Clustering function: a measure of social influence

Mindaugas Bloznelis[1,2] and Valentas Kurauskas[1]

Abstract

A commonly used characteristic of statistical dependence of adjacency relations in real networks, the clustering coefficient, evaluates chances that two neighbours of a given vertex are adjacent. Another characteristic is obtained by considering conditional probabilities that two randomly chosen vertices are adjacent given that they have r common neighbours. We denote such probabilities $cl(r)$ and call $r \to cl(r)$ the clustering function. We compare clustering functions of several networks having non-negligible clustering coefficient. They show similar patterns and surprising regularity. We also provide mathematically rigorous analysis of the clustering function of related random intersection graph models aimed at explaining the empirical results.

key words: clustering coefficient, power law, social network, intersection graph

2000 Mathematics Subject Classifications: 91D30, 05C80, 05C07, 91C20

arXiv:1207.4941v1 [stat.AP] 20 Jul 2012

1 Introduction

Our study is motivated by the question: given two vertices of a network, the presence of how many common neighbours would imply with certainty that these two vertices are adjacent. A "softer" question is about the probability that two vertices with (at least) r common neighbours establish a link. The answer is given by the clustering functions (1) and (2).

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a finite graph on the vertex set $\mathcal{V}$ and with the edge set $\mathcal{E}$. The number of neighbours of a vertex v is denoted $d(v)$. The number of common neighbours of vertices v_i and v_j is denoted $d(v_i, v_j)$. We are interested in the fraction of adjacent pairs $v_i \sim v_j$ among all pairs $\{v_i, v_j\} \subset \mathcal{V}$ having (at least) r common neighbours. Here and below '$\sim$' denotes the adjacency relation of $\mathcal{G}$. More formally, let us consider the random pair of distinct vertices $\{v_1^*, v_2^*\}$ drawn from $\mathcal{V}$ uniformly at random. Define the clustering functions of $\mathcal{G}$

$$r \to cl_{\mathcal{G}}(r) := \mathbf{P}(v_1^* \sim v_2^* | \, d(v_1^*, v_2^*) = r), \tag{1}$$

$$r \to Cl_{\mathcal{G}}(r) := \mathbf{P}(v_1^* \sim v_2^* | \, d(v_1^*, v_2^*) \geq r). \tag{2}$$

In the case of a social network (1), (2) could be interpreted as measures of social influence or pressure exercised by the neighbours on a pair of actors to establish a communication link. We remark that characteristics (1) and (2) are related to the clustering coefficient of $\mathcal{G}$. We recall its definition for convenience. Let (v_1^*, v_2^*, v_3^*) be an ordered triple of distinct vertices drawn from $\mathcal{V}$ uniformly at random. The conditional probability that v_1^* is adjacent to v_2^*, given that v_1^* and v_2^* are both adjacent to v_3^*, is called the (global) clustering coefficient ([3], [16], [17], [24]). We denote it $C_{\mathcal{G}} = \mathbf{P}(v_1^* \sim v_2^* | v_1^* \sim v_3^*, v_2^* \sim v_3^*)$.

[1]Faculty of Mathematics and Informatics, Vilnius University, 03225 Vilnius, Lithuania
[2]Corresponding author mindaugas.bloznelis@mif.vu.lt

In this paper we study clustering functions first by considering empirical data and then by a rigorous analysis of related random graph models.
We consider clustering function (1) of real networks admitting positive clustering coefficient: the actor network, where two actors are declared adjacent whenever they have acted in the same film ([27]), and the Facebook network ([1], [12], [25]). We remark that empirical plots show similar pattern and surprising regularity of the clustering function $r \to cl_{\mathcal{G}}(r)$ (see Sect. 2 below).
Our choice of the random graph model is motivated by an observation of Newman et al. [18] that the clustering property of some social networks (so called affiliation networks) could be explained by the presence of a bipartite graph structure. For example, the bipartite graph, where actors are linked to films, defines the actor network. It seems reasonable that a bipartite graph structure might also be helpful in explaining (at least to some extent) the adjacency relations of Facebook network: two members become adjacent because they share some common interests/attributes.
We secondly consider clustering function (1) of a random intersection graph, where vertices (actors) are prescribed attribute sets independently at random and two vertices are declared adjacent whenever they share at least one common attribute ([15], [13], see also [2], [14]). The random intersection graphs are relatively simple objects and for them rigorous mathematical results can be obtained. We evaluate the probabilities $\mathbf{P}(v_1^* \sim v_2^* | \, d(v_1^*, v_2^*) = r)$, $r = 0, 1, 2, \ldots$ for a random intersection graph in Sect. 3 below. Our theoretical results are then used to interpret empirical findings of Sect 2.

2 Clustering functions: empirical results

In Fig.1 we plot clustering functions (1) and (2) of three drama actor networks: English actor network with $n = 402622$ actors, $m = 66127$ films and the clustering coefficient $C = 0.32$ (clustering coefficient here and below is rounded up to 2 decimal places), French actor network with $n = 43204$ actors, $m = 5629$ films and the clustering coefficient $C = 0.30$ and Russian actor network with $n = 9880$ actors, $m = 2459$ films and $C = 0.44$. Data is obtained from [27]. In Fig.2 we plot clustering function (1) of three networks describing relations between community members at three different universities (data from [25]): the first network has $n = 17425$ vertices and the clustering coefficient $C = 0.16$ (• blue graph); the second network has $n = 9414$ vertices and the clustering coefficient $C = 0.15$ ($\star$ green graph); the third network has $n = 6596$ vertices and the clustering coefficient $C = 0.16$ (■ red graph).

3 Clustering functions of random intersection graphs

Vertices $v_1, \ldots, v_n$ of an intersection graph are represented by subsets $D_1, \ldots, D_n$ of a given ground set $W = \{w_1, \ldots, w_m\}$. Elements of W are called attributes or keys. Vertices v_i and v_j are declared adjacent if $D_i \cap D_j \neq \emptyset$. The adjacency relations of such an intersection graph resemble that of some real networks, e.g., the collaboration network, where authors are declared adjacent whenever they have coauthored a paper, or the actor network, where two actors are linked by an edge whenever they have acted in the same film. Random intersection graph have attracted considerable attention in the recent literature, see, e.g., [4], [5], [7] [9], [11], [20], [19], [26]. They admit a power law degree distribution and tunable clustering. We consider two models of random intesection graphs: the active graph and the inhomogeneous graph.
Active graph. In the *active* random intersection graph $G_1(n, m, P)$ every vertex $v_i \in V = \{v_1, \ldots, v_n\}$ selects its attribute set D_i independently at random ([13], [15]). Here we assume for simplicity that independent random sets $D_1, \ldots, D_n$ have the same probability distribution

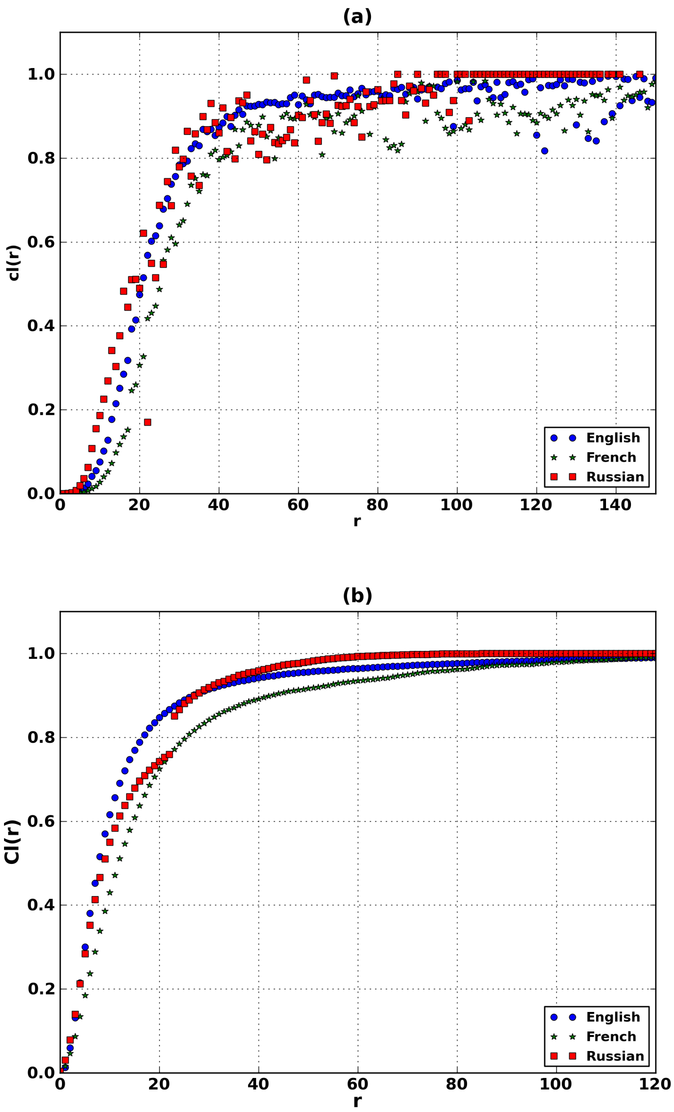

Figure 1: Clustering functions for three actor networks: (a) $cl(r)$, (b) $Cl(r)$.

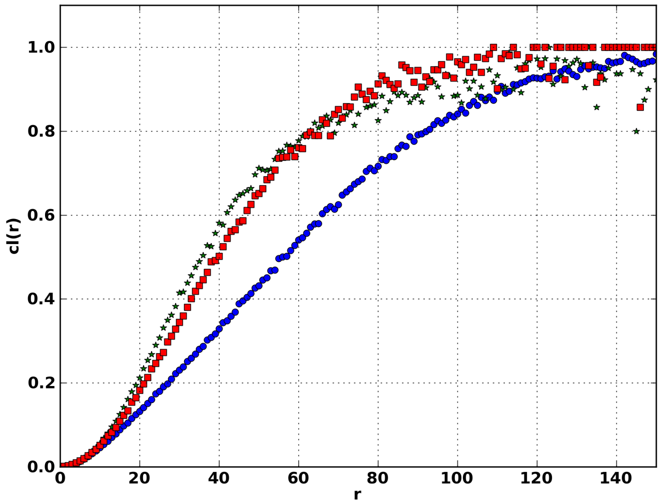

Figure 2: Clustering functions of three university networks.

such that all attributes have equal probabilities to be selected. In particular, we have

$$\mathbf{P}(D_i = A) = \binom{m}{|A|}^{-1} P(|A|), \qquad \text{for} \quad \text{any} \quad A \subset W. \tag{3}$$

