Abstract

Let $S(1), \ldots, S(n), T(1), \ldots, T(n)$ be random subsets of the set $[m] = \{1, \ldots, m\}$. We consider the random digraph $D$ on the vertex set $[n]$ defined as follows: the arc $i \rightarrow j$ is present in $D$ whenever $S(i) \cap T(j) \neq \emptyset$. Assuming that the pairs of sets $(S(i), T(i)), 1 \leq i \leq n,$ are independent and identically distributed, we study the in- and outdegree distributions of a typical vertex of $D$ as $n, m \to \infty$.

key words: random intersection digraph, degree distribution, clustering, random intersection graph

1 Introduction

Given two collections of subsets $S(1), \ldots, S(n)$ and $T(1), \ldots, T(n)$ of a set $W = \{w_1, \ldots, w_m\}$, define the intersection digraph on the vertex set $V = \{v_1, \ldots, v_n\}$ such that the arc $v_i \rightarrow v_j$ is present in the digraph whenever $S(i) \cap T(j) \neq \emptyset$ for $i \neq j$. Assuming that the sets $S(i)$ and $T(i), i = 1, \ldots, n,$ are drawn at random, we obtain a random intersection digraph.

We consider the class of random intersection digraphs where the pairs of random subsets $(S(i), T(i)), i = 1, \ldots, n,$ are independent and identically distributed. In addition, we assume that the distributions of $S(i)$ and $T(i)$ are mixtures of uniform distributions. That is, for every $k$, conditionally on the event $|S(i)| = k$, the random set $S(i)$ is uniformly distributed in the class $W_k$ of all subsets of $W$ of size $k$. Similarly, conditionally on the event $|T(i)| = k$, the random set $T(i)$ is uniformly distributed in $W_k$. In particular, with $P_{S*}$ and $P_{T*}$ denoting the distributions of $|S(i)|$ and $|T(i)|$, we have that, for every $A \subset W$, $P(S(i) = A) = \binom{m}{|A|}^{-1} P_{S*}(|A|)$ and $P(T(i) = A) = \binom{m}{|A|}^{-1} P_{T*}(|A|)$. By $D(n, m, P_*)$ we denote the random intersection digraph generated by independent and identically distributed pairs of random subsets $(S(i), T(i)), 1 \leq i \leq n,$ where $P_*$ denotes the common distribution of pairs $(S(i), T(i))$.

Up to our best knowledge, the intersection digraphs with possibly infinite “ground” set $W$ were first studied by Beineke and Zamfirescu [2] and Harary, Kabell, and McMorris [10]. Since then, several tens of papers related to geometric intersection digraphs (interval digraphs, etc.) have appeared in the literature, see, e.g., [1] and references therein. Random intersection digraphs

---

1Research supported by the Research Council of Lithuania Grant MIP-10088
D(n, m, P_s) differ much from geometric intersection digraphs. We show that they are flexible enough to model random digraphs with in- and outdegrees having some desired statistical properties, such as, e.g., a power-law outdegree distribution and a bounded-support indegree distribution. Assuming, for example, that S(i) and T(i) intersect with positive probability, we can obtain a random digraph with a clustering property; see Example 2 in Section 2 below.

Note that the related random intersection graph model introduced by Karoński, Scheinerman, and Singer-Cohen [13] and Singer-Cohen [14] (see also Godzhe and Jaworski [9]), has received considerable attention in recent literature ([8], [7], [15], [12], [6], [3], etc.). The increasing number of applications of this model motivated our interest in its directed counterpart D(n, m, P_s).

The paper is organized as follows: the results are formulated in Section 2 and proved in Section 3.

2 Results

We describe conditions on P_{S, a} and P_{T, a} that make the in- and outdegrees of D(n, m, P_s) stochastically bounded and converging in distribution as n, m → ∞. Our motivation to consider the case of stochastically bounded degrees separately is that random digraphs with stochastically bounded indegrees (or outdegrees) are sparse.

