Degree distribution of an inhomogeneous random intersection graph

Mindaugas Bloznelis*
Faculty of mathematics and informatics
Vilnius University
Vilnius, Lithuania
mindaugas.bloznelis@mif.vu.lt

Julius Damarackas
Faculty of mathematics and informatics
Vilnius University
Vilnius, Lithuania
julius.damarackas@mif.stud.vu.lt

Submitted: Oct 8, 2012; Accepted: Jul 11, 2013; Published: Jul 19, 2013
Mathematics Subject Classifications: 05C80, 05C07, 05C82

Abstract

We show the asymptotic degree distribution of the typical vertex of a sparse inhomogeneous random intersection graph.

Keywords: degree distribution; random graph; random intersection graph; power law

1 Introduction

Let $X_1, \dots, X_m, Y_1, \dots, Y_n$ be independent non-negative random variables such that each X_i has the probability distribution P_1 and each Y_j has the probability distribution P_2. Given realized values $X = \{X_i\}_{i=1}^m$ and $Y = \{Y_j\}_{j=1}^n$ we define the random bipartite graph $H_{X,Y}$ with the bipartition $V = \{v_1, \dots, v_n\}$, $W = \{w_1, \dots, w_m\}$, where edges $\{w_i, v_j\}$ are inserted with probabilities $p_{ij} = \min\{1, X_i Y_j (nm)^{-1/2}\}$ independently for each $\{i, j\} \in [m] \times [n]$. The *inhomogeneous* random intersection graph $G(P_1, P_2, n, m)$ defines the adjacency relation on the vertex set V: vertices $v', v'' \in V$ are declared adjacent (denoted $v' \sim v''$) whenever v' and v'' have a common neighbour in $H_{X,Y}$.

The degree distribution of the typical vertex of the random graph $G(P_1, P_2, n, m)$ has been first considered by Shang [21]. The proof of the main result of [21] contains gaps and the result is incorrect in the regime where $m/n \to \beta \in (0, +\infty)$ as $m, n \to +\infty$. We remark that this regime is of particular importance, because it leads to inhomogeneous graphs with the clustering property: the clustering coefficient $\mathbf{P}(v_1 \sim v_2 | v_1 \sim v_3, v_2 \sim v_3)$ is bounded away from zero provided that $0 < \beta < +\infty$ and $\mathbf{E}X_1^3 < \infty$, $\mathbf{E}Y_1^2 < \infty$, see [8].

*Supported by the Research Council of Lithuania Grant MIP-053/2011.

The aim of the present paper is to show the asymptotic degree distribution in the case where $m/n \to \beta$ for some $\beta \in [0, +\infty]$.

We consider a sequence of graphs $\{G_n = G(P_1, P_2, n, m)\}$, where $m = m_n \to +\infty$ as $n \to \infty$, and where P_1, P_2 do not depend on n. We denote $a_k = \mathbf{E}X_1^k$, $b_k = \mathbf{E}Y_1^k$. By $d(v) = d_{G_n}(v)$ we denote the degree of a vertex v in G_n (the number of vertices adjacent to v in G_n). We remark that for every n the random variables $d_{G_n}(v_1), \dots, d_{G_n}(v_n)$ are identically distributed. In Theorem 1 below we show the asymptotic distribution of $d(v_1)$.

Theorem 1. *Let $m, n \to \infty$.*

(i) Assume that $m/n \to 0$. Suppose that that $\mathbf{E}X_1 < \infty$. Then $\mathbf{P}(d(v_1) = 0) = 1 - o(1)$.

(ii) Assume that $m/n \to \beta$ for some $\beta \in (0, +\infty)$. Suppose that $\mathbf{E}X_1^2 < \infty$ and $\mathbf{E}Y_1 < \infty$. Then $d(v_1)$ converges in distribution to the random variable

$$d_* = \sum_{j=1}^{\Lambda_1} \tau_j, \tag{1}$$

where $\tau_1, \tau_2, \dots$ are independent and identically distributed random variables independent of the random variable Λ_1. They are distributed as follows. For $r = 0, 1, 2, \dots$, we have

$$\mathbf{P}(\tau_1 = r) = \frac{r+1}{\mathbf{E}\Lambda_2} \mathbf{P}(\Lambda_2 = r+1) \qquad \text{and} \qquad \mathbf{P}(\Lambda_i = r) = \mathbf{E}\, e^{-\lambda_i} \frac{\lambda_i^r}{r!}, \qquad i = 1, 2. \tag{2}$$

Here $\lambda_1 = Y_1 a_1 \beta^{1/2}$ and $\lambda_2 = X_1 b_1 \beta^{-1/2}$.

(iii) Assume that $m/n \to +\infty$. Suppose that $\mathbf{E}X_1^2 < \infty$ and $\mathbf{E}Y_1 < \infty$. Then $d(v_1)$ converges in distribution to a random variable Λ_3 having the probability distribution

$$\mathbf{P}(\Lambda_3 = r) = \mathbf{E}e^{-\lambda_3} \frac{\lambda_3^r}{r!}, \qquad r = 0, 1, \dots. \tag{3}$$

Here $\lambda_3 = Y_1 a_2 b_1$.

Remark 2. The probability distributions P_{Λ_i} of Λ_i, $i = 1, 2, 3$, are Poisson mixtures. One way to sample from the distribution P_{Λ_i} is to generate random variable λ_i and then, given λ_i, to generate Poisson random variable with the parameter λ_i. The realized value of the Poisson random variable has the distribution P_{Λ_i}.

Remark 3. The asymptotic distributions (1) and (3) admit heavy tails. In the case (ii) we obtain a power law asymptotic degree distribution (1) provided that the heavier of the tails $t \to \mathbf{P}(\tau_1 > t)$ and $t \to \mathbf{P}(\Lambda_1 > t)$ has a power law, see, e.g., [13]. Similarly, in the case (iii) we obtain a power law asymptotic degree distribution (3) provided that P_2 has a power law.

Remark 4. Since the second moment a_2 does not show up in (1), (2) we expect that in the case (ii) the second moment condition $\mathbf{E}X_1^2 < \infty$ is redundant and could perhaps be replaced by the weaker first moment condition $\mathbf{E}X_1 < \infty$.

