RESEARCH ARTICLE

ScienceAsia 35 (2009): **203–210**

doi: 10.2306/scienceasia1513-1874.2009.35.203

On the normal approximation of the number of vertices in a random graph

Kritsana Neammanee*, **Angkana Suntadkarn**

Department of Mathematics, Faculty of Science, Chulalongkorn University, Phyathai Road, Patumwan, Bangkok 10330, Thailand

*Corresponding author, e-mail: kritsana.n@chula.ac.th

Received 22 Sep 2008
Accepted 9 May 2009

ABSTRACT: In this paper, we use Stein-Chen's method to give a uniform bound on the normal approximation of the number of vertices of a fixed degree in a random graph. This work corrects our results published previously.

KEYWORDS: Stein-Chen method, uniform bound, vertex

INTRODUCTION AND MAIN RESULTS

A random graph is a collection of points or vertices with lines or degrees connecting pairs of them at random. The study of random graphs has a long history. A systematic study of random graphs began with the influential work of Erdös and Rényi[1–3] and it has since developed into a significant area of study in modern discrete mathematics. There are many applications of random graphs (see Refs. 4–7).

Let $\mathbb{G}(n,p)$ be a random graph on n labelled vertices $\{1,2,\dots,n\}$ with the edges added randomly such that each of the $\binom{n}{2}$ possible edges exists with probability p, $0<p<1$. In this study, the assumption has been made that the presence of an edge between two vertices is independent of the others.

Let $G=(V(G),E(G))$ be a graph where $V(G)$ and $E(G)$ are the sets of vertices and edges of G, respectively. The degree of a vertex v in graph G, denoted by $\deg(v)$, is the number of edges incident to v, i.e., $\deg(v)=|\{w\in V(G)\mid vw\in E(G)\}|$. Any vertex of degree zero is called an isolated vertex.

Let S_n be the number of vertices of a fixed degree $d\geqslant 0$ in $\mathbb{G}(n,p)$. Then $S_n=Y_1+Y_2+\dots+Y_n$ where

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

for $i=1,2,\dots,n$. Note that $a_n:=E[Y_i]$ is given by

$$a_n=P(Y_i=1)=\binom{n-1}{d}p^d q^{n-1-d}$$

where $q=1-p$ and $E[S_n]=na_n$. Also[8],

$$\sigma_n^2:=\operatorname{Var}(S_n)=\frac{n}{n-1}\binom{n-1}{d}^2(d-(n-1)p)^2 \times p^{2d-1}(1-p)^{2(n-d)-3}+E[S_n]-\frac{(E[S_n])^2}{n}.$$

Stein introduced a new powerful technique for obtaining the rate of convergence to the standard normal distribution[9]. His approach was subsequently extended to cover the convergence to the Poisson distribution[10]. Stein's method was applied to random graphs by Barbour[11].

The Poisson convergence has since been widely taken up (see, e.g., Refs. 12–15). Barbour et al[16] proved that the distribution of S_n converges to the Poisson distribution with parameter $\lambda=na_n$ if either $np\to 0$ and $d\geqslant 2$, or np is bounded away from 0 and $(np)^{-\frac{1}{2}}|d-np|\to\infty$. Later, Suntadkarn and Neammanee[17] gave the rates of convergence in case of $d\geqslant 1$ and $p=1/n^\delta$ for some $\delta>0$. They showed that for $A\subset\{1,2,\dots,n\}$, $|P(S_n\in A)-\mathrm{Poi}_\lambda(A)|$, where

$$\mathrm{Poi}_\lambda(A)=\sum_{k\in A}\frac{\mathrm{e}^{-\lambda}\lambda^k}{k!},$$

is $O(n^{-(\delta-1)(d-1)})$ if $\delta>1$ and is $O(n^{-d(1-\delta)})$ for $0<\delta<1$.

For the normal approximation, Barbour et al[8] proved that the distribution function of

$$W_n:=\frac{S_n-na_n}{\sigma_n}$$

converges to the standard normal distribution function

$$\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}\mathrm{e}^{-t^2/2}\,\mathrm{d}t.$$

For $E[S_n] \to \infty$ they gave the rate of convergence expressed in terms of $E[S_n]$ with respect to the Wasserstein metric d_1 which is defined by $d_1(X,Y) = \sup\{|E[h(X)] - E[h(Y)]| : \sup_{x\in\mathbb{R}} |h(x)| + \sup_{x\in\mathbb{R}} |h^{'}(x)| \leqslant 1$ for all bounded test functions h with bounded derivative$\}$ for any random variables X and Y. They showed that

$$d_1(W_n, \mathcal{N}(0,1)) \leqslant \frac{C(d)}{\sqrt{E[S_n]}}$$

where $\mathcal{N}(0,1)$ is the standard normal random variable and $C(d)$ is a positive constant depending on d. In this article we consider the uniform distance

$$\delta_n := \sup_{z\in\mathbb{R}} |P(W_n \leqslant z) - \Phi(z)|.$$

Note that $\delta_n = O(\{d_1(W_n, \mathcal{N}(0,1))\}^{1/2})$, in general[18]. In Barbour et al[8] the authors explicitly restricted their considerations to smooth test functions but Raič[19] used the Lipschitz test function to modify the proof of Barbour et al[8] for non-smooth test functions at the cost of a boundedness condition.

In this paper, we use Stein's method and a Lipschitz test function to correct the work on normal approximation of Neammanee and Suntadkarn[20]. We give a uniform bound of the normal approximation of the number of vertices of a fixed degree in a random graph. The following theorem is our main result.