We consider a sequence of random intersection digraphs D_n = D(n, m_n, P_n), where m_n → ∞ as n → ∞. We assume without loss of generality that there are two countable sets V = \{v_1, v_2, \ldots \} and W = \{w_1, w_2, \ldots \} such that, for every n, the random digraph D_n is defined on the vertex set V_n = \{v_1, \ldots , v_n\} and the ground set W_n = \{w_1, \ldots , w_{m_n}\}. Given n, let \{S_n(v), T_n(v), v \in V_n\} denote the collection of subsets of W_n that defines the intersection digraph D_n. We denote X_{n, i} = |S_n(v_i)| and Y_{n} = |T_n(v_i)|, v \in V_n. In particular, (X_{n, 1}, Y_{n, 1}), \ldots , (X_{n, n}, Y_{n, n}) are independent and identically distributed bivariate random vectors with nonnegative integer coordinates taking values in [0, m_n]^2. Given a vertex v \in V_n, let

\[ I_n(v) = \sum_{u \in V_n \setminus \{v\}} \mathbb{I}_{\{u \to v\}}, \quad O_n(v) = \sum_{u \in V_n \setminus \{v\}} \mathbb{I}_{\{v \to u\}} \]  

denote the indegree and outdegree of v in D_n that do not count the possible loop v \to v. Note that, by symmetry, the random variables I_n(v), v \in V_n have the same probability distribution. Similarly, all O_n(v), v \in V_n, have the same probability distribution.

Introduce the random variables

\[ \hat{O}_n = X_{n, 1} Y_{n, 2} + \cdots + Y_{n, m_n} m_n^{-1}, \quad \hat{I}_n = Y_{n, 1} X_{n, 2} + \cdots + X_{n, m_n} m_n^{-1}, \]  
\[ \hat{O}_n^* = X_{n, 1}^2 Y_{n, 2}^2 + \cdots + Y_{n, m_n}^2 m_n^{-2}, \quad \hat{I}_n^* = Y_{n, 1}^2 X_{n, 2}^2 + \cdots + X_{n, m_n}^2 m_n^{-2}. \]

Recall that a sequence of random variables \{Z_n\} is called stochastically bounded if, for every ε > 0, there exists B = B_ε > 0 such that P(|Z_n| ≥ B) < ε for every n. We write Z_n = o_P(1) if, in addition, P(|Z_n| ≥ δ) = o(1) as n → ∞ for every δ > 0.

**Theorem 1.** (i) Assume that the sequence \{\hat{O}_n\} is stochastically bounded. Then the sequence \{O_n(v_1)\} is stochastically bounded.

(ii) Assume that the sequence \{\hat{I}_n\} is stochastically bounded. Then the sequence \{I_n(v_1)\} is stochastically bounded.

(iii) Assume that the sequence \{\hat{O}_n\} converges in distribution to a random variable \hat{Y} and \hat{O}_n^* = o_P(1) as n → ∞. Then the sequence \{O_n(v_1)\} converges in distribution to a random variable \hat{O}_∞ with the distribution

\[ P(O_∞ = k) = (k!)^{-1} E(\hat{Y}^k e^{-\hat{Y}}), \quad k = 0, 1, 2, \ldots \]  

2
(ii') Assume that the sequence \( \{I_n\} \) converges in distribution to a random variable \( \tilde{Z} \) and \( I_n' = o_P(1) \) as \( n \to \infty \). Then the sequence \( \{I_n(v_i)\} \) converges in distribution to a random variable \( \tilde{I}_\infty \) with the distribution

\[
P(I_\infty = k) = (k!)^{-1} \mathbb{E}(\tilde{Z}^k e^{-\tilde{Z}}), \quad k = 0, 1, 2, \ldots.
\]

Note that (2) and (3) are mixed Poisson distributions (i.e., Poisson distributions \( P(\lambda) \) with random parameters \( \lambda = \tilde{Y} \) and \( \lambda = \tilde{Z} \), respectively).

Example 1. Let \( X_{n1} \equiv 1 \), and let \( m_n^{-1}Y_{nk} \) have Bernoulli distribution with success probability \( n^{-1} \) for every \( 1 \leq k \leq n \). Then \( \tilde{O}_n = \tilde{O}_n' \) and \( O_n(v) \) have binomial distribution \( Bi(n-1, n^{-1}) \). Therefore, \( \tilde{O}_n' \neq o_P(1) \), and the limiting distribution of the outdegree sequence \( \{O_n(v)\} \) is the Poisson distribution with mean 1, which now differs from (2).

The example suggests that the convergence in distribution of \( \{\tilde{O}_n\} \) alone is not sufficient for the convergence of distributions of \( O_n(v) \) to the distribution (2).

Remark 1. The statements (i), (i') of Theorem 1 hold for the indegrees \( I_n'(v) = I_n(v) + \mathbb{I}_{\{v\to v\}} \) and the outdegrees \( O_n'(v) = O_n(v) + \mathbb{I}_{\{v\to v\}} \), which now count the possible loop \( v \to v \). Since the probability \( P(v \to v) \) may not vanish as \( n \to \infty \), in order to obtain limit theorems for distributions of \( I_n'(v) \) and \( O_n'(v) \), one needs to impose an extra conditions on the joint distribution \( P_n \) of the pair of random sets \( (S_n(v), T_n(v)) \). For example, if \( S_n(v) \) and \( T_n(v) \) are independent, then the results (ii), (ii') of Theorem 1 and that of Corollary 1 below extend to the sequences \( \{I_n'(v)\} \) and \( \{O_n'(v)\} \).

