Clustering coefficient of random intersection graphs with infinite degree variance

Mindaugas Bloznelis and Valentas Kurauskas

Vilnius university, Lithuania

Dedicated to Professor Michał Karoński on the occasion of his 70th birthday

Abstract

For a random intersection graph with a power law degree sequence having a finite mean and an infinite variance we show that the global clustering coefficient admits a tunable asymptotic distribution.

key words: clustering coefficient, power law, infinite variance, random intersection graph, affiliation network.

1 Introduction

The global clustering coefficient C_G of a graph G is the ratio $C_G = 3\Delta/\Lambda$, where Δ is the number of triangles and Λ is the number of paths of length 2. Another way to represent the global clustering coefficient is by the conditional probability that a randomly chosen triple of vertices makes up a triangle given that the first two vertices are adjacent to the third one. Formally,

$$C_G = \mathbf{P}^*(v_1^* \sim v_2^* \,|v_1^* \sim v_3^*,\, v_2^* \sim v_3^*),$$

where (v_1^*, v_2^*, v_3^*) is an ordered triple of vertices sampled uniformly at random and the probability $\mathbf{P}^*$ refers to the sampling. By $\sim$ we denote the adjacency relation.

In this paper we study the relation between the clustering coefficient and the tail of the degree sequence in large complex networks. We focus on random intersection graph models of real affiliation networks (mode two networks), [12], [9], [3]. They admit tunable degree distribution and non-vanishing clustering coefficient [14], [6], [1], [4]. Definition of a random intersection graph is recalled below in this section.

The global clustering coefficient C_G of a realised instance G of a random graph is a random variable. We note that generally this random variable behaves differently depending on whether the degree variance is finite or infinite [1], [11], [17]. When the degree variance is finite the global clustering coefficient C_G can be approximated by the corresponding numerical characteristic of the underlying random intersection graph model, the conditional probability $\alpha_C := \mathbf{P}(v_1^* \sim v_2^* \,|v_1^* \sim v_3^*,\, v_2^* \sim v_3^*)$, [13]. Here and below $\mathbf{P}$ refers to all the sources of randomness defining the events considered (these are the uniform sampling of vertices (v_1^*, v_2^*, v_3^*) and random graph generation mechanism in the present context). We remark that α_C admits a simple asymptotic expression in terms of the first and second moment of the degree sequence [1], [10], [4], [5].

The question about the behaviour of the clustering coefficient C_G when the degree variance is infinite remained open. We address this question in the present paper. Our study is analytical.

For an infinite degree variance we show that C_G admits a non-degenerate asymptotic distribution with tunable characteristics in the case where the weights defining the underlying random intersection graph achieve a certain balance. In this way our theoretical findings contribute to the discussion about whether and when a power law network model with an infinite degree variance can have a non-vanishing global clustering coefficient, cf. [17], where a negative result was obtained.

The paper is organized as follows. In this section we introduce random intersection graphs, formulate and discuss our results. Proofs are given in section 2. Technical lemmas are postponed to Section 3.

1.1 Random intersection graphs

Random intersection graphs model social networks, where the actors establish communication links provided that they share some common attributes (collaboration networks, actor networks, etc.). A random intersection graph G on the vertex set $V = \{v_1, \dots, v_n\}$ is defined by a random bipartite graph, denoted by H, with the bipartition $V \cup W$, where $W = \{w_1, \dots, w_m\}$ is an auxiliary set of attributes. Two vertices in G are adjacent whenever they have a common neighbour in H. This neighbour is called a witness of the adjacency relation.

In the *active* graph, denoted by $G(n, m, \mathrm{P})$, vertices $v \in V$ select their neighbourhoods $S_v \subset W$ in H independently at random according to the probability distribution $\mathbf{P}(S_v = A) = \mathrm{P}(|A|)\binom{m}{|A|}^{-1}$, $A \subset W$. Here P is the probability distribution modeling the size $|S_v|$ of the neighbourhood of v in H. Given the size $|S_v|$, the elements of S_v are selected uniformly at random. Two vertices u, v are adjacent in G whenever the random sets S_u and S_v (called attribute sets of u and v) intersect.

In the *passive* graph, denoted by $G^\star(n, m, \mathrm{P}^\star)$, attributes $w \in W$ select their neighbourhoods $D_w \subset V$ in H independently at random according to the probability distribution $\mathbf{P}(D_w = A) = \mathrm{P}^\star(|A|)\binom{n}{|A|}^{-1}$, $A \subset V$. Two vertices u, v are adjacent in $G^\star(n, m, \mathrm{P}^\star)$ whenever $u, v \in D_w$ for some $w \in W$.

The *inhomogeneous* graph, denoted by $G(n, m, \mathrm{P}_X, \mathrm{P}_Y)$, interpolates between the active and passive models. It is defined by the random bipartite graph, where attributes $w_i \in W$ and vertices $v_j \in V$ are assigned independent random weights X_i and Y_j respectively. The weights model the attractiveness of attributes and activity of actors. Every pair $(w_i, v_j) \in W \times V$ is linked in H with probability $p_{ij} = \min\{1, X_i Y_j / \sqrt{mn}\}$ independently of the other pairs. Here $X_1, \dots, X_m$ and $Y_1, \dots, Y_n$ are non-negative independent random variables with the distributions P_X and P_Y respectively.

In what follows we assume that n/m is bounded and it is bounded away from zero as $m, n \to +\infty$, denoted by $n = \Theta(m)$. The rationale behind this assumption is that in the range $n = \Theta(m)$ the active, passive and inhomogeneous models admit non-degenerate asymptotic degree distributions including power laws [1, 2, 4, 6]. More importantly, in this range these random graph models admit tunable global clustering coefficient $C_G \approx \alpha_C$, provided that the degree variance is finite [13]. Therefore it is reasonable to consider the range $n = \Theta(m)$, also when studying the global clustering coefficient of a power law intersection graph with an infinite degree variance.

1.2 Results

Let $d(v_i)$ denote the degree of a vertex $v_i \in V = \{v_1, \dots, v_n\}$ in a random intersection graph. We note that the random variables $d(v_1), \dots, d(v_n)$ are identically distributed for each particular model: active, passive and inhomogeneous. When speaking about the asymptotic degree distribution below we think about the limit in distribution of the random variable $d(v_1)$ as $n, m \to +\infty$.

Active graph $G = G(n, m, P)$. In Theorem 1 below we show that an active graph with an infinite degree variance has the global clustering coefficient $C_G \approx 0$.

Theorem 1. *Let* $\beta > 0$. *Let* $m, n \to +\infty$. *Assume that* $m/n \to \beta$. *Let* Z *be a non-negative random variable such that* $\mathbf{E}Z < \infty$ *and* $\mathbf{E}Z^2 = \infty$. *Let* P *denote the distribution of* $\min\{Z, m\}$. *The global clustering coefficient of the active random graph* $G(n, m, P)$ *satisfies* $C_G = o_P(1)$.

Under conditions of Theorem 1 the active graph has a mixed Poisson asymptotic degree distribution assigning probabilities $\mathbf{E}e^{-\lambda}\frac{\lambda^k}{k!}$ to the integers $k = 0, 1, \ldots$, see [1]. Here $\lambda = (\mathbf{E}Z)\beta^{-1}Z$ is a random variable. In the case where Z has a power law with the tail index $\alpha > 1$, i.e., for some $c_z > 0$ we have

$$\mathbf{P}(Z > t) = c_z t^{-\alpha} + o(t^{-\alpha}) \qquad \text{as} \qquad t \to +\infty, \tag{1}$$

the asymptotic degree distribution described above is a power law with the same tail index α. For $1 < \alpha \leq 2$ it has a finite first moment, infinite variance and the clustering coefficient $C_G \approx 0$.

Passive graph $G^\star = G^\star(n, m, P^\star)$. In Theorem 2 below we show that a passive graph with an infinite degree variance has the global clustering coefficient $C_{G^\star} \approx 1$. By X we denote a random variable with the distribution $\mathrm{P}^\star$.

Theorem 2. *Let* $\beta > 0$. *Let* $m, n \to \infty$. *Assume that* $mn^{-1} \to \beta$ *and*
(i) X *converges in distribution to a random variable* Z;
(ii) $\mathbf{E}Z^2 < \infty$ *and* $\lim_{m,n\to\infty} \mathbf{E}X^2 = \mathbf{E}Z^2$;
(iii) $\mathbf{E}Z^3 = \infty$.
Then the clustering coefficient $C_{G^\star} = 1 - o_P(1)$.

We mention that under conditions of Theorem 2, the degree $d(v_1)$ converges in distribution to the compound Poisson random variable $d_* = \sum_{j=1}^{\zeta} \tilde{Z}_j$, see [1]. Here $\tilde{Z}_1, \tilde{Z}_2, \ldots$ are independent random variables with the common probability distribution $\mathbf{P}(\tilde{Z}_1 = r) = (r+1)\mathbf{P}(Z = r + 1)/\mathbf{E}Z$, $r = 0, 1, 2, \ldots$. The random variable ζ is independent of the sequence $\tilde{Z}_1, \tilde{Z}_2, \ldots$ and has Poisson distribution with mean $\mathbf{E}\zeta = \beta^{-1}\mathbf{E}Z$. Assuming that for some $\alpha \in (3, 4)$ and $c > 0$

$$\mathbf{P}(Z = r) = cr^{-\alpha}(1 + o(1)) \qquad \text{as} \qquad r \to +\infty, \tag{2}$$

we obtain, by Theorem 4.30 of [8], that

$$\mathbf{P}(d_* = r) = \mathbf{P}(\tilde{Z}_1 = r)(\mathbf{E}\zeta)(1 + o(1)) = c'\beta^{-1}r^{1-\alpha}(1 + o(1)) \qquad \text{as} \qquad r \to +\infty,$$

for some constant $c' > 0$. In this case G^* has asymptotic power law degree distribution with a finire first moment, infinite variance and the clustering coefficient $C_{G^*} \approx 1$.

Inhomogeneous graph $G(n, m, P_X, P_Y)$. In Theorem 3 below we show that the global clustering coefficient of an inhomogeneous graph with an infinite degree variance is highly determined by the ratio of the random variables

$$S_X = \sum_{i=1}^{m} X_i^3 \qquad \text{and} \qquad S_Y = \sum_{j=1}^{n} Y_j^2.$$

We denote $a_i = \mathbf{E}X_1^i$, $i = 1, 2$, and $b_1 = \mathbf{E}Y_1$.

Theorem 3. *Let $\beta > 0$. Let $m, n \to \infty$. Assume that $mn^{-1} \to \beta$. Suppose that $\mathbf{E}X_1^2 < \infty$, $\mathbf{E}X_1^3 = \infty, \mathbf{E}Y_1 < \infty, \mathbf{E}Y_1^2 = \infty$. Denote $\kappa = \beta^{3/2} a_2^2 b_1^{-1}$. We have $C_G = \bigl(1+\kappa S_Y/S_X\bigr)^{-1} + o_P(1)$.*

In the case where S_X and S_Y grow to infinity at the same rate we can obtain a non-trivial limit of C_G. The next remark addresses the case where the distributions of X_1^3 and Y_1^2 belong to the domain of attraction of stable distributions having the same characteristic exponent $\alpha \le 1$.

Remark 1. *Let $\alpha, \beta > 0$. Let $m, n \to \infty$. Assume that $mn^{-1} \to \beta$. Suppose that for some $c_x, c_y > 0$ we have*

$$\mathbf{P}(X_1 > t) = c_x t^{-3\alpha} + o(t^{-3\alpha}), \qquad \mathbf{P}(Y_1 > t) = c_y t^{-2\alpha} + o(t^{-2\alpha}) \qquad as \qquad t \to +\infty. \qquad (3)$$

(i) For $0 < \alpha < 1$ the ratio S_Y/S_X converges in distribution to the random variable $c^* Z_\alpha / Z'_\alpha$, where Z_α, Z'_α are independent stable random variables with the Laplace transform $\mathbf{E}e^{-sZ_\alpha} = \mathbf{E}e^{-sZ'_\alpha} = e^{-s^\alpha}$ and $c^* = (c_y/(c_x\beta))^{1/\alpha}$.
(ii) For $\alpha = 1$ the ratio $S_Y/S_X = c_y(c_x\beta)^{-1} + o_P(1)$.

Let us apply Theorem 3 to power law random weights (3). We observe that $\mathbf{E}X_1^2, \mathbf{E}Y_1 < \infty$ and $\mathbf{E}X_1^3, \mathbf{E}Y_1^2 = \infty$ imply $2/3 < \alpha \le 1$. For $\alpha = 1$ the result of Theorem 3 implies that $C_G \approx \bigl(1 + \kappa c_y/(c_x\beta)\bigr)^{-1}$ is asymptotically constant. For $2/3 < \alpha < 1$ it implies that C_G *converges in distribution* to the random variable $\bigl(1 + \kappa (c_y/(c_x\beta))^{1/\alpha} Z_\alpha/Z'_\alpha\bigr)^{-1}$.

Finally, we mention that for $m/n \to \beta \in (0, +\infty)$ and $2/3 < \alpha \le 1$, the inhomogeneous graph defined by power law weights (3) has a power law asymptotic degree distribution with the tail index $3\alpha - 1$, see [2]. In particular, the asymptotic degree distribution has a finite first moment and an infinite variance.

