Chin. Ann. Math. Ser. B
43(1), 2022, 63–78
DOI: 10.1007/s11401-022-0305-x

Chinese Annals of Mathematics, Series B

Range-Renewal Processes: SLLNs and Power Laws*

Xinxing CHEN[1] Jiansheng XIE[2] Jiangang YING[2]

Abstract Given n samples (viewed as an n-tuple) of a γ-regular discrete distribution π, in this article the authors concern with the weighted and unweighted graphs induced by the n samples. They first prove a series of SLLN results (of Dvoretzky-Erdös' type). Then they show that the vertex weights of the graphs under investigation obey asymptotically power law distributions with exponent $1+\gamma$. They also give a conjecture that the degrees of unweighted graphs would exhibit asymptotically power law distributions with constant exponent 2. This exponent is obviously independent of the parameter $\gamma \in (0,1)$, which is a surprise to us at first sight.

Keywords Range renewal process, Strong law of large numbers, Power law
2000 MR Subject Classification 60F15, 05C80

1 Introduction

Let $\xi := (\xi_n : n \geq 1)$ be a random symbol sequence. Let R_n be the number of distinct values among the first n elements of the process ξ. We call $(R_n : n \geq 1)$ the range-renewal process of ξ. An interesting problem is to investigate the growth rate of R_n.

In the autumn of 2010, the second author reported a classic result of Dvoretzky and Erdös [13] in a seminar at Fudan University and was fascinated by their neat and beautiful result that, for a simple symmetric random walk (SSRW for short) on $\mathbb{Z}^d$ with $d \geq 2$, (R_n) satisfies the following strong law of large numbers (SLLN for short): $\frac{R_n}{\mathbb{E}R_n} \xrightarrow{\text{a.s.}} 1$. We then try to find out the more recent results concerning R_n for more general process ξ. For Markov chains, Chosid and Isaac [9–10], and Athreya [3] obtained that, under a suitable integrable condition, $\frac{R_n}{n} \xrightarrow{\text{a.s.}} 0$. Derriennic [11] extended Dvoretzky-Erdös' result to simple random walks ξ on arbitrarily discrete Abelian groups, and showed that $\lim\limits_n \frac{R_n}{n} = 0$ a.s. when ξ is recurrent; otherwise the limit is the escape probability. In addition, the central limit theorem of R_n of SSRW on $\mathbb{Z}^d$ can be found in Jain and Pruitt [21, 23] ($d \geq 3$) and Le Gall [25] ($d = 2$); the corresponding laws of the iterated logarithm are discussed by Jain and Pruitt [22] ($d \geq 4$) and Bass and Kumagai [7] ($d = 2$ or 3). More discussions on R_n of null recurrent or transient Markov chains can be found in [14, 16–19, 27–29] and references therein.

However, there are relatively few results concerning R_n of positive recurrent Markov chains (or of stationary processes). What will be the accurate order of R_n tending to $+\infty$? This

Manuscript received April 28, 2020. Revised April 3, 2021.
[1]School of Mathematical Sciences, Shanghai Jiaotong University, Shanghai 200240, China.
E-mail: chenxinx@sjtu.edu.cn
[2]School of Mathematical Sciences, Fudan University, Shanghai 200433, China.
E-mail: jsxie@fudan.edu.cn jgying@fudan.edu.cn
*This work was supported by the National Natural Science Foundation of China (Nos. 11871162, 11771286, 11271255, 11271077, 11001173, 11790273) and the Key Laboratory of Mathematics for Nonlinear Science, Fudan University.

problem has rather few investigation even for independent and identically distributed (i.i.d. for short) sequence ever since Dvoretzky and Erdös' work in [13].

So, in this paper, we focus on the simple (but nontrivial) case that $\xi = (\xi_n : n \geq 1)$ is an sequence of i.i.d. random variables with a common discrete distribution π. In such situation, after our main work of the current paper being done, we recently found that, Bahadur [4] had proved $\lim\limits_{n\to\infty} \frac{R_n}{\mathbb{E}R_n} = 1$ in probability; while we will prove as a preliminary result that the convergence indeed holds almost surely. We next consider the random digraph $\overrightarrow{G}_n$ (and undigraph $\overline{G}_n$ respectively) formed by the first n-steps of ξ, and establish a system of SLLNs for the numbers of different kinds of vertices in $\overrightarrow{G}_n$ (and in $\overline{G}_n$ respectively) under the condition of π being γ-regular (see Definition 2.2). These results lead us to the discovery of an interesting phenomenon that the vertex weights of $\overrightarrow{G}_n$ (and of $\overline{G}_n$ respectively) obey asymptotically power-law distributions with exponent $1+\gamma$. Based on these results, we also propose a conjecture that the un-weighted (directed and undirected) graphs, with their degrees concerned, would exhibit power law distributions with exponent 2, independent of the regular index $\gamma \in (0,1)$. As it is well known, there are thousands of works (see [1–2, 5–6] and references therein and thereafter) concerning power-laws of different kinds of random graph models; a significant part of our current work was influenced and inspired by them.

The paper is organized in the following way. Section 2 is devoted to the presentation of the main settings and the main results. In Section 3 we present some necessary estimates for our model. Section 4 is devoted to the proof of the main Theorems 2.1–2.4. In Section 5 we discuss the critical case of $\gamma = 0$ and $\gamma = 1$. There, for the non-critical case of $\gamma \in (0,1)$, we also propose the conjecture just mentioned above with a heuristic deduction.

2 Main Settings and Main Results

2.1 Main settings

Let π be a probability measure on $\mathbb{N}$ with $\pi_i \geq \pi_{i+1} > 0$ for all $i \in \mathbb{N}$. Let $\xi = (\xi_n : n \geq 1)$ be a sequence of i.i.d. random variables with common law π. Let R_n be the number of distinct values achieved by the first n samples from ξ, i.e.,

$$R_n := \#\{\xi_k : 1 \leq k \leq n\}. \tag{2.1}$$

We define $\overrightarrow{G}_n := (V_n, \overrightarrow{E}_n, W_n)$ to be the random weighted directed graph formed by the first n steps of ξ. Here, vertex set V_n, edge set $\overrightarrow{E}_n$ and weight function W_n are defined as follows: $V_n := \{\xi_i : 1 \leq i \leq n\}$, $\overrightarrow{E}_n := \{\overrightarrow{\xi_i\xi_{i+1}} : 1 \leq i \leq n-1\}$ and $W_n(x,y) := \sum\limits_{i=1}^{n-1} 1_{\{\xi_i = x, \xi_{i+1} = y\}}$. Let $W_n(x) := \sum\limits_{y} W_n(x,y)$ be the weight of vertex x in $\overrightarrow{G}_n$. Write

$$\overrightarrow{\mathcal{V}}_n(x) := \{y \in V_n : W_n(x,y) \geq 1\} \tag{2.2}$$

for the set of outgoing neighbors of x in $\overrightarrow{G}_n$, thus $\#\overrightarrow{\mathcal{V}}_n(x)$ is the out-degree of vertex x. We also define $\overline{G}_n := (V_n, \overline{E}_n)$ for the un-directed graph corresponding to $\overrightarrow{G}_n$, so that, $\overline{G}_n$ is a connected graph without multiedge but possibly has loops. Put $\overline{\mathcal{V}}_n(x) := \{y \in V_n : W_n(x,y) + W_n(y,x) \geq 1\}$, it is the set of neighbors of x in $\overline{G}_n$.

We will investigate the numbers of different kinds of vertices in $\overrightarrow{G}_n$ and $\overline{G}_n$, which are introduced as the following. First, put

$$N_n(x) := \sum_{k=1}^{n} 1_{\{\xi_k = x\}}, \tag{2.3}$$

which is the number of visit times (visit intensity) of ξ at vertex x up to time n. Then $W_n(x) = N_{n-1}(x)$. For each $\ell \geq 1$, set $R_{n,\,\ell} := \sum\limits_x 1_{\{N_n(x)=\ell\}}$, this is the number of vertices with visit intensities being exactly ℓ in $\overrightarrow{G}_n$. Define $R_{n,\,\ell+} := \sum\limits_{k=\ell}^{n} R_{n,\,k}$. Then $R_n = R_{n,1+}$. Similarly, we define $\overrightarrow{R}_{n,\,\ell} := \sum\limits_x 1_{\{\#\overrightarrow{\mathcal{V}}_n(x)=\ell\}}$ to be the number of vertices whose out-degrees are exactly ℓ in $\overrightarrow{G}_n$. Then $\sum\limits_{\ell} \overrightarrow{R}_{n,\,\ell} = R_{n-1}$. Noting that $\sum\limits_y 1_{\{W_n(x,y)\geq 1\}}$ and $\sum\limits_y 1_{\{W_n(y,x)\geq 1\}}$ have the same law, we only consider the out degree $\#\overrightarrow{\mathcal{V}}_n(x)$ here. Set $\overrightarrow{R}_{n,\,\ell+} := \sum\limits_{k=\ell}^{n} \overrightarrow{R}_{n,\,k}$. Define $\overline{R}_{n,\ell}$ and $\overline{R}_{n,\ell+}$ in a similar way for undigraphs $\overline{G}_n$.

2.2 Main results

For the above range-renewal processes induced by i.i.d. samples ξ with a common distribution π, an SLLN of Dvoretzky-Erdös' type holds true as is indicated below.

