*Adv. Appl. Prob.* **42,** 577–604 (2010)
*Printed in Northern Ireland*


# ASYMPTOTIC ANALYSIS FOR PERSONALIZED WEB SEARCH

YANA VOLKOVICH * AND
NELLY LITVAK,** *University of Twente*

**Abstract**

PageRank with personalization is used in Web search as an importance measure for Web documents. The goal of this paper is to characterize the tail behavior of the PageRank distribution in the Web and other complex networks characterized by power laws. To this end, we model the PageRank as a solution of a stochastic equation $R \stackrel{\mathrm{D}}{=} \sum_{i=1}^{N} A_i R_i + B$, where the $R_i$s are distributed as $R$. This equation is inspired by the original definition of the PageRank. In particular, $N$ models the number of incoming links to a page, and $B$ stays for the user preference. Assuming that $N$ or $B$ are heavy tailed, we employ the theory of regular variation to obtain the asymptotic behavior of $R$ under quite general assumptions on the involved random variables. Our theoretical predictions show good agreement with experimental data.

*Keywords:* PageRank; Web; regular variation; stochastic equation; Tauberian theorem

2010 Mathematics Subject Classification: Primary 68P10; 90B15
Secondary 40E05

## 1. Introduction

Today the World Wide Web is an important part of our lives. Hence, understanding properties of the Web is one of the most essential research needs. The Web has a complex structure with some notable features. Cardinally, it is huge. By some estimations, the indexed Web contains at least 27.5 billion pages (source: www.worldwidewebsize.com/ (accessed in July 2008)), and it continues to grow very fast. The Web has a linking or, more precisely, a hyperlinking structure. A convenient way to analyze the Web structure is to consider the Web as a graph, where pages are nodes, and links are edges. Then we can assign different characteristics for each node in such a graph. The terms *in-degree* and *out-degree* are used for the number of incoming and outgoing links of a page, respectively. Furthermore, *PageRank* is a widely accepted notion for characterizing the importance of each node in the graph. It is worth noting that in- and out-degrees are natural characteristics of the graph structure, while PageRank is a popularity measure designed to enhance Web search. The PageRank as originally introduced by Google is one of significant characteristics that affects the listing of Web pages returned by a search engine in response to a query. We provide a formal definition of the PageRank in Section 1.1.

Most experimental studies of the Web agree that the in-degree, the out-degree, and the PageRank of the Web follow power laws. In simple words, a random variable $X$ has a power law distribution with exponent $\alpha > 0$ if its probability of obtaining a value grater than $x$ is

Received 14 October 2008; revision received 20 October 2009.
* Current address: Barcelona Media – Innovation Center, Av. Diagonal 177, 9th Floor, 08018 Barcelona, Spain.
Email address: yana.volkovich@barcelonamedia.org
** Postal address: Faculty of EEMCS, University of Twente, 7500 AE Enschede, The Netherlands.
Email address: n.litvak@ewi.utwente.nl

proportional to $x^{-\alpha}$. In the Web, the power law exponents can deviate depending on a data set and an estimator, but are believed to satisfy $\alpha = 1.1$ for the in-degree and PageRank, and $\alpha \approx 2$ for the out-degree [9], [32], [38].

The goal of this paper is to provide mathematical evidence for the power law behavior of the PageRank and its relation to the different characteristics of the underlying graph. To this end, we propose a stochastic model that is a considerable extension of our previous work [26], [37]. The PageRank is modeled as a solution of a distributional identity, and the tail behavior of such a solution is obtained under various assumptions on the involved parameters. The generality of our analytical model allows us to take into account many different factors affecting the PageRank, such as personalization of the PageRank, as defined in the next section, and a possible dependence between personalized preference scores and in-degrees of the Web pages. The analyzed stochastic equation, as described in Section 1.3, is of independent mathematical interest.

### 1.1. Personalized PageRank

With the evolution of the Web, the first search engines quickly became insufficient because the underlying techniques were developed for document collections, in which all documents were assumed to have a high quality and be homogeneous. This assumption holds, for example, for collections of papers or books where the number of citations is a good measure of their popularity. However, the homogeneity assumption is definitely violated in a representative collection of Web pages, where the best text match does not imply the highest relevance, and the large number of incoming links can often indicate a spam. To resolve the problem, Brin and Page, using the PageRank algorithm [8], [31], and Kleinberg, using the HITS algorithm [22], proposed employing link analysis to measure the importance of pages in a Web search. The idea turns out to be very successful, and both of the algorithms are widely used today not only in search engines, but in various ranking related problems. Hence, PageRank is successfully used for spam detection [15], graph partitioning [1], and finding gems in scientific citations [10], to name just a few. In this work we focus only on PageRank.

The PageRank is defined as the stationary distribution of an 'easily bored surfer' random walk on the graph. At each step, with probability $c$, such a random walk follows a randomly chosen outgoing link of a page, and, with probability $(1-c)$, the walk starts afresh from a page chosen at random according to some *teleportation* distribution. In other words, at each step the surfer makes a *teleportation jump* to a random page with probability $(1-c)$. The constant $c$ is called a *damping factor*, and takes values between 0 and 1. Traditionally, the value of $c$ is chosen as 0.85, and it appears that this value provides a reasonable ranking for Web pages. In [4], [7], and [11] the authors study other values of the damping factor.

If a page is a *dangling* node, i.e. it has no outgoing link, then we follow the approach of [31], and assume that this page has links to all pages in the network. Let the total number of nodes in the Web graph be denoted by $w$. Then the probability of following a particular link from such a page becomes $1/w$, and it is almost 0 for large $w$.

We can summarize the PageRank definition as

$$\mathrm{PR}(i) = c \sum_{j \to i} \frac{1}{d_j} \mathrm{PR}(j) + \frac{c}{w} \sum_{j \in \mathcal{D}} \mathrm{PR}(j) + (1-c)T(i), \qquad i = 1, \ldots, w, \tag{1.1}$$

where $\mathrm{PR}(i)$ is the PageRank of page $i$, $d_j$ is the out-degree of page $j$, the sum is taken over all pages $j$ that link to page $i$, $\mathcal{D}$ is a set of dangling nodes, and $T(i)$ is the probability that the walk starts afresh from page $i$.

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

The theoretical foundation [23], [33] for including a probability distribution, as suggested by PageRank, into the overall scoring of a page for a given query can be briefly explained as follows. We are interested in the probability $\mathrm{P}(d \mid q)$ that a document $d$ is relevant for a given query $q$. Using Bayes' rule, we can rewrite this probability as $\mathrm{P}(d \mid q) = \mathrm{P}(d)\,\mathrm{P}(q \mid d)/\mathrm{P}(q)$. For page-ranking purposes, $\mathrm{P}(q)$ is irrelevant, since it does not depend on the document. The term $\mathrm{P}(q \mid d)$ is one of the main interests to the information retrieval community. Various heuristics are used to estimate the relevance of a query to a document. The $\mathrm{P}(d)$ term has a natural interpretation in the PageRank model (and similar models) as the likelihood that a document would be relevant independent of the query. One of the aims of this work is to provide a better understanding of what the $\mathrm{P}(d)$ term might look like, and to examine how it is distributed under the PageRank model.

In the definition of the *standard PageRank* [31], the teleportation distribution is assumed to be uniform, i.e. $T(i) = 1/w$ for every $i = 1, \ldots, w$. In the original paper, Page *et al.* [31] also suggested modifying the PageRank so that the teleportation jumps favor trusted nodes and are the same for all users, or to favor specific nodes for each user with respect to the individual user's tastes. The knowledge of user preferences can be based on the usage data, such as browsing histories or search engine logs; or/and on the user data, such as information about personal characteristics of the user, e.g. name, age, or geographic location [30]. However, individually personalized PageRank is computationally infeasible in practice. Therefore, the idea is to build an approximation of such an individual PageRank, which still allows us to achieve a good level of personalization. Below we list several approaches for this approximation [17]. The topic-sensitive PageRank [16] restricts the interests of a user to a small number of topics, say $K = 20$. Then the teleportation jump can be defined as follows: $T(i) = \sum_{i \in J} p_J p_{i,J}$, where $p_J$ is the teleportation probability of the topic $J$, $J = 1, \ldots, K$, and $p_{i,J}$ is the probability of teleporting into particular page $i$ within topic $J$. Intuitively, if some individuals like to surf for pages about sport then their search result can be improved by enlarging the $T(i)$s in (1.3), below, for the pages with sport content. Then, the topic-sensitive PageRank represents user preferences for the beneficial topics choice. Modular PageRank, which was proposed by Jeh and Widom in [18], is similar to the above approach. However, in this case the surfer teleports to those pages with high rank instead of set of topic-related pages.

In BlockRank [21] the Web is considered to be an aggregation of blocks, where, for example, each block represents a host. Then the teleportation jump can be defined as follows: $T(i) = p_J \mathrm{PR}_J(i)$, where $p_J$ is the probability of jumping into block $J$ and $\mathrm{PR}_J(i)$ is the local PageRank of page $i$ in block $J$. We now also mention two approaches that personalize PageRank, but do not do so through teleportation. The first approach, query-dependent PageRank [35], follows by replacing $1/d_j$ in (1.1) with $p_q(j \to i)$, the probability that the random walk follows the link to page $i$ given that it is on page $j$ and is searching for query $q$. For the second approach, Constantine and Gleich [11] suggested modifying the damping factor $c$ according to the user surfing properties.

In the recent literature the term personalized PageRank of page $i$ often means a PageRank distribution computed under the assumption that the random walk always restarts from $i$, i.e. $T(i) = 1$. Such a definition is important, for instance, in graph clustering [2] and in sampling from large graph data [25]. In this paper we follow an original approach [24] that defines personalization as taking into account user's preferences for better search results. Therefore, we assume that the support of the distribution $T(\cdot)$ is not limited by a small number of pages. In other words, we assume that, for every $i$, $T(i)$ scales as $1/w$ as $w \to \infty$.

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

It is clear that the PageRank values in (1.1) scale as $1/w$ with the number of pages. In our analysis, it is more convenient to deal with the corresponding *scale-free PageRank* scores,

$$R(i) = w\mathrm{PR}(i), \qquad i = 1, \dots, w, \tag{1.2}$$

assuming that $w$ goes to $\infty$. In this setting, it is easier to compare the probabilistic properties of PageRank and in- and out-degrees, which are also scale free. In the remainder of the paper, by PageRank we mean the scale-free PageRank scores (1.2). Then the original definition (1.1) can be written as

$$R(i) = c \sum_{j \to i} \frac{1}{d_j} R(j) + \frac{c}{w} \sum_{j \in \mathcal{D}} R(j) + (1-c) w T(i), \qquad i = 1, \dots, w. \tag{1.3}$$

With any of the abovementioned approaches to personalized ranking, the resulting distribution of the PageRank scores given by (1.3) for a given Web graph depends on local graph characteristics, such as the in-degree and out-degree. In the next section we discuss the tail behavior of the PageRank distribution, and its relations to different parameters in the Web.

### 1.2. Power law distributions in the Web

It has become common knowledge that the in-degree and PageRank of the Web follow a power law with the same exponent (see [14], [26], [32], and [37]). From the definition of the PageRank we can see that the PageRank should be related to the in-degree. However, the main idea behind PageRank is that it depends not only on quantity but also on the quality of incoming links to a page. Moreover, we emphasize that PageRank is a global characteristic of the Web, while the in-degree is a local characteristic. Thus, the phenomena of asymptotic similarity between the in-degree and PageRank is not trivial to justify. In [3] and [13] the authors verified asymptotic properties of the PageRank distribution for the case of preferential attachment models [5], which are often used for simulating graphs with power law distributed in-degree. In this work, as in [26] and [37], we explain the asymptotic behavior of the PageRank distribution by modeling a personalized PageRank as the solution of a stochastic equation.

To obtain the asymptotic behavior of PageRank, we employ the theory of regular variation, which provides natural mathematical formalism for analyzing power laws. A nonnegative random variable $X$ is said to be *regularly varying* with index $\alpha$ if $\mathrm{P}(X > x) \sim x^{-\alpha} L(x)$ as $x \to \infty$ for some positive, *slowly varying* function $L(x)$ (that is, by definition, for every $y > 0$, we have $L(yx)/L(x) \to 1$ as $x \to \infty$). Here, as in the remainder of this paper, the notation $a(x) \sim b(x)$ means that $a(x)/b(x) \to 1$. We provide all necessary preliminaries on the theory of regular variation in Appendix A.

### 1.3. Stochastic equations

From a mathematical point of view, this paper presents the analysis of the following distributional identity:

$$R \stackrel{\mathrm{D}}{=} \sum_{j=1}^{N} A_j R_j + B, \tag{1.4}$$

where '$\stackrel{\mathrm{D}}{=}$' denotes equality in distribution and we assume that all random variables are positive; the $R_j$s are independent and distributed as $R$, and the $A_j$s are independent and distributed as some random variable $A$ with $\mathrm{E}(A) = [1 - \mathrm{E}(B)]/\mathrm{E}(N) < 1$. We also set the $R_j$s and $A_j$s to be independent, and independent of $N$ and $B$. Moreover, it is essential that $\mathrm{E}(B) < 1$. We emphasize that $N$ and $B$ can be dependent.

