Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications xxx (xxxx) xxx

stochastic processes and their applications

www.elsevier.com/locate/spa

PageRank on inhomogeneous random digraphs

Jiung Lee[a], Mariana Olvera-Cravioto[b,*]

[a] Department of Industrial Engineering and Operations Research, University of California, Berkeley, United States
[b] Department of Statistics and Operations Research, University of North Carolina at Chapel Hill, United States

Received 9 July 2017; received in revised form 9 January 2019; accepted 3 July 2019
Available online xxxx

Abstract

We study the typical behavior of Google's PageRank algorithm on inhomogeneous random digraphs, including directed versions of the Erdős–Rényi model, the Chung–Lu model, the Poissonian random graph and the generalized random graph. Specifically, we show that the rank of a randomly chosen vertex converges weakly to the attracting endogenous solution to the stochastic fixed-point equation

$$\mathcal{R} \stackrel{\mathcal{D}}{=} \sum_{i=1}^{\mathcal{N}} \mathcal{C}_i \mathcal{R}_i + \mathcal{Q},$$

where $(\mathcal{N}, \mathcal{Q}, \{\mathcal{C}_i\}_{i\geq 1})$ is a real-valued vector with $\mathcal{N} \in \mathbb{N}$, and the $\{\mathcal{R}_i\}$ are i.i.d. copies of $\mathcal{R}$, independent of $(\mathcal{N}, \mathcal{Q}, \{\mathcal{C}_i\}_{i\geq 1})$;

$\stackrel{\mathcal{D}}{=}$ denotes equality in distribution. This result provides further evidence of the power-law behavior of PageRank on graphs whose in-degree distribution follows a power law.

MSC: 05C80; 60J80; 68P20; 41A60; 37A30; 60B10

Keywords: PageRank; Ranking algorithms; Directed random graphs; Weighted branching processes; Stochastic fixed-point equations; Power laws

1. Introduction

In the recent decades, a growing amount of data and computer power has motivated the development of algorithms capable of efficiently organizing and analyzing large data sets. In many cases, this data is highly interconnected, and can be represented in the form

* Corresponding author.
E-mail address: molvera@email.unc.edu (M. Olvera-Cravioto).

https://doi.org/10.1016/j.spa.2019.07.002
0304-4149/

of complex networks. Some important examples include the Internet and the World Wide Web, telecommunication networks, electrical power grids, protein–protein interactions, and the various social networks that have become an integral part of our society. Interestingly, many of these networks share some basic characteristics that we have learned to expect, such as short typical distances between nodes, known as the *small-world property*, and highly variable degrees whose distributions follow a power-law, known as the *scale-free* property. Of special interest is the problem of identifying relevant or central nodes in these networks.

We focus on the analysis of a general form of Google's PageRank algorithm [6], which was originally created to rank webpages in the World Wide Web. PageRank is a popular algorithm for ranking nodes in complex networks due to its ability to efficiently identify important/relevant nodes. Its typical behavior on scale-free directed complex networks has also been an important research topic, since abundant empirical evidence suggests that the distribution of the ranks produced by PageRank follows a power-law distribution with the same tail index as the in-degree distribution [2,23,27,30,34,35]. The first rigorous proof of why this power-law behavior is observed was given in [8], where it was shown that the rank of a randomly chosen node in a graph generated via the directed configuration model [9] has a limiting distribution exhibiting power-law tails whenever the in-degree distribution is scale-free. Here, we extend this analysis to a different class of random graph models that includes as special cases directed versions of classical models such as the Erdős–Rényi graph [1,4,15,16,18,22], the Chung–Lu model [11–14,28], the Poissonian random graph [17,20,29] and the generalized random graph [7,17,20]. Since the scale-free property of the degrees in these models (i.e., their power-law behavior) is due to node-specific attributes which can also be used to influence the rankings produced by PageRank, we believe they provide a more natural way of modeling and understanding the behavior of ranking algorithms on complex networks than the directed configuration model.

As is the case for the directed configuration model, the power-law behavior of the ranks produced by PageRank can be explained by arguing that the limiting distribution of the rank of a randomly chosen node can be written in terms of the attracting endogenous solution to a stochastic-fixed point equation (SFPE) of the form

$$\mathcal{R} \overset{\mathcal{D}}{=} \sum_{i=1}^{\mathcal{N}} \mathcal{C}_i \mathcal{R}_i + \mathcal{Q},$$

where the $\{\mathcal{R}_i\}$ are i.i.d. copies of $\mathcal{R}$, independent of the vector $(\mathcal{N}, \mathcal{Q}, \{\mathcal{C}_i\}_{i\geq 1})$, with the $\{\mathcal{C}_i\}$ i.i.d. and independent of $(\mathcal{N}, \mathcal{Q})$. The random variable $\mathcal{N}$ corresponds to the in-degree distribution of the network being analyzed, and the power law behavior of the rank distribution follows from the known asymptotic equivalence

$$P(\mathcal{R} > x) \sim K\, P(\mathcal{N} > x), \qquad x \to \infty,$$

for some constant $0 < K < \infty$, when $\mathcal{N}$ has a power-law distribution [23,34]. Theorem 3.3 in this paper shows that the same type of representation holds for the family of inhomogeneous random digraphs studied here, provided the in-degree and out-degree of the same vertex are asymptotically independent. Therefore, the results in this paper provide further evidence of the power-law behavior of PageRank in scale-free directed networks.

The remainder of the paper is organized as follows. Section 2 describes the family of inhomogeneous random digraphs mentioned above, and includes some of its most basic properties, in particular, its ability to generate inhomogeneous directed graphs with a wide range of degree distributions, including scale-free ones. In Section 3 we state our main result

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

on the distribution of the ranks produced by a generalized form of PageRank, including the description of a three-step approach towards its proof, and in Section 4 we provide all the proofs.

2. A family of inhomogeneous random digraphs

As mentioned in the introduction, the fact that many of the complex networks in the real world exhibit highly variable degrees, often with tails that appear to follow a power law, motivates our interest in random graph models capable of generating inhomogeneous degrees. One model that produces graphs from any prescribed (graphical) degree sequence is the configuration or pairing model [3,20], which assigns to each vertex in the graph a number of half-edges equal to its target degree and then randomly pairs half-edges to connect vertices. The resulting graph, when the pairing process does not create self-loops or multiple edges, is known to have the distribution of a uniformly chosen graph among all graphs having the prescribed degree sequence. If one chooses this degree sequence according to a power-law, one immediately obtains a scale-free graph.

Alternatively, one could think of obtaining the scale-free property (power-law degree distribution) as a consequence of how likely different nodes are to have an edge between them. In the spirit of the classical Erdős–Rényi graph [1,4,15,16,18,22], we assume that whether there is an edge between vertices i and j is determined by a coin-flip, independently of all other edges. Unfortunately, this elegant and simple rule is known to produce highly homogeneous degrees, Poisson distributed in the limit, making it inappropriate for modeling most real-world networks. Several models capable of producing graphs with inhomogeneous degrees while preserving the independence among edges have been suggested in the recent literature, including: the Chung–Lu model [11–14,28], the Norros–Reittu model (or Poissonian random graph) [17,20,29], and the generalized random graph [7,17,20], to name a few. In all of these models, the inhomogeneity of the degrees is created by allowing the success probability of each coin-flip to depend on the "attributes" of the two vertices being connected; the scale-free property can then be obtained by choosing the attributes according to a power-law. We briefly mention that it was shown in [17] that all these models also exhibit the small-world property, i.e., small typical distances between vertices, hence, we expect the same to be true of their directed counterparts.

We now give a precise description of the family of directed random graphs that we study in this paper, which includes as special cases the directed versions of all the models mentioned above. Throughout the paper we refer to a directed graph $\mathcal{G}(V_n, E_n)$ on the vertex set $V_n = \{1, 2, \ldots, n\}$ simply as a *random digraph* if the event that edge (i, j) belongs to the set of edges E_n is independent of all other edges.

In order to obtain inhomogeneous degree distributions, to each vertex $i \in V_n$ we assign a *type* $\mathbf{W}_i = (W_i^-, W_i^+) \in \mathbb{R}_+^2$, which will be used to determine how likely vertex i is to have inbound/outbound neighbors.[1] The sequence of types $\{\mathbf{W}_i : i \geq 1\}$ is assumed to have a limiting behavior, in the sense that the empirical joint distribution satisfies:

$$F_n(u, v) = \frac{1}{n} \sum_{i=1}^{n} 1(W_i^- \leq u,\ W_i^+ \leq v) \xrightarrow{P} F(u, v), \qquad \text{as } n \to \infty, \tag{1}$$

[1] The $-$ and $+$ superscripts refer to the inbound or outbound nature of edges in the graph, and are not related to the positive and negative parts of a real number.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

for all continuity points of some distribution F, where F is defined on the space $\mathcal{S} = \mathbb{R}_+^2$ and $\xrightarrow{P}$ denotes convergence in probability. Let $\mathscr{F}_n = \sigma(\mathbf{W}_i : 1 \le i \le n)$, and define $\mathbb{P}_n(\cdot) = P(\cdot|\mathscr{F}_n)$ and $\mathbb{E}_n[\cdot] = E[\cdot|\mathscr{F}_n]$ to be the conditional probability and conditional expectation, respectively, given the type sequence. Later in Section 3 we will enlarge the type vectors to include additional vertex attributes.

Remark 2.1. In general, depending on the nature of the type sequence (e.g., a deterministic sequence of numbers), it may be necessary to consider a double sequence $\{\mathbf{W}_i^{(n)} : i \ge 1,\ n \ge 1\}$ in order to satisfy (1). In practice, an easy way to avoid the need for considering double sequences is to assume that the type sequence $\{\mathbf{W}_i : i \ge 1\}$ consists of i.i.d. observations from distribution F.

We now define our family of random digraphs using the conditional probability, given the type sequence, that edge $(i, j) \in E_n$,

$$p_{ij}^{(n)} \triangleq \mathbb{P}_n((i,j) \in E_n) = 1 \wedge \frac{W_i^+ W_j^-}{\theta n}(1 + \varphi_n(\mathbf{W}_i, \mathbf{W}_j)), \qquad 1 \le i \ne j \le n, \tag{2}$$

where $-1 < \varphi_n(\mathbf{W}_i, \mathbf{W}_j)$ a.s., and $n^{-1}\sum_{i=1}^n (W_i^- + W_i^+) \xrightarrow{P} \theta > 0$ as $n \to \infty$.

We point out that the term $\varphi_n(\mathbf{W}_i, \mathbf{W}_j) = \varphi(n, \mathbf{W}_i, \mathbf{W}_j, \mathscr{W}_n)$ may depend on the entire sequence $\mathscr{W}_n \triangleq \{\mathbf{W}_i : 1 \le i \le n\}$, on the types of the vertices (i, j), or exclusively on n. Here and in the sequel, $x \wedge y = \min\{x, y\}$ and $x \vee y = \max\{x, y\}$. In the context of [5], definition (2) corresponds to the so-called rank-1 kernel, i.e., $\kappa(\mathbf{W}_i, \mathbf{W}_j) = \kappa_+(\mathbf{W}_i)\kappa_-(\mathbf{W}_j)$, with $\kappa_+(\mathbf{W}) = W^+/\sqrt{\theta}$ and $\kappa_-(\mathbf{W}) = W^-/\sqrt{\theta}$.

Example 2.2. Directed versions of some known random graph models covered by (2):

- Directed Erdős–Rényi model:

$$p_{ij}^{(n)} = \frac{\lambda}{2n}, \qquad 1 \le i \ne j \le n,$$

  for $\lambda > 0$, which corresponds to taking $W_i^- = W_i^+ = \lambda$ for all $1 \le i \le n$. This graph produces homogeneous graphs with Poisson(λ) degrees. Here, $\varphi_n(\mathbf{W}_i, \mathbf{W}_j) \equiv 0$ for all $1 \le i \ne j \le n$.
- Directed Chung–Lu model:

$$p_{ij}^{(n)} = \frac{W_i^+ W_j^-}{L_n} \wedge 1, \qquad 1 \le i \ne j \le n,$$

  where $L_n = \sum_{i=1}^n (W_i^- + W_i^+)$. This model is defined for any nonnegative sequences $\{W_i^- : i \ge 1\}$ and $\{W_i^+ : i \ge 1\}$ possessing some limiting distributions, e.g., power-laws, usually with finite second moments. Here, $\varphi_n(\mathbf{W}_i, \mathbf{W}_j) = \theta n/L_n - 1$ for all $1 \le i \ne j \le n$.
- Directed generalized random graph:

$$p_{ij}^{(n)} = \frac{W_i^+ W_j^-}{L_n + W_i^+ W_j^-}, \qquad 1 \le i \ne j \le n,$$

  where L_n is defined as above. Since the ratios in the definition of $p_{ij}^{(n)}$ are self-normalized, it provides a more natural model for graphs with infinite variance degrees. Here, $\varphi_n(\mathbf{W}_i, \mathbf{W}_j) = \theta n/(L_n + W_i^+ W_j^-) - 1$ for $1 \le i \ne j \le n$.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

- Directed Poissonian random graph or Norros–Reittu model:

$$p_{ij}^{(n)} = 1 - e^{-W_i^+ W_j^- / L_n}, \qquad 1 \le i \ne j \le n,$$

where L_n is defined as above. Here, $\varphi_n(\mathbf{W}_i, \mathbf{W}_j) = (1 - e^{-W_i^+ W_j^- / L_n})\theta n/(W_i^+ W_j^-) - 1$ for $1 \le i \ne j \le n$.

From a modeling perspective, one can think of W_i^+ as an attribute of vertex i that determines how likely it is for it to have outbound neighbors, and W_i^- as an attribute that indicates its popularity, or likelihood that other vertices may have edges pointing towards it. In other words, W_i^+ controls the out-degree of vertex i and W_i^- its in-degree. In applications, e.g., the World Wide Web, these two attributes can be used to model how trustworthy a webpage is, how valuable/relevant is its content, or how carefully it chooses the webpages it references.

2.1. Degree distributions

Our first result in the paper establishes that the family of random digraphs defined via (2) produces inhomogeneous graphs whose degree distribution can be modeled through that of the type distribution. The ArXiv version of this paper [26] includes a verification of all our assumptions for each of the models in Example 2.2 when the type sequence $\{\mathbf{W}_i : i \ge 1\}$ consists of i.i.d. random vectors.

Assumption 2.3. Let $\mathcal{G}(V_n, E_n)$ be a random digraph having type sequence $\{\mathbf{W}_i : i \ge 1\}$ and edge probabilities given by (2). Suppose further that:

(a) The type sequence $\{\mathbf{W}_i : i \ge 1\}$ satisfies (1).
(b) The following limits hold in probability:

$$E[W^-] = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n W_i^- \qquad \text{and} \qquad E[W^+] = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n W_i^+,$$

with $\theta = E[W^- + W^+] < \infty$.
(c) $\mathcal{E}_n \triangleq \frac{1}{n}\sum_{i=1}^n \sum_{1\le j\le n, j\ne i} |p_{ij}^{(n)} - (r_{ij}^{(n)} \wedge 1)| \xrightarrow{P} 0$ as $n \to \infty$, where $r_{ij}^{(n)} = W_i^+ W_j^-/(\theta n)$.
(d) $\frac{1}{n}\sum_{i=1}^n \sum_{1\le j\le n, j\ne i} \sum_{1\le k\le n, k\ne i} p_{ji}^{(n)} p_{ik}^{(n)} \xrightarrow{P} E[W^-W^+]E[W^+]E[W^-]/\theta^2$ as $n \to \infty$, with $E[W^-W^+] < \infty$.

We point out that Assumption 2.3 implies that $|\varphi_n(\mathbf{W}_i, \mathbf{W}_j)| \xrightarrow{P} 0$ as $n \to \infty$ for any $i, j \ge 1$.

We now define the in-degree and out-degree of vertex $i \in V_n$ according to

$$D_i^- = \sum_{j\in V_n,\, j\ne i} X_{ji} \qquad \text{and} \qquad D_i^+ = \sum_{j\in V_n,\, j\ne i} X_{ij},$$

respectively, where $X_{ij} = 1((i, j) \in E_n)$ is the indicator function of whether edge (i, j) is present in the graph. Note that from the independent edges assumption, we have that the $\{X_{ij} : 1 \le i \ne j \le n\}$ form a sequence of independent Bernoulli random variables, with $\mathbb{P}_n(X_{ij} = 1) = p_{ij}^{(n)}$.

The following theorem provides the distribution of the in-degree and out-degree of a randomly chosen vertex, i.e., a typical vertex, in a graph generated via our model. Its proof is given in Section 4.2.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

Theorem 2.4. *Under Assumption 2.3(a)–(c), the degrees (D_ξ^-, D_ξ^+) of a randomly chosen vertex in $\mathcal{G}(V_n, E_n)$ satisfy, as $n \to \infty$,*

$$\sup_{A \subseteq \mathbb{N}^2} \left| \mathbb{P}_n \left((D_\xi^-, D_\xi^+) \in A \right) - P((Z^-, Z^+) \in A) \right| \xrightarrow{P} 0,$$

and

$$\mathbb{E}_n[D_\xi^\pm] \xrightarrow{P} E[Z^\pm] = \frac{E[W^-]E[W^+]}{\theta},$$

where Z^- and Z^+ are mixed Poisson random variables with mixing parameters, $\frac{E[W^+]}{\theta} W^-$ and $\frac{E[W^-]}{\theta} W^+$, respectively, with Z^- and Z^+ conditionally independent given (W^-, W^+). Moreover, if Assumption 2.3(a)–(d) holds then, as $n \to \infty$,

$$\mathbb{E}_n[D_\xi^- D_\xi^+] \xrightarrow{P} E[Z^- Z^+] = \frac{E[W^- W^+]E[W^-]E[W^+]}{\theta^2}.$$

Remark 2.5. To relate this result with scale-free graphs where at least one of the degree distributions, usually the in-degree, follows a power-law, we point out that when W^- (W^+) has a regularly varying distribution with index $-\alpha < -1$, i.e., $P(W^- > x) = x^{-\alpha} L(x)$ for some slowly varying function L, then, by Proposition 8.4 in [19], we have that Z^- (Z^+) is also regularly varying with the same index. Furthermore, it can be shown that if (W^-, W^+) is jointly regularly varying (possibly in the non-standard sense defined in [32]), then so is (Z^-, Z^+), however, for our analysis of PageRank we will impose that W^- and W^+ be independent, so only the marginal distributions of Z^- and Z^+ are relevant to our main result.

3. Generalized PageRank

We now move on to the analysis of the typical behavior of the PageRank algorithm on the family of inhomogeneous random digraphs described in Section 2. Our main result shows that the distribution of the ranks produced by the algorithm converges to that of the attracting endogenous solution, $\mathcal{R}$, to a linear SFPE. Moreover, since the behavior of $\mathcal{R}$ is known to follow a power-law when the limiting in-degree distribution does, our theorem provides further evidence of the universality of the so-called "power-law hypothesis" on scale-free complex networks [8]. For completeness, we give below a brief description of the algorithm, which is well-defined for any directed graph $\mathcal{G}(V_n, E_n)$ on the vertex set $V_n = \{1, 2, \ldots, n\}$ with edges in the set E_n.

