Rate of Convergence to the Poisson Law of the Numbers of Cycles in the Generalized Random Graphs

Sergey G. Bobkov, Maria A. Danshina, and Vladimir V. Ulyanov

Abstract Convergence of order $O(1/\sqrt{n})$ is obtained for the distance in total variation between the Poisson distribution and the distribution of the number of fixed size cycles in generalized random graphs with random vertex weights. The weights are assumed to be independent identically distributed random variables which have a power-law distribution. The proof is based on the Chen–Stein approach and on the derived properties of the ratio of the sum of squares of random variables and the sum of these variables. These properties can be applied to other asymptotic problems related to generalized random graphs.

Keywords Generalized random graphs · Poisson law · Rate of convergence · Cycles

S. G. Bobkov
University of Minnesota, Minneapolis, MN, USA

National Research University Higher School of Economics, Moscow, Russia
e-mail: bobkov@math.umn.edu

M. A. Danshina
Moscow Center for Fundamental and Applied Mathematics, Lomonosov Moscow State University, Moscow, Russia
e-mail: danschina.maria@yandex.ru

V. V. Ulyanov (✉)
National Research University Higher School of Economics, Moscow, Russia

Moscow Center for Fundamental and Applied Mathematics, Lomonosov Moscow State University, Moscow, Russia
e-mail: vulyanov@cs.msu.ru

A. N. Karapetyants et al. (eds.), *Operator Theory and Harmonic Analysis*,
Springer Proceedings in Mathematics & Statistics 358,
https://doi.org/10.1007/978-3-030-76829-4_5

1 Introduction

Complex networks attract increasing attention of researchers in various fields of science. In the last years, numerous network models have been proposed. With the uncertainty and the lack of regularity in real-world networks, these models are usually random graphs. Random graphs were first defined by Paul Erdős and Alfred Rényi in their 1959 paper "On Random Graphs" (see [10]) and independently by Gilbert in [12]. The suggested models are closely related: there are n isolated vertices, and every possible edge occurs independently with probability $p : 0 < p < 1$. It is assumed that there are no self-loops. Later, the models were generalized. A natural generalization of the Erdös–Rényi random graph is that the equal edge probabilities are replaced by probabilities depending on the vertex weights. Vertices with high weights are more likely to have more neighbors than vertices with small weights. Vertices with extremely high weights could act as the hubs observed in many real-world networks.

The following generalized random graph (GRG) model was first introduced by Britton et al.; see [5]. Let $V = \{1, 2, .., n\}$ be the set of vertices and $W_i > 0$ be the weight of vertex i, $1 \le i \le n$. The edge probability of the edge between any two vertices i and j, for $i \neq j$, is equal to

$$p_{ij} = \frac{W_i W_j}{L_n + W_i W_j} \tag{1}$$

and $p_{ii} = 0$ for all $i \le n$. Here, $L_n = \sum_{i=1}^{n} W_i$ denotes the total weight of all vertices. The weights W_i, $i = 1, 2, ..., n$ can be taken to be deterministic or random. If we take all $W_i - s$ as the same constant $W_i \equiv n\lambda/(n - \lambda)$ for some $0 \le \lambda < n$, it is easy to see that $p_{ij} = \lambda/n$ for all $1 \le i < j \le n$. That is, the ErdHos–Rényi random graph with $p = \lambda/n$ is a special case of the GRG.

There are many versions of the GRG, such as Poissonian random graph (introduced by Norros and Reittu in [19] and studied by Bhamidi et al. [3]), rank-1 inhomogeneous random graph (see [4]), random graph with given prescribed degrees (see [8]), and Chung–Lu model of heterogeneous random graph (see [7]). The Chung–Lu model is the closest to the model of generalized random graph. Two vertices i and j are connected with probability $p_{ij} = W_i W_j / L_n$ and independently of other pairs of vertices, where $W = (W_1, W_2, ..., W_n)$ is a given sequence. It is necessary to assume that $W_i^2 \le L_n$, for all i. Under some common conditions (see [15]), all of the abovementioned versions of the GRG are asymptotically equivalent, meaning that all events have asymptotically equal probabilities. The updated review on the results about these inhomogeneous random graphs can be seen in Chapter 6 in [21].

One of the problems that arise in real networks of various nature is the spread of the virus. In [6], the authors proposed an approach called nonlinear dynamic system (NLDS) for modeling such processes. Consider a network of n vertices represented by an undirected graph G. Assume an infection rate $\beta > 0$ for each connected edge that is connected to an infected vertex and a recovery rate of $\delta > 0$ for each infected individual. Define the epidemic threshold τ as a value such that

$$\beta/\delta < \tau \Rightarrow \text{infection dies out over time}$$
$$\beta/\delta > \tau \Rightarrow \text{infection survives and becomes an epidemic.}$$

τ is related to the adjacency matrix A of the graph. The matrix $A = [a_{ij}]$ is an $n \times n$ symmetric matrix defined as $a_{ij} = 1$ if vertices i and j are connected by an edge and $a_{ij} = 0$ otherwise. Define a walk of length k in G from vertex v_0 to v_k to be an ordered sequence of vertices $(v_0, v_1, ..., v_k)$, with $v_i \in V$, such that v_i and v_{i+1} are connected for $i = 0, 1, ..., k-1$. If $v_0 = v_k$, then the walk is closed. A closed walk with no repeated vertices (with the exception of the first and last vertices) is called a cycle. For example, triangles, quadrangles, and pentagons are cycles of length three, four, and five, respectively. In the following, the cycle will be denoted by the first k vertices, without specifying the vertex v_k, which is the same as v_0: $(v_0, v_1, ..., v_{k-1})$.

In Theorem 1 in [6], it has been stated that τ is equal to $1/\lambda_1$, where λ_1 is the largest eigenvalue of the adjacency matrix A. The following lower bound for $\lambda_1(A)$ was shown in [20]

$$\lambda_1(A) \geq \frac{6\triangle + \sqrt{36\triangle^2 + 32e^3/n}}{4e},$$

where n, e, and $\triangle$ are the number of vertices, edges, and triangles in G, resp. Moreover, using information about the cycle numbers of higher orders, one can get more precise upper bounds for τ.

In [13], the central limit theorems were proved for the total number of edges in GRG. There are also many results on asymptotic properties of the number of triangles in homogeneous cases. For example, for the ErdHos–Rényi random graph, the upper tails for the distribution of the triangle number had been studied in [2, 9, 14, 16]. Recently, in [18], it was shown for GRG model that asymptotic distribution of the triangle number converges to a Poisson distribution under strong assumption that the vertex weights are bounded random variables.

A lot of real-world networks such as social or computer networks in the Internet (see, e.g. [11]) follow a so-called scale-free graph model; see Chapter 1 in [21]. In Chapter 6, in [21], it was shown that when the vertex weights have approximately a power-law distribution, the GRG model leads to scale-free random graph.

In the present paper, we prove not only the convergence, but we get the convergence rate of order $O(1/\sqrt{n})$ for the distance in total variation between the Poisson distribution and the distribution of the number of fixed size cycles in GRG with random vertex weights. The weights are assumed to be independent identically distributed random variables which have a power-law distribution. The proof is based on the Chen–Stein approach and on the derived properties of the ratio of the sum of squares of random variables and the sum of these variables. These properties can be applied to other asymptotic problems related to GRG.

The main results are formulated in Sect. 2. For their proofs, see Sect. 4. Section 3 contains auxiliary lemmas, some of which are of independent interest.

2 Main Results

Let $\{1, 2, ..., n\}$ be the set of vertices and W_i be a weight of vertex $i : 1 \leq i \leq n$. The probability of the edge between vertices i and j is defined in (1). Let W_i, $i = 1, 2, ..., n$, be independent identically distributed random variables distributed as a random variable W. For $k \geq 3$, denote by $I(k)$ the set of potential cycles of length k. We have that the number of elements in $I(k)$ is equal to $(n)_k/(2k)$, where $(n)_k = n(n-1)...(n-k+1)$ is the number of ways to select k distinct vertices in order, and the factor $1/(2k)$ appears since, for $k > 2$, a permutation of k vertices corresponds to a choice of a cycle in $I(k)$ together with a choice of any of two orientations and k starting points. For example, all six cycles $\{1, 3, 4\}$, $\{3, 4, 1\}$, $\{4, 1, 3\}$, $\{4, 3, 1\}$, $\{1, 4, 3\}$, and $\{3, 1, 4\}$ are, in fact, one cycle of length 3. For $\alpha \in I(k)$, let Y_α be the indicator that α occurs as a cycle in GRG. For example, $\mathbb{P}(Y_{\{1,3,4\}} = 1) = p_{13}p_{34}p_{41}$.

For any integer-valued nonnegative random variables Y and Z, denote the total variation distance between their distributions $\mathrm{L}(Y)$ and $\mathrm{L}(Z)$ by

$$\| \mathrm{L}(Y) - \mathrm{L}(Z) \| \equiv sup_{\|h\|=1} |\mathbb{E}h(Y) - \mathbb{E}h(Z)|, \tag{2}$$

where h is any real function defined on $\{0, 1, 2, ...\}$ and $\| h \| \equiv \sup_{m\geq 0} |h(m)|$.

For $k \geq 3$, put $S_n(k) = \sum_{\alpha \in I(k)} Y_\alpha$, that is, $S_n(k)$ is the number of cycles of length k. Let Z_k be a random variable having Poisson distribution with parameter $\lambda(k) = (\mathbb{E}W^2/\mathbb{E}W)^k/(2k)$.