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

Equations similar to (1.4) are well known in the literature. For instance, such an equation can also describe the distribution of the busy period in the M/G/1 queue:

$$R \stackrel{\mathrm{D}}{=} \sum_{j=1}^{N(S_1)} R_j + S_1,$$

where $R$ is the duration of the busy period (the time interval during which the queue is nonempty), $S_1$ is the service time of the customer that initiated the busy period, $N(S_1)$ is the number of Poisson arrivals during this service time, and the $R_j$s are independent and distributed as $R$. We refer the reader to [12] and [40] for more details on the asymptotics of a busy period in queues with heavy tails.

Another version of (1.4) arises in the theory of branching processes. For $B = 0$, we can obtain the following equation:

$$R \stackrel{\mathrm{D}}{=} \sum_{j=1}^{N} A_j R_j,$$

which has been analyzed in detail by Liu in [28] and [29].

Our model, as presented in (1.4), was also further studied in [19] in the interesting context of weighted branching processes. The authors suggested an alternative probabilistic proof to some of our results. Furthermore, they also showed that the solution $R$ of the stochastic equation (1.4) can be heavy tailed, while $N$ and $B$ are not.

The rest of the paper is organized as follows. In Section 2 we describe the model for the in- and out-degrees, and provide the stochastic equation for the PageRank in the form (1.4), where each random variable represents a certain parameter of the Web. In Section 3 we use a probabilistic approach to show that the proposed equation has a unique nontrivial solution with a finite mean. We introduce a recurrent stochastic model for the power iteration algorithm commonly used in PageRank computations [24], and we obtain the PageRank asymptotics after each iteration in Section 3.3. The tail behavior of the PageRank in our model is obtained in Section 4.3. To this end, we use Laplace–Stieltjes transforms and apply the Tauberian theorem; see Theorem 8.1.6 of [6] or Theorem A.1 in Appendix A.

Our analysis reveals that the in-degree distribution is not the only determining factor for the asymptotic behavior of the personalized PageRank. It turns out that the teleportation distribution can play a significant role as well. In fact, the asymptotic properties of PageRank as a solution of (1.4) are defined by the distribution with the heaviest tail. We are also able to explicitly derive the constant multiplicative factor that quantifies the difference between the tail asymptotics of PageRank, in-degree, and teleportation distributions. In Section 5 we show that analytical results are in agreement with Web data.

## 2. Model

We develop the idea suggested in [26] and [37] for the personalized PageRank. We start with models for in- and out-degree distributions in the Web. Then we define the PageRank of a random page in the network as the solution of a stochastic equation in Section 2.2.

### 2.1. In- and out-degrees

We set the in-degree of a randomly chosen page in the network to be an integer-valued random variable $N$. In the Web graph, as well in some other graphs, where we observe power law behavior of the in-degree distribution, we set $N$ to be an integer-valued, regularly varying

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

random variable with index $\alpha_N > 1$. One of the ways to model $N$ is as follows: we assume that $N = N(X)$, where $X$ is regularly varying with index $\alpha_N$ and $N(x)$ is the number of Poisson arrivals during the time interval $[0, x]$, when the arrival rate is 1. Thus, if $X$ is regularly varying then $N(X)$ is also regularly varying and asymptotically identical to $X$ (see, e.g. [26]):

$$\mathrm{P}(X > x) \sim x^{-\alpha_N} L_N(x) \quad \Longleftrightarrow \quad \mathrm{P}(N(X) > x) \sim x^{-\alpha_N} L_N(x) \quad \text{as } x \to \infty. \tag{2.1}$$

Then $N(X)$ is indeed an integer and obeys the power law. We use this representation for $N$ in Section 4.

Next, we model the weights $1/d_j$ in (1.3). Recall that $d_j$ is the out-degree of page $j$ that has a link to page $i$. As in [37], we consider a random variable $D$ that represents the out-degree of a page that links to a particular randomly chosen page $i$. Note that $D$ is *not* the same random variable as an out-degree of a random page since the additional information that a page has a link to $i$ alters the out-degree distribution. This phenomenon is known as the inspection paradox [36]. Thus, the number of out-links from a page containing a random link is stochastically larger than an out-degree of a random page. If $p_j$ is a fraction of the pages with out-degree $j \geq 0$ then we can obtain

$$\lim_{w\to\infty} \mathrm{P}(D = j) = \frac{jp_j}{\mathrm{E}(N)}, \qquad j \geq 1, \tag{2.2}$$

where $\mathrm{E}(N)$ is the average in/out-degree, and $w$ is the number of pages in the Web. For sufficiently large networks, we may assume that the distribution of $D$ is equal to its limiting distribution as defined by (2.2). We refer to $D$ as an *effective out-degree*. The term is motivated by the fact that the distribution of $D$ is the one that participates in the PageRank formula (1.3).

### 2.2. Stochastic equation for the PageRank

Now we are ready to model the PageRank distribution. We view the PageRank of a random page as a random variable $R$ with $\mathrm{E}(R) = 1$. Furthermore, we assume that the PageRank of a random page does not depend on whether the page is dangling. We note that such independence immediately implies that, in large networks, the fraction of the total PageRank mass concentrated in dangling nodes is equal to the fraction of dangling nodes $p_0$, simply by the law of large numbers: $p_0 = (1/w) \sum_{j\in\mathcal{D}} R(j)$.

Our goal is to analyze to what extent the tail probability $\mathrm{P}(R > x)$ for large enough $x$ depends on the in-degree $N$, the effective out-degree $D$, the teleportation jump $T$, and the fraction of dangling nodes $p_0$. To this end, we model the PageRank $R$ as a solution of a stochastic equation involving $N$, $T$, and $D$. Inspired by the original formula (1.3), the stochastic equation for the PageRank is as follows:

$$R \overset{\mathrm{D}}{=} c \sum_{j=1}^{N} \frac{1}{D_j} R_j + cp_0 + (1 - c)wT. \tag{2.3}$$

Here the $R_j$s and $D_j$s are independent and distributed as $R$ and $D$, respectively. Moreover, we need to assume that the $R_j$s and $D_j$s are independent and independent of $N$ and $T$. As before, $c \in (0, 1)$ is a damping factor. We emphasize that $N$ and $T$ are allowed to be dependent, which is often the case for the personalized PageRank.

We note that the assumption that $N$ and the $R_j$s are independent is obviously not true in general. However, it is also not the case that the PageRank values of the pages linking to the same page $i$ are directly related to each other or the in-degree of $i$, so we may assume independence

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

in this study. Empirical and analytical characterizations of the dependencies between power law network parameters can be obtained using the extreme value theory [34, Chapters 6 and 9]. In [39] we applied these methods to experimentally compute the dependencies between the in-degree and PageRank of a page. For analytical results on the tail dependence between $R$ and $N$ in (2.3) and further discussion on the dependence structure in complex networks, we refer the reader to [27].

In stochastic equation (2.3) we generalize models from [26] and [37] for the cases of the random out-degree and the random teleportation jump. Moreover, here we allow this personalization jump to be dependent on the in-degree. In Section 3.2 we will show that (2.3) has a unique solution $R$ such that $\mathrm{E}(R) = 1$.

## 3. Probabilistic analysis

In the next two sections we will analyze the following stochastic equation:

$$R \stackrel{\mathrm{D}}{=} \sum_{j=1}^{N} A_j R_j + B, \tag{3.1}$$

where we assume that all random variables are positive; the $R_j$s are independent and distributed as $R$, and the $A_j$s are independent and distributed as some random variable with $\mathrm{E}(A) = [1 - \mathrm{E}(B)]/\mathrm{E}(N)$. We also set the $R_j$s and $A_j$s to be independent, and to be independent of $N$ and $B$. Moreover, it is essential that $\mathrm{E}(B) < 1$. We emphasize that $N$ and $B$ can be dependent. It is easy to see that the above equation corresponds to (2.3) for $A \stackrel{\mathrm{D}}{=} c/D$ and $B \stackrel{\mathrm{D}}{=} cp_0 + (1 - c)nT$.

In Sections 3.2 and 3.3 we establish the existence and the asymptotic properties of $R$ in (3.1) using an iterative procedure defined in the next section.

### 3.1. Iterations

We use the following notation adopted from [29]. Let $\{(N_u, A_{u_1}, A_{u_2}, \ldots)\}_u$ be a family of independent copies of $(N, A_1, A_2, \ldots)$ indexed by all finite sequences $u = u_1 \cdots u_i$, where $u_j \in \{1, 2, \ldots\}$, $j = 1, \ldots, i$. Let $\mathbb{T}$ be *the Galton–Watson tree* with defining elements $\{N_u\}$: we have $\varnothing \in \mathbb{T}$ and, if $u \in \mathbb{T}$ and $j \in \{1, 2, \ldots\}$, then concatenation $uj \in \mathbb{T}$ if and only if $1 \le j \le N_u$. In other words, we indexed the nodes of the tree with root $\varnothing$ and the first level nodes $1, 2, \ldots, N_\varnothing$, and at every subsequent level, the $j$th offspring of $u$ is termed $uj$ (see Figure 1).

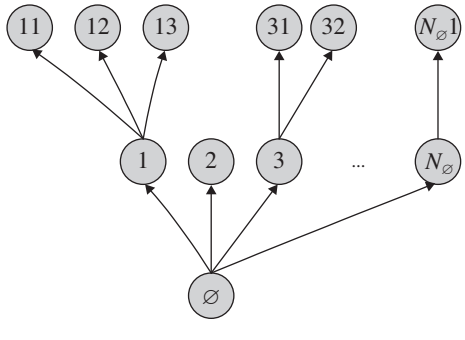


FIGURE 1: An example of a Galton–Watson tree.

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

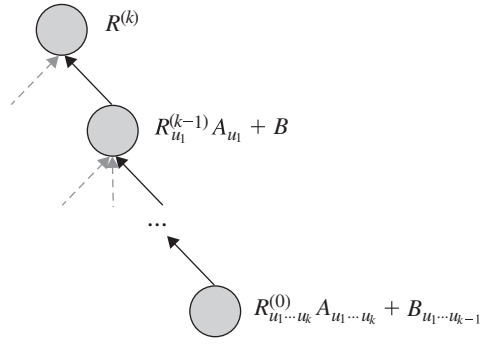


FIGURE 2: The $k$th iteration.

We start with an initial distribution $R^{(0)}$, and, for every $k \geq 1$, we define the result of the $k$th iteration of (3.1) through the distributional identity

$$R^{(k)} = \sum_{j=1}^{N} A_j R_j^{(k-1)} + B, \tag{3.2}$$

where $R_j^{(k-1)}$ and $A_j$, $j \geq 1$, are independent and distributed as $R^{(k-1)}$ and $A$, respectively.

Repeatedly applying (3.2), we obtain the following representation for $R^{(k)}$, $k \geq 1$:

$$R^{(k)} = \sum_{u_1 \cdots u_k \in \mathbb{T}} A_{u_1} \cdots A_{u_1 \cdots u_k} R^{(0)}_{u_1 \cdots u_k} + \sum_{i=0}^{k-1} \sum_{u_1 \cdots u_i \in \mathbb{T}} A_{u_1} \cdots A_{u_1 \cdots u_i} B_{u_1 \cdots u_i}, \tag{3.3}$$

where $\mathbb{T}$ is the notation for the Galton–Watson tree. In Figure 2 we display the graphic interpretation of $R^{(k)}$.

### 3.2. Existence and uniqueness of the solution

We use the next lemma to prove the existence of the solution (3.1). This lemma is a result mentioned in [29].

**Lemma 3.1.** *If* $\mathrm{E}(\sum_{j=1}^{N} A_j) = 1$ *then the sequence* $\sum_{u_1 \cdots u_i \in \mathbb{T}} A_{u_1} \cdots A_{u_1 \cdots u_i}$ *is a martingale.*

In the next theorem we show that iterations $R^{(k)}$, $k \geq 1$, converge to the unique solution of (3.1).

**Theorem 3.1.** *Equation (3.1) has the unique nontrivial solution with mean* 1 *given by*

$$R^{(\infty)} = \lim_{k \to \infty} R^{(k)} = \sum_{i=0}^{\infty} \sum_{u_1 \cdots u_i \in \mathbb{T}} A_{u_1} \cdots A_{u_1 \cdots u_i} B_{u_1 \cdots u_i}. \tag{3.4}$$

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

*Proof.* It is easy to verify that $R^{(\infty)}$ in (3.4) is a well-defined solution of (3.1). In particular, because all random variables are positive, we apply Fubini's theorem to obtain

$$\begin{aligned}\mathrm{E}(R^{(\infty)}) &= \mathrm{E}\Bigg(\sum_{i=0}^{\infty}\sum_{u_1\cdots u_i\in\mathbb{T}} A_{u_1}\cdots A_{u_1\cdots u_i}B_{u_1\cdots u_i}\Bigg)\\ &= \mathrm{E}(B)\sum_{i=0}^{\infty}(1-\mathrm{E}(B))^n\,\mathrm{E}\Bigg(\sum_{u_1\cdots u_i\in\mathbb{T}}\frac{1}{1-\mathrm{E}(B)}A_{u_1}\cdots\frac{1}{1-\mathrm{E}(B)}A_{u_1\cdots u_i}\Bigg)\\ &= 1,\end{aligned}$$

where the final equation holds since $\sum_{u_1\cdots u_i\in\mathbb{T}}(A_{u_1}/(1-\mathrm{E}(B)))\cdots(A_{u_1\cdots u_i}/(1-\mathrm{E}(B)))$ is a martingale with mean 1 according to Lemma 3.1. In the second equality we can take $\mathrm{E}(B)$ outside of the summation since $B_{u_1\cdots u_i}$ comes from the $(i-1)$th step, and is independent of the number of incoming links to level $i$. We refer the reader to Figure 2 for an illustration.

