Probab. Theory Relat. Fields (2008) 141:445–470
DOI 10.1007/s00440-007-0090-1

Large deviations for sums indexed by the generations of a Galton–Watson process

Klaus Fleischmann · Vitali Wachtel

Received: 11 August 2006 / Revised: 5 June 2007 / Published online: 21 July 2007

Abstract In this paper, we study the large deviation behavior of sums S_{Z_n} of i.i.d. random variables X_i, where Z_n is the nth generation of a supercritical Galton–Watson process. We assume the finiteness of the moments EX_1^2 and $EZ_1 \log Z_1$. The underlying interplay of large deviation probabilities of partial sums of the X_i and of lower deviation probabilities of Z is clarified. Here, we heavily use lower deviation probability results on Z we recently published in [7].

Keywords Large deviation probabilities · Supercritical Galton–Watson processes · Lower deviation probabilities · Schröder case · Böttcher case · Lotka–Nagaev estimator

Mathematics Subject Classification (2000) Primary: 60J80; Secondary: 60F10

Contents

Supported by the German Science Foundation.
This paper has been written during the time the second author was a staff member of the WIAS Berlin.

K. Fleischmann (✉)
Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstr. 39,
10117 Berlin, Germany
e-mail: fleischm@wias-berlin.de

V. Wachtel
Technische Universität München Zentrum Mathematik, Bereich M5,
85747 Garching bei München, Germany
e-mail: wachtel@ma.tum.de

1 Introduction and results

1.1 Motivation

Let $Z = (Z_n)_{n\geq 0}$ denote a *Galton–Watson process* with offspring law $\{p_k;\ k \geq 0\}$. We will assume that Z is *supercritical*: $m := \sum_{k=1}^{\infty} kp_k \in (1, \infty)$. As a rule we start with $Z_0 = 1$.

A basic task in statistical inference of Galton–Watson processes is the estimation of the offspring mean m. Let us recall at this place the well-known Lotka–Nagaev estimator Z_{n+1}/Z_n of m due to Nagaev [10]. If $\varsigma := (\mathbf{V}\mathrm{ar} Z_1)^{1/2} \in (0, \infty)$, then for every $x \in \mathbb{R}$,

$$\lim_{n\uparrow\infty} \mathbf{P}\left(m^{n/2}\left(\tfrac{Z_{n+1}}{Z_n} - m\right) < x;\ Z_n > 0\right) = \int_0^{\infty} \Phi\left(\tfrac{xu^{1/2}}{\varsigma}\right) w(u)\, \mathrm{d}u, \tag{1}$$

where w denotes the continuous density function of the a.s. limit variable $W := \lim_{n\uparrow\infty} m^{-n} Z_n$ restricted to $\{W > 0\}$, and Φ is the standard normal distribution function,

$$\Phi(y) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} \mathrm{e}^{-z^2/2}\, \mathrm{d}z, \quad y \in \mathbb{R}. \tag{2}$$

The study of the ratio Z_{n+1}/Z_n has attracted the attention of several researchers in recent years, since it can also be used for estimating important parameters such as the amplification rate and the initial size in a quantitative polymerase chain reaction experiment; see Jacob and Peccoud [8,9].

Fix $k \geq 0$. Aimed to a finer description of the Galton–Watson model, let $Z_n(k)$ denote the number of particles in the nth generation having exactly k children. Then, on the event $\{Z_n > 0\}$, results for the estimator $\tilde{p}_k(n) := Z_n(k)/Z_n$ of p_k, which hold analogously to (1), had been provided by Pakes [15, Theorems 5 and 6].

The mentioned results from [10] and [15] can be seen from a unified point of view as follows. Independently of Z, let $X = (X_n)_{n\geq 1}$ denote a family of *i.i.d.* (*real-valued*) *random variables with mean zero and variance in* $(0, \infty)$. Let $n \geq 0$. Put $S_n := X_1 + \cdots + X_n$. On the event $\{Z_n > 0\}$, the random variable

$$R_n := S_{Z_n}/Z_n \tag{3}$$

is well-defined. For convenience, *we agree* that an event involving R_n is always tacitly assumed to be included in $\{Z_n > 0\}$. For instance, $\mathbf{P}(R_n < x)$ means $\mathbf{P}(R_n < x;\; Z_n > 0)$ more carefully written. If now X_1 coincides in law with $Z_1 - m$, then, for n fixed, R_n coincides in law with $Z_{n+1}/Z_n - m$ on the event $\{Z_n > 0\}$. On the other hand, if X_1 takes on the value $1 - p_k$ with probability p_k (for k fixed) and $-p_k$ otherwise, then for n fixed, we have $R_n = \tilde{p}_k(n) - p_k$ in law on the event $\{Z_n > 0\}$.

Sums such as S_{Z_n} arise also in models of polymerase chain reactions with mutations, see Piau [17].