Here P is the common probability distribution of the sizes of selected sets $X_i := |D_i|$ (for each $i = 1, \dots, n$ we have $\mathbf{P}(X_i = k) = P(k)$, $k = 0, 1, \dots m$). We remark that $X_1, \dots, X_n$ are independent random variables taking values in $\{0, 1, \dots, m\}$.
Here we study the clustering function

$$r \to cl(r) = \mathbf{P}(v_1^* \sim v_2^* |\, d(v_1^*, v_2^*) = r) = \mathbf{P}(v_1 \sim v_2 |\, d(v_1, v_2) = r) \tag{4}$$

of a sparse random intersection graph with large number of vertices (by sparse we mean that the number of edges scales as the number of vertices n as $n \to +\infty$). It is convenient to consider a sequence of random intersection graphs $\{G_{(n)}\}_n$, where $G_{(n)} = G_1(n, m, P)$ and where $m = m_n$ and $P = P_n$ both depend on n. We remark that $\{G_{(n)}\}_n$ is a sequence of sparse random graphs whenever the size X_1 of the typical random set scales as $(m/n)^{1/2}$ as $m, n \to \infty$ ([6]). Assuming, in addition, that
(i) $X_1 \sqrt{n/m}$ converges in distribution to some random variable Z;
(ii) $\mathbf{E}Z < \infty$ and $\mathbf{E}X_1 \sqrt{n/m}$ converges to $\mathbf{E}Z$
one obtains the asymptotic degree distribution of $\{G_{(n)}\}$

$$\lim_{n \to +\infty} \mathbf{P}(d(v_1)) = \mathbf{E}e^{-z_1 Z}(z_1 Z)^k / k!, \qquad \text{for} \qquad k = 0, 1, \dots, \tag{5}$$

see [6], [7], [10], [22]. Here $d(v)$ denotes the degree of a vertex v and z_i denotes the $i-$th moment of Z, $z_i = \mathbf{E}Z^i$. Along with the first moment condition (ii) we shall also consider the $r-$th moment condition
(ii-r) $\mathbf{E}Z^r < \infty$ and $\mathbf{E}(X_1 \sqrt{n/m})^r$ converges to $\mathbf{E}Z^r$.
We remark that the adjacency relations in a random intersection graph are statistically dependent events. In particular, the clustering coefficient $\alpha = \alpha(G_{(n)}) = \mathbf{P}(v_1 \sim v_2 | v_1 \sim v_3, v_2 \sim v_3)$

of a sparse random intersection graph $G_{(n)}$ is bounded away from zero as n$\to +\infty$ provided that the second moment of the degree distribution is finite and the ratio $\beta_n = m/n$ is bounded ([7], [10]). More precisely, assuming that (i) and (ii-2) hold one obtains as $n \to +\infty$

$$\alpha = \beta_n^{-1/2}\delta_1^{3/2}(\delta_2 - \delta_1)^{-1} + o(1). \tag{6}$$

Here δ_r denotes the $r-$th moment of the asymptotic degree distribution. That is, we write $\delta_r = \mathbf{E}d_*^r$ for a random variable d_* having the distribution $\mathbf{P}(d_* = k) = \mathbf{E}e^{-z_1 Z}(z_1 Z)^k/k!$, $k = 0, 1, \ldots$. From (6) we see that $\beta_n \to +\infty$ implies $\alpha = o(1)$. For comparison, the (unconditional) edge probability $p_e = \mathbf{P}(v_1 \sim v_2) = \delta_1 n^{-1} + o(n^{-1})$ is of order $O(n^{-1})$ no matter whether β_n is bounded or not ([7]).
Theorems 1 and 2 show a first order asymptotics as $n \to +\infty$ of $cl(r)$ in the cases where β_n is bounded and $\beta_n \to +\infty$, respectively.

Theorem 1. *Let $m, n \to \infty$. Assume that (i), (ii-2) hold. Suppose that $\beta_n \to \beta \in (0, +\infty)$. Denote $\Lambda = \sqrt{\delta_1/\beta}$. We have*

$$cl(r) = \begin{cases} p_e^{-1}e^{-\Lambda}(1 + o(1)), & r = 0; \\ \frac{\alpha}{\alpha + (1-\alpha)e^{\Lambda}}(1 + o(1)), & r = 1; \\ 1 - o(1), & r \geq 2. \end{cases} \tag{7}$$

Here $p_e = n^{-1}\delta_1 + o(n^{-1})$ denotes the edge probability, $p_e = \mathbf{P}(v_1 \sim v_2)$.

We remark that for $r \geq 2$ the convergence to 1 in (7) can be quite slow, especially in the cases where the (asymptotic) average degree δ_1 is large. This can be seen from a more detailed expression for $cl(r)$, $r \geq 2$, which is obtained from the proof of Theorem 1,

$$cl(r) = \left(1 + p_e^{-1}f_r^{-1}(\Lambda)\mathbf{E}f_r(\Lambda')\right)^{-1} + O(n^{-1}). \tag{8}$$

Here $\Lambda' = n^{-1}(z_2 - \Lambda)Z_1 Z_2$ and $f_r(\lambda) = e^{-\lambda}\lambda^r/r!$ denotes the Poisson probability. We note that when Λ is large, the probability $f_r(\Lambda)$ is small for $r << \Lambda$ and, in this case, $cl(r)$ may deviate substantially from 1 even for comparatively large values of n, see also Fig. 4 and 5 below.

Theorem 2. *Let $m, n \to \infty$. Assume that (i), (ii-2) hold. Suppose that $\beta_n \to +\infty$. We have*

$$cl(r) = \begin{cases} p_e(1 + o(1)), & r = 0; \\ \alpha(1 + o(1)) + O(n^{-1}), & r = 1. \end{cases} \tag{9}$$

In particular, $cl(0) = O(n^{-1})$ and $cl(1) = o(1)$, see (6). Furthermore, we have

$$cl(2) = \begin{cases} 1 + o(1), & \textit{for} \quad \beta_n/n \to 0; \\ \frac{1}{1 + \beta_*(\delta_2 - \delta_1)^2\delta^{-4}} + o(1), & \textit{for} \quad \beta_n/n \to \beta_* \in (0, +\infty); \\ o(1), & \textit{for} \quad \beta_n/n \to +\infty. \end{cases} \tag{10}$$

Assuming, in addition, that $\beta_n^k = o(n)$ for each $k = 1, 2, 3 \ldots$, we obtain

$$cl(r) = 1 + o(1), \qquad \textit{for} \qquad r = 2, 3, \ldots. \tag{11}$$

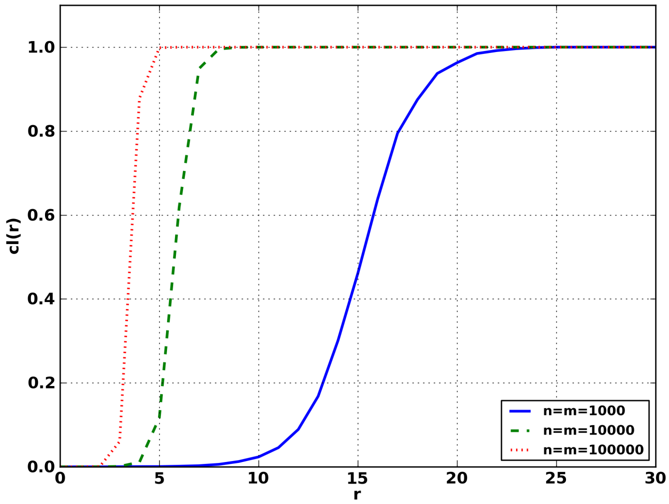

Figure 3: Convergence to the step function for random intersection graphs with all sets of size 10.

We conclude from (7), (9) that the edge dependence measures $cl(1)$ and α are, in fact, equivalent in the case of a random intersection graph. In particular, we have $cl(1) = 1 - o(1) \Leftrightarrow \alpha = 1 - o(1)$ and $cl(1) = o(1) \Leftrightarrow \alpha = o(1)$. Furthermore, (10) tels us that the parameter $cl(2)$ is able to distinguish between the cases $\beta_n = o(n)$ and $n = o(\beta_n)$. Finally, (11) tels us that any $c(r)$, $r = 1, 2, \ldots$ can't distinguish between sequences $\{\beta_n\}$ and $\{\beta'_n\}$ growing slower that any power of n (take $\beta_n = \ln n$ and $\beta'_n = \ln^2 n$, for example).
Remark 1. It is likely that (10) can be extended to an arbitrary r as follows

$$cl(r) = \begin{cases} 1 + o(1), & \text{for} \quad \beta_n / n^{4-2r^{-1}} \to 0; \\ c(r, \beta_*) + o(1), & \text{for} \quad \beta_n / n^{4-2r^{-1}} \to \beta_* \in (0, +\infty); \\ o(1), & \text{for} \quad \beta_n / n^{4-2r^{-1}} \to +\infty. \end{cases}$$

Here $c(r, \beta_*) = (\beta_*^{r/2} z_1^{-r-2} z_2^r z_r^2 + 1)^{-1}$. We note that numbers $z_i = \mathbf{E} Z^i$ can be expressed in terms of moments of the asymptotic degree distribution (5).
Proofs of Theorems 1 and 2 are given in the Appendix.
Fig. 3 illustrates the convergence to a step function shown by Theorem 1. Here we plot clustering function (1) of random intersection graphs $G_i = G(n_i, m_i, P)$, where $n_i = m_i = 10^3 5^{i-1}$, $i = 1, 2, 3$, and $P(10) = 1$.
Fig. 4 illustrates the influence of the expected degree δ_1 on the slope of the clustering function (1): the larger is δ_1 the more gradual is the slope. In Fig. 4 we plot (1) for random intersection graphs $G_i = G(n, m, P_i)$, where $n = m = 10^4$ and $P_i(3^i) = 1$, $i = 1, 2, 3$.
Inhomogeneous graph. The *inhomogeneous* random intersection graph $G_1(n, m, P_1, P_2)$ on the vertex set $V = \{v_1, \ldots, v_n\}$ is obtained as follows. We first generate independent random variables $A_1, \ldots, A_n, B_1, \ldots, B_m$ such that each A_i has the probability distribution P_1 and each B_j has the probability distribution P_2. Then, conditionally on the realized values $\{A_i, B_j\}_{i,j=1}^{n,m}$, we include the attribute $w_j \in W$ in the set D_i with probability $p_{ij} = \min\{1, A_i B_j (nm)^{-1/2}\}$ independently for each i and j (see [2], [8], [21]). We consider a sequence of inhomogeneous intersection graphs where P_1, P_2 remain fixed and $m = m_n$ and n tend to infinity. We remark

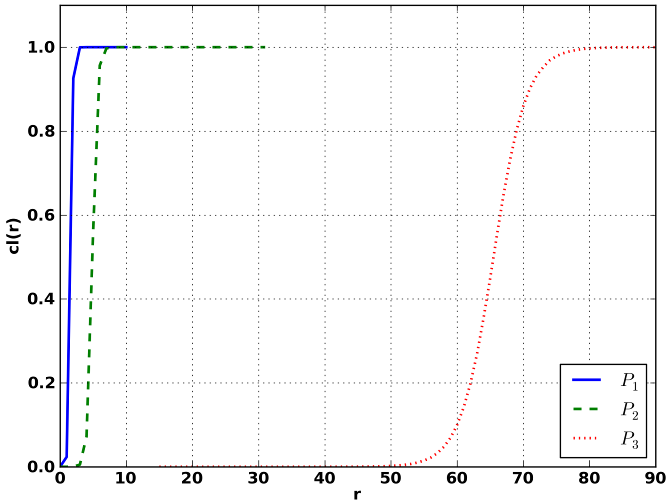

Figure 4: Clustering function of random intersection graphs with $n = m = 10000$ and $P_i(3^i) = 1$, $i = 1, 2, 3$.

that $\tilde{G}_n = G_1(n, m, P_1, P_2)$ is sparse and the edge probability $p_e := \mathbf{P}(v_1 \sim v_2) = a_1^2 b_2 n^{-1} + o(n^{-1})$. Here we denote $a_k = \mathbf{E}A_1^k$ and $b_k = \mathbf{E}B_1^k$.
Our next result shows that $\tilde{G}_n$ admits a nonvanishing clustering coefficient $\alpha = \alpha(\tilde{G}_n) = \mathbf{P}(v_1 \sim v_2 | v_1 \sim v_3, v_2 \sim v_3)$ in the case where $\beta_n = m/n$ is bounded and it is bounded away from zero as $n, m \to +\infty$. In addition, we show a first order asymptotics of the clustering function $cl(\cdot)$.

Theorem 3. *Let $m, n \to \infty$. Assume that $0 < \mathbf{E}A_1^2 < \infty$ and $0 < \mathbf{E}B_1^3 < \infty$. Suppose that $\beta_n \to \beta \in (0, +\infty)$. Then we have*

$$\alpha = \frac{b_3 \kappa}{b_3 \kappa + 0.5\sqrt{\beta}} + o(1) \tag{12}$$

and

$$cl(r) = \begin{cases} a_1^2 b_2^* n^{-1}(1 + o(1)), & r = 0; \\ \frac{b_3^* \kappa}{b_3^* \kappa + 0.5\sqrt{\beta}}(1 + o(1)), & r = 1; \\ 1 - o(1), & r \geq 2. \end{cases} \tag{13}$$

Here $\kappa = a_1 a_2^{-1} b_2^{-2}$ and $b_k^* = \mathbf{E}B_1^k e^{-a_1 B_1/\sqrt{\beta}}$.

