Applied Probability Trust (27 October 2009)

INFORMATION RANKING AND POWER LAWS ON TREES

PREDRAG R. JELENKOVIĆ,[*] *Columbia University*

MARIANA OLVERA-CRAVIOTO,[**] *Columbia University*

Abstract

We consider the stochastic analysis of information ranking algorithms of large interconnected data sets, e.g. Google's PageRank algorithm for ranking pages on the World Wide Web. The stochastic formulation of the problem results in a recursion of the form

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

where $N, Q, \{R_i\}_{i\geq 1}, \{C, C_i\}_{i\geq 1}$ are independent non-negative random variables, $\{C, C_i\}_{i\geq 1}$ are identically distributed, and $\{R_i\}_{i\geq 1}$ are independent copies of R; $\overset{\mathcal{D}}{=}$ stands for equality in distribution. We study the asymptotic properties of the distribution of R that, in the context of PageRank, represents the frequencies of highly ranked pages. The preceding recursion is interesting in its own right since it belongs to a more general class of weighted branching processes that have been found useful in the analysis of many other algorithms. Our first main result shows that if $ENE[C^\alpha] = 1, \alpha > 0$ and Q, N have higher moments than α, then R has a power law distribution of index α. This result is obtained using a new approach based on an extension of Goldie's (1991) implicit renewal theorem. Furthermore, when N is regularly varying of index $\alpha > 1$, $ENE[C^\alpha] < 1$ and Q, C have higher moments than α, then the distributions of R and N are tail equivalent. The latter result is derived via a novel sample path large deviation method for recursive random sums. Similarly, we characterize the situation when the distribution of R is determined by the tail of Q. The preceding approaches may be of independent interest, as they can be used for analyzing other recursions on trees. We also discuss the engineering implications of our results throughout the paper.

Keywords: Information ranking; stochastic recursions; weighted branching processes; power laws; regular variation; implicit renewal theory; large deviations

2000 Mathematics Subject Classification: Primary 60H25
Secondary 60J80,60F10,60K05

1. Introduction

We consider a problem of ranking large interconnected information (data) sets, e.g., ranking pages on the World Wide Web (Web). A solution to the preceding problem is given by Google's PageRank algorithm, the details of which are presented in Section 1.1. Given the large scale of these information sets, we adopt a stochastic approach to the page ranking problem, e.g. Google's PageRank algorithm. The stochastic

---

[*] Postal address: Department of Electrical Engineering, Columbia University, New York, NY 10027

[**] Postal address: Department of Industrial Engineering and Operations Research, Columbia University, New York, NY 10027

formulation naturally results in a recursion of the form

$$R \stackrel{\mathcal{D}}{=} Q + \sum_{i=1}^{N} C_i R_i, \tag{1.1}$$

where $N, Q, \{R_i\}_{i\geq 1}, \{C, C_i\}_{i\geq 1}$ are independent non-negative random variables, $\{C, C_i\}_{i\geq 1}$ are identically distributed, and $\{R_i\}_{i\geq 1}$ are independent copies of R; $\stackrel{\mathcal{D}}{=}$ stands for equality in distribution. We study the asymptotic properties of the distribution of R that, in the context of PageRank, represents the frequencies of highly ranked pages. In somewhat smaller generality, the preceding stochastic setup was first introduced and analyzed in [33] for the PageRank algorithm; the formulation given in (1.1) was later studied in [32].

The canonical representation given by recursion (1.1) is also of independent interest since it belongs to a more general class of weighted branching processes (WBPs) [28, 24, 22]; the connection to WBPs is discussed in more detail in Section 2.3. With a small abuse of notation, we also refer to our more restrictive processes as WBPs. These processes have been found useful in the average-case analysis of many algorithms [29], e.g. quicksort algorithm [14], and thus, our study of recursion (1.1), may be useful in these types of applications. Furthermore, when $Q = 1, C_i \equiv 1$, the steady state solution to (1.1) represents the total number of individuals born in an ordinary branching process. Also, by letting N be a Poisson random variable and fixing $Q = 1, C_i \equiv 1$, equation (1.1) reduces to the recursion that is satisfied by the busy period of an M/G/1 queue. Similarly, selecting $N = 1$ yields the recursion of the first order autoregressive process; see Section 2.3 for a more thorough discussion on related processes.

In Section 2 we connect the iterations of recursion (1.1) to an explicit construction of a WBP on a tree, such that the sum of all the weights of the first n generations of the tree are directly related to the nth iteration of the recursion. Then, in Section 3 we present explicit estimates for the moments of the total weight, W_n, of the nth generation in the corresponding WBP. Using these moment estimates and the WBP representation, we show in Section 3.1 that under mild conditions the iterations of (1.1) converge in distribution to a unique and finite steady state random variable R. Hence, under the stated assumptions, this limiting distribution $P(R \leq x)$ is the unique solution to (1.1). The steady state variable R represents the sum of all the weights in the corresponding branching tree.

Studying the asymptotic tail properties of the steady state solution R to (1.1) represents the main focus of this paper. In particular, we study the possible causes that can result in power tail asymptotics for $P(R > x)$. We discover that the tail behavior of R can be determined/dominated by the statistical properties of any of the three variables C, N and Q. The corresponding results are presented in Sections 4, 5 and 6, respectively. Our emphasis on power law asymptotics is motivated by the well established empirical fact that the number of pages that point to a specific page (in-degree) on the Web, represented by N in recursion (1.1), follows a power law distribution; other complex data sets, e.g. citations, are found to posses similar power law properties as well.

Our first main result on the tail behavior of $P(R > x)$ is presented in Theorem 4.2, showing that if $ENE[C^\alpha] = 1, \alpha > 0$ and Q, N have higher moments than α, then R has a power law distribution of index α, with an explicitly characterized constant of proportionality. In particular, when α is an integer, the constant of proportionality of the power law distribution is explicitly computable, see Corollary 4.1. This result is obtained by an extension of Goldie's (1991) implicit renewal theorem that we present in Theorem 4.1. This extension may be of independent interest since R and C in the statement of Theorem 4.1 can be any two independent random variables that may satisfy a different recursion. In the context of the broader literature on WBPs, our results are related to the studies in [28] (see Theorem 6), and more recently in [24], that use transform methods to characterize the distribution of R, under appropriate conditions, as stable distributions. In the case of positive R that we consider, these results are restricted to positive stable laws that allow power law tails only for $0 < \alpha \leq 1$ (see Section 5.3.5 in [7]). Also, related results for the homogeneous case ($Q \equiv 0$) and $\alpha > 1$ can be found in Theorem 2.2 of [25] and Proposition 7 of [16]. Interestingly, our approach for the non-homogeneous case ($P(Q > 0) > 0$) shows that the distribution of R can have a uniform treatment for any $\alpha > 0$. For additional comments on results related to our Theorem 4.2

see the remarks following its statement. Furthermore, this result may provide a new explanation of why power laws are so commonly found in the distribution of wealth since weighted branching processes appear to be reasonable models for the total wealth of a family tree.

Section 5 studies the case when N is power law and dominates the tail behavior of R. This is the case that more closely relates to the original formulation of PageRank and the structure of the Web graph since the in-degree N is well accepted to be a power law. Our main result in this case, stated in Theorem 5.1, shows that, when N is regularly varying of index $\alpha > 1$, $ENE[C^\alpha] < 1$ and Q, C have higher moments than α, then the distribution of R is tail equivalent to that of N. Our approach in deriving this result is based on a new sample path heavy-tailed large deviation method for weighted recursions on trees. The key technical result is given by Proposition 5.1 that provides a uniform bound (in n and x) on the distribution of the total weight of the nth generation $P(W_n > x)$. We would also like to point out that Proposition 5.1 resembles to some extent a classical result by Kesten (see Lemma 7 on p. 149 of [4]), which provides a uniform bound for the sum of heavy-tailed (subexponential) random variables. The main difference between the latter result and our uniform bound is that n refers to the depth of the recursion in our case, while in Lemma 7 of [4], n is the number of terms in the sum. This makes the derivation of Proposition 5.1 considerably more complicated, and perhaps implausible, if it were not for the fact that we restrict our attention to regularly varying distributions, as opposed to the general subexponential class.

Section 6 investigates a third possible source of heavy tails for R, the one that arises from the innovation, Q, being power law, see Theorem 6.1. For $N = 1$, this result is consistent with a corresponding result for the first order autoregressive process in Lemma A.3 of [26]. The proofs of more technical results are postponed to Section 7.

Finally, from a mathematical perspective, we would like to emphasize that our sample path large deviation approach as well as the extension of the implicit renewal theory, provide a new set of tools that can be of potential use in other applications, as well as in studying the broader class of recursions on trees, e.g., one can readily characterize the asymptotic behavior of the distribution that solves $R = Q + \max_{1 \le i \le N} C_i R_i$. Furthermore, from an engineering perspective, our Theorem 5.1 shows that for highly ranked pages, the PageRank algorithm basically reflects the popularity vote given by the number of references N, implying that overly inflated referencing may be advantageous. A more detailed discussion on the engineering implications of the performance and design of ranking algorithms, e.g. PageRank, can be found at the end of Section 5.

1.1. Google's algorithm: PageRank

PageRank is an algorithm trademarked by Google, the Internet search engine, to assign to each page a numerical weight that measures its relative importance with respect to other pages. We think of the Web as a very large interconnected graph where nodes correspond to pages. The Google trademarked algorithm PageRank defines the page rank as:

$$R(p_i) = \frac{1-d}{n} + d \sum_{p_j \in M(p_i)} \frac{R(p_j)}{L(p_j)}, \tag{1.2}$$

where, using Google's notation, $p_1, p_2, \ldots, p_n$ are the pages under consideration, $M(p_i)$ is the set of pages that link to p_i, $L(p_j)$ is the number of outbound links on page p_j, n is the total number of pages on the Web, and d is a damping factor, usually $d = 0.85$. As noted in the original paper by Brin and Page (1998) [10] PageRank "can be calculated using a simple iterative algorithm, and corresponds to the principal eigenvector of the normalized link matrix of the Web. Also, a PageRank for 26 million web pages can be computed in a few hours on a medium size workstation." PageRank is based on citation analysis that was developed in the 1950s by Eugene Garfield at the University of Pennsylvania. Web link analysis was first developed by Kleinberg [21].

While in principle the solution to (1.2) reduces to the solution of a large system of linear equations, due to the ever increasing size of the Web and similar highly interconnected data sets, we believe that the "brute force" deterministic approach might be impractical, as well as non insightful. Specifically, if one obtains the principal eigenvector of the normalized link matrix, it is hard to obtain from the solution qualitative insights about the relationship between highly ranked pages and the in-degree/out-degree statistical properties of the graph.

In particular, the division by the out-degree ($L(p_j)$ in equation (1.2)) was meant to decrease the contribution of pages with highly inflated referencing, i.e., those pages that basically point/reference possibly indiscriminately to other documents. However, the stochastic approach (to be described in the following sections) reveals that highly ranked pages are essentially insensitive to the parameters of the out-degree distribution, implying that the PageRank algorithm may not reduce the effects of overly inflated referencing (citations, voting) as originally intended, i.e., it may lead to possibly unjustifiable highly ranked pages. The observation that the tail of the rank distribution is dominated by N was also made in [33] and [32]. More discussions on this topic are provided at the end of Section 5.

A stochastic approach to analyze (1.2) is to consider the recursion

$$R \stackrel{\mathcal{D}}{=} \gamma + c \sum_{i=1}^{N} \frac{R_i}{D_i}, \tag{1.3}$$

where $\gamma, c > 0$ are constants, $cE[1/D] < 1$, N is a random variable independent of the R_i's and D_i's, the D_i's are iid random variables satisfying $D_i \geq 1$, and the R_i's are iid random variables having the same distribution as R. In terms of recursion (1.2), R is the rank of a random page, N corresponds to the in-degree of that node, the R_i's are the ranks of the pages pointing to it, and the D_i's correspond to the out-degrees of each of these pages. The assumption that the in-degree of a page is independent of the ranks and out-degrees of the pages pointing to it is justified by the massive size and sparse nature of the underlying graph. Furthermore, the experimental justification of these independence assumptions can be found in [31]. This setup was also considered in [33], where a partial analysis of recursion (1.3) was carried out. Recently, the following more general recursion for the cases when N or Q dominate was analyzed via Tauberian theorems, under some more restrictive assumptions, in [32]. In this paper we develop a novel sample path approach for the analysis of

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

as defined previously in (1.1). Recall that $N, Q, \{R_i\}_{i\geq 1}, \{C, C_i\}_{i\geq 1}$ are independent non-negative random variables, $\{C, C_i\}_{i\geq 1}$ are identically distributed, and $\{R_i\}_{i\geq 1}$ are independent copies of R.

2. Model Description

As outlined above, we study the sequence of random variables that are obtained by iterating (1.1). Specifically, we consider

$$R^*_{n+1} = Q_{n+1} + \sum_{i=1}^{N_n} C_i^{(n+1)} R^*_{n,i}, \tag{2.1}$$

where $\{R^*_{n,i}\}_{i\geq 1}$ are iid copies of R^*_n from the previous iteration, and $\{N_n\}, \{C_i^{(n+1)}\}, \{Q_{n+1}\}$ are mutually independent iid sequences of random variables; for $n = 0$, $R^*_{0,i}$ are iid copies of the initial value R^*_0.

In this section we will discuss the weak convergence of R^*_n to a finite random variable R, independently of the initial condition R^*_0. In other words, R is the unique solution to (1.1). In particular, we will construct a process R_n on a tree that converges a.s. to R. These convergence results may be of practical interest as well since ranking algorithms are implemented recursively. The actual proofs are postponed until Section 3.1.

2.1. Construction of R on a Tree

To better understand the dynamics of our recursion, we give below a sample path construction of the random variable R on a tree. Consider the branching process $\{Z_n\}_{n\geq 0}$ given by recursion

$$Z_n = \sum_{i=1}^{Z_{n-1}} N^{(n-1)}_{i_1,\dots,i_{n-1}}, \qquad Z_0 = 1,$$

where $\{N^{(n)}_{i_1,\dots,i_n}\}_{n\geq 0}$ is a sequence of iid random variables having the same distribution as N. Here, $N^{(n)}_{i_1,\dots,i_n}$ is the number of offspring that individual $(i_1,\dots,i_n)$ from the nth generation has.

Suppose now that individual $(i_1,\dots,i_n)$ in the tree has a weight $\mathbf{C}^{(n)}_{i_1,\dots,i_n}$ defined via the recursion

$$\mathbf{C}^{(0)} = 1, \qquad \mathbf{C}^{(n)}_{i_1,\dots,i_n} = C^{(n)}_{i_1,\dots,i_n}\mathbf{C}^{(n-1)}_{i_1,\dots,i_{n-1}}, \quad n \geq 1,$$

where the random variables $\{C^{(n)}_{i_1,\dots,i_n} : n \geq 0, i_k \geq 1\}$ are iid with the same distribution as C. Note that $\mathbf{C}^{(n)}_{i_1,\dots,i_n}$ is equal to the product of all the weights $C^{(\cdot)}_{\cdot}$ along the branch leading to node $(i_1,\dots,i_n)$, as depicted on the figure below. Define now the process

$$W_n = \sum_{(i_1,\dots,i_n)\in A_n} Q^{(n)}_{i_1,\dots,i_n}\mathbf{C}^{(n)}_{i_1,\dots,i_n}, \qquad n \geq 0,$$

where A_n is the set of all individuals in the nth generation and $\{Q^{(n)}_{i_1,\dots,i_n}\}_{n\geq 0}$ is a sequence of iid random variables having the same distribution as Q (see Figure 1).

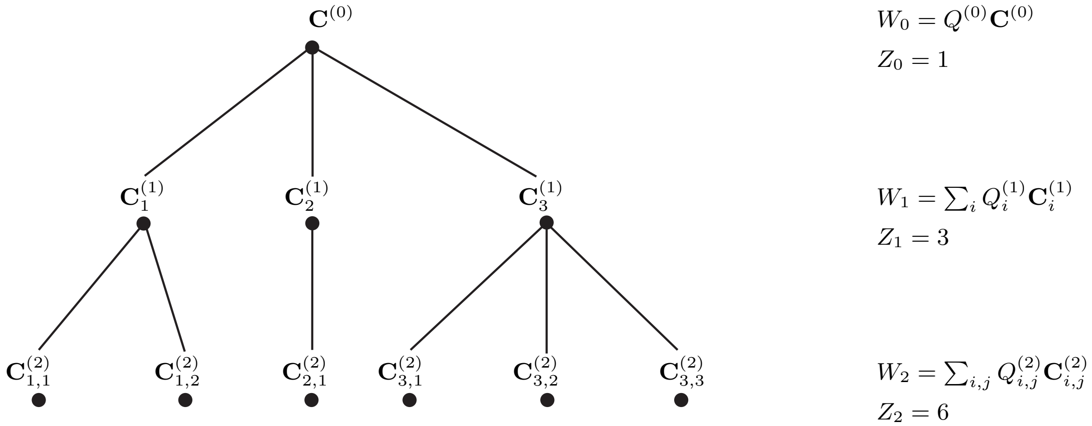

Figure 1: Construction on a tree

Observe that when $C^{(\cdot)}_{\cdot} \equiv 1$ and $Q^{(\cdot)}_{\cdot} \equiv 1$, W_n is equal to the number of individuals in the nth generation of the corresponding branching process, and in particular $Z_n = W_n$. Otherwise, W_n represents the sum of the weights of all the individuals in the nth generation. Related processes known as weighted branching processes have been considered in the existing literature [28, 24, 22] and are discussed in more detail in Section 2.3. With a small abuse of notation we also refer to our more restrictive processes as WBPs.