In the remaining part of this section, we consider the important particular case where \( n = O(m_n) \) and the random sets \( S_n(v) \) and \( T_n(v) \) are of the same scale. In particular, we assume that \( n \leq cm_n \) for some absolute constant \( c > 0 \) and that \( |S_n(v)|, |T_n(v)| = O_P(\sqrt{m_n/n}) \).

The next result is formulated for the outdegree sequence \( \{O_n(v)\} \) only. Obviously, the analogous result holds for the indegree sequence \( \{I_n(v)\} \) as well.

Corollary 1. Assume that

(i) \( \{X_{n1}(n/m_n)^{1/2}\} \) converges in distribution to a random variable \( X_\infty \);

(ii) \( \{Y_{n1}(n/m_n)^{1/2}\} \) converges in distribution to a random variable \( Y_\infty \);

(iii) \( EY_\infty < \infty \), and \( \lim_n EY_{n1}(n/m_n)^{1/2} = EY_\infty \).

Then \( \{O_n(v)\} \) converges in distribution to \( O_\infty \); see (2), where \( \tilde{Y} = X_\infty EY_\infty \).

Vertex degree distribution of a random intersection graph was studied in [15], [12], [6], and [4]. Theorem 1 and Corollary 1 extend related results of these papers to the random intersection digraph \( D(n, m, P_s) \). Let us mention that our proof differs from those of [15], [12], [6], and [4] and leads to more general and precise results.

Next, we give an example of random intersection digraph with a clustering property.

Example 2 (cf. [6]). Fix \( a > 0 \) and let \( m = \lfloor an \rfloor \). Let \( X_1, X_2, X_3 \geq 0 \) be independent integer-valued random variables with finite first moments \( \alpha_i = EX_i \), and let \( Z_1, Z_2, Z_3 \) be independent random subsets of \( W = \{w_1, \ldots, w_m\} \) such that, given \( X_i \), the random set \( Z_i \) is uniformly distributed in the class of subsets of \( W \) of size \( X_i \cap m, 1 \leq i \leq 3 \). Here we denote \( x \land y = \min\{x, y\} \). Put \( S = Z_1 \cup Z_3 \) and \( T = Z_2 \cup Z_3 \), and let \( D_n \) be the random intersection digraph defined by the sequence \( \{(S(v_i), T(v_i))\}, 1 \leq i \leq n \) of independent copies of \( (S, T) \). Note that, by Corollary 1, the in- and outdegree distributions of \( D_n \) converge to nondegenerate limits, provided that \( \alpha_1 + \alpha_3, \alpha_2 + \alpha_3 > 0 \). If, in addition, \( \alpha_3 > 0 \) (i.e., \( \lim_n P(S \cap T \neq \emptyset) > 0 \)) and the second moments \( \beta_i = EX_i^2 \) are finite, then the conditional probabilities of a triangle, given any two of its sides, are positive and bounded away from zero as \( n \to \infty \). In particular, we have, as
\( n \to \infty, \)
\[
p_{13|12,23} = \alpha_3(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3 + \beta_3)^{-1} + o(1),
\]
\[
p_{31|12,23} = \alpha_3(\alpha_1 + \alpha_3)^{-1}(\alpha_2 + \alpha_3)^{-1}(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3 + \beta_3)^{-1} + o(1),
\]
\[
p_{13|12,32} = \alpha_3(\alpha_2 + \alpha_3)(\alpha_1 + \alpha_3)^{-1}(\beta_2 + 2\alpha_2\alpha_3 + \beta_3)^{-1} + o(1),
\]
\[
p_{13|21,23} = \alpha_3(\alpha_1 + \alpha_3)(\alpha_2 + \alpha_3)^{-1}(\beta_1 + 2\alpha_1\alpha_3 + \beta_3)^{-1} + o(1).
\]
Here \( p_{ij|st,fg} = P(v_i \to v_j | v_s \to v_t, v_f \to v_g) \) denotes the conditional probability of the arc \( v_i \to v_j \), given the event that arcs \( v_s \to v_t \) and \( v_f \to v_g \) are present in \( D_n \).

### 3 Proof

We start with auxiliary Lemma 1. Then we prove Theorem 1 and Corollary 1. Relations (4) are shown at the end of the section.

