RANDOM ASSOCIATION OF SYMMETRIC ARRAYS

A.D. Barbour
Universität Zürich, Institut für Angewandte Mathematik
Rämistrasse 74, 8001 ZUERICH, Switzerland

G.K. Eagleson
CSIRO, Division of Mathematics and Statistics
P.O.Box 218, LINDFIELD NSW 2070, Australia

ABSTRACT

A study is made of the asymptotic behaviour of quantities of the form
\[ U = \sum_{i,j} d_{ij} y_{\pi(i)\pi(j)}, \]
where \( \pi \) is randomly chosen from the uniform distribution over the set of permutations of \( \{1, 2, \ldots, n\} \). \( U \) can always be decomposed into the sum of two uncorrelated parts, one degenerate and the other non-degenerate. When the non-degenerate part dominates asymptotically, the limit law for \( U \) is typically normal. When the degenerate part dominates, the limit law is sometimes normal and sometimes a quadratic form in correlated normal variables. Applications to random vertex colourings of graphs are discussed.

1. INTRODUCTION.

Let \( \mathcal{G} \) and \( \mathcal{H} \) be graphs on \( n \) vertices, and let \( U \) denote the number of edges in common to the two graphs, when the vertices of the graph \( \mathcal{G} \) are identified with those of the graph \( \mathcal{H} \) at random. Thus, if the incidence matrices of the graphs are denoted by \( G \) and \( H \), \( U \) can be represented as
\[ U = \frac{1}{2} \sum_{i,j}^{n} G_{ij} H_{\pi(i)\pi(j)}, \]
where \( \pi \) is a random element of the permutation group \( S_n \). A description of the distribution of quantities of this form, and of the more general form where \( G \) and \( H \) are arbitrary symmetric matrices, is frequently needed in
the statistical analysis of spatial autocorrelation and space-time clustering, and has been discussed in particular in [3] and [4] and by Barton (unpublished lecture notes). A serious difficulty arising in practice is that a normal approximation is often inappropriate, even when $n$ is large. Sufficient conditions for the asymptotic normality of the distribution of $U$ as $n \to \infty$ were proposed in [1] and [5]. Barton showed these conditions to be incorrect, and established in their place a set of sufficient conditions described in Section 3.

In this paper, the approximation of the distribution of $U$, for arbitrary symmetric matrices $G$ and $H$, is developed further. Theorems 1 and 2 imply two sets of conditions under which the distribution of $U$ is asymptotically normal as $n \to \infty$. The first of these is derived very simply, using a projection technique, but is inadequate in many cases of practical interest, in which at least one of the matrices is the incidence matrix of a graph with average vertex degree of order 1 as $n \to \infty$. The second set is very much sharper, but the proof of sufficiency is unfortunately long and messy. A comparison of the two sets of conditions, with regard to statistical applications, is made in [2]. In Theorem 3, conditions under which convergence to distributions other than the normal may occur are derived. The occurrence of such limits is shown to be closely related to the phenomenon of degeneracy for U-statistics, and their existence helps to explain why a normal approximation to $U$ may not always be suitable in practice.

The method used to establish Theorems 1 and 2 involves making a direct estimate, for any fixed $n$, $G$ and $H$, of the distance between the distribution of $U$ and the normal distribution with the same mean and variance. The distance $d_1$ used is based on expectations of smooth functions, and is defined by

$$d_1(F, G) := \sup_{h \in \mathcal{C}_1} \{| \int h dF - \int h dG | / \| h \|_1 \},$$

(1.1)

where $\| h \|_1 := \sup_x |h(x)| + \sup_x |h'(x)|$. The advantage of such an approach is that it yields not only limit theorems, but also orders of magnitude for the rate of convergence to the limit.

The main results are stated in Section 2, and their proofs are given in Sections 4 and 5. Section 3 discusses some applications of the results to random intersections of graphs.
2. RESULTS.

Let \((d_{ij})_{1 \leq i \neq j \leq n}\) and \((y_{ij})_{1 \leq i \neq j \leq n}\) be real symmetric arrays, and consider the distribution of

\[
U := U(\pi) := \sum_{i,j=1 \atop i \neq j}^{n} d_{ij} y_{\pi(i) \pi(j)},
\]

where \(\pi\) is uniformly distributed over all permutations of \(\{1, 2, \ldots, n\}\). Direct calculation shows that

\[
\text{EU} = n(n - 1)DY
\]

and that

\[
s^2 := \text{var} U = \frac{4n^2(n - 2)^2}{n - 1} D_1 Y_1 + \frac{2n(n - 1)^2}{n - 3} D_2 Y_2,
\]

where

\[
D := \frac{1}{n(n - 1)} \sum_{i,j} d_{ij};
\]

\[
D_1 := \frac{1}{n} \sum_{i=1}^{n} d_{i}^2; \quad D_2 := \frac{1}{n(n - 1)} \sum_{i,j} (d_{ij} - d_i^* - d_j^* - D)^2
\]

and

\[
d_i^* := \frac{1}{n - 2} \sum_{j \neq i} (d_{ij} - D),
\]

and where similar formulae hold for the \(y\)'s. The notation \(\sum_{i,j}\) is used, here and subsequently, to denote \(\sum_{i,j=1 \atop i \neq j}^{n}\), and \(\sum_{i \neq j}^{n}\) denotes \(\sum_{j=1 \atop j \neq i}^{n}\). We shall be interested in finding approximations to the distribution of \(U\), based on distributional limit theorems for sequences \(s_n^{-1}(U^{(n)} - \text{EU}^{(n)})\), \(n \geq 1\), where \(U^{(n)}\) is derived from arrays \(d^{(n)} := (d_{ij}^{(n)})\) and \(y^{(n)} := (y_{ij}^{(n)})\), and \(s_n^2 = \text{var} U^{(n)}\).

The expression \(U\) can be projected (c.f. [7]) in the following way. Let

\[
u_{ij} = d_{ij} - d_i^* - d_j^* - D; \quad v_{ij} = y_{ij} - y_i^* - y_j^* - Y.
\]
Then it follows that

\[ \sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} y_i^2 = 0; \quad \sum_{j \neq i} u_{ij} = \sum_{j \neq i} v_{ij} = 0, \]

and so

\[ U - \mathbf{E}U = X + Z, \quad (2.4) \]

where

\[ X = \sum_{i,j} u_{ij} v_{\pi(i)\pi(j)}, \quad (2.5) \]

and

\[ Z = 2(n - 2) \sum_{i=1}^{n} d_i^2 y_{\pi(i)}^2. \quad (2.6) \]

The approximations to the distribution of \( U \) that are obtained depend in character upon the relative magnitudes of the components \( X \) and \( Z \). These magnitudes are easily compared using their variances, for which we have the formulae

\[ \text{var} X = 2n(n - 1)^2(n - 3)^{-1} D_2 Y_2, \]

\[ \text{var} Z = 4n^2(n - 2)^2(n - 1)^{-1} D_1 Y_1, \quad (2.7) \]

and

\[ \text{cov}(X, Z) = 0. \]

Consider first the case when \( \text{var} Z \gg \text{var} X \); that is, \( n D_1 Y_1 \gg D_2 Y_2 \). Then the distribution of \( W := s^{-1}(U - \mathbf{E}U) \) is dominated by that of \( s^{-1} Z \), and \( Z \) is of the classical Wald-Wolfowitz form. It is thus easy to use Stein's [10] method to estimate the \( d_1 \) distance between \( \mathcal{L}(W) \), the distribution of \( W \), and the standard normal distribution \( \mathcal{N} \), in the following third moment form.

**Theorem 1.** There exists a universal constant \( K_1 \) such that

\[ d_1(\mathcal{L}(W), \mathcal{N}) \leq K_1 \{ \delta + \eta \}, \]

where

\[ \delta := n^2 \sigma^{-3} \sum_{i=1}^{n} |d_i| \sum_{i=1}^{n} |y_i|^3 \]
\[ \eta := \left\{ D_2 Y_2 / n D_1 Y_1 \right\}^{\frac{1}{2}}. \]

**Remark.** \( \eta \) represents the cost of ignoring the contribution of \( X \) to \( U \), and \( \delta \) bounds the error in the Wald-Wolfowitz central limit theorem.

**Corollary 1.1.** If \( U_n \) is a sequence of quantities of the form (2.1) such that \( \delta_n \to 0 \) and \( \eta_n \to 0 \), then \( U_n \) is asymptotically normally distributed.

Because of the extra factor \( n \), it would be usual to expect that \( n D_1 Y_1 \gg D_2 Y_2 \), so that a normal approximation to the statistic \( U \) should be appropriate. However, both \( D_1 \) and \( Y_1 \) may be small, since they are sums of squares of the centred row averages \( d_i^* \) and \( y_i^* \) respectively: in particular, if all the row sums \( \sum_{j \neq i} d_{ij} \) are equal, \( D_1 \) degenerates to zero. It seems not to be unusual for the row sums of one or both of the arrays \( d \) and \( y \), arising from statistics of practical interest, to be sufficiently uniform for the inequality \( n D_1 Y_1 \gg D_2 Y_2 \) to be violated, and there is no a priori reason to suppose in such circumstances that \( W \) should be approximately normally distributed. However, Theorem 2 shows that \( W \) may in some cases be approximately normally distributed, even when \( X \) makes a significant contribution to \( U \).

In order to state Theorem 2, it is first necessary to define certain absolute moments of an array \( (a_{ij})_{1 \leq i \neq j \leq n} \). In these definitions, \((n)_k\) denotes \( n(n-1) \cdots (n-k+1) \), and sums of the form \( \sum_{i,j} (\sum_{i,j,k}, \sum_{i,j,k,m}) \) are always taken over all pairs (triples, quadruples) of distinct integers from \( \{1, 2, \ldots, n\} \). Let

\[
\begin{align*}
a_1 &:= \frac{1}{(n)_2} \sum_{i,j} |a_{ij}|; & a_5 &:= \frac{1}{(n)_3} \sum_{i,j,k} |a_{ij}^2 a_{ik}|; \\
a_2 &:= \frac{1}{(n)_3} \sum_{i,j,k} |a_{ij} a_{ik}|; & a_6 &:= \frac{1}{(n)_2} \sum_{i,j} |a_{ij}|^3; \\
a_3 &:= \frac{1}{(n)_4} \sum_{i,j,k,m} |a_{ij} a_{ik} a_{jm}|; & a_7 &:= \frac{1}{(n)_3} \sum_{i,j,k} |a_{ij} a_{ik} a_{jk}|; \\
a_4 &:= \frac{1}{(n)_4} \sum_{i,j,k,m} |a_{ij} a_{ik} a_{jm}|; & a_8 &:= \frac{1}{(n)_2} \sum_{i,j} a_{ij}^2.
\end{align*}
\]
Thus, if $A$ is the incidence matrix of a graph $\mathcal{A}$, $a_1 = a_6 = a_8$ and $a_2 = a_5$: in this case, $a_1$ describes the relative frequency of occurrence of edges in $\mathcal{A}$, $a_2$ the frequency of trees of order 3, $a_7$ the frequency of triangles, and $a_3$ and $a_4$ the frequencies of the two types of trees of order 4.

Define also

$$\epsilon := s^{-3}\{n^4(d_1^3 + d_1d_2 + d_3 + d_4)(y_1^3 + y_1y_2 + y_3 + y_4) + n^3(d_5 + d_1d_6)(y_5 + y_1y_6) + n^2d_6y_6\}, \quad (2.9)$$

and let $\epsilon'$ be defined in the same way, but based on the centred arrays $d'$ and $y'$ given by $d'_{ij} = d_{ij} - D; y'_{ij} = y_{ij} - Y$. Note that if $a$ is the incidence matrix of a graph $\mathcal{A}$, the quantities $a_1^2, a_1a_2, a_3$ and $a_4$ are related to the relative frequencies of the different configurations of three edges, excluding the triangle; $a_5$ and $a_1a_8$ to those of configurations of two edges; and $a_6$ to the single edge. Indeed, letting $v_i$ denote the degree of vertex $i$ in $\mathcal{A}$, and defining the degree moments $v_m := \sum_{i=1}^{n} v_i^m$, it follows that the first group of quantities are of order $n^{-3}v_3$, the second group of order $n^{-2}v_2$ and the third of order $n^{-1}v_1$, though the first two estimates may be rather coarse in extremely sparse graphs. Note also that, if $|a_{ij}| \leq A$ for all $i$ and $j$, all these quantities are bounded by $A^3$.

**Theorem 2.** There exist universal constants $K_2$ and $K'_2$ such that

$$d_1(\mathcal{L}(W), \mathcal{N}) \leq K'_2\epsilon' \leq K_2\epsilon. \quad (2.10)$$

**Corollary 2.1.** If $U_n$ is a sequence of quantities of the form (2.1) such that $\epsilon_n \to 0$ or $\epsilon'_n \to 0$, then $U_n$ is asymptotically normally distributed.

**Remark.** $\epsilon$ may be easier to evaluate than $\epsilon'$ when many array elements are zero. It is also easy to show that $\epsilon \leq 21s^{-3}n^4d_6y_6$ for all $n \geq 1$, leading to a simpler, but less useful, estimate than that of Theorem 2.

The statement of Theorem 2 does not explicitly involve the relative orders of magnitude of var $X$ and var $Z$. However, it can be shown that $\delta_n = O(\epsilon'_n)$ as $n \to \infty$, whereas arrays $d^{(n)}$ and $y^{(n)}$ can be constructed in such a way that $\delta_n \to 0$, $\eta_n \to 0$ and $\epsilon'_n \not\to 0$. Thus $\delta_n \to 0$ is a slightly weaker condition for convergence to the normal than $\epsilon'_n \to 0$ when $\eta_n \to 0$. 

holds, so that Corollary 2.1 is principally intended for use when \( \text{var } X^{(n)} \) is at least comparable with \( \text{var } Z^{(n)} \). This, however, tends to be precisely the case that occurs in practical applications.

