REGULAR VARIATION IN A FIXED-POINT PROBLEM FOR SINGLE- AND MULTI-CLASS BRANCHING PROCESSES AND QUEUES

BY SØREN ASMUSSEN AND SERGEY FOSS

Abstract

Tail asymptotics of the solution R to a fixed-point problem of the type $R \overset{\mathrm{D}}{=} Q + \sum_1^N R_m$ are derived under heavy-tailed conditions allowing both dependence between Q and N and the tails to be of the same order of magnitude. Similar results are derived for a K-class version with applications to multi-type branching processes and busy periods in multi-class queues.

Keywords: Busy period; Galton–Watson process; intermediate regular variation; multivariate regular variation; random recursion; random sum

2010 Mathematics Subject Classification: Primary 60H25
Secondary 60J80; 60K25; 60F10

1. Introduction

In this paper we study the tail asymptotics of the solution R to the fixed-point problem

$$R \overset{\mathrm{D}}{=} Q + \sum_{m=1}^{N} R_m \tag{1}$$

under suitable regular variation (RV) conditions and the similar problem in a multidimensional setting stated below at (6). Here in (1), Q and N are (possibly dependent) nonnegative, nondegenerate random variables, N being integer valued, $R_1, R_2, \ldots$ are independent and identically distributed (i.i.d.) and distributed as R, and $\bar{n} = \mathbb{E}[N] < 1$ (similar notation for expected values is used in the following).

In a classical example R is the M/G/1 busy period (cf. [9] and [28]), where Q is the service time of the first customer in the busy period and N the number of arrivals during his service. Here Q and N are indeed heavily dependent, with tails of the same order of magnitude when Q has a regularly varying distribution; more precisely, N is Poisson(λq) given $Q = q$. Another example is the total progeny of a subcritical branching process, where $Q \equiv 1$ and N is the number of children of the ancestor. More generally, R could be the total life span of the individuals in a Crump–Mode–Jagers process [19], corresponding to Q being the lifetime of the ancestor and N the number of her children. Related examples are weighted branching processes (see [20] for references). Note that connections between branching processes and RV have a long history; for some early work, see [5], [6], [24], and [25].

Recall some definitions of classes of heavy-tailed distributions. A distribution F on the real line is *long tailed*, $F \in \mathcal{L}$, if, for some $y > 0$,

$$\frac{\bar{F}(x+y)}{\bar{F}(x)} \to 1 \quad \text{as } x \to \infty; \tag{2}$$

*doi:*10.1017/apr.2018.69 © *Applied Probability Trust* 2018

https://doi.org/10.1017/apr.2018.69 Published online by Cambridge University Press

a function $L \geq 0$ is *slowly varying* at ∞ if $L(\alpha x)/L(x) \to 1$ for all finite $\alpha > 0$; F is *regularly varying*, $F \in \mathcal{RV}$, if, for some $\beta > 0$, $\bar{F}(x) = x^{-\beta} L(x)$, where $L(x)$ is slowly varying at ∞; F is *intermediate regularly varying*, $F \in \mathcal{IRV}$, if

$$\lim_{\alpha \uparrow 1} \limsup_{x \to \infty} \frac{\bar{F}(\alpha x)}{\bar{F}(x)} = 1. \tag{3}$$

It is known that $\mathcal{L} \supset \mathcal{IRV} \supset \mathcal{RV}$, and if F has a finite mean then $\mathcal{L} \supset \mathcal{S}^* \supset \mathcal{IRV}$, where $\mathcal{S}^*$ is the class of so-called *strong subexponential distributions*; see, e.g. [13] or [18] for further definitions and properties of heavy-tailed distributions.

Tail asymptotics of quantities related to R have been studied earlier in [20] and [27] under RV conditions (see also [7]). Our main result is the following.

Theorem 1. *Assume that $\bar{n} < 1$ and $\bar{q} < \infty$. Then*

(i) *there is only one nonnegative solution R to (1) with finite mean; for this solution, $\bar{r} = \bar{q}/(1 - \bar{n})$;*

(ii) *if further*

(C) *the distribution of $Q + cN$ is intermediate regularly varying for all $c > 0$ in the interval $(\bar{r} - \epsilon, \bar{r} + \epsilon)$, where $\bar{r}$ is as in (i) and $\epsilon > 0$ is any small number,*

then

$$\mathbb{P}(R > x) \sim \frac{1}{1 - \bar{n}} \mathbb{P}(Q + \bar{r}N > x) \quad \textit{as } x \to \infty; \tag{4}$$

(iii) *in particular, condition (C) holds in the following three cases:*

(a) *(Q, N) has a two-dimensional regularly varying distribution;*

(b) *Q has an intermediate regularly varying distribution and $\mathbb{P}(N > x) = o(\mathbb{P}(Q > x))$;*

(c) *N has an intermediate regularly varying distribution and $\mathbb{P}(Q > x) = o(\mathbb{P}(N > x))$.*

Part (i) is well known from several sources and not deep (see the proof of the more general Proposition 1 below and the references at the end of the section for more general versions). Part (ii) generalizes and unifies results of [20] and [27] in several ways. Motivated by Google's PageRank algorithm, both of these papers consider the more general recursion

$$R \stackrel{\mathrm{D}}{=} Q + \sum_{m=1}^{N} A_m R_m. \tag{5}$$

However, [20] does not allow dependence and/or the tails of Q and N to be equally heavy. These features are incorporated in [27], but, on the other hand, that paper requires strong conditions on the A_i which do not allow us to take $A_i \equiv 1$ when dealing with sharp asymptotics. To remove all of these restrictions is essential for the applications to queues and branching processes that we have in mind. Also, our proofs are considerably simpler and shorter than those in [20] and [27]. The key tool is a general result of [17] giving the tail asymptotics of the maximum of a random walk up to a (generalised) stopping time.

Remark 1. Theorem 1 considers only the case in which $A_i \equiv 1$. However, our approach may work in the more general setting of (5) with i.i.d. positive $\{A_m\}$ that do not depend on Q, N, and $\{R_m\}$. For example, if we assume, in addition to $\bar{n} < 1$, that $\mathbb{P}(0 < A_1 \leq 1) = 1$, then

https://doi.org/10.1017/apr.2018.69 Published online by Cambridge University Press

the exact tail asymptotics for $\mathbb{P}(R > x)$ may be easily found using the upper bound (4) and the principle of a single big jump. However, the formula for the tail asymptotics in this case is much more complicated than (4).

The multivariate version involves a set $(R(1), \ldots, (R(K))$ of random variables satisfying

$$R(i) \stackrel{\mathrm{D}}{=} Q(i) + \sum_{k=1}^{K} \sum_{m=1}^{N^{(k)}(i)} R_m(k). \tag{6}$$

In the branching process setting, this relates to K-type processes by thinking of $N^{(k)}(i)$ as the number of type-k children of a type-i ancestor. One example is the total progeny where $Q(i) \equiv 1$, others relate as above to the total life span and weighted branching processes. A queueing example concerns the busy periods $R(i)$ in the multi-class queue in [14], with i the class of the first customer in the busy period and $Q(i)$ the service time of a class-i customer; the model states that during service of a class-i customer, class-k customers arrive at rate λ_{ik}. Note that, for this example, [14] gives only lower asymptotic bounds, whereas here we provide sharp asymptotics.

The treatment of (6) is considerably more involved than for (1), and we defer the details of the assumptions and results to Section 3. We remark here only that the concept of multivariate regular variation (MRV) plays a key role; that the analogue of the crucial assumption $\bar{n} < 1$ above is subcriticality, $\rho = \operatorname{spr}(\boldsymbol{M}) < 1$, where spr denotes the spectral radius and $\boldsymbol{M}$ is the offspring mean matrix with elements $m_{ik} = \mathbb{E}N^{(k)}(i)$; and that the argument involves a recursive procedure from [16], reducing K to $K - 1$ so that ultimately we recover the $K = 1$ case of (1) and Theorem 1.