From now on, as a rule we work with the more general meaning of R_n, based on (X, Z), as introduced in (3). Clearly, we have the following *strong law of large numbers*:

$$R_n \to 0 \quad \text{a.s.} \quad \text{as} \quad n \uparrow \infty. \tag{4}$$

Moreover, using methods from [10] and [15], one can easily verify the following "*normal deviation probabilities*" for R_n :

$$\lim_{n\uparrow\infty} \mathbf{P}\left(m^{n/2} R_n < x\right) = \int_0^\infty \Phi\left(\tfrac{x u^{1/2}}{\sigma}\right) w(u)\, \mathrm{d}u, \quad x \in \mathbb{R}, \tag{5}$$

where $\sigma := (\mathbf{E}X_1^2)^{1/2}$ from now on. Let $\varepsilon_n > 0$ and consider $\mathbf{P}(R_n \geq \varepsilon_n)$. In the case $\varepsilon_n m^{n/2} \to \infty$, statement (5) implies the following simple *large deviation probabilities* for R_n :

$$\lim_{n\uparrow\infty} \mathbf{P}(R_n \geq \varepsilon_n) = 0. \tag{6}$$

But the main task of large deviation theory is to determine the *rate* of such convergence. Clearly, one of the reasons to be interested in large deviation probabilities comes from statistical applications. On the one hand, these probabilities describe the quality (error probabilities) of many tests. On the other hand, a question concerning the Bahadur efficiency of estimators leads also to a large deviation problem.

For the particular model $X_1 \stackrel{\mathcal{L}}{=} Z_1 - m$ mentioned above, the special case $\varepsilon_n \equiv \varepsilon$ is more or less studied in the literature. In fact, Athreya [3] proved that if $p_0 = 0$, $p_1 > 0$, and $\mathbf{E}Z_1^{2\alpha+\delta} < \infty$ for some $\delta > 0$, where $\alpha \in (0, \infty)$ denotes the so-called Schröder constant [see (8) below], then

$$\lim_{n\uparrow\infty} m^{\alpha n}\, \mathbf{P}\big(|R_n| \geq \varepsilon\big) \;\; \text{exists finitely.} \tag{7}$$

On the other hand, using asymptotic properties of harmonic moments of Z_n , Ney and Vidyashankar [12] found the rate of $\mathbf{P}\big(|R_n| \geq \varepsilon\big)$ under the weaker assumption that $\mathbf{P}(Z_1 \geq j) \sim a j^{1-\eta}$ as $j \uparrow \infty$, for some $\eta > 2$ and $a > 0$. The same authors proved in [13] a version of a large deviation principle for R_n conditioned on $Z_n \geq v_n$ with numbers $v_n \to \infty$; see also Rouault [18].

The *purpose of the present paper* is to study the rate of convergence of (large deviation) probabilities of $R_n \geq \varepsilon_n$ in the more interesting case $\varepsilon_n \to 0$ as $n \uparrow \infty$ (working with our more general setting of R_n). For this we heavily relay on results on lower deviation probabilities of Z, we recently established in [7]. In the next section we briefly recall what we need from that paper.

Note that large deviation probabilities in the case $\varepsilon_n \to 0$ are needed, for instance, for testing two close hypotheses, i.e. when the distance between the hypotheses tends to zero as the size of the sample gets larger and larger.

1.2 Lower deviation probabilities for Z

We start with recalling the following basic notation, reflecting a crucial dichotomy for supercritical Galton–Watson processes.

Definition 1 (Schröder and Böttcher case). For our supercritical offspring distribution we distinguish between the *Schröder* and the *Böttcher* case, in dependence on whether $p_0 + p_1 > 0$ or $= 0$, respectively.

Write f for the generating function of our supercritical offspring law: $f(s) = \sum_{j\geq 0} p_j s^j$, $0 \leq s \leq 1$. Let q denote the extinction probability of Z,

$$\text{set } \gamma := f'(q), \quad \text{and define } \alpha \text{ by } \quad \gamma = m^{-\alpha}. \tag{8}$$

Note that $\gamma \in [0, 1)$ and $\alpha \in (0, \infty]$. Obviously, we are in the Schröder case if and only if $\gamma > 0$, if and only if $\alpha < \infty$. In the latter case, α is said to be the *Schröder constant.* We also need the following notion.

Definition 2 (Type (d, μ)). We say the offspring distribution *is of type* (d, μ), if $d \geq 1$ is the greatest common divisor of the set $\{j - \ell : \ j \neq \ell, \ p_j p_\ell > 0\}$, and $\mu \geq 0$ is the minimal j for which $p_j > 0$.

In the present paper, (d, μ) *always refers* to the type of our offspring law. Recall that $\mu \geq 2$ in the Böttcher case. Here the *Böttcher constant* $\beta \in (0, 1)$ is defined by $\mu = m^\beta$.