Let D_i^- and D_i^+ denote the in-degree and out-degree, respectively, of vertex i in $\mathcal{G}(V_n, E_n)$. We refer to the sequence $\{(D_i^-, D_i^+) : 1 \le i \le n\}$ as the bi-degree sequence of the graph $\mathcal{G}(V_n, E_n)$. The generalized PageRank vector $\mathbf{r} = (r_1, \ldots, r_n)$ is the unique solution to the following system of equations:

$$r_i = \sum_{(j,i) \in E_n} \frac{\zeta_j}{D_j^+} \cdot r_j + q_i, \quad i = 1, \ldots, n, \tag{3}$$

where $\mathbf{q} = (q_1, \ldots, q_n)$ is known as the personalization or teleportation vector (usually, a probability vector), and the $\{\zeta_i\}$ are referred to as the damping factors. In the original formulation of PageRank [6], the personalization values and the damping factors are given, respectively, by $q_i = (1-c)/n$ and $\zeta_i = c$ for all $1 \le i \le n$; the constant $c \in (0, 1)$ is known as the "damping factor". The formulation given in [8] is more general, and it allows any choice for both the personalization values and the damping factors, provided that $\max_{1 \le i \le n} |\zeta_i| \le c < 1$.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

We refer the reader to §1.1 in [8] for further details on the history of PageRank, its applications, and a matrix representation of the solution $\mathbf{r}$ to (3).

In order to analyze $\mathbf{r}$ on directed complex networks, we first eliminate the dependence on the size of the graph by computing the scale free ranks $(R_1, \dots, R_n) = \mathbf{R} \triangleq n\mathbf{r}$, which corresponds to solving:

$$R_i = \sum_{(j,i)\in E_n} C_j R_j + Q_i, \quad i = 1, \dots, n, \tag{4}$$

where $Q_i = q_i n$ and $C_j = \zeta_j / D_j^+$. We refer to the $\{C_j\}$ as the weights.

On scale-free graphs, i.e., where the in-degree sequence (or both the in-degree and out-degree sequences) follow a power law distribution, the power law hypothesis states that the distribution of the ranks $\{R_i : 1 \le i \le n\}$ will also have a power-law with the same index as that of the in-degrees. The first approach towards a proof of this phenomenon was given in [23,27,35], where the tree heuristic commonly used in the analysis of locally tree-like random graphs yields a stochastic fixed-point equation of the form

$$\mathcal{R} \overset{\mathcal{D}}{=} \sum_{j=1}^{\mathcal{N}} \mathcal{C}_j \mathcal{R}_j + \mathcal{Q}, \tag{5}$$

where $\mathcal{N}$ is a random variable distributed according to the limiting in-degree distribution of the graph, $\mathcal{Q}$ has the limiting distribution of the personalization values, the weights $\{\mathcal{C}_j\}$ are i.i.d. and independent of $(\mathcal{N}, \mathcal{Q})$, and are size-biased versions of the weights $\{C_j\}$ in (4), and the $\{\mathcal{R}_i\}$ are i.i.d. copies of $\mathcal{R}$. The connection between (4) and (5) can be understood by interpreting $\mathcal{R}$ as the rank of a randomly chosen vertex, with $(\mathcal{N}, \mathcal{Q})$ denoting its in-degree and personalization value, respectively, and then arguing that, provided the neighborhood of the chosen vertex looks locally like a tree, the ranks of its inbound neighbors should have the same distribution as $\mathcal{R}$. That the weights $\{\mathcal{C}_j\}$ in (5) are different from the $\{C_j\}$ appearing in (4), and are instead size-biased versions of them, follows from the observation that vertices with high out-degrees are more likely to be the neighbors of the randomly chosen vertex.

The heuristic described above was first made rigorous in [8], where it was shown that on graphs generated via the directed configuration model [9], the rank of a randomly chosen vertex converges in distribution, as the size of the graph grows to infinity, to a random variable

$$\mathcal{R}^* = \sum_{i=1}^{\mathcal{N}_0} \mathcal{C}_i \mathcal{R}_i + \mathcal{Q}_0,$$

where the $\{\mathcal{R}_i\}$ are i.i.d. copies of the attracting endogenous solution to (5), and are independent of $(\mathcal{N}_0, \mathcal{Q}_0, \{\mathcal{C}_i\}_{i\ge 1})$. The vector $(\mathcal{N}_0, \mathcal{Q}_0)$ may have a different distribution from that of $(\mathcal{N}, \mathcal{Q})$ in (5) depending on how we choose the first vertex (it has the same distribution as $(\mathcal{N}, \mathcal{Q})$ when the first vertex is chosen uniformly at random and the in-degree and out-degree are asymptotically independent, which is consistent with the approach we take here). That the solution $\mathcal{R}$ to (5) has a power-law distribution when $\mathcal{N}$ does has been the topic of a number of papers [23–25,34], and together with the results in [8] (see Theorems 6.4 and 6.6) provides the first proof of the power-law hypothesis on a complex network. We now show that a similar result also holds for the family of inhomogeneous random digraphs considered here.

3.1. PageRank on inhomogeneous random digraphs

As with the analysis done in [8] on the directed configuration model, the key idea is to couple the rank of a randomly chosen vertex with the rank of the root node of a tree, in

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

this case, a multi-type branching process. In order to incorporate vertex information used by the algorithm, as described by (3), we expand the type of vertex i to be of the form $\mathbf{W}_i = (W_i^-, W_i^+, Q_i, \zeta_i) \in \mathbb{R}_+^2 \times \mathbb{R}^2 \triangleq \mathcal{S}$, where the sequence $\{\mathbf{W}_i : i \geq 1\}$ satisfies

$$H_n(u, v, q, t) = \frac{1}{n} \sum_{i=1}^{n} 1(W_i^- \leq u, W_i^+ \leq v, Q_i \leq q, \zeta_i \leq t) \xrightarrow{P} H(u, v, q, t), \tag{6}$$

as $n \to \infty$, for all continuity points of some distribution H. With some abuse of notation, we continue using $\mathscr{F}_n = \sigma(\mathbf{W}_i : 1 \leq i \leq n)$ to denote the sigma-algebra generated by $\{\mathbf{W}_i : 1 \leq i \leq n\}$, along with the corresponding conditional probability and expectation $\mathbb{P}_n(\cdot) = P(\cdot|\mathscr{F}_n)$ and $\mathbb{E}_n[\cdot] = E[\cdot|\mathscr{F}_n]$.

We now impose some assumptions on the extended type sequence $\{\mathbf{W}_i : i \geq 1\}$.

Assumption 3.1. Let $\mathcal{G}(V_n, E_n)$ be a random digraph having type sequence $\{\mathbf{W}_i : i \geq 1\}$ and edge probabilities given by (2). Suppose further that:

(a) The extended type sequence $\{\mathbf{W}_i : i \geq 1\}$ satisfies (6).
(b) The following limits hold in probability:

$$E[W^-] = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} W_i^-, \qquad E[W^+] = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} W_i^+,$$

$$E[|Q|] = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} |Q_i|, \qquad \text{and} \qquad E[|\zeta|] = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} |\zeta_i|,$$

with $\theta = E[W^- + W^+] < \infty$ and $E[|Q| + |\zeta|] < \infty$.
(c) $\mathcal{E}_n = \frac{1}{n} \sum_{i=1}^{n} \sum_{1\leq j\leq n, j\neq i} |p_{ij}^{(n)} - (r_{ij}^{(n)} \wedge 1)| \xrightarrow{P} 0$ as $n \to \infty$, where $r_{ij}^{(n)} = W_i^+ W_j^- / (\theta n)$.
(d) $|\zeta_i| \leq c < 1$ for all $i = 1, \ldots, n$.
(e) The following limits hold in probability:

$$E[W^+W^-] = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1} W_i^- W_i^+ \qquad \text{and} \qquad E[W^+|Q|] = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1} W_i^+ |Q_i|,$$

with $E[W^+W^- + W^+|Q|] < \infty$.
(f) The vectors (W^-, Q) and (W^+, ζ) are independent.

Remark 3.2. Note that Assumption 3.1(a)–(c) implies Assumption 2.3(a)–(c).

Our main result on the distribution of the rank of a randomly chosen vertex in the inhomogeneous random digraph from Section 2 is given below. To avoid repetition, we refer the reader to [8] or [25] for a detailed description of the attracting endogenous solution $\mathcal{R}$ to (5), as well as its asymptotic behavior in terms of that of $\mathcal{N}, \mathcal{Q}, \mathcal{C}$; $\Rightarrow$ denotes weak convergence.

Theorem 3.3. *Suppose that Assumption 3.1 holds, and let R_ξ denote the rank of a uniformly chosen vertex in the inhomogeneous random digraph $\mathcal{G}(V_n, E_n)$. Then, as $n \to \infty$,*

$$R_\xi \Rightarrow \mathcal{R}, \tag{7}$$

where $\mathcal{R}$ is the attracting endogenous solution to (5). The distributions of all the random variables involved in (5) are given below:

$$P(\mathcal{N} = m, \mathcal{Q} \in dq) = E\left[1(Q \in dq) \cdot \frac{e^{-E[W^+]W^-/\theta} (E[W^+]W^-/\theta)^m}{m!} \right],$$

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

$$m = 0, 1, \ldots,$$
$$P(\mathcal{C}_1 \in dt) = \frac{E[1(\zeta/(Z^+ + 1) \in dt)W^+]}{E[W^+]},$$

and Z^+ *is a mixed Poisson random variable with parameter* $E[W^-]W^+/\theta$.

The proof of Theorem 3.3 is based on a coupling argument between a graph exploration process and a multi-type branching process, which is similar to the techniques used in [8] for the analysis of generalized PageRank on the directed configuration model. Together with the results in [8] and Remark 2.5, Theorem 3.3 provides further evidence of the "universality" of the power-law hypothesis on scale-free directed complex networks. The ArXiv version of this paper [26] contains some numerical examples illustrating the convergence of $R_\xi \Rightarrow \mathcal{R}$ for all the models in Example 2.2.

In the following section we explain the main steps involved in the proof of Theorem 3.3, postponing all the technical proofs to Section 4.

3.2. Deriving the SFPE approximation

To make the proof of Theorem 3.3 easier to follow, we have divided it into three main steps: (1) approximating the rank using the local neighborhood, (2) coupling with a branching process, and (3) proving convergence to the attracting endogenous solution.

3.2.1. Approximating the rank using the local neighborhood

The first step towards proving Theorem 3.3 consists in showing that it is enough to consider only the local neighborhood of each vertex in the graph to compute its rank. The first observation we make is that the system of linear equations given by (4) can be written in matrix notation as

$$\mathbf{R} = \mathbf{R}\mathbf{M} + \mathbf{Q},$$

where $\mathbf{R} = (R_1, \ldots, R_n)$, $\mathbf{Q} = (Q_1, \ldots, Q_n)$ and the matrix $\mathbf{M}$ has (i, j)th component

$$M_{ij} = s_{ij}C_i,$$

where s_{ij} is the number of edges from i to j. Recall that $C_j = \zeta_j/D_j^+$, where D_j^+ is the out-degree of vertex j and $|\zeta_j| \le c < 1$ for all $j \ge 1$. Since the graphs we consider here are simple, we have $s_{ij} \in \{0, 1\}$, however, the definition of matrix $\mathbf{M}$ also applies to multigraphs. It follows that the rank vector $\mathbf{R}$ can be written as

$$\mathbf{R} = \mathbf{R}^{(n,\infty)} = \sum_{i=0}^{\infty} \mathbf{Q}\mathbf{M}^i.$$

Next, define $(R_1^{(n,k)}, \ldots, R_n^{(n,k)}) = \mathbf{R}^{(n,k)} = \sum_{i=0}^{k} \mathbf{Q}\mathbf{M}^i$, and note that the i.i.d. nature of the type sequence implies that all the coordinates of the vector $\mathbf{R}^{(n,\infty)} - \mathbf{R}^{(n,k)}$ are identically distributed (they are not identically distributed given $\mathscr{F}_n$). It follows from the exact arguments used in Section 4.2 in [8] that for a randomly chosen vertex ξ,

$$\mathbb{P}_n\left(\left|R_\xi^{(n,\infty)} - R_\xi^{(n,k)}\right| > x^{-1}\right) \le \frac{xc^k}{1-c} \cdot \frac{1}{n}\sum_{i=1}^{n} |Q_i| \tag{8}$$

for any $x \ge 1$.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

Note that the calculation of each of the $R_i^{(n,k)}$, $i = 1, \dots, n$, requires only information about the vertices in the graph having a directed path to vertex i of length at most k, i.e., it can be computed using only the local (inbound) neighborhood of each vertex.

3.2.2. Coupling with a branching process

Now that we have reduced the problem of analyzing a randomly chosen component of the vector $\mathbf{R}^{(n,\infty)}$ to that of analyzing the corresponding component of the vector $\mathbf{R}^{(n,k)}$, the next step is to couple $R_\xi^{(n,k)}$ with the rank of the root node of a branching process. For the directed configuration model analyzed in [8], the coupling was done with a marked Galton–Watson process, referred to as a "thorny branching process" in [8], that was then used to define a weighted branching process [31]. The same idea works also for the inhomogeneous random digraphs considered here, although the coupling is more easily understood if instead of using from the beginning a marked Galton–Watson process we first consider a marked multi-type branching process. The marks include the number of outbound neighbors, the damping factor and the personalization value of each vertex discovered during the graph exploration process.

As it is usual when analyzing trees, we index the nodes with a label that allows us to trace their entire path from the root. More precisely, denote the root node $\emptyset$, and label its offspring as $\{1, 2, \dots, \hat{N}_\emptyset\}$, where $\hat{N}_\emptyset$ is the number of offspring that $\emptyset$ has. Set $\hat{A}_0 = \{\emptyset\}$ and $\hat{A}_1 = \{1, 2, \dots, \hat{N}_\emptyset\}$ to be the sets of individuals in generation zero and generation one of the tree, respectively. In general, we use $\hat{A}_k$ to denote the set of individuals in the kth generation of the tree, and a node/individual in $\hat{A}_k$ has a label of the form $\mathbf{i} = (i_1, \dots, i_k) \in \mathbb{N}_+^k$. Moreover, the set $\hat{A}_{k+1}$ can be constructed recursively according to

$$\hat{A}_{k+1} = \{(\mathbf{i}, j) : \mathbf{i} \in \hat{A}_k,\ 1 \le j \le \hat{N}_\mathbf{i}\},$$

where $\hat{N}_\mathbf{i}$ is the number of offspring of node $\mathbf{i}$, and we use $(\mathbf{i}, j) = (i_1, \dots, i_k, j)$ to denote the index concatenation operation; if $\mathbf{i} = \emptyset$, then $(\mathbf{i}, j) = j$. We use throughout the paper $\mathcal{U} = \bigcup_{k=0}^\infty \mathbb{N}_+^k$, with the convention that $\mathbb{N}_+^0 = \{\emptyset\}$.

To describe the multi-type branching process used in the coupling, we assume that each node in the tree has a type from the set $\mathscr{W}_n = \{\mathbf{W}_i : 1 \le i \le n\}$, where $\mathbf{W}_i = (W_i^-, W_i^+, Q_i, \zeta_i)$. Individuals in the tree have a random number of offspring, potentially of various types, independently of all other nodes. More precisely, if we let Z_{ji} denote the number of offspring of type $\mathbf{W}_j$ that an individual of type $\mathbf{W}_i$ has, we have that for $(m_1, \dots, m_n) \in \mathbb{N}^n$,

$$\mathbb{P}_n\left(Z_{1i} = m_1, \dots, Z_{ni} = m_n\right) = \prod_{j=1}^n \frac{e^{-q_{ji}^{(n)}} (q_{ji}^{(n)})^{m_j}}{m_j!}, \tag{9}$$

where

$$q_{ji}^{(n)} = \frac{(W_j^+ \wedge a_n)(W_i^- \wedge b_n)}{\theta n}, \qquad 1 \le i, j \le n,$$

and $a_n, b_n \ge 1$ are sequences to be determined later. To simplify the notation, we write $\bar{W}_i^+ = W_i^+ \wedge a_n$ and $\bar{W}_i^- = W_i^- \wedge b_n$. Note that the random variables $\{Z_{ji} : 1 \le j \le n\}$ are conditionally independent (given $\mathscr{F}_n$) Poisson random variables with the mean of Z_{ji} equal to $q_{ji}^{(n)}$. To avoid the label of a node from giving us any information about its type, we assume that all $\hat{N}_\mathbf{i}$ offspring of node $\mathbf{i}$ are permuted uniformly at random before being assigned a label of the form $(\mathbf{i}, j)$, $j = 1, \dots, \hat{N}_\mathbf{i}$.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

To make this a marked multi-type branching process, we give to each node $\mathbf{i}$ in the tree a mark $\hat{D}_{\mathbf{i}}$, such that if $\mathbf{i}$ has type $\mathbf{W}_s$, then

$$\mathbb{P}_n\left(\hat{D}_{\mathbf{i}} = m \,\middle|\, \mathbf{i} \text{ has type } \mathbf{W}_s\right) = \frac{e^{-\bar{W}_s^+ \Lambda_n^-/(\theta n)}(\bar{W}_s^+ \Lambda_n^-/(\theta n))^m}{m!}, \qquad m = 0, 1, 2, \ldots, \tag{10}$$

independently of all other nodes. Here and in the sequel, $\Lambda_n^- = \sum_{i=1}^n \bar{W}_i^-$ and $\Lambda_n^+ = \sum_{i=1}^n \bar{W}_i^+$. We refer to this marked multi-type branching process as a Poisson branching tree (PBT).

As mentioned earlier, it turns out that the PBT we just described can also be thought of as a marked Galton–Watson process. To see this, note that the properties of the Poisson distribution imply that the type of a node $\mathbf{i}$ in the tree is independent of the type of its parent, as the following result shows (its proof is given in Section 4.1).

Lemma 3.4. *For any node* $\mathbf{i}$ *in the PBT and any* $1 \le r, s \le n$, *we have*

$$\mathbb{P}_n(\mathbf{i} \text{ has type } \mathbf{W}_s | \text{parent has type } \mathbf{W}_r) = \frac{\bar{W}_s^+}{\Lambda_n^+}.$$

This means that we could construct the PBT by assigning to each node $\mathbf{i}$ in the tree a number of offspring $\hat{N}_{\mathbf{i}}$ and then sampling their types according to Lemma 3.4, independently of everything else. The marks $\hat{D}_{(\mathbf{i},j)}$ of each of these offspring would then be sampled according to (10). Since the type of the root node is chosen uniformly at random from the set $\mathscr{W}_n = \{\mathbf{W}_i : 1 \le i \le n\}$, the distribution of $\hat{N}_\emptyset$ may be different from that of all other nodes. This effect will disappear in the limit due to Assumption 3.1(f).

We now explain how to construct a coupling of the inhomogeneous random digraph $\mathcal{G}(V_n, E_n)$ and a PBT. We start by choosing uniformly at random a vertex in the graph, call it ξ, and then exploring its in-component using a breadth-first exploration process. The coupled PBT is constructed to be in perfect agreement with the graph exploration process for a number of generations large enough to ensure that the rank of the randomly chosen node can be accurately approximated by its rank computed up to that point. The exploration process will have even and odd steps: in Step $2k-1$ we will discover the set of vertices that have a directed path of length k to the randomly chosen vertex, in Step $2k$ we will uncover all the outbound neighbors of the vertices discovered in Step $2k-1$. To keep track of this process, each vertex in the graph exploration will be assigned one of three labels: {active, inactive, dead}; vertices that have not been uncovered have no label. Active vertices will be those that are currently most distant from the randomly chosen vertex, and all we know about them is that they have an outbound edge connecting them to the in-component of the first vertex. The vertices that have already been added to the exploration process, and whose inbound neighbors have been discovered, will be labeled dead. Vertices that have been discovered as additional outbound neighbors of active vertices are labeled inactive, and all we know about them is that they have an inbound edge connecting them to a vertex in the in-component we are exploring. Fig. 1 illustrates this process.