**Lemma 1.** Let \( S_1, S_2 \) be independent random subsets of the set \( W = \{1, \ldots, m\} \) such that \( S_1 \) (respectively \( S_2 \)) is uniformly distributed in the class of subsets of \( W \) of size \( j \) (respectively \( k \)). Then the probability \( p' := P(S_1 \cap S_2 = \emptyset) = (m-k)/m \) satisfies, for \( j+k < m \),
\[
1 - \frac{j+k}{1 - (j+k)/m} \leq p' \leq 1 - \frac{j+k}{m} + \left(\frac{j+k}{m}\right)^2.
\]
Here we denote \( (m)_j = m(m-1) \cdots (m-j+1) \). For \( 0 < \alpha < 1 \) and \( j+k \leq \alpha m \), we have
\[
\frac{j+k}{m} + \frac{2}{1-\alpha}\left(\frac{j+k}{m}\right)^2 \geq P(S_1 \cap S_2 = \emptyset) \geq \frac{j+k}{m} - \left(\frac{j+k}{m}\right)^2.
\]
We also have
\[
P(|S_1 \cap S_2| \geq 2) \leq 2^{-1}(jk)^2m^{-2}.
\]

**Proof of Lemma 1.** Inequalities (5) and (6) are shown in [12]; see also [4]. Let us prove (7). Write \( S_1 = \{u_1, \ldots, u_j\} \) and let \( i_1 = \min\{i : u_i \in S_1 \cap S_2\} \) and \( i_2 = \min\{i > i_1 : u_i \in S_1 \cap S_2\} \). We have \( P(|S_1 \cap S_2| \geq 2) = \sum_{1 \leq s < \ell \leq j} P(i_1 = s, i_2 = \ell) \). Invoking the identity \( P(i_1 = s, i_2 = \ell) = \frac{(m-k)(m-2k)}{(m)_j^2} \), we obtain (7). \( \square \)

**Proof of Theorem 1.** We prove (i) and (ii). The proof of (i') and (ii') is much the same. Write, for short, \( O_n = O_n(v_1) \). Given a vector \( \vec{x} = (x, y_2, \ldots, y_n) \) with integer coordinates, we denote
\[
\lambda(\vec{x}) = (x/m_n) \sum_{2 \leq k \leq n} y_k, \quad \kappa(\vec{x}) = (x/m_n)^2 \sum_{2 \leq k \leq n} y_k^2.
\]
Let us prove (i). Split \( O_n = \xi_n + \eta_n \), where
\[
\xi_n := \sum_{2 \leq i \leq n} I_{\{v_i \to v_j\}} I_{\{4X_nY_{ni} \leq m_n\}}, \quad \eta_n := \sum_{2 \leq i \leq n} I_{\{v_i \to v_j\}} I_{\{4X_nY_{ni} > m_n\}}.
\]
We shall show that both sequences \( \{\xi_n\} \) and \( \{\eta_n\} \) are stochastically bounded. Fix \( 0 < \varepsilon < 1 \). Choose a (nonrandom) number \( B_\varepsilon > \varepsilon^{-1} \) such that the event \( \mathcal{H}_n := \{\tilde{O}_n < B_\varepsilon\} \) has the probability \( P(\mathcal{H}_n) \geq 1 - \varepsilon \) for every \( n \). The sequence \( \{\eta_n\} \) is stochastically bounded because of the inequality
\[
P(\eta_n \geq 4B_\varepsilon) < \varepsilon \Leftrightarrow P(|\{\eta_n \geq 4B_\varepsilon\} \cap \mathcal{H}_n) = \varepsilon.
\]
Indeed, we have $P(\{\eta_n \geq 4B_\varepsilon\} \cap \mathcal{H}_n) = 0$, since on the event $\mathcal{H}_n$, the number of summands $I_{\{4X_{n1}Y_{ni} > m_n\}}$ taking value 1 is less than $4B_\varepsilon$.

In order to show that $\{\xi_n\}$ is stochastically bounded, we use the fact that, given

$$X_{n1} = x, \quad Y_{n2} = y_2, \ldots, \quad Y_{nn} = y_n,$$

the random variable $\xi_n$ is the sum $\tau_2 + \cdots + \tau_n$ of (conditionally) independent Bernoulli random variables $\tau_i$ with success probabilities

$$p_i = P\left( S_n(v_i) \cap T_n(v_i) \neq \emptyset \bigg| X_{n1} = x, Y_{ni} = y_i \right) \mathbb{I}_{\{4xy\leq m_n\}}.$$ 

In addition, we have

$$p_i \leq (2/m_n)xy_\varepsilon \mathbb{I}_{\{4xy\leq m_n\}}.$$ \hspace{1cm} (9)

For $xy = 0$, inequality (9) is trivial. Indeed, in this case, at least one of the sets $S_n(v_1), T_n(v_1)$ is empty. For $x, y \geq 1$ satisfying $4xy \leq m_n$, we have $x + y \leq 2xy \leq 2^{-1}m_n$. Now (9) follows from (6).

