

# A Central Limit Theorem for Decomposable Random Variables with Applications to Random Graphs

A. D. BARBOUR

*Institut für Angewandte Mathematik, Universität Zürich,  
Rämistrasse 74, CH-8001 Zürich, Switzerland*

AND

MICHAŁ KAROŃSKI AND ANDRZEJ RUCIŃSKI

*Institute of Mathematics, Adam Mickiewicz University,  
Matejki 48/49, 60-769 Poznań, Poland*

*Communicated by W. R. Pulleyblank*

Received August 11, 1987

The application of Stein's method of obtaining rates of convergence to the normal distribution is illustrated in the context of random graph theory. Problems which exhibit a dissociated structure and problems which do not are considered. Results are obtained for the number of copies of a given graph  $G$  in  $K(n, p)$ , for the number of induced copies of  $G$ , for the number of isolated trees of order  $k \geq 2$ , for the number of vertices of degree  $d \geq 1$ , and for the number of isolated vertices.

© 1989 Academic Press, Inc.

## 1. INTRODUCTION

In 1970, when investigating the central limit theorem for stationary sequences of random variables, Stein [24] introduced a powerful new technique for obtaining estimates of the rate of convergence to the standard normal distribution. His approach was subsequently extended to cover convergence to Poisson distributions by Chen [6]. Both methods were illustrated, in the context of random graph theory, in Barbour [1]. The method for proving Poisson convergence has since been widely taken up (Karoński [14], Karoński and Ruciński [15], Nowicki [20], Janson [12]), but results for random graphs subsequently obtained by the method for normal convergence seem to be limited to examples in Barbour and Eagleson [2, 3]. This paper is intended to make the approach more

generally accessible. In particular, it is shown how, by using Stein's method, excellent results can be obtained for the distributions of the counts of small subgraphs, tree components, and vertices of a given degree, in the Bernoulli random graph  $K(n, p)$ .

The Stein approach has several advantages over the method of moments, which has been the most popular technique for proving convergence in random graph theory. The principal advantage is that a rate of convergence is automatically obtained, but the computations are also often easier, and fewer moment assumptions are required. The latter two properties frequently lead to conditions for convergence weaker than those obtainable by the method of moments. However, in some counting problems in random graph theory, the existence and computation of moments present few problems, and an example is given where better conditions for convergence can be obtained by the traditional approach.

The main reason why Stein's method for normal convergence has been less readily exploited is that the argument in Barbour [1] is much more difficult than that for Poisson convergence. This is partly inevitable, as the examples in this paper show. However, the argument in Barbour [1] is mostly complicated by the effort involved in obtaining the sharpest possible estimate of the rate of convergence, expressed in the traditional form

$$\delta_n := \sup_x |F_n(x) - \Phi(x)|, \quad (1.1)$$

where  $F_n$  is the distribution function being approximated, and  $\Phi$  that of the standard normal distribution. This way of expressing a rate of convergence, natural in a statistical context, is not natural when considering convergence in distribution more generally, nor does it arise naturally from Stein's method. Instead, one obtains direct estimates of the form

$$\left| \int h(x) dF_n(x) - \int h(x) d\Phi(x) \right| \leq \varepsilon_n \|h\|, \quad (1.2)$$

for all bounded test functions  $h$  with bounded derivative, where  $\|h\| := \sup_x |h(x)| + \sup_x |h'(x)|$ . The quantity  $\varepsilon_n$  in (1.2) provides an upper estimate of the distance between  $F_n$  and  $\Phi$ , but in a metric  $d_1$  different from that in (1.1): see also Barbour and Hall [4]. In general,  $\delta_n = O(\varepsilon_n^{1/2})$  when convergence to  $\Phi$  is being considered; very often, it can also be proved that  $\delta_n \asymp \varepsilon_n$ , but only at the cost of much greater effort (see Chen [8]). However, the Stein argument leading to estimates of the form (1.2) is often tractable, and frequently yields optimal conditions for convergence, making it a profitable technique to employ. In particular, the example used in Barbour [1] to illustrate normal approximation is reconsidered here in terms only of establishing (1.2), leading to considerable simplification.

In Section 2, Stein’s method is used to prove a normal approximation theorem in a rather general setting. Use of the theorem is then illustrated by application to some problems in random graph theory in Sections 3 and 4.

2. A CENTRAL LIMIT THEOREM FOR DECOMPOSABLE RANDOM VARIABLES

In this section we establish some sufficient conditions for convergence to the standard normal distribution of a sequence of random variables, using Stein’s method. Suppose that a random variable  $W$  is decomposed using finite index sets  $I$  and  $K_i \subset I, i \in I$ , and sets of square integrable random variables  $\{X_i\}, \{W_i\}, \{Z_i\}, \{Z_{ik}\}, \{W_{ik}\}, \{V_{ik}\}, i \in I, k \in K_i$ , in the following way:

$$W = \sum_{i \in I} X_i; \tag{2.1}$$

$$\mathbb{E}X_i = 0, i \in I; \quad \mathbb{E}W^2 = 1; \tag{2.2}$$

$$W = W_i + Z_i, i \in I, \quad \text{where } W_i \text{ is independent of } X_i; \tag{2.3}$$

$$Z_i = \sum_{k \in K_i} Z_{ik}, \quad i \in I; \tag{2.4}$$

$$W_i = W_{ik} + V_{ik}, i \in I, k \in K_i, \\ \text{where } W_{ik} \text{ is independent of the pair } (X_i, Z_{ik}). \tag{2.5}$$

A simple example of such a decomposition, used for counting the number of induced subgraphs of a given kind, is given in (3.1), (3.2) below. Then the following lemma may be proved:

LEMMA 1. *Let  $W$  be decomposed as in (2.1)–(2.5). Then for every bounded function  $f: \mathbb{R} \rightarrow \mathbb{R}$  with bounded first and second derivatives,*

$$|\mathbb{E}(Wf(W) - f'(W))| \leq C\varepsilon,$$

where  $C = \sup_x |f''(x)|$  and

$$\varepsilon := \frac{1}{2} \sum_{i \in I} \mathbb{E}(|X_i| Z_i^2) + \sum_{i \in I} \sum_{k \in K_i} (\mathbb{E}|X_i Z_{ik} V_{ik}| + \mathbb{E}|X_i Z_{ik}| \mathbb{E}|Z_i + V_{ik}|). \tag{2.6}$$

*Remark.* 1. In Theorem 1 below, the quantity  $\varepsilon$  is shown to measure the  $d_1$  distance between the distribution of  $W$  and the standard normal distribution. Hence the decomposition (2.1)–(2.5) is to be chosen to make  $\varepsilon$  as small as possible.

2. Infinite index sets  $I$  and  $K_i$  can also be allowed, provided that the sums in (2.1) and (2.4) are  $L_2$ -convergent.

*Proof.* The proof relies largely on Taylor's expansion. First, write

$$\begin{aligned} \mathbb{E}(Wf(W)) - \mathbb{E}f'(W) &= \left\{ \mathbb{E}(Wf(W)) - \sum_{i \in I} \mathbb{E}(X_i Z_i f'(W_i)) \right\} \\ &\quad + \left\{ \sum_{i \in I} \mathbb{E}(X_i Z_i f'(W_i)) - \sum_{i \in I} \sum_{k \in K_i} \mathbb{E}(X_i Z_{ik}) \mathbb{E}f'(W_{ik}) \right\} \\ &\quad + \sum_{i \in I} \sum_{k \in K_i} \left\{ \mathbb{E}(X_i Z_{ik}) [\mathbb{E}f'(W_{ik}) - \mathbb{E}f'(W)] \right\}, \end{aligned}$$

which is possible, since, by (2.1)–(2.4),

$$\sum_{i \in I} \sum_{k \in K_i} \mathbb{E}(X_i Z_{ik}) = \sum_{i \in I} \mathbb{E}(X_i Z_i) = \sum_{i \in I} \mathbb{E}(X_i W) = \mathbb{E}W^2 = 1.$$

By (2.1) and (2.3), we have

$$Wf(W) = \sum_{i \in I} X_i f(W) = \sum_{i \in I} X_i \left\{ f(W_i) + Z_i f'(W_i) + \frac{1}{2} Z_i^2 f''(W_i + \theta_i Z_i) \right\}$$

for some  $\theta_i \in [0, 1]$ . Thus, applying (2.2) and (2.3),

$$\left| \mathbb{E}(Wf(W)) - \sum_{i \in I} \mathbb{E}(X_i Z_i f'(W_i)) \right| \leq \frac{1}{2} C \sum_{i \in I} \mathbb{E}(|X_i| Z_i^2).$$

Moreover, by (2.4) and (2.5),

$$\begin{aligned} X_i Z_i f'(W_i) &= \sum_{k \in K_i} X_i Z_{ik} f'(W_i) \\ &= \sum_{k \in K_i} X_i Z_{ik} \{ f'(W_{ik}) + V_{ik} f''(W_{ik} + \theta_{ik} V_{ik}) \}. \end{aligned}$$