Theorem 2.1 *For any discrete probability measure π, we have $\lim\limits_{n\to\infty} \frac{R_n}{\mathbb{E}R_n} = 1$ almost surely.*

To present the other main results in a neat way, we introduce the following definitions.

Definition 2.1 *Let $\zeta : [1,\infty) \to [1,\infty)$ be a strictly increasing function. Let $\gamma \in [0,1]$. We say that ζ is γ-regular if*

$$\lim_{x\to\infty} \frac{\zeta(\lambda x)}{\zeta(x)} = \lambda^{\gamma}, \quad \forall \lambda > 0. \tag{2.4}$$

Note that if ζ_i is γ_i-regular for each $i \in \{1,2\}$ and $\lim\limits_{x\to\infty} \frac{\zeta_1(x)}{\zeta_2(x)} = 1$, then $\gamma_1 = \gamma_2$. So, the following definitions of regular distributions on $\mathbb{N}$ are well-posed.

Definition 2.2 *Let $\gamma \in (0,1)$. A distribution π on $\mathbb{N}$ is said to be γ-regular if there is a γ-regular function ζ satisfying*

$$\lim_{n\to\infty} \pi_n \cdot \zeta^{-1}(n) = 1, \tag{2.5}$$

where ζ^{-1} is the inverse function of ζ.

Definition 2.3 *A distribution π on $\mathbb{N}$ is said to be 0-regular if: (1) There is a 0-regular function ζ satisfying (2.5); (2) the function ζ is continuously differentiable and*

$$\lim_{x\to\infty} \frac{\zeta'(\lambda x)}{\zeta'(x)} = \lambda^{-1}, \quad \forall \lambda > 0. \tag{2.6}$$

Definition 2.4 *A distribution π on $\mathbb{N}$ is said to be 1-regular if: (1) There is a 1-regular function ζ satisfying (2.5); (2) there exists an increasing function $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ and a continuous and integrable function $g : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\lim\limits_{x\to\infty} \psi(x) = \infty$ and*

$$\lim_{x\to\infty} \frac{\zeta(x \cdot \mathrm{e}^{\lambda\psi(x)})}{\zeta(x) \cdot \mathrm{e}^{\lambda\psi(x)}} = g(\lambda) \tag{2.7}$$

holds uniformly on each compact λ-set in $(0,\infty)$.

For any γ-regular distribution π on $\mathbb{N}$ with $\gamma \in [0,1]$, we will refer $\gamma(\pi) := \gamma$ to be the regular index of π. In the following we will show some backgrounds for the regular functions.

Remark 2.1 (2.4) is just the definition of regularly varying function with exponent γ, which is originally introduced by Karamata [24]. See for example [15, pp. 241–250], [30, pp. 13], [20, pp. 321–324], or [8] for the definition and related properties.

Remark 2.2 If ζ is γ-regular for some $\gamma \in [0,1]$, then $\lim_{x\to\infty} \frac{\log(\zeta(x))}{\log x} = \gamma$.

Remark 2.3 In the previous versions of this paper, we remarked that we need suitable dominations (without an explicit presentation of those dominations) in limits (2.4) and (2.7) so that the Lebesgue's Dominated Convergence theorem (DCT for short) can be applied in the proof of Lemma 3.2 (and in other similar estimates). Thanks to the anonymous referee for reminding us the following important fact: For regularly varying function φ with exponent $\gamma \in \mathbb{R}$, the limit $\lim_{x\to\infty} \frac{\varphi(\lambda x)}{\varphi(x)} = \lambda^\gamma$ holds uniformly on each compact λ-set in $(0,\infty)$ (see [8, Theorem 1.5.1]). It is also clear that, for any γ-regular function ζ (with $0 \le \gamma \le 1$) and any $\varepsilon \in (0,1)$, there exist constants $C_\varepsilon, M_\varepsilon \ge 1$ such that

$$\frac{\zeta(\lambda x)}{\zeta(x)} \le C_\varepsilon \cdot \lambda^{\gamma+\varepsilon}, \quad \forall x \ge M_\varepsilon,\ \lambda \ge 1. \tag{2.8}$$

So the exchangeability of the limit $z \to \infty$ and the integral with respect to ds in the proof of Lemma 3.2 (and in other similar estimates) holds for regular distributions π on $\mathbb{N}$ with index $\gamma = \gamma(\pi) \in [0,1)$. The requirement of uniformly convergence of the limit (2.7) is to serve the same end in our deductions for the case of 1-regular distributions.

As follows, we show that the regular distributions on $\mathbb{N}$ contain many interesting examples.

Example 2.1 Let $C > 0$, $a > 0$ and $b > 0$.

(1) If $\pi_n = \frac{C+o(1)}{n^{1+a}}$, then choosing $\zeta(x) = (Cx)^\gamma$, one can prove that π is γ-regular with $\gamma = \frac{1}{1+a} \in (0,1)$.

(2.a) If $\pi_n = C \cdot \mathrm{e}^{-an^b} \cdot [1+o(1)]$, then one can choose $\zeta(x) = \big[\frac{\log(Cx)}{a}\big]^{\frac{1}{b}}$ and prove that π is 0-regular.

(2.b) If $\pi_n = C \cdot \mathrm{e}^{-a(\log n)^{1+b}} \cdot [1+o(1)]$, then one can choose $\zeta(x) = \exp\big\{\big[\frac{\log(Cx)}{a}\big]^{\frac{1}{1+b}}\big\}$ and prove that π is 0-regular.

(3) If $\pi_n = \frac{C}{n\cdot(\log n)^{1+a}} \cdot [1+o(1)]$, then one can choose $\zeta(x) = \frac{Cx}{(\log x)^{1+a}}$, $\psi(x) = \log x$ and $g(\lambda) = \frac{1}{(1+\lambda)^{1+a}}$ and prove that π is 1-regular.

Now we assume the common distribution π to be γ-regular for some $\gamma \in (0,1)$. Then we have a series of SLLNs as Theorems 2.2–2.4. For completeness, the interested readers can find in Section 5 the corresponding results for π being 0-regular or 1-regular.

Theorem 2.2 *For each $k \in \mathbb{N}$, we have $\lim_{n\to\infty} \frac{R_{n,k+}}{\mathbb{E}R_{n,k+}} = 1$ almost surely. The same result holds when we replace $R_{n,k+}$ by $\overrightarrow{R}_{n,k+}, \overline{R}_{n,k+}$ respectively.*

Theorem 2.3 *For each $k \ge 1$, we have $\lim_{n\to\infty} \frac{R_{n,k}}{R_n} = r_k(\gamma) := \frac{\gamma\Gamma(k-\gamma)}{\Gamma(1-\gamma)\Gamma(k+1)}$ almost surely, where $\Gamma(\cdot)$ is the usual Gamma function. Consequently, we have almost surely*

$$\lim_{n\to\infty} \frac{R_{n,k+1}}{R_{n,k}} = \frac{k-\gamma}{k+1}, \tag{2.9}$$

$$\lim_{n\to\infty} \frac{R_{n,k}}{R_{n,k+}} = \frac{\gamma}{k}. \tag{2.10}$$

Remark 2.4 (2.10) means that the proportion of the relatively 'new' vertices at visit intensity level k is approximately $\frac{\gamma}{k}$; this is a kind of average escape rate (at intensity level k). In the case of SSRW on $\mathbb{Z}^d$ with $d \geq 3$, the limit in (2.10) is always γ_d, the usual escape rate (escape probability), see for example [26, pp. 220].

Theorem 2.4 *We have almost surely*

$$\lim_{n\to\infty} \frac{\overrightarrow{R}_{n,k}}{R_n} = \overrightarrow{r}_k(\pi) := \sum_{\ell=k}^{\infty} r_\ell(\gamma)\mathbb{P}(R_\ell = k), \tag{2.11}$$

$$\lim_{n\to\infty} \frac{\overline{R}_{n,k}}{R_n} = \overline{r}_k(\pi) := \sum_{\ell=\left\lfloor \frac{k+1}{2} \right\rfloor}^{\infty} r_\ell(\gamma)\mathbb{P}(R_{2\ell} = k). \tag{2.12}$$

Note $\sum_{\ell=1}^{\infty} r_k(\gamma) = 1$ and $r_k(\gamma) = \frac{\gamma}{\Gamma(1-\gamma)} \cdot k^{-(1+\gamma)} \cdot [1 + O(k^{-1})]$ as $k \to \infty$. We remark here that, Theorem 2.3 indeed tells us that $\left\{\frac{R_{n,k}}{R_n}\right\}_{k=1}^{\infty}$ is asymptotically a power law distribution with exponent $1+\gamma$.

With a heuristic argument in Section 5, we also conjecture that $\{\overrightarrow{r}_k(\pi)\}_{k=1}^{\infty}$ and $\{\overline{r}_k(\pi)\}_{k=1}^{\infty}$ (defined in Theorem 2.4) are power law distributions with exponent 2.

In the rest of this section, we show that Theorems 2.2–2.4 may fail to be true without any regular condition. See the following counter example.

Example 2.2 Let $\alpha_1 > \alpha_2 > 1$. There exist some distribution π and an increasing sequence $\{n_j\}_{j=1}^{\infty} \subset \mathbb{N}$ such that almost surely

$$\lim_{j\to\infty} \frac{R_{n_{2j-1},1}}{R_{n_{2j-1}}} = \frac{1}{\alpha_1} \quad \text{and} \quad \lim_{j\to\infty} \frac{R_{n_{2j},1}}{R_{n_{2j}}} = \frac{1}{\alpha_2}. \tag{2.13}$$

For the convenience of readers, we now present a proof of the above example based on Theorems 2.1–2.4, while their own proofs are postponed to the successive sections.