Define the process $\{R_n\}_{n\geq 1}$ according to

$$R_n = \sum_{k=0}^{n} W_k, \qquad n \geq 0,$$

that is, R_n is the sum of the weights of all the individuals on the tree. Clearly, when $Q_\cdot \equiv 1$ and $C_\cdot^{(\cdot)} \equiv 1$, R_n is simply the number of individuals in a branching process up to the nth generation. We define the random variable R according to

$$R \triangleq \lim_{n\to\infty} R_n = \sum_{k=0}^{\infty} W_k. \tag{2.2}$$

Furthermore, it is not hard to see that R_n satisfies the recursion

$$R_n = \sum_{j=1}^{N^{(0)}} C_j^{(1)} R_j^{(n-1)} + Q^{(0)}, \qquad R^{(0)}_{(i_1,\dots,i_{n+1})} = 0,$$

where $\{R_j^{(n-1)}\}$ are independent copies of R_{n-1} corresponding to the tree starting with individual j in the first generation. Therefore, R_n satisfies the recursion

$$R_n \stackrel{\mathcal{D}}{=} \sum_{i=1}^{N_{n-1}} C_i^{(n)} R_i^{(n-1)} + Q_n, \qquad n \geq 1. \tag{2.3}$$

Moreover, since the tree structure repeats itself after the first generation, W_n satisfies

$$\begin{aligned} W_n &= \sum_{(i_1,\dots,i_n)\in A_n} Q^{(n)}_{i_1,\dots,i_n} \mathbf{C}^{(n)}_{i_1,\dots,i_n} \\ &= \sum_{k=1}^{N^{(0)}} C_k^{(1)} \sum_{(k,\dots,i_n)\in A_n} Q^{(n)}_{k,\dots,i_n} \mathbf{C}^{(n)}_{k,\dots,i_n} / C_k^{(1)} \\ &\stackrel{\mathcal{D}}{=} \sum_{k=1}^{N} C_k W_{(n-1),k}, \end{aligned}$$

where N, C_k, $W_{(n-1),k}$ are independent of each other and of all other random variables, and $W_{(n-1),k}$ has the same distribution of W_{n-1}.

2.2. Connection between R_n^* and R_n

We now connect the two processes R_n^* and R_n, the one obtained by iterating (1.1) and the one obtained from the tree construction, respectively. To do this define

$$W_n(R_0^*) = \sum_{(i_1,\dots,i_n)\in A_n} R^*_{0,(i_1,\dots,i_n)} \mathbf{C}^{(n)}_{i_1,\dots,i_n},$$

where $R^*_{0,(\cdot)}$ are iid copies of the initial condition R_0^*, and the weights $\mathbf{C}_\cdot^{(n)}$ are the ones defined in Section 2.1. In words, $W_n(R_0)$ is the sum of all the weights in the nth generation of the tree with the coefficients $Q_\cdot^{(n)}$ substituted by the corresponding $R^*_{0,(\cdot)}$. We claim that

$$R_n^* \stackrel{\mathcal{D}}{=} R_{n-1} + W_n(R_0^*).$$

To see this note that for $n = 1$,

$$R_1^* = Q_1 + \sum_{i=1}^{N_0} C_i^{(1)} R^*_{0,i} \stackrel{\mathcal{D}}{=} Q^{(0)}\mathbf{C}^{(0)} + \sum_{i=1}^{N^{(0)}} \mathbf{C}_i^{(1)} R^*_{0,i} = W_0 + W_1(R_0^*) \qquad (\text{recall } \mathbf{C}^{(0)} = 1),$$

and by induction in n,

$$\begin{aligned} R^*_{n+1} &= Q_{n+1} + \sum_{i=1}^{N_n} C_i^{(n+1)} R^*_{n,i} \\ &\overset{\mathcal{D}}{=} Q_{n+1} + \sum_{i=1}^{N_n} C_i^{(n+1)} (R_{(n-1),i} + W_{n,i}(R^*_0)) \qquad \text{(by induction)} \\ &\overset{\mathcal{D}}{=} Q^{(0)} + \sum_{i=1}^{N^{(0)}} C_i^{(1)} \left(R_i^{(n-1)} + \sum_{(i,i_1,\ldots,i_n)\in A_{n+1}} R^*_{0,(i,i_1,\ldots,i_n)} \mathbf{C}^{(n+1)}_{i,i_1,\ldots,i_n} / C_i^{(1)} \right) \\ &= R_n + \sum_{i=1}^{N^{(0)}} R^*_{0,(i,i_1,\ldots,i_n)} \mathbf{C}^{(n+1)}_{i,i_1,\ldots,i_n} \\ &= R_n + W_{n+1}(R^*_0), \end{aligned}$$

where $R_i^{(n-1)}$ corresponds to the process R_{n-1} obtained from the tree starting with individual i in the first generation (a descendent of the root). Since $R_{n-1} \to R$ a.s., it will follow from Slutsky's Theorem (see Theorem 1, p. 254 in [11]) that if $W_n(R^*_0) \Rightarrow 0$, then

$$R^*_n \Rightarrow R,$$

where $\Rightarrow$ denotes convergence in distribution. The proof of this convergence and that of the finiteness of R are given in Section 3.1. Understanding the asymptotic properties of the distribution of R, as defined by (2.2), is the main objective of this paper.

2.3. Related Processes

As we mentioned above, the stochastic recursion defined in (1.1) leads to the analysis of a process known in the literature as a weighted branching process (WBP). WBPs were introduced by Rösler [28] in a construction that is more general than ours. More precisely, each individual in the tree has potentially an infinite number of offsprings, and each offspring inherits a certain (nonnegative) weight from its parent and multiplies it by a factor T_i, where the index i refers to his birth order (i.e., a first born multiplies his inheritance by T_1, a second born by T_2, etc.). Each individual branches independently, using an independent copy of the sequence $T_1, T_2, \ldots$. However, within the sequence, $T_1, T_2, \ldots$ can be dependent. Only individuals whose weight is different than zero are considered to be alive. The construction we give in this paper would correspond to having

$$T_i = C_i 1_{(N \geq i)}.$$

The definition of a WBP described above leads to the following stochastic recursion for the total weight of the nth generation,

$$W_n \overset{\mathcal{D}}{=} \sum_{i=1}^{\infty} T_i W_{(n-1),i} \tag{2.4}$$

and a corresponding non-homogeneous recursion of the form

$$R \overset{\mathcal{D}}{=} \sum_{i=1}^{\infty} T_i R_i + Q. \tag{2.5}$$

In the construction given in [28], the $\{T_i\}$ and Q are allowed to be dependent as well.

We now briefly describe some of the existing literature on WBPs that more directly relates to our problem. In [28], the martingale structure of W_n/m^n ($m = E[\sum_i T_i]$) was used to point out the existence of $W =$

$\lim_{n\to\infty} W_n/m^n$, and it was shown that some positive stable distributions are solutions to (2.5) when Q and the T_i's are deterministic constants satisfying $\sum_j T_j^\alpha = 1$ for some $\alpha \in (0, 1]$, or when $Q = 0$ and the T_i's are random; also general stable distributions ($0 < \alpha < 2$) are shown to arise when the T_i's are random. Furthermore, for a detailed analysis of the case when W follows a positive stable distribution ($0 < \alpha \leq 1$) see [24]. The convergence of W_n/m^n to W was studied in [30], and conditions for W to belong to the domain of attraction of an α-stable law ($1 < \alpha < 2$) were given in [30], along with an analysis of the rate of convergence. In the case of positive R that we consider, these results are restricted to positive stable laws that allow power law tails only for $0 < \alpha \leq 1$ (see Section 5.3.5 in [7]). A generalization of the WBP described in [28] to a random environment was given in [22], where necessary and sufficient conditions for W to be nondegenerate were derived. More results about the solutions to (2.5) for the case when Q and the T_i's are random are given in [2]. The existence of moments of W was studied in [1]. The power law tail of W for $\alpha > 1$ was derived in Theorem 2.2 of [25] and Proposition 7 of [16]. Note that in this case we study the solution to the non-homogeneous equation (1.1) ($P(Q > 0) > 0$), where the results from [25], [16] do not apply; see the remarks after Theorem 4.2 for additional comments. For an even longer list of references to WBPs and related work see [22] and [2].

From the discussion above it is clear that the prior literature on WBPs is extensive, but we point out that the more specific structure of our model, given by (1.1), as well as our novel analysis via implicit renewal theory, allow us to characterize the asymptotic power law behavior of the distribution of R for all $\alpha > 0$ when the $\{C_i\}$ dominate the tail. In addition, we study the non-homogeneous equation (2.5), while the preceding work primarily focuses on the homogeneous case (2.4). The case when N dominates the tail, which is important for the page ranking problem, has not been considered until very recently in [33] and [32]. In reference to the latter work, our analysis is based on a new sample path approach, while the studies in [33], [32] use transforms and tauberian theorems as well as some more restrictive assumptions. We will provide more details on these connections throughout the paper in remarks after the corresponding theorems.

From a different mathematical perspective, our model also constitutes a generalization of several important types of processes. For instance, by setting $N_n \equiv 1$ and fixing $C_i^{(n)}$ to be a constant, (2.1) reduces to an autoregressive process of order one. Also, by letting N be a Poisson random variable and fixing $C_i \equiv 1$, $Q \equiv 1$, (1.1) becomes the recursion that the number of customers in a busy period of an M/G/1 queue satisfies. Recursion (1.2) and its connection to the busy period when the weights D_i are equal to a deterministic constant was exploited in [23].

It is worth noting that probabilistic sample path approaches for the busy period ($C_i^{(n)} \equiv 1$) were developed in [34, 17, 6]; the work in [34, 17] is also relying on the theory of cycle maximum [3]. However, for our more general model (random $C_i^{(n)}$'s) it is not clear if there is a tractable way of generalizing this analysis. Instead of pursuing the preceding directions, we develop a direct sample path large deviation analysis for recursive random sums that provides greater generality.

3. Moments of W_n

In this section we provide explicit estimates for the moments of the total weight, W_n, of the nth generation that will be used throughout the paper. In particular, we apply these estimates in Section 3.1 to prove that $R_n^* \Rightarrow R$ where $R < \infty$ a.s. Our estimates may be of independent interest due to their explicit nature.

A simple calculation shows that provided $E[N], E[Q], E[C] < \infty$, then $E[W_n] < \infty$ and is given by

$$E[W_n] = E[N]E[C]E[W_{n-1}] = (E[N]E[C])^n E[W_0] = (E[N]E[C])^n E[Q].$$

We give below upper bounds on the general moments of W_n.

Throughout the paper we will use K to denote a large positive constant that may be different in different places, say $K = K/2$, $K = K^2$, etc.

Lemma 3.1. *Suppose* $E[Q^\alpha]E[N]E[C^\alpha] < \infty$ *for* $0 < \alpha \leq 1$*, then*

$$E[W_n^\alpha] \leq (E[C^\alpha]E[N])^n E[Q^\alpha]$$

for all $n \geq 0$*.*

Proof. Simply note that

$$\begin{aligned}
E[W_n^\alpha] &= E\left[\left(\sum_{i=1}^{N} C_i W_{(n-1),i}\right)^\alpha\right] \\
&= \sum_{k=1}^{\infty} E\left[\left(\sum_{i=1}^{k} C_i W_{(n-1),i}\right)^\alpha\right] P(N=k) \\
&\leq \sum_{k=1}^{\infty} E\left[\sum_{i=1}^{k} \left(C_i W_{(n-1),i}\right)^\alpha\right] P(N=k) \qquad (3.1) \\
&= E[C^\alpha]E[N]E[W_{n-1}^\alpha] \\
&\leq (E[C^\alpha]E[N])^n E[W_0^\alpha] \qquad \text{(iterating } n \text{ times)} \\
&= (E[C^\alpha]E[N])^n E[Q^\alpha],
\end{aligned}$$

where for (3.1) we used the well known inequality $\left(\sum_{i=1}^k y_i\right)^\alpha \leq \sum_{i=1}^k y_i^\alpha$ for $0 < \alpha \leq 1$, $y_i \geq 0$ (see e.g., Exercise 4.2.1, p. 102, in [11]).

The lemma for moments greater than one is given below.

Lemma 3.2. *Suppose* $E[Q^\alpha] < \infty$*,* $E[N^\alpha] < \infty$*, and* $E[N]\max\{E[C^\alpha], E[C]\} < 1$ *for some* $\alpha > 1$*. Then, there exists a constant* $K_\alpha > 0$ *such that*

$$E[W_n^\alpha] \leq K_\alpha (E[N]\max\{E[C^\alpha], E[C]\})^n$$

for all $n \geq 0$*.*

The proof of Lemma 3.2 is given in Section 7.1.

REMARK: Recall that when $C \equiv 1$ and $Q \equiv 1$ then $E[W_n^\alpha]$ is the α-moment of a subcritical branching process Z_n and our result reduces to $E[Z_n^\alpha] \leq K_\alpha (E[N])^n$, which is in agreement with the classical results from branching processes, e.g. see Corollary 1 on p. 18 of [5]. Moreover, from the proof of the integer α case (given in Section 7.1), it is clear that W_n^α scales as $\rho^{\alpha n}$ if $(E[C])^\alpha < E[C^\alpha]$ and as ρ_α^n if $(E[C])^\alpha > E[C^\alpha]$, where $\rho = E[N]E[C]$ and $\rho_\alpha = E[N]E[C^\alpha]$. Note that this is not quite the same as our upper bounds, and the reason we choose the geometric term $(\rho \vee \rho_\alpha)^n$ instead is that it makes the proofs simpler and is sufficient for our purposes. Similar techniques to those used in proving the preceding lemmas can yield, with some additional work, lower bounds for the α-moments of W_n, showing that the correct leading term is $(\rho^\alpha \vee \rho_\alpha)^n$. More technical results dealing with the existence of the α-moments of $W \triangleq \lim_{n\to\infty} W_n/\rho^n$ can be found in [1]. There, necessary and sufficient conditions are given for the finiteness of $E[W^\alpha L(W)]$ when $\alpha \geq 1$ and $L(\cdot)$ is slowly varying (see Theorems 1.2 and 1.3). In particular, the approach the authors take is to first normalize the process so that $\rho = E[W_1] = 1$, and then impose a condition that in our case reduces to $\rho_\alpha = E[N]E[C^\alpha] < 1$, that is, they preclude the situation where W_n^α might scale as ρ_α^n when $\rho^\alpha < \rho_\alpha$. An example where W_n^α scales as ρ_α^n is when $N \equiv 1$, since then $W_n^\alpha \overset{\mathcal{D}}{=} Q^\alpha \prod_{i=1}^n C_i^\alpha$.

3.1. Convergence of R_n^* and finiteness of R

As discussed in Section 2.2, there are two issues regarding the process R_n^* that remain to be addressed. One, is the proof that

$$R_n^* \Rightarrow R = \sum_{k=0}^{\infty} W_k$$

for any initial condition R_0^*; the other one is the finiteness of R. The lemma below shows that $R < \infty$ a.s.

Lemma 3.3. *Suppose $E[Q^\beta] < \infty$, $E[N^\beta] < \infty$, and $E[N]E[C^\beta] < 1$ for some $\beta > 0$. Then, $E[R^\beta] < \infty$, and in particular, $R < \infty$ a.s. Moreover, if $\beta \geq 1$, $R_n \overset{L_\beta}{\to} R$, where L_β stands for convergence in $(E|\cdot|^\beta)^{1/\beta}$ norm.*

Proof. Let

$$\eta = \begin{cases} E[N]E[C^\beta], & \text{if } \beta \leq 1 \\ E[N]\min\{E[C], E[C^\beta]\}, & \text{if } \beta > 1. \end{cases}$$

Then by Lemmas 3.1 and 3.2,

$$E[W_n^\beta] \leq K\eta^n \tag{3.2}$$

for some $K > 0$. Suppose $\beta \geq 1$, then, by monotone convergence and Minkowski's inequality,

$$\begin{aligned} E[R^\beta] &= E\left[\lim_{n\to\infty}\left(\sum_{k=0}^n W_k\right)^\beta\right] = \lim_{n\to\infty} E\left[\left(\sum_{k=0}^n W_k\right)^\beta\right] \\ &\leq \lim_{n\to\infty}\left(\sum_{k=0}^n E[W_k^\beta]^{1/\beta}\right)^\beta \leq K\left(\sum_{k=0}^\infty \eta^{k/\beta}\right)^\beta < \infty. \end{aligned}$$

This implies that $R < \infty$ a.s. When $0 < \beta < 1$ use the inequality $\left(\sum_{k=0}^n y_k\right)^\beta \leq \sum_{k=0}^n y_k^\beta$ for any $y_i \geq 0$ instead of Minkowski's inequality. That $R_n \overset{L_\beta}{\to} R$ whenever $\beta \geq 1$ follows from noting that $E[|R_n - R|^\beta] = E\left[\left(\sum_{k=n+1}^\infty W_k\right)^\beta\right]$ and applying the same arguments used above to obtain the bound $E[|R_n - R|^\beta] \leq K\eta^{n+1}/(1-\eta^{1/\beta})^\beta$.

Next, it is easy to verify, by conditioning on N_{n-1}, $\{C_i^{(n-1)}\}_{i\geq 1}$ and Q_{n-1} in equation (2.3) and applying dominated convergence, that R must solve

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

where $\{R_i\}_{i\geq 1}$ are iid copies of R.