So, using (2.5) again, we obtain

$$\left| \sum_{i \in I} \mathbb{E}(X_i Z_i f'(W_i)) - \sum_{i \in I} \sum_{k \in K_i} \mathbb{E}(X_i Z_{ik}) \mathbb{E}f'(W_{ik}) \right| \leq C \sum_{i \in I} \sum_{k \in K_i} \mathbb{E}|X_i Z_{ik} V_{ik}|.$$

In turn, by (2.3) and (2.5),  $W_{ik} = W_i - V_{ik} = W - Z_i - V_{ik}$ , so that

$$f'(W_{ik}) = f'(W) - (Z_i + V_{ik}) f''(W - \theta(Z_i + V_{ik})),$$

and, as a result,

$$|\mathbb{E}(f'(W_{ik})) - \mathbb{E}f'(W)| \leq C\mathbb{E}|Z_i + V_{ik}|,$$

which completes the proof. ■

**THEOREM 1.** *If  $\{W^{(n)}\}_{n=1}^\infty$  is a sequence of random variables decomposed as in (2.1)–(2.5),  $d_1(\mathcal{L}(W^{(n)}), \mathcal{N}(0, 1)) \leq K\varepsilon^{(n)}$ , for a universal constant  $K$ , where  $\varepsilon^{(n)}$  is as in (2.6).*

*Proof.* It is well known (cf. Billingsley [5, p. 345], remark following the proof of Theorem 25.8) that  $W^{(n)} \rightarrow^{\mathcal{D}} \mathcal{N}(0, 1)$  if and only if, for every bounded function  $h: \mathbb{R} \rightarrow \mathbb{R}$  with bounded first derivative and such that  $\mathbb{E}h(N) = 0$ ,  $\mathbb{E}h(W^{(n)}) \rightarrow 0$  as  $n \rightarrow \infty$ , where  $N$  denotes a standard normal random variable. Here, we wish to obtain an analogous rate of convergence using (1.2), by showing that  $|\mathbb{E}h(W^{(n)})| \leq K\varepsilon^{(n)}\|h\|$ . It is easy to check that every such function  $h$  can be expressed as  $h(x) = f'(x) - xf(x)$ , where

$$f(x) = -e^{x^2/2} \int_x^\infty e^{-t^2/2} h(t) dt.$$

To conclude the proof, it remains to be shown that  $\sup_x |f''(x)| \leq K[\sup_x |h(x)| + \sup_x |h'(x)|]$ , and to apply Lemma 1: this can be accomplished as in the proof of Theorem A of Barbour and Eagleson [2]. ■

The decomposition (2.1)–(2.5) is chosen explicitly to match the argument used in Lemma 1, which is typical of those used in the exploitation of Stein’s method. The notion of finite dependence used by Chen [7, 8] corresponds to a decomposition analogous to (2.1)–(2.5), but with the added restriction that  $Z_{ik} = X_k$  in (2.4), and with a similar modification to (2.5). Dissociated random variables, as introduced by McGinley and Sibson [18], exhibit finite dependence: their indices  $i$  are  $r$ -tuples  $\{i_1, \dots, i_r\}$  of positive integers, and two collections of random variables  $\{X_j, j \in J\}$  and  $\{X_l, l \in L\}$  are independent whenever  $(\cup_{j \in J} \{j_1, \dots, j_r\}) \cap (\cup_{l \in L} \{l_1, \dots, l_r\}) = \emptyset$ . Thus, taking  $I = \{1, 2, \dots, n\}^r$  and  $K_i = L_i := \{k \in I: \{k_1, \dots, k_r\} \cap \{i_1, \dots, i_r\} \neq \emptyset\}$  in (2.1)–(2.5), we obtain a decomposition with  $Z_{ik} := X_k$  and  $V_{ik} := \sum_{l \in K_k \setminus K_i} X_l$ . A version of Theorem 1 for dissociated random variables, with an estimate of  $\varepsilon$  simplified so as to incorporate only the moments  $\mathbb{E}|X_i|^3$ , appears as Theorem 2.1 in Barbour and Eagleson [2]: here, in this case, we use the sharper estimate

$$\varepsilon \leq 2 \sum_{i \in I} \sum_{k, l \in L_i} \{ \mathbb{E}|X_i X_k X_l| + \mathbb{E}|X_i X_k| \mathbb{E}|X_l| \}, \tag{2.7}$$

which is better suited to random graph applications.

Further specialization leads to dissociated random variables expressible in the form  $X_i = \phi(Y_{i_1}, \dots, Y_{i_r})$ , where  $(Y_j)_{j \geq 1}$  is an underlying sequence of

independent random variables. When the  $Y_j$  are identically distributed, and the same sequence and the same function  $\phi$  (up to a multiplicative constant) are used to generate  $W^{(n)}$  for each  $n$ , the resulting structure is that of  $U$ -statistics (Hoeffding [10, 11]; see also Serfling [23]): when  $I^{(n)}$  is allowed to be only a part of  $\{1, 2, \dots, n\}^r$ , partial  $U$ -statistics are obtained, and have been used by Nowicki [19] in a random graph context. Finally, specializing to the case  $r = 1$ , independent random variables  $X_1, \dots, X_n$  are obtained, and the estimate (2.7) gives  $\varepsilon = O(\sum_{i=1}^n \mathbb{E}|X_i|^3)$ , the classical Lyapounov estimate. As in the classical theory of partial sums, judicious use of truncation can weaken the moment assumptions required.

### 3. APPLICATIONS TO RANDOM GRAPHS: DISSOCIATED PROPERTIES

Let  $K(n, p)$  be a binomial random graph on the vertex set  $N_n := \{1, 2, \dots, n\}$  in which edges appear independently with the same probability  $p = p(n)$ . In this section, we consider examples of numerical characteristics of  $K(n, p)$  which have Chen's finite dependence. First, let  $S^{(n)}$  be the number of induced subgraphs of  $K(n, p)$  isomorphic to a given graph  $G$  with  $r$  vertices, and let  $I_n := \{(i_1, \dots, i_r) : 1 \leq i_1 < i_2 < \dots < i_r \leq n\}$  index the  $r$ -element subsets of the set  $N_n$ . Define the indicator random variables  $Y_i^{(n)}$ ,  $i \in I_n$ , as follows:

$$Y_i^{(n)} = \begin{cases} 1 & \text{if the subgraph of } K(n, p) \text{ induced by } i \text{ is isomorphic to } G; \\ 0 & \text{otherwise.} \end{cases} \quad (3.1)$$

Then

$$W^{(n)} := (S^{(n)} - \mathbb{E}S^{(n)})/\sigma_n = \sum_{i \in I} X_i^{(n)}, \quad (3.2)$$

where  $X_i^{(n)} = (Y_i^{(n)} - \mathbb{E}Y_i^{(n)})/\sigma_n$  and  $\sigma_n^2 = \text{var } S^{(n)}$ , and the family  $(X_i^{(n)})_{i \in I_n}$  is dissociated, since the collections  $\{X_j, j \in J\}$  and  $\{X_l, l \in L\}$  are generated by distinct sets of (independent) edges whenever  $(\cup_{j \in J} \{j_1, \dots, j_r\}) \cap (\cup_{l \in L} \{l_1, \dots, l_r\}) = \emptyset$ . Thus Theorem 1 can be applied to  $W^{(n)}$  using estimate (2.7) for  $\varepsilon$ .

However, it is to be noted that the family  $(X_i^{(n)})_{i \in I_n}$  enjoys a stronger independence structure than dissociation, in that, for instance,  $X_i$  and  $X_j$  are independent also if  $|\{i_1, \dots, i_r\} \cap \{j_1, \dots, j_r\}| = 1$ . Thus more terms than are necessary appear in the sums in (2.7), and it is advantageous to return to (2.1)–(2.5) and define

$$K_i := L'_i := \{j : |\{i_1, \dots, i_r\} \cap \{j_1, \dots, j_r\}| \geq 2\}; \quad (3.3)$$

$$Z_i := \sum_{k \in K_i} X_k; \quad V_{ik} := \sum_{l \in K_i \setminus K_k} X_l,$$

whence an estimate

$$\varepsilon \leq 2 \sum_{i \in I} \sum_{k, l \in L_i} \{ \mathbb{E}|X_i X_k X_l| + \mathbb{E}|X_i X_k| \mathbb{E}|X_l| \} \tag{3.4}$$

analogous to (2.7) is obtained, but with fewer terms in the sums because  $L_i$  is replaced by  $L'_i$ . Note that the extra independence leads to a smaller  $\sigma$  than might typically be expected for dissociated random variables, and since  $\sigma^3$  appears in the denominator of  $\varepsilon$ , the improvement over (2.7) obtained in (3.4) can be of critical importance.