Proof of Example 2.2 For any distribution π on $\mathbb{N}$, write $\mathbb{P}_\pi$ for the probability measure induced by the i.i.d. sequence of $\{\xi_n : n \geq 1\}$ with common distribution π.

Let $\alpha_{2k-1} = \alpha_1, \alpha_{2k} = \alpha_2$ for each $k \geq 1$. Let $\pi^{(1)}$ be a probability measure on $\mathbb{N}$ with $\pi_x^{(1)} := \frac{1}{Z_1 \cdot x^{\alpha_1}}$, $x \in \mathbb{N}$, where $Z_1 := \sum_x \frac{1}{x^{\alpha_1}}$ is the normalizing constant. We know that $\pi^{(1)}$ is $\frac{1}{\alpha_1}$-regular by Example 2.1(1). Applying Theorem 2.3, we have $\frac{R_{n,1}}{R_n} \xrightarrow{\text{a.s.}} \frac{1}{\alpha_1}$ with respect to $\mathbb{P}_{\pi^{(1)}}$. Thus there exist large enough integers $n_1, m_1 \geq 1$ such that

$$\mathbb{P}_{\pi^{(1)}}\Big(\Big|\frac{R_{n_1,1}}{R_{n_1}} - \frac{1}{\alpha_1}\Big| \leq \frac{1}{2}, V_{n_1} \subset [1, m_1)\Big) \geq 1 - \frac{1}{2}$$

and $\sum_{x\geq m_1} \pi_x^{(1)} \leq \frac{1}{2}$. Inductively, we can define a sequence of distributions $(\pi^{(k)}, k \geq 1)$ and two strictly increasing integer sequences $(n_k, k \geq 1)$ and $(m_k, k \geq 1)$ such that

$$\pi_x^{(k+1)} = \begin{cases} \pi_x^{(k)}, & 1 \leq x < m_k; \\ \dfrac{1}{Z_{k+1} \cdot x^{\alpha_{k+1}}}, & x \geq m_k; \end{cases}$$

and for each $k \geq 1$, with $A_k := \left\{\left|\frac{R_{n_k,1}}{R_{n_k}} - \frac{1}{\alpha_k}\right| \leq \frac{1}{2k}\right\} \cap \{V_{n_k} \subset [1, m_k)\}$,

$$\mathbb{P}_{\pi^{(k)}}(A_k) \geq 1 - 2^{-k} \quad \text{and} \quad \sum_{x \geq m_k} \pi_x^{(k)} \leq \frac{1}{2k}.$$

$\big($Here, $Z_{k+1} > 0$ is the constant which satisfies $\sum\limits_x \pi_x^{(k+1)} = 1.\big)$

Let $\pi_x^{(\infty)} := \lim\limits_{n\to\infty} \pi_x^{(n)}$ for each $x \in \mathbb{N}$. Clearly $\pi^{(\infty)}$ is still a probability measure on $\mathbb{N}$. Since $\{x \in \mathbb{N} : \pi_x^{(\infty)} \neq \pi_x^{(k)}\} \subset [m_k, \infty)$ for each $k \geq 1$, we get

$$\mathbb{P}_{\pi^{(\infty)}}(A_k) = \mathbb{P}_{\pi^{(k)}}(A_k) \geq 1 - 2^{-k}.$$

Therefore, (2.13) holds $\mathbb{P}_{\pi^{(\infty)}}$-almost surely.

3 Preliminary Estimates

3.1 Expectation-variance estimate for R_n

For any discrete distribution π, we have the following estimation.

Lemma 3.1 *For each $n \in \mathbb{N}$, we have*

$$\mathbb{E}R_n = \sum_x [1 - (1 - \pi_x)^n], \tag{3.1}$$

and $\mathrm{Var}\,(R_n) \leq \mathbb{E}R_n$. *As $n \to \infty$, $\frac{\mathbb{E}R_n}{n} = o(1)$.*

Proof (3.1) is obvious since $R_n = \sum\limits_x 1_{\{N_n(x)\geq 1\}}$ and $\mathbb{P}(N_n(x) = 0) = (1 - \pi_x)^n$. Next, since $\mathbb{E}[R_n(R_n - 1)] = \sum\limits_{x\neq y} \mathbb{P}(N_n(x) \geq 1, N_n(y) \geq 1)$, we have

$$\begin{aligned}\mathbb{E}[R_n(R_n-1)] &= \sum_{x\neq y}[1 - (1-\pi_x)^n - (1-\pi_y)^n + (1-\pi_x-\pi_y)^n] \\ &\leq \sum_{x,y}[1-(1-\pi_x)^n][1-(1-\pi_y)^n] = [\mathbb{E}R_n]^2.\end{aligned}$$

It follows $\mathrm{Var}\,(R_n) \leq \mathbb{E}R_n$ immediately. Furthermore, noting $\lim\limits_{n\to\infty} \frac{1-(1-\pi_x)^n}{n} = 0$ and $\frac{1-(1-\pi_x)^n}{n} \leq \pi_x$, we have $\frac{\mathbb{E}R_n}{n} = o(1)$ in view of the DCT.

Fix $\gamma \in (0, 1)$ now. We always assume that π is γ-regular and ζ is a γ-regular function which satisfies (2.5). In the forthcoming subsection, we will give some necessary estimates for the related quantities as $n \to \infty$, which will be helpful for the proofs of Theorems 2.2–2.4.

3.2 Expectation estimates for visit intensity statistics

Fix $a \in \left(\frac{1}{2}, \frac{3}{2}\right)$. For each $z > 0$, let $E(z) := E(z; 1)$ with

$$E(z; a) := \sum_{j>\zeta(2)} \left[1 - \left(1 - \frac{a}{\zeta^{-1}(j)}\right)^z\right]. \tag{3.2}$$

Lemma 3.2 $E(z; a) = \Gamma(1-\gamma) \cdot a^\gamma \cdot \zeta(z) \cdot [1 + o(1)]$ *as $z \to \infty$.*

Proof Since $\zeta(\cdot)$ is strictly increasing, the discrete sum in (3.2) can be approximated by the integral $\int_{\zeta(2)}^{\infty}\big[1-\big(1-\frac{a}{\zeta^{-1}(x)}\big)^z\big]\mathrm{d}x$ with the error term bounded by $\big[1-\big(1-\frac{a}{\zeta^{-1}(\zeta(2))}\big)^z\big]\le 1$. Hence

$$\begin{aligned}E(z;a) &= O(1)+\int_{\zeta(2)}^{\infty}\Big[1-\Big(1-\frac{a}{\zeta^{-1}(x)}\Big)^z\Big]\mathrm{d}x = O(1)-\int_0^{\frac{1}{2}}[1-(1-at)^z]\mathrm{d}\zeta\Big(\frac{1}{t}\Big)\\ &= O(1)-\Big[1-\Big(1-\frac{a}{2}\Big)^z\Big]\cdot\zeta(2)+\lim_{t\to 0}[1-(1-at)^z]\cdot\zeta\Big(\frac{1}{t}\Big)\\ &\quad +\int_0^{\frac{1}{2}} az\cdot\zeta\Big(\frac{1}{t}\Big)\cdot(1-at)^{z-1}\mathrm{d}t\\ &= O(1)+\lim_{t\to 0} azt\cdot\zeta\Big(\frac{1}{t}\Big)+a\zeta(z)\cdot\int_0^{\frac{z}{2}}\frac{\zeta(\frac{z}{s})}{\zeta(z)}\cdot\Big(1-\frac{as}{z}\Big)^{z-1}\mathrm{d}s.\end{aligned}$$

For the second term on the right side of the above equation, we have $\lim\limits_{t\to 0}\big[t\cdot\zeta\big(\frac{1}{t}\big)\big]=0$ based on Remark 2.2. So that, we are left to estimate the third term. Applying (2.8) with $\varepsilon=\frac{1-\gamma}{2}$ and $\lambda=s^{-1}$, we get

$$\frac{\zeta(\frac{z}{s})}{\zeta(z)}\le C_{\frac{1-\gamma}{2}}\cdot s^{-\frac{\gamma+1}{2}},\quad \forall\, z\ge M_{\frac{1-\gamma}{2}},\ s\in(0,1).$$

It is easy to get

$$\Big(1-\frac{x}{z}\Big)^{z-1}\le\exp\Big\{-\frac{x(z-1)}{z}\Big\}\le\exp\Big\{-\frac{x}{2}\Big\},\quad \forall\, z>2,\ x>0.$$

Hence $\frac{\zeta(\frac{z}{s})}{\zeta(z)}\cdot\big(1-\frac{as}{z}\big)^{z-1}$ is bounded by $\max\{1,C_{\frac{1-\gamma}{2}}\cdot s^{\frac{-(\gamma+1)}{2}}\}\mathrm{e}^{\frac{-as}{2}}$ for all $z>2+M_{\frac{1-\gamma}{2}}$ and all $s>0$. Therefore, we can apply the DCT to get

$$\begin{aligned}\lim_{z\to\infty}\int_0^{\frac{z}{2}}\frac{\zeta(\frac{z}{s})}{\zeta(z)}\cdot\Big(1-\frac{as}{z}\Big)^{z-1}\mathrm{d}s &= \int_0^{\infty}\lim_{z\to\infty}\frac{\zeta(\frac{z}{s})}{\zeta(z)}\cdot\Big(1-\frac{as}{z}\Big)^{z-1}\cdot 1_{\{s<\frac{z}{2}\}}\mathrm{d}s\\ &= \int_0^{\infty}s^{-\gamma}\mathrm{e}^{-as}\mathrm{d}s=\Gamma(1-\gamma)\cdot a^{\gamma-1}.\end{aligned}$$

This proves the lemma.