For the coupled PBT we will need to keep track of the active vertices at the end of Step $2k-1$ of the graph exploration process, which will constitute the kth generation of nodes/individuals in the tree. We will also keep track of the inactive vertices by defining a similar set composed of all the types sampled during the creation of the marks $\{\hat{D}_{\mathbf{i}}\}$ (note that in the graph each type appears only once, since each type can be identified with one of the n vertices, while on the

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

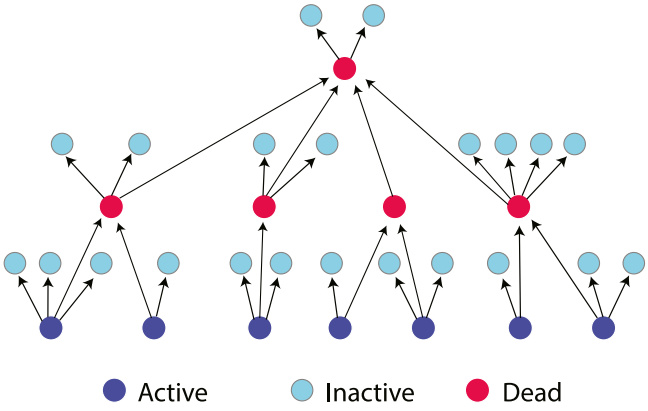

Fig. 1. Graph exploration process after completing Step 4.

tree types can appear repeatedly). The notation below will help us with the construction of the coupling.

For $k = 1, 2, \ldots,$ let

$$A_k = \text{set of "active" vertices at the end of Step } 2k-1.$$
$$T_k = \bigcup_{m=0}^{k-1} A_m = \text{set of "dead" vertices after Step } 2k-1;\ T_0 = \varnothing.$$
$$I_k = \text{set of "inactive" vertices after Step } 2k.$$
$$\hat{A}_k = \text{set of nodes in the PBT at distance } k \text{ from the root.}$$
$$\hat{T}_k = \bigcup_{m=0}^{k-1} \hat{A}_m = \text{set of nodes in the PBT at distance at most } k-1 \text{ from the root;}\ \hat{T}_0 = \varnothing.$$
$$\hat{I}_k = \text{set of "inactive" types in the PBT belonging to nodes at distance at most } k \text{ from the root.}$$

Note: The sets T_k and $\hat{T}_k$ grow in odd steps of the exploration process, while the sets I_k and $\hat{I}_k$ do so in even steps.

We now describe the coupling, for which we will require a sequence $\{U_{ij} : 1 \le i, j \le n\}$ of i.i.d. Uniform(0, 1) random variables that will be the same for the graph exploration process and the construction of the PBT. To make the role that the choice of the first node plays in the coupling explicit, we state our coupling results in terms of the first node we choose. The coupling has odd steps and even steps, during odd steps we discover new nodes in the inbound component of the first node, in even steps we explore the outbound neighbors (which will become the marks) of the nodes discovered in the previous step. Throughout the paper we will use $g^{-1}(u) = \inf\{x \in \mathbb{R} : g(x) \ge u\}$ to denote the generalized inverse of function g and $|A|$ to denote the cardinality of set A.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

Construction of the graph:

Step 0: Choose the vertex whose neighborhood will be explored, say vertex i. Set $A_0 = \{i\}$ and label vertex i as "active". To reveal all its outbound edges realize $X_{it} = 1(U_{it} > 1 - p_{it}^{(n)})$, $t = 1, 2, \ldots, n$, $t \neq i$. If $X_{it} = 1$, label node t as "inactive", so $I_0 = \{t \in \{1, 2, \ldots, n\} \setminus \{i\} : X_{it} = 1\}$.

In Step $2k-1$, $k \geq 1$, we explore the inbound neighbors of nodes in the set A_{k-1}. For each $i \in A_{k-1}$:

(1) For all $j = 1, 2, \ldots, n$, $j \neq i$ and $j \notin T_{k-1}$:

i. Realize $X_{ji} = 1(U_{ji} > 1 - p_{ji}^{(n)})$.

ii. If $X_{ji} = 1$ and node j was previously labeled "inactive", relabel it as "active".

(2) Label i as "dead".

In Step $2k$, $k \geq 1$, we explore the outbound neighbors of all the nodes in the set A_k. For each $j \in A_k$ and all $t = 1, 2, \ldots, n$, $t \neq j$ and $t \notin T_k$:

(1) Realize $X_{jt} = 1(U_{jt} > 1 - p_{jt}^{(n)})$.

(2) If $X_{jt} = 1$, label node t as "inactive".

Step $2k$ ends when we have uncovered all the nodes in A_k as well as their outbound neighbors.

Coupled construction of the PBT:

To each node $\mathbf{i}$ in the tree we will also determine its *mark* $\hat{D}_{\mathbf{i}}$ (the type of vertex $\mathbf{i}$ includes the values of $Q_{\mathbf{i}}$ and $\boldsymbol{\zeta}_{\mathbf{i}}$ so we can ignore those in the coupling). This value $\hat{D}_{\mathbf{i}}$ will be created independently for each node in the PBT according to (10), but will be coupled with the creation of "inactive" vertices the first time that a type appears. As long as the coupling holds, we choose nodes in the tree in the same order as in the graph, thus $\hat{D}_{\mathbf{i}}$ represents the out-degree of the corresponding node in the graph.

Step 0: Set $\hat{A}_0 = \{\emptyset\}$ and set the root of the PBT, $\emptyset$, to have type $\mathbf{W}_i$, where i is the vertex chosen in Step 0 of the graph construction. Define $G_{ij}(x) = \sum_{t=0}^{\lfloor x \rfloor} e^{-q_{ij}^{(n)}} (q_{ij}^{(n)})^t / t!$ to be the distribution of Z_{ij}, where Z_{ij} has the interpretation of being the number of offspring of type $\mathbf{W}_i$ that a node of type $\mathbf{W}_j$ has. Next, realize all the $Z_{it} = G_{it}^{-1}(U_{it})$ for $t = 1, 2, \ldots, n$, $t \neq i$, and $Z_{ii}^* \sim \text{Poisson}(q_{ii}^{(n)})$, independent of U_{it} and of any other Z_{it}, $t \neq i$. Set

$$\hat{D}_{\emptyset} = Z_{ii}^* + \sum_{1 \leq t \leq n,\, t \neq i} Z_{it}$$

and for each $Z_{it} \geq 1$ (or $Z_{ii}^* \geq 1$) add type $\mathbf{W}_t$ (or $\mathbf{W}_i$) to $\hat{I}_0$.

In Step $2k-1$, $k \geq 1$, we identify the individuals, and their types, in the kth generation of the PBT. For each node $\mathbf{i} \in \hat{A}_{k-1}$:

(a) If node $\mathbf{i}$ is the first node in the PBT to have type $\mathbf{W}_i$ proceed as follows:

(1) For $j = 1, 2, \ldots, n$, $j \neq i$:

i. Realize $Z_{ji} = G_{ji}^{-1}(U_{ji})$.

ii. If $Z_{ji} \geq 1$, add Z_{ji} nodes of type $\mathbf{W}_j$ to the active set.

(2) Realize $Z_{ii}^* \sim$Poisson$(q_{ii}^{(n)})$, independently of anything else, and assign a number Z_{ii}^* of type $\mathbf{W}_i$ offspring.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

(b) If node $\mathbf{i}$ is not the first node in the PBT to have type $\mathbf{W}_i$, sample a vector $(V_1, V_2, \ldots, V_n)$ of i.i.d. Uniform(0, 1) random variables, independent of the sequence $\{U_{ij} : 1 \le i, j \le n\}$, and of any other V_i's sampled before, and assign to node $\mathbf{i}$ a number $G_{ji}^{-1}(V_j)$ of type $\mathbf{W}_j$ offspring, for $j = 1, 2, \ldots, n$.

In Step $2k$, $k \ge 1$, we sample the marks of all the nodes in $\hat{A}_k$. For each node $\mathbf{i} \in \hat{A}_k$:

(a) If node $\mathbf{i}$ is the first node in the PBT to have type $\mathbf{W}_j$ proceed as follows:

i. Realize all the $Z_{jt} = G_{jt}^{-1}(U_{jt})$ for $t = 1, 2, \ldots, n$, and $t \notin T_k$.

ii. Sample $Z_{jt}^* \sim$Poisson$(q_{jt}^{(n)})$ for t already "dead", independently of everything else.

iii. Set

$$\hat{D}_{\mathbf{i}} = \sum_{1 \le t \le n,\, t \ne j,i,\, t \notin T_k} Z_{jt} + Z_{jj}^* + Z_{ji}^* + \sum_{t \in T_k} Z_{jt}^*,$$

and add all the corresponding types (i.e., add type $\mathbf{W}_t$ if $Z_{jt} \ge 1$ or $Z_{jt}^* \ge 1$) to the "inactive" set.

(b) If node $\mathbf{i}$ is not the first node in the PBT to have type $\mathbf{W}_j$, sample a vector $(V_1, V_2, \ldots, V_n)$ of i.i.d. Uniform(0,1) random variables, independent of the sequence $\{U_{ij} : 1 \le i, j \le n\}$, and of any other V_i's sampled before, and assign to node $\mathbf{i}$ a number $G_{ji}^{-1}(V_j)$ of type $\mathbf{W}_j$ offspring, for $j = 1, 2, \ldots, n$. Set $\hat{D}_{\mathbf{i}} = \sum_{t=1}^{n} Z_{jt}^*$ and add the corresponding types to the "inactive" set.

Definition 3.5. We say that the coupling of the graph and the PBT holds up to Step $2k$ if the graph exploration process up to a distance k from the first (root) vertex is identical to that of the PBT, i.e., $|A_m| = |\hat{A}_m|$ and $I_m = \hat{I}_m$ for all $0 \le m \le k$. Let τ be the step in the graph exploration process during which the coupling breaks.

Before stating the main result obtained from this step, we need to define:

$$\Delta_n = d_1(F_n, F), \tag{11}$$

where $d_1(F, G)$ is the Kantorovich–Rubinstein distance (or Wasserstein distance of order one) between distributions F and G. In particular,

$$d_1(F, G) = \inf_{\mathbf{X} \sim F, \mathbf{Y} \sim G} E[\|\mathbf{X} - \mathbf{Y}\|_1],$$

where the infimum is taken over all possible couplings of $\mathbf{X}$ and $\mathbf{Y}$, where $\mathbf{X}$ has distribution F and $\mathbf{Y}$ has distribution G. Since convergence in d_1 is equivalent to weak convergence and convergence of the first absolute moments (see Theorem 6.9 and Definition 6.8(i) in [33]), Assumption 2.3(a)–(b) implies that

$$\Delta_n \xrightarrow{P} 0 \qquad \text{as } n \to \infty.$$

Note that Δ_n is measuring the distance between the empirical distribution of the extended types and its limiting distribution, while $\mathcal{E}_n$ is measuring the error between the edge probabilities in the graph and their asymptotic equivalents.

Let $\mathbb{P}_{n,i}(\cdot) = \mathbb{P}_n(\cdot | A_0 = \{i\})$ or $\mathbb{P}_{n,i}(\cdot) = \mathbb{P}_n(\cdot | \emptyset$ has type $\mathbf{W}_i)$, depending on whether the event refers to the graph or to the PBT, respectively. The main result obtained from this step is given below, and its proof is given in Section 4.1. The theorem provides an explicit upper bound for the probability that the coupling breaks before step $2k$; we will later choose the sequences a_n, b_n, c_n, s_n in such a way that the bound converges to zero as $n \to \infty$.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

Theorem 3.6. *Fix $a_n, b_n, c_n, s_n \geq 1$ such that $s_n \leq a_n \wedge b_n$. Then, for any $k \in \mathbb{N}_+$,*

$$\begin{aligned}\mathbb{P}_{n,i}(\tau \leq 2k) \leq\ & 1(W_i^+ > a_n) + \mathcal{P}_n^+(i) + (\mathcal{H}_n/\theta)\bar{W}_i^+(\Delta_n + g_-(b_n) + a_n b_n/n) \\ & + \mathcal{H}_n k s_n\,(g_+(c_n) + c_n\,(\mathcal{E}_n + \Delta_n + g_+(a_n) + g_-(b_n) + a_n b_n/n)) \\ & + \mathbb{P}_{n,i}(|\hat{T}_k| \vee |\hat{I}_k| > s_n),\end{aligned}$$

where $\mathcal{P}_n^+(i) = \sum_{1\leq j\leq n, j\neq i} |p_{ij}^{(n)} - (r_{ij}^{(n)} \wedge 1)|$, $g_-(x) = E[(W^- - x)^+]$, $g_+(x) = E[(W^+ - x)^+]$, and

$$\mathcal{H}_n = \frac{2n}{\Lambda_n^+}(1 + \Delta_n/\theta)^2(3 + a_n^{-1} + b_n^{-1} + c_n^{-1}).$$

3.2.3. Convergence to the attracting endogenous solution

In view of Theorem 3.6, computing $R_\xi^{(n,k)}$ requires us to analyze only the first k generations of the PBT, provided $\tau > 2k$. In order to do so we first explain how to use the marks $\{\hat{D}_\mathbf{i}\}$ to compute the generalized PageRank of the root node of the PBT. For each node $\mathbf{i}$ in the PBT having type $\mathbf{W}_s$, we define its weight and personalization value according to

$$\hat{C}_\mathbf{i} = \frac{\zeta_s}{\hat{D}_\mathbf{i} + 1} \qquad \text{and} \qquad \hat{Q}_\mathbf{i} = Q_s.$$

Using the tree-indexing notation introduced in Section 3.2.2, we iteratively compute the rank of the root node of the PBT, denoted $\hat{R}_\emptyset^{(n,k)}$, according to

$$\hat{R}_\mathbf{i}^{(n,k)} = \sum_{j=1}^{\hat{N}_\mathbf{i}} \hat{C}_{(\mathbf{i},j)}\hat{R}_{(\mathbf{i},j)}^{(n,k-1)} + \hat{Q}_\mathbf{i}, \qquad k \geq 1, \qquad \hat{R}_\mathbf{j}^{(n,0)} = 0, \tag{12}$$

where $\hat{N}_\mathbf{i}$ is the total number of offspring that node $\mathbf{i}$ has. In view of Lemma 3.4 and the observation that the type of the root node will be chosen uniformly at random, we have that the distribution of $(\hat{N}_\emptyset, \hat{Q}_\emptyset)$ is given by

$$\mathbb{P}_n\left(\hat{N}_\emptyset = m,\ \hat{Q}_\emptyset = q\right) = \sum_{s=1}^n 1(Q_s = q)\cdot \frac{e^{-\frac{\Lambda_n^+}{\theta n}\bar{W}_s^-}(\Lambda_n^+\bar{W}_s^-/(\theta n))^m}{m!}\cdot\frac{1}{n}, \tag{13}$$

for $m \in \mathbb{N}$ and $q \in \mathbb{R}$. Moreover, for any node $\mathbf{i} \neq \emptyset$, we have that

$$\begin{aligned}&\mathbb{P}_n\left(\hat{N}_\mathbf{i} = m,\ \hat{Q}_\mathbf{i} = q,\ \hat{C}_\mathbf{i} = t\right) \\ &= \sum_{s=1}^n \mathbb{P}_n\left(\hat{N}_\mathbf{i} = m,\ \hat{Q}_\mathbf{i} = q,\ \hat{C}_\mathbf{i} = t \,\middle|\, \mathbf{i} \text{ has type } \mathbf{W}_s\right)\frac{\bar{W}_s^+}{\Lambda_n^+} \\ &= \sum_{s=1}^n 1(Q_s = q)\cdot\frac{e^{-\frac{\Lambda_n^+}{\theta n}\bar{W}_s^-}(\Lambda_n^+\bar{W}_s^-/(\theta n))^m}{m!}\cdot\mathbb{P}_n(\zeta_s/(\hat{D}_\mathbf{i}+1) = t|\mathbf{i} \text{ has type } \mathbf{W}_s)\cdot\frac{\bar{W}_s^+}{\Lambda_n^+} \\ &= \sum_{s=1}^n 1(Q_s = q, \zeta_s/t - 1 \in \mathbb{N})\cdot p(m; \Lambda_n^+\bar{W}_s^-/(\theta n))\cdot p(\zeta_s/t - 1; \Lambda_n^-\bar{W}_s^+/(\theta n))\cdot\frac{\bar{W}_s^+}{\Lambda_n^+},\end{aligned} \tag{14}$$

for $m \in \mathbb{N}$ and $t, q \in \mathbb{R}$, where $p(m;\lambda) = e^{-\lambda}\lambda^m/m!$. Note that the independence of the edges implies that the sequence $\{(\hat{N}_\mathbf{i}, \hat{Q}_\mathbf{i}, \hat{C}_\mathbf{i}) : \mathbf{i} \in \mathcal{U}\}$ consists of conditionally independent vectors given $\mathcal{F}_n$.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

Now that we have explained how to compute generalized PageRank on the PBT, we obtain, as a consequence of Theorem 3.6, the following result for $R_\xi^{(n,k)}$; its proof is given in Section 4.2.

Theorem 3.7. *Let ξ be the index of a uniformly chosen vertex in $\mathcal{G}(V_n, E_n)$. Under Assumption 3.1(a)–(c) we have that for any fixed $k \in \mathbb{N}_+$,*

$$\mathbb{P}_n\left(R_\xi^{(n,k)} \neq \hat{R}_\emptyset^{(n,k)}\right) \leq \frac{1}{n}\sum_{i=1}^n \mathbb{P}_{n,i}(\tau \leq 2k) \xrightarrow{P} 0,$$

as $n \to \infty$.

To make the connection with the SFPE, note that since we assume that (W^-, Q) is independent of (W^+, ζ), the vectors $\{(\hat{N}_\mathbf{i}, \hat{Q}_\mathbf{i}, \{\hat{C}_{(\mathbf{i},j)}\}_{j\geq 1}) : \mathbf{i} \in \mathcal{U}\}$ will be asymptotically independent, and therefore can be used to define a weighted branching process (WBP) with generic branching vector $(\mathcal{N}, \mathcal{Q}, \{\mathcal{C}_j\}_{j\geq 1})$, where the latter is the distributional limit of $(\hat{N}_\mathbf{i}, \hat{Q}_\mathbf{i}, \{\hat{C}_{(\mathbf{i},j)}\}_{j\geq 1})$, $\mathbf{i} \neq \emptyset$. Moreover, the $\{\mathcal{C}_j\}_{j\geq 1}$ will be i.i.d. and independent of $(\mathcal{N}, \mathcal{Q})$. We refer the reader to [8,23] for more details on the description and basic properties of WBPs of this form. The proof of this convergence in the Kantorovich–Rubinstein metric (see, e.g., Chapter 6 in [33]) is given in Section 4.4. Once this convergence is established, the convergence of $\hat{R}_\emptyset^{(n,k)}$ will follow from Theorem 2 in [10]. The precise statement of this last step in the proof of Theorem 3.3 is given below.

Theorem 3.8. *Under Assumption 3.1(a)–(b) & (d)–(f), we have that for any fixed $k \in \mathbb{N}_+$ the rank of the root node in the PBT computed up to generation k satisfies*

$$\hat{R}_\emptyset^{(n,k)} \Rightarrow \mathcal{R}^{(k)} \qquad \text{and} \qquad \mathbb{E}_n\left[|\hat{R}_\emptyset^{(n,k)}|\right] \xrightarrow{P} E[|\mathcal{R}^{(k)}|], \qquad n \to \infty,$$

where $\mathcal{R}^{(k)} \to \mathcal{R}$ a.s. as $k \to \infty$, with $\mathcal{R}$ defined as in Theorem 3.3.