1.3 Discussion

One motivation of our study was the recent paper [17], which claims that “if the degree distribution obeys the power law with an infinite variance, then the global clustering coefficient tends to zero with high probability as the size of a graph grows.” This may look a bit confusing in view of the fact that some large social networks with quite substantial global clustering coefficients are believed to have a power law degree distribution with an infinite variance. The present study could be viewed as an attempt to resolve this seemingly contradiction with the aid of a known theoretical model of an affiliation network.
We observe that random intersection graphs considered in this paper admit *asymptotic* power law degree distributions, but their degree sequence is not an iid sample from a power law. We mention that some real affiliation networks are believed to have a power law degree sequence, but with an exponential cutoff, [15], [16], [18].
In what follows we discuss the relation between the result of [17] and our Theorems 1, 2, 3 in some detail. To this aim we briefly recall the argument of [17]. We call a path $x \sim y \sim z$ a cherry produced by vertex y. For example, a vertex v_j of degree $d_j = d(v_j)$ produces $\binom{d_j}{2}$ cherries. Ostroumova and Samosvat [17] observed that cherries produced by vertices of large degrees highly outnumber the triangles of the graph. Indeed, among the iid degrees $d_1, \dots, d_n$ obeying a power law with the tail index $1 < \alpha < 2$, the largest few roughly scale as $n^{1/\alpha}$. Consequently, the number of cherries produced by the largest vertices roughly scale as $n^{2/\alpha}$. On the other hand, the number of triangles incident to any vertex v_j does not exceed the number of cherries $\binom{d_j}{2}$. More importantly, this number is bounded by the total number of edges of the graph (edges needed to close cherries produced by v_j). But for $1 < \alpha$ the average degree is

bounded and the total number of edges scales as n. This implies that only a negligible fraction $n^{1-(2/\alpha)}$ of cherries produced by the largest vertices are closed. Putting things together one can show that $3\Delta \le c\sum_j n \wedge \binom{d_j}{2}$ is negligible compared to $\Lambda = \sum_j \binom{d_j}{2}$. Hence $C_G = o_P(1)$.

In a random intersection graph G the triadic closure of a cherry is explained by a common attribute shared by all three vertices of the cherry (triangles whose edges are witnessed by distinct attributes are rare and can be neglected). We exploit this clustering mechanism while evaluating the global clustering coefficient C_G: When counting triangles we focus on cliques of G induced by the neighbourhoods $D_i = D_{w_i} \subset V$ of attributes $w_i \in W$ in the underlying bipartite graph H. Every set D_i of size $\tilde{X}_i := |D_i|$ covers $\binom{\tilde{X}_i}{3}$ triangles of G and the total number of triangles obtained in this way scales as $\tilde{S}_X = \sum_i \binom{\tilde{X}_i}{3}$ (overlaps can be neglected). In fact, this number dominates the total number of triangles in each of random intersection graphs considered in Theorems 1–3.

In the active graph (with bounded average degree) the random variables $\tilde{X}_i$ have the same asymptotic Poisson distribution. Hence $\tilde{S}_X$ scales as m. Furthermore, the degrees $\{d_j\}$ of vertices $\{v_j\}$ can be approximated by asymptotically independent Poisson random variables having means $\lambda_j = Z_j\beta^{-1}\mathbf{E}Z_j$. Here $Z_1, \dots, Z_n$ are iid copies of Z. Hence $\Lambda = \sum_j \binom{d_j}{2}$ scales as $\Theta(S_Z)$, where $S_Z = \sum_j Z_j^2$. For $\mathbf{E}Z^2 = \infty$ the sum S_Z is super-linear in n and for $n = \Theta(m)$ we obtain $\tilde{S}_X/S_Z = o_P(1)$. Thus $C_G = o_P(1)$. We note that similarly to the case of iid degrees considered in [17] the number of cherries of active intersection graph scales as a sum of iid random variables having an infinite mean. One difference from [17] is that in our Theorem 1 we have relaxed the structural "power law degree" condition of [17].

The passive graph is a union of independently located cliques induced by the sets $D_{w_i} \subset V$, $w_i \in W$. Since $|D_{w_i}| = \tilde{X}_i$ converges in distribution to a random variable having infinite third moment, we have that $\tilde{S}_X$ is super-linear in m. Furthermore, we show that Λ is dominated by the number of cherries covered by the cliques. This number scales as $3\sum_i \binom{X_i}{3} = 3\tilde{S}_X$ (we neglect overlaps again). Hence, $C_G^* = 1 + o_P(1)$.

The inhomogeneous graph interpolates between the active and passive graphs. The number of triangles $\tilde{S}_X$ scales as $\Theta(S_X)$ as in the passive graph, while Λ is approximately the sum of the number of cherries covered by large cliques (as in the passive graph) and the number of cherries produced by the largest vertices (as in the active graph). These numbers scale as $3\tilde{S}_X$ and $\Theta(S_Y)$ respectively. In this way we obtain the approximation $C_G \approx (1+\Theta(S_Y/S_X))^{-1}$. Finally, we note that the inhomogeneous graph is a fitness model of a real affiliation network, where activity of vertices is modeled by the distribution P_Y and attractiveness of attributes is modeled by the distribution P_X. We summarize the result of Theorem 3 as follows: The global clustering coefficient is non-vanishing whenever the attractiveness "outweighs" the activity.

2 Proofs

We begin by establishing some notation. Detailed proofs are given afterwards.

Notation. By $\mathbf{E}_{\mathbb{X}}$ and $\mathbf{P}_{\mathbb{X}}$ (respectively $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{P}}$) we denote the conditional expectation and conditional probability given $\mathbb{X} = (X_1, \dots, X_m)$ (respectively $\mathbb{X}$ and $\mathbb{Y} = (Y_1, \dots, Y_n)$). We use the notation $[k]$ for the set $\{1, 2, \dots, k\}$ and the shorthand notation $\sum_{\Lambda}$ for the double sum $\sum_{x\in V}\sum_{\{y,z\}\subset V\setminus\{x\}}$. Denote empirical means $\hat{a}_r = m^{-1}\sum_{i\in[m]} X_i^r$ and $\hat{b}_r = n^{-1}\sum_{j\in[n]} Y_j^r$.

Let G be the intersection graph defined by a bipartite graph H with the bipartition $V \cup W$. For $x, y \in V$ and $w \in W$ we denote by $\mathbb{I}_{x\sim y}$ and $\mathbb{I}_{xw}$ the indicators of the events that x, y are adjacent in G and x, w are adjacent in H. For $v = v_j \in V$ and $w = w_i \in W$ we write interchangeably Y_j or Y_v and X_i or X_w also p_{ij} or p_{wv}. For $v \in V$ and $w \in W$ we denote $\lambda_{vw} = Y_v X_w (nm)^{-1/2}$.

For $w \in W$, let $D_w \subset V$ denote the set of neighbours of w in H. Note that each D_w induces a clique in G. Given a subgraph $G' \subset G$ and a subset $W' \subset W$ we say that the collection of sets $\{D_w,\, w \in W'\}$ is a cover of G' if every edge of G' is witnessed by some $w \in W'$ and for every $w \in W'$ there is an edge in G' having no other witness from W', but w (any proper subset of W' can't be a cover of G').
A subgraph of G is labeled "lucky" if it has a cover consisting of a single set D_w, for some $w \in W$. A subgraph is labeled "unlucky" if it has a cover consisting of two or more sets. We note that a subgraph can be labeled "lucky" and "unlucky" simultaneously.
The numbers of lucky and unlucky triangles (2-paths) are denoted by Δ_L and Δ_U (Λ_L and Λ_U). The number of triangles (2-paths) receiving both lucky and unlucky labels is denoted Δ_{LU} (Λ_{LU}). Clearly, we have

$$\Delta = \Delta_L + \Delta_U - \Delta_{LU}, \qquad \Lambda = \Lambda_L + \Lambda_U - \Lambda_{LU}. \tag{4}$$

Proof of Theorem 1. In the proof we use some ideas of [17]. Before the proof we collect notation and auxiliary facts. Let $Z_1, Z_2, \dots$ be iid copies of Z. We denote by $\mathbf{E}_Z$ ($\mathbf{Var}_Z$) the conditional expectation (variance) given the sequence $\{Z_i, i \geq 1\}$. Furthermore, we denote $z_1 = \mathbf{E}Z$ and $S_Z = \sum_{i \in [n]} Z_i^2$. Given $A \subset [n]$ we denote $S_{Z,A} = \sum_{i \in A} Z_i^2$. By $d_{i,A} = \sum_{j \in A \setminus \{i\}} \mathbb{I}_{v_i \sim v_j}$ we denote the number of neighbours from the set $\{v_j,\, j \in A\} \subset V$ of a vertex v_i in the intersection graph G. In the proof we use the following inequalities for the intersection probability of two independent uniformly distributed random subsets $\mathcal{S}, \mathcal{T} \subset W$ (see, e.g., Lemma 6 of [1])

$$stm^{-1}(1 - st/(m-s)) \leq \mathbf{P}\Big(\mathcal{S} \cap \mathcal{T} \neq \emptyset \Big| \, |\mathcal{S}| = s, |\mathcal{T}| = t\Big) \leq stm^{-1}. \tag{5}$$

We recall that every vertex $v_i \in V = \{v_1, \dots, v_n\}$ is prescribed a subset $S_i \subset W = \{w_1, \dots, w_m\}$ of size $|S_i| = \min\{m, Z_i\}$. Furthermore, the condition $\mathbf{E}Z < \infty$ ensures the existence of a positive sequence $\varepsilon_n \downarrow 0$ such that

$$\mathbf{P}(\max_{i \in [n]} Z_i < n\varepsilon_n) = 1 - o(1), \tag{6}$$

see Lemma 3. Note that (6) implies $\mathbf{P}(\max_{i \in [n]} Z_i < m) = 1 - o(1)$.
Now we prove the theorem. For this purpose we show that there is a constant $c^* > 0$ and a sequence $\varkappa_n \downarrow 0$ both depending on the distribution of Z and on β such that

$$\mathbf{P}(\Lambda > c^* S_Z) = 1 - o(1), \tag{7}$$
$$\mathbf{P}(\Delta \leq n^{3/2}\varkappa_n) = 1 - o(1), \tag{8}$$
$$\Delta = O_P\big(n + n^{-3}S_Z^3\big). \tag{9}$$

Let us show that (7), (8), (9) imply $C_G = o_P(1)$. Introduce the event $B = \{S_Z \leq n^{3/2}\sqrt{\varkappa_n}\}$ and let $\bar{B}$ denote the complement event. We have

$$C_G = \frac{3\Delta}{\Lambda} = \frac{3\Delta}{\Lambda}\mathbb{I}_B + \frac{3\Delta}{\Lambda}\mathbb{I}_{\bar{B}} = O_P\Big(\frac{n}{S_Z}\Big) + O_P(\varkappa_n) + O_P(\sqrt{\varkappa_n}) = o_P(1). \tag{10}$$

Here on the event B we have bounded Δ using (9) and on the event $\bar{B}$ we have applied (8). In the final step we invoked the bound $n/S_Z = o_P(1)$, which follows by Lemma 1. It remains to prove (7), (8) and (9).
Proof of (7). Fix $0 < a < b$ such that $p := \mathbf{P}(a < Z < b) > 0$. Define random subsets of $[n]$

$$R = \{i : \, a < Z_i < b\}, \qquad T = \{i : \, Z_i \leq \ln^2 n\}, \qquad \Theta = \{i : \, \ln^2 n < Z_i \leq n\varepsilon_n\}.$$

Note that for any $i \in [n]$ and $A \subset [n]$ the degree d_i of a vertex v_i is larger or equal to $d_{i,A}$. Therefore, we have

$$\Lambda = \sum_{i \in [n]} \binom{d_i}{2} \geq \Lambda_T + \Lambda_\Theta, \qquad \Lambda_T = \sum_{i \in T} \binom{d_{i,T}}{2}, \qquad \Lambda_\Theta = \sum_{i \in \Theta} \binom{d_{i,R}}{2}. \tag{11}$$

In order to prove (7) we show below that

$$\Lambda_T = (1 + o_P(1)) 2^{-1} \beta^{-2} z_1^2 S_{Z,T}, \qquad \mathbf{P}\Big(\Lambda_\Theta \geq \Big(\frac{ap}{4\beta}\Big)^2 S_{Z,\Theta}\Big) = 1 - o(1). \tag{12}$$

Indeed, (11), (12) combined with the identity $S_{Z,T} + S_{Z,\Theta} = S_Z$, which holds with probability $1 - o(1)$ (see (6)), imply (7).

