf{Y}_{i,\,1} \in \hat{C}_i)],
\end{aligned}$$

$$\begin{aligned}
M_{i,2}^{(2)}(m) &= m(m-1)E[\zeta^2(\mathbf{X}_{i,\,1}, \mathbf{X}_{i,\,2}) I(1* = 2,\ 2* = 1,\ \mathbf{Y}_{i,\,1}, \mathbf{Y}_{i,\,2} \in \hat{C}_i)],\\
M_{i,3}^{(2)}(m) &= m(m-1)(m-2)E\big[\zeta(\mathbf{X}_{i,\,1}, \mathbf{X}_{i,\,3})\zeta(\mathbf{X}_{i,\,2}, \mathbf{X}_{i,\,3})\\
&\qquad\qquad \times I(1* = 3,\ 2* = 3,\ \mathbf{Y}_{i,\,1}, \mathbf{Y}_{i,\,2} \in \hat{C}_i)\big],\\
M_{i,4}^{(2)}(m) &= m(m-1)(m-2)E\big[\zeta(\mathbf{X}_{i,\,1}, \mathbf{X}_{i,\,2})\zeta(\mathbf{X}_{i,\,2}, \mathbf{X}_{i,\,3})\\
&\qquad\qquad \times I(1* = 2,\ 2* = 3,\ \mathbf{Y}_{i,\,1}, \mathbf{Y}_{i,\,2} \in \hat{C}_i)\big],\\
M_{i,5}^{(2)}(m) &= m(m-1)(m-2)E\big[\zeta(\mathbf{X}_{i,\,1}, \mathbf{X}_{i,\,3})\zeta(\mathbf{X}_{i,\,2}, \mathbf{X}_{i,\,1})\\
&\qquad\qquad \times I(1* = 3,\ 2* = 1,\ \mathbf{Y}_{i,\,1}, \mathbf{Y}_{i,\,2} \in \hat{C}_i)\big],\\
M_{i,6}^{(2)}(m) &= m(m-1)(m-2)(m-3)E\big[\zeta(\mathbf{X}_{i,\,1}, \mathbf{X}_{i,\,3})\zeta(\mathbf{X}_{i,\,2}, \mathbf{X}_{i,\,4})\\
&\qquad\qquad \times I(1* = 3,\ 2* = 4,\ \mathbf{Y}_{i,\,1}, \mathbf{Y}_{i,\,2} \in \hat{C}_i)\big]
\end{aligned}$$

and

$$M_i^{(1)}(m) = \big(m(m-1)E[\zeta(\mathbf{X}_{i,1}, \mathbf{X}_{i,2})I(1* = 2,\ \mathbf{Y}_{i,1} \in \hat{C}_i)]\big)^2.$$

Clearly,

$$M_{i,1}^{(2)}(m) = m(m-1)\int \zeta^2(\mathbf{x}_1, \mathbf{x}_2)I(\mathbf{y}_1 \in \hat{C}_i)f_i(\mathbf{z}_1)f_i(\mathbf{z}_2) \\ \times \bar{G}_i^{m-2}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\, d\mathbf{z}_1\, d\mathbf{z}_2.$$

Changing variables from $(\mathbf{y}_1, \mathbf{y}_2)$ to $(\mathbf{y}_1, \boldsymbol{\varepsilon})$ where $\boldsymbol{\varepsilon} = (mg_i(\mathbf{y}_1))^{1/d_{\mathrm{y}}}(\mathbf{y}_2 - \mathbf{y}_1)$, the above becomes

$$\begin{aligned} M_{i,1}^{(2)}(m) &= (m-1)\int \zeta^2(\mathbf{x}_1, \mathbf{x}_2)I(\mathbf{y}_1 \in \hat{C}_i)f_i(\mathbf{x}_1|\mathbf{y}_1) \\ &\qquad \times f_i(\mathbf{x}_2, \mathbf{y}_1 + \boldsymbol{\varepsilon}/(mg_i(\mathbf{y}_1))^{1/d_{\mathrm{y}}}) \\ &\qquad \times \bar{G}_i^{m-2}(S(\mathbf{y}_1; |\boldsymbol{\varepsilon}|/(mg_i(\mathbf{y}_1))^{1/d_{\mathrm{y}}}))\, d\mathbf{x}_1\, d\mathbf{y}_1\, d\mathbf{x}_2\, d\boldsymbol{\varepsilon} \\ &= \frac{m-1}{G(C_i)}\int I(\mathbf{y}_1 \in \hat{C}_i, \mathbf{y}_1 + \boldsymbol{\varepsilon}/(mg_i(\mathbf{y}_1))^{1/d_{\mathrm{y}}} \in C_i) \\ &\qquad \times \zeta^2(\mathbf{x}_1, \mathbf{x}_2)f(\mathbf{x}_1|\mathbf{y}_1)f(\mathbf{x}_2, \mathbf{y}_1 + \boldsymbol{\varepsilon}/(mg_i(\mathbf{y}_1))^{1/d_{\mathrm{y}}}) \\ &\qquad \times \bar{G}_i^{m-2}(S(\mathbf{y}_1; |\boldsymbol{\varepsilon}|/(mg_i(\mathbf{y}_1))^{1/d_{\mathrm{y}}}))\, d\mathbf{x}_1\, d\mathbf{y}_1\, d\mathbf{x}_2\, d\boldsymbol{\varepsilon}. \end{aligned}$$