Let $P_{x,y}$ and $E_{x,y}$ denote the conditional distribution and the conditional expectation given the event (8). In view of (9), we have on the event $\mathcal{H}_n$ that

$$E_{x,y}\xi_n = \sum_{2 \leq i \leq n} p_i \leq 2B_\varepsilon.$$

In addition, by Chebyshev’s inequality,

$$P_{x,y}(\xi_n > 3B_\varepsilon) = P_{x,y}(\xi_n - E_{x,y}\xi_n > 3B_\varepsilon - E_{x,y}\xi_n) \leq \sum_{2 \leq i \leq n} p_i B_\varepsilon^{-2} \leq 2B_\varepsilon^{-1} < 2\varepsilon.$$ \hspace{1cm} (10)

Finally, we obtain

$$P(\xi_n > 3B_\varepsilon) < \varepsilon + P(\{\xi_n > 3B_\varepsilon\} \cap \mathcal{H}_n) = \varepsilon + E\left( I_{\mathcal{H}_n} P(\xi_n > 3B_\varepsilon \bigg| X_{n1}, Y_{n2}, \ldots, Y_{nn}) \right) \leq \varepsilon + E(I_{\mathcal{H}_n} 2\varepsilon) \leq 3\varepsilon.$$

Here, in the first step, we invoke the inequality $P(\mathcal{H}_n) \geq 1 - \varepsilon$. In the second step, we apply bound (10) to the conditional probability $P\left( \xi_n > 3B_\varepsilon \bigg| X_{n1}, Y_{n2}, \ldots, Y_{nn} \right)$ of the event $\{\xi_n > 3B_\varepsilon\}$ given $X_{n1}, Y_{n2}, \ldots, Y_{nn}$. Hence, the sequence $\{\xi_n\}$ is stochastically bounded. The proof of statement (i) is complete.

Let us prove (ii). Let $f_n(t) = E e^{it\tilde{O}_n}$ and $f_\infty(t) = E e^{it\tilde{O}_\infty}$ denote the Fourier transforms of the probability distributions of $\tilde{O}_n$ and $\tilde{O}_\infty$. In order to prove (ii), we show that $\lim_n f_n(t) = f_\infty(t)$ for every real $t$.

Given $0 < \delta < 0.01$ and integer $n$, introduce the event $\mathcal{A}_n = \{\tilde{O}_n^* < \delta\}$. Note that $P(\mathcal{A}_n) = 1 - o(1)$ as $n \to \infty$. Therefore, we have

$$f_n(t) = E(e^{it\tilde{O}_n} \mathbb{I}_{\mathcal{A}_n}) + o(1).$$ \hspace{1cm} (11)
On the event \( A_n \), we approximate the conditional characteristic function
\[
f_n(t; \overline{x} \overline{y}) := \mathbb{E}(e^{itO_n} | X_{n1} = x, Y_{n2} = y_2, \ldots, Y_{nn} = y_n)
\]
by the Fourier transform of the Poisson distribution with mean \( \lambda(\overline{x} \overline{y}) \),
\[
g_n(t; \overline{x} \overline{y}) = \exp\{\lambda(\overline{x} \overline{y})(e^{it} - 1)\}.
\]
Since the conditional distribution of \( O_n \), given (8), is that of the sum of independent Bernoulli random variables with success probabilities
\[
q_k = P\left(S_n(v_1) \cap T_n(v_k) \neq \emptyset \mid X_{n1} = x, Y_{nk} = y_k\right), \quad 2 \leq k \leq n,
\]
we write
\[
f_n(t; \overline{x} \overline{y}) = \prod_{2 \leq k \leq n} (1 + q_k(e^{it} - 1)) = \exp\left\{ \sum_{2 \leq k \leq n} \ln(1 + q_k(e^{it} - 1)) \right\}.
\]
Note that, on the event \( A_n \), we have \( x^2y_k^2/m_n^2 \leq \delta \leq 0.01 \). Therefore, (6) implies
\[
\frac{x^2y_k^2}{m_n} - \frac{(x^2y_k^2)}{m_n} \leq q_k \leq \frac{x^2y_k^2}{m_n} + 4\left(\frac{x^2y_k^2}{m_n}\right)^2, \quad 2 \leq k \leq n.
\]
In particular, for each \( k \), we have \( |q_k(e^{it} - 1)| < 0.5 \). Invoking the inequality \( |\ln(1+z) - z| \leq |z|^2 \) for complex numbers \( z \) satisfying \( |z| \leq 0.5 \) (see, e.g., Proposition 8.46 of [5]), we obtain
\[
f_n(t; \overline{x} \overline{y}) = \exp\{\lambda(\overline{x} \overline{y})(e^{it} - 1) + r(t)\},
\]
where \( |r(t)| \leq \kappa(\overline{x} \overline{y}) \). Here and below, \( \kappa \) and \( c' \) denote absolute constants. Now, the bound \( \kappa(\overline{x} \overline{y}) < \delta \), which holds on the event \( A_n \), implies that \( |f_n(t; \overline{x} \overline{y}) - g_n(t; \overline{x} \overline{y})| \leq c' \delta \). Invoking this inequality in (11), we obtain
\[
|f_n(t) - \mathbb{E}\exp\{\hat{O}_n(e^{it} - 1)\}| \leq c' \delta + o(1) \quad \text{as} \quad n \to \infty.
\]
Finally, the convergence of distributions of \( \hat{O}_n \) implies the convergence of corresponding expectations of bounded continuous functions. Therefore, \( \lim_n \mathbb{E}e^{O_n(e^{it} - 1)} = f_\infty(t) \). We obtain the inequality \( \lim_n |f_n(t) - f_\infty(t)| \leq c' \delta \), which holds for arbitrarily small \( \delta > 0 \). The proof of (ii) is complete. \( \square \)