Proof of the first relation of (12). In view of Lemma 2 it suffices to show that

$$\mathbf{E}_Z \Lambda_T = (1 + o_P(1)) 2^{-1} \beta^{-2} z_1^2 S_{Z,T}, \qquad \mathbf{Var}_Z \Lambda_T = o_P(S_{Z,T}^2). \tag{13}$$

We note that the sum $S_{Z,T} = \sum_{i \in [n]} Z_i^2 \mathbb{I}_{Z_i < \ln^2 n}$ is superlinear in n as $n \to +\infty$, see Lemma 1. To prove the first relation of (13) we write

$$\Lambda_T = \sum_{i \in T} \sum_{\{j,k\} \subset T \setminus \{i\}} \mathbb{I}_{v_i \sim v_j} \mathbb{I}_{v_i \sim v_k}$$

and evaluate the expectation

$$\mathbf{E}_Z \Lambda_T = \sum_{i \in T} \sum_{\{j,k\} \subset T \setminus \{i\}} \bar{p}_{ij} \bar{p}_{ik}, \qquad \bar{p}_{ij} := \mathbf{P}_Z(v_i \sim v_j) = \mathbf{P}_Z(S_i \cap S_j \neq \emptyset).$$

Invoking the inequalities that follow from (5)

$$Z_i Z_j m^{-1} (1 - 2m^{-1} \ln^4 n) \leq \bar{p}_{ij} \leq Z_i Z_j m^{-1} \tag{14}$$

we obtain

$$\mathbf{E}_Z \Lambda_T = \Big(1 + O\Big(\frac{\ln^4 n}{m}\Big)\Big) \sum_{i \in T} \sum_{\{j,k\} \subset T \setminus \{i\}} \frac{Z_i^2 Z_j Z_k}{m^2} = (1 + o_P(1)) S_{Z,T} \frac{1}{2} \frac{\hat{z}_{1,T}^2}{\beta^2}.$$

Here we denote $\hat{z}_{1,T} := n^{-1} \sum_{i \in T} Z_i$. Finally, the law of large numbers implies $\hat{z}_{1,T} = z_1 + o_P(1)$. To prove the second relation of (13) we write Λ_T in the form $\Lambda_T = \mathbf{E}_Z \Lambda_T + L_T + Q_T$, where

$$\begin{aligned} L_T &= \sum_{\{i,j\} \subset T} \big(\mathbb{I}_{v_i \sim v_j} - \bar{p}_{ij}\big) \sum_{k \in T \setminus \{i,j\}} \big(\bar{p}_{ik} + \bar{p}_{jk}\big), \\ Q_T &= \sum_{i \in T} \sum_{\{j,k\} \subset T \setminus \{i\}} \big(\mathbb{I}_{v_i \sim v_j} - \bar{p}_{ij}\big)\big(\mathbb{I}_{v_i \sim v_k} - \bar{p}_{ik}\big). \end{aligned}$$

We observe that L_T and Q_T are conditionally uncorrelated (given $\{Z_n\}$). Therefore

$$\mathbf{Var}_Z \Lambda_T = \mathbf{Var}_Z L_T + \mathbf{Var}_Z Q_T. \tag{15}$$

We bound the summands on the right using (14). A simple calculation shows that

$$\begin{aligned}
\mathbf{Var}_Z L_T &= \sum_{\{i,j\}\subset T} \bar{p}_{ij}(1-\bar{p}_{ij})\Big(\sum_{k\in T\setminus\{i,j\}} \big(\bar{p}_{ik}+\bar{p}_{jk}\big)\Big)^2 \\
&\le \sum_{\{i,j\}\subset T} \frac{Z_iZ_j}{m}\Big(Z_i\beta^{-1}\hat{z}_{1,T}+Z_j\beta^{-1}\hat{z}_{1,T}\Big)^2 \\
&\le 2\beta^{-3}\hat{z}_{1,T}^3\sum_{i\in T} Z_i^3.
\end{aligned}$$

Now, invoking the inequality $\sum_{i\in T} Z_i^3 \le S_{Z,T}\max_{i\in T} Z_i \le S_{Z,T}^{3/2}$ and the bound $\hat{z}_{1,T} = O_P(1)$ we obtain $\mathbf{Var}_Z L_T = O_P(S_{Z,T}^{3/2}) = o_P(S_{Z,T}^2)$. Furthermore, we have

$$\mathbf{Var}_Z Q_T = \sum_{i\in T}\sum_{\{j,k\}\subset T\setminus\{i\}} \bar{p}_{ij}(1-\bar{p}_{ij})\bar{p}_{ik}(1-\bar{p}_{ik}) \le \sum_{i\in T}\sum_{\{j,k\}\subset T\setminus\{i\}} \bar{p}_{ij}\bar{p}_{ik}.$$

Invoking the inequality $\bar{p}_{ij}\bar{p}_{ik} \le Z_i^2Z_jZ_km^{-2}$ (which follows from (14)) we obtain

$$\mathbf{Var}_Z Q_T \le S_{Z,T}\hat{z}_{1,T}^2\beta^{-2} = O_P(S_{Z,T}) = o_P(S_{Z,T}^2).$$

Finally, (15) implies $\mathbf{Var}_Z\Lambda_T = o_P(S_{Z,T}^2)$.

Proof of the second relation of (12). For every $i\in\Theta$ and $j\in R$ we have, by (5),

$$\mathbf{P}_Z(v_i\sim v_j) \ge 0.9am^{-1}Z_i =: q_i.$$

Here 0.9 is a lower bound for the number $1-Z_iZ_j/(m-Z_j)$ valid for sufficiently large m,n. We note that conditionally, given $\{Z_i,\, i\ge 1\}$ and $|R|$, the random variable $d_{i,R}$ is a sum of independent indicators (their number is $|R|$) each having success probability at least q_i. Furthermore, $|R|$ has binomial distribution with mean np. Given $t\ge 0$ we have

$$\mathbf{P}(d_{i,R}\ge t) \ge \mathbf{P}\big(d_{i,R}\ge t\big|\,|R|\ge np/2\big) - r_1 \ge \mathbf{P}(L_i\ge t) - r_1. \tag{16}$$

Here $r_1 = \mathbf{P}(|R| < np/2)$ and L is the sum of $n' := \lfloor np/2\rfloor$ independent indicators with the same success probability q_i. Chernoff's inequality implies

$$\mathbf{P}(|R|<np/2) \le e^{-np/4} = O(n^{-9}), \qquad \mathbf{P}(L<n'q_i/2) \ge e^{-n'q_i/4} = O(n^{-9}). \tag{17}$$

Note that the second bound holds uniformly in $i\in\Theta$, since $Z_i\ge \ln^2 n$ for $i\in\Theta$. Choosing $t_i = n'q_i/2$ in (16) we obtain

$$\mathbf{P}(d_{i,R}\ge t_i,\, i\in\Theta) \ge 1 - O(n^{-8}).$$

This bound implies the second relation of (12).

Proof of (9). We recall that $\max_{i\in[n]} Z_i \le m$ with probability $1-o(1)$. Assuming that this inequality holds we prove below that $\mathbf{E}_Z\Delta \le O_P(n+n^{-3}S_Z^3)$. Now (9) follows by Lemma 2 (ii). We have $\Delta\le\Delta_L+\Delta_U$, where the numbers Δ_L and Δ_U of lucky and unlucky triangles satisfy

$$\begin{aligned}
\Delta_L &\le \sum_{w\in W}\sum_{\{i,j,k\}\subset[n]} \mathbb{I}_{w\in S_i}\mathbb{I}_{w\in S_j}\mathbb{I}_{w\in S_k}, \\
\Delta_U &\le \sum_{\substack{w,\tau,\varkappa\in W\\ w\ne\tau\ne\varkappa}}\sum_{\{i,j,k\}\subset[n]} \mathbb{I}_{w\in S_i}\mathbb{I}_{w\in S_j}\mathbb{I}_{\tau\in S_i}\mathbb{I}_{\tau\in S_k}\mathbb{I}_{\varkappa\in S_j}\mathbb{I}_{\varkappa\in S_k}.
\end{aligned}$$

Invoking the identity $\mathbf{P}_Z(w \in S_i) = m^{-1}Z_i$ and inequality $\mathbf{P}_Z(w, \tau \in S_i) \le m^{-2}Z_i^2$ we obtain

$$\begin{aligned}\mathbf{E}_Z\Delta_L &\le m^{-2}\sum_{\{i,j,k\}\subset[n]} Z_iZ_jZ_k \le \beta^{-2}\hat{z}_1^3 n = O_P(n),\\ \mathbf{E}_Z\Delta_U &\le m^{-3}\sum_{\{i,j,k\}\subset[n]} Z_i^2Z_j^2Z_k^2 \le \beta^{-3}n^{-3}S_Z^3.\end{aligned}$$

Proof of (8). By Lemma 3, we can find an increasing positive function $\psi(t) \uparrow +\infty$ as $t \to +\infty$ such that $\mathbf{E}Z\psi(Z) < \infty$. We can assume that $\psi(t) < t^{1/4}$, for $t \ge 1$. Denote $\delta_n = 1/\psi(n^{1/4})$ and $\tau_n = \mathbf{E}Z\psi(Z)\mathbb{I}_{\{Z \ge n^{1/4}\}}$. Put $\varkappa_n = \min\{\delta_n^{1/4}, \tau_n^{1/4}\}$. Clearly, $\delta_n \downarrow 0$, $\tau_n \downarrow 0$ and $\varkappa_n \downarrow 0$. We observe that

$$\mathbf{E}Z^2\mathbb{I}_{\{Z<\sqrt{n}\delta_n\}} \le \sqrt{n}\delta_n z_1, \qquad \mathbf{P}(Z \ge \sqrt{n}\delta_n) \le \frac{\mathbf{E}Z\psi(Z)\mathbb{I}_{\{Z\ge\sqrt{n}\delta_n\}}}{\sqrt{n}\delta_n\psi(\sqrt{n}\delta_n)} \le \frac{\tau_n}{\sqrt{n}}. \tag{18}$$

Now we estimate Δ. We observe that the number Δ_i of triangles incident to a given vertex $v_i \in V$ is at most $\binom{d_i}{2}$. Furthermore, Δ_i is always less than the total number of edges in the graph, denoted by $\mathcal{E}$. Therefore, we have

$$3\Delta = \sum_{i\in[n]}\Delta_i \le \sum_{i\in[n]:\, Z_i<\sqrt{n}\delta_n}\binom{d_i}{2} + \mathcal{E}\sum_{i\in[n]:\, Z_i\ge\sqrt{n}\delta_n} 1 =: U_1 + \mathcal{E}U_2. \tag{19}$$

We show below that $\mathbf{E}U_1 = O(n^{3/2}\delta_n)$, $\mathbf{E}U_2 = O(\sqrt{n}\tau_n)$ and $\mathbf{E}\mathcal{E} = O(n)$. These bounds together with (19) imply (8).
For $\mathcal{E} = \sum_{\{u,v\}\subset V}\mathbb{I}_{u\sim v}$ we have, by (5),

$$\mathbf{E}\mathcal{E} = \binom{n}{2}\mathbf{P}(v_1 \sim v_2) \le \binom{n}{2}\mathbf{E}(Z_1Z_2/m) \le \frac{n^2}{2m}(\mathbf{E}Z_1)^2 = O(n).$$

For $U_2 = \sum_{i\in[n]}\mathbb{I}_{Z_i>\sqrt{n}\delta_n}$ we have, see (18),

$$\mathbf{E}U_2 \le n\mathbf{P}(Z_1 > \sqrt{n}\delta_n) \le \sqrt{n}\tau_n.$$

It remains to bound $\mathbf{E}U_1$. For every i we have, by (5),

$$\mathbf{E}\Bigl(\binom{d_i}{2}\Big|Z_i\Bigr) = \sum_{\{k,r\}\subset[n]\setminus\{i\}}\mathbf{P}(v_k \sim v_i|Z_i)\mathbf{P}(v_r \sim v_i|Z_i) \le \binom{n-1}{2}Z_i^2(\mathbf{E}Z_1)^2m^{-2}. \tag{20}$$

Invoking the first inequality of (18) we obtain

$$\mathbf{E}\binom{d_i}{2}\mathbb{I}_{\{Z_i\le\sqrt{n}\delta_n\}} \le \binom{n-1}{2}\frac{z_1^2}{m^2}\mathbf{E}Z_i^2\mathbb{I}_{\{Z_i\le\sqrt{n}\delta_n\}} \le \frac{z_1^3}{2\beta^2}\sqrt{n}\delta_n.$$

Finally, we have

$$\mathbf{E}U_1 = \sum_{i\in[n]}\binom{d_i}{2}\mathbb{I}_{Z_i\le\sqrt{n}\delta_n} \le \frac{z_1^3}{2\beta^2}n^{3/2}\delta_n.$$

□

Proof of Theorem 2. In the proof we use the notation $X_i = |D_{w_i}|$, $w_i \in W$, and $S_X = \sum_{i \in [n]} X_i^3$. We firstly count triangles. For every $w \in W$ there are $N_w := \binom{|D_w|}{3}$ lucky triangles covered by D_w. We have, by inclusion-exclusion, that

$$N - N^* \le \Delta_L \le N, \qquad \text{where} \qquad N = \sum_{w \in W} \binom{|D_w|}{3}, \qquad N^* = \sum_{\{w,\tau\} \subset W} \binom{|D_w \cap D_\tau|}{3}. \quad (21)$$

Here $\binom{|D_w \cap D_\tau|}{3}$ counts triangles covered by D_w and D_τ simultaneously. Every unlucky triangle has its edges covered by distinct sets. Therefore, Δ_U is at most the sum

$$N^{**} := \sum_{\{x,y,z\} \subset V} \sum_{1 \le i \ne j \ne k \le m} \mathbb{I}_{\{x,y\} \subset D_i} \mathbb{I}_{\{x,z\} \subset D_j} \mathbb{I}_{\{y,z\} \subset D_k}.$$