The structure of the problem, when \( \text{var } Z = O(\text{var } X) \) and the normal approximation is inappropriate, is in general very complicated. However, by analogy with the behaviour of degenerate U-statistics, one would expect that \( U \) would often be well approximated by a weighted sum of chi-squared random variables. Theorem 3 gives conditions under which the sequence \( n^{-1}(U^{(n)} - \mathbb{E}U^{(n)}) \) converges to a limit belonging to a somewhat broader class of distributions.

The simplest way to study the behaviour of \( U \), in this case, is to use eigenfunction expansions for the matrices \( d \) and \( y \). This is a standard approach, and has been used by a number of authors in similar circumstances ([8],[9]). In order to obtain a limit theorem, some sort of regularity conditions must be imposed on the eigenfunction expansions. Either one has to assume that the eigenvalues and eigenfunctions for different \( n \) are related ([8]) or that the matrices \( d \) and \( y \) are obtained from fixed, well-behaved kernels. We shall adopt the latter approach.

We assume that \( d(u, v) \) and \( y(u, v) \) are two symmetric kernels defined on \( [0, 1] \times [0, 1] \), such that

\[
\|d\|_2^2 := \int_0^1 \int_0^1 d^2(u, v) \, du \, dv < \infty;
\]

\[
\|y\|_2^2 := \int_0^1 \int_0^1 y^2(u, v) \, du \, dv < \infty.
\]

Let \( I_{ni} = \{ u : \frac{i-1}{n} < u \leq \frac{i}{n} \} \), \( 1 \leq i \leq n \), and define \( d_{ij} = d^{(n)}_{ij} \) \( (y_{ij} = y^{(n)}_{ij}) \) to be the average of \( d(u, v) \) \( (y(u, v)) \) over \( I_{ni} \times I_{nj} \). Since the symmetric kernels \( d(u, v) \) and \( y(u, v) \) are square integrable, they can be expressed in terms of eigenfunction expansions

\[
d(u, v) = \sum_{r=0}^{\infty} \alpha_r \phi_r(u) \phi_r(v),
\]

\[
y(u, v) = \sum_{s=0}^{\infty} \beta_s \psi_s(u) \psi_s(v),
\]

(2.13) (2.14)
where the sums converge in $L_2$,
\[
\sum_{r=0}^{\infty} \alpha_r^2 = \|d\|_2^2, \quad \sum_{s=0}^{\infty} \beta_s^2 = \|y\|_2^2
\]
and
\[
\int_0^1 \phi_r(x) \phi_s(x) \, dx = \int_0^1 \psi_r(x) \psi_s(x) \, dx = \delta_{rs}.
\]

The kernel $d$ is said to be degenerate if one of its eigenvectors, without
loss of generality $\phi_0$, is constant. In this case,
\[
\int_0^1 d(u, v) \, du = \int_0^1 \int_0^1 d(u, v) \, dudv := \bar{d}
\]
is constant in $v$, and
\[
\bar{\phi}_r := \int_0^1 \phi_r(x) \, dx = 0, \quad r = 1, 2, \ldots.
\]
Thus, if the kernel $d$ is degenerate, the row sums of $d^{(n)}$ are almost constant,
in that
\[
d^{(n)}_{i} = \frac{1}{n-2} \sum_{j \neq i} (d^{(n)}_{ij} - D^{(n)})
\]
\[
= -\frac{n}{n-2} \left\{ n \int_{l_{ni}^2} d(u, v) \, dudv - \int_{\Delta_n} d(u, v) \, dudv \right\},
\]
where $\Delta_n = \cup_{i=1}^{n} l_{ni}^2$, and so
\[
D^{(n)}_1 = \frac{1}{n} \sum_{i=1}^{n} (d^{(n)}_{i})^2 \leq \frac{1}{n(n-2)^2} \sum_{i=1}^{n} \left\{ n^2 \int_{l_{ni}^2} d(u, v) \, dudv \right\}^2
\]
\[
\leq \frac{n}{(n-2)^2} \int_{\Delta_n} d^2(u, v) \, dudv.
\]
Hence
\[
n D^{(n)}_1 Y^{(n)}_1 = O\left\{ \|y\|_2^3 \int_{\Delta_n} d^2(u, v) \, dudv \right\} = o(1) \quad \text{as} \quad n \to \infty,
\]
whereas
\[
\lim_{n \to \infty} D^{(n)}_2 Y^{(n)}_2 =
\left( \int_0^1 \int_0^1 \{d(u, v) - \bar{d}\}^2 \, dudv \right) \left( \int_0^1 \int_0^1 \{y(u, v) - \bar{y} - \bar{y} + \bar{y}\}^2 \, dudv \right),
\]
where \( y^*(u) := \int_0^1 \{ y(u, v) - y \} \, dv \), and is strictly positive unless \( d(u, v) = d \) for all \( u \) and \( v \) or \( y(u, v) = y + y^*(u) + y^*(v) \). Hence, excluding these trivial cases, if the kernel \( d \) is degenerate, the variance of \( X^{(n)} \) is strictly of order \( n^2 \), as compared with that of \( Z^{(n)} \) which is \( o(n^2) \), and Condition (i) of Theorem 1 is not satisfied: nor is Theorem 2 useful, since \( \epsilon_n \) is then of strict order \( n \). On the other hand, if neither \( d \) nor \( y \) is degenerate,

\[
D_1^{(n)} Y_1^{(n)} \rightarrow \int_0^1 \{ d^*(u) \}^2 \, du \cdot \int_0^1 \{ y^*(u) \}^2 \, du > 0,
\]

and Theorem 1 is applicable. Thus we are particularly interested in theorems for the case when one at least of the kernels \( d \) and \( y \) is degenerate, such as Theorem 3 below. The proof of Theorem 3 is presented in Section 5: of course, the assumption that \( d \) rather than \( y \) be degenerate is without loss of generality.

**Theorem 3.** Assume that the kernel \( d \) is degenerate. Then

\[
n^{-1} \{ U^{(n)} - \mathbb{E} U^{(n)} \} \overset{\mathbb{D}}{\longrightarrow} V,
\]

where

\[
V := \sum_{r \geq 1} \sum_{s=0}^{\infty} \alpha_r \beta_s \{ Z_{rs}^2 - \mathbb{E} Z_{rs}^2 \},
\]

the sums converging in \( L_2 \), and where the \( Z_{rs} \) are jointly normally distributed with zero mean and covariance

\[
\text{cov} ( Z_{rs}, Z_{r's'} ) = \delta_{rr'} \int_0^1 \{ \psi_s(v) - \bar{\psi}_s \} \{ \psi_{s'}(v) - \bar{\psi}_{s'} \} \, dv.
\]

If both \( d \) and \( y \) are degenerate, \( V \) takes the simpler form

\[
\sum_{r \geq 1} \sum_{s \geq 1} \alpha_r \beta_s \{ Z_{rs}^2 - 1 \},
\]

where the \( Z_{rs}^2 \) are independent \( \chi_1^2 \) random variables.

**Remark.** If \( d \) is degenerate and \( y(u, v) = y + y^*(u) + y^*(v) \), \( V = 0 \) a.s.

Theorem 3 can be extended to include sequences \( U^{(n)} \) based on matrices \( d^{(n)} \) and \( g^{(n)} \) which are only approximately of the kernel form prescribed in Theorem 3. Suppose, indeed, that

\[
d_{ij}^{(n)} = \delta_{ij}^{(n)} + \epsilon_{ij}^{(n)}; \quad y_{ij}^{(n)} = \gamma_{ij}^{(n)} + \eta_{ij}^{(n)} ,
\]
where $\hat{d}^{(n)}$ and $\hat{y}^{(n)}$ are based on symmetric $L_2$-finite kernels $\hat{d}$ and $\hat{y}$, such that $\hat{d}$ is degenerate.

**Corollary 3.1.** Suppose that, as $n \to \infty$,

\begin{align*}
(i) \quad & \sum_{i=1}^{n} (\varepsilon_i^{(n)})^2 \to 0; \\
(ii) \quad & n^{-2} \sum_{i,j} (\xi_i^{(n)})^2 \to 0; \quad n^{-2} \sum_{i,j} (\eta_i^{(n)})^2 \to 0.
\end{align*}

Then $n^{-1}(U^{(n)} - \mathbb{E}U^{(n)}) \xrightarrow{d} V$, where $V$ has the limit distribution of Theorem 3 derived from the kernels $\hat{d}$ and $\hat{y}$. If both $\hat{d}$ and $\hat{y}$ are degenerate, condition (i) may be replaced by

\begin{align*}
(iii) \quad & n^{-1} \sum_{i=1}^{n} (\varepsilon_i^{(n)})^2 \sum_{i=1}^{n} (\eta_i^{(n)})^2 \to 0.
\end{align*}

Theorem 3 and Corollary 3.1 improve upon the results of [9], by allowing any symmetric $L_2$ kernel $\mathcal{y}$ instead of the kernel $\mathcal{I} \{ u \leq \lambda \} \mathcal{I} \{ v \leq \lambda \}$ appropriate to the two sample problem, and by removing the condition that $\hat{d}$ be of trace-class.

### 3. GRAPH COLOURING.

The graph colouring problems studied in [4] and by Barton (unpublished lecture notes) using the method of moments can be more completely analysed by use of Theorem 2. The setting is as follows. The $n$ vertices of a graph $\mathcal{G}$ are randomly assigned coloured labels, by sampling without replacement from a set of labels of $c$ different colours: there are $n_i$ labels of colour $i$, $1 \leq i \leq c$, and $\sum_{i=1}^{c} n_i = n$. The quantity $U_n$ is taken to be

$$U_n := \sum_{(i,j) \in \mathcal{G}^{(n)}} f(C_n(i), C_n(j)),$$

where $C_n(i)$ denotes the colour assigned to vertex $i$ and the function $f$ is non-constant. Barton proved that the distribution of $U_n$ was asymptotically
normal, provided that

(i) \( \lim_{n \to \infty} n_i/n = p_i > 0, \quad 1 \leq i \leq c; \)
(ii) \( \limsup_{n \to \infty} n^{-1} M_n(\Gamma) < \infty \) for all connected graphs \( \Gamma; \)
(iii) \( \lim_{n \to \infty} n^{-1} M_n(A) = a; \)
(iv) \( \lim_{n \to \infty} n^{-1} M_n(C) = g, \)

for some \( a, g \in \mathbb{R}^+ \) such that not both are zero, where \( M_n(\Gamma) \) denotes
the number of distinct sub-graphs of \( \mathcal{G}^{(n)} \) isomorphic to \( \Gamma \), and \( A \) and \( C \)
denote respectively a single edge and a tree of order three: it should also
be assumed that \( f(u, v) \) cannot be expressed in the form \( h(u) + h(v) \).

To show that Corollary 2.1 can be used to prove Barton’s theorem, take

\[ y_{ij}^{(n)} = f(C_n(i), C_n(j)) \]  \hspace{1cm} (3.2)

and let \( d^{(n)} \) be the incidence matrix of \( \mathcal{G}^{(n)} \). It is then immediate from the
assumptions on \( f \) that there exist \( n \)-independent constants \( K \) and \( L \) such that

\[ \max_{i,j} |y_{ij}^{(n)}| \leq K; \quad Y_2^{(n)} \geq L > 0, \]  \hspace{1cm} (3.3)

from which it follows readily that

\[ y_i^{(n)} \leq \min(K, K^3), \quad 1 \leq i \leq 8, \]

and that \( n^2 D_2^{(n)} = O(s_n^2) \). Furthermore, applying assumption (ii) to edges
and to trees of orders three and four, it follows that

\[ d_1^{(n)} = d_0^{(n)} = d_8^{(n)} = O(n^{-1}); \quad d_2^{(n)} = d_5^{(n)} = O(n^{-2}); \]
\[ d_3^{(n)} = O(n^{-3}); \quad d_4^{(n)} = O(n^{-3}), \]

and so \( s_n^3 \varepsilon_n = O(n) \). Finally, if \( v_i^{(n)} \) denotes the degree of vertex \( i \) in \( \mathcal{G}^{(n)} \),
and the degree moments are denoted by

\[ v_m^{(n)} := \frac{1}{n} \sum_{i=1}^{n} \left\{ v_i^{(n)} \right\}^m, \quad m \geq 1, \]
assumption (ii) applied to trees of order 4 implies that

\[ v_1^{(n)} \leq M \quad \text{for some constant } M, \quad (3.4) \]

whereas Hölder’s inequality shows that \( v_1^{(n)} v_3^{(n)} \geq \{v_2^{(n)}\}^2 \), from which it follows that \( a \) cannot be zero in assumption (iii). Then it is easy to show, again using (3.4), that \( D_2^{(n)} \sim 2an^{-1} \), and hence that \( n = O(\varepsilon_n^2) \). Thus, under Barton’s assumptions, \( \varepsilon_n = O(n^{-1/2}) \), and Corollary 2.1 can be applied. Note also that \( n^3 D_1^{(n)} Y_1^{(n)} = O(n) \) and \( n^2 D_2^{(n)} Y_2^{(n)} \sim n \), so that Corollary 1.1 is of no use here.

Although the definition of \( \varepsilon_n \) is quite explicit, it can be advantageous to have sufficient conditions for convergence to the normal which are easier to use in practice than that of Corollary 2.1. The degree moments are particularly appealing in this respect: indeed, it has already been observed that, if \( d \) is the incidence matrix of a graph \( G \) with degree moments \( \bar{v}_m \), its contribution to the numerator of \( \varepsilon \) is easily estimated in terms of \( \bar{v}_m \), \( 1 \leq m \leq 3 \). Universal estimates of the factor \( s^{-3} \) are somewhat more difficult, and, since it is not unusual for one of \( D_1 \) and \( Y_1 \) to be zero, it is best to restrict attention to lower bounds for \( D_2 \). Here, the natural lower bound is of the form \( n D_2 \geq k \bar{v}_1 \) for a suitable constant \( k \), but some restriction has to be imposed upon \( G \) to make it hold. However, the relation

\[
1 \geq (n - 1) D_2 / \bar{v}_1 = 1 - 2(\bar{v}_2 / \bar{v}_1)(n - 2)^{-1} + n \bar{v}_1 (n - 2)^{-2}
\]

implies that \( n D_2^{(n)} \geq k \bar{v}_1^{(n)} \) for some \( k > 0 \) and for all \( n \) sufficiently large, provided that

\[
\alpha(G^{(i)}) := \limsup_{n \to \infty} \bar{v}_2^{(n)} / \{n \bar{v}_1^{(n)}(1 + n^{-1} \bar{v}_1^{(n)})\} < \frac{1}{2}. \quad (3.5)
\]

Combining these considerations with Corollary 2.1 leads to the following improvement to and simplification of Barton’s theorem, with the conditions on \( d \) expressed only in terms of the first and third degree moments.