We now turn our attention to the proof of the convergence of R_n^* to R. Recall from Section 2.2 that

$$R_n^* \overset{\mathcal{D}}{=} R_{n-1} + W_n(R_0^*), \tag{3.3}$$

where

$$W_n(R_0^*) = \sum_{(i_1,\ldots,i_n)\in A_n} R_{0,(i_1,\ldots,i_n)}^* \mathbf{C}_{i_1,\ldots,i_n}^{(n)}.$$

The following lemma shows that $R_n^* \Rightarrow R$ for any initial condition R_0^* satisfying a moment assumption.

Lemma 3.4. *For any* $R_0^* \geq 0$*, if* $E[Q^\beta], E[(R_0^*)^\beta] < \infty$ *and* $E[N]E[C^\beta] < 1$ *for some* $0 < \beta \leq 1$*, then*

$$R_n^* \Rightarrow R.$$

Furthermore, the distribution of R *is the unique solution to recursion* (1.1).

Proof. In view of (3.3), and since $R_n \to R$ a.s., the result will follow from Slutsky's Theorem (see Theorem 1, p. 254 in [11]) once we show that $W_n(R_0^*) \Rightarrow 0$. Recall that $W_n(R_0^*)$ is the same as W_n if we substitute the $Q_{i_1,\dots,i_n}$ by the $R^*_{0,(i_1,\dots,i_n)}$. Fix $\epsilon > 0$, then

$$\begin{aligned} P(W_n(R_0^*) > \epsilon) &\leq \epsilon^{-\beta} E[W_n(R_0^*)^\beta] \\ &\leq \epsilon^{-\beta} (E[C^\beta]E[N])^n E[(R_0^*)^\beta] \qquad \text{(by Lemma 3.1).} \end{aligned}$$

Since by assumption the right hand side converges to zero as $n \to \infty$, then $R_n^* \Rightarrow R$. Clearly, the distribution of R represents the unique solution to (1.1), since any other possible solution would have to converge to the same limit.

Remarks: (i) Note that when $E[N] < 1$, then the branching tree is a.s. finite and no conditions on the C's are necessary for $R < \infty$ a.s. This corresponds to the second condition in Theorem 1 of [9]. (ii) In view of the same theorem from [9], one could possibly establish the convergence of $R_n^* \Rightarrow R < \infty$ under milder conditions. However, since in this paper we only study the power tails of R, the assumptions of Lemma 3.4 are not restrictive.

4. The case when the C's dominate: Implicit renewal theory

In this section we study the power law phenomenon that arises from the multiplicative effects of the weights $\{C_i\}$ in (1.1).

4.1. Implicit Renewal Theorem on Trees

One observation that will help gain some intuition about (2.3) is to consider the case when $N \equiv 1$. The process $\{R_n\}$ then reduces to

$$R_n \overset{\mathcal{D}}{=} Q_{n-1} + C_{n-1} R_{n-1},$$

also known as a (random coefficient) autoregressive process of order one. The steady state solution to this recursion satisfies

$$R \overset{\mathcal{D}}{=} Q + CR,$$

where R is independent of C and Q. This is precisely one of the stochastic recursions considered in [15] (see also [20]), where it is shown that under the assumption that $E[C^\alpha] = 1$ and some other technical conditions on the distribution of C, we have that

$$P(R > x) \sim H x^{-\alpha} \tag{4.1}$$

for some (computable) constant $H > 0$ (see Theorem 4.1 in [15]). The fact that the index of the power law depends on the distribution of the weights is already promising in terms of our goal of identifying other sources of power law behavior.

Informally speaking, the recursions studied in [15] are basically multiplicative away from the boundary. However, (1.1) always has an additive component given by $\sum_{i=1}^N C_i R_i$ regardless of how far from the boundary one may be. Fortunately, due to the heavy-tailed nature of R, our intuition says that it is only one of the additive $C_i R_i$ components that determines the behavior of (1.1), thus the sum will behave as the maximum term, simplifying to

$$P\left(Q + \sum_{i=1}^N C_i R_i > x\right) \sim E[N] P(CR > x), \tag{4.2}$$

assuming that Q has a light enough tail. This heuristic suggests the following generalization of Theorem 2.3 from [15].

Here, we would like to emphasize that R and C in the following theorem can be *any* two independent random variables that satisfy the stated conditions, i.e., they do not have to be related by recursion (1.1). Hence, the theorem may be of potential use in other applications. Note that we prove the theorem for a general constant m, that in our application refers to $E[N]$, as suggested by (4.2).

Theorem 4.1. *Suppose $C \geq 0$ a.s., $0 < E[C^\alpha \log C] < \infty$ for some $\alpha > 0$, and that the conditional distribution of $\log C$ given $C \neq 0$ is nonarithmetic. Suppose further that R is independent of C, $mE[C^\alpha] = 1$, and that $E[R^\beta] < \infty$ for some $0 < \beta < 1$. If*

$$\int_0^\infty |P(R > t) - mP(CR > t)| \, t^{\alpha-1} dt < \infty, \tag{4.3}$$

then

$$P(R > t) \sim Ht^{-\alpha}, \qquad t \to \infty,$$

where

$$H = \frac{1}{mE[C^\alpha \log C]} \int_0^\infty v^{\alpha-1}(P(R > v) - mP(CR > v))\, dv.$$

The proof of this theorem follows the same steps as Theorem 2.3 from [15], and is presented in Section 7.2.

REMARKS: (i) As pointed out in [15], the statement of the theorem has content only when R has infinite moment of order α, since otherwise the constant $H = (\alpha E[N]E[C^\alpha \log C])^{-1}(E[R^\alpha] - E[N]E[(CR)^\alpha])$ will be zero by independence of R and C. (ii) Note that some of the assumptions of Theorem 4.1 are stronger than the corresponding ones from Theorem 2.3 in [15]. In particular, it is no longer the case that $E[C^\alpha \log C] > 0$ whenever α solves $mE[C^\alpha] = 1$, since if $m > 1$ it is possible to construct counterexamples, hence the need to include this as an assumption. Another difference is our requirement that $E[R^\beta] < \infty$ for some $0 < \beta < 1$. In the case of applying Theorem 4.1 to recursion (1.1), the condition on $E[R^\beta]$ is not restrictive since we readily obtain the moments of R from the computed moments of W_n from Section 3.

In what follows we will use the preceding theorem to derive the asymptotic behavior of $P(R > x)$ when R satisfies (1.1). Here, the main difficulty will be to show that

$$|P(R > x) - E[N]P(CR > x)| = o(x^{\alpha-1}).$$

Theorem 4.2. *Suppose that $0 < E[C^\alpha \log C] < \infty$ for some $\alpha > 0$, the conditional distribution of $\log C$ given $C \neq 0$ is nonarithmetic, and that C and R are independent where R satisfies recursion (1.1). Then, if $E[N]E[C^\alpha] = 1$, and $E[Q^\alpha], E[N^{\alpha \vee (1+\epsilon)}] < \infty$ for some $0 < \epsilon < 1$, we have*

$$P(R > t) \sim Ht^{-\alpha}, \qquad t \to \infty,$$

where

$$\begin{aligned} H &= \frac{1}{E[N]E[C^\alpha \log C]} \int_0^\infty v^{\alpha-1}(P(R > v) - E[N]P(CR > v))\, dv \\ &= \frac{E\left[\left(\sum_{i=1}^N C_i R_i + Q\right)^\alpha - \sum_{i=1}^N (C_i R_i)^\alpha\right]}{\alpha E[N]E[C^\alpha \log C]}. \end{aligned}$$

REMARKS: (i) Note that the second expression for H is more suitable for actually computing it, especially in the case of α being an integer, as will be stated in the forthcoming corollary. (ii) When α is not an integer we can derive an explicit bound on H by using the forthcoming Lemma 4.3 and (4.5). (ii) For the homogeneous equation ($Q \equiv 0$) and $\alpha > 1$, closely related results to our theorem can be found in Theorem 2.2 of [25] and Proposition 7 of [16]; the latter result covers the lattice case. The approach from [25] transforms the recursion $W \overset{\mathcal{D}}{=} \sum_{i=1}^N C_i W_i$ for the critical case $E[W] = 1$, $E\left[\sum_{i=1}^N C_i\right] = 1$ to a first order difference (autoregressive) equation on a different probability space, see Lemma 4.1 in [25]. Note that the tail behavior of W does not imply that of R. Furthermore, it appears that the method from [25] does not extend to the non-homogeneous case since the proof of Lemma 4.1 in [25] critically depends on having both $E[W] = 1$ and $E\left[\sum_{i=1}^N C_i\right] = 1$, which is only possible when $Q \equiv 0$. (iii) Related results for stable laws have been considered in [28] (see Theorem 6), and more recently, in [24] for $0 < \alpha \leq 1$. (iv) Moreover, the results obtained in the references cited above appear to be less explicit in the expression for H than the statement of Theorem 4.2, as Corollary 4.1 below illustrates. (v) Furthermore, Theorem 4.1 and the preceding technique of Theorem 4.2 can be adapted to analyze other, possibly non-linear, recursions on trees, e.g., one can characterize the asymptotic behavior of $P(R > x)$ that solves

$$R = Q + \max_{1 \leq i \leq N} C_i R_i.$$

(vi) We also want to point out that one can obtain the *logarithmic asymptotics* of R, that is, the behavior of $\log P(R > x)$, much easier and under less restrictive conditions, e.g. the C_i's need not be nonarithmetic (this condition is required because the use of the Renewal Theorem). An upper bound can be obtained from our moment estimates and the union bound. For the lower bound, we can inductively use

$$\begin{aligned} P(R > x) &\geq P(W_n > x) \geq P\left(\max_{1 \leq i \leq N} C_i W_{n-1,i} > x, N \leq k\right) \\ &\geq E\left[N P(C W_{n-1} \leq x)^N 1_{(N \leq k)}\right] P(C W_{n-1} > x) \end{aligned}$$

and $P(CW_{n-1} \leq x) \geq P(R \leq x)$, for all x, to show that for any $0 < \epsilon < 1$, all $n \geq 0$ and x large enough,

$$P(R > x) \geq (1-\epsilon)^n (E[N])^n P\left(\prod_{i=1}^n C_i > x\right).$$

The proof can be completed by optimizing the choice of n and using standard (Cramér type) large deviations arguments, more precisely, by choosing the change of measure $\eta(du) = e^{\alpha u} E[N] P(\log C \in du)$, and setting $n = \log x / \mu_\alpha$, where μ_α is the expected value of $\log C$ under the new measure η. Hence, one can derive with a considerably smaller effort and more general conditions

$$\log P(R > x) \sim -\alpha \log x,$$

where α satisfies $E[N]E[C^\alpha] = 1$. Therefore, the majority of the work in proving Theorem 4.2 goes into the derivation of the exact asymptotic. Furthermore, it is worth noting that the logarithmic approach, although less precise, can be obtained in a more general setting. For example, one can have $C_\cdot^{(\cdot)}$ to be dependent across different generations, as in the so called WBP in a random environment. Here, one could derive the asymptotics of $\log P(R > x)$ if $E\left[\left(\prod_{i=1}^n C_{(1,1,\ldots,1)}^{(n)}\right)^\alpha\right]$ satisfies the polynomial type Gärtner-Ellis conditions that were recently considered in [18].

Corollary 4.1. *For integer $\alpha \geq 1$, and under the same assumptions of Theorem 4.2, the constant H can be explicitly computed as a function of $E[R^k], E[C^k], E[Q^k]$, $0 \leq k \leq \alpha - 1$. In particular, for $\alpha = 1$,*

$$H = \frac{E[Q]}{E[N]E[C \log C]},$$

and for $\alpha = 2$,

$$H = \frac{E[Q^2] + 2E[Q]E[C]E[N]E[R] + 2E[N(N-1)](E[C]E[R])^2}{2E[N]E[C^2 \log C]}, \qquad E[R] = \frac{E[Q]}{1 - E[N]E[C]}.$$

Proof. The proof follows directly from multinomial expansions of the second expression for H in Theorem 4.2.

Before giving the proof of Theorem 4.2 we state the following three preliminary lemmas. Their proofs are given in Section 7.2.

Lemma 4.1. *Suppose that* $0 < E[C^\alpha \log C] < \infty$ *for some* $\alpha > 0$ *and* $E[N]E[C^\alpha] = 1$*. Assume also that* $E[Q^\alpha], E[N^{\alpha \vee 1}] < \infty$*. Then,*

$$E[R^\beta] < \infty$$

for all $0 < \beta < \alpha$*.*

Proof. The result follows from Lemma 3.3 by simply noting that the assumptions imply that $E[N]E[C^\beta] < 1$ for all $\beta < \alpha$.

Lemma 4.2. *Let* $\alpha > 1$ *and* $p = \lceil \alpha \rceil \in \{2, 3, 4, \dots\}$*.*

1. *For any sequence of nonnegative numbers* $\{y_i\}_{i \geq 1}$ *and any* $k \in \{1, 2, 3, \dots\}$ *we have*

$$\left(\sum_{i=1}^k y_i \right)^\alpha \geq \sum_{i=1}^k y_i^\alpha.$$

2. *For any sequence of iid nonnegative random variables* $\{Y, Y_i\}_{i \geq 1}$ *and any* $k \in \{1, 2, 3, \dots\}$ *we have*

$$E\left[\left(\sum_{i=1}^k Y_i \right)^\alpha - \sum_{i=1}^k Y_i^\alpha \right] \leq k^\alpha E[Y^{p-1}]^{\alpha/(p-1)}.$$

Lemma 4.3. *Suppose* $\{C, C_i\}$ *and* $\{R, R_i\}$ *are iid sequences of nonnegative random variables independent of each other and of* N*. Assume that* $E[C^\alpha] < \infty$*,* $E[N^{1+\epsilon}] < \infty$ *for some* $0 < \epsilon < 1$*, and* $E[R^\beta] < \infty$ *for any* $0 < \beta < \alpha$*. Then,*

$$0 \leq \int_0^\infty \left(E[N]P(CR > t) - P\left(\max_{1 \leq i \leq N} C_i R_i > t \right) \right) t^{\alpha-1} \, dt = \frac{1}{\alpha} E\left[\sum_{i=1}^N (C_i R_i)^\alpha - \left(\max_{1 \leq i \leq N} C_i R_i \right)^\alpha \right] < \infty.$$

Proof of Theorem 4.2. By Lemma 4.1 we know that $E[R^\beta] < \infty$ for any $0 < \beta < \alpha$. The statement of the theorem with the first expression for H will follow from Theorem 4.1 once we prove condition (4.3) for $m = E[N]$. Define

$$R^* = \sum_{i=1}^N C_i R_i + Q.$$

Then,

$$\begin{aligned} |P(R > t) - E[N]P(CR > t)| \leq & \left| P(R > t) - P\left(\max_{1 \leq i \leq N} C_i R_i > t \right) \right| \\ & + \left| P\left(\max_{1 \leq i \leq N} C_i R_i > t \right) - E[N]P(CR > t) \right|. \end{aligned}$$

Since $R \stackrel{\mathcal{D}}{=} R^* \geq \max_{1\leq i\leq N} C_i R_i$, the first absolute value disappears. For the second one note that by the union bound

$$E[N]P(CR>t) - P\left(\max_{1\leq i\leq N} C_i R_i > t\right) = E\left[NP(CR>t) - 1 + P(CR\leq t)^N\right] \geq 0.$$

It follows that

$$\begin{aligned} |P(R>t) - E[N]P(CR>t)| &\leq P(R>t) - P\left(\max_{1\leq i\leq N} C_i R_i > t\right) \\ &\quad + E[N]P(CR>t) - P\left(\max_{1\leq i\leq N} C_i R_i > t\right). \end{aligned} \tag{4.4}$$

Note that we only need to verify that

$$\int_0^\infty \left(P(R>t) - P\left(\max_{1\leq i\leq N} C_i R_i > t\right)\right) t^{\alpha-1}\, dt < \infty,$$

since the integral corresponding to (4.4) is finite by Lemma 4.3. To see this note that $R \stackrel{\mathcal{D}}{=} R^*$ and $1_{(R^*>t)} - 1_{(\max_{1\leq i\leq N} C_i R_i > t)} \geq 0$, thus, by Fubini's Theorem, we have

$$\int_0^\infty \left(P(R>t) - P\left(\max_{1\leq i\leq N} C_i R_i > t\right)\right) t^{\alpha-1}\, dt = \frac{1}{\alpha} E\left[(R^*)^\alpha - \left(\max_{1\leq i\leq N} C_i R_i\right)^\alpha\right].$$

If $0<\alpha\leq 1$ we apply the inequality $\left(\sum_{i=1}^k x_i\right)^\beta \leq \sum_{i=1}^k x_i^\beta$ for $0<\beta\leq 1$, $x_i \geq 0$, to obtain

$$E\left[(R^*)^\alpha - \left(\max_{1\leq i\leq N} C_i R_i\right)^\alpha\right] \leq E\left[Q^\alpha + \sum_{i=1}^N (C_i R_i)^\alpha - \left(\max_{1\leq i\leq N} C_i R_i\right)^\alpha\right],$$

which is finite by Lemma 4.3 and the assumption $E[Q^\alpha]<\infty$. If $\alpha>1$ we have $\left(\sum_{i=1}^k x_i\right)^\alpha \geq \sum_{i=1}^k x_i^\alpha$ by Lemma 4.2, so we can split the expectation as follows

$$E\left[(R^*)^\alpha - \left(\max_{1\leq i\leq N} C_i R_i\right)^\alpha\right] = E\left[(R^*)^\alpha - \sum_{i=1}^N (C_i R_i)^\alpha\right] + E\left[\sum_{i=1}^N (C_i R_i)^\alpha - \left(\max_{1\leq i\leq N} C_i R_i\right)^\alpha\right],$$

which can be done since both expressions inside the expectations on the right hand side are nonnegative. The second expectation is again finite by Lemma 4.3. To see that the first expectation is finite let $S = \sum_{i=1}^N C_i R_i$ and note that $R^* = S + Q$, where S and Q are independent. Let $p = \lceil \alpha \rceil$ and note that $1 \leq p-1 < \alpha$. Then, by Lemma 4.2,

$$\begin{aligned} E\left[(R^*)^\alpha - \sum_{i=1}^N (C_i R_i)^\alpha\right] &= E\left[(S+Q)^\alpha - S^\alpha\right] + E\left[\left(\sum_{i=1}^N C_i R_i\right)^\alpha - \sum_{i=1}^N (C_i R_i)^\alpha\right] \\ &\leq E\left[(S+Q)^\alpha - S^\alpha\right] + E\left[N^\alpha\right]\left(E[(CR)^{p-1}]\right)^{\alpha/(p-1)}. \end{aligned}$$

The second expectation is finite since by Lemma 3.3 $E[R^\beta]<\infty$ for any $0<\beta<\alpha$. For the first expectation we use the inequality

$$(x+t)^\kappa \leq \begin{cases} x^\kappa + t^\kappa, & 0<\kappa\leq 1, \\ x^\kappa + (x+t)^{\kappa-1} t, & \kappa > 1, \end{cases}$$

for any $x, t \geq 0$. We apply the second expression $p-1$ times and then the first one to obtain

$$(x+t)^\alpha \leq x^\alpha + (x+t)^{\alpha-1}t \leq \cdots \leq x^\alpha + \sum_{i=1}^{p-2} x^{\alpha-i}t^i + (x+t)^{\alpha-p+1}t^{p-1} \leq x^\alpha + t^\alpha + \sum_{i=1}^{p-1} x^{\alpha-i}t^i.$$

We conclude that

$$E\left[(S+Q)^\alpha - S^\alpha\right] \leq E[Q^\alpha] + \sum_{i=1}^{p-1} E[S^{\alpha-i}]E[Q^i], \tag{4.5}$$

where $E[S^{\alpha-i}] \leq E[(R^*)^{\alpha-i}] < \infty$ for any $1 \leq i \leq p-1$ by Lemma 3.3.