There is another way of looking at such problems. Instead of considering  $S^{(n)}$  as a sum of components indexed by subsets of vertices, one can consider it as a sum indexed by subsets of edges. From this viewpoint, if  $(e_j)_{1 \leq j \leq \binom{n}{2}}$  denote the edges of the complete graph on  $n$  vertices, we define

$$I_n := \left\{ i = \{i_1, \dots, i_R\} : 1 \leq i_1 < i_2 < \dots < i_R \leq \binom{n}{2}, \right. \\ \left. \{e_{i_1}, \dots, e_{i_R}\} \text{ form a complete graph on } r \text{ vertices} \right\},$$

where  $R = \binom{r}{2}$ . Then  $W^{(n)}$  takes the form

$$W^{(n)} = \sum_{i \in I_n} X_i = \sum_{i \in I_n} \phi(E_{i_1}^{(n)}, \dots, E_{i_R}^{(n)}), \tag{3.5}$$

where  $(E_i^{(n)})_{1 \leq i \leq R}$  are independent Bernoulli random variables with  $\mathbb{P}[E_i^{(n)} = 1] = p(n)$ , and therefore, for each  $n$ , have the structure of a partial  $U$ -statistic: see Nowicki [19]. With this choice of index set, the family  $(X_i)_{i \in I_n}$  is dissociated in precisely the usual sense, and (2.7) can be used directly.

We illustrate these considerations by examining the problem recently settled by Ruciński [22] (see also Nowicki and Wierman [21]), of when the number of copies (not necessarily induced) of a given subgraph  $G$  in  $K(n, p)$  is asymptotically normally distributed. The following theorem strengthens his result, by giving a rate of convergence in the  $d_1$  metric. We suppress further explicit mention of  $n$  where possible. For any graph  $H$ , let  $v(H)$  and  $e(H)$  denote, respectively, the number of its vertices and edges.

**THEOREM 2.** *Let  $S$  denote the number of copies (not necessarily induced) of  $G$  in  $K(n, p)$ . Then*

$$d_1(\mathcal{L}(W), \mathcal{N}(0, 1)) \leq k(G) \begin{cases} \psi^{-1/2} & \text{if } p \leq \frac{1}{2} \\ n^{-1}(1-p)^{-1/2} & \text{if } p > \frac{1}{2}, \end{cases}$$

where  $W = (S - \mathbb{E}S) / \sqrt{\text{var } S}$ ,  $\psi = \min_{H \subset G, e(H) > 0} \{n^{v(H)} p^{e(H)}\}$ , attained at  $H^*$ , say, and  $k(G)$  is a constant depending on  $G$ .

*Proof.* Let  $v = v(G)$ ,  $e = e(G)$ . We apply the dissociated representation (3.5) with

$$I := \left\{ i: 1 \leq i_1 < \dots < i_e \leq \binom{n}{2}, \{e_{i_1}, \dots, e_{i_e}\} \text{ is a copy of } G \right\}; \quad (3.6)$$

$$Y_i := \prod_{l=1}^e E_{i_l}; \quad \phi(x_1, \dots, x_e) = \sigma^{-1} \left( \prod_{l=1}^e x_l - p^e \right),$$

where  $\sigma^2 = \text{var } S$ . Let  $G_i := \{e_{i_1}, \dots, e_{i_e}\}$  for  $i \in I_n$ , and note that, as in Ruciński [22],

$$\begin{aligned} \sigma^2 &= \sum_{i \in I} \sum_{j \in L_i} \text{cov}(Y_i, Y_j) = \sum_{\substack{H \subset G \\ e(H) \geq 1}} \sum_{\substack{i, j \in I \\ G_i \cap G_j = H}} p^{2e-e(H)}(1-p^{e(H)}) \\ &\sim \sum_{\substack{H \subset G \\ e(H) \geq 1}} c_H n^{2v-v(H)} p^{2e-e(H)}(1-p^{e(H)}) \\ &\geq (1-p) n^{2v} p^{2e} \sum_{\substack{H \subset G \\ e(H) \geq 1}} c_H n^{-v(H)} p^{-e(H)} \geq (1-p) n^{2v} p^{2e} c_H \psi^{-1}, \end{aligned} \quad (3.7)$$

where  $c_H$  is a combinatorial constant depending on  $G$  and  $H$ . On the other hand, using (2.7),

$$\varepsilon \leq \left\{ \frac{32}{\sigma^3} \sum_{i \in I} \sum_{k, l \in L_i} \mathbb{E}(Y_i Y_k Y_l) \right\} \wedge \left\{ \frac{8}{\sigma^3} \sum_{i \in I} \sum_{k, l \in L_i} \mathbb{E}(1 - Y_i) \right\}, \quad (3.8)$$

where the first estimate arises from the 16 possible terms in the expansion of the summand in (2.7) in terms of the  $Y_i$ , and the second arises from the inequalities

$$\sigma |X_i| \leq 1 \quad \text{and} \quad \sigma \mathbb{E}|X_i| = \mathbb{E}|1 - Y_i - \mathbb{E}(1 - Y_i)| \leq 2\mathbb{E}(1 - Y_i).$$

For  $p \geq \frac{1}{2}$ , the latter term directly yields the estimate

$$\varepsilon \leq c\sigma^{-3} n^{3v-4} (1-p) \leq cn^{3v-4} (1-p) (n^{2v-2} (1-p))^{-3/2} \asymp n^{-1} (1-p)^{-1/2}, \quad (3.9)$$

where the generic constants  $c$  are uniform in  $p \geq \frac{1}{2}$ . The former term gives

$$\begin{aligned} \varepsilon &\leq 32\sigma^{-3} \sum_{\substack{H \subset G \\ e(H) \geq 1}} \sum_{\substack{i, k \in I \\ G_i \cap G_k \cong H}} \sum_{\substack{K \subset (G_i \cup G_k) \\ e(K) \geq 1}} \sum_{\substack{l \in I \\ G_l \cap (G_i \cup G_k) = K}} p^{3e-e(H)-e(K)} \\ &\leq 32\sigma^{-3} \sum_{\substack{H \subset G \\ e(H) \geq 1}} \sum_{\substack{i, k \in I \\ G_i \cap G_k \cong H}} \sum_{\substack{K \subset (G_i \cup G_k) \\ K \subset G_l \text{ for some } l \\ e(K) \geq 1}} c p^{2e-e(H)} n^{v-v(K)} p^{e-e(K)} \\ &\leq c\sigma^{-1} \psi^{-1} n^v p^e, \end{aligned} \quad (3.10)$$

where the generic constant  $c$  is uniform in  $p \leq \frac{1}{2}$ . Hence, from (3.7),  $\varepsilon \leq c\psi^{-1/2}$  uniformly in  $p \leq \frac{1}{2}$ , and the theorem follows.  $\blacksquare$

*Remarks.* 1. Note that  $\psi^{-1/2} \asymp n^{-1}$  for all fixed  $p$ ,  $0 < p < 1$ , and that the value  $p = \frac{1}{2}$  is chosen arbitrarily to separate the two cases in the statement of the theorem.

2. The estimate (3.10) of  $\varepsilon$  and (3.7) can still be used if  $G = G^{(n)}$  is allowed to vary with  $n$ , provided that, for asymptotic calculations, the dependence on  $n$  of the  $G$ -dependent constants is taken into account.

The problem mentioned at the beginning of the section, of letting  $S$  be the number of induced subgraphs isomorphic to a given  $G$ , has been previously considered under a variety of circumstances by Maehara [17], Nowicki [19], and Janson [13]. Here we give an essentially complete solution to the convergence problem, and add a rate of convergence in (3.11) below, valid away from the critical value  $p^* := e/(\binom{v}{2})$  of  $p$ . The problem is not exactly equivalent to that of Theorem 2, unless  $G$  is empty or a complete subgraph. It is clear that, for induced subgraphs, it is enough to consider  $p \leq \frac{1}{2}$ , since the complement of  $G$  can be considered if  $p > \frac{1}{2}$ . For small  $p$ , the arguments used in Theorem 2 lead to the same conclusion, but for  $p$  of order 1 the variance estimates are different. In particular, for any  $\delta > 0$ , one can establish

$$d_1(\mathcal{L}(W^{(n)}), \mathcal{N}(0, 1)) \leq k_\delta(G) \psi^{-1/2} \tag{3.11}$$

(in the notation of Theorem 2) uniformly for all  $p$  such that  $0 \leq p \leq \frac{1}{2}$  and  $|p - p^*| \geq \delta$ , but for  $p = p^*$  the variance of  $S$  is  $O(n^{2v-3})$  instead of  $\asymp n^{2v-2}$ , and the estimate of  $d_1$  from Theorem 1 is  $O(n^{1/2})$  instead of  $O(n^{-1})$ . This critical case is interesting, and deserves elaboration.