Theorem 1 *For any $k \geq 3$, one has*

$$\| \mathrm{L}(S_n(k)) - \mathrm{L}(Z_k) \| = O(n^{-1/2}), \tag{3}$$

provided that

$$\mathbb{P}(W > x) = o(x^{-2k-1}), \ \text{ as } \ x \to +\infty. \tag{4}$$

Remark 1 Relation (3) holds under condition that W has power-law distribution. The condition on the tail behavior of the distribution of W can be replaced by stronger moment condition: the finiteness of expectation $\mathbb{E}W^{2k+1}$.

Remark 2 Recently in [18], the convergence in distribution of the number of triangles $S_n(3)$ in a generalized random graphs to the Poisson random variable Z_3 was proved by method of moments under assumption that the vertex weights W_i-s are bounded random variables. In Theorem 1, we have used the Chen–Stein approach; see, e.g., [1] and [2]. This allows us not only to extend the

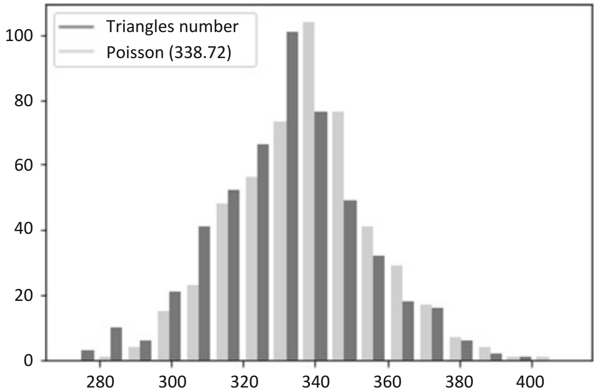

Fig. 1 Histogram of the number of triangles in GRG with 2000 vertices. The distribution of vertex weights $W_i \sim Uni(10, 15)$, forall $i \leq 2000$. The number of realizations is 500

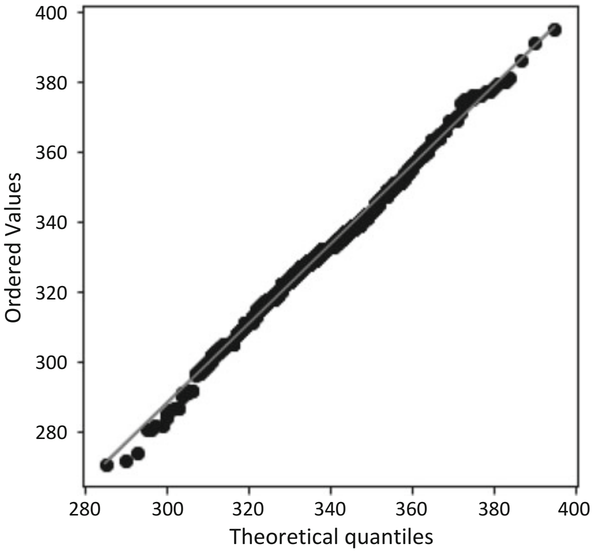

Fig. 2 Q–Q plot for the number of triangles in GRG with 2000 vertices and the Poisson variable $Pois(338.72)$. $W_i \sim Uni(10, 15)$, forall $i \leq 2000$. The number of realizations is 500

convergence result to cycles of any fixed length k but also to get the rate of convergence. Moreover, we replace the assumption about the boundness of W_i-s with the condition that W_i has a power-law distribution. As we noted in Introduction, this condition better matches real-world networks.

Figures 1 and 2 illustrate the results of Theorem 1, with the example of the number of triangles distribution.

The next results are not directly connected with number of cycles in GRG. They are an important part of the proof of Theorem 1. At the same time, the results are of independent interest. They describe the asymptotic properties of ratio of a sum of the squares of n i.i.d. random variables and a sum of these random variables. These properties can be applied to other asymptotic problems related to GRG.

Given i.i.d. positive random variables $X, X_1, \dots, X_n$, define the statistics

$$T_n = \frac{X_1^2 + \cdots + X_n^2}{X_1 + \cdots + X_n}.$$

Assume that X has a finite second moment, so that, by the law of large numbers, with probability one

$$\lim_{n \to \infty} T_n^p = \left(\frac{\mathbb{E}X^2}{\mathbb{E}X}\right)^p$$

for any $p \geq 1$. Here, we describe the tail-type and moment-type conditions which ensure that this convergence also holds on average.

Theorem 2 *Given an integer $p \geq 2$, the convergence*

$$\lim_{n \to \infty} \mathbb{E}T_n^p = (\mathbb{E}X^2/\mathbb{E}X)^p \tag{5}$$

is equivalent to the tail condition

$$\mathbb{P}\{X \geq x\} = o(x^{-p-1}) \text{ as } x \to \infty. \tag{6}$$

Moreover, if $\mathbb{P}\{X \geq x\} = O(x^{-p-3/2})$ as $x \to \infty$, then

$$\mathbb{E}T_n^p - (\mathbb{E}X^2/\mathbb{E}X)^p = O(n^{-1/2}) \tag{7}$$

The finiteness of the moment $\mathbb{E}X^{p+1}$ is sufficient for (5) to hold, while the finiteness of the moments $\mathbb{E}X^q$ is necessary for any real value $1 \leq q < p + 1$.

Let $M_n = \max_{1 \leq i \leq n} X_i$. For $p \geq 2$, define

$$R_n^{(p)} = T_n^p M_n^2/(X_1 + X_2 + \dots + X_n). \tag{8}$$

By the law of large numbers, $R_n^{(p)} \to 0$ as $n \to \infty$ a.s., under mild moment assumptions. The next theorem gives the order of convergence of $\mathbb{E}R_n^{(p)}$ to zero under tail-type and moment-type conditions.

Theorem 3 *Given an integer $p \geq 2$, if $\mathbb{P}(X \geq x) = O(x^{-p-7/2})$ as $x \to +\infty$, then*

$$\mathbb{E}R_n^{(p)} = O(n^{-1/2}). \tag{9}$$

When $p > 8$ *and* $\mathbb{E}X^{p+4}$ *is finite, the rate can be improved to*

$$\mathbb{E}R_n^{(p)} = O(n^{-(p-2)/(p+4)}). \tag{10}$$

Moreover, if $\mathbb{E}\, e^{\varepsilon X} < \infty$ *for some* $\varepsilon > 0$*, then*

$$\mathbb{E}R_n^{(p)} = O\Big(\frac{(\log n)^2}{n}\Big). \tag{11}$$

3 Auxiliary Lemmas

Lemma 1 *Let* $S_n = \eta_1 + \cdots + \eta_n$ *be the sum of independent random variables* $\eta_k \geq 0$ *with finite second moment, such that* $\mathbb{E}S_n = n$ *and* $\mathrm{Var}(S_n) = \sigma^2 n$*. Then, for any* $0 < \lambda < 1$*, one has*

$$\mathbb{P}\{S_n \leq \lambda n\} \leq \exp\bigg\{ - \frac{(1-\lambda)^2}{2\left[\sigma^2 + \max_k\, (\mathbb{E}\eta_k)^2\right]}\, n \bigg\}. \tag{12}$$

Proof We use here the standard arguments. Fix a parameter $t > 0$. We have

$$\mathbb{E}\, e^{-tS_n} \geq e^{-\lambda t n}\, \mathbb{P}\{S_n \leq \lambda n\}.$$

Every function $u_k(t) = \mathbb{E}\, e^{-t\xi_k}$ is positive and convex and admits Taylor's expansion near zero up to the quadratic form, which implies that

$$u_k(t) \leq 1 - t\, \mathbb{E}\xi_k + \frac{t^2}{2}\, \mathbb{E}\xi_k^2 \leq \exp\Big\{ - t\, \mathbb{E}\xi_k + \frac{t^2}{2}\, \mathbb{E}\xi_k^2 \Big\}.$$

Multiplying these inequalities, we get

$$\mathbb{E}\, e^{-tS_n} \leq \exp\Big\{ - tn + \frac{bt^2}{2} \Big\}, \quad b = \sum_{k=1}^n \mathbb{E}\xi_k^2.$$

The two bounds yield

$$\mathbb{P}\{S_n \leq \lambda n\} \ \leq\ \exp\Big\{ - (1-\lambda)nt + bt^2/2 \Big\},$$

and after optimization over t (in fact, $t = \frac{1-\lambda}{b}\, n$), we arrive at the exponential bound

$$\mathbb{P}\{S_n \leq \lambda n\} \leq \exp\Big\{ - \frac{(1-\lambda)^2}{2b}\, n^2 \Big\}.$$

Note that

$$b = \mathrm{Var}(S_n) + \sum_{k=1}^{n} (\mathbb{E}\xi_k)^2 \le \Big(\sigma^2 + \max_k\, (\mathbb{E}\xi_k)^2\Big)\, n,$$

and (12) follows. □

For further lemmas, we need additional notation.