Random intersection graphs have attracted considerable attention in the recent literature, see, e.g., [1, 2, 3, 9, 10, 12, 19]. Starting with the paper by Karoński et al [17], see also [22], where the case of degenerate distributions $P_1 = P_{1n}$, $P_2 = P_{2n}$ depending on n was considered (i.e., $\mathbf{P}_{1n}(c_n) = \mathbf{P}_{2n}(c_n) = 1$, for some $c_n > 0$), several more complex random intersection graph models were later introduced by Godehardt and Jaworski [14], Spirakis et al. [18], Shang [21]. The asymptotic degree distribution for various random intersection graph models was shown in [4, 5, 6, 7, 11, 15, 16, 20, 23].

2 Proof

Before the proof of Theorem 1 we introduce some notation and give two auxiliary lemmas.

The event that the edge $\{w_i, v_j\}$ is present in $H = H_{X,Y}$ is denoted $w_i \to v_j$. We denote

$$\mathbb{I}_{ij} = \mathbb{I}_{\{w_i \to v_j\}}, \qquad \mathbb{I}_i = \mathbb{I}_{i1}, \qquad u_i = \sum_{2 \leqslant j \leqslant n} \mathbb{I}_{ij}, \qquad L = \sum_{i=1}^{m} u_i \mathbb{I}_i$$

and remark that u_i counts all neighbours of w_i in H belonging to the set $V \setminus \{v_1\}$. Denote

$$\hat{a}_k = m^{-1} \sum_{1 \leqslant i \leqslant m} X_i^k, \qquad \hat{b}_k = n^{-1} \sum_{2 \leqslant j \leqslant n} Y_j^k,$$

$$\lambda_{ij} = \frac{X_i Y_j}{\sqrt{mn}}, \qquad S_{XY} = \sum_{i=1}^{m} \sum_{j=2}^{n} \frac{X_i}{\sqrt{mn}} \lambda_{ij} \min\{1, \lambda_{ij}\}, \tag{4}$$

and introduce the event $\mathcal{A}_1 = \{\lambda_{i1} < 1,\ 1 \leqslant i \leqslant m\}$. We denote $\beta_n = m/n$. By $\mathbf{P}_1$ and $\mathbf{E}_1$ we denote the conditional probability and conditional expectation given Y_1. By $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{E}}$ we denote the conditional probability and conditional expectation given X, Y. By $d_{TV}(\zeta, \xi)$ we denote the total variantion distance between the probability distributions of random variables ζ and ξ. In the case where ζ, ξ, X, Y are defined on the same probability space we denote by $\tilde{d}_{TV}(\zeta, \xi)$ the total variation distance between the conditional distributions of ζ and ξ given X, Y.

In the proof below we use the following simple fact about the convergence of a sequence of random variables $\{\varkappa_n\}$:

$$\varkappa_n = o_P(1), \quad \mathbf{E} \sup_n \varkappa_n < \infty \quad \Rightarrow \quad \mathbf{E}\varkappa_n = o(1). \tag{5}$$

We remark that the condition $\mathbf{E} \sup_n \varkappa_n < \infty$ (which assumes implicitly that all $\varkappa_n$ are defined on the same probability space) can be replaced by a more restrictive condition that there exists a constant $c > 0$ such that $|\varkappa_n| \leqslant c$, for all n. In the latter case we do not need all $\varkappa_n$ to be defined on the same probability space. In particular, given a sequence of bivariate random vectors $\{(\eta_n, \theta_n)\}$ such that, for every n and $m = m_n$ random variables η_n, θ_n and $\{X_i\}_{i=1}^m$, $\{Y_j\}_{j=1}^n$ are defined on the same probability space, we have

$$\tilde{d}_{TV}(\eta_n, \theta_n) = o_P(1) \ \Rightarrow \ d_{TV}(\eta_n, \theta_n) \leqslant \mathbf{E}\tilde{d}_{TV}(\eta_n, \theta_n) = o(1).$$

Here and below all limits are taken as $n \to +\infty$ (if not stated otherwise).

Lemma 5. *Assume that* $\mathbf{E}X_1^2 < \infty$ *and* $\mathbf{E}Y_1 < \infty$. *We have as* $n \to +\infty$

$$\mathbf{P}(d(v_1) \neq L) = o(1), \tag{6}$$

$$\mathbf{P}(\mathcal{A}_1) = 1 - o(1), \tag{7}$$

$$S_{XY} = o_P(1), \qquad \mathbf{E}S_{XY} = o(1). \tag{8}$$

Proof. Proof of (6). We observe that the event $\mathcal{A} = \{d(v_1) \neq L\}$ occurs in the case where for some $2 \leqslant j \leqslant n$ and some distinct $i_1, i_2 \in [m]$ the event $\mathcal{A}_{i_1,i_2,j} = \{w_{i_1} \to v_1,\, w_{i_1} \to v_j,\, w_{i_2} \to v_1,\, w_{i_2} \to v_j\}$ occurs. From the union bound and the inequality $\tilde{\mathbf{P}}(\mathcal{A}_{i_1,i_2,j}) \leqslant m^{-2}n^{-2}X_{i_1}^2 X_{i_2}^2 Y_1^2 Y_j^2$ we obtain

$$\tilde{\mathbf{P}}(\mathcal{A}) = \tilde{\mathbf{P}}\left(\bigcup_{\{i_1,i_2\}\subset[m]} \bigcup_{2\leqslant j\leqslant n} \mathcal{A}_{i_1,i_2,j}\right) \leqslant n^{-1}\hat{b}_2 Y_1^2 Q_X. \tag{9}$$

Here $Q_X = m^{-2}\sum_{\{i_1,i_2\}\subset[m]} X_{i_1}^2 X_{i_2}^2$. We note that Q_X and Y_1^2 are stochastically bounded and $n^{-1}\hat{b}_2 = o_P(1)$ as $n \to +\infty$. Therefore, $\tilde{\mathbf{P}}(\mathcal{A}) = o_P(1)$. Now (5) implies (6).

Proof of (7). Let $\overline{\mathcal{A}}_1$ denote the complement event to $\mathcal{A}_1$. We have, by the union bound and Markov's inequality,

$$\mathbf{P}_1(\overline{\mathcal{A}}_1) \leqslant \sum_{i\in[m]} \mathbf{P}_1(\lambda_{i1} \geqslant 1) \leqslant (nm)^{-1} Y_1^2 \sum_{i\in[m]} \mathbf{E}X_i^2 = n^{-1}a_2 Y_1^2.$$

We conclude that $\mathbf{P}_1(\overline{\mathcal{A}}_1) = o_P(1)$, and now (5) implies $\mathbf{P}(\overline{\mathcal{A}}_1) = \mathbf{E}\mathbf{P}_1(\overline{\mathcal{A}}_1) = o(1)$.