The proof of Theorem 3 is given in the Appendix. We remark that the approach used in the proof applies to the case of $\beta_n \to +\infty$ as well. In particular, for $\beta_n \to +\infty$, we have $cl(0) = p_e(1+o(1))$ and $\alpha = o(1)$, $cl(1) = o(1)$. One can show, in addition, that

$$cl(2) = \begin{cases} 1 + o(1), & \text{for} \quad \beta_n/n \to 0; \\ \frac{b_4 \kappa'}{b_4 \kappa' + 0.25\beta} + o(1), & \text{for} \quad \beta_n/n \to \beta \in (0, +\infty); \\ o(1), & \text{for} \quad \beta_n/n \to +\infty. \end{cases} \tag{14}$$

Here $\kappa' = a_2^{-2} b_2^{-4}$.

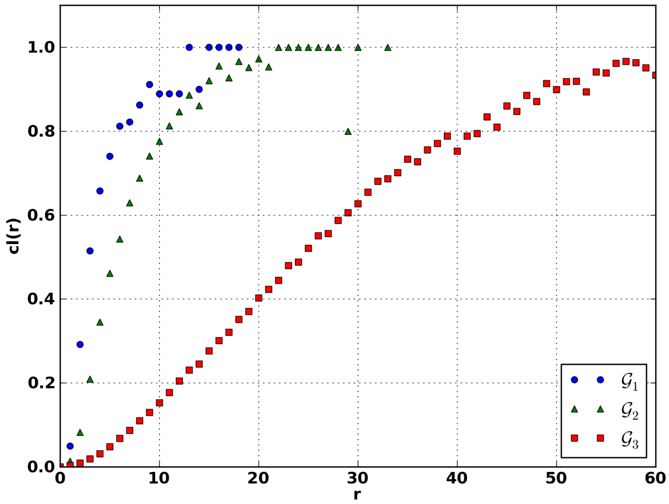

Figure 5: Sampling subgraphs with degree constraints.

4 Discussion

The first order asymptotics (7), (10), (11) and (13) suggests that the clustering function $cl(\cdot)$ of a (very) large affiliation network can be approximated by a step - like function. Furthermore, $cl(1)$ is closely related to the clustering coefficient.
Simulations in Fig. 3 and 4 show that the convergence in (7) can be rather slow and we observe a sigmoid function approximation of the step function. Furthermore, the larger is the average degree, the more remote is the "step" from the origin and the more gradual is the slope of the clustering function.
In order to learn about the influence of the inhomogeneity of the degree sequence on the slope of the clustering function $r \to cl(r)$ we select various subnetworks of real networks according to certain regularity conditions satisfied by their degree sequences. We observe that the inhomogeneity (heavy tail) of the degree sequence affects the slope of the clustering function: the heavier the tail the more gradual is the slope of the clustering function. We illustrate this observations in Fig. 5 and 6.
Fig. 5 plots clustering function (1) of subgraphs of the first university network (see Sect 2.) sampled as follows. $\mathcal{G}_1$ is the subgraph that includes all vertices of degree not larger than 50. It has $n_0 = 7165$ vertices. $\mathcal{G}_2$ is a subgraph induced by n_0 vertices drawn uniformly at random (without replacement) from the vertices of degree not larger than 150. $\mathcal{G}_3$ is a subgraph of induced by n_0 vertices drawn uniformly at random (without replacement) from the set of all vertices. Now all three graphs have the same number of vertices.
In Fig. 6 we plot two subgraphs of French actor network (data from [27]). The subgraph $\mathcal{G}_4$ is induced by the set of marked vertices obtained as follows: we put a mark on each vertex v with probability $d^{-\tau}(v)$ and independently of the other vertices. Choosing $\tau = 0.5$ we obtain a random subgraph denoted $\mathcal{G}_4$. In our case the realized number of marked vertices $n_1 = 8871$. $\mathcal{G}_5$ is the subgraph of French actor network induced by n_1 vertices drawn uniformly at random (without replacement) from the set of all vertices. Now both subgraphs have the same number of vertices, but the degree sequence of $\mathcal{G}_4$ is much more regular than that of $\mathcal{G}_5$.

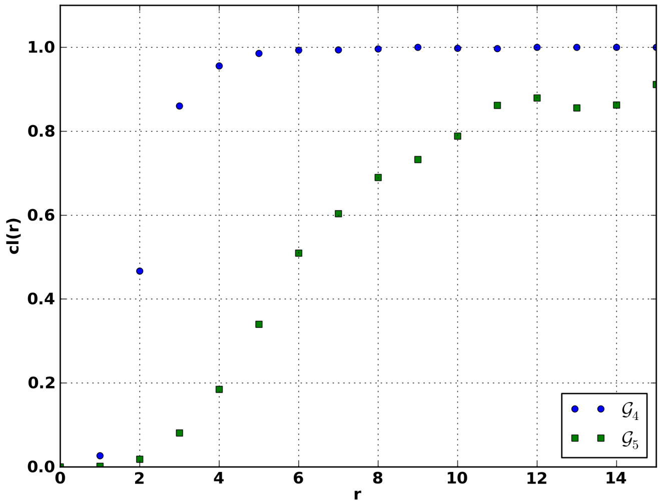

Figure 6: Sampling a subgraph with a random degree constraint.

Finally, we examine how well a random intersection graph fits the real data. For this purpose we consider a memoryless actor network obtained as follows. Assume every actor of a given actor graph has forgotten about the titles of movies he or she acted in, but remembers the number of movies.

We first simulate an instance of the *active* memoryless graph where each actor chooses films independently and uniformly at random from a given set of $\tilde{m}$ films so that the number of films chosen by each actor is the same as in the true actor graph. In the active memoryless graph all films have equal chances to be selected by any of actors. We remark that in the case where $\tilde{m} = m$, i.e., the number of films in the active memoryless graph is the same as in the real underlying actor network, the expected degree of the memoryless graph does not match the average degree of the real network. We can easily adjust the number of films (of the memoryless graph) so that these degrees match. We denote this number m' and call the active memoryless graph with $\tilde{m} = m'$ *adjusted* one. In Fig. 7. we plot clustering function (1) of two instances of memoryless graphs for comparison with the underlying French actor network: one with the true number of films and another with the adjusted number of films.

We secondly simulate an instance of the *inhomogeneous* memoryless graph where an actor v_i chooses the film w_j with probability $a_i b_j M^{-1}$ independently for each i and j. Here the numbers a_i, b_j are observed characteristics of the underlying actor network: v_i acted in a_i films; b_j actors acted in the film w_j. $M = \sum_{1 \le i \le n} a_i = \sum_{1 \le j \le m} b_j$ is the total number of links of the bipartite graph where actors are linked to films. In Fig. 8 we plot clustering function (1) of an instance of the inhomogeneous memoryless graph of the French actor network. Here we observe a remarkable accuracy of the approximation of the real clustering function by that of the memoryless graph.

5 Appendix

The section is organized as follows: we first collect some notation, then we prove Theorems 1, 2 and 3. At the very end we present two auxiliary results used in proofs: Lemmas 2 and 3.

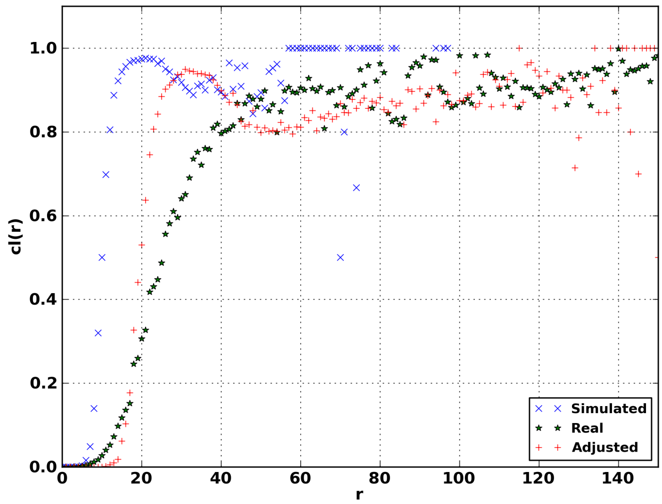

Figure 7: Real French actor network and two simulated active memoryless networks.

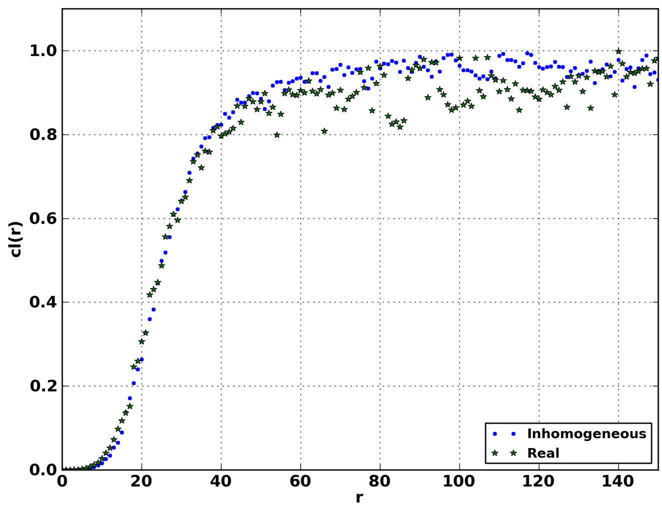

Figure 8: Real French actor network and a simulated inhomogeneous memoryless network.

By X_{ni} we denote the size of the set D_i in $G_{(n)}$. Furthermore, we denote $Z_{n1} = \beta_n^{-1/2} X_{n1}$ and $Z_{01} := Z$. Introduce the function

$$t \to \varphi(t) = \sup_{n \geq 0} \mathbf{E} Z_{n1}^2 \mathbb{I}_{\{Z_{n1} \geq t\}}.$$

We remark that conditions (i), (ii-2) imply $\varphi(t) = o(1)$ as $t \to +\infty$ (see [7]). By $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{E}}$ we denote the conditional probability and the conditional expectation given D_1, D_2. We introduce events $\mathcal{A} = \{v_1 \sim v_2\}$, $\mathcal{A}_i = \{|D_1 \cap D_2| = i\}$ and probabilities $p_i(r) = \mathbf{P}(\mathcal{A}_i \cap \{d_{12} = r\})$. By $f_r(\lambda) = e^{-\lambda}\lambda^r/r!$ we denote the Poisson probability.