Denote by $F(x) = \mathbb{P}\{X \le x\}$ ($x \in \mathbb{R}$) the distribution function of the random variable X, and put

$$\varepsilon_q(x) = x^q\,(1 - F(x)), \quad x \ge 0,\ q > 0.$$

Raising the sum $U_n = X_1^2 + \dots + X_n^2$ to the power p with $n \ge 2p$, we have

$$U_n^p \,=\, \sum X_{i_1}^2 \dots X_{i_p}^2, \tag{13}$$

where the summation is performed over all collections of numbers $i_1, \dots, i_p \in \{1, \dots, n\}$. For $r = 1, \dots, p$, we denoted by $\mathrm{C}(p, r)$ the collection of all tuples $\gamma = (\gamma_1, \dots, \gamma_r)$ of positive integers such that $\gamma_1 + \dots + \gamma_r = p$. For any $\gamma \in \mathrm{C}(p, r)$, there are $n(n-1)\dots(n-r+1)$ sequences $X_{i_1}, \dots, X_{i_p}$ with r distinct terms that are repeated $\gamma_1, \dots, \gamma_r$ times, resp. Therefore, by the i.i.d. assumption,

$$\mathbb{E}T_n^p \,=\, \sum_{r=1}^{p} \frac{n(n-1)\dots(n-r+1)}{n^p} \sum_{\gamma \in \mathrm{C}(p,r)} \mathbb{E}\xi_n(\gamma), \tag{14}$$

where

$$\xi_n(\gamma) = X_1^{2\gamma_1} \dots X_r^{2\gamma_r} / (\frac{1}{n}\, S_r + \frac{1}{n}\, S_{n,r})^p$$

and

$$S_r = X_1 + \dots + X_r, \quad S_{n,r} = X_{r+1} + \dots + X_n.$$

In the following lemmas, without loss of generality, let $\mathbb{E}X = 1$.

Lemma 2 *For the boundedness of the sequence $\mathbb{E}T_n^p$, it is necessary that the moment $\mathbb{E}X^p$ be finite. Moreover, for the particular collection $\gamma = (p)$ with $r = 1$, we have*

$$\mathbb{E}\xi_n(\gamma) \,\ge\, 2^{-p}\, n^p\, \mathbb{E}X^p\, 1_{\{X \ge n\}}. \tag{15}$$

Proof Since $\xi_n(\gamma) = X_1^{2p}/(\frac{1}{n}\,X_1 + \frac{1}{n}\,S_{n,1})^p$, applying Jensen's inequality, we get

$$\begin{aligned}\mathbb{E}\xi_n(\gamma) &\geq \mathbb{E}_{X_1}\,\frac{X_1^{2p}}{(\frac{1}{n}\,X_1 + \frac{1}{n}\,\mathbb{E}_{S_{n,1}}\,S_{n,1})^p}\\ &= \mathbb{E}\,\frac{X^{2p}}{(\frac{1}{n}\,X + \frac{n-1}{n})^p}\ \geq\ 2^{-p}\,n^p\,\mathbb{E}X^p\,1_{\{X\geq n\}}.\end{aligned}$$

□

In the sequel, we use the events

$$A_{n,r} = \Big\{S_{n,r} \leq \frac{n-r}{2}\Big\} \quad \text{and} \quad B_{n,r} = \Big\{S_{n,r} > \frac{n-r}{2}\Big\}. \tag{16}$$

By Lemma 1, whenever $n \geq 2p$, for some constant $c > 0$ independent of n,

$$\mathbb{P}(A_{n,r}) \leq e^{-c(n-r)} \leq e^{-cn/2}. \tag{17}$$

Lemma 3 *If $\mathbb{E}X^p$ is finite, then $\mathbb{E}\xi_n \to (\mathbb{E}X^2)^p$ as $n \to \infty$, where*

$$\xi_n\ =\ X_1^2 \ldots X_p^2/(\frac{1}{n}\,S_p + \frac{1}{n}\,S_{n,p})^p. \tag{18}$$

Proof Using $X_1 \ldots X_p \leq S_p^p$, we have $\xi_n\ \leq\ S_p^{2p}/(\frac{1}{n}\,S_n)^p\ \leq\ n^p\,S_p^p$. Hence,

$$\mathbb{E}\,\xi_n\,1_{A_{n,p}}\ \leq\ n^p\,\mathbb{E}\,S_p^p\,\mathbb{P}(A_{n,p})\ =\ o(e^{-cn}) \tag{19}$$

for some constant $c > 0$ independent of n. Here, we applied (17) with $r = p$ and Lemma 2 which ensures that $\mathbb{E}\,S_p^p < \infty$. Further, $\xi_n\,1_{B_{n,p}}\ \leq\ 2^p X_1^2 \ldots X_p^2$.Hence, the random variables $\xi_n\,1_{B_{n,p}}$ have an integrable majorant. Since also $\xi_n \to X_1^2 \ldots X_p^2$ (the law of large numbers) and $1_{B_{n,p}} \to 1$ a.s. (implied by (17)), one may apply the Lebesgue dominated convergence theorem, which gives $\mathbb{E}\xi_n 1_{B_n} \to (\mathbb{E}X^2)^p$. Together with (19), we get $\mathbb{E}\xi_n \to (\mathbb{E}X^2)^p$. □

Lemma 4 *If the moment $\mathbb{E}X^p$ is finite, then for any $\gamma = (\gamma_1, \ldots, \gamma_r) \in \mathrm{C}(p, r)$,*

$$\mathbb{E}\,\xi_n(\gamma)\ =\ 4^p\,\mathbb{E}\,\frac{X_1^{2\gamma_1} \ldots X_r^{2\gamma_r}}{(\frac{1}{n}\,S_r + 1)^p} + o(1).$$

Proof Using an elementary bound $X_1^{2\gamma_1} \dots X_r^{2\gamma_r} \le (X_1 + \dots + X_r)^{2\gamma_1+\dots+2\gamma_r} = S_r^{2p}$ and applying Jensen's inequality, we see that $\xi_n(\gamma) \le n^p\, S_r^p \le n^p\, r^{p-1}\,(X_1^p + \dots + X_r^p)$. Hence,

$$\mathbb{E}\xi_n(\gamma)\,1_{A_{n,r}} \le n^p\, r^{p-1} \sum_{k=1}^{r} \mathbb{E}X_k^p\, 1_{A_{n,r}} = n^p\, r^p\, \mathbb{E}X^p\, \mathbb{P}(A_{n,r}) \;=\; o(e^{-c'n}). \tag{20}$$

On the other hand, on the set $B_{n,r}$, there is a point-wise bound

$$\xi_n(\gamma)\,1_{B_{n,r}} \le \frac{X_1^{2\gamma_1} \dots X_r^{2\gamma_r}}{(\frac{1}{n}\,S_r + \frac{n-r}{2n})^p} \le 4^p\, \frac{X_1^{2\gamma_1} \dots X_r^{2\gamma_r}}{(\frac{1}{n}\,S_r + 1)^p}. \tag{21}$$

□

Our task is reduced to the estimation of the expectation on the right-hand side of (21). Let us first consider the shortest collection $\gamma = (p)$ of length $r = 1$.

Lemma 5 *Under the condition* (6)*,*

$$\mathbb{E}\, \frac{X_1^{2p}}{(\frac{1}{n}\,X_1 + 1)^p} = o(n^{p-1}). \tag{22}$$

In addition, if $\mathbb{P}\{X \ge x\} = O(x^{-q})$ *for some real value* q *in the interval* $p < q < 2p$*, then*

$$\mathbb{E}\, \frac{X_1^{2p}}{(\frac{1}{n}\,X_1 + 1)^p} = O(n^{2p-q}). \tag{23}$$

Proof We have

$$\begin{aligned}\mathbb{E}\, \frac{X_1^{2p}}{(\frac{1}{n}\,X_1 + 1)^p} &= \mathbb{E}\, \frac{X_1^{2p}}{(\frac{1}{n}\,X_1 + 1)^p}\,1_{\{X_1 \ge n\}} + \mathbb{E}\, \frac{X_1^{2p}}{(\frac{1}{n}\,X_1 + 1)^p}\,1_{\{X_1 < n\}} \\ &\le n^p\, \mathbb{E}\, X^p\, 1_{\{X \ge n\}} + \mathbb{E}\, X^{2p}\, 1_{\{X < n\}}.\end{aligned}$$

In view of (6), to derive (22), it remains to be bound to the last expectation by $o(n^{p-1})$. Integrating by parts and assuming that $x = n$ is the point of continuity of $F(x)$, we have, using $\varepsilon_{p+1}(x) \to 0$ as $x \to \infty$,

$$\begin{aligned}\mathbb{E}\, X^{2p}\, 1_{\{X < n\}} &= -n^{2p}\,(1 - F(n)) + 2p \int_0^n x^{2p-1}\,(1 - F(x))\,dx \\ &\le 2p \int_0^n x^{p-2}\, \varepsilon_{p+1}(x)\,dx \;=\; o(n^{p-1}),\end{aligned} \tag{24}$$

For the second assertion (23), we similarly have

$$2p\int_0^n x^{2p-1-q}\,\varepsilon_q(x)\,dx \;=\; O(n^{2p-q}),$$

$$\mathbb{E}\,X^p\,1_{\{X\ge n\}} = O(n^{p-q}) + p\int_n^\infty x^{p-q-1}\,\varepsilon_q(x)\,dx \;=\; O(n^{p-q}).$$

□

Lemma 6 *Let* $\gamma = (\gamma_1, \dots, \gamma_r) \in \mathrm{C}(p, r)$, $2 \le r \le p-1$. *Under* (6)*, we have*

$$\mathbb{E}\,\frac{X_1^{2\gamma_1}\dots X_r^{2\gamma_r}}{(\frac{1}{n}\,S_r+1)^p} = o(n^{p-r-1}\log n). \tag{25}$$