The proof of Theorem 3.3 is obtained by combining (8), Theorems 3.7, and 3.8. All the proofs are given in Section 4.

4. Proofs

This section includes the proofs of Theorem 2.4, Lemma 3.4, Theorem 3.6, Theorem 3.7, Theorem 3.8, and ends with the proof of Theorem 3.3. Since some of the proofs are rather technical and require some preliminary results, we have organized them in subsections. We start with the proof of Theorem 3.6 followed by that of Theorem 3.7, since their proofs can be used to give a short proof of Theorem 2.4.

4.1. Proofs of Lemma 3.4 and Theorem 3.6

The proof of Theorem 3.6 is rather long, so we split some of the technical steps into three preliminary results to ease its reading. We point out that all of the results in this section are proven conditionally on the type sequence $\mathscr{W}_n = \{\mathbf{W}_i : 1 \leq i \leq n\}$, and therefore, all the expectations that appear throughout the section are finite. In some of our results related to the coupling, we use the notation $\mathbb{P}_{n,i}(\cdot) = \mathbb{P}_n(\cdot|A_0 = \{i\})$ or $\mathbb{P}_{n,i}(\cdot) = \mathbb{P}_n(\cdot|\emptyset \text{ has type } \mathbf{W}_i)$, depending on whether the event occurs on the graph or on the PBT, respectively. Similarly, we use $\mathbb{E}_{n,i}[\cdot] = \mathbb{E}_n[\cdot|A_0 = \{i\}]$ or $\mathbb{E}_{n,i}[\cdot] = \mathbb{E}_n[\cdot|\emptyset \text{ has type } \mathbf{W}_i]$ to denote the corresponding conditional expectations.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

Proof of Lemma 3.4. We start by noting that

$$\mathbb{P}_n(\mathbf{i} \text{ has type } \mathbf{W}_s | \text{parent has type } \mathbf{W}_r) = \mathbb{E}_n\left[\frac{Z_{sr}}{Z_{1r}+\cdots+Z_{nr}}\middle| Z_{1r}+\cdots+Z_{nr}\geq 1\right],$$

where the $\{Z_{jr} : 1 \leq j \leq n\}$ are independent Poisson random variables with $\mathbb{E}_n[Z_{jr}] = q_{jr}^{(n)}$. Since for two independent Poisson random variables X and Y with means μ and λ, respectively, we have that $X|X+Y=n$ has a Binomial$(n, \mu/(\mu+\lambda))$ distribution, then

$$\begin{aligned} E\left[\frac{X}{X+Y}\middle| X+Y\geq 1\right] &= \frac{1}{P(X+Y\geq 1)}E\left[\frac{X}{X+Y}\cdot 1(X+Y\geq 1)\right] \\ &= \frac{1}{P(X+Y\geq 1)}\sum_{n=1}^{\infty}\frac{1}{n}E[X|X+Y=n]\cdot\frac{e^{-\mu-\lambda}(\mu+\lambda)^n}{n!} \\ &= \frac{\mu}{(\mu+\lambda)P(X+Y\geq 1)}\sum_{n=1}^{\infty}\frac{e^{-\mu-\lambda}(\mu+\lambda)^n}{n!} = \frac{\mu}{\mu+\lambda}. \end{aligned}$$

It follows that

$$\mathbb{P}_n(\mathbf{i} \text{ has type } \mathbf{W}_s | \text{parent has type } \mathbf{W}_r) = \frac{q_{sr}^{(n)}}{q_{1r}^{(n)}+\cdots+q_{nr}^{(n)}} = \frac{\bar{W}_s^+}{A_n^+}. \quad \square$$

The proof of Theorem 3.6 is divided into two parts, one that computes the probability that the coupling breaks in Step 0 and another that computes the probability that it breaks in Step m, $m \geq 1$. In both cases, the idea behind the proofs is to identify the possible ways in which the coupling can break in Step m, and carefully estimate their corresponding probabilities. To help explain the steps in the proofs that follow, it may be helpful to list the events that can lead to the coupling breaking in Step m.

Remark 4.1. The coupling breaks at time τ for the following reasons:

- If $A_0 = \{i\}$, then $\tau = 0$ if: $Z_{it} \neq X_{it}$ for any $t = 1, 2, \ldots, n$, $t \neq i$, or $Z_{ii} \geq 1$.
- $\tau = 2m-1$, $m \geq 1$, if for some $i \in A_{m-1}$ any of the following happen:
  a. $X_{ji} = 1$ for any $j \in I_{m-1}$ (in which case a cycle is created).
  b. $Z_{ji} \geq 1$ for $j = i$ or $j \in I_{m-1}$.
  c. $X_{ji} \neq Z_{ji}$ for some $j = 1, 2, \ldots, n$, $j \neq i$, $j \notin I_{m-1}$.
- $\tau = 2m$, $m \geq 1$, if for some $j \in A_m$ either:
  d. $X_{jt} \neq Z_{jt}$ for some $t = 1, 2, \ldots, n$, $t \neq j, i$, $t \notin T_m$.
  e. $Z_{jt}^* \geq 1$ for $t = i$ or $t \in T_m$.

A first step in the derivation of the bounds we seek is the following preliminary result bounding the probabilities of having edge discrepancies between the exploration of the graph and of the coupled PBT, both inbound and outbound. Recall that $\Delta_n = d_1(F_n, F)$ was defined in (11).

Lemma 4.2. *For any $1 \leq i \leq n$ we have*

$$\mathbb{P}_n\left(\max_{1\leq j\leq n, j\neq i}|X_{ji}-Z_{ji}|\geq 1\right) \leq \min\left\{1, 1(W_i^- > b_n) + \mathcal{P}_n^-(i) + \bar{W}_i^-\eta_n^-\right\},$$

$$\mathbb{P}_n\left(\max_{1\le j\le n, j\neq i}|X_{ij}-Z_{ij}|\ge 1\right)\le \min\left\{1, 1(W_i^+ > a_n)+\mathcal{P}_n^+(i)+\bar{W}_i^+\eta_n^+\right\},$$

where

$$\mathcal{P}_n^-(i)=\sum_{1\le j\le n, j\neq i}\left|p_{ji}^{(n)}-(r_{ji}^{(n)}\wedge 1)\right|, \qquad \mathcal{P}_n^+(i)=\sum_{1\le j\le n, j\neq i}\left|p_{ij}^{(n)}-(r_{ij}^{(n)}\wedge 1)\right|,$$

$$\eta_n^- = (\Delta_n + g_+(a_n) + a_nb_n/n + a_nb_n\Delta_n/(\theta n))/\theta,$$
$$\eta_n^+ = (\Delta_n + g_-(b_n) + a_nb_n/n + a_nb_n\Delta_n/(\theta n))/\theta,$$

$$g_-(x)=E[(W^- - x)^+], \qquad \textit{and} \qquad g_+(x)=E[(W^+ - x)^+].$$

Proof. The analysis of the two probabilities is essentially the same, so we only prove the result for outbound edges. Let $R_{ij}=1(U_{ij} > 1-r_{ij}^{(n)})$ with $r_{ij}^{(n)}=W_i^+W_j^-/(\theta n)$. The union bound gives:

$$\mathbb{P}_n\left(\max_{1\le j\le n, j\neq i}|X_{ij}-Z_{ij}|\ge 1\right)\le 1(W_i^+ > a_n) + 1(W_i^+\le a_n)\sum_{1\le j\le n, j\neq i}\mathbb{P}_n(|X_{ij}-Z_{ij}|\ge 1).$$

Now note that

$$\begin{aligned}\mathbb{P}_n(|X_{ij}-Z_{ij}|\ge 1) &= \mathbb{P}_n(|X_{ij}-Z_{ij}|\ge 1, |X_{ij}-R_{ij}|\ge 1)\\ &\quad + \mathbb{P}_n(|X_{ij}-Z_{ij}|\ge 1, |X_{ij}-R_{ij}|= 0)\\ &\le \mathbb{P}_n(|X_{ij}-R_{ij}|\ge 1)+\mathbb{P}_n(|R_{ij}-Z_{ij}|\ge 1).\end{aligned}$$

The first probability can be computed to be:

$$\mathbb{P}_n(|X_{ij}-R_{ij}|\ge 1)=|p_{ij}^{(n)}-(r_{ij}^{(n)}\wedge 1)|.$$

To analyze each of probabilities involving R_{ij} and Z_{ij}, note that

$$\begin{aligned}\mathbb{P}_n\left(|R_{ij}-Z_{ij}|\ge 1\right) &= \mathbb{P}_n(R_{ij}=0, Z_{ij}\ge 1)+\mathbb{P}_n(R_{ij}=1, Z_{ij}=0)\\ &\quad + \mathbb{P}_n(R_{ij}=1, Z_{ij}\ge 2)\\ &= \left(1-(1\wedge r_{ij}^{(n)})-e^{-q_{ij}^{(n)}}\right)^+ + \left(e^{-q_{ij}^{(n)}}-1+(1\wedge r_{ij}^{(n)})\right)^+\\ &\quad + \min\left\{1-e^{-q_{ij}^{(n)}}(1+q_{ij}^{(n)}),\ (1\wedge r_{ij}^{(n)})\right\}\\ &= \left|1-(1\wedge r_{ij}^{(n)})-e^{-q_{ij}^{(n)}}\right|\\ &\quad + \min\left\{(1\wedge r_{ij}^{(n)}),\ e^{-q_{ij}^{(n)}}(e^{q_{ij}^{(n)}}-1-q_{ij}^{(n)})\right\}.\end{aligned}$$

Now use the inequalities $e^{-x}\ge 1-x$, $e^{-x}-1+x\le x^2/2$ and $e^x-1-x\le x^2e^x/2$ for $x\ge 0$, to obtain that

$$\begin{aligned}\mathbb{P}_n\left(|R_{ij}-Z_{ij}|\ge 1\right) &\le r_{ij}^{(n)}-q_{ij}^{(n)}+\left|1-q_{ij}^{(n)}-e^{-q_{ij}^{(n)}}\right|+e^{-q_{ij}^{(n)}}(e^{q_{ij}^{(n)}}-1-q_{ij}^{(n)})\\ &= r_{ij}^{(n)}-q_{ij}^{(n)}+e^{-q_{ij}^{(n)}}-1+q_{ij}^{(n)}+e^{-q_{ij}^{(n)}}(e^{q_{ij}^{(n)}}-1-q_{ij}^{(n)})\\ &\le r_{ij}^{(n)}-q_{ij}^{(n)}+(q_{ij}^{(n)})^2.\end{aligned}$$

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

It follows that

$$\begin{aligned}
&1(W_i^+ \le a_n) \sum_{1\le j\le n, j\ne i} \mathbb{P}_n(|X_{ij} - Z_{ij}| \ge 1) \\
&\le 1(W_i^+ \le a_n) \sum_{1\le j\le n, j\ne i} \left(|p_{ij}^{(n)} - (r_{ij}^{(n)} \wedge 1)| + r_{ij}^{(n)} - q_{ij}^{(n)} + (q_{ij}^{(n)})^2 \right) \\
&\le \mathcal{P}_n^+(i) + \sum_{1\le j\le n, j\ne i} \frac{\bar{W}_i^+(W_j^- - \bar{W}_j^-)}{\theta n} + \frac{(\bar{W}_i^+)^2}{(\theta n)^2} \sum_{1\le j\le n, j\ne i} (\bar{W}_j^-)^2 \\
&\le \mathcal{P}_n^+(i) + \frac{\bar{W}_i^+}{\theta n} \sum_{j=1}^n (W_j^- - b_n)^+ + \frac{(\bar{W}_i^+)^2 b_n \Lambda_n^-}{(\theta n)^2}.
\end{aligned}$$

To further bound the second term note that if we let $(W^{(-,n)}, W^{(+,n)})$ denote a random vector distributed according to F_n and (W^-, W^+) a random vector distributed according to F, then

$$\frac{1}{n}\sum_{j=1}^n (W_j^- - b_n)^+ = \mathbb{E}_n\left[(W^{(-,n)} - b_n)^+\right] \le d_1(F_n, F) + E\left[(W^- - b_n)^+\right] = \Delta_n + g_-(b_n).$$

And for the last term,

$$\frac{(\bar{W}_i^+)^2 b_n \Lambda_n^-}{(\theta n)^2} \le \frac{\bar{W}_i^+ a_n b_n}{\theta^2 n} \cdot \mathbb{E}_n\left[W^{(-,n)}\right] \le \frac{\bar{W}_i^+ a_n b_n}{\theta^2 n}\left(\Delta_n + E[W^-]\right).$$

We conclude that for η_n^+ as defined in the statement of the lemma,

$$\begin{aligned}
1(W_i^+ \le a_n) \sum_{1\le j\le n, j\ne i} \mathbb{P}_n(|X_{ij} - Z_{ij}| \ge 1) &\le \mathcal{P}_n^+(i) + \frac{\bar{W}_i^+}{\theta}(\Delta_n + g_-(b_n)) \\
&\quad + \frac{\bar{W}_i^+ a_n b_n}{\theta^2 n}(\Delta_n + E[W^-]) \\
&\le \mathcal{P}_n^+(i) + \eta_n^+ \bar{W}_i^+,
\end{aligned}$$

which in turn yields

$$\mathbb{P}_n\left(\max_{1\le j\le n, j\ne i} |X_{ij} - Z_{ij}| \ge 1\right) \le \min\left\{1, 1(W_i^+ > a_n) + \mathcal{P}_n^+(i) + \eta_n^+ \bar{W}_i^+\right\}. \quad \square$$

We now give an upper bound for the probability that the coupling breaks on Step 0 when the starting vertex is i.

Lemma 4.3. *We have*

$$\mathbb{P}_{n,i}(\tau = 0) \le 1(W_i^+ > a_n) + \mathcal{P}_n^+(i) + \bar{W}_i^+(\eta_n^+ + b_n/(\theta n)).$$

Proof. By the union bound followed by Lemma 4.2 we have,

$$\begin{aligned}
\mathbb{P}_{n,i}(\tau = 0) &\le \mathbb{P}_n\left(\max_{1\le t\le n, t\ne i} |X_{it} - Z_{it}| > 0\right) + \mathbb{P}_n\left(Z_{ii}^* \ge 1\right) \\
&\le 1(W_i^+ > a_n) + \mathcal{P}_n^+(i) + \bar{W}_i^+\eta_n^+ + 1 - e^{-q_{ii}^{(n)}} \\
&\le 1(W_i^+ > a_n) + \mathcal{P}_n^+(i) + \bar{W}_i^+\eta_n^+ + q_{ii}^{(n)} \\
&\le 1(W_i^+ > a_n) + \mathcal{P}_n^+(i) + \bar{W}_i^+(\eta_n^+ + b_n/(\theta n)). \quad \square
\end{aligned}$$

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

We now give an upper bound for the probability that the coupling breaks in Step m for $m \geq 1$.

Proposition 4.4. *Fix $c_n, s_n \geq 1$ with $s_n \leq a_n \wedge b_n$ and define for $m \geq 0$ the event $M_m = \left\{|\hat{T}_m| \vee |\hat{I}_m| \leq s_n\right\}$. Then, for any $m \geq 1$,*

$$\mathbb{P}_{n,i}(\tau = 2m-1, M_{m-1}) \leq \frac{s_n}{\Lambda_n^+/n}\left(g_+(c_n) + \Delta_n + c_n\mathcal{E}_n + c_n\gamma_n^-\right),$$
$$\mathbb{P}_{n,i}(\tau = 2m, M_m) \leq \frac{s_n}{\Lambda_n^+/n}\left(g_+(c_n) + \Delta_n + c_n\mathcal{E}_n + c_n\gamma_n^+\right),$$

where g_+ and g_- are defined as in Lemma 4.2,

$$\mathcal{E}_n = \frac{1}{n}\sum_{i=1}^n \sum_{1\leq j\leq n, j\neq i} |p_{ij}^{(n)} - (r_{ij}^{(n)} \wedge 1)|,$$

and

$$\gamma_n^- = \frac{(E[W^-] + \Delta_n)}{\theta}(\Delta_n + g_+(a_n) + 2a_nb_n/n + a_nb_n\Delta_n/(\theta n) + a_n/n),$$
$$\gamma_n^+ = \frac{(E[W^+] + \Delta_n)}{\theta}(\Delta_n + g_-(b_n) + 2a_nb_n/n + a_nb_n\Delta_n/(\theta n) + b_n/n).$$

Proof. We start by defining the following events:

$$F_i(I) = \left\{\max_{1\leq j\leq n, j\neq i} |X_{ji} - Z_{ji}| = 0,\ Z_{ii} + \sum_{1\leq j\leq n,\ j\neq i,\ j\in I} Z_{ji} = 0\right\},$$
$$G_j(D) = \left\{\max_{1\leq t\leq n, t\neq j,\ t\in D^c} |X_{jt} - Z_{jt}| = 0,\ Z_{jj}^* + \sum_{1\leq t\leq n,\ t\neq j,\ t\in D} Z_{jt}^* = 0\right\},$$
$$H_k = \bigcap_{i\in A_{k-1}} F_i(I_{k-1}),$$
$$J_k = \bigcap_{j\in A_k} G_j(D_k).$$

Now note that for any $m \geq 1$,

$$\mathbb{P}_{n,i}\left(\tau = 2m-1, M_{m-1}\right) = \mathbb{P}_{n,i}\left(M_{m-1} \cap J_0 \cap \bigcap_{k=1}^{m-1}(H_k \cap J_k) \cap H_m^c\right),$$
$$\mathbb{P}_{n,i}(\tau = 2m, M_m) = \mathbb{P}_{n,i}\left(M_m \cap J_0 \cap \bigcap_{k=1}^{m-1}(H_k \cap J_k) \cap H_m \cap J_m^c\right),$$

with the convention that $\bigcap_{k=1}^0(H_k \cap J_k) = \Omega$.