1.1. Bibliographical remarks

Any R, or its distribution, satisfying (5) is often called a fixed point of the smoothing transform (going back to [11]). There is an extensive literature on this topic, but rather than on tail asymptotics, the emphasis is most often on existence and uniqueness questions (these are easy in our context with all random variables nonnegative with finite mean and we give short self-contained proofs). Also, the assumption $A_i \neq 1$ is crucial for most of this literature. See further [1], [2], and [3], and the references therein.

It should be noted that the term 'multivariate smoothing transform' (e.g. [8]) refers to a recursion of vectors, that is, a version of (1) with R and $Q \in \mathbb{R}^K$. This differs from our setup because in (6) we are interested only in the one-dimensional distributions of the $R(i)$. In fact, for our applications, there is no interpretation of a vector with ith marginal having the distribution of $R(i)$.

In [26], tail asymptotics for the total progeny of a multi-type branching process are studied by different techniques in the critical case $\rho = 1$.

2. One-dimensional case: (1)

The heuristic underlying (4) is the principle of a single large jump: for R to exceed x, either one or both elements of (Q, N) must be large, or the independent event occurs that $R_m > x$ for some $m \le N$, in which case N is small or moderate. If N is large, $\sum_1^N R_m$ is approximately $\bar{r}N$, so roughly the probability of the first possibility is $\mathbb{P}(Q + \bar{r}N > x)$. On the other hand, results for compound heavy-tailed sums suggest that the approximate probability of the second possibility is $\bar{n}\mathbb{P}(R > x)$. We thus arrive at (4) via

$$\mathbb{P}(R > x) \approx \mathbb{P}(Q + \bar{r}N > x) + \bar{n}\mathbb{P}(R > x).$$

https://doi.org/10.1017/apr.2018.69 Published online by Cambridge University Press

In the proof of Theorem 1, let $(Q_1, N_1), (Q_2, N_2), \ldots$ be an i.i.d. sequence of pairs distributed as the (possibly dependent) pair (Q, N) in (1). Then $S_n = \sum_{i=1}^{n} \xi_i$, $i = 0, 1, \ldots$, where $\xi_i = N_i - 1$ is a random walk. Clearly, $\mathbb{E}\xi_i < 0$. Let

$$\tau = \min\{n \geq 1 \colon S_n < 0\} = \min\{n \geq 1 \colon S_n = -1\}.$$

Note that, from $S_\tau = -1$ and Wald's identity, $\mathbb{E}S_\tau = \mathbb{E}\tau \cdot \mathbb{E}(N-1)$, we have

$$\mathbb{E}\tau = \frac{1}{1 - \mathbb{E}N}.$$

Now either $N_1 = 0$, in which case $\tau = 1$, or $N_1 > 0$ so that $S_1 = N_1 - 1$ and to proceed to level -1, the random walk must go down one level N_1 times. This shows that (in obvious notation)

$$\tau \stackrel{\mathrm{D}}{=} 1 + \sum_{i=1}^{N} \tau_i. \tag{7}$$

That is, τ is a solution to (1) with $Q \equiv 1$. On the other hand, the total progeny in a Galton–Watson process with the number of offspring of an individual distributed as N obviously also satisfies (7), and, hence, by uniqueness, must have the same distribution as τ. This result first occurs as Equation (4) of [12], but we remark that an alternative representation (1) in that paper appears to have received the most attention in the literature.

Now define

$$\varphi_i = k_0 + k_1 Q_i \quad \text{and} \quad V = \sum_{i=1}^{\tau} \varphi_i,$$

where k_0 and k_1 are nonnegative constants with $k_0 + k_1 > 0$. In particular, if $(k_0, k_1) = (1, 0)$ then $V = \tau$, while

$$(k_0, k_1) = (0, 1) \quad \text{implies that} \quad V \stackrel{\mathrm{D}}{=} R.$$

Indeed, arguing as before, we conclude that the equation $V \stackrel{\mathrm{D}}{=} \varphi + \sum_1^N V_i$ has only one integrable positive solution, and, clearly,

$$V \stackrel{\mathrm{D}}{=} \varphi + \sum_1^N V_i \stackrel{\mathrm{D}}{=} \varphi + \sum_1^N \varphi_i + \sum_1^N \sum_1^{N_i} \varphi_{i,j} + \sum_1^N \sum_1^{N_i} \sum_1^{N_{i,j}} \varphi_{i,j,k} + \cdots \stackrel{\mathrm{D}}{=} \sum_1^{\tau} \varphi_i,$$

where, as before, (φ, N), (φ_i, N_i), $(\varphi_{i,j}, N_{i,j})$, etc. are i.i.d. vectors. In particular, V becomes R on replacing φ by Q.

Proof of Theorem 1. It remains to find the asymptotics of $\mathbb{P}(V > x)$ as $x \to \infty$. Throughout the proof, we assume that $k_1 > 0$.

Let r_0 be the solution to the equation

$$\mathbb{E}\varphi_1 + r_0 \mathbb{E}\xi_1 = 0.$$

Note that in the particular case that $(k_0, k_1) = (0, 1)$,

$$r_0 = \frac{\mathbb{E}Q}{1 - \mathbb{E}N} = \bar{r}. \tag{8}$$

https://doi.org/10.1017/apr.2018.69 Published online by Cambridge University Press

Choose $r > r_0$ and as close to r_0 as needed (see before (9)), and let

$$\psi_i = \varphi_i + r\xi_i.$$

We find upper and lower bounds for the asymptotics of $\mathbb{P}(V > x)$ and show that they are asymptotically equivalent. Since $k_1 > 0$ and $Q + Nr/k_1$ has an IRV distribution, the distribution of $k_1 Q + rN$ is IRV, too.

Upper bound. The key is to apply the main result of [17] to obtain the following upper bound:

$$\begin{aligned}
\mathbb{P}(V > x) &= \mathbb{P}\bigg(\sum_{i=1}^{\tau} \varphi_i > x\bigg) \\
&= \mathbb{P}\bigg(\sum_{i=1}^{\tau} \psi_i > x + rS_\tau\bigg) \\
&= \mathbb{P}\bigg(\sum_{i=1}^{\tau} \psi_i > x - r\bigg) \\
&\le \mathbb{P}\bigg(\max_{1\le k\le \tau} \sum_{i=1}^{k} \psi_i > x - r\bigg) \\
&\sim \mathbb{E}\tau \mathbb{P}(\psi_1 > x - r) \\
&\sim \mathbb{E}\tau \mathbb{P}(\psi_1 > x - r + k_0) \\
&= \mathbb{E}\tau \mathbb{P}(k_1 Q + rN > x).
\end{aligned}$$

Here the first equivalence follows from [17], noting that the distribution of ψ_1 belongs to the class S^* and that [17] only requires $\varphi_1, \varphi_2, \dots$ to be i.i.d. with respect to some filtration with respect to which τ is a stopping time. For the second equivalence, we use the long-tail property (2) of the distribution of ψ_1.

Let F be the distribution function of $k_1 Q + r_0 N$. Then, as $x \to \infty$,

$$\bar{F}(x) \le \mathbb{P}(k_1 Q + rN > x) \le \mathbb{P}\bigg(\frac{rk_1 Q}{r_0} + rN > x\bigg) \le \bar{F}(\alpha x) \le (1 + o(1))c(\alpha)\bar{F}(x),$$

where $\alpha = r_0/r < 1$ and $c(\alpha) = \limsup_{y\to\infty} \bar{F}(\alpha y)/\bar{F}(y)$.

Now we assume that the IRV condition holds, let $r \downarrow r_0$, and apply (3) to obtain the upper bound

$$\mathbb{P}(R > x) \le (1 + o(1))\mathbb{E}\tau \mathbb{P}(k_1 Q + r_0 N > x). \tag{9}$$

In particular, if $(k_0, k_1) = (0, 1)$ then $r_0 = \bar{r}$ is as in (8).

Lower bound. Here we set $\psi_n = \varphi_n + r\xi_n$, where $r > 0$ and is strictly smaller than r_0. Then the ψ_n are i.i.d. random variables with common mean $\mathbb{E}\psi_1 < 0$.

