Applied Probability Trust (13 November 2018)

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

SØREN ASMUSSEN,* *Aarhus University*

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

. *Novosibirsk State University*

Abstract

Tail asymptotics of the solution R to a fixed-point problem of type $R \stackrel{\mathcal{D}}{=} Q+\sum_1^N R_m$ is 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 multitype branching processes and busy periods in multiclass queues.

Keywords: Busy period; Galton-Watson process; intermediate regular variation; multivariate regular variation; random recursion; random sums

2010 Mathematics Subject Classification: Primary 60H25

Secondary 60J80; 60K25; 60F10

1. Introduction

This paper is concerned with the tail asymptotics of the solution R to the fixed-point problem

$$R \stackrel{\mathcal{D}}{=} Q + \sum_{m=1}^{N} R_m \tag{1.1}$$

under suitable regular variation (RV) conditions and the similar problem in a multidimensional setting stated below as (1.6). Here in (1.1) Q, N are (possibly dependent) non-negative non-degenerate r.v.'s where N is integer-valued, $R_1, R_2, \ldots$ are i.i.d. and

* Postal address: Department of Mathematics, Aarhus University, Ny Munkegade, 8000 Aarhus C, Denmark

** Postal address: School of Mathematical and Computer Sciences, Heriot-Watt University, EH14 4AS, Edinburgh, United Kingdom. Research supported by RSF grant No. 17-11-01173.

distributed as R, and $\overline{n} = \mathbb{E}N < 1$ (similar notation for expected values is used in the following).

A classical example is R being the M/G/1 busy period, cf. [21], [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 (RV) 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 ([17]), corresponding to Q being the lifetime of the ancestor and N the number of her children. Related examples are weighted branching processes, see [19] for references. Note that connections between branching processes and RV have a long history, going back at least to [24], [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 $y > 0$

$$\frac{\overline{F}(x+y)}{\overline{F}(x)} \to 1 \quad \text{as } x \to \infty; \tag{1.2}$$

F is *regularly varying*, $F \in \mathcal{RV}$, if, for some $\beta > 0$, $\overline{F}(x) = x^{-\beta} L(x)$, where $L(x)$ is a *slowly varying* (at infinity) function;

F is *intermediate regularly varying*, $F \in \mathcal{IRV}$, if

$$\lim_{\alpha \uparrow 1} \limsup_{x \to \infty} \frac{\overline{F}(\alpha x)}{\overline{F}(x)} = 1. \tag{1.3}$$

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. [16] or [15] for further definitions and properties of heavy-tailed distributions.

Tail asymptotics of quantities related to R have earlier been studied in [19], [27] under RV conditions (see also [7]). Our main result is the following:

Theorem 1. *Assume* $\overline{n} < 1$ *and* $\overline{q} < \infty$. *Then:*

(i) *There is only one non-negative solution* R *to equation* (1.1) *with finite mean. For this solution,* $\overline{r} = \overline{q}/(1 - \overline{n})$.

(ii) *If further*

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

$$\mathbb{P}(R > x) \sim \frac{1}{1-\overline{n}}\mathbb{P}(Q + \overline{r}N > x) \qquad \text{as } x \to \infty\,. \tag{1.4}$$

(iii) *In particular, condition* **(C)** *holds in the following three cases:*
(a) (Q, N) *has a 2-dimensional regularly varying distribution;*
(b) Q *has an intermediate regularly varying distribution and* $\mathbb{P}(N > x) = \mathrm{o}(\mathbb{P}(Q > x))$*;*
(c) N *has an intermediate regularly varying distribution and* $\mathbb{P}(Q > x) = \mathrm{o}(\mathbb{P}(N > x))$.

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 [19], [27] in several ways. Motivated from the Google page rank algorithm, both of these papers consider a more general recursion

$$R \stackrel{\mathcal{D}}{=} Q + \sum_{m=1}^{N} A_m R_m\,. \tag{1.5}$$

However, [19] 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 require strong conditions on the A_i which do not allow 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 we have in mind. Also, our proofs are considerably simpler and shorter than those of [19], [27]. The key tool is a general result of [14] giving the tail asymptotics of the maximum of a random walk up to a (generalised) stopping time.

Remark 1. In Theorem 1, we considered the case $A_i \equiv 1$ only. However, our approach may work in the more general setting of (1.5) 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 $\overline{n} < 1$, that $\mathbb{P}(0 < A_1 \le 1) = 1$, then the exact tail asymptotics for $\mathbb{P}(R > x)$ may be easily found using the upper bound (1.4) and the principle of a single big jump. However, the formula for the tail asymptotics in this case is much more complicated that (1.4).

The multivariate version involves a set $\bigl(R(1),\ldots,(R(K)\bigr)$ of r.v.'s satisfying

$$R(i) \ \stackrel{\mathcal{D}}{=}\ Q(i)+\sum_{k=1}^{K}\sum_{m=1}^{N^{(k)}(i)} R_m(k) \tag{1.6}$$

In the branching 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 is the busy periods $R(i)$ in the multiclass queue in [4], 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}. One should note for this example [4] gives only lower asymptotic bounds, whereas we here provide sharp asymptotics.