To prove the uniqueness, assume that there is another solution with mean 1 and take this solution as an initial distribution $R^{(0)}$ with $\mathrm{E}(R^{(0)})=1$. Consider $R^{(k)}$. Then the first part of (3.3) has mean

$$\mathrm{E}\Bigg(\sum_{u_1\cdots u_k\in\mathbb{T}}A_{u_1}\cdots A_{u_1\cdots u_k}R^{(0)}_{u_1\cdots u_k}\Bigg) = (\mathrm{E}(N))^k\Bigg(\frac{(1-\mathrm{E}(B))}{\mathrm{E}(N)}\Bigg)^k = (1-\mathrm{E}(B))^k,$$

and, hence, this part converges in probability to 0 as $k\to\infty$, because, by the Markov inequality, the probability that this term is greater than some $\epsilon>0$ is at most $(1-\mathrm{E}(B))^k/\epsilon\to 0$ as $k\to\infty$. Moreover, the second part of (3.3) converges almost surely to $R^{(\infty)}$ as $k\to\infty$. It follows that (3.3) converges to $R^{(\infty)}$ in probability. We conclude that there is no other fixed point of (3.1) with mean 1 except $R^{(\infty)}$.

### 3.3. Asymptotics for iterations

Our main goal is to show how the asymptotics of $R$ in (3.1) depend on the distribution of $N$ and $B$. We divide this problem into three possible cases. In the first case, we assume that $N$ is a regularly varying random variable, and $B$ has some distribution with lighter tail, that is, $\mathrm{P}(B>x)=o(\mathrm{P}(N>x))$ as $x\to\infty$. Then we recall that $N$ is an integer-valued, regularly varying random variable

$$\mathrm{P}(N>x)\sim x^{-\alpha_N}L_N(x)\quad\text{as } x\to\infty,$$

where $L_N(x)$ is a slowly varying function. In the second case, we take $B$ to be regularly varying and $N$ to have a lighter tail. Then we have

$$\mathrm{P}(B>x)\sim x^{-\alpha_B}L_B(x)\quad\text{as } x\to\infty,$$

where $L_B(x)$ is a slowly varying function. In the final case, we consider both variables to be regularly varying with the same indices.

At this point, we assume that $\mathrm{E}(N)\,\mathrm{E}(A^\alpha)<1$, where $\alpha=\min(\alpha_N,\alpha_B)$.

In Theorem 3.2, below, we consider the case when the initial distribution $R^{(0)}$ has a lighter tail than $N$ or $B$. This assumption makes sense since iterations usually start with $R^{(0)}\equiv 1$. For the other types of distribution of $R^{(0)}$, we refer the reader to Proposition 3.1, below.

In short, Theorem 3.2 states that the tail behavior of $R^{(k)}$ is determined by the asymptotics of the random variable with the heaviest tail among $N$ and $B$. Moreover, if the tails of $N$ and $B$ are equally heavy, then in fact we get the sum of two asymptotic expressions.

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

**Theorem 3.2.** (i) *If* $\mathrm{P}(B > x) = o(\mathrm{P}(N > x))$ *and* $\mathrm{P}(R^{(0)} > x) = o(\mathrm{P}(N > x))$, *then, for all* $k \geq 1$,

$$\mathrm{P}(R^{(k)} > x) \sim C_N^{(k)} \mathrm{P}(N > x) \quad \text{as } x \to \infty,$$

*where* $C_N^{(k)} = (\mathrm{E}(A))^{\alpha_N} \sum_{i=0}^{k-1} [\mathrm{E}(N)\,\mathrm{E}(A^{\alpha_N})]^i$.

(ii) *If* $\mathrm{P}(N > x) = o(\mathrm{P}(B > x))$ *and* $\mathrm{P}(R^{(0)} > x) = o(\mathrm{P}(B > x))$, *then, for all* $k \geq 1$,

$$\mathrm{P}(R^{(k)} > x) \sim C_B^{(k)} \mathrm{P}(B > x) \quad \text{as } x \to \infty,$$

*where* $C_B^{(k)} = \sum_{i=0}^{k-1} [\mathrm{E}(N)\,\mathrm{E}(A^{\alpha_B})]^i$.

(iii) *If* $\mathrm{P}(B > x) \sim C_{BN}\,\mathrm{P}(N > x)$ *for some constant* $C_{BN}$, $\mathrm{P}(R^{(0)} > x) = o(\mathrm{P}(N > x))$, *and* $\mathrm{P}(N > x,\ B > x) = o(\mathrm{P}(N > x))$, *then, for all* $k \geq 1$,

$$\mathrm{P}(R^{(k)} > x) \sim C^{(k)} \mathrm{P}(N > x) \quad \text{as } x \to \infty,$$

*where* $C^{(k)} = [C_{BN} + (\mathrm{E}(A))^{\alpha_N}] \sum_{i=0}^{k-1} [\mathrm{E}(N)\,\mathrm{E}(A^{\alpha_N})]^i$.

*Proof.* (i) We will use induction. For $k = 1$, we apply Lemma A.1(i) and (iv) to obtain

$$\begin{aligned}
\mathrm{P}(R^{(1)} > x) &= \mathrm{P}\left(\sum_{j=1}^{N} A_j R_j^{(0)} + B > x\right) \\
&\sim \mathrm{P}\left(\sum_{j=1}^{N} A_j R_j^{(0)} > x\right) \\
&\sim (\mathrm{E}(A))^{\alpha_N} \mathrm{P}(N > x) \quad \text{as } x \to \infty,
\end{aligned}$$

since $\mathrm{E}(N) < \infty$, $\mathrm{E}(A_1 R_1^{(0)}) = \mathrm{E}(A) < \infty$, and $\mathrm{P}(A_1 R_1^{(0)} > x) = o(\mathrm{P}(N > x))$. Now, assume that the result has been shown for the $(k-1)$th iteration, $k \geq 2$. Then Lemma A.1(iii) yields

$$\mathrm{P}(A_1 R_1^{(k-1)} > x) \sim C_N^{(k-1)} \mathrm{E}(A^{\alpha_N})\,\mathrm{P}(N > x). \tag{3.5}$$

Because of (3.5) and the fact that $\mathrm{E}(A_1 R_1^{(k-1)}) = \mathrm{E}(A) < \infty$, we can apply Lemma A.1(i) and (vi) to obtain

$$\begin{aligned}
\mathrm{P}(R^{(k)} > x) &\sim \mathrm{P}\left(\sum_{j=1}^{N} A_j R_j^{(k-1)} + B > x\right) \\
&\sim [C_N^{(k-1)} \mathrm{E}(A^{\alpha_N})\,\mathrm{E}(N) + (\mathrm{E}(A))^{\alpha_N}]\,\mathrm{P}(N > x) \\
&= C_N^{(k)} \mathrm{P}(N > x) \quad \text{as } x \to \infty.
\end{aligned}$$

(ii) From Lemma A.1(i) we have

$$\mathrm{P}(R^{(1)} > x) \sim \mathrm{P}\left(\sum_{j=1}^{N} A_j R_j^{(0)} + B > x\right) \sim \mathrm{P}(B > x) \quad \text{as } x \to \infty.$$

Assume that the statement holds for $(k-1)$, where $k \geq 2$. Then, from Lemma A.1(iii) we obtain

$$\mathrm{P}(A_1 R_1^{(k-1)} > x) \sim C_B^{(k-1)} \mathrm{E}(A^{\alpha_B})\,\mathrm{P}(B > x).$$

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

Because $\mathrm{E}(N) < \infty$, we apply Lemma A.1(ii) and (v) to obtain

$$\begin{aligned}\mathrm{P}(R^{(k)} > x) &\sim \mathrm{P}\Bigg(\sum_{j=1}^{N} A_j R_j^{(k-1)} + B > x\Bigg)\\ &\sim [\mathrm{E}(N) C_B^{(k-1)} \mathrm{E}(A^{\alpha_B}) + 1]\,\mathrm{P}(B > x)\\ &= C_B^{(k)}\,\mathrm{P}(B > x) \quad \text{as } x \to \infty.\end{aligned}$$

(iii) We start the induction with $k = 1$ as follows:

$$\begin{aligned}\mathrm{P}(R^{(1)} > x) &\sim \mathrm{P}\Bigg(\sum_{j=1}^{N} A_j R_j^{(0)} + B > x\Bigg)\\ &\sim (\mathrm{E}(A))^{\alpha_N}\,\mathrm{P}(N > x) + \mathrm{P}(B > x)\\ &\sim [(\mathrm{E}(A))^{\alpha_N} + C_{BN}]\,\mathrm{P}(N > x) \quad \text{as } x \to \infty,\end{aligned}$$

where we have used Lemma A.1(ii) and (iv). Next, from (3.5), $\mathrm{E}(A_1 R_1^{(k-1)}) = \mathrm{E}(A) < \infty$, and using Lemma A.1(ii) and (vi) we obtain, for any $k \geq 2$,

$$\begin{aligned}\mathrm{P}(R^{(k)} > x) &\sim \mathrm{P}\Bigg(\sum_{j=1}^{N} A_j R_j^{(k-1)} + B > x\Bigg)\\ &\sim [\mathrm{E}(N) C^{(k-1)} \mathrm{E}(A^{\alpha_N}) + (\mathrm{E}(A))^{\alpha_N} + C_{BN}]\,\mathrm{P}(N > x)\\ &= C^{(k)}\,\mathrm{P}(N > x) \quad \text{as } x \to \infty.\end{aligned}$$

With $R^{(k)}$ for $A \stackrel{\mathrm{D}}{=} c/D$ and $B \stackrel{\mathrm{D}}{=} cp_0 + (1-c)wT$, the random variable $R^{(k)}$ serves as a stochastic model for the result of the $k$th matrix iteration [24] in the PageRank computation. Since the PageRank vector is always a result of a finite number of iterations, we can conclude that the distribution of the PageRank should follow a power law with exponent $\alpha = \min(\alpha_N, \alpha_B)$. However, if the initial distribution $R^{(0)}$ has one of the heaviest tails, then the following results hold.

**Proposition 3.1.** *Let $R^{(0)}$ be a regularly varying random variable with index $\alpha_R > 0$. Then the following statements hold.*

(i) *If $\mathrm{P}(N > x) = o(\mathrm{P}(R^{(0)} > x))$ and $\mathrm{P}(B > x) = o(\mathrm{P}(R^{(0)} > x))$, then, for all $k \geq 1$,*

$$\mathrm{P}(R^{(k)} > x) \sim C_R^{(k)}\,\mathrm{P}(R^{(0)} > x) \quad \text{as } x \to \infty,$$

*where $C_R^{(k)} = \prod_{i=0}^{k} [\mathrm{E}(N)\,\mathrm{E}(A^{\alpha_R})]^i$.*

(ii) *If $\mathrm{P}(R^{(0)} > x) \sim C_{RN}\,\mathrm{P}(N > x)$ and $\mathrm{P}(B > x) = o(\mathrm{P}(R^{(0)} > x))$, then, for all $k \geq 1$,*

$$\mathrm{P}(R^{(k)} > x) \sim C_{RN}^{(k)}\,\mathrm{P}(N > x) \quad \text{as } x \to \infty,$$

*where $C_{RN}^{(k)} = [\mathrm{E}(N)\,\mathrm{E}(A^{\alpha_N})]^k C_{RN} + [\mathrm{E}(A)]^{\alpha_N} \sum_{i=0}^{k-1} [\mathrm{E}(N)\,\mathrm{E}(A^{\alpha_N})]^i$.*

(iii) *If $\mathrm{P}(N > x) = o(\mathrm{P}(R^{(0)} > x))$, $\mathrm{P}(R^{(0)} > x) \sim C_{RB}\,\mathrm{P}(B > x)$, and $\mathrm{P}(R^{(0)} > x,\ B > x) = o(\mathrm{P}(B > x))$, then, for all $k \geq 1$,*

$$\mathrm{P}(R^{(k)} > x) \sim C_{RB}^{(k)}\,\mathrm{P}(B > x) \quad \text{as } x \to \infty,$$

*where $C_{RB}^{(k)} = [\mathrm{E}(N)\,\mathrm{E}(A^{\alpha_B})]^k C_{RB} + \sum_{i=0}^{k-1} [\mathrm{E}(N)\,\mathrm{E}(A^{\alpha_B})]^i$.*

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

(iv) *If* $\mathrm{P}(R^{(0)} > x) \sim C_{RN}\,\mathrm{P}(N > x)$, $\mathrm{P}(B > x) \sim C_{BN}\,\mathrm{P}(N > x)$, $\mathrm{P}(R^{(0)} > x,\ N > x) = o(\mathrm{P}(N > x))$, *and* $\mathrm{P}(B > x,\ N > x) = o(\mathrm{P}(N > x))$, *then, for all* $k \geq 1$,

$$\mathrm{P}(R^{(k)} > x) \sim C_{RBN}^{(k)}\,\mathrm{P}(N > x) \quad \text{as } x \to \infty,$$

*where* $C_{RBN}^{(k)} = [\mathrm{E}(N)\,\mathrm{E}(A^{\alpha_N})]^k C_{RN} + [C_{BN} + [\mathrm{E}(A)]^{\alpha_N}] \sum_{i=0}^{k-1} [\mathrm{E}(N)\,\mathrm{E}(A^{\alpha_N})]^i$.