Since g is bounded away from 0 and ∞ on $C = \bigcup_i C_i$, we have

$$\sup_{1 \le i \le p,\, m \ge 1,\, \mathbf{y} \in C} \bar{G}_i^{m-2}(S(\mathbf{y}; |\boldsymbol{\varepsilon}|/(mg_i(\mathbf{y}))^{1/d_{\mathrm{y}}})) \le \exp(-\delta|\boldsymbol{\varepsilon}|^2), \tag{51}$$

where

$$\delta := \frac{\inf_{\mathbf{y} \in C} g(\mathbf{y})}{\sup_{\mathbf{y} \in C} g(\mathbf{y})} > 0.$$

Now replace m by $N_n(C_i)$ in $M_{i,1}^{(2)}(m)$ and sum $M_{i,1}^{(2)}(N_n(C_i))$ over $i = 1, \dots, p$. Taking account of the event $\bigcap_{i=1}^p E_i^{(2)}$, it follows from (51) and dominated convergence that

$$\sum_{i=1}^p M_{i,1}^{(2)}(N_n(C_i)) \sim n\omega_1(A)\int \exp(-\|(S(0; |\boldsymbol{\varepsilon}|)\|)\, d\boldsymbol{\varepsilon} = n\omega_1(A) \tag{52}$$

as $n \to \infty$ uniformly on $\bigcap_{i=1}^p E_i^{(2)}$. Similarly, uniformly on $\bigcap_{i=1}^p E_i^{(2)}$, we have, as $n \to \infty$,

$$\sum_{i=1}^p M_{i,2}^{(2)}(N_n(C_i)) \sim n\omega_1(A)\int \exp(-\|S(0, \boldsymbol{\varepsilon}; |\boldsymbol{\varepsilon}|, |\boldsymbol{\varepsilon}|)\|)\, d\boldsymbol{\varepsilon}, \tag{53}$$

$$\begin{aligned} &\sum_{i=1}^p M_{i,3}^{(2)}(N_n(C_i)) \\ &\quad \sim n\omega_2(A)\int I(|\boldsymbol{\varepsilon}_{13} - \boldsymbol{\varepsilon}_{23}| > |\boldsymbol{\varepsilon}_{13}| \vee |\boldsymbol{\varepsilon}_{23}|) \\ &\qquad\qquad \times \exp(-\|S(\boldsymbol{\varepsilon}_{13}, \boldsymbol{\varepsilon}_{23}; |\boldsymbol{\varepsilon}_{13}|, |\boldsymbol{\varepsilon}_{23}|)\|)\, d\boldsymbol{\varepsilon}_{13}\, d\boldsymbol{\varepsilon}_{23} \end{aligned} \tag{54}$$

and

$$\sum_{i=1}^{p} M_{i,4}^{(2)}(N_n(C_i)) = \sum_{i=1}^{p} M_{i,5}^{(2)}(N_n(C_i)) \tag{55}$$
$$\sim n\omega_2(A) \int I(|\boldsymbol{\varepsilon}_{12} - \boldsymbol{\varepsilon}_{23}| > |\boldsymbol{\varepsilon}_{12}| > |\boldsymbol{\varepsilon}_{23}|) \times \exp(-\|S(0, \boldsymbol{\varepsilon}_{12}; |\boldsymbol{\varepsilon}_{12}|, |\boldsymbol{\varepsilon}_{23}|)\|)\, d\boldsymbol{\varepsilon}_{12}\, d\boldsymbol{\varepsilon}_{23}.$$

Now let's consider the most crucial term, $M_{i,6}^{(2)}$. Observe that

$$\begin{aligned} M_{i,6}^{(2)}(m) &= m(m-1)(m-2)(m-3) \\ &\quad \times \int \zeta(\mathbf{x}_1, \mathbf{x}_2)\zeta(\mathbf{x}_3, \mathbf{x}_4) I(\mathbf{y}_1, \mathbf{y}_2 \in \hat{C}_i) f_i(\mathbf{z}_1) f_i(\mathbf{z}_2) f_i(\mathbf{z}_3) f_i(\mathbf{z}_4) \\ &\quad \times I\big(|\mathbf{y}_1 - \mathbf{y}_3| < |\mathbf{y}_1 - \mathbf{y}_2| \wedge |\mathbf{y}_1 - \mathbf{y}_4|, \\ &\qquad |\mathbf{y}_2 - \mathbf{y}_4| < |\mathbf{y}_2 - \mathbf{y}_1| \wedge |\mathbf{y}_2 - \mathbf{y}_3|\big) \\ &\quad \times \bar{G}_i^{m-4}\big(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_3|) \cup S(\mathbf{y}_2; |\mathbf{y}_2 - \mathbf{y}_4|)\big)\, d\mathbf{z}_1\, d\mathbf{z}_2\, d\mathbf{z}_3\, d\mathbf{z}_4 \\ &= M_{i,6,1}^{(2)}(m) + M_{i,6,2}^{(2)}(m), \end{aligned}$$