**Proof of Corollary 1.** Denote \( \mathbb{E}Y_\infty = a \). For \( \varepsilon > 0 \) and \( 1 \leq k \leq n \), denote
\[
Z_{nk} = Y_{nk}(n/m_n)^{1/2}, \quad Z_{nkek} = Z_{nk}1_{Z_{nk} \leq n\varepsilon}, \quad a_n = \mathbb{E}Z_{n1}, \quad a_{n\varepsilon} = \mathbb{E}Z_{n1\varepsilon}.
\]
In view of statement (ii) of Theorem 1 and the identities
\[
\hat{O}_n = X_{n1}(n/m_n)^{1/2}(Z_{n2} + \cdots + Z_{nn})n^{-1}, \quad \hat{O}_n^* = X_{n1}^2(n/m_n)(Z_{n2}^2 + \cdots + Z_{nn}^2)n^{-2},
\]
the corollary would follow if we show that
\[
(Z_{n2} + \cdots + Z_{nn})n^{-1} - a_n = o_P(1); \quad (Z_{n2}^2 + \cdots + Z_{nn}^2)n^{-2} = o_P(1).
\]
The proof of (13) and (14) is obtained by a routine application of the truncation argument. In the proof, we also use the observation that conditions (ii) and (iii) imply the uniform integrability of the sequence of random variables \( \{Z_{n1}\} \); see, e.g., [4]. That is,
\[
\forall \varepsilon > 0, \exists \Delta > 0 \quad \text{such that} \quad \forall n \geq 1, \quad \text{we have} \quad \mathbb{E}(Z_{n1}1_{Z_{n1} > \Delta}) < \varepsilon.
\]
In order to prove (13), we shall show that, for every $0 < \delta < 1$,
\[
\lim_{n} P\left(\left|Z_{n2} + \cdots + Z_{nn}(n-1)^{-1} - a_n\right| > 2\delta\right) \leq a\delta.
\] (16)

Fix $0 < \delta < 1$ and put $\varepsilon = \delta^3$. Introduce the events
\[
B_n = \left\{\left|Z_{n2} + \cdots + Z_{nn}(n-1)^{-1} - a_n\right| > \delta\right\}, \quad D_n = \left\{\max_{2 \leq k \leq n} Z_{nk} \leq n\varepsilon\right\}.
\]

It follows from (15) that, as $n \to \infty$,
\[
0 \leq a_n - a_n\varepsilon = E\left(Z_{n1}\mathbb{1}_{\{Z_{1n} > \varepsilon\}}\right) = o(1),
\] (17)
\[
1 - P(D_n) \leq (n-1)P(Z_{nk} > n\varepsilon) \leq \varepsilon^{-1}E\left(Z_{n1}\mathbb{1}_{\{Z_{1n} > \varepsilon\}}\right) = o(1).
\] (18)

In view of (17) and (18), we can replace $a_n$ by $a_n\varepsilon$ and $Z_{nk}$ by $Z_{nk\varepsilon}$ in (16). In particular, (16) follows from the inequality
\[
\lim_{n} P\left(B_n\right) \leq a\delta.
\] (19)

Let us prove (19). By Chebyshev’s inequality and symmetry,
\[
P\left(B_n\right) \leq \delta^{-2}(n-1)^{-1}E\left(Z_{n1\varepsilon} - a_n\varepsilon\right)^2.
\]

Invoking the simple inequalities
\[
E\left(Z_{n1\varepsilon} - a_n\varepsilon\right)^2 \leq EZ_{n1\varepsilon}^2 \leq n\varepsilon EZ_{n1\varepsilon} \leq n\varepsilon a_n,
\]
we obtain $P\left(B_n\right) \leq (n/(n-1))a_n\delta$. This inequality implies (19). The proof of (13) is complete.