We estimate the total number of triangles Δ from the inequalities $\Delta_L \le \Delta \le \Delta_L + \Delta_U$. Hence

$$|\Delta - N| \le N^* + N^{**}. \quad (22)$$

We secondly count 2-paths. We have $\Lambda = \Lambda_L + \Lambda_U - \Lambda_{LU}$, where Λ_{LU} is the number of paths labeled both lucky and unlucky. For the number of lucky paths $\Lambda_L = 3\Delta_L$, we can evaluate Λ_L using (21). Furthermore, the number Λ_U of unlucky paths is at most the sum

$$M^* := \sum_{\{w,\tau\} \subset W} |D_w \cap D_\tau| \times |D_w| \times |D_\tau|.$$

Here $|D_w \cap D_\tau| \times |D_w| \times |D_\tau|$ is an upper bound for the number of 2-paths with the central vertex belonging to $D_w \cap D_\tau$ and with the endpoints belonging to $D_w \setminus D_\tau$ and $D_\tau \setminus D_w$ respectively. From the inequalities $\Lambda_L \le \Lambda \le \Lambda_L + \Lambda_U$ we obtain

$$|\Lambda - 3N| \le 3N^* + M^*. \quad (23)$$

Finally, we derive the relation $C_{G^*} = 3\Delta/\Lambda = 1 + o_P(1)$ from (22), (23) and the bounds $N^*, N^{**}, M^* = o_P(N)$ shown below.
Let us bound N^*, N^{**}, M^*. We note that the sum S_X is superlinear in m. Indeed, Lemma 1 implies that $\mathbf{P}(S_X > m\phi_m) = 1 + o(1)$ for some $\phi_m \uparrow +\infty$. A simple consequence of this fact is that $6N = (1 + o_P(1))S_X$ is superlinear in m as well. Furthermore, the bounds $N^*, N^{**}, M^* = o_P(N)$ are equivelaent to the bounds $N^*, N^{**}, M^* = o_P(S_X)$. In order to show these we prove that

$$\mathbf{E}_{\mathbb{X}} N^* = o_P(S_X), \qquad \mathbf{E}_{\mathbb{X}} N^{**} = o_P(S_X), \qquad \mathbf{E}_{\mathbb{X}} M^* = o_P(S_X), \quad (24)$$

and apply Lemma 2. To prove the first bound of (24) we write $\binom{|D_w \cap D_\tau|}{3}$ in the form

$$\binom{|D_w \cap D_\tau|}{3} = \sum_{\{x,y,z\} \subset V} \mathbb{I}_{\{x,y,z\} \subset D_w} \mathbb{I}_{\{x,y,z\} \subset D_\tau},$$

evaluate the conditional expectation

$$\mathbf{E}_{\mathbb{X}} N^* = \binom{n}{3} \sum_{\{i,j\} \subset [m]} \binom{X_i}{3} \binom{X_j}{3} \binom{n}{3}^{-2},$$

and invoke (37) of Lemma 4. To prove the second bound of (24) we evaluate

$$\mathbf{E}_{\mathbb{X}} N^{**} = \binom{n}{3} \sum_{1 \le i \ne j \ne k \le m} \binom{X_i}{2} \binom{X_j}{2} \binom{X_k}{2} \binom{n}{2}^{-3}$$

and invoke (38) Lemma 4. To prove the third bound of (24) we evaluate

$$\mathbf{E}_{\mathbb{X}} M^* = \mathbf{E}_{\mathbb{X}} \sum_{\{i,j\} \subset [m]} X_i X_j \sum_{x \in V} \mathbb{I}_{x \in D_i} \mathbb{I}_{x \in D_j} = n^{-1} \sum_{\{i,j\} \subset [m]} X_i^2 X_j^2$$

and invoke (39) of Lemma 4. □

Proof of Theorem 3. Before the proof we introduce some notation. We fix positive sequences $\varepsilon \downarrow 0$ and $t_n \uparrow +\infty$ such that $\mathbf{P}\big(\max_{i\in[n]} Y_i < \varepsilon_n t_n^{-1} n\big) = 1 - o(1)$, see Lemma 3. Note that $\mathbf{E} Y_1 \mathbb{I}_{Y_1 \ge t_n} = o(1)$ implies $n^{-1}\sum_{i\in[n]} Y_i \mathbb{I}_{Y_i \ge t_n} = o_P(1)$. We recall that the inhomogeneous graph G is defined by a bipartite graph H with the bipartition $V \cup W$. We color vertices in V white and those in W black. Given a bipartite graph $H' = (V', W'; E')$ with the bipartition $V' \cup W'$ and the edge set E', we color vertices in V' white and those in W' black. Define the bipartite graphs

$$\begin{aligned}
H_1 &= \Big(\{1,2,3\}, \{a\}; \big\{\{1,a\},\{2,a\},\{3,a\}\big\}\Big),\\
H_2 &= \Big(\{1,2,3\}, \{a,b\}; \big\{\{1,a\},\{2,a\},\{2,b\},\{3,b\}\big\}\Big),\\
H_3 &= \Big(\{1,2,3\}, \{a,b,c\}; \big\{\{1,a\},\{2,a\},\{2,b\},\{3,b\},\{1,c\},\{3,c\}\big\}\Big),\\
H_4 &= \Big(\{1,2,3\}, \{a,b,c\}; \big\{\{1,a\},\{2,a\},\{2,b\},\{3,b\},\{1,c\},\{2,c\}\big\}\Big),\\
H_5 &= \Big(\{1,2,3\}, \{a,b\}; \big\{\{1,a\},\{2,a\},\{3,a\},\{1,b\},\{2,b\}\big\}\Big).
\end{aligned}$$

For $1 \le i \le 5$ we denote by $\mathcal{H}_i$ the set of copies of H_i in H. The number of copies is denoted $N_i = |\mathcal{H}_i|$. We note that every $H' \in \mathcal{H}_1$ defines a lucky triangle in G, $H'' \in \mathcal{H}_2$ defines an unlucky path in G, and $H''' \in \mathcal{H}_3$ defines an unlucky triangle in G. In particular, we have $\Delta_L \le N_1$, $\Lambda_L \le 3N_1$, $\Delta_U \le N_3$, and $\Lambda_U \le N_2$. We call an edge $v_i \sim v_j$ of G heavy if $Y_i Y_j > \varepsilon_n n$. A subgraph of G is called heavy if it contains a heavy edge. Otherwise it is called light. The number of heavy (light) copies of H_i is denoted N_i^+ (N_i^-).

The theorem follows from (4) and the relations

$$N_1 = (1 + o_P(1))\tilde{\mathbf{E}} N_1 = 6^{-1}\beta^{-3/2} b_1^3 S_X + o_P(S_X), \tag{25}$$
$$N_2 = (1 + o_P(1))\tilde{\mathbf{E}} N_2 = 2^{-1} a_2^2 b_1^2 S_Y + o_P(S_Y), \tag{26}$$
$$\Delta_L = N_1 + o_P(S_X), \tag{27}$$
$$\Delta_U = o_P(S_Y), \tag{28}$$
$$\Lambda_U = N_2 + o_P(S_Y) + o_P(S_X), \tag{29}$$
$$\Lambda_{LU} = o_P(S_X) + o_P(S_Y). \tag{30}$$

Relations (25), (26) follow from Lemmas 2, 5, 6. It remains to prove (27-30).
We begin with establishing auxiliary facts. Denote

$$L_n := n^{-1} \sum_{\{i,j\}\subset[n]} Y_i Y_j (Y_i + Y_j) \mathbb{I}_{Y_i Y_j > \varepsilon_n n}, \qquad L_n' := n^{-1} \sum_{\{i,j\}\subset[n]} Y_i Y_j \mathbb{I}_{Y_i Y_j > \varepsilon_n n}.$$

We have

$$L_n = o_P(S_Y), \qquad L_n' = o_P(n) \tag{31}$$
$$N_1^+ = o_P(S_X), \qquad N_2^+ = o_P(S_Y), \tag{32}$$
$$N_3^- = o_P(S_Y), \qquad N_4^- = o_P(S_Y), \qquad N_5^- = o_P(S_X). \tag{33}$$

Proof of (31). On the event $\{\max_{i\in[n]} Y_i \le \varepsilon_n t_n^{-1} n\}$ which has probability $1-o(1)$ we have

$$Y_iY_j(Y_i+Y_j)\mathbb{I}_{Y_iY_j>\varepsilon_n n} \le Y_i^2Y_j\mathbb{I}_{Y_j>t_n} + Y_j^2Y_i\mathbb{I}_{Y_i>t_n}.$$

Hence $L_n \le S_Y n^{-1}\sum_{i\in[n]} Y_i\mathbb{I}_{Y_i>t_n}$. The bound $n^{-1}\sum_{i\in[n]} Y_i\mathbb{I}_{Y_i>t_n} = o_P(1)$ implies the first bound of (31). The second bound is obtained in a similar way.
Proof of (32). We combine Lemma 3 with the inequalities

$$\begin{aligned}
\tilde{\mathbf{E}}N_1^+ &\le \sum_{\{x,y,z\}\subset V}\sum_{w\in W}\frac{Y_xY_yY_zX_w^3}{(nm)^{3/2}}\big(\mathbb{I}_{\{Y_xY_y>\varepsilon_n n} + \mathbb{I}_{\{Y_xY_z>\varepsilon_n n} + \mathbb{I}_{\{Y_yY_z>\varepsilon_n n}\big)\\
&\le 3\beta^{-3/2}\hat{b}_1 n^{-1}L_n' S_X = o_P(S_X),
\end{aligned} \tag{34}$$

$$\begin{aligned}
\tilde{\mathbf{E}}N_2^+ &\le \sum_{\Lambda}\sum_{w,\tau\in W:\, w\ne\tau}\frac{Y_yY_x^2Y_zX_w^2X_\tau^2}{n^2m^2}\big(\mathbb{I}_{Y_xY_y>\varepsilon_n n} + \mathbb{I}_{Y_xY_z>\varepsilon_n n}\big)\\
&= \sum_{x\in V}\sum_{y,z\in V\setminus\{x\}:\, y\ne z}\sum_{w,\tau\in W:\, w\ne\tau}\frac{Y_yY_x^2Y_zX_w^2X_\tau^2}{n^2m^2}\mathbb{I}_{Y_xY_y>\varepsilon_n n}\\
&\le \hat{a}_2^2\hat{b}_1 L_n = o_P(S_Y).
\end{aligned} \tag{35}$$

In the last steps of (34) and (35) we have used (31).
Proof of (33). We combine Lemma 3 with the inequalities

$$\begin{aligned}
\tilde{\mathbf{E}}N_3^- &\le 3!\sum_{\{x,y,z\}\subset V}\sum_{\{w,\tau,\eta\}\subset W}\frac{Y_x^2Y_y^2Y_z^2X_w^2X_\tau^2X_\eta^2}{(nm)^3}\mathbb{I}_{Y_xY_y\le\varepsilon_n n}\\
&\le \varepsilon_n\hat{a}_2^3\sum_{x,y,z\in V:\, x\ne y\ne z}\frac{Y_xY_yY_z^2}{n^2}\\
&\le \varepsilon_n\hat{a}_2^3\hat{b}^2 S_Y = o_P(S_Y),
\end{aligned}$$

$$\begin{aligned}
\tilde{\mathbf{E}}N_4^- &\le \sum_{x,y,z\in V:\, x\ne y\ne z}\sum_{w,\tau,\eta\in W:\, w\ne\tau\ne\eta}\frac{Y_x^2Y_y^3Y_zX_w^2X_\tau^2X_\eta^2}{(nm)^3}\mathbb{I}_{Y_xY_y\le\varepsilon_n n}\\
&\le \varepsilon_n\hat{a}_2^3\sum_{x,y,z\in V:\, x\ne y\ne z}\frac{Y_xY_y^2Y_z}{n^2}\\
&\le \varepsilon_n\hat{a}_2^3\hat{b}_1^2 S_Y = o_P(S_Y),
\end{aligned}$$

$$\begin{aligned}
\tilde{\mathbf{E}}N_5^- &\le \sum_{x,y,z\in V:\, x\ne y\ne z}\sum_{w,\tau\in W:\, w\ne\tau}\frac{Y_x^2Y_y^2Y_zX_w^3X_\tau^2}{(nm)^{5/2}}\mathbb{I}_{Y_xY_y\le\varepsilon_n n}\\
&\le \varepsilon_n\hat{a}_2 S_X\sum_{x,y,z\in V:\, x\ne y\ne z}\frac{Y_xY_yY_z}{\beta^{3/2}n^3}\\
&\le \varepsilon_n\hat{a}_2\hat{b}_1^3\beta^{-3/2}S_X = o_P(S_X).
\end{aligned}$$

Now we are ready to prove (27-30).