Theorem 1 *For $d \geqslant 0$, $0 < \beta < 1$, and a positive integer $r_0 \geqslant \beta/(1-\beta)$ there exists a positive constant $C(d, r_0)$ such that, for large n,*

$$\sup_{z\in\mathbb{R}} |P(W_n \leqslant z) - \Phi(z)| \leqslant \frac{C(d,r_0)(1+(np)^{r_0+1})}{\sigma_n^{\beta}} + \frac{C(d)(1+(np)^{r_0+1})\sqrt{n}}{\sigma_n^{r_0(1-\beta)+1}}.$$

Furthermore, if $\sigma^2 = O(n)$ then

$$\sup_{z\in\mathbb{R}} |P(W_n \leqslant z) - \Phi(z)| \leqslant \frac{C(d,r_0)(1+(np)^{r_0+1})}{\sigma_n^{\beta}}.$$

In the case of $np = O(1)$ we have $\sigma_n^2 = O(n)$. Hence, for $\beta > \frac{1}{2}$ and $r_0 \geqslant \beta/(1-\beta)$, there exists a constant $C(d, r_0)$ such that $\delta_n = C(d, r_0)/\sigma_n^{\beta}$. Here, and throughout the paper, $C(c_1, c_2, \ldots, c_k)$ stands for an absolute constant depending on $c_1, c_2, \ldots, c_k$.

In the next section we prove auxiliary results and in the final section we introduce the Stein's method for normal approximation which we use in the proof of main result in the last section.

AUXILIARY RESULTS

For each $i \in \{1, 2, \ldots, n\}$, let

$$X_i = \frac{Y_i - E[Y_i]}{\sigma_n}$$

and for any $\Lambda \subset \{1, 2, \ldots, n\}$ we define

$$Y_i^{(\Lambda)} = \begin{cases} 1, & \text{if vertex } i \text{ has degree } d \text{ in } G(n,p)-\{\Lambda\}, \\ 0, & \text{otherwise,} \end{cases}$$

where $G(n,p) - \{\Lambda\}$ is the random graph obtained from $G(n,p)$ by removing the vertices in Λ.

For $i, j = 1, 2, \ldots, n$, let

$$Z_{ij} = \begin{cases} Y_i/\sigma_n, & i = j, \\ (Y_j - Y_j^{(i)})/\sigma_n, & i \neq j, \end{cases}$$

$$Z_i = \sum_{l=1}^{n} Z_{il},$$

$$W^{(i)} = \sum_{\substack{l=1 \\ l\neq i}}^{n} \frac{1}{\sigma_n}(Y_l^{(i)} - E[Y_l^{(i)}]) - E[Z_i] = W_n - Z_i, \quad (1)$$

$$V_{ij} = \begin{cases} 0, & i = j, \\ \dfrac{1}{\sigma_n}\Big\{Y_j^{(i)} + \sum\limits_{\substack{l=1 \\ l\neq i,j}}^{n} (Y_l^{(i)} - Y_l^{(i,j)})\Big\} & ; i \neq j, \end{cases}$$

$$W_{ij} = \sum_{\substack{l=1 \\ l\neq i,j}}^{n} \frac{1}{\sigma_n}(Y_l^{(i,j)} - E[Y_l^{(i,j)}]) - E[V_{ij}] - E[Z_i] = W^{(i)} - V_{ij} \quad (2)$$

where $Y_j^{(i)} := Y_j^{(\{i\})}$ and $Y_l^{(i,j)} := Y_l^{(\{i,j\})}$. Note that $W^{(i)}$ is independent of X_i, and W_{ij} is independent of the pair (X_i, Z_{ij}) (see Ref. 8) and

$$W_n = \sum_{i=1}^{n} X_i.$$

The following propositions improve the results of Propositions 2.1–2.3 in Ref. 20 for the case of arbitrary p.

Proposition 1 *For large n, we have $\sigma_n^2 \geqslant \frac{1}{2} n a_n$.*

Proof: From Proposition 2.1 of Ref. 20 we know that $\sigma_n^2 \geqslant n a_n(1 - a_n)$. If we can show that $a_n \leqslant \frac{1}{2}$, then the proposition is proved. From the fact that $1 - p \leqslant \mathrm{e}^{-p}$, we have

$$a_n = \binom{n-1}{d} p^d q^{n-1-d} \leqslant \frac{\mathrm{e}^{(1+d)p}}{d!} \cdot \frac{(np)^d}{\mathrm{e}^{np}} \leqslant \frac{1}{2}$$

for large n. □

Proposition 2 *For* $i, j \in \{1, 2, \ldots, n\}$ *and* $r_1, r_2, r_3 \in \mathbb{N}$*, there exists a positive constant* $C(d, r_2, r_3)$ *such that for large* n*,*

$$E[|X_i^{r_1} Z_i^{r_2} V_{ij}^{r_3}|] \leqslant \frac{C(d, r_2, r_3) a_n (1 + (np)^{r_2+r_2-1})}{\sigma^{r_1+r_2+r_3}}.$$

Proof: We follow the proof of Proposition 2.2 of Ref. 20 to show that for r, r_1, $r_2, \ldots, r_m \in \mathbb{N}$ and for distinct $j_1, j_2, \ldots, j_m$ which are not equal to i we have

$$\begin{aligned} &E[|Y_{j_1} - Y_{j_1}^{(i)}|^{r_1} |Y_{j_2} - Y_{j_2}^{(i)}|^{r_2} \cdots |Y_{j_m} - Y_{j_m}^{(i)}|^{r_m}] \\ &\quad \leqslant \frac{p^{m-1} a_n}{(n-1)} \left[d + \frac{np}{q} \right] \leqslant C(d)(1 + np) \frac{p^{m-1} a_n}{n} \end{aligned} \tag{3}$$

and

$$\begin{aligned} &E\left[\left| \sum_{\substack{j=1 \\ j \neq i}}^{n} (Y_j - Y_j^{(i)}) \right|^r \right] \\ &= C(d)(1 + np) a_n \Big(1 + np + \cdots + (np)^{r-1} \Big) \\ &\leqslant C(d, r)(1 + (np)^{r-1}) a_n. \end{aligned} \tag{4}$$

Hence

$$\begin{aligned} E[|Z_i^r|] &= \frac{1}{\sigma_n^r} E\left[\left| (Y_i + \sum_{\substack{j=1 \\ j \neq i}}^{n} (Y_j - Y_j^{(i)}))^r \right| \right] \\ &\leqslant \frac{2^{r-1}}{\sigma_n^r} \left\{ E[|Y_i^r|] + E\left[\left| \sum_{\substack{j=1 \\ j \neq i}}^{n} (Y_j - Y_j^{(i)}) \right|^r \right] \right\} \\ &\leqslant \frac{2^{r-1}}{\sigma_n^r} \Big(E[Y_i] + C(d, r)(1 + (np)^{r-1}) a_n \Big) \\ &= C(d, r)(1 + (np)^{r-1}) \frac{a_n}{\sigma_n^r}. \end{aligned} \tag{5}$$