For any fixed $C > 0$, $L > 0$, $n = 1, 2, \dots$, and $x \ge 0$, we have

$$\mathbb{P}(V > x) \ge \mathbb{P}\bigg(\sum_{i=1}^{\tau} \psi_i > x\bigg) \ge \sum_{i=1}^{n} \mathbb{P}(D_i \cap A_i), \tag{10}$$

where

$$D_i = \bigg\{\sum_{j=1}^{i-1} |\psi_j| \le C,\ \tau \ge i,\ \psi_i > x + C + L\bigg\} \quad \text{and} \quad A_i = \bigcap_{\ell\ge 1} \bigg\{\sum_{j=1}^{\ell} \psi_{i+j} \ge -L\bigg\}.$$

https://doi.org/10.1017/apr.2018.69 Published online by Cambridge University Press

Indeed, the first inequality in (10) holds since S_τ is nonpositive. Next, the events D_i are disjoint and, given D_i, we have $\sum_1^i \psi_j > x + L$. Then, given $D_i \cap A_i$, we have $\sum_1^k \psi_j \ge x$ for all $k \ge i$ and, in particular, $\sum_{j=1}^{\tau} \psi_j > x$. Thus, (10) holds.

The events $\{A_i\}$ form a stationary sequence. Appealing to the strong law of large numbers (SLLN), for any $\varepsilon > 0$, we can choose $L = L_0$ so large that $\mathbb{P}(A_i) \ge 1 - \varepsilon$. For this ε, choose n_0 and C_0 such that

$$\sum_{i=1}^{n_0} \mathbb{P}\biggl(\sum_{j=1}^{i-1} |\psi_j| \le C_0,\ \tau \ge i\biggr) \ge (1-\varepsilon)\mathbb{E}\tau.$$

Since the random variables $(\{\psi_j\}_{j<i}, \mathbb{I}(\tau \le i))$ are independent of $\{\psi_j\}_{j \ge i}$, we further obtain, for any $\varepsilon \in (0, 1)$ and any $n \ge n_0$, $C \ge C_0$, and $L \ge L_0$,

$$\begin{aligned}
\mathbb{P}(V > x) &\ge \sum_{i=1}^{n} \mathbb{P}\biggl(\sum_{j=1}^{i-1} |\psi_j| \le C,\ \tau \ge i\biggr)\mathbb{P}(\psi_i > x + C + L)\mathbb{P}(A_i) \\
&\ge (1-\varepsilon)^2 \mathbb{P}(\psi_1 > x + C + L) \sum_{i=1}^{n} \mathbb{P}(\tau \ge i) \\
&\sim (1-\varepsilon)^2 \mathbb{P}(\psi_1 > x) \sum_{i=1}^{n} \mathbb{P}(\tau \ge i) \quad \text{as } x \to \infty.
\end{aligned}$$

Here the final equivalence follows from the long tailedness of ψ_1. Letting first $n \to \infty$ and then $\varepsilon \to 0$, we obtain $\liminf_{x\to\infty}(\mathbb{P}(V > x)/\mathbb{E}\tau\mathbb{P}(\psi_1 > x)) \ge 1$. Then we let $r \uparrow r_0$ and use the IRV property (3). In the particular case $(k_0, k_1) = (0, 1)$, we obtain an asymptotic lower bound that is equivalent to the upper bound derived above. $\square$

Remark 2. A slightly more intuitive approach to the lower bound is to bound $\mathbb{P}(R > x)$ below by the sum of the contributions from the disjoint events B_1, B_2, and B_3, where

$$B_1 = B \cap \{\bar{r}N > \epsilon x\}, \qquad B_2 = B \cap \{A < \bar{r}N \le \epsilon x\}, \qquad B_3 = \{\bar{r}N \le A\},$$

and $B = \{Q + \bar{r}N > (1+\epsilon)x\}$. Here, for large x and A, and small ϵ,

$$\begin{aligned}
\mathbb{P}(R > x;\, B_1) &\sim \mathbb{P}(Q + \bar{r}N > x,\ \bar{r}N > \epsilon x), \\
\mathbb{P}(R > x;\, B_2) &\ge \mathbb{P}(Q > x,\ \bar{r}N \le \epsilon x) \sim \mathbb{P}(Q + \bar{r}N > x,\ \bar{r}N \le \epsilon x), \\
\mathbb{P}(R > x;\, B_3) &\ge \sum_{n=0}^{A/\bar{r}} \mathbb{P}(R_1 + \cdots + R_n > x)\mathbb{P}(N = n) \\
&\ge \sum_{n=0}^{A/\bar{r}} \mathbb{P}(\max(R_1, \dots, R_n) > x)\mathbb{P}(N = n) \\
&\sim \sum_{n=0}^{A/\bar{r}} n\mathbb{P}(R > x)\mathbb{P}(N = n) \\
&\sim \mathbb{E}\biggl[\frac{N \wedge A}{\bar{r}}\biggr]\mathbb{P}(R > x) \\
&\sim \bar{n}\mathbb{P}(R > x).
\end{aligned}$$

We omit further detail because the arguments are close those given in Section 5 for the multivariate case.

https://doi.org/10.1017/apr.2018.69 Published online by Cambridge University Press

3. Multivariate version

The assumptions for (6) are that all the $R_m(k)$ are independent of the vector

$$\boldsymbol{V}(i) = (Q(i), N^{(1)}(i), \dots, N^{(K)}(i)), \tag{11}$$

that they are mutually independent, and that $R_m(k) \overset{\mathrm{D}}{=} R(k)$. Recall that we are interested only in the one-dimensional distributions of the $R(i)$. Accordingly, for a solution to (6), we require the validity only for each fixed i.

Denote the offspring mean matrix by $\boldsymbol{M} = (m_{ik}),\ m_{ik} = \mathbb{E}N^{(k)}(i)$. Recall that $\rho = \operatorname{spr}(\boldsymbol{M})$, where ρ is the Perron–Frobenius root if $\boldsymbol{M}$ is irreducible, but we do not need to assume this. No restrictions on the dependence structure of the vectors in (11) need to be imposed for the following result to hold (but later we do need MRV!).

Proposition 1. *Assume that $\rho < 1$. Then*

(i) *the fixed-point problem (6) has a unique nonnegative solution with $\bar{r}_i = \mathbb{E}R(i) < \infty$ for all i; and*

(ii) *the $\bar{r}_i = \mathbb{E}R(i) < \infty$ constitute the unique solution to the set of linear equations*

$$\bar{r}_i = \bar{q}_i + \sum_{k=1}^{K} m_{ik}\bar{r}_k, \qquad i = 1, \dots, K. \tag{12}$$

Proof. (i) Assume first that $Q(i) \equiv 1,\ i = 1, \dots, K$. The existence of a solution to (6) is then clear since we may take $R(i)$ as the total progeny of a type-i ancestor in a K-type Galton–Watson process, where the vector of children of a type-j individual is distributed as $(N^{(1)}(j), \dots, N^{(K)}(j))$. For uniqueness, let $(R(1), \dots, R(K))$ be any solution and consider the K-type Galton–Watson trees $\mathcal{G}(i),\ i = 1, \dots, K$, where $\mathcal{G}(i)$ corresponds to an ancestor of type i. If we define $R^{(0)}(i) = 1$,

$$R^{(n)}(i) \overset{\mathrm{D}}{=} 1 + \sum_{k=1}^{K} \sum_{m=1}^{N^{(k)}(i)} R_m^{(n-1)}(k),$$

with similar conventions as for (6), then $R^{(n)}(i)$ is the total progeny of a type-i ancestor under the restriction that the depth of the tree is at most n. Induction easily gives $R^{(n)}(i) \preceq_{\mathrm{st}} R(i)$ ('$\preceq_{\mathrm{st}}$' denotes the stochastic order) for each i. Since also $R^{(n)}(i) \preceq R^{(n+1)}(i)$, limits $R^{(\infty)}(i)$ exist, $R^{(\infty)}(i)$ must simply be the unrestricted vector of the total progeny of different types, and $R^{(\infty)}(i) \preceq_{\mathrm{st}} R(i)$. Assuming that the $R(i)$ have finite mean, (12) clearly holds with $\bar{q}_i = 1$, and so the $\Delta_i = \bar{r}_i - \mathbb{E}R^{(\infty)}(i)$ satisfy $\Delta_i = \sum_1^K m_{ik}\Delta_k$. But $\rho < 1$ implies that $\boldsymbol{I} - \boldsymbol{M}$ is invertible, so the only solution is $\Delta_i = 0$, which in view of $R^{(\infty)}(i) \preceq_{\mathrm{st}} R(i)$ implies that $R^{(\infty)}(i) \overset{\mathrm{D}}{=} R(i)$ and the stated uniqueness when $Q(i) \equiv 1$.

