University of Ljubljana
Institute of Mathematics, Physics and Mechanics
Department of Mathematics
Jadranska 19, 1000 Ljubljana, Slovenia

Preprint series, Vol. 40 (2002), 852

A MULTIVARIATE CLT FOR DECOMPOSABLE RANDOM VECTORS WITH FINITE SECOND MOMENTS

Martin Raič

ISSN 1318-4865

November 29, 2002

Ljubljana, November 29, 2002

A multivariate CLT for decomposable random vectors with finite second moments

Martin Raič

Department of Mathematics
Faculty of Mathematics and Physics
University of Ljubljana, Slovenia

November 26, 2002

Abstract

We derive a generalization of the result of Barbour et al. [4] to random vectors with finite second moments. This allows us to derive Lindeberg–Feller type theorems for sums of random vectors with certain dependency structures. We apply the main theorem to the study of three problems: local dependency, random graph degree statistics and linear rank statistics.

Key words: Stein's method; Multivariate Central Limit Theorems; Random Graphs; Dependency Graph; Linear Rank Statistics

1 Introduction

Stein's method has turned out to be an efficient tool for estimating the error in the CLT under various dependency structures: see Stein [22], Bolthausen [5], Barbour et al. [4], Goldstein and Rinott [13], Goldstein and Reinert [12] and references therein. Excellent results have been obtained for numerous problems, such as local dependence, random graphs, simple random sampling, sums of nonlinear functions and many others.

In most applications of Stein's method to normal approximation, the error in the CLT is expressed in terms of third moments. The aim of this paper is to estimate it in terms of quantities which only require finite second moments. The first result of this kind seems to be due to Katz [16] for i. i. d. random variables, and to Petrov [18] for independent but not necessarily identically distributed random variables. The following variant of their result appears in Chen and Shao [10] (see also Feller [11]):

Theorem 1.1. *Let $X_1, \dots X_n$ be independent random variables with zero means and finite variances. Define $W = X_1 + \dots + X_n$ and assume that $\mathrm{var}(W) = 1$. Then for all $x \in \mathbb{R}$,*

$$\left|\mathbf{P}[W \leq x] - \Phi(x)\right| \leq 4.1 \sum_{i=1}^{n} \mathbf{E}\, X_i^2 \min\{1, |X_i|\} \tag{1.1}$$

where Φ denotes the standard normal distribution function.

Remark. Note that Theorem 1.1 implies the celebrated Lindeberg–Feller theorem.

Chen and Shao use Stein's method to prove Theorem 1.1 (and its extention to non-uniform bounds). In the present paper, we prove an extension of Theorem 1.1 to the multivariate case and random vectors with a certain type of dependency structure described in Barbour et al. [4], but with the rate of convergence expressed in terms of expectations of *smooth* test functions, naturally arising from Stein's method. In other words, we extend the result of Barbour et al. [4] to random vectors with finite second moment only.

Clearly, Berry–Esseen type bounds are more natural in a statistical context. However, by way of Stein's method, the derivation of the correct rate of convergence requires much greater effort; to the best of the author's knowledge, no result comparable in generality to Barbour et al. [4] has so far been obtained (see Chen and Ho [9], Barbour [1], Bolthausen [5], Stein [22], Chen [8], Götze [14], Bolthausen and Götze [6], Rinott [19], Rinott and Rotar [20] and Rinott and Rotar [21]). On the other hand, expectations of smooth test functions are much easier to handle and they suffice to prove weak convergence to the normal distribution.

The elegant and powerful approach suggested by Barbour et al. [4] can be applied to numerous dependency structures. In Section 3, we apply our result to the study of local dependency; we prove a modification of Rinott's [19] result, where bounds analogous to the r. h. s. of (1.1) are derived for the multivariate case and smooth test functions. As an illustration, we investigate the rate of convergence in the CLT for non-degenerate U-statistics. Section 4 is concerned with random graph degree statistics. Finally, in Section 5, we consider linear rank statistics and prove an analogue of Theorem 1.1 in the sense described above.

2 Main result

Similarly as in Barbour et al. [4], let W be a $\mathbb{R}^m$-valued random vector decomposed in the following way:

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

$$\mathbf{E}\, X_i = 0,\ i \in I; \quad \mathrm{var}(W) = \mathbf{I}_m \tag{2.2}$$

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

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

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

I and K_i are index sets, and $\mathbf{I}_m$ denotes the identity matrix in $\mathbb{R}^m$. We also assume that:

$$\sum_{i \in I} \big(\mathbf{E}\,|X_i|^2\big)^{1/2} < \infty, \quad \sum_{i \in I} \sum_{k \in K_i} \mathbf{E}\,|X_i|\,|Z_{ik}| < \infty \tag{2.6}$$

Before formulating the main theorem, we introduce some notation. By $f^{(r)}(x) u_1 u_2 \ldots u_r$, we denote the r-th derivative of a function $f\colon\ \mathbb{R}^m \to \mathbb{R}$ in directions $u_1, \ldots u_r$. Thus, $f^{(r)}(x)$ can be regarded as an r-linear form on $\mathbb{R}^m$, while $u_1 \ldots u_r$ can be regarded as an r-tensor. Note that a 2-tensor uv can be identified with the matrix uv^T.

For a $(r-1)$-times differentiable function f, define:

$$M_r(f) := \sup_{\substack{|u_1|, \ldots |u_r| \le 1 \\ x \ne y}} \frac{\big|\big\{f^{(r-1)}(x) - f^{(r-1)}(y)\big\} u_1 \ldots u_{r-1}\big|}{|x - y|}$$

where $|\cdot|$ denotes the euclidean norm of a vector (as well as the absolute value of a number).

The following theorem is our main result.

Theorem 2.1. *For every $\mathbb{R}^m$-valued random vector W decomposed as in* (2.1)–(2.5) *and for every differentiable test function f, we have:*

$$\begin{aligned} \big|\mathbf{E}\, f(W) - \mathcal{N}(0, \mathbf{I}_m)\{f\}\big| &\le M_2(f) \sum_{i \in I} \sum_{k \in K_i} \mathbf{E}\,|X_i||Z_{ik}| \times \\ \times \Bigg[\mathbf{E} \min\Bigg\{1, \sqrt{\frac{\pi}{8}} |Z_i + V_{ik}|\Bigg\} &+ \min\Bigg\{1, \sqrt{\frac{\pi}{32}} \big(|Z_i + V_{ik}| + |V_{ik}|\big)\Bigg\}\Bigg] \end{aligned} \tag{2.7}$$

As already mentioned, we shall apply Stein's method. The following modification of Lemma 2.1 in Götze [14] (see also Barbour [2], equations (2.19)–(2.24)) concerns the solutions of the Stein equation:

$$\Delta g(x) - g'(x)x = f(x) - \mathcal{N}(0, \mathbf{I}_m)\{f\} \tag{2.8}$$

where Δ denotes the Laplace operator. We postpone the proof of the lemma until the end of the section.