Lemma 3.3 $\mathbb{E}R_n=\Gamma(1-\gamma)\cdot\zeta(n)\cdot[1+o(1)]$.

Proof Fix $0<\varepsilon<\frac{1}{2}$. By (2.5), there exists sufficiently large integer i_0, such that $\pi_i\le\frac{1+\varepsilon}{\zeta^{-1}(i)}$ for all $i\ge i_0$. Hence $\mathbb{E}R_n\le O(1)+E(n,1+\varepsilon)$. By Lemma 3.2,

$$\overline{\lim_{n\to\infty}}\frac{\mathbb{E}R_n}{\zeta(n)}\le\lim_{n\to\infty}\frac{E(n,1+\varepsilon)}{\zeta(n)}=\Gamma(1-\gamma)\cdot(1+\varepsilon)^{\gamma}.$$

Similarly, $\underline{\lim}_{n\to\infty}\frac{\mathbb{E}R_n}{\zeta(n)}\ge\Gamma(1-\gamma)\cdot(1-\varepsilon)^{\gamma}$. We prove the lemma by letting $\varepsilon\downarrow 0$.

It is easy to see that, for any $n\ge\ell\ge 1$,

$$\mathbb{E}R_{n,\ell}=\sum_x C_n^{\ell}\cdot\pi_x^{\ell}(1-\pi_x)^{n-\ell}. \tag{3.3}$$

In order to estimate $\mathbb{E}R_{n,\ell}$, we define

$$S_\ell(n):=\sum_x\pi_x^{\ell}(1-\pi_x)^{n-\ell}. \tag{3.4}$$

Then $\mathbb{E}R_{n,\ell}=\frac{n^\ell}{\ell!}S_\ell(n)[1+O(\frac{1}{n})]$ as n goes to infinity.

Lemma 3.4 *Fix $\ell \geq 1$ and $d \geq 1$. As n goes to infinity,*

$$S_\ell(n) = \gamma\Gamma(\ell-\gamma) \cdot \frac{\zeta(n)}{n^\ell} \cdot [1+o(1)],$$
$$\mathbb{E}R_{n,\ell} = \frac{\gamma\Gamma(\ell-\gamma)}{\ell!} \cdot \zeta(n) \cdot [1+o(1)]$$

and

$$S_\ell(n-d) = \Big[1+O\Big(\frac{1}{n}\Big)\Big] \cdot S_\ell(n). \tag{3.5}$$

Proof Fix $\ell \geq 1$. For each $z > \ell$ and $a \in (\frac{1}{2}, \frac{3}{2})$, let

$$S_\ell(z;a) := \sum_{n>\zeta(2)} \Big(\frac{a}{\zeta^{-1}(n)}\Big)^\ell \Big(1-\frac{a}{\zeta^{-1}(n)}\Big)^{z-\ell}.$$

Following the proofs of Lemmas 3.2–3.3, we can prove first $S_\ell(z;a) = a^\gamma \cdot \gamma\Gamma(\ell-\gamma) \cdot \frac{\zeta(z)}{z^\ell} \cdot [1+o(1)]$, and then $S_\ell(n) = \gamma\Gamma(\ell-\gamma) \cdot \frac{\zeta(n)}{n^\ell} \cdot [1+o(1)]$. Hence we have the equation for $\mathbb{E}R_{n,\ell}$.

In proving (3.5), we assume $d=1$ for simplicity. Clearly $S_\ell(n-1) - S_\ell(n) = S_{\ell+1}(n)$ and $\frac{S_{\ell+1}(n)}{S_\ell(n)} = \frac{\gamma\Gamma(\ell+1-\gamma)\cdot\frac{\zeta(n)}{n^{\ell+1}}\cdot[1+o(1)]}{\gamma\Gamma(\ell-\gamma)\cdot\frac{\zeta(n)}{n^\ell}\cdot[1+o(1)]} = \frac{\ell-\gamma+o(1)}{n}$. This proves (3.5).

3.3 Variance estimates for visit intensity statistics

Lemma 3.5 *Fix $\ell \geq 1$. As n goes to infinity,* $\mathrm{Var}\,(R_{n,\ell}) \leq [1+o(1)] \cdot \mathbb{E}R_{n,\ell}$.

Proof Mimicking the proof of Lemma 3.1, we have

$$\mathbb{E}[R_{n,\,\ell}^2 - R_{n,\,\ell}] = \sum_{x\neq y} \mathbb{P}(N_n(x)=\ell, N_n(y)=\ell).$$

By (3.3) and (3.5),

$$\begin{aligned}\mathbb{E}[R_{n,\,\ell}^2 - R_{n,\,\ell}] &= \sum_{x\neq y} \frac{n!}{(\ell!)^2(n-2\ell)!} \cdot \pi_x^\ell \pi_y^\ell \cdot (1-\pi_x-\pi_y)^{n-2\ell} \\ &\leq \frac{n!}{(\ell!)^2(n-2\ell)!} S_\ell(n-\ell)^2 = \Big[\frac{n!}{\ell!(n-\ell)!}\Big]^2 \cdot \Big[1+O\Big(\frac{1}{n}\Big)\Big] \cdot S_\ell(n)^2 \\ &= \Big[1+O\Big(\frac{1}{n}\Big)\Big] \cdot (\mathbb{E}R_{n,\,\ell})^2.\end{aligned}$$

It follows $\mathrm{Var}\,(R_{n,\,\ell}) \leq \big[1+O\big(\frac{\mathbb{E}R_{n,\,\ell}}{n}\big)\big] \cdot \mathbb{E}R_{n,\,\ell}$ immediately. This proves the lemma by $\mathbb{E}R_{n,\ell} \leq \mathbb{E}R_n$ and $\frac{\mathbb{E}R_n}{n} \to 0$.

3.4 Estimates for out-degree statistics

Recall the definition of $\overrightarrow{\mathcal{V}}_n(x)$ in (2.2) and recall also $W_n(x) = N_{n-1}(x)$. In order to investigate $\overrightarrow{R}_{n,k}$, we define

$$\overrightarrow{R}_{n,\,k,\,\ell} := \sum_x 1_{\{W_n(x)=\ell,\, \#\overrightarrow{\mathcal{V}}_n(x)=k\}}. \tag{3.6}$$

Then $\overrightarrow{R}_{n,k} = \sum\limits_{\ell\geq k} \overrightarrow{R}_{n,\,k,\,\ell}$.

Lemma 3.6 *Fix $\ell \geq k \geq 1$. As n goes to infinity,*

$$\mathbb{E}\overrightarrow{R}_{n,\,k,\,\ell} = \frac{\gamma\Gamma(\ell-\gamma)}{\ell!}\cdot \mathbb{P}(R_\ell = k)\cdot \zeta(n)\cdot [1+o(1)].$$

Proof Fix $x \in \mathbb{N}$ and $n \geq \ell \geq k \geq 1$. Define random set $\mathcal{W}_n(x) := \{i \leq n-1 : \xi_i = x\}$. Then $W_n(x) = \#\mathcal{W}_n(x)$. Define $\mathcal{A}^{(k)} := \{A \subset \mathbb{N} : |A| = k\}$ and $\mathcal{A}^{(k)}_x := \{A \subset \mathbb{N} : |A| = k, x \notin A\}$. Write

$$\begin{aligned}\mathcal{B}_\ell &:= \{B \subset [1, n-1]\cap \mathbb{N} : |B| = \ell\},\\ \mathcal{B}^{(1)}_\ell &:= \{B \in \mathcal{B}_\ell : |i-j| \geq 2 \text{ for all } i, j \in B \text{ with } i \neq j\}\end{aligned}$$

and $\mathcal{B}^{(2)}_\ell := \mathcal{B}_\ell \setminus \mathcal{B}^{(1)}_\ell$. Then according to the decompositions $\mathcal{B}_\ell = \mathcal{B}^{(1)}_\ell \cup \mathcal{B}^{(2)}_\ell$ and $\mathcal{A}^{(k)} = \mathcal{A}^{(k)}_x \cup [\mathcal{A}^{(k)} \setminus \mathcal{A}^{(k)}_x]$, we have

$$\begin{aligned}\{W_n(x) = \ell, \overrightarrow{\mathcal{V}}_n(x) \in \mathcal{A}^{(k)}_x\} &= \{\mathcal{W}_n(x) \in \mathcal{B}^{(1)}_\ell, \overrightarrow{\mathcal{V}}_n(x) \in \mathcal{A}^{(k)}_x\},\\ \{W_n(x) = \ell, \overrightarrow{\mathcal{V}}_n(x) \in \mathcal{A}^{(k)} \setminus \mathcal{A}^{(k)}_x\} &= \{\mathcal{W}_n(x) \in \mathcal{B}^{(2)}_\ell, \overrightarrow{\mathcal{V}}_n(x) \in \mathcal{A}^{(k)}\}\\ &\quad \cup \{\mathcal{W}_n(x) \in \mathcal{B}^{(1)}_\ell, \xi_{n-1} = \xi_n = x, \overrightarrow{\mathcal{V}}_n(x) \in \mathcal{A}^{(k)}\}.\end{aligned}$$

So we can write $\mathbb{P}(W_n(x) = \ell, \#\overrightarrow{\mathcal{V}}_n(x) = k) = I_1(x) + I_2(x)$ with

$$\begin{aligned}I_1(x) &:= \mathbb{P}(\mathcal{W}_n(x) \in \mathcal{B}^{(1)}_\ell, \overrightarrow{\mathcal{V}}_n(x) \in \mathcal{A}^{(k)}_x),\\ I_2(x) &:= \mathbb{P}(\mathcal{W}_n(x) \in \mathcal{B}^{(2)}_\ell, \overrightarrow{\mathcal{V}}_n(x) \in \mathcal{A}^{(k)}) + \mathbb{P}(\mathcal{W}_n(x) \in \mathcal{B}^{(1)}_\ell, \xi_{n-1} = \xi_n = x, \overrightarrow{\mathcal{V}}_n(x) \in \mathcal{A}^{(k)}).\end{aligned}$$

Then we have $\mathbb{E}\overrightarrow{R}_{n,\,k,\,\ell} = \sum_x I_1(x) + \sum_x I_2(x)$.