Proof If all $\gamma_i \le p/2$, there is nothing to prove, since then

$$\mathbb{E}\,\frac{X_1^{2\gamma_1}\dots X_r^{2\gamma_r}}{(\frac{1}{n}\,S_r+1)^p} \;\le\; \mathbb{E}\,X_1^{2\gamma_1}\dots\mathbb{E}\,X_r^{2\gamma_r} \;\le\; (\mathbb{E}X^p)^r.$$

In the other case, suppose for definiteness that γ_1 is the largest number among all γ_is. Necessarily, $\gamma_1 > p/2$ and $\gamma_i < p/2$ for all $i \ge 2$. Since $S_r < n$ implies $X_1 < n$, we similarly have

$$\mathbb{E}\,\frac{X_1^{2\gamma_1}\dots X_r^{2\gamma_r}}{(\frac{1}{n}\,S_r+1)^p}\,1_{\{S_r<n\}} \;\le\; (\mathbb{E}X^p)^{r-1}\,\mathbb{E}\,X^{2\gamma_1}\,1_{\{X<n\}}.$$

To bound the last expectation, note that $r-1 \le \gamma_2+\dots+\gamma_r < p/2$, so that $p \ge 2r-1$. Hence, if $x = n$ is the point of continuity of $F(x)$, similarly to (24), we get

$$\mathbb{E}\,X^{2\gamma_1}\,1_{\{X<n\}} \le 2\gamma_1\int_0^n x^{2\gamma_1-p-2}\,\varepsilon_{p+1}(x)\,dx. \tag{26}$$

But since $\gamma_1 \le p-r+1$,

$$\int_1^n x^{2\gamma_1-p-2}\,\varepsilon_{p+1}(x)\,dx \;\le\; \int_1^n x^{p-2r}\,\varepsilon_{p+1}(x)\,dx \;=\; o(n^{p-2r+1}), \tag{27}$$

if $p \ge 2r$ or $p \le 2r-2$, which is even stronger than the rate $o(n^{p-r-1})$. In the remaining case $p = 2r-1$, the last integral is $o(\log n)$. This proves (25) for the part of the expectation restricted to the set $S_r < n$, that is,

$$\mathbb{E}\,\frac{X_1^{2\gamma_1}\dots X_r^{2\gamma_r}}{(\frac{1}{n}\,S_r+1)^p}\,1_{\{S_r<n\}} \;=\; o(n^{p-r-1}\log n). \tag{28}$$

Note that the logarithmic term cannot be removed in the special situation where $p = 3$, $r = 2$, $\gamma_1 = 2$, and $\gamma_2 = 1$, in which case the last integral in (27) becomes $\int_1^n x^{-1}\,\varepsilon_4(x)\,dx$.

Turning to the expectation over the complementary set $S_r \geq n$, introduce the events

$$\Omega_i = \Big\{X_i \geq \max_{j \neq i} X_j\Big\}, \quad i = 1, \dots, r.$$

On every such set, $X_i \leq S_r \leq rX_i$. In particular, $S_r \geq n$ implies $X_i \geq n/r$. Hence, together with (28), (25) would follow from the stronger assertion

$$\mathbb{E}\, \frac{X_1^{2\gamma_1} \dots X_r^{2\gamma_r}}{X_i^p}\, 1_{\{X_i \geq n\} \cap \Omega_i} \, = \, o(n^{-r-1}) \tag{29}$$

with an arbitrary index $1 \leq i \leq r$.

Case 1 $i \geq 2$. If we fix any values $X_1 = x_1$ and $X_i = x_i$, then the expectation with respect to X_j, $j \neq i$, in (29) will yield a bounded quantity (since the p-moment is finite). Hence, (29) is simplified to

$$\mathbb{E}\, X_1^{2\gamma_1} X_i^{2\gamma_i - p}\, 1_{\{X_i \geq n\} \cap \{X_i \geq X_1\}} \, = \, o(n^{-r-1}). \tag{30}$$

Here, the expectation over X_1 may be carried out and estimated similarly to (26), by replacing n with x_i. Namely,

$$\mathbb{E}\, X_1^{2\gamma_1} 1_{\{X_1 \leq x_i\}} \, \leq \, 2\gamma_1 \int_0^{x_i} x^{2\gamma_1 - p - 2}\, \varepsilon_{p+1}(x)\, dx \, = \, \delta(x_i)\, x_i^{2\gamma_1 - p}$$

with some $\delta(x_i) \to 0$ as $x_i \to \infty$ (this assertion may be strengthened when $2\gamma_1 - p = 1$). Hence, the expectation in (30) is bounded by

$$\begin{aligned}
\mathbb{E}\, X_i^{2\gamma_i + 2\gamma_1 - 2p}\, \delta(X_i)\, 1_{\{X_i \geq n\}} &\leq \delta_n\, \mathbb{E}\, X_i^{2\gamma_i + 2\gamma_1 - 2p}\, 1_{\{X_i \geq n\}} \\
&= \, \delta_n\, n^{2\gamma_i + 2\gamma_1 - 2p}\, (1 - F(n)) + c_i \delta_n \int_n^\infty x^{2\gamma_i + 2\gamma_1 - 2p - 1}\, (1 - F(x))\, dx \\
&= \, o(n^{2\gamma_i + 2\gamma_1 - 3p - 1}) + c_i \delta_n \int_n^\infty x^{2\gamma_i + 2\gamma_1 - 3p - 2}\, \varepsilon_{p+1}(x)\, dx \\
&= \, o(n^{2\gamma_i + 2\gamma_1 - 3p - 1}),
\end{aligned}$$

where $\delta_n = \sup_{x \geq n} \delta(x) \to 0$. To obtain (30), it remains to check that $2\gamma_i + 2\gamma_1 - 3p - 1 \leq -r - 1$. And indeed, since $p = \gamma_i + \gamma_1 + \sum_{j \neq i,1} \gamma_j \geq \gamma_i + \gamma_1 + (r-2)$, the desired relation would follow from $2\,(p - (r-2)) - 3p - 1 \leq -r - 1$, that is, $p + r \geq 4$ (which is true).

Case 2 $i = 1$. If we fix any value $X_1 = x_1$, the expectation with respect to X_j, $j \neq 1$, will yield a bounded quantity (since the p-moment is finite). Hence, (29) is simplified to

$$\mathbb{E}\, X^{2\gamma_1 - p}\, 1_{\{X \geq n\}} = o(n^{-r-1}).$$

Here, the expectation may be estimated similarly. Namely,

$$\begin{aligned}\mathbb{E}\, X^{2\gamma_1 - p}\, 1_{\{X \geq n\}} &= \int_n^\infty x^{2\gamma_1 - p}\, dF(x) \\ &= o(n^{2\gamma_1 - 2p - 1}) + \int_n^\infty x^{2\gamma_1 - 2p - 2}\, \varepsilon_{p+1}(x)\, dx \;=\; o(n^{2\gamma_1 - 2p - 1}).\end{aligned}$$

It remains to be seen that $2\gamma_1 - 2p - 1 \leq -r - 1$. Again, since $\gamma_1 \leq p - (r-1)$, the latter would follow from $2(p - r + 1) - 2p - 1 \leq -r - 1$, which is the same as $r \geq 2$. $\square$

We now consider the lemmas which enable us to get a bound for $\mathbb{E} R_n^{(p)}$; see (8). Without loss of generality, let $\mathbb{E} X = 1$ and $n \geq 2p$.

Introduce additional notation: $M_{n,r} = \max_{r < i \leq n} X_i and\ \ (1 \leq r \leq p)$.

Recall that there is the representation (13) but instead of (14) we write now

$$\mathbb{E} R_n^{(p)} \;=\; \sum_{r=1}^p \frac{n(n-1)\dots(n-r+1)}{n^{p+1}} \sum_{\gamma \in \mathrm{C}(p,r)} \mathbb{E}\, \eta_n(\gamma) M_n^2, \tag{31}$$

where

$$\eta_n(\gamma) = \frac{X_1^{2\gamma_1} \dots X_r^{2\gamma_r}}{(\frac{1}{n}\, S_r + \frac{1}{n}\, S_{n,r})^{p+1}}.$$

In order to bound the expectations on the right-hand side of (31), we use again the events $A_{n,r}$ and $B_{n,r}$; see (16). From elementary inequalities $M_n \leq S_n$ and

$$X_1^{2\gamma_1} \dots X_r^{2\gamma_r} \leq (X_1 + \dots + X_r)^{2\gamma_1 + \dots + 2\gamma_r} \leq S_n^{2p},$$

it follows that $\eta_n(\gamma) M_n^2 \;\leq\; n^{p+1}\, S_n^{p+1} \;\leq\; 2^p n^{p+1}\, (S_r^{p+1} + S_{n,r}^{p+1})$, implying

$$\mathbb{E}\, \eta_n(\gamma)\, M_n^2\, 1_{A_{n,r}} \leq 2^p n^{p+1} \Big(\mathbb{E}\, S_r^{p+1}\, \mathbb{P}(A_{n,r}) + \mathbb{E} S_r^{p+1}\, \mathbb{E}\, S_{n,r}^{p+1}\, 1_{A_{n,r}} \Big). \tag{32}$$

Here, by Lemma 1 with $\lambda = 1/2$ and using $n - r \geq \frac{1}{2}\, n$, we have

$$\mathbb{P}(A_{n,r}) \leq \exp\Big\{ -\frac{1}{16\, b^2}\, n \Big\}, \qquad b^2 = \mathbb{E} X^2. \tag{33}$$

Also, assuming that the moment $\mathbb{E}X^{p+2}$ is finite and applying the Hölder inequality with exponents $(p+2)/(p+1)$ and $p+2$, one may bound the last expectation in (32) as

$$\mathbb{E}\, S_{n,r}^{p+1}\, 1_{A_{n,r}} \leq \big(\mathbb{E}\, S_{n,r}^{p+2}\big)^{\frac{p+1}{p+2}}\, (\mathbb{P}(A_{n,r}))^{\frac{1}{p+2}}.$$

By Jensen inequality, $\mathbb{E}\, S_{n,r}^{p+2} \leq r^{p+1}\, \mathbb{E}X^{p+2}$. Applying this in (32), the inequality (33) yields an exponential bound

$$\mathbb{E}\, \eta_n(\gamma)\, M_n^2\, 1_{A_{n,r}} \leq e^{-cn} \tag{34}$$

with some constant $c > 0$ which does not depend on n.