The treatment of (1.6) is considerably more involved than for (1.1), and we defer the details of assumptions and results to Section 3. We remark here only that the concept of multivariate regular variation (MRV) will play a key role; that the analogue of the crucial assumption $\overline{n}<1$ above is subcriticality, $\rho=$spr$(\boldsymbol{M})<1$ where spr means spectral radius and $\boldsymbol{M}$ is the offspring mean matrix with elements $m_{ik}=\mathbb{E}N^{(k)}(i)$; and that the argument will involve a recursive procedure from [12, 13], reducing K to $K-1$ so that in the end we are back to the case $K=1$ of (1.1) and Theorem 1.

Bibliographical remarks An R, or its distribution, satisfying (1.5) is often called a fixed point of the smoothing transform (going back to [9]). 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 r.v.'s non-negative 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], [3] and references there.

It should be noted that the term "multivariate smoothing transform" (e.g. [6]) means to a recursion of vectors, that is, a version of (1.1) with $R,Q\in\mathbb{R}^K$. This is different from our set-up because in (1.6) we are only interested 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 multitype branching process is studied by different techniques in the critical case $\rho=1$.

2. One-dimensional case: equation (1.1)

The heuristics behind (1.4) 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 $\overline{r}N$, so roughly the probability of the first possibility is $\mathbb{P}(Q + \overline{r}N > x)$. On the other hand, results for compound heavy-tailed sums suggest that the approximate probability of the second possibility is $\overline{n}\mathbb{P}(R > x)$. We thus arrive at

$$\mathbb{P}(R > x) \approx \; \mathbb{P}(Q + \overline{r}N > x) \, + \, \overline{n}\mathbb{P}(R > x)$$

and (1.4).

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 (1.1). 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

$$\tau = \min\{n \ge 1 \;:\; S_n < 0\} = \min\{n \ge 1 \;:\; S_n = -1\}\,. \tag{2.1}$$

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

$$\mathbb{E}\tau = \frac{1}{1 - \mathbb{E}N} \tag{2.2}$$

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)

$$\tau \; \stackrel{\mathcal{D}}{=} \; 1 + \sum_{i=1}^{N} \tau_i \tag{2.3}$$

That is, τ is a solution to (1.1) with $Q \equiv 1$. On the other hand, the total progeny in a Galton-Watson process with the number of offsprings of an individual distributed as N obviously also satisfies (2.3), and hence by uniqueness must have the same distribution as τ. This result first occurs as equation (4) in [10], but note that an alternative representation (1) in that paper appears to have been the one receiving the most attention in the literature.

Now define $\varphi_i = k_0 + k_1 Q_i$,

$$V = \sum_{i=1}^{\tau} \varphi_i \tag{2.4}$$

Here the k_0, k_1 are non-negative constants, $k_0 + k_1 > 0$. In particular, if $k_0 = 1, k_1 = 0$, then $V = \tau$, and further

$$k_0 = 0, k_1 = 1 \qquad \Rightarrow \qquad V \stackrel{\mathcal{D}}{=} R. \tag{2.5}$$

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

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

where, like before, (φ, N), (φ_i, N_i), $(\varphi_{i,j}, N_{i,j})$, etc. are i.i.d. vectors. In particular, V becomes R when 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 $k_1 > 0$.

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 where $k_0 = 0$ and $k_1 = 1$,

$$r_0 = \frac{\mathbb{E}Q}{1 - \mathbb{E}N} = \overline{r}. \tag{2.6}$$

Choose $r > r_0$ as close to r_0 as needed and let

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

We will 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 [14] to obtain the following upper bound.

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

Here the first equivalence follows from [14], noting that the distribution of ψ_1 belongs to the class S^* and that [14] only requires $\varphi_1, \varphi_2, \ldots$ to be i.i.d. w.r.t. some filtration w.r.t. which τ is a stopping time. For the second, we used the long-tail property (1.2) of the distribution of ψ_1.

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

$$\overline{F}(x) \le \mathbb{P}(k_1 Q + rN > x) \le \mathbb{P}(rk_1 Q/r_0 + rN > x) \le \overline{F}(\alpha x) \le (1+o(1))c(\alpha)\overline{F}(x)$$

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

Now we assume the IRV condition to hold, let $r \downarrow r_0$ and apply (1.3) to obtain the upper bound

$$\mathbb{P}(R > x) \le (1+o(1))\mathbb{E}\tau\mathbb{P}(k_1 Q + r_0 N > x) \tag{2.7}$$

In particular, if $k_0 = 0$ and $k_1 = 1$, then $r_0 = \overline{r}$ is as in (2.6).

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

We have, for any fixed $C > 0$, $L > 0$, $n = 1, 2, \ldots$ and $x \ge 0$ that

$$\mathbb{P}(V > x) \;\ge\; \mathbb{P}\Big(\sum_{i=1}^{\tau} \psi_i > x\Big) \;\ge\; \sum_{i=1}^{n} \mathbb{P}(D_i \cap A_i) \tag{2.8}$$

where

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

Indeed, the first inequality in (2.8) holds since S_τ is non-positive. 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, (2.8) holds.