Again for distinct $l_1, l_2, \ldots, l_m$ which are not equal to i, j, we have

$$E\left[\left| \prod_{k=1}^{m} |Y_{l_k}^{(i)} - Y_{l_k}^{(i,j)}|^{r_k} \right| \right] \leqslant C(d)(1+np) \frac{p^{m-1} a_n}{n}$$

and

$$E\left[\left| \sum_{\substack{l=1 \\ l \neq i,j}}^{n} (Y_l^{(i)} - Y_l^{(i,j)})^r \right| \right] \leqslant C(d, r)(1+(np)^{r-1}) a_n.$$

From this and the fact that $E[Y_j^{(i)}] = P(Y_j^{(i)} = 1) \leqslant C(d) a_n$ (see Ref. 20) we have,

$$E[|V_{ij}^r|] \leqslant C(d, r)(1 + (np)^{r-1}) \frac{a_n}{\sigma_n^r}. \tag{6}$$

Since

$$|X_i| = \left| \frac{(Y_i - \mu)}{\sigma_n} \right| \leqslant \frac{1}{\sigma_n}, \tag{7}$$

we get

$$\begin{aligned} &E[|X_i^{r_1} Z_i^{r_2} V_{ij}^{r_3}|] \leqslant \frac{1}{\sigma_n^{r_1}} E[|Z_i^{r_2} V_{ij}^{r_3}|] \\ &\leqslant \frac{1}{\sigma_n^{r_1}} \Big\{ E[|Z_i|^{r_2+r_3}] \Big\}^{\frac{r_2}{r_2+r_3}} \Big\{ E[|V_{ij}|^{r_2+r_3}] \Big\}^{\frac{r_3}{r_2+r_3}} \\ &\leqslant C(d, r_2, r_3)(1 + (np)^{r_2+r_3-1}) \frac{a_n}{\sigma_n^{r_1+r_2+r_3}} \end{aligned}$$

from (5) and (6). □

Proposition 3 *For* $r_1 \in \mathbb{N} \cup \{0\}$, $r_2, r_3 \in \mathbb{N}$ *and* $i, j \in \{1, 2, \ldots, n\}$*, there exists a positive constant* $C(d, r_3)$ *such that*

$$E[|X_i^{r_1} Z_{ij}^{r_2} V_{ij}^{r_3}|] \leqslant C(d, r_3)(1+(np)^{r_3}) \frac{a_n}{n \sigma_n^{r_1+r_2+r_3}}.$$

Proof: From the proof of Proposition 2.3 of Ref. 20 we can see that

$$E[|(Y_j^{(i)})^{r_1} (Y_j - Y_j^{(i)})^{r_2}|] \leqslant C(d) a_n p$$

and

$$\begin{aligned} &E[|(Y_j - Y_j^{(i)})^{r_1} (Y_{l_1}^{(i)} - Y_{l_1}^{(i,j)})^{r_2} (Y_{l_2}^{(i)} - Y_{l_2}^{(i,j)})^{r_3} \\ &\quad \cdots (Y_{l_m}^{(i)} - Y_{l_m}^{(i,j)})^{r_{m+1}}|] \leqslant C(d)(1 + np) \frac{p^m a_n}{n}. \end{aligned}$$

Hence

$$\begin{aligned} &E[|X_i^{r_1} Z_{ij}^{r_2} V_{ij}^{r_3}|] \\ &\leqslant \frac{1}{\sigma_n^{r_1+r_2+r_3}} E\left[\left| (Y_j - Y_j^{(i)})^{r_2} \right.\right. \\ &\qquad \left.\left. \Big\{ Y_j^{(i)} + \sum_{\substack{l=1 \\ l \neq i,j}}^{n} (Y_l^{(i)} - Y_l^{(i,j)}) \Big\}^{r_3} \right| \right] \\ &\leqslant \frac{C(d, r_3)}{\sigma_n^{r_1+r_2+r_3}} \left\{ E[|(Y_j - Y_j^{(i)})^{r_2} (Y_j^{(i)})^{r_3}|] \right. \\ &\qquad \left. + E\left[\left| (Y_j - Y_j^{(i)})^{r_2} [\sum_{\substack{l=1 \\ l \neq i,j}}^{n} (Y_l^{(i)} - Y_l^{(i,j)})]^{r_3} \right| \right] \right\} \\ &\leqslant \frac{C(d, r_3)}{\sigma_n^{r_1+r_2+r_3}} \left(a_n p + (1 + (np)^{r_3}) \frac{a_n}{n} \right) \\ &\leqslant C(d, r_3)(1 + (np)^{r_3}) \frac{a_n}{n \sigma_n^{r_1+r_2+r_3}}. \end{aligned}$$

□

STEIN'S METHOD FOR NORMAL APPROXIMATION

Stein's method[9] relies on the elementary differential equation

$$f'(w) - wf(w) = h(w) - \mathcal{N}h \tag{8}$$

where $f : \mathbb{R} \to \mathbb{R}$ is a continuous and piecewise continuously differentiable function, h is a bounded test function with bounded derivative, and

$$\mathcal{N}h := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(t) \mathrm{e}^{-t^2/2} \, \mathrm{d}t.$$

In order to use (8) to obtain the bound of normal approximation, many authors[20–22], choose the test function $h = I_z$ where

$$I_z(w) = \begin{cases} 1, & w \leqslant z, \\ 0, & w > z. \end{cases}$$

Hence (8) becomes

$$f'(w) - wf(w) = I_z(w) - \Phi(z) \tag{9}$$

and the unique solution f_z of (9) is

$$f_z(w) = \begin{cases} \sqrt{2\pi} \mathrm{e}^{w^2/2} \Phi(w)[1 - \Phi(z)], & w \leqslant z, \\ \sqrt{2\pi} \mathrm{e}^{w^2/2} \Phi(z)[1 - \Phi(w)], & w > z. \end{cases} \tag{10}$$