For more general $Q(i)$, we equip each individual of type j in $\mathcal{G}(i)$ with a weight distributed as $Q(j)$, such that the dependence between her $Q(j)$ and her offspring vector has the given structure. The argument is then a straightforward generalization and application of what was done above for $Q(i) \equiv 1$.

(ii) Taking expectations in (6) yields (12), which in matrix notation reads $\boldsymbol{r} = \boldsymbol{q} + \boldsymbol{M}\boldsymbol{r}$. Note as before that $\boldsymbol{I} - \boldsymbol{M}$ is invertible. □

For tail asymptotics, we need an MRV assumption. The definition of MRV exists in some equivalent variants (cf. [4], [21], [22], and [23]); we use the definition in polar L_1-coordinates adapted to deal with several random vectors at a time as in (11). Fix here and in the following

https://doi.org/10.1017/apr.2018.69 Published online by Cambridge University Press

a reference randomly varying tail $\bar{F}(x) = L(x)/x^\alpha$ on $(0, \infty)$. For $\boldsymbol{v} = (v_1, \dots, v_p)$, define $\|\boldsymbol{v}\| = \|\boldsymbol{v}\|_1 = |v_1| + \dots + |v_p|$ and let $\mathcal{B} = \mathcal{B}_p = \{\boldsymbol{v} \colon \|\boldsymbol{v}\| = 1\}$. We then say that a random vector $\boldsymbol{V} = (V_1, \dots, V_p)$ satisfies MRV(F) (or, has property MRV(F)) if $\mathbb{P}(\|\boldsymbol{V}\| > x) \sim b\bar{F}(x)$, where either $b = 0$ or $b > 0$ and the angular part $\boldsymbol{\Theta}_V = \boldsymbol{V}/\|\boldsymbol{V}\|$ satisfies

$$\mathbb{P}(\boldsymbol{\Theta}_V \in \cdot \mid \|\boldsymbol{V}\| > x) \xrightarrow{\mathrm{D}} \mu \quad \text{as } x \to \infty$$

for some measure μ on $\mathcal{B}$ (the angular measure). Our basic condition is then that, for the given reference randomly varying tail $\bar{F}(x)$, the following holds:

(MRV) for any $i = 1, \dots, K$, the vector $\boldsymbol{V}(i)$ in (11) satisfies MRV(F), where $b = b(i) > 0$ for at least one i.

Note that F is the same for all i, but the angular measures μ_i not necessarily so. We also assume that the mean $\bar{z}$ of F is finite; this ensures that all expected values occurring below are finite.

Assumption MRV(F) implies the RV of linear combinations, in particular of marginals. More precisely (see the appendix),

$$\mathbb{P}(a_0 Q(i) + a_1 N^{(1)}(i) + \dots + a_K N^{(K)}(i) > x) \sim c_i(a_0, \dots, a_K)\bar{F}(x), \tag{13}$$

where $c_i(a_0, \dots, a_K) = b(i) \int_{\mathcal{B}} (a_0\theta_0 + \dots + a_K\theta_K)^\alpha \mu_i(\mathrm{d}\theta_0, \dots, \mathrm{d}\theta_K)$.

Theorem 2. *Assume that $\rho < 1$, $\bar{z} < \infty$, and that (MRV) holds. Then there are constants $d_1, \dots, d_K$ such that*

$$\mathbb{P}(R(i) > x) \sim d_i \bar{F}(x) \quad \textit{as } x \to \infty. \tag{14}$$

Here the d_i constitute the unique solution to the set of linear equations

$$d_i = c_i(1, \bar{r}_1, \dots, \bar{r}_K) + \sum_{k=1}^{K} m_{ik} d_k, \qquad i = 1, \dots, K, \tag{15}$$

where the $\bar{r}_i$ are as in Proposition 1 and the c_i as in (13).

The proof follows in Sections 4–7.

4. Outline of the proof of Theorem 2

When $K > 1$, we have not found a random walk argument extending the proof in Section 2. Instead, we use a recursive procedure, going back to [16] in a queueing setting, for eventually being able to infer (14). Identification (15) of the d_i then follows immediately from the following result to be proved in Section 5 (the $p = 1$ case is Lemma 4.7 of [15]).

Proposition 2. *Let $\boldsymbol{N} = (N_1, \dots, N_p)$ be MRV with $\mathbb{P}(\|\boldsymbol{N}\| > x) \sim c_{\boldsymbol{N}}\bar{F}(x)$, and let the random variables $Z_m^{(i)}$, $i = 1, \dots, p$, $m = 1, 2, \dots$, be independent with distribution F_j for $Z_i^{(j)}$, independent of $\boldsymbol{N}$ and with finite means $\bar{z}_j = \mathbb{E}Z_m^{(j)}$. Define $S_m^{(j)} = Z_1^{(j)} + \dots + Z_m^{(j)}$. If $\bar{F}_j(x) \sim c_j \bar{F}(x)$ then*

$$\mathbb{P}(S_{N_1}^{(1)} + \dots + S_{N_p}^{(p)} > x) \sim \mathbb{P}(\bar{z}_1 N_1 + \dots + \bar{z}_1 N_p > x) + c_0 \bar{F}(x),$$

where $c_0 = c_1 \mathbb{E}N_1 + \dots + c_p \mathbb{E}N_p$.

The recursion idea in [16] amounts in a queueing context to letting all class-K customers be served first. We implement it here in the branching process setting. Consider the multi-type Galton–Watson tree $\mathcal{G}$. For an ancestor of type $i < K$ and any of her daughters $m = 1, \dots, N^{(K)}(i)$ of type K, consider the family tree $\mathcal{G}_m(i)$ formed by m and all her type-K

descendants in *direct* line. For a vertex $g \in \mathcal{G}_m(i)$ and $k < K$, let $N_g^{(k)}(K)$ denote the number of type-k daughters of g.

Note that $\mathcal{G}_m(i)$ is simply a one-type Galton–Watson tree starting from a single ancestor and with the number of daughters distributed as $N^{(K)}(K)$. In particular, the expected size of $\mathcal{G}_m(i)$ is $1/(1 - m_{KK})$. We further have

$$R(i) \stackrel{\mathrm{D}}{=} \tilde{Q}(i) + \sum_{k=1}^{K-1} \sum_{m=1}^{\tilde{N}^{(k)}(i)} R_m(k), \qquad i = 1, \ldots, K-1, \tag{16}$$

where

$$\tilde{Q}(i) = Q(i) + \sum_{m=1}^{N^{(K)}(i)} \sum_{g \in \mathcal{G}_m(i)} Q_g(K), \qquad \tilde{N}^{(k)}(i) = N^{(k)}(i) + \sum_{m=1}^{N^{(K)}(i)} \sum_{g \in \mathcal{G}_m(i)} N_g^{(k)}(K),$$

that is, a fixed-point problem with one type less.

Example 1. Let $K = 2$, and consider the two-type family tree in Figure 1, where type-1 individuals are denoted by either filled squares or filled triangles, type-2 descendants of the ancestor in direct line are denoted by open squares, and the remaining type-2 individuals are denoted by open triangles. Type-1 individuals denoted by filled triangles are those that are counted as extra type-1 children in the reduced recursion (16). We have $N^{(2)}(1) = 2$, and if m is the upper open-square individual of type 2 then $\mathcal{G}_m(2)$ has size 4. Furthermore, $\sum_{g \in \mathcal{G}_m(1)} N_g^{(1)} = 2$, with m herself and her upper daughter each contributing one individual.

The offspring mean in the reduced one-type family tree is $\tilde{m} = m_{11} + m_{12}m_{21}/(1 - m_{22})$. Indeed, the first term is the expected number of original type-1 offspring of the ancestor, and in the second term, m_{12} is the expected number of type-2 offspring of the ancestor, $1/(1 - m_{22})$ the size of the direct line type-2 family tree produced by each of them, and m_{21} the expected number of type-1 offspring of each individual in this tree.