Proof of Theorems 1 and 2. We have

$$cl(r) = \mathbf{P}(\mathcal{A}|d_{12} = r) = \frac{\mathbf{P}(\mathcal{A} \cap \{d_{12} = r\})}{\mathbf{P}(d_{12} = r)}. \tag{15}$$

In order to evaluate the numerator we write $\mathcal{A} = \cup_{i \geq 1} \mathcal{A}_i$ and apply the total probability formula

$$\mathbf{P}(\mathcal{A} \cap \{d_{12} = r\}) = \sum_{i \geq 1} p_i(r) = \sum_{1 \leq i \leq k} p_i(r) + R_k(r). \tag{16}$$

Here $R_k(r) = \sum_{i > k} p_i(r) \leq \mathbf{P}(|D_1 \cap D_2| \geq k+1)$. Similarly we obtain

$$\mathbf{P}(d_{12} = r) = \sum_{i \geq 0} p_i(r) = \sum_{0 \leq i \leq k} p_i(r) + R_k(r). \tag{17}$$

In order to prove Theorem 1 we choose $k = 1$ in (16), (17) and invoke the asymptotic expressions of $p_i(r)$ and the upper bound for $\mathbf{P}(|D_1 \cap D_2| \geq k+1)$ shown in Lemma 1.
Theorem 2 is obtained in the same way, but now we choose $k = 2$. □

Lemma 1. *Assume that $\beta_n \to \beta \in (0, +\infty]$. Suppose that (i), (ii-2) hold. Denote $\Lambda_1 = \beta_n^{-1/2} z_1$ and $\Lambda_2 = z_2 - \beta_n^{-1/2} z_1$. We have as $n \to +\infty$*

$$p_0(0) = 1 - o(1), \qquad p_0(r) = o(n^{-2}), \qquad r \geq 3, \tag{18}$$

$$p_0(r) = n^{-r}(r!)^{-1}\Lambda_2^r z_r^2 + o(n^{-r}), \qquad r = 1, 2, \tag{19}$$

$$p_1(r) = n^{-1} z_1^2 f_r(\Lambda_1) + \tau_r, \qquad r \geq 0 \tag{20}$$

$$p_2(r) = 2^{-1} n^{-2} f_r(2\Lambda_1)\Lambda_2^2 + \tau_2. \qquad r \geq 0. \tag{21}$$

Here $\tau_r = O(n^{-2})$, for $r = 0, 1$, and $\tau_r = O(n^{-2}\beta_n^{-1/2}) + o(n^{-2})$, for $r \geq 2$. Furthermore,

$$\mathbf{P}(|D_1 \cap D_2| \geq 3) = o(n^{-2}) \qquad \text{and} \qquad \mathbf{P}(|D_1 \cap D_2| \geq k) = O(n^{-k}), \quad k = 1, 2. \tag{22}$$

Proof of Lemma 1. Before the proof we introduce some notation. By $\tilde{\mathbf{P}}_i(\cdot) = \mathbf{P}(\cdot|\mathcal{A}_i, D_1, D_2)$ we denote the conditional probability given $\mathcal{A}_i$ and D_1, D_2. By $\mathbb{I}_{\mathcal{B}}$ we denote the indicator of an event $\mathcal{B}$ and write $\bar{\mathbb{I}}_{\mathcal{B}} = 1 - \mathbb{I}_{\mathcal{B}}$. In the proof we use several indicators

$$\mathbb{I} = \mathbb{I}_{\{X_1 + X_2 < \varepsilon^2 n \beta_n^{1/2}\}}, \quad \mathbb{I}_j = \mathbb{I}_{\{X_j < 0.5\varepsilon^2 n \beta_n^{1/2}\}}, \quad \mathbb{I}_{*j} = \mathbb{I}_{\{X_j \leq \beta_n^{1/2}\varepsilon^{-1}\}},$$
$$\mathbb{I}_{\star j} = \mathbb{I}_{\{X_j \leq \varepsilon m\}}, \qquad \mathbf{I}_j = \mathbb{I}_{\{X_j < m^{1/2} n^{-1/4}\}}, \quad \mathbf{I}_{*j} = \mathbb{I}_{\{X_j \leq 0.5m\}}.$$

Some of them depend on $\varepsilon > 0$, value of which will be clear from the context. We will use the following simple properties of the Poisson probability $\lambda \to f_r(\lambda)$. It follows from the mean value theorem $f_r(t) - f_r(s) = f_r'(\xi)(t-s)$, where $0 < s \le \xi \le t$, and the inequalities $|f_r'(\xi)| \le 1$ and $|f_{2+r}'(\xi)| \le \xi$ that

$$|f_r(s) - f_r(t)| \le |s-t| \qquad \text{and} \qquad |f_{2+r}(s) - f_{2+r}(t)| \le (s+t)|s-t|. \tag{23}$$

Now we outline the proof. In order to evaluate $p_i(r)$ we write

$$p_i(r) = \mathbf{E}\tilde{\mathbf{P}}(\mathcal{A}_i \cap \{d_{12} = r\}) = \mathbf{E}\tilde{p}_i(r)\tilde{\mathbf{P}}(\mathcal{A}_i), \tag{24}$$

where $\tilde{p}_i(r) = \tilde{\mathbf{P}}_i(d_{12} = r)$. Next we approximate $\tilde{\mathbf{P}}(\mathcal{A}_i)$ using (73) and apply the Poisson approximation to $\tilde{p}_i(r)$. We observe that, given $\mathcal{A}_i$, D_1, D_2, the random variable

$$d_{12} = \sum_{3 \le j \le n} \mathbb{I}_{\{v_1 \sim v_j\}} \mathbb{I}_{\{v_2 \sim v_j\}}$$

has binomial distribution $\mathbf{Bin}(n-2, q_i)$, where q_i is the probability that D_3 intersects with both sets D_1 and D_2. Hence, $\tilde{p}_i(r)$ can be approximated by the Poisson probability $f_r(\lambda_i)$, where $\lambda_i = (n-2)q_i$. Finally, we approximate λ_i by $\tilde{\lambda}_i = n\tilde{q}_i$ and $f_r(\lambda_i)$ by $f_r(\tilde{\lambda}_i)$. Here

$$\tilde{q}_0 = n^{-2}(z_2 - \beta_n^{-1/2} z_1) Z_1 Z_2, \qquad \tilde{q}_1 = n^{-1}\beta_n^{-1/2} z_1, \qquad \tilde{q}_2 = 2n^{-1}\beta_n^{-1/2} z_1 \tag{25}$$

are approximations of q_1, q_2 and q_3 respectively. In order to obtain an upper bound for the error of such approximation we write

$$\Delta_{r,i} := \tilde{p}_i(r) - f_r(\tilde{\lambda}_i) = \Delta'_{r,i} + \Delta''_{r,i}, \tag{26}$$

where $\Delta'_{r,i} = \tilde{p}_i(r) - f_r(\lambda_i)$ and $\Delta''_{r,i} = f_r(\lambda_i) - f_r(\tilde{\lambda}_i)$, and estimate

$$|\Delta_{r,i}| \le 2nq_i^2, \qquad i, r = 0, 1, 2, \dots, \tag{27}$$

using LeCam's lemma (see Lemma 2), and estimate $\Delta''_{r,i}$ using the mean value theorem, see (23). Now we briefly explain approximations (25). Let $\{w_1^*, \dots, w_i^*\}$ denote the intersection $D_1 \cap D_2$ provided it is non empty. We write $q_i = \tilde{\mathbf{P}}_i(n_1 \ge 1, n_2 \ge 1)$, where $n_j = |D_3 \cap D_j|$, $j = 1, 2$, and split

$$q_0 = q_{01} + q_{02}, \qquad q_1 = q_{11} + q_{12}, \qquad q_2 = q_{21} + q_{22} + q_{23} + q_{24},$$

where

$$\begin{aligned}
q_{01} &= \tilde{\mathbf{P}}_0(n_1 = 1, n_2 = 1), & q_{02} &= \tilde{\mathbf{P}}_0(n_1 + n_2 \ge 3, n_1 \ge 1, n_2 \ge 1),\\
q_{11} &= \tilde{\mathbf{P}}_1(w_1^* \in D_3), & q_{12} &= \tilde{\mathbf{P}}_1(w_1^* \notin D_3, n_1 \ge 1, n_2 \ge 1),\\
q_{21} &= \tilde{\mathbf{P}}_2(w_1^* \in D_3, w_2^* \notin D_3), & q_{22} &= \tilde{\mathbf{P}}_2(w_1^* \notin D_3, w_2^* \in D_3),\\
q_{23} &= \tilde{\mathbf{P}}_2(w_1^*, w_2^* \in D_3), & q_{24} &= \mathbf{P}_2^*(w_1^*, w_2^* \notin D_3, n_1 \ge 1, n_2 \ge 1).
\end{aligned} \tag{28}$$

Approximations $q_i \approx \tilde{q}_i$, see (25), are obtained as follows. We have $q_0 \approx q_{01} \approx \tilde{q}_0$, $q_1 \approx q_{11} = \tilde{q}_1$ and $q_2 \approx q_{21} + q_{22} \approx \tilde{q}_2$.

Proof of (18), (19). In order to prove (18), (19) we show that

$$\mathbf{E}\Delta_{r,0}\tilde{\mathbf{P}}(\mathcal{A}_0) = o(n^{-r\wedge 2}), \qquad r \ge 0 \tag{29}$$

$$\mathbf{E}f_r(\tilde{\lambda}_0)\tilde{\mathbf{P}}(\mathcal{A}_0) = (r!)^{-1}\mathbf{E}\tilde{\lambda}_0^r + o(n^{-r\wedge 2}). \qquad r = 0, 1, 2, \tag{30}$$

$$\mathbf{E}f_r(\tilde{\lambda}_0)\tilde{\mathbf{P}}(\mathcal{A}_0) = o(n^{-2}), \qquad r \ge 3. \tag{31}$$

We firstly prove (29). In the case where $\beta < \infty$ we find $n_0 > 0$ such that $\beta < 2\beta_n$ for $n \geq n_0$. In the case where $\beta = +\infty$ we find n_0 such that $\beta_n > 1$ for $n \geq n_0$. In order to prove (29) we show that for any $0 < \varepsilon < \min\{0.5\beta^{1/2}, 0.1\}$ and $n \geq n_0$ we have

$$\mathbf{E}|\Delta_{r,0}| \leq c_* n^{-3} + c_* n^{-r\wedge 2} R_1(\varepsilon) + 4n^{-2}\varepsilon^{-4} R_2(\varepsilon), \tag{32}$$

where

$$R_1(\varepsilon) = 2\varphi(\varepsilon^{-1}) + 4z_2(\varepsilon + m^{-1} + n^{-1}) \qquad \text{and} \qquad R_2(\varepsilon) = \varphi(\varepsilon n)(1 + 4\varepsilon^{-4} n^{-2} z_2).$$

Here and below c_* denotes a constant independent of n, m and ε. Observe that $\lim_{t\to+\infty}\varphi(t) = 0$ implies that the left-hand side of (32) is $o(n^{-r\wedge 2})$. Now we fix ε and prove (32). For this purpose we write $\Delta_{r,0} = \Delta_{r,0}\mathbb{I} + \Delta_{r,0}\bar{\mathbb{I}}$ and show the inequalities

$$\mathbf{E}|\Delta_{r,0}\mathbb{I}| \leq c_* n^{-3} + c_* n^{-r\wedge 2} R_1(\varepsilon), \tag{33}$$

$$\mathbf{E}|\Delta_{r,0}\bar{\mathbb{I}}| \leq \mathbf{E}\bar{\mathbb{I}} \leq 4n^{-2}\varepsilon^{-4} R_2(\varepsilon). \tag{34}$$

The first inequality of (34) is obvious. In order to prove the second one we combine the inequalies

$$\varepsilon^4 n^2 \mathbf{E}\bar{\mathbb{I}} \leq \beta_n^{-1}\mathbf{E}(X_1 + X_2)^2\bar{\mathbb{I}} \leq 2\beta_n^{-1}\mathbf{E}(X_1^2 + X_2^2)\bar{\mathbb{I}} = 4\beta_n^{-1}\mathbf{E}X_1^2\bar{\mathbb{I}},$$

which follow from Markov's inequality, with the inequalities

$$\beta_n^{-1}\mathbf{E}X_1^2\bar{\mathbb{I}} \leq \beta_n^{-1}\mathbf{E}X_1^2(\bar{\mathbb{I}}_1 + \bar{\mathbb{I}}_2) \leq \beta_n^{-1}\mathbf{E}X_1^2(\bar{\mathbb{I}}_1 + 4\varepsilon^{-4} n^{-2}\beta_n^{-1} X_2^2 \bar{\mathbb{I}}_2) \leq R_2(\varepsilon).$$

Here we applied the inequality $\bar{\mathbb{I}} \leq \bar{\mathbb{I}}_1 + \bar{\mathbb{I}}_2$ and then Markov's inequality.
In order to prove (33) we write $\Delta_{r,0}\mathbb{I} = \Delta'_{r,0}\mathbb{I} + \Delta''_{r,0}\mathbb{I}$, see (26), and invoke the inequalities

$$\mathbf{E}|\Delta'_{r,0}|\mathbb{I} \leq c_* n^{-3} \qquad \text{and} \qquad \mathbf{E}|\Delta''_{r,0}|\mathbb{I} \leq c_* n^{-r\wedge 2} R_1(\varepsilon). \tag{35}$$