Substituting any random variable W for w in (9), we get

$$E[f_z'(W) - Wf_z(W)] = P(W \leqslant z) - \Phi(z).$$

Hence, to bound $|P(W \leqslant z) - \Phi(z)|$, it suffices to bound $E[f_z'(W) - Wf_z(W)]$. But in our work, we choose the Lipschitz test function

$$I_{z,\varepsilon}(w) = \begin{cases} 1, & w < z - \varepsilon, \\ -\dfrac{w - z - \varepsilon}{2\varepsilon}, & z - \varepsilon \leqslant w < z + \varepsilon, \\ 0, & w \geqslant z + \varepsilon, \end{cases}$$

where $\varepsilon > 0$ is fixed. Observe that

$$I_{z,\varepsilon}(w) = \frac{1}{2\varepsilon} \int_{z-\varepsilon}^{z+\varepsilon} I_t(w) \, \mathrm{d}t$$

and

$$\mathcal{N}I_{z,\varepsilon} = \frac{1}{2\varepsilon} \int_{z-\varepsilon}^{z+\varepsilon} \mathcal{N}I_t \, \mathrm{d}t = \frac{1}{2\varepsilon} \int_{z-\varepsilon}^{z+\varepsilon} \Phi(t) \, \mathrm{d}t.$$

From (8) we have

$$E[I_{z,\varepsilon}(W) - \mathcal{N}I_{z,\varepsilon}] = E[f_{z,\varepsilon}'(W) - Wf_{z,\varepsilon}(W)]$$

where

$$f_{z,\varepsilon}(w) = \frac{1}{2\varepsilon} \int_{z-\varepsilon}^{z+\varepsilon} f_t(w) \, \mathrm{d}t$$

and

$$f_{z,\varepsilon}'(w) = \frac{1}{2\varepsilon} \int_{z-\varepsilon}^{z+\varepsilon} f_t'(w) \, \mathrm{d}t.$$

Note that[19]

$$\sup_{\substack{x \in \mathbb{R} \\ \varepsilon > 0}} |f_{z,\varepsilon}(x)| \leqslant \frac{\sqrt{2\pi}}{4},$$

$$\sup_{\substack{x,y \in \mathbb{R} \\ x \neq y \\ \varepsilon > 0}} \left| \frac{f_{z,\varepsilon}(x) - f_{z,\varepsilon}(y)}{x - y} \right| \leqslant 1.$$

Raič[19] used a test function $I_{z,\varepsilon}$ to give a bound in the following theorem. However, his work cannot be applied to a random graph since our V_{ij} is not bounded.

Theorem 2 (Raič[19]) *Let W_n be a decomposed random variable defined by*

$$W_n = \sum_{i \in I} X_i,$$

$E[X_i] = 0$, $i \in I$, $E[W_n^2] = 1$, $W_n = W^{(i)} + Z_i$, where $W^{(i)}$ is independent of X_i, and

$$Z_i = \sum_{j \in K_i} Z_{ij}, \qquad i \in I, \quad K_i \subset I$$

$$W^{(i)} = W_{ij} + V_{ij}, \qquad i \in I, \quad j \in K_i,$$

where W_{ij} is independent of the pair (X_i, Z_{ij}). Suppose that $|X_i| \leqslant A_i$, $|Z_{ik}| \leqslant B_{ik}$, $|V_{ik}| \leqslant C_{ik}$, $|Z_i + V_{ik}| \leqslant C'_{ik}$ for some constants A_i, B_{ik}, C_{ik}, and C'_{ik}. Then

$$\sup_{z \in \mathbb{R}} |P(W_n \leqslant z) - \Phi(z)| \leqslant 13.7 \sum_{i \in I} A_i B_i^2 + \sum_{i \in I} \sum_{k \in K_i} A_i B_{ik} (6.8 C_{ik} + 9.3 C'_{ik})$$

where $B_i := \sum_{k \in K_i} B_{ik}$.

To prove this theorem, Raič[19] showed that for all $\varepsilon > 0$,

$$|P(W_n \leqslant z) - \Phi(z)| \leqslant A_1 + A_2 + A_3 + B_1 + B_2 + B_3 + \frac{\varepsilon}{\sqrt{2\pi}}, \tag{11}$$

where

$$A_1 := \sum_{i\in I} E[|g_{z,\varepsilon}(W^{(i)} + \theta_1 Z_i) - g_{z,\varepsilon}(W^{(i)})| \\ |X_i Z_i|],$$

$$A_2 := \sum_{i\in I}\sum_{k\in K_i} E[|g_{z,\varepsilon}(W^{(i)}) - g_{z,\varepsilon}(W_{ik})| \\ |X_i Z_{ik}|],$$

$$A_3 := \sum_{i\in I}\sum_{k\in K_i} E[|g_{z,\varepsilon}(W_n) - g_{z,\varepsilon}(W_{ik})| \\ E[|X_i Z_{ik}|]],$$

$$B_1 := \sum_{i\in I} E[|I_{z,\varepsilon}(W^{(i)}) - I_{z,\varepsilon}(W^{(i)} + \theta_1 Z_i)| \\ |X_i Z_i|],$$

$$B_2 := \sum_{i\in I}\sum_{k\in K_i} E[|I_{z,\varepsilon}(W_{ij}) - I_{z,\varepsilon}(W_{ik} + V_{ik})| \\ |X_i Z_{ik}|],$$

$$B_3 := \sum_{i\in I}\sum_{k\in K_i} E[|I_{z,\varepsilon}(W_n) - I_{z,\varepsilon}(W_{ik})| \\ E[|X_i Z_{ik}|]],$$

in which $g_{z,\varepsilon}(x) = f_{z,\varepsilon}(x)x$, and θ_1 is a random variable whose value is in $[0,1]$. Then he used the boundness of X_i, Z_{ik}, V_{ik} and $Z_i + V_{ik}$ to complete his result. In our work, we also prove our result by using (11).