Since the original two-type tree is finite, the reduced one-type tree must necessarily also be so, so that $\tilde{m} \leq 1$. A direct verification of this is instructive. First note that

$$\begin{aligned} \tilde{m} \leq 1 &\iff m_{11} - m_{11}m_{22} + m_{12}m_{21} \leq 1 - m_{22} \\ &\iff \operatorname{tr}(\boldsymbol{M}) - \det(\boldsymbol{M}) \leq 1. \end{aligned}$$

However, the characteristic polynomial of the two-type offspring mean matrix $\boldsymbol{M}$ equals $\lambda^2 - \lambda \operatorname{tr}(\boldsymbol{M}) + \det(\boldsymbol{M})$, and the dominant eigenvalue ρ of $\boldsymbol{M}$ satisfies $\rho < 1$ so that

$$\operatorname{tr}(\boldsymbol{M}) - \det(\boldsymbol{M}) \leq \rho \operatorname{tr}(\boldsymbol{M}) - \det(\boldsymbol{M}) = \rho^2 < 1.$$

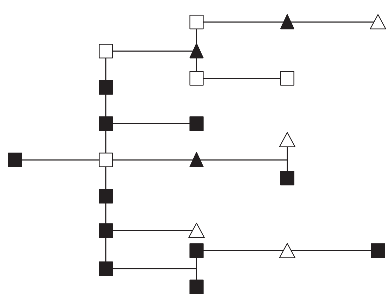

FIGURE 1: Reducing from a two-type to a one-type family tree.

https://doi.org/10.1017/apr.2018.69 Published online by Cambridge University Press

5. Proof of Proposition 2

We need the following result of Nagaev *et al.* (see the discussion in [15] around Equation (4.2) therein for references).

Lemma 1. *Let $Z_1, Z_2, \dots$ be i.i.d. and regularly varying with finite mean $\bar{z}$, and define $S_k = Z_1 + \cdots + Z_k$. Then, for any $\delta > 0$,*

$$\sup_{y \geq \delta k} \left| \frac{\mathbb{P}(S_k > k\bar{z} + y)}{k\bar{F}(y)} - 1 \right| \to 0 \quad \text{as } k \to \infty.$$

Corollary 1. *Under the assumptions of Lemma 1,*

$$d(F, \epsilon) = \limsup_{x \to \infty} \sup_{k < \epsilon x} \frac{\mathbb{P}(S_k > x)}{k\bar{F}(x)} < \infty \quad \text{for } 0 < \epsilon < \frac{1}{\bar{z}}.$$

Proof. Define $\delta = (1 - \epsilon\bar{z})/\epsilon$. We can write $x = k\bar{z} + y$, where

$$y = y(x, k) = x - k\bar{z} \geq x(1 - \epsilon\bar{z}) = x\epsilon\delta \geq \delta k.$$

Lemma 1 therefore implies that, for all large x, we can bound $\mathbb{P}(S_k > x)$ by $Ck\bar{F}(y)$, where C does not depend on x. Now just note that, by RV,

$$\bar{F}(y) \leq \bar{F}(x\epsilon\delta) \sim (\epsilon\delta)^{-\alpha}\bar{F}(x). \qquad \square$$

Proof of Proposition 2. For ease of exposition, we start with the $p = 2$ case. We split the probability in question into four parts:

$$\begin{aligned}
p_1(x) &= \mathbb{P}(S_{N_1}^{(1)} + S_{N_2}^{(2)} > x,\ N_1 \leq \epsilon x,\ N_2 \leq \epsilon x), \\
p_{21}(x) &= \mathbb{P}(S_{N_1}^{(1)} + S_{N_2}^{(2)} > x,\ N_1 > \epsilon x,\ N_2 \leq \epsilon x), \\
p_{22}(x) &= \mathbb{P}(S_{N_1}^{(1)} + S_{N_2}^{(2)} > x,\ N_1 \leq \epsilon x,\ N_2 > \epsilon x), \\
p_3(x) &= \mathbb{P}(S_{N_1}^{(1)} + S_{N_2}^{(2)} > x,\ N_1 > \epsilon x,\ N_2 > \epsilon x).
\end{aligned}$$

Here

$$p_1(x) = \sum_{k_1, k_2 = 0}^{\epsilon x} \mathbb{P}(S_{k_1}^{(1)} + S_{k_2}^{(2)} > x)\mathbb{P}((N_1, N_2) = (k_1, k_2)).$$

Since $S_{k_1}^{(1)}$ and $S_{k_2}^{(2)}$ are independent, standard RV theory implies that

$$\mathbb{P}(S_{k_1}^{(1)} + S_{k_2}^{(2)} > x) \sim (k_1 c_1 + k_2 c_2)\bar{F}(x) \quad \text{as } x \to \infty.$$

Furthermore, Corollary 1 shows that, for $k_1, k_2 \leq \epsilon x$ and all large x,

$$\begin{aligned}
\mathbb{P}(S_{k_1}^{(1)} + S_{k_2}^{(2)} > x) &\leq \mathbb{P}\big(S_{k_1}^{(1)} > \tfrac{1}{2}x\big) + \mathbb{P}\big(S_{k_2}^{(2)} > \tfrac{1}{2}x\big) \\
&\leq 2(d(F_1, 2\epsilon)k_1 + d(F_2, 2\epsilon)k_2)\bar{F}(x).
\end{aligned}$$

Hence, by dominated convergence, as $x \to \infty$,

$$\frac{p_1(x)}{\bar{F}(x)} \to \sum_{k_1, k_2 = 0}^{\infty} (k_1 c_1 + k_2 c_2)\mathbb{P}((N_1, N_2) = (k_1, k_2)) = c_1\mathbb{E}N_1 + c_2\mathbb{E}N_2.$$

https://doi.org/10.1017/apr.2018.69 Published online by Cambridge University Press

For $p_3(x)$, let $A_j(m)$ denote the event that $S^{(j)}_{k_j}/k_j \le \bar{z}_j/(1-\epsilon)$ for all $k_j > m$. Then, by the SLLN, there are constants $r(m)$ converging to 0 as $m \to \infty$ such that $\mathbb{P}(A_j(m)^{\mathrm{c}}) \le r(m)$ for $j = 1, 2$. It follows that, as $x \to \infty$,

$$\begin{aligned}
p_3(x) &\le (\mathbb{P}(A_1(\epsilon x)^{\mathrm{c}}) + \mathbb{P}(A_2(\epsilon x)^{\mathrm{c}}))\mathbb{P}(N_1 > \epsilon x,\ N_2 > \epsilon x) \\
&\qquad + \mathbb{P}(S^{(1)}_{N_1} + S^{(2)}_{N_2} > x,\ N_1 > \epsilon x,\ N_2 > \epsilon x, A_1(\epsilon x), A_2(\epsilon x)) \\
&\le r(\epsilon x) O(\bar{F}(x)) + \mathbb{P}\bigg(\frac{\bar{z}_1 N_1 + \bar{z}_2 N_2}{1-\epsilon} > x,\ N_1 > \epsilon x,\ N_2 > \epsilon x\bigg) \\
&\le o(\bar{F}(x))\mathbb{P}(\bar{z}_1 N_1 + \bar{z}_2 N_2 > \eta x,\ N_1 > \epsilon x,\ N_2 > \epsilon x),
\end{aligned}$$

where $\eta < 1 - \epsilon$ is specified shortly.