Since $\#\mathcal{B}^{(1)}_\ell = C^\ell_{n-\ell} \leq n^\ell$ and $\#\mathcal{B}^{(2)}_\ell = C^\ell_{n-1} - C^\ell_{n-\ell} \leq \lambda_\ell \cdot n^{\ell-1}$ where λ_ℓ is a constant depending only on ℓ, we have

$$\begin{aligned}\sum_x I_2(x) &\leq \sum_x [\mathbb{P}(\mathcal{W}_n(x) \in \mathcal{B}^{(2)}_\ell) + \mathbb{P}(\mathcal{W}_n(x) \in \mathcal{B}^{(1)}_\ell, \xi_{n-1} = \xi_n = x)]\\ &\leq \#\mathcal{B}^{(2)}_\ell \cdot \sum_x \pi^\ell_x (1-\pi_x)^{n-1-\ell} + \#\mathcal{B}^{(1)}_\ell \cdot \sum_x \pi^{\ell+1}_x (1-\pi_x)^{n-1-\ell}\\ &\leq \lambda_\ell \cdot n^{\ell-1} S_\ell(n-1) + n^\ell S_{\ell+1}(n).\end{aligned}$$

By Lemma 3.4, we have $S_\ell(n-1) = O\big(\frac{\zeta(n)}{n^\ell}\big)$ and $S_{\ell+1}(n) = O\big(\frac{\zeta(n)}{n^{\ell+1}}\big)$. So

$$\sum_x I_2(x) = O\Big(\frac{\zeta(n)}{n}\Big).$$

Now we turn to estimate $\sum_x I_1(x)$. Let $B \in \mathcal{B}^{(1)}_\ell$. By our construction, $\mathcal{W}_n(x) = B$ and $\overrightarrow{\mathcal{V}}_n(x) \in \mathcal{A}^{(k)}_x$ are just to say: $\xi_i = x, \forall i \in B$ and $\{\xi_{i+1} : i \in B\} \in \mathcal{A}^{(k)}_x$, $\xi_j \neq x, \forall j \in \{i \leq n-1 : i \notin B, i-1 \notin B\}$. Note that $\{\xi_{i+1}, i \in B\}$ has the same distribution with $\{\xi_i, i \leq \ell\}$ and that $\#\{j \in [1, n-1]\cap \mathbb{N} : j \notin B, j-1 \notin B\} = n-1-2\ell+1_{\{n-1\in B\}}$. Hence we have

$$1 \geq \frac{\mathbb{P}(\mathcal{W}_n(x) = B,\ \overrightarrow{\mathcal{V}}_n(x) \in \mathcal{A}^{(k)}_x)}{\pi^\ell_x \cdot (1-\pi_x)^{n-1-2\ell} \cdot \mathbb{P}(\{\xi_i : i \leq \ell\} \in \mathcal{A}^{(k)}_x)} \geq 1-\pi_x. \tag{3.7}$$

Write $\Delta_x := \mathbb{P}(\{\xi_i : i \leq \ell\} \in \mathcal{A}^{(k)}) - \mathbb{P}(\{\xi_i : i \leq \ell\} \in \mathcal{A}_x^{(k)})$. Then by $\mathbb{P}(R_\ell = k) = \mathbb{P}(\{\xi_i : i \leq \ell\} \in \mathcal{A}^{(k)})$, we have

$$\mathbb{P}(\{\xi_i : i \leq \ell\} \in \mathcal{A}_x^{(k)}) = \mathbb{P}(R_\ell = k) - \Delta_x.$$

Furthermore, by $\#\mathcal{B}_\ell^{(1)} = C_{n-\ell}^\ell$ and the first inequality of (3.7), we get

$$\begin{aligned} I_1(x) &= \sum_{B \in \mathcal{B}_\ell^{(1)}} \mathbb{P}(\mathcal{W}_n(x) = B,\ \overrightarrow{\mathcal{V}}_n(x) \in \mathcal{A}_x^{(k)}) \\ &\leq C_{n-\ell}^\ell \cdot \pi_x^\ell (1-\pi_x)^{n-1-2\ell} \cdot \mathbb{P}(\{\xi_i : i \leq \ell\} \in \mathcal{A}_x^{(k)}) \\ &= C_{n-\ell}^\ell \cdot \pi_x^\ell (1-\pi_x)^{n-1-2\ell} [\mathbb{P}(R_\ell = k) - \Delta_x]. \end{aligned}$$

Since $0 \leq \Delta_x \leq \mathbb{P}(x \in \{\xi_i : i \leq \ell\}) \leq \ell \pi_x$, we have

$$0 \leq \sum_x C_{n-\ell}^\ell \cdot \pi_x^\ell (1-\pi_x)^{n-1-2\ell} \Delta_x \leq \ell C_{n-\ell}^\ell \sum_x \pi_x^{\ell+1} (1-\pi_x)^{n-1-2\ell}.$$

By Lemma 3.4, we have $\sum\limits_x \pi_x^{\ell+1}(1-\pi_x)^{n-1-2\ell} = O(\frac{\zeta(n)}{n^{\ell+1}})$. So

$$\sum_x C_{n-\ell}^\ell \cdot \pi_x^\ell (1-\pi_x)^{n-1-2\ell} \Delta_x = O\Big(\frac{\zeta(n)}{n}\Big).$$

On the other hand, applying Lemma 3.4 again, we get

$$\sum_x \pi_x^\ell (1-\pi_x)^{n-1-2\ell} = \gamma\Gamma(\ell-\gamma) \cdot \frac{\zeta(n)}{n^\ell} \cdot [1+o(1)].$$

Therefore,

$$\begin{aligned} \sum_x I_1(x) &\leq O\Big(\frac{\zeta(n)}{n}\Big) + \sum_x C_{n-\ell}^\ell \cdot \pi_x^\ell (1-\pi_x)^{n-1-2\ell} \mathbb{P}(R_\ell = k) \\ &\leq \frac{\gamma\Gamma(\ell-\gamma)}{\ell!} \cdot \mathbb{P}(R_\ell = k) \cdot \zeta(n) \cdot [1+o(1)]. \end{aligned}$$

In the same way, we can apply the second inequality of (3.7) to get

$$\sum_x I_1(x) \geq \frac{\gamma\Gamma(\ell-\gamma)}{\ell!} \cdot \mathbb{P}(R_\ell = k) \cdot \zeta(n) \cdot [1+o(1)].$$

Hence we finish the proof of the lemma.

Lemma 3.7 *For fixed* $1 \leq k \leq \ell$, $\mathrm{Var}\,(\overrightarrow{R}_{n,k,\ell}) \leq [1+o(1)] \cdot \mathbb{E}\overrightarrow{R}_{n,k,\ell}$.

Proof The proof is similar to that of Lemmas 3.5–3.6. Here we just give a sketched proof by outlining the main differences.