The events $\{A_i\}$ form a stationary sequence. Due to the SLLN, for any $\varepsilon > 0$, one 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

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

Since the random variables $(\{\psi_j\}_{j<i}, \mathbf{I}(\tau \le i))$ are independent of $\{\psi_j\}_{j \ge i}$, we obtain

further that, for any $\varepsilon \in (0,1)$ and for any $n \ge n_0$, $C \ge C_0$ and $L \ge L_0$,

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

as $x \to \infty$. Here the final equivalence follows from the long-tailedness of ψ_1. Letting first n go to infinity and then ε to zero, we get $\liminf_{x\to\infty} \mathbb{P}(V > x)/\mathbb{E}\tau\mathbb{P}(\psi_1 > x) \ge 1$. Then we let $r \uparrow r_0$ and use the IRV property (1.3). In the particular case $k_0 = 0, k_1 = 1$ we obtain an asymptotic lower bound that is equivalent to the upper bound derived above □

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, B_3 where

$$B_1 = B \cap \{\overline{r}N \; > \epsilon x\}, \quad B_2 = B \cap \{A < \overline{r}N \le \epsilon x\}, \quad B_3 = \{\overline{r}N \le A\}$$

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

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

We omit the full details since they are close to arguments given in Section 5 for the multivariate case.

3. Multivariate version

The assumptions for (1.6) are that all $R_m(k)$ are independent of the vector

$$\boldsymbol{V}(i) \;=\; \big(Q(i), N^{(1)}(i), \ldots, N^{(K)}(i)\big)\,, \tag{3.1}$$

that they are mutually independent and that $R_m(k) \stackrel{\mathcal{D}}{=} R(k)$. Recall that we are only interested in the one-dimensional distributions of the $R(i)$. Accordingly, for a solution to (1.6) we only require the validity for each fixed i.

Recall that the offspring mean matrix is denoted $\boldsymbol{M}$ where $m_{ik} = \mathbb{E}N^{(k)}(i)$, and that $\rho = \text{spr}(\boldsymbol{M})$; ρ is the Perron-Frobenius root if $\boldsymbol{M}$ is irreducible which it is not necessary to assume. No restrictions on the dependence structure of the vectors in (3.1) need to be imposed for the following result to hold (but later we need MRV!):

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

(i) *the fixed-point problem* (1.6) *has a unique non-negative solution with $\overline{r}_i = \mathbb{E}R(i) < \infty$ for all i;*

(ii) *the $\overline{r}_i = \mathbb{E}R(i) < \infty$ are given as the unique solution to the set*

$$\overline{r}_i \;=\; \overline{q}_i + \sum_{k=1}^{K} m_{ik}\overline{r}_k\,, \qquad i = 1, \ldots, K, \tag{3.2}$$

of linear equations.

Proof. (i) Assume first $Q(i) \equiv 1$, $i = 1, \ldots, K$. The existence of a solution to (1.6) 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 $\big(N^{(1)}(j), \ldots, N^{(K)}(j)\big)$. For uniqueness, let $\big(R(1), \ldots, R(K)\big)$ be any solution and consider the K-type Galton-Watson trees $\mathcal{G}(i)$, $i = 1, \ldots, K$, where $\mathcal{G}(i)$ corresponds to an ancestor of type i. If we define $R^{(0)}(i) = 1$,

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

with similar conventions as for (1.6), 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 that $R^{(n)}(i) \preceq_{\text{st}} R(i)$ ($\preceq_{\text{st}} =$ 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 total progeny of different types, and $R^{(\infty)}(i) \preceq_{\text{st}} R(i)$. Assuming the $R(i)$ to have finite mean, (3.2) clearly holds with $\overline{q}_i = 1$, and so the $\Delta_i = \overline{r}_i - \mathbb{E}R^{(\infty)}(i)$ 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 $R^{(\infty)}(i) \preceq_{\rm st} R(i)$ implies $R^{(\infty)}(i) \stackrel{\mathcal{D}}{=} R(i)$ 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) Just take expectations in (1.6) and 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. [22], [20], [5], [23], but we shall use the one in polar L_1-coordinates adapted to deal with several random vectors at a time as in (3.1). Fix here and in the following a reference RV tail $\overline{F}(x) = L(x)/x^\alpha$ on $(0,\infty)$, for $\boldsymbol{v} = (v_1,\dots,v_p)$ define $\|\boldsymbol{v}\| = \|\boldsymbol{v}\|_1 = |v_1| + \cdots + |v_p|$ and let $\mathcal{B} = \mathcal{B}_p = \{\boldsymbol{v} : \|\boldsymbol{v}\| = 1\}$. We then say that a random vector $\boldsymbol{V} = (V_1,\dots,V_p)$ satisfies MRV(F) or has property MRV(F) if $\mathbb{P}\big(\|\boldsymbol{V}\| > x\big) \sim b\overline{F}(x)$ where either (1) $b = 0$ or (2) $b > 0$ and the angular part $\boldsymbol{\Theta_V} = \boldsymbol{V}/\|\boldsymbol{V}\|$ satisfies

$$\mathbb{P}\big(\boldsymbol{\Theta_V} \in \cdot \,\big|\, \|\boldsymbol{V}\| > x\big) \ \stackrel{\mathcal{D}}{\longrightarrow}\ \mu \text{ as } x \to \infty$$

for some measure μ on $\mathcal{B}$ (the angular measure). Our basic condition is then that for the given reference RV tail $\overline{F}(x)$ we have that