**Theorem 4.** Suppose that the matrices \( y^{(n)} \) satisfy (3.3) and that the matrices \( d^{(n)} \) are incidence matrices of graphs \( G^{(n)} \). Then \( U_n \) is asymptotically normally distributed if

\[
n^{-1/2} \varepsilon_3^{(n)} \{v_1^{(n)}\}^{-3/2} \to 0.
\]
Remarks. 1. The condition \( n^{-1/2} \varphi_3^{(n)}/\varphi_1^{(n)} \rightarrow 0 \) implies that \( n^{-1/3} \varphi_2^{(n)}/\varphi_1^{(n)} \rightarrow 0 \), so that (3.5) is easily satisfied.

2. Note that, under Barton's conditions, the quantities \( \psi_m^{(n)} \) are constrained to be of order 1 for all \( m \geq 1 \), and \( \psi_1^{(n)} \propto 1 \).

3. If \( f(u, v) = h(u) + h(v) \), the matrices \( y^{(n)} \) defined by (3.2) need not satisfy (3.3). For example, in the case \( c = 2 \), define \( h(1) = 0 \) and \( h(2) = 1 \), and suppose that \( n_1 = n_2 = n/2 \) for even integers \( n \): then \( Y_2^{(n)} = 0 \) for all even \( n \), in violation of (3.3). In this case, if \( G^{(n)} \) is a graph with \( (n - 4)/2 \) isolated edges together with a cycle of order 4 with vertex set \( \{1, 2, 3, 4\} \), \( \varphi_1^{(n)} = 1 + 4/n \) and \( \varphi_3^{(n)} = 1 + 28/n \), so that \( n^{-1/2} \varphi_2^{(n)}/\varphi_1^{(n)} \rightarrow 0 \). However, \( U_n = n + 2 \sum_{i=1}^{n} h(C_n(i)) \) has a distribution concentrated on five distinct values, and is hence not asymptotically normally distributed.

4. Barton considered an explicitly calculable example, with \( G^{(n)} \) a star graph with \( u_n + 1 \) vertices together with \( [(n - u_n - 1)/2] \) isolated edges. For suitably chosen matrices \( y^{(n)} \), the distribution of \( U_n \) was shown to be asymptotically normal for \( u_n = o(n^{1/2}) \) and non-normal for \( u_n \propto n^{1/2} \), by direct argument. Barton's conditions imply asymptotic normality here only if \( \psi_m^{(n)} = O(n) \) for all \( m \geq 1 \), whereas the condition of Theorem 4 is sharp.

Bloemena [4], Theorems 4.1.1 and 4.1.2, considered the case \( c = 3 \), with the functions \( f \) given by

\[
f(1, 1) = 1, \quad f(i, j) = 0 \quad \text{otherwise},
\]

and

\[
f(2, 1) = f(1, 2) = 1, \quad f(i, j) = 0 \quad \text{otherwise},
\]

corresponding to counting the numbers of black-black and black-white joins in \( G^{(n)} \) respectively. He restricted \( G^{(n)} \) by requiring that

\[
1 \leq \psi_i^{(n)} \leq d, \quad \text{for all } n \text{ and } 1 \leq i \leq n,
\]  

(3.6)

but allowed \( n_1, n_2 \) and \( n_3 \) to vary with \( n \) more freely that (3.3) would necessarily permit, so that Theorem 4 is not always applicable. These problems are however equivalent to counting the number of edges, under random association of vertices, common to \( G^{(n)} \) and \( H_1^{(n)} \) or \( H_2^{(n)} \), where
\( \mathcal{K}_1^{(n)} \) is a complete graph on \( n_1 \) vertices together with \( n - n_1 \) isolated points, and \( \mathcal{K}_2^{(n)} \) is a graph in which each member of a set of \( n_1 \) vertices is joined to each member of a disjoint set of \( n_2 \) vertices, and no other edges are present. Thus, in this context, it is more natural to look for conditions under which the number of edges in common to randomly associated graphs \( \mathcal{G}^{(n)} \) and \( \mathcal{H}^{(n)} \) is asymptotically normally distributed, expressed in terms of the vertex moments \( \psi_m^{(n)} \) of \( \mathcal{G}^{(n)} \) and \( \psi_m^{(n)} \) of \( \mathcal{H}^{(n)} \). The following theorem is then a consequence of Corollary 2.1.

**Theorem 5.** Suppose that \( d^{(n)} \) and \( y^{(n)} \) are incidence matrices corresponding to graphs \( \mathcal{G}^{(n)} \) and \( \mathcal{H}^{(n)} \) respectively. Then \( U_n \) is asymptotically normally distributed if

\[
\alpha(\mathcal{G}^{(1)}) < \frac{1}{2}, \quad \alpha(\mathcal{H}^{(1)}) < \frac{1}{2}
\]

and

\[
\eta_n := \max(n^{-2}\psi_3^{(n)}\psi_3^{(n)}, \psi_1^{(n)}\psi_1^{(n)}) \{\psi_1^{(n)}\psi_1^{(n)}\}^{-3/2} \to 0.
\]

**Remark.** Since, in this setting, \( \mathbb{E}U_n = \psi_1^{(n)}\psi_1^{(n)} \) and \( U_n \) takes only non-negative integer values, the condition \( \psi_1^{(n)}\psi_1^{(n)} \to \infty \) is clearly necessary for \( U_n \) to be asymptotically non-trivially normally distributed.

Now suppose that \( \mathcal{G}^{(n)} \) satisfies (3.6) and that \( \mathcal{H}^{(n)} = \mathcal{K}_1^{(n)} \). Then \( \psi_1^{(n)} \asymp 1, \psi_2^{(n)} \asymp 1, \psi_3^{(n)} \asymp 1, \psi_1^{(n)} \sim n_1^2/n, \psi_2^{(n)} \sim n_1^2/n \) and \( \psi_3^{(n)} \sim n_1^2/n \). Hence \( \alpha(\mathcal{G}^{(1)}) = 0 \), and \( \alpha(\mathcal{H}^{(1)}) < \frac{1}{2} \) provided that \( \limsup_{n \to \infty} (n_1/n) < 1 \). Since also \( \eta_n \asymp n_1^{1/2}/n_1 \), it follows from Theorem 5 that \( U_n \) is asymptotically normally distributed if \( n_1^2/n \to \infty \) and \( \limsup_{n \to \infty} (n_1/n) < 1 \), yielding Bloemena's Theorem 4.1.1. If now \( \mathcal{K}_2^{(n)} \) is taken instead of \( \mathcal{K}_1^{(n)} \), and \( 1 \leq n_1 \leq n_2 \) without loss of generality, \( \psi_1^{(n)} \sim 2n_1n_2/n, \psi_2^{(n)} \sim n_1n_2(n_1 + n_2)/n \) and \( \psi_3^{(n)} \sim n_1n_2(n_1^2 + n_2^2)/n \), yielding \( \eta_n \asymp (n/n_1n_2)^{1/2} \) and \( \alpha(\mathcal{H}^{(1)}) < \frac{1}{2} \) whenever \( \limsup_{n \to \infty} (n_2/n) < 1 \), and hence Bloemena's Theorem 4.1.2. Of course, the conclusions still remain true if (3.6) is relaxed to \( \psi_1^{(n)} \asymp 1, \psi_3^{(n)} \asymp 1 \).

As in the remark following Theorem 5, the conditions \( n_1^2/n \to \infty \) and \( n_1n_2/n \to \infty \) in these examples are necessary for \( U_n \) to have a non-trivial normal approximation. The remaining conditions, however, merely indicate
that the natural lower bound on $s_n^2$ does not hold. Sharp results can in fact be devised, under a global set of conditions on $G^{(n)}$ which is significantly broader than (3.6) imposed by Bloemena, and which in particular allows vertices to be isolated.

**Theorem 6.** Suppose that the graphs $G^{(n)}$ satisfy

$$v_1^{(n)} \asymp 1, \quad v_3^{(n)} \asymp 1,$$

(3.7)

and the Lindeberg condition

$$\lim_{\eta \to \infty} \limsup_{n \to \infty} \frac{1}{n \sigma_n^2} \sum_{i=1}^{n} |v_i^{(n)} - v_1^{(n)}|^2 I[|v_i^{(n)} - v_1^{(n)}| > \eta] = 0,$$

(3.8)

where $\sigma_n^2 = v_2^{(n)} - \{v_1^{(n)}\}^2$.

(i) If $\lambda^{(n)} = \lambda_1^{(n)}$, then $U_n$ is non-trivially asymptotically normally distributed as $n \to \infty$ if and only if

$$n_1^2/n \to \infty \quad \text{and} \quad (n - n_1)^2/n + (n - n_1)\sigma_n^2 \to \infty.$$  

(3.9)

(ii) If $\lambda^{(n)} = \lambda_2^{(n)}$ and $1 \leq n_1 \leq n_2$, then $U_n$ is non-trivially asymptotically normally distributed as $n \to \infty$ if and only if

$$n_1 n_2/n \to \infty \quad \text{and} \quad n_1 (n - n_2)/n + n_1 \sigma_n^2 \to \infty.$$  

(3.10)

**Remarks.** 1. The norming sequence $\{\text{var} \, U_n\}^{1/2}$ is appropriate in either case.

2. Theorem 6(i) generalises Bloemena's Theorem 4.1.1 for black-black joins, and Theorem 6(ii) his Theorem 4.1.2 for black-white joins.

3. Condition (3.8) is only used in case (i) for $n - n_1 = O(n^{3/4})$, and in case (ii) for $n_1 = O(n^{3/4})$ and $n - n_2 = o(n)$.

**Proof.** We consider case (ii) in detail; case (i) is similar, but easier. The index $n$ is, where possible, suppressed, and $n_3$ is used to denote $n - n_1 - n_2$.

Note first that, as $n \to \infty$,

$$\text{var} \, U \sim v_1 w_1 f(v_1, v_2) f(w_1, w_2) + n^{-1} \sigma^2 r^2,$$

(3.11)
where
\[ f(x, y) := 1 - \frac{2y}{(n-2)x} + \frac{nx}{(n-2)^2} \]
and \( r^2 := \bar{w}_2 - \bar{w}_1^2 \). It follows from (3.7) that \( f(\bar{v}_1, \bar{v}_2) \sim 1 \). Then, since
\[ \bar{w}_1 = 2n_1n_2/n; \quad \bar{w}_2 = n_1n_2(n_1 + n_2)/n, \tag{3.12} \]
the relation
\[ f(\bar{w}_1, \bar{w}_2) \times (n_1 + n_3)/n \tag{3.13} \]
holds, except when \( n_1 = 1 = n - n_2 \), when \( f(\bar{w}_1, \bar{w}_2) \asymp n^{-2} \); and also
\[ r^2 = n_1n_2\{n(n_1 + n_2) - 4n_1n_2\}/n^2 \leq n_1n_2. \tag{3.14} \]

For \( U \) to be non-trivially asymptotically normal, it is necessary that \( \mathbb{E}U \to \infty \) and \( \text{var} U \to \infty \), because \( U \) takes only non-negative integer values. If \( n_1n_2/n \not\to \infty \), \( \mathbb{E}U \not\to \infty \), and so \( n_1n_2/n \to \infty \) is a necessary condition. On the other hand, using (3.7), (3.11),(3.13) and (3.14),
\[ \text{var} U = O(n_1n_2(n_1 + n_3)/n^2 + \sigma^2 r^2/n) = O(n_1(n_1 + n_3)/n + n_1\sigma^2), \]
from which it follows that \( n_1(n - n_2)/n + n_1\sigma^2 \to \infty \) is also necessary for \( U \) to be asymptotically normal.

Since also \( s^3 \epsilon \), as in (2.9), is easily estimated by
\[ \bar{v}_1 \bar{w}_1 + n^{-2} \bar{v}_3 \bar{w}_3 \asymp n_1n_2/n, \]
it follows from (3.11) and Theorem 2 that
\[ d_1(\mathcal{L}(W), \mathcal{N}) = O((n_1n_2/n)\{n_1n_2(n_1 + n_3)/n^2 + \sigma^2 r^2/n\}^{-3/2}), \tag{3.15} \]
where, as usual, \( W = \delta^{-1}(U - \mathbb{E}U) \). Thus \( d_1(\mathcal{L}(W), \mathcal{N}) \) certainly tends to zero when \( n_1n_2/n \to \infty \), unless \( n - n_2 = o(n) \), in which case
\[ n_1^{1/3}(1 - n_2/n) \to \infty \tag{3.16} \]
is still sufficient. Since also \( r^2 \sim n_1n \) for \( n - n_2 \leq n/4 \), \( d_1(\mathcal{L}(W), \mathcal{N}) \) also tends to zero if \( n - n_2 = o(n) \) and
\[ n_1^{1/3}\sigma^2 \to \infty. \tag{3.17} \]
However, these considerations based on (3.15) alone are not quite sharp enough to establish the theorem.