Let

$$\tilde{Y}_i := Y_i - \mathbb{E}Y_i, \quad i \in I := \{(i_1, \dots, i_v) : 1 \leq i_1 < \dots < i_v \leq n\},$$

and  $\tilde{S} := S - \mathbb{E}S$ , where  $Y_i$  and  $S$  are defined as in (3.1), and set

$$E_{ij} := I[\text{the edge } (i, j) \text{ is present in } K(n, p)].$$

Then we can write

$$\sigma^2 := \text{var } S = \sum_{r=2}^v \sum_{|i \cap j|=r} \mathbb{E}(\tilde{Y}_i \tilde{Y}_j), \tag{3.12}$$

in which the only term of leading order  $n^{2v-2}$ , that with  $r = 2$ , contributes  $\binom{n}{2} \binom{n-2}{v-2} \binom{n-2}{v-2} \mathbb{E}(\mathbb{E}(\tilde{Y}_{(1, \dots, v)} | E_{12})^2)$ . However, if  $Y_{(1, \dots, v)}$  and  $E_{12}$  are independ-

ent, as is the case when  $p = p^*$ , it follows that  $\mathbb{E}(\tilde{Y}_{(1, \dots, v)} | E_{12}) \equiv 0$ , and thus  $\sigma^2 = O(n^{2v-3})$ . The corresponding phenomenon is observed in the Hoeffding [10] projection,

$$\tilde{S} = \sum_{i_1 < i_2} Z_{i_1 i_2} + \left\{ \tilde{S} - \sum_{i_1 < i_2} Z_{i_1 i_2} \right\}, \tag{3.13}$$

where  $Z_{i_1 i_2} := \mathbb{E}(\tilde{S} | E_{i_1 i_2})$ ; the first term, a sum of independent identically distributed random variables, each with the distribution of  $\binom{n-2}{v-2} \mathbb{E}(\tilde{Y}_{(1, \dots, v)} | E_{12})$ , dominates  $\tilde{S}$ , unless  $p = p^*$ , when  $Z_{i_1 i_2} \equiv 0$ .

If  $p = p^*$ , the next orthogonal projection yields

$$\tilde{S} = \sum_{i_1 < i_2 < i_3} Z_{i_1 i_2 i_3} + \left\{ \tilde{S} - \sum_{i_1 < i_2 < i_3} Z_{i_1 i_2 i_3} \right\}, \tag{3.14}$$

where the random variables  $Z_{i_1 i_2 i_3} := \mathbb{E}(\tilde{S} | E_{i_1 i_2}, E_{i_1 i_3}, E_{i_2 i_3})$  are pairwise independent and identically distributed, with the distribution of  $\binom{n-3}{v-3} \mathbb{E}(\tilde{Y}_{(1, \dots, v)} | E_{12}, E_{13}, E_{23})$ . In such “degenerate” cases, in the usual  $U$ -statistic setting, normal limits cannot occur, but here, with very incomplete  $U$ -statistics, they can and indeed do, as the following result shows.

**THEOREM 3.** *With the above notation, suppose that  $p = p^* := e/\binom{v}{2}$  and that  $\tau^2 := \mathbb{E}\{\mathbb{E}(\tilde{Y}_{(1, \dots, v)} | E_{12}, E_{13}, E_{23})^2\} > 0$ . Then  $\sigma^{-1} \tilde{S} \rightarrow^{\mathcal{D}} \mathcal{N}(0, 1)$ .*

*Proof.* The only term of leading order  $n^{2v-3}$  in (3.12) is that with  $r = 3$ , yielding

$$\sigma^2 = \binom{n}{3} \binom{n-3}{v-3} \binom{n-v}{v-3} \tau^2 \{1 + O(n^{-1})\} \asymp n^{2(v-3/2)}, \tag{3.15}$$

whereas

$$\text{var} \left( \sum_{i_1 < i_2 < i_3} Z_{i_1 i_2 i_3} \right) = \binom{n}{3} \binom{n-3}{v-3}^2 \tau^2 = \sigma^2 \{1 + O(n^{-1})\}. \tag{3.16}$$

Hence  $\sigma^{-1} \tilde{S}$  and  $\sigma^{-1} \sum_{i_1 < i_2 < i_3} Z_{i_1 i_2 i_3}$  have the same asymptotic behaviour.

To prove that  $\sum_{i_1 < i_2 < i_3} Z_{i_1 i_2 i_3}$  is asymptotically normally distributed, we resort to the method of moments. The key observation is that  $\mathbb{E}(\prod_{j=1}^k Z_{i_{j_1} i_{j_2} i_{j_3}}) = 0$  if, for any  $1 \leq l \leq k$ ,  $|U'_l \cap U_l| \leq 1$ , where  $U'_l := \{(i_{l1}, i_{l2}), (i_{l1}, i_{l3}), (i_{l2}, i_{l3})\}$  and  $U_l := \bigcup_{j \neq l} \{(i_{j1}, i_{j2}), (i_{j1}, i_{j3}), (i_{j2}, i_{j3})\}$ . This is because  $\mathbb{E}(\prod_{j=1}^k Z_{i_{j_1} i_{j_2} i_{j_3}} | \{E_{i_1 i_2}, (i_1, i_2) \in U_l\})$  contains either  $\mathbb{E}(\tilde{S}) = 0$  or  $\mathbb{E}(\tilde{S} | E_{j_1 j_2}) \equiv 0$  as a factor, where  $(j_1, j_2) \in U'_l \cap U_l$ . In particular, if any component of an index appears exactly once in the product, the expectation is zero. Hence, in computing  $\mu_k := \mathbb{E}\{(\sum_{i_1 < i_2 < i_3} Z_{i_1 i_2 i_3})^k\}$ , only products involving

$3k/2$  or fewer distinct index components need be considered. Thus, for  $k = 2m + 1$ ,  $\mu_k = O(n^{3m+1+k(v-3)}) = O(n^{-1/2} \cdot n^{k(v-3/2)})$ , whereas, for  $k = 2m$ ,  $\mu_k = O(n^{3m+k(v-3)}) = O(n^{k(v-3/2)})$ , so that only the even standardized moments are non-negligible. Finally, the term of leading order in the  $2m$ th moment arises exclusively from products in which each  $Z_{i_1 i_2 i_3}$  occurs exactly twice and the index sets are otherwise mutually disjoint, since, for each  $l$ ,  $|U_l' \cap U_l| \geq 2$  can only be attained with each of  $i_{l1}, i_{l2}$ , and  $i_{l3}$  appearing exactly twice among the index components if  $(i_{l1}, i_{l2}, i_{l3}) = (i_{j1}, i_{j2}, i_{j3})$  for some  $j \neq l$ , and the contribution from all products with a greater degree of index overlap is of relative order  $n^{-1}$ . ■

This analysis is, however, as yet incomplete: it can be the case, for some choices of  $G$ , that  $\tau^2 = 0$  also. An example is given by the graph

$$G := \{(1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (1, 7), (2, 3), (2, 4), (2, 5), (3, 5), (4, 5), (5, 8), (6, 8), (7, 8)\}$$

on 8 vertices, for which  $p^* = \frac{1}{2}$ : note that the numbers of triangles in  $G$  with  $(0, 1, 2, 3)$  edges present are  $(7, 21, 21, 7)$ , in precisely the binomial  $(p^3, 3p^2(1-p), 3p(1-p)^2, (1-p)^3)$  proportions at  $p = p^*$ . Here one must use the projection

$$\tilde{S} = \sum_{i_1 < i_2 < i_3 < i_4} Z_{i_1 i_2 i_3 i_4} + R$$

with

$$Z_{i_1 i_2 i_3 i_4} := \mathbb{E}\{\tilde{S} | (E_{ijk}, 1 \leq j < k \leq 4)\},$$

where  $Z_{i_1 i_2 i_3 i_4}$  has the distribution of  $\binom{n-4}{v-4} \mathbb{E}\{\tilde{Y}_{(1, \dots, v)} | (E_{jk}, 1 \leq j < k \leq 4)\}$ . In this case,

$$\text{var } S \sim \text{var} \left\{ \sum_{i_1 < i_2 < i_3 < i_4} Z_{i_1 i_2 i_3 i_4} \right\} = O(n^{2v-4}). \tag{3.17}$$

We can then complete the investigation with the following theorem.

**THEOREM 4.** *If, in the above setting,  $p = p^* = e/\binom{v}{2}$  and  $\tau^2 = 0$ , the distribution of  $\sigma^{-1} \tilde{S}$  converges in distribution to a non-trivial non-normal limit.*

*Proof.* It is easy to check, in a fashion analogous to the proof of Theorem 3, that the moments of  $n^{-(v-2)} \tilde{S}$  converge, and that the  $2m$ th moment is of order  $2^{-m}(2m)!$  in  $m$ , so that the limiting moments define a distribution. Hence, by the Fréchet–Shohat theorem,  $n^{-(v-2)}$  converges in distribution.