Lemma 2.2. *For every differentiable function $f\colon \mathbb{R}^m \to \mathbb{R}$ with $M_2(f) < \infty$, there is a solution g of the Stein equation* (2.8) *satisfying:*

$$M_2(g) \le \frac{1}{2} M_2(f) \tag{2.9}$$

$$M_3(g) \le \sqrt{\frac{\pi}{8}} M_2(f) \tag{2.10}$$

Proof of Theorem 2.1. Let g be the solution from Lemma 2.2. Identifying the 2-tensor $e_1^2 + \ldots + e_m^2$ with the matrix $\mathbf{I}_m$ and the random matrix WW^T with the random 2-tensor W^2, observe that:

$$\Delta g(x) = g''(x)\mathbf{I}_m = g''(x)\operatorname{var}(W) = g''(x)\,\mathbf{E}\,W^2 \tag{2.11}$$

Combining this with (2.8) and (2.1)–(2.3), we obtain:

$$\begin{aligned} \mathbf{E}\,f(W) - \mathcal{N}(0, \mathbf{I}_m)\{f\} &= \sum_{i \in I} \mathbf{E}\Big[g''(W)\,\mathbf{E}\,X_i W - g'(W)X_i\Big] \\ &= \sum_{i \in I} \mathbf{E}\Big[g''(W)\,\mathbf{E}\,X_i Z_i + g'(W_i)X_i - g'(W)X_i\Big] \end{aligned} \tag{2.12}$$

Now let θ be a random variable uniformly distributed over $[0,1]$ and independent of all other variates. Applying the Leibniz formula and combining with (2.4)–(2.5), we obtain:

$$\begin{aligned} \mathbf{E}\,f(W) - \mathcal{N}(0, \mathbf{I}_m)\{f\} &= \sum_{i \in I} \mathbf{E}\Big[g''(W)\,\mathbf{E}\,X_i Z_i - g''(W_i + \theta Z_i)X_i Z_i\Big] = \\ &= \sum_{i \in I}\sum_{k \in K_i} \mathbf{E}\Big[g''(W)\,\mathbf{E}\,X_i Z_{ik} - g''(W_i + \theta Z_i)X_i Z_{ik}\Big] = \\ &= \sum_{i \in I}\sum_{k \in K_i} \mathbf{E}\Big[\big\{g''(W) - g''(W_{ik})\big\}\hat{X}_i \hat{Z}_{ik} - \\ &\qquad - \big\{g''(W_i + \theta Z_i) - g''(W_{ik})\big\}X_i Z_{ik}\Big] \end{aligned} \tag{2.13}$$

where $(\hat{X}_i, \hat{Z}_{ik})$ is an independent copy of (X_i, Z_{ik}). Using (2.9) and (2.10), we estimate:

$$\begin{aligned} \left|\mathbf{E}\, f(W) - \mathcal{N}(0, \mathbf{I}_m)\{f\}\right| \leq M_2(f) \sum_{i \in I} \sum_{k \in K_i} \mathbf{E}\Big[|\hat{X}_i|\, |\hat{Z}_{ik}| \times \\ \times \min\left\{1, \sqrt{\frac{\pi}{8}} |Z_i + V_{ik}|\right\} + |X_i|\, |Z_{ik}| \min\left\{1, \sqrt{\frac{\pi}{8}} |\theta Z_i + V_{ik}|\right\}\Big] \end{aligned} \tag{2.14}$$

and observe that $|\theta Z_i + V_{ik}| \leq \theta |Z_i + V_{ik}| + (1-\theta)|V_{ik}|$. Theorem 2.1 now follows. □

Proof of Lemma 2.2. We shall use the standard argument based on the fact that the operator on the l. h. s. of (2.8) is the infinitesimal generator of the Ornstein–Uhlenbeck semigroup. Without loss of generality, we may assume that $\mathcal{N}(0, \mathbf{I}_m)\{f\} = 0$. Denoting by ϕ_m the density of the standard normal distribution in $\mathbb{R}^m$, one can follow Barbour [2] to verify that the function:

$$g(x) := -\int_0^\infty \int_{\mathbb{R}^m} f\big(e^{-t} x + \sqrt{1 - e^{-2t}}\, z\big) \phi_m(z)\, dz\, dt \tag{2.15}$$

solves (2.8). Differentiation with respect to x, combined with the dominated convergence theorem, yields:

$$g'(x)u = -\int_0^\infty e^{-t} \int_{\mathbb{R}^m} f'\big(e^{-t} x + \sqrt{1 - e^{-2t}}\, z\big) u\, \phi_m(z)\, dz\, dt \tag{2.16}$$

Consequently,

$$|\{g'(x) - g'(y)\}u| \leq M_2(f)|x-y| \int_0^\infty e^{-2t}\, dt \int_{\mathbb{R}^m} \phi_m(z)\, dz = \frac{1}{2} M_2(f)|x-y| \tag{2.17}$$

which proves (2.9). On the other hand, integration by parts in the inner integral of (2.16), in the direction of the vector u, leads to:

$$g'(x)u = \int_0^\infty \frac{e^{-t}}{\sqrt{1 - e^{-2t}}} \int_{\mathbb{R}^m} f\big(e^{-t} x + \sqrt{1 - e^{-2t}}\, z\big) \phi_m'(z) u\, dz\, dt \tag{2.18}$$

Similarly, further differentiation yields:

$$\begin{aligned} M_3(g) &\leq M_2(f) \int_0^\infty \frac{e^{-3t}}{\sqrt{1 - e^{-2t}}}\, dt \sup_{\substack{u \in \mathbb{R}^m \\ |u|=1}} \int_{\mathbb{R}^m} |\phi_m'(z) u|\, dz = \\ &= \frac{\pi}{4} M_2(f) \int_{-\infty}^\infty |\phi_1'(z)|\, dz = \\ &= \sqrt{\frac{\pi}{8}}\, M_2(f) \end{aligned} \tag{2.19}$$

proving (2.10). □

3 Local dependency

In the context of normal approximation by Stein's method, local dependency has been investigated, among others, by Stein [22], Chen [8], Barbour et al. [4], Barbour [1], Rinott [19], and Rinott and Rotar [20]. We shall consider local dependency according to Rinott's [19] definition.

Definition. Let I be an index set. A graph Γ with the vertex set I is said to be a *dependency graph* for a collection of random variables X_i, $i \in I$, if for every disjoint subsets $K, L \subset I$ which are not connected by an edge of Γ, the collections $\{X_k : k \in K\}$ and $\{X_l : l \in L\}$ are independent.

We shall prove the following analogue of Theorem 1.1.

Theorem 3.1. *Let Γ be a dependency graph for a collection of $\mathbb{R}^m$-valued random vectors X_i, $i \in I$. Suppose that $\mathbf{E}\, X_i = 0$ for all $i \in I$. In addition, suppose that the sum $W := \sum_{i \in I} X_i$ is L^2-convergent and $\operatorname{var}(W) = \mathbf{I}_m$. Denoting by $\deg(i; \Gamma)$ the degree of the vertex i with respect to Γ, put $D := 1 + \max_{i \in I} \deg(i; \Gamma)$. Then for every differentiable function $f \colon \mathbb{R}^m \to \mathbb{R}$, we have:*

$$\begin{aligned} &\left|\mathbf{E}\, f(W) - \mathcal{N}(0, \mathbf{I}_m)\{f\}\right| \le \\ &\quad \le \frac{247}{92} M_2(f) D \sum_{i \in I} \mathbf{E}\, |X_i|^2 \min\left\{1, \sqrt{\frac{\pi}{2}}\, D|X_i|\right\} \end{aligned} \tag{3.1}$$

The following corollary provides a sufficient condition for the convergence to the standard normal distribution.

Corollary 3.2. *Let $X_i^{(n)}$, $n \in \mathbb{N}$, $i \in I_n$, be a triangular array of $\mathbb{R}^m$-valued random vectors. For each n, let Γ_n be a dependency graph of the family $X_i^{(n)}$, $i \in I_n$. Define $W^{(n)}$ and D_n analogously to the quanities in Theorem 3.1; similarly, assume that $\mathbf{E}\, X_i^{(n)} = 0$ and $\operatorname{var}(W^{(n)}) = \mathbf{I}_m$ for all $n \in \mathbb{N}, i \in I_n$. In addition, suppose that for every $\varepsilon > 0$,*

$$\lim_{n \to \infty} D_n \sum_{i \in I_n} \mathbf{E} \bigl|X_i^{(n)}\bigr|^2 \mathbf{1}\left[\bigl|X_i^{(n)}\bigr| > \frac{\varepsilon}{v_n}\right] = 0 \tag{3.2}$$

where:

$$v_n := D_n^2 \sum_{i \in I_n} \mathbf{E} \bigl|X_i^{(n)}\bigr|^2 \tag{3.3}$$

Then the sequence $W^{(n)}$ converges weakly to $\mathcal{N}(0, \mathbf{I}_m)$.