For $p_{21}(x)$, write $p_{21}(x) = p'_{21}(x) + p''_{21}(x)$, where, with $\gamma = 2\epsilon\bar{z}_2$,

$$\begin{aligned}
p'_{21}(x) &= \mathbb{P}(S^{(1)}_{N_1} + S^{(2)}_{N_2} > x, S^{(2)}_{N_2} \le \gamma x,\ N_1 > \epsilon x,\ N_2 \le \epsilon x), \\
p''_{21}(x) &= \mathbb{P}(S^{(1)}_{N_1} + S^{(2)}_{N_2} > x, S^{(2)}_{N_2} > \gamma x,\ N_1 > \epsilon x,\ N_2 \le \epsilon x).
\end{aligned}$$

Here

$$\begin{aligned}
p''_{21}(x) &\le \mathbb{P}(S^{(1)}_{N_1} + S^{(2)}_{\epsilon x} > x,\ S^{(2)}_{\epsilon x} > \gamma x,\ N_1 > \epsilon x,\ N_2 \le \epsilon x) \\
&\le \mathbb{P}(S^{(2)}_{\epsilon x} > \gamma x,\ N_1 > \epsilon x) \\
&= \mathbb{P}(S^{(2)}_{\epsilon x} > \gamma x)\mathbb{P}(N_1 > \epsilon x) \\
&= o(1) O(\bar{F}(x)) \\
&= o(\bar{F}(x)),
\end{aligned}$$

using the LLN in the fourth step. Furthermore, as in the estimates above,

$$\begin{aligned}
p'_{21}(x) &\le \mathbb{P}(S^{(1)}_{N_1} > x(1-\gamma),\ N_1 > \epsilon x,\ N_2 \le \epsilon x) \\
&\le o(\bar{F}(x)) + \mathbb{P}(\bar{z}_1 N_1 > x(1-\gamma)(1-\epsilon),\ N_1 > \epsilon x,\ N_2 \le \epsilon x) \\
&\le \mathbb{P}(\bar{z}_1 N_1 + \bar{z}_2 N_2 > x(1-\gamma)(1-\epsilon),\ N_1 > \epsilon x,\ N_2 \le \epsilon x).
\end{aligned}$$

We can now finally put the above estimates together. For ease of notation, write $\eta = \eta(\epsilon) = (1-\gamma)(1-\epsilon)$ and note that $\eta \uparrow 1$ as $\epsilon \downarrow 0$. Using a similar estimate for $p_{12}(x)$ as for $p_{21}(x)$ and noting that, for small enough ϵ,

$$\mathbb{P}(\bar{z}_1 N_1 + \bar{z}_2 N_2 > \eta x,\ N_1 \le \epsilon x,\ N_2 \le \epsilon x) = 0,$$

we get

$$\begin{aligned}
\limsup_{x\to\infty} \frac{\mathbb{P}(S^{(1)}_{N_1} + S^{(2)}_{N_2} > x)}{\bar{F}(x)} &= c_1 \mathbb{E} N_1 + c_2 \mathbb{E} N_2 + \limsup_{x\to\infty} \frac{\mathbb{P}(\bar{z}_1 N_1 + \bar{z}_2 N_2 > \eta x)}{\bar{F}(x)} \\
&= c_1 \mathbb{E} N_1 + c_2 \mathbb{E} N_2 + c(\bar{z}_1, \bar{z}_2) \limsup_{x\to\infty} \frac{\bar{F}(\eta x)}{\bar{F}(x)} \\
&= c_1 \mathbb{E} N_1 + c_2 \mathbb{E} N_2 + c(\bar{z}_1, \bar{z}_2) \frac{1}{\eta^\alpha}.
\end{aligned}$$

Letting $\epsilon \downarrow 0$ shows that the lim sup is bounded by $c_0 + c(\bar{z}_1, \bar{z}_2)$. Similar estimates for the lim inf complete the proof for $p = 2$.

If $p > 2$, the only essential difference is that $p_{21}(x)$ and $p_{22}(x)$ must be replaced by $2^p - 2$ terms corresponding to all combinations of some N_i being smaller or equal to ϵx and others being greater than ϵx, with the two exceptions where either all are smaller or equal to ϵx or all are greater than ϵx. However, for each of these two cases, estimates similar to those above for $p_{21}(x)$ apply. □

6. Preservation of MRV under sum operations

Before giving our main auxiliary result, Proposition 4, it is instructive to recall two extremely simple examples of MRV. The first is two i.i.d. RV(F) random variables X_1 and X_2, where a large value of X_1+X_2 can occur only if one variable is large and the other small; this gives MRV with the angular measure concentrated on the points $(1, 0), (0, 1) \in \mathcal{B}_2$ with mass $\frac{1}{2}$ for each.

The second example is slightly more complicated as follows.

Proposition 3. *Let $N, Z, Z_1, Z_2, \dots$ be nonnegative random variables such that $N \in \mathbb{N}$ and $Z, Z_1, Z_2, \dots$ are i.i.d., nonnegative, and independent of N. Assume that $\mathbb{P}(N > x) \sim c_N \bar{F}(x)$ and $\mathbb{P}(Z > x) \sim c_Z \bar{F}(x)$ for some regularly varying tail $\bar{F}(x) = L(x)/x^{\alpha}$, and write $S = \sum_1^N Z_i$, $\bar{n} = \mathbb{E}N$, and $\bar{z} = \mathbb{E}Z$, where $c_N + c_Z > 0$. Then*

(i) $\mathbb{P}(S > x) \sim (c_N \bar{z}^{\alpha} + c_Z \bar{n})\bar{F}(x)$; *and*

(ii) *the random vector (N, S) is MRV with*

$$\mathbb{P}(\|(N, S)\| > x) \sim c_{N,S}\bar{F}(x), \quad \textit{where} \quad c_{N,S} = c_N(1 + \bar{z}^{\alpha}) + c_Z\bar{n},$$

and angular measure $\mu_{N,S}$ concentrated on the points $\boldsymbol{b}_1 = (1/(1 + \bar{z}), \bar{z}/(1 + \bar{z}))$ and $\boldsymbol{b}_2 = (0, 1)$, with

$$\mu_{N,S}(\boldsymbol{b}_1) = \frac{c_N}{c_N + c_Z\bar{n}}, \qquad \mu_{N,S}(\boldsymbol{b}_2) = \frac{c_Z\bar{n}}{c_N + c_Z\bar{n}}.$$

Proof. Part (i) is Lemma 4.7 of [15] (see also [10]). The proof in [15] also shows that if $S > x$ then, approximately, either $N\bar{z} > x$, occurring with probability $c_N\bar{F}(x/\bar{z}) \sim c_N\bar{z}^{\alpha}\bar{F}(x)$, or $N \le \epsilon x$ and $Z_i > x$, occurring with probability $c_Z\mathbb{E}[N \wedge \epsilon x]\bar{F}(x)$. The first possibility gives the atom of $\mu_{N,S}$ at $\boldsymbol{b}_1$ and the second gives the atom at $\boldsymbol{b}_2$ since $\mathbb{E}[N \wedge \epsilon x] \uparrow \bar{n}$. □

Proposition 4. *Let $\boldsymbol{V} = (\boldsymbol{T}, N) \in [0, \infty)^p \times \mathbb{N}$ satisfy* MRV(F), *let $\boldsymbol{Z}, \boldsymbol{Z}_1, \boldsymbol{Z}_2, \dots \in [0, \infty)^q$ be i.i.d., independent of $(\boldsymbol{T}, N)$, and satisfying* MRV(F), *and define $\boldsymbol{S} = \sum_1^N \boldsymbol{Z}_i$. Then $\boldsymbol{V}^* = (\boldsymbol{T}, N, \boldsymbol{S})$ satisfies* MRV(F).