where

$$\begin{aligned} M_{i,6,1}^{(2)}(m) &= m(m-1)(m-2)(m-3) \\ &\quad \times \int \zeta(\mathbf{x}_1, \mathbf{x}_2)\zeta(\mathbf{x}_3, \mathbf{x}_4) I(\mathbf{y}_1, \mathbf{y}_2 \in \hat{C}_i) \\ &\quad \times \Big[I\big(|\mathbf{y}_1 - \mathbf{y}_3| < |\mathbf{y}_1 - \mathbf{y}_2| \wedge |\mathbf{y}_1 - \mathbf{y}_4|, \\ &\qquad |\mathbf{y}_2 - \mathbf{y}_4| < |\mathbf{y}_2 - \mathbf{y}_1| \wedge |\mathbf{y}_2 - \mathbf{y}_3|, \\ &\qquad |\mathbf{y}_1 - \mathbf{y}_2| > |\mathbf{y}_1 - \mathbf{y}_3| + |\mathbf{y}_2 - \mathbf{y}_4|\big) - 1 \Big] \\ &\quad \times f_i(\mathbf{z}_1) f_i(\mathbf{z}_2) f_i(\mathbf{z}_3) f_i(\mathbf{z}_4) \Big(1 - G_i(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_3|)) \\ &\qquad - G_i(S(\mathbf{y}_2; |\mathbf{y}_2 - \mathbf{y}_4|))\Big)^{m-4} d\mathbf{z}_1\, d\mathbf{z}_2\, d\mathbf{z}_3\, d\mathbf{z}_4 \\ &\quad + m(m-1)(m-2)(m-3) \int \zeta(\mathbf{x}_1, \mathbf{x}_2)\zeta(\mathbf{x}_3, \mathbf{x}_4) I(\mathbf{y}_1, \mathbf{y}_2 \in \hat{C}_i) \\ &\quad \times I\big(|\mathbf{y}_1 - \mathbf{y}_3| < |\mathbf{y}_1 - \mathbf{y}_2| \wedge |\mathbf{y}_1 - \mathbf{y}_4|, \\ &\qquad |\mathbf{y}_2 - \mathbf{y}_4| < |\mathbf{y}_2 - \mathbf{y}_1| \wedge |\mathbf{y}_2 - \mathbf{y}_3|, \\ &\qquad |\mathbf{y}_1 - \mathbf{y}_2| \le |\mathbf{y}_1 - \mathbf{y}_3| + |\mathbf{y}_2 - \mathbf{y}_4|\big) f_i(\mathbf{z}_1) f_i(\mathbf{z}_2) f_i(\mathbf{z}_3) f_i(\mathbf{z}_4) \\ &\quad \times \bar{G}_i^{m-4}\big(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_3|) \cup S(\mathbf{y}_2; |\mathbf{y}_2 - \mathbf{y}_4|)\big)\, d\mathbf{z}_1\, d\mathbf{z}_2\, d\mathbf{z}_3\, d\mathbf{z}_4 \end{aligned}$$

and

$$\begin{aligned} M_{i,6,2}^{(2)}(m) &= m(m-1)(m-2)(m-3) \\ &\quad \times \int \zeta(\mathbf{x}_1, \mathbf{x}_2)\zeta(\mathbf{x}_3, \mathbf{x}_4) I(\mathbf{y}_1, \mathbf{y}_2 \in \hat{C}_i) f_i(\mathbf{z}_1) f_i(\mathbf{z}_2) f_i(\mathbf{z}_3) f_i(\mathbf{z}_4) \\ &\qquad \times \Big(1 - G_i(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_3|)) - G_i(S(\mathbf{y}_2; |\mathbf{y}_2 - \mathbf{y}_4|))\Big)^{m-4} \\ &\qquad \times d\mathbf{z}_1\, d\mathbf{z}_2\, d\mathbf{z}_3\, d\mathbf{z}_4. \end{aligned}$$

As before, it is straightforward to show that

$$
\begin{aligned}
&\sum_{i=1}^{p} M_{i,6,1}^{(2)}(N_n(C_i)) \\
&\quad \sim n\omega_3(A)\biggl[-\int_{\boldsymbol{\varepsilon}_{12}}\int_{\boldsymbol{\varepsilon}_{13}}\int_{\boldsymbol{\varepsilon}_{24}} I\bigl(|\boldsymbol{\varepsilon}_{13}| > |\boldsymbol{\varepsilon}_{12}| \wedge |\boldsymbol{\varepsilon}_{12} - \boldsymbol{\varepsilon}_{24}| \\
&\qquad\qquad \text{or } |\boldsymbol{\varepsilon}_{24}| > |\boldsymbol{\varepsilon}_{12}| \wedge |\boldsymbol{\varepsilon}_{12} - \boldsymbol{\varepsilon}_{13}| \text{ or } |\boldsymbol{\varepsilon}_{12}| < |\boldsymbol{\varepsilon}_{13}| + |\boldsymbol{\varepsilon}_{24}|\bigr) \\
&\qquad\qquad \times \exp(-\|S(0;|\boldsymbol{\varepsilon}_{13}|)\| - \|S(0;|\boldsymbol{\varepsilon}_{24}|)\|)\, d\boldsymbol{\varepsilon}_{12}\, d\boldsymbol{\varepsilon}_{13}\, d\boldsymbol{\varepsilon}_{24} \\
&\qquad\qquad + \int_{\boldsymbol{\varepsilon}_{12}}\int_{\boldsymbol{\varepsilon}_{13}}\int_{\boldsymbol{\varepsilon}_{24}} I\bigl(|\boldsymbol{\varepsilon}_{13}| < |\boldsymbol{\varepsilon}_{12}| \wedge |\boldsymbol{\varepsilon}_{12} - \boldsymbol{\varepsilon}_{24}|, \\
&\qquad\qquad\qquad |\boldsymbol{\varepsilon}_{24}| < |\boldsymbol{\varepsilon}_{12}| \wedge |\boldsymbol{\varepsilon}_{12} - \boldsymbol{\varepsilon}_{13}|, \\
&\qquad\qquad\qquad\qquad |\boldsymbol{\varepsilon}_{12}| \le |\boldsymbol{\varepsilon}_{13}| + |\boldsymbol{\varepsilon}_{24}|\bigr) \\
&\qquad\qquad \times \exp(-\|S(0,\boldsymbol{\varepsilon}_{12};|\boldsymbol{\varepsilon}_{13}|,|\boldsymbol{\varepsilon}_{24}|)\|)\, d\boldsymbol{\varepsilon}_{12}\, d\boldsymbol{\varepsilon}_{13}\, d\boldsymbol{\varepsilon}_{24}\biggr].
\end{aligned}
\tag{56}
$$