Proof of (29). Given a light unlucky path $x \sim y \sim z$ of G, let $\mathcal{H}_2^{x,y,z} \subset \mathcal{H}_2$ denote the set of copies of H_2 defining this path. Fix an element $H_2^* \in \mathcal{H}_2^{x,y,z}$. All the other elements of $\mathcal{H}_2^{x,y,z}$ are called duplicates. We do this for each light unlucky path. We claim that the total number of duplicates is at most $N_3^- + N_5^-$. Indeed, given $H_2^* \in \mathcal{H}_2^{x,y,z}$ with bipartition denoted by $V' = \{x, y, z\} \subset V$ and $W' = \{w, \tau\} \subset W$, one potential duplicate is the distinct element of $\mathcal{H}_2^{x,y,z}$ with the same attribute set W'. The union of both copies of H_2 defines the complete bipartite graph on $V' \cup W'$ and hence a copy of H_5 on $V' \cup W'$. The duplicates of this kind are counted by N_5^-. Remaining possible duplicates of H_2^* have attribute sets different from W'. We note that a duplicate H_2'' whose attribute set $W'' \neq W'$ defines a copy of H_3. Indeed, for $W'' \cap W' = \{w\}$ the union $H_2^* \cup H_2''$ is a copy of H_3. Furthermore, for $W'' \cap W' = \emptyset$, the union $H_2^* \cup H_2''$ with deleted vertex w is a copy of H_3. Note that distinct duplicates H_2'' define distinct copies of H_3. Hence, their total number is at most N_3^-. Our claim is established. It implies that the number of light unlucky paths is at least $N_2^- - N_3^- - N_5^-$. Hence the total number of unlucky paths

$$\Lambda_U \geq N_2^- - N_3^- - N_5^- = N_2 - N_2^+ - N_3^- - N_5^-.$$

These inequalities in combination with (32), (33) and the simple inequality $\Lambda_U \leq N_2$ imply (29).
Proof of (30). A light path $x \sim y \sim z$ receives both labels lucky and unlucky whenever H has a light copy of H_1 with the vertex set $\{x, y, z\} \cup \{w\}$ and it has a light copy of H_2 with the vertex set $\{x, y, z\} \cup \{w', \tau\}$. Here w and $w' \neq \tau$ are arbitrary elements of W not necessarily all distinct. The union of these two copies contains a light copy of H_5. Hence the number of light paths which are both lucky and unlucky is at most N_5^-. The number of heavy unlucky paths is at most N_2^+. Putting things together we obtain $\Lambda_{LU} \leq N_5^- + N_2^+$. Now (32), (33) imply (30).
Proof of (28). Every heavy unlucky triangle contains at least two heavy unlucky paths. Hence the number of such triangles is at most $N_2^+/2$. The number of light unlucky triangles is at most N_3^-. Hence $\Delta_U \leq N_3^- + N_2^+/2 = o_P(S_Y)$.
Proof of (27). Given a light lucky triangle $x \sim y \sim z \sim x$ of G, let $\mathcal{H}_1^{x,y,z} \subset \mathcal{H}_1$ denote the set of copies of H_1 defining this triangle. Fix an element $H_1^* \in \mathcal{H}_1^{x,y,z}$. It is the complete bipartite graph on the bipartition $\{x, y, z\} \cup \{w\}$ for some $w \in W$. All the other elements of $\mathcal{H}_1^{x,y,z}$ are called duplicates. We claim that the total number of duplicates is at most N_5^-. Indeed, for any duplicate $H_1' \in \mathcal{H}_1^{x,y,z}$ with bipartition denoted by $\{x, y, z\} \cup \{w'\}$, the union $H_1^* \cup H_1'$ is the complete bipartite graph on $\{x, y, z\} \cup \{w, w'\}$. We remove the edge $\{z, w'\}$ and obtain a copy of H_5. We conclude that the number of light lucky triangles is at least $N_1^- - N_5^-$. Hence $\Delta_L \geq N_1^- - N_5^- = N_1 - N_1^+ - N_5^-$. These inequalities in combination with (32), (33) and the simple inequality $\Delta_L \leq N_1$ imply (27).
In the proof we use the fact that $n = o_P(S_Y)$ and $m = o_P(S_X)$.

□

Proof of Remark 1. For $\alpha < 1$, random variables $S_X(c_x\Gamma(1-\alpha)m)^{-1/\alpha}$ and $S_Y(c_y\Gamma(1-\alpha)n)^{-1/\alpha}$ converge in distribution to independent and identically distributed α stable random variables, say Z_1, Z_2, having the Laplace transform $s \to \mathbf{E}e^{-sZ_1} = e^{-s^\alpha}$, see Theorem 2 of Section 6 of Chapter XIII of [7]. Here Γ is Euler's Gamma function. Hence the statement (i).
For $\alpha = 1$, there exist deterministic sequences $b_{m,x} = (c_x + o(1)) \ln m$ and $b_{n,y} = (c_y + o(1)) \ln n$ such that the random variables $m^{-1}S_x - b_{m,x}$ and $n^{-1}S_Y - b_{n,y}$ converge in distribution to independent asymmetric stable random variables with the characteristic exponent $\alpha = 1$, see Theorem 3 of Section 5 of Chapter XVII of [7]. Hence the statement (ii). □

3 Appendix

In Appendix A we place auxiliary lemmas. Proofs are given in Appendix B. We remark that Lemmas 4 and 5, 6 refer to the notation of the proofs of Theorems 2 and 3 respectively.

3.1 Appendix A

Lemma 1. *Let $X_1, X_2, \ldots$ be a sequence of non-negative random variables converging in distribution to a random variable X. Assume that $\mathbf{E}X = \infty$. Then for some positive nonrandom sequence $\{\phi_n\}$ converging to $+\infty$ we have*

$$\mathbf{P}\big(X_{n,1} + \cdots + X_{n,n} > \phi_n n\big) = 1 - o(1). \tag{36}$$

Here $X_{n,1}, \ldots, X_{n,n}$ are iid copies of X_n.

Lemma 2. *Let $\{Z_n\}$ and $\eta = \{\eta_n\}$ be sequences of random variables defined on the same probability space. Let $\mathbf{E}_\eta$ denote the conditional expectation given η. Assume that $\mathbf{E}_\eta Z_n = 0$ implies $Z_n = 0$. Then*
(i) $\mathbf{E}_\eta(Z_n - \mathbf{E}_\eta Z_n)^2 = o_P\big((\mathbf{E}_\eta Z_n)^2\big)$ implies $Z_n = (1 + o_P(1))\mathbf{E}_\eta Z_n$;
(ii) $Z_n = O_P(\mathbf{E}_\eta Z_n)$.

Lemma 3. *Let $t > 0$. Let Z be a non-negative random variable with $\mathbf{E}Z < \infty$.*
(i) There exists a positive increasing function $\psi(\cdot)$ such $\psi(t) \uparrow +\infty$ as $t \uparrow +\infty$ and $\mathbf{E}Z\psi(Z) < \infty$. Furthermore, there exists a positive decreasing function $\varepsilon(\cdot)$ such that $\varepsilon(s) \downarrow 0$ as $s \uparrow +\infty$ and $\mathbf{P}\big(Z > s\varepsilon(s)\big) = o(s^{-1})$ for $s \to +\infty$.
(ii) Let $Z_1, Z_2, \ldots$ be iid copies of Z. Let $n \to +\infty$. Then $Z_1^{1+t} + \cdots + Z_n^{1+t} = o_P(n^{1+t})$. Furthermore, for $\varepsilon(\cdot)$ of statement (i), we have $\mathbf{P}(\max_{1 \le i \le n} Z_i > n\varepsilon(n)) = o(1)$.

We remark that the functions $\psi(\cdot)$, $\varepsilon(\cdot)$ depend on the probability distribution of Z.

Lemma 4. *Let $X_1, X_2, \ldots$ be a sequence of non-negative random variables converging in distribution to a random variable X. Assume that $\mathbf{E}X^3 = \infty$ and $0 < \mathbf{E}X^2 < \infty$. Assume that $\mathbf{E}X_n^2 < \infty$, for each $n = 1, 2, \ldots$, and $\lim_n \mathbf{E}X_n^2 = \mathbf{E}X^2$. Let $\{m_n, n \ge 1\}$ be an integer sequence and, for every n, let $X_{n,1}, \ldots, X_{n,m_n}$ be iid copies of X_n. Let $n \to +\infty$. Assume that $m_n \uparrow +\infty$. Denote $S_{X,n} = \sum_{j \in [m_n]} X_{n,j}^3$. We have*

$$m_n^{-3} \sum_{\{j,k\} \subset [m_n]} X_{n,j}^3 X_{n,k}^3 = o_P\big(S_{X,n}\big), \tag{37}$$

$$m_n^{-3} \sum_{\{j,k,r\} \subset [m_n]} X_{n,j}^2 X_{n,k}^2 X_{n,r}^2 = o_P\big(S_{X,n}\big), \tag{38}$$

$$m_n^{-1} \sum_{\{j,k\} \subset [m_n]} X_{n,j}^2 X_{n,k}^2 = o_P\big(S_{X,n}\big). \tag{39}$$

Lemma 5. *Assume that $\mathbf{E}X_1^2 < \infty$ and $\mathbf{E}Y_1 < \infty$.*
(i) For $\mathbf{E}X_1^3 = \infty$ we have $\tilde{\mathbf{E}}\big(N_1 - \tilde{\mathbf{E}}N_1\big)^2 = o_P(S_X^2)$.
(ii) We have $\tilde{\mathbf{E}}N_1 = 6^{-1}\beta^{-3/2}b_1^3 S_X + o_P(S_X)$.

Lemma 6. *Assume that $\mathbf{E}X_1^2 < \infty$, $\mathbf{E}Y_1 < \infty$.*
(i) For $\mathbf{E}Y_1^2 = \infty$ we have $\tilde{\mathbf{E}}\big(N_2 - \tilde{\mathbf{E}}N_2\big)^2 = o_P(S_Y^2)$.
(ii) We have $\tilde{\mathbf{E}}N_2 = 2^{-1}a_2^2 b_1^2 S_Y + o_P(S_Y)$.

3.2 Appendix B

Proof of Lemma 1. We need some notation. Given random variable Z and sequence $Z_1, \dots, Z_N$ of iid copies of Z, we denote $S_N(Z) = Z_1 + \cdots + Z_N$. For a constant $A > 0$ we denote the truncated random variable $Z[A] = Z\mathbb{I}_{\{Z \le A\}}$ and $S_N(Z[A])$ denotes the sum of truncated iid copies of Z.
Let us prove (36). Choose a sequence $\{A_n\}$ of positive constants converging to $+\infty$ (slowly enough) such that

$$\mathbf{Var}(X[A_n]) = o(n) \qquad \text{and} \qquad \mathbf{E}X_n^i[A_n] - \mathbf{E}X^i[A_n] = o(1), \quad i = 1, 2. \tag{40}$$

In particular, we have $\mathbf{E}X_n[A_n] \to +\infty$ and $\mathbf{Var}X_n[A_n] = o(n)$ as $n \to +\infty$. Now Chebyshev's inequality implies

$$\mathbf{P}\big(S_n(X_n[A_n]) < (n/2)\mathbf{E}X_n[A_n]\big) \le 4n^{-1}\big(\mathbf{E}X_n[A_n]\big)^{-2}\mathbf{Var}X_n[A_n] = o(1).$$

Hence, for $S_n(X_n) \ge S_n(X_n[A_n])$ and $\phi_n = 0.5\mathbf{E}X_n[A_n]$ we obtain $\mathbf{P}(S_n(X_n) < \phi_n n) = o(1)$. □

Proof of Lemma 2. Let $\mathbf{P}_\eta$ denote the conditional probability given η and let z_n denote $\mathbf{E}_\eta Z_n$. We obtain (i) by Chebyshev's inequality: $\forall\ \varepsilon > 0$

$$\begin{aligned} \mathbf{P}(|Z_n - z_n| > \varepsilon z_n) &= \mathbf{E}\min\big\{1,\ \mathbf{P}_\eta(|Z_n - z_n| > \varepsilon z_n)\big\}\mathbb{I}_{\{z_n \ne 0\}} \\ &\le \mathbf{E}\min\big\{1, (\varepsilon z_n)^{-2}\mathbf{E}_\eta(Z_n - z_n)^2\big\}\mathbb{I}_{\{z_n \ne 0\}} \\ &= o(1). \end{aligned}$$

In the last step we used the fact that $\mathbf{P}(\mathbf{E}_\eta(Z_n - z_n)^2 > \delta z_n^2) = o(1)$ for any $\delta > 0$.
We obtain (ii) by Markov's inequality: $\forall\ \varepsilon > 0$

$$\mathbf{P}(Z_n > \varepsilon^{-1}z_n) = \mathbf{E}\Big(\mathbf{P}_\eta(Z_n > \varepsilon^{-1}z_n)\mathbb{I}_{\{z_n \ne 0\}}\Big) \le \mathbf{E}\big(\varepsilon z_n^{-1}\mathbf{E}_\eta Z_n\big)\mathbb{I}_{\{z_n \ne 0\}} \le \varepsilon.$$