It remains to be shown that  $\text{var}(n^{-(v-2)} \tilde{S}) \asymp 1$  and that the limit is not normal. The former follows because, when  $\tau^2 = 0$  and  $p = p^*$ ,

$$\mathbb{E}(\tilde{Y}_{(1,\dots,v)} | E_{12}, E_{34}) \begin{cases} < 0 & \text{if } E_{12} = E_{34} \\ > 0 & \text{if } E_{12} \neq E_{34}. \end{cases} \tag{3.18}$$

To see this, note that the fraction of non-incident pairs of edges in the complete graph on  $G$ 's vertices such that both edges are present in  $G$  is

$$f := \left[ \frac{1}{2} e(e-1) - n_\nu \right] / \left[ \frac{1}{2} \binom{v}{2} \left( \binom{v}{2} - 1 \right) - \binom{v}{2} (v-2) \right],$$

where  $n_\nu$  denotes the number of incident pairs of edges in  $G$ . If  $\tau^2 = 0$  at the critical probability  $p = p^*$ ,  $G$  is balanced with respect to triangles, so that

$$n_\nu = \{3p^3 + 3p^2(1-p)\} \binom{v}{3} = \binom{v}{2} p^2(v-2),$$

and hence

$$f = p^2 \left\{ 1 - \left( \frac{1}{p} - 1 \right) / \left[ \binom{v}{2} - 2v + 3 \right] \right\} < p^2.$$

A similar argument holds for neither edge present, and the remaining assertion follows because the unconditional expectation  $\mathbb{E} \tilde{Y}_{(1,\dots,v)} = 0$ . Thus  $\mathbb{E} \{ \tilde{Y}_{(1,\dots,v)} | (E_{jk}, 1 \leq j < k \leq 4) \}$  cannot have zero variance, and so  $\text{var}(n^{-(v-2)} \tilde{S}) \asymp 1$  from (3.17).

Note also that, using (3.18), the limit distribution necessarily has negative skewness, because of the contribution to  $\mathbb{E} \tilde{S}^3$  from products of the form  $\mathbb{E}(Z_{1234} Z_{3456} Z_{1256})$ , and so cannot be normal. ■

#### 4. APPLICATIONS TO RANDOM GRAPHS: SEMI-INDUCED PROPERTIES

In Karoński and Ruciński [15], the number of subsets of  $K(n, p)$  with a given property, determined by those edges with at least one endpoint in the subset, is considered. Such properties are called semi-induced, and the most natural examples are “being an isolated tree” and “having a given degree.” Unfortunately, random variables counting subsets of  $K(n, p)$  with a given semi-induced property typically do not possess the structure of Chen’s finite dependence. However, using the notion of decomposability, we are able to prove asymptotic normality for such random variables. Formally, let  $\mathcal{A}$  be a family of subsets of  $N$ . Let

$$E_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are joined in } K(n, p); \\ 0 & \text{otherwise.} \end{cases}$$

For all  $\alpha \in \mathcal{A}$ , define the  $N$ -hedgehog  $H_\alpha(N)$  to be the graph with vertex set  $N$  and edge set  $\{(i, j): |\{i, j\} \cap \alpha| \geq 1; i, j \in N\}$ . Let  $f_\alpha$  be a function defined over all graphs whose vertex set contains  $\alpha$ , and, for any  $\delta \subset N$  such that  $\delta \cap \alpha = \emptyset$ , let

$$Y_\alpha^{(\delta)} := f_\alpha(H_\alpha(N \setminus \delta) \cap K(n, p)), \quad Y_\alpha := Y_\alpha^{(\emptyset)}. \tag{4.1}$$

Define

$$S := \sum_{\alpha \in \mathcal{A}} Y_\alpha, \quad \mu_\alpha := \mathbb{E}Y_\alpha, \quad \sigma^2 := \text{var } S, \tag{4.2}$$

$$X_\alpha := (Y_\alpha - \mu_\alpha)/\sigma, \quad W = \sum_{\alpha \in \mathcal{A}} X_\alpha,$$

so that (2.1) and (2.2) are satisfied. Now define

$$Z_\alpha = \sigma^{-1} \left\{ \sum_{\beta \cap \alpha \neq \emptyset} Y_\beta + \sum_{\beta \cap \alpha = \emptyset} (Y_\beta - Y_\beta^{(\alpha)}) \right\} = \sum_{\beta \in \mathcal{A}} Z_{\alpha\beta}, \tag{4.3}$$

$$W_\alpha = \sum_{\beta \cap \alpha = \emptyset} X_\beta^{(\alpha)} - \mathbb{E}Z_\alpha, \quad X_\beta^{(\alpha)} = \sigma^{-1}(Y_\beta^{(\alpha)} - \mathbb{E}(Y_\beta^{(\alpha)})).$$

It is clear that  $W = W_\alpha + Z_\alpha$ , and  $W_\alpha$  is independent of  $X_\alpha$  for all  $\alpha \in \mathcal{A}$ , since the random variables  $Y_\beta^{(\alpha)}$  are constructed only from edges not belonging to  $H_\alpha(N)$ : hence (2.3) and (2.4) are satisfied, with  $K_x = \mathcal{A}$ . In order to complete the decomposition of  $W$ , we have to perform a “second removal” in the following way:

$$V_{\alpha\beta} = \sigma^{-1} \left\{ \sum_{\substack{\gamma \cap \beta \neq \emptyset \\ \gamma \cap \alpha = \emptyset}} Y_\gamma^{(\alpha)} + \sum_{\substack{\gamma \cap \beta \neq \emptyset \\ \gamma \cap \alpha = \emptyset}} (Y_\gamma^{(\alpha)} - Y_\gamma^{(\alpha \cup \beta)}) \right\}; \tag{4.4}$$

$$W_{\alpha\beta} = \sum_{\gamma \cap (\alpha \cup \beta) = \emptyset} X_\gamma^{(\alpha \cup \beta)} - \mathbb{E}V_{\alpha\beta} - \mathbb{E}Z_\alpha;$$

it follows in a similar way that, for all  $\alpha, \beta \in \mathcal{A}$ ,  $W_\alpha = W_{\alpha\beta} + V_{\alpha\beta}$  and  $W_{\alpha\beta}$  is independent of  $(X_\alpha, Z_{\alpha\beta})$ , so that (2.5) is also satisfied. Thus we are able to formulate the following version of Theorem 1.

**THEOREM 5.** *If random variables  $S, W, X_\alpha, Z_\alpha, Z_{\alpha\beta}$ , and  $V_{\alpha\beta}$  are defined as above,  $d_1(\mathcal{L}(W), \mathcal{N}(0, 1)) \leq K\varepsilon$ , where*

$$\varepsilon = \frac{1}{2} \sum_{\alpha \in \mathcal{A}} \mathbb{E}(|X_\alpha| Z_\alpha^2) + \sum_{\alpha \in \mathcal{A}} \sum_{\beta \in \mathcal{A}} (\mathbb{E}|X_\alpha Z_{\alpha\beta} V_{\alpha\beta}| + \mathbb{E}|X_\alpha Z_{\alpha\beta}| \mathbb{E}|Z_\alpha + V_{\alpha\beta}|), \tag{4.5}$$

and  $K$  is a universal constant. In particular, if  $\varepsilon \rightarrow 0$  as  $n \rightarrow \infty$ ,

$$W = (S - \mathbb{E}S) / \sqrt{\text{var } S} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

We now give some applications of this general result.

**A. Isolated Trees**

Let  $S = S^{(n)}(k)$  be the number of isolated trees of size  $k$  in  $K(n, p)$ . In their seminal paper, Erdős and Rényi [9] proved by the method of moments that, for  $k \geq 2$ ,

$$W = (S - \mathbb{E}S) / \sqrt{\mathbb{E}S} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1), \quad \text{if } \mathbb{E}S \rightarrow \infty \text{ not too fast as } n \rightarrow \infty.$$

Barbour [1] used Stein’s method to show that  $W \rightarrow^{\mathcal{D}} \mathcal{N}(0, 1)$  provided only that  $\mathbb{E}S \rightarrow \infty$  as  $n \rightarrow \infty$ , and to determine the rate of convergence. Theorem 5 is developed from his approach. We use it here to prove convergence in a slightly more general setting.

**THEOREM 6.** *Let  $\mathcal{T}_k$  be a family of  $k$ -vertex trees for some fixed  $k \geq 2$ , and let  $S = S(\mathcal{T}_k)$  count the components of  $K(n, p)$  isomorphic to an element of  $\mathcal{T}_k$ . Then*

$$d_1(\mathcal{L}(W), \mathcal{N}(0, 1)) = O((\mathbb{E}S)^{-1/2}),$$

where  $W := (S - \mathbb{E}S) / \sqrt{\text{var } S}$ . In particular,  $W \rightarrow^{\mathcal{D}} \mathcal{N}(0, 1)$  if  $\mathbb{E}S \rightarrow \infty$  as  $n \rightarrow \infty$ , i.e., if  $n^k p^{k-1} \rightarrow \infty$  but  $kn p - \log n - (k - 1) \log \log n \rightarrow -\infty$ .