(MRV) For any $i = 1,\dots,K$ the vector $\boldsymbol{V}(i)$ in (3.1) 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 $\overline{z}$ of F is finite, which will ensure that all expected values coming up in the following are finite.

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

$$\mathbb{P}\big(a_0Q(i) + a_1N^{(1)}(i) + \cdots + a_KN^{(K)}(i) > x\big) \ \sim\ c_i(a_0,\dots,a_K)\overline{F}(x) \tag{3.3}$$

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

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

$$\mathbb{P}(R(i) > x) \sim d_i \overline{F}(x) \quad \text{as } x \to \infty. \tag{3.4}$$

Here the d_i *are given as the unique solution to the set*

$$d_i = c_i(1, \overline{r}_1, \dots, \overline{r}_K) + \sum_{k=1}^{K} m_{ik} d_k\,, \qquad i = 1, \dots, K, \tag{3.5}$$

of linear equations where the $\overline{r}_i$ *are as in Proposition 1 and the* c_i *as in* (3.3)*.*

The proof follows in Sections 4–7.

4. Outline of proof

When $K > 1$, we did not manage to find a random walk argument extending Section 2. Instead, we shall use a recursive procedure, going back to [12, 13] in a queueing setting, for eventually being able to infer (3.4). The identification (3.5) of the d_i then follows immediately from the following result to be proved in Section 5 (the case $p = 1$ is Lemma 4.7 of [11]):

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

$$\mathbb{P}\big(S_{N_1}^{(1)} + \cdots + S_{N_p}^{(p)} > x\big) \sim \mathbb{P}(\overline{z}_1 N_1 + \cdots + \overline{z}_1 N_p > x) + c_0 \overline{F}(x)$$

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

The recursion idea in [12, 13] amounts in a queueing context to let all class K customers be served first. We implement it here in the branching setting. Consider the multitype Galton-Watson tree $\mathcal{G}$. For an ancestors 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 descendant 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 with the number of daughters distributed as $N^{(K)}(K)$ and starting from a single ancestor. In particular, the expected size of $\mathcal{G}_m(i)$ is $1/(1-m_{KK})$. We further have

$$R(i) \stackrel{\mathcal{D}}{=} \widetilde{Q}(i) + \sum_{k=1}^{K-1} \sum_{m=1}^{\widetilde{N}^{(k)}(i)} R_m(k)\,, \quad i = 1, \dots, K-1, \tag{4.1}$$

where

$$\widetilde{Q}(i) = Q(i) + \sum_{m=1}^{N^{(K)}(i)} \sum_{g \in \mathcal{G}_m(i)} Q_g(K)\,, \tag{4.2}$$

$$\widetilde{N}^{(k)}(i) = N^{(k)}(i) + \sum_{m=1}^{N^{(K)}(i)} \sum_{g \in \mathcal{G}_m(i)} N_g^{(k)}(K)\,. \tag{4.3}$$

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

Example 1. Let $K = 2$ and consider the 2-type family tree in Fig. 1, where type $i = 1$ has green color, the type 2 descendants of the ancestor in direct line red, and the remaining type 2 individuals blue. The green type 1 individuals marked with a triangle are the ones that are counted as extra type 1 children in the reduced recursion (4.1). We have $N^{(2)}(1) = 2$ and if m is the upper red individual of type 2, then $\mathcal{G}_m(2)$ has size 4. Further $\sum_{g\in\mathcal{G}_m(1)} N_g^{(1)} = 2$, with m herself and her upper daughter each contributing with one individual.

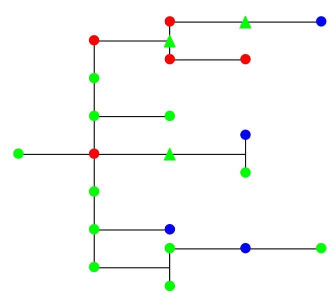

FIGURE 1: Reducing from 2 types to 1

The offspring mean in the reduced 1-type family tree is $\widetilde{m} = m_{11}+m_{12}m_{21}/(1-m_{22})$. 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 2-type tree is finite, the reduced 1-type tree must necessarily also be so, so that $\widetilde{m} \le 1$. A direct verification of this is instructive. First note that

$$\widetilde{m} \le 1 \iff m_{11} - m_{11}m_{22} + m_{12}m_{21} \le 1 - m_{22} \iff \mathrm{tr}(\boldsymbol{M}) - \det(\boldsymbol{M}) \le 1$$