Let us prove (14). For this purpose, given $\varepsilon > 0$, we show that
\[
\lim_{n} P\left(Z_{n2}^2 + \cdots + Z_{nn}^2 > \varepsilon n^2\right) \leq a\varepsilon.
\] (20)

Introduce the events $K_n = \{\max_{2 \leq k \leq n} Z_{nk} \leq \varepsilon^2 n\}$. By symmetry and Markov’s inequality, we obtain from (15) that
\[
1 - P(K_n) \leq (n-1)P(Z_{nk} > \varepsilon^2 n) \leq \varepsilon^{-2}E\left(Z_{n1}\mathbb{1}_{\{Z_{1n} > \varepsilon^2 n\}}\right) = o(1)
\] (21)
as $n \to \infty$. Similarly, by symmetry and Markov’s inequality, we have
\[
P\left(\{Z_{n2}^2 + \cdots + Z_{nn}^2 > \varepsilon^2 n\} \cap K_n\right) \leq P\left(Z_{n2} + \cdots + Z_{nn} > n\varepsilon^{-1}\right) \leq \varepsilon E\left(Z_{n2}\right).
\]

This inequality in combination with (21) shows (20). We have arrived at (14).

**Proof of (4).** We sketch the proof of the first identity of (4). The remaining identities are obtained in much the same way.

Given $v \in V_n = \{v_1, \ldots, v_n\}$, let $Z_1(v), Z_2(v), Z_3(v)$ denote the random sets that define the pair $(S(v), T(v))$, i.e., $S(v) = Z_1(v) \cup Z_3(v)$ and $T(v) = Z_2(v) \cup Z_3(v)$. We write $X_{ij} = |Z_i(v_j)|$ and $\bar{X} = \{X_{ij}, 1 \leq i, j \leq 3\}$. Given $\bar{x} = \{x_{ij}, 1 \leq i, j \leq 3\}$, we use the shorthand notation for the conditional probability $P_{\bar{x}}(\cdot) := P(\cdot|\bar{X} = \bar{x})$. The conditional probabilities of the events
\[
I_1 = \{v_1 \to v_2, v_2 \to v_3\}, \quad I_2 = \{v_1 \to v_2, v_2 \to v_3, v_1 \to v_3\}
\]
are denoted $p_k(\bar{x}) = P_{\bar{x}}(I_k), k = 1, 2$. We shall show that, as $m \to \infty$,
\[
P_{\bar{x}}(\bar{X}) = m^{-2}E_{\bar{x}}(\bar{X}) + o(m^{-2}), \quad k = 1, 2.
\] (22)
Here \( a_1(\bar{x}) = (x_{11} + x_{31})(x_{22} + x_{32})(x_{12} + x_{32})(x_{23} + x_{33}) \) and \( a_2(\bar{x}) = (x_{11} + x_{31})x_{32}(x_{23} + x_{33}) \). Observe that substitution of (22) into the identity \( p_{13|12,23} = E_{P_2(X)} / E_{P_1(X)} \) gives (4).

In order to prove (22), we show that, for every \( c \geq 1 \), we have, uniformly in \( \bar{x} \subset [0,c] \),

\[
p_k(\bar{x}) = m^{-2}a_k(\bar{x}) + o(m^{-2}), \quad k = 1, 2.
\]

(23)

Clearly, for bounded random variables \( X_{ij} \in \bar{X} \), (22) is an immediate consequence of (23). If \( X_{ij} \) are not bounded but have finite second moments, we can safely replace \( X_{ij} \) by the truncated random variables \( X_{ij} \wedge (m/4) \) and then apply (23) to the truncated random variables. We omit the details.