Remark. For independent random vectors, the graph Γ may be taken to be empty, so that $D = 1$ and the r. h. s. of (3.1) reduces (up to a constant factor) to the r. h. s. of (1.1), while Corollary 3.2 reduces to the Lindeberg–Feller theorem.

Before proving Theorem 3.1 and Corollary 3.2, we apply Theorem 3.1 to the study of non-degenerate U-statistics. Let $\xi_1, \xi_2, \dots$ be independent and identically distributed random variables taking values in a measurable space $(\Xi, \mathcal{X})$. Furthermore, let $h\colon \Xi^r \to \mathbb{R}^m$ be a symmetric measurable function with $\mathbf{E}\,|h(\xi_1, \dots \xi_r)|^2 < \infty$ and such that $\operatorname{var}\bigl(\mathbf{E}(h(\xi_1, \dots \xi_r) \mid \xi_1)\bigr) > 0$. Without loss of generality, we can assume that $\mathbf{E}\,h(\xi_1, \dots \xi_r) = 0$ and that $\operatorname{var}\bigl(\mathbf{E}(h(\xi_1, \dots \xi_r) \mid \xi_1)\bigr) = \mathbf{I}_m$. Denote by I_n the set of all subsets of $\{1, \dots n\}$ with exactly r elements and define:

$$
\begin{aligned}
Y_i &:= h(\xi_{i_1}, \dots \xi_{i_r}); \quad i = \{i_1, \dots i_r\} \in \bigcup_{n \ge r} I_n \\
S^{(n)} &:= \sum_{i \in I_n} Y_i \\
W^{(n)} &:= \operatorname{var}\bigl(S^{(n)}\bigr)^{-1/2} S^{(n)}
\end{aligned}
\tag{3.4}
$$

Asymptotic normality of $W^{(n)}$ was proved by Hoeffding [15]. For finite third moment, (univariate) Berry–Esseen type bounds were derived by Callaert and Janssen [7]. We prove the following theorem.

Theorem 3.3. *For every $r \in \mathbb{N}$, there is a universal constant C_r, such that for every differentiable function $f\colon \mathbb{R}^m \to \mathbb{R}$,*

$$
\begin{aligned}
&\bigl|\mathbf{E}\,f\bigl(W^{(n)}\bigr) - \mathcal{N}(0, \mathbf{I}_m)\{f\}\bigr| \le \\
&\quad \le C_r M_2(f)\ \mathbf{E}\,|h(\xi_1, \dots \xi_r)|^2 \min\bigl\{1, n^{-1/2}|h(\xi_1, \dots \xi_r)|\bigr\}
\end{aligned}
\tag{3.5}
$$

Consequently, the sequence $W^{(n)}$ converges weakly to $\mathcal{N}(0, \mathbf{I}_m)$.

Proof. One can write:

$$
W^{(n)} = \sum_{i \in I_n} X_i^{(n)} \tag{3.6}
$$

where:

$$
X_i^{(n)} := \operatorname{var}\bigl(S^{(n)}\bigr)^{-1/2} Y_i \tag{3.7}
$$

Now let Γ_n be the graph on I_n in which the vertices i and j are connected in Γ_n when $i \cap j \neq \emptyset$. Clearly, Γ_n is a dependency graph for the family $\bigl\{X_i^{(n)} : i \in I_n\bigr\}$. Hence,

$$
D_n = \binom{n}{r} - \binom{n-r}{r} \asymp n^{r-1} \tag{3.8}
$$

writing $a_n \asymp b_n$ for sequences of positive real numbers with $0 < \liminf_{n\to\infty} a_n/b_n \le \limsup_{n\to\infty} a_n/b_n < \infty$. Furthermore, observe that:

$$
\operatorname{var}\bigl(S^{(n)}\bigr) \ge \sum_{|i \cap j| = 1} \operatorname{cov}(Y_i, Y_j) = c_n \operatorname{var}\bigl(\mathbf{E}(h(\xi_1, \dots \xi_r) \mid \xi_1)\bigr) = c_n \mathbf{I}_m \tag{3.9}
$$

where:

$$c_n := |\{(i,j) \in I_n \times I_n : |i \cap j| = 1\}| \asymp n^{2r-1} \tag{3.10}$$

and where $|\cdot|$ denotes cardinality. The proof is now completed by Theorem 3.1 and a straightforward calculation. ☐

Proof of Theorem 3.1. In order to satisfy (2.3)–(2.5), define K_i to be the set of all vertices adjacent to i with respect to Γ and put:

$$Z_{ik} := X_k, \quad V_{ik} := \sum_{l \in K_k \setminus K_i} X_l \tag{3.11}$$

By Theorem 2.1, we have:

$$\begin{aligned} &\left|\mathbf{E}\, f(W) - \mathcal{N}(0, \mathbf{I}_m)\{f\}\right| \leq \\ &\leq M_2(f) \sum_{i \in I} \sum_{k \in K_i} \mathbf{E}\,|X_i|\,|X_k| \Bigg[\mathbf{E} \min\Bigg\{1, \sqrt{\frac{\pi}{8}} \sum_{l \in K_i \cup K_k} |X_l|\Bigg\} + \\ &\quad + \min\Bigg\{1, \sqrt{\frac{\pi}{32}} \Bigg(\sum_{l \in K_i \cup K_k} |X_l| + \sum_{l \in K_k \setminus K_i} |X_l|\Bigg)\Bigg\}\Bigg] \end{aligned} \tag{3.12}$$

The proof is now completed by the following lemma and a routine calculation. ☐

Lemma 3.4. *For any non-negative numbers $x, y, z_1, \ldots z_r, a_1, \ldots a_r$ and for every $a \geq a_1 + \ldots a_r$, $a > 0$, we have:*

$$\begin{aligned} &xy \min\Bigg\{1, \sum_{j=1}^{r} a_j z_j\Bigg\} \leq L_r(x, y; z_1, \ldots z_r; a_1, \ldots a_r; a) := \\ &:= \frac{1}{2}x^2 \min\{1, ax\} + \frac{1}{2}y^2 \min\{1, ay\} + \frac{9}{23a} \sum_{j=1}^{r} a_j z_j^2 \min\{1, az_j\} \end{aligned} \tag{3.13}$$

Consequently, for any non-negative random variables $X, Y, Z_1, \ldots Z_r$ and numbers $a, a_1, \ldots a_r$ as before, we have:

$$\mathbf{E}\, XY \min\Bigg\{1, \sum_{j=1}^{r} a_j Z_j\Bigg\} \leq \mathbf{E}\, L_r(X, Y; Z_1, \ldots Z_r; a_1, \ldots a_r; a) \tag{3.14}$$

$$\mathbf{E}\, XY \cdot \mathbf{E} \min\Bigg\{1, \sum_{j=1}^{r} a_j Z_j\Bigg\} \leq \mathbf{E}\, L_r(X, Y; Z_1, \ldots Z_r; a_1, \ldots a_r; a) \tag{3.15}$$

Proof. Observe first that it suffices to prove (3.13); (3.14) is then immediate. To derive (3.15), we can suppose without loss of generality that (X, Y) is independent of $(Z_1, \ldots Z_r)$. Then (3.15) follows immediately from (3.14).