*Remark.* Since  $\mathbb{P}[S \notin \{0, 1, \dots, l\}] \leq \mathbb{E}S/l$ ,  $l \geq 1$ , there is no hope of non-trivial asymptotic normality unless  $\mathbb{E}S \rightarrow \infty$ . Erdős and Rényi [9] established Poisson convergence in both ranges where  $\mathbb{E}S \rightarrow c$ .

*Proof.* Let  $\mathcal{A}$  be the family of all  $k$ -vertex subsets of the set  $\{1, \dots, n\}$ . For each  $\alpha \in \mathcal{A}$ , let

$$Y_\alpha = \begin{cases} 1 & \text{if } \alpha \text{ spans in } K(n, p) \text{ a component isomorphic to an element of } \mathcal{T}_k, \\ 0 & \text{otherwise.} \end{cases}$$

Then the indicator  $Y_\alpha$  is derived from a function  $f_\alpha$  as in (4.1). Making use of Theorem 5, we must show that  $\varepsilon = O((\mathbb{E}S)^{-1/2})$ . We first calculate the expectation and variance of the random variable  $S = \sum_{\alpha \in \mathcal{A}} Y_\alpha$ . Observe that

$$\mu_\alpha = \mathbb{E}Y_\alpha = \sum_{T \in \mathcal{T}_k} \frac{k!}{\text{aut}(T)} p^{k-1} (1-p)^{\binom{n-k}{k} + \binom{k}{2} - k + 1} \sim c(\mathcal{T}_k) p^{k-1} e^{-npk}. \tag{4.6}$$

Moreover,

$$\sigma^2 = \text{var } S = \sum_{\alpha \in \mathcal{A}} \sum_{\beta \in \mathcal{A}} \text{cov}(Y_\alpha, Y_\beta) \geq \mathbb{E}S - \sum_{\beta \cap \alpha \neq \emptyset} \mu_\alpha \mu_\beta,$$

since

$$\text{cov}(Y_\alpha, Y_\beta) = \begin{cases} \mu_\alpha - \mu_\alpha^2 & \text{if } \alpha = \beta \\ -\mu_\alpha \mu_\beta & \text{if } \alpha \neq \beta, \alpha \cap \beta \neq \emptyset, \\ \mu_\alpha \mu_\beta ((1-p)^{-k^2} - 1) & \text{if } \alpha \cap \beta = \emptyset. \end{cases}$$

For  $\beta_0 \in \mathcal{A}$ ,

$$\sum_{\alpha \cap \beta_0 \neq \emptyset} \mu_\alpha = \sum_{r=1}^k \binom{k}{r} \binom{n-k}{k-r} \mu_\alpha \sim \frac{kc(\mathcal{T}_k)}{(k-1)!} f(x) := Q_n,$$

where  $f(x) := x^{k-1} e^{-kx}$  and  $x = np$ . Since  $f$  takes its maximum at  $(k-1)/k$ , since  $c(\mathcal{T}_k) \leq k^{k-2}$  and since, for all  $k \geq 1$ ,  $k! > (k/e)^k \sqrt{2\pi k}$ , we conclude that  $Q_n < (2\pi(k-1))^{-1/2} < 1$  for all  $n$  and for all  $k \geq 2$ . Thus

$$\sigma^2 \geq \mathbb{E}S(1 - Q_n) > c\mathbb{E}S,$$

for some  $c > 0$  and for  $n$  large enough.

The proof is now completed by showing that each of the three terms in expression (4.5) for  $\varepsilon$  is of order  $O(\sigma^{-3}\mathbb{E}S)$ , which, with the above lower bound for  $\sigma$ , gives the desired result. To start with, we show that

$$\begin{aligned} & \sum_{\alpha \in \mathcal{A}} \sum_{\beta \in \mathcal{A}} \mathbb{E}|X_\alpha Z_{\alpha\beta} V_{\alpha\beta}| \\ &= \sigma^{-1} \sum_{\beta \cap \alpha \neq \emptyset} \mathbb{E}|X_\alpha Y_\beta V_{\alpha\beta}| + \sigma^{-1} \sum_{\beta \cap \alpha = \emptyset} \mathbb{E}|X_\alpha (Y_\beta - Y_\beta^{(\alpha)}) V_{\alpha\beta}| \\ &= O(\sigma^{-3}\mathbb{E}S). \end{aligned} \tag{4.7}$$

Take first the terms with  $\alpha \cap \beta \neq \emptyset$ . Then  $Y_\alpha Y_\beta = 0$  if  $\alpha \neq \beta$ , whereas  $V_{\alpha\beta} = 0$  if  $\alpha = \beta$ . Thus

$$\mathbb{E}|X_\alpha Y_\beta V_{\alpha\beta}| \leq \sigma^{-1} \mathbb{E}[(Y_\alpha + \mu_\alpha) | Y_\beta V_{\alpha\beta} ] = \sigma^{-1} \mu_\alpha \mathbb{E}|Y_\beta V_{\alpha\beta}|.$$

Moreover,  $Y_\beta Y_\gamma^{(\alpha)} = 0$  if  $\gamma \cap \beta \neq \emptyset$  and  $\gamma \cap \alpha = \emptyset$ , whereas  $Y_\beta (Y_\gamma^{(\alpha)} - Y_\gamma^{(\alpha \cup \beta)}) = 0$  if  $\gamma \cap \beta = \emptyset$  and  $\gamma \cap \alpha = \emptyset$ . Hence  $\mathbb{E}|Y_\beta Y_{\alpha\beta}| = 0$ . Similarly, taking the terms with  $\alpha \cap \beta = \emptyset$ , it follows that

$$\mathbb{E}|X_\alpha (Y_\beta - Y_\beta^{(\alpha)}) V_{\alpha\beta}| = \sigma^{-1} \mathbb{E}|X_\alpha (Y_\beta - Y_\beta^{(\alpha)}) Y_\beta^{(\alpha)}|$$

from the term in  $V_{\alpha\beta}$  with  $\gamma = \beta$ , and hence that

$$\sigma^{-1} \sum_{\beta \cap \alpha = \emptyset} \mathbb{E}|X_\alpha (Y_\beta - Y_\beta^{(\alpha)}) V_{\alpha\beta}| = O(n^{2k} p^{2k-1} e^{-2npk} \sigma^{-3}) = (\sigma^{-3}\mathbb{E}S).$$

Next, we observe that, since  $Y_\beta - Y_\beta^{(\alpha)} \leq 0$ ,

$$\begin{aligned} & \sum_{\alpha \in \mathcal{A}} \mathbb{E}(|X_\alpha| Z_\alpha^2) \\ & \leq \sigma^{-3} \sum_{\alpha \in \mathcal{A}} \mathbb{E} \left\{ (Y_\alpha + \mu_\alpha) \left[ \left( \sum_{\beta \cap \alpha \neq \emptyset} Y_\beta \right)^2 + \left( \sum_{\beta \cap \alpha = \emptyset} (Y_\beta - Y_\beta^{(\alpha)}) \right)^2 \right] \right\}. \end{aligned} \quad (4.8)$$

For the first term in (4.8) we have

$$\sum_{\alpha \in \mathcal{A}} \sum_{\beta \cap \alpha \neq \emptyset} \sum_{\gamma \cap \alpha \neq \emptyset} \mathbb{E}(Y_\alpha Y_\beta Y_\gamma) = \mathbb{E}S, \quad (4.9)$$

since  $Y_\alpha Y_\beta Y_\gamma = 1$  if and only if  $\alpha = \beta = \gamma$ , and

$$\sum_{\alpha \in \mathcal{A}} \mu_\alpha \sum_{\beta \cap \alpha \neq \emptyset} \sum_{\gamma \cap \alpha \neq \emptyset} \mathbb{E}(Y_\beta Y_\gamma) = \sum_{\alpha \in \mathcal{A}} \mu_\alpha O(n^{k-1} p^{k-1} e^{-npk}) = O(\mathbb{E}S). \quad (4.10)$$