The first inequality of (35) follows from (27) and the inequalities $q_0^2 \leq 2\tilde{q}_0^2 + 2(q_0 - \tilde{q}_0)^2$ and

$$|q_0\mathbb{I} - \tilde{q}_0\mathbb{I}| \leq n^{-1} m^{-1} X_1 X_2 R_1(\varepsilon). \tag{36}$$

The second inequality of (35) follows from (23) and (36).
We complete the proof of (29) by showing (36). We note that for $n \geq n_0$ we have $\varepsilon\beta_n^{-1/2} < 1$. In particular, the inequality $X_1 + X_2 < \varepsilon^2 n \beta_n^{1/2}$ implies $X_1 + X_2 < \varepsilon m$. We shall show that

$$(1 - 3m^{-1})\tilde{q}_0 \leq q_{01} \leq (1 + 2\varepsilon)\tilde{q}_0, \tag{37}$$

$$q_{02} \leq 2n^{-1} m^{-1} X_1 X_2 (\varphi(\varepsilon^{-1}) + 2\varepsilon z_2). \tag{38}$$

These inequalities imply (36). In order to prove (37) we write $q_{01} = \tilde{\mathbf{E}} q_{01}^*$, where $q_{01}^* = \tilde{\mathbf{P}}_0(n_1 = 1, n_2 = 1 | X_3)$, and show that

$$\varkappa_1 \varkappa_2 (1 - 3m^{-1}) \leq q_{01}^* \leq \varkappa_1 \varkappa_2 (1 + 2\varepsilon), \tag{39}$$

where $\varkappa_1 = m^{-1} X_1 X_3$ and $\varkappa_2 = m^{-1} X_2 (X_3 - 1)$. We have

$$q_{01}^* = \tau_1 \tau_2, \qquad \tau_1 = \tilde{\mathbf{P}}_0(n_1 = 1 | X_3), \qquad \tau_2 = \tilde{\mathbf{P}}_0(n_2 = 1 | n_1 = 1, X_3). \tag{40}$$

Combining the inequalities, which follow from (73),

$$\varkappa_1(1 - (m - X_1)^{-1}) \leq \tau_1 \leq \varkappa_1,$$

$$\varkappa_2(1 - (m - X_1 - X_2)^{-1}) \leq \tau_2 \leq \varkappa_2 m (m - X_1)^{-1}$$

with the inequality $X_1+X_2<\varepsilon m$ we obtain (39).
Now we prove (38). We split $q_{02}=q_{03}+q_{04}$, where $q_{03}=\tilde{\mathbf{P}}_0(n_1\ge 2, n_2\ge 1)$ and $q_{04}=\tilde{\mathbf{P}}_0(n_1\ge 1, n_2\ge 2)$ and construct upper bounds for q_{03} and q_{04}. Both quantities are estimated in the same way. We only consider q_{03}. We denote $p_{k*}=\tilde{\mathbf{P}}(n_1\ge k, n_2\ge 1|X_3)$, $k=1,2$, and write

$$q_{03}=\tilde{\mathbf{E}}p_{2*}=\tilde{\mathbf{E}}p_{2*}(\mathbb{I}_{*3}+\bar{\mathbb{I}}_{*3})\le\tilde{\mathbf{E}}p_{2*}\mathbb{I}_{*3}+\tilde{\mathbf{E}}p_{1*}\bar{\mathbb{I}}_{*3}. \tag{41}$$

Here we use the inequality $p_{2*}\le p_{1*}$. Next, proceeding as in the proof of the right hand side inequality of (39) we obtain $p_{1*}\le\varkappa_1\varkappa_2(1+2\varepsilon)\le 2\varkappa_1\varkappa_2$. Hence, we have $\tilde{\mathbf{E}}p_{1*}\bar{\mathbb{I}}_{*3}\le 2n^{-1}m^{-1}X_1X_2\varphi(\varepsilon^{-1})$. In order to estimate $\tilde{\mathbf{E}}p_{2*}\mathbb{I}_{*3}$ we write

$$p_{2*}=\tau_{1*}\tau_{2*},\qquad \tau_{1*}=\tilde{\mathbf{P}}_0(n_1=2|X_3),\qquad \tau_{2*}=\tilde{\mathbf{P}}_0(n_2=1|n_1=2,X_3)$$

and invoke the inequalities, which follow from (73),

$$\tau_{1*}\le 2^{-1}m^{-2}X_1^2X_3^2,\quad \tau_{2*}\le(m-X_1)^{-1}X_2X_3.$$

We note that $X_1\le X_1+X_2<\varepsilon^2n\beta_n^{1/2}<\varepsilon m$ imply $\tau_{1*}\tau_{2*}\le\varepsilon^2n\beta_n^{1/2}m^{-3}X_1X_2X_3^3$. Furthermore, using the inequality $X_3\mathbb{I}_{*3}\le\varepsilon^{-1}\beta_n^{1/2}$ we write

$$\tilde{\mathbf{E}}p_{2*}\mathbb{I}_{*3}\le\tilde{\mathbf{E}}\tau_{1*}\tau_{2*}\mathbb{I}_{*3}\le\varepsilon\beta_nnm^{-3}X_1X_2\mathbf{E}X_3^2.$$

Invoking in (41) the upper bounds for $\tilde{\mathbf{E}}p_{2*}\mathbb{I}_{*3}$ and $\tilde{\mathbf{E}}p_{1*}\bar{\mathbb{I}}_{*3}$ we obtain (38). Proof of (36) is complete.
We secondly prove (30). For this purpose we write

$$f_r(\tilde{\lambda}_0)\tilde{\mathbf{P}}(\mathcal{A}_0)=(r!)^{-1}\tilde{\lambda}_0^r\tilde{\mathbf{P}}(\mathcal{A}_0)+R_{01}=(r!)^{-1}\tilde{\lambda}_0^r+R_{01}+R_{02}. \tag{42}$$

Here in the first step we apply the inequality $1-e^{-\tilde{\lambda}_0}\le\tilde{\lambda}_0$ and in the second step we apply the inquality, see (73),

$$1-\tilde{\mathbf{P}}(\mathcal{A}_0)=\tilde{\mathbf{P}}(D_1\cap D_2\ne\emptyset)\le X_1X_2m^{-1}. \tag{43}$$

Hence, $|R_{01}|\le\tilde{\lambda}_0^{r+1}$ and $|R_{02}|\le\tilde{\lambda}_0^rX_1X_2m^{-1}$. Now, for $r=0,1$, we obtain (30) from (42) using the simple bounds $\mathbf{E}|R_{0i}|=O(n^{-r-1})$, $i=1,2$. In the case where $r=2$ we invoke the truncation argument. It follows from the inequalities

$$\mathbf{I}_1\mathbf{I}_2\le 1\le\mathbf{I}_1\mathbf{I}_2+\bar{\mathbf{I}}_1+\bar{\mathbf{I}}_2 \tag{44}$$

that

$$f_2(\tilde{\lambda}_0)\tilde{\mathbf{P}}(\mathcal{A}_0)=\mathbf{I}_1\mathbf{I}_2f_2(\tilde{\lambda}_0)\tilde{\mathbf{P}}(\mathcal{A}_0)+R_{03},\qquad\tilde{\lambda}_0^2=\mathbf{I}_1\mathbf{I}_2\tilde{\lambda}_0^2+R_{04}, \tag{45}$$

where $|R_{0j}|\le\tilde{\lambda}_0^2(\bar{\mathbf{I}}_1+\bar{\mathbf{I}}_2)$ are negligibly small, i.e., $\mathbf{E}|R_{0j}|\le c_*n^{-2}\varphi(n^{1/4})=o(n^{-2})$, for $j=3,4$. Now combining (42) and (45), and invoking the bounds

$$\mathbf{E}|R_{0j}\mathbf{I}_1\mathbf{I}_2|\le c_*\mathbf{E}(X_1X_2m^{-1})^3\mathbf{I}_1\mathbf{I}_2\le c_*n^{-1/2}\mathbf{E}(X_1X_2m^{-1})^2=O(n^{-5/2}),\qquad j=1,2$$

we obtain (30) for $r=2$.
Let us prove (31). We write

$$\mathbf{E}f_r(\tilde{\lambda}_0)\tilde{\mathbf{P}}(\mathcal{A}_0)\le\mathbf{E}f_r(\tilde{\lambda}_0)\le\mathbf{E}f_r(\tilde{\lambda}_0)(\mathbf{I}_1\mathbf{I}_2+\bar{\mathbf{I}}_1+\bar{\mathbf{I}}_2)$$

and apply the inequalities $f_r(t)\le t^jf_{r-j}(t)\le t^j$, $0\le j\le r$. For $r\ge 3$ we obtain

$$\mathbf{E}f_r(\tilde{\lambda}_0)\mathbf{I}_1\mathbf{I}_2\le\mathbf{E}\tilde{\lambda}_0^3\mathbf{I}_1\mathbf{I}_2\le c_*n^{-1/2}\mathbf{E}\tilde{\lambda}_0^2=O(n^{-5/2}),$$
$$\mathbf{E}f_r(\tilde{\lambda}_0)(\bar{\mathbf{I}}_1+\bar{\mathbf{I}}_2)\le\mathbf{E}\tilde{\lambda}_0^2(\bar{\mathbf{I}}_1+\bar{\mathbf{I}}_2)\le c_*n^{-2}\varphi(n^{1/4})=o(n^{-2}).$$

Proof of (20), (21). Given a sequence of random variables $\{Y_n\}$ and $r \geq 0$ we write $Y_n \prec \tau_r$ to denote the fact that $\mathbf{E}|Y_n| = O(n^{-2})$, for $r \in \{0, 1\}$, and $\mathbf{E}|Y_n| = O(n^{-2}\beta_n^{-1/2}) + o(n^{-2})$, for $r \geq 2$. In the proof we use the following inequalities, which are obtained from (44) and (73),

$$\tilde{\mathbf{P}}(\mathcal{A}_i) \leq \varkappa_i, \tag{46}$$

$$\tilde{\mathbf{P}}(\mathcal{A}_i) \leq \tilde{\mathbf{P}}(\mathcal{A}_i)(\mathbf{I}_1\mathbf{I}_2 + \bar{\mathbf{I}}_1 + \bar{\mathbf{I}}_2) \leq \varkappa_i \mathbf{I}_1\mathbf{I}_2 + \bar{\mathbf{I}}_1 + \bar{\mathbf{I}}_2 \leq n^{-i/2} + \bar{\mathbf{I}}_1 + \bar{\mathbf{I}}_2. \tag{47}$$

Here $\varkappa_i = (X_1)_i(X_2)_i/(i!(m)_i)$ and $i = 1, 2$. We remark that (20), (21) follows from the bounds

$$\Delta_{r,i}\tilde{\mathbf{P}}(\mathcal{A}_i) \prec \tau_{r\vee i}, \tag{48}$$

$$f_r(\tilde{\lambda}_i)\tilde{\mathbf{P}}(\mathcal{A}_i) - f_r(\tilde{\lambda}_i)\varkappa_i \prec \tau_{r\vee i}, \qquad i = 1, 2. \tag{49}$$

Let us show (48). For this purpose we write

$$\Delta_{r,i} = \Delta_{r,i}\mathbf{I}_{*1} + \Delta_{r,i}\bar{\mathbf{I}}_{*1} = \Delta'_{r,i}\mathbf{I}_{*1} + \Delta''_{r,i}\mathbf{I}_{*1} + \Delta_{r,i}\bar{\mathbf{I}}_{*1}$$

and invoke the bounds

$$\Delta'_{r,i}\mathbf{I}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_i) \prec \tau_{r\vee i}, \qquad \Delta''_{r,i}\mathbf{I}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_i) \prec \tau_{r\vee i}, \qquad \Delta_{r,i}\bar{\mathbf{I}}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_i) \prec \tau_2. \tag{50}$$