Finally, applying Theorem 4.1 gives

$$P(R > t) \sim Ht^{-\alpha},$$

where $H = (E[N]E[C^\alpha \log C])^{-1} \int_0^\infty v^{\alpha-1}(P(R > v) - E[N]P(CR > v))\, dv$.

To obtain the second expression for H note that

$$\begin{aligned}
&\int_0^\infty v^{\alpha-1}(P(R > v) - E[N]P(CR > v))\, dv \\
&= \int_0^\infty v^{\alpha-1} \left(E\left[1_{(\sum_{i=1}^N C_iR_i + Q > v)}\right] - E\left[\sum_{i=1}^N 1_{(C_iR_i > v)}\right] \right) dv \\
&= E\left[\int_0^\infty v^{\alpha-1} \left(1_{(\sum_{i=1}^N C_iR_i + Q > v)} - \sum_{i=1}^N 1_{(C_iR_i > v)} \right) dv \right] &\quad (4.6) \\
&= E\left[\int_0^{\sum_{i=1}^N C_iR_i + Q} v^{\alpha-1} dv - \sum_{i=1}^N \int_0^{C_iR_i} v^{\alpha-1} dv \right] &\quad (4.7) \\
&= \frac{1}{\alpha} E\left[\left(\sum_{i=1}^N C_iR_i + Q \right)^\alpha - \sum_{i=1}^N (C_iR_i)^\alpha \right],
\end{aligned}$$

where (4.6) is justified by Fubini's Theorem and the absolute integrability of $v^{\alpha-1}(P(R > v) - E[N]P(CR > v))$, and (4.7) is justified from the observation that

$$v^{\alpha-1} 1_{(\sum_{i=1}^N C_iR_i + Q > v)} \qquad \text{and} \qquad v^{\alpha-1} \sum_{i=1}^N 1_{(C_iR_i > v)}$$

are each almost surely absolutely integrable as well. This completes the proof.

5. The case when the N dominates

We now turn our attention to the distributional properties of R_n and R when N has a heavy-tailed distribution (in particular, regularly varying) that is heavier than the potential power law effect arising from the multiplicative weights $\{C_i\}$. This case is particularly important for understanding the behavior of Google's PageRank algorithm since the C_i's are smaller than one and the in-degree distribution of the Web graph is well accepted to be a power law. We start this section by stating the corresponding lemma that describes the asymptotic behavior of R_n. The main technical difficulty of extending this lemma to steady state ($R = R_\infty$) is to develop a uniform bound for $R - R_n$, which is enabled by our main technical result of this section, Proposition 5.1. The proof of the lemma is given in Section 7.3.

Lemma 5.1. *Suppose N is regularly varying with index $\alpha > 1$ and suppose $E[Q^{\alpha+\epsilon}] < \infty$, $E[C^{\alpha+\epsilon}] < \infty$ for some $\epsilon > 0$. Let $\rho = E[N]E[C]$ and $\rho_\alpha = E[N]E[C^\alpha]$ and assume that $R_{0,i} = Q_i$ for all i (so that $R_0 = W_0$*

coincides with the tree construction of R_n). Then, for any fixed $n \in \{1,2,3,\dots\}$,

$$P(R_n > x) \sim \frac{(E[C]E[Q])^\alpha}{(1-\rho)^\alpha} \sum_{k=0}^{n} \rho_\alpha^k (1-\rho^{n-k})^\alpha P(N > x) \tag{5.1}$$

as $x \to \infty$.

Lemma 5.2. *Suppose N is regularly varying with index $\alpha > 1$ and suppose $E[Q^{\alpha+\epsilon}] < \infty$, $E[C^{\alpha+\epsilon}] < \infty$ for some $\epsilon > 0$. Let $\rho = E[N]E[C]$ and $\rho_\alpha = E[N]E[C^\alpha]$. Then, for any fixed $n \in \{1,2,3,\dots\}$,*

$$P(W_n > x) \sim (E[C]E[Q])^\alpha \sum_{k=0}^{n-1} \rho_\alpha^k \rho^{(n-1-k)\alpha} P(N > x)$$

as $x \to \infty$.

From this result it is to be expected that a bound of the form

$$P(W_n > x) \leq K\eta^n P(N > x)$$

might hold for all n and $x \geq 1$, for some $\rho \vee \rho_\alpha < \eta < 1$. Such a bound will provide the necessary tools to ensure that $R - R_n$ is negligible for large enough n, allowing the exchange of limits in Lemma 5.1. Proving this result is the main technical contribution of this section; the actual proof is given in Section 7.3. This bound may be of independent interest for computing the distributional properties of other recursions on branching trees, e.g. it is straightforward to apply our method to study the solution to

$$R = Q + \max_{1 \leq i \leq N} C_i R_i.$$

Proposition 5.1. *Suppose $P(N > x) = x^{-\alpha} L(x)$, with $L(\cdot)$ slowly varying and $\alpha > 1$, $E[C^{\alpha+\nu}] < \infty$, $E[Q^{\alpha+\nu}] < \infty$ for some $\nu > 0$, and let $E[N]\max\{E[C^\alpha], E[C]\} < \eta < 1$. Then, there exists a constant $K = K(\eta,\nu) > 0$, that does not depend on n, such that for all $n \geq 1$ and all $x \geq 1$,*

$$P(W_n > x) \leq K\eta^n P(N > x). \tag{5.2}$$

We would also like to point out that a bound of type (5.2) resembles a classical result by Kesten (see Lemma 7 on p. 149 of [4]) stating that the sum of heavy-tailed (subexponential) random variables satisfies

$$P(X_1 + \cdots + X_n > x) \leq K(1+\epsilon)^n P(X_1 > x),$$

uniformly for all n and x, for any $\epsilon > 0$ (see also [13] for more recent work). The main difference between this result and (5.2) is that while n above refers to the number of terms in the sum, in (5.2) it refers to the *depth of the recursion*. This makes the derivation of (5.2) considerably more complicated, and perhaps implausible if it were not for the fact that we restrict our attention to regularly varying distributions, as opposed to the general subexponential class.

In view of (5.2), we can now prove the main theorem of this section.

Theorem 5.1. *Suppose $P(N > x) = x^{-\alpha} L(x)$, with $L(\cdot)$ slowly varying and $\alpha > 1$. Let $\rho = E[N]E[C]$ and $\rho_\alpha = E[N]E[C^\alpha]$. Suppose $\rho \vee \rho_\alpha < 1$, and $E[C^{\alpha+\epsilon}], E[Q^{\alpha+\epsilon}] < \infty$ for some $\epsilon > 0$. Then,*

$$P(R > x) \sim \frac{(E[C]E[Q])^\alpha}{(1-\rho)^\alpha (1-\rho_\alpha)} P(N > x)$$

as $x \to \infty$.

REMARKS: (i) A related result was derived very recently in [32] using transform methods and tauberian theorems under the more restrictive conditions $E[Q] < 1$, $E[C] = (1 - E[Q])/E[N]$. (ii) Note that this result implies the classical result on the busy period of an M/G/1 queue derived in [12]. Specifically, the total number of customers in a busy period B satisfies the recursion $B \stackrel{\mathcal{D}}{=} 1 + \sum_{i=1}^{N(S)} B_i$, where the B_i's are iid copies of B, $N(t)$ is a Poisson process of rate λ and S is the service distribution; $\{B_i\}$, $N(t)$ and S are mutually independent and $\rho = E[N(S)] < 1$. Now, the recursion for B is obtained from our theorem by setting $C \equiv 1$ and $Q \equiv 1$, implying that $P(B > x) \sim P(N(S) > x)/(1-\rho)^{\alpha+1}$. Next, one can obtain the asymptotics for the length of the busy period P by using the identity $B = N(P)$. This can be easily derived, in spite of the fact that $N(t)$ and P are correlated, since $N(t)$ is highly concentrated around its mean. For recent work on the polynomial asymptotics of the GI/GI/1 busy period see [34]. (iii) In view of Lemma 5.1, the theorem shows that the limits $\lim_{x\to\infty} \lim_{n\to\infty} P(R_n > x)/P(N > x)$ are interchangeable.

Proof of Theorem 5.1. Fix $0 < \delta < 1$ and $n_0 \geq 1$. Choose $\rho \vee \rho_\alpha < \eta < 1$ and use Proposition 5.1 to obtain that for some constant $K_0 > 0$,

$$P(W_n > x) \leq K_0 \eta^n P(N > x)$$

for all $n \geq 1$ and all $x \geq 1$. Let $H_\alpha^{(n)} = (E[C]E[Q])^\alpha (1-\rho)^{-\alpha} \sum_{k=0}^n \rho_\alpha^k (1 - \rho^{n-k})^\alpha$ and $H_\alpha = H_\alpha^{(\infty)}$. Then,

$$\begin{aligned}
&|P(R > x) - H_\alpha P(N > x)| \\
&\leq |P(R > x) - P(R_{n_0} > x)| && (5.3)\\
&\quad + \left| P(R_{n_0} > x) - H_\alpha^{(n_0)} P(N > x) \right| && (5.4)\\
&\quad + \left| \sum_{k=0}^{n_0} \rho_\alpha^k (1 - \rho^{n_0-k})^\alpha - \sum_{k=0}^{\infty} \rho_\alpha^k \right| (1-\rho_\alpha) H_\alpha P(N > x). && (5.5)
\end{aligned}$$

By Lemma 5.1, there exists a function $\varphi(x) \downarrow 0$ as $x \to \infty$ such that

$$\left| P(R_{n_0} > x) - H_\alpha^{(n_0)} P(N > x) \right| \leq \varphi(x) H_\alpha P(N > x).$$

To bound (5.3) let $\beta = \eta^{1/(2\alpha+2)} < 1$ and note that

$$\begin{aligned}
|P(R > x) - P(R_{n_0} > x)| &\leq P\left(R_{n_0} + (R - R_{n_0}) > x,\, R - R_{n_0} \leq \delta x\right) - P(R_{n_0} > x) \\
&\quad + P\left(R - R_{n_0} > \delta x\right) \\
&\leq P(R_{n_0} > (1-\delta)x) - P(R_{n_0} > x) + P\left(\sum_{n=n_0+1}^{\infty} W_n > \delta x \right) \\
&\leq P(R_{n_0} > (1-\delta)x) - H_\alpha^{(n_0)} P(N > (1-\delta)x) \\
&\quad + H_\alpha^{(n_0)} P(N > x) - P(R_{n_0} > x) \\
&\quad + H_\alpha^{(n_0)} P(N > (1-\delta)x) - H_\alpha^{(n_0)} P(N > x) \\
&\quad + \sum_{n=n_0+1}^{\infty} P\left(W_n > \delta x (1-\beta)\beta^{n-n_0-1}\right) \\
&\leq \left\{ 2\varphi((1-\delta)x) \frac{P(N > (1-\delta)x)}{P(N > x)} \right. \\
&\quad \left. + \left(\frac{P(N > (1-\delta)x)}{P(N > x)} - 1 \right) \right\} H_\alpha P(N > x) \\
&\quad + \sum_{n=n_0+1}^{\infty} K_0 \eta^n P\left(N > \delta x (1-\beta)\beta^{n-n_0-1}\right),
\end{aligned}$$

where in the last inequality we applied the uniform bound from Proposition 5.1. The expression in curly brackets is bounded by

$$2\varphi((1-\delta)x)(1-\delta)^{-\alpha}\frac{L((1-\delta)x)}{L(x)} + \left((1-\delta)^{-\alpha}\frac{L((1-\delta)x}{L(x)} - 1\right) \to (1-\delta)^{-\alpha} - 1$$

as $x \to \infty$. By Potter's Theorem (see Theorem 1.5.6 (ii) on p. 25 in [7]), there exists a constant $A > 1$ such that

$$\begin{aligned}
&\sum_{n=n_0+1}^{\infty} K_0 \eta^n P\left(N > \delta x (1-\beta)\beta^{n-n_0-1}\right) \\
&\le K_0 A \sum_{n=n_0+1}^{\infty} \eta^n \left(\delta(1-\beta)\beta^{n-n_0-1}\right)^{-\alpha-1} P(N > x) \\
&= K_0 A (\delta(1-\beta))^{-\alpha-1}(1-\eta^{1/2})^{-1}\eta^{n_0+1} P(N > x) \\
&\le K\delta^{-\alpha-1}\eta^{n_0} P(N > x).
\end{aligned}$$

Next, for (5.5) simply note that

$$\begin{aligned}
&(1-\rho_\alpha)\left| \sum_{k=0}^{n_0} \rho_\alpha^k (1-\rho^{n_0-k})^\alpha - \sum_{k=0}^{\infty} \rho_\alpha^k \right| \\
&= (1-\rho_\alpha)\sum_{k=0}^{n_0} \rho_\alpha^k (1-(1-\rho^{n_0-k})^\alpha) + (1-\rho_\alpha)\sum_{k=n_0+1}^{\infty} \rho_\alpha^k \\
&\le (1-\rho_\alpha)\sum_{k=0}^{n_0} \rho_\alpha^k \alpha \rho^{n_0-k} + \rho_\alpha^{n_0+1} \\
&\le [\alpha(1-\rho_\alpha)(n_0+1) + \rho_\alpha](\rho_\alpha \vee \rho)^{n_0} \\
&\le K\eta^{n_0}.
\end{aligned}$$

It follows that

$$\lim_{x\to\infty}\left| \frac{P(R > x)}{H_\alpha P(N > x)} - 1 \right| \le (1-\delta)^{-\alpha} - 1 + K\delta^{-\alpha-1}\eta^{n_0}.$$

Since the right hand side can be made arbitrarily small by choosing δ and n_0 appropriately, the result of the theorem follows.

Engineering implications. Recall that for Google's PageRank algorithm the weights are given by $C_i = c/D_i < 1$, where $0 < c < 1$ is a constant related to the damping factor and the number of nodes in the Web graph, and D_i corresponds to the out-degree of a page. We point out that dividing the ranks of neighboring pages by their out-degree has the purpose of decreasing the contribution of pages with highly inflated referencing. However, Theorem 5.1 reveals that the page rank is essentially insensitive to the parameters of the out-degree distribution, which means that PageRank basically reflects the popularity vote given by the number of references N.

Furthermore, Theorem 4.1 clearly shows that the choice of weights C_i in the ranking algorithm can determine the distribution of R as well. Note that for the PageRank algorithm the weights $C_i = c/D_i < 1$ can never dominate the asymptotic behavior of R when N is a power law. Therefore, Theorem 4.1 suggests a potential development of new ranking algorithms where the ranks will be much more sensitive to the weights.

6. The case when the Q's dominate

This section of the paper treats the case when the heavy-tailed behavior of R arises from the $\{Q_i\}$, known in the autoregressive processes literature as innovations. The results presented here are very similar to those in Section 5, and so are their proofs. We will therefore only present the statements of the results and skip most of the proofs. We start with the equivalent of Lemma 5.1 in this context; its proof is given in Section 7.4.

Lemma 6.1. *Suppose Q is regularly varying with index $\alpha > 1$ and suppose $E[N^{\alpha+\epsilon}] < \infty$, $E[C^{\alpha+\epsilon}] < \infty$ for some $\epsilon > 0$. Let $\rho_\alpha = E[N]E[C^\alpha]$ and assume that $R_{0,i} = Q_i$ for all i (so that $R_0 = W_0$ coincides with the tree construction of R_n). Then, for any fixed $n \in \{1, 2, 3, \dots\}$,*

$$P(R_n > x) \sim \sum_{k=0}^{n} \rho_\alpha^k P(Q > x)$$

as $x \to \infty$.