But the characteristic polynomial of the 2-type offspring mean matrix $\boldsymbol{M}$ is $\lambda^2 - \lambda\,\mathrm{tr}(\boldsymbol{M}) + \det(\boldsymbol{M})$. Further the dominant eigenvalue ρ of $\boldsymbol{M}$ satisfies $\rho < 1$ so that

$$\mathrm{tr}(\boldsymbol{M}) - \det(\boldsymbol{M}) \ \le\ \rho\,\mathrm{tr}(\boldsymbol{M}) - \det(\boldsymbol{M}) \ =\ \rho^2 \ <\ 1.$$

5. Proof of Proposition 2

We shall need the following result of Nagaev et al. (see the discussion in [11] around equation (4.2) there for references):

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

$$\sup_{y \ge \delta k} \Bigl| \frac{\mathbb{P}(S_k > k\overline{z} + y)}{k\overline{F}(y)} - 1 \Bigr| \ \to\ 0,\ k \to \infty.$$

Corollary 1. *Under the assumptions of Lemma 1, it holds for $0 < \epsilon < 1/\overline{z}$ that*

$$d(F, \epsilon) \ =\ \limsup_{x\to\infty} \sup_{k<\epsilon x} \frac{\mathbb{P}(S_k > x)}{k\overline{F}(x)} \ <\ \infty$$

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

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

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

$$\overline{F}(y) \ \le\ \overline{F}(x\epsilon\delta) \ \sim (\epsilon\delta)^{-\alpha}\overline{F}(x)\,.$$

□

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

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

Here

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

Since $S^{(1)}_{k_1}$, $S^{(2)}_{k_2}$ are independent, we have by standard RV theory that

$$\mathbb{P}\big(S^{(1)}_{k_1} + S^{(2)}_{k_2} > x\big) \sim (k_1c_1 + k_2c_2)\overline{F}(x)$$

as $x \to \infty$. Further Corollary 1 gives that for $k_1, k_2 \le \epsilon x$ and all large x we have

$$\begin{aligned} \mathbb{P}\big(S^{(1)}_{k_1} + S^{(2)}_{k_2} > x\big) &\le \mathbb{P}\big(S^{(1)}_{k_1} > x/2\big) + \mathbb{P}\big(S^{(2)}_{k_2} > x/2\big) \\ &\le 2\big(d(F_1, 2\epsilon)k_1 + d(F_2, 2\epsilon)k_2\big)\overline{F}(x)\,. \end{aligned}$$

Hence by dominated convergence

$$\frac{p_1(x)}{\overline{F}(x)} \to \sum_{k_1,k_2=0}^{\infty} (k_1c_1 + k_2c_2)\mathbb{P}(N_1 = k_1, N_2 = k_2) = c_1\mathbb{E}N_1 + c_2\mathbb{E}N_2\,.$$

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

$$\begin{aligned} p_3(x) &\le \big(\mathbb{P}\big(A_1(\epsilon x)^{\mathrm{c}}\big) + \mathbb{P}\big(A_2(\epsilon x)^{\mathrm{c}}\big)\big)\mathbb{P}(N_1 > \epsilon x,\ N_2 > \epsilon x) \\ &\quad + \mathbb{P}\big(S^{(1)}_{N_1} + S^{(2)}_{N_2} > x,\ N_1 > \epsilon x,\ N_2 > \epsilon x, A_1(\epsilon x), A_2(\epsilon x)\big) \\ &\le r(\epsilon x)\mathrm{O}\big(\overline{F}(x)\big) + \mathbb{P}\big((\overline{z}_1N_1 + \overline{z}_2N_2)/(1-\epsilon) > x,\ N_1 > \epsilon x,\ N_2 > \epsilon x\big) \\ &\le \mathrm{o}\big(\overline{F}(x)\big)\mathbb{P}\big((\overline{z}_1N_1 + \overline{z}_2N_2) > \eta x,\ N_1 > \epsilon x,\ N_2 > \epsilon x\big) \end{aligned}$$

as $x \to \infty$, where $\eta < 1 - \epsilon$ will be specified later.

For $p_{21}(x)$, we write $p_{21}(x) = p'_{21}(x) + p''_{21}(x)$ where

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

with $\gamma = 2\epsilon\overline{z}_2$. Here

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

using the LLN in the fourth step. Further as in the estimates above

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

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

$$
\mathbb{P}\big(\overline{z}_1 N_1 + \overline{z}_2 N_2 > \eta x,\, N_1 \leq \epsilon x,\, N_2 \leq \epsilon x\big) \;=\; 0
$$

for ϵ small enough, we get

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

Letting $\epsilon \downarrow 0$ gives that the $\limsup$ is bounded by $c_0 + c(\overline{z}_1, \overline{z}_2)$. Similar estimates for the $\liminf$ complete the proof for $p = 2$.