Let $\mathcal{F}_t$ denote the sigma-algebra that contains the history of the inbound exploration process in the graph as well as that of the PBT, up to the end of Step t of the graph exploration process. It follows that we can write:

$$\mathbb{P}_{n,i}(\tau = 2m-1, M_{m-1}) = \mathbb{E}_{n,i}\left[1\left(M_{m-1} \cap J_0 \cap \bigcap_{k=1}^{m-1}(H_k \cap J_k)\right)\mathbb{P}_n(H_m^c|\mathcal{F}_{2(m-1)})\right],$$

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

$$\mathbb{P}_{n,i}(\tau = 2m, M_m) = \mathbb{E}_{n,i}\left[1\left(M_m \cap J_0 \cap \bigcap_{k=1}^{m-1}(H_k \cap J_k) \cap H_m\right)\mathbb{P}_n(J_m^c|\mathcal{F}_{2m-1})\right].$$

To analyze the two conditional probabilities inside the expectations above note that conditionally on $\mathcal{F}_{2(m-1)}$, the types of the nodes in I_{m-1} are known. Therefore, by the union bound and the independence among the edges, we have:

$$\begin{aligned}\mathbb{P}_n(H_m^c|\mathcal{F}_{2(m-1)}) &= \mathbb{P}_n\left(\bigcup_{i\in A_{m-1}} F_i(I_{m-1})^c \,\middle|\, \mathcal{F}_{2(m-1)}\right)\\ &\le \sum_{i\in A_{m-1}} \mathbb{P}_n\left(F_i(I_{m-1})^c \,\middle|\, \mathcal{F}_{2(m-1)}\right)\\ &\le \sum_{i\in A_{m-1}} \min\left\{1,\ \mathbb{P}_n\left(\max_{1\le j\le n, j\ne i}|X_{ji} - Z_{ji}| \ge 1 \,\middle|\, \mathcal{F}_{2(m-1)}\right)\right.\\ &\quad \left. + \mathbb{P}_n\left(Z_{ii} + \sum_{1\le j\le n, j\ne i, j\in I_{m-1}} Z_{ji} \ge 1 \,\middle|\, \mathcal{F}_{2(m-1)}\right)\right\}.\end{aligned}$$

Now use the independence of the edges from the rest of the exploration process and Lemma 4.2 to obtain that

$$\begin{aligned}\mathbb{P}_n\left(\max_{1\le j\le n, j\ne i}|X_{ji} - Z_{ji}| \ge 1 \,\middle|\, \mathcal{F}_{2(m-1)}\right) &= \mathbb{P}_n\left(\max_{1\le j\le n, j\ne i}|X_{ji} - Z_{ji}| \ge 1\right)\\ &\le 1(W_i^- > b_n) + \mathcal{P}_n^-(i) + \bar{W}_i^-\eta_n^-.\end{aligned}$$

The independence from the exploration process also yields

$$\begin{aligned}\mathbb{P}_n\left(Z_{ii} + \sum_{1\le j\le n, j\ne i, j\in I_{m-1}} Z_{ji} \ge 1 \,\middle|\, \mathcal{F}_{2(m-1)}\right) &= 1 - e^{-q_{ii}^{(n)} - \sum_{1\le j\le n, j\ne i, j\in I_{m-1}} q_{ji}^{(n)}}\\ &= 1 - e^{-\frac{\bar{W}_i^-}{\theta n}\left(\bar{W}_i^+ + \sum_{1\le j\le n, j\ne i, j\in I_{m-1}} \bar{W}_j^+\right)}\\ &\le 1 - e^{-\frac{a_n\bar{W}_i^-}{\theta n}(1+|I_{m-1}|)}\\ &\le \frac{a_n\bar{W}_i^-}{\theta n}(1 + |I_{m-1}|),\end{aligned}$$

where in the last inequality we used $1 - e^{-x} \le x$ for $x \ge 0$.

It follows that

$$\begin{aligned}\mathbb{P}_{n,i}(\tau = 2m-1, M_{m-1}) \le \mathbb{E}_{n,i}\left[1\left(M_{m-1} \cap J_0 \cap \bigcap_{k=1}^{m-1}(H_k \cap J_k)\right)\sum_{j\in A_{m-1}} \min\left\{1,\ \mathcal{P}_n^-(j)\right.\right.\\ \left.\left. + \eta_n^-\bar{W}_j^- + \frac{a_n\bar{W}_j^-}{\theta n}(1 + |I_{m-1}|)\right\}\right].\end{aligned}$$

Almost the exact arguments can be used to obtain

$$\mathbb{P}_{n,i}(\tau = 2m, M_m) \le \mathbb{E}_{n,i}\left[1\left(M_m \cap J_0 \cap \bigcap_{k=1}^{m-1}(H_k \cap J_k) \cap H_m\right) \sum_{j\in A_m} 1(\bar{W}_i^+ \le c_n) \times \min\left\{1,\ \mathcal{P}_n^+(j) + \eta_n^+ \bar{W}_j^+ + \frac{b_n \bar{W}_j^+}{\theta n}(1 + |T_m|)\right\}\right].$$

To analyze these two remaining expectations we note that on the events $\{J_0 \cap \bigcap_{k=1}^{m-1}(H_k \cap J_k)\}$ and $\{J_0 \cap \bigcap_{k=1}^{m-1}(H_k \cap J_k) \cap H_m\}$ the coupling has not broken yet, and therefore we can replace A_{m-1}, I_{m-1}, A_m and T_m with their tree counterparts $\hat{A}_{m-1}$, $\hat{I}_{m-1}$, $\hat{A}_m$ and $\hat{T}_m$. Also, note that by Lemma 3.4 we have that the types of the nodes in each of the active sets $\hat{A}_k$ are independent of the type of their parents. We will then identify the nodes in $\hat{A}_{m-1}$ (or $\hat{A}_{m-1}$) as $\{Y_1, \dots, Y_{|\hat{A}_{m-1}|}\}$ (or $\{Y_1, \dots, Y_{|\hat{A}_m|}\}$), where for any $t \ge 1$,

$$\mathbb{P}_n(Y_t = j) = \frac{\bar{W}_j^+}{\Lambda_n^+}, \quad j = 1, 2, \dots, n.$$

It follows that

$$\mathbb{P}_{n,i}(\tau = 2m-1, M_{m-1}) \le \mathbb{E}_{n,i}\left[1(M_{m-1}) \sum_{t=1}^{|\hat{A}_{m-1}|} \min\left\{1,\ \mathcal{P}_n^-(Y_t) + \eta_n^- \bar{W}_{Y_t}^- + \frac{a_n \bar{W}_{Y_t}^-}{\theta n}(1 + |\hat{I}_{m-1}|)\right\}\right]$$

and

$$\mathbb{P}_{n,i}(\tau = 2m, M_m) \le \mathbb{E}_{n,i}\left[1(M_m) \sum_{t=1}^{|\hat{A}_m|} \min\left\{1,\ \mathcal{P}_n^+(Y_t) + \eta_n^+ W_{Y_t}^+ + \frac{b_n \bar{W}_{Y_t}^+}{\theta n}(1 + |\hat{T}_m|)\right\}\right].$$

Since on the event M_k we have $|\hat{I}_k| \le s_n \le b_n$ and $|\hat{T}_k| \le s_n \le a_n$, we further obtain that

$$\begin{aligned}\mathbb{P}_{n,i}(\tau = 2m-1, M_{m-1}) &\le \mathbb{E}_{n,i}\left[\sum_{t=1}^{\lfloor s_n \rfloor} \min\left\{1,\ \mathcal{P}_n^-(Y_t) + \eta_n^- \bar{W}_{Y_t}^- + \frac{a_n \bar{W}_{Y_t}^-}{\theta n}(1 + b_n)\right\}\right] \\ &\le s_n \mathbb{E}_n\left[\min\left\{1,\ \mathcal{P}_n^-(Y_1) + (\eta_n^- + a_n(1+b_n)/(\theta n))\bar{W}_{Y_1}^-\right\}\right]\end{aligned}$$

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

and

$$\mathbb{P}_{n,i}(\tau = 2m, M_m) \le \mathbb{E}_{n,i}\left[\sum_{t=1}^{\lfloor s_n \rfloor} \min\left\{1,\ \mathcal{P}_n^+(Y_t) + \eta_n^+ \bar{W}_{Y_t}^+ + \frac{b_n \bar{W}_{Y_t}^+}{\theta n}(1+a_n)\right\}\right]$$
$$\le s_n \mathbb{E}_n\left[\min\left\{1,\ \mathcal{P}_n^+(Y_1) + (\eta_n^+ + b_n(1+a_n)/(\theta n))\bar{W}_{Y_1}^+\right\}\right].$$

It only remains to compute the last two expectations. Throughout the rest of the proof, let $(W^{(-,n)}, W^{(+,n)}, W^-, W^+)$ be constructed according to the optimal coupling for F_n and F, i.e., $\mathbb{E}_n\left[|W^{(-,n)} - W^-| + |W^{(+,n)} - W^+|\right] = \Delta_n$. Let $(\bar{W}^{(-,n)}, \bar{W}^{(+,n)}) = (W^{(-,n)} \wedge b_n, W^{(+,n)} \wedge a_n)$. Now let $\gamma_n^- = (E[W^-] + \Delta_n)(\eta_n^- + a_n(1+b_n)/(\theta n))$ and note that for any $c_n \ge 1$,

$$s_n \mathbb{E}_n\left[\min\left\{1,\ \mathcal{P}_n^-(Y_1) + \frac{\gamma_n^-}{E[W^-]+\Delta_n}\bar{W}_{Y_1}^-\right\}\right]$$
$$= s_n \sum_{j=1}^n \frac{\bar{W}_j^+}{\Lambda_n^+} \min\left\{1,\ \mathcal{P}_n^-(j) + \frac{\gamma_n^-}{E[W^-]+\Delta_n}\bar{W}_j^-\right\}$$
$$\le \frac{s_n n}{\Lambda_n^+} \cdot \frac{1}{n}\sum_{j=1}^n (\bar{W}_j^+ - c_n)^+ + \frac{s_n n}{\Lambda_n^+} \cdot \frac{1}{n}\sum_{j=1}^n c_n \min\left\{1,\ \mathcal{P}_n^-(j) + \frac{\gamma_n^-}{E[W^-]+\Delta_n}\bar{W}_j^-\right\}$$
$$\le \frac{s_n}{\Lambda_n^+/n} \cdot \mathbb{E}_n[(\bar{W}^{(+,n)} - c_n)^+] + \frac{s_n c_n}{\Lambda_n^+/n}\left(\mathcal{E}_n + \frac{\gamma_n^-}{E[W^-]+\Delta_n}\mathbb{E}_n[\bar{W}^{(-,n)}]\right)$$
$$\le \frac{s_n}{\Lambda_n^+/n}\left(g_+(c_n) + \Delta_n + c_n\mathcal{E}_n + c_n\gamma_n^-\right).$$

Essentially the same arguments also yield for $\gamma_n^+ = (E[W^+] + \Delta_n)(\eta_n^+ + b_n(1+a_n)/(\theta n))$,

$$s_n \mathbb{E}_n\left[\min\left\{1,\ \mathcal{P}_n^+(Y_1) + \frac{\gamma_n^+}{E[W^+]+\Delta_n}\bar{W}_{Y_1}^+\right\}\right]$$
$$\le \frac{s_n}{\Lambda_n^+/n}\left(g_+(c_n) + \Delta_n + c_n\mathcal{E}_n + c_n\gamma_n^+\right).$$

This completes the proof. □

We are now ready to prove Theorem 3.6.

Proof of Theorem 3.6. Fix $c_n, s_n \ge 1$ with $s_n \le a_n \wedge b_n$, and the event M_m as in Proposition 4.4. Now write

$$\mathbb{P}_{n,i}(\tau \le 2k) \le \mathbb{P}_{n,i}(\tau \le 2k, M_k) + \mathbb{P}_{n,i}(M_k^c)$$
$$= \mathbb{P}_{n,i}(\tau = 0, M_k) + \sum_{m=1}^k \left\{\mathbb{P}_{n,i}(\tau = 2m-1, M_k) + \mathbb{P}_{n,i}(\tau = 2m, M_k)\right\}$$
$$+ \mathbb{P}_{n,i}(M_k^c)$$
$$\le \mathbb{P}_{n,i}(\tau = 0) + \sum_{m=1}^k \left\{\mathbb{P}_{n,i}(\tau = 2m-1, M_{m-1}) + \mathbb{P}_{n,i}(\tau = 2m, M_m)\right\}$$
$$+ \mathbb{P}_{n,i}(M_k^c),$$

where in the last inequality we used the observation that $M_{m+1} \subseteq M_m$ for all $m \geq 1$. Now use Lemma 4.3 and Proposition 4.4 to obtain that

$$\begin{aligned}\mathbb{P}_{n,i}(\tau \leq 2k) &\leq 1(W_i^+ > a_n) + \mathcal{P}_n^+(i) + \bar{W}_i^+(\eta_n^+ + b_n/(\theta n)) \\ &\quad + \sum_{m=1}^{k} \frac{s_n}{\Lambda_n^+/n}\left(2g_+(c_n) + 2\Delta_n + 2c_n\mathcal{E}_n + c_n(\gamma_n^- + \gamma_n^+)\right) + \mathbb{P}_{n,i}(M_k^c) \\ &\leq 1(W_i^+ > a_n) + \mathcal{P}_n^+(i) + (\mathcal{H}_n/\theta)\bar{W}_i^+(\Delta_n + g_-(b_n) + a_n b_n/n) \\ &\quad + \mathcal{H}_n k s_n\left(g_+(c_n) + c_n\left(\mathcal{E}_n + \Delta_n + g_+(a_n) + g_-(b_n) + a_n b_n/n\right)\right) \\ &\quad + \mathbb{P}_{n,i}(|\hat{T}_k| \vee |\hat{I}_k| > s_n),\end{aligned}$$

where

$$\mathcal{H}_n = \frac{2}{\Lambda_n^+/n}(1 + \Delta_n/\theta)^2(3 + a_n^{-1} + b_n^{-1} + c_n^{-1}). \quad \square$$

4.2. Proof of Theorem 3.7

In view of Theorem 3.6, the proof of Theorem 3.7 reduces to showing that we can choose a_n, b_n, c_n, s_n such that the bound in Theorem 3.6 converges to zero. The only term that is not yet explicit is

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{P}_{n,i}\left(|\hat{T}_k| \vee |\hat{I}_k| > s_n\right),$$

which we will first write in terms of a marked Galton–Watson process that does not depend on the type sequence $\mathscr{W}_n = \{\mathbf{W}_i : 1 \leq i \leq n\}$. To do this we need two preliminary results, the first of which shows the convergence of the degree vectors $(\hat{N}_\emptyset, \hat{D}_\emptyset)$ and $(\hat{N}_1, \hat{D}_1)$ in the total variation distance.

Lemma 4.5. *Define* $(\bar{W}^-, \bar{W}^+) = (W^- \wedge b_n, W^+ \wedge a_n)$ *and consider the joint distributions*

$$\begin{aligned}\mathbb{P}_n\left(\hat{N}_\emptyset = m,\ \hat{D}_\emptyset = k\right) &= \sum_{s=1}^{n} p(m; \Lambda_n^+ \bar{W}_s^-/(\theta n)) \cdot p(k; \Lambda_n^- \bar{W}_s^+/(\theta n)) \cdot \frac{1}{n}, \\ \mathbb{P}_n\left(\hat{N}_1 = m, \hat{D}_1 = k\right) &= \sum_{s=1}^{n} p(m; \Lambda_n^+ \bar{W}_s^-/(\theta n)) \cdot p(k; \Lambda_n^- \bar{W}_s^+/(\theta n)) \cdot \frac{\bar{W}_s^+}{\Lambda_n^+}, \\ P\left(\bar{\mathcal{N}}_0 = m,\ \bar{\mathcal{D}}_0 = k\right) &= E\left[p(m; E[\bar{W}^+]\bar{W}^-/\theta) \cdot p(k; E[\bar{W}^-]\bar{W}^+/\theta)\right], \\ P\left(\bar{\mathcal{N}} = m, \bar{\mathcal{D}} = k\right) &= \frac{1}{E[W^+]} E\left[W^+ p(m; E[\bar{W}^+]\bar{W}^-/\theta) \cdot p(k; E[\bar{W}^-]\bar{W}^+/\theta)\right],\end{aligned}$$

for $m, k \in \mathbb{N}$, *where* $p(m; \lambda) = e^{-\lambda}\lambda^m/m!$. *Then, under Assumption 2.3(a)–(b), and for any* $c_n \geq 1$,

$$\begin{aligned}\sup_{A \subseteq \mathbb{N}^2} \left|\mathbb{P}_n\left((\hat{N}_\emptyset, \hat{D}_\emptyset) \in A\right) - P((\bar{\mathcal{N}}_0, \bar{\mathcal{D}}_0) \in A)\right| &\leq \frac{\Delta_n^2}{\theta} + 2\Delta_n, \\ \sup_{A \subseteq \mathbb{N}^2} \left|\mathbb{P}_n\left((\hat{N}_1, \hat{D}_1) \in A\right) - P((\bar{\mathcal{N}}, \bar{\mathcal{D}}) \in A)\right| &\leq \frac{1}{E[\bar{W}^+]}\left(4c_n\Delta_n + 2g_+(a_n) + g_+(c_n)\right. \\ &\quad \left. + c_n\Delta_n^2/\theta\right),\end{aligned}$$

where g_+ *is defined as in Lemma 4.2.*

Proof. By Assumption 2.3(a)–(b) and the observations following Definition 3.5, we know that F_n converges to F in the Kantorovich–Rubinstein distance, and therefore we can pick a random vector

$$\left(W^{(-,n)}, W^{(+,n)}, W^-, W^+\right)$$

so that $(W^{(-,n)}, W^{(+,n)})$ has distribution F_n, (W^-, W^+) has distribution F, and

$$\mathbb{E}_n\left[|W^{(-,n)} - W^-| + |W^{(+,n)} - W^+|\right] = d_1(F_n, F) = \Delta_n.$$

Next, using this optimal coupling define $(\bar{W}^{(-,n)}, \bar{W}^{(+,n)}) = (W^{(-,n)} \wedge b_n, W^{(+,n)} \wedge a_n)$, $X^{(-,n)} = \Lambda_n^+ \bar{W}^{(-,n)}/(\theta n)$, $X^{(+,n)} = \Lambda_n^- \bar{W}^{(+,n)}/(\theta n)$, $X^- = E[\bar{W}^+]\bar{W}^-/\theta$, and $X^+ = E[\bar{W}^-]\bar{W}^+/\theta$. Let $\mathscr{G}_n = \sigma\left(\mathscr{F}_n \cup \sigma(W^{(-,n)}, W^{(+,n)}, W^-, W^+)\right)$.

Now note that for any $A \subseteq \mathbb{N}^2$,

$$\begin{aligned}
&\left|\mathbb{P}_n\left((\hat{N}_\emptyset, \hat{D}_\emptyset) \in A\right) - P\left((\mathcal{N}_0, \mathcal{D}_0) \in A\right)\right| \\
&\le \mathbb{E}_n\left[\left|\mathbb{P}_n\left((\hat{N}_\emptyset, \hat{D}_\emptyset) \in A \middle| \mathscr{G}_n\right) - \mathbb{P}_n\left((\bar{\mathcal{N}}_0, \bar{\mathcal{D}}_0) \in A \middle| \mathscr{G}_n\right)\right|\right] \\
&\le \mathbb{E}_n\left[\sup_{B\subseteq\mathbb{N}} \left|\mathbb{P}_n(\hat{N}_\emptyset \in B|\mathscr{G}_n) - \mathbb{P}_n(\bar{\mathcal{N}}_0 \in B|\mathscr{G}_n)\right|\right. \\
&\quad \left. + \sup_{B\subseteq\mathbb{N}} \left|\mathbb{P}_n(\hat{D}_\emptyset \in B|\mathscr{G}_n) - \mathbb{P}_n(\bar{\mathcal{D}}_0 \in B|\mathscr{G}_n)\right|\right] \\
&\le \mathbb{E}_n\left[\min\{1, |X^{(-,n)} - X^-|\} + \min\{1, |X^{(+,n)} - X^+|\}\right],
\end{aligned}$$

where in the first inequality we used the conditional independence of $(\hat{N}_\emptyset, \bar{\mathcal{N}}_0)$ and $(\hat{D}_\emptyset, \bar{\mathcal{D}}_0)$ given $\mathscr{G}_n$, and in the second one we used the observation that if $\mathrm{Poi}(\lambda)$ denotes a Poisson random variable with mean λ, then

$$\sup_{A\in\mathbb{N}} |P(\mathrm{Poi}(\mu) \in A) - P(\mathrm{Poi}(\lambda) \in A)| \le P(\mathrm{Poi}(|\mu - \lambda|) \ge 1) \le \min\{1, |\mu - \lambda|\}.$$

Moreover, since $\Lambda_n^\pm/n = \mathbb{E}_n[\bar{W}^{(\pm,n)}]$ and $\mathbb{E}_n[|\bar{W}^{(\pm,n)} - \bar{W}^\pm|] \le \mathbb{E}_n[|W^{(\pm,n)} - W^\pm|]$, we have that

$$\begin{aligned}
\mathbb{E}_n\left[|X^{(-,n)} - X^-|\right] &\le \frac{\Lambda_n^+}{\theta n}\mathbb{E}_n\left[|\bar{W}^{(-,n)} - \bar{W}^-|\right] + E[\bar{W}^-]\left|\frac{\Lambda_n^+}{\theta n} - \frac{E[\bar{W}^+]}{\theta}\right| \\
&\le \left(\frac{\mathbb{E}_n[\bar{W}^{(+,n)} - \bar{W}^+]}{\theta} + \frac{E[\bar{W}^+]}{\theta}\right)\mathbb{E}_n\left[|\bar{W}^{(-,n)} - \bar{W}^-|\right] \\
&\quad + \frac{E[\bar{W}^-]}{\theta}\left|\mathbb{E}_n[\bar{W}^{(+,n)}] - E[\bar{W}^+]\right| \\
&\le \left(\frac{\mathbb{E}_n[|W^{(+,n)} - W^+|]}{\theta} + \frac{E[W^+]}{\theta}\right)\mathbb{E}_n[|W^{(-,n)} - W^-|] \\
&\quad + \frac{E[W^-]}{\theta}\mathbb{E}_n[|W^{(+,n)} - W^+|],
\end{aligned}$$

and similarly,

$$\begin{aligned}
\mathbb{E}_n\left[|X^{(+,n)} - X^+|\right] &\le \left(\frac{\mathbb{E}_n[|W^{(-,n)} - W^-|]}{\theta} + \frac{E[W^-]}{\theta}\right)\mathbb{E}_n[|W^{(+,n)} - W^+|] \\
&\quad + \frac{E[W^-]}{\theta}\mathbb{E}_n[|W^{(+,n)} - W^+|].
\end{aligned}$$

Combining the two bounds we obtain

$$\mathbb{E}_n\left[|X^{(-,n)} - X^-| + |X^{(+,n)} - X^+|\right] \le \frac{\Delta_n^2}{\theta} + 2\Delta_n.$$

Taking the supremum over all A gives the first result.