Let $x, y \in \mathbb{N}$ with $x \neq y$. Recall $\mathcal{A}^{(k)} := \{A \subset \mathbb{N} : |A| = k\}$ in the proof of Lemma 3.6. Define $\mathcal{A}_{x,y}^{(k)} := \{A \subset \mathbb{N} : |A| = k, \{x, y\} \cap A = \emptyset\}$. Write

$$\begin{aligned} \mathcal{D}_\ell &:= \{(B, C) : B \cup C \subset [1, n-1] \cap \mathbb{N}, B \cap C = \emptyset, |B| = |C| = \ell\}, \\ \mathcal{D}_\ell^{(1)} &:= \{(B, C) \in \mathcal{D}_\ell : |i - j| \geq 2 \text{ for all } i, j \in B \cup C \text{ with } i \neq j\} \end{aligned}$$

and $\mathcal{D}_\ell^{(2)} := \mathcal{D}_\ell \setminus \mathcal{D}_\ell^{(1)}$. Then we have

$$\mathbb{P}(W_n(x) = W_n(y) = \ell, \#\overrightarrow{\mathcal{V}}_n(x) = \#\overrightarrow{\mathcal{V}}_n(y) = k) = J_1(x, y) + J_2(x, y),$$

where

$$
\begin{aligned}
J_1(x,y) &= \mathbb{P}((\mathcal{W}_n(x),\mathcal{W}_n(y)) \in \mathcal{D}_\ell^{(1)}, \{\overrightarrow{\mathcal{V}}_n(x), \overrightarrow{\mathcal{V}}_n(y)\} \subset \mathcal{A}_{x,y}^{(k)}),\\
J_2(x,y) &= \mathbb{P}((\mathcal{W}_n(x),\mathcal{W}_n(y)) \in \mathcal{D}_\ell^{(2)}, \{\overrightarrow{\mathcal{V}}_n(x), \overrightarrow{\mathcal{V}}_n(y)\} \subset \mathcal{A}^{(k)})\\
&\qquad + \mathbb{P}((\mathcal{W}_n(x),\mathcal{W}_n(y)) \in \mathcal{D}_\ell^{(1)}, \{\xi_{n-1},\xi_n\} \subset \{x,y\}, \{\overrightarrow{\mathcal{V}}_n(x), \overrightarrow{\mathcal{V}}_n(y)\} \subset \mathcal{A}^{(k)}).
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
\mathbb{E}(\overrightarrow{R}^2_{n,\,k,\,\ell} - \overrightarrow{R}_{n,\,k,\,\ell}) &= \sum_{x\neq y} \mathbb{P}(W_n(x) = W_n(y) = \ell, \#\overrightarrow{\mathcal{V}}_n(x) = \#\overrightarrow{\mathcal{V}}_n(y) = k)\\
&= \sum_{x\neq y} J_1(x,y) + \sum_{x\neq y} J_2(x,y).
\end{aligned}
$$

A direct calculation gives

$$
\#\mathcal{D}_\ell^{(1)} = C_{n-2\ell}^{2\ell} C_{2\ell}^{\ell} = \Big[1 + O\Big(\frac{1}{n}\Big)\Big] \cdot \frac{n^{2\ell}}{(\ell!)^2} \quad \text{and} \quad \#\mathcal{D}_\ell^{(2)} = C_{n-1}^{2\ell} C_{2\ell}^{\ell} - C_{n-2\ell}^{2\ell} C_{2\ell}^{\ell} \le \lambda_\ell n^{2\ell-1},
$$

where $\lambda_\ell > 0$ is a constant depending only on ℓ. So in view of Lemmas 3.4 and 3.6,

$$
\begin{aligned}
\sum_{x\neq y} J_2(x,y) &\le \#\mathcal{D}_\ell^{(2)} \cdot \sum_{x\neq y} \pi_x^\ell \pi_y^\ell (1-\pi_x-\pi_y)^{n-1-2\ell}\\
&\qquad + \#\mathcal{D}_\ell^{(1)} \cdot \sum_{x\neq y} \pi_x^\ell \pi_y^\ell (\pi_x+\pi_y)(1-\pi_x-\pi_y)^{n-1-2\ell}\\
&\le \lambda_\ell n^{2\ell-1} \cdot S_\ell(n-1-\ell)^2 + 2n^{2\ell} \cdot S_{\ell+1}(n-\ell) S_\ell(n-1-\ell)\\
&\le O\Big(\frac{\zeta(n)^2}{n}\Big) = o(1) \cdot \mathbb{E}\overrightarrow{R}_{n,\,k,\,\ell}.
\end{aligned}
$$

Since

$$
0 \le \mathbb{P}(\{\xi_i : i \le \ell\} \in \mathcal{A}^{(k)}) - \mathbb{P}(\{\xi_i : i \le \ell\} \in \mathcal{A}_{x,y}^{(k)}) \le \ell(\pi_x + \pi_y),
$$

we write $\Delta_{x,y} := \mathbb{P}(\{\xi_i : i \le \ell\} \in \mathcal{A}^{(k)})^2 - \mathbb{P}(\{\xi_i : i \le \ell\} \in \mathcal{A}_{x,y}^{(k)})^2$ and get

$$
\begin{aligned}
0 \le \sum_{x\neq y} \pi_x^\ell \pi_y^\ell (1-\pi_x-\pi_y)^{n-1-4\ell} \Delta_{x,y} &\le \sum_{x\neq y} \pi_x^\ell \pi_y^\ell (1-\pi_x-\pi_y)^{n-1-4\ell} \cdot 2\ell(\pi_x+\pi_y)\\
&\le 4\ell S_{\ell+1}(n-3\ell) S_\ell(n-1-3\ell)\\
&= O\Big(\frac{\zeta(n)^2}{n^{2\ell+1}}\Big).
\end{aligned}
$$

So, we can apply Lemmas 3.4 and 3.6 again to get

$$
\begin{aligned}
\sum_{x\neq y} J_1(x,y) &\le \#\mathcal{D}_\ell^{(1)} \cdot \sum_{x\neq y} \pi_x^\ell \pi_y^\ell (1-\pi_x-\pi_y)^{n-1-4\ell} [\mathbb{P}(\{\xi_i : i \le \ell\} \in \mathcal{A}^{(k)})^2 - \Delta_{x,y}]\\
&= \frac{1 + O(n^{-1})}{(\ell!)^2} n^{2\ell} \cdot \Big[S_\ell(n-1-3\ell)^2 \cdot \mathbb{P}(R_\ell = k)^2 + O\Big(\frac{\zeta(n)^2}{n^{2\ell+1}}\Big)\Big]\\
&= [1 + o(1)](\mathbb{E}\overrightarrow{R}_{n,\,k,\,\ell})^2,
\end{aligned}
$$

and

$$
\sum_{x\neq y} J_1(x,y) \ge \#\mathcal{D}_\ell^{(1)} \cdot \sum_{x\neq y} \pi_x^\ell \pi_y^\ell (1-\pi_x-\pi_y)^{n-4\ell} [\mathbb{P}(\{\xi_i : i \le \ell\} \in \mathcal{A}^{(k)})^2 - \Delta_{x,y}]
$$

$$= [1+o(1)](\mathbb{E}\overrightarrow{R}_{n,\,k,\,\ell})^2.$$

Therefore, we have $\mathbb{E}(\overrightarrow{R}^2_{n,\,k,\,\ell} - \overrightarrow{R}_{n,\,k,\,\ell}) = [1+o(1)](\mathbb{E}\overrightarrow{R}_{n,\,k,\,\ell})^2$ and complete the proof of the lemma.

4 Proofs for the Main Theorems 2.1–2.4

First we state the following lemma. It plays an important role in the proof of the main theorems.

Lemma 4.1 *Let S_n be a sequence of non-negative increasing random variables. Suppose $\mathbb{E}S_n \to +\infty$ and $M := \sup\{\mathbb{E}(S_{n+1}-S_n) : n \ge 1\} < +\infty$. Suppose that there exist constants $C, \delta > 0$ such that*

$$\mathrm{Var}\,(S_n) \le C \cdot (\mathbb{E}S_n)^{2-\delta}, \quad n \ge 1 \tag{4.1}$$

or even more weakly

$$\mathrm{Var}\,(S_n) \le \frac{C \cdot (\mathbb{E}S_n)^2}{(\log \mathbb{E}S_n)^{1+\delta}}, \quad n \ge 1. \tag{4.2}$$

Then $\lim\limits_{n\to\infty} \frac{S_n}{\mathbb{E}S_n} = 1$ almost surely.

Since the proof of the above lemma is in fact contained implicitly in [13] and is indeed an easy application of Borel-Cantelli lemma, we omit the details here.

Now we present the proofs of Theorems 2.1–2.4. The main idea is to exploit Lemma 4.1. We split the proofs into 5 parts.

(A) By the construction of the range-renewal process $(R_n : n \ge 1)$, we have $0 \le R_{n+1} - R_n \le 1$ for all $n \ge 1$. Using Lemmas 3.1 and 4.1, we prove Theorem 2.1.

(B) Since $R_{n,k}$ is not increasing in n, we turn to $R_{n,k+}$. By Lemmas 3.3–3.4,

$$\mathbb{E}R_{n,k+} = \mathbb{E}R_n - \sum_{\ell=1}^{k-1} \mathbb{E}R_{n,\ell} = \big[1+o(1)\big] \cdot \frac{\Gamma(k-\gamma)}{(k-1)!} \cdot \zeta(n).$$

In view of Cauchy's inequality and Lemmas 3.1, 3.3–3.4 and 3.7, we have

$$\begin{aligned}\mathrm{Var}\,(R_{n,\,k+}) &= \mathrm{Var}\,\Big(R_n - \sum_{\ell=1}^{k-1} R_{n,\,\ell}\Big) \le k \cdot \Big[\mathrm{Var}\,(R_n) + \sum_{\ell=1}^{k-1} \mathrm{Var}\,(R_{n,\,\ell})\Big] \\ &\le k \cdot \Big[\mathbb{E}R_n + [1+o(1)] \cdot \sum_{\ell=1}^{k-1} \mathbb{E}R_{n,\,\ell}\Big] = O(\zeta(n)) = O(\mathbb{E}R_{n,k+}).\end{aligned} \tag{4.3}$$

So $\frac{R_{n,k+}}{\mathbb{E}R_{n,k+}} \xrightarrow[n\to\infty]{\text{a.s.}} 1$ by Lemma 4.1. This proves the first result of Theorem 2.2.