*Proof.* We again use induction. We start with $k = 1$, for which all statements are valid. Next, we assume that the result has been shown for the $(k-1)$th iteration, where $k > 2$. Then we consider every case respectively.

(i) We apply Lemma A.1(i), (iii), and (v) to obtain

$$\begin{aligned}
\mathrm{P}(R^{(k)} > x) &= \mathrm{P}\left(\sum_{j=1}^{N} A_j R_j^{(k-1)} + B > x\right) \\
&\sim \mathrm{P}\left(\sum_{j=1}^{N} A_j R_j^{(k-1)} > x\right) \\
&\sim \mathrm{E}(N)\,\mathrm{E}(A^{\alpha_R})\,\mathrm{P}(R^{(k-1)} > 0) \\
&= C_R^{(k)}\,\mathrm{P}(R^{(0)} > 0).
\end{aligned}$$

(ii) In this case we have

$$\begin{aligned}
\mathrm{P}(R^{(k)} > x) &= \mathrm{P}\left(\sum_{j=1}^{N} A_j R_j^{(k-1)} + B > x\right) \\
&\sim \mathrm{P}\left(\sum_{j=1}^{N} A_j R_j^{(k-1)} > x\right) \\
&\sim [\mathrm{E}(A^{\alpha_N})\,\mathrm{E}(N) C_{RN}^{(k-1)} + (\mathrm{E}(A))^{\alpha_N}]\,\mathrm{P}(N > x) \\
&= C_{RN}^{(k)}\,\mathrm{P}(N > x),
\end{aligned}$$

where we have used Lemma A.1(i), (iii), and (vi).

(iii) From Lemma A.1(ii), (iii), and (v), we obtain the statement

$$\begin{aligned}
\mathrm{P}(R^{(k)} > x) &= \mathrm{P}\left(\sum_{j=1}^{N} A_j R_j^{(k-1)} + B > x\right) \\
&\sim \mathrm{P}\left(\sum_{j=1}^{N} A_j R_j^{(k-1)} > x\right) + \mathrm{P}(B > x) \\
&\sim [\mathrm{E}(A^{\alpha_B})\,\mathrm{E}(N) C_{RB}^{(k-1)} + 1]\,\mathrm{P}(B > x) \\
&= C_{RB}^{(k)}\,\mathrm{P}(B > x).
\end{aligned}$$

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

(iv) Here we use Lemma A.1(ii), (iii), and (vi) to obtain

$$\begin{aligned}\mathrm{P}(R^{(k)} > x) &= \mathrm{P}\Bigg(\sum_{j=1}^{N} A_j R_j^{(k-1)} + B > x\Bigg)\\ &\sim \mathrm{P}\Bigg(\sum_{j=1}^{N} A_j R_j^{(k-1)} > x\Bigg) + \mathrm{P}(B > x)\\ &\sim [\mathrm{E}(A^{\alpha_N})\,\mathrm{E}(N) C_{RBN}^{(k-1)} + (\mathrm{E}(A))^{\alpha_N} + C_{BN}]\,\mathrm{P}(N > x)\\ &= C_{RBN}^{(k)}\,\mathrm{P}(N > x).\end{aligned}$$

### 3.4. Asymptotics: from $R^{(k)}$ to $R^{(\infty)}$

Combining the results from Theorems 3.1 and 3.2, we can assume the following asymptotic similarities for $R^{(\infty)}$, the unique nontrivial solution of (3.1).

(i) If $\mathrm{P}(B > x) = o(\mathrm{P}(N > x))$ then $\mathrm{P}(R^{(\infty)} > x) \sim C_N\,\mathrm{P}(N > x)$ as $x \to \infty$, where $C_N = \lim_{k\to\infty} C_N^{(k)} = (\mathrm{E}(A))^{\alpha_N}[1 - \mathrm{E}(N)\,\mathrm{E}(A^{\alpha_N})]^{-1}$.

(ii) If $\mathrm{P}(N > x) = o(\mathrm{P}(B > x))$ then $\mathrm{P}(R^{(\infty)} > x) \sim C_B\,\mathrm{P}(B > x)$ as $x \to \infty$, where $C_B = \lim_{k\to\infty} C_B^{(k)} = [1 - \mathrm{E}(N)\,\mathrm{E}(A^{\alpha_B})]^{-1}$.

(iii) If $\mathrm{P}(B > x) \sim C_{BN}\,\mathrm{P}(N > x)$ for some constant $C_{BN}$, and $\mathrm{P}(N > x,\ B > x) = o(\mathrm{P}(N > x))$, then $\mathrm{P}(R^{(\infty)} > x) \sim C\,\mathrm{P}(N > x)$ as $x \to \infty$, where $C = \lim_{k\to\infty} C^{(k)} = [C_{BN} + (\mathrm{E}(A))^{\alpha_N}][1 - \mathrm{E}(N)\,\mathrm{E}(A^{\alpha_N})]^{-1}$.

Proving these results by probabilistic methods requires an exchange of limits in $x$ and $k$, which is usually a difficult technical problem. Indeed, if we assume that $\mathrm{P}(R^{(k)} > x) \sim h_k(x)$ as $x \to \infty$ for every $k$ and some function $h_k(x)$, then $\mathrm{P}(R^{(\infty)} > x) \sim \lim_{k\to\infty} h_k(x)$ is not true in general. For instance, from Proposition 3.1 we know that the asymptotics of $R^{(k)}$ can be defined by the asymptotics of $R^{(0)}$, whereas representation (3.4) clarifies that $R^{(\infty)}$ does not depend on the distribution of $R^{(0)}$. In the next section we prove the above similarities using a Laplace–Stieltjes transform analysis.

## 4. Laplace–Stieltjes transform analysis

As in our previous work [26], we follow the technique of [12]. We start with an equation for the Laplace–Stieltjes transforms of $N$, $B$, and $R$. The idea is to use this equation and the Tauberian theorem (see Theorem A.1) to classify the asymptotic behavior of $R$. To this end, we first show that the conditions of Theorem A.1 are satisfied. In particular, in Lemmas 4.1 and 4.2 we justify the fact that the existence of the $k$th moments of $N$ and $B$ implies the existence of the $k$th moment of $R$, and vice versa. Then we define the necessary equivalences for the Laplace–Stieltjes transforms of $N$, $B$, and $R$ in Corollary 4.1, and obtain the main result in Theorem 4.1.

In this section we need to assume that $A < 1$ and that $\alpha = \min(\alpha_N, \alpha_B) > 1$ is a noninteger. Moreover, we model the in-degree $N$ as the number of Poisson(1) events on $[0, X]$, where $X$ is a regular varying random variable with index $\alpha_N$. The asymptotic behavior of $N(X)$ is given by (2.1).

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

### 4.1. Equation for Laplace–Stieltjes transforms

Define $f(s)$ and $\phi(s)$ to be the Laplace–Stieltjes transforms of $X$ and $N = N(X)$, respectively, where $X$ is regularly varying with index $\alpha_N$ and $N(x)$ is the number of Poisson arrivals on the time interval $[0, x]$, as before. Then we can write the following expression:

$$\phi(s) = \mathrm{E}(\mathrm{e}^{-sN}) = f(1 - \mathrm{e}^{-s}). \tag{4.1}$$

Moreover, since the corresponding moments of $X$ and $N$ always exist together [26], we use only moments of $X$, and we denote them by $\xi_0 = 1, \xi_1 = \mathrm{E}(N), \xi_2, \ldots, \xi_n$. Then, provided that $\xi_n$ is finite, we define

$$f_n(s) = (-1)^{n+1}\left(f(s) - \sum_{i=0}^{n} \frac{\xi_i}{i!}(-s)^i\right).$$

Next, we denote the first $m$ moments of $B$ by $\beta_1, \beta_2, \ldots, \beta_m$, and let $\beta_0 = 1$. Then, provided that $\beta_m$ is finite, we define

$$b_m(s) = (-1)^{m+1}\left(b(s) - \sum_{i=0}^{m} \frac{\beta_i}{i!}(-s)^i\right),$$

where $b(s)$ is the Laplace–Stieltjes transform of $B$.

We also introduce the following function:

$$G(t, s) = \mathrm{E}(\mathrm{e}^{-tX}\mathrm{e}^{-sB}),$$

where it is easy to see that $G(t, 0) = f(t)$ and $G(0, s) = b(s)$. Moreover, if $X$ and $B$ are independent, implying that $N$ and $B$ are independent, then we have

$$G(t, s) = f(t)b(s).$$

Let $r(s)$ be the Laplace–Stieltjes transform of $R$. Then, by (3.1) and (4.1), the following holds:

$$\begin{aligned} r(s) &= \mathrm{E}(\mathrm{e}^{-sR}) \\ &= \mathrm{E}\left(\exp\left(-s\sum_{j=1}^{N} A_j R_j\right)\mathrm{e}^{-sB}\right) \\ &= \mathrm{E}\left(\mathrm{E}\left(\exp\left(-s\sum_{j=1}^{N} A_j R_j\right)\mathrm{e}^{-sB} \;\middle|\; N, B\right)\right) \\ &= G[1 - \mathrm{E}(r(As)), s]. \end{aligned}$$

Thus, we derive the equation

$$r(s) = G[1 - \mathrm{E}(r(As)), s]. \tag{4.2}$$

Define

$$t(s) = 1 - \mathrm{E}(r(As)), \tag{4.3}$$

and write (4.2) as

$$r(s) = G(t(s), s). \tag{4.4}$$

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

### 4.2. Auxiliary results

We define $\rho_1, \dots, \rho_k$ to be the first $k$ moments of $R$. If $\rho_k < \infty$, we have

$$r_k(s) = (-1)^{k+1}\left(r(s) - \sum_{i=0}^{k} \frac{\rho_i}{i!}(-s)^i\right), \tag{4.5}$$

as in Lemma A.2.

We define $k = \min(m, n)$, where $m$ and $n$ are integers, and such that $\beta_m = \mathrm{E}(B^m) < \infty$ and $\xi_n = \mathrm{E}(X^n) < \infty$. Next, we assume that $\mathrm{E}(X^j B^{k+1-j}) < \infty$ for all $0 < j < k+1$. We note that this assumption is always true in the case of the independent $N$ and $B$. Then we can prove the following lemma.

**Lemma 4.1.** *If $\xi_n < \infty$ and $\beta_m < \infty$ for some integers $m, n \geq 1$, and $\mathrm{E}(X^j B^{k+1-j}) < \infty$ for all $0 < j < k+1$, where $k = \min(m, n)$, then $\rho_k < \infty$.*

*Proof.* We use induction, starting from $k = 1$, for which the statement is valid. Assume that, for $i = 1, 2, \dots, k-1$, the lemma has been proved, so we can use the extension

$$r(s) = 1 - s + \sum_{i=2}^{k-1} \frac{\rho_i}{i!}(-s)^i + o(s^{k-1})$$

to present $t(s)$ as a sum:

$$t(s) = -\mathrm{E}\left(\sum_{i=1}^{k-1} \frac{\rho_i}{i!} A^i (-s)^i + o(s^{k-1})\right) = -\sum_{i=1}^{k-1} \frac{\rho_i}{i!}\,\mathrm{E}(A^i)(-s)^i + o(s^{k-1}).$$

As a result of this, we can actually obtain $t^i(s)$:

$$t^i(s) = \sum_{j=i}^{k+i-2} \zeta_{i,j} s^j + o(s^{k+i-2}) \tag{4.6}$$

for $i \geq 1$ and some appropriate constants $\zeta_{i,j},\ j = i, \dots, k+i-2$.