So it remains to show how $M_{i,6,2}^{(2)}(m)$ interacts with $M_i^{(1)}(m)$. Write

$$
\begin{aligned}
M_{i,6,2}^{(2)}(m) = m^4 \int & \zeta(\mathbf{x}_1,\mathbf{x}_2)\zeta(\mathbf{x}_3,\mathbf{x}_4) I(\mathbf{y}_1,\mathbf{y}_2 \in \hat{C}_i) \\
& \times \exp\bigl[-mG_i(S(\mathbf{y}_1;|\mathbf{y}_1-\mathbf{y}_3|)) - mG_i(S(\mathbf{y}_2;|\mathbf{y}_2-\mathbf{y}_4|))\bigr] \\
& \times f_i(\mathbf{z}_1)f_i(\mathbf{z}_2)f_i(\mathbf{z}_3)f_i(\mathbf{z}_4)\, d\mathbf{z}_1\, d\mathbf{z}_2\, d\mathbf{z}_3\, d\mathbf{z}_4 + R_i^{(2)}(m),
\end{aligned}
$$

where

$$
\begin{aligned}
R_i^{(2)}(m) &= m(m-1)(m-2)(m-3) \\
&\quad \times \int \zeta(\mathbf{x}_1,\mathbf{x}_2)\zeta(\mathbf{x}_3,\mathbf{x}_4) I(\mathbf{y}_1,\mathbf{y}_2 \in \hat{C}_i) \\
&\quad\quad \times \bigl(1 - G_i(S(\mathbf{y}_1;|\mathbf{y}_1-\mathbf{y}_3|)) - G_i(S(\mathbf{y}_2;|\mathbf{y}_2-\mathbf{y}_4|))\bigr)^{m-4} \\
&\quad\quad \times f_i(\mathbf{z}_1)f_i(\mathbf{z}_2)f_i(\mathbf{z}_3)f_i(\mathbf{z}_4)\, d\mathbf{z}_1\, d\mathbf{z}_2\, d\mathbf{z}_3\, d\mathbf{z}_4 \\
&\quad\quad - m^4 \int \zeta(\mathbf{x}_1,\mathbf{x}_2)\zeta(\mathbf{x}_3,\mathbf{x}_4) I(\mathbf{y}_1,\mathbf{y}_2 \in \hat{C}_i) \\
&\quad\quad\quad \times \exp\bigl[-mG_i(S(\mathbf{y}_1;|\mathbf{y}_1-\mathbf{y}_3|)) - mG_i(S(\mathbf{y}_2;|\mathbf{y}_2-\mathbf{y}_4|))\bigr] \\
&\quad\quad\quad \times f_i(\mathbf{z}_1)f_i(\mathbf{z}_2)f_i(\mathbf{z}_3)f_i(\mathbf{z}_4)\, d\mathbf{z}_1\, d\mathbf{z}_2\, d\mathbf{z}_3\, d\mathbf{z}_4.
\end{aligned}
$$

Also write

$$
\begin{aligned}
M_i^{(1)}(m) = m^4 \int & \zeta(\mathbf{x}_1,\mathbf{x}_2)\zeta(\mathbf{x}_3,\mathbf{x}_4) I(\mathbf{y}_1,\mathbf{y}_2 \in \hat{C}_i) \\
& \times \exp\bigl[-mG_i(S(\mathbf{y}_1;|\mathbf{y}_1-\mathbf{y}_3|)) - mG_i(S(\mathbf{y}_2;|\mathbf{y}_2-\mathbf{y}_4|))\bigr] \\
& \times f_i(\mathbf{z}_1)f_i(\mathbf{z}_2)f_i(\mathbf{z}_3)f_i(\mathbf{z}_4)\, d\mathbf{z}_1\, d\mathbf{z}_2\, d\mathbf{z}_3\, d\mathbf{z}_4 + R_i^{(1)}(m),
\end{aligned}
$$

where

$$R_i^{(1)}(m) = \Big(m(m-1)\int \zeta(\mathbf{x}_1, \mathbf{x}_2) I(\mathbf{y}_1 \in \hat{C}_i) f_i(\mathbf{z}_1) f_i(\mathbf{z}_2) \times \bar{G}_i^{m-2}(S(\mathbf{y}_1; |\mathbf{y}_2 - \mathbf{y}_1|))\, d\mathbf{z}_1\, d\mathbf{z}_2\Big)^2$$
$$- m^4 \int \zeta(\mathbf{x}_1, \mathbf{x}_2)\zeta(\mathbf{x}_3, \mathbf{x}_4) I(\mathbf{y}_1, \mathbf{y}_2 \in \hat{C}_i) \times \exp\big[-mG_i(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_3|)) - mG_i(S(\mathbf{y}_2; |\mathbf{y}_2 - \mathbf{y}_4|))\big] \times f_i(\mathbf{z}_1) f_i(\mathbf{z}_2) f_i(\mathbf{z}_3) f_i(\mathbf{z}_4)\, d\mathbf{z}_1\, d\mathbf{z}_2\, d\mathbf{z}_3\, d\mathbf{z}_4.$$