(C) By Lemmas 3.3–3.4, $\lim\limits_{n\to\infty} \frac{\mathbb{E}R_{n,\,k}}{\mathbb{E}R_n} = r_k(\gamma)$ and $\lim\limits_{n\to\infty} \frac{\mathbb{E}R_{n,\,k+}}{\mathbb{E}R_n} = r_{k+}(\gamma) := \sum\limits_{\ell \ge k} r_\ell(\gamma)$. Combining this with the SLLNs for R_n and $R_{n,k+}$, we have

$$\lim_{n\to\infty} \frac{R_{n,k}}{R_n} = \lim_{n\to\infty} \frac{R_{n,k+} - R_{n,(k+1)+}}{R_n} = \lim_{n\to\infty} \frac{\mathbb{E}R_{n,k+}}{\mathbb{E}R_n} - \lim_{n\to\infty} \frac{\mathbb{E}R_{n,(k+1)+}}{\mathbb{E}R_n} = r_k(\gamma)$$

almost surely, which implies Theorem 2.3 immediately.

(D) Write $\overrightarrow{R}_{n,\,k+,\,\ell+} := R_{n-1} - \sum\limits_{i<k}\sum\limits_{j<\ell} \overrightarrow{R}_{n,\,i,\,j}$. Then using Lemmas 3.6–3.7, we can get $\mathrm{Var}\,(\overrightarrow{R}_{n,\,k+,\,\ell+}) = O(\mathbb{E}\overrightarrow{R}_{n,\,k+,\,\ell+})$ as (4.3). Hence $\lim\limits_{n\to\infty} \frac{\overrightarrow{R}_{n,\,k+,\,\ell+}}{\mathbb{E}\overrightarrow{R}_{n,\,k+,\,\ell+}} = 1$ almost surely by Lemma 4.1. Combining this with Lemmas 3.3 and 3.6 and noting

$$\overrightarrow{R}_{n,\,k,\,\ell} = \overrightarrow{R}_{n,\,k+,\,\ell+} - \overrightarrow{R}_{n,\,(k+1)+,\,\ell+} - \overrightarrow{R}_{n,\,k+,\,(\ell-1)+} + \overrightarrow{R}_{n,\,(k+1)+,\,(\ell-1)+},$$

we obtain almost surely

$$\lim_{n\to\infty} \frac{\overrightarrow{R}_{n,\,k,\,\ell}}{R_{n-1}} = \lim_{n\to\infty} \frac{\mathbb{E}\overrightarrow{R}_{n,\,k,\,\ell}}{\mathbb{E}R_{n-1}} = r_\ell(\gamma)\cdot\mathbb{P}(R_\ell = k).$$

Note that $\sum\limits_{k=1}^{\infty}\sum\limits_{\ell=k}^{\infty} \frac{\overrightarrow{R}_{n,\,k,\,\ell}}{R_{n-1}} = 1$ and $\sum\limits_{k=1}^{\infty}\sum\limits_{\ell=k}^{\infty} r_\ell(\gamma)\cdot\mathbb{P}(R_\ell = k) = 1$. By Scheffés' theorem (see [12, pp. 101]), we have almost surely

$$\lim_{n\to\infty} \frac{\overrightarrow{R}_{n,k}}{R_n} = \lim_{n\to\infty} \sum_{\ell=k}^{\infty} \frac{\overrightarrow{R}_{n,\,k,\,\ell}}{R_{n-1}} = \sum_{\ell=k}^{\infty} \lim_{n\to\infty} \frac{\overrightarrow{R}_{n,\,k,\,\ell}}{R_{n-1}} = \sum_{\ell=k}^{\infty} r_\ell(\gamma)\cdot\mathbb{P}(R_\ell = k),$$

which implies the first result of Theorem 2.4. Then we can establish an SLLN of Dvoretzky-Erdös' type for $(\overrightarrow{R}_{n,k}, n \geq 1)$, and prove the second statement of Theorem 2.2.

(E) The proof for the results of the induced undirected graph $\overline{G}_n$ is omitted here since it is rather similar to the directed graph case.

By **(A)**–**(E)**, we finish the proofs of Theorems 2.1–2.4.

5 Discussions

In the first two subsections, we will discuss the critical cases of the distribution π being 0-regular and 1-regular respectively. Then in the last subsection, we will propose a conjecture concerning the asymptotic degree distributions of our unweighted random graphs in the case of π being γ-regular with $\gamma \in (0,1)$.

Recall (3.4) for the definition of $S_\ell(n)$. Similarly we have the following lemma.

Lemma 5.1 *Let $\ell, d \in \mathbb{N}$ be given.*
(1) *If $\gamma = 0$ then $\mathbb{E}R_n = \zeta(n)[1+o(1)]$ and $S_\ell(n) = \frac{(\ell-1)!}{n^{\ell-1}}\cdot\zeta'(n)\cdot[1+o(1)]$.*
(2) *If $\gamma = 1$, letting g and ψ be as introduced in Definition 2.4, then*

$$\begin{aligned}
\mathbb{E}R_n &= \|g\|_1\cdot\zeta(n)\cdot\psi(n)\cdot[1+o(1)],\\
S_1(n) &= \frac{\|g\|_1}{n}\cdot\zeta(n)\cdot\psi(n)\cdot[1+o(1)],\\
S_\ell(n) &= \frac{(\ell-2)!}{n^\ell}\cdot\zeta(n)\cdot[1+o(1)], \quad \ell \geq 2.
\end{aligned}$$

(3) *For $\gamma = 0$ or 1, we always have $S_\ell(n-d) = S_\ell(n)\cdot\left[1+O\left(\frac{1}{n}\right)\right]$.*

When $\gamma = 1$, we have $\mathbb{E}R_n = \|g\|_1\cdot\zeta(n)\cdot\psi(n)\cdot[1+o(1)]$. Here $\lim\limits_{n} \frac{\log\zeta(n)}{\log n} = 1$ and $\lim\limits_{n} \frac{\mathbb{E}R_n}{n} = 0$. Therefore, in this critical case, $\lim\limits_{n\to\infty} \frac{\log\psi(n)}{\log n} = 0$.

It is also notable that Lemmas 3.5 and 3.7 still hold for the critical cases $\gamma \in \{0,1\}$.

5.1 0-Regular case: $\gamma(\pi) = 0$

In the same spirit, we can prove the following SLLNs:

$$\lim_{n\to\infty} \frac{R_{n,\,\ell+}}{\mathbb{E}R_{n,\,\ell+}} = 1 \quad \text{and} \quad \lim_{n\to\infty} \frac{\overrightarrow{R}_{n,\,k+,\,\ell+}}{\mathbb{E}\overrightarrow{R}_{n,\,k+,\,\ell+}} = 1.$$

A careful calculation reveals that

$$\mathbb{E}R_{n,\,\ell+} = [1 + o(1)] \cdot \mathbb{E}R_n \quad \text{and} \quad \mathbb{E}\overrightarrow{R}_{n,\,k+,\,\ell+} = [1 + o(1)] \cdot \mathbb{E}R_n,$$

which implies $\lim\limits_{n\to\infty} \frac{R_{n,\,\ell+}}{R_n} = 1$ and $\lim\limits_{n\to\infty} \frac{\overrightarrow{R}_{n,\,k+,\,\ell+}}{R_n} = 1$. Hence almost surely we have for any fixed $\ell \in \mathbb{N}$,

$$\lim_{n\to\infty} \frac{R_{n,\,\ell}}{R_n} = 0, \quad \lim_{n\to\infty} \frac{\overrightarrow{R}_{n,\,\ell}}{R_n} = 0 \quad \text{and} \quad \lim_{n\to\infty} \frac{\overline{R}_{n,\,\ell}}{R_n} = 0.$$

5.2 1-Regular case: $\gamma(\pi) = 1$

Since $\mathbb{E}R_n = O(\zeta(n) \cdot \psi(n))$ and $O(\zeta(n)) = O(\mathbb{E}R_{n,\,\ell+})$ for $\ell \geq 2$, we have

$$\lim_{n\to\infty} \frac{\log \mathbb{E}R_n}{\log n} = 1, \tag{5.1}$$

$$\lim_{n\to\infty} \frac{\log \mathbb{E}R_{n,\,\ell+}}{\log n} = 1, \quad \ell \geq 2. \tag{5.2}$$

Therefore, we can derive easily that $\mathrm{Var}\,(R_{n,\,\ell+}) \leq C \cdot (\mathbb{E}R_{n,\,\ell+})^{\frac{3}{2}}$ for some constant $C > 0$ (depending possibly on $\ell \geq 2$), which yields an SLLN for $R_{n,\,\ell+}$ by our Lemma 4.1. In the same spirit, we have also an SLLN for $\overrightarrow{R}_{n,\,k+,\,\ell+}$.