Proof of (8). Since the first bound of (8) follows from the second one, we only prove the latter. Denote $\hat{X}_1 = \max\{X_1, 1\}$ and $\hat{Y}_1 = \max\{Y_1, 1\}$. We observe that $\mathbf{E}X_1^2 < \infty$, $\mathbf{E}Y_1 < \infty$ implies

$$\lim_{t\to+\infty} \mathbf{E}\hat{X}_1^2 \mathbb{I}_{\{\hat{X}_1 > t\}} = 0, \qquad \lim_{t\to+\infty} \mathbf{E}\hat{Y}_1 \mathbb{I}_{\{\hat{Y}_1 > t\}} = 0.$$

Hence one can find a strictly increasing function $\varphi : [1, +\infty) \to [0, +\infty)$ with $\lim_{t\to+\infty} \varphi(t) = +\infty$ such that

$$\mathbf{E}\hat{X}_1^2 \varphi(\hat{X}_1) < \infty, \qquad \mathbf{E}\hat{Y}_1 \varphi(\hat{Y}_1) < \infty. \tag{10}$$

We remark that one can find φ which satisfies, in addition, the inequalities

$$\varphi(t) < t \qquad \text{and} \qquad \varphi(st) \leqslant \varphi(t)\varphi(s), \qquad \forall s, t \geqslant 1. \tag{11}$$

For this purpose we choose φ of the form $\varphi(t) = e^{\psi(\ln(t))}$, where $\psi : [0, +\infty) \to [0, +\infty)$ is a differentiable concave function, which grows slowly enough to satisfy (10) and the first inequality of (11), and takes value 0 at the origin. We note that the second inequality of (11) follows from the concavity property of ψ. We remark, that (11) implies

$$t/(st) = s^{-1} \leqslant 1/\varphi(s) \leqslant \varphi(t)/\varphi(st), \qquad s, t \geqslant 1. \tag{12}$$

Let $\hat{S}_{XY}$ be defined as in (4) above, but with λ_{ij} replaced by $\hat{\lambda}_{ij} = \hat{X}_i\hat{Y}_j/\sqrt{mn}$. We note, that $S_{XY} \leqslant \hat{S}_{XY}$. Furthermore, from the inequalities

$$\min\{1, \hat{\lambda}_{ij}\} \leqslant \min\Big\{1, \frac{\varphi(\hat{X}_i\hat{Y}_j)}{\varphi(\sqrt{mn})}\Big\} \leqslant \frac{\varphi(\hat{X}_i\hat{Y}_j)}{\varphi(\sqrt{mn})} \leqslant \frac{\varphi(\hat{X}_i)\varphi(\hat{Y}_j)}{\varphi(\sqrt{mn})} \tag{13}$$

we obtain $S_{XY} \leqslant \hat{S}_{XY} \leqslant S^*_{XY}/\varphi(\sqrt{nm})$, where

$$S^*_{XY} = (mn)^{-1}\left(\sum_{1\leqslant i\leqslant m}\hat{X}_i^2\varphi(\hat{X}_i)\right)\left(\sum_{2\leqslant j\leqslant n}\hat{Y}_j\varphi(\hat{Y}_j)\right).$$

We remark that (11) and (12) and imply the third and the first inequality of (13), respectively.

Finally, the bound $\mathbf{E}S_{XY} = o(1)$ follows from the inequality $S_{XY} \leqslant S^*_{XY}/\varphi(\sqrt{nm})$ and the fact that $\mathbf{E}S^*_{XY}$ remains bounded as $n, m \to +\infty$, see (10). □

In the proof of Theorem 1 we use the following inequality referred to as LeCam's inequality, see e.g., [24].

Lemma 6. *Let $S = \mathbb{I}_1 + \mathbb{I}_2 + \cdots + \mathbb{I}_n$ be the sum of independent random indicators with probabilities $\mathbf{P}(\mathbb{I}_i = 1) = p_i$. Let Λ be Poisson random variable with mean $p_1 + \cdots + p_n$. The total variation distance between the distributions P_S of P_Λ of S and Λ*

$$\sup_{A\subset\{0,1,2\ldots\}} |\mathbf{P}(S \in A) - \mathbf{P}(\Lambda \in A)| = \frac{1}{2}\sum_{k\geqslant 0}|\mathbf{P}(S = k) - \mathbf{P}(\Lambda = k)| \leqslant \sum_{1\leqslant i\leqslant n} p_i^2. \tag{14}$$

Proof of Theorem 1. The case (i). Since $\mathbf{P}(d(v_1) \leqslant L) = 1$ it suffices to show that $\mathbf{P}(L > \varepsilon) = o(1)$, for any $\varepsilon \in (0, 1)$. For this purpose we write $\mathbf{P}(L > \varepsilon) = \mathbf{E}\mathbf{P}_1(L > \varepsilon)$ and prove that $\mathbf{P}_1(L > \varepsilon) = o_P(1)$, see (5). Here we estimate $\mathbf{P}_1(L > \varepsilon)$ using the union bound and Markov's inequality,

$$\mathbf{P}_1(L > \varepsilon) \leqslant \sum_{1\leqslant k\leqslant m}\mathbf{P}_1(\mathbb{I}_k = 1) \leqslant \mathbf{E}_1\sum_{1\leqslant k\leqslant m}\lambda_{k1} = \sqrt{m/n}Y_1\mathbf{E}X_1 = o_P(1).$$