As for the case when N dominates the asymptotic behavior of R, we can here expect that

$$P(R > x) \sim (1 - \rho_\alpha)^{-1} P(Q > x),$$

and the only technical difficulty is justifying the exchange of limits. The same techniques used in Section 5 can be used in this case as well. Therefore, we give a sketch of the arguments in Section 7.4 but omit the proof. The following is the equivalent of Lemma 5.2.

Lemma 6.2. *Suppose Q is regularly varying with index $\alpha > 1$ and suppose $E[N^{\alpha+\epsilon}] < \infty$, $E[C^{\alpha+\epsilon}] < \infty$ for some $\epsilon > 0$. Let $\rho_\alpha = E[N]E[C^\alpha]$. Then, for any fixed $n \in \{1, 2, 3, \dots\}$,*

$$P(W_n > x) \sim \rho_\alpha^n P(Q > x)$$

as $x \to \infty$.

The corresponding version of Proposition 5.1 is given below.

Proposition 6.1. *Suppose $P(Q > x) = x^{-\alpha} L(x)$, with $L(\cdot)$ slowly varying and $\alpha > 1$. Suppose $E[N] \max\{E[C^\alpha], E[C]\} < 1$, and $E[C^{\alpha+\nu}], E[N^{\alpha+\nu}] < \infty$ for some $\nu > 0$. Then, for any $E[N]E[C^\alpha] < \eta < 1$ there exists a constant $K = K(\eta, \nu) > 0$, that does not depend on n, such that for all $n \geq 1$ and all $x \geq 1$,*

$$P(W_n > x) \leq K \eta^n P(Q > x).$$

A sketch of the proof can be found in Section 7.4.

And finally, the main theorem of this section. The proof again greatly resembles that of Theorem 5.1 and is therefore omitted.

Theorem 6.1. *Suppose $P(Q > x) = x^{-\alpha} L(x)$, with $L(\cdot)$ slowly varying and $\alpha > 1$. Let $\rho_\alpha = E[N]E[C^\alpha]$. Suppose $\rho \vee \rho_\alpha < 1$, and $E[C^{\alpha+\epsilon}], E[N^{\alpha+\epsilon}] < \infty$ for some $\epsilon > 0$. Then,*

$$P(R > x) \sim (1 - \rho_\alpha)^{-1} P(Q > x)$$

as $x \to \infty$.

Compare this result with Lemma A.3 in [26], where the autoregressive process of order one with regularly varying innovations is shown to be tail-equivalent to Q. In particular, if we set $N \equiv 1$ in Theorem 6.1 and let $A_k = \prod_{i=1}^{k-1} C_i$, our result reduces to

$$P\left(\sum_{k=0}^{\infty} A_k Q_k > x\right) \sim \sum_{k=0}^{\infty} E[A_k^\alpha] P(Q > x),$$

which is in line with the commonly accepted intuition about heavy-tailed large deviations where large sums are due to one large summand Q_k.

7. Proofs

This section contains the proofs to most of the results presented in the paper, along with some auxiliary lemmas that are needed along the way. The section is divided into four subsections, each corresponding to the content of Sections 3, 4, 5, and 6, respectively.

7.1. Moments of W_n

Here we give the proof of the moment bound for the α-moment, $\alpha > 1$, of the sum of the weights, W_n of the nth generation. As an intermediate step, we present a lemma for the integer moments of W_n, but first we give the proof of Lemma 4.2.

Proof of Lemma 4.2. Let $\{y_i\}_{i\geq 0}$ be a sequence of nonnegative numbers. When $k = 2$ we have

$$\frac{y_1^\alpha + y_2^\alpha}{(y_1+y_2)^\alpha} = \left(\frac{y_1}{y_1+y_2}\right)^\alpha + \left(\frac{y_2}{y_1+y_2}\right)^\alpha \leq \frac{y_1}{y_1+y_2} + \frac{y_2}{y_1+y_2} = 1.$$

The first statement of the lemma follows by induction on k.

For the second statement let $p = \lceil \alpha \rceil \in \{2, 3, \dots\}$ and $\beta = \alpha/p \in (0, 1]$. Define $A_p(k) = \{(j_1, \dots, j_k) \in \mathbb{Z}^k : j_1 + \cdots + j_k = p, 0 \leq j_i < p\}$. Then,

$$\begin{aligned}
\left(\sum_{i=1}^{k} y_i\right)^\alpha &= \left(\sum_{i=1}^{k} y_i\right)^{p\beta} \\
&= \left(\sum_{i=1}^{k} y_i^p + \sum_{(j_1,\dots,j_k)\in A_p(k)} \binom{p}{j_1,\dots,j_k} y_1^{j_1}\cdots y_k^{j_k}\right)^\beta \\
&\leq \sum_{i=1}^{k} y_i^{p\beta} + \left(\sum_{(j_1,\dots,j_k)\in A_p(k)} \binom{p}{j_1,\dots,j_k} y_1^{j_1}\cdots y_k^{j_k}\right)^\beta,
\end{aligned}$$

where for the last step we used the well known inequality $\left(\sum_{i=1}^k x_i\right)^\beta \leq \sum_{i=1}^k x_i^\beta$ for $0 < \beta \leq 1$ and $x_i \geq 0$

(see the proof of Lemma 3.1). We now use Jensen's inequality to obtain

$$\begin{aligned} E\left[\left(\sum_{i=1}^{k} Y_i\right)^{\alpha} - \sum_{i=1}^{k} Y_i^{\alpha}\right] &\leq E\left[\left(\sum_{(j_1,\dots,j_k)\in A_p(k)} \binom{p}{j_1,\dots,j_k} Y_1^{j_1}\cdots Y_k^{j_k}\right)^{\beta}\right] \\ &\leq \left(E\left[\sum_{(j_1,\dots,j_k)\in A_p(k)} \binom{p}{j_1,\dots,j_k} Y_1^{j_1}\cdots Y_k^{j_k}\right]\right)^{\beta} \\ &= \left(\sum_{(j_1,\dots,j_k)\in A_p(k)} \binom{p}{j_1,\dots,j_k} E\left[Y_1^{j_1}\cdots Y_k^{j_k}\right]\right)^{\beta}. \end{aligned}$$

Since the $\{Y_i\}$ are iid, we have

$$E\left[Y_1^{j_1}\cdots Y_k^{j_k}\right] = ||Y||_{j_1}^{j_1}\cdots ||Y||_{j_k}^{j_k},$$

where $||Y||_\kappa = E[|Y|^\kappa]^{1/\kappa}$ for $\kappa \geq 1$ and $||Y||_0 \triangleq 1$. Since $||Y||_\kappa$ is increasing for $\kappa \geq 1$ it follows that $||Y||_{j_i}^{j_i} \leq ||Y||_{p-1}^{j_1}$. It follows that

$$||Y||_{j_1}^{j_1}\cdots ||Y||_{j_k}^{j_k} \leq ||Y||_{p-1}^{p},$$

which in turn implies that

$$\begin{aligned} E\left[\left(\sum_{i=1}^{k} Y_i\right)^{\alpha} - \sum_{i=1}^{k} Y_i^{\alpha}\right] &\leq \left(\sum_{(j_1,\dots,j_k)\in A_p(k)} \binom{p}{j_1,\dots,j_k} ||Y||_{p-1}^{p}\right)^{\beta} \\ &= ||Y||_{p-1}^{\alpha}(k^p - k)^{\beta} \\ &\leq ||Y||_{p-1}^{\alpha} k^{\alpha}. \end{aligned}$$

Lemma 7.1. *Suppose* $E[Q^p] < \infty$, $E[N^p] < \infty$, *and* $E[N]\max\{E[C^p], E[C]\} < 1$ *for some* $p \in \{2,3,\dots\}$. *Then, there exists a constant* $K_p > 0$ *such that*

$$E[W_n^p] \leq K_p \left(E[N]\max\{E[C], E[C^p]\}\right)^n$$

for all $n \geq 0$.

Proof. Let $Y = CW_{(n-1)}$, where C is independent of W_{n-1} and let $\{Y_i\}$ be independent copies of Y. We will give an induction proof in p. For $p = 2$ we have

$$\begin{aligned} E[W_n^2] &= E\left[\left(\sum_{i=1}^{N} Y_i\right)^2\right] \\ &= E\left[\sum_{i=1}^{N} Y_i^2 + \sum_{i\neq j} Y_i Y_j\right] \\ &= E[N]E[Y^2] + E[N(N-1)]E[Y]^2 \\ &= E[N]E[C^2]E[W_{n-1}^2] + E[N(N-1)](E[C]E[W_{n-1}])^2. \end{aligned}$$

Using the preceding recursion, letting $\rho = E[N]E[C]$, $\rho_2 = E[N]E[C^2]$, and noting that,

$$E[W_{n-1}] = \rho^{n-1}E[Q],$$

we obtain

$$E[W_n^2] = \rho_2 E[W_{n-1}^2] + K\rho^{2(n-1)}, \tag{7.1}$$

where $K = E[N(N-1)](E[C]E[Q])^2$. Now, iterating (7.1) gives

$$\begin{aligned} E[W_n^2] &= \rho_2 \left(\rho_2 E[W_{n-2}^2] + K\rho^{2(n-2)} \right) + K\rho^{2(n-1)} \\ &= \rho_2^{n-1} \left(\rho_2 E[W_0^2] + K \right) + K \sum_{i=0}^{n-2} \rho_2^i \, \rho^{2(n-1-i)} \\ &= \rho_2^n E[Q^2] + K \sum_{i=0}^{n-1} \rho_2^i \, \rho^{2(n-1-i)} \\ &\leq (\rho_2 \vee \rho)^n E[Q^2] + K(\rho_2 \vee \rho)^n \sum_{i=0}^{n-1} (\rho_2 \vee \rho)^{n-2-i} \\ &\leq \left(E[Q^2] + \frac{K}{\rho_2 \vee \rho} \sum_{j=0}^{\infty} (\rho_2 \vee \rho)^j \right) (\rho_2 \vee \rho)^n \\ &= K_2 (\rho_2 \vee \rho)^n. \end{aligned}$$

Next, for any $p \in \{2, 3, \dots\}$ let $\rho_p = E[N]E[C^p]$. Suppose now that there exists a constant $K_{p-1} > 0$ such that

$$E[W_n^{p-1}] \leq K_{p-1} \left(\rho_{p-1} \vee \rho \right)^n \tag{7.2}$$

for all $n \geq 0$. By Lemma 4.2 we have

$$\begin{aligned} E[W_n^p] &= \sum_{k=1}^{\infty} E\left[\left(\sum_{i=1}^{k} Y_i \right)^p \right] P(N = k) \\ &\leq \sum_{k=1}^{\infty} \left(kE\left[Y^p\right] + k^p (E[Y^{p-1}])^{p/(p-1)} \right) P(N = k) \\ &= E[N]E[C^p]E[W_{n-1}^p] + E[N^p](E[C^{p-1}])^{p/(p-1)} (E[W_{n-1}^{p-1}])^{p/(p-1)} \\ &\leq \rho_p E[W_{n-1}^p] + E[N^p](E[C^{p-1}])^{p/(p-1)} (K_{p-1})^{p/(p-1)} (\rho_{p-1} \vee \rho)^{(n-1)p/(p-1)}, \end{aligned}$$

where the last inequality corresponds to the induction hypothesis. We then obtain the recursion

$$E[W_n^p] \leq \rho_p E[W_{n-1}^p] + K(\rho_{p-1} \vee \rho)^{\frac{(n-1)p}{p-1}}, \tag{7.3}$$

where $K = E[N^p](E[C^{p-1}])^{p/(p-1)}(K_{p-1})^{p/(p-1)}$. Iterating (7.3) as for the case $p = 2$ gives

$$\begin{aligned} E[W_n^p] &\leq \rho_p^n E[Q^p] + K \sum_{i=0}^{n-1} \rho_p^i \left(\rho_{p-1} \vee \rho \right)^{\frac{(n-1-i)p}{p-1}} \\ &\leq (\rho_p \vee \rho)^n E[Q^p] + K \sum_{i=0}^{n-1} (\rho_p \vee \rho)^{\frac{(n-1)p-i}{p-1}} \\ &= (\rho_p \vee \rho)^n E[Q^p] + K(\rho_p \vee \rho)^n \sum_{i=0}^{n-1} (\rho_p \vee \rho)^{\frac{n-i-p}{p-1}} \\ &\leq \left(E[Q^p] + K(\rho_p \vee \rho)^{-1} \sum_{j=0}^{\infty} (\rho_p \vee \rho)^{\frac{j}{p-1}} \right) (\rho_p \vee \rho)^n \\ &= K_p (\rho_p \vee \rho)^n. \end{aligned}$$

The proof for the general α-moment, $\alpha > 1$, is given below.

Proof of Lemma 3.2. Set $p = \lceil \alpha \rceil \geq \alpha > 1$. Since the result when $\beta = 0$ follows from Lemma 7.1, we assume that $\beta > 0$. Let $Y = CW_{(n-1)}$, where C is independent of W_{n-1} and $\{Y_i\}$ are independent copies of Y. Also, recall that $\rho = E[N]E[C]$ and $\rho_\alpha = E[N]E[C^\alpha]$. Then, by Lemma 4.2,

$$\begin{aligned} E[W_n^\alpha] &= E\left[\left(\sum_{i=1}^N Y_i \right)^\alpha \right] \\ &= \sum_{k=1}^\infty E\left[\left(\sum_{i=1}^k Y_i \right)^\alpha \right] P(N = k) \\ &= \sum_{k=1}^\infty \left(E\left[\left(\sum_{i=1}^k Y_i \right)^\alpha - \sum_{i=1}^k Y_i^\alpha \right] + E\left[\sum_{i=1}^k Y_i^\alpha \right] \right) P(N = k) \\ &\leq \sum_{k=1}^\infty \left(k^\alpha E[Y^{p-1}]^{\alpha/(p-1)} + kE\left[Y^\alpha\right] \right) P(N = k) \\ &= \rho_\alpha E[W_{n-1}^\alpha] + E[N^\alpha](E[C^{p-1}])^{\alpha/(p-1)}(E[W_{n-1}^{p-1}])^{\alpha/(p-1)}. \end{aligned}$$

By Lemma 7.1,

$$\begin{aligned} E[W_n^\alpha] &\leq \rho_\alpha E[W_{n-1}^\alpha] + E[N^\alpha](E[C^{p-1}])^{\alpha/(p-1)}(K_{p-1}(\rho_{p-1} \vee \rho)^{n-1})^{\alpha/(p-1)} \\ &= \rho_\alpha E[W_{n-1}^\alpha] + K(\rho_{p-1} \vee \rho)^{(n-1)\gamma}, \end{aligned}$$

where $\gamma = \alpha/(p-1) > 1$. Finally, iterating the preceding bound $n-1$ times gives

$$\begin{aligned} E[W_n^\alpha] &\leq \rho_\alpha^n E[W_0^\alpha] + K \sum_{i=0}^{n-1} \rho_\alpha^i (\rho \vee \rho_{p-1})^{\gamma(n-1-i)} \\ &\leq E[W_0^\alpha](\rho \vee \rho_\alpha)^n + K \sum_{i=0}^{n-1} (\rho \vee \rho_\alpha)^{\gamma(n-1-i)+i} \\ &= E[Q^\alpha](\rho \vee \rho_\alpha)^n + K(\rho \vee \rho_\alpha)^{n-1} \sum_{i=0}^{n-1} (\rho \vee \rho_\alpha)^{(\gamma-1)i} \\ &\leq K_\alpha (\rho \vee \rho_\alpha)^n. \end{aligned}$$

This completes the proof.

7.2. The case when the C's dominate: Implicit renewal theory

In this section we give the proofs to Theorems 4.1 and Lemma 4.3.