Hence

$$M_{i,6,2}^{(2)}(m) - M_i^{(1)}(m) = R_i^{(2)}(m) - R_i^{(1)}(m).$$

Treating this with a similar approach to before yields

$$\sum_{i=1}^{p} M_{i,6,2}^{(2)}(N_n(C_i)) - \sum_{i=1}^{p} M_i^{(1)}(N_n(C_i)) \sim -n\omega_3(A). \tag{57}$$

The proof of (49) follows from (52)–(57). □

LEMMA 8. *Under the conditions of Proposition* 6,

$$\lim_{n\to\infty} n^{-1} \sum_{i=1}^{p} \operatorname{var}(V_{n,i}) = \sigma_2^2,$$

where σ_2^2 *is defined by* (39).

PROOF. For convenience, denote $N_n(C_i)$ by $N_{n,i}$ and use the notation in Proposition 6 and Lemma 7. Thus,

$$E[\mathscr{M}(\hat{C}_i; \mathscr{P}_n)|N_{n,i}] = N_{n,i}(N_{n,i}-1)\int \zeta(\mathbf{x}_1, \mathbf{x}_2) I(\mathbf{y}_1 \in \hat{C}_i) f_i(\mathbf{z}_1) f_i(\mathbf{z}_2) \times \bar{G}_i^{N_{n,i}-2}(S(\mathbf{y}_1; |\mathbf{y}_2 - \mathbf{y}_1|))\, d\mathbf{z}_1\, d\mathbf{z}_2.$$

Let N be a Poisson random variable with mean λ and let

$$\phi(s) = Es^N = e^{\lambda(s-1)}, \qquad s > 0.$$

Then for any $s > 0$,

$$\begin{aligned} E(N^2(N-1)^2 s^{N-2}) &= s^2 E(N(N-1)(N-2)(N-3)s^{N-4}) \\ &\quad + 4sE(N(N-2)(N-2)s^{N-3}) + 2E(N(N-1)s^{N-2}) \\ &= s^2\phi^{(4)}(s) + 4s\phi^{(3)}(s) + 2\phi^{(2)}(s) \\ &= e^{\lambda(s-1)}(s^2\lambda^4 + 4s\lambda^3 + 2\lambda^2) \\ &=: \gamma(s, \lambda). \end{aligned}$$

Hence,

$$
\begin{aligned}
&E(E^2[\mathscr{M}(\hat{C}_i;\mathscr{P}_n)|N_{n,i}])\\
&\quad=\int \zeta(\mathbf{x}_1,\mathbf{x}_2)\zeta(\mathbf{x}_3,\mathbf{x}_4)I(\mathbf{y}_1,\mathbf{y}_3\in\hat{C}_i)f_i(\mathbf{z}_1)f_i(\mathbf{z}_2)f_i(\mathbf{z}_3)f_i(\mathbf{z}_4)\\
&\qquad\times E\big[N_{n,i}^2(N_{n,i}-1)^2(\bar{G}_i(S(\mathbf{y}_1;|\mathbf{y}_2-\mathbf{y}_1|))\bar{G}_i(S(\mathbf{y}_3;|\mathbf{y}_3-\mathbf{y}_4|)))^{N_{n,i}-2}\big]\\
&\qquad\times d\mathbf{z}_1\,d\mathbf{z}_2\,d\mathbf{z}_3\,d\mathbf{z}_4\\
&\quad=\int \zeta(\mathbf{x}_1,\mathbf{x}_2)\zeta(\mathbf{x}_3,\mathbf{x}_4)I(\mathbf{y}_1,\mathbf{y}_3\in\hat{C}_i)f_i(\mathbf{z}_1)f_i(\mathbf{z}_2)f_i(\mathbf{z}_3)f_i(\mathbf{z}_4)\\
&\qquad\times\gamma(\bar{G}_i(S(\mathbf{y}_1;|\mathbf{y}_2-\mathbf{y}_1|))\bar{G}_i(S(\mathbf{y}_3;|\mathbf{y}_3-\mathbf{y}_4|)),nG(C_i))\,d\mathbf{z}_1\,d\mathbf{z}_2\,d\mathbf{z}_3\,d\mathbf{z}_4.
\end{aligned}
$$