For the second term in (4.8), let

$$L_{\alpha, \beta} = \begin{cases} 1 & \text{if there is an edge between } \alpha \text{ and } \beta \text{ in } K(n, p), \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\sum_{\alpha \in \mathcal{A}} \sum_{\beta \cap \alpha = \emptyset} \sum_{\gamma \cap \alpha = \emptyset} \mathbb{E}[Y_\alpha (Y_\beta^{(\alpha)} - Y_\beta)(Y_\gamma^{(\alpha)} - Y_\gamma)] = 0, \quad (4.11)$$

since  $Y_\beta^{(\alpha)} - Y_\beta = 1$  only if  $L_{\alpha\beta} = 1$ , which in turn implies that  $Y_\alpha = 0$ . Furthermore,

$$\sum_{\alpha \in \mathcal{A}} \mu_\alpha \sum_{\beta \cap \alpha = \emptyset} \sum_{\gamma \cap \alpha = \emptyset} \mathbb{E}[(Y_\beta^{(\alpha)} - Y_\beta)(Y_\gamma^{(\alpha)} - Y_\gamma)] = O(\mathbb{E}S), \quad (4.12)$$

since

$$\begin{aligned} & \mathbb{E}[(Y_\beta^{(\alpha)} - Y_\beta)(Y_\gamma^{(\alpha)} - Y_\gamma)] \\ & \leq \mathbb{E}[L_{\alpha\beta} Y_\beta^{(\alpha)} L_{\alpha\gamma} Y_\gamma^{(\alpha)}] = \begin{cases} O(p^{2k} e^{-2npk}), & \beta \cap \gamma = \emptyset \\ 0, & \beta \cap \gamma \neq \emptyset, \beta \neq \gamma \\ O(p^k e^{-npk}), & \beta = \gamma. \end{cases} \end{aligned}$$

Thus, from (4.9)–(4.12),

$$\sum_{\alpha \in \mathcal{A}} \mathbb{E}(|X_\alpha| Z_\alpha^2) = O(\sigma^{-3}\mathbb{E}S).$$

Finally,

$$\begin{aligned} & \sum_{\alpha \in \mathcal{A}} \sum_{\beta \in \mathcal{A}} \mathbb{E}|X_\alpha Z_{\alpha\beta}| \mathbb{E}|Z_\alpha + V_{\alpha\beta}| \\ & \leq \sum_{\alpha \in \mathcal{A}} \sum_{\beta \in \mathcal{A}} \mathbb{E}|X_\alpha Z_{\alpha\beta}| [\mathbb{E}|Z_\alpha| + \mathbb{E}|V_{\alpha\beta}|]. \end{aligned} \tag{4.13}$$

Arguing much as above,

$$\mathbb{E}|Z_\alpha| \leq \sum_{\beta \in \mathcal{A}} \mathbb{E}|Z_{\alpha\beta}| = O(\sigma^{-1})$$

and  $\mathbb{E}|V_{\alpha\beta}|$  likewise: also,  $\sum_{\beta \in \mathcal{A}} Y_\alpha |Z_{\alpha\beta}| = \sigma^{-1} Y_\alpha$ . Thus, immediately,

$$\sigma^{-1} \sum_{\alpha \in \mathcal{A}} \sum_{\beta \in \mathcal{A}} \mu_\alpha \mathbb{E}|Z_{\alpha\beta}| [\mathbb{E}|Z_\alpha| + \mathbb{E}|V_{\alpha\beta}|] = O(\sigma^{-3}\mathbb{E}S),$$

and

$$\sigma^{-1} \sum_{\alpha \in \mathcal{A}} \sum_{\beta \in \mathcal{A}} \mathbb{E}(Y_\alpha |Z_{\alpha\beta}|) [\mathbb{E}|Z_\alpha| + \mathbb{E}|V_{\alpha\beta}|] = O(\sigma^{-3}\mathbb{E}S),$$

so that the third part of  $\varepsilon$  is  $O(\sigma^{-3}\mathbb{E}S)$  also. ■

### B. Vertex Degrees

Let  $S = S_n(d)$  be the number of vertices of degree  $d$  in  $K(n, p)$ . Karoński and Ruciński [15] proved that, for  $d \geq 1$ ,  $W := (S - \mathbb{E}S)/\sqrt{\text{var } S} \rightarrow^{\mathcal{D}} \mathcal{N}(0, 1)$  if and only if  $\mathbb{E}S \rightarrow \infty$  and either  $np \rightarrow 0$  or  $np \rightarrow \infty$ . Here we give a full description of when  $S_n(d)$  is asymptotically normal.

**THEOREM 7.** *If  $d \geq 1$  then  $d_1(\mathcal{L}(W), \mathcal{N}(0, 1)) = O\{(\mathbb{E}S)^{-1/2}\}$ . In particular,  $W \rightarrow^{\mathcal{D}} \mathcal{N}(0, 1)$  if  $\mathbb{E}S \rightarrow \infty$ , i.e., if  $n^{d+1}p^d \rightarrow \infty$  but  $np - \log n - d \log \log n \rightarrow -\infty$ .*

*Remark.* As before, there can be no non-trivial normal limit unless  $\mathbb{E}S \rightarrow \infty$ . In both ranges where  $\mathbb{E}S \rightarrow c$ , Erdős and Rényi [9] established Poisson convergence.

*Proof.* Let  $\mathcal{A}$  be the set  $\{1, \dots, n\}$ . For each  $i \in \mathcal{A}$ , let

$$Y_i := \begin{cases} 1 & \text{if vertex } i \text{ has degree } d \text{ in } K(n, p), \\ 0 & \text{otherwise.} \end{cases}$$

Now  $S = \sum_{i=1}^n Y_i$  and

$$\mathbb{E}S = n\mu = n\mathbb{E}Y_1 = n \binom{n-1}{d} p^d (1-p)^{n-1-d} \sim \frac{1}{d!} n^{d+1} p^d e^{-np}. \quad (4.14)$$

Moreover,

$$\begin{aligned} \sigma^2 = \text{var } S &= \frac{n}{n-1} \binom{n-1}{d}^2 (d - (n-1)p)^2 p^{2d-1} (1-p)^{2(n-d)-3} \\ &\quad + \mathbb{E}S - n^{-1}(\mathbb{E}S)^2 \\ &= O(\mathbb{E}S) + (1 - n^{-1}\mathbb{E}S) \mathbb{E}S \sim c\mathbb{E}S, \quad c > 0, \end{aligned} \quad (4.15)$$

since  $n^{-1}\mathbb{E}S < (2\pi d)^{-1/2} < 1$ .

We show that  $c = O(\sigma^{-3}\mathbb{E}S)$ . Recall that  $\sigma Z_i = Y_i + \sum_{j \neq i} (Y_j - Y_j^{(i)})$ , so that

$$\sigma^2 Z_i^2 = Y_i^2 + 2Y_i \sum_{j \neq i} (Y_j - Y_j^{(i)}) + \sum_{j \neq i} \sum_{k \neq i} (Y_j - Y_j^{(i)})(Y_k - Y_k^{(i)}).$$

Observe that  $|Y_j - Y_j^{(i)}| \leq E_{ij} I[\text{deg}(j) = d \text{ or } d+1]$ . Therefore

$$\begin{aligned} \sigma^2 \mathbb{E}Z_i^2 &\leq \mu + O\{n^{2d-1} p^{2d-1} e^{-2np} (1+np)\} + O\{n^{2d} p^{2d} e^{-2np} (1+np)^2\} \\ &\quad + O\{n^{2d-2} p^{2d-1} e^{-2np} (1+np)^2\} + O\{n^d p^d e^{-np} (1+np)\} = O(1), \end{aligned}$$

where the three contributions from the double sum arise from the cases  $\{j \neq k, E_{jk} = 0\}$ ,  $\{j \neq k, E_{jk} = 1\}$ , and  $\{j = k\}$ . Similarly,

$$\begin{aligned} \sigma^2 \mathbb{E}(Y_i Z_i^2) &\leq \mu \{1 + O((np)^{d-1} e^{-np} (1+np)) + O((np)^{2(d-1)} e^{-2np} (1+np)^2) \\ &\quad + O((np)^{2d-4} p e^{-2np} (1+np)^2) + O((np)^{d-1} e^{-np} (1+np))\} \\ &= O(\mu(1 + n^{-2} p^{-1})) = O(\mu), \end{aligned}$$

since  $d \geq 1$  and  $n^2 p \rightarrow \infty$ . Hence

$$\sum_{i=1}^n \mathbb{E}(|X_i| Z_i^2) \leq \sigma^{-1} \sum_i \mathbb{E}[(Y_i + \mu) Z_i^2] = O(\sigma^{-3}\mathbb{E}S).$$