As for the set $B_{n,r}$, we use on it a point-wise upper bound $\eta_n(\gamma) \leq 2^{p+1}\, X_1^{2\gamma_1} \dots X_r^{2\gamma_r}/(\frac{1}{n}\, S_r + 1)^{p+1}$. One may also use $M_n \leq M_r + M_{n,r} \leq S_r + M_{n,r}$, implying, by Jensen's inequality, $M_n^2 \leq (r+1)\,(X_1^2 + \dots + X_r^2 + M_{n,r}^2)$. It gives

$$\begin{aligned}\mathbb{E}\, \eta_n(\gamma)\, M_n^2\, 1_{B_{n,r}} \leq\ & 2^{p+1}(r+1) \sum_{k=1}^{r} \mathbb{E}\, \frac{X_1^{2\gamma_1} \dots X_r^{2\gamma_r}}{(\frac{1}{n}\, S_r + 1)^{p+1}}\, X_k^2 \\ & + 2^{p+1}(r+1) \sum_{k=1}^{r} \mathbb{E}\, \frac{X_1^{2\gamma_1} \dots X_r^{2\gamma_r}}{(\frac{1}{n}\, S_r + 1)^{p+1}}\, \mathbb{E}M_{n,r}^2\end{aligned} \tag{35}$$

Without an essential loss, the last expectation $\mathbb{E}M_{n,r}^2$ may be replaced with $\mathbb{E}M_n^2$. The second last expectation was considered in Lemmas 5–6 under the condition (6), which holds as long as the moment $\mathbb{E}X^{p+1}$ is finite. The third last expectation in (35), due to an additional factor X_k^2, dominates the second last and needs further consideration under stronger moment assumptions. Recalling (34) and returning to (31), let us summarize in the following statement.

Lemma 7 *If the moment $\mathbb{E}X^{p+2}$ is finite, then*

$$\begin{aligned}c\, \mathbb{E}R_n^{(p)} \leq e^{-cn} &+ \max_{1 \leq k \leq r}\ \max_{\gamma \in \mathrm{C}(p,r)}\ \frac{1}{n^{p-r+1}}\, \mathbb{E}\, \frac{X_1^{2\gamma_1} \dots X_r^{2\gamma_r}}{(\frac{1}{n}\, S_r + 1)^{p+1}}\, X_k^2 \\ &+ \max_{\gamma \in \mathrm{C}(p,r)}\ \frac{1}{n^{p-r+1}}\, \mathbb{E}\, \frac{X_1^{2\gamma_1} \dots X_r^{2\gamma_r}}{(\frac{1}{n}\, S_r + 1)^{p+1}}\, \mathbb{E}M_n^2\end{aligned} \tag{36}$$

with some constant $c > 0$ which does not depend on $n \geq 2p$.

In order to obtain polynomial bounds for the expectations in (36) under suitable moment or tail assumptions, we need to develop corresponding analogs of Lemmas 5–6. We will consider separately the cases $r = 1$, $r = p$, and $2 \leq r \leq p-1$

under the tail condition

$$\mathbb{P}\{X \geq x\} = O(1/x^{p+\alpha}) \quad \text{as } x \to \infty, \tag{37}$$

where $\alpha > 0$ is a parameter. It implies that the moments $\mathbb{E}X^q$ are finite for all $q < p + \alpha$ and is fulfilled as long as the moment $\mathbb{E}X^{p+\alpha}$ is finite. Put $\varepsilon(x) = x^{p+\alpha}\,(1 - F(x))$, where F denotes the distribution function of the random variable X.

Lemma 8 *Under* (37) *with* $1 < \alpha \leq p + 2$*,*

$$\mathbb{E}\,\frac{X_1^{2p}}{(\frac{1}{n}\,X_1 + 1)^{p+1}} \;=\; O(n^{p-\alpha+2}). \tag{38}$$

Moreover, for any index $1 \leq k \leq r$*,*

$$\mathbb{E}\,\frac{X_1^{2p}}{(\frac{1}{n}\,X_1 + 1)^{p+1}}\;X_k^2 \;=\; O(n^{p-\alpha+2}\log n). \tag{39}$$

Proof The expectation in (38) is equal to and satisfies

$$\mathbb{E}\,\frac{X_1^{2p}}{(\frac{1}{n}\,X_1 + 1)^{p+1}}\,1_{\{X_1 \geq n\}} + \mathbb{E}\,\frac{X_1^{2p}}{(\frac{1}{n}\,X_1 + 1)^{p+1}}\,1_{\{X_1 < n\}}$$
$$\leq\; n^{p+1}\,\mathbb{E}X^{p-1}\,1_{\{X \geq n\}} + \mathbb{E}\,X^{2p}\,1_{\{X<n\}}$$

Similarly to (24), we get

$$\mathbb{E}\,X^{2p}\,1_{\{X<n\}} \leq 2p\int_0^n x^{p-\alpha-1}\,\varepsilon(x)\,dx \;=\; O(n^{p-\alpha}),$$

provided that $\alpha < p$. In the case $\alpha = p$, the last integral is bounded by $O(\log n)$. In addition,

$$\mathbb{E}\,X^p\,1_{\{X \geq n\}} = O(n^{p-\alpha}) + p\int_n^\infty x^{p-\alpha-1}\,\varepsilon_q(x)\,dx \;=\; O(n^{p-\alpha}).$$

This proves (38) for $\alpha \leq p$. If $p < \alpha \leq p + 2$, then (38) holds automatically, since then $2p < p + \alpha$ and therefore the expectation in (38) does not exceed the finite moment $\mathbb{E}X_1^{2p}$, while the right-hand side is bounded away from zero.

For the second assertion, one may assume that $k = 1$, in which case the expectation in (39) is equal to and satisfies

$$\mathbb{E}\,\frac{X_1^{2p+2}}{(\frac{1}{n}\,X_1 + 1)^{p+1}} = \mathbb{E}\,\frac{X_1^{2p+2}}{(\frac{1}{n}\,X_1 + 1)^{p+1}}\,1_{\{X_1 \geq n\}} + \mathbb{E}\,\frac{X_1^{2p+2}}{(\frac{1}{n}\,X_1 + 1)^{p+1}}\,1_{\{X_1 < n\}}$$
$$\leq n^{p+1}\,\mathbb{E}X^{p+1}\,1_{\{X \geq n\}} + \mathbb{E}\,X^{2p+2}\,1_{\{X<n\}}.$$

Here, similarly to the previous step, if $\alpha < p + 2$,

$$\mathbb{E}\, X^{2p+2}\, 1_{\{X<n\}} \;\le\; 2p \int_0^n x^{p-\alpha+1}\, \varepsilon(x)\, dx \;=\; O(n^{p-\alpha+2}).$$

In the case $\alpha = p + 2$, the last integral is bounded by $O(\log n)$. In addition,

$$\mathbb{E}\, X^{p+1}\, 1_{\{X\ge n\}} = O(n^{-\alpha+1}) + p \int_n^\infty x^{-\alpha}\, \varepsilon(x)\, dx \;=\; O(n^{-\alpha+1}).$$

$\square$

Lemma 9 *If the moment* $\mathbb{E}X^4$ *is finite, then*

$$\mathbb{E}\, \frac{X_1^2 \dots X_p^2}{(\frac{1}{n}\, S_p + 1)^{p+1}} \;=\; O(1).$$

Moreover, for any index $1 \le k \le p$,

$$\mathbb{E}\, \frac{X_1^2 \dots X_p^2}{(\frac{1}{n}\, S_p + 1)^{p+1}}\, X_k^2 \;=\; O(1).$$

This statement is clear. The last expectation does not exceed $\mathbb{E}X^4\, (\mathbb{E}X^2)^{p-1}$ which is finite and does not depend on n.