If $p > 2$, the only essential difference is that $p_{21}(x), p_{22}(x)$ need to be replaced by the $2^p - 2$ terms corresponding to all combinations of some N_i being $\leq \epsilon x$ and the others $> \epsilon x$, with the two exceptions being the ones where either all are $\leq \epsilon x$ or all are $> \epsilon x$. However, to each of these similar estimates as the above ones 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 example of MRV. The first is two i.i.d. RV(F) r.v.'s X_1, X_2, where a big value of the $X_1 + X_2$ can only occur if one variable is big and the other small, which gives MRV with the angular measure concentrated on the points $(1,0)$, $(0,1) \in \mathcal{B}_2$ with mass 1/2 for each. Slightly more complicated:

Proposition 3. *Let $N, Z, Z_1, Z_2, \ldots$ be non-negative r.v.'s such that $N \in \mathbb{N}$, $Z, Z_1, Z_2, \ldots$ are i.i.d., non-negative and independent of N. Assume that $\mathbb{P}(N > x) \sim c_N \overline{F}(x)$, $\mathbb{P}(Z > x) \sim c_Z \overline{F}(x)$ for some RV tail $\overline{F}(x) = L(x)/x^\alpha$ and write $S = \sum_1^N Z_i$, $\overline{n} = \mathbb{E}N$, $\overline{z} = \mathbb{E}Z$, where $c_N + c_Z > 0$. Then:*

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

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

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

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

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

Proof. Part (i) is Lemma 4.7 of [11] (see also [8]). The proof in [11] also shows that is $S > x$, then either approximately $N\overline{z} > x$, occuring w.p. $c_N \overline{F}(x/\overline{z}) \sim c_N \overline{z}^\alpha \overline{F}(x)$, or $N \le \epsilon x$ and $Z_i > x$, occuring w.p. $c_Z \mathbb{E}[N \wedge \epsilon x]\overline{F}(x)$. The first possibility is what gives the atom of $\mu_{N,S}$ at b_1 and the second gives the atom at b_2 since $\mathbb{E}[N \wedge \epsilon x] \uparrow \overline{n}$. □

Proposition 4. *Let $\boldsymbol{V} = (\boldsymbol{T}, N) \in [0,\infty)^p \times \mathbb{N}$ satisfy* MRV(F)*, let $\boldsymbol{Z}, \boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots \in [0,\infty)^q$ be i.i.d. and 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 $\overline{\boldsymbol{z}} \in [0,\infty)^q$ be the mean of $\boldsymbol{Z}$. Similar arguments as in Section 5 show that $\|V^*\| > x$ basically occurs when either $\|\boldsymbol{T}\| + N + N\|\overline{\boldsymbol{z}}\| > x$ or when $\|\boldsymbol{V}\| \le \epsilon x$ and some $\|\boldsymbol{Z}_i\| > x$. The probabilities of these events are approximately of the form $c'\overline{F}(x)$ and $c''\overline{F}(x)$, so the radial part of $\boldsymbol{V}^*$ is RV with asymptotic tail $c_{\boldsymbol{V}^*}\overline{F}(x)$ where

$c_{\boldsymbol{V}^*} = c' + c''$. Now

$$\mathbb{P}\Big(\frac{(\boldsymbol{T}, N)}{\|(\boldsymbol{T}, N)\|} \in \cdot \,\Big|\, \|\boldsymbol{T}\| + N + N\|\overline{\boldsymbol{z}}\| > x\Big) \;\to\; \mu'$$

for some probability measure μ' on the $(p{+}1)$-dimensional unit sphere $\mathcal{B}_{p+1}$; this follows since $\|\boldsymbol{T}\| + N + N\|\overline{\boldsymbol{z}}\|$ is a norm and the MRV property of a vector is independent of the choice of norm. Letting δ_0' be Dirac measure at $(0, \ldots, 0) \in \mathbb{R}^q$, δ_0'' be Dirac measure at $(0, \ldots, 0) \in \mathbb{R}^{p+1}$ and $\mu'' = \mu_{\boldsymbol{Z}}$ 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

$$\mu_{\boldsymbol{V}^*} \;=\; \frac{c'}{c' + c''}\,\mu' \otimes \delta_0'' + \frac{c''}{c' + c''}\,\delta_0' \otimes \mu''$$

□

In calculations to follow (Lemma 2), extending some $\boldsymbol{V}$ to some $\boldsymbol{V}^*$ in a number of steps, expressions for $c_{\boldsymbol{V}^*}, \mu_{\boldsymbol{V}^*}$ can be deduced along the lines of the proof of Propositions 3–4 but the expression and details become extremely tedious. Fortunately, they won't be needed and are therefore omitted — all that matters is existence. If α is not an even integer, the MRV alone of $\boldsymbol{V}^*$ can alternatively (and slightly easier) be obtained from Theorem 1.1(iv) of [5], stating that by non-negativity it suffices to verify MRV of any linear combination.