The case (ii). Before the proof we introduce some notation. Given X, Y, let

$$\{\Delta_{lk}\}_{k=1}^m, \quad l = 1, 2, 3, \qquad \{\xi_{rk}\}_{k=1}^m, \quad r = 1, 2, 3, 4, \quad \text{and} \quad \{\xi'_{3k}\}_{k=1}^m, \quad \{\xi'_{4k}\}_{k=1}^m, \tag{15}$$

and $\{\eta_k\}_{k=1}^m$ be sequences of conditionally independent (within each sequence) Poisson random variables with mean values

$$\tilde{\mathbf{E}}\xi_{1k} = \sum_{2\leqslant j\leqslant n} p_{kj}, \qquad \tilde{\mathbf{E}}\xi_{2k} = \sum_{2\leqslant j\leqslant n}\lambda_{kj}, \qquad \tilde{\mathbf{E}}\xi_{3k} = \frac{1}{\sqrt{\beta_n}}b_1X_k, \qquad \tilde{\mathbf{E}}\xi_{4k} = \frac{1}{\sqrt{\beta}}b_1X_k,$$

$$\tilde{\mathbf{E}}\Delta_{1k} = \sum_{2\leqslant j\leqslant n}(\lambda_{kj} - p_{kj}), \qquad \tilde{\mathbf{E}}\Delta_{2k} = \frac{1}{\sqrt{\beta_n}}\delta_2X_k, \qquad \tilde{\mathbf{E}}\Delta_{3k} = \frac{1}{\sqrt{\beta_n}}\delta_3X_k,$$

$$\tilde{\mathbf{E}}\xi'_{3k} = \beta_n^{-1/2}bX_k, \qquad \tilde{\mathbf{E}}\xi'_{4k} = \beta' b_1X_k, \qquad \tilde{\mathbf{E}}\eta_k = \lambda_{k1}.$$

Here we denote $b := \min\{\hat{b}_1, b_1\}$, $\delta_2 := \hat{b}_1 - b$, $\delta_3 := b_1 - b$ and $\beta' := \min\{\beta^{-1/2}, \beta_n^{-1/2}\}$. We recall that $\tilde{\mathbf{E}}$ denotes the conditional expectation given X, Y. We assume that, given X, Y, the sequences $\{\xi'_{3k}\}_{k=1}^m$, $\{\Delta_{2k}\}_{k=1}^m$, $\{\Delta_{3k}\}_{k=1}^m$ are conditionally independent, the sequences $\{\mathbb{I}_k\}_{k=1}^m$, $\{\xi_{3k}\}_{k=1}^m$ are conditionally independent, the sequences $\{\xi_{1k}\}_{k=1}^m$, $\{\Delta_{1k}\}_{k=1}^m$ are conditionally independent, and the sequence $\{\eta_k\}_{k=1}^m$ is conditionally independent of the sequences (15). We assume, in addition, that for $1 \leqslant k \leqslant m$

$$\xi_{2k} := \xi_{1k} + \Delta_{1k}, \qquad \xi_{3k} := \xi'_{4k} + (\beta_n^{-1/2} - \beta')b_1 X_k, \qquad \xi_{4k} := \xi'_{4k} + (\beta^{-1/2} - \beta')b_1 X_k.$$

Furthermore, we suppose that, given X, Y_1, the sequence $\{\eta_k\}_{k=1}^m$ is conditionally independent of the subgraph of $H_{X,Y}$ spanned by the vertices $\{v_2, \dots, v_n\} \cup W$.

Next we introduce several random variables which are intermediate between L and d_*:

$$L_0 = \sum_{1 \leqslant k \leqslant m} \eta_k u_k, \qquad L_r = \sum_{1 \leqslant k \leqslant m} \eta_k \xi_{rk}, \qquad r = 1, 2, 3, 4,$$
$$L'_2 = \sum_{1 \leqslant k \leqslant m} \eta_k (\xi'_{3k} + \Delta_{2k}), \qquad L'_3 = \sum_{1 \leqslant k \leqslant m} \eta_k (\xi'_{3k} + \Delta_{3k}).$$

Finally, we denote

$$f_\tau(t) = \mathbf{E} e^{it\tau_1}, \qquad \bar{f}_\tau(t) = \sum_{r \geqslant 0} e^{itr} \bar{p}_r, \qquad \bar{p}_r = \bar{\lambda}^{-1} \sum_{1 \leqslant k \leqslant m} \lambda_{k1} \mathbb{I}_{\{\xi_{4k} = r\}}, \qquad \bar{\lambda} = \sum_{k=1}^m \lambda_{k1},$$
$$\delta = (\bar{f}_\tau(t) - 1)\bar{\lambda} - (f_\tau(t) - 1)\lambda_1, \qquad f(t) = \mathbf{E}_1 e^{itd_*}, \qquad \bar{f}(t) = \bar{\mathbf{E}} e^{itL_4}.$$

Here $\bar{\mathbf{E}}$ denotes the conditional expectation given X, Y and $\xi_{41}, \dots, \xi_{4m}$. We recall that $\mathbf{E}_1$ denotes the conditional expectation given Y_1.

Let us prove the statement (ii). In view of (6) the random variables $d(v_1)$ and L have the same asymptotic distribution (if any). We shall show that L converges in distribution to (1). We proceed in two steps. Firstly, we approximate the distribution of L by that of L_4 using LeCam's inequality, see Lemma 6, and a coupling argument. Secondly, we prove the convergence of the characteristic functions $\mathbf{E} e^{itL_4} \to \mathbf{E} e^{itd_*}$. Here and below $i = \sqrt{-1}$ denotes the imaginary unit.

Step 1. Let us show that L and L_4 have the same asymptotic probability distribution (if any). To this aim we prove that

$$d_{TV}(L, L_0) = o(1), \qquad d_{TV}(L_0, L_1) = o(1), \tag{16}$$
$$\mathbf{E}|L_1 - L_2| = o(1), \qquad L'_2 - L'_3 = o_P(1), \qquad \mathbf{E}|L_3 - L_4| = o(1), \tag{17}$$

and observe that L_2 (respectively L_3) has the same distribution as L'_2 (respectively L'_3).

Let us prove the first bound of (16). In view of (5) it suffices to show that $\tilde{d}_{TV}(L_0, L) = o_P(1)$. In order to prove the latter bound we apply the inequality

$$\tilde{d}_{TV}(L_0, L)\mathbb{I}_{\mathcal{A}_1} \leqslant n^{-1} Y_1^2 \hat{a}_2 \tag{18}$$

shown below. We remark that (18) implies

$$\tilde{d}_{TV}(L_0, L) \leqslant \tilde{d}_{TV}(L_0, L)\mathbb{I}_{\mathcal{A}_1} + \mathbb{I}_{\overline{\mathcal{A}}_1} \leqslant n^{-1}Y_1^2\hat{a}_2 + \mathbb{I}_{\overline{\mathcal{A}}_1} = o_P(1).$$

Here $n^{-1}Y_1^2\hat{a}_2 = o_P(1)$, because $Y_1^2\hat{a}_2$ is stochastically bounded. Furthermore, the bound $\mathbb{I}_{\overline{\mathcal{A}}_1} = o_P(1)$ follows from (7). It remains to prove (18). We denote $\bar{L}_k = \sum_{l=1}^{k} \mathbb{I}_l u_l + \sum_{l=k+1}^{m} \eta_l u_l$ and write, by the triangle inequality,

$$\tilde{d}_{TV}(L_0, L) \leqslant \sum_{k=1}^{m} \tilde{d}_{TV}(\bar{L}_{k-1}, \bar{L}_k).$$

Then we estimate $\tilde{d}_{TV}(\bar{L}_{k-1}, \bar{L}_k) \leqslant \tilde{d}_{TV}(\eta_k, \mathbb{I}_k) \leqslant (nm)^{-1}Y_1^2X_k^2$. Here the first inequality follows from the properties of the total variation distance. The second inequality follows from Lemma 6 and the fact that on the event $\mathcal{A}_1$ we have $p_{k1} = \lambda_{k1}$.