Now,

$$
\begin{aligned}
&\gamma(\bar{G}_i(S(\mathbf{y}_1;|\mathbf{y}_2-\mathbf{y}_1|))\bar{G}_i(S(\mathbf{y}_3;|\mathbf{y}_3-\mathbf{y}_4|)),nG(C_i))\\
&\quad=(nG(C_i))^4\exp\big(-nG(C_i)\big[G_i(S(\mathbf{y}_1;|\mathbf{y}_2-\mathbf{y}_1|))\\
&\qquad\qquad+G_i(S(\mathbf{y}_3;|\mathbf{y}_3-\mathbf{y}_4|))\big]\big)\\
&\qquad\times\Big(1-2G_i(S(\mathbf{y}_1;|\mathbf{y}_2-\mathbf{y}_1|))-2G_i(S(\mathbf{y}_3;|\mathbf{y}_3-\mathbf{y}_4|))\\
&\qquad\quad+nG(C_i)G_i(S(\mathbf{y}_1;|\mathbf{y}_2-\mathbf{y}_1|))G_i(S(\mathbf{y}_3;|\mathbf{y}_3-\mathbf{y}_4|))\Big)\\
&\qquad+4(nG(C_i))^3\exp\big(-nG(C_i)\big[G_i(S(\mathbf{y}_1;|\mathbf{y}_2-\mathbf{y}_1|))\\
&\qquad\qquad+G_i(S(\mathbf{y}_3;|\mathbf{y}_3-\mathbf{y}_4|))\big]\big)\\
&\qquad+\text{smaller order terms.}
\end{aligned}\tag{58}
$$

Similarly,

$$E(N(N-1)s^{N-2})=\lambda^2e^{\lambda(s-1)}:=\delta(s,\lambda)$$

and hence

$$
\begin{aligned}
&E^2(E[\mathscr{M}(\hat{C}_i;\mathscr{P}_n)|N_{n,i}])\\
&\quad=\int \zeta(\mathbf{x}_1,\mathbf{x}_2)\zeta(\mathbf{x}_3,\mathbf{x}_4)I(\mathbf{y}_1,\mathbf{y}_3\in\hat{C}_i)f_i(\mathbf{z}_1)f_i(\mathbf{z}_2)f_i(\mathbf{z}_3)f_i(\mathbf{z}_4)\\
&\qquad\times E\Big[N_{n,i}(N_{n,i}-1)\bar{G}_i^{N_{n,i}-2}(S(\mathbf{y}_1;|\mathbf{y}_2-\mathbf{y}_1|))\Big]\\
&\qquad\times E\Big[N_{n,i}(N_{n,i}-1)\bar{G}_i^{N_{n,i}-2}(S(\mathbf{y}_3;|\mathbf{y}_3-\mathbf{y}_4|))\Big]\,d\mathbf{z}_1\,d\mathbf{z}_2\,d\mathbf{z}_3\,d\mathbf{z}_4\\
&\quad=\int \zeta(\mathbf{x}_1,\mathbf{x}_2)\zeta(\mathbf{x}_3,\mathbf{x}_4)I(\mathbf{y}_1,\mathbf{y}_3\in\hat{C}_i)f_i(\mathbf{z}_1)f_i(\mathbf{z}_2)f_i(\mathbf{z}_3)f_i(\mathbf{z}_4)\\
&\qquad\times\delta(\bar{G}_i(S(\mathbf{y}_1;|\mathbf{y}_2-\mathbf{y}_1|)),nG(C_i))\\
&\qquad\times\delta(\bar{G}_i(S(\mathbf{y}_3;|\mathbf{y}_3-\mathbf{y}_4|)),nG(C_i))\,d\mathbf{z}_1\,d\mathbf{z}_2\,d\mathbf{z}_3\,d\mathbf{z}_4,
\end{aligned}
$$

where

$$\delta(\bar{G}_i(S(\mathbf{y}_1;|\mathbf{y}_2-\mathbf{y}_1|)), nG(C_i))\delta(\bar{G}_i(S(\mathbf{y}_3;|\mathbf{y}_3-\mathbf{y}_4|)), nG(C_i))$$
$$= (nG(C_i))^4 \exp(-nG(C_i)[G_i(S(\mathbf{y}_1;|\mathbf{y}_2-\mathbf{y}_1|)) + G_i(S(\mathbf{y}_3;|\mathbf{y}_3-\mathbf{y}_4|))]).$$

Note that this cancels the first term in (58). By an approach similar to that in the proof of Lemma 7, it is now easy to show from (58) that

$$\begin{aligned}\operatorname{var}\left(\sum_{i=1}^{p} V_{n,i}\right) &= \sum_{i=1}^{p}\{E(E^2[\mathscr{M}(\hat{C}_i;\mathscr{P}_n)|N_{n,i}]) - E^2(E[\mathscr{M}(\hat{C}_i;\mathscr{P}_n)|N_{n,i}])\}\\
&\sim n\omega_3(A)\biggl[-4\int \exp(-\|S(0,|\boldsymbol{\varepsilon}|)\|)\,d\boldsymbol{\varepsilon}\\
&\qquad\times\int \|S(0,|\boldsymbol{\varepsilon}|)\|\exp(-\|S(0,|\boldsymbol{\varepsilon}|)\|)\,d\boldsymbol{\varepsilon}\\
&\qquad+\left(\int \|S(0,|\boldsymbol{\varepsilon}|)\|\exp(-\|S(0,|\boldsymbol{\varepsilon}|)\|)\,d\boldsymbol{\varepsilon}\right)^2\\
&\qquad\qquad+4\left(\int \exp(-\|S(0,|\boldsymbol{\varepsilon}|)\|)\,d\boldsymbol{\varepsilon}\right)^2\biggr]\\
&= n\omega_3(A).\end{aligned}$$

□

7. Other technical details.

PROPOSITION 9. *Under the conditions of Proposition* 6, *there exists some finite constant M such that for all m, n,*

$$\operatorname{var}\bigl(\mathscr{K}(A,A;\mathscr{I}_{n+m}) - \mathscr{K}(A,A;\mathscr{I}_n)\bigr) \le Mm.$$