Lemma 10 *Let* $\gamma = (\gamma_1, \dots, \gamma_r) \in \mathrm{C}(p, r)$, $2 \le r \le p$. *Under the condition* (37) *with* $2 < \alpha \le 4$, *for any index* $1 \le k \le r$,

$$\mathbb{E}\, \frac{X_1^{2\gamma_1} \dots X_r^{2\gamma_r}}{(\frac{1}{n}\, S_r + 1)^{p+1}}\, X_k^2 \;=\; O(n^{p-r-\alpha+4}). \tag{40}$$

Proof One may reformulate (40) as the statement

$$\mathbb{E}\, \frac{X_1^{2\gamma_1} \dots X_r^{2\gamma_r}}{(\frac{1}{n}\, S_r + 1)^{p+1}} \;=\; O(n^{p-r-\alpha+4}) \tag{41}$$

in which $\gamma = (\gamma_1, \dots, \gamma_r) \in \mathrm{C}(p+2, r)$. If all $\gamma_i \le \frac{p+2}{2}$, there is nothing to prove, since then

$$\mathbb{E}\, \frac{X_1^{2\gamma_1} \dots X_r^{2\gamma_r}}{(\frac{1}{n}\, S_r + 1)^{p+1}} \le \mathbb{E}\, X_1^{2\gamma_1} \dots \mathbb{E}\, X_r^{2\gamma_r} \;\le\; (\mathbb{E}X^{p+2})^r.$$

In the other case, we repeat the arguments used in the proof of Lemma 6. Suppose for definiteness that γ_1 is the largest number among all γ_is. Necessarily, $\gamma_1 > \frac{p+2}{2}$, and therefore, $\gamma_i < \frac{p+2}{2}$ for all $i \geq 2$. Since $S_r < n \Rightarrow X_1 < n$, we similarly have

$$\mathbb{E}\,\frac{X_1^{2\gamma_1}\dots X_r^{2\gamma_r}}{(\frac{1}{n}\,S_r+1)^{p+1}}\,1_{\{S_r<n\}} \;\leq\; (\mathbb{E}X^{p+2})^{r-1}\,\mathbb{E}\,X^{2\gamma_1}\,1_{\{X<n\}}. \tag{42}$$

To bound the last expectation, note that, if $x = n$ is the point of continuity of $F(x)$,

$$\begin{aligned}\mathbb{E}\,X^{2\gamma_1}\,1_{\{X<n\}} &= -n^{2\gamma_1}\,(1-F(n)) + 2\gamma_1\int_0^n x^{2\gamma_1-1}\,(1-F(x))\,dx\\ &\leq 2\gamma_1\int_0^n x^{2\gamma_1-1-p-\alpha}\,\varepsilon(x)\,dx.\end{aligned}$$

But since $\gamma_1 \leq p-r+3$ (which follows from $\gamma_1+\gamma_2+\dots+\gamma_r = p+2$ and $\gamma_i \geq 1$), we have

$$\int_1^n x^{2\gamma_1-p-\alpha-1}\,\varepsilon_{p+\alpha}(x)\,dx \;\leq\; \int_1^n x^{p-2r-\alpha+5}\,\varepsilon(x)\,dx.$$

The last integral grows at the desired rate $O(n^{p-r-\alpha+4})$ as the worst case, if and only if $p-2r-\alpha+5 \leq p-r-\alpha+3$, that is, $r \geq 2$ (which is true). Thus,

$$\mathbb{E}\,X^{2\gamma_1}\,1_{\{X<n\}} \;=\; O(n^{p-r-\alpha+4})$$

In view of (42), this proves (41) for the part of the expectation restricted to the set $S_r < n$, that is,

$$\mathbb{E}\,\frac{X_1^{2\gamma_1}\dots X_r^{2\gamma_r}}{(\frac{1}{n}\,S_r+1)^{p+1}}\,1_{\{S_r<n\}} \;=\; O(n^{p-r-\alpha+4}). \tag{43}$$

Here, the worst situation is attained in the case $r = 2$, $\gamma_1 = p+1$, $\gamma_2 = 1$.

Turning to the expectation over the complementary set $S_r \geq n$, introduce the events

$$\Omega_i = \Big\{X_i \geq \max_{j\neq i} X_j\Big\}, \quad i = 1,\dots,r.$$

On every such set, $X_i \leq S_r \leq rX_i$. In particular, $S_r \geq n$ implies $X_i \geq n/r$. Hence, together with (43), (41) would follow from the inequality

$$\mathbb{E}\,\frac{X_1^{2\gamma_1}\dots X_r^{2\gamma_r}}{X_i^{p+1}}\,1_{\{X_i\geq n\}\cap\Omega_i} \;=\; O(n^{-r-\alpha+3}) \tag{44}$$

with an arbitrary index $1 \le i \le r$.

Case 1 $i \ge 2$. If we fix any values $X_1 = x_1$ and $X_i = x_i$, then the expectation with respect to X_j, $j \ne i$, in (44) will yield a bounded quantity (since the $(p+2)$ moment is finite). Hence, (44) is simplified to

$$\mathbb{E}\, X_1^{2\gamma_1} X_i^{2\gamma_i - p - 1}\, 1_{\{X_i \ge n\} \cap \{X_i \ge X_1\}} \;=\; O(n^{-r-\alpha+3}). \tag{45}$$

Here, the expectation over X_1 may be estimated similarly to the previous step, by replacing n with x_i. Recall that $\gamma_1 > \frac{p+2}{2}$ and hence $2\gamma_1 \ge p + 3$.

Case 1.1 $2\gamma_1 > p + \alpha$. Then, we have

$$\mathbb{E}\, X_1^{2\gamma_1} 1_{\{X_1 \le x_i\}} \;\le\; 2\gamma_1 \int_0^{x_i} x^{2\gamma_1 - 1 - p - \alpha}\, \varepsilon(x)\, dx \;\le\; C x_i^{2\gamma_1 - p - \alpha}$$

with some constant $C > 0$. Hence, up to a constant, the expectation in (44) is bounded by

$$\begin{aligned}
\mathbb{E}\, X_i^{2\gamma_i + 2\gamma_1 - 2p - \alpha - 1}\, 1_{\{X_i \ge n\}} &= \int_n^\infty x^{2\gamma_i + 2\gamma_1 - 2p - \alpha - 1}\, dF(x) \\
&= n^{2\gamma_i + 2\gamma_1 - 2p - \alpha - 1}\, (1 - F(n)) \\
&\quad + c_i \int_n^\infty x^{2\gamma_i + 2\gamma_1 - 2p - \alpha - 2}\, (1 - F(x))\, dx \\
&= O(n^{2\gamma_i + 2\gamma_1 - 3p - 2\alpha - 1}) \\
&\quad + c_i \int_n^\infty x^{2\gamma_i + 2\gamma_1 - 3p - 2\alpha - 2}\, \varepsilon(x)\, dx \\
&= O(n^{2\gamma_i + 2\gamma_1 - 3p - 2\alpha - 1}).
\end{aligned}$$

To obtain (45), it remains to check that $2\gamma_i + 2\gamma_1 - 3p - 2\alpha - 1 \le -r - \alpha + 3$. And indeed, since

$$p = \gamma_i + \gamma_1 + \sum_{j \ne i, 1} \gamma_j \ge \gamma_i + \gamma_1 + (r - 2),$$

the desired relation would follow from $2\,(p - (r-2)) - 3p - 2\alpha - 1 \le -r - \alpha + 3$, that is, $p + r \ge \alpha$ (which is true since $\alpha \le 4$ while $p, r \ge 2$).

Case 1.2 $2\gamma_1 \le p+\alpha$. Then, $\mathbb{E}\, X_1^{2\gamma_1} 1_{\{X_1 \le x_i\}} \le \mathbb{E}\, X^{p+\alpha}$ which is bounded in x_i, and the expectation in (44) does not exceed up to a constant

$$
\begin{aligned}
\mathbb{E}\, X_i^{2\gamma_i+2\gamma_1-p-1}\, 1_{\{X_i \ge n\}} &= \int_n^\infty x^{2\gamma_i+2\gamma_1-p-1}\, dF(x) \\
&= n^{2\gamma_i+2\gamma_1-p-1}\,(1-F(n)) \\
&\quad + c_i \int_n^\infty x^{2\gamma_i+2\gamma_1-p-2}\,(1-F(x))\, dx \\
&= O(n^{2\gamma_i+2\gamma_1-2p-\alpha-1}) \\
&\quad + c_i \int_n^\infty x^{2\gamma_i+2\gamma_1-2p-\alpha-2}\,\varepsilon(x)\, dx \\
&= O(n^{2\gamma_i+2\gamma_1-2p-\alpha-1}).
\end{aligned}
$$

To obtain (45), it remains to check that

$$2\gamma_i + 2\gamma_1 - 2p - \alpha - 1 \le -r - \alpha + 3.$$

And indeed, by (45), the desired relation would follow from $2\,(p-(r-2)) - 2p - \alpha - 1 \le -r - \alpha + 3$, that is, $r \ge 0$.

Case 2 $i = 1$. If we fix any value $X_1 = x_1$ in (44), the expectation with respect to X_j, $j \ne 1$, yields a bounded quantity. Hence, (44) is simplified to

$$\mathbb{E}\, X^{2\gamma_1-p-1}\, 1_{\{X \ge n\}} = O(n^{-r-\alpha+3}).$$

We have

$$
\begin{aligned}
\mathbb{E}\, X^{2\gamma_1-p-1}\, 1_{\{X \ge n\}} &= \int_n^\infty x^{2\gamma_1-p-1}\, dF(x) \\
&= n^{2\gamma_1-p-1}\,(1-F(n)) \\
&\quad + (2\gamma_1 - p - 1) \int_n^\infty x^{2\gamma_1-p-2}\,(1-F(x))\, dx \\
&= O(n^{2\gamma_1-2p-\alpha-1}) + \int_n^\infty x^{2\gamma_1-2p-\alpha-2}\,\varepsilon(x)\, dx \\
&= O(n^{2\gamma_1-2p-\alpha-1}).
\end{aligned}
$$

It remains to be seen that $2\gamma_1 - 2p - \alpha - 1 \le -r - \alpha + 3$, that is, $2\gamma_1 + r \le 2p + 4$. But this follows from $p + 2 = \gamma_1 + \cdots + \gamma_r \ge \gamma_1 + (r-1) \ge \gamma_1 + \frac{r}{2}$.