We note that the third bound of (50) follows by Markov's inequality

$$\mathbf{E}\Delta_{r,i}\bar{\mathbf{I}}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_i) \leq \mathbf{E}\bar{\mathbf{I}}_{*1} \leq (4nm)^{-1}\varphi(0.5\sqrt{nm}) = o(n^{-2}).$$

Next we establish the first and second bound of (50) in the case where $i = 1$. The first bound of is obtained from (27) using the simple inequality $q_1^2 \leq 2q_{11}^2 + 2q_{12}^2$ (recall that $q_{11} = \tilde{q}_1$) and the inequality, which is shown below,

$$q_{12}\mathbf{I}_{*1} \leq 2n^{-1}m^{-1}z_2X_1X_2. \tag{51}$$

More precisely, we note that (51) implies $\mathbf{E}nq_{12}^2\mathbf{I}_{*1} = O(n^{-3})$ and proceed as follows

$$\Delta'_{r,1}\mathbf{I}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_1) \leq 2nq_1^2\mathbf{I}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_1) \leq 4n\tilde{q}_1^2\tilde{\mathbf{P}}(\mathcal{A}_1) + 4nq_{12}^2\mathbf{I}_{*1} \prec \tau_r.$$

In the last step we used the inequality, see (46), $\mathbf{E}n\tilde{q}_1^2\tilde{\mathbf{P}}(\mathcal{A}_1) \leq c_*n^{-2}\beta_n^{-1}$.
In the proof of the second bound of (50) we combine (23), (51) and apply the simple inequalities

$$|\lambda_1 - \tilde{\lambda}_1| \leq 2\tilde{q}_1 + nq_{12}, \qquad \lambda_1 + \tilde{\lambda}_1 \leq 2n\tilde{q}_1 + nq_{12}.$$

In particular, for $r = 0, 1$ we have

$$\mathbf{E}\Delta''_{r,1}\mathbf{I}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_1) \leq \mathbf{E}|\lambda_1 - \tilde{\lambda}_1|\mathbf{I}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_1) \leq 2\mathbf{E}\tilde{q}_1\tilde{\mathbf{P}}(\mathcal{A}_1) + \mathbf{E}nq_{12}\mathbf{I}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_1) \prec \tau_r.$$

In the last step we invoked the bounds $\mathbf{E}\tilde{q}_1\tilde{\mathbf{P}}(\mathcal{A}_1) = O(n^{-2})$ and $\mathbf{E}nq_{12}\mathbf{I}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_1) = O(n^{-2})$, which follow from (46), (51). For $r \geq 2$ we have

$$\begin{aligned}\Delta''_{r,1}\mathbf{I}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_1) &\leq |\lambda_1 - \tilde{\lambda}_1|(\lambda_1 + \tilde{\lambda}_1)\mathbf{I}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_1) \\ &\leq 4(n\tilde{q}_1^2 + n^2\tilde{q}_1q_{12} + n^2q_{12}^2)\mathbf{I}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_1) \\ &\prec \tau_r.\end{aligned}$$

Here in the last step we invoked the bounds, which follow from (46), (51),

$$\mathbf{E}n\tilde{q}_1^2\tilde{\mathbf{P}}(\mathcal{A}_1) = O(n^{-2}\beta_n^{-1}), \qquad \mathbf{E}n^2\tilde{q}_1q_{12}\mathbf{I}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_1) = O(n^{-2}\beta_n^{-1/2})$$

and the inequalities, see (51), (47),

$$\mathbf{E}n^2q_{12}^2\mathbf{I}_{*1}\tilde{\mathbf{P}}(\mathcal{A}_1) \le \mathbf{E}n^2q_{12}^2\mathbf{I}_{*1}(n^{-1/2}+\bar{\mathbf{I}}_1+\bar{\mathbf{I}}_2) \le c_*n^{-5/2}+c_*n^{-2}\varphi(n^{1/4}).$$

We complete the proof of (50), in the case where $i=1$, by showing (51). We obtain (51) from the inequalities, see (73),

$$\begin{aligned} q_{12} &\le \tilde{\mathbf{P}}_1(n_1\ge 1, n_2\ge 1) = \tilde{\mathbf{E}}\tilde{\mathbf{P}}_1(n_1\ge 1, n_2\ge 1|X_3) = \tilde{\mathbf{E}}\tau_1'\tau_2', \\ \tau_1' : &= \tilde{\mathbf{P}}_1(n_1\ge 1|X_3) \le m^{-1}X_1X_3, \\ \tau_2' : &= \tilde{\mathbf{P}}_1(n_2\ge 1|n_1\ge 1, X_3) \le (m-X_1)^{-1}X_2X_3. \end{aligned}$$

Now we establish the first two bounds of (50) in the case where $i=2$. The first bound is obtained from (27) using the simple inequality $q_2^2 \le 2q_{21}^2+2q_{22}^2+2q_{23}^2+2q_{24}^2$ and the upper bounds for q_{2j}, $j=1,2,3,4$ shown below. In particular, we have $q_{21}=q_{22}$ and

$$q_{22} = \tilde{\mathbf{E}}\tilde{\mathbf{P}}_2(w_2^*\in D_3|w_1^*\notin D_3, X_3)\tilde{\mathbf{P}}(w_1^*\notin D_3, X_3) = \tilde{\mathbf{E}}\frac{X_3}{m-1}\left(1-\frac{X_3}{m}\right) = \frac{\tilde{q}_2}{2}+R_{21}, \qquad (52)$$

where $|R_{21}|\le c_*n^{-2}(\beta_n^{-3/2}+\beta_n^{-1})$. Furthermore, we have, see (73),

$$\begin{aligned} q_{23} &= \tilde{\mathbf{E}}(X_3)_2/(m)_2 \le c_*n^{-2}, && (53)\\ q_{24}\mathbf{I}_{*1} &\le c_*(nm)^{-1}X_1X_2. && (54)\end{aligned}$$

Here (54) is shown in the same way as (51) above. From (27), (46) we obtain

$$\Delta'_{r,2}\tilde{\mathbf{P}}(\mathcal{A}_2)\mathbf{I}_{*1} \le 2nq_2^2\tilde{\mathbf{P}}(\mathcal{A}_2)\mathbf{I}_{*1} \le 4n(q_{21}^2+q_{22}^2+q_{23}^2)\varkappa_2+4nq_{24}^2\mathbf{I}_{*1} \prec \tau_2.$$

In the last step we used the bounds, which follow from (52), (53) and (54),

$$n(q_{21}^2+q_{22}^2+q_{23}^2)\varkappa_2 \prec \tau_2 \qquad \text{and} \qquad q_{24}^2\mathbf{I}_{*1}\prec\tau_2.$$

Let us show the second bound of (50). Since the absolute value of

$$\lambda_2-\tilde{\lambda}_2 = (n-2)(q_{23}+q_{24}+(q_{21}+q_{22}-\tilde{q}_2))-2\tilde{q}_2$$

is bounded from above by $n(q_{23}+q_{24}+2R_{21})+2\tilde{q}_2$, see (53), we obtain from (23) that

$$\Delta''_{r,2}\tilde{\mathbf{P}}(\mathcal{A}_2)\mathbf{I}_{*1} \le n(q_{23}+2R_{21})\varkappa_2+2\tilde{q}_2\varkappa_2+nq_{24}\mathbf{I}_{*1}\varkappa_2^{1/2}(n^{-1}+\bar{\mathbf{I}}_1+\bar{\mathbf{I}}_2)\prec\tau_2.$$

Here in the first step we applied (46) and the inequality $(\tilde{\mathbf{P}}(\mathcal{A}_2))^2\le\varkappa_2(n^{-1}+\bar{\mathbf{I}}_1+\bar{\mathbf{I}}_2)$, which follows from (46) and (47). In the second step we used (52), (53), (54).
Finally, we prove (49). In the case where $i=1$ it suffices to show that

$$0\le\mathbf{E}|m^{-1}X_1X_2-\tilde{\mathbf{P}}(\mathcal{A}_1)|\le c_*n^{-2}(1+\beta_n^{-1/2}). \qquad (55)$$

We derive (55) from the inequalities

$$m^{-1}X_1X_2\ge\tilde{\mathbf{P}}(\mathcal{A}_1)\ge\tilde{\mathbf{P}}(\mathcal{A}_1)\mathbf{I}_{*1}\ge m^{-1}X_1X_2(1-R_{11}). \qquad (56)$$

Here

$$R_{11}=\bar{\mathbf{I}}_{*1}+(m-X_1)^{-1}X_1X_2\mathbf{I}_{*1}\le 2m^{-1}X_1+2m^{-1}X_1X_2.$$

We note that the first and third inequality of (56) follow from (73). Indeed (73) implies

$$\tilde{\mathbf{P}}(\mathcal{A}_1) \geq m^{-1} X_1 X_2 (1 - (m - X_1)^{-1} X_1 X_2)$$

and we have $(1 - (m - X_1)^{-1} X_1 X_2)\mathbf{I}_{*1} = 1 - R_{11}$.
Now we prove (49) in the case where $i = 2$. For this purpose we show that for every $0 < \varepsilon < 0.5$

$$\mathbf{E}|\varkappa_2 - \tilde{\mathbf{P}}(\mathcal{A}_2)| \leq \varepsilon \mathbf{E}\varkappa_2 + c_* n^{-2} \varphi(\varepsilon\sqrt{n}). \tag{57}$$

We note that (57) follows from the inequalities

$$\varkappa_2 \geq \tilde{\mathbf{P}}(\mathcal{A}_2) \geq \tilde{\mathbf{P}}(\mathcal{A}_2)\mathbb{I}_{\star 1}\mathbb{I}_{\star 2} \geq \varkappa_2 \left(1 - \frac{X_1 X_2}{m - X_1}\right) \mathbb{I}_{\star 1}\mathbb{I}_{\star 2} \geq \varkappa_2 (1 - \varepsilon - \bar{\mathbb{I}}_{\star 1} - \bar{\mathbb{I}}_{\star 2}).$$

We obtain these inequalities combining the following ones

$$\begin{aligned} &\mathbb{I}_{*1}\mathbb{I}_{*2} \leq 1 \leq \mathbb{I}_{*1}\mathbb{I}_{*2} + \bar{\mathbb{I}}_{*1} + \bar{\mathbb{I}}_{*2}, \\ &\varkappa_2 \geq \tilde{\mathbf{P}}(\mathcal{A}_2) \geq \varkappa_2 \left(1 - \frac{X_1 X_2}{m - X_1}\right). \end{aligned}$$

Here the first line is obvious, the second line follows from (73).
Proof of (22). In the proof we apply (73). We have, see (46),

$$\tilde{\mathbf{P}}(|D_1 \cap D_2| \geq k) \leq \varkappa_k \leq Z_1^k Z_2^k n^{-k}. \tag{58}$$

Taking the expected values in (58) we obtain (22) for $k = 1, 2$. For $k = 3$ we write, see (44), $\tilde{\mathbf{P}}(|D_1 \cap D_2| \geq 3) \leq n^{-3/2} + \bar{\mathbf{I}}_1 + \bar{\mathbf{I}}_2$ and

$$\tilde{\mathbf{P}}(|D_1 \cap D_2| \geq 3) = \tilde{\mathbf{P}}^{2/3}(|D_1 \cap D_2| \geq 3)\tilde{\mathbf{P}}^{1/3}(|D_1 \cap D_2| \geq 3) \leq \varkappa_3^{2/3}(n^{-3/2} + \bar{\mathbf{I}}_1 + \bar{\mathbf{I}}_2)^{1/3}.$$

Hence, $\mathbf{P}(|D_1 \cap D_2| \geq 3) \leq \mathbf{E}\varkappa_3^{2/3}(n^{-3/2} + \bar{\mathbf{I}}_1 + \bar{\mathbf{I}}_2)^{1/3} = o(n^{-2})$. □