To prove (3.13), observe that $uv \leq \frac{1}{2}(u^2 + v^2)$ for any $u, v \geq 0$. Hence it suffices to prove:

$$x^2 \min\left\{1, \sum_{j=1}^{r} a_j z_j\right\} \leq x^2 \min\{1, ax\} + \frac{9}{23a} \sum_{j=1}^{r} a_j z_j^2 \min\{1, az_j\} \tag{3.16}$$

To prove (3.16), we consider two cases. Suppose first that $ax \geq 1$. In this case, we simply estimate:

$$x^2 \min\left\{1, \sum_{j=1}^{r} a_j z_j\right\} \leq x^2 = x^2 \min\{1, ax\} \tag{3.17}$$

Now suppose that $ax < 1$. In this case, we first apply the inequality $u^2 v \leq \frac{2}{3}u^3 + \frac{1}{3}v^3$ for $u = a^{1/3}x$, followed by the Cauchy–Schwarz inequality:

$$\begin{aligned} x^2 \min\left\{1, \sum_{j=1}^{r} a_j z_j\right\} &\leq \frac{2a}{3}x^3 + \frac{1}{3a^2} \min\left\{1, \left(\sum_{j=1}^{r} a_j z_j\right)^3\right\} \leq \\ &\leq x^2 \min\{1, ax\} + \frac{1}{3a^2} \min\left\{1, \left(a \sum_{j=1}^{r} a_j z_j^2\right)^{3/2}\right\} \end{aligned} \tag{3.18}$$

It can be proved by straightforward calculation that $\min\{1, s^{3/2}\} \leq \frac{27}{23}(t^3 + s - t^2)$ for every $s \geq 0$, $0 \leq t \leq \sqrt{s}$. Consequently, we have $\min\{1, (u+v)^{3/2}\} \leq \frac{27}{23}(u^{3/2} + v)$ for all $u, v \geq 0$. Splitting the sum in the second term and applying this inequality, we obtain:

$$\begin{aligned} &\frac{1}{3a^2} \min\left\{1, \left(a \sum_{j=1}^{r} a_j z_j^2\right)^{3/2}\right\} = \\ &= \frac{1}{3a^2} \min\left\{1, a^{3/2}\left(\sum_{j: az_j < 1} a_j z_j^2 + \sum_{j: az_j \geq 1} a_j z_j^2\right)^{3/2}\right\} \leq \\ &\leq \frac{9}{23a^2}\left\{\left(a \sum_{j: az_j < 1} a_j z_j^2\right)^{3/2} + a \sum_{j: az_j \geq 1} a_j z_j^2\right\} \end{aligned} \tag{3.19}$$

Applying Hölder's inequality in the first term, we conclude that:

$$\frac{1}{3a^2}\min\left\{1,\left(a\sum_{j=1}^{r}a_j z_j^2\right)^{3/2}\right\}\le\frac{9}{23a}\left\{a\sum_{j:az_j<1}a_j z_j^3+\sum_{j:az_j\ge 1}a_j z_j^2\right\}= \\ =\frac{9}{23a}\sum_{j=1}^{r}a_j z_j^2\min\{1,az_j\} \tag{3.20}$$

This completes the proof. □

Proof of Corollary 3.2. Observe that for every $\varepsilon>0$ and $n\in\mathbb{N}$,

$$\min\left\{1,\sqrt{\frac{\pi}{2}}\,D_n\big|X_i^{(n)}\big|\right\}\le\mathbf{1}\left[\big|X_i^{(n)}\big|>\frac{\varepsilon}{v_n}\right]+\sqrt{\frac{\pi}{2}}D_n\frac{\varepsilon}{v_n} \tag{3.21}$$

Since every bounded continuous function can be approximated by smooth functions, it suffices to prove that $\lim_{n\to\infty}\mathbf{E}\,f\big(W^{(n)}\big)=\mathcal{N}(0,\mathbf{I}_m)\{f\}$ for every differentiable f with $M_2(f)\le 1$. Combining (3.21) with (3.1) and (3.3), we obtain:

$$\big|\mathbf{E}\,f\big(W^{(n)}\big)-\mathcal{N}(0,\mathbf{I}_m)\{f\}\big|\le \\ \le\frac{247}{92}\left\{D_n\sum_{i\in I_n}\mathbf{E}\big|X_i^{(n)}\big|^2\,\mathbf{1}\left[\big|X_i^{(n)}\big|>\frac{\varepsilon}{v_n}\right]+\sqrt{\frac{\pi}{2}}\,\varepsilon\right\} \tag{3.22}$$

Letting $n\to\infty$, applying (3.2) and then letting $\varepsilon\to 0$, we obtain the desired result. □

4 Random graph degree statistics

Stein's method has turned out to be particularly suitable for dependency structures appearing in random graphs in the context of Poisson (see Barbour et al. [3]) as well as normal approximation (see Barbour et al. [4], Goldstein and Rinott [13], Rinott and Rotar [20] and references therein). In particular, Goldstein and Rinott [13] investigate random graph degree counts. In this paper, we replace indicators by arbitrary multivariate statistics; normal convergence is proved under the assumption of finite second moments with respect to a suitable Poisson distribution in the sense of (4.14).

Consider an index set I and a symmetric matrix p_{ij}, $i,j\in I$, with $p_{ii}=0$. Let Γ be a random graph on the vertex set I, where each unordered pair $\{i,j\}$ of vertices makes up an edge with probability p_{ij}, independently of all other such pairs. For each $i\in I$, take a measurable function $h_i\colon\ \mathbb{Z}_+\to\mathbb{R}^m$, and define:

$$X_i:=h_i(\delta_i);\quad W:=\sum_{i\in I}X_i \tag{4.1}$$

where δ_i denotes the degree of the vertex i with respect to Γ. Without loss of generality, we may (and will) assume that $\mathbf{E}\, X_i = 0$. The covariance matrix of W can then be expressed in the following way: denote by $\delta_k^{(i_1 i_2 \ldots i_r)}$ the degree of the vertex k, ignoring all edges with an endpoint in any of the vertices $i_1, \ldots i_r$. Since $\delta_k^{(i)}$ is independent of X_i, we have:

$$\operatorname{var}(W) = \sum_{i \in I} \mathbf{E}\, X_i X_i^T + \sum_{i \in I} \sum_{k \in I \setminus \{i\}} \mathbf{E}\, X_i Z_{ik} \tag{4.2}$$

where:

$$Z_{ik} := X_k - X_{ik}, \quad X_{ik} := h_k\big(\delta_k^{(i)}\big); \quad i \neq k \tag{4.3}$$

Noting that:

$$X_i Z_{ik} = \mathbf{1}(J_{ik}) h_i\big(1 + \delta_i^{(k)}\big) \Big[h_k\big(1 + \delta_k^{(i)}\big) - h_k\big(\delta_k^{(i)}\big)\Big] \tag{4.4}$$

where:

$$J_{ik} := \{i \text{ and } k \text{ are adjacent in } \Gamma\} \tag{4.5}$$

and using the independence of J_{ik}, $\delta_i^{(k)}$ and $\delta_k^{(i)}$, we obtain:

$$\begin{aligned} \operatorname{var}(W) = &\sum_{i \in I} \mathbf{E}\, h_i(\delta_i) h_i(\delta_i)^T + \\ &+ \sum_{i \in I} \sum_{k \in I \setminus \{i\}} p_{ik}\ \mathbf{E}\, h_i\big(1 + \delta_i^{(k)}\big) \Big[\mathbf{E}\, h_k\big(1 + \delta_k^{(i)}\big) - \mathbf{E}\, h_k\big(\delta_k^{(i)}\big)\Big]^T \end{aligned} \tag{4.6}$$