Let us prove the second bound of (16). In view of (5) it suffices to show that $\tilde{d}_{TV}(L_0, L_1) = o_P(1)$. We denote $L_k^* = \sum_{l=1}^{k} \eta_l u_l + \sum_{l=k+1}^{m} \eta_l \xi_{1l}$ and write, by the triangle inequality,

$$\tilde{d}_{TV}(L_0, L_1) \leqslant \sum_{k=1}^{m} \tilde{d}_{TV}(L_{k-1}^*, L_k^*). \tag{19}$$

Here

$$\tilde{d}_{TV}(L_{k-1}^*, L_k^*) \leqslant \tilde{d}_{TV}(\eta_k u_k, \eta_k \xi_{1k}) \leqslant \tilde{\mathbf{P}}(\eta_k \neq 0)\tilde{d}_{TV}(u_k, \xi_{1k}). \tag{20}$$

Now, invoking the inequalities $\tilde{\mathbf{P}}(\eta_k \neq 0) = 1 - e^{-\lambda_{k1}} \leqslant \lambda_{k1}$ and $\tilde{d}_{TV}(u_k, \xi_{1k}) \leqslant \sum_{j=2}^{n} p_{kj}^2$, we obtain the desired bound from (8), (19) and (20)

$$\tilde{d}_{TV}(L_0, L_1) \leqslant \sum_{k=1}^{m} \lambda_{k1} \sum_{j=2}^{n} p_{kj}^2 \leqslant Y_1 S_{XY} = o_P(1).$$

Let us prove the first bound of (17). We observe that

$$|L_2 - L_1| = L_2 - L_1 = \sum_{1 \leqslant k \leqslant m} \eta_k \Delta_{1k}$$

and

$$\tilde{\mathbf{E}} \sum_{1 \leqslant k \leqslant m} \eta_k \Delta_{1k} = \sum_{1 \leqslant k \leqslant m} \lambda_{k1} \sum_{2 \leqslant j \leqslant n} (\lambda_{kj} - 1)\mathbb{I}_{\{\lambda_{kj} > 1\}} \leqslant Y_1 S_{XY}.$$

Hence, we obtain $\mathbf{E}|L_2 - L_1| \leqslant \mathbf{E}Y_1\mathbf{E}S_{XY} = o(1)$, see (8).

Let us prove the second bound of (17). It suffices to show that $\tilde{\mathbf{E}}|L_2' - L_3'| = o_P(1)$. We have

$$\tilde{\mathbf{E}}|L_2' - L_3'| = (\delta_2 + \delta_3)Y_1\hat{a}_2 = |\hat{b}_1 - b_1|Y_1\hat{a}_2 = o_P(1). \tag{21}$$

In the last step we used the fact that $Y_1\hat{a}_2 = O_P(1)$ and $\hat{b}_1 - b_1 = o_P(1)$.

The third bound of (17) is straightforward

$$\mathbf{E}|L_3 - L_4| \leqslant \Big|\frac{1}{\sqrt{\beta_n}} - \frac{1}{\sqrt{\beta}}\Big| b_1 \sum_{1\leqslant k\leqslant m} \mathbf{E}\eta_k X_k = \Big|1 - \frac{\sqrt{\beta_n}}{\sqrt{\beta}}\Big| b_1^2 a_2 = o(1).$$

Step 2. Here we show that for every real t we have $\mathbf{E}\Delta(t) = o(1)$, where $\Delta(t) = e^{itL_4} - e^{itd_*}$. To this aim we prove that for any realized value Y_1 there exists a positive constant $c = c(t, Y_1)$ such that for every $0 < \varepsilon < 0.5$ we have

$$\limsup_{n,m\to+\infty} |\mathbf{E}(\Delta(t)|Y_1)| < c\,\varepsilon. \tag{22}$$

Clearly, (22) implies $\mathbf{E}(\Delta(t)|Y_1) = o(1)$. This fact together with the simple inequality $|\Delta(t)| \leqslant 2$ yields the desired bound $\mathbf{E}\Delta(t) = o(1)$, by Lebesgue's dominated convergence theorem.

We fix $0 < \varepsilon < 0.5$ and prove (22). Introduce event $\mathcal{D} = \{|\hat{a}_1 - a_1| < \varepsilon \min\{1, a_1\}\}$ and let $\overline{\mathcal{D}}$ denote the complement event. Furthermore, select a number $T > 1/\varepsilon$ such that $\mathbf{P}(\tau_1 \geqslant T) < \varepsilon$. By $c_1, c_2, \dots$ we denote positive numbers which do not depend on n, m.

We observe that, given Y_1, the conditional distribution of d_* is the compound Poisson distribution with the characteristic function $f(t) = e^{\lambda_1(f_\tau(t)-1)}$. Similarly, given X, Y and $\xi_{41}, \dots, \xi_{4m}$, the conditional distribution of L_4 is the compound Poisson distribution with the characteristic function $\bar{f}(t) = e^{\bar{\lambda}(\bar{f}_\tau(t)-1)}$. In the proof of (22) we exploit the convergence $\bar{\lambda} \to \lambda_1$ and $\bar{f}_\tau(t) \to f_\tau(t)$. In what follows we assume that m, n are so large that $\beta \leqslant 2\beta_n \leqslant 4\beta$.

Let us show (22). For this purpose we write

$$\mathbf{E}(\Delta(t)|Y_1) = \mathbf{E}_1\Delta(t) = I_1 + I_2, \qquad I_1 = \mathbf{E}_1\Delta(t)\mathbb{I}_{\mathcal{D}}, \qquad I_2 = \mathbf{E}_1\Delta(t)\mathbb{I}_{\overline{\mathcal{D}}}$$

and estimate I_1 and I_2. We have $|I_2| \leqslant 2\mathbf{P}_1(\overline{\mathcal{D}}) = 2\mathbf{P}(\overline{\mathcal{D}}) = o(1)$, by the law of large numbers.