Proof of Theorem 3. Given $w_j, w_k, w_l \in W$ we denote the events

$$\begin{aligned} \mathcal{H}_j \; &= \{w_j \in D_1 \cap D_2\}, \qquad \mathcal{H}_j^1 = \{w_j \in D_1 \cap D_2 \cap D_3\}, \\ \mathcal{H}_{jk} &= \{w_j \in D_1 \cap D_3,\, w_k \in D_2 \cap D_3\}, \\ \mathcal{H}_{jkl} &= \{w_j \in D_1 \cap D_2,\, w_k \in D_1 \cap D_3,\, w_l \in D_2 \cap D_3\}, \end{aligned}$$

and their unions

$$\mathcal{H}^1 = \bigcup_{j \in [m]} \mathcal{H}_j^1, \qquad \mathcal{H}^2 = \bigcup_{\{j,k\} \subset [m]} \mathcal{H}_{jk}, \qquad \mathcal{H}^3 = \bigcup_{\{j,k,l\} \subset [m]} \mathcal{H}_{jkl}.$$

Furthermore, we denote $\tilde{\mathcal{H}}^1 = \cup_{j \in [m-1]} \mathcal{H}_j^1$ and $\tilde{\mathcal{H}}^2 = \cup_{\{j,k\} \subset [m-1]} \mathcal{H}_{jk}$.
Let us prove (13). To this aim we write $cl(r) = p_r^* / (\overline{p}_r^* + p_r^*)$, where

$$p^*(r) := \mathbf{P}(v_1 \sim v_2, d_{12} = r), \qquad \overline{p}_r^* := \mathbf{P}(v_1 \not\sim v_2, d_{12} = r),$$

and show that

$$p_r^* = \mathbf{E} f_\Lambda(r) a_1^2 B_m^2 n^{-1} + o(n^{-1}), \tag{59}$$

$$\overline{p}_r^* = f_{\Lambda_1}(r) + O(n^{-\tau_r}). \tag{60}$$

Here we denote $\Lambda = a_1 B_m \beta_n^{-1/2}$ and $\Lambda_1 = 2^{-1} a_1^2 a_2 b_2^2 n^{-1}$, and $\tau_r = 1 + \mathbb{I}_{\{r \ge 1\}}$.
We first show (59). Expanding the event $\{v_1 \sim v_2\} = \cup_{j \in [m]} \mathcal{H}_j$ we obtain, by inclusion-exclusion, $0 \le S_1 - p_r^* \le S_2$. Here

$$S_1 = \sum_{j \in [m]} \mathbf{P}(\mathcal{H}_j \cap \{d_{12} = r\}) = m \mathbf{P}(\mathcal{H}_m \cap \{d_{12} = r\}),$$
$$S_2 = \sum_{\{j,k\} \subset [m]} \mathbf{P}(\mathcal{H}_j \cap \mathcal{H}_k) = \binom{m}{2} \mathbf{P}(\mathcal{H}_1 \cap \mathcal{H}_2) = O(n^{-2}).$$

We complete the proof of (59) by showing that

$$\mathbf{P}(\mathcal{H}_m \cap \{d_{12} = r\}) = \mathbf{E} f_\Lambda(r) a_1^2 B_m^2 n^{-2} \beta_n^{-1} + o(n^{-2}). \tag{61}$$

We denote, for short, $\varkappa := \mathbf{P}(\mathcal{H}_m \cap \{d_{12} = r\})$ and write

$$\varkappa = \mathbf{E}\overline{\mathbf{P}}(\mathcal{H}_m \cap \{d_{12} = r\}) = \mathbf{E}\overline{\mathbf{P}}(d_{12} = r | \mathcal{H}_m) \overline{\mathbf{P}}(\mathcal{H}_m) = \mathbf{E}\overline{\mathbf{P}}(d_{12} = r | \mathcal{H}_m) p_{1m} p_{2m},$$

where $\overline{\mathbf{P}}$ denotes the conditionall probability given A_1, A_2, B_m, and approximate $\overline{\mathbf{P}}(d_{12} = r | \mathcal{H}_m)$ by a Poisson probability $f_\lambda(r)$. Indeed, $d_{12} = \sum_{i=3}^n \mathbb{I}_i$ is the sum of indicators of events $\mathcal{L}_i = \{D_i \cap D_1 \ne \emptyset, D_i \cap D_2 \ne \emptyset\}$ which are conditionally independent, given $A_1, A_2, B_m, \mathcal{H}_m$. Hence, for $\lambda = (n-2)\tilde{p}$ and $\tilde{p} = \overline{\mathbf{P}}(\mathcal{L}_3)$, Lemma 2 implies

$$|\overline{\mathbf{P}}(d_{12} = r | \mathcal{H}_m) - f_\lambda(r)| \le 2(n-2)\tilde{p}^2. \tag{62}$$

Next we evaluate the probability $\tilde{p}$. We observe that $\mathcal{L}_3 = \{w_m \in D_3\} \cup \tilde{\mathcal{H}}^1 \cup \tilde{\mathcal{H}}^2$. Denoting $p_\star := \overline{\mathbf{P}}(w_m \in D_3)$ we write

$$0 \le \tilde{p} - p_\star \le \overline{\mathbf{P}}(\tilde{\mathcal{H}}^1) + \overline{\mathbf{P}}(\tilde{\mathcal{H}}^2). \tag{63}$$

Since $p_\star \le a_1 B_m (nm)^{-1/2}$ and

$$\overline{\mathbf{P}}(\tilde{\mathcal{H}}^1) = (m-1)\overline{\mathbf{P}}(\mathcal{H}_1^1) \le A_1 A_2 a_1 b_3 n^{-3/2} m^{-1/2},$$
$$\overline{\mathbf{P}}(\tilde{\mathcal{H}}^2) = \binom{m-1}{2} \overline{\mathbf{P}}(\mathcal{H}_{12}) \le A_1 A_2 a_2 b_2^2 n^{-2}$$

we obtain from (63) that

$$\tilde{p}^2 \le 3\overline{\mathbf{P}}(\tilde{\mathcal{H}}^1) + 3\overline{\mathbf{P}}(\tilde{\mathcal{H}}^2) + 3p_\star^2 \le c\left(A_1 A_2 + B_m(n^{1/4} + n\overline{\mathbb{I}}_\star)\right) n^{-2}. \tag{64}$$

Here c is a constant which does not depend on n and m. In the last step we used the simple inequality $p_\star \le p_\star \mathbb{I}_\star + \overline{\mathbb{I}}_\star$, where $\mathbb{I}_\star := \mathbb{I}_{\{B_m < n^{1/4}\}}$ and $\overline{\mathbb{I}}_\star := 1 - \mathbb{I}_\star$. Now, (62), (64) imply

$$\varkappa = \mathbf{E} f_\lambda(r) p_{1m} p_{2m} + o(n^{-2}).$$

Next, using the inequality $|f_\lambda(r) - f_\Lambda(r)| \le |\lambda - \Lambda|$ we replace λ by Λ. We have arrived to (61)

$$\varkappa = \mathbf{E} f_\Lambda(r) p_{1m} p_{2m} + o(n^{-2}) = \mathbf{E} f_\Lambda(r) a_1^2 B_m^2 n^{-2} \beta_n^{-1} + o(n^{-2}).$$

We secondly show (60). We denote $\mathcal{C} = \{v_1 \not\sim v_2\}$ and write

$$\overline{p}_r^* = \mathbf{P}(d_{12} = r | \mathcal{C}) \mathbf{P}(\mathcal{C}). \tag{65}$$

By the inequalities $0 \le 1 - \mathbf{P}(\mathcal{C}) \le \mathbf{E} \sum_{3 \le j \le m} p_{1j} p_{2j}$, the second probability $\mathbf{P}(\mathcal{C}) = 1 + O(n^{-1})$. For the first probability we apply the Poisson approximation as in the proof of (59) above. In particular, Lemma 2 implies

$$|\mathbf{P}(d_{12} = r | \mathcal{C}) - f_{\lambda_1}(r)| \le 2(n-2)\hat{p}^2, \tag{66}$$

where $\lambda_1 = (n-2)\hat{p}$ and $\hat{p} = \mathbf{P}(\mathcal{L}_3|\mathcal{C})$. Next we evaluate $\hat{p} = \mathbf{P}(\mathcal{L}_3 \cap \mathcal{C})/\mathbf{P}(\mathcal{C})$. We observe that $\mathcal{L}_3 \cap \mathcal{C} = \mathcal{C} \cap \mathcal{H}^2$. Hence, by inclusion-exclusion, we have $0 \le \hat{S}_1 - \mathbf{P}(\mathcal{L}_3 \cap \mathcal{C}) \le \hat{S}_2$, where

$$\hat{S}_1 = \sum_{\{s,t\}\subset[m]} \mathbf{P}(\mathcal{H}_{st} \cap \mathcal{C}) = \binom{m}{2}\mathbf{P}(\mathcal{H}_{12} \cap \mathcal{C}), \tag{67}$$

$$\hat{S}_2 = \sum_{*} \mathbf{P}(\mathcal{H}_{st} \cap \mathcal{H}_{xy}) = \binom{m}{2}\left(\binom{m-2}{2}p' + 2(m-2)p''\right) = O(n^{-3}).$$

Here the sum $\sum_*$ runs over distinct pairs $\big\{\{s,t\},\{x,y\}\big\}$ of subsets of $[m]$ of size 2 and

$$p' = \mathbf{P}(\mathcal{H}_{12} \cap \mathcal{H}_{34}) = O(n^{-7}), \qquad p'' = \mathbf{P}(\mathcal{H}_{12} \cap \mathcal{H}_{13}) = O(n^{-6}).$$

Therefore, $\mathbf{P}(\mathcal{L}_3 \cap \mathcal{C}) = \hat{S}_1 + O(n^{-3})$. Furthermore, combining (67) with relations

$$\mathbf{P}(\mathcal{H}_{12} \cap \mathcal{C}) = \mathbf{E} p_{11}p_{31}p_{22}p_{32}q = a_1^2 a_2 b_2^2 (nm)^{-2} + O(n^{-5}),$$

here we denote $q = (1-p_{12})(1-p_{21})\prod_{j=3}^m (1-p_{1j}p_{2j})$ and use

$$0 \le 1 - q \le p_{12} + p_{21} + \sum_{3\le j\le m} p_{1j}p_{2j}, \tag{68}$$

we obtain $\hat{S}_1 = 2^{-1}a_1^2a_2b_2^2n^{-2} + O(n^{-3})$. We have shown that

$$\hat{p} = \mathbf{P}(\mathcal{L}_3 \cap \mathcal{C})/\mathbf{P}(\mathcal{C}) = 2^{-1}a_1^2a_2b_2^2n^{-2} + O(n^{-3}).$$

Invoking this expression of $\hat{p}$ in (66) we derive (60) from (65), (66).