In order to complete the proof, realise \( G \) and \( H \) together as follows. Choose \( n_1 \) indices from \( N := \{1, 2, \ldots, n\} \) uniformly and at random, and call this subset \( N_1 \); then choose a further \( n_3 \) uniformly and at random from \( N \setminus N_1 \), and call this subset \( N_3 \); \( H \) is defined to be \( \{(i, j), \quad i \in N_1, \quad j \in N \setminus (N_1 \cup N_3)\} \). \( G \) is then associated with \( H \) by assigning labels to the \( n \) vertices of \( G \) at random. \( U \) can now be rewritten as

\[
U = n_1 v + \sum_{i \in N_1} (v_i - v) - \sum_{i \in N_1} I[(i, j) \in G] - \sum_{i, j \in N_1} I[(i, j) \in G] \tag{3.18}
\]

\[
= n_1 v + U_1 - U_2 - U_3,
\]
say. Now \( \mathbb{E} U_1 = 0 \), \( \text{var} U_1 \sim \sigma_1^2 n_1 (1 - n_1/n) \), and, from Hajék’s [6] version of the Wald-Wolfowitz theorem, using the Lindeberg condition (3.8), \( U_1/\{\text{var} U_1\}^{\frac{1}{2}} \) is asymptotically distributed as \( \mathcal{N}(0, 1) \) if \( n_1 \sigma_1^2 \to \infty \). On the other hand, for \( n - n_2 \leq n/4 \),

\[
\text{var} (U_2 + U_3) \propto n_1 (n_1 + n_3)/n, \tag{3.19}
\]

and \( \epsilon \) from (2.9) computed for \( U_2 + U_3 \) satisfies

\[
\epsilon = O\left(\{n_1 (n_1 + n_3)/n\}-\frac{1}{2}\right). \tag{3.20}
\]

Hence, using Theorem 2, if \( n_1 (n_1 + n_3)/n + n_1 \sigma_1^2 \to \infty \) and either \( n \sigma_2/(n_1 + n_3) \to 0 \) or \( n \sigma_2/(n_1 + n_3) \to \infty \), so that one of \( U_1 \) or \( U_2 + U_3 \) dominates, then \( W \) is asymptotically \( \mathcal{N}(0, 1) \).

When \( \sigma^2 \propto (n_1 + n_3)/n \), it is necessary to refine (3.18) by an appropriate coupling. The case \( \sigma^2 \propto (n_1 + n_3)/n \propto 1 \), \( n_1 n_2/n \to \infty \) is already covered by (3.15), and so we consider only \( \sigma^2 \to 0 \), in which case it follows from the definition of \( \sigma \), and because the \( v_i \)'s take only integer values, that there exists an integer \( v^* \) such that, if \( M := \{i : v_i \neq v^*\} \) and \( m := |M| \), \( m/n \to 0 \) as \( n \to \infty \). Now define

\[
N'_1 := (N_1 \cap M^c) \cup N_1^*; \quad N'_3 := (N_3 \cap M^c) \cup N_3^*,
\]
where $N_1^*$ is formed by taking $|N_1 \cap M|$ points uniformly and at random from $M^c \setminus (N_1 \cup N_3)$, and $N_3^*$ by taking $|N_3 \cap M|$ points uniformly and at random from $M^c \setminus (N_1 \cup N_3 \cup N_1^*)$. Let

$$U' = U_2 + U_3' := \sum_{i \in N_1, j \in N_3} I[(i, j) \in G] + \sum_{i, j \in N_1^*} I[(i, j) \in G].$$

(3.21)

Since $U_1 = \sum_{i \in N_1 \cap M} (v_i - v^*) + n_1 (v^* - \bar{v})$, it is independent of $U_2' + U_3'$. It therefore suffices to show that

$$\{U_2' + U_3' - \mathbb{E}(U_2' + U_3')\}/\{\text{var}(U_2' + U_3')\}^{1/2} \rightarrow \mathcal{N}(0, 1) \quad (3.22)$$

when $n_1(n_1 + n_3)/n \rightarrow \infty$ and $\sigma^2 \times (n_1 + n_3)/n \rightarrow 0$, and that var $(U_2' - U_2)$ and var $(U_3' - U_3)$ are negligible in comparison to var $U$.

In order to establish (3.22), consider the graph $G'$ on $n - m$ vertices obtained by removing the vertex set $M$ and all edges incident to vertices of $M$ from $G$; let its degree moments be denoted by $\bar{v}_r'$, $r \geq 1$. Since, for any $\eta > 0$,

$$0 \leq \sum_{i \in M} v_i \leq m\bar{v}_1 + \sum_{i \in M} (v_i - \bar{v}_1) \leq m(\bar{v}_1 + \eta) + \frac{1}{2} \sum_{i \in M} |v_i - \bar{v}_1|^2 I[|v_i - \bar{v}_1| > \eta], \quad (3.23)$$

it follows that $n^{-1} \sum_{i \in M} v_i = O(\sigma^2)$, and hence that, if $\sigma^2 \to 0$, $\bar{v}_r' \to \bar{v}_r \to 1$ as $n \to \infty$; also, clearly, $\bar{v}_3' = O(\bar{v}_3) = O(1)$. Define the symmetric $(n - m) \times (n - m)$ matrix $H'$ by setting $H'_{ij} = 0$ for all $i$, and, for $i < j$, setting

$$H'_{ij} = \begin{cases} 1, & 1 \leq i < j \leq n_1 \\ \frac{1}{2}, & 1 \leq i \leq n_1, \quad n_1 + 1 \leq j \leq n_1 + n_3 \\ 0, & \text{otherwise.} \end{cases}$$

Then the distribution of $U_2' + U_3'$ is that of the quantity in (2.1) obtained by taking $d = H'$ and $y = C'$, the incidence matrix of $G'$. Calculations similar to those leading to (3.20) now show that the quantity $\epsilon$ from (2.9) for $U_2' + U_3'$ satisfies

$$\epsilon = O\left(\{n_1(n_1 + n_3)/n\}^{-\frac{1}{2}}\right),$$
whenever \( n_1 + n_3 \leq n/4 \), and (3.22) follows from Theorem 2.

Now consider the variance of \( U_2 - U_2' \). We can write

\[
U_2 - U_2' = \left\{ \sum_{i \in M} Z_{ij} + \sum_{J \in M} Z_{ij} \right\} + \sum_{i \in M} Z_{ij} - \sum_{i \in N_1} Z_{ij} - \sum_{i \in N_2} Z_{ij},
\]

where \( Z_{ij} := I[(i, j) \in \mathcal{G}] \). Take the term in braces. Define \( \mathcal{G}' \) to consist of those edges of \( \mathcal{G} \) connecting a vertex of \( M \) with a vertex of \( M^c \), and let \( \nu^n_t \) denote the corresponding degree moments. Note that, from (3.23), \( \nu^n_t = O(\sigma^2) \). Let \( \mathcal{H}'_2 \) denote a graph of the form \( \mathcal{H}_2 \) with \( n_1 \) and \( n_3 \) (instead of \( n_2 \)) elements in the special vertex sets. Then the quantity in braces is the number of edges in common to \( \mathcal{G}' \) and \( \mathcal{H}'_2 \), and its variance, from (3.11)-(3.14) with \( n_3 \) for \( n_2 \), is found to be of order

\[
O(n_1 n_3 \nu^n_t / n + n_1 n_3 (n_1 + n_3) / n^2) = o(\text{var } U),
\]

when \( \sigma^2 \asymp (n_1 + n_3) / n \to 0 \). The variance of the next term in \( U_2 - U_2' \) is shown to be negligible in a similar manner. For \( U_4 := \sum_{i \in N_1, j \in N_3} Z_{ij} \), suppose first that \( |N_1| = l \). Then its distribution is that of the intersection of \( \mathcal{G}' \) and an \( \mathcal{H}'_2 \)-graph on \( n - m \) vertices, where the special vertex sets have cardinality \( l \) and \( n_3 \). Hence

\[
\mathbb{IE}(U_4 | l) = l n_3 \nu^n_t / (n - m - 1)
\]

and

\[
\text{var}(U_4 | l) = O(n^{-1} l n_3 \nu^n_t + n^{-2} l n_3 (l + n_3) (\sigma^2)^t),
\]

whence, using the known hypergeometric distribution of \( l \), it follows easily that

\[
\text{var } U_4 = O(m n_1 (n_1 + n_3) / n^2) = o(\text{var } U)
\]

when \( \sigma^2 \asymp (n_1 + n_3) / n \to 0 \). The remaining term in \( U_2 - U_2' \) is treated similarly, and the variance of \( U_3 - U_3' \) is estimated in an analogous fashion. Combining the various cases using a standard argument, part (ii) of the theorem follows.

In part (i), it is advantageous to replace \( \mathcal{H}_1 \) by its complement when \( \binom{n_1}{2} > \frac{1}{2} \binom{n}{2} \); the remaining argument is similar to that used to prove part (ii), but simpler. \( \square \)
The preceding results have concentrated on normal approximation, and the tool found to be appropriate was Theorem 2. The following is an example of the application of Theorem 3 to graph colouring problems. Let $U_n$ be the quantity defined in (3.1), where $n$ runs through $c\mathbb{Z}$ for some fixed $c$ and $n_i = n/c$ for each $i$, and let

$$f(i, j) = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

Thus $U_n$ counts twice the number of edges in $G^{(n)}$ joining two vertices of the same colour, and the numbers of vertices of each colour are the same. Take $G^{(n)}$ to be the complete graph on vertices $\{1, 2, \ldots, [an]\}$, so that the associated matrix $d^{(n)}$ can be expressed as

$$d_{ij}^{(n)} = n^2 \int_{I_{n_i} \times I_{n_j}} a.a^{-\frac{1}{2}} I[0 \leq u \leq a].a^{-\frac{1}{2}} I[0 \leq v \leq a] \, dudv + \epsilon_{ij}^{(n)},$$

where $n^{-2} \sum_{i,j} \epsilon_{ij}^{(n)} = 0$. The matrix $y^{(n)}$ for the colour graph is given by

$$y_{ij}^{(n)} = n^2 \int_{I_{n_i} \times I_{n_j}} y(u, v) \, dudv,$$

where

$$y(u, v) := \sum_{j=1}^{c} I[(j - 1)/c \leq u \leq j/c] I[(j - 1)/c \leq v \leq j/c].$$

Note that all the eigenfunctions in the expansion of $y$ have eigenvalue $c^{-1}$, and that $y(u, v)$ can also be expressed in the form

$$y(u, v) = \frac{1}{c} \{ 1 + \sum_{r=1}^{c-1} \phi_r(u)\phi_r(v) \}$$

for suitable orthonormal $\phi_r$, making it clear that $y$ is degenerate. Applying Corollary 3.1 (with $d$ and $y$ exchanged), it follows that, as $n \to \infty$,

$$\frac{k}{na(1-a)} \{ U^{(n)} - EU^{(n)} \} \xrightarrow{D} \chi_{k-1}^2 - (k-1).$$
4. PROOF OF THEOREM 2.

Before beginning to prove Theorem 2, we require some subsidiary results.

For any permutation \( \pi \in S_n \), let \( \phi_{ij} \pi \in S_n \) be defined by

\[
(\phi_{ij} \pi)(i) = j; \quad (\phi_{ij} \pi)(\pi^{-1}(j)) = \pi(i);
\]

\[
(\phi_{ij} \pi)(t) = \pi(t), \quad t \notin \{i, \pi^{-1}(j)\}.
\]

Fix \( k \geq 0 \) and \( k \)-tuples \((s_1, \ldots, s_k)\) and \((t_1, \ldots, t_k)\) from \( \{1, \ldots, n\} \).

**Lemma 4.1.** Let \( \pi \) be uniformly distributed over \( D := \{ \pi \in S_n; \pi(s_m) = t_m, 1 \leq m \leq k \} \), and suppose that \( i \notin \{s_1, \ldots, s_k\} \) and \( j \notin \{t_1, \ldots, t_k\} \). Then \( \phi_{ij} \pi \) is uniformly distributed over all permutations that, in addition, satisfy \( \pi(i) = j \).

**Proof.** Let \( \rho \in S_n \) satisfy \( \rho(s_m) = t_m, 1 \leq m \leq k \) and \( \rho(i) = j \): there are \((n - k - 1)!\) such permutations. Then \( \text{Card}(\phi_{ij}^{-1}(\rho) \cap D) = n - k \), and so \( \mathbb{P}[\phi_{ij} \pi = \rho] = \mathbb{P}[\pi \in \phi_{ij}^{-1}(\rho)] = 1/(n - k - 1)! \). \( \square \)

**Corollary 4.1.1.** Let \( i \neq j, k \neq m \), and let \( \pi \) be uniformly distributed on \( S_n \). Then \( \phi_{km} \phi_{ij} \pi \) is uniformly distributed over \( \{ \pi \in S_n : \pi(i) = j, \pi(k) = m \} \).

**Proof.** Immediate. \( \square \)

**Lemma 4.2.** Fix \( k \geq 0 \) and \( k \)-tuples \( S := (s_1, \ldots, s_k) \) and \( T := (t_1, \ldots, t_k) \) from \( \{1, 2, \ldots, n\} \). Suppose also that \( i \notin S \) and \( j \notin T \). Then, if

\[
W := W(\pi) := s^{-1} \sum_{u,v} d_{uv} y_{\pi(u)\pi(v)}, \text{ where } \pi \text{ is uniformly distributed over } S_n, \text{ and if } g \in C_1,
\]

\[
|\mathbb{E}\{g(W) \mid \pi(s_m) = t_m, 1 \leq m \leq k; \pi(i) = j\} - \mathbb{E}\{g(W) \mid \pi(s_m) = t_m, 1 \leq m \leq k\}| \leq K\|g\|_0 s^{-1}\left\{ \frac{1}{m=1} \sum_{m=1}^k |d_{i,s_m}y_{j,t_m}| + n^{-1}\left[ d_{i,y} + \sum_{m=1}^k (|d_{s_m,y_{t_m}}| + |d_{s_m,y_{j,t_m}}|) \right] \right. \\
\left. + n^{-2}\left[ d_{i,y} + d_{..y} + \sum_{m=1}^k d_{s_m,y_{t_m}} \right] + n^{-3} d_{..y} \right\},
\]

where

\[
d_{ij} = \alpha_{ij} + \beta_{ij}, \quad \text{and } \alpha_{ij} = \sum_{i \neq j} \sum_{k \neq i} \frac{d_{ik}}{d_{kk}}, \quad \beta_{ij} = \sum_{k \neq i} \frac{d_{ik}}{d_{kk}}.
\]
where $K$ is a constant, $a_{uv} := \sum_{u \neq v} |a_{uv}|$ and $a_{..} := \sum_{u,v} |a_{uv}|$.