Proof. Let $\bar{z} \in [0, \infty)^q$ be the mean of $\boldsymbol{Z}$. Arguments similar to those in Section 5 show that $\|V^*\| > x$ basically occurs when either $\|\boldsymbol{T}\| + N + N\|\bar{z}\| > x$ or $\|\boldsymbol{V}\| \le \epsilon x$ and some $\|\boldsymbol{Z}_i\| > x$. The probabilities of these events are approximately of the form $c'\bar{F}(x)$ and $c''\bar{F}(x)$, so the radial part of $\boldsymbol{V}^*$ is randomly varying with asymptotic tail $c_{\boldsymbol{V}^*}\bar{F}(x)$, where $c_{\boldsymbol{V}^*} = c' + c''$. Now

$$\mathbb{P}\biggl(\frac{(\boldsymbol{T}, N)}{\|(\boldsymbol{T}, N)\|} \in \cdot \biggm| \|\boldsymbol{T}\| + N + N\|\bar{z}\| > x\biggr) \to \mu'$$

for some probability measure μ' on the $(p + 1)$-dimensional unit sphere $\mathcal{B}_{p+1}$; this follows from the facts that $\|\boldsymbol{T}\| + N + N\|\bar{z}\|$ is a norm and the MRV property of a vector is independent of the choice of norm. Letting δ_0' be the Dirac measure at $(0, \dots, 0) \in \mathbb{R}^q$, δ_0'' be the Dirac measure at $(0, \dots, 0) \in \mathbb{R}^{p+1}$, and $\mu'' = \mu_{\boldsymbol{Z}}$ be the angular measure of $\boldsymbol{Z}$, we obtain the desired conclusion with $c_{\boldsymbol{V}^*} = c' + c''$ and the angular measure of $\boldsymbol{V}^*$ given by

$$\mu_{\boldsymbol{V}^*} = \frac{c'}{c' + c''}\mu' \otimes \delta_0'' + \frac{c''}{c' + c''}\delta_0' \otimes \mu''. \qquad \square$$

In calculations that follow in Lemma 2, extending some $\boldsymbol{V}$ to some $\boldsymbol{V}^*$ in a number of steps, expressions for $c_{\boldsymbol{V}^*}$ and $\mu_{\boldsymbol{V}^*}$ can be deduced along the lines of the proof of Propositions 3–4, but the expression and details become extremely tedious. Fortunately, they are not needed and we therefore omit them—all that matters is existence. If α is not an even integer, the MRV alone of $\boldsymbol{V}^*$ can be obtained alternatively (and slightly more easily) from Theorem 1.1(iv) of [4], stating that, by nonnegativity, it suffices to verify MRV of any linear combination.

7. Proof of Theorem 2 completed

Lemma 2. *In the setting of (16), the random vector*

$$\boldsymbol{V}^*(i) = (\tilde{Q}(i), \tilde{N}^{(1)}(i), \dots, \tilde{N}^{(K-1)}(i))$$

satisfies MRV(F) *for all* i.

Proof. Let $|\mathcal{G}_m(i)|$ be the number of elements of $\mathcal{G}_m(i)$, and let

$$M_1(i) = \sum_{m=1}^{N^{(K)}(i)} |\mathcal{G}_m(i)|, \qquad M_2(i) = \sum_{m=1}^{N^{(K)}(i)} \sum_{g \in \mathcal{G}_m(i)} (Q_g(K), N_g^{(1)}(K), \dots, N_g^{(1)}(K-1)).$$

Recall our basic assumption that the

$$\boldsymbol{V}^*(i) = (Q(i), N^{(1)}(i), \dots, N^{(K)}(i)) \tag{17}$$

satisfy MRV(F). The connection to a Galton–Watson tree and Theorem 1 with $Q \equiv 1$ and $N = N^{(K)}(i)$ therefore implies that so does any $|\mathcal{G}_m(i)|$, and since these random variables are i.i.d. and independent of $N^{(K)}(i)$, it follows from Proposition 4 that $\boldsymbol{V}_1(i) = (\boldsymbol{V}(i), M_1(i))$ satisfies MRV(F). Now the MRV(F) property of (17) with $i = K$ implies that the vectors $(Q_g(K), N_g^{(1)}(K), \dots, N_g^{(K-1)}(K))$, being distributed as $(Q(K), N^{(1)}(K), \dots, N^{(K-1)}(K))$ again satisfy MRV(F). But $M_2(i)$ is a sum of $M_1(i)$ such vectors that are i.i.d. given $M_1(i)$. Using Proposition 4 again shows that $\boldsymbol{V}_2(i) = (\boldsymbol{V}(i), M_1(i), M_2(i))$ satisfies MRV(F). But $\boldsymbol{V}^*(i)$ is a function of $\boldsymbol{V}_2(i)$. Since this function is linear, property MRV(F) of $\boldsymbol{V}_2(i)$ carries over to $\boldsymbol{V}^*(i)$. □

Proof of Theorem 2. We use induction in K. The $K = 1$ case is just Theorem 1, so assume we have shown Theorem 2 for $K - 1$.

The induction hypothesis and Lemma 2 imply that $\mathbb{P}(R(i) > x) \sim d_i \bar{F}(x)$ for $i = 1, \dots, K - 1$. Rewriting (6) for $i = K$ as

$$R(K) \stackrel{\mathrm{D}}{=} Q^*(K) + \sum_{m=1}^{N^{(K)}(K)} R_m(K), \quad \text{where} \quad Q^*(K) = \sum_{k=1}^{K-1} \sum_{m=1}^{N^{(k)}(K)} R_m(k),$$

we have a fixed-point problem of type (1) and can then use Theorem 2 to also conclude that $\mathbb{P}(R(K) > x) \sim d_K \bar{F}(x)$, noting that the MRV condition needed on $(Q^*(K), N^{(k)}(K))$ follows by another application of Proposition 4.

Finally, to identify the d_i via (15), appeal to Proposition 2 with $\boldsymbol{N} = (Q(i), N^{(1)}(i), \dots, N^{(K)}(i))$, writing the right-hand side of (6) as

$$O(1) + \sum_{m=1}^{\lfloor Q(i) \rfloor} 1 + \sum_{k=1}^{K} \sum_{m=1}^{N^{(k)}(i)} R_m(k).$$

Existence and uniqueness of a solution to (15) follows by again noting that $\rho < 1$ implies that $\boldsymbol{I} - \boldsymbol{M}$ is invertible. □

https://doi.org/10.1017/apr.2018.69 Published online by Cambridge University Press

Appendix A. Proof of (13)

The RV of linear combinations subject to MRV assumptions has received considerable attention (see, e.g. [4]), but we could not find explicit formulae like (13) for the relevant constants, so we give a self-contained proof; the formula is a special case of the following. If $\boldsymbol{X} = (X_1, \dots, X_n) \in \mathbb{R}^n$ is a random vector such that $\mathbb{P}(\|\boldsymbol{X}\| > t) \sim L(t)/t^\alpha$ and $\boldsymbol{\Theta} = \boldsymbol{X}/\|\boldsymbol{X}\|$ has conditional limit distribution μ in $\mathcal{B}_1$ given $\|\boldsymbol{X}\| > t$ as $t \to \infty$, then

$$\mathbb{P}(\boldsymbol{a} \cdot \boldsymbol{X} > t) = \mathbb{P}(a_1 X_1 + \dots + a_n X_n > t) \sim \frac{L(t)}{t^\alpha} \int_{\mathcal{B}_1} \mathbb{I}(\boldsymbol{a} \cdot \boldsymbol{\theta} > 0)(\boldsymbol{a} \cdot \boldsymbol{\theta})^\alpha \mu(\mathrm{d}\boldsymbol{\theta}).$$

To see this, note that, given $\boldsymbol{\Theta} = \boldsymbol{\theta} \in \mathcal{B}_1$, $\boldsymbol{a} \cdot \boldsymbol{X} = \|\boldsymbol{X}\|(\boldsymbol{a} \cdot \boldsymbol{\theta})$ exceeds $t > 0$ precisely when $\boldsymbol{a} \cdot \boldsymbol{\theta} > 0$ and $\|\boldsymbol{X}\| > t/(\boldsymbol{a} \cdot \boldsymbol{\theta})$. Thus, we expect that

$$\begin{aligned}\mathbb{P}(\boldsymbol{a} \cdot \boldsymbol{X} > t) &\sim \int_{\mathcal{B}_1} \mathbb{I}(\boldsymbol{a} \cdot \boldsymbol{\theta} > 0)\mathbb{P}\bigg(\|\boldsymbol{X}\| > \frac{t}{\boldsymbol{a} \cdot \boldsymbol{\theta}}\bigg)\mu(\mathrm{d}\boldsymbol{\theta}) \\ &\sim \int_{\mathcal{B}_1} \mathbb{I}(\boldsymbol{a} \cdot \boldsymbol{\theta} > 0)\frac{L(t/(\boldsymbol{a} \cdot \boldsymbol{\theta}))}{(t/(\boldsymbol{a} \cdot \boldsymbol{\theta}))^\alpha}\mu(\mathrm{d}\boldsymbol{\theta}) \\ &\sim \frac{L(t)}{t^\alpha} \int_{\mathcal{B}_1} \mathbb{I}(\boldsymbol{a} \cdot \boldsymbol{\theta} > 0)(\boldsymbol{a} \cdot \boldsymbol{\theta})^\alpha \mu(\mathrm{d}\boldsymbol{\theta}),\end{aligned}$$

which is the same as asserted.