Let us prove (12). To this aim we write $\alpha = \mathbf{P}(\mathcal{B})/\mathbf{P}(\mathcal{D})$, where $\mathcal{D}$ denotes the event $\{v_1 \sim v_3, v_2 \sim v_3\}$ and $\mathcal{B} = \mathcal{D} \cap \{v_1 \sim v_2\}$, and show that

$$\mathbf{P}(\mathcal{B}) = \varkappa_1 + o(n^{-2}), \qquad \mathbf{P}(\mathcal{D}) = \varkappa_1 + \varkappa_2 + o(n^{-2}). \tag{69}$$

Here $\varkappa_1 := a_1^3 b_3 n^{-3/2} m^{-1/2}$ and $\varkappa_2 := 2^{-1}a_1^2a_2b_2^2n^{-2}$. To show the first relation of (69) we observe that the event $\mathcal{H}^1$ implies $\mathcal{B}$, and the event $\mathcal{B}$ implies $\mathcal{H}^1 \cup \mathcal{H}^3$. In particular, we have $0 \le \mathbf{P}(\mathcal{B}) - \mathbf{P}(\mathcal{H}^1) \le \mathbf{P}(\mathcal{H}^3)$. Here

$$\mathbf{P}(\mathcal{H}^3) = \binom{m}{3}\mathbf{P}(\mathcal{H}_{123}) = O(n^{-3}).$$

Hence, $\mathbf{P}(\mathcal{B}) = \mathbf{P}(\mathcal{H}^1) + O(n^{-3})$. Next we approximate $\mathbf{P}(\mathcal{H}^1)$ using inclusion-exclusion

$$\sum_{s\in[m]} \mathbf{P}(\mathcal{H}_s^1) - \sum_{\{s,t\}\subset[m]} \mathbf{P}(\mathcal{H}_s^1 \cap \mathcal{H}_t^1) \le \mathbf{P}(\mathcal{H}^1) \le \sum_{s\in[m]} \mathbf{P}(\mathcal{H}_s^1) \tag{70}$$

and obtain $\mathbf{P}(\mathcal{H}^1) = \varkappa_1 + o(n^{-2})$. Here we invoke the bound

$$\sum_{\{s,t\}\subset[m]} \mathbf{P}(\mathcal{H}_s^1 \cap \mathcal{H}_t^1) = \binom{m}{2}\mathbf{P}(\mathcal{H}_1^1 \cap \mathcal{H}_2^1) = O(n^{-3})$$

and $\sum_{s\in[m]} \mathbf{P}(\mathcal{H}_s^1) = m\mathbf{P}(\mathcal{H}_s^1) = \varkappa_1 + o(n^{-2})$.
Let us prove the second relation of (69). We observe that $\mathcal{D} = \mathcal{H}^1 \cup \mathcal{H}^2$ and approximate

$$\mathbf{P}(\mathcal{D}) \approx \mathbf{P}(\mathcal{H}^1) + \mathbf{P}(\mathcal{H}^2) \approx m\mathbf{P}(\mathcal{H}_1^1) + \binom{m}{2}\mathbf{P}(\mathcal{H}_{12}) = \varkappa_1 + \varkappa_2 + o(n^{-2}).$$

The rigorous argument is a bit more involved since we operate under minimal moment conditions. Introduce event $\mathcal{A}^* = \{A_3 < n^{1/4}\}$ and its indicator function $\mathbb{I} = \mathbb{I}_{\mathcal{A}^*}$. We construct upper and lower bounds for $\mathbf{P}(\mathcal{D})$ using the inequalities

$$\mathbf{P}(\mathcal{H}^1 \cap \mathcal{A}^*) + \mathbf{P}(\mathcal{H}^2 \cap \mathcal{A}^*) - \mathbf{P}(\mathcal{H}^1 \cap \mathcal{H}^2 \cap \mathcal{A}^*) \le \mathbf{P}(\mathcal{D} \cap \mathcal{A}^*) \le \mathbf{P}(\mathcal{D}) \le \mathbf{P}(\mathcal{H}^1) + \mathbf{P}(\mathcal{H}^2).$$

By the union bound, the right hand side is bounded by $m\mathbf{P}(\mathcal{H}_1^1) + \binom{m}{2}\mathbf{P}(\mathcal{H}_{12}) = \varkappa_1 + \varkappa_2 + o(n^{-2})$. Next we show a matching lower bound for $\mathbf{P}(\mathcal{D})$. Proceeding as in (70) we write

$$\mathbf{P}(\mathcal{H}^1 \cap \mathcal{A}^*) = m\mathbf{P}(\mathcal{H}_1^1 \cap \mathcal{A}^*) + O(n^{-3}),$$

where

$$\mathbf{P}(\mathcal{H}_1^1 \cap \mathcal{A}^*) = \mathbf{E}p_{11}p_{21}p_{31}\mathbb{I} = \mathbf{E}p_{11}p_{21}p_{31} + o(n^{-3}) = a_1^3 b_3 (nm)^{-3/2} + o(n^{-3}).$$

Hence, we have $\mathbf{P}(\mathcal{H}^1 \cap \mathcal{A}^*) = \varkappa_1 + o(n^{-2})$. It remains to show that

$$\mathbf{P}(\mathcal{H}^2 \cap \mathcal{A}^*) \ge \varkappa_2 + o(n^{-2}), \qquad \mathbf{P}(\mathcal{H}^1 \cap \mathcal{H}^2 \cap \mathcal{A}^*) = o(n^{-2}). \tag{71}$$

For the first inequality we write, by inclusion-exclusion,

$$\mathbf{P}(\mathcal{H}^2 \cap \mathcal{A}^*) \ge \sum_{\{s,t\} \subset [m]} \mathbf{P}(\mathcal{H}_{st} \cap \mathcal{A}^*) - R = \binom{m}{2}\mathbf{P}(\mathcal{H}_{12} \cap \mathcal{A}^*) - R,$$

where $\mathbf{P}(\mathcal{H}_{12} \cap \mathcal{A}^*) = a_1^2 a_2 b_2^2 (nm)^{-2} + o(n^{-4})$ and where

$$R = \sum_{*} \mathbf{P}(\mathcal{H}_{st} \cap \mathcal{H}_{xy} \cap \mathcal{A}^*) = \binom{m}{2}\binom{m-2}{2} R_1 + 2\binom{m}{2}(m-2)R_2 = o(n^{-2}).$$

Here we invoked the bounds

$$R_1 = \mathbf{P}(\mathcal{H}_{12} \cap \mathcal{H}_{34} \cap \mathcal{A}^*) \le (nm)^{-4}\mathbf{E}(A_1A_2)^2 A_*^4 (B_1B_2B_3B_4)^2 = O(n^{-3.5}m^{-4}),$$
$$R_2 = \mathbf{P}(\mathcal{H}_{12} \cap \mathcal{H}_{13} \cap \mathcal{A}^*) \le (nm)^{-3}\mathbf{E}A_1A_2^2 A_*^3 (B_1B_2B_3)^2 = O(n^{-2.75}m^{-3}).$$

In the last step we used inequalities $A_3^4\mathbb{I} < A_3^2 n^{1/2}$ and $A_3^3\mathbb{I} < A_3^2 n^{1/4}$.
Finally, in order to show the second bound of (71) we apply the union bound

$$\mathbf{P}(\mathcal{H}^1 \cap \mathcal{H}^2 \cap \mathcal{A}^*) \le \sum_{\{s,t\} \subset [m]} \mathbf{P}(\mathcal{H}^1 \cap \mathcal{H}_{st} \cap \mathcal{A}^*) \le \sum_{\{s,t\} \subset [m]} (r_s + r_t + \sum_{u \in [m] \setminus \{s,t\}} r'_u)$$

and invoke the bounds

$$r_s = \mathbf{P}(\mathcal{H}_s^1 \cap \mathcal{H}_{st} \cap \mathcal{A}^*) \le (nm)^{-5/2}\mathbf{E}A_1A_2^2A_3^2B_s^3B_t^2 = O((nm)^{-5/2}),$$
$$r'_u = \mathbf{P}(\mathcal{H}_u^1 \cap \mathcal{H}_{st} \cap \mathcal{A}^*) \le (nm)^{-7/2}\mathbf{E}A_1^2A_2^2A_3^3B_s^2B_t^2B_u^3\mathbb{I} = O((nm)^{-7/2}n^{1/4}).$$

□

The following inequality is refered to as LeCam's lemma, see e.g., [23].

Lemma 2. *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*

$$\sup_{A \subset \{0,1,2\dots\}} |\mathbf{P}(S \in A) - \mathbf{P}(\Lambda \in A)| \le 2\sum_i p_i^2. \tag{72}$$

Lemma 3. ([7]) *Given integers $1 \le s \le d_1 \le d_2 \le m$, let D_1, D_2 be independent random subsets of the set $W = \{1, \dots, m\}$ such that D_1 (respectively D_2) is uniformly distributed in the class of subsets of W of size d_1 (respectively d_2). The probabilities $p' := \mathbf{P}(|D_1 \cap D_2| = s)$ and $p'' := \mathbf{P}(|D_1 \cap D_2| \ge s)$ satisfy*

$$\left(1 - \frac{(d_1 - s)(d_2 - s)}{m + 1 - d_1}\right) p^*_{d_1,d_2,s} \ \le \ p' \ \le \ p'' \le \ p^*_{d_1,d_2,s}, \tag{73}$$

Here we denote $p^*_{d_1,d_2,s} = \binom{d_1}{s}\binom{d_2}{s}\binom{m}{s}^{-1}$.

Acknowledgement. Research was supported in part by the Research Council of Lithuania grant MIP-053/2011.

References

[1] E. Bakshy, I. Rosenn, C. Marlow and L. Adamic, The role of social networks in information diffusion, in: Proceedings of the 21st international conference on World Wide Web, WWW 2012, April 1620, 2012, Lyon, France. (2012), 519-528. ACM 978-1-4503-1229-5/12/04.

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

[3] A. Barrat and M. Weigt, On the properties of small-world networks, The European Physical Journal B 13 (2000), 547–560.

[4] M. Behrisch, Component evolution in random intersection graphs, The Electronical Journal of Combinatorics 14 (2007), #R17.

[5] S. Blackburn and S. Gerke, Connectivity of the uniform random intersection graph, Discrete Mathematics, 309 (2009), 5130-5140.

[6] M. Bloznelis, Degree distribution of a typical vertex in a general random intersection graph, Lithuanian Mathematical Journal 48 (2008), 38–45.

[7] M. Bloznelis, Degree and clustering coefficient in sparse random intersection graphs, to appear in The Annals of Applied Probability.

[8] M. Bradonjic, A. Hagberg, N. W. Hengartner, A. G. Percus, Component Evolution in General Random Intersection Graphs, The 7th Workshop on Algorithms and Models for the Web Graph, WAW2010. Lecture Notes in Computer Science (Springer-Verlag, Berlin, 2010), Vol. 6516, pp. 36-49.

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

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

[11] 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 (2002), 41–47.

[12] I. Foudalis, K. Jain, C. Papadimitriou, and M. Sideri, Modeling social networks through user background and behavior, in: Algorithms and Models for the Web Graph. Proceedings of the 8th International Workshop,WAW 2011, Lecture notes in computer science 6732, (2011), 85–102.

[13] E. Godehardt and J. Jaworski, Two models of random intersection graphs for classification, in: Studies in Classification, Data Analysis and Knowledge Organization, Springer, Berlin–Heidelberg–New York, 2003, 67–81.

[14] J. L. Guillaume, M. Latapy, Bipartite structure of all complex networks, Inform. Process. Lett. 90 (2004) 215–221.

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

[16] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Random graphs with arbitrary degree distributions and their applications, Physical Review E 64 (2001) 026118.

[17] M. E. J. Newman, Properties of highly clustered networks, Physical Review E 68 (2003) 026121.

[18] M. E. J. Newman, D. J. Watts, and S. H. Strogatz, Random graph models of social networks, Proc. Natl. Acad. Sci. USA, 99 (Suppl. 1) (2002), 2566–2572.

[19] S. Nikoletseas, C. Raptopoulos, and P. G. Spirakis, On the independence number and Hamiltonicity of uniform random intersection graphs, Theoretical Computer Science 412 (2011), 6750–6760.

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

[21] Y. Shang, Degree distributions in general random intersection graphs, The Electronical Journal of Combinatorics 17 (2010), #R23.

[22] D. Stark, The vertex degree distribution of random intersection graphs, Random Structures and Algorithms 24 (2004), 249–258.

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

[24] S. H. Strogatz and D. J. Watts, Collective dynamics of small-world networks, Nature, 393 (1998), 440–442.

[25] A. L. Traud, E. D. Kelsic, P. J Muchta and M. A. Porter, Community structure in online collegiate social networks. tech. rep. arXiv:0809.0690 (September 2008), http://www.amath.unc.edu/Faculty/mucha/reprints/facebook.pdf.

[26] O. Yagan and A. M. Makowski, Random key graphs – can they be small worlds? in: 2009 First International Conference on Networks & Communications (2009) 313–318.

[27] Information courtesy of the internet movie database, http://www.imdb.com.