PROOF. For $i = 1,\dots,m+n$, let $i*$ be the index for which $\mathbf{Y}_{i*}$ is the nearest neighbor of $\mathbf{Y}_i$ in $\mathscr{I}_{n+m}$ and for $1 \le i \le n$, let $i\circledast$ be the index for which $\mathbf{Y}_{i\circledast}$ is the nearest neighbor of $\mathbf{Y}_i$ in $\mathscr{I}_n$. Write

$$\mathscr{K}(A,A;\mathscr{I}_{n+m}) - \mathscr{K}(A,A;\mathscr{I}_n) =: E_1 + E_2,$$

where, with ψ defined in Theorem 3,

$$R_1 = \sum_{i=n+1}^{n+m} \psi(\mathbf{X}_i,\mathbf{X}_{i*})I(\mathbf{Z}_i,\mathbf{Z}_{i*}\in A),$$
$$R_2 = \sum_{i=1}^{n}[\psi(\mathbf{X}_i,\mathbf{X}_{i*})I(\mathbf{Z}_i,\mathbf{Z}_{i*}\in A) - \psi(\mathbf{X}_i,\mathbf{X}_{i\circledast})I(\mathbf{Z}_i,\mathbf{Z}_{i\circledast}\in A)].$$

The first of these can be handled in a way that is very similar to what is in the proof of Theorem 2 or Lemma 7 to give

$$\operatorname{var}(R_1) \le Mm.$$

Now write

$$D_i = \psi(\mathbf{X}_i, \mathbf{X}_{i*}) I(\mathbf{Z}_i, \mathbf{Z}_{i*} \in A) - \psi(\mathbf{X}_i, \mathbf{X}_{i\circledast}) I(\mathbf{Z}_i, \mathbf{Z}_{i\circledast} \in A)$$

and hence by symmetry,

$$\operatorname{var}(R_2) = nED_1^2 + n(n-1)ED_1D_2 - n^2E^2D_1.$$

First,

$$nED_1^2 = nED_1^2 I(1* \neq 1_\circledast) \leq MnP(1* \neq 1_\circledast) = Mn\frac{m}{n-1}.$$

Next,

$$\begin{aligned} n(n-1)ED_1D_2 &= n(n-1)ED_1D_2I(1* \neq 1_\circledast, 2* \neq 2_\circledast) \\ &= n(n-1)\big(ED_1D_2I[(1* \neq 1_\circledast, 2* \neq 2_\circledast) \cap F] \\ &\qquad + ED_1D_2I[(1* \neq 1_\circledast,\ 2* \neq 2_\circledast) \cap F^c]\big), \end{aligned}$$

where

$$F = (1_\circledast \neq 2,\ 1_\circledast \neq 2_\circledast,\ 2_\circledast \neq 1,\ 1* \neq 2*).$$

Keep in mind that the D_i are bounded. Hence, taking an event in F^c, say $(1_\circledast = 2)$, the contribution of it to $n(n-1)ED_1D_2$ is

$$\begin{aligned} &n(n-1)E|D_1D_2|I(1* \neq 1_\circledast,\ 2* \neq 2_\circledast,\ 1_\circledast = 2) \\ &\quad \leq Mn(n-1)P(1* \neq 1_\circledast,\ 1_\circledast = 2) \\ &\quad = Mn(n-1)m\int I(|\mathbf{y}_1 - \mathbf{y}_2| > |\mathbf{y}_1 - \mathbf{y}_{n+1}|)\bar{G}^{n-1}(S(\mathbf{y}_1\colon |\mathbf{y}_1 - \mathbf{y}_2|)) \\ &\qquad\qquad \times \bar{G}^{m-1}(S(\mathbf{y}_1\colon |\mathbf{y}_1 - \mathbf{y}_{n+1}|))\, dG(\mathbf{y}_1)\, dG(\mathbf{y}_2)\, dG(\mathbf{y}_{n+1}) \\ &\quad = Mn(n-1)\frac{m}{n}\frac{1}{n+m}. \end{aligned}$$

The contribution of other events in F^c to $n(n-1)ED_1D_2$ can be dealt with using the same principle to give $O(m)$. So it remains to consider $n(n-1)ED_1D_2I[(1* \neq 1_\circledast,\ 2* \neq 2_\circledast) \cap F^c] - n^2E^2D_1$. Clearly,

$$\begin{aligned} &n(n-1)ED_1D_2I[(1* \neq 1_\circledast,\ 2* \neq 2_\circledast) \cap F] \\ &\quad = n(n-1)(n-2)(n-3)m(m-1)ED_1D_2 \\ &\qquad \times I(1_\circledast = 3,\ 2_\circledast = 4,\ 1* = n+1,\ 2* = n+2), \end{aligned}$$

whereas

$$n^2E^2D_1 = (n(n-1)mED_1I(1* = n+1,\ 1_\circledast = 2))^2$$

As in the proof of Lemma 7, the two leading terms here cancel and the remainders are of $O(m)$. This concludes the proof. □

PROPOSITION 10. *Under the conditions of Proposition* 6,

$$n^{-1/2}(\tau(P_n) - \tau(n)) \longrightarrow_d \text{Normal } (0, \sigma_3^2)$$

where σ_3^2 *is defined by* (39).