Without loss of generality, we may (and will) assume that $\operatorname{var}(W) = \mathbf{I}_m$. Now define:

$$\begin{aligned} \lambda &:= \max_{i \in I} \sum_{k \in I \setminus \{i\}} p_{ik} \\ T_{1i} &:= \max_{\substack{r=0,1,2 \\ k,l \in I}} \mathbf{E}\big|h_i\big(r + \delta_i^{(kl)}\big)\big| \\ T_{2i} &:= \max_{\substack{r=0,1 \\ k \in I}} \mathbf{E}\big|h_i\big(r + \delta_i^{(k)}\big)\big|^2 \\ T_{3i} &:= \mathbf{E}\, |X_i|^2 \min\left\{1, \sqrt{\frac{\pi}{32}}\, |X_i|\right\} \quad ; \quad i \in I \\ T_r &:= \sum_{i \in I} T_{ri}^{3/r} \quad ; \quad r = 1, 2, 3 \end{aligned} \tag{4.7}$$

Theorem 4.1. *Under the conditions given above, we have:*

$$\begin{aligned}&\big|\mathbf{E}\, f(W) - \mathcal{N}(0, \mathbf{I}_m)\{f\}\big| \le \\ &\le M_2(f)\left[\sqrt{\frac{\pi}{2}}\Big\{7\lambda^2 T_1 + 2\lambda T_1 + 3\lambda T_1^{1/3} T_2^{2/3}\Big\} + 3T_3\right]\end{aligned} \tag{4.8}$$

Remark. Like in Corollary 3.2, a Lindeberg–Feller type theorem could also be derived from Theorem 4.1.

To illustrate how Theorem 4.1 works, we consider the case where all p_{ik}'s and h_i's are equal. Fix $\lambda > 0$ and a function $h\colon \mathbb{Z}_+ \to \mathbb{R}^m$ and let $\Gamma^{(n)} := K(n+1, \lambda/n)$, the random graph on vertices $0, 1, \dots n$ where any two distinct vertices are connected with probability λ/n. Define:

$$Y_i^{(n)} := h\big(\deg(i; \Gamma^{(n)})\big); \qquad S^{(n)} := \sum_{i=1}^{n} Y_i^{(n)} \tag{4.9}$$

and, provided that $\operatorname{var}\big(S^{(n)}\big)$ is non-degenerate,

$$W^{(n)} := \operatorname{var}\big(S^{(n)}\big)^{-1/2}\big(S^{(n)} - \mathbf{E}\, S^{(n)}\big) \tag{4.10}$$

In order to apply Theorem 4.1, we shall need the following lemma, which will be proved at the end of the section.

Lemma 4.2. *For every $\lambda > 0$ and $d, s \in \mathbb{Z}_+$, there is a constant $C_{\lambda,d,s}$, such that for all $n \in \mathbb{N}$ and $r \in \mathbb{Z}_+$,*

$$\operatorname{Bi}\left(n - d, \frac{\lambda}{n}\right)\{r - s\} \le C_{\lambda,d,s} \min\{n^{s/2}, r^s\} \operatorname{Po}(\lambda)\{r\} \tag{4.11}$$

provided that $n \ge d$.

Firstly, we shall examine the asymptotic behavior of $\operatorname{var}\big(S^{(n)}\big)$. Take random variables:

$$B_k^{(n)} \sim \operatorname{Bi}\left(k, \frac{\lambda}{n}\right); \quad \Pi \sim \operatorname{Po}(\lambda) \tag{4.12}$$

By (4.6), we have:

$$\begin{aligned}\frac{1}{n+1}\operatorname{var}\big(S^{(n)}\big) &= \mathbf{E}\, h\big(B_n^{(n)}\big) h\big(B_n^{(n)}\big)^T + \\ &+ \lambda\, \mathbf{E}\, h\big(1 + B_{n-1}^{(n)}\big)\Big[\mathbf{E}\, h\big(1 + B_{n-1}^{(n)}\big) - \mathbf{E}\, h\big(B_{n-1}^{(n)}\big)\Big]^T\end{aligned} \tag{4.13}$$

By the Poisson law of small numbers, the point probabilities $\mathbf{P}\big[B_{n+k}^{(n)} = r\big]$ converge to $\mathbf{P}\big[\Pi = r\big]$ as $n \to \infty$ for all $k, r \in \mathbb{Z}$. Suppose that:

$$\mathbf{E}\,|h(\Pi)|^2 < \infty \tag{4.14}$$

Furthermore, without loss of generality, we may assume that:

$$\mathbf{E}\,h(\Pi) = 0 \tag{4.15}$$

The dominated convergence theorem together with Lemma 4.2 now implies:

$$\lim_{n\to\infty} \frac{1}{n+1} \operatorname{var}\big(S^{(n)}\big) = \mathbf{E}\,h(\Pi)h(\Pi)^T + \lambda\,\mathbf{E}\,h(1+\Pi)h(1+\Pi)^T \tag{4.16}$$

Without loss of generality, we may assume that:

$$\mathbf{E}\,h(\Pi)h(\Pi)^T + \lambda\,\mathbf{E}\,h(1+\Pi)h(1+\Pi)^T = \mathbf{I}_m \tag{4.17}$$

Theorem 4.3. *Under the assumption of* (4.14), (4.15) *and* (4.17), *we have:*

$$\begin{aligned}\big|\mathbf{E}\,f\big(W^{(n)}\big) - \mathcal{N}(0,\mathbf{I}_m)\{f\}\big| &\le C(\lambda)\Big[\mathbf{E}\,|h(\Pi)|^2 \min\{1, n^{-1/2}|h(\Pi)|\} + \\ &+ n^{-1/2}\big(\mathbf{E}\,|h(\Pi)|^2\big)^{3/2} + \big(\mathbf{E}\,|h(\Pi)|^2\big)^{1/2}\,\mathbf{E}\,|h(\Pi)|^2 \min\{1, n^{-1/2}\Pi\}\Big]\end{aligned} \tag{4.18}$$

Consequently, the sequence $W^{(n)}$ converges weakly to $\mathcal{N}(0,\mathbf{I}_m)$.

Proof. One can write $W^{(n)} = \sum_{i=0}^{n} X_i^{(n)}$, where:

$$\begin{aligned} X_i^{(n)} &:= \operatorname{var}\big(S^{(n)}\big)^{-1/2}\big(Y_i^{(n)} - \mathbf{E}\,Y_i^{(n)}\big) = h^{(n)}\big(\deg\big(i;\Gamma^{(n)}\big)\big) \\ h^{(n)}(r) &:= \operatorname{var}\big(S^{(n)}\big)^{-1/2}\big(h(r) - \mathbf{E}\,Y_i^{(n)}\big)\end{aligned} \tag{4.19}$$

Letting $T_k^{(n)}$ be defined as T_k in (4.7), with X_i replaced by $X_i^{(n)}$ and h_i by $h^{(n)}$, we have:

$$\begin{aligned} T_1^{(n)} &= (n+1) \max_{\substack{d=1,2\\ s=0,1,2}} \Big[\mathbf{E}\big|h^{(n)}\big(s + B_{n-d}\big)\big|\Big]^3 \\ T_2^{(n)} &= (n+1) \max_{r=0,1} \Big[\mathbf{E}\big|h^{(n)}\big(r + B_{n-1}^{(n)}\big)\big|\Big]^{3/2} \\ T_3^{(n)} &= (n+1)\,\mathbf{E}\big|h^{(n)}\big(B_n^{(n)}\big)\big|^2 \min\left\{1, \sqrt{\frac{\pi}{32}}\big|h^{(n)}\big(B_n^{(n)}\big)\big|\right\}\end{aligned} \tag{4.20}$$