Now we show that $I_1 \leqslant c\,\varepsilon + o(1)$. Combining the identity $\mathbf{E}_1\Delta(t) = \mathbf{E}_1 f(t)(e^{\delta} - 1)$ with the inequalities $|f(t)| \leqslant 1$ and $|e^s - 1| \leqslant |s|e^{|s|}$ we obtain

$$|I_1| \leqslant \mathbf{E}_1|\delta|e^{|\delta|}\mathbb{I}_{\mathcal{D}} \leqslant c_1\mathbf{E}_1|\delta|\mathbb{I}_{\mathcal{D}}. \tag{23}$$

Here we estimated $e^{|\delta|} \leqslant e^{8\lambda_1} =: c_1$ using the inequalities

$$|\delta| \leqslant 2\bar{\lambda} + 2\lambda_1, \qquad \bar{\lambda} = Y_1\beta_n^{1/2}\hat{a}_1 \leqslant 3\lambda_1.$$

We remark that the last inequality holds for $\beta_n \leqslant 2\beta$ provided that the event $\mathcal{D}$ occurs. We shall show below that

$$\mathbf{E}_1|\delta|\mathbb{I}_{\mathcal{D}} \leqslant (c_2 + \lambda_1 c_3 + \lambda_1 c_4)\varepsilon + o(1). \tag{24}$$

This inequality together with (23) yields the desired upper bound for I_1.

Let us prove (24). We write

$$\delta = (\bar{f}_\tau(t) - 1)(\bar{\lambda} - \lambda_1) + (\bar{f}_\tau(t) - f_\tau(t))\lambda_1$$

and estimate

$$|\delta| \leqslant 2|\bar{\lambda} - \lambda_1| + \lambda_1 |\bar{f}_\tau(t) - f_\tau(t)|.$$

Now, from the inequalities $|\hat{a}_1 - a_1| < \varepsilon$ and $\beta_n \leqslant 2\beta$ we obtain

$$|\bar{\lambda} - \lambda_1| \leqslant Y_1 |\hat{a}_1 - a| \beta_n^{1/2} + Y_1 a_1 |\beta_n^{1/2} - \beta^{1/2}| \leqslant 2Y_1 \beta^{1/2} \varepsilon + o(1).$$

Hence, $\mathbf{E}_1 |\bar{\lambda} - \lambda_1| \mathbb{I}_{\mathcal{D}} \leqslant c_2 \varepsilon + o(1)$, where $c_2 = 2Y_1 \beta^{1/2}$. We complete the proof of (24) by showing that

$$\mathbf{E}_1 |\bar{f}_\tau(t) - f_\tau(t)| \mathbb{I}_{\mathcal{D}} \leqslant (c_3 + c_4)\varepsilon + o(1). \tag{25}$$

In order to prove (25) we split

$$\bar{f}_\tau(t) - f_\tau(t) = \sum_{r \geqslant 0} e^{itr} (\bar{p}_r - p_r) = R_1 - R_2 + R_3,$$

and estimate separately the terms

$$R_1 = \sum_{r \geqslant T} e^{itr} \bar{p}_r, \qquad R_2 = \sum_{r \geqslant T} e^{itr} p_r, \qquad R_3 = \sum_{0 \leqslant r < T} e^{itr} (\bar{p}_r - p_r).$$

Here we denote $p_r = \mathbf{P}(\tau_1 = r)$. The upper bound for R_2 follows by the choice of T

$$|R_2| \leqslant \sum_{r \geqslant T} p_r = \mathbf{P}(\tau_1 \geqslant T) < \varepsilon.$$

Next, combining the identity $\bar{p}_r = (\hat{a}_1 m)^{-1} \sum_{1 \leqslant k \leqslant m} X_k \mathbb{I}_{\{\xi_{4k} = r\}}$ with the inequalities

$$|R_1| \leqslant (\hat{a}_1 m)^{-1} \sum_{r \geqslant T} \sum_{1 \leqslant k \leqslant m} X_k \mathbb{I}_{\{\xi_{4k} = r\}} = (\hat{a}_1 m)^{-1} \sum_{1 \leqslant k \leqslant m} X_k \mathbb{I}_{\{\xi_{4k} \geqslant T\}} \tag{26}$$

and $\hat{a}_1^{-1} \mathbb{I}_{\mathcal{D}} \leqslant 2a_1^{-1}$, we estimate

$$\mathbf{E}_1 |R_1| \mathbb{I}_{\mathcal{D}} \leqslant 2a_1^{-1} \mathbf{E}_1 (X_k \mathbb{I}_{\{\xi_{4k} \geqslant T\}}) \leqslant 2a_1^{-1} T^{-1} \mathbf{E}_1 (X_k \xi_{4k}) = 2a_1^{-1} a_2 b_1 \beta^{-1/2} \varepsilon.$$

Hence, we obtain $\mathbf{E}_1 |R_1| \mathbb{I}_{\mathcal{D}} \leqslant c_4 \varepsilon$, where $c_4 = 2a_1^{-1} a_2 b_1 \beta^{-1/2}$.

Now we estimate R_3. We denote $p_r' = (\hat{a}_1 / a_1) \bar{p}_r$ and observe that the inequality $|\hat{a}_1 - a| \leqslant \varepsilon a_1$ implies $|\hat{a}_1 a_1^{-1} - 1| \leqslant \varepsilon$ and

$$|\sum_{0 \leqslant r \leqslant T} e^{itr} (\bar{p}_r - p_r')| \leqslant \varepsilon \sum_{0 \leqslant r \leqslant T} \bar{p}_r \leqslant \varepsilon.$$

In the last inequality we use the fact that the probabilities $\{\bar{p}_r\}_{r \geqslant 0}$ sum up to 1. It follows now that

$$|R_3| \mathbb{I}_{\mathcal{D}} \leqslant \varepsilon + \sum_{0 \leqslant r \leqslant T} |p_r' - p_r|.$$

Furthermore, observing that $\mathbf{E}_1 p'_r = a^{-1}\mathbf{E}_1 X_k \mathbb{I}_{\{\xi_{4k}=r\}} = p_r$, for $1 \leqslant k \leqslant m$, we obtain

$$\mathbf{E}_1|p'_r - p_r|^2 = m^{-1}\mathbf{E}_1|a^{-1}X_k\mathbb{I}_{\{\xi_{4k}=r\}} - p_r|^2 \leqslant m^{-1}a_1^{-2}\mathbf{E}X_k^2.$$

Hence, $\mathbf{E}_1|p'_r - p_r| = O(m^{-1/2})$. We conclude that

$$\mathbf{E}_1|R_3|\mathbb{I}_{\mathcal{D}} \leqslant \varepsilon + O(|T|m^{-1/2}) = \varepsilon + o(1),$$

thus completing the proof of (25).