□

Proof of Lemma 3. The poof is elementary. We present it for reader's convenience.
Proof of (i). $\mathbf{E}Z < \infty$ implies that the function $\phi(t) = \mathbf{E}Z\mathbb{I}_{\{Z>t\}}$ is non-increasing and $\phi(t) \to 0$ as $t \to +\infty$. Choose an increasing positive sequence $\{s_k\}_{k\ge1}$ such that $s_k \uparrow +\infty$ and $\phi(s_k) \le 2^{-k}$ and $s_1 \ge 1$. Put $s_0 = 0$. Consider the non-decreasing function $\psi : [0, +\infty) \to [0, +\infty)$ attaining value k on the interval $[s_{k-1}, s_k]$, for $k = 1, 2 \dots$. Clearly, $\psi(t) \to +\infty$ as $t \uparrow +\infty$ and we have $\mathbf{E}Z\psi(Z) < \infty$. Furthermore, we can easily modify $\psi(\cdot)$ in order to obtain a strictly increasing function satisfying the requirements of statement (i). Now we choose $\varepsilon(\cdot)$ decaying slowly enough ($\varepsilon(s) \downarrow 0$ as $s \uparrow +\infty$) so that $s\varepsilon(s) \to +\infty$ and $\varepsilon(s)\psi\big(s\varepsilon(s)\big) \to +\infty$ as $s \to +\infty$. Finally, Markov's inequality implies

$$\mathbf{P}\big(Z > s\varepsilon(s)\big) \le \big(s\varepsilon(s)\psi(s\varepsilon(s))\big)^{-1}\mathbf{E}Z\psi(Z)\mathbb{I}_{Z\psi(Z) > s\varepsilon(s)} = o(s^{-1}).$$

Proof of (ii). We write $Z_1^{1+t} + \cdots + Z_n^{1+t} \le AB$, where $A := Z_1 + \cdots + Z_n$ and $B := \max_{i\in[n]} Z_i^t$, and invoke the bounds $A = O_P(n)$ and $B = o_P(n^t)$. The first bound follows by the law of large numbers. The second one follows by Markov's inequality and the union bound: $\forall\delta > 0$ we have

$$\mathbf{P}(B > \delta^t n^t) = \mathbf{P}\Big(\max_{i\in[n]} Z_i > \delta n\Big) \le n\mathbf{P}(Z_1 > \delta n) \le \delta^{-1}\mathbf{E}Z_1\mathbb{I}_{Z_1 > \delta n} = o(1).$$

Similarly, from (i) we obtain $\mathbf{P}\big(\max_{i\in[n]} Z_i > n\varepsilon(n)\big) \le n\mathbf{P}\big(Z_1 > n\varepsilon(n)\big) = o(1)$. □

Proof of Lemma 4. Proof of (37). The event $\mathcal{A}_n = \{\max_{j\in[m_n]} X_j \geq m_n^{4/7}\}$ has probability

$$\mathbf{P}(\mathcal{A}_n) \leq m_n \mathbf{P}(X_{n,j} \geq m_n^{4/7}) \leq m_n^{-1/7}\mathbf{E}X_{n,j}^2 = o(1).$$

On the complement event, the left side of (37) is less than

$$m_n^{-3} \sum_{\{j,k\}\subset[m_n]} m_n^{12/7} X_{n,k}^3 \leq m_n^{-2/7} S_{X,n}.$$

Proof of (38). Denote $S_{*n} = m_n^{-1}\sum_{i\in[m_n]} X_{n,j}^2$. The relation $\mathbf{E}S_{*n} = \mathbf{E}X_n^2 \to \mathbf{E}X^2$ implies $S_{*n} = O_P(1)$. The left side of (38) is less than $S_{*n}^3 = O_P(1)$. The right side is superlinear in m, by Lemma 1.
Proof of (39). The left side of (39) is less than $m_n S_{*n}^2 = O_P(m_n)$. The right side is superlinear in m, by Lemma 1. □

Proof of Lemma 5. It is convenient to write N_1 in the form

$$N_1 = \sum_{w\in W} U_w, \qquad U_w := \sum_{\{x,y,z\}\subset V} \mathbb{I}_{xw}\mathbb{I}_{yw}\mathbb{I}_{zw}.$$

Proof of (i). Given $\mathbb{X}$ and $\mathbb{Y}$, the random variables U_w, $w \in W$ are conditionally independent. Hence

$$\tilde{\mathbf{E}}(N_1 - \tilde{\mathbf{E}}N_1)^2 = \sum_{w\in W} \tilde{\mathbf{E}}\tilde{U}_w^2, \qquad \text{where} \qquad \tilde{U}_w := U_w - \tilde{\mathbf{E}}U_w. \tag{41}$$

We bound every expectation $\tilde{\mathbf{E}}\tilde{U}_w^2$ using conditional Hoeffding's decomposition $\tilde{U}_w = L_w + Q_w + K_w$, where the components

$$L_w = \sum_{x\in V} \tilde{\mathbf{E}}(\tilde{U}_w|\mathbb{I}_{xw}), \qquad Q_w = \sum_{\{x,y\}\subset V} \tilde{\mathbf{E}}\big(\tilde{U}_w - L_w\big|\mathbb{I}_{xw}, \mathbb{I}_{yw}\big), \tag{42}$$
$$K_w = \sum_{\{x,y,z\}\subset V} \tilde{\mathbf{E}}\big(\tilde{U}_w - L_w - Q_w\big|\mathbb{I}_{xw}, \mathbb{I}_{yw}, \mathbb{I}_{zw}\big),$$

called the linear, quadratic and cubic part of the decomposition, are conditionally uncorrelated. We have in particular that

$$\tilde{\mathbf{E}}\tilde{U}_w^2 = \tilde{\mathbf{E}}L_w^2 + \tilde{\mathbf{E}}Q_w^2 + \tilde{\mathbf{E}}K_w^2.$$

Moreover the summands of all three sums of (42) are conditionally uncorrelated (given $\mathbb{X}, \mathbb{Y}$. Now (i) follows from (41) and the bounds shown below

$$\sum_{w\in W} \tilde{\mathbf{E}}L_w^2 = o_P(S_X^2), \qquad \sum_{w\in W} \tilde{\mathbf{E}}Q_w^2 = o_P(S_X^2), \qquad \sum_{w\in W} \tilde{\mathbf{E}}K_w^2 = o_P(S_X^2). \tag{43}$$

Let us prove (43). Denote, for $x, y, z \in V$ and $w \in W$,

$$s_{x|w} = \sum_{\{y,z\}\subset V\setminus\{x\}} p_{yw}p_{zw}, \qquad s_{xy|w} = \sum_{z\in V\setminus\{x,y\}} p_{zw}.$$

A straightforward calculation shows that

$$\tilde{\mathbf{E}}(\tilde{U}_w|\mathbb{I}_{xw}) = (\mathbb{I}_{xw} - p_{xw})s_{x|w} =: l_w(x),$$
$$\tilde{\mathbf{E}}\big(\tilde{U}_w - L_w\big|\mathbb{I}_{xw}, \mathbb{I}_{yw}\big) = (\mathbb{I}_{xw} - p_{xw})(\mathbb{I}_{yw} - p_{yw})s_{xy|w} =: q_w(x, y),$$
$$\tilde{\mathbf{E}}\big(\tilde{U}_w - L_w - Q_w\big|\mathbb{I}_{xw}, \mathbb{I}_{yw}, \mathbb{I}_{zw}\big) = (\mathbb{I}_{xw} - p_{xw})(\mathbb{I}_{yw} - p_{yw})(\mathbb{I}_{zw} - p_{zw}) =: k_w(x, y, z).$$

Invoking the simple inequalities $s_{x|w} \le X_w^2 \beta_n^{-1} \hat{b}_1^2$ and $s_{xy|w} \le X_w \beta_n^{-1/2} \hat{b}_1$ we obtain

$$\begin{aligned} \tilde{\mathbf{E}} l_w^2(x) &= \tilde{\mathbf{E}}\big((\mathbb{I}_{xw} - p_{xw}) s_{x|w}\big)^2 \le p_{xw} s_{x|w}^2 \le X_w^5 \beta_n^{-5/2} Y_x n^{-1} \hat{b}_1^4, \\ \tilde{\mathbf{E}} q_w^2(x,y) &= \tilde{\mathbf{E}}\big((\mathbb{I}_{xw} - p_{xw})(\mathbb{I}_{yw} - p_{yw}) s_{xy|w}\big)^2 \le p_{xw} p_{yw} s_{xy|w}^2 \le X_w^4 \beta_n^{-2} n^{-2} Y_x Y_y \hat{b}_1^2, \\ \tilde{\mathbf{E}} k_w^2(x,y,z) &\le p_{xw} p_{yw} p_{zw} \le X_w^3 \beta_n^{-3/2} n^{-3} Y_x Y_y Y_z. \end{aligned} \tag{44}$$

We note that for $x, y, z \in V$ the random variables $l_w(x)$, $q_w(x,y)$ and $k_w(x,y,z)$ are uncorrelated. Hence

$$\tilde{\mathbf{E}} L_w^2 = \sum_{x \in V} \tilde{\mathbf{E}} l_w^2(x), \qquad \tilde{\mathbf{E}} Q_w^2 = \sum_{\{x,y\} \subset V} \tilde{\mathbf{E}} q_w^2(x,y), \qquad \tilde{\mathbf{E}} K_w^2 = \sum_{\{x,y,z\} \subset V} \tilde{\mathbf{E}} k_w^2(x,y,z)$$

Now from (44) we obtain the bounds

$$\tilde{\mathbf{E}} L_w^2 = O_P\Big(\sum_{w \in W} X_w^5\Big), \qquad \tilde{\mathbf{E}} Q_w^2 = O_P\Big(\sum_{w \in W} X_w^4\Big), \qquad \tilde{\mathbf{E}} K_w^2 = O_P\Big(\sum_{w \in W} X_w^3\Big).$$

Next, we apply Hölder's inequality. For $r = 3, 4, 5$, we have

$$\Big(\sum_{w \in W} 1 \cdot X_w^r\Big)^{6/r} \le \Big(\sum_{w \in W} 1\Big)^{(6-r)/r} \sum_{w \in W} X_w^6 \le m^{(6-r)/r} S_X^2.$$

Finally, from the bound $m = o_P(S_X)$, which holds for $\mathbf{E} X_1^3 = \infty$, see Lemma 1, we obtain

$$\sum_{w \in W} X_w^r \le m^{(6-r)/6} S_X^{r/3} = o_P\big(S_X^{(6-r)/6}\big) S_X^{r/3} = o_P(S_X^2).$$

Proof of (ii). Denote $H_w = \sum_{x \in V} \lambda_{xw}$ and $R_w = H_w^3 - 6\tilde{\mathbf{E}} U_w$. We have

$$\tilde{\mathbf{E}} N = \sum_{w \in W} \tilde{\mathbf{E}} U_w = 6^{-1} \sum_{w \in W} H_w^3 - 6^{-1} \sum_{w \in W} R_w.$$

A straightforward calculation shows that

$$\sum_{w \in W} H_w^3 = \beta_n^{-3/2} \hat{b}_1^3 S_X = (1 + o_P(1)) \beta^{-3/2} b_1^3 S_X.$$

Hence, it remains to prove that $\sum_{w \in W} R_w = o_P(S_X)$. To show this bound we write $R_w = R_{1,w} + R_{2,w}$, where

$$R_{1,w} = H_w^3 - Z_w^3, \qquad R_{2,w} = Z_w^3 - 6\tilde{\mathbf{E}} U_w, \qquad \text{and} \qquad Z_w = \sum_{x \in V} p_{xw},$$

and establish the bounds

$$\sum_{w \in W} R_{1,w} = o_P(S_X) \qquad \text{and} \qquad \sum_{w \in W} R_{2,w} = o_P(S_X). \tag{45}$$

We first prove the second bound of (45). We have

$$0 \le R_{2,w} = \sum_{x \in V} p_{xw}^3 + 3 \sum_{x \in V} \sum_{y \in V \setminus \{x\}} p_{xw}^2 p_{yw} \le \beta_n^{-3/2} X_w^3 \big(n^{-2} \hat{b}_3 + 3n^{-1} \hat{b}_1 \hat{b}_2\big). \tag{46}$$

In the last step we used $p_{xw} \le \lambda_{xw}$. Next, invoking the bounds $n^{-2}\hat{b}_1^3, n^{-1}\hat{b}_2 = o_P(1)$, which hold for $\mathbf{E}Y_1 < \infty$, by Lemma (3), we obtain

$$0 \le \sum_{w \in W} R_{2,w} \le \beta_n^{-3/2} S_X \big(n^{-2}\hat{b}_3 + 3n^{-1}\hat{b}_1\hat{b}_2\big) = o_P(S_X).$$

Let us prove the first bound of (45). We note that $\mathbf{E}X_1^2 < \infty$, $\mathbf{E}Y_1 < \infty$ imply that

$$\mathbf{E}X_1^2 \mathbb{I}_{\{X_1 > \sqrt{m}\}} \to 0, \qquad \mathbf{E}Y_1 \mathbb{I}_{\{Y_1 > \sqrt{n}\}} \to 0. \tag{47}$$

We select a sequence $\delta_n \downarrow 0$ such that $\mathbf{E}Y_1 \mathbb{I}_{\{Y_1 > \sqrt{n}\}} = o(\delta_n)$ and introduce events

$$\mathcal{A} = \Big\{\max_{i \in [m]} X_i \le \sqrt{m}\Big\}, \qquad \mathcal{B} = \Big\{n^{-1} \sum_{j \in [n]} Y_j \mathbb{I}_{\{Y_j > \sqrt{n}\}} \le \delta_n\Big\}.$$

We claim that $\mathbf{P}(\mathcal{A}), \mathbf{P}(\mathcal{B}) = 1 - o(1)$. Indeed, by Markov's inequality and (47)

$$1 - \mathbf{P}(\mathcal{A}) \le \sum_{i \in [m]} \mathbf{P}(X_i > \sqrt{m}) \le \mathbf{E}X_1^2 \mathbb{I}_{\{X_1 > \sqrt{m}\}} \to 0,$$
$$1 - \mathbf{P}(\mathcal{B}) \le (n\delta_n)^{-1} \sum_{j \in [n]} \mathbf{E}Y_j \mathbb{I}_{\{Y_j > \sqrt{n}\}} = \delta_n^{-1} \mathbf{E}Y_1 \mathbb{I}_{\{Y_1 > \sqrt{n}\}} \to 0.$$

Assuming that events $\mathcal{A}$ and $\mathcal{B}$ hold we estimate the difference

$$H_w - Z_w = \sum_{x \in V} (\lambda_{xw} - 1)\mathbb{I}_{\{\lambda_{xw} > 1\}} \le \sum_{x \in V} \lambda_{xw} \mathbb{I}_{\{Y_x > \sqrt{n}\}} \le X_w \beta_n^{-1/2} \delta_n. \tag{48}$$

Here we used the inequality $\mathbb{I}_{\{\lambda_{xw} > 1\}} \le \mathbb{I}_{\{Y_x > \sqrt{n}\}}$, which holds for $X_w \le \sqrt{m}$. Invoking (48) in the inequalities

$$0 \le R_{1,w} = (H_w - Z_w)(Z_w^2 + Z_w H_w + H_w^2) \le 3(H_w - Z_w)H_w^2,$$

and using the identity $H_w^2 = X_w^2 \beta_n^{-1} \hat{b}_1^2$, we obtain

$$\sum_{w \in W} R_{1,w} \le S_X \beta_n^{-3/2} \hat{b}_1^2 \delta_n.$$

For the latter inequality holds with probability $1 - o(1)$ and $\delta_n = o(1)$, we conclude that $\sum_{w \in W} R_{1,w} = o_P(S_X)$.