Proof of Theorem 4.1. For any $k \in \mathbb{N}$ define $\Pi_k = \prod_{i=1}^k C_i$ and $V_k = \sum_{n=1}^k \log C_i$, where the C_i's are independent copies of C. Then, for any $t \in \mathbb{R}$,

$$\begin{aligned} P(R > e^t) &= \sum_{k=1}^n \left(m^{k-1} P(\Pi_{k-1} R > e^t) - m^k P(\Pi_k R > e^t) \right) + m^n P(\Pi_n R > e^t) \\ &= \sum_{k=1}^n \left(m^{k-1} P(e^{V_{k-1}} R > e^t) - m^k P(e^{V_{k-1}} CR > e^t) \right) + m^n P(e^{V_n} R > e^t) \\ &= \sum_{k=0}^{n-1} m^k \int_{-\infty}^\infty \left(P(R > e^{t-v}) - mP(CR > e^{t-v}) \right) P(V_k \in dv) + m^n P(e^{V_n} R > e^t). \end{aligned}$$

Next, define

$$\nu_n(dt) = e^{\alpha t} \sum_{k=0}^{n} m^k P(V_k \in dt), \qquad g(t) = e^{\alpha t}(P(R > e^t) - mP(CR > e^t)),$$

$$r(t) = e^{\alpha t} P(R > e^t) \qquad \text{and} \qquad \delta_n(t) = m^n P(e^{V_n} R > e^t).$$

Then, for any $t \in \mathbb{R}$ and $n \in \mathbb{N}$,

$$r(t) = (g * \nu_{n-1})(t) + \delta_n(t).$$

Next, define the operator

$$\breve{f}(t) = \int_{-\infty}^{t} e^{-(t-u)} f(u)\, du$$

and note that

$$\begin{aligned}
\breve{r}(t) &= \int_{-\infty}^{t} e^{-(t-u)} (g * \nu_{n-1})(u)\, du + \breve{\delta}_n(t) \\
&= \int_{-\infty}^{t} e^{-(t-u)} \int_{-\infty}^{\infty} g(u-v) \nu_{n-1}(dv)\, du + \breve{\delta}_n(t) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{t} e^{-(t-u)} g(u-v)\, du\, \nu_{n-1}(dv) + \breve{\delta}_n(t) \\
&= \int_{-\infty}^{\infty} \breve{g}(t-v)\, \nu_{n-1}(dv) + \breve{\delta}_n(t) \\
&= (\breve{g} * \nu_{n-1})(t) + \breve{\delta}_n(t).
\end{aligned} \tag{7.4}$$

Next, we will show that one can pass $n \to \infty$ in the preceding identity. To this end, let $\eta(du) = e^{\alpha u} m P(\log C \in du)$, and note that by assumption $\mu = E[C^\alpha \log C] \in (0, \infty)$, so $\eta(\cdot)$ is a nonarithmetic measure on $\mathbb{R}$ that places no mass at $-\infty$. Also,

$$\int_{-\infty}^{\infty} \eta(du) = mE[e^{\alpha \log C}] = mE[C^\alpha] = 1$$

and

$$\int_{-\infty}^{\infty} u\, \eta(du) = mE[e^{\alpha \log C} \log C] = mE[C^\alpha \log C] = m\mu$$

imply that $\eta(\cdot)$ is a probability measure with mean $m\mu$. Moreover,

$$\nu(dt) = \sum_{k=0}^{\infty} m^k e^{\alpha t} P(V_k \in dt)$$

is its renewal measure $\sum_{n=0}^{\infty} \eta^{*n}$. Since $m\mu \neq 0$, then $(|f| * \nu)(t) < \infty$ for all t whenever f is directly Riemann integrable. By (4.3) and Lemma 9.2 from [15], $\breve{g}$ is directly Riemann integrable, resulting in $(|\breve{g}| * \nu)(t) < \infty$ for all t. Thus, $(|\breve{g}| * \nu)(t) = E\left[\sum_{k=0}^{\infty} m^k e^{\alpha V_k} |\breve{g}(t - V_k)|\right] < \infty$. By Fubini's theorem, $E\left[\sum_{k=0}^{\infty} m^k e^{\alpha V_k} \breve{g}(t - V_k)\right]$ exists and

$$(\breve{g} * \nu)(t) = E\left[\sum_{k=0}^{\infty} m^k e^{\alpha V_k} \breve{g}(t - V_k)\right] = \sum_{k=0}^{\infty} E\left[m^k e^{\alpha V_k} \breve{g}(t - V_k)\right] = \lim_{n \to \infty} (\breve{g} * \nu_n)(t).$$

The fact that $\breve{\delta}_n(t) \to 0$ as $n \to \infty$ for all fixed t follows from

$$\begin{aligned}
\breve{\delta}_n(t) &= \int_{-\infty}^{t} e^{-(t-u)} m^n P(e^{V_n} R > e^u)\, du \\
&\leq e^{-t} \int_{-\infty}^{t} e^u m^n \frac{E[e^{\beta V_n} R^\beta]}{e^{\beta u}}\, du \qquad \text{(for some } 0 < \beta < 1) \\
&= \frac{e^{-\beta t}}{1-\beta} E[R^\beta] (m E[C^\beta])^n \to 0
\end{aligned}$$

as $n \to \infty$. Hence, the preceding arguments allow us to pass $n \to \infty$ in (7.4), and obtain

$$\breve{r}(t) = (\breve{g} * \nu)(t).$$

Now, by the key renewal theorem for two-sided random walks in [4],

$$e^{-t} \int_0^{e^t} v^\alpha P(R > v)\, dv = \breve{r}(t) \to \frac{1}{m\mu} \int_{-\infty}^{\infty} \breve{g}(u)\, du \triangleq H, \qquad t \to \infty,$$

while Lemma 9.3 in [15] implies

$$P(R > t) \sim H t^{-\alpha}, \qquad t \to \infty.$$

Finally,

$$\begin{aligned}
H &= \frac{1}{m\mu} \int_{-\infty}^{\infty} \int_{-\infty}^{t} e^{-(t-u)} g(u)\, du\, dt \\
&= \frac{1}{m\mu} \int_{-\infty}^{\infty} g(u)\, du \\
&= \frac{1}{m\mu} \int_0^\infty v^{\alpha-1} (P(R > v) - mP(CR > v))\, dv.
\end{aligned}$$

We end this section with the proof of Lemma 4.3.

Proof of Lemma 4.3. That the integral is positive follows from the union bound. That

$$\int_0^\infty \left(E[N]P(CR > t) - P\left(\max_{1\leq i\leq N} C_i R_i > t \right) \right) t^{\alpha-1} dt = \frac{1}{\alpha} E\left[\sum_{i=1}^{N} (C_i R_i)^\alpha - \left(\max_{1\leq i\leq N} C_i R_i \right)^\alpha \right]$$

follows from similar arguments to those used to derive the alternative expression for H in the proof of Theorem 4.2. The rest of the proof shows that the integral is finite.

Clearly

$$\int_0^1 \left(E[N]P(CR > t) - P\left(\max_{1\leq i\leq N} C_i R_i > t \right) \right) t^{\alpha-1}\, dt \leq E[N] \int_0^1 t^{\alpha-1}\, dt < \infty.$$

Hence, it remains to prove that the remaining part of the integral $\left(\int_1^\infty \cdots dt \right)$ is finite. To do this, we start by letting $Y = CR$ and $F(y) = P(Y \leq y)$. Then

$$\begin{aligned}
E[N]P(CR > t) - P\left(\max_{1\leq i\leq N} C_i R_i > t \right) &= \sum_{k=1}^{\infty} \left(F(t)^k - 1 + k\overline{F}(t) \right) P(N = k) \\
&= E\left[(1 - \overline{F}(t))^N - 1 + N\overline{F}(t) \right].
\end{aligned}$$

Use the inequality $1-x \leq e^{-x}$ for $x>0$ to obtain

$$E\left[(1-\overline{F}(t))^N - 1 + N\overline{F}(t)\right] \leq E\left[e^{-\overline{F}(t)N} - 1 + N\overline{F}(t)\right].$$

Choose $0<\delta<\alpha\epsilon/(1+\epsilon)$ (recall that $0<\epsilon<1$) and let $\beta=\alpha-\delta$. By Markov's inequality and Lemma 3.3

$$\overline{F}(t) \leq t^{-\beta} E[Y^\beta] = t^{-\beta} E[R^\beta] E[C^\beta] \triangleq c t^{-\beta} < \infty$$

for any $t>0$. Note that the function $h(x)=e^{-x}-1+x$ is increasing on $[0,\infty)$, so $h(N\overline{F}(t)) \leq h(cNt^{-\beta})$. By Fubini's Theorem (the integrand is nonnegative),

$$\int_1^\infty \left(E[N]P(CR>t) - P\left(\max_{1\leq i\leq N} C_i R_i > t\right)\right) t^{\alpha-1}\, dt \leq E\left[\int_1^\infty \left(e^{-cNt^{-\beta}} - 1 + Nt^{-\beta}\right) t^{\alpha-1}\, dt\right].$$

Using the change of variables $u=cNt^{-\beta}$ gives

$$\begin{aligned}\int_1^\infty \left(e^{-cNt^{-\beta}} - 1 + Nt^{-\beta}\right) t^{\alpha-1}\, dt &= (cN)^{\alpha/\beta} \int_0^{cN} \left(e^{-u}-1+u\right) u^{-\alpha/\beta-1}\, du \\ &\leq (cN)^{\alpha/\beta} \int_0^{\infty} \left(e^{-u}-1+u\right) u^{-\alpha/\beta-1}\, du.\end{aligned}$$

Our choice of $\beta=\alpha-\delta$ guarantees that $1<\alpha/\beta<1+\epsilon$, so $E[(cN)^{\alpha/\beta}]<\infty$. It only remains to show that the last (non-random) integral is finite. To see this note that $e^{-x}-1+x \leq x^2/2$ and $e^{-x}-1\leq 0$ for any $x\geq 0$, so

$$\begin{aligned}\int_0^\infty \left(e^{-u}-1+u\right) u^{-\alpha/\beta-1}\, du &\leq \frac{1}{2}\int_0^1 u^{1-\alpha/\beta}\, du + \int_1^\infty u^{-\alpha/\beta}\, du \\ &= \frac{1}{2(2-\alpha/\beta)} + \frac{1}{\alpha/\beta - 1} < \infty.\end{aligned}$$

This completes the proof.

7.3. The case when the N dominates

This section contains the proofs of Lemma 5.1 and Proposition 5.1. We also present in Lemma 7.2 a result for sums of iid truncated random variables that may be of independent interest in the context of heavy-tailed asymptotics, since it provides bounds that do not depend on the distribution of the summands. Most of the work involved in its proof of Proposition 5.1 goes into obtaining a bound for one iteration of the recursion satisfied by W_n, and for the convenience of the reader is presented separately in Lemma 7.3.

Proof of Lemma 5.1. We proceed by induction in n. For $n=1$ fix $\alpha/(\alpha+\epsilon)<\delta<1$ and note that

$$\begin{aligned}P(R_1>x) &= P\left(\sum_{i=1}^N C_i R_{0,i} + Q > x\right) \\ &= P\left(\sum_{i=1}^N C_i Q_i > x - Q,\, Q\leq x^\delta\right) + P\left(Q > x^\delta\right) \\ &\sim P\left(\sum_{i=1}^N C_i Q_i > x\right) + O\left(x^{-\delta(\alpha+\epsilon)}\right) \\ &\sim P(N > x/E[CQ]) + o(P(N>x)) \\ &\sim (E[C]E[Q])^\alpha P(N>x),\end{aligned}$$

where the fourth step is justified by Lemma 3.7(2) from [19]. Now suppose that we have

$$P(R_n > x) \sim \frac{(E[C]E[Q])^\alpha}{(1-\rho)^\alpha} \sum_{k=0}^{n} \rho_\alpha^k (1-\rho^{n-k})^\alpha P(N > x).$$

Note that since $E[C^{\alpha+\epsilon}] < \infty$, then by Lemma 4.2 from [19],

$$P(CR_n > x) \sim E[C^\alpha] P(R_n > x).$$

Let $c^{-1} = E[C^\alpha](E[C]E[Q])^\alpha (1-\rho)^{-\alpha} \sum_{k=0}^{n} \rho_\alpha^k (1-\rho^{n-k})^\alpha$, then

$$P(N > x) \sim c P(CR_n > x),$$

and by Lemma 3.7(5) from [19] we have

$$\begin{aligned} P(R_{n+1} > x) &= P\left(\sum_{i=1}^{N} C_i R_{n,i} + Q > x \right) \\ &\sim P\left(\sum_{i=1}^{N} C_i R_{n,i} > x \right) \\ &\sim (E[N] + c(E[CR_n])^\alpha) P(CR_n > x) \\ &\sim (E[N] + c(E[CR_n])^\alpha) c^{-1} P(N > x). \end{aligned}$$

Next, observing that $E[R_n] = \sum_{i=0}^{n} E[W_i] = E[Q]\sum_{i=0}^{n} \rho^i = E[Q](1-\rho^{n+1})/(1-\rho)$, we obtain

$$\begin{aligned} (E[N] + c(E[CR_n])^\alpha)c^{-1} &= \left(\rho_\alpha + \frac{E[R_n]^\alpha (1-\rho)^\alpha}{E[Q]^\alpha \sum_{k=0}^{n} \rho_\alpha^k (1-\rho^{n-k})^\alpha} \right) \frac{(E[C]E[Q])^\alpha}{(1-\rho)^\alpha} \sum_{k=0}^{n} \rho_\alpha^k (1-\rho^{n-k})^\alpha \\ &= \left(\rho_\alpha \sum_{k=0}^{n} \rho_\alpha^k (1-\rho^{n-k})^\alpha + (1-\rho^{n+1})^\alpha \right) \frac{(E[C]E[Q])^\alpha}{(1-\rho)^\alpha} \\ &= \frac{(E[C]E[Q])^\alpha}{(1-\rho)^\alpha} \sum_{k=0}^{n+1} \rho_\alpha^k (1-\rho^{n+1-k})^\alpha. \end{aligned}$$

Lemma 7.2 below is based on traditional heavy-tailed techniques such as those used in [27] and [8], to name some references. The reason why we need to give complete proofs here and cannot simply use existing results is our need to guarantee that the bounds do not depend on the distribution of the summands, which will be key when we apply them to W_n. The corollary we obtain from this lemma will be used in the proof of Lemma 7.3.

Lemma 7.2. *Suppose that $Y_1, Y_2, \ldots$ are nonnegative iid random variables with the same distribution as Y, where $E[Y^\beta] < \infty$ for some $\beta > 0$. Fix $0 < \epsilon < 1$. Then,*

1. *for $0 < \beta < 1$, $1 \le k \le x^\beta / E[Y^\beta]$, and $x \ge e^{(Ke)^{1/(1-\beta)}}$,*

$$P\left(\sum_{i=1}^{k} Y_i > x,\ \max_{1\le i\le k} Y_i \le x/\log x \right) \le e^{-(1-\beta)(\log x)(\log\log x)\left(1 - \frac{\log(eK)}{(1-\beta)\log\log x}\right)},$$

2. *for $\beta > 1$, $1 \le k \le (1-\epsilon)x/(E[Y] \vee E[Y^\beta])$, and $x \ge e^2 \vee (Ke/\epsilon)^{2/(\beta-1)}$,*

$$P\left(\sum_{i=1}^{k} Y_i > x,\ \max_{1\le i\le k} Y_i \le x/\log x \right) \le e^{-\epsilon(\beta-1)(\log x)^2\left(1 - \frac{\log\log x}{\log x} - \frac{\log(Ke/\epsilon)}{(\beta-1)\log x}\right) + (\beta-1)^2},$$

where $K > 1$ is a constant that does not depend on Y, ϵ or k.

Proof. Let $F(t) = P(Y \leq t)$, set $y = x/\log x$ and note that

$$P\left(\sum_{i=1}^{k} Y_i > x, \max_{1\leq i\leq k} Y_i \leq y\right) = P\left(\sum_{i=1}^{k} Y_i^{(y)} > x\right) F(y)^k,$$

where $P(Y^{(y)} \leq t) = F(t \wedge y)/F(y)$. Fix $\theta \geq 1/y$ and use the standard Chernoff's bound method for truncated heavy tailed sums (see, e.g. [27, 8]) to obtain

$$P\left(\sum_{i=1}^{k} Y_i^{(y)} > x\right) \leq e^{-\theta x} E\left[e^{\theta Y_1^{(y)}}\right]^k = e^{-\theta x} E\left[e^{\theta Y} 1_{(Y\leq y)}\right]^k F(y)^{-k}.$$

From where it follows that

$$P\left(\sum_{i=1}^{k} Y_i^{(y)} > x\right) F(y)^k \leq e^{-\theta x} E\left[e^{\theta Y} 1_{(Y\leq y)}\right]^k.$$

To analyze the preceding truncated exponential moment suppose first that $\beta > 1$. Then, by using the identity

$$E[Y^\eta] = \int_0^\infty \eta t^{\eta-1}\overline{F}(t)dt \tag{7.5}$$

we obtain

$$\begin{aligned} E\left[e^{\theta Y} 1_{(Y\leq y)}\right] &= \overline{F}(0) - e^{\theta y}\overline{F}(y) + \theta\int_0^y e^{\theta t}\overline{F}(t)\,dt \\ &\leq 1 + \theta\int_0^{1/\theta}\overline{F}(t)\,dt + \theta\int_0^{1/\theta}(e^{\theta t}-1)\overline{F}(t)\,dt + \theta\int_{1/\theta}^y e^{\theta t}\overline{F}(t)\,dt \\ &\leq 1 + \theta E[Y] + e\theta^2\int_0^{1/\theta} t\overline{F}(t)\,dt + \theta\int_{1/\theta}^y e^{\theta t}\overline{F}(t)\,dt \\ &\leq 1 + \theta E[Y] + \frac{e\theta^{2\wedge\beta}}{2\wedge\beta}E[Y^{2\wedge\beta}] + \theta\int_{1/\theta}^y e^{\theta t}\overline{F}(t)\,dt, \end{aligned} \tag{7.6}$$

where in the second inequality we use $e^x - 1 \leq xe^x$, $x \geq 0$, and in the last inequality we use $t^{2-(2\wedge\beta)} \leq \theta^{-2+(2\wedge\beta)}$ and (7.5) with $\eta = 2\wedge\beta$. Similarly, if $0 < \beta \leq 1$, then

$$\begin{aligned} E\left[e^{\theta Y} 1_{(Y\leq y)}\right] &\leq 1 + \theta\int_0^{1/\theta}\overline{F}(t)\,dt + \theta\int_{1/\theta}^y e^{\theta t}\overline{F}(t)\,dt \\ &\leq 1 + \frac{e\theta^\beta}{\beta}E[Y^\beta] + \theta\int_{1/\theta}^y e^{\theta t}\overline{F}(t)\,dt. \end{aligned} \tag{7.7}$$

Next, by Markov's inequality we have

$$\overline{F}(t) \leq E[Y^\beta]t^{-\beta},$$

which, in combination with (7.6) and (7.7), gives

$$E\left[e^{\theta Y} 1_{(Y\leq y)}\right] \leq \begin{cases} 1 + \theta E[Y] + \frac{e\theta^2}{2}E[Y^2] + E[Y^\beta]\theta\int_{1/\theta}^y e^{\theta t}t^{-\beta}dt, & \beta > 2, \\ 1 + \theta E[Y] + \frac{e\theta^\beta}{\beta}E[Y^\beta] + E[Y^\beta]\theta\int_{1/\theta}^y e^{\theta t}t^{-\beta}dt, & 1 < \beta \leq 2, \\ 1 + \frac{e\theta^\beta}{\beta}E[Y^\beta] + E[Y^\beta]\theta\int_{1/\theta}^y e^{\theta t}t^{-\beta}dt, & 0 < \beta \leq 1. \end{cases} \tag{7.8}$$