The case (iii). Before the proof we introduce several random variables. Given $\varepsilon \in (0,1)$, denote

$$\mathbb{I}'_k = \mathbb{I}_{\{X_k\beta_n^{-1/2}b_1<\varepsilon\}}, \quad \gamma_k = X_k\beta_n^{-1/2}b_1\mathbb{I}'_k, \quad 1 \leqslant k \leqslant m, \qquad \text{and} \qquad \gamma = \sum_{1\leqslant k\leqslant m}\lambda_{k1}\gamma_k.$$

Given X, Y, let $\tilde{\mathbb{I}}_1, \ldots, \tilde{\mathbb{I}}_m$ be conditionally independent Bernoulli random variables with success probabilities

$$\tilde{\mathbf{P}}(\tilde{\mathbb{I}}_k = 1) = 1 - \tilde{\mathbf{P}}(\tilde{\mathbb{I}}_k = 0) = \gamma_k.$$

We assume that, given X, Y, the sequences $\{\mathbb{I}_k\}_{k=1}^m$, $\{\tilde{\mathbb{I}}_k\}_{k=1}^m$ and $\{\xi_{3k}\}_{k=1}^m$ (see (15)) are conditionally independent. Furthermore, introduce random variables

$$L_5 = \sum_{1\leqslant k\leqslant m}\mathbb{I}_k\xi_{3k}, \qquad L_6 = \sum_{1\leqslant k\leqslant m}\mathbb{I}_k\mathbb{I}'_k\xi_{3k}, \qquad L_7 = \sum_{1\leqslant k\leqslant m}\mathbb{I}_k\tilde{\mathbb{I}}_k$$

and let L_8 be a random variable with the distribution $\mathbf{P}(L_8 = r) = \mathbf{E}e^{-\gamma}\gamma^r/r!$, for $r = 0, 1, \ldots$. We remark that the probability distribution of L_8 is a Poisson mixture, i.e., the Poisson distribution with random parameter γ.

We note that by (16) and the first two bounds of (17), the random variables L and L_3 have the same asymptotic distribution (if any). Now we prove that L_3 converges in distribution to Λ_3. For this purpose we show that for any $\varepsilon \in (0,1)$

$$d_{TV}(L_3, L_5) = o(1), \qquad \mathbf{E}|L_5 - L_6| = o(1), \tag{27}$$

$$d_{TV}(L_6, L_7) \leqslant a_2 b_1^2\varepsilon, \qquad d_{TV}(L_7, L_8) = o(1), \tag{28}$$

$$\mathbf{E}e^{itL_8} - \mathbf{E}e^{it\Lambda_3} = o(1). \tag{29}$$

Let us prove (27). The first bound of (27) is obtained in the same way as the first bound of (16). To show the second bound of (27) we write

$$\tilde{\mathbf{E}}|L_5 - L_6| = \sum_{1\leqslant k\leqslant m}(1 - \mathbb{I}'_k)\tilde{\mathbf{E}}\mathbb{I}_k\tilde{\mathbf{E}}\xi_{3k} \leqslant Y_1 b_1 m^{-1}\sum_{1\leqslant k\leqslant m}(1 - \mathbb{I}'_k)X_k^2$$

and obtain

$$\mathbf{E}|L_5 - L_6| = \mathbf{E}\tilde{\mathbf{E}}|L_5 - L_6| \leqslant b_1^2\mathbf{E}X_1^2\mathbb{I}_{\{X_1\beta_n^{-1/2}b_1\geqslant\varepsilon\}} = o(1).$$

We note that the right hand side tends to zero since $\beta_n \to +\infty$.

Let us prove the first inequality of (28). Proceeding as in (19), (20) and using the identity $\tilde{\mathbb{I}}_k = \tilde{\mathbb{I}}_k \mathbb{I}'_k$ we write

$$\tilde{d}_{TV}(L_6, L_7) \leqslant \sum_{1 \leqslant k \leqslant m} \mathbb{I}'_k \tilde{\mathbf{P}}(\mathbb{I}_k \neq 0) \tilde{d}_{TV}(\xi_{3k}, \tilde{\mathbb{I}}_k).$$

Next, we estimate $\mathbb{I}'_k \tilde{d}_{TV}(\xi_{3k}, \tilde{\mathbb{I}}_k) \leqslant \gamma_k^2$, by LeCam's inequality (14), and invoke the inequality $\tilde{\mathbf{P}}(\mathbb{I}_k \neq 0) \leqslant \lambda_{k1}$. We obtain

$$\tilde{d}_{TV}(L_6, L_7) \leqslant \sum_{1 \leqslant k \leqslant m} \mathbb{I}'_k \lambda_{k1} \gamma_k^2 \leqslant \varepsilon \sum_{1 \leqslant k \leqslant m} \mathbb{I}'_k \lambda_{k1} \gamma_k \leqslant \varepsilon Y_1 b_1 \hat{a}_2.$$

Here we estimated $\gamma_k^2 \leqslant \varepsilon \gamma_k$. Now the inequalities $d_{TV}(L_6, L_7) \leqslant \mathbf{E} \tilde{d}_{TV}(L_6, L_7) \leqslant a_2 b_1^2 \varepsilon$ imply the first relation of (28).