For the second result we start by noting that for any $A \subseteq \mathbb{N}^2$,

$$\mathbb{P}_n\left((\hat{N}_1, \hat{D}_1) \in A\right) = \mathbb{E}_n\left[\frac{\bar{W}^{(+,n)}}{\mathbb{E}_n[\bar{W}^{(+,n)}]}\mathbb{P}_n((\hat{N}_\emptyset, \hat{D}_\emptyset) \in A|\mathscr{G}_n)\right]$$

and

$$P\left((\bar{\mathcal{N}}, \bar{\mathcal{D}}) \in A\right) = E\left[\frac{W^+}{E[W^+]}P((\bar{\mathcal{N}}_0, \bar{\mathcal{D}}_0) \in A|W^-, W^+)\right].$$

Hence, for any $A \subseteq \mathbb{N}^2$ and any $c_n \ge 1$,

$$\begin{aligned}
&\left|\mathbb{P}_n\left((\hat{N}_1, \hat{D}_1) \in A\right) - P\left((\bar{\mathcal{N}}, \bar{\mathcal{D}}) \in A\right)\right| \\
&\le \left|\mathbb{E}_n\left[\frac{\bar{W}^{(+,n)}}{\mathbb{E}_n[\bar{W}^{(+,n)}]}\mathbb{P}_n((\hat{N}_\emptyset, \hat{D}_\emptyset) \in A|\mathscr{G}_n)\right] - \mathbb{E}_n\left[\frac{\bar{W}^+}{E[\bar{W}^+]}\mathbb{P}_n((\hat{N}_\emptyset, \hat{D}_\emptyset) \in A|\mathscr{G}_n)\right]\right| \\
&\quad + \left|\mathbb{E}_n\left[\frac{\bar{W}^+}{E[\bar{W}^+]}\mathbb{P}_n((\hat{N}_\emptyset, \hat{D}_\emptyset) \in A|\mathscr{G}_n)\right] - \mathbb{E}_n\left[\frac{\bar{W}^+}{E[\bar{W}^+]}\mathbb{P}_n\left((\bar{\mathcal{N}}_0, \bar{\mathcal{D}}_0) \in A\middle|\mathscr{G}_n\right)\right]\right| \\
&\quad + \left|E\left[\frac{\bar{W}^+}{E[\bar{W}^+]}P\left((\bar{\mathcal{N}}_0, \bar{\mathcal{D}}_0) \in A\middle|W^-, W^+\right)\right]\right. \\
&\quad \left. - E\left[\frac{W^+}{E[W^+]}P((\bar{\mathcal{N}}_0, \bar{\mathcal{D}}_0) \in A|W^-, W^+)\right]\right| \\
&\le \mathbb{E}_n\left[\left|\frac{\bar{W}^{(+,n)}}{\mathbb{E}_n[\bar{W}^{(+,n)}]} - \frac{\bar{W}^+}{E[\bar{W}^+]}\right|\right] \\
&\quad + \mathbb{E}_n\left[\frac{\bar{W}^+}{E[\bar{W}^+]}\left|\mathbb{P}_n((\hat{N}_\emptyset, \hat{D}_\emptyset) \in A|\mathscr{G}_n) - \mathbb{P}_n\left((\bar{\mathcal{N}}_0, \bar{\mathcal{D}}_0) \in A\middle|\mathscr{G}_n\right)\right|\right] \\
&\quad + E\left[\left|\frac{\bar{W}^+}{E[\bar{W}^+]} - \frac{W^+}{E[W^+]}\right|\right] \\
&\le \frac{\mathbb{E}_n\left[|E[\bar{W}^+]\bar{W}^{(+,n)} - \mathbb{E}_n[\bar{W}^{(+,n)}]\bar{W}^+|\right]}{\mathbb{E}_n[\bar{W}^{(+,n)}]E[\bar{W}^+]} + \frac{E[|\bar{W}^+E[W^+] - W^+E[\bar{W}^+]|]}{E[\bar{W}^+]E[W^+]} \\
&\quad + \mathbb{E}_n\left[\frac{\bar{W}^+}{E[\bar{W}^+]}\left(\min\{1, |X^{(-,n)} - X^-|\} + \min\{1, |X^{(+,n)} - X^+|\}\right)\right] \\
&\le \frac{2\mathbb{E}_n[|\bar{W}^{(+,n)} - \bar{W}^+|]}{E[\bar{W}^+]} + \frac{2E[(W^+ - a_n)^+]}{E[W^+]} \\
&\quad + \mathbb{E}_n\left[\frac{\bar{W}^+}{E[\bar{W}^+]}\left(\min\{1, |X^{(-,n)} - X^-|\} + \min\{1, |X^{(+,n)} - X^+|\}\right)\right] \\
&\le \frac{2\Delta_n}{E[\bar{W}^+]} + \frac{2g_+(a_n)}{E[W^+]} + \frac{c_n}{E[\bar{W}^+]}\mathbb{E}_n\left[|X^{(-,n)} - X^-| + |X^{(+,n)} - X^+|\right]
\end{aligned}$$

$$+\frac{E[(\bar{W}^+ - c_n)^+]}{E[\bar{W}^+]}$$

$$\leq \frac{1}{E[\bar{W}^+]}\left(2\Delta_n + 2g_+(a_n) + g_+(c_n) + 2c_n\Delta_n + c_n\Delta_n^2/\theta\right).$$

This completes the proof. □

The second technical result prior to the proof of Theorem 3.7 states the convergence in total variation of the processes $|\hat{\mathcal{T}}_k|$ and $|\hat{\mathcal{I}}_k|$. Note that the delayed marked Galton–Watson process appearing in the lemma still depends on n via the truncation of $W^+ \wedge a_n$ and $W^- \wedge b_n$, but does not depend on the type sequence $\mathscr{W}_n = \{\mathbf{W}_i : 1 \leq i \leq n\}$. In particular, by monotonicity of the Poisson distribution in its parameter, we have that under Assumption 2.3(a)–(b),

$$(\bar{\mathcal{N}}_0, \bar{\mathcal{D}}_0) \nearrow (\mathcal{N}_0, \mathcal{D}_0) \quad \text{a.s.} \qquad \text{and} \qquad (\bar{\mathcal{N}}, \bar{\mathcal{D}}) \nearrow (\mathcal{N}, \mathcal{D}) \quad \text{a.s,}$$

as $n \to \infty$, for well-defined random vectors $(\mathcal{N}_0, \mathcal{D}_0)$ and $(\mathcal{N}, \mathcal{D})$. Moreover, under Assumption 2.3(a)–(b) we have that $E[\mathcal{N}_0 + \mathcal{D}_0] < \infty$, although it is possible to have $E[\mathcal{N} + \mathcal{D}] = \infty$. If the latter happens, the probability $P\left(|\bar{\mathcal{T}}_k| \vee |\bar{\mathcal{I}}_k| > s_n\right)$ will still converge to zero as $s_n \to \infty$ for any fixed $k \in \mathbb{N}_+$, however, it may do so very slowly.

Lemma 4.6. *Under Assumption 2.3(a)–(b) we have that for any fixed $k \geq 1$ and any $c_n, s_n \geq 1$,*

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{P}_{n,i}\left(|\hat{\mathcal{T}}_k| \vee |\hat{\mathcal{I}}_k| > s_n\right) \leq P\left(|\bar{\mathcal{T}}_k| \vee |\bar{\mathcal{I}}_k| > s_n\right) + \frac{\Delta_n^2}{\theta} + 2\Delta_n$$

$$+ \frac{ks_n}{E[W^+ \wedge a_n]}\left(4c_n\Delta_n + 2g_+(a_n) + g_+(c_n) + c_n\Delta_n^2/\theta\right),$$

where $|\bar{\mathcal{T}}_k| = 1 + \sum_{m=1}^{k-1} |\bar{\mathcal{A}}_m|$, $|\bar{\mathcal{I}}_k| = \sum_{m=0}^{k} \sum_{\mathbf{i} \in \bar{\mathcal{A}}_m} \bar{\mathcal{D}}_{\mathbf{i}}$, and $\bar{\mathcal{A}}_m$ is the set of individuals in the mth generation of a delayed marked Galton–Watson process whose root has offspring/mark distributed as $(\bar{\mathcal{N}}_0, \bar{\mathcal{D}}_0)$ and all other nodes have offspring/mark distributed as $(\bar{\mathcal{N}}, \bar{\mathcal{D}})$, as defined in Lemma 4.5.

Proof. We start by noting that under the measure $\mathbb{P}_n(\cdot) = n^{-1}\sum_{i=1}^{n} \mathbb{P}_{n,i}(\cdot)$, $|\hat{\mathcal{T}}_k|$ and $|\hat{\mathcal{I}}_k|$ denote the total population and the sum of all the marks, up to generation k, on a marked Galton–Watson process whose offspring/mark distribution is that of $(\hat{N}_\emptyset, \hat{D}_\emptyset)$ for the root node and $(\hat{N}_1, \hat{D}_1)$ for all other nodes, as defined in Lemma 4.5. Next, let $(\hat{N}_\emptyset, \hat{D}_\emptyset, \bar{\mathcal{N}}_0, \bar{\mathcal{D}}_0)$ and $(\hat{N}_1, \hat{D}_1, \bar{\mathcal{N}}, \bar{\mathcal{D}})$ be couplings satisfying

$$\sup_{A \subseteq \mathbb{N}^2} \left|\mathbb{P}_n((\hat{N}_\emptyset, \hat{D}_\emptyset) \in A) - P((\bar{\mathcal{N}}_0, \bar{\mathcal{D}}_0) \in A)\right| = \mathbb{P}_n((\hat{N}_\emptyset, \hat{D}_\emptyset) \neq (\bar{\mathcal{N}}_0, \bar{\mathcal{D}}_0))$$

and

$$\sup_{A \subseteq \mathbb{N}^2} \left|\mathbb{P}_n((\hat{N}_1, \hat{D}_1) \in A) - P((\bar{\mathcal{N}}, \bar{\mathcal{D}}) \in A)\right| = \mathbb{P}_n((\hat{N}_1, \hat{D}_1) \neq (\bar{\mathcal{N}}, \bar{\mathcal{D}})),$$

which are guaranteed to exist (see, e.g., Theorem 2.12 in [21]). Construct the two marked Galton–Watson processes simultaneously using this optimal coupling of the degree/mark vectors and let $\sigma = \inf\{m \geq 0 : (\hat{N}_\mathbf{i}, \hat{D}_\mathbf{i}) \neq (\bar{\mathcal{N}}_\mathbf{i}, \bar{\mathcal{D}}_\mathbf{i})$ for some $\mathbf{i} \in \bar{\mathcal{A}}_m\}$.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

Now note that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\mathbb{P}_{n,i}\left(|\hat{\mathcal{T}}_k|\vee|\hat{\mathcal{I}}_k|>s_n\right)&=\mathbb{P}_n\left(|\hat{\mathcal{T}}_k|\vee|\hat{\mathcal{I}}_k|>s_n\right)\\&\le P\left(|\bar{\mathcal{T}}_k|\vee|\bar{\mathcal{I}}_k|>s_n\right)+\mathbb{P}_n\left(|\hat{\mathcal{T}}_k|\vee|\hat{\mathcal{I}}_k|>s_n\ge|\bar{\mathcal{T}}_k|\vee|\bar{\mathcal{I}}_k|\right)\\&\le P\left(|\bar{\mathcal{T}}_k|\vee|\bar{\mathcal{I}}_k|>s_n\right)+\sum_{m=0}^{k}\mathbb{P}_n\left(|\bar{\mathcal{T}}_k|\vee|\bar{\mathcal{I}}_k|\le s_n,\ \sigma=m\right).\end{aligned}$$

To analyze each of the probabilities in the sum let $\mathcal{F}_m = \sigma\left((\hat{N}_{\mathbf{i}}, \hat{D}_{\mathbf{i}}, \bar{\mathcal{N}}_{\mathbf{i}}, \bar{\mathcal{D}}_{\mathbf{i}}) : \mathbf{i}\in\hat{A}_l,\ 0\le l\le m\right)$ and note that for $0\le m\le k$:

$$\begin{aligned}\mathbb{P}_n\left(|\bar{\mathcal{T}}_k|\vee|\bar{\mathcal{I}}_k|\le s_n,\ \sigma=m\right)&\le\mathbb{P}_n\left(|\bar{\mathcal{A}}_m|\le s_n,\ \sigma=m\right)\\&=\mathbb{E}_n\left[1(|\bar{\mathcal{A}}_m|\le s_n)\,\mathbb{P}_n(\sigma=m|\mathcal{F}_{m-1})\right]\\&\le\mathbb{E}_n\left[1(|\bar{\mathcal{A}}_m|\le s_n)\,\mathbb{P}_n\right.\\&\quad\left.\times\left(\bigcup_{\mathbf{i}\in\bar{\mathcal{A}}_m}\{(\hat{N}_{\mathbf{i}},\hat{D}_{\mathbf{i}})\ne(\bar{\mathcal{N}}_{\mathbf{i}},\bar{\mathcal{D}}_{\mathbf{i}})\}\,\middle|\,\mathcal{F}_{m-1}\right)\right]\\&\le\mathbb{E}_n\left[1(|\bar{\mathcal{A}}_m|\le s_n)\sum_{\mathbf{i}\in\bar{\mathcal{A}}_m}\mathbb{P}_n\left((\hat{N}_{\mathbf{i}},\hat{D}_{\mathbf{i}})\ne(\bar{\mathcal{N}}_{\mathbf{i}},\bar{\mathcal{D}}_{\mathbf{i}})\right)\right]\\&=\mathbb{E}_n\left[1(|\bar{\mathcal{A}}_m|\le s_n)(|\bar{\mathcal{A}}_m|-1)\mathbb{P}_n\left((\hat{N}_1,\hat{D}_1)\ne(\bar{\mathcal{N}},\bar{\mathcal{D}})\right)\right]\\&\quad\times 1(m\ge1)\\&\quad+\mathbb{P}_n\left((\hat{N}_\emptyset,\hat{D}_\emptyset)\ne(\bar{\mathcal{N}}_0,\bar{\mathcal{D}}_0)\right)\\&\le s_n\mathbb{P}_n\left((\hat{N}_1,\hat{D}_1)\ne(\bar{\mathcal{N}},\bar{\mathcal{D}})\right)1(m\ge1)\\&\quad+\mathbb{P}_n\left((\hat{N}_\emptyset,\hat{D}_\emptyset)\ne(\bar{\mathcal{N}}_0,\bar{\mathcal{D}}_0)\right).\end{aligned}$$

It follows that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\mathbb{P}_{n,i}\left(|\hat{\mathcal{T}}_k|\vee|\hat{\mathcal{I}}_k|>s_n\right)&\le P\left(|\bar{\mathcal{T}}_k|\vee|\bar{\mathcal{I}}_k|>s_n\right)+\mathbb{P}_n\left((\hat{N}_\emptyset,\hat{D}_\emptyset)\ne(\bar{\mathcal{N}}_0,\bar{\mathcal{D}}_0)\right)\\&\quad+ks_n\mathbb{P}_n\left((\hat{N}_1,\hat{D}_1)\ne(\bar{\mathcal{N}},\bar{\mathcal{D}})\right).\end{aligned}$$

The conclusion now follows from Lemma 4.5. □

We now use these two results to prove Theorem 3.7.

Proof of Theorem 3.7. Let $(W^{(-,n)}, W^{(+,n)})$ be distributed according to F_n. Then, by Theorem 3.6 and Lemma 4.6 we have that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\mathbb{P}_{n,i}(\tau\le2k)&\le\mathbb{P}_n(W^{(-,n)}>a_n)+\mathcal{E}_n+(\mathcal{H}_n/\theta)\frac{\Lambda_n^+}{n}(\Delta_n+g_-(b_n)+a_nb_n/n)\\&\quad+\mathcal{H}_nks_n\left(g_+(c_n)+c_n(\mathcal{E}_n+\Delta_n+g_+(a_n)+g_-(b_n)+a_nb_n/n)\right)\end{aligned}$$

$$
\begin{aligned}
&+ P\left(|\bar{\mathcal{T}}_k| \vee |\bar{\mathcal{I}}_k| > s_n\right) + \frac{\Delta_n^2}{\theta} + 2\Delta_n \\
&+ \frac{k s_n}{E[W^+ \wedge a_n]} \left(4c_n\Delta_n + 2g_+(a_n) + g_+(c_n) + c_n\Delta_n^2/\theta\right) \\
&\leq \mathbb{P}_n(W^{(-,n)} > a_n) + P\left(|\bar{\mathcal{T}}_k| \vee |\bar{\mathcal{I}}_k| > s_n\right) \\
&\quad + \mathcal{K}_n k s_n \left(g_+(c_n) + c_n(\mathcal{E}_n + \Delta_n + g_+(a_n) + g_-(b_n) + a_n b_n/n)\right),
\end{aligned}
$$

with

$$
\mathcal{K}_n = \frac{2}{\mathbb{E}_n[W^{(-,n)} \wedge a_n] \wedge E[W^+ \wedge a_n]} (1 + \Delta_n/\theta)^2 (3 + a_n^{-1} + b_n^{-1} + c_n^{-1}),
$$

and $\bar{\mathcal{T}}_k, \bar{\mathcal{I}}_k$ defined as in Lemma 4.6. As argued right before the statement of Lemma 4.6, $\lim_{n\to\infty} |\bar{\mathcal{T}}_k| \vee |\bar{\mathcal{I}}_k| < \infty$ a.s. for any fixed $k \in \mathbb{N}_+$, and therefore,

$$
\lim_{n\to\infty} P\left(|\bar{\mathcal{T}}_k| \vee |\bar{\mathcal{I}}_k| > s_n\right) = 0
$$

provided $s_n \to \infty$. Clearly, $\lim_{n\to\infty} \mathbb{P}_n(W^{(-,n)} > a_n) \leq \lim_{n\to\infty} a_n^{-1}\mathbb{E}_n[W^{(-,n)}] = 0$ for any $a_n \to \infty$, so it only remains to show that we can pick a_n, b_n, c_n, s_n such that $\min\{a_n, b_n, c_n, s_n\} \xrightarrow{P} \infty$ and

$$
s_n \left(g_+(c_n) + c_n(\mathcal{E}_n + \Delta_n + g_+(a_n) + g_-(b_n) + a_n b_n/n)\right) \xrightarrow{P} 0 \tag{15}
$$

as $n \to \infty$. To this end, choose $a_n = b_n = n^{(1-\epsilon)/2}$ for some $0 < \epsilon < 1$,

$$
c_n = (\mathcal{E}_n + \Delta_n + g_+(a_n) + g_-(b_n) + a_n b_n/n)^{-1/2},
$$

and $s_n = (g_+(c_n) + c_n^{-1})^{-1/2}$. Assumption 2.3(a)–(c) guarantee that $c_n \xrightarrow{P} \infty$ as $n \to \infty$ and our choice of s_n ensures (15) holds. □

4.3. Proof of Theorem 2.4

We now give a short proof of Theorem 2.4 using Theorem 3.6 and Lemma 4.5. A direct proof is possible, but would involve repeating some of the arguments used earlier.