PROOF OF MAIN RESULT

In this section, we give the proof of Theorem 1. Using the argument of Raič[19], we can show that (11) holds.

Step 1 We will show that

$$A_1 + A_2 + A_3 \leqslant \frac{C(d)(1+(np)^2)}{\sigma_n^\beta} + \frac{C(d)(1+(np)^{r_0+1})\sqrt{n}}{\sigma_n^{r_0(1-\beta)+1}}.$$

Following the argument of equations (5.17) and (5.18) of Raič[19], we can show that

$$A_1 \leqslant C\sum_{i=1}^n \Big(E[|X_i|Z_i^2] + E[|W_n X_i|Z_i^2]\Big),$$

$$A_2 \leqslant C\sum_{i=1}^n\sum_{\substack{j=1\\ j\neq i}}^n \Big(E[|X_i Z_{ij} V_{ij}|] \\ + \frac{\sqrt{2\pi}}{2} E[|W_{ij} + V_{ij}||X_i Z_{ij} V_{ij}|]\Big)$$

$$A_3 \leqslant C\sum_{i=1}^n\sum_{\substack{j=1\\ j\neq i}}^n \Big(E[|Z_i + V_{ij}|]E[|X_i Z_{ij}|] \\ + E[|Z_i + V_{ij}||W_n|]E[|X_i Z_{ij}|]\Big).$$

From (1), (5), (7), and the facts that W_{ij} is independent of (X_i, Z_{ij}) and that $\sigma_n^2 E[X_i^2] = E[(Y_i - a_n)^2] \leqslant a_n$, we have

$$\begin{aligned}
E[|W_n X_i|Z_i^2] &= E[|(W^{(i)} + Z_i)X_i|Z_i^2] \\
&\leqslant E[|W^{(i)}X_i|Z_i^2] + E[|X_i Z_i^3|] \\
&\leqslant \Big\{E[(W^{(i)})^2 X_i^2]\Big\}^{\frac12}\Big\{E[Z_i^4]\Big\}^{\frac12} \\
&\quad + \frac{C(d)(1+(np)^2)a_n}{\sigma_n^4} \\
&= \Big\{E[(W^{(i)})^2]E[X_i^2]\Big\}^{\frac12}\Big\{E[Z_i^4]\Big\}^{\frac12} \\
&\quad + \frac{C(d)(1+(np)^2)a_n}{\sigma_n^4} \\
&\leqslant C\Big\{E[W_n^2 + Z_i^2]E[X_i^2]\Big\}^{\frac12}\Big\{E[Z_i^4]\Big\}^{\frac12} \\
&\quad + \frac{C(d)(1+(np)^2)a_n}{\sigma_n^4} \\
&\leqslant \frac{C(d)(1+(np)^2)a_n}{\sigma_n^3}.
\end{aligned}$$

Hence, by (5) and (7),

$$A_1 \leqslant \frac{C(d)(1+(np)^2)}{\sigma_n}. \tag{12}$$

From (1) and (5), we have $E[|W_{ij}|] \leqslant E[|W_n|] + E[|V_{ij}|] + E[|Z_i|] \leqslant C(d)$. Therefore,

$$\begin{aligned}
&E[|W_{ij} X_i Z_{ij} V_{ij}|] \\
&\leqslant E[|W_{ij} X_i Z_{ij} V_{ij} I(|V_{ij}| \leqslant 1/\sigma_n^\beta)|] \\
&\quad + E[|W_{ij} X_i Z_{ij} V_{ij} I(|V_{ij}| > 1/\sigma_n^\beta)|] \\
&\leqslant \frac{1}{\sigma_n^\beta} E[|W_{ij}|]E[|X_i Z_{ij}|] + \sigma_n^{r_0\beta} E[|W_{ij} X_i Z_{ij} V_{ij}^{r_0+1}|] \\
&\leqslant \frac{C(d)a_n}{n\sigma_n^{2+\beta}} \\
&\quad + \sigma_n^{r_0\beta}\Big\{E[W_{ij}^2]E[X_i^2]\Big\}^{\frac12}\Big\{E[|Z_{ij}^2 V_{ij}^{2(r_0+1)}|]\Big\}^{\frac12} \\
&\leqslant \frac{C(d)a_n}{n\sigma_n^{2+\beta}} + \frac{C(d)(1+(np)^{r_0+1})a_n}{\sqrt{n}\sigma_n^{r_0(1-\beta)+3}}.
\end{aligned}$$

Hence,

$$A_2 \leqslant \frac{C(d)(1+(np)^2)}{\sigma_n^\beta} + \frac{C(d)(1+(np)^{r_0+1})\sqrt{n}}{\sigma_n^{r_0(1-\beta)+1}}. \tag{13}$$

From (3), (5), (6), (7), and $E[W_n^2] = 1$,

$$A_3 \leqslant \frac{C(d)(1 + (np)^2)}{\sigma_n}. \tag{14}$$

From (12–14),

$$A_1 + A_2 + A_3 \leqslant \frac{C(d)(1 + (np)^2)}{\sigma_n^\beta} + \frac{C(d)(1 + (np)^{r_0+1})\sqrt{n}}{\sigma_n^{r_0(1-\beta)+1}}.$$

In the remainder of the proof we let $\varepsilon > 0$.