Let us prove the second relation of (28). In view of (5) it suffices to show that $\tilde{d}_{TV}(L_7, L_8) = o_P(1)$. For this purpose we write

$$\tilde{d}_{TV}(L_7, L_8) \leqslant \mathbb{I}_{\mathcal{A}_1} \tilde{d}_{TV}(L_7, L_8) + \mathbb{I}_{\overline{\mathcal{A}}_1},$$

where $\mathbb{I}_{\overline{\mathcal{A}}_1} = o_P(1)$, see (7), and estimate using LeCam's inequality (14)

$$\mathbb{I}_{\mathcal{A}_1} \tilde{d}_{TV}(L_7, L_8) \leqslant \mathbb{I}_{\mathcal{A}_1} \sum_{1 \leqslant k \leqslant m} \tilde{\mathbf{P}}^2(\mathbb{I}_k \tilde{\mathbb{I}}_k = 1) \mathbb{I}'_k = \mathbb{I}_{\mathcal{A}_1} \sum_{1 \leqslant k \leqslant m} \lambda_{k1}^2 \gamma_k^2 \leqslant b_1^2 Y_1^2 m^{-1} \hat{a}_4 = o_P(1).$$

In the last step we used the fact that $\mathbf{E} X_1^2 < \infty$ implies $m^{-1} \hat{a}_4 = o_P(1)$.

Finally, we show (29). We write $\tilde{\mathbf{E}} e^{itL_8} = e^{\gamma(e^{it}-1)}$ and observe that

$$Y_1 b_1 a_2 - \gamma = o_P(1). \tag{30}$$

Furthermore, since for any real t the function $z \to e^{z(e^{it}-1)}$ is bounded and uniformly continuous for $z \geqslant 0$, we conclude that (30) implies the convergence

$$\mathbf{E} e^{itL_8} = \mathbf{E} e^{\gamma(e^{it}-1)} \to \mathbf{E} e^{Y_1 b_1 a_2 (e^{it}-1)} = \mathbf{E} e^{it\Lambda_3}.$$

It remains to prove (30). We write $Y_1 b_1 a_2 - \gamma = Y_1 b_1 (a_2 - \hat{a}_2) + Y_1 b_1 \hat{a}_2 - \gamma$ and note that $a_2 - \hat{a}_2 = o_P(1)$, by the law of large numbers, and

$$0 \leqslant \mathbf{E}(Y_1 b_1 \hat{a}_2 - \gamma) = b_1^2 \mathbf{E} X_1^2 \mathbb{I}_{\{X_1 \beta_n^{-1/2} b_1 \geqslant \varepsilon\}} = o(1).$$

□

Acknowledgements

We thank the anonymous referee for remarks and suggestions that have improved the presentation.

References

[1] A. D. Barbour and G. Reinert. The shortest distance in random multi-type intersection graphs. *Random Structures and Algorithms*, 39:179–209, 2011.

[2] M. Behrisch. Component evolution in random intersection graphs. *Electron. J. Combin.*, 14:Research Paper 17, 12, 2007.

[3] S. Blackburn and S. Gerke. Connectivity of the uniform random intersection graph. *Discrete Mathematics*, 309: 5130–5140, 2009.

[4] M. Bloznelis. Degree distribution of a typical vertex in a general random intersection graph. *Lithuanian Mathematical J.*, 48:38–45, 2008.

[5] M. Bloznelis. A random intersection digraph: Indegree and outdegree distributions. *Discrete Mathematics*, 310:2560–2566, 2010.

[6] M. Bloznelis. The largest component in an inhomogeneous random intersection graph with clustering. *Electron. J. Combin.*, 17:Research Paper 110, 17, 2010.

[7] M. Bloznelis. Degree and clustering coefficient in sparse random intersection graphs. *Annals of Applied Probability*, 23:1254–1289, 2013.

[8] M. Bloznelis, V. Kurauskas. Clustering function: a measure of social influence. arXiv:1207.4941, 2012

[9] M. Bradonjic, A. Hagberg, N. W. Hengartner, and A. G. Percus. Component Evolution in General Random Intersection Graphs. In *The 7th Workshop on Algorithms and Models for the Web Graph, WAW2010*, volume 6516 of *Lecture Notes in Comput. Sci.*, pages 36–49. Springer, 2010.

[10] T. Britton, M. Deijfen, M. Lindholm, and N. A. Lageras. Epidemics on random graphs with tunable clustering. *J. Appl. Prob.*, 45:743–756, 2008.

[11] M. Deijfen and W. Kets. Random intersection graphs with tunable degree distribution and clustering. *Probab. Engrg. Inform. Sci.*, 23:661–674, 2009.

[12] L. Eschenauer and V. D. Gligor. A key-management scheme for distributed sensor networks. In *Proceedings of the 9th ACM Conference on Computer and Communications Security*, pages 41–47. ACM New York, 2002.

[13] S. Foss, D. Korshunov, and S. Zachary. *An Introduction to Heavy-Tailed and Subexponential Distributions.* ACM, New York, 2011.

[14] E. Godehardt and J. Jaworski. Two models of random intersection graphs and their applications. *Electronic Notes in Discrete Mathematics*, 10:129–132, 2001.

[15] J. Jaworski, M. Karoński, and D. Stark. The degree of a typical vertex in generalized random intersection graph models. *Discrete Mathematics*, 306:2152–2165, 2006.

[16] J. Jaworski and D. Stark. The vertex degree distribution of passive random intersection graph models. *Combinatorics, Probability and Computing*, 17:549–558, 2008.

[17] M. Karoński, E. R. Scheinerman, and K. B. Singer-Cohen. On random intersection graphs: The subgraph problem. *Combinatorics, Probability and Computing*, 8:131–159, 1999.

[18] S. Nikoletseas, C. Raptopoulos, and P. G. Spirakis. The existence and efficient construction of large independent sets in general random intersection graphs. In *ICALP (2004), J. Daz, J. Karhumki, A. Lepist, and D. Sannella, Eds.*, volume 3142 of *Lecture Notes in Comput. Sci.*, pages 1029–1040. Springer, 2004.

[19] K. Rybarczyk. Diameter, connectivity, and phase transition of the uniform random intersection graph. *Discrete Mathematics*, 311:1998–2019, 2011.

[20] K. Rybarczyk. The degree distribution in random intersection graphs. In *Studies in Classification, Data Analysis and Knowledge Organization*, pages 291–299. Springer, 2012.

[21] Y. Shang. Degree distributions in general random intersection graphs. *Electron. J. Combin.*, 17:Research Paper 23, 8, 2010.

[22] K. B. Singer-Cohen. Random intersection graphs. PhD thesis, Department of Mathematical Sciences, The Johns Hopkins University, 1995.

[23] D. Stark. The vertex degree distribution of random intersection graphs. *Random Structures and Algorithms*, 24:249–258, 2004.

[24] J. M. Steele. Le Cam's inequality and Poisson approximations. *The American Mathematical Monthly*, 101:48–54, 1994.