Proof of Theorem 2.4. Start by noting that $(Z^-, Z^+) \overset{\mathcal{D}}{=} (\mathcal{N}_0, \mathcal{D}_0)$, as defined in Lemma 4.5, and therefore,

$$
\begin{aligned}
&\sup_{A\subseteq\mathbb{N}^2} \left|\mathbb{P}_n((D_\xi^-, D_\xi^+) \in A) - P((Z^-, Z^+) \in A)\right| \\
&\leq \sup_{A\subseteq\mathbb{N}^2} \left|\mathbb{P}_n((D_\xi^-, D_\xi^+) \in A) - \mathbb{P}_n((\hat{N}_\emptyset, \hat{D}_\emptyset) \in A)\right| \\
&\quad + \sup_{A\subseteq\mathbb{N}^2} \left|\mathbb{P}_n((\hat{N}_\emptyset, \hat{D}_\emptyset) \in A) - P((\mathcal{N}_0, \mathcal{D}_0) \in A)\right| \\
&\leq \frac{1}{n}\sum_{i=1}^n \mathbb{P}_{n,i}(\tau \leq 1) + \sup_{A\subseteq\mathbb{N}^2} \left|\mathbb{P}_n((\hat{N}_\emptyset, \hat{D}_\emptyset) \in A) - P((\mathcal{N}_0, \mathcal{D}_0) \in A)\right|.
\end{aligned}
$$

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

Now use Lemma 4.3 and Proposition 4.4 as in the proof of Theorem 3.6 to obtain that

$$\begin{aligned}\mathbb{P}_{n,i}(\tau \le 1) &\le \mathbb{P}_{n,i}(\tau = 0) + \mathbb{P}_{n,i}(\tau = 1, M_0) + \mathbb{P}_{n,i}(M_0^c)\\ &\le 1(W_i^+ > a_n) + \mathcal{P}_{n,i}(i) + (\eta_n^+ + b_n/(\theta n))\bar{W}_i^+\\ &\quad + \frac{s_n}{\Lambda_n^+/n}(g_+(c_n) + \Delta_n + c_n\mathcal{E}_n + c_n\gamma_n^-) + \mathbb{P}_{n,i}(|\hat{I}_0| > s_n).\end{aligned}$$

Note that under $\mathbb{P}_n(\cdot) = n^{-1}\sum_{i=1}^n \mathbb{P}_{n,i}(\cdot)$ we have that $|\hat{I}_0| \overset{\mathcal{D}}{=} \hat{D}_\emptyset$, so following the same steps as in the proof of Theorem 3.7, we obtain

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n \mathbb{P}_{n,i}(\tau \le 1) &\le \mathbb{P}_n(W^{(-,n)} > a_n) + \mathbb{P}_n(\hat{D}_\emptyset > s_n)\\ &\quad + \mathcal{K}_n s_n\left(g_+(c_n) + c_n(\mathcal{E}_n + \Delta_n + g_+(a_n) + g_-(b_n) + a_n b_n/n)\right),\end{aligned}$$

where $(W^{(-,n)}, W^{(+,n)})$ is distributed according to F_n and $\mathcal{K}_n$ is bounded. Moreover, this bound converges to zero as $n \to \infty$ for the same choice of a_n, b_n, c_n, s_n used in the proof of Theorem 3.7. Finally, Lemma 4.5 gives that

$$\sup_{A\subseteq\mathbb{N}^2}\left|\mathbb{P}_n((\hat{N}_\emptyset, \hat{D}_\emptyset) \in A) - P((\mathcal{N}_0, \mathcal{D}_0) \in A)\right| \le \frac{\Delta_n^2}{\theta} + 2\Delta_n \overset{P}{\to} 0$$

as $n \to \infty$. We conclude that

$$\sup_{A\subseteq\mathbb{N}^2}\left|\mathbb{P}_n((D_\xi^-, D_\xi^+) \in A) - P((Z^-, Z^+) \in A)\right| \overset{P}{\to} 0 \qquad n \to \infty.$$

To obtain the convergence of the means, let $(W^{(-,n)}, W^{(+,n)})$ and $(\hat{W}^{(-,n)}, \hat{W}^{(+,n)})$ be conditionally i.i.d. (given $\mathscr{F}_n$) vectors have distribution F_n and note that

$$\begin{aligned}\left|\mathbb{E}_n[D_\xi^\pm] - E[Z^\pm]\right| &\le \left|\frac{1}{n}\sum_{i=1}^n \sum_{1\le j\le n, j\ne i}\left(p_{ij}^{(n)} - (r_{ij}^{(n)} \wedge 1)\right)\right|\\ &\quad + \frac{1}{n}\sum_{i=1}^n \sum_{1\le j\le n, j\ne i} r_{ij}^{(n)} 1(r_{ij}^{(n)} > 1)\\ &\quad + \left|\frac{1}{n}\sum_{i=1}^n \sum_{1\le j\le n, j\ne i} r_{ij}^{(n)} - E[Z^\pm]\right|\\ &\le \mathcal{E}_n + \frac{1}{\theta}\mathbb{E}_n\left[W^{(+,n)}\hat{W}^{(-,n)} 1(W^{(+,n)}\hat{W}^{(-,n)} > \theta n)\right]\\ &\quad + \frac{1}{\theta}\left|\mathbb{E}_n[W^{(-,n)}]\mathbb{E}_n[W^{(+,n)}] - E[W^-]E[W^+]\right|,\end{aligned}$$

where $\mathcal{E}_n \overset{P}{\to} 0$ by Assumption 2.3(c) and $\mathbb{E}_n[W^{(-,n)}]\mathbb{E}_n[W^{(+,n)}] \overset{P}{\to} E[W^+]E[W^-]$ by Assumption 2.3(b). For the middle term note that

$$\begin{aligned}&\mathbb{E}_n\left[W^{(+,n)}\hat{W}^{(-,n)} 1(W^{(+,n)}\hat{W}^{(-,n)} > \theta n)\right]\\ &\le \mathbb{E}_n\left[W^{(+,n)}\hat{W}^{(-,n)}\left(1(W^{(+,n)} > \sqrt{\theta n}) + 1(\hat{W}^{(-,n)} > \sqrt{\theta n})\right)\right]\\ &= \mathbb{E}_n[W^{(-,n)}]\mathbb{E}_n\left[W^{(-,n)} 1(W^{(-,n)} > \sqrt{\theta n})\right] + \mathbb{E}_n[W^{(+,n)}]\mathbb{E}_n\left[W^{(+,n)} 1(W^{(+,n)} > \sqrt{\theta n})\right],\end{aligned}$$

which also converges in probability to zero as $n \to \infty$ since $E[W^- + W^+] < \infty$.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

The result for the mixed expectation is a consequence of Assumption 2.3(d) since

$$\mathbb{E}_n\left[D_\xi^- D_\xi^+\right] = \frac{1}{n}\sum_{i=1}^{n}\sum_{1\le j\le n, j\neq i}\sum_{1\le k\le n, k\neq i} p_{ji}^{(n)} p_{ik}^{(n)} \xrightarrow{P} \frac{E[W^-W^+]E[W^-]E[W^+]}{\theta^2}$$
$$= E[Z^-Z^+],$$

as $n \to \infty$. This completes the proof. □

4.4. Proof of Theorem 3.8

In this section we prove Theorem 3.8, which establishes the convergence of $\hat{R}_\emptyset^{(n,k_n)}$ to the attracting endogenous solution to (5), under Assumption 3.1. The main step in the proof of Theorem 3.8 consists in showing that the vectors $\{(\hat{N}_\mathbf{i}, \hat{\mathcal{Q}}_\mathbf{i}, \hat{\mathcal{C}}_\mathbf{i}) : \mathbf{i} \in \mathcal{U}\}$ converge, in the Kantorovich–Rubinstein metric, to the distribution of $(\mathcal{N}, \mathcal{Q}, \mathcal{C})$ defined in Theorem 3.3, with $\mathcal{C}$ independent of $(\mathcal{N}, \mathcal{Q})$. To simplify the proof of Theorem 3.8, we show this convergence separately.

Throughout the section, for probability measures ϕ, χ in $\mathbb{R}^d$, we interchangeably use the notation $d_1(\phi, \chi) = d_1(F, G)$ to denote the Kantorovich–Rubinstein distance between ϕ and χ, where F and G are the cumulative distribution functions of ϕ and χ, respectively.

Theorem 4.7. *Define $G_n^*(m,q) = \mathbb{P}_n(\hat{N}_\emptyset \le m,\ \hat{\mathcal{Q}}_\emptyset \le q)$ and $G_n(m,q,t) = \mathbb{P}_n(\hat{N}_\mathbf{i} \le m,\ \hat{\mathcal{Q}}_\mathbf{i} \le q,\ \hat{\mathcal{C}}_\mathbf{i} \le t)$ for $\mathbf{i} \neq \emptyset$, according to (13) and (14), respectively. Define $G^*(m,q) = P(\mathcal{N}_0 \le m, \mathcal{Q}_0 \le q)$ and $G(m,q,t) = P(\mathcal{N} \le m, \mathcal{Q} \le q, \mathcal{C} \le t)$, for $m \in \mathbb{N}$, $q, t \in \mathbb{R}$, according to:*

$$P(\mathcal{N}_0 = m, \mathcal{Q}_0 \in dq) = E\left[1(Q \in dq)\, p(m; E[W^+]W^-/\theta)\right],$$
$$P(\mathcal{N} = m, \mathcal{Q} \in dq, \mathcal{C} \in dt) = E\left[\frac{W^+}{E[W^+]} 1\left(Q \in dq, \zeta/(Z^+ + 1) \in dt\right) \times p(m; E[W^+]W^-/\theta)\right],$$

where Z^+ is a mixed Poisson random variable with mixing parameter $E[W^-]W^+/\theta$ and $p(m;\lambda) = e^{-\lambda}\lambda^m/m!$. Then, under Assumption 3.1(a)–(b), we have that

$$d_1(G_n^*, G^*) \xrightarrow{P} 0 \qquad \text{and} \qquad (\hat{N}_1,\ \hat{\mathcal{Q}}_1, \hat{\mathcal{C}}_1) \Rightarrow (\mathcal{N}, \mathcal{Q}, \mathcal{C})$$

as $n \to \infty$. Moreover, if Assumption 3.1(a), (b), (e) hold, then

$$d_1(G_n, G) \xrightarrow{P} 0 \quad \text{as } n \to \infty.$$

Proof. We start by showing the weak convergence of the vectors $(\hat{N}_\emptyset, \hat{\mathcal{Q}}_\emptyset)$ and $(\hat{N}_1, \hat{D}_1, \hat{\mathcal{Q}}_1, \hat{\zeta}_1)$; recall that $\hat{\mathcal{C}}_1 = \hat{\zeta}_1/(\hat{D}_1 + 1)$. Let $(W^{(-,n)}, W^{(+,n)}, Q^{(n)}, \zeta^{(n)})$ be distributed according to H_n and define $(\bar{W}^{(-,n)}, \bar{W}^{(+,n)}) = (W^{(-,n)} \wedge b_n, W^{(+,n)} \wedge a_n)$. Now note that if $f : \mathbb{N} \times \mathbb{R} \to \mathbb{R}$ is bounded and continuous, then the function $J(x,q) = \sum_{m=0}^{\infty} f(m,q)p(m;x)$ is also bounded and continuous, and by Assumption 3.1(a)–(b) we have

$$\mathbb{E}_n\left[f(\hat{N}_\emptyset, \hat{\mathcal{Q}}_\emptyset)\right] = \frac{1}{n}\sum_{s=1}^{n}\sum_{m=0}^{\infty} f(m, Q_s)p(m; \Lambda_n^+ \bar{W}_s^-/(\theta n))$$
$$= \mathbb{E}_n\left[J\left(Q^{(n)}, \Lambda_n^+ \bar{W}^{(-,n)}/(\theta n)\right)\right]$$
$$\xrightarrow{P} E\left[J(Q, E[W^+]W^-/\theta)\right] = E[f(\mathcal{N}_0, \mathcal{Q}_0)],$$

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

as $n \to \infty$. For the second vector let $g : \mathbb{N} \times \mathbb{R}^2 \to \mathbb{R}$ be a bounded and continuous function and note that $J_M(x, y, q, z) = (y \wedge M) \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} p(m; x)p(k; y)g(m, q, z/(k+1))$ is bounded and continuous for any fixed $M > 0$. Then,

$$\begin{aligned}\mathbb{E}_n\left[g(\hat{N}_1, \hat{Q}_1, \hat{C}_1)\right] &= \sum_{s=1}^{n} \frac{\bar{W}_s^+}{\Lambda_n^+} \sum_{m=0}^{\infty}\sum_{k=0}^{\infty} p(m; \Lambda_n^+ \bar{W}_s^-/(\theta n))p(k; \Lambda_n^- \bar{W}_s^+/(\theta n)) \\ &\quad \times g(m, Q_s, \zeta_s/(k+1)) \\ &\le \frac{\theta n^2}{\Lambda_n^+ \Lambda_n^-} \mathbb{E}_n\left[J_M\left(\Lambda_n^+ \bar{W}^{(-,n)}/(\theta n), \Lambda_n^- \bar{W}^{(+,n)}/(\theta n), Q^{(n)}, \zeta^{(n)}\right)\right] \\ &\quad + \frac{n}{\Lambda_n^+} \mathbb{E}_n\left[(\bar{W}^{(+,n)} - M)^+\right] \sup_{m,q,z} |g(m, q, z)|,\end{aligned}$$

and, similarly,

$$\begin{aligned}\mathbb{E}_n\left[g(\hat{N}_1, \hat{Q}_1, \hat{C}_1)\right] &\ge \frac{\theta n^2}{\Lambda_n^+ \Lambda_n^-} \mathbb{E}_n\left[J_M\left(\Lambda_n^+ \bar{W}^{(-,n)}/(\theta n), \Lambda_n^- \bar{W}^{(+,n)}/(\theta n), Q^{(n)}, \zeta^{(n)}\right)\right] \\ &\quad - \frac{n}{\Lambda_n^+} \mathbb{E}_n\left[(\bar{W}^{(+,n)} - M)^+\right] \sup_{m,q,z} |g(m, q, z)|.\end{aligned}$$

It follows from Assumption 3.1(a)–(b) that the limits

$$\begin{aligned}&\frac{\theta}{E[W^+]E[W^-]} E\left[J_M(E[W^+]W^-/\theta, E[W^-]W^+/\theta, Q, \zeta)\right] \\ &\quad - \frac{E[(W^+ - M)^+]}{E[W^+]} \sup_{m,q,z} |g(m, q, z)| \\ &\le \liminf_{n\to\infty} \mathbb{E}_n\left[g(\hat{N}_1, \hat{Q}_1, \hat{C}_1)\right] \le \limsup_{n\to\infty} \mathbb{E}_n\left[g(\hat{N}_1, \hat{Q}_1, \hat{C}_1)\right] \\ &\le \frac{\theta}{E[W^+]E[W^-]} E\left[J_M(E[W^+]W^-/\theta, E[W^-]W^+/\theta, Q, \zeta)\right] \\ &\quad + \frac{E[(W^+ - M)^+]}{E[W^+]} \sup_{m,q,z} |g(m, q, z)|\end{aligned}$$

hold in probability. Taking the limit as $M \to \infty$ now yields (via the dominated convergence theorem)

$$\begin{aligned}\mathbb{E}_n\left[g(\hat{N}_1, \hat{Q}_1, \hat{C}_1)\right] &\xrightarrow{P} \frac{\theta}{E[W^+]E[W^-]} \lim_{M\to\infty} E\left[J_M(E[W^+]W^-/\theta, E[W^-]W^+/\theta, Q, \zeta)\right] \\ &= E\left[\frac{W^+}{E[W^+]} \sum_{m=0}^{\infty}\sum_{k=0}^{\infty} p(m; E[W^+]W^-/\theta)p(k; E[W^-]W^+/\theta)\right. \\ &\quad \left. \times g(m, Q, \zeta/(k+1))\right] \\ &= E\left[g(\mathcal{N}, \mathcal{Q}, \mathcal{C})\right].\end{aligned}$$

This establishes the weak convergence for both G_n^* and G_n.

To prove the convergence in the Kantorovich–Rubinstein distance recall that it suffices to show that the first absolute moments converge (see Theorem 6.9 and Definition 6.8(i) in [33]).

Under Assumption 3.1(a)–(b) we have that

$$\mathbb{E}_n\left[\hat{N}_\emptyset + |\hat{\mathcal{Q}}_\emptyset|\right] = \mathbb{E}_n\left[\frac{\Lambda_n^+ \bar{W}^{(-,n)}}{\theta n} + |Q^{(n)}|\right] \xrightarrow{P} E\left[\frac{E[W^+]W^-}{\theta} + |Q|\right]$$
$$= E\left[\mathcal{N}_0 + |\mathcal{Q}_0|\right]$$

as $n \to \infty$. We conclude that $d_1(G_n^*, G^*) \xrightarrow{P} 0$ as $n \to \infty$.