7. Proof of Theorem 2 completed

Lemma 2. *In the setting of* (4.1)*, the random vector*

$$\boldsymbol{V}^*(i) \;=\; \big(\widetilde{Q}(i), \widetilde{N}^{(1)}(i), \ldots, \widetilde{N}^{(K-1)}(i)\big)$$

satisfies MRV(F) *for all i.*

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

$$M_1(i) \;=\; \sum_{m=1}^{N^{(K)}(i)} \big|\mathcal{G}_m(i)\big|\,,$$

$$M_2(i) \;=\; \sum_{m=1}^{N^{(K)}(i)} \sum_{g \in \mathcal{G}_m(i)} \big(Q_g(K), N_g^{(1)}(K), \ldots, N_g^{(1)}(K-1)\big)$$

Recall that our basic assumption is that the

$$\boldsymbol{V}^*(i) \;=\; \bigl(Q(i), N^{(1)}(i), \dots, N^{(K)}(i)\bigr) \tag{7.1}$$

satisfy MRV(F). The connection to a Galton-Watson tree and Theorem 1 with $Q \equiv 1$, $N = N^{(K)}(i)$ therefore imply that so does any $\bigl|\mathcal{G}_m(i)\bigr|$, and since these r.v.'s are i.i.d. and independent of $N^{(K)}(i)$, Proposition 4 gives that $\boldsymbol{V}_1(i) = \bigl(\boldsymbol{V}(i), M_1(i)\bigr)$ satisfies MRV(F). Now the MRV(F) property of (7.1) with $i = K$ implies that the vectors $\bigl(Q_g(K), N_g^{(1)}(K), \dots, N_g^{(K-1)}(K)\bigr)$, being distributed as $\bigl(Q(K), N^{(1)}(K), \dots, N^{(K-1)}(K)\bigr)$ 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 once more gives that $\boldsymbol{V}_2(i) = \bigl(\boldsymbol{V}(i), M_1(i), M_2(i)\bigr)$ 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 case $K = 1$ is just Theorem 1, so assume Theorem 2 shown for $K - 1$.

The induction hypothesis and Lemma 2 implies that $\mathbb{P}(R(i) > x) \sim d_i\overline{F}(x)$ for $i = 1, ..., K-1$. Rewriting (1.6) for $i = K$ as

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

we have a fixed-point problem of type (1.1) and can then use Theorem 2 to conclude that also $\mathbb{P}(R(K) > x) \sim d_K\overline{F}(x)$, noting that the needed MRV condition on $\bigl(Q^*(K), N^{(k)}(K)\bigr)$ follows by another application of Proposition 4.

Finally, to identify the d_i via (3.5), appeal to Proposition 2 with $\boldsymbol{N} = \bigl(Q(i), N^{(1)}(i), \dots, N^{(K)}(i)\bigr)$, writing the r.h.s. of (1.6) as

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

Existence and uniqueness of a solution to (3.5) follows by once more noticing that $\rho < 1$ implies that $\boldsymbol{I} - \boldsymbol{M}$ is invertible. □

References

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

[2] Alsmeyer, G., Biggins, J. and Meiners, M. (2012). The functional equation of the smoothing transform. *Ann. Probab.* **40**, 2069–2105.

[3] Alsmeyer, G. and Meiners, M. (2013). Fixed points of the smoothing transform: two-sided solutions. *Probab. Th. Rel. Fields* **155**, 165–199.

[4] Asmussen, S., Ernst, P. and Hasenbein, J. (2017). Stability and tail asymptotics in a multiclass queue with state-dependent arrival rates. In revision for *QUESTA*.

[5] Basrak, B., Davis, R.A. and Mikosch, T. (2002). A characterization of multivariate regular variation. *Ann. Appl. Probab.* **12**, 908–920.

[6] Buraczewski, D., Damek, E., Mentemeier, S. and Mirek, M. (2013). Heavy tailed solutions of multivariate smoothing transforms. *Stoch. Proc. Appl.* **123**, 1947–1986.

[7] Buraczewski, D., Damek, E. and Mikosch, T. (2016). *Stochastic Models with Power-Law Tails.* Springer-Verlag.

[8] Denisov, D., Foss, S. and Korshunov, D. (2010). Asymptotics of randomly stopped sums in the presence of heavy tails. *Bernoulli* **16**, 971–994.

[9] Durrett, R. and Liggett, T.M. (1983). Fixed points of the smoothing transformation. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **64**, 275–301.

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

[11] Faÿ, G., González-Arévalo, B., Mikosch, T. and Samorodnitsky, G. (2006). Modeling teletraffic arrivals by a Poisson cluster process. *QUESTA* **54**, 121–140.

[12] Foss, S. (1980). On systems with several classes of customers. *Theory of Probability and Its Applications* **25**, 135–137.

[13] Foss, S. (1984). Queues with customers of several types. In: *Limit Theorems and Related Problems*, 348–377. Optimization Software,

[14] 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. Probab.* **13**, 37–53.

[15] Foss, S., Korshunov, D. and Zachary, S. (2013). *An Introduction to Heavy-Tailed and Subexponential Distributions.* Springer, 2nd Edition.