To analyze the remaining integral we split it as follows,

$$\begin{aligned}\theta\int_{1/\theta}^{y} e^{\theta t}t^{-\beta}dt &\leq \theta^{1+\beta}\int_{1/\theta}^{y/2} e^{\theta t}dt + \theta\int_{y/2}^{y} e^{\theta t}t^{-\beta}\,dt \\ &\leq \theta^{\beta}e^{\theta y/2} + \theta y^{1-\beta}\int_{1/2}^{1} e^{\theta y u}u^{-\beta}\,du \\ &\leq \theta^{\beta}e^{\theta y/2} + \theta y^{1-\beta}2^{\beta}\int_{1/2}^{1} e^{\theta y u}\,du \\ &\leq \theta^{\beta}e^{\theta y/2} + 2^{\beta}e^{\theta y}y^{-\beta},\end{aligned}$$

from where it follows that

$$\begin{aligned}&2e\theta^{\beta}E[Y^{\beta}] + E[Y^{\beta}]\theta\int_{1/\theta}^{y} e^{\theta t}t^{-\beta}dt \\ &\leq 2e\theta^{\beta}E[Y^{\beta}] + E[Y^{\beta}]\theta^{\beta}e^{\theta y/2} + E[Y^{\beta}]2^{\beta}e^{\theta y}y^{-\beta} \\ &\leq 2^{\beta}E[Y^{\beta}]e^{\theta y}y^{-\beta}\left(1 + e2^{1-\beta}(\theta y)^{\beta}e^{-\theta y} + 2^{-\beta}(\theta y)^{\beta}e^{-\theta y/2}\right) \\ &\leq 2^{\beta}E[Y^{\beta}]e^{\theta y}y^{-\beta}\left(1 + 2e\sup_{t\geq 1} t^{\beta}e^{-t} + \sup_{t\geq 1/2} t^{\beta}e^{-t}\right).\end{aligned}$$

Hence, we have shown that

$$2e\theta^{\beta}E[Y^{\beta}] + E[Y^{\beta}]\theta\int_{1/\theta}^{y} e^{\theta t}t^{-\beta}dt \leq KE[Y^{\beta}]e^{\theta y}y^{-\beta},$$

where $K = 2^{\beta}\left(1 + (2e+1)\sup_{t\geq 1/2} t^{\beta}e^{-t}\right)$ does not depend on Y or θ. Replacing the preceding inequality in (7.8) and using $1+t \leq e^{t}$ give,

$$e^{-\theta x}E\left[e^{\theta Y}1_{(Y\leq y)}\right]^{k} \leq \begin{cases} e^{-\theta(x-kE[Y])+ek\theta^{2}E[Y^{2}]+KkE[Y^{\beta}]e^{\theta y}y^{-\beta}}, & \beta > 2, \\ e^{-\theta(x-kE[Y])+KkE[Y^{\beta}]e^{\theta y}y^{-\beta}}, & 1 < \beta \leq 2, \\ e^{-\theta x+KkE[Y^{\beta}]e^{\theta y}y^{-\beta}}, & 0 < \beta \leq 1. \end{cases} \tag{7.9}$$

Now, to complete the proof, we optimize the choice of θ in the preceding bounds. For $0 < \beta < 1$, choose $\theta = \frac{1}{y}\log\left(\frac{x}{KkE[Y^{\beta}]y^{1-\beta}}\right)$ and note that for all $1 \leq k \leq x^{\beta}/E[Y^{\beta}]$ and $x \geq e^{(Ke)^{1/(1-\beta)}}$,

$$\theta y \geq \log\left(\frac{(\log x)^{1-\beta}}{K}\right) \geq 1.$$

Then,

$$\begin{aligned}e^{-\theta x+KkE[Y^{\beta}]e^{\theta y}y^{-\beta}} &= e^{-(\log x)\log\left(\frac{x^{\beta}(\log x)^{1-\beta}}{KekE[Y^{\beta}]}\right)} \\ &\leq e^{-(\log x)\log\left(\frac{(\log x)^{1-\beta}}{Ke}\right)} \\ &= e^{-(1-\beta)(\log x)(\log\log x)\left(1-\frac{\log(eK)}{(1-\beta)\log\log x}\right)}.\end{aligned}$$

Now, for $\beta > 1$, set $\theta = \frac{1}{y}\log\left(\frac{(x-kE[Y])y^{\beta-1}}{Kx}\right)$ and note that for and $x \geq e^{2} \vee (Ke/\epsilon)^{2/(\beta-1)}$,

$$\theta y \geq \log\left(\frac{\epsilon y^{\beta-1}}{K}\right) \geq \log\left(\frac{\epsilon x^{(\beta-1)/2}}{K}\right) \geq 1.$$

Then, for $1 < \beta \leq 2$ and all $1 \leq k \leq (1-\epsilon)x/(E[Y] \vee E[Y^\beta])$,

$$\begin{aligned} e^{-\theta(x-kE[Y])+KkE[Y^\beta]e^{\theta y}y^{-\beta}} &= e^{-\frac{(x-kE[Y])}{y}\log\left(\frac{(x-kE[Y])y^{\beta-1}}{Kx}\right)+kE[Y^\beta]\frac{(x-kE[Y])}{xy}} \\ &\leq e^{-\frac{(x-kE[Y])}{y}\log\left(\frac{\epsilon y^{\beta-1}}{Ke}\right)} \\ &\leq e^{-\epsilon(\beta-1)(\log x)^2\left(1-\frac{\log\log x}{\log x}-\frac{\log(Ke/\epsilon)}{(\beta-1)\log x}\right)}. \end{aligned}$$

In addition, for $\beta > 2$ note that

$$\sup_{x\geq e} ek\theta^2 E[Y^2] \leq \sup_{x\geq e} \frac{ex}{y^2}\left(\log\left(\frac{y^{\beta-1}}{K}\right)\right)^2 \leq \sup_{x\geq e} \frac{e(\beta-1)^2(\log x)^4}{x} = (\beta-1)^2.$$

Finally, by combining the preceding two bounds with the first two inequalities in (7.9), we derive

$$P\left(\sum_{i=1}^k Y_i > x,\ \max_{1\leq i\leq k} Y_i \leq y\right) \leq e^{-\epsilon(\beta-1)(\log x)^2\left(1-\frac{\log\log x}{\log x}-\frac{\log(Ke/\epsilon)}{(\beta-1)\log x}\right)+(\beta-1)^2}$$

for any $\beta > 1$.

As an immediate corollary to the preceding lemma we obtain:

Corollary 7.1. *Suppose that* $Y_1, Y_2, \ldots$ *are nonnegative iid random variables with the same distribution as* Y*, where* $E[Y^\beta] < \infty$ *for some* $\beta > 0$*. Then, for any* $\kappa > 0$ *there exists a constant* $x_0 > 0$ *that does not depend on* k *or the distribution of* Y *such that*

$$\sup_{1\leq k\leq m_\beta(x)} P\left(\sum_{i=1}^k Y_i > x,\ \max_{1\leq i\leq k} Y_i \leq x/\log x\right) \leq x^{-\kappa}$$

for all $x \geq x_0$*, where*

$$m_\beta(x) = \begin{cases} \frac{x^\beta}{E[Y^\beta]}, & 0 < \beta < 1, \\ \frac{(1-\epsilon)x}{E[Y]\vee E[Y^\beta]}, & \beta > 1, 0 < \epsilon < 1. \end{cases}$$

Lemma 7.3 below gives a bound for the distribution of W_{n+1} in terms of that of W_n. This lemma can also be used to prove the corresponding uniform bound for W_n in the case when the Q's dominate recursion (1.1).

Lemma 7.3. *Suppose that* $P(N > x) \leq x^{-\alpha}L(x)$*, with* $\alpha > 1$ *and* $L(\cdot)$ *slowly varying. Suppose* $E[N]\max\{E[C^\alpha], E[C]\} < 1$*. Then, for any* $c > 0$*,* $0 < \epsilon < 1$*,* $0 < \delta < 1 \wedge (\alpha-1)/2$*, and* $E[N]\max\{E[C^\alpha], E[C]\} < \eta < 1$ *there exist constants* $K = K(\delta, \epsilon, c, \eta) > 0$ *and* $x_0 = x_0(\delta, \epsilon, c, \eta) > 0$*, that do not depend on* n*, such that for all* $1 \leq n \leq c\log x/|\log\eta|$ *and all* $x \geq x_0$*,*

$$P(W_{n+1} > x) \leq K\eta^{(2\wedge(\alpha-\delta))n}x^{-\alpha}L(x) + E[N]P(CW_n > (1-\epsilon)x).$$

REMARK: Note that condition $E[N]\max\{E[C^\alpha], E[C]\} < 1$ is natural since it is needed for the finiteness of $E[R^\beta]$ for any $\beta < \alpha$. It is also in agreement with Lemma 5.1 in the sense that it is a necessary condition for the convergence (as $n \to \infty$) of the sum appearing in (5.1). The choice of η is also suggested by the fact that for $\beta < \alpha$ one can obtain a weaker uniform bound by applying the moment estimate on $E[W_n^\beta]$ from Lemma 3.2, i.e., $P(W_n > x) \leq E[W_n^\beta]x^{-\beta} \leq K_\beta(E[N]\max\{E[C], E[C^\beta]\})^n x^{-\beta}$.

Before going into the proof, we would like to emphasize that special care goes into making sure that K and x_0 in the statement of the lemma do not depend on n. This is important since Lemma 7.3 will be applied iteratively in the proof of Proposition 5.1, where one does not want K and x_0 to grow from one iteration to the next.

Proof of Lemma 7.3. By convexity of $f(\theta) = E[C^\theta]$, $\max\{E[C^\alpha], E[C]\} \geq \max\{E[C^{\alpha-\delta}], E[C]\}$, implying

$$\epsilon' \triangleq \frac{\eta}{E[N]\max\{E[C^{\alpha-\delta}], E[C]\}} - 1 > 0.$$

Next, recall that $W_{n+1} \overset{\mathcal{D}}{=} \sum_{i=1}^N C_i W_{n,i}$ where $W_{n,i}$ are iid copies of W_n, let $Y \overset{\mathcal{D}}{=} Y_i = C_i W_{n,i}$ and $\beta = \alpha - \delta > 1$. Note that by Lemma 3.2 there exists a constant $K_1 > 0$ (that does not depend on n) such that,

$$\begin{aligned} E[Y^\beta] &= E[C^\beta]E[W_n^\beta] \\ &\leq K_1 (E[N]\max\{E[C^{\alpha-\delta}], E[C]\})^n \\ &= K_1(1+\epsilon')^{-n}\eta^n, \end{aligned} \tag{7.10}$$

where the last equality comes from the definition of ϵ'. And since $E[Y] = E[Q](E[N]E[C])^n \leq E[Q](E[N]\max\{E[C^{\alpha-\delta}], E[C]\})^n$, then

$$E[Y^\beta] \vee E[Y] \leq K_2(1+\epsilon')^{-n}\eta^n \tag{7.11}$$

for some constant $K_2 > 0$ that does not depend on n. With the intent of applying Corollary 7.1, we define $y = \epsilon x$, and

$$m_n(x) \triangleq \lfloor \epsilon^2 x / (E[Y^\beta] \vee E[Y]) \rfloor.$$

Then,

$$\begin{aligned} P(W_{n+1} > x) &= P\left(\sum_{i=1}^N Y_i > x\right) \\ &\leq P\left(\sum_{i=1}^N Y_i > x,\, N \leq m_n(x)\right) + P(N > m_n(x)) \\ &\leq P\left(\sum_{i=1}^N Y_i > x,\, M_N^{(N)} \leq (1-\epsilon)x,\, N \leq m_n(x)\right) \\ &\quad + P\left(M_N^{(N)} > (1-\epsilon)x,\, N \leq m_n(x)\right) + P(N > m_n(x)) \\ &\leq P\left(\sum_{i=1}^N Y_i > x,\, M_N^{(N)} \leq (1-\epsilon)x,\, M_N^{(N-1)} \leq y/\log y,\, N \leq m_n(x)\right) && (7.12) \\ &\quad + P\left(M_N^{(N-1)} > y/\log y,\, N \leq m_n(x)\right) && (7.13) \\ &\quad + P\left(M_N^{(N)} > (1-\epsilon)x,\, N \leq m_n(x)\right) + P(N > m_n(x)), && (7.14) \end{aligned}$$

where $M_k^{(i)}$ is the ith order statistic of $\{Y_1, \ldots, Y_k\}$, with $M_k^{(k)}$ being the largest. Note that the term in (7.12) can be bounded as follows

$$\begin{aligned} &P\left(\sum_{i=1}^N Y_i > x,\, M_N^{(N)} \leq (1-\epsilon)x,\, M_N^{(N-1)} \leq y/\log y,\, N \leq m_n(x)\right) \\ &\leq P\left(\sum_{i=1}^N Y_i - M_N^{(N)} > y,\, M_N^{(N-1)} \leq y/\log y,\, N \leq m_n(x)\right) \\ &\leq P\left(\sum_{i=1}^N Y_i > y,\, M_N^{(N)} \leq y/\log y,\, N \leq m_n(x)\right) \\ &\leq P\left(\sum_{i=1}^{m_n(x)} Y_i > y,\, \max_{1\leq i< m_n(x)} Y_i \leq y/\log y\right). \end{aligned}$$

Fix $\nu = \alpha + \delta + c(\alpha - \delta)$, then, by Corollary 7.1, there exists a constant $x_1 \geq e$, that does not depend on the distribution of Y (and therefore, does not depend on n), such that

$$\begin{aligned}
P\left(\sum_{i=1}^{m_n(x)} Y_i > y, \max_{1 \leq i < m_n(x)} Y_i \leq y/\log y \right) &\leq y^{-\nu} \\
&= \epsilon^{-\nu} \eta^{\frac{c(\alpha-\delta)}{|\log \eta|} \cdot \log x} x^{-\alpha-\delta} \\
&\leq \epsilon^{-\nu} \eta^{(\alpha-\delta)n} x^{-\alpha-\delta} \\
&\leq \epsilon^{-\nu} \sup_{t \geq 1} \frac{t^{-\delta}}{L(t)} \eta^{\beta n} x^{-\alpha} L(x)
\end{aligned}$$

for all $y \geq x_1$, where the second inequality follows from the assumption $n \leq c \log x / |\log \eta|$, and in the last inequality we use the definition $\beta = \alpha - \delta$. To bound (7.13), we condition on N,

$$\begin{aligned}
P\left(M_N^{(N-1)} > y/\log y, N \leq m_n(x) \right) &= \sum_{k=1}^{m_n(x)} P\left(M_k^{(k-1)} > y/\log y \right) P(N = k) \\
&\leq \sum_{k=1}^{m_n(x)} \binom{k}{2} P(Y > y/\log y)^2 P(N = k) \\
&\leq E\left[N^2 1_{(N \leq m_n(x))} \right] P(Y > y/\log y)^2 \\
&\leq E\left[N^{2\wedge\beta} \right] m_n(x)^{(2-\beta)^+} P(Y > y/\log y)^2,
\end{aligned}$$

where in the last inequality we use $N \leq m_n(x)$ in case N does not have a second moment. Now, by Markov's inequality and the definition of $m_n(x)$,

$$\begin{aligned}
m_n(x)^{(2-\beta)^+} P(Y > y/\log y)^2 &\leq m_n(x)^{(2-\beta)^+} \left(\frac{E[Y^\beta](\log y)^\beta}{y^\beta} \right)^2 \\
&\leq \left(\frac{E[Y^\beta]}{E[Y^\beta] \vee E[Y]} \right)^{(2-\beta)^+} \frac{\epsilon^{(2-\beta)^+} E[Y^\beta]^{2\wedge\beta} (\log y)^{2\beta}}{y^{2\beta \wedge (3\beta-2)}} \\
&\leq \frac{\epsilon^{(2-\beta)^+} E[Y^\beta]^{2\wedge\beta} (\log y)^{2\beta}}{y^{2\beta \wedge (3\beta-2)}} \\
&\leq \frac{\epsilon^{(2-\beta)^+} (K_1 (1+\epsilon')^{-n} \eta^n)^{2\wedge\beta} (\log y)^{2\beta}}{y^{2\beta \wedge (3\beta-2)}} \qquad \text{(by (7.10))}.
\end{aligned}$$

Our choice of δ guarantees that $2\beta \wedge (3\beta - 2) > \alpha + \delta$ and $\beta = \alpha - \delta > 1$, and therefore,

$$\begin{aligned}
P\left(M_N^{(N-1)} > y/\log y, N \leq m_n(x) \right) &\leq K_3 \frac{\eta^{(2\wedge\beta)n}}{(1+\epsilon')^{(2\wedge\beta)n}} x^{-\alpha-\delta} \\
&\leq K_3 \sup_{t \geq 1} \frac{t^{-\delta}}{L(t)} \eta^{(2\wedge\beta)n} x^{-\alpha} L(x)
\end{aligned}$$

for all $x \geq x_2 = \epsilon^{-1} e$, where

$$K_3 = K_3(\epsilon, \delta) = E\left[N^{2\wedge\beta} \right] \epsilon^{(2-\beta)^+ - \alpha - \delta} K_1^{2\wedge\beta} \sup_{t \geq e} \frac{(\log t)^{2\beta}}{t^{2\beta \wedge (3\beta-2) - \alpha - \delta}}.$$

To bound (7.14), we first note that by Potter's Theorem (see Theorem 1.5.6 (ii) on p. 25 in [7]), there exists a constant $x_3 = x_3(\epsilon', \delta)$ such that

$$\begin{aligned}\frac{P(N > m_n(x))}{P(N > x)} &\leq (1+\epsilon') \max\left\{ \left(\frac{m_n(x)}{x}\right)^{-\alpha+\delta}, \left(\frac{m_n(x)}{x}\right)^{-\alpha-\delta} \right\} \\ &= (1+\epsilon') \max\left\{ \left(\frac{E[Y^\beta] \vee E[Y]}{\epsilon^2}\right)^{\alpha-\delta}, \left(\frac{E[Y^\beta] \vee E[Y]}{\epsilon^2}\right)^{\alpha+\delta} \right\} \\ &\leq \frac{(1+\epsilon')}{\epsilon^{2(\alpha+\delta)}} (E[Y^\beta] \vee E[Y])^\beta \\ &\leq \frac{K_2^\beta}{\epsilon^{2(\alpha+\delta)}} \cdot \frac{\eta^{\beta n}}{(1+\epsilon')^{\beta n - 1}} \qquad \text{(by (7.11))}.\end{aligned}$$

Now, from the last estimate, it follows that

$$P(N > m_n(x)) \leq \frac{K_2^\beta}{\epsilon^{2(\alpha+\delta)}} \cdot \eta^{\beta n} P(N > x) \leq K_4 \eta^{\beta n} x^{-\alpha} L(x).$$

Finally, for the second term in (7.14),

$$\begin{aligned}P\left(M_N^{(N)} > (1-\epsilon)x, \, N \leq m_n(x)\right) &\leq P\left(M_N^{(N)} > (1-\epsilon)x\right) \\ &\leq E[N] P(Y > (1-\epsilon)x).\end{aligned}$$

Combining the preceding bounds for (7.12) - (7.14) and setting $x_0 = \max\{x_1, x_2, x_3\}$ and $K = (\epsilon^{-\nu} + K_3) \sup_{t \geq 1} \frac{t^{-\delta}}{L(t)} + K_4$ completes the proof.