Let us prove (23). Let \( A_i \) denote the event that \( Z_1(v_i), Z_2(v_i), Z_3(v_i) \) are pairwise disjoint. Write \( A = A_1 \cap A_2 \cap A_3 \) and let \( \bar{A} \) denote the event complement to \( A \). Denoting the conditional probabilities \( p_k^i(\bar{x}) = P_{\bar{x}}(I_k|A) \) and \( p_k^i(\bar{x}) = P_{\bar{x}}(I_k|\bar{A}) \), we write

\[
p_k(\bar{x}) = p_k^i(\bar{x}) + (p_k^{ii}(\bar{x}) - p_k^i(\bar{x}))P_{\bar{x}}(\bar{A}).
\]

(24)

Note that (6) implies \( P_{\bar{x}}(\bar{A}) \sim O(m^{-1}) \). Now (23) follows from (24) and from the bounds

\[
p_k^i(\bar{x}) = m^{-2}a_k(\bar{x}) + O(m^{-3}), \quad p_k^{ii}(\bar{x}) \leq m^{-2}a_k(\bar{x}) + O(m^{-3}).
\]

(25)

In order to prove (25), for \( k = 1 \), we apply (6) to the pairs of random sets \( S(v_1), T(v_2) \) and \( S(v_2), T(v_3) \). In particular, the inequalities \( |S(v_1)| \leq x_{11} + x_{31} \) and \( |T(v_1)| \leq x_{21} + x_{31} \) imply the second bound of (25). Similarly, the identities

\[
|S(v_1)| = x_{11} + x_{31}, \quad |T(v_1)| = x_{21} + x_{31},
\]

(26)

which hold on the event \( A \), imply the first bound of (25).

Let us prove (25) for \( k = 2 \). We only prove the first bound. The proof of the second bound is much the same. Denote \( H = S(v_2) \cap T(v_2) \), \( U = S(v_2) \cap T(v_2) \) and introduce the events

\[
C' = \{|H \cap S(v_1)| = 1\}, \quad C'' = \{|H \cap S(v_1)| \geq 2\},
\]

\[
D' = \{|H \cap T(v_3)| = 1\}, \quad D'' = \{|H \cap T(v_3)| \geq 2\},
\]

\[
E = \{S(v_1) \cap T(v_3) \cap (W \setminus H) = \emptyset\}.
\]

Write \( C = \{S(v_1) \cap H \neq \emptyset\} = C' \cup C'' \) and \( D = \{T(v_3) \cap H \neq \emptyset\} = D' \cup D'' \), and observe that \( I_2 = I_2 \cap C \cap D \). We are going to replace \( p_2(\bar{x}) = P_{\bar{x}}(I_2, \bar{A}) \) by \( p_2^{i}(\bar{x}) = P_{\bar{x}}(I_2 \cap H, \bar{A}) \), where \( H = C' \cap D' \cap E \). For this purpose, we first split

\[
C \cap D = (C \cap D'^*) \cup (C' \cap D') \cup (C'' \cap D'^*)
\]

and replace \( p_2^{i}(\bar{x}) \) by \( p_2^{ii}(\bar{x}) = P_{\bar{x}}(I_2 \cap C' \cap D', \bar{A}) \). The error of this replacement

\[
0 \leq p_2^{ii}(\bar{x}) - p_2^{ii}(\bar{x}) \leq \bar{p}_1 + \bar{p}_2,
\]

where \( \bar{p}_1 = P_{\bar{x}}(C \cap D'^*, A) \) and \( \bar{p}_2 = P_{\bar{x}}(C'' \cap D'^*, A) \). Secondly, we replace \( p_2^{ii}(\bar{x}) \) by \( p_3(\bar{x}) \). We have

\[
0 \leq p_2^{ii}(\bar{x}) - p_3(\bar{x}) \leq \bar{p}_3, \quad \bar{p}_3 = P_{\bar{x}}(C' \cap D' \cap E, \bar{A}).
\]

Here \( E \) denotes the event complement to \( E \). Invoking the simple bounds, which follow from Lemma 1,

\[
\bar{p}_1, \bar{p}_2, \bar{p}_3 = O(m^{-3}),
\]

(27)

we obtain \( p_2^{ii}(\bar{x}) = p_2^{ii}(\bar{x}) + O(m^{-3}) \). Observe that the event \( I_2 \cap H \) holds whenever

\[
|S(v_1) \cap T(v_3) \cap U| = 1, \quad (S(v_1) \cap T(v_3)) \cap (W_n \setminus U) = \emptyset.
\]

(28)

Finally, it follows from Lemma 1 that the event (28) has the probability

\[
|S(v_1)| \times |T(v_3)| \times |U|/m^2 + O(m^{-3}) = a_2(\bar{x}) + O(m^{-3}).
\]

In the last step, we invoke (26) and the identity \( |U| = x_{32} \).
Acknowledgement. I would like to thank anonymous referees for their valuable comments and suggestions.

References