[16] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events.* Springer.

[17] Jagers, P. (1975) *Branching Processes with Biological Applications.* Wiley.

[18] Jelenković, P. and Momcilović, P. (2004). Large deviations of square root insensitive random sums. *Math. Oper. Res.* **29**, 398–406.

[19] Jelenković, P.R. and Olvera-Cravioto, M. (2010). Information ranking and power laws on trees. *Adv. Appl. Probab.* **42**, 577–604.

[20] Meerschaert, M.M. and Scheffler, H.-P. (2001) *Limit Distributions for Sums of Independent Random Vectors. Heavy Tails in Theory and Practice.* Wiley

[21] 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. Probab.* **17**, 802–813.

[22] Resnick, S. (1987). *Extreme Values, Regular Variation, and Point Processes.* Springer-Verlag.

[23] Resnick, S. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.* Springer-Verlag.

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

[25] Seneta, E. (1974). Regularly varying functions in the theory of simple branching processes. *Adv. Appl. Probab.* **6**, 408–820.

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

[27] Volkovich, Y. and Litvak, N. (2010) Asymptotic analysis for personalized web search. *Adv. Appl. Probab.* **42**, 577–604.

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

Appendix A. Proof of (3.3)

The RV of linear combinations subject to MRV assumptions has received considerable attention, see e.g. [5], but we could not find explicit formulas like (3.3) 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 $\mathbb{P}\big(\|\boldsymbol{X}\| > t\big) \sim L(t)/t^\alpha$ 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

$$\mathbb{P}(\boldsymbol{a}\cdot\boldsymbol{X} > x) \;=\; \mathbb{P}(a_1X_1 + \cdots + a_nX_n > t) \;\sim\; \frac{L(t)}{t^\alpha}\int_{\mathcal{B}_1} \mathbb{I}(\boldsymbol{a}\cdot\boldsymbol{\theta} > 0)(\boldsymbol{a}\cdot\boldsymbol{\theta})^\alpha \mu(\mathrm{d}\boldsymbol{\theta})$$

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

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

which is the same as asserted.

For the rigorous proof, assume $\|\boldsymbol{a}\| = 1$. Then $\mathcal{B}_1$ is the disjoint union of the sets $B_{1,n}, \dots, B_{n,n}$ where $B_{i,n} = \{\boldsymbol{\theta} \in \mathcal{B}_1 : (i-1)/n < \boldsymbol{a}\cdot\boldsymbol{\theta} \le i/n\}$ for $i = 2, \dots, n$ and $B_{1,n} = \{\boldsymbol{\theta} \in \mathcal{B}_1 : \boldsymbol{a}\cdot\boldsymbol{\theta} \le 1/n\}$. Assuming $\mathbb{P}(\boldsymbol{\Theta} = i/n) = 0$ for all integers i, n, we get

$$\begin{aligned}\mathbb{P}(\boldsymbol{a}\cdot\boldsymbol{X} > t) \;&=\; \sum_{i=1}^n \mathbb{P}\big(\boldsymbol{a}\cdot\boldsymbol{X} > t, \boldsymbol{\Theta} \in B_{i,n}\big) \;\le\; \sum_{i=1}^n \mathbb{P}\big(\|\boldsymbol{X}\| > ti/n, \boldsymbol{\Theta} \in B_{i,n}\big) \\ &\sim\; \sum_{i=1}^n \frac{L(ti/n)}{(ti/n)^\alpha}\mathbb{P}\big(\boldsymbol{\Theta} \in B_{i,n} \,\big|\, |\boldsymbol{X}\| > t\big) \;\sim\; \frac{L(t)}{t^\alpha}\sum_{i=1}^n (i/n)^\alpha \mathbb{P}\big(\boldsymbol{\Theta} \in B_{i,n}\big) \\ &=\; \frac{L(t)}{t^\alpha}\int_{\mathcal{B}_1} f_{+,n}(\boldsymbol{\theta})\,\mu(\mathrm{d}\boldsymbol{\theta})\end{aligned}$$

where $f_{+,n}$ is the step function taking value $(i/n)^\alpha$ on $B_{i,n}$. A similar argument gives the asymptotic lower bound $\int f_{-,n}\,\mathrm{d}\mu\, L(t)/t^\alpha$ for $\mathbb{P}(\boldsymbol{a}\cdot\boldsymbol{X} > t)$ where $f_{-,n}$ equals $\big((i-1)/n)\big)^\alpha$ on $B_{i,n}$ for $i > 1$ and 0 on $B_{1,n}$. But $f_{\pm,n}(\boldsymbol{\theta})$ both have limits $\big[(\boldsymbol{a}\cdot\boldsymbol{\theta})^+\big]^\alpha$ as $n \to \infty$ and are bounded by 1. Letting $n \to \infty$ and using dominated convergence completes the proof.

The case $\mathbb{P}(\boldsymbol{\Theta} = i/n) > 0$ for some i, n is handled by a trivial redefinition of the $B_{i,n}$.