Finally, we give the proof of Proposition 5.1, the main technical contribution of Section 5.

Proof of Proposition 5.1. Note that it is enough to prove the proposition for all $x \geq x_0$ for some $x_0 = x_0(\eta, \nu) > 1$, since for all $1 \leq x \leq x_0$ and $n \geq 1$,

$$\begin{aligned}P(W_n > x) &= \frac{P(W_n > x)}{\eta^n P(N > x)} \eta^n P(N > x) \\ &\leq \frac{E[Q](E[N]E[C])^n x^{-1}}{\eta^n P(N > x)} \eta^n P(N > x) \qquad \text{(by Markov's inequality)} \\ &\leq \sup_{1 \leq t \leq x_0} \frac{E[Q]}{t P(N > t)} \cdot \eta^n P(N > x).\end{aligned}$$

Next, choose $0 < \epsilon < 1$ such that

$$E[N]E[C^\alpha]\left((1-\epsilon)^{-\alpha-1} + 2\epsilon\right) \leq \eta, \tag{7.15}$$

define $c = \nu/2$,

$$\gamma = \frac{1}{|\log \eta|} \log\left(\frac{\eta}{E[N]\max\{E[C^\alpha], E[C]\}}\right),$$

and select $0 < \delta < \min\{1, (\alpha-1)/2, c\gamma\}$. Now, by Lemma 7.3, there exist constants $K_1, x_1 > 0$ (that do not depend on n) such that

$$P(W_{n+1} > x) \leq K_1 \eta^{(2 \wedge (\alpha-\delta))n} P(N > x) + E[N] P(C W_n > (1-\epsilon)x)$$

for all $x \geq x_1$. Hence, by defining $n_0 = (2 \wedge (\alpha-\delta) - 1)^{-1} (\log \eta)^{-1} \log(\epsilon E[N] E[C^\alpha])$, we obtain

$$P(W_{n+1} > x) \leq K_1 E[N] E[C^\alpha] \epsilon \eta^n P(N > x) + E[N] P(C W_n > (1-\epsilon)x) \tag{7.16}$$

for all $n \geq n_0$, and all $x \geq x_1$.

Next, in order to derive an explicit bound for $P(W_n > x)$, we need the following two estimates (7.17) and (7.18). In this regard, choose $x_0 \geq 1 \vee x_1$ such that

$$P(CN > (1-\epsilon)x) \leq E[C^\alpha](1-\epsilon)^{-\alpha-1} P(N > x) \tag{7.17}$$

for all $x \geq x_0$. This is possible since by Lemma 4.2 from [19] $P(CN > (1-\epsilon)x) \sim E[C^\alpha](1-\epsilon)^{-\alpha} P(N > x)$. Also, by Markov's inequality, we have that for all $1 \leq n \leq c \log x / |\log \eta|$,

$$\begin{aligned} P(C > (1-\epsilon)x/x_0) &\leq E[C^{\alpha+\nu}](1-\epsilon)^{-\alpha-\nu} x_0^{\alpha+\nu} x^{-\alpha-\nu} \\ &\leq \frac{E[C^{\alpha+\nu}] x_0^{\alpha+\nu}}{(1-\epsilon)^{\alpha+\nu} x^{\nu/2} L(x)} \, \eta^n P(N > x), \end{aligned} \tag{7.18}$$

where in the second inequality we use $x^{-\nu/2} = x^{-c} = \eta^{\frac{c \log x}{|\log \eta|}} \leq \eta^n$. Now, define

$$K_2 = \max \left\{ 1, \, K_1, \, \sup_{x \geq x_0} \frac{E[C^{\alpha+\nu}] x_0^{\alpha+\nu}}{\epsilon E[C^\alpha](1-\epsilon)^{\alpha+\nu} x^{\nu/2} L(x)} \right\}.$$

Now we proceed to derive bounds for $P(W_n > x)$ for different ranges of n. For all $1 \leq n \leq n_0$ and all $x \geq x_0$, by Lemma 5.1, there exists a constant $K_0 \geq K_2$ such that

$$P(W_n > x) \leq K_0 \, \eta^n P(N > x). \tag{7.19}$$

Next, for the values $n_0 \leq n \leq c \log x / |\log \eta|$ we proceed by induction using (7.16). To this end, suppose (7.19) holds for some n in the specified range. Then, note that by (7.18) and the induction hypothesis (7.19), we have for all $x \geq x_0$,

$$\begin{aligned} P(CW_n > (1-\epsilon)x) &\leq P(CW_n > (1-\epsilon)x, C \leq (1-\epsilon)x/x_0) + P(C > (1-\epsilon)x/x_0) \\ &\leq \int_0^{(1-\epsilon)x/x_0} P(W_n > (1-\epsilon)x/y) P(C \in dy) + K_2 E[C^\alpha] \epsilon \eta^n P(N > x) \\ &\leq K_0 \eta^n \int_0^\infty P(N > (1-\epsilon)x/y) P(C \in dy) + K_2 E[C^\alpha] \epsilon \eta^n P(N > x) \\ &= K_0 \eta^n P(CN > (1-\epsilon)x) + K_2 E[C^\alpha] \epsilon \eta^n P(N > x) \\ &\leq K_0 E[C^\alpha] \left((1-\epsilon)^{-\alpha-1} + \epsilon \right) \eta^n P(N > x), \end{aligned}$$

where in the last inequality we used (7.17). Then, by replacing the preceding bound in (7.16) and using (7.15), we derive

$$\begin{aligned} P(W_{n+1} > x) &\leq K_0 E[N] E[C^\alpha] \left((1-\epsilon)^{-\alpha-1} + 2\epsilon \right) \eta^n P(N > x) \\ &\leq K_0 \eta^{n+1} P(N > x) \end{aligned}$$

for all $x \geq x_0$ and all $1 \leq n \leq c \log x / |\log \eta|$.

Finally, for $n \geq c \log x / |\log \eta|$, we follow a different approach that comes from our moment estimates for W_n. Let

$$\epsilon' = \frac{\eta}{E[N] \max\{E[C^\alpha], E[C]\}} - 1 > 0$$

and note that by convexity

$$E[N] \max\{E[C^{\alpha-\delta}], E[C]\} \leq E[N] \max\{E[C^\alpha], E[C]\} = (1+\epsilon')^{-1} \eta.$$

Then, by Markov's inequality and Lemma 3.2, we have

$$\begin{aligned} P(W_n > x) &\leq E[W_n^{\alpha-\delta}] x^{-\alpha+\delta} \\ &\leq K_{\alpha-\delta} (E[N] \max\{E[C^{\alpha-\delta}], E[C]\})^n x^{-\alpha+\delta} \\ &= K_{\alpha-\delta} (1+\epsilon')^{-n} \eta^n x^{-\alpha+\delta} \\ &\leq K_{\alpha-\delta} x^{-\log(1+\epsilon')c/|\log\eta|} \eta^n x^{-\alpha+\delta} \end{aligned} \tag{7.20}$$

for all $x > 0$. Note that the preceding bound,

$$\frac{\log(1+\epsilon')}{|\log\eta|} = \frac{1}{|\log\eta|} \log\left(\frac{\eta}{E[N]\max\{E[C^\alpha], E[C]\}} \right) = \gamma,$$

and (7.20) yield

$$P(W_n > x) \leq K_{\alpha-\delta} \eta^n x^{-c\gamma-\alpha+\delta} \leq K_{\alpha-\delta} \eta^n x^{-\alpha+\delta-c\gamma} \leq K_{\alpha-\delta} \sup_{t\geq 1} \frac{t^{\delta-c\gamma}}{L(t)} \, \eta^n P(N > x)$$

for all $x \geq 1$; recall that $\delta < c\gamma$. Thus, setting $K = \max\{K_0, K_{\alpha-\delta} \sup_{t\geq 1} t^{\delta-c\gamma}(L(t))^{-1}\}$ completes the proof.

7.4. The case when the Q's dominate

We end the paper with the proof of Lemma 6.1 and a sketch of the proof of Proposition 6.1. As mentioned before, the proofs of the other results presented in Section 6 have been omitted since they are very similar to those from Section 5.

Proof of Lemma 6.1. We proceed by induction in n. By Lemma 4.2 from [19],

$$P(CQ > x) \sim E[C^\alpha] P(Q > x),$$

by Lemma 3.7(1) from the same source,

$$P\left(\sum_{i=1}^N C_i Q_i > x \right) \sim E[N] P(CQ > x) \sim E[N] E[C^\alpha] P(Q > x),$$

and by Lemma 3.1, again from the same source, we have

$$\begin{aligned} P(R_1 > x) &= P\left(\sum_{i=1}^N C_i Q_i + Q > x \right) \\ &\sim P\left(\sum_{i=1}^N C_i Q_i > x \right) + P(Q > x) \\ &\sim (\rho_\alpha + 1) P(Q > x). \end{aligned}$$

Now suppose that we have

$$P(R_n > x) \sim \sum_{k=0}^n \rho_\alpha^k \, P(Q > x).$$

Then,

$$\begin{aligned}
P(R_{n+1} > x) &= P\left(\sum_{i=1}^{N} C_i R_{n,i} + Q > x\right) \\
&\sim P\left(\sum_{i=1}^{N} C_i R_{n,i} > x\right) + P(Q > x) \\
&\sim E[N]E[C^\alpha]P(R_n > x) + P(Q > x) \\
&\sim \left(\rho_\alpha \sum_{k=0}^{n} \rho_\alpha^k + 1\right) P(Q > x) \\
&= \sum_{k=0}^{n+1} \rho_\alpha^k \, P(Q > x).
\end{aligned}$$

Sketch of the proof of Proposition 6.1. By Markov's inequality

$$P(N > x) \leq E[N^{\alpha+\nu}]x^{-\alpha-\nu}$$

for all $x > 0$. Use Lemma 7.3 to obtain

$$P(W_{n+1} > x) \leq K_1 E[N]E[C^\alpha]\epsilon\eta^n P(Q > x) + E[N]P(CW_n > (1-\epsilon)x)$$

for all $n_0 \leq n \leq \kappa \log x$ and all $x \geq x_1$ (for suitably chosen constants ϵ, n_0, κ). Choose $x_0 \geq 1 \vee x_1$ such that

$$P(CQ > (1-\epsilon)x) \leq E[C^\alpha](1-\epsilon)^{-\alpha-1}P(Q > x).$$

The rest of the proof continues as in Proposition 5.1 with some modifications.

Acknowledgement

The authors are grateful to Professor Charles Goldie for pointing out a reference.

References

[1] G. Alsmeyer and D. Kuhlbusch. Double martingale structure and existence of ϕ-moments for weighted branching processes. *Münster Journal of Mathematics*, 1, 2008.

[2] G. Alsmeyer and U. Rösler. A stochastic fixed point equation related to weighted branching with deterministic weights. *Electron. J. Probab.*, 11:27–56, 2005.

[3] S. Asmussen. Subexponential asymptotics for stochastic processes: Extremal behavior, stationary distributions and first passage probabilities. *Ann. Appl. Probab.*, 8(2):354–474, 1998.

[4] K. B. Athreya, D. McDonald, and P. Ney. Limit theorems for semi-Markov processes and renewal theory for Markov chains. *Ann. Probab.*, 6(5):788–797, 1978.

[5] K. B. Athreya and P. E. Ney. *Branching Processes.* Dover, New York, 2004.

[6] A. Baltrūnas, D.J. Daley, and C. Klüppelberg. Tail behavior of the busy period of a GI/GI/1 queue with subexponential service times. *Stochastic Process. Appl.*, 111(2):237–258, 2004.

[7] N. H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular variation.* Cambridge University Press, Cambridge, 1987.

[8] A. Borovkov. Estimates for the distribution of sums and maxima of sums of random variables without the Cramér condition. *Siberian Math J.*, 41(5):997–1038, 2000.

[9] A. Brandt. The stochastic equation $y_{n+1} = a_n y_n + b_n$ with stationary coefficients. *Adv. Appl. Prob.*, 18:211–220, 1986.

[10] S. Brin and L. Page. The anatomy of a large-scale hypertextual Web search engine. *Comput. Networks ISDN Systems*, 30(1-7):107–117, 1998.

[11] Y.S. Chow and H. Teicher. *Probability Theory.* Springer-Verlag, New York, 1988.

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

[13] D. Denisov, S. Foss, and D. Korshunov. Asymptotics of randomly stopped sums in the presence of heavy tails. *arXiv:0808.3697v3.*

[14] J.A. Fill and S. Janson. Approximating the limiting Quicksort distribution. *Random Structures Algorithms*, 19(3-4):376–406, 2001.

[15] C. M. Goldie. Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.*, 1(1):126–166, 1991.

[16] A. M. Iksanov. Elementary fixed points of the BRW smoothing transforms with infinite number of summands. *Stochastic Process. Appl.*, 114:27–50, 2004.

[17] P.R. Jelenković and P. Momčilović. Large deviations of square root insensitive random sums. *Math. Oper. Res.*, 29(2):398–406, 2004.

[18] P.R. Jelenković and J. Tan. Modulated branching processes, origins of power laws and queueing duality. *arXiv: 0709.4297*, 2007.

[19] A. H. Jessen and T. Mikosch. Regularly varying functions. *Publications de L'Institut Mathematique, Nouvelle Serie*, 80(94):171–192, 2006.

[20] H. Kesten. Random difference equations and renewal theory for products of random matrices. *Acta Math.*, 131:207–248, 1973.

[21] J. Kleinberg. Authoritative sources in a hyperlinked environment. *ACM-SIAM Symposium on Discrete Algorithms*, 1998.

[22] D. Kuhlbusch. On weighted branching processes in random environment. *Stochastic Process. Appl.*, 109:113–144, 2004.

[23] N. Litvak, W.R.W. Scheinhardt, and Y. Volkovich. In-degree and PageRank: Why do they follow similar power laws? *Internet Math.*, 4(2-3):175–198, 2007.

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

[25] Q. Liu. On generalized multiplicative cascades. *Stochastic Process. Appl.*, 86:263–286, 2000.

[26] T. Mikosch and G. Samorodnitsky. The supremum of a negative drift random walk with dependent heavy-tailed steps. *Ann. Appl. Probab.*, 10(3):1025–1064, 2000.

[27] S.V. Nagaev. On the asymptotic behavior of one-sided large deviation probabilities. *Theory Probab. Appl.*, 26(2):362–366, 1982.

[28] U. Rösler. The weighted branching process. *Dynamics of complex and irregular systems (Bielefeld, 1991)*, pages 154–165, 1993. Bielefeld Encounters in Mathematics and Physics VIII, World Science Publishing, River Edge, NJ.

[29] U. Rösler and L. Rüschendorf. The contraction method for recursive algorithms. *Algorithmica*, 29(1-2):3–33, 2001.

[30] U. Rösler, V.A. Topchii, and V.A. Vatutin. Convergence conditions for the weighted branching process. *Discrete Math. Appl.*, 10(1):5–21, 2000.

[31] Y. Volkovich. Stochastic analysis of web page ranking. Ph.D. Thesis, University of Twente, 2009.

[32] Y. Volkovich and N. Litvak. Asymptotic analysis for personalized web search. *Memorandum 1884, Department of Applied Mathematics, University of Twente*, 2008.

[33] Y. Volkovich, N. Litvak, and D. Donato. Determining factors behind the pagerank log-log plot. *Proceedings of the 5th International Workshop on Algorithms and Models for the Web-Graph, WAW 2007*, 2007.

[34] B. Zwart. Tail asymptotics for the busy period in the GI/G/1 queue. *Math. Oper. Res.*, 26(3):485–493, 2001.