By (4.16) and (4.17), we have $\operatorname{var}\big(S^{(n)}\big) \geq 2^{-2/3}(n+1)\mathbf{I}_m$ for sufficiently large n, so that, by Lemma 4.2 and the Cauchy–Schwarz inequality,

$$\begin{aligned}
T_1^{(n)} &\leq 16n^{-1/2} \max_{d,s=0,1,2} \Big[\mathbf{E}\big|h\big(s+B_{n-d}^{(n)}\big)\big|\Big]^3 \leq \\
&\leq 16n^{-1/2} \max_{d,s=0,1,2} \big[C_{\lambda,d,s}\ \mathbf{E}\,|\Pi^s h(\Pi)|\big]^3 \leq \\
&\leq 16n^{-1/2} \max_{d,s=0,1,2} C_{\lambda,d,s}^3 \big(\mathbf{E}\,\Pi^{2s}\big)^{3/2}\big(\mathbf{E}\,|h(\Pi)|^2\big)^{3/2}
\end{aligned} \tag{4.21}$$

Similarly, applying $(x+y)^2 \leq 2(x^2+y^2)$ and $|\,\mathbf{E}\,Y|^2 \leq \mathbf{E}\,|Y|^2$, we have:

$$\begin{aligned}
T_2^{(n)} &\leq 4\sqrt{2}\,n^{-1/2}\Big[\max_{r=0,1}\mathbf{E}\big|h\big(r+B_{n-1}^{(n)}\big)\big|^2 + \mathbf{E}\big|h\big(B_n^{(n)}\big)\big|^2\Big]^{3/2} \leq \\
&\leq 4\sqrt{2}\,n^{-1/2}\Big[(C_{\lambda,0,0}+C_{\lambda,1,0})\,\mathbf{E}\,|h(\Pi)|^2 + \\
&\qquad + C_{\lambda,1,1}\ \mathbf{E}\,|h(\Pi)|^2 \min\{n^{1/2},\Pi\}\Big]^{3/2}
\end{aligned} \tag{4.22}$$

Finally, we estimate:

$$\begin{aligned}
T_3^{(n)} &\leq 4\,\mathbf{E}\Big(\big|h\big(B_n^{(n)}\big)\big|^2 + \mathbf{E}\big|h\big(B_n^{(n)}\big)\big|^2\Big)\times \\
&\qquad \times \min\Big\{1, n^{-1/2}\Big(\big|h\big(B_n^{(n)}\big) + \mathbf{E}\big|h\big(B_n^{(n)}\big)\big|\Big)\Big\} \leq \\
&\leq 4\,\mathbf{E}\big|h\big(B_n^{(n)}\big)\big|^2 \min\{1, n^{-1/2}\big|h\big(B_n^{(n)}\big)\big|\} + 12n^{-1/2}\Big(\mathbf{E}\big|h\big(B_n^{(n)}\big)\big|^2\Big)^{3/2} \leq \\
&\leq 4C_{\lambda,0,0}\ \mathbf{E}\,|h(\Pi)|^2 \min\{1, n^{-1/2}|h(\Pi)|\} + 12C_{\lambda,0,0}^{3/2} n^{-1/2}\big(\mathbf{E}\,|h(\Pi)|^2\big)^{3/2}
\end{aligned} \tag{4.23}$$

Combining these estimates with (4.8) yields the desired result. □

Proof of Theorem 4.1. As one can easily check, (2.3)–(2.5) can be satisfied by recalling (4.3) and putting:

$$\begin{aligned}
W_i &:= \sum_{k\in I\setminus\{i\}} X_{ik} & K_i &:= I \\
Z_{ii} &:= X_i & W_{ii} &:= W_i \\
W_{ik} &:= \sum_{l\in I\setminus\{i,k\}} X_{ikl} & X_{ikl} &:= h_l\big(\delta_l^{(ik)}\big); \qquad k \neq i
\end{aligned} \tag{4.24}$$

Noting that:

$$
\begin{aligned}
Z_i &= X_i + \sum_{k \in I \setminus \{i\}} Z_{ik} \\
V_{ii} = 0, \quad V_{ik} &= X_{ik} + \sum_{l \in I \setminus \{i,k\}} V_{ikl} && ; \; k \neq i \\
Z_{ik} + V_{ik} &= X_i + X_k + \sum_{l \in I \setminus \{i,k\}} (Z_{il} + V_{ikl}) && ; \; k \neq i
\end{aligned}
\tag{4.25}
$$

where:

$$
V_{ikl} := X_{il} - X_{ikl} \tag{4.26}
$$

Theorem 2.1 yields:

$$
\left| \mathbf{E}\, f(W) - \mathcal{N}(0, \mathbf{I}_m)\{f\} \right| \leq M_2(f) R \tag{4.27}
$$

where:

$$
\begin{aligned}
R := & \sum_{i \in I} \mathbf{E}\, |X_i|^2 \left(\mathbf{E} \min\left\{ 1, \sqrt{\frac{\pi}{8}} |X_i| \right\} + \min\left\{ 1, \sqrt{\frac{\pi}{32}} |X_i| \right\} \right) + \\
& + \sqrt{\frac{\pi}{8}} \sum_{i \in I} \sum_{k \in I \setminus \{i\}} \mathbf{E}\, |X_i|^2 \big(\mathbf{E}\, |Z_{ik}| + Z_{ik} \big) + \\
& + \sqrt{\frac{\pi}{8}} \sum_{i \in I} \sum_{k \in I \setminus \{i\}} \mathbf{E}\, |X_i|\, |Z_{ik}| \Big(\mathbf{E}\, |X_i| + \mathbf{E}\, |X_k| + \frac{1}{2} |X_k| + \frac{1}{2} |X_{ik}| \Big) + \\
& + \sqrt{\frac{\pi}{8}} \sum_{i \in I} \sum_{k \in I \setminus \{i\}} \sum_{l \in I \setminus \{i,k\}} \mathbf{E}\, |X_i|\, |Z_{ik}| \Big(\mathbf{E}\, |Z_{il}| + \mathbf{E}\, |V_{ikl}| + \frac{1}{2} |Z_{il}| + |V_{ikl}| \Big)
\end{aligned}
\tag{4.28}
$$

To estimate the first term, note that $|X_i|^2$ and $\min\{2, \sqrt{\pi/8}\, |X_i|\}$ are positively correlated because the second random variable is an non-decreasing function of the first one. Hence,

$$
\mathbf{E}\, |X_i|^2 \cdot \mathbf{E} \min\left\{ 1, \sqrt{\frac{\pi}{8}}\, |X_i| \right\} \leq \mathbf{E}\, |X_i|^2 \cdot \mathbf{E} \min\left\{ 2, \sqrt{\frac{\pi}{8}}\, |X_i| \right\} \leq 2 T_{3i} \tag{4.29}
$$