Step 2 We will show that

$$B_1 \leqslant \frac{C(d, r_0)(1 + (np)^{r_0+1})}{\varepsilon\sigma_n}\left(\delta_n + \frac{1}{\sigma_n^\beta}\right) + \frac{C(d)(1 + np)}{\sigma_n}$$

where $\delta_n = \sup_{z\in\mathbb{R}} |P(W_n \leqslant z) - \Phi(z)|$. From the fact that[19],

$$I_{z,\varepsilon}(x) - I_{z,\varepsilon}(y) = \frac{y - x}{2\varepsilon}\int_0^1 I_{[z-\varepsilon\leqslant(1-\theta)x+\theta y\leqslant z+\varepsilon]}\,\mathrm{d}\theta \tag{15}$$

where

$$I_A(w) = \begin{cases} 1, & w \in A \\ 0, & \text{otherwise} \end{cases}$$

we have

$$\begin{aligned} B_1 &= \sum_{i=1}^n E[|I_{z,\varepsilon}(W^{(i)}) - I_{z,\varepsilon}(W^{(i)} + \theta_1 Z_i)||X_i Z_i|] \\ &\leqslant \frac{1}{2\varepsilon}\sum_{i=1}^n E\left[\left|\int_0^1 I_{[\Xi_i(\theta)]}\,\mathrm{d}\theta\right| |X_i Z_i^2|\right] \\ &\leqslant \frac{1}{2\varepsilon}\sum_{i=1}^n E\left[|X_i Z_i^2| \left|\int_0^1 I_{[\Xi_i(\theta), |V_{ij}|>\frac{1}{\sigma_n^\beta}]}\,\mathrm{d}\theta\right|\right] \\ &\quad + \frac{1}{2\varepsilon}\sum_{i=1}^n E\left[|X_i Z_i^2| \left|\int_0^1 I_{[\Xi_i(\theta), |V_{ij}|\leqslant\frac{1}{\sigma_n^\beta}]}\,\mathrm{d}\theta\right|\right] \end{aligned}$$

where $\Xi_i(\theta)$ denotes $z - \varepsilon \leqslant W^{(i)} + \theta\theta_1 Z_i \leqslant z + \varepsilon$, and we denote the two terms on the last right-hand side by B_{11} and B_{12}, respectively. By Proposition 2,

$$\begin{aligned} B_{11} &\leqslant \frac{\sigma_n^{r_0\beta}}{2\varepsilon}\sum_{i=1}^n E\left[|X_i Z_i^2 V_{ij}^{r_0}| \times \left|\int_0^1 I_{[\Xi_i(\theta), |V_{ij}|>\frac{1}{\sigma_n^\beta}]}\,\mathrm{d}\theta\right|\right] \\ &\leqslant \frac{\sigma_n^{r_0\beta}}{2\varepsilon}\sum_{i=1}^n E[|X_i Z_i^2 V_{ij}^{r_0}|] \\ &\leqslant \frac{C(d, r_0)(1 + (np)^{r_0+1}) n a_n}{\varepsilon\sigma_n^{r_0(1-\beta)+3}} \\ &\leqslant \frac{C(d, r_0)(1 + (np)^{r_0+1})}{\varepsilon\sigma_n^{\beta+1}}. \end{aligned} \tag{16}$$

To bound B_{12} we use the concentration inequality

$$P(a \leqslant W_n \leqslant b) \leqslant 2\delta_n + \frac{b - a}{\sqrt{2\pi}} \tag{17}$$

(see Raič[19]). Note that from (1), (2), (17) and Chebyshev's inequality we have

$$\begin{aligned} &P(a \leqslant W_{ij} \leqslant b) \\ &= P(a \leqslant W_n - (Z_i + V_{ij}) \leqslant b,\ |Z_i + V_{ij}| \leqslant \frac{1}{\sigma_n^\beta}) \\ &+ P(a \leqslant W_n - (Z_i + V_{ij}) \leqslant b,\ |Z_i + V_{ij}| > \frac{1}{\sigma_n^\beta}) \\ &\leqslant P(a - \frac{1}{\sigma_n^\beta} \leqslant W_n \leqslant b + \frac{1}{\sigma_n^\beta}) \\ &\quad + P(|Z_i + V_{ij}| > \frac{1}{\sigma_n^\beta}) \\ &\leqslant 2\delta_n + \frac{b - a}{\sqrt{2\pi}} + \frac{2}{\sqrt{2\pi}\sigma_n^\beta} + E|Z_i + V_{ij}|^{r_0}\sigma_n^{r_0\beta} \\ &\leqslant 2\delta_n + \frac{b - a}{\sqrt{2\pi}} + \frac{2}{\sqrt{2\pi}\sigma_n^\beta} \\ &\quad + C(d, r_0)(1 + (np)^{r_0-1})\frac{a_n}{\sigma_n^\beta} \\ &\leqslant 2\delta_n + \frac{b - a}{\sqrt{2\pi}} + \frac{C(d, r_0)(1 + (np)^{r_0-1})}{\sigma_n^\beta}. \end{aligned} \tag{18}$$

From (5), (7), (18), and the fact that W_{ij} is independent of (X_i, Z_{ij}), we have

$$\begin{aligned} B_{12} &= \frac{1}{2\varepsilon}\sum_{i=1}^n E\Bigg[|X_i Z_i^2| \\ &\times \left|\int_0^1 I_{[z-\varepsilon\leqslant W_{ij}+V_{ij}+\theta\theta_1 Z_i\leqslant z+\varepsilon, |Z_i|\leqslant\frac{1}{\sigma_n^\beta}, |V_{ij}|\leqslant\frac{1}{\sigma_n^\beta}]}\,\mathrm{d}\theta\right. \\ &+ \left.\int_0^1 I_{[z-\varepsilon\leqslant W_{ij}+V_{ij}+\theta\theta_1 Z_i\leqslant z+\varepsilon, |Z_i|>\frac{1}{\sigma_n^\beta}, |V_{ij}|\leqslant\frac{1}{\sigma_n^\beta}]}\,\mathrm{d}\theta\right|\Bigg] \end{aligned}$$

$$
\begin{aligned}
&\leqslant \frac{1}{2\varepsilon}\sum_{i=1}^{n} E\left[|X_i Z_i^2|\int_0^1 I_{[z-\varepsilon-\frac{2}{\sigma_n^\beta}\leqslant W_{ij}\leqslant z+\varepsilon+\frac{2}{\sigma_n^\beta}]}\,\mathrm{d}\theta\right]\\
&\quad+\frac{\sigma_n^{r_0\beta}}{2\varepsilon}\sum_{i=1}^{n}E[|X_iZ_i^{r_0+2}|]\\
&\leqslant \frac{1}{2\varepsilon}\sum_{i=1}^{n} E\Bigg[|X_i Z_i^2|\\
&\quad\times P\Big(z-\varepsilon-\frac{2}{\sigma_n^\beta}\leqslant W_{ij}\leqslant z+\varepsilon+\frac{2}{\sigma_n^\beta}\Big)\Bigg]\\
&\quad+\frac{C(d,r_0)(1+(np)^{r_0+1})\sigma_n^{r_0\beta}na_n}{\varepsilon\sigma_n^{r_0+3}}\\
&\leqslant \frac{C(d)(1+np)}{\varepsilon\sigma_n}\Big[\delta_n+\varepsilon+\frac{C(d,r_0)(1+(np)^{r_0})}{\sigma_n^\beta}\Big]\\
&\leqslant \frac{C(d,r_0)(1+(np)^{r_0+1})}{\varepsilon\sigma_n}\Big(\delta_n+\frac{1}{\sigma_n^\beta}\Big)\\
&\quad+\frac{C(d)(1+np)}{\sigma_n}.
\end{aligned}
\tag{19}
$$