Now, we consider the Taylor expansion of $G(t(s), s)$:

$$\begin{aligned} G(t(s), s) = {} & \left[\sum_{i=0}^{k} \frac{\xi_i}{i!}(-t(s))^i + (-1)^{k+1} f_k(t(s))\right] + \left[\sum_{i=0}^{k} \frac{\beta_i}{i!}(-s)^i + (-1)^{k+1} b_k(s)\right] \\ & - 1 + \sum_{i=0}^{k+1} \frac{(-1)^i}{i!} \sum_{j=1}^{i-1} \binom{i}{j} \mathrm{E}(X^j B^{i-j}) t^j(s) s^{i-j} + o(s^{k+1}), \end{aligned} \tag{4.7}$$

where $t(s) \sim \mathrm{E}(A)s$. Here we used the fact that

$$G'_{t^j s^{i-j}}(0, 0) = (-1)^i\,\mathrm{E}(X^j B^{i-j}) < \infty \quad \text{for all } 0 \leq i \leq k+1,\ 0 < j < k+1.$$

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

Then, from (4.3), (4.4), and (4.7), we obtain

$$\begin{aligned} r(s) &= 1 - \mathrm{E}(N)t(s) + \left[\sum_{i=2}^{k} \frac{\xi_i}{i!}(-t(s))^i + (-1)^{k+1} f_k(t(s))\right] \\ &\quad + \left[\sum_{i=0}^{k} \frac{\beta_i}{i!}(-s)^i + (-1)^{k+1} b_k(s)\right] - 1 + \sum_{i=0}^{k+1} \frac{(-1)^i}{i!} \sum_{j=1}^{i-1} \binom{i}{j} \mathrm{E}(X^j B^{i-j}) t^j(s) s^{i-j} \\ &\quad + o(s^{k+1}) \\ &= 1 - \mathrm{E}(N)[1 - \mathrm{E}(r(As))] + \sum_{i=1}^{k} \eta_i s^i + o(s^k), \end{aligned}$$

where we have used (4.6), and the facts that $f_k(t(s)) = o(s^k)$ and $b_k(s) = o(s^k)$ to find appropriate constants $\eta_1, \ldots, \eta_k$. Next, we rewrite the last equation as

$$r(s) - \mathrm{E}(N)\,\mathrm{E}(r(As)) = 1 - \mathrm{E}(N) + \sum_{i=1}^{k} \eta_i s^i + o(s^k),$$

and apply (4.5) to obtain

$$\begin{aligned} &r_{k-1}(s) - \mathrm{E}(N)\,\mathrm{E}(r_{k-1}(As)) + (-1)^k \sum_{i=0}^{k-1} \frac{\rho_i}{i!}(1 - \mathrm{E}(A^i))(-s)^i \\ &\qquad = 1 - \mathrm{E}(N) + \sum_{i=0}^{k} \eta_i s^i + o(s^k). \end{aligned}$$

Because $r_{k-1}(s) = o(s^{k-1})$, $\mathrm{E}(r_{k-1}(As)) = o(s^{k-1})$, and the uniqueness of the series expansion, we can remove all powers up to $k$:

$$r_{k-1}(s) - \mathrm{E}(N)\,\mathrm{E}(r_{k-1}(As)) = \eta_k s^k + o(s^k). \tag{4.8}$$

Now, we let $A_1, A_2, \ldots$ be independent and distributed as $A$. We consider the partial sums

$$\begin{aligned} &\sum_{j=0}^{M} (\mathrm{E}(N))^j [\mathrm{E}(r_{k-1}(A_1 \cdots A_j s)) - \mathrm{E}(N)\,\mathrm{E}(r_{k-1}(A_1 \cdots A_{j+1} s))] \\ &\qquad = r_{k-1}(s) - (\mathrm{E}(N))^{M+1}\,\mathrm{E}(r_{k-1}(A_1 \cdots A_{M+1} s)). \end{aligned}$$

We claim that the second term converges to 0 as $M \to \infty$. From the induction hypothesis and the definition of $o(s^{k-1})$, for all $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon)$ such that $|r_{k-1}(s)| < \varepsilon s^{k-1}$ whenever $0 < s \le \delta$. Fix some $\varepsilon$ and take $\delta = \delta(\varepsilon)$. Then the following holds:

$$\mathrm{E}\,|r_{k-1}(A_1 \cdots A_{M+1} s)| < \varepsilon s^{k-1}\,\mathrm{E}(A_1^{k-1} \cdots A_{M+1}^{k-1}) = \varepsilon s^{k-1} (\mathrm{E}(A^{k-1}))^{M+1},$$

where the final equality holds due to the independence of the $A$s. Taking the limit as $M \to \infty$, since $\mathrm{E}(B) < 1$, $A < 1$, $\mathrm{E}(A) = (1 - \mathrm{E}(B))/\mathrm{E}(N)$, and $\mathrm{E}(A^{n-1}) \le \mathrm{E}(A)$, we have

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

$\lim_{M\to\infty} \mathrm{E}(N)^{M+1}\,\mathrm{E}(r_{k-1}(A_1\cdots A_{M+1}s)) = 0$. It follows that we can express $r_{k-1}(s)$ as an infinite sum:

$$r_{k-1}(s) = \sum_{j=0}^{\infty} (\mathrm{E}(N))^j [\mathrm{E}(r_{k-1}(A_1\cdots A_j s)) - \mathrm{E}(N)\,\mathrm{E}(r_{k-1}(A_1\cdots A_{j+1}s))], \tag{4.9}$$

where we can apply (4.8) to each of the terms. From the definition of $o(s^k)$, for every $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon)$ such that

$$|r_{k-1}(s) - \mathrm{E}(N)\,\mathrm{E}(r_{k-1}(As)) - \eta_k s^k| < \varepsilon s^k$$

whenever $0 < s \le \delta$. Moreover, for this $\varepsilon$ and $0 < s \le \delta$, we also have

$$\begin{aligned}
&|\mathrm{E}(r_{k-1}(A_1\cdots A_j s)) - \mathrm{E}(N)\,\mathrm{E}(r_{k-1}(A_1\cdots A_{j+1}s)) - \eta_k s^k\,\mathrm{E}(A_1^k\cdots A_j^k)| \\
&\quad \le \mathrm{E}\,|\mathrm{E}(r_{k-1}(A_1\cdots A_j s) - \mathrm{E}(N) r_{k-1}(A_1\cdots A_{j+1}s) - \eta_k s^k A_1^k\cdots A_j^k \mid A_1,\ldots,A_j)| \\
&\quad < \varepsilon s^k (\mathrm{E}(A^k))^j
\end{aligned}$$

for every $j \ge 0$ and $A_1,\ldots,A_{j+1}$, which are independent and distributed as $A$. Here the last inequality holds because $A < 1$, and then $0 < A_1\cdots A_{j+1}s \le s < \delta$ for every $j \ge 0$. Using the representation of $r_{k-1}(s)$ as an infinite sum, (4.9), we obtain

$$\begin{aligned}
&\left|r_{k-1}(s) - \eta_k \sum_{j=0}^{\infty} (\mathrm{E}(N))^j\,\mathrm{E}(A_1^k\ldots A_j^k)s^k\right| \\
&\quad = \left|\sum_{j=0}^{\infty} (\mathrm{E}(N))^j [\mathrm{E}(r_{k-1}(A_1\cdots A_j s)) - \mathrm{E}(N)\,\mathrm{E}(r_{k-1}(A_1\ldots A_{j+1}s))]\right. \\
&\qquad \left. - \eta_k \sum_{j=0}^{\infty} (\mathrm{E}(N))^j\,\mathrm{E}(A_1^k\cdots A_j^k)s^k\right| \\
&\quad \le \varepsilon s^k \sum_{j=1}^{\infty} (\mathrm{E}(N)\,\mathrm{E}(A^k))^j \\
&\quad = \varepsilon[1 - \mathrm{E}(N)\,\mathrm{E}(A^k)]^{-1} s^k.
\end{aligned}$$

Thus, we have shown that $r_{k-1}(s) - \eta_k[1 - \mathrm{E}(N)\,\mathrm{E}(A^k)]^{-1}s^k = o(s^k)$. Taking $\rho_k = -\eta_k[1 - \mathrm{E}(N)\,\mathrm{E}(A^k)]^{-1}$, from Lemma A.2 and the last equation, we conclude that $\rho_k$ is the $k$th moment of $R$ and that it is finite.

We can also prove the converse of Lemma 4.1.

**Lemma 4.2.** *If $\rho_k < \infty$, $k \ge 1$, then $\xi_k < \infty$ and $\beta_k < \infty$.*

*Proof.* Let $R$ be a nonnegative random variable that satisfies (3.1) and has finite $k$th moment. Equation (3.1) implies that $R$ is stochastically greater than $B$, and, thus, $R$ is also stochastically greater than $B(AN(X)+1)$. Hence, the existence of the $k$th moment of $R$ ensures the existence of the $k$th moment of $B$ and $N(X)$, which in turn ensures the existence of the $k$th moment of $X$.

The next corollary follows from the proof of Lemma 4.1.

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

**Corollary 4.1.** *It follows from Lemma 4.1 that*

(i) *if* $n < m$ *then* $r_n(s) - \mathrm{E}(N)\,\mathrm{E}(r_n(As)) = f_n(t(s)) + O(s^{n+1})$,

(ii) *if* $n > m$ *then* $r_m(s) - \mathrm{E}(N)\,\mathrm{E}(r_m(As)) = b_m(s) + O(s^{m+1})$,

(iii) *if* $n = m$ *then* $r_n(s) - \mathrm{E}(N)\,\mathrm{E}(r_n(As)) = f_n(t(s)) + b_n(s) + O(s^{n+1})$.

*Proof.* Recall that $k = \min(m, n)$. Because $r_k(s) = o(s^k)$, we can consider the following expansion of (4.6):

$$t^i(s) = \sum_{j=l}^{k+i-1} \zeta_{i,j} s^j + o(s^{k+i-1}) \tag{4.10}$$

for $i \geq 1$ and appropriate constants $\zeta_{i,j},\ j = i, \ldots, k+i-1$.

From (4.4), (4.7), (4.10), the definitions of $r_k(s)$, $b_k(t)$, and $t(s)$, and Lemma 4.1, it follows that

$$\begin{aligned}
&(-1)^{k+1} r_k(s) + \sum_{i=0}^{k} \frac{\rho_i}{i!}(-s)^i \\
&\quad = (-1)^{k+1} f_k(t(s)) + \sum_{i=2}^{k} \frac{\xi_i}{i!}(-t(s))^i + 1 \\
&\qquad - \mathrm{E}(N)\left[1 - \mathrm{E}\left((-1)^{k+1} r_k(As) + \sum_{i=0}^{k} \frac{\rho_i}{i!}(-As)^i\right)\right] - 1 \\
&\qquad + \sum_{i=0}^{k} \frac{\beta_i}{i!}(-s)^i + (-1)^{k+1} b_k(s) \\
&\qquad + \sum_{i=0}^{k+1} \frac{(-1)^i}{i!} \sum_{j=1}^{i-1} \binom{i}{j} \mathrm{E}(X^j B^{i-j}) t^j(s) s^{i-j} + o(s^{k+1}) \\
&\quad = (-1)^{k+1}[b_k(s) + f_k(t) + \mathrm{E}(N)\,\mathrm{E}(r_k(As))] + \sum_{i=0}^{k+1} \varsigma_i s^i + o(s^{k+1}),
\end{aligned}$$

where $\varsigma_0, \ldots, \varsigma_{k+1}$ are appropriate constants. Due to the uniqueness of the series expansion, we can reduce the above formula to

$$r_k(s) = b_k(s) + f_k(t) + \mathrm{E}(N)\,\mathrm{E}(r_k(As)) + (-1)^{k+1} \varsigma_{k+1} s^{k+1} + o(s^{k+1}).$$

The corollary follows because $t(s) \sim \mathrm{E}(A)s$ as $s \to 0$.

Now we are ready to prove our main result.

### 4.3. Main theorem

In the next theorem we obtain our main result that establishes the tail behavior of the PageRank distribution under various assumptions on the distribution of the in-degree and the teleportation.

**Theorem 4.1.** (i) *If* $\mathrm{P}(B > x) = o(\mathrm{P}(N > x))$ *then the following statements are equivalent:*

(a) $\mathrm{P}(N > x) \sim x^{-\alpha_N} L_N(x)$ *as* $x \to \infty$,

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

(b) $\mathrm{P}(R > x) \sim C_N x^{-\alpha_N} L_N(x)$ *as* $x \to \infty$, *where* $C_N = (\mathrm{E}(A))^{\alpha_N}[1 - \mathrm{E}(N)\,\mathrm{E}(A^{\alpha_N})]^{-1}$.

(ii) *If* $\mathrm{P}(N > x) = o(\mathrm{P}(B > x))$ *then the following statements are equivalent:*

(a) $\mathrm{P}(B > x) \sim x^{-\alpha_B} L_B(x)$ *as* $x \to \infty$,

(b) $\mathrm{P}(R > x) \sim C_B x^{-\alpha_B} L_B(x)$ *as* $x \to \infty$, *where* $C_B = [1 - \mathrm{E}(N)\,\mathrm{E}(A^{\alpha_B})]^{-1}$.