□

Proof of Lemma 6. Proof of (i). In the proof we make use of Hoeffding's decomposition. Let $\mathbb{I}_j$, $j \in [4]$ be independent Bernoulli random variables with positive success probabilities p_j, $j \in [4]$. Hoeffding's decomposition represents the random variable $T = \mathbb{I}_1\mathbb{I}_2\mathbb{I}_3\mathbb{I}_4 - p_1p_2p_3p_4$ by the sum of uncorrelated U statistics of increasing order

$$T = U_1 + U_2 + U_3 + U_4, \qquad U_1 = \sum_{j \in [4]} T_j, \qquad U_2 = \sum_{\{i,j\} \subset [4]} T_{ij}, \qquad U_3 = \sum_{\{i,j,k\} \subset [4]} T_{ijk}. \tag{49}$$

The first, second, and third order terms T_i, T_{ij}, and T_{ijk} are defined iteratively as follows

$$T_i = \mathbf{E}(T|\mathbb{I}_i), \qquad T_{ij} = \mathbf{E}\big(T - U_1\big|\mathbb{I}_i, \mathbb{I}_j\big), \qquad T_{ijk} = \mathbf{E}\big(T - U_1 - U_2\big|\mathbb{I}_i, \mathbb{I}_j, \mathbb{I}_k\big). \tag{50}$$

Denoting $p = p_1 p_2 p_3 p_4$ and $p_i^* = p/p_i$, $p_{ij}^* = p/(p_i p_j)$, $p_{ijk}^* = p/(p_i p_j p_k)$ we have

$$T_i = (\mathbb{I}_i - p_i)p_i^*, \qquad T_{ij} = (\mathbb{I}_i \mathbb{I}_j - p_i p_j)p_{ij}^* - T_i - T_j,$$
$$T_{ijk} = (\mathbb{I}_i \mathbb{I}_j \mathbb{I}_k - p_i p_j p_k)p_{ijk}^* - T_{ij} - T_{ik} - T_{jk} - T_i - T_j - T_k.$$

The fourth order term $U_4 = T_{1234} := T - U_1 - U_2 - U_3$. We note that various terms of Hoeffding's decomposition are mutually uncorrelated.
Let us prove the lemma. Denote $\Lambda_* = N_2 - \tilde{\mathbf{E}} N_2$ and $T_{(w,\tau)}^{yxz} = \mathbb{I}_{yw}\mathbb{I}_{xw}\mathbb{I}_{x\tau}\mathbb{I}_{z\tau} - \tilde{\mathbf{E}}\mathbb{I}_{yw}\mathbb{I}_{xw}\mathbb{I}_{x\tau}\mathbb{I}_{z\tau}$. We have

$$\Lambda_* = \sum_{x \in V} \sum_{\{y,z\} \subset V \setminus \{x\}} \sum_{w \in W} \sum_{\tau \in W \setminus \{w\}} T_{(w,\tau)}^{yxz}. \tag{51}$$

We decompose every $T_{(w,\tau)}^{yxz}$ using (49) and invoke these decompositions in (51). We then group the first order terms, the second order terms, etc. and obtain Hoeffding's decomposition of Λ_*,

$$\Lambda_* = U_1^* + U_2^* + U_3^* + U_4^*.$$

We specify the linear part U_1^* (the sum of the first order terms), quadratic part U_2^* (the sum of the second order terms), etc. in (52) below. For this purpose we introduce some more notation. Consider the complete bipartite graph $\mathcal{K}_{V,W}$ with the bipartition $V \cup W$. Let $\mathcal{E} = \{(y, w) : y \in V,\, w \in W\}$ denote the set of edges of $\mathcal{K}_{V,W}$. Let $\mathcal{E}^*$ denote the set of paths of length 4 which start from V. After we remove an edge of such a path we obtain a triple of edges, which we call trunk. The set of trunks is denoted $\mathcal{E}^{**}$. For any edge $a = (yw) \in \mathcal{E}$ we denote $\mathbb{I}_a = \mathbb{I}_{yw}$ the indicator of the event that vertex y is linked to the attribute w in the random bipartite graph H. We also denote $p_a = \tilde{\mathbf{E}}\mathbb{I}_a$. Furthermore for distinct edges $a, b, c, d \in \mathcal{E}$ we denote

$$\begin{aligned} t_a &= \mathbb{I}_a - p_a, \qquad t_{ab} = \big(\mathbb{I}_a \mathbb{I}_b - p_a p_b\big) - (\mathbb{I}_a - p_a)p_b - (\mathbb{I}_b - p_b)p_a, \\ t_{abc} &= \big(\mathbb{I}_a \mathbb{I}_b \mathbb{I}_c - p_a p_b p_c\big) - t_{ab}p_c - t_{ac}p_b - t_{bc}p_a - t_a p_b p_c - t_b p_a p_c - t_c p_a p_b. \end{aligned}$$

Finally, t_{abcd} is defined as T_{1234} above, but for $T = \mathbb{I}_a \mathbb{I}_b \mathbb{I}_c \mathbb{I}_d - p_a p_b p_c p_d$.
A calculation shows that

$$\begin{aligned} U_1^* &= \sum_{a \in \mathcal{E}} t_a Q_a, & U_2^* &= \sum_{\{a,b\} \subset \mathcal{E}} t_{ab} Q_{ab}, \\ U_3^* &= \sum_{\{a,b,c\} \in \mathcal{E}^{**}} t_{abc} Q_{abc}, & U_4^* &= \sum_{\{a,b,c,d\} \in \mathcal{E}^*} t_{abcd}, \end{aligned} \tag{52}$$

where coefficients Q_a, Q_{ab} and Q_{abc} are given below. For any $a = (y, w)$ we have

$$\begin{aligned} &Q_a = Q_{a1} + Q_{a2}, \\ &Q_{a1} = \sum_{x \in V \setminus \{y\}} p_{xw} \sum_{\tau \in W \setminus \{w\}} p_{x\tau} \sum_{z \in V \setminus \{x,y\}} p_{z\tau}, \qquad Q_{a2} = \sum_{x \in V \setminus \{y\}} p_{xw} \sum_{\tau \in W \setminus \{w\}} p_{y\tau} \sum_{z \in V \setminus \{x,y\}} p_{z\tau}. \end{aligned} \tag{53}$$

We note that sums Q_{a1} and Q_{a2} represent 4-paths, where y has degree 1 and degree 2 respectively (e.g., paths $y \sim w \sim x \sim \tau \sim z$ and $x \sim w \sim y \sim \tau \sim z$). Furthermore, for a non incident pair $a = (y, w)$ and $c = (x, \tau)$ we have

$$\begin{aligned} &Q_{ac} = Q_{ac1} + Q_{ac2} + Q_{ac3} \\ &Q_{ac1} = p_{y\tau} \sum_{z \in V \setminus \{y,x\}} p_{z\omega}, \qquad Q_{ac2} = p_{xw} \sum_{z \in V \setminus \{y,x\}} p_{z\tau}, \qquad Q_{ac3} = \sum_{z \in V \setminus \{y,x\}} p_{zw} p_{z\tau}, \end{aligned} \tag{54}$$

The sum Q_{ac1} (Q_{ac2}) represents 4-paths, where y (x) has degree 2 (e.g., paths $x \sim \tau \sim y \sim w \sim z$ and $y \sim w \sim x \sim \tau \sim z$). The sum Q_{ac3} represents 4-paths, where y and x has degree 1 (e.g., paths $y \sim w \sim z \sim \tau \sim x$). Similarly, for incident pairs $a = (y, w)$, $b = (x, w)$ and $b = (x, w)$, $c = (x, \tau)$ we have

$$Q_{ab} = \sum_{z \in V \setminus \{x,y\}} \sum_{\tau \in W \setminus \{w\}} \bigl(p_{x\tau} p_{z\tau} + p_{y\tau} p_{z\tau}\bigr) \qquad Q_{bc} = \sum_{\{y,z\} \subset V \setminus \{x\}} \bigl(p_{yw} p_{z\tau} + p_{zw} p_{y\tau}\bigr). \tag{55}$$

Finally, for a trunk $\{a, b, c\}$ which makes up a 3-path, say, $a = (yw)$, $b = (xw)$, $c = (x\tau)$, we have $Q_{a,b,c} = \sum_{z \in V \setminus \{x,y\}} p_{z\tau}$. For a trunk $\{a, b, d\}$ which is not a path (a union of 2-path and an edge), say, $a = (yw)$, $b = (xw)$ and $d = (z\tau)$, we have $Q_{a,b,d} = p_{x\tau}$.
Now we estimate $\tilde{\mathbf{E}}\Lambda_*^2$. From the fundamental property of Hoeffding's decomposition that various terms are uncorrelated we obtain that

$$\tilde{\mathbf{E}}\Lambda_*^2 \;=\; \sum_{a \in \mathcal{E}} Q_a^2 \tilde{E} t_a^2 + \sum_{\{a,b\} \subset \mathcal{E}} Q_{ab}^2 \tilde{E} t_{ab}^2 + \sum_{\{a,b,c\} \in \mathcal{E}^{**}} Q_{abc}^2 \tilde{E} t_{abc}^2 + \sum_{\{a,b,c,d\} \subset \mathcal{E}^*} \tilde{E} t_{abcd}^2.$$

It remains to show that the sums in the right, which we denote by Z_1, Z_2, Z_3, Z_4, are of order $o_P(S_Y^2)$. For this purpose we combine the expressions of $Q_{a\dots c}$ obtained above with the simple inequalities

$$\tilde{\mathbf{E}} t_a^2 \le p_a, \qquad \tilde{\mathbf{E}} t_{ab}^2 \le C p_a p_b, \qquad \tilde{\mathbf{E}} t_{abc}^2 \le C p_a p_b p_c, \qquad \tilde{\mathbf{E}} t_{abcd}^2 \le C p_a p_b p_c p_d.$$

Here C is an absolute constant. We also use the inequalities $p_{xw} \le (nm)^{-1/2} Y_x X_w$.
Proof of the bound $Z_1 = o_P(S_Y^2)$. We have

$$Z_1 \le \sum_{y \in V} \sum_{w \in W} p_{yw} (Q_{(yw)1} + Q_{(yw)2})^2.$$

Invoking the inequalities $(Q_{(yw)1} + Q_{(yw)2})^2 \le 2Q_{(yw)1}^2 + 2Q_{(yw)2}^2$ and

$$\begin{aligned}
Q_{(yw)1} &\le \sum_{x \in V \setminus \{y\}} \sum_{\tau \in W \setminus \{w\}} \sum_{z \in V \setminus \{x,y\}} \frac{Y_x X_w}{\sqrt{nm}} \frac{Y_x X_\tau}{\sqrt{nm}} \frac{Y_z X_\tau}{\sqrt{nm}} \le \frac{X_w}{\sqrt{m}} \frac{S_Y}{\sqrt{n}} \hat{a}_2 \hat{b}_1, \\
Q_{(yw)2} &\le \sum_{x \in V \setminus \{y\}} \sum_{\tau \in W \setminus \{w\}} \sum_{z \in V \setminus \{x,y\}} \frac{Y_x X_w}{\sqrt{nm}} \frac{Y_y X_\tau}{\sqrt{nm}} \frac{Y_z X_\tau}{\sqrt{nm}} \le \frac{X_w}{\sqrt{m}} Y_y \sqrt{n} \hat{a}_2 \hat{b}_1^2,
\end{aligned}$$

we obtain

$$Z_1 \le 2\hat{a}_2^2 \hat{b}_1^3 \frac{S_Y^2}{\sqrt{n}} \frac{S_X}{m^{3/2}} + 2\hat{a}_2^2 \hat{b}_1^4 \frac{S_X}{m^{3/2}} \sqrt{n} \sum_{y \in V} Y_y^3.$$