**Proof.** From Lemma 4.1, we can write

$$\left| \mathbb{E}\{g(W) \mid \pi(s_m) = t_m, 1 \leq m \leq k; \pi(i) = j\} - \mathbb{E}\{g(W) \mid \pi(s_m) = t_m, 1 \leq m \leq k\} \right|$$

$$= \left| \mathbb{E}\{g(W(\phi_{ij} \rho))\} - \mathbb{E}\{g(W(\rho))\} \right| \leq \|g'\| \mathbb{E}|W(\phi_{ij} \rho) - W(\rho)|,$$

where $\rho$ is uniformly distributed on $\{\rho \in S_n : \rho(s_m) = t_m, 1 \leq m \leq k\}$. Now

$$W(\phi_{ij} \rho) - W(\rho) = s^{-1} \sum_{(i,j) \neq (i',j')} (d_{ij} - d_{j^{-1}(j)ij}) (y_{ij}(z) - y_{j^{-1}(j)ij}(z)),$$

and so

$$\mathbb{E}|W(\phi_{ij} \rho) - W(\rho)|$$

$$\leq K s^{-1} \sum_{\alpha, \beta} \sum_{p,z} \left[ |d_{iz} y_{j\rho}| + |d_{iz} y_{\alpha \rho}| + |d_{pz} y_{j\rho}| + |d_{pz} y_{\alpha \rho}| \right]$$

$$+ \frac{1}{n^2} \sum_{m=1}^{k} \sum_{\alpha \in \{T, j\}} \sum_{p \in \{S, i\}} \left[ |d_{iz} y_{j\alpha m}| + |d_{iz} y_{\alpha \alpha m}| + |d_{pz} y_{j\alpha m}| + |d_{pz} y_{\alpha \alpha m}| \right],$$

where $\sum_{\alpha, \beta}$ denotes a sum over pairs of distinct integers $(\alpha, \beta)$ such that $\{\alpha, \beta\} \cap \{T, j\} = \emptyset$, and $\sum_{p,z}$ over pairs of distinct integers $(p, z)$ such that $\{p, z\} \cap \{S, i\} = \emptyset$. The statement of Lemma 4.2 is now obtained by overestimating each sum, by taking in addition terms excluded by the restrictions on the ranges of $\alpha, \beta, p$ and $z$. $\square$

**Corollary 4.2.1.** For each $k \geq 1$ and for each pair of $k$-tuples $(s_1, \ldots, s_k)$ and $(t_1, \ldots, t_k)$ from $\{1, 2, \ldots, n\}$,

$$\left| \mathbb{E}\{f'(W) \mid \pi(s_m) = t_m, 1 \leq m \leq k\} - \mathbb{E}f'(W) \right|$$

$$\leq K \|f''\| s^{-1} \left\{ \sum_{1 \leq m < p \leq k} |d_{s_m s_p y_{t_m t_p}} + n^{-1} \sum_{m=1}^{k} d_{s_m y_{t_m}} + n^{-2} \sum_{m=1}^{k} d_{s_m} + d_{..} \sum_{m=1}^{k} y_{t_m} + n^{-3} d_{..} y_{..} \right\},$$
where $K$ is a constant.

**Proof.** Immediate, from iteration of Lemma 4.2. \( \square \)

**Lemma 4.3.** The following relationships hold between the absolute array moments introduced in (2.8):

\[
n^2 a_1 a_8 \geq a_6 \geq \max\{a_5, a_1 a_8\}; \quad n a_4 a_1 a_2 \geq \max\{a_3, a_4\};
\]

\[
n^2 a_1 a_2 \geq a_5 \geq \max\{a_3, a_4, a_7\}; \quad a_8 \geq \max\{a_2, a_7\}.
\]

**Proof.** We illustrate the argument by proving that $n^2 a_1 a_2 \geq a_5 \geq a_3$. Since $|a_{ij} a_{ik} a_{lm}| \leq \frac{1}{2}(|a_{ij} a_{lk}| + |a_{lm} a_{ik}|)$, it is immediate that $a_3 \leq a_5$. On the other hand, $|a_{ij} a_{ik}| \leq |a_{ij} a_{ik}| \sum_{m,l} |a_{ml}|$, proving that $a_5 \leq n^2 a_1 a_2$. \( \square \)

We now prove Theorem 2 by establishing estimate (2.11). Using Stein’s method, given any function $h \in C_1$, we can write $h(w) - \mathbb{E}h(N) = f'(w) - w f'(w)$, where

\[
||f'|| \leq (1 + \sqrt{2\pi e})||h||_1
\]

(\textit{c.f.} [10]). Thus

\[
\mathbb{E}h(W) - \mathbb{E}h(N) = \mathbb{E}\{f'(W)\} - \mathbb{E}\{W f(W)\}
\]

\[
= \mathbb{E}\{f'(W)\} - \frac{1}{sn(n-1)} \sum_{i,k} \sum_{j,m} d_{ik} y_{jm} \mathbb{E}\left\{ f(W) \mid \pi(i) = j, \pi(k) = m \right\}.
\]

Now, if $\pi$ is uniformly distributed over $S_n$, $\phi_{km} \phi_{ij} \pi$ is distributed as $\pi$ conditional on $\pi(i) = j$ and $\pi(k) = m$, because of Corollary 4.1.1: hence we can write

\[
\mathbb{E}\left\{ f(W) \mid \pi(i) = j, \pi(k) = m \right\} = \mathbb{E}\{f(W(\phi_{km} \phi_{ij} \pi))\} = \mathbb{E}\{f(W + \Delta_{ijkm})\},
\]

where $\Delta_{ijkm}(\pi) := W(\phi_{km} \phi_{ij} \pi) - W(\pi)$. Hence, using Taylor’s expansion,

\[
\mathbb{E}h(W) - \mathbb{E}h(N) = \mathbb{E}f'(W) - \frac{1}{sn(n-1)} \sum_{i,k} \sum_{j,m} d_{ik} y_{jm} \mathbb{E}\{\Delta_{ijkm} f'(W)\} + \epsilon_1,
\]
where
\[ |\epsilon_1| \leq \frac{1}{n^2 s} \sum_{i,k} \sum_{j,m} |d_{ijk}y_{jm}| \mathbb{E}\{\Delta^2_{ijkm}\} \|f''\|. \]  
\hspace{1cm} (4.2)

Now
\[ \frac{1}{sn(n-1)} \sum_{i,k} \sum_{j,m} d_{ijk}y_{jm} \mathbb{E}\{\Delta_{ijkm}\} \]
\[ \hspace{1cm} = \frac{1}{sn(n-1)} \sum_{i,k} \sum_{j,m} d_{ijk}y_{jm} \mathbb{E}\{W \mid \pi(i) = j, \pi(k) = m\}, \]
because \(\mathbb{E}W = 0\), and this is in turn equal to
\[ s^{-1} \sum_{i,k} d_{ik} \mathbb{E}\{W y_{\pi(i)\pi(k)}\} = \mathbb{E}W^2 = 1. \]

Thus
\[ \mathbb{E}h(W) - \mathbb{E}h(N) = \epsilon_1 - \epsilon_2, \]  
\hspace{1cm} (4.3)

where
\[ \epsilon_2 := \mathbb{E}\{\frac{1}{sn(n-1)} \sum_{i,k} \sum_{j,m} d_{ijk}y_{jm} \Delta_{ijkm}(f'(W) - \mathbb{E}f'(W))\}. \]  
\hspace{1cm} (4.4)

It now remains to estimate \(\epsilon_1\) and \(\epsilon_2\) in terms of \(\epsilon\).

Taking \(\epsilon_2\) first, and writing \(\tilde{\pi}\) as shorthand for \(\phi_{km}\phi_{ij}\pi\), we have
\[ \epsilon_2 = \frac{1}{sg^2n(n-1)} \sum_{i,k} \sum_{j,m} d_{ijk}y_{jm} d_{rs} \]
\[ \hspace{2cm} \mathbb{E}\{(y_{\tilde{\pi}(r)\tilde{\pi}(z)} - y_{\pi(r)\pi(z)})(f'(W) - \mathbb{E}f'(W))\}, \]
so that
\[ |\epsilon_2| \leq \frac{1}{s^2 n(n-1)} \sum_{i,k} \sum_{j,m} |d_{ijk}y_{jm}| \left\{ \sum_{r,s}^{(1)} |d_{rz}| C_{ikrs}^{jm} \right. \]
\[ \hspace{3cm} + 2 \sum_{z \neq i,k} |d_{iz}| D_{ikz}^{jm} + 2 \sum_{z \neq i,k} |d_{kz}| F_{ikz}^{jm} + 2 |d_{ik}| G_{ik}^{jm} \left. \right\}, \]  
\hspace{1cm} (4.5)

where \(\sum^{(1)}\) is over all pairs \((r, z)\) such that \(\{r, z\} \cap \{i, k\} = \emptyset\), and where
\[ C_{ikrs}^{jm} := \frac{1}{(n)_4} \sum_{w, u, i, u} f^{\left[ \pi(i) = w, \pi(r) = u, \pi(k) = t, \pi(z) = v \right]} |y_{\tilde{\pi}(r)\tilde{\pi}(z)} - y_{uv}| \]
\[ \hspace{2cm} \mathbb{E}\{f'(W) \mid \pi(i) = w, \pi(r) = u, \pi(k) = t, \pi(z) = v\} - \mathbb{E}f'(W) \}; \]  
\hspace{1cm} (4.6)
\[ D^m_{ikz} := \frac{1}{(n)^3} \sum_{w,t,u} I \left[ \pi(i) = w, \pi(k) = t, \pi(z) = v \right] |y_{j\bar{x}}(x) - y_{wv}|. \]  

\[ \frac{1}{(n)^2} \sum_{w,t,i} |y_{jm} - y_{wm}| \left| \mathbb{E} \left\{ f'(W) \right| \pi(i) = w, \pi(k) = t \right\} - \mathbb{E} f'(W) \right|; \]  

(4.7)

\[ F^m_{ikz} \text{ has the same formula as } D^m_{ikz}, \text{ but with the pairs } i, k \text{ and } j, m \text{ exchanged; and} \]

\[ G^m_{ik} := \frac{1}{(n)^2} \sum_{w,t,i} |y_{jm} - y_{wm}| \left| \mathbb{E} \left\{ f'(W) \right| \pi(i) = w, \pi(k) = t \right\} - \mathbb{E} f'(W) \right| \]  

(4.8)

Note that, in the formulae for \( G^m_{ikz} \) and \( D^m_{ikz} \), the unresolved images under \( \bar{x} \) are determined by the arguments of the corresponding indicator functions. These images also depend on \( j \) and \( m \). Corollary 4.2.1 is now used to estimate the last factor in each of these expressions, and the inequality \(|y_{\alpha\beta} - y_{\gamma\delta}| \leq |y_{\alpha\beta}| + |y_{\gamma\delta}|\) is used whenever \( \{\alpha, \beta\} \neq \{\gamma, \delta\} \), which in (4.6) is only the case when \( \{u, v\} \cap \{j, m\} \neq \emptyset \). the rest of the argument, showing that \(|\epsilon_2| \leq K_2 \epsilon', ||f'||\), involves careful book-keeping and the use of Lemma 4.3: more details are given in the appendix.

For \( \epsilon_1 \), observe that

\[ \mathbb{E}(\Delta^2_{izkm}) = \frac{1}{\sigma^2} \sum_{r,z,u,v} dr_z du_v \mathbb{E}\{ (y_{\bar{x}}(r)\bar{x}(z) - y_{\pi}(r)\pi(z))(y_{\bar{x}}(u)\bar{x}(v) - y_{\pi}(u)\pi(v)) \} \]

which implies that

\[ |\epsilon_1| \leq (||f'||/n^2 \sigma^2) \sum_{i,k} \sum_{r,z,u,v} |d_{ik}| dr_z du_v \]

\[ \cdot \left\{ \sum_{j,m} |y_{jm}| \mathbb{E}\left[ (y_{\bar{x}}(r)\bar{x}(z) - y_{\pi}(r)\pi(z))(y_{\bar{x}}(u)\bar{x}(v) - y_{\pi}(u)\pi(v)) \right] \right\}. \]  

(4.9)

Now the term in braces takes one of seventeen different values, depending on the number and arrangement of the intersections of \( \{r, z, u, v\} \) with \( \{i, k\} \), each of which, in the estimate of \(|\epsilon_1|\), is multiplied by a corresponding sum of products of \( d \)'s. Again, the computation of this estimate of \( \epsilon_1 \), showing that \(|\epsilon_1| \leq K_1 \epsilon' ||f'||\), involves laborious book-keeping and the use of Lemma 4.3; more details are given in the appendix. Estimate (2.11) now follows from (4.1) and (4.3), and hence Theorem 2 holds under the assumption that \( \epsilon_n \to 0 \).
It remains to be shown that \( \epsilon' \leq K \epsilon \) as \( n \to \infty \), where, here and subsequently, \( K \) denotes a generic constant. Since, from Hölder's inequality, 
\[
|D| \leq d_1 \leq d_1^{1/2} \leq d_6^{1/3},
\]
it follows easily that
\[
(d_1')^3 + d_1' d_2' + d_3' + d_4' \leq K(d_1^3 + d_1 d_2 + d_3 + d_4),
\]
\[
d_5' + d_1' d_6' \leq K(d_5 + d_1 d_6) \quad \text{and} \quad d_6' \leq K d_6.
\]

Similar estimates are valid for the \( y \)'s, proving the remaining assertion.

\[\Box\]

5. **Proof of Theorem 3.**

To avoid technical difficulties, we shall first assume that the expansions (2.13) and (2.14) have only a finite number of non-zero terms.