□

4 Proofs of Main Results

Proof of Theorem 1 Fix $k \geq 3$. In the following, we omit k in notation when it is not necessary. So we write $S_n = S_n(k)$, $\lambda = \lambda(k)$, $Z = Z_k$, and $I = I(k)$. We say that a random variable V has a mixed Poisson distribution with mixing distribution F when, for every integer $m \geq 0$,

$$\mathbb{P}(V = m) = \mathbb{E}\left(e^{-\Lambda}\frac{\Lambda^m}{m!}\right),$$

where Λ is a random variable with distribution F.

Put $\Lambda = \sum_{\alpha \in I} \mathbb{E}_{W_1, W_2, ..., W_n} Y_\alpha$.

We have for any real function $h : \{0, 1, 2, ...\} \to \mathbb{R}$

$$|\mathbb{E}h(S_n) - \mathbb{E}h(Z)| \leq \mathbb{E}|\mathbb{E}_{W_1,...,W_n} h(S_n) - \mathbb{E}_{W_1,...,W_n} h(V)| + |\mathbb{E}h(V) - \mathbb{E}h(Z)|. \tag{46}$$

For each $\alpha \in I$, define $B_\alpha \equiv \{\beta \in I : \alpha \text{ and } \beta \text{ have at least one edge in common}\}$. Put

$$b_1 = \sum_{\alpha \in I} \sum_{\beta \in B_\alpha} p_\alpha p_\beta,$$

where $p_\alpha = \mathbb{E}_{W_1,...,W_n} Y_\alpha$:

$$b_2 = \sum_{\alpha \in I} \sum_{\alpha \neq \beta \in B_\alpha} p_{\alpha\beta},$$

where $p_{\alpha\beta} = \mathbb{E}_{W_1,...,W_n} Y_\alpha Y_\beta$.

Note that, for any $\alpha \in I$ and $\beta \in I \setminus B_\alpha$, the cycles α and β may have joint vertices but they do not have any edge in common. Therefore, for such α and β, the random variables Y_α and Y_β are conditionally independent given weights $W_1, ..., W_n$. Thus, by Theorem 1 in [2], proved with the Chen–Stein method and relations (2) and (46), we get

$$\| \operatorname{L}(S_n(k)) - \operatorname{L}(Z) \| \lesssim \mathbb{E}(b_1 + b_2) + |\mathbb{E}\Lambda - \lambda(k)|, \tag{47}$$

where we write here and in the following that $A_n \lesssim B_n$ or $A_n \gtrsim B_n$ when there exists a positive constant c not depending on n such that $A_n \leq cB_n$ or $A_n \geq cB_n$.

For random variables b_1 and b_2, we get, cf. (14), by the i.i.d. assumption and simple inequality for positive c and d : $2cd \leq c^2 + d^2$

$$\mathbb{E}(b_1 + b_2) \lesssim \sum_{p=k+2}^{2k} \sum_{r=k}^{p-1} \frac{n(n-1)\dots(n-r+1)}{n^p} \sum_{\gamma \in \mathrm{C}(p,r)} \mathbb{E}\psi_n(\gamma), \tag{48}$$

where

$$\psi_n(\gamma) = W_1^{2\gamma_1} \dots W_r^{2\gamma_r} / (\frac{1}{n} L_r + \frac{1}{n} L_{n,r})^p$$

and

$$L_r = W_1 + \cdots + W_r, \quad L_{n,r} = W_{r+1} + \cdots + W_n.$$

For example, we have minimal values $p = k + 2$ and $r = k + 1$ for the cycles $\alpha = (1, 2, \dots, k)$ and $\beta = (1, 2, \dots, k - 1, k + 1)$. Then,

$$\mathbb{E} p_{\alpha\beta} \lesssim \mathbb{E} W_1^4 W_2^2 \dots W_{k-1}^2 W_k^2 W_{k+1}^2 / L_n^{k+2}.$$

We have maximal values $p = 2k$ and $r = k$ for the cycle $\alpha = (1, 2, \dots, k)$. Then,

$$\mathbb{E} p_\alpha^2 \le \mathbb{E} W_1^4 \dots W_k^4 / L_n^{2k}.$$

Lemmas 4 and 6 and inequality (48) under condition (4) imply

$$\mathbb{E}(b_1 + b_2) = o\left(\frac{\log n}{n}\right). \tag{49}$$

Now, we construct an upper bound for the last summand in (47).

It is clear that

$$\mathbb{E}\Lambda \le \frac{1}{2k} \mathbb{E}\left(\frac{(W_1^2 + \cdots + W_n^2)^k}{L_n^k}\right). \tag{50}$$

On the other hand, note that for a positive a and positive sequence $\{x_i\}, i = 1, 2, \dots k,$ we have (see, e.g., Lemma 8 in [17])

$$\prod_{i=1}^{k} \frac{1}{a + x_i} \ge \frac{1}{a^k} - \frac{\sum_{i=1}^k x_i}{a^{k+1}}.$$

Therefore, by the i.i.d. assumption, we get

$$\begin{aligned} \mathbb{E}\Lambda \ge \frac{1}{2k} \mathbb{E}\left(\frac{(W_1^2 + \cdots + W_n^2)^k}{L_n^k}\right) \\ -c_1 \sum_{r=1}^{k-1} \frac{n(n-1)\dots(n-r+1)}{n^k} \sum_{\gamma \in \mathrm{C}(k,r)} \mathbb{E}\psi_n(\gamma) \\ - \frac{c_2}{n} \sum_{\gamma \in \mathrm{C}(k+1,k)} \mathbb{E}\left(\frac{W_1^{2\gamma_1} \dots W_r^{2\gamma_r}}{(L_n/n)^{k+1}}\right), \end{aligned} \tag{51}$$

where c_1 and c_2 do not depend on n.

Combining Lemmas 4 and 6 and relations (7), (47), (49), (50), and (51), we finish the proof of Theorem 1. □

Proof of Theorem 2 We split the proof of the theorem into several steps. Without loss of generality, let $\mathbb{E}X = 1$.

Necessity By Lemma 3, for the convergence $\mathbb{E}T_n^p \to (\mathbb{E}X^2)^p$, it is necessary that all summands in (14) with $r < p$ should be vanishing at infinity. In particular, for the shortest tuple γ with $r = 1$ as in Lemma 2, it should be required that $n^{1-p}\,\mathbb{E}\,\xi_n(\gamma) \to 0$ as $n \to \infty$. Hence, from the inequality (15), it follows that

$$\mathbb{E}X^p\,1_{\{X\geq n\}} = o(1/n).$$

This relation may be simplified in terms of the tails of the distribution of X. Indeed, $\mathbb{E}\,X^p\,1_{\{X\geq n\}} \geq n^p\,\mathbb{P}\{X \geq n\}$, so that the property (6) is necessary for the convergence $\mathbb{E}T_n^p \to (\mathbb{E}X^2)^p$.

Sufficiency and Rate of Convergence First, note that the condition (6) ensures that the moment $\mathbb{E}X^p$ is finite. For the convergence part of Theorem 2, we apply Lemmas 4–6, which imply that $\mathbb{E}\xi_n(\gamma) = o(n^{p-r})$ for any collection $\gamma = (\gamma_1, \dots, \gamma_r)$ with $r < p$. It remains to take into account Lemma 3 about the longest tuple $\tilde{\gamma} = (1, \dots, 1)$ of length $r = p$ and to recall the representation (14).

Turning to the rate of convergence, first, note that by Lemmas 4 and 6, for any $\gamma \in \mathrm{C}(p,r)$ with $2 \leq r \leq p-1$,

$$\frac{n(n-1)\dots(n-r+1)}{n^p}\,\mathbb{E}\xi_n(\gamma) = o\Big(\frac{\log n}{n}\Big) \tag{52}$$

For the shortest tuple $\gamma = (p)$ with $r = 1$, we apply Lemma 5 with $q = p + \frac{3}{2}$ and thus assume that $\mathbb{P}\{X \geq x\} = O(x^{-p-\frac{3}{2}})$. Together with Lemma 4, this gives

$$\frac{n}{n^p}\,\mathbb{E}\xi_n(\gamma) = O\Big(\frac{1}{\sqrt{n}}\Big). \tag{53}$$

Note that with this tail hypothesis, necessarily, $\mathbb{E}X^\beta < \infty$ for any $\beta < p + \frac{3}{2}$. Since $p \geq 2$, we have that the third moment $\mathbb{E}X^3$ is finite. Applying both (52) and (53) in the representation (14) and using (20), we thus obtain that

$$\mathbb{E}T_n^p = \mathbb{E}\xi_n 1_{B_{n,p}} + O(1/\sqrt{n}), \tag{54}$$

with ξ_n defined in (18).