(iii) *If* $\mathrm{P}(B > x) \sim C_{BN}\,\mathrm{P}(N > x)$ *then the following statements are equivalent:*

(a) $\mathrm{P}(N > x) \sim x^{-\alpha_N} L_N(x)$ *and* $\mathrm{P}(B > x) \sim x^{-\alpha_N} L_B(x) \sim C_{BN} x^{-\alpha_N} L_N(x)$ *as* $x \to \infty$,

(b) $\mathrm{P}(R > x) \sim C x^{-\alpha_N} L_N(x)$ *as* $x \to \infty$, *where*

$$C = [C_{BN} + (\mathrm{E}(A))^{\alpha_N}][1 - \mathrm{E}(N)\,\mathrm{E}(A^{\alpha_N})]^{-1}.$$

The results of Theorem 4.1 describe the tail behavior of $R$ under various assumptions on the distribution of the Web parameters. First of all, we observe that the power law exponent is defined by the random variable with the heaviest tail among $N$ and $B$, representing the in-degree and the user preference, respectively. Next, we see that the obtained multiplicative constants agree with the results of Section 3.4. When $B$ has a lighter tail than $N$, we observe that the distribution of $B$ has no influence on the asymptotics of the PageRank. In the next case we find that $C_B$ depends only on the mean value of the in-degree $\mathrm{E}(N)$, and in the case of the similar tails of $N$ and $B$ we have the effects from both of them. We also note that if $A$ is defined as $c/D$ then $\mathrm{E}(A) = c(1 - p_0)/\mathrm{E}(N)$. So, the obtained constants also depend on the damping factor $c$ and the fraction of the dangling nodes $p_0$. The distribution of the effective out-degree $D$ has a negligible effect.

*Proof of Theorem 4.1((i)(a), (ii)(a), (iii)(a)) ⇒ ((i)(b), (ii)(b), (iii)(b)).* It follows from parts (a) of (i), (ii), and (iii), and Theorem A.1 that

(i) $f_n(t) \sim (-1)^n \Gamma(1 - \alpha_N) t^{\alpha_N} L_N(1/t)$ as $t \to 0$,

(ii) $b_m(s) \sim (-1)^m \Gamma(1 - \alpha_B) s^{\alpha_B} L_B(1/s)$ as $s \to 0$,

(iii) both statements under (i) and (ii) hold,

where $m$ and $n$ are the largest integer values not exceeding $\alpha_B$ and $\alpha_N$, respectively.

Recall that $t(s) \sim \mathrm{E}(A)s$ as $s \to 0$, because of (4.3) and $r(s) = 1 - s + o(s)$. Then, by applying Corollary 4.1 we can obtain, as $s \to 0$,

(i) $r_n(s) - \mathrm{E}(N)\,\mathrm{E}(r_n(As)) \sim (-1)^n \Gamma(1 - \alpha_N)(\mathrm{E}(A))^{\alpha_N} L_N(1/s) s^{\alpha_N}$,

(ii) $r_m(s) - \mathrm{E}(N)\,\mathrm{E}(r_m(As)) \sim (-1)^m \Gamma(1 - \alpha_B) L_B(1/s) s^{\alpha_B}$,

(iii) $r_n(s) - \mathrm{E}(N)\,\mathrm{E}(r_n(As)) \sim (-1)^n \Gamma(1 - \alpha_N)[(\mathrm{E}(A))^{\alpha_N} L_N(1/s) + L_B(1/s)] s^{\alpha_N}$.

Let $V_N$ and $V_B$ be constants that are defined as follows:

- $V_N = (\mathrm{E}(A))^{\alpha}$ and $V_B = 0$,
- $V_N = 0$ and $V_B = 1$,
- $V_N = (\mathrm{E}(A))^{\alpha}$ and $V_B = 1$.

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

Next, we define

$$Z(s) = r_k(s) - \mathrm{E}(N)\,\mathrm{E}(r_k(As)),$$
$$Y(s) = (-1)^k\Gamma(1-\alpha)\left[V_N L_N\left(\frac{1}{s}\right) + V_B L_B\left(\frac{1}{s}\right)\right]s^{\alpha},$$

where $\alpha = \min(\alpha_N, \alpha_B)$ and $k = \min(n, m)$. We note that $Y(s) \geq 0$ for every $s > 0$.

We prove the statement of the theorem in two steps. First, we use the representation (4.9) for $r_k(s)$, and show that the following asymptotic similarity holds:

$$\sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,\mathrm{E}(Z(A_1\cdots A_i s)) \sim \sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,\mathrm{E}(Y(A_1\cdots A_i s)) \tag{4.11}$$

as $s \to 0$. Second, we demonstrate that the right-hand side of (4.11) has the desired asymptotics.

As we saw above, $Z(s) \sim Y(s)$ as $s \to 0$. Then, for every $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon)$ such that $|Z(s)/Y(s) - 1| < \varepsilon$ whenever $0 < s \leq \delta$. We fix some $\varepsilon$ and take $\delta = \delta(\varepsilon)$. Now, again, let $A_1, A_2, \ldots$ be independent random variables, which are distributed as $A$. Because $A < 1$, and then $0 < A_1\cdots A_i s \leq s \leq \delta$ for every $i \geq 0$, we have

$$\left|\frac{Z(A_1\cdots A_i s)}{Y(A_1\cdots A_i s)} - 1\right| < \varepsilon. \tag{4.12}$$

From (4.12) we obtain

$$\begin{aligned}
&\left|\frac{\sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,\mathrm{E}(Z(A_1\cdots A_i s))}{\sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,\mathrm{E}(Y(A_1\cdots A_i s))} - 1\right| \\
&\quad\leq \frac{\sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,|\mathrm{E}(Z(A_1\cdots A_i s) - Y(A_1\cdots A_i s))|}{|\sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,\mathrm{E}(Y(A_1\cdots A_i s))|} \\
&\quad\leq \frac{\sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,\mathrm{E}(|Z(A_1\cdots A_i s)/Y(A_1\cdots A_i s) - 1|Y(A_1\cdots A_i s))}{\sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,\mathrm{E}(Y(A_1\cdots A_i s))} \\
&\quad< \frac{\varepsilon\sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,\mathrm{E}(Y(A_1\cdots A_i s))}{\sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,\mathrm{E}(Y(A_1\cdots A_i s))} \\
&\quad= \varepsilon,
\end{aligned}$$

which implies (4.11).

Next, we use Lemma A.3, and then, for every $\vartheta > 1$ and $\delta > 0$, we can find finite constants $s_B$ and $s_N$ such that, for all $i > 0$ and $0 < s < \min(s_B, s_N)$,

$$\vartheta^{-1}(A_1\cdots A_i)^{\delta} \leq \frac{L_B(1/(A_1\cdots A_i s))}{L_B(1/s)} \leq \vartheta(A_1\cdots A_i)^{-\delta}$$

and

$$\vartheta^{-1}(A_1\cdots A_i)^{\delta} \leq \frac{L_N(1/(A_1\cdots A_i s))}{L_N(1/s)} \leq \vartheta(A_1\cdots A_i)^{-\delta}. \tag{4.13}$$

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

We divide the right-hand side of equation (4.11) by $L_B(1/s)L_N(1/s)$, and apply (4.13) to $Y(A_1 \cdots A_i s)/L_B(1/s)L_N(1/s)$ to obtain

$$\begin{aligned}
&\vartheta^{-1}(-1)^k\Gamma(1-\alpha)\left(\frac{V_N}{L_B(1/s)}+\frac{V_B}{L_N(1/s)}\right)s^\alpha\sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,\mathrm{E}((A_1\cdots A_i)^{\alpha+\delta})\\
&\quad\le\frac{\sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,\mathrm{E}(Y(A_1\cdots A_i s))}{L_B(1/s)L_N(1/s)}\\
&\quad\le\vartheta(-1)^k\Gamma(1-\alpha)\left(\frac{V_N}{L_B(1/s)}+\frac{V_B}{L_N(1/s)}\right)s^\alpha\sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,\mathrm{E}((A_1\cdots A_i)^{\alpha-\delta}).
\end{aligned}$$

Because $A_1, A_2, \ldots$ are independent and identically distributed as $A$, we can conclude that

$$\begin{aligned}
&\vartheta^{-1}(-1)^k\Gamma(1-\alpha)\left(\frac{V_N}{L_B(1/s)}+\frac{V_B}{L_N(1/s)}\right)s^\alpha\frac{1}{1-\mathrm{E}(N)\,\mathrm{E}(A^{\alpha+\delta})}\\
&\quad\le\frac{\sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,\mathrm{E}(Y(A_1\cdots A_i s))}{L_B(1/s)L_N(1/s)}\\
&\quad\le\vartheta(-1)^k\Gamma(1-\alpha)\left(\frac{V_N}{L_B(1/s)}+\frac{V_B}{L_N(1/s)}\right)s^\alpha\frac{1}{1-\mathrm{E}(N)\,\mathrm{E}(A^{\alpha-\delta})}.
\end{aligned}$$

Taking $\vartheta \to 1$ and $\delta \to 0$ by the dominated convergence we obtain

$$\begin{aligned}
\sum_{i=0}^{\infty}(\mathrm{E}(N))^i\,\mathrm{E}(Y(A_1\cdots A_i s)) &\sim (-1)^k\Gamma(1-\alpha)[1-\mathrm{E}(N)\,\mathrm{E}(A^\alpha)]^{-1}\\
&\quad\times\left(\frac{V_N}{L_B(1/s)}+\frac{V_B}{L_N(1/s)}\right)L_B\left(\frac{1}{s}\right)L_N\left(\frac{1}{s}\right)s^\alpha \quad \text{as } s\to 0.
\end{aligned}$$

Combining the last equivalence, (4.11), and the infinite-sum representation (4.9) for $r_k(s)$, we obtain

$$r_k(s)=\sum_{i=0}^{\infty}(\mathrm{E}(N))^i[\mathrm{E}(r_k(A_1\cdots A_i s))-\mathrm{E}(N)\,\mathrm{E}(r_k(A_1\cdots A_i s))].$$

Then

$$r_k(s)\sim(-1)^k\Gamma(1-\alpha)\left[V_N L_N\left(\frac{1}{s}\right)+V_B L_B\left(\frac{1}{s}\right)\right][1-\mathrm{E}(N)\,\mathrm{E}(A^\alpha)]^{-1}s^\alpha \tag{4.14}$$

as $s \to 0$. Now, we again apply Theorem A.1, leading to the statement of the theorem.

*((i)(a), (ii)(a), (iii)(a)) ⇐ ((i)(b), (ii)(b), (iii)(b)).* We define $V_N$ and $V_B$, $k = \min(n, m)$, and $\alpha \in (k, k+1)$ as before. Then, from parts (b) of (i), (ii), and (iii), and Theorem A.1, we can obtain (4.14), leading to the asymptotic equivalence

$$r_k(s)-\mathrm{E}(N)\,\mathrm{E}(r_k(As))\sim(-1)^k\Gamma(1-\alpha)L\left(\frac{1}{s}\right)[1-\mathrm{E}(N)\,\mathrm{E}(A^\alpha)]^{-1}s^\alpha \tag{4.15}$$

as $s \to 0$, where we define

$$\begin{aligned}
L\left(\frac{1}{s}\right)&=V_N\left[L_N\left(\frac{1}{s}\right)-\mathrm{E}(N)\,\mathrm{E}\left(A^\alpha L_N\left(\frac{1}{As}\right)\right)\right]\\
&\quad+V_B\left[L_B\left(\frac{1}{s}\right)-\mathrm{E}(N)\,\mathrm{E}\left(A^\alpha L_B\left(\frac{1}{As}\right)\right)\right].
\end{aligned}$$

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

Next, we again use the bounds in (4.13) to obtain

$$\begin{aligned}\left[\frac{V_N}{L_B(1/s)}+\frac{V_B}{L_N(1/s)}\right][1-\vartheta^{-1}\,\mathrm{E}(N)\,\mathrm{E}(A^{\alpha+\delta})]\\ \le \frac{L(1/s)}{L_N(1/s)L_B(1/s)}\\ \le \left[\frac{V_N}{L_B(1/s)}+\frac{V_B}{L_N(1/s)}\right][1-\vartheta\,\mathrm{E}(N)\,\mathrm{E}(A^{\alpha-\delta})].\end{aligned}$$

Thus, by the dominated convergence for $\vartheta \to 1$ and $\delta \to 0$, we have

$$L\left(\frac{1}{s}\right) \sim [1-\mathrm{E}(N)\,\mathrm{E}(A^{\alpha})]\left[C_N L_N\left(\frac{1}{s}\right)+C_B L_B\left(\frac{1}{s}\right)\right].$$

From the last similarity and (4.15), we obtain

$$r_k(s)-\mathrm{E}(N)\,\mathrm{E}(r(As)) \sim (-1)^k\Gamma(1-\alpha)\left[V_N L_N\left(\frac{1}{s}\right)+V_B L_B\left(\frac{1}{s}\right)\right]s^{\alpha}$$

as $s \to 0$, from where, by applying Corollary 4.1, we show parts (a) of (i), (ii), and (iii).