Integrating the expansions (2.13) and (2.14), we obtain

\[
d_{ij}^{(n)} = \sum_{r=0}^{\infty} \alpha_r \phi_{nr}(i) \phi_{nr}(j), \quad y_{ij}^{(n)} = \sum_{s=0}^{\infty} \beta_s \psi_{ns}(i) \psi_{ns}(j),
\]

where

\[
\phi_{nr}(i) = n \int_{I_{ni}} \phi_r(x) \, dx \quad \text{and} \quad \psi_{ns}(i) = n \int_{I_{ni}} \psi_s(x) \, dx.
\]

The statistic \( U^{(n)} \) can then be written as

\[
U^{(n)} = \sum_{r=0}^{\infty} \sum_{s=0}^{\infty} \alpha_r \beta_s \left\{ \left[ \sum_{i=1}^{n} \phi_{nr}(i) \psi_{ns}(\pi(i)) \right]^2 - \sum_{i=1}^{n} \phi_{nr}^2(i) \psi_{ns}^2(\pi(i)) \right\}.
\]

To analyse the first term within the braces, define

\[
S_{r,s}^{(n)} = \sum_{i=1}^{n} \phi_{nr}(i) \psi_{ns}(\pi(i)).
\]

By Theorem V.1.6a together with Lemma V.1.6b of [7], \( S_{r,s}^{(n)} \) is asymptotically normal \((\mu_n, \sigma_n^2)\) with

\[
\mu_n = n^{-1} \left( \sum_{i=1}^{n} \phi_{nr}(i) \right) \left( \sum_{i=1}^{n} \psi_{ns}(i) \right) = n \bar{\phi}_{nr} \bar{\psi}_{ns}
\]
and
\[ \sigma_n^2 = \left[ \sum_{i=1}^{n} (\phi_{nr}(i) - \bar{\phi}_{nr})^2 \right] \int_{0}^{1} (\psi_s(x) - \bar{\psi}_s)^2 \, dx, \]
where
\[ \bar{\phi}_{nr} = \frac{1}{n} \sum_{i=1}^{n} \phi_{nr}(i), \quad \bar{\phi}_r = \int_{0}^{1} \phi_r(x) \, dx, \]
and similar formulae hold for \( \bar{\psi}_{ns} \) and \( \bar{\psi}_s \). However, from the definition of \( \phi_{nr}(i) \) and \( \psi_{ns}(i) \), \( \bar{\phi}_{nr} = \bar{\phi}_r \) and \( \bar{\psi}_{ns} = \bar{\psi}_s \). Thus \( n^{-\frac{1}{2}} (S^{(n)}_{r,s} - n\bar{\phi}_r \bar{\psi}_s) \) is asymptotically normally distributed with mean zero and variance
\[ \int_{0}^{1} (\phi_r(x) - \bar{\phi}_r)^2 \, dx \int_{0}^{1} (\psi_s(x) - \bar{\psi}_s)^2 \, dx. \]
In fact, by considering linear combinations,
\[ \{n^{-\frac{1}{2}} (S^{(n)}_{r,s} - n\bar{\phi}_r \bar{\psi}_s); r, s = 1, 2, \ldots \} \]
are asymptotically jointly normally distributed with
\[ \lim_{n \to \infty} \text{cov} \left( n^{-\frac{1}{2}} S^{(n)}_{r,s}, n^{-\frac{1}{2}} S^{(n)}_{r',s'} \right) \]
\[ = \int_{0}^{1} (\phi_r(x) - \bar{\phi}_r)(\phi_{r'}(x) - \bar{\phi}_{r'}) \, dx \int_{0}^{1} (\psi_s(x) - \bar{\psi}_s)(\psi_{s'}(x) - \bar{\psi}_{s'}) \, dx. \]  
(5.2)

For the second term in (5.1), we need a weak law of large numbers for the permutation statistic \( P_n := \frac{1}{n} \sum_{i=1}^{n} \phi_{nr}(i) \psi_{ns}(\pi(i)) \), showing that it converges in probability to the constant \( \int_{0}^{1} \phi_r^2(u) \, du \int_{0}^{1} \psi_s^2(v) \, dv = 1 \) as \( n \to \infty \). This can be achieved by truncation, as follows. Choose a sequence of constants \( C_n \to \infty \) such that \( C_n = o(n^{1/4}) \), and define
\[ a_{in} := \phi_{nr}^2(i); \quad b_{in} = \psi_{ns}^2(i); \]
\[ \tilde{a}_{in} := \{n \int_{I_n} n \phi_r(u) \, du\}^2; \quad \tilde{b}_{in} := \{n \int_{I_n} n \psi_s(u) \, du\}^2; \]
\[ \epsilon_{in} := a_{in} - \tilde{a}_{in}; \quad \eta_{in} = b_{in} - \tilde{b}_{in}, \]
where, for any function \( \theta \), we define \( n \theta(u) := \theta(u) I[|\theta(u)| \leq C_n] \). Setting \( \tilde{P}_n = \frac{1}{n} \sum_{i=1}^{n} \tilde{a}_{in} \tilde{b}_{n(i)} \), we show that
\[ \lim_{n \to \infty} \mathbb{E}[P_n - \tilde{P}_n] = 0, \]  
(5.3)
\[
\lim_{n \to \infty} \mathbb{E} P_n = 1 \quad (5.4)
\]

and
\[
\lim_{n \to \infty} \text{var } \tilde{P}_n = 0, \quad (5.5)
\]

from which it follows that \( P_n \to 1 \) in \( L_1 \).

Equation (5.4) follows immediately from Lemma V.1.6b of [7]. Then note that, for any function \( \theta \),
\[
\frac{1}{n} \sum_{i=1}^{n} \left\{ n \int_{I_{ni}} |\theta(u)| \, du \right\}^2 \leq \sum_{i=1}^{n} \int_{I_{ni}} \theta^2(u) \, du = \int_{0}^{1} \theta^2(u) \, du. \quad (5.6)
\]

Thus, for (5.5),
\[
\text{var } \tilde{P}_n \leq \frac{1}{n-1} \left\{ \frac{1}{n} \sum_{i=1}^{n} \tilde{a}^2_{in} \right\} \left\{ \frac{1}{n} \sum_{i=1}^{n} \tilde{b}^2_{in} \right\} \leq \frac{C_n^4}{n-1} \left\{ \frac{1}{n} \sum_{i=1}^{n} \tilde{a}_{in} \right\} \left\{ \frac{1}{n} \sum_{i=1}^{n} \tilde{b}_{in} \right\} \leq \frac{C_n^4}{n-1} \to 0,
\]

using (5.6). Next,
\[
\mathbb{E}|P_n - \tilde{P}_n| \leq \frac{1}{n} \mathbb{E} \left\{ \sum_{i=1}^{n} \tilde{a}_{in} |\eta_{\pi(i)n}| + \sum_{i=1}^{n} |\epsilon_{in}| \tilde{b}_{\pi(i)n} + \sum_{i=1}^{n} |\epsilon_{in} \eta_{\pi(i)n}| \right\},
\]

and we estimate the first term, the others being similarly treated. Now, for each \( i \),
\[
\mathbb{E}|\eta_{\pi(i)n}| = \frac{1}{n} \sum_{i=1}^{n} |\eta_{in}| \leq \frac{1}{n} \sum_{i=1}^{n} \alpha_{in} \beta_{in} \leq \left\{ \frac{1}{n} \sum_{i=1}^{n} \alpha_{in}^2 \right\}^{\frac{1}{2}} \left\{ \frac{1}{n} \sum_{i=1}^{n} \beta_{in}^2 \right\}^{\frac{1}{2}},
\]

where
\[
\alpha_{in} = n \int_{I_{ni}} |\psi_\theta(u)| I[|\psi_\theta(u)| > C_n] \, du
\]

and
\[
\beta_{in} = 2n \int_{I_{ni}} |\psi_\theta(u)| \, du.
\]

Hence, from (5.6),
\[
\mathbb{E}|\eta_{\pi(i)n}| \leq 2 \left\{ \int_{0}^{1} |\psi_\theta(u)|^2 I[|\psi_\theta(u)| > C_n] \, du \right\}^{\frac{1}{2}} \to 0
\]
as \( n \to \infty \). Since also \( \frac{1}{n} \sum_{i=1}^{n} \bar{a}_{in} \leq 1 \), (5.3) is established.

Suppose now that \( d(x, y) \) is degenerate, so that \( \phi_0 \equiv 1 \) and \( \bar{\phi}_r = 0 \), \( r = 1, 2, \ldots \). Then \( \mathbb{E} \bar{S}_r^{(n)} = 0 \) unless \( r = 0 \), when \( \bar{S}_0^{(n)} = \sum_{i=1}^{n} \psi_{ne}(i) = n\bar{\psi}_e \) is almost surely constant. Thus

\[
n^{-1}(U^{(n)} - \mathbb{E} U^{(n)}) = \sum_{r \geq 1} \sum_{s = 0}^{\infty} a_r \beta_s \left[ \left( n^{\frac{1}{2}} S_r^{(n)} \right)^2 - \mathbb{E} \left( n^{\frac{1}{2}} S_r^{(n)} \right)^2 \right] \\
- \left\{ \frac{1}{n} \sum_{i=1}^{n} \phi_{nr}^2(i) \psi_{ns}^2(\pi(i)) - \left( \frac{1}{n} \sum_{i=1}^{n} \phi_{nr}^2(i) \right) \left( \frac{1}{n} \sum_{i=1}^{n} \psi_{ns}^2(i) \right) \right\}
\]

\[\overset{D}{\to} \sum_{r \geq 1} \sum_{s = 0}^{\infty} a_r \beta_s (Z_{rs}^2 - \mathbb{E} Z_{rs}^2),\]

where the \( Z_{rs} \) are jointly normally distributed with \( \mathbb{E} Z_{rs} = 0 \) and

\[\text{cov} (Z_{rs}, Z_{r's'}) = \delta_{rr'} \int_{0}^{1} \int_{0}^{1} (\psi_s(v) - \bar{\psi}_s)(\psi_{s'}(v) - \bar{\psi}_{s'}) \, dv.
\]

When both \( d(u, v) \) and \( y(u, v) \) are degenerate, \( \int_{0}^{1} \psi_s(v) \, dv = 0 \) for \( s \neq 0 \) also. It then follows that

\[n^{-1}(U^{(n)} - \mathbb{E} U^{(n)}) \overset{D}{\to} \sum_{r \geq 1} \sum_{s = 1}^{\infty} a_r \beta_s (Z_{rs}^2 - 1),\]

where the \( Z_{rs} \) are independent standard normal variables.

We now relax the condition that the expansions (2.13) and (2.14) have only finitely many non-zero terms. Suppose, without loss of generality, that the kernel \( d \) is degenerate, and that

\[\int_{0}^{1} \int_{0}^{1} d(u, v) \, du \, dv = \int_{0}^{1} \int_{0}^{1} y(u, v) \, du \, dv = 0;\]

the latter relationship may be achieved by subtracting the appropriate constant from each of \( d \) and \( y \), which does not affect the value of \( U^{(n)} - \mathbb{E} U^{(n)} \). Now define

\[V := \sum_{r \geq 1} \sum_{s = 0}^{\infty} a_r \beta_s \{ Z_{rs}^2 - \mathbb{E} Z_{rs}^2 \},\]

where the \( Z_{rs} \) are jointly normally distributed with zero mean and covariance

\[\text{cov} (Z_{rs}, Z_{r's'}) = \delta_{rr'} \int_{0}^{1} (\psi_s(v) - \bar{\psi}_s)(\psi_{s'}(v) - \bar{\psi}_{s'}) \, dv = \delta_{rr'} (\delta_{ss'} - \bar{\psi}_s \bar{\psi}_{s'}).
\]
and the infinite sums are to be interpreted in $L_2$: $V$ has the distribution to which $n^{-1}(U(n) - \mathbb{E}U(n))$ is to be shown to converge. It is not immediately clear that $V$ so defined exists. However, for any finite set of index pairs $R$, we have

$$
\theta_R := \mathbb{E}\left\{ \left( \sum_R \alpha_r \beta_s (Z_{r,s}^2 - \mathbb{E}Z_{r,s}^2) \right)^2 \right\} = \sum_R \alpha_r^2 \beta_s^2 \mathbb{E}(Z_{r,s}^2 - \mathbb{E}Z_{r,s}^2)^2 
+ \sum_R \sum_{r' \in R, s' \neq s} \alpha_r^2 \beta_s \beta_{s'} \mathbb{E}\{(Z_{r,s}^2 - \mathbb{E}Z_{r,s}^2)(Z_{r',s'}^2 - \mathbb{E}Z_{r',s'}^2)\},
$$

where $R_r := \{ s : (r, s) \in R \}$, because $Z_{r,s}$ and $Z_{r',s'}$ are independent whenever $r \neq r'$. Since also $\mathbb{E}(Z_{r,s}^2 - \mathbb{E}Z_{r,s}^2)^2 = 2(1 - \bar{\psi}_s^2)^2$ and $\mathbb{E}\{(Z_{r,s}^2 - \mathbb{E}Z_{r,s}^2)(Z_{r',s'}^2 - \mathbb{E}Z_{r',s'}^2)\} = 2\bar{\psi}_s^2 \bar{\psi}_{s'}^2$, $s \neq s'$, it follows that

$$
\theta_R \leq 2 \sum_R \alpha_r^2 \beta_s^2 + 2 \sum_{r \geq 1} \left( \sum_{s \in R_r} \beta_s \bar{\psi}_s^2 \right)^2.
$$

Now

$$
\left( \sum_{s \in R_r} \beta_s \bar{\psi}_s^2 \right)^2 = \left( \int_0^1 \int_0^1 \sum_{s \in R_r} \beta_s \psi_s(u) \psi_s(v) \, du \, dv \right)^2 
\leq \int_0^1 \int_0^1 \left( \sum_{s \in R_r} \beta_s \psi_s(u) \psi_s(v) \right)^2 \, du \, dv = \sum_{s \in R_r} \beta_s^2.
$$

Hence

$$
\theta_R \leq 4 \sum_R \alpha_r^2 \beta_s^2;
$$

and, since $\sum_{r \geq 1} \alpha_r^2 \sum_{s \geq 0} \beta_s^2 < \infty$, $V$ is well defined.