A rigorous proof can be carried out either by involving Riemann sums, or by using the so-called *limit measure* or *exponent measure* ν (see [23]), defined by

$$a(t)\mathbb{P}\bigg(\frac{\boldsymbol{X}}{t} \in \cdot\bigg) \to \nu \tag{18}$$

for some suitable function $a(t)$, with ν nondegenerate. If $\boldsymbol{X}$ is MRV as above then (18) holds and we can take

$$a(t) = \frac{1}{\mathbb{P}(\|\boldsymbol{X}\| > t)}. \tag{19}$$

Furthermore, ν is a product measure in polar coordinates $(\|\boldsymbol{x}\|, \boldsymbol{\theta})$ and subject to (19), ν can be expressed in terms of α and μ as

$$\nu\{\|\boldsymbol{x}\| \in \mathrm{d}z,\ \boldsymbol{\theta} \in \mathrm{d}\theta\} = \alpha z^{-\alpha-1}\mathrm{d}z \times \mu(\mathrm{d}\theta).$$

For the proof of (13), observe that, by (18) and (19),

$$\mathbb{P}(\boldsymbol{a} \cdot \boldsymbol{X} > t) \sim \frac{1}{a(t)}\nu\{\boldsymbol{a} \cdot \boldsymbol{x} > 1\} \sim \frac{L(t)}{t^\alpha}\nu\{\boldsymbol{a} \cdot \boldsymbol{x} > 1\}.$$

But $\int_0^\infty \mathbb{I}(z > b)\alpha z^{-\alpha-1}\,\mathrm{d}z = b^{-\alpha}$ for $b > 0$, and so taking $b = 1/\boldsymbol{\theta} \cdot \boldsymbol{a}$ gives

$$\nu\{\boldsymbol{a} \cdot \boldsymbol{x} > 1\} = \int_{\mathcal{B}_1} \mathrm{d}\boldsymbol{\theta} \int_0^\infty \mathbb{I}(z\boldsymbol{\theta} \cdot \boldsymbol{a} > 1)\alpha z^{-\alpha-1}\,\mathrm{d}z = \int_{\mathcal{B}_1} \mathbb{I}(\boldsymbol{a} \cdot \boldsymbol{\theta} > 0)(\boldsymbol{a} \cdot \boldsymbol{\theta})^\alpha \mu(\mathrm{d}\boldsymbol{\theta}).$$

References

[1] Aldous, D. J. and Bandyopadhyay, A. (2005). A survey of max-type recursive distributional equations. *Ann. Appl. Prob.* **15,** 1047–1110.

[2] Alsmeyer, G. and Meiners, M. (2013). Fixed points of the smoothing transform: two-sided solutions. *Prob. Theory Relat. Fields* **155,** 165–199.

https://doi.org/10.1017/apr.2018.69 Published online by Cambridge University Press

[3] ALSMEYER, G., BIGGINS, J. D. AND MEINERS, M. (2012). The functional equation of the smoothing transform. *Ann. Prob.* **40,** 2069–2105.

[4] BASRAK, B., DAVIS, R. A. AND MIKOSCH, T. (2002). A characterization of multivariate regular variation. *Ann. Appl. Prob.* **12,** 908–920.

[5] BINGHAM, N. H. AND DONEY, R. A. (1974). Asymptotic properties of supercritical branching processes. I. The Galton-Watson process. *Adv. Appl. Prob.* **6,** 711–731.

[6] BINGHAM, N. H. AND DONEY, R. A. (1975). Asymptotic properties of supercritical branching processes. II. Crump-Mode and Jirina processes. *Adv. Appl. Prob.* **7,** 66–82.

[7] BURACZEWSKI, D., DAMEK, E. AND MIKOSCH, T. (2016). *Stochastic Models with Power-Law Tails.* Springer.

[8] BURACZEWSKI, D., DAMEK, E., MENTEMEIER, S. AND MIREK, M. (2013). Heavy tailed solutions of multivariate smoothing transforms. *Stoch. Process. Appl.* **123,** 1947–1986.

[9] DE MEYER, A. AND TEUGELS, J. L. (1980). On the asymptotic behaviour of the distributions of the busy period and service time in $M/G/1$. *J. Appl. Prob.* **17,** 802–813.

[10] DENISOV, D., FOSS, S. AND KORSHUNOV, D. (2010). Asymptotics of randomly stopped sums in the presence of heavy tails. *Bernoulli* **16,** 971–994.

[11] DURRETT, R. AND LIGGETT, T. M. (1983). Fixed points of the smoothing transformation. *Z. Wahrscheinlichkeitsth.* **64,** 275–301.

[12] DWASS, M. (1969). The total progeny in a branching process and a related random walk. *J. Appl. Prob.* **6,** 682–686.

[13] EMBRECHTS, P., KLÜPPELBERG, C. AND MIKOSCH, T. (1997). *Modelling Extremal Events*. Springer, Berlin.

[14] ERNST, P. A., ASMUSSEN, S. AND HASENBEIN, J. J. (2018). Stability and tail asymptotics in a multiclass queue with state-dependent arrival rates. *Queueing Systems* **90,** 207–224.

[15] FAŸ, G., GONZÁLEZ-ARÉVALO, B., MIKOSCH, T. AND SAMORODNITSKY, G. (2006). Modeling teletraffic arrivals by a Poisson cluster process. *Queueing Systems* **54,** 121–140.

[16] FOSS, S. (1984). Queues with customers of several types. In *Advances in Probability Theory: Limit Theorems & Related Problems*, ed. A. A. Borovkov, Optimization Software, Publications Division, New York, pp. 348–377.

[17] FOSS, S. AND ZACHARY, S. (2003). The maximum on a random time interval of a random walk with long-tailed increments and negative drift. *Ann. Appl. Prob.* **13,** 37–53.

[18] FOSS, S., KORSHUNOV, D. AND ZACHARY, S. (2013). *An Introduction to Heavy-Tailed and Subexponential Distributions*, 2nd edn. Springer, New York.

[19] JAGERS, P. (1975). *Branching Processes with Biological Applications*. John Wiley, New York.

[20] JELENKOVIĆ, P. R. AND OLVERA-CRAVIOTO, M. (2010). Information ranking and power laws on trees. *Adv. Appl. Prob.* **42,** 1057–1093.

[21] MEERSCHAERT, M. M. AND SCHEFFLER, H.-P. (2001). *Limit Distributions for Sums of Independent Random Vectors: Heavy Tails in Theory and Practice.* John Wiley, New York.

[22] RESNICK, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes* (Ser. Appl. Prob. Trust **4**). Springer, New York.

[23] RESNICK, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.* Springer, New York.

[24] SENETA, E. (1969). Functional equations and the Galton–Watson process. *Adv. Appl. Prob.* **1,** 1–42.

[25] SENETA, E. (1974). Regularly varying functions in the theory of simple branching processes. *Adv. Appl. Prob.* **6,** 408–420.

[26] VATUTIN, V. A. (1977). Limit theorems for critical Markov branching processes with several types of particles and infinite second moments. *Mat. Sbornik* **32,** 215–225.

[27] VOLKOVICH, Y. AND LITVAK, N. (2010). Asymptotic analysis for personalized web search. *Adv. Appl. Prob.* **42,** 577–604.

[28] ZWART, A. P. (2001). Tail asymptotics for the busy period in the $GI/G/1$ queue. *Math. Operat. Res.* **26,** 485–493.

SØREN ASMUSSEN, *Aarhus University*

Department of Mathematics, Aarhus University, Ny Munkegade, 8000 Aarhus C, Denmark.
Email address: asmus@math.au.dk

SERGEY FOSS, *Heriot-Watt University and Novosibirsk State University*

School of Mathematical and Computer Sciences, Heriot-Watt University, EH14 4AS, Edinburgh, UK.

Research supported by RSF grant No. 17-11-01173.

https://doi.org/10.1017/apr.2018.69 Published online by Cambridge University Press