Note that $\mathbf{E}X_1^2 < \infty$ implies $S_X m^{-3/2} = o_P(1)$. Furthermore, we have $\hat{a}_2, \hat{b}_1 = O_P(1)$. Hence the first summand is $o_P(S_Y^2)$. To show that the second summand is $o_P(S_Y^2)$ we use the fact (which follows from $\mathbf{E}Y_1^2 = \infty$ by Lemma 1) that $n = o_P(S_Y)$ and invoke inequalities

$$\sum_{y \in V} Y_y^3 \le S_Y \max_{y \in V} Y_y \le S_Y (S_Y)^{1/2}. \tag{56}$$

We obtain $\sqrt{n} \sum_{y \in V} Y_y^3 = o_P(\sqrt{S_Y}) S_Y^{3/2} = o_P(S_Y^2)$. We conclude that $Z_1 = o_P(S_Y^2)$.
Proof of the bound $Z_2 = o_P(S_Y^2)$. We split $Z_2 = Z_{21} + Z_{22} + Z_{23}$, where the sum

$$Z_{21} = \sum_{\{x,y\} \subset V} \sum_{w \in W} \sum_{\tau \in W \setminus \{w\}} t_{(yw)(x\tau)}^2 Q_{(yw)(x\tau)}^2$$

accounts for pairs of non incident edges $a=(yw)$ and $c=(x\tau)$, while the sums

$$Z_{22}=\sum_{\{x,y\}\subset V}\sum_{w\in W}t^2_{(yw)(xw)}Q^2_{(yw)(xw)}\qquad\text{and}\qquad Z_{23}=\sum_{x\in V}\sum_{\{w,\tau\}\subset W}t^2_{(xw)(x\tau)}Q^2_{(xw)(x\tau)}$$

account for pairs of incident edges $a=(y,w)$, $b=(x,w)$ and $b=(x,w)$, $c=(x,\tau)$ respectively. To estimate Z_{21} we use (54) and obtain that

$$Q_{(yw)(x\tau)}\le\frac{X_wX_\tau}{m}\Big(Y_y\hat b_1+Y_x\hat b_1+\frac{S_Y}{n}\Big).$$

Hence,

$$Z_{21}\le C\frac{S_X^2}{m^3}\Big(2\hat b_1^3\sum_{x\in V}Y_x^3+\hat b_1^2n^{-1}S_Y^2\Big).$$

From (56) and the fact that $n=o_P(S_Y)$ we conclude that $Z_{21}=o_P(S_Y^2)$. To estimate Z_{22} and Z_{23} we use the first and second identities of (55). We obtain

$$Q_{(yw)(xw)}\le(Y_x+Y_y)\hat a_2\hat b_1\qquad\text{and}\qquad Q_{(xw)(x\tau)}\le 2\hat b_1^2X_wX_\tau nm^{-1}.$$

Hence,

$$Z_{22}\le C\hat a_2^3\hat b_1^3\sum_{x\in V}Y_x^3,\qquad Z_{23}\le C\hat b_1^4S_X^2m^{-3}S_Yn.$$

We note that both quantities on the right are of order $o_P(S_Y^2)$, since $n=o_P(S_Y)$ by Lemma 1 and $S_X^2=o_P(m^3)$. We conclude that $Z_2=o_P(S_Y^2)$.
Proof of the bound $Z_3=o_P(S_Y^2)$. We split $Z_3=Z_{31}+Z_{32}$, where

$$\begin{aligned}Z_{31}&=\sum_{y\in V}\sum_{w\in W}\sum_{x\in V\setminus\{y\}}\sum_{\tau\in W\setminus\{w\}}t^2_{(yw)(xw)(x\tau)}\Big(\sum_{z\in V\setminus\{x,y\}}p_{z\tau}\Big)^2,\\ Z_{32}&=\sum_{\{x,y\}\subset V}\sum_{w\in W}\sum_{\tau\in W\setminus\{w\}}\sum_{z\in V\setminus\{x,y\}}t^2_{(yw)(xw)(z\tau)}\big(p_{y\tau}+p_{x\tau}\big)^2.\end{aligned}$$

We have

$$Z_{31}\le C\hat a_2\hat b_1^3\frac{S_X}{m^{3/2}}\sqrt{n}S_Y,\qquad Z_{32}\le C\hat a_2\hat b_1^2\frac{S_X}{m^{3/2}}n^{-1/2}\sum_{x\in V}Y_x^3.$$

By the same argument as above we obtain that $Z_3=o_P(S_Y^2)$.
Finally, we have

$$Z_4\le\sum_{x\in V}\sum_{y\in V\setminus\{x\}}\sum_{z\in V\setminus\{x,y\}}\sum_{w\in W}\sum_{\tau\in W\setminus\{w\}}t^2_{(yw)(xw)(x\tau)(z\tau)}\le C\hat a_2^2\hat b_1^2S_Y=O_P(S_Y)=o_P(S_Y^2).$$

The proof of the bound $\tilde{\mathbf{E}}\big(N_2-\tilde{\mathbf{E}}N_2\big)^2=o_P\big(S_Y^2\big)$ is completed.

Proof of (ii). Denoting the sum $\sum_\Lambda\sum_{w\in W}\sum_{\tau\in W\setminus\{w\}}$ by $\sum_*$ and using the shorthand notation

$$\begin{aligned}&\mathbb{I}^*=\mathbb{I}_{xw}\mathbb{I}_{yw}\mathbb{I}_{x\tau}\mathbb{I}_{z\tau},\qquad p^*=p_{xw}p_{yw}p_{x\tau}p_{z\tau},\qquad \lambda^*=\lambda_{xw}\lambda_{yw}\lambda_{x\tau}\lambda_{z\tau},\\ &\delta_1^*=(\lambda_{xw}-p_{xw})\lambda_{yw}\lambda_{x\tau}\lambda_{z\tau},\qquad \delta_2^*=p_{xw}(\lambda_{yw}-p_{yw})\lambda_{x\tau}\lambda_{z\tau},\\ &\delta_3^*=p_{xw}p_{yw}(\lambda_{x\tau}-p_{x\tau})\lambda_{z\tau},\qquad \delta_4^*=p_{xw}p_{yw}p_{x\tau}(\lambda_{z\tau}-p_{z\tau}),\end{aligned}$$

we have $N_2=\sum_*\mathbb{I}^*$ and $\tilde{\mathbf{E}}\mathbb{I}^*=p^*$, and $\tilde{\mathbf{E}}N_2=\sum_*p^*$.

We derive (ii) from the relations shown below

$$\sum_* p^* = (1 + o_P(1)) \sum_* \lambda^* \qquad \text{and} \qquad \sum_* \lambda^* = 2^{-1} \hat{a}_2^2 \hat{b}_1^2 S_Y + o_P(S_Y). \tag{57}$$

To prove the second relation we regroup the sum

$$\begin{aligned} \sum_* \lambda^* &= \sum_{\Lambda} Y_x^2 \frac{Y_y Y_z}{n^2} \Big(\hat{a}_2^2 - \sum_{w \in W} \frac{X_w^4}{m^2} \Big) \\ &= \frac{1}{2} \sum_{x \in V} Y_x^2 \left(\Big(\hat{b}_1 - \frac{Y_x}{n} \Big)^2 - \sum_{z \in V \setminus \{x\}} \frac{Y_z^2}{n^2} \right) \Big(\hat{a}_2^2 - \sum_{w \in W} \frac{X_w^4}{m^2} \Big). \end{aligned} \tag{58}$$

For $\mathbf{E} X_1^2 < \infty$ and $\mathbf{E} Y_1 < \infty$, we obtain from Lemma 3 that

$$\sum_{w \in W} \frac{X_w^4}{m^2} = o_P(1), \quad \sum_{z \in V} \frac{Y_z^2}{n^2} = o_P(1), \quad \sum_{x \in V} \frac{Y_x^3}{n} = o_P(S_Y), \quad \sum_{x \in V} \frac{Y_x^4}{n^2} = o_P(S_Y).$$

Invoking these bounds in (58) we obtain the second relation of (57).
To prove the first bound of (57) we write

$$0 \le \sum_* \lambda^* - \sum_* p_* = \sum_* (\delta_1^* + \delta_2^* + \delta_3^* + \delta_4^*) \le \sum_* \lambda^* (\mathbb{I}_1^* + \mathbb{I}_2^* + \mathbb{I}_3^* + \mathbb{I}_4^*),$$
$$\mathbb{I}_1^* := \mathbb{I}_{Y_x X_w > \sqrt{nm}}, \quad \mathbb{I}_2^* := \mathbb{I}_{Y_y X_w > \sqrt{nm}}, \quad \mathbb{I}_3^* := \mathbb{I}_{Y_x X_\tau > \sqrt{nm}}, \quad \mathbb{I}_4^* := \mathbb{I}_{Y_z X_\tau > \sqrt{nm}}$$

and estimate $\sum_* \lambda^* \mathbb{I}_i^* = o_P(S_Y)$, for $i \in [4]$. We only show this bound for $i = 1$. Let $\varepsilon(\cdot)$ be the function associated with the distribution of Y_1 by Lemma 3. So that $\varepsilon(n) = o(1)$ and with probability $1 - o_P(1)$ we have $\max_{x \in V} Y_x \le n\varepsilon(n)$. If the latter inequality holds, then every event $\mathbb{I}_{Y_x X_w > \sqrt{nm}} = 1$ implies $X_w > \beta_n^{1/2} \varepsilon^{-1}(n)$. We denote the indicator of the latter event $\mathbb{I}_w$. We have with probability $1 - o(1)$

$$\sum_* \lambda^* \mathbb{I}_1^* \le \sum_{\Lambda} Y_x^2 \frac{Y_y Y_z}{n^2} \sum_{w \in W} \frac{X_w^2}{m} \mathbb{I}_w \sum_{\tau \in W \setminus \{w\}} \frac{X_\tau^2}{m} \le 2^{-1} \hat{a}_2 \hat{b}_1^2 S_Y \sum_{w \in W} \frac{X_w^2}{m} \mathbb{I}_w.$$

Finally, $\mathbf{E} X_1^2 < \infty$ implies $\mathbf{E} \sum_{w \in W} \frac{X_w^2}{m} \mathbb{I}_w = o(1)$. Hence $\sum_{w \in W} \frac{X_w^2}{m} \mathbb{I}_w = o_P(1)$.

□

References

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

[2] Bloznelis, M., Damarackas, J.: Degree distribution of an inhomogeneous random intersection graph. Electron. J. Comb. 20(3) (2013), #P3.

[3] Bloznelis, M., Godehardt, E., Jaworski, J., Kurauskas, V., Rybarczyk, K.: Recent Progress in Complex Network Analysis - Models of Random Intersection Graphs. In: Lausen, B., Krolak-Schwerdt, S., Böhmer, M. (eds) Data Science, Learning by Latent Structures, and Knowledge Discovery, Springer, Berlin, (2015), 69–78.

[4] Bloznelis, M., Godehardt, E., Jaworski, J., Kurauskas, V., Rybarczyk, K.: Recent Progress in Complex Network Analysis - Properties of Random Intersection Graphs. In: Lausen, B., Krolak-Schwerdt, S., Böhmer, M. (eds) Data Science, Learning by Latent Structures, and Knowledge Discovery, Springer, Berlin, (2015), 79–88.

[5] Bloznelis, M., Kurauskas, K.: Clustering function: another view on clustering coefficient. Journal of Complex Networks 4 (2016), 61–86.

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

[7] Feller, W.: An introduction to probability theory and its applications. Vol. II. 2nd Edition. John Wiley & Sons, New York, (1971).

[8] Foss, S., Korshunov, D., Zachary, S.: An Introduction to Heavy-Tailed and Subexponential Distributions. 2nd Edition. Springer, New York, (2013).

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

[10] Godehardt, E., Jaworski, J., Rybarczyk, K.: Clustering coefficients of random intersection graphs, in: Studies in Classification, Data Analysis and Knowledge Organization, Springer, Berlin, (2012), 243–253.

[11] Jacob, E., Mörters, P.: A spatial preferential attachment model with local clustering, In: Bonato, A., Mitzenmacher, M., Prałat, P. (eds) Algorithms and Models for the Web Graphs, WAW2013, LNCS 8305, Springer, Berlin, (2013), 14–25.

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

[13] Kurauskas, V.: On local weak limit and subgraph counts for sparse random graphs, arXiv:1504.08103v2, (2015).

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

[15] Newman, M. E. J.: The structure of scientific collaboration networks, Proc. Natl.Acad. Sci. USA 98 (2001), 404–409.

[16] Newman, M. E. J.: Scientific collaboration networks. I and II, Physical Review E 64 (2001) 016131, 016132.

[17] Ostroumova Prokhorenkova, L., Samosvat, E.: Global clustering coefficient in scale-free networks, In: Bonato, A., Chung Graham, F., Prałat, P. (eds) Algorithms and Models for the Web Graphs, WAW2014, LNCS 8882, Springer, Berlin, (2014), 47–58.

[18] Redner, S.: How popular is your paper? An empirical study of the citation distribution, Eur.Phys. J. B 4. (1998), 131–134.