To estimate  $\sum_{i=1}^n \sum_{j=1}^n \mathbb{E}|X_i Z_{ij} V_{ij}|$  note that  $V_{ij} = 0$  if  $i = j$ . Moreover, if  $i \neq j$ ,  $\sigma V_{ij}$  reduces to  $Y_j^{(i)} + \sum_{l \neq i, j} (Y_l^{(i)} - Y_l^{(i, j)})$ . Thus we obtain

$$\begin{aligned} \sum_{i, j} \mathbb{E}|X_i Z_{ij} V_{ij}| &\leq \sigma^{-2} \sum_{j \neq i} \mathbb{E}[(Y_i + \mu) | V_{ij} | | Y_j - Y_j^{(i)} |] \\ &\leq \sigma^{-3} \sum_{j \neq i} \left\{ \mathbb{E}[|Y_i| | Y_j - Y_j^{(i)} | | Y_j^{(i)} |] + \mu \mathbb{E}[|Y_j - Y_j^{(i)}| | Y_j^{(i)} |] \right. \\ &\quad \left. + \sum_{l \neq i, j} (\mathbb{E}[|Y_i| | Y_j - Y_j^{(i)} | | Y_l^{(i)} - Y_l^{(i, j)} |]) \right. \\ &\quad \left. + \mu \mathbb{E}[|Y_j - Y_j^{(i)}| | Y_l^{(i)} - Y_l^{(i, j)} |]) \right\} \\ &\leq \sigma^{-3} n \mu O\{(np)^{d-1} e^{-np}(1+np) + e^{-2np}(np)^{2(d-1)}(1+np)^3\} \\ &= O(\sigma^{-3} \mathbb{E}S). \end{aligned}$$

Finally, similar arguments yield

$$\sigma \mathbb{E}|Z_i + V_{ij}| = O(\mu(1+np)) = O(1)$$

and

$$\sigma^2 \sum_j \mathbb{E}|X_i Z_{ij}| = O(\mu\{1 + (np)^{d-1} e^{-np}[1 + np]^2\}) = O(\mu),$$

whence the last term of  $\varepsilon$  is estimated as

$$\sum_{i=1}^n \sum_{j=1}^n \mathbb{E}|X_i Z_{ij}| \mathbb{E}|Z_i + V_{ij}| = O(\sigma^{-3} \mathbb{E}S),$$

and the theorem follows. ■

### C. Isolated Vertices

Let  $S = S^{(n)}(0) = S^{(n)}(\mathcal{F}_1)$ , the number of isolated vertices in  $K(n, p)$ . This characteristic, in contrast to the cases  $d \geq 1$  and  $k \geq 2$ , no longer has two Poisson phases with a normal phase between, and must be treated separately.

**THEOREM 8.**  $W := (S - \mathbb{E}S) / \sqrt{\text{var } S} \rightarrow^{\mathcal{D}} \mathcal{N}(0, 1)$  if and only if  $n^2 p \rightarrow \infty$  and  $np - \log n \rightarrow -\infty$  as  $n \rightarrow \infty$ .

*Proof.* In the case when  $np \rightarrow \infty$ , the theorem was proved in Barbour [1]. Assume now that  $np = O(1)$ . One can easily check that the estimate  $\varepsilon = O(\sigma^{-3} \mathbb{E}S)$  shown in the proof of Theorem 6 above holds also for  $k = 1$ .

However, from (4.14) and (4.15), we now have  $\mathbb{E}S \sim ne^{-np}$  and  $\text{var } S \asymp n^2pe^{-2np}$ , so  $\varepsilon \rightarrow 0$  provided that  $n^4p^3 \rightarrow \infty$ . Hence it remains to prove the theorem for  $p = O(n^{-4/3})$ .

Set

$$R = n - S - 2S(\mathcal{T}_2) - 3S(\mathcal{T}_3). \quad (4.16)$$

Then

$$\begin{aligned} \frac{S - \mathbb{E}S}{\sqrt{\text{var } S(\mathcal{T}_2)}} &= -2 \frac{S(\mathcal{T}_2) - \mathbb{E}S(\mathcal{T}_2)}{\sqrt{\text{var } S(\mathcal{T}_2)}} - 3 \frac{S(\mathcal{T}_3) - \mathbb{E}S(\mathcal{T}_3)}{\sqrt{\text{var } S(\mathcal{T}_2)}} - \frac{R - \mathbb{E}R}{\sqrt{\text{var } S(\mathcal{T}_2)}} \\ &= -2Z_2 - 3Z_3 - Z_R. \end{aligned}$$

Now  $Z_2 \rightarrow^{\mathcal{D}} \mathcal{N}(0, 1)$  by the central limit theorem, and  $Z_3 \rightarrow^{\mathcal{D}} 0$  for  $p = o(n^{-1})$ , since  $\text{var } S(\mathcal{T}_2) \sim \frac{1}{2}n^2p$  and  $\text{var } S(\mathcal{T}_3) \sim \frac{1}{2}n^3p^2$ . Finally,  $Z_R \rightarrow^{\mathcal{D}} 0$ , since  $R \geq 0$  and, from (4.6) and (4.16),

$$\begin{aligned} \mathbb{E}R &= n - n(1-p)^{n-1} - 2 \binom{n}{2} p(1-p)^{2n-4} - 3 \binom{n}{3} \frac{3!}{2} p^2(1-p)^{3n-8} \\ &= O(n^4p^3) = O(1); \end{aligned}$$

to conclude the proof, note that

$$\text{var } S \sim 2n^2p \sim 4 \text{var } S(\mathcal{T}_2). \quad \blacksquare$$

*Remark.* This argument would also yield  $d_1(\mathcal{L}(W), \mathcal{N}(0, 1)) = O((\text{var } S)^{-1/2})$  for  $n^3p^2 = O(1)$  and for  $1 = O(np)$ , but gives a less precise estimate otherwise. The more refined treatment of Kordecki [16], also using Stein's method, yields the correct  $O((\text{var } S)^{-1/2})$  convergence rate throughout the range, using the traditional measure (1.1).

## REFERENCES

1. A. D. BARBOUR, Poisson convergence and random graphs, *Math. Proc. Cambridge Philos. Soc.* **92** (1982), 349–359.
2. A. D. BARBOUR AND G. K. EAGLESON, Multiple comparisons and sums of dissociated random variables, *Adv. Appl. Probab.* **17** (1985), 147–162.
3. A. D. BARBOUR AND G. K. EAGLESON, Random association of symmetric arrays, *Stochastic Anal. Appl.* **4** (1986), 239–281.
4. A. D. BARBOUR AND P. HALL, On bounds to the rate of convergence in the central limit theorem, *Bull. London Math. Soc.* **17** (1985), 151–156.
5. P. BILLINGSLEY, "Probability and measure," 2nd ed., Wiley, New York, 1986.
6. L. H. Y. CHEN, Poisson approximation for dependent trials, *Ann. Probab.* **3** (1975), 534–545.

7. L. H. Y. CHEN, Two central limit problems for dependent random variables, *Z. Wahrsch. Verw. Gebiete* **43** (1978), 223–243.
8. L. H. Y. CHEN, The rate of convergence in a central limit theorem for dependent random variables with arbitrary index set, *Ann. Probab.*, in press (1987).
9. P. ERDŐS AND A. RÉNYI, On the evolution of random graphs, *Publ. Math. Inst. Hungar. Acad. Sci.* **5** (1960), 17–61.
10. W. Hoeffding, A class of statistics with asymptotically normal distribution, *Ann. Math. Stat.* **19** (1948), 293–325.
11. W. Hoeffding, A non-parametric test of independence, *Ann. Math. Stat.* **19** (1948), 546–557.
12. S. JANSON, Poisson convergence and Poisson processes with applications to random graphs, *Stochastic Process. Appl.*, in press (1987).
13. S. JANSON, Normal convergence by higher semi-invariants with applications to sums of dependent random variables and random graphs, *Ann. Probab.*, in press (1987).
14. M. KAROŃSKI, “Balanced Subgraphs of Large Random Graphs,” Adam Mickiewicz, Univ. Press, Poznań.
15. M. KAROŃSKI AND A. RUCIŃSKI, Poisson convergence and semi-induced properties of random graphs, *Math. Proc. Cambridge Philos. Soc.* **101** (1987), 291–300.
16. W. KORDECKI, Normal approximation and isolated vertices in random graphs, *Proc. Random Graphs '87*, in press.
17. H. MAEHARA, On the number of induced subgraphs of a random graph, *Discrete Math.* **64** (1987), 309–312.
18. W. G. MCGINLEY AND R. SIBSON, Dissociated random variables, *Math. Proc. Cambridge Philos. Soc.* **77** (1975), 185–188.
19. K. NOWICKI, Asymptotic normality of graph statistics, *J. Stat. Plan. Inf.*, in press (1987).
20. K. NOWICKI, Asymptotic Poisson distributions with applications to statistical analysis of graphs, *Appl. Probab.*, in press (1987).
21. K. NOWICKI AND J. WIERMAN, Subgraph counts in random graphs using incomplete  $U$ -statistic methods, *Proc. 1st Jap. Conf. Graph Theory*, in press (1987).
22. A. RUCIŃSKI, When small subgraphs of a random graph are normally distributed, *Probab. Theory Related Fields*, in press (1987).
23. R. J. SERFLING, “Approximation Theorems of Mathematical Statistics,” Wiley, New York, 1980.
24. C. STEIN, A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, *Proc. VIth Berk. Symp. Math. Stat. Probab.* **2** (1970), 583–602.