Next, given any $\delta > 0$, choose $N = N_\delta$ so large that $\sum_{r \geq N+1} \alpha_r^2 < \delta$ and $\sum_{s \geq N+1} \beta_s^2 < \delta$. Let

$$
d^\delta(u, v) := \sum_{r=1}^N \alpha_r \phi_r(u) \phi_r(v); \quad \epsilon(u, v) := d(u, v) - d^\delta(u, v)
$$

$$
y^\delta(u, v) := \sum_{s=1}^N \beta_s \psi_s(u) \psi_s(v); \quad \eta(u, v) := y(u, v) - y^\delta(u, v).
$$
Then we have
\[ \int_0^1 d^\delta(u, v) \, dv = \int_0^1 \epsilon(u, v) \, dv = 0, \quad 0 \leq u \leq 1, \]
\[ \int_0^1 \int_0^1 \{d^\delta(u, v)\}^2 \, dudv = \sum_{r=1}^N \alpha_r^2 \leq \|d\|^2_2 \]
and
\[ \int_0^1 \int_0^1 \epsilon^2(u, v) \, dudv = \sum_{r \geq N+1} \alpha_r^2 < \delta, \]
and similar quadratic inequalities hold for \( y^\delta \) and \( \eta \). We approximate \( n^{-1} U(n) \) by the scheme \( n^{-1} U(n) = n^{-1} \{U_n^\delta + E_n^\delta + H_n^\delta\} \), where
\[ E_n^\delta := \sum_{i,j} \epsilon_{ij} y_{\pi(i)\pi(j)}; \quad H_n^\delta := \sum_{i,j} d_{ij}^\delta \eta_{\pi(i)\pi(j)} \quad \text{and} \quad U_n^\delta := \sum_{i,j} d_{ij}^\delta y_{\pi(i)\pi(j)}, \]
and where, for example,
\[ \epsilon_{ij} := n^2 \int_{I_{ni} \times I_{nj}} \epsilon(u, v) \, dudv. \]
It is then possible to show that
\[ \text{var} (n^{-1} E_n^\delta) \leq K \delta \|y\|_2^2 \]  
(5.8)
and
\[ \text{var} (n^{-1} H_n^\delta) \leq K \delta \|d\|_2^2 \]  
(5.9)
uniformly in \( n \).

Now \( U_n^\delta \) is constructed from symmetric square integrable kernels \( d^\delta \) and \( y^\delta \) with only finitely many non-zero eigenfunctions, and \( d^\delta \) is degenerate. Hence \( n^{-1}(U_n^\delta - \mathbb{E}U_\delta^\delta) \overset{D}{\to} V^\delta \) as \( n \to \infty \), for each fixed \( \delta \), by the first stage of the argument, where
\[ V^\delta := \sum_{r=1}^N \sum_{\sigma=0}^N \alpha_r \beta_{\sigma}(Z_{r\sigma}^2 - \mathbb{E}Z_{r\sigma}^2). \]
Then, if \( f \in C_1 \),

\[
\left| \mathbb{E}f(n^{-1}(U^{(n)} - \mathbb{E}U^{(n)})) - \mathbb{E}f(V) \right| \\
\leq \mathbb{E}\left| f(n^{-1}(U^{(n)} - \mathbb{E}U^{(n)})) - f(n^{-1}(U_n^\delta - \mathbb{E}U_n^\delta)) \right| \\
+ \left| \mathbb{E}f(n^{-1}(U_n^\delta - \mathbb{E}U_n^\delta)) - \mathbb{E}f(V^\delta) \right| + \mathbb{E}|f(V^\delta) - f(V)| \\
\leq \left| \mathbb{E}f(n^{-1}(U_n^\delta - \mathbb{E}U_n^\delta)) - \mathbb{E}f(V^\delta) \right| \\
+ K\|f\|_1 \{n^{-1} \left[ \var{U^{(n)} - U_n^\delta} \right]^{\frac{1}{2}} + \left[ \var{V - V^\delta} \right]^{\frac{1}{2}} \}.
\]

Now, since the sum defining \( V \) converges in \( L_2 \), we can use (5.7) with an arbitrary index set \( R \), and so \( \var{V - V^\delta} \leq \delta \{\|d\|_2^2 + \|y\|_2^2 \} \). Hence, provided that (5.8) and (5.9) hold,

\[
\limsup_{n \to \infty} |\mathbb{E}f(n^{-1}(U^{(n)} - \mathbb{E}U^{(n)})) - \mathbb{E}f(V)| \leq K\|f\|_1 \delta^{\frac{1}{2}} \{\|d\|_2^2 + \|y\|_2^2 \}^{\frac{1}{2}},
\]

and this proves that \( n^{-1}(U^{(n)} - \mathbb{E}U^{(n)}) \overset{D}{\to} V \), since \( \delta > 0 \) was chosen arbitrarily.

To prove (5.8) and (5.9), we use (2.3). Consider first (5.8). Because \( \int_0^1 \epsilon(u, v) \, du = 0 \) for all \( v \), we have

\[
\sum_{j \neq i} \epsilon_{ij} = -n^2 \int_{I_n^2} \epsilon(u, v) \, du \, dv \quad \text{and} \quad \sum_{i,j} \epsilon_{ij} = -n^2 \int_{\Delta_n} \epsilon(u, v) \, du \, dv,
\]

where \( \Delta_n := \cap_{i=1}^n I_n^2 \). Hence

\[
\frac{1}{n} \sum_{i=1}^n (\epsilon_i^2) \leq \left( \frac{n}{n-2} \right)^2 \frac{1}{n} \sum_{i=1}^n \left( n \int_{I_n^2} \epsilon(u, v) \, du \, dv \right)^2 \\
\leq Kn^{-3} \sum_{i=1}^n n^2 \int_{I_n^2} \epsilon^2(u, v) \, du \, dv = Kn^{-1} \int_{\Delta_n} \epsilon^2(u, v) \, du \, dv.
\]

On the other hand, \( Y_1 \leq n(n-1)(n-2)^{-2} \|y\|_2^2 \), and so the first term from (2.3) in the evaluation of \( \var{n^{-1}E_n^2} \) is no greater than

\[
K\|y\|_2^2 \int_{\Delta_n} \epsilon^2(u, v) \, du \, dv,
\]

which is bounded by \( K\delta \|y\|_2^2 \). Since also \( Y_2 \leq n^{-1}(n-1)\|y\|_2^2 \), and a similar inequality holds for \( \epsilon \), the second term is also at most \( K\delta \|y\|_2^2 \), proving (5.8).
The proof of (5.9) is similar: replace $\epsilon$ by $d^6$ and $y$ by $\eta$, obtaining
\[
\var(n^{-1} H_n^\delta) \leq K d^6 n^{-1} \|\eta\|^2 \leq K \delta d n^{-1}.
\]
Corollary 3.1 is proved by writing
\[
U_n^{(n)} = \hat{U}_n + \sum_{i,j} \{\epsilon_{ij} \hat{y}_{\pi(i)\pi(j)} + \hat{d}_{ij} \eta_{\pi(i)\pi(j)} + \epsilon_{ij} \eta_{\pi(i)\pi(j)}\},
\]
and using (2.3) on each of the last three terms. \qed

APPENDIX.

We distinguish four groups of third moments of an array $(a_{ij})$:

\[
G_1(a) := \{a_1, a_2, a_3, a_4, a_7\}; \quad G_2(a) := \{a_7\};
\]
\[
G_3(a) := \{a_5, a_1 a_8\}; \quad G_4(a) := \{a_6\}.
\]

From the inequalities in Lemma 4.3, any member of $G_1$ or $G_2$ is less than or equal to at least one member of $G_3$, and both members of $G_3$ are less than or equal to the member of $G_4$. Members of $G_3$ are only obtained from sums of the form $\sum |a_{\alpha\beta} a_{\gamma\delta} a_{\epsilon\eta}|$ when two of the three sets $\{\alpha, \beta\}, \{\gamma, \delta\}, \{\epsilon, \eta\}$ are equal, and all three must be equal to give the member of $G_4$. The estimates that follow involve such sums for the arrays $(d_{ij})$ and $(y_{ij})$, and expressions such as $O(G_1(a) + n^{-1} G_2(a) + n^{-2} G_4(a))$ have their natural interpretation. In particular, we are to prove that $|\epsilon_1| + |\epsilon_2| = \|f''\| O(\epsilon)$, where

\[
\epsilon \sim s^{-3}(n^4 G_1(d) G_1(y) + n^3 G_3(d) G_3(y) + n^2 G_4(d) G_4(y)).
\]

Estimation of $|\epsilon_2|$. We start with (4.6). The factor $|\tilde{y}_{(r)}(z) - y_{uv}|$ is zero unless $\{u, v\} \cap \{j, m\} \neq \emptyset$. Consider the case $u = j$, for which we have

\[
\begin{align*}
&u = j, \ w = m : \quad |\tilde{y}_{(r)}(z) - y_{uv}| \leq |y_{iv}| + |y_{ju}| \\
&u = j, \ t = m : \quad |\tilde{y}_{(r)}(z) - y_{uv}| \leq |y_{uv}| + |y_{ju}| \\
&u = j, \ v = m : \quad |\tilde{y}_{(r)}(z) - y_{uv}| \leq |y_{uw}| + |y_{jv}| \\
&u = j, \ m \notin \{w, t, v\} : \quad |\tilde{y}_{(r)}(z) - y_{uv}| \leq |y_{uv}| + |y_{jv}|.
\end{align*}
\]

(A.1)
From the \( |y_{jv}| \) terms, using Corollary 4.2.1, we get a contribution to \( |\varepsilon_2| \) no greater than
\[
K\|f''\|s^{-3}n^{-6} \sum_{i,k,r,z,j,m} |d_{ik}d_{rz}y_{jm}| \sum_{w,t,u\neq j} |y_{jv}|
\]
\[
\{ |d_{ik}y_{wt}| + |d_{ir}y_{wz}| + |d_{iz}y_{wu}| + |d_{kr}y_{tz}| + |d_{kz}y_{tw}| + |d_{rz}y_{ju}| \}
\]
\[
+ n^{-1} [d_i y_{w*} + d_k y_{tu} + d_r y_{j*} + d_z y_{u*}] + n^{-3} [y_{..d..}]
\]
\[
+ n^{-2} [y_{..(d_i + d_k + d_r + d_z)} + d_{..(y_{w*} + y_{tu} + y_{j*} + y_{u*})}] \).
\] (A.2)

We now compute the following \( y \)-estimates:
\[
\sum |y_{jm}y_{jv}y_{wt}| \leq K(n^5 y_1 y_2 + n^4 y_1 y_3)
\]
\[
\sum |y_{jm}y_{jv}y_{wz}| \leq K(n^5 y_3 + n^4 y_3)
\]
\[
\sum |y_{jm}y_{jv}y_{wu}| \leq K(n^5 y_4 + n^4 y_5)
\]
\[
\sum |y_{jm}y_{jv}y_{tz}| \leq K(n^5 y_3 + n^4 y_5)
\]
\[
\sum |y_{jm}y_{jv}y_{tw}| \leq K(n^5 y_4 + n^4 y_5)
\]
\[
\sum |y_{jm}y_{jv}y_{j*}| \leq K(n^5 y_5 + n^4 y_6),
\] (A.3)

which are all of order \( O\{n^5 G_3(y) + n^4 G_4(y)\}; \)
\[
\sum |y_{jm}y_{jv}y_{w*}| \leq K(n^6 y_1 y_2 + n^5 y_1 y_3)
\]
\[
\sum |y_{jm}y_{jv}y_{tu}| \leq K(n^6 y_1 y_2 + n^5 y_1 y_3)
\]
\[
\sum |y_{jm}y_{jv}y_{j*}| \leq K(n^6 y_3 + n^5 y_5 + n^4 y_6)
\]
\[
\sum |y_{jm}y_{jv}y_{u*}| \leq K(n^6 y_4 + n^5 y_5 + n^4 y_6),
\] (A.4)

which are all of order \( O\{n^6 G_1(y) + n^5 G_3(y) + n^4 G_4(y)\}; \)
\[
\sum |y_{jm}y_{jv}|y_{..} \leq K(n^7 y_1 y_2 + n^6 y_1 y_3) = O\{n^7 G_1(y) + n^6 G_3(y)\},
\] (A.5)

where all sums are of the form \( \sum_{i,w,t,u} \sum_{m} \), and the \( d \)-estimates:
\[
\sum |d_{ik}^2 d_{rz}| \leq K n^4 d_1 d_8 \quad \sum |d_{ik} d_{rz} d_{kr}| \leq K n^4 d_4
\]
\[
\sum |d_{ik} d_{rz} d_{ir}| \leq K n^4 d_4 \quad \sum |d_{ik} d_{rz} d_{kz}| \leq K n^4 d_4
\]
\[
\sum |d_{ik} d_{rz} d_{iz}| \leq K n^4 d_4 \quad \sum |d_{ik} d_{rz|^2}| \leq K n^4 d_1 d_8
\] (A.6)
which are all of order $O\{n^4G_3(d)\}$;

$$
\sum |d_{ik}d_{rs}d_{i,}| \leq K(n^5d_1d_2 + n^4d_1d_3) \\
\sum |d_{ik}d_{rs}d_{k,}| \leq K(n^5d_1d_2 + n^4d_1d_3) \\
\sum |d_{ik}d_{rs}d_{r,}| \leq K(n^5d_1d_2 + n^4d_1d_3) \\
\sum |d_{ik}d_{rs}d_{s,}| \leq K(n^5d_1d_2 + n^4d_1d_3),
$$

(A.7)

which are all of order $O\{n^5G_1(d) + n^4G_3(d)\}$;

$$
\sum |d_{ik}d_{rs}d_{..,}| \leq Kn^6d^3_1 = O\{n^6G_1(d)\},
$$

(A.8)

where all sums are of the form $\sum_{i,k,r,s}$. Putting these estimates into (A.2) gives a contribution of order $\|f''\| O(\epsilon)$.