From (16) and (19) we have

$$
\begin{aligned}
B_1 &\leqslant \frac{C(d,r_0)(1+(np)^{r_0+1})}{\varepsilon\sigma_n}\Big(\delta_n+\frac{1}{\sigma_n^\beta}\Big)\\
&\quad+\frac{C(d)(1+np)}{\sigma_n}.
\end{aligned}
$$

Step 3 We will show that

$$
\begin{aligned}
B_2 &\leqslant \frac{C(d,r_0)(1+(np)^{r_0+1})}{\varepsilon\sigma_n^\beta}\Big(\delta_n+\frac{1}{\sigma_n^\beta}\Big)\\
&\quad+\frac{C(d)(1+np)}{\sigma_n^\beta}.
\end{aligned}
$$

By using (15), Proposition 1, Proposition 3 and (18) we have

$$
\begin{aligned}
B_2 &= \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E[|I_{z,\varepsilon}(W_{ij})-I_{z,\varepsilon}(W_{ij}+V_{ij})||X_iZ_{ij}|]\\
&\leqslant \frac{1}{2\varepsilon}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E\Bigg[|X_iZ_{ij}V_{ij}|\\
&\quad\times\int_0^1 I_{[z-\varepsilon\leqslant W_{ij}+\theta V_{ij}\leqslant z+\varepsilon]}\,\mathrm{d}\theta\Bigg]\\
&\leqslant \frac{1}{2\varepsilon}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E\Bigg[|X_iZ_{ij}V_{ij}|\\
&\quad\times\int_0^1 I_{[z-\varepsilon\leqslant W_{ij}+\theta V_{ij}\leqslant z+\varepsilon,|V_{ij}|>\frac{1}{\sigma_n^\beta}]}\,\mathrm{d}\theta\Bigg]\\
&\quad+\frac{1}{2\varepsilon}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E\Bigg[X_iZ_{ij}V_{ij}\\
&\quad\times\int_0^1 I_{[z-\varepsilon\leqslant W_{ij}+\theta V_{ij}\leqslant z+\varepsilon,|V_{ij}|\leqslant\frac{1}{\sigma_n^\beta}]}\,\mathrm{d}\theta\Bigg]\\
&\leqslant \frac{\sigma_n^{r_0\beta}}{2\varepsilon}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E[|X_iZ_{ij}V_{ij}^{r_0+1}|]\\
&\quad+\frac{1}{2\varepsilon\sigma_n^\beta}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E\Bigg[X_iZ_{ij}\\
&\quad\times\int_0^1 I_{[z-\varepsilon-\frac{1}{\sigma_n^\beta}\leqslant W_{ij}\leqslant z+\varepsilon+\frac{1}{\sigma_n^\beta}]}\,\mathrm{d}\theta\Bigg]\\
&\leqslant \frac{C(d,r_0)(1+(np)^{r_0+1})}{\varepsilon\sigma_n^{\beta+1}}\\
&\quad+\frac{1}{2\varepsilon\sigma_n^\beta}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E\Bigg[|X_iZ_{ij}|\\
&\quad\times P\Big(z-\varepsilon-\frac{1}{\sigma_n^\beta}\leqslant W_{ij}\leqslant z+\varepsilon+\frac{1}{\sigma_n^\beta}\Big)\Bigg]\\
&\leqslant \frac{C(d)(1+np)}{\varepsilon\sigma_n^\beta}\Big[\delta_n+\varepsilon+\frac{C(d,r_0)(1+(np)^{r_0})}{\sigma_n^\beta}\Big]\\
&\leqslant \frac{C(d,r_0)(1+(np)^{r_0+1})}{\varepsilon\sigma_n^\beta}\Big(\delta_n+\frac{1}{\sigma_n^\beta}\Big)\\
&\quad+\frac{C(d)(1+np)}{\sigma_n^\beta}.
\end{aligned}
$$

Step 4 We will show that

$$
B_3 \leqslant \frac{C(d)(1+np)}{\varepsilon\sigma_n^\beta}\Big(\delta_n+\varepsilon+\frac{C(d,r_0)}{\sigma_n^\beta}\Big).
$$

From Proposition 1 we have

$$
\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E[|X_iZ_{ij}|]\leqslant C(d)(1+np). \tag{20}
$$