Furthermore, recalling (4.4) and (4.5), noting that:

$$
\begin{aligned}
|Z_{ik}| &\le \mathbf{1}(J_{ik})\Big[\big|h_k\big(\delta_k^{(i)}\big)\big| + \big|h_k\big(1+\delta_k^{(i)}\big)\big|\Big] \\
|V_{ikl}| &\le \mathbf{1}(J_{kl})\Big[\big|h_l\big(\delta_l^{(ik)}\big)\big| + \big|h_l\big(1+\delta_l^{(ik)}\big)\big|\Big] \\
X_i = h_i\big(1+\delta_i^{(k)}\big) \quad &\text{and} \quad X_{ik} = h_k\big(\delta_k^{(i)}\big) \qquad \text{on the event } J_{ik} \\
X_i &= h_i\big(2+\delta_i^{(kl)}\big) \qquad \text{on the event } J_{ik}\cap J_{il} \\
\mathbf{1}(J_{kl})|Z_{ik}| &\le \mathbf{1}(J_{ik}\cap J_{kl})\Big[\big|h_k\big(1+\delta_k^{(il)}\big)\big| + \big|h_k\big(2+\delta_k^{(il)}\big)\big|\Big] \\
\big|h_k\big(\delta_k^{(i)}\big)\big|\cdot\big|h_k\big(1+\delta_k^{(i)}\big)\big| &\le \frac{1}{2}\Big[\big|h_k\big(\delta_k^{(i)}\big)\big|^2 + \big|h_k\big(1+\delta_k^{(i)}\big)\big|^2\Big]
\end{aligned} \tag{4.30}
$$

Using Hölder's inequality and applying (4.7), Theorem 4.1 is now immediate. □

Proof of Lemma 4.2. Writing:

$$\mathrm{Bi}\left(n,\frac{\lambda}{n}\right)\{r\} = \frac{\lambda^r}{r!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{r-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n-r} \tag{4.31}$$

and applying $1+x\le e^x$, we have:

$$\mathrm{Bi}\left(n,\frac{\lambda}{n}\right)\{r\} \le \mathrm{Po}(\lambda)\{r\}\exp\left(\frac{r(2\lambda+1-r)}{2n}\right) \tag{4.32}$$

For $r\le n$, this inequality implies:

$$\mathrm{Bi}\left(n,\frac{\lambda}{n}\right)\{r\} \le e^{\lambda+1/2}\,\mathrm{Po}(\lambda)\{r\}e^{-r^2/(2n)} \tag{4.33}$$

which is trivial for $r>n$. Next, observe that for $0\le\delta\le\lambda$,

$$\mathrm{Po}(\lambda-\delta)\{r\} \le e^{\delta}\,\mathrm{Po}(\lambda)\{r\} \tag{4.34}$$

Furthermore, observe that:

$$\mathrm{Po}(\lambda)\{r-s\} \le \left(\frac{r}{\lambda}\right)^s \mathrm{Po}(\lambda)\{r\} \tag{4.35}$$

(assuming $0^0=1$). Combining (4.33)–(4.35), we obtain:

$$\mathrm{Bi}\left(n-d,\frac{\lambda}{n}\right)\{r-s\} \le \lambda^{-s}e^{\lambda(1+d)+1/2}\,\mathrm{Po}(\lambda)\{r\}r^se^{-r^2/(2n)} \tag{4.36}$$

Applying $x^pe^{-x}\le p^pe^{-p}$ for all $p,x\ge 0$, we have:

$$r^se^{-r^2/(2n)} \le (ns)^{s/2}e^{-s/2} \tag{4.37}$$

so that:

$$r^se^{-r^2/(2n)} \le \min\left\{s^{s/2}e^{-s/2}n^{s/2}, r^s\right\} \tag{4.38}$$

Lemma 4.2 now follows. □

5 Linear rank statistics

In this section, we shall investigate random vectors of the form:

$$W := \sum_{i=1}^{n} a(i, \pi(i)) \tag{5.1}$$

where $a(i,k) \in \mathbb{R}^m$, $i \in \mathbb{N}_n$, $k \in \mathbb{N}_N$, $N \geq n$ and where π is a random mapping drawn from the uniform distribution over all injections from $\mathbb{N}_n$ to $\mathbb{N}_N$; for $r \in \mathbb{N}$, we denote $\mathbb{N}_r := \{1, 2, \ldots r\}$. Without loss of generality, we can (and will) assume that for all $i \in \mathbb{N}_n$,

$$\sum_{j=1}^{n} a(i,j) = 0 \tag{5.2}$$

In addition, we can assume that:

$$\operatorname{var}(W) = \frac{1}{N-1} \sum_{i=1}^{n} \sum_{j=1}^{N} a(i,j)\, a(i,j)^T - \frac{1}{N(N-1)}\, c(j)\, c(j)^T = \mathbf{I}_m \tag{5.3}$$

where:

$$c(j) = \sum_{i=1}^{n} a(i,j) \tag{5.4}$$

A univariate Lindeberg–Feller type theorem for random variables of this kind was first proved by Motoo [17]. Schneller [23] gives a shorter proof, using Stein's method and Bolthausen's [5] combinatorial argument. In the latter paper, Berry–Esseen type bounds in the CLT are derived. This result was generalized to the multivariate case by Bolthausen and Götze [6]. Barbour [2] proves a functional version of Bolthausen's result.

In all three papers mentioned above, bounds in the CLT are expressed in terms of third moments. Here we shall prove the following result.

Theorem 5.1. *Let W be as in* (5.1)*, satisfying* (5.2) *and* (5.3)*. Then we have:*

$$\left|\mathbf{E}\, f(W) - \mathcal{N}(0, \mathbf{I}_m)\{f\}\right| \leq \frac{8}{N}\, M_2(f) \sum_{i=1}^{n} \sum_{j=1}^{N} |a(i,j)|^2 \min\left\{1, \sqrt{8\pi}\, |a(i,j)|\right\} \tag{5.5}$$

Remark. Theorem 5.1 implies Motoo's [17] result in a similar way as Theorem 3.1 implies Corollary 3.2.

Similarly as in Barbour [2], the proof of Theorem 5.1 will be based on the following fact. Firstly, for any two finite sets $I, J \subset \mathbb{N}$ of the same cardinality, let $\tau_{I,J} \colon \mathbb{N} \to \mathbb{N}$ be the map which maps the r-th element of $I \setminus J$ to the r-th element of $J \setminus I$ and vice versa; the other elements are left unchanged. Thus $\tau_{I,J}$ maps the set I bijectively onto J and $\mathbb{N} \setminus I$ onto $\mathbb{N} \setminus J$. The following lemma is straightforward and is therefore left without proof.

Lemma 5.2. *Let $A, B \subset \mathbb{N}$ with $|A| \le |B| < \infty$, where $|\cdot|$ denotes cardinality, and let π be a uniformly distributed random injection $A \to B$. Then for any subset $I \subset A$ and any random set $J \subset B$ drawn from the uniform distribution over $\{J \subset B : |J| = |I|\}$ and independent of π, the random mapping:*

$$\tau_{\pi(I),J} \circ \pi \colon \; A \setminus I \to B \tag{5.6}$$

is independent of the family $\{\pi(i) : i \in I\}$.

Proof of Theorem 5.1. Let J_1 and J_2 be random sets drawn from the uniform distribution over all subsets of $\mathbb{N}_N$ with one, respectively two elements, independent of each other and jointly independent of π. For any distinct indices $i, k, l \in I := \mathbb{N}_n$, define:

$$\begin{aligned} X_i &:= a\big(i, \pi(i)\big) \\ X_{ik} &:= a\big(k, \tau_{\{\pi(i)\},J_1}(\pi(k))\big) \\ X_{ikl} &:= a\big(l, \tau_{\{\pi(i),\pi(k)\},J_2}(\pi(l))\big) \end{aligned} \tag{5.7}$$