For G_n we have that

$$\mathbb{E}_n\left[\hat{N}_1 + |\hat{\mathcal{Q}}_1| + |\hat{\mathcal{C}}_1|\right]$$
$$= \mathbb{E}_n\left[\hat{N}_1 + |\hat{\mathcal{Q}}_1| + \frac{|\hat{\zeta}_1|}{\hat{D}_1 + 1}\right]$$
$$= \mathbb{E}_n\left[\frac{n\bar{W}^{(+,n)}}{\Lambda_n^+}\left(\frac{\Lambda_n^+ \bar{W}^{(-,n)}}{\theta n} + |Q^{(n)}| + \frac{|\zeta^{(n)}|}{\Lambda_n^- \bar{W}^{(+,n)}/(\theta n)}\left(1 - e^{-\Lambda_n^- \bar{W}^{(+,n)}/(\theta n)}\right)\right)\right]$$
$$= \mathbb{E}_n\left[\frac{\bar{W}^{(+,n)}\bar{W}^{(1,n)}}{\theta} + \frac{n\bar{W}^{(+,n)}|Q^{(n)}|}{\Lambda_n^+} + \frac{\theta n^2 |\zeta^{(n)}|}{\Lambda_n^- \Lambda_n^+}\left(1 - e^{-\Lambda_n^- \bar{W}^{(+,n)}/(\theta n)}\right)\right],$$

where we used the observation that if Z is Poisson with mean λ then $E[1/(Z+1)] = \frac{1}{\lambda}(1-e^{-\lambda})$. The third summand inside the expectation converges under Assumption 3.1(a)–(b), however, the first two require part (e) of the assumption. Hence, under Assumption 3.1(a),(b),(e) we have

$$\mathbb{E}_n\left[\hat{N}_1 + |\hat{\mathcal{Q}}_1| + |\hat{\mathcal{C}}_1|\right] \xrightarrow{P} E\left[\frac{W^+W^-}{\theta} + \frac{W^+|Q|}{E[W^+]}\right.$$
$$\left. + \frac{\theta|\zeta|}{E[W^-]E[W^+]}\left(1 - e^{-E[W^-]W^+/\theta}\right)\right]$$
$$= E\left[\frac{W^+}{E[W^+]}\left(\frac{E[W^+]W^-}{\theta} + |Q|\right.\right.$$
$$\left.\left. + \frac{|\zeta|}{E[W^-]W^+/\theta}\left(1 - e^{-E[W^-]W^+/\theta}\right)\right)\right]$$
$$= E\left[\mathcal{N} + |\mathcal{Q}| + |\mathcal{C}|\right]$$

as $n \to \infty$. This completes the proof. □

The proof of Theorem 3.8 will now follow from Theorem 2 in [10].

Proof of Theorem 3.8. Recall that the sequence $\{(\hat{N}_\mathbf{i}, \hat{\mathcal{Q}}_\mathbf{i}, \hat{\mathcal{C}}_\mathbf{i}) : \mathbf{i} \in \mathcal{U}, \mathbf{i} \neq \emptyset\}$ consists of conditionally i.i.d. vectors, given $\mathscr{F}_n$, with $(\hat{N}_\emptyset, \hat{\mathcal{Q}}_\emptyset, \hat{\mathcal{C}}_\emptyset)$ conditionally independent of this sequence. To simplify the notation let $(\hat{\mathcal{Q}}, \hat{N}, \hat{\mathcal{C}}) \overset{\mathcal{D}}{=} (\hat{\mathcal{Q}}_1, \hat{N}_1, \hat{\mathcal{C}}_1)$. Define ϕ_n to be the probability measure of the vector

$$(\hat{\mathcal{C}}\hat{\mathcal{Q}}, \hat{\mathcal{C}}1(\hat{N} \geq 1), \hat{\mathcal{C}}1(\hat{N} \geq 2), \dots)$$

and let ϕ_n^* denote the probability measure of $(\hat{\mathcal{Q}}_\emptyset, \hat{N}_\emptyset)$. Similarly, define ϕ and ϕ^* to be the probability measures of vectors

$$(\mathcal{C}\mathcal{Q}, \mathcal{C}1(\mathcal{N} \geq 1), \mathcal{C}1(\mathcal{N} \geq 2), \dots) \qquad \text{and} \qquad (\mathcal{Q}, \mathcal{N}),$$

respectively, where $\mathcal{C}$ and $(\mathcal{N}, \mathcal{Q})$ are distributed as in Theorem 4.7, with $\mathcal{C}$ independent of $(\mathcal{Q}, \mathcal{N})$.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

Next, let d_1 denote the Kantorovich–Rubinstein distance on $\mathcal{S}$, with $\mathcal{S}$ either $\mathbb{R}^\infty$ or $\mathbb{R}^2$ as needed, defined conditionally on $\mathscr{F}_n$. More precisely, if we let $\|\mathbf{x}\|_1 = \sum_i |x_i|$ for $\mathbf{x} \in \mathcal{S}$, then, for any two probability measures ϕ and χ on $\mathbb{R}^\infty$,

$$d_1(\phi, \chi) = \inf_{\mathbf{U},\mathbf{V}} \mathbb{E}_n \left[\|\mathbf{U} - \mathbf{V}\|_1\right],$$

where $\mathbf{U}$ is distributed according to ϕ and $\mathbf{V}$ is distributed according to χ, and the infimum is taken over all couplings of ϕ and χ.

Let $\mathcal{R}^{(k)} = \sum_{r=0}^{k} \sum_{\mathbf{i}\in\mathcal{A}_r} \Pi_{\mathbf{i}} \mathcal{Q}_{\mathbf{i}}$ denote the rank of the root node in the delayed weighted branching process constructed using the i.i.d. vectors $\{(\mathcal{N}_{\mathbf{i}}, \mathcal{Q}_{\mathbf{i}}, \{\mathcal{C}_{(\mathbf{i},j)}\}_{j\geq 1}) : \mathbf{i} \in \mathcal{U}\}$ (see [10] for more details). By Theorem 2 (Case 2) in [10], the convergence of $\hat{R}_{\emptyset}^{(n,k)}$ to $\mathcal{R}^{(k)}$ in the Kantorovich–Rubinstein distance will follow once we show that

$$d_1(\phi_n^*, \phi^*) + d_1(\phi_n, \phi) \xrightarrow{P} 0 \quad \text{as } n \to \infty. \tag{16}$$

That convergence in d_1 is equivalent to weak convergence plus convergence of the first absolute moments follows from Theorem 6.9 in [33].

Now let $G_n^*(m, q) = \mathbb{P}_n(\hat{N}_\emptyset \leq m, \hat{\mathcal{Q}}_\emptyset \leq q)$, $G_n(m, q, x) = \mathbb{P}_n(\hat{N} \leq m, \hat{\mathcal{Q}} \leq q, \hat{\mathcal{C}} \leq x)$, $G^*(m, q) = P(\mathcal{N} \leq m, \mathcal{Q} \leq q)$, and $G(m, q, x) = P(\mathcal{N} \leq m, \mathcal{Q} \leq q)P(\mathcal{C} \leq x)$. Note that by Theorem 4.7 we have

$$d_1(G_n, G) + d_1(G_n^*, G^*) \xrightarrow{P} 0 \qquad n \to \infty.$$

Moreover, $d_1(G_n^*, G^*) = d_1(\phi_n^*, \phi^*)$. To see that $d_1(G_n, G) \xrightarrow{P} 0$ implies that $d_1(\phi_n, \phi) \xrightarrow{P} 0$, choose $(\hat{N}, \hat{\mathcal{Q}}, \hat{\mathcal{C}}, \mathcal{N}, \mathcal{Q}, \mathcal{C})$ and $(\hat{N}_\emptyset, \hat{\mathcal{Q}}_\emptyset, \hat{\mathcal{C}}_\emptyset, \mathcal{N}, \mathcal{Q}, \mathcal{C})$ such that $\mathbb{E}_n\left[\|(\hat{N}, \hat{\mathcal{Q}}, \hat{\mathcal{C}}) - (\mathcal{N}, \mathcal{Q}, \mathcal{C})\|_1\right] = d_1(G_n, G)$, which can be done since optimal couplings always exist (see Theorem 4.1 in [33]). Next, note that since $|\hat{\mathcal{C}}| \leq c$ and $|\mathcal{C}| \leq c$ with $c < 1$,

$$\begin{aligned}
d_1(\phi_n, \phi) &\leq \mathbb{E}_n\Big[\| \left(\hat{\mathcal{C}}\hat{\mathcal{Q}}, \hat{\mathcal{C}}1(\hat{N} \geq 1), \hat{\mathcal{C}}1(\hat{N} \geq 2), \dots\right) \\
&\qquad -\left(\mathcal{C}\mathcal{Q}, \mathcal{C}1(\mathcal{N} \geq 1), \mathcal{C}1(\mathcal{N} \geq 2), \dots\right) \|_1 \Big] \\
&= \mathbb{E}_n\left[|\hat{\mathcal{C}}\hat{\mathcal{Q}} - \mathcal{C}\mathcal{Q}| + \sum_{i=1}^{\infty} |\hat{\mathcal{C}}1(\hat{N} \geq i) - \mathcal{C}1(\mathcal{N} \geq i)|\right] \\
&\leq \mathbb{E}_n\Bigg[|\hat{\mathcal{C}}||\hat{\mathcal{Q}} - \mathcal{Q}| + |\mathcal{Q}||\hat{\mathcal{C}} - \mathcal{C}| + \sum_{i=1}^{\infty} |\hat{\mathcal{C}}||1(\hat{N} \geq i) - 1(\mathcal{N} \geq i)| \\
&\qquad + |\hat{\mathcal{C}} - \mathcal{C}|1(\mathcal{N} \geq i)\Bigg] \\
&\leq c\mathbb{E}_n\left[|\hat{\mathcal{Q}} - \mathcal{Q}|\right] + \mathbb{E}_n[|\mathcal{Q}||\hat{\mathcal{C}} - \mathcal{C}|] + c\sum_{i=1}^{\infty} \mathbb{E}_n\left[|1(\hat{N} \geq i) - 1(\mathcal{N} \geq i)|\right] \\
&\qquad + \mathbb{E}_n\left[|\hat{\mathcal{C}} - \mathcal{C}| \sum_{i=1}^{\infty} 1(\mathcal{N} \geq i)\right] \\
&= c\mathbb{E}_n\left[|\hat{\mathcal{Q}} - \mathcal{Q}|\right] + c\sum_{i=1}^{\infty} \mathbb{E}_n\left[|1(\hat{N} < i \leq \mathcal{N}) - 1(\mathcal{N} < i \leq \hat{N})|\right] \\
&\qquad + \mathbb{E}_n\left[|\hat{\mathcal{C}} - \mathcal{C}|(\mathcal{N} + |\mathcal{Q}|)\right]
\end{aligned}$$

$$\leq c\mathbb{E}_n\left[|\hat{\mathcal{Q}}-\mathcal{Q}|\right]+c\mathbb{E}_n\left[|\hat{\mathcal{N}}-\mathcal{N}|\right]+\mathbb{E}_n\left[|\hat{\mathcal{C}}-\mathcal{C}|(\mathcal{N}+|\mathcal{Q}|)\right]$$
$$\leq cd_1(G_n,G)+\mathbb{E}_n\left[|\hat{\mathcal{C}}-\mathcal{C}|(\mathcal{N}+|\mathcal{Q}|)\right].$$

Since $\mathbb{E}_n[|\hat{\mathcal{C}}-\mathcal{C}|] \leq d_1(G_n,G) \xrightarrow{P} 0$, then dominated convergence gives that $\mathbb{E}_n\left[|\hat{\mathcal{C}}-\mathcal{C}|(\mathcal{N}+|\mathcal{Q}|)\right] \xrightarrow{P} 0$ as well.

Finally, it is well known that provided $E\left[\sum_{i=1}^{\mathcal{N}}|\mathcal{C}_i|\right] = E[\mathcal{N}]E[\mathcal{C}_1] < 1$, we have $\mathcal{R}^{(k)} \to \mathcal{R} = \sum_{r=0}^{\infty}\sum_{\mathbf{i}\in\mathcal{A}_r} \Pi_{\mathbf{i}}\mathcal{Q}_{\mathbf{i}}$ a.s. as $k\to\infty$ (see, e.g., Lemma 4.1 in [25]). To see that the required condition is satisfied note that

$$\begin{aligned}
E[\mathcal{N}]E[|\mathcal{C}_1|] &= E\left[\frac{W^+}{E[W^+]}\cdot\frac{E[W^+]W^-}{\theta}\right]E\left[\frac{W^+}{E[W^+]}\int_{-\infty}^{\infty}|t|1(\zeta/(Z^++1)\in dt)\right]\\
&= \frac{E[W^+W^-]}{\theta}\cdot E\left[\frac{W^+}{E[W^+]}\frac{|\zeta|}{Z^++1}\right]\\
&= \frac{E[W^-]}{\theta}\cdot E\left[\frac{W^+|\zeta|}{E[W^-]W^+/\theta}\left(1-e^{-E[W^-]W^+/\theta}\right)\right]\\
&= E\left[|\zeta|\left(1-e^{-E[W^-]W^+/\theta}\right)\right] \leq c < 1,
\end{aligned}$$

where we used the observation that if Z is Poisson with mean λ then $E[1/(Z+1)] = \lambda^{-1}(1-e^{-\lambda})$. This completes the proof. □

4.5. Proof of Theorem 3.3

Finally, we combine the inequality in (8), Theorem 3.7, and Theorem 3.8 to prove Theorem 3.3.

Proof. Fix $\epsilon > 0$ and $k \in \mathbb{N}_+$. Next, construct $(R_\xi^{(n,\infty)}, R_\xi^{(n,k)}, \hat{R}_\emptyset^{(n,k)})$ using the graph exploration and coupling described in Section 3.2.2. Then, by (8) and Theorem 3.7, the following limit holds in probability:

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{P}_n\left(|R_\xi^{(n,\infty)}-\hat{R}_\emptyset^{(n,k)}|>\epsilon\right)\\
&\leq \lim_{n\to\infty}\mathbb{P}_n\left(|R_\xi^{(n,\infty)}-R_\xi^{(n,k)}|>\epsilon/2\right)+\lim_{n\to\infty}\mathbb{P}_n\left(|\hat{R}_\xi^{(n,k)}-\hat{R}_\emptyset^{(n,k)}|>\epsilon/2\right)\\
&= \frac{2c^k}{\epsilon(1-c)}E[|Q|].
\end{aligned}$$

Also, by Theorem 3.8 there exists $\mathcal{R}^{(k)}$ such that

$$\lim_{n\to\infty}\mathbb{P}_n\left(|\hat{R}_\emptyset^{(n,k)}-\mathcal{R}^{(k)}|>\epsilon\right)=0$$

in probability, with $\mathcal{R}^{(k)}\to\mathcal{R}$ a.s. as $k\to\infty$. Choosing k sufficiently large yields

$$R_\xi^{(n,\infty)} \Rightarrow \mathcal{R}$$

as $n\to\infty$. □

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] T.L. Austin, R.E. Fagen, W.F. Penney, J. Riordan, The number of components in random linear graphs, Ann. Math. Stat. 30 (1959) 747–754.
[2] K. Avrachenkov, D. Lebedev, PageRank of scale-free growing networks, Internet Math. 3 (2006) 207–231.
[3] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, European J. Combin. (1980) 311–316.
[4] B. Bollobás, Random graphs, Cambridge University Press, 2001.
[5] B. Bollobás, S. Janson, O. Riordan, The phase transition in inhomogeneous random graphs, Random Struct. Algorithms 31 (2007) 3–122.
[6] S. Brin, L. Page, The anatomy of a large-scale hypertextual web search engine, Comput. Netw. 30 (1998) 107–117.
[7] T. Britton, M. Deijfen, A. Martin-Läf, Generating simple random graphs with prescribed degree distribution, J. Stat. Phys. 124 (2006) 1377–1397.
[8] N. Chen, N. Livtak, M. Olvera-Cravioto, Generalized PageRank on directed configuration networks, Random Struct. Algorithms 51 (2) (2017) 237–274.
[9] N. Chen, M. Olvera-Cravioto, Directed random graphs with given degree distributions, Stoch. Syst. 3 (2013) 147–186.
[10] N. Chen, M. Olvera-Cravioto, Coupling on weighted branching trees, Adv. Appl. Probab. 48 (2) (2016) 499–524.
[11] F. Chung, L. Lu, Connected components in random graphs with given expected degree sequences, Ann. Comb. 6 (2002) 125–145.
[12] F. Chung, L. Lu, The average distances in random graphs with given expected degrees, in: Proceedings of National Academy of Sciences, Vol. 99, 2002a, pp. 15879–15882.
[13] F. Chung, L. Lu, The volume of the giant component of a random graph with given expected degrees, SIAM J. Discrete Math. 20 (2006a) 395–411.
[14] F. Chung, L. Lu, Complex Graphs and Networks, in: CBMS Regional Conference Series in Mathematics, vol. 107, 2006b.
[15] R. Durrett, Random Graph Dynamics, in: Cambridge Series in Statistics and Probabilistic Mathematics, Cambridge University Press, 2007.
[16] P. Erdös, A. Rényi, On random graphs, Publ. Math. (Debrecen) 6 (1959) 290–297.
[17] H. van den Esker, R. van der Hofstad, G. Hooghiemstra, Universality for the distance in finite variance random graphs, J. Stat. Phys. 133 (2008) 169–202.
[18] E.N. Gilbert, Random graphs, Ann. Math. Stat. 30 (1959) 1141–1144.
[19] J. Grandell, Mixed Poisson Processes, CRC Press, 1997.
[20] R. van der Hofstad, Random Graphs and Complex Networks, Vol. 1, Cambridge U. Press, 2017, URL http://www.win.tue.nl/rhofstad/NotesRGCN.pdf.
[21] F. den Hollander, Probability theory: The coupling method, 2012. URL http://websites.math.leidenuniv.nl/probability/lecturenotes/CouplingLectures.pdf.
[22] S. Janson, T. Luczak, A. Rucinski, Random graphs, Wiley-Interscience, 2000.
[23] P.R. Jelenković, M. Olvera-Cravioto, Information ranking and power laws on trees, Adv. Appl. Probab. 42 (2010) 1057–1093.
[24] P.R. Jelenković, M. Olvera-Cravioto, Implicit renewal theorem and power tails on trees, Adv. Appl. Probab. 44 (2012) 528–561.
[25] P.R. Jelenković, M. Olvera-Cravioto, Implicit renewal theorem for trees with general weights, Stochastic Process. Appl. 122 (2012) 3209–3238.
[26] J. Lee, M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, ArXiv:1707.02492 (2019) 43 p.
[27] N. Litvak, W.R.W. Scheinhardt, Y. Volkovich, In-degree and PageRank: Why do they follow similar power laws? Internet Math. 4 (2007) 175–198.
[28] L. Lu, Probabilistic Methods in Massive Graphs and Internet Computing (Ph.D. Thesis), University of California, San Diego, 2002.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.

[29] I. Norros, H. Reittu, On a conditionally Poissonian graph process, Adv. Appl. Probab. 38 (2006) 59–75.
[30] G. Pandurangan, P. Raghavan, E. Upfal, Using pagerank to characterize web structure, in: The Eighth Annual International Computing and Combinatorics Conference (COCOON), 2002, pp. 330–339.
[31] U. Rösler, The weighted branching process, in: Dynamics of Complex and Irregular Systems, Bielefeld Encounters in Mathematics and Physics VIII, World Science Publishing, River Edge, NJ, 1993, pp. 154–165.
[32] G. Samorodnitsky, S. Resnick, D. Towsley, R. Davis, A. Willis, P. Wan, Nonstandard regular variation of in-degree and out-degree in the preferential attachment model, J. Appl. Probab. 53 (1) (2016) 146–161.
[33] C. Villani, Optimal Transport, Old and New, Springer, New York, 2009.
[34] Y. Volkovich, N. Litvak, Asymptotic analysis for personalized web search, Adv. Appl. Probab. 42 (2010) 577–604.
[35] Y. Volkovich, N. Litvak, B. Zwart, Determining factors behind the PageRank log-log plot, in: Proceedings of the 5th International Conference on Algorithms and Models for the Web-Graph, San Diego, CA, 2007, pp. 108–123.

Please cite this article as: J. Lee and M. Olvera-Cravioto, PageRank on inhomogeneous random digraphs, Stochastic Processes and their Applications (2019), https://doi.org/10.1016/j.spa.2019.07.002.