Now a detailed calculation of $\mathbb{E}R_{n,\,\ell+}$ and $\mathbb{E}\overrightarrow{R}_{n,\,k+,\,\ell+}$ reveals the following result:

$$\lim_{n\to\infty} \frac{R_{n,\,1}}{R_n} = 1 \quad \text{and} \quad \lim_{n\to\infty} \frac{R_{n,\,\ell}}{R_n} = 0, \quad \ell \geq 2.$$

And a re-scaling yields

$$\lim_{n\to\infty} \frac{R_{n,\,\ell}}{R_{n,2+}} = \frac{1}{\ell \cdot (\ell - 1)}, \quad \ell \geq 2. \tag{5.3}$$

Furthermore, for any $k \geq 2$,

$$\lim_{n\to\infty} \frac{\overrightarrow{R}_{n,\,k}}{R_{n,\,2+}} = \overrightarrow{r}_k(\pi) := \sum_{\ell=k}^{\infty} \frac{\mathbb{P}(R_\ell = k)}{\ell \cdot (\ell - 1)},$$

$$\lim_{n\to\infty} \frac{\overline{R}_{n,\,k}}{R_{n,\,2+}} = \overline{r}_k(\pi) := \sum_{\ell=\left\lfloor \frac{k+1}{2} \right\rfloor}^{\infty} \frac{\mathbb{P}(R_{2\ell} = k)}{\ell \cdot (\ell - 1)}.$$

5.3 A conjecture

As mentioned in Section 2, we conjecture here that, when π is γ-regular with $\gamma \in (0, 1)$, the out degrees of $\overrightarrow{G}_n$ (and of $\overline{G}_n$ respectively) exhibit asymptotically power law distributions with exponent 2. More precisely, $\{\overrightarrow{r}_k(\pi)\}_{k=1}^{\infty}$ and $\{\overline{r}_k(\pi)\}_{k=1}^{\infty}$ should be $\frac{1}{2}$-regular and

$$\lim_{k\to\infty} \frac{\log \overrightarrow{r}_k(\pi)}{\log(\frac{1}{k})} = \lim_{k\to\infty} \frac{\log \overline{r}_k(\pi)}{\log(\frac{1}{k})} = 2, \tag{5.4}$$

whence the graphs $\overrightarrow{G}_n$ and $\overline{G}_n$ are asymptotically scale-free (see [30, pp. 12]). A heuristic deduction is as the following. Define for each $k \geq 1$,

$$T_k := \inf\{n \geq 1 : R_n = k\}. \tag{5.5}$$

Clearly $R_n < k$ if and only if $T_k > n$; here we always have $T_k > k-1$. Hence

$$\begin{aligned}\overrightarrow{r}_k(\pi) &= \sum_{\ell=k}^{\infty} r_\ell(\gamma) \cdot \mathbb{P}(R_\ell = k) = \sum_{\ell=k}^{\infty} r_\ell(\gamma) \cdot \mathbb{E}[1_{\{T_k \leq \ell < T_{k+1}\}}] \\ &= \mathbb{E} \sum_{\ell=k}^{\infty} r_\ell(\gamma) \cdot 1_{\{T_k \leq \ell < T_{k+1}\}} = \mathbb{E}[r_{{T_k}+}(\gamma) - r_{{T_{k+1}}+}(\gamma)].\end{aligned}$$

Therefore

$$\overrightarrow{r}_k(\pi) = \mathbb{E}[r_{{T_k}+}(\gamma)] - \mathbb{E}[r_{{T_{k+1}}+}(\gamma)]. \tag{5.6}$$

Now we assume $\pi_x = \frac{C}{x^\alpha}, x \in \mathbb{N}$ for simplicity, where $\alpha = \frac{1}{\gamma}$. We know that R_n is of order n^γ as $n \to \infty$; it is natural to believe that T_k should be of order k^α as $k \to \infty$; and the difference $\Delta T_k := T_{k+1} - T_k$ should be of the same order as $T_k^{1-\gamma}$, which comes from the following observation: $(n+d)^\gamma - n^\gamma = 1$ implies $d \approx \alpha \cdot n^{1-\gamma}$ as $n \to \infty$. It is easy to know

$$r_{\ell+} = \frac{\prod_{j=1}^{\ell-1}(j-\gamma)}{(\ell-1)!} = O(\ell^{-\gamma})$$

for large enough ℓ. And

$$T_k^{-\gamma} - T_{k+1}^{-\gamma} = T_k^{-\gamma} \cdot \left[1 - \left(1 + \frac{\Delta T_k}{T_k}\right)^{-\gamma}\right] \approx \gamma \cdot T_k^{-\gamma} \cdot \left(\frac{\Delta T_k}{T_k}\right)$$

would be of the same order as $T_k^{-2\gamma}$. Whence $\overrightarrow{r}_k(\pi)$ should be of order $(k^\alpha)^{-2\gamma} = k^{-2}$ as $k \to \infty$. The same heuristic deduction also works for $\overline{r}_k(\pi)$.

Acknowledgements We are in debt to the anonymous referee for his valuable suggestions which make the current version of our manuscript more readable. The second author would like to thank Prof. De-Jun Feng and his colleagues Professors Yi-Jun Yao and Yun-Xin Zhang for helpful discussions and comments. He also expresses the gratitude to Prof. Derriennic for sending him a copy of [11].

References

[1] Adamic, L. A. and Huberman, B. A., Power-law distribution of the world wide web, *Science*, **287**, 2000, 2115.

[2] Albert, R. and Barabási, A.-L., Statistical mechanics of complex networks, *Rev. Mod. Phys.*, **74**, 2002, 47–97.

[3] Athreya, K. B., On the range of recurrent Markov chains, *Statist. Probab. Lett.*, **3**, 1985, 143–145.

[4] Bahadur, R. R., On the number of distinct values in a large sample from an infinite discrete distribution, *Proc. Nat. Inst. Sci. India Part A*, **26**, 1960, supplement II, 67–75.

[5] Barabási, A.-L., Scale-free networks: A decade and beyond, *Science*, **325**, 2009, 412.

[6] Barabási, A.-L. and Albert, R., Emergence of scaling in random networks, *Science*, **286**, 1999, 509.

[7] Bass, R. F. and Kumagai, T., Laws of the iterated logarithm for the range of random walks in two and three dimensions, *Ann. Probab.*, **30**, 2002, 1369–1396.

[8] Bingham, N. H., Goldie, C. M. and Teugels, J. L., Regular Variation, Encyclopedia of Mathematics and its Applications, 27, Cambridge University Press, Cambridge, 1987.

[9] Chosid, L. and Isaac, R., On the range of recurrent Markov chains, *Ann. Probab.*, **6**, 1978, 680–687.

[10] Chosid, L. and Isaac, R., Correction to: "On the range of recurrent Markov chains", [Ann. Probab. **6**(4), 1978, 680–687; MR 57 #14146], *Ann. Probab.*, **8**(5), 1980, 1000.

[11] Derriennic, Y., Quelques applications du théorème ergodique sous-additif, (French. English summary) Conference on Random Walks (Kleebach, 1979) (French), 183–201, **4**, *Astérisque*, **74**, 1980, Soc. Math. France, Paris.

[12] Durrett, R., Probability: Theory and Examples, 5th ed., Cambridge University Press, Cambridge, 2019.

[13] Dvoretzky, A. and Erdös, P., Some problems on random walk in space, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950, 353–367, University of California Press, Berkeley and Los Angeles, 1951.

[14] Erdös, P. and Taylor, S. J., Some problems concerning the structure of random walk paths, *Acta Math. Acad. Sci. Hungar.*, **11**, 1960, 137–162.

[15] Feller, W., An Introduction to Probability Theory and Its Applications, 2nd ed., **2**, Wiley Publishing, Inc. (Chinese translation edition, Posts & Telecom Press, 2008.)

[16] Glynn, P. W., On the range of a regenerative sequence, *Stoch. Proc. Appl.*, **20**, 1985, 105–113.

[17] Glynn, P. W. and Sigman, K., Uniform cesàro limit theorems for synchronous processes with applications to queues, *Stoch. Proc. Appl.*, **40**, 1992, 29–43.

[18] Hamana, Y., The fluctuation result for the multiple point range of two-dimensional recurrent random walks, *Ann. Probab.*, **25**, 1997, 598–639.

[19] Hamana, Y., An almost sure invariance principle for the range of random walks, *Stoch. Proc. Appl.,* **78**, 1998, 131–143.

[20] Iosifescu, M. and Kraaikamp, C., Metrical Theory of Continued Fractions, Mathematics and Its Applications, **547**, Kluwer Academic Publishers, Dordrecht, 2002.

[21] Jain, N. C. and Pruitt, W. E., The range of transient random walk, *J. Anal. Math.*, **24**, 1971, 369–393.

[22] Jain, N. C. and Pruitt, W. E., The law of the iterated logarithm for the range of random walk, *Ann. Math. Statist.*, **43**, 1972, 1692–1697.

[23] Jain, N. C. and Pruitt, W. E., Further limit theorems for the range of random walk, *J. Anal. Math.*, **27**, 1974, 94–117.

[24] Karamata, J., Sur un mode de croissance régulière, Théorèmes fondamentaux, (French) *Bull. Soc. Math. France*, **61**, 1933, 55–62.

[25] Le Gall, J.-F., Propriétés d'intersection des marches aléatoires, I, Convergence vers le temps local d'intersection, *Comm. Math. Phys.*, **104**, 1986, 471–507.

[26] Révész, P., Random Walk in Random and Non-Random Environments, 2nd ed., World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2005, xvi+380 pp.

[27] Salminen, P. and Vallois, P., On first range times of linear diffusions., *J. Theoret. Probab.*, **18**, 2005, 567–593.

[28] Vallois, P., The range of a simple random walk on $\mathbb{Z}$, *Adv. in Appl. Probab.*, **28**, 1996, 1014–1033.

[29] Vallois, P. and Tapiero, C. S., Range reliability in random walks, *Math. Methods Oper. Res.*, **45**, 1997, 325–345.

[30] Hofstad, R., Random Graphs and Complex Networks (Cambridge Series in Statistical and Probabilistic Mathematics), Cambridge University Press, Cambridge, 2016.