An asymptotic behavior of the last expectation in (54) remains to be studied. Note that $\frac{1}{n}\,S_n \geq \frac{1}{n}\,S_{n,p} \geq \frac{1}{2}$ on the set $B_{n,p}$ as long as $n \geq 2p$. Applying the Taylor formula, we use an elementary inequality $|x^{-p} - 1| \leq p\,2^{p+1}\,|x-1|$ for $x \geq \frac{1}{2}$. In particular, on the set $B_{n,p}$, one has $\Big|(\frac{1}{n}\,S_n)^{-p} - 1\Big| \leq p\,2^{p+1}\,\Big|\frac{1}{n}\,S_n - 1\Big|$.

This gives

$$\left|\xi_n - X_1^2 \dots X_p^2 \right| 1_{B_{n,p}} \le p\, 2^{p+1}\, X_1^2 \dots X_p^2 \left|1 - \frac{1}{n}\, S_p - \frac{1}{n}\, S_{n,p}\right|$$
$$\le p\, 2^{p+1}\, X_1^2 \dots X_p^2 \left|1 - \frac{1}{n}\, S_{n,p}\right| + \frac{p\, 2^{p+1}}{n}\, X_1^2 \dots X_p^2\, S_p,$$

so, taking the expected values,

$$\left|\mathbb{E}\xi_n 1_{B_{n,p}} - \mathbb{E}X_1^2 \dots X_p^2\, 1_{B_{n,p}}\right| \le p\, 2^{p+1}\, (\mathbb{E}X^2)^p\, \mathbb{E}\left|1 - \frac{1}{n}\, S_{n,p}\right|$$
$$+ \frac{p\, 2^{p+1}}{n}\, (\mathbb{E}X^2)^{p-1}\, \mathbb{E}X^3.$$

In view of (17),

$$\mathbb{E}X_1^2 \dots X_p^2\, 1_{B_{n,p}} = \mathbb{E}X_1^2 \dots X_p^2 + e^{-cn} = (\mathbb{E}X^2)^p + o(e^{-cn}).$$

for some constant $c > 0$. Recalling (20), we thus get that

$$\left|\mathbb{E}\xi_n - (\mathbb{E}X^2)^p\right| \le p\, 2^{p+1}\, (\mathbb{E}X^2)^p\, \mathbb{E}\left|1 - \frac{1}{n}\, S_{n,p}\right|$$
$$+ \frac{p\, 2^{p+1}}{n}\, (\mathbb{E}X^2)^{p-1}\, \mathbb{E}X^3 + o(e^{-cn}).$$

Finally,

$$\mathbb{E}\left|\frac{1}{n}\, S_{n,p} - 1\right| = \frac{1}{n}\, \mathbb{E}\, |S_{n,p} - n| \le \frac{1}{n}\, \mathbb{E}\, |S_{n,p} - (n-p)| + \frac{p}{n}$$
$$\le \frac{1}{n}\sqrt{\mathrm{Var}(S_{n,p})} + \frac{p}{n} \le \frac{1}{\sqrt{n}}\sqrt{\mathbb{E}X^2} + \frac{p}{n}.$$

It remains to refer to (54). □

Proof of Theorem 3 Let us apply Lemmas 8–10 in the inequality (36). Using the bounds for the cases $r = 1$, $r = p$, and $2 \le r \le p - 1$ and assuming that (37) is fulfilled for an integer $p \ge 2$ and a real number $2 < \alpha \le 4$, they imply that

$$\mathbb{E}R_n^{(p)} \le e^{-cn} + \Big(\frac{1}{n^{\alpha-2}} + \frac{1}{n} + \frac{1}{n^{\alpha-3}}\Big) + \Big(\frac{\log n}{n^{\alpha-2}} + \frac{1}{n} + \frac{\log n}{n^2}\Big)\, \mathbb{E}M_n^2,$$

where the constant $c > 0$ does not depend on n. To simplify, we have to assume that $\alpha \ge 3$ leading to

$$c\, \mathbb{E}R_n \le \frac{1}{n^{\alpha-3}} + \frac{\log n}{n}\, \mathbb{E}M_n^2. \tag{55}$$

The last expectation in (55) may also be estimated in a polynomial way. Namely, since, for any $q \geq 2$, one has $M_n^2 \leq (X_1^2 + \cdots + X_n^q)^{2/q}$, we get, by Jensen's inequality,

$$\mathbb{E}M_n^2 \leq (\mathbb{E}X_1^q + \cdots + \mathbb{E}X_n^q)^{\frac{2}{q}} = n^{\frac{2}{q}} (\mathbb{E}X^q)^{\frac{2}{q}}.$$

Therefore, choosing $2 < q < p + \alpha$ to be sufficiently close to $p + \alpha$ and using $\alpha = 7/2$, from (55), we obtain (9).

When $\mathbb{E}X^{p+4}$ is finite, we get (10).

At last, to prove (11), note that the finiteness of the exponential moment of X is actually equivalent to the family of moment bounds $(\mathbb{E}X^q)^{1/q} \leq cq$, for $q \geq 1$, which for $q \geq 2$ give

$$\mathbb{E}M_n^2 \leq \mathbb{E}\,(X_1^q + \cdots + X_n^q)^{2/q} \leq (\mathbb{E}X_1^q + \cdots + \mathbb{E}X_n^q)^{2/q} \leq (cq)^2\, n^{2/q}.$$

Choosing here q to be of order $\log n$, we arrive at $\mathbb{E}M_n^2 \leq C\,(\log n)^2$ with a constant C independent of n. Applying this bound in (55) with $\alpha = 4$, we then obtain the much better rate as in (11). □

Acknowledgments Theorem 1 has been obtained under support of the Ministry of Education and Science of the Russian Federation as part of the program of the Moscow Center for Fundamental and Applied Mathematics under the agreement No. 075-15-2019-1621. Theorems 2 and 3 were proved within the framework of the HSE University Basic Research Program. Research of S.Bobkov was supported by the NSF grant DMS-1855575.

References

1. Arratia, R., Goldstein, L., Gordon, L.: Two moments suffice for Poisson approximations: the Chen-Stein method. Ann. Probab. **17**, 9–25 (1989)
2. Arratia, R., Goldstein, L., Gordon, L.: Poisson approximation and the Chen-Stein method. Stat. Sci. **5**, 403–434 (1990)
3. Bhamidi, S., van der Hofstad, R., van Leeuwaarden, J.S.H.: Novel scaling limits for critical inhomogeneous random graphs. Ann. Probab. **40**, 2299–2361 (2012)
4. Bollobas, B., Janson, S., Riordan, O.: The phase transition in inhomogeneous random graphs. Random Struct. Algorithms **31**, 3–122 (2007)
5. Britton, T., Deijfen, M., Martin-Löf, A.: Generating simple random graphs with prescribed degree distribution. J. Stat. Phys. **124**, 1377–1397 (2006)
6. Chakrabarti, D., et al.: Epidemic thresholds in real networks. ACM Trans. Inf. Syst. Secur. (TISSEC) **10**(4), 1–26 (2008)
7. Chung, F., Lu, L.: Connected components in random graphs with given expected degree sequences. Ann. Combinat. **6**(2), 125–145 (2002)
8. Chung, F., Lu, L.: The volume of the giant component of a random graph with given expected degrees. SIAM J. Discrete Math. **20**, 395–411 (2006) (electronic)
9. Demarco, B., Kahn, J.: Upper tails for triangles. Random Struct. Algorithms **40**(4), 452–459 (2012)
10. Erdos, P., Renyi, A.: On random graphs. Publ. Math. Debrecen **6**, 290–297 (1959)

11. Faloutsos, C., Faloutsos, P., Faloutsos, M.: On power-law relationships of the internet topology. Comput. Commun. Rev. **29**, 251–262 (1999)
12. Gilbert, E.N.: Random graphs. Ann. Math. Stat. **30**, 1141–1144 (1959)
13. Hu, Z., Ulyanov, V., Feng, Q.: Limit theorems for number of edges in the generalized random graphs with random vertex weights. J. Math. Sci. **218**(2), 231–237 (2016)
14. Janson, S., Oleszkiewicz, K., Rucinski, A.: Upper tails for subgraph counts in random graphs. Israel J. Math. **142**, 61–92 (2004)
15. Janson, S.: Asymptotic equivalence and contiguity of some random graphs. Random Struct. Algorithms **36**, 26–45 (2010)
16. Kim, J.H., Vu, V.H.: Divide and conquer martingales and the number of triangles in a random graph. Random Struct. Algorithms **24**(2), 166–174 (2004)
17. Liu, Q., Dong, Z.: Moment-based spectral analysis of large-scale generalized random graphs. IEEE Access **5**, 9453–9463 (2017)
18. Liu, Q., Dong, Z.: Limit laws for the number of triangles in the generalized random graphs with random node weights. Stat. Probab. Lett. **161**, 108733 (2020)
19. Norros, I., Reittu, H.: On a conditionally Poisson graph process. Adv. Appl. Probab. **38**, 59–75 (2006)
20. Preciado, V.M., Jadbabaie, A.: Moment-based spectral analysis of large-scale networks using local structural information. IEEE/ACM Trans. Netw. **21**(2), 373–382 (2013)
21. van der Hofstad, R.: Random Graphs and Complex Networks. Cambridge Series in Statistical and Probabilistic Mathematics, vol. 1. Cambridge University Press, Cambridge (2017)