PROOF. The proof is based on an application of the "delta-method," as follows. Clearly,

$$\tau(m) = m(m-1)\int \psi(\mathbf{x}_1, \mathbf{x}_2) I(\mathbf{z}_1, \mathbf{z}_2 \in A) \times \bar{G}^{m-2}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\, dF(\mathbf{z}_1)\, dF(\mathbf{z}_2).$$

Now define

$$\rho(x) = \tau(nx).$$

Then

$$\begin{aligned}\rho'(x) = n(2nx-1)&\int \psi(\mathbf{x}_1, \mathbf{x}_2) I(\mathbf{z}_1, \mathbf{z}_2 \in A) \\ &\times \bar{G}^{nx-2}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|))\, dF(\mathbf{z}_1)\, dF(\mathbf{z}_2) \\ &+ nx(nx-1)\int \psi(\mathbf{x}_1, \mathbf{x}_2) I(\mathbf{z}_1, \mathbf{z}_2 \in A) \bar{G}^{nx-2}(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)) \\ &\times \log(G^n(S(\mathbf{y}_1; |\mathbf{y}_1 - \mathbf{y}_2|)))\, dF(\mathbf{z}_1)\, dF(\mathbf{z}_2)\end{aligned}$$

By dominated convergence,

$$\rho'(1) \sim n\int \psi(\mathbf{x}_1, \mathbf{x}_2) I(\mathbf{x}_1, \mathbf{x}_2 \in B,\ \mathbf{y} \in C) f(\mathbf{x}_1|\mathbf{y}) f(\mathbf{x}_2|\mathbf{y}) f(\mathbf{y})\, d\mathbf{x}_1\, d\mathbf{x}_2\, d\mathbf{y}.$$

By the "delta-method" based on the fact that $(P_n - n)/\sqrt{n} \to_d$ Normal $(0, 1)$, we obtain

$$\begin{aligned}\frac{\tau(P_n) - \tau(n)}{\sqrt{n}} &= \frac{\rho(P_n/n) - \rho(1)}{\sqrt{n}} \\ &= \rho'(1)\frac{(P_n/n) - 1}{\sqrt{n}} + o_P(1) \longrightarrow_d \text{Normal } (0, \omega_4(A)).\end{aligned}$$ □

Acknowledgments. I am grateful to the referee for pointing out the flaws in the previous version of this paper. I also wish to thank Zhidong Bai for helping me with the proof of Lemma 4 and Sungchul Lee for making me aware of the relevant results in geometric probability. Special thanks also go to Ray Carroll who motivated me to write this paper and helped me see through things in more ways than one.

REFERENCES

ALDOUS, D. and STEELE, M. (1992). Asymptotics for Euclidean minimal spanning trees on random graphs. *Probab. Theory Related Fields* **92** 247–258.

AVRAM, F. and BERTSIMAS, D. (1993). On central limit theorems in geometrical probability. *Ann. Appl. Probab.* **3** 1033–1046.

BAI, Z. D., MIAO, B. Q. and RADHAKRISHNA, R. (1991). Estimation of directions of arrival of signals: asymptotic results. In *Advances in Spectral Analysis and Array Processing* (S. Haykin, ed.) **2** 327–347. Prentice-Hall, Englewood Cliffs, NJ.

BICKEL, P. J. and BREIMAN, L. (1983). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. *Ann. Probab.* **11** 185–214.

BREIMAN, L., FRIEDMAN, J. H., OLSHEN, R. and STONE, C. (1984). *Classification of Regression Trees*. Wadsworth, Belmont, CA.

CHEN, H. (1991). Estimation of a projection-pursuit type regression model. *Ann. Statist.* **19** 142–157.

COOK, R. D. (1995). Graphics for studying net effects of regression predictors. *Statist. Sinica* **5** 689–708.

COOK, R. D. (1998). Principal Hessian directions revisited. *J. Amer. Statist. Assoc.* **93** 84–100.

FRIEDMAN, J. H. and STUETZLE, W. (1981). Projection pursuit regression. *J. Amer. Statist. Assoc.* **76** 817–823.

HALL, P. (1989). On projection pursuit regression. *Ann. Statist.* **17** 573–588.

HALL, P. and LI, K. C. (1993). On almost linearity of low-dimensional projections from high-dimensional data. *Ann. Statist.* **21** 867–889.

HÄRDLE, W. and STOKER, T. M. (1989). Investigation smooth multiple regression by the method of average derivatives. *J. Amer. Statist. Assoc.* **84** 986–995.

HASTIE, T. and TIBSHIRANI, R. (1986). Generalized additive models. *Statist. Sci.* **1** 297–318.

HUBER, P. (1985). Projection pursuit (with discussion). *Ann. Statist.* **13** 435–526.

LI, K. C. (1991). Sliced inverse regression for dimension reduction (with discussion). *J. Amer. Statist. Assoc.* **86** 316–342.

LI, K. C. (1992). On principal Hessian directions for data visualization and dimension reduction: another application of Stein's lemma. *J. Amer. Statist. Assoc.* **87** 1025–1039.

LI, K. C., ARAGON, Y. and THOMAS-AGNAN, C. (1994). Analysis of multivariate outcome data: SIR and a non-linear theory of Hotelling's most predictable variates. Preprint.

SAMOROV, A. M. (1993). Exploring regression structure using nonparametric functional estimation. *J. Amer. Statist. Assoc.* **88** 836–847.

DEPARTMENT OF STATISTICS
AND APPLIED PROBABILITY
NATIONAL UNIVERSITY OF SINGAPORE
LOWER KENT RIDGE ROAD
SINGAPORE 119260