By (5), (6), (15) and (17) we have

$$
\begin{aligned}
&E[|I_{z,\varepsilon}(W_n)-I_{z,\varepsilon}(W_{ij})|]\\
&\leqslant \frac{1}{2\varepsilon}E\left[|Z_i+V_{ij}|\int_0^1 I_{[z-\varepsilon\leqslant W_n-\theta(Z_i+V_{ij})\leqslant z+\varepsilon]}\,\mathrm{d}\theta\right]\\
&\leqslant \frac{1}{2\varepsilon}E\Bigg[|Z_i+V_{ij}|\\
&\quad\times\int_0^1 I_{[z-\varepsilon\leqslant W_n-\theta(Z_i+V_{ij})\leqslant z+\varepsilon,|Z_i+V_{ij}|\leqslant\frac{1}{\sigma_n^\beta}]}\,\mathrm{d}\theta\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{2\varepsilon}E\Bigg[|Z_i+V_{ij}| \\
&\quad\times\int_0^1 I_{[z-\varepsilon\leqslant W_n-\theta(Z_i+V_{ij})\leqslant z+\varepsilon,|Z_i+V_{ij}|>\frac{1}{\sigma_n^\beta}]}\,\mathrm{d}\theta\Bigg] \\
&\leqslant\frac{1}{2\varepsilon\sigma_n^\beta}E\Bigg[\int_0^1 I_{[z-\varepsilon-\frac{1}{\sigma_n^\beta}\leqslant W_n\leqslant z+\varepsilon+\frac{1}{\sigma_n^\beta}]}\,\mathrm{d}\theta\Bigg] \\
&\quad+\frac{\sigma_n^{r_0\beta}}{2\varepsilon}E[|Z_i+V_{ij}|^{r_0+1}] \\
&=\frac{1}{2\varepsilon\sigma_n^\beta}P\Big(z-\varepsilon-\frac{1}{\sigma_n^\beta}\leqslant W_n\leqslant z+\varepsilon+\frac{1}{\sigma_n^\beta}\Big) \\
&\quad+\frac{C(d,r_0)}{\varepsilon\sigma_n^{\beta+1}} \\
&\leqslant\frac{C}{\varepsilon\sigma_n^\beta}\Big(\delta_n+\varepsilon+\frac{1}{\sigma_n^\beta}\Big)+\frac{C(d,r_0)}{\varepsilon\sigma_n^{\beta+1}} \\
&\leqslant\frac{C}{\varepsilon\sigma_n^\beta}\Big(\delta_n+\varepsilon+\frac{C(d,r_0)}{\sigma_n^\beta}\Big).
\end{aligned}
$$

From this fact, and (20) we have

$$
B_3\leqslant\frac{C(d)(1+np)}{\varepsilon\sigma_n^\beta}\Big(\delta_n+\varepsilon+\frac{C(d,r_0)}{\sigma_n^\beta}\Big).
$$

We now deduce our main result. From (11) and Steps 1–4 we have

$$
\begin{aligned}
\delta_n&=\sup_{z\in\mathbb{R}}|P(W_n\leqslant z)-\Phi(z)| \\
&\leqslant\frac{C(d,r_0)(1+(np)^{r_0+1})}{\varepsilon\sigma_n^\beta}\Big(\delta_n+\frac{1}{\sigma_n^\beta}\Big) \\
&\qquad+\frac{C(d)(1+(np)^2)}{\sigma_n^\beta}+\frac{C}{\sigma_n^\beta}+\frac{\varepsilon}{\sqrt{2\pi}}.
\end{aligned}
$$

for all $\varepsilon>0$. Hence we can choose $\varepsilon=C(d,r_0)(1+(np)^{r_0+1})/2\sigma_n^\beta$ such that for large n,

$$
B_1+B_2+B_3\leqslant\frac{C(d,r_0)(1+(np)^{r_0+1})}{\sigma_n^\beta}.
$$

This fact and Step 1 complete the proof.

Acknowledgements: The authors would like to thank the referee for many helpful remarks which have led us to improve the presentation and the Thailand Research Fund for financial support (RGJ2550).

REFERENCES

1. Erdös P, Rény A (1959) On random graphs. I. *Publ Math Debrecen* **6**, 290–7.
2. Erdös P, Rény A (1960) On the evolution of random graphs. *Publ Math Inst Hung Acad Sci* **5**, 17–61.
3. Erdös P, Rény A (1961) On the strength of connectedness of a random graph. *Acta Math Hung* **12**, 261–7.
4. Foster CC, Rapoport A, Orwant C (1963) A study of large sociogram: Elimination of free parameters. *Behav Sci* **8**, 56–65.
5. Heckatorn DD (1997) Respondent-driven sampling: A new approach to the study of hidden population. *Soc Probl* **44**, 174–99.
6. Bernard HR, Kilworth P, Evans M, McCarty C, Selley G (1998) Studying social relations crossculturally. *Ethnology* **2**, 155–79.
7. Fararo TJ, Sunshine M (1964) *A study of a biased friendship network*, Syracuse Univ Press.
8. Barbour AD, Karoński M, Rucińsk A (1989) A central limit theorem for decomposable random variables with application to random graphs. *J Combin Theor B* **47**, 125–45.
9. Stein CM (1972) A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In: Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability, vol 3, 583-602.
10. Chen LHY (1975) Poisson approximation for dependent trials. *Ann Probab* **3**, 534–45.
11. Barbour AD (1982) Poisson convergence and random graphs. *Math Proc Camb Phil Soc* **92**, 349–59.
12. Karoński M (1984) *Balanced subgraphs of large random graphs*, Adam Mickiewicz, Univ. Press, Poznan.
13. Karoński M, Ruciński A (1987) Poisson convergence and semi-induced properties random graphs. *Math Proc Camb Phil Soc* **101**, 291–300.
14. Nowicki K (1988) Asymtotic Poisson distributions with applications to statistical analysis of graphs. *Adv Appl Probab* **20**, 315–30.
15. Janson S (1987) Poisson convergence and Poisson processes with applications to random graphs. *Stoch Process Appl* **26**, 1–30.
16. Barbour AD, Holst L, Janson S (1992) *Poisson Approximation*, Clarendon Press, Oxford.
17. Neammanee K, Santadkarn A (2008) Poisson approximation of the number of vertices in a random graph. *Int Math Forum* **3**, 331–40.
18. Barbour AD, Hall P (1985) On bounds to the rate of convergence in the central limit theorem. *Bull Lond Math Soc* **17**, 151–6.
19. Raič M (2003) Normal approximation by Stein's method. In: Proceedings of the 7th Young Statisticians Meeting, pp 71–97.
20. Neammanee K, Suntadkarn A (2007) A bound on normal approximation of number of vertices of a fixed degree in a random graph. *Int J Pure Appl Math* **2**, 241–55.
21. Chen LHY, Shao QM (2001) A non-uniform Berry-Esseen bound via Stein's method. *Probab Theor Relat Field* **120**, 236–54.
22. Chen LHY, Shao Q-M (2005) Stein's method for normal approximation. In: Barbour AD, Chen LHY (eds) *An Introduction to Stein's Method*, World Scientific, pp 1–59.