## 5. Numerical results

In order to illustrate the results of Theorem 4.1, we perform a number of small-scale experiments. More numerical results can be found in [37], where we considered a simpler model of the standard PageRank with uniform teleportation. Here we use the Stanford data set (see www.kamvar.org/personalized_search (accessed in April 2009)) with $w = 281\,903$ pages and $2\,312\,497$ links. It is a relatively small Web sample; however, it is known to possess the basic properties of the Web. In particular, in this data set, the in-degree shows typical power law behavior with exponent $\alpha_N = 1.1$.

We create the teleportation distribution by using the inverse transformation method. First, we generate random numbers $u_1, \dots, u_w$ from the standard uniform distribution, and then we set $t_i = (1-u_i)^{-1/\alpha_B}$, where $i = 1, \dots, w$. These $t_i$s are random numbers that are Pareto distributed with exponent $\alpha_B$. We choose $\alpha_B = 0.5$, 1.1, and 3.0. Second, we denote $\bar{t}$ as the mean value of $t_1, \dots, t_w$, and define the teleportation probability of a jump to page $i$ as $T(i) = t_i/(w\bar{t})$. Next, we use (1.3) to obtain personalized PageRanks. We also compute the PageRank with uniform teleportation jumps. The calculation of the PageRank is done by applying the matrix power iteration method (see [24] for more details).

In Figure 3(a)–(d) we present cumulative log-log plots for the in-degree, the teleportation, and the PageRanks for damping factors $c = 0.5$ and $c = 0.85$. Here we consider a scale-free teleportation, so we plot the complementary cumulative distribution function $\mathrm{P}(wT > x) = (\bar{t}x)^{-\alpha_B}$. Then, $y = -\alpha_B x - \alpha_B * \log_{10}(\bar{t})$ is the straight line that corresponds to the teleportation log-log plot. We also fit the in-degree plot with the straight line $y = -1.1x + 0.08$.

First, we consider the log-log plots of the standard PageRank with uniform teleportation (see Figure 3(a)). In this case we use Theorem 4.1(i) for $A \overset{\mathrm{D}}{=} c/D$ to obtain the distance between the in-degree and PageRank log-log plots as

$$\log_{10}(C_N) = \log_{10}\left[\frac{c^{\alpha_N}(1-p_0)^{\alpha_N}}{(\mathrm{E}(N))^{\alpha_N}(1-c^{\alpha_N}\,\mathrm{E}(N)\,\mathrm{E}(1/D^{\alpha_N}))}\right], \tag{5.1}$$

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

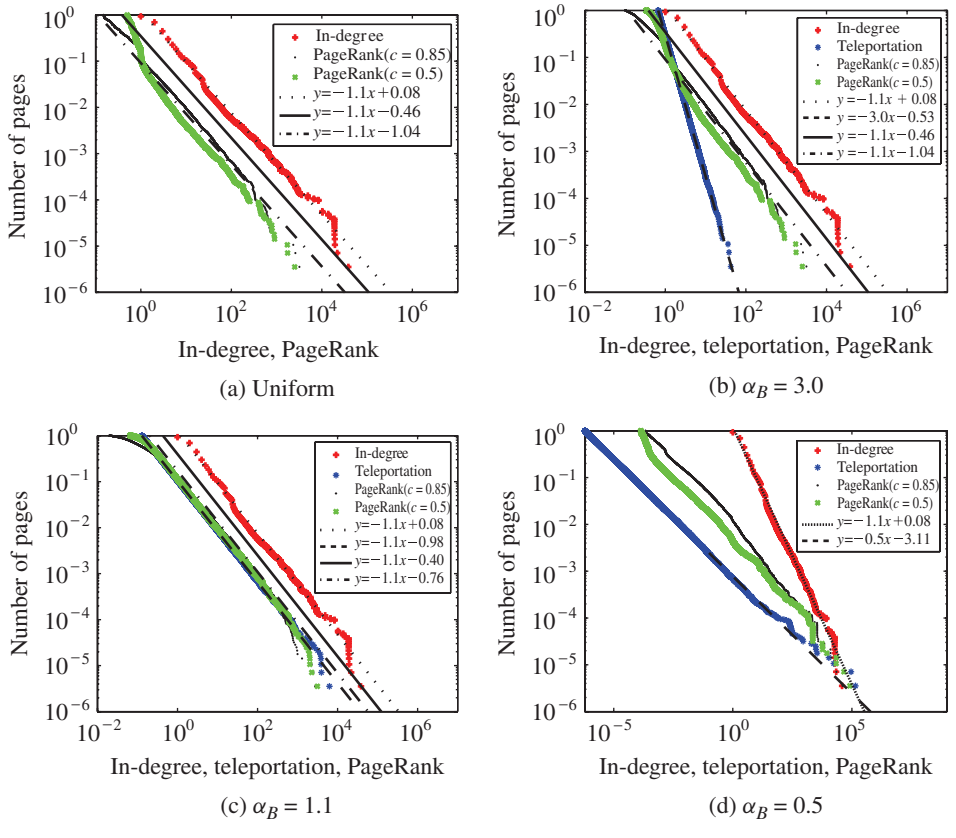


FIGURE 3: Number of pages with in-degree/teleportation/PageRank greater than the values on the $x$-axes, on a log-log scale.

where, as before, $N$ is the in-degree and $D$ is the effective out-degree. From $\mathrm{E}(N) = 8.2032$, $p_0 = 0.006$, and $\mathrm{E}(1/D^{1.1}) = 0.1043$, we predict the PageRank log-log plots: $y = -1.1x - 0.46$ for $c = 0.85$ and $y = -1.1x - 1.04$ for $c = 0.5$. In the plot we show these theoretically predicted lines and experimental PageRank log-log plots. We see that both lines perfectly match the slopes of the PageRanks, and they trace the direction of changes in the PageRank distribution with respect to changes in the damping factor. Indeed, the plot of the PageRank with $c = 0.5$ is further from the in-degree log-log plot than the plot of the PageRank with $c = 0.85$. We note that we underestimate the predicted distance in the case of $c = 0.85$, due to some assumptions of our model. We refer the reader to Section 6 for a discussion.

We again use the results of Theorem 4.1(i) for the case of the PageRank with teleportation that follows a power law with exponent $\alpha_B = 3.0$. Then we end up with the same constant as in (5.1), and, therefore, we obtain the same predicted lines for the PageRank log-log plots: $y = -1.1x - 0.46$ for $c = 0.85$ and $y = -1.1x - 1.04$ for $c = 0.5$. In Figure 3(b) we plot the distributions of the teleportation and the PageRanks along with the predicted straight lines. The results are similar to the previous case. Thus, we can see that the distribution of the teleportation has no influence on the tail behavior of the PageRank in the case when the teleportation has a lighter tail than the in-degree.

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

Next, we consider the $T(i)$s with $\alpha_B = 1.1$ and define $B(i) = cp_0 + (1-c)wT(i)$, where $i = 1, \dots, w$. Then, $\mathrm{P}(B > x) \sim (1-c)^{\alpha_B} \mathrm{P}(wT > x) \sim C_{NB} \mathrm{P}(N > x)$ as $x \to \infty$. Because $y = -1.1x + 0.08$ and $y = -1.1x - 0.98$ are the fitted lines for the log-log plots of the in-degree and teleportation, respectively, we find that $C_{NB} = 0.0108$ for $c = 0.85$, and $C_{NB} = 0.4063$ for $c = 0.5$. So, in the case when the in-degree and the teleportation are regularly varying with the same index $\alpha_N = \alpha_B = 1.1$, we can define the distance in the following way:

$$\log_{10}(C) = \log_{10}\left[\frac{(\mathrm{E}(N))^{\alpha_N} C_{NB} + c^{\alpha_N}(1-p_0)^{\alpha_N}}{(\mathrm{E}(N))^{\alpha_N}(1 - c^{\alpha_N}\,\mathrm{E}(N)\,\mathrm{E}(1/D^{\alpha_N}))}\right]. \tag{5.2}$$

We apply these constants in the above formula to obtain $y = -1.1x - 0.41$ and $y = -1.1x - 0.76$ for the PageRank plots for $c = 0.85$ and $c = 0.5$, respectively. We plot these lines in Figure 3(c). Compared to Figure 3(a) and (b), here the teleportation distribution smoothes the log-log plots of the PageRanks. Thus, we can hardly see the difference between the plots for $c = 0.5$ and $c = 0.85$. The slopes of the experimental PageRanks again correspond to the predicted power law exponent 1.1. The differences between the log-log plots of the in-degree and the PageRanks agree better than in the previous cases.

Finally, we present results for the teleportation with power law exponent $\alpha_B = 0.5$ in Figure 3(d). Note that we cannot find the distance in this case, because the first moment of $B$ does not exist. However, we can clearly see that the PageRank tends to follow a power law with the same exponent as the teleportation distribution.

Note that the constant in (5.1) is the same as the predicted constant from [37], where we assumed that the out-degree is random and the teleportation is uniform. Furthermore, from Jensen's inequality, $\mathrm{E}(1/D^{\alpha_N}) \geq (\mathrm{E}(1/D))^{\alpha_N} = [(1-p_0)/\mathrm{E}(N)]^{\alpha_N}$, it follows that

$$C_N \geq \frac{c^{\alpha_N}(1-p_0)^{\alpha_N}}{(\mathrm{E}(N))^{\alpha_N}[1 - c^{\alpha_N}(1-p_0)^{\alpha_N}(\mathrm{E}(N))^{1-\alpha_N}]}. \tag{5.3}$$

The last expression is the value of $C_N$ in the case when the out-degree of all nondangling nodes is a constant $\mathrm{E}(N)/(1-p_0)$, as in [26]. If $\alpha_N = 1.1$ then the difference between the left- and right-hand sides of (5.3) is small for any reasonable out-degree distribution. We can also ignore the term $c^{\alpha_N}(1-p_0)^{\alpha_N}(\mathrm{E}(N))^{1-\alpha_N}$ in (5.1). Then $C_N$ can be approximated from above as

$$C_N \geq \frac{c^{\alpha_N}(1-p_0)^{\alpha_N}}{(\mathrm{E}(N))^{\alpha_N}} = c^{\alpha_N}\left[\mathrm{E}\left(\frac{1}{D}\right)\right]^{\alpha_N} = C'_N.$$

Note that the asymptotic equivalence $\mathrm{P}(R > x) \sim C'_N \mathrm{P}(N > x)$ as $x \to \infty$ holds if we assume that the values of the PageRank $R$ can be approximated by $cN\,\mathrm{E}(1/D)$, as proposed in [14]. Furthermore, we can repeat a similar reasoning for (5.2) to obtain

$$C \geq \frac{(\mathrm{E}(N))^{\alpha_N} C_{NB} + c^{\alpha_N}(1-p_0)^{\alpha_N}}{(\mathrm{E}(N))^{\alpha_N}[1 - c^{\alpha_N}(1-p_0)^{\alpha_N}(\mathrm{E}(N))^{1-\alpha_N}]} \geq C_{NB} + c^{\alpha_N}\left[\mathrm{E}\left(\frac{1}{D}\right)\right]^{\alpha_N}.$$

## 6. Conclusions

This work has proposed a generalized stochastic model that characterizes the distribution of the personalized PageRank scores. Under various assumptions on the distribution of the Web parameters and teleportation, the model captures essential features of the PageRank tail behavior, and reveals which properties of the Web graph influence this behavior the most. In particular, the results show that the in-degree and, sometimes, the teleportation play an

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

important role, while the influence of the out-degree distribution is minimal. The results have been obtained by means of analyzing the asymptotic properties of the solution of a stochastic equation that is related to branching processes and, to the best of the authors' knowledge, has not been studied to this extent before.

Our results are in good agreement with the Web data. The differences between the model and the data depend on many factors, in particular, on the choice of data set, as observed in [37]. Furthermore, the assumption of the branching structure of the Web implicitly made in (2.3) is probably not justified. Future work could try to investigate how to improve the model in that respect, mainly by studying the dependencies amongst the $R_i$s in (2.3), or between the $R_i$s on the one hand and $N$ on the other hand.

## Appendix A. Regular variation preliminaries

The theory of regular variation is a natural mathematical formalism for analyzing power laws. In this section we provide the main definitions and some facts that have been used throughout this paper. For more details, we refer the reader to the classic book by Bingham *et al.* [6], and to the recent review by Jessen and Mikosch [20].

The next lemma describes the asymptotic behavior of the product, sum, and random sums of regularly varying random variables. We use these results for defining the asymptotic properties of the PageRank, when the PageRank is a result of a finite number of the iteration steps (see Section 3). In the lemma, relation (iii) is known as Breiman's theorem (see, e.g. Lemma 4.2(1) of [20]). Properties (iv), (v), and (vi) are statements (2), (1), and (5) of Lemma 3.7 of [20], respectively. The similarity for sums (i) and (ii) follows from Lemmas 3.12 and 3.1 of [20], respectively.