The next most common terms from (A.1), those involving $|y_{uw}|$, may be estimated from a formula similar to (A.2), but with the first $|y_{ju}|$ factor replaced by $|y_{uw}|$. This leads to slightly different first estimates from those in (A.3-A.5), but to the same overall orders of magnitude. Thus these terms give a total contribution of order $\|f''\| O(\epsilon)$. The remaining terms from (A.1) occur only when two of $w, t, u$ and $v$ are fixed, and not surprisingly contribute only order $n^{-1}\|f''\| O(\epsilon)$. Finally, the case $u = m$, and the cases where $v$ and $u$ are interchanged, all give similar contributions, because of symmetry. Thus the contribution to $|\epsilon_2|$ from the terms involving $C_{ikrs}^{im}$ is of the required order.

Next we turn to (4.7). The terms with $|y_{uw}|$ contribute at most

$$
K\|f''\|\theta^{-3}n^{-5} \sum_{i,k,z,j,m} |d_{ik}d_{iz}y_{jm}| \sum_{w,t,v} |y_{uw}|
$$

\[
\left\{ |d_{iz}y_{uw}| + |d_{kz}y_{tv}| + |d_{ik}y_{wt}| + n^{-1}[d_{x,y,v} + d_{k,y,t} + d_{i,y,w}] \right. \\
+ n^{-2}[d_{..}(y_{vu} + y_{tv} + y_{tw}) + y_{.}(d_{x} + d_{k} + d_{i})] + n^{-3}d_{..}y_{.}\right\},
\]

(A.9)

and the $d$-estimates

$$
\sum |d_{ik}d_{iz}|\{ |d_{iz}|, |d_{kz}|, |d_{ik}| \} = O\{n^3G_3(d)\} \\
\sum |d_{ik}d_{iz}|\{ |d_{x,z}|, |d_{k,x}|, |d_{i,x}| \} = O\{n^4G_1(d) + n^3G_3(d)\} \\
\sum |d_{ik}d_{iz}|d_{..} = O\{n^5G_1(d)\}
$$

(A.10)
(sums $\sum_{i,k,z}$), and the $y$-estimates

$$
\sum |y_{jm}y_{wv}| \{ |ywv|, |ywv|, |ywv| \} = O\{ n^5 G_3(y) \}
$$

$$
\sum |y_{jm}y_{wv}| \{ y_v, y_t, y_v \} = O\{ n^6 G_4(y) + n^5 G_3(y) \}
$$

$$
\sum |y_{jm}y_{wv}| y_v = O\{ n^7 G_4(y) \}
$$

(sums $\sum_{j,m} \sum_{w,t,u}$) show that the total contribution is no larger than $\|f''\| O(\varepsilon)$. The other common term has $|y_{ju}|$ instead of $|y_{uw}|$, occurring only when $v \neq j, m$ and $w \neq j$, yielding $y$-estimates

$$
\sum |y_{jm}y_{ju}| \{ |ywv|, |ywv|, |ywv| \} = O\{ n^5 G_3(y) \}
$$

$$
\sum |y_{jm}y_{ju}| \{ y_v, y_t, y_v \} = O\{ n^6 G_4(y) + n^5 G_3(y) \}
$$

$$
\sum |y_{jm}y_{ju}| y_v = O\{ n^7 G_4(y) \}
$$

(sums $\sum_{j,m} \sum_{u,w,t,v}$), which also yields a contribution of the appropriate order. The remaining terms involve $y_j(x) = y_{jt}$ ($v = m, t \neq j$ and $v = j, w = m$) and $y_{jw}(x) = y_{jw}$ ($v = m, t = j$), and both yield contributions of order $n^{-1}\|f''\| O(\varepsilon)$.

Finally, consider the quantity (4.8). The contribution from the terms with $|y_{jm}|$ is no greater than

$$
K\|f''\| \omega^{-3} n^{-4} \sum_{i,k} \sum_{j,m} |d_{ik}y_{jm}| \sum_{w,t} |y_{jm}|
$$

$$
\cdot \left\{ |d_{ik}y_{wt}| + n^{-1}[d_i.y_w + d_k.y_t] + n^{-2}[d_i.y_w + y_t + y_i(d_i + d_k)] + n^{-3}d..y.. \right\},
$$

(A.13)

and, using the $d$-estimates

$$
\sum |d_{ik}|^3 = O\{ n^2 G_4(d) \}
$$

$$
\sum |d_{ik}|^2 \{ d_i, d_k \} = O\{ n^3 G_3(d) + n^2 G_4(d) \}
$$

$$
\sum |d_{ik}|^2 d.. = O\{ n^4 G_3(d) \}
$$

(A.14)

(sums $\sum_{i,k}$) and the $y$-estimates

$$
\sum |y_{jm}y_{wt}| = O\{ n^4 G_4(y) \}
$$

$$
\sum |y_{jm}|^2 \{ y_w, y_t \} = O\{ n^5 G_3(y) \}
$$

$$
\sum |y_{jm}|^2 y.. = O\{ n^6 G_3(y) \}
$$

(A.15)
(sums $\sum_{i,m} \sum_{w,t}$), it can be estimated to be no greater than $\|f^w\| O(\epsilon)$. The corresponding term with $|y_{wt}|$ has $y$-estimates

\[
\sum |y_{jm} y_{wt}| = O\{n^4 G_4(y)\}
\]
\[
\sum |y_{jm} y_{wt}||y_{w\cdot} y_{t\cdot}| = O\{n^5 G_3(y)\}
\]
\[
\sum |y_{jm} y_{wt}| y_{\cdot\cdot} = O\{n^6 G_3(y)\}
\]

(sums $\sum_{i,m} \sum_{w,t}$), leading to a similar bound on the contribution to $|\epsilon_2|$. Thus, in all, $|\epsilon_2| = \|f^w\| O(\epsilon)$.

**Estimation of $|\epsilon_1|$**. We distinguish the following twelve arrangements of the suffices $\{r, z; u, v\}$ with respect to $\{i, k\}$:

1: $\{i, k; i, k\}$
2: $\{i, z; i, k\}$
3: $\{i, k; r, z\}$
4: $\{i, z; i, z\}$
5: $\{i, z; i, v\}$
6: $\{i, z; k, z\}$
7: $\{i, z; k, v\}$
8: $\{i, z; u, v\}$
9: $\{i, z; z, v\}$
10: $\{r, z; r, z\}$
11: $\{r, z; r, v\}$
12: $\{r, z; u, v\}$.

The symmetries of $i$ and $k$, and of $r, z, u$ and $v$, in the estimate (4.9) of $|\epsilon_1|$, ensure that all others are equivalent, for our purposes, to one of these. In each set of braces, distinct letters denote distinct integers. We determine the orders of the sums $\{C_x := \sum_{i,k} |d_{ik}| \sum(z) d_{rx} d_{uv}, 1 \leq x \leq 12\}$, where the sum $\sum(z)$ is taken over those variables, excluding $i$ and $k$, appearing in index set $x$: thus, $C_0 = \sum_{i,k} |d_{ik}| \sum_{z \neq i,k} d_{iz} d_{kz}$. We then, for each $x$, estimate the quantity

\[
E_x := \frac{1}{n^2} \sum_{j,m} |y_{jm}| \left| \mathbb{E}\left\{ (y_{x}(r) \tilde{y}(x) - y_{\pi}(r) \pi(x))(y_{x}(u) \tilde{y}(v) - y_{\pi}(u) \pi(v)) \right\} \right|
\]

for any choice of $\{r, z, u, v\}$ in index set $x$ (it is the same for any such choice).

Direct calculation and the use of the inequalities of Lemma 4.3 show
that

\[ C_1 = O\{n^2 G_4(d)\}; \quad C_7 = O\{n^4 G_1(d)\}; \]
\[ C_2 = O\{n^3 G_3(d)\}; \quad C_8 = O\{n^5 G_1(d)\}; \]
\[ C_3 = O\{n^4 G_3(d)\}; \quad C_9 = O\{n^4 G_1(d)\}; \]
\[ C_4 = O\{n^3 G_3(d)\}; \quad C_{10} = O\{n^4 G_3(d)\}; \]
\[ C_5 = O\{n^4 G_1(d)\}; \quad C_{11} = O\{n^5 G_1(d)\}; \]
\[ C_6 = O\{n^3 G_2(d)\}; \quad C_{12} = O\{n^6 G_1(d)\}. \]

The quantities \( E_x \) are more tedious to compute, because of the awkward way in which \( \tilde{\pi} \) depends on \( j \) and \( m \). However, their orders of magnitude in terms of the groups \( \{G_\alpha(y), 1 \leq \alpha \leq 4\} \) can be relatively simply deduced. Thus, for \( \pi \) uniformly distributed on \( S_n \),

\[ E_1 = \frac{1}{n^2} \sum_{j,m} |y_{jm}| \mathbb{E}\{(y_{jm} - y_{\pi(i)\pi(k)})^2\} = O\{G_4(y)\}. \]

Then

\[ E_2 = \frac{1}{n^2} \sum_{j,m} |y_{jm}| \left| \mathbb{E}\{ (y_{j\tilde{\pi}(z)} - y_{\pi(i)\pi(z)}) (y_{jm} - y_{\pi(i)\pi(k)}) \} \right|, \]

in which \( \pi(i), \pi(k) \) and \( \pi(z) \) are distinct, and \( \tilde{\pi}(z) \notin (j, m) \). Hence there can be no \( G_4(y) \) terms, except those arising when \( \{\pi(i), \pi(z)\} = \{j, m\} \), and there is a contribution of \( O(G_3(y)) \) from the terms with \( |y_{jm}|^2 \); thus \( E_2 = O\{G_3(y) + n^{-2}G_4(y)\} \). Similar arguments yield

\[ E_3 = \frac{1}{n^2} \sum_{j,m} |y_{jm}| \left| \mathbb{E}\{ (y_{jm} - y_{\pi(i)\pi(k)}) (y_{\tilde{\pi}(r)\tilde{\pi}(z)} - y_{\pi(r)\pi(z)}) \} \right| \]
\[ = O\{n^{-1}G_3(y) + n^{-2}G_4(y)\} \]

(note that the second factor in the expectation is zero unless \( \{\pi(r), \pi(z)\} \cap \{j, m\} \neq \emptyset \));

\[ E_4 = \frac{1}{n^2} \sum_{j,m} |y_{jm}| \mathbb{E}\{(y_{j\tilde{\pi}(z)} - y_{\pi(i)\pi(z)})^2\} = O\{G_3(y) + n^{-2}G_4(y)\}; \]

\[ E_5 = \frac{1}{n^2} \sum_{j,m} |y_{jm}| \left| \mathbb{E}\{ (y_{j\tilde{\pi}(z)} - y_{\pi(i)\pi(z)}) (y_{j\tilde{\pi}(v)} - y_{\pi(i)\pi(v)}) \} \right| \]
\[ = O\{G_1(y) + n^{-1}G_3(y)\}; \]
\[ E_6 = \frac{1}{n^2} \sum_{j,m} |y_{jm}| |E\{(y_j\tilde{\pi}(x) - \gamma \pi(x))(y_m\tilde{\pi}(x) - \gamma \pi(x))\}| \\
= O\{G_1(y) + G_2(y) + n^{-1}G_3(y)\}; \]

\[ E_7 = \frac{1}{n^2} \sum_{j,m} |y_{jm}| |E\{(y_j\tilde{\pi}(x) - \gamma \pi(x))(y_m\tilde{\pi}(v) - \gamma \pi(v))\}| \\
= O\{G_1(y) + n^{-2}G_3(y)\}; \]

\[ E_8 = \frac{1}{n^2} \sum_{j,m} |y_{jm}| |E\{(y_j\tilde{\pi}(x) - \gamma \pi(x))(y_j\tilde{\pi}(v) - \gamma \pi(v))\}| \\
= O\{n^{-1}G_1(y) + n^{-2}G_3(y)\}; \]

\[ E_9 = \frac{1}{n^2} \sum_{j,m} |y_{jm}| |E\{(y_j\tilde{\pi}(x) - \gamma \pi(x))(y_j\tilde{\pi}(v) - \gamma \pi(v))\}| \\
= O\{n^{-1}G_3(y)\}; \]

\[ E_{10} = \frac{1}{n^2} \sum_{j,m} |y_{jm}| |E\{(y_{\tilde{\pi}}(x) - y_{\pi}(x))^2\}| \\
= O\{n^{-1}G_3(y) + n^{-2}G_4(y)\}; \]

\[ E_{11} = \frac{1}{n^2} \sum_{j,m} |y_{jm}| |E\{(y_{\tilde{\pi}}(x) - y_{\pi}(x))(y_{\tilde{\pi}}(v) - y_{\pi}(v))\}| \\
= O\{n^{-1}G_1(y) + n^{-2}G_3(y)\}; \]

\[ E_{12} = \frac{1}{n^2} \sum_{j,m} |y_{jm}| |E\{(y_{\tilde{\pi}}(x) - y_{\pi}(x))(y_{\tilde{\pi}}(v) - y_{\pi}(v))\}| \\
= O\{n^{-2}G_1(y)\}. \]

Finally, from (4.9), \(|\epsilon_1| = \|f''\| O\{\sum_{x=1}^{12} C_x E_x\} = \|f''\| O(\epsilon). \]

REFERENCES.