Using Lemma 5.2, observe that (2.3)–(2.5) can be satisfied by putting:

$$\begin{aligned} &W_i := \sum_{k \in I \setminus \{i\}} X_{ik}, && K_i := \mathbb{N}_n && \\ &Z_{ii} := X_i, && Z_{ik} := X_k - X_{ik} && ;\ k \ne i \\ &W_{ii} := W_i, && W_{ik} := \sum_{l \in I \setminus \{i,k\}} X_{ikl} && ;\ k \ne i \end{aligned} \tag{5.8}$$

and notice that:

$$\begin{aligned} Z_i &= X_i + \sum_{k \in I \setminus \{i\}} Z_{ik} \\ V_{ii} = 0, \quad V_{ik} &= X_{ik} + \sum_{l \in I \setminus \{i,k\}} (V_{ikl} - Z_{il}) \qquad ;\ k \ne i \\ Z_{ik} + V_{ik} &= X_i + X_k + \sum_{l \in I \setminus \{i,k\}} V_{ikl} \qquad ;\ k \ne i \end{aligned} \tag{5.9}$$

where (in contrast to (4.26)):

$$V_{ikl} := X_l - X_{ikl} \tag{5.10}$$

so that, by Theorem 2.1,

$$\left|\mathbf{E}\, f(W) - \mathcal{N}(0, \mathbf{I}_m)\{f\}\right| \le M_2(f) R \tag{5.11}$$

where:

$$R := \sum_{i\in\mathbb{N}_n} \mathbf{E}\,|X_i|^2 R_i + \sum_{i\in\mathbb{N}_n}\sum_{k\in\mathbb{N}_n\setminus\{i\}} \mathbf{E}\,|X_i||Z_{ik}|R_{ik} \tag{5.12}$$

and:

$$\begin{aligned}
R_i &:= \mathbf{E}\min\Big\{1, \sqrt{\frac{\pi}{8}}\Big(|X_i| + \sum_{l\in\mathbb{N}_n\setminus\{i\}} |Z_{il}|\Big)\Big\} + \min\Big\{1, \sqrt{\frac{\pi}{32}}|X_i|\Big\} \\
R_{ik} &:= \mathbf{E}\min\Big\{1, \sqrt{\frac{\pi}{8}}\Big(|X_i| + |X_k| + \sum_{l\in\mathbb{N}_n\setminus\{i,k\}} |V_{ikl}|\Big)\Big\} + \\
&\quad + \min\Big\{1, \sqrt{\frac{\pi}{32}}\Big(|X_i| + |X_k| + |X_{ik}| + \sum_{l\in\mathbb{N}_n\setminus\{i,k\}} \big(|Z_{il}| + 2|V_{ikl}|\big)\Big)\Big\}
\end{aligned} \tag{5.13}$$

Noting that for any distinct $i, k, l \in \mathbb{N}_n$,

$$\begin{gathered}
|Z_{ik}| \le \mathbf{1}[\pi(k) \in J_1]\big(|a(k, \pi(i))| + |a(k, \pi(k))|\big), \quad Z_{ik}Z_{il} = 0, \\
|V_{ikl}| \le \mathbf{1}[\pi(l) \in J_2]\big(|a(l, \pi(i))| + |a(l, \pi(k))| + |a(l, \pi(l))|\big), \\
\mathbf{P}[\pi(k) \in J_1] = \frac{1}{N}, \quad \mathbf{P}[\pi(l) \in J_2] \le \frac{2}{N}
\end{gathered} \tag{5.14}$$

conditioning on π and using independence, we find that R can be estimated by two sums of the following form:

$$\sum_{i\in\mathbb{N}_n}\sum_r \alpha_{ir}\,\mathbf{E}\big|a(i, \pi(i))\big|\,\big|a(k'_{ir}, \pi(k''_{ir}))\big| \cdot \mathbf{E}\min\Big\{1, \sum_s \beta_{irs}\big|a(l'_{irs}, \pi(l''_{irs}))\big|\Big\} \tag{5.15}$$

and

$$\sum_{i\in\mathbb{N}_n}\sum_r \alpha_{ir}\,\mathbf{E}\big|a(i, \pi(i))\big|\,\big|a(k'_{ir}, \pi(k''_{ir}))\big| \min\Big\{1, \sum_s \beta_{irs}\big|a(l'_{irs}, \pi(l''_{irs}))\big|\Big\} \tag{5.16}$$

where $\alpha_{ir}, \beta_{irs} > 0$ and $\sum_s \beta_{irs} \le \sqrt{8\pi}$ for all i and r. Theorem 5.1 now follows from Lemma 3.4 and a routine calculation. □

References

[1] Barbour, A. D.: Poisson convergence and random graphs. *Math. Proc. Cambridge Philos. Soc* **92** (1982), 349–359.

[2] Barbour, A. D.: Stein's method for diffusion approximations. *Probab. Theory Related Fields* **84** (1990), 297–322.

[3] Barbour, A. D., Holst, L., and Janson, S.: *Poisson Approximation.* Clarendon Press, Oxford, 1992.

[4] Barbour, A. D., Karoński, M., and Ruciński, A.: A central limit theorem for decomposable random variables with applications to random graphs. *J. Combin. Theory Ser. B* **47** (1989), 125–145.

[5] Bolthausen, E.: An estimate of the remainder in a combinatorial central limit theorem. *Z. Wahrsch. Verw. Gebiete* **66** (1984), 379–386.

[6] Bolthausen, E., and Götze, F.: The rate of convergence for multivariate sampling statistics. *Ann. Statist.* **21**, No. 4 (1993), 1692–1710.

[7] Callaert, H., and Janssen, P.: The Berry–Esseen theorem for U-statistics. *Ann. Statist.* **6** (1978), 417–421.

[8] Chen, L. H. Y.: The rate of convergence in a central limit theorem for dependent random variables with arbitrary index set. IMA Preprint Series #243, 1986.

[9] Chen, L. H. Y., and Ho, S.–T: An L_p bound for the remainder in a combinatorial central limit theorem. *Ann. Probab.* **6** (1978), 231–249.

[10] Chen, L. H. Y., and Shao, Q.–M.: A non-uniform Berry–Esseen bound via Stein's method. *Probab. Theory Related Fields* **120**, No. 3 (2001), 236–254.

[11] Feller, W.: On the Berry–Esseen theorem. *Z. Wahrsch. Verw. Gebiete* **10** (1968), 261–268.

[12] Goldstein, L., and Reinert, G.: Stein's method and the zero bias transformation with applications to simple random sampling. *Ann. Appl. Probab.* **7**, No. 4 (1997), 935–952.

[13] Goldstein, L., and Rinott, Y.: Multivariate normal approximations by Stein's method and size biased couplings. *J. Appl. Probab.* **33** (1996), 1–17.

[14] Götze, F.: On the rate of convergence in the multivariate CLT. *Ann. Probab.* **19** (1991), 724–739.

[15] Hoeffding, W.: A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** (1948), 293–325.

[16] Katz, M. L.: Note on the Berry–Esseen theorem. *Ann. Math. Statist.* **34** (1963), 1107–1108.

[17] Motoo, M.: On the Hoeffding's combinatorial central limit theorem. *Ann. Inst. Statist. Math.* **8** (1957), 145–154.

[18] Petrov, V. V.: An estimate of the deviation of the distribution of a sum of independent random variables from the normal law. *Dokl. Akad. Nauk SSSR n. Ser.* **160** (1965), 1013–1015.

[19] Rinott, Y.: On normal approximation rates for certain sums of dependent random variables. *J. Appl. Comp. Math.* **55** (1994), 135–143.

[20] Rinott, Y., and Rotar, V.: A multivariate CLT for local dependence with $n^{-1/2}\log n$ rate and applications to multivariate graph related statistics. *J. Multivariate Anal.* **56** (1996), 333–350.

[21] Rinott, Y., and Rotar, V.: On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted U-statistics. *Ann. Appl. Probab.* **7** (1997), 1080–1105.

[22] Stein, C.: *Approximate Computation of Expectations*. IMS, Hayward, California, 1986.

[23] Schneller, W.: A short proof of Motoo's combinatorial central limit theorem using Stein's method. *Probab. Theory Related Fields* **78**, No. 2 (1988), 249–252.