**Lemma A.1.** (i) *Assume that $X_1$ is a nonnegative, regularly varying random variable with index $\alpha \geq 0$. If the random variable $X_2 > 0$ is such that $\mathrm{P}(X_2 > x) = o(\mathrm{P}(X_1 > x))$ then*

$$\mathrm{P}(X_1 + X_2 > x) \sim \mathrm{P}(X_1 > x) \quad \text{as } x \to \infty.$$

(ii) *Assume that $X_1$ is a nonnegative, regularly varying random variable with index $\alpha \geq 0$. If the random variable $X_2 > 0$ satisfies $\mathrm{P}(X_2 > x) \sim C\,\mathrm{P}(X_1 > x)$ for some $C > 0$, and $\mathrm{P}(X_1 > x,\ X_2 > x) = o(\mathrm{P}(X_1 > x))$, then*

$$\mathrm{P}(X_1 + X_2 > x) \sim (1 + C)\,\mathrm{P}(X_1 > x) \quad \text{as } x \to \infty.$$

(iii) *Assume that $X_1$ and $X_2$ are two independent, nonnegative random variables such that $X_1$ is regularly varying with index $\alpha$ and $\mathrm{E}(X_2^{\alpha+\epsilon}) < \infty$ for some $\epsilon > 0$. Then*

$$\mathrm{P}(X_1 X_2 > x) \sim \mathrm{E}(X_2^{\alpha})\,\mathrm{P}(X_1 > x) \quad \text{as } x \to \infty.$$

(iv) *Assume that $N$ is regularly varying with index $\alpha \geq 0$; if $\alpha = 1$ then assume that $\mathrm{E}(N) < \infty$. Moreover, let $(X_i)$ be an independent and identically distributed (i.i.d.) sequence such that $\mathrm{E}(X_1) < \infty$ and $\mathrm{P}(X_1 > x) = o(\mathrm{P}(N > x))$. Then, as $x \to \infty$,*

$$\mathrm{P}\left(\sum_{i=1}^{N} X_i > x\right) \sim (\mathrm{E}(X_1))^{\alpha}\,\mathrm{P}(N > x) \quad \text{as } x \to \infty.$$

https://doi.org/10.1239/aap/1275055243 Published online by Cambridge University Press

(v) *Assume* $(X_i)$ *is an i.i.d. sequence of regular varying random variables with index* $\alpha > 0$, $\mathrm{E}(N) < \infty$, *and* $\mathrm{P}(N > x) = o(\mathrm{P}(X_1 > x))$. *Then*

$$\mathrm{P}\left(\sum_{i=1}^{N} X_i > x\right) \sim \mathrm{E}(N)\,\mathrm{P}(X_1 > x) \quad \text{as } x \to \infty.$$

(vi) *Assume that* $\mathrm{P}(X_1 > x) \sim C\,\mathrm{P}(N > x)$ *for some* $C > 0$, *that* $X_1$ *is regularly varying with index* $\alpha \geq 1$, *and that* $\mathrm{E}(X_1) < \infty$. *Then*

$$\mathrm{P}\left(\sum_{i=1}^{N} X_i > x\right) \sim (C\,\mathrm{E}(N) + (\mathrm{E}(X_1))^{\alpha})\,\mathrm{P}(N > x) \quad \text{as } x \to \infty.$$

In this paper we presented the PageRank as the solution of a stochastic equation. In order to define its asymptotics, we need to use the *Laplace–Stieltjes transform* analysis (see Section 4). We denote by $f(s) = \mathrm{E}\,\mathrm{e}^{-sX}$, $s > 0$, the Laplace–Stieltjes transform of $X$, and let $\xi_i = \int_0^\infty x^i\,\mathrm{d}F_X(x)$ be the $i$th moment of $X$, where $F_X$ is the distribution function of $X$. The successive moments of $X$ can be obtained by expanding $f(s)$ in a series at $s = 0$. More precisely, we write the following.

**Lemma A.2.** *The nth moment of* $X$ *is finite if and only if there exist finite numbers* $\xi_0 = 1$ *and* $\xi_1, \ldots, \xi_n$ *such that*

$$f_n(s) = (-1)^{n+1}\left(f(s) - \sum_{i=0}^{n} \frac{\xi_i}{i!}(-s)^i\right) = o(s^n) \quad \text{as } s \to 0.$$

*In this case,* $\xi_i$ *is the* $i$*th moment of* $X$.

The following theorem establishes the relation between the asymptotic behavior of a regularly varying distribution and its Laplace–Stieltjes transform. We use this result in the proof of Theorem 4.1.

**Theorem A.1.** (Tauberian theorem.) *If* $n \in \mathbb{N}$, $\xi_n < \infty$, *and* $\alpha \in (n, n+1)$, *then the following statements are equivalent:*

(i) $f_n(s) \sim (-1)^n \Gamma(1-\alpha) s^{\alpha} L(1/s)$ *as* $s \to 0$,

(ii) $\mathrm{P}(X > x) \sim x^{-\alpha} L(x)$ *as* $x \to \infty$.

The next lemma provides a useful bound for slowly varying functions.

**Lemma A.3.** (Potter bounds.) *Let* $L$ *be a slowly varying function. Then, for any fixed* $\vartheta > 1$ *and* $\delta > 0$, *there exists a finite constant* $s_0 < 1$ *such that, for all* $s_1, s_2 < s_0$,

$$\frac{L(1/s_1)}{L(1/s_2)} \leq \vartheta \max\left\{\left(\frac{s_1}{s_2}\right)^{\delta}, \left(\frac{s_1}{s_2}\right)^{-\delta}\right\}.$$

**Acknowledgements**

We would like to thank Bert Zwart for useful discussions. This work was supported by NWO Meervoud, under grant number 632.002.401. Part of this research was funded by the Dutch BSIK/BRICKS project. This paper was also the result of joint research with the 3TU Centre of Competence NIRICT (Netherlands Institute for Research on ICT) within the Federation of Three Universities of Technology in The Netherlands.

## References


[1] Andersen, R., Chung, F. and Lang, K. (2006). Local graph partitioning using PageRank vectors. In *Proc. FOCS 2006*, IEEE Computer Society, Washington, DC, pp. 475–486.

[2] Andersen, R., Chung, F. and Lang, K. (2007). Local partitioning for directed graphs using PageRank. In *Algorithms and Models for the Web-Graph* (Lecture Notes Comput. Sci. **4863**), Springer, Berlin, pp. 166–178.

[3] Avrachenkov, K. and Lebedev, D. (2006). PageRank of scale-free growing networks. *Internet Math.* **3,** 207–231.

[4] Avrachenkov, K., Litvak, N. and Pham, K. S. (2007). Distribution of PageRank mass among principle components of the web. In *Algorithms and Models for the Web-Graph* (Lecture Notes Comput. Sci. **4863**), Springer, Berlin, pp. 16–28.

[5] Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science* **286,** 509–512.

[6] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1989). *Regular Variation*. Cambridge University Press.

[7] Boldi, P., Santini, M. and Vigna, S. (2005). PageRank as a function of the damping factor. In *Proc. 14th Internat. Conf. World Wide Web*, ACM, New York, pp. 557–566.

[8] Brin, S. and Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. *Comput. Networks ISDN Systems* **30,** 107–117.

[9] Broder, A. *et al.* (2000). Graph structure in the Web. *Comput. Networks* **33,** 309–320.

[10] Chen, P., Xie, H., Maslov, S. and Redner, S. (2007). Finding scientific gems with Google's PageRank algorithm. *J. Informetrics* **1,** 8–15.

[11] Constantine, P. G. and Gleich, D. F. (2007). Using polynomial chaos to compute the influence of multiple random surfers in the PageRank model. In *Algorithms and Models for the Web-Graph* (Lecture Notes Comput. Sci. **4863**), Springer, Berlin, pp. 82–95.

[12] De Meyer, A. and Teugels, J. L. (1980). On the asymptotic behaviour of the distributions of the busy period and service time in M/G/1. *J. Appl. Prob.* **17,** 802–813.

[13] Fortunato, S. and Flammini, A. (2007). Random walks on directed networks: the case of PageRank. *Internat. J. Bifur. Chaos Appl. Sci. Eng.* **17,** 2343–2353.

[14] Fortunato, S., Boguñá, M., Flammini, A. and Menczer, F. (2006). Approximating PageRank from in-degree. In *Algorithms and Models for the Web-Graph* (Lecture Notes Comput. Sci. **4936**), Springer, Berlin, pp. 59–71.

[15] Gyöngyi, Z., Garcia-Molina, H. and Pedersen, J. (2004). Combating Web spam with TrustRank. In *Proc. 13th Internat. Conf. Very Large Data Bases*, VLDB Endowment, pp. 576–587.

[16] Haveliwala, T. H. (2003). Topic-sensitive PageRank: a context-sensitive ranking algorithm for Web search. *IEEE Trans. Knowledge Data Eng.* **15,** 784–796.

[17] Haveliwala, T. H., Kamvar, A. and Jeh, G. (2003). An analytical comparison of approaches to personalizing PageRank. Tech. Rep., Stanford University.

[18] Jeh, G. and Widom, J. (2003). Scaling personalized Web search. In *Proc. 12th Internat. Conf. World Wide Web*, ACM, New York, pp. 271–279.

[19] Jelenkovic, P. R. and Olvera-Cravioto, M. (2009). Information ranking and power laws on trees. Preprint. Available at http://arxiv.org/abs/0905.1738.

[20] Jessen, A. H. and Mikosch, T. (2006). Regularly varying functions. *Publ. Inst. Math. (Beograd) (N.S.)* **80,** 171–192.

[21] Kamvar, S. D., Haveliwala, T. H., Manning, C. D. and Golub, G. H. (2003). Exploiting the block structure of the Web for computing. Tech. Rep., Stanford University.

[22] Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *J. ACM* **46,** 604–632.

[23] Kraaij, W., Westerveld, T. and Hiemstra, D. (2002). The importance of prior probabilities for entry page search. In *Proc. 25th Annual Internat. ACM SIGIR Conf. Research and Development in Information Retrieval*, ACM, New York, pp. 27–34.

[24] Langville, A. N. and Meyer, C. D. (2006). *Google's PageRank and Beyond: the Science of Search Engine Rankings*. Princeton University Press.

[25] Leskovec, J. and Faloutsos, C. (2006). Sampling from large graphs. In *Proc. 12th ACM SIGKDD Internat. Conf. Knowledge Discovery and Data Mining*, ACM, New York, pp. 631–636.

[26] Litvak, N., Scheinhardt, W. R. W. and Volkovich, Y. (2007). In-degree and PageRank: why do they follow similar power laws? *Internet Math.* **4,** 175–198.

[27] Litvak, N., Scheinhardt, W., Volkovich, Y. and Zwart, B. (2009). Characterization of tail dependence for in-degree and PageRank. In *Algorithms and Models for the Web-Graph* (Lecture Notes Comput. Sci. **5427**), Springer, Berlin, pp. 90–103.

[28] Liu, Q. (1998). Fixed points of a generalized smoothing transformation and applications to the branching random walk. *Adv. Appl. Prob.* **30,** 85–112.

[29] Liu, Q. (2001). Asymptotic properties and absolute continuity of laws stable by random weighted mean. *Stoch. Process. Appl.* **95,** 83–107.

[30] Micarelli, A., Gasparetti, F., Sciarrone, F. and Gauch, S. (2007). Personalized search on the World Wide Web. In *The Adaptive Web* (Lecture Notes Comput. Sci. **4321**), Springer, Berlin, pp. 195–230.
[31] Page, L., Brin, S., Motwani, R. and Winograd, T. (1998). The PageRank citation ranking: bringing order to the Web. Tech. Rep., Stanford Digital Library Technologies Project.
[32] Pandurangan, G., Raghavan, P. and Upfal, E. (2002). Using PageRank to characterize Web structure. In *Computing and Combinatorics* (Lecture Notes Comput. Sci. **2387**), Springer, Berlin, pp. 330–339.
[33] Ponte, J. M. and Croft, W. B. (1998). A language modeling approach to information retrieval. In *Proc. 21st Annual Internat. ACM SIGIR Conf. Research and Development in Information Retrieval*. ACM, New York, pp. 275–281.
[34] Resnick, S. I. (2007). *Heavy-tail Phenomena*. Springer, New York.
[35] Richardson, M. and Domingos, P. (2002). The intelligent surfer: probabilistic combination of link and content information in PageRank. *Adv. NIPS* **14,** 1441–1448.
[36] Ross, S. M. (2003). The inspection paradox. *Prob. Eng. Inform. Sci.* **17,** 47–51.
[37] Volkovich, Y., Litvak, N. and Donato, D. (2007). Determining factors behind the PageRank log-log plot. In *Algorithms and Models for the Web-Graph* (Lecture Notes Comput. Sci. **4863**), Springer, Berlin, pp. 108–123.
[38] Volkovich, Y., Litvak, N. and Zwart, B. (2008). A framework for evaluating statistical dependencies and rank correlations in power law graphs. Memorandum 1868. University of Twente Enschede.
[39] Volkovich, Y., Litvak, N. and Zwart, B. (2009). Extremal dependencies and rank correlations in power law networks. In *Complex Sciences*, Springer, Berlin, pp. 1642–1653.
[40] Zwart, A. P. (2001). Queueing systems with heavy tails. Doctoral Thesis, Eindhoven University of Technology.