We also *always assume* that the moment $\mathbf{E}Z_1 \log Z_1$ is finite. Under this moment condition, the lower deviation results of [7, Corollary 5 and Theorem 6] can be specified to the following two propositions.

Proposition 3 (Schröder case). *In the Schröder case, for $k_n \leq m^n$ satisfying $k_n \to \infty$ as $n \uparrow \infty$, we have*

$$\sup_{k\in[k_n,m^n] \text{ with } k\equiv\mu \,(\mathrm{mod}\, d)} \left| \frac{m^n}{d\, w\,(k/m^n)} \mathbf{P}(Z_n = k) - 1 \right| \underset{n\uparrow\infty}{\longrightarrow} 0 \tag{9}$$

and

$$\sup_{k\in[k_n,m^n]} \left| \frac{\mathbf{P}(0 < Z_n \leq k)}{\mathbf{P}\,(0 < W < k/m^n)} - 1 \right| \underset{n\uparrow\infty}{\longrightarrow} 0. \tag{10}$$

Proposition 4 (Böttcher case). *Suppose the Böttcher case. Then there exist positive constants B_1 and B_2 such that for all $k_n \geq \mu^n$ with $k_n = o(m^n)$ as $n \uparrow \infty$,*

$$-B_1 \leq \liminf_{n\uparrow\infty} (k_n/m^n)^{\beta/(1-\beta)} \log \mathbf{P}(Z_n \leq k_n) \tag{11a}$$

$$\leq \limsup_{n\uparrow\infty} (k_n/m^n)^{\beta/(1-\beta)} \log \mathbf{P}(Z_n \leq k_n) \leq -B_2 . \tag{11b}$$

Inequalities (11) *remain true if* $\mathbf{P}(Z_n \leq k_n)$ *is replaced by* $m^n\,\mathbf{P}(Z_n = k_n)$, *provided that additionally* $k_n \equiv \mu^n \pmod d$.

In order to explain the influence of lower deviation probabilities of Z_n on $R_n = S_{Z_n}/Z_n$, look at the decomposition,

$$\mathbf{P}(R_n \geq \varepsilon_n) = \sum_{k=1}^{\infty} \mathbf{P}(Z_n = k)\,\mathbf{P}(S_k \geq \varepsilon_n k). \tag{12}$$

Thus, in order to find the asymptotics of $\mathbf{P}(R_n \geq \varepsilon_n)$, we need to determine the range of values of k, which give the main contribution in decomposition (12). As we will see, this depends on parameters of the offspring law (as α, for instance) and, on the other hand, on the tail behavior of X_1 . Here we mention several possibilities of the interplay. If k is of order m^n (the regime of normal deviations for Z_n) and $\varepsilon_n^2 m^n \to \infty$, then $\varepsilon_n k$ has to be in the domain of large deviations of S_k . On the other hand, if k is of order ε_n^{-2} (regime of normal deviations for S_k), then k has to be in the domain of lower deviations for Z_n . And finally, if $k/m^n \to 0$ and $\varepsilon_n^2 k \to \infty$, then simultaneously we need lower deviations for Z_n and large deviations for S_k .

1.3 Large deviations in the Schröder case

In the remainder of the paper we consider

$$\varepsilon_n > 0 \quad \text{with} \quad \varepsilon_n \to 0 \quad \text{and} \quad \varepsilon_n^2 m^n \to \infty \text{ as } n \uparrow \infty. \tag{13}$$

Recall that we always assume $\mathbf{E}Z_1 \log Z_1 < \infty$ and $\mathbf{E}X_1^2 < \infty$. As usual, we set $X_1^+ := X_1 \vee 0$. We say that X_1^+ *has a tail of index* θ, if for some constant $a > 0$,

$$\mathbf{P}(X_1 \geq x) \sim a\,x^{-\theta} \quad \text{as} \quad x \uparrow \infty. \tag{14}$$

(Here the involved constant is always denoted by a.) Define

$$\varkappa := \frac{1+\alpha-\theta}{2\alpha-\theta}\,. \tag{15}$$

Here is the *main result* of our paper.

Theorem 5 (Schröder case). *Suppose the Schröder case* (*i.e.* $0 < \alpha < \infty$).

(a) *If*

$$\mathbf{E}(X_1^+)^{1+\alpha} \ < \ \infty, \tag{16}$$

or if X_1^+ *has a tail of index* $\theta \in (2, 1+\alpha)$ $($*with* $1 < \alpha < \infty)$ *as well as* $\varepsilon_n m^{\varkappa n} \to 0$ *as* $n \uparrow \infty$, *then*

$$0 \ < \ V_* \Gamma_\alpha \leq \liminf_{n\uparrow\infty} \ \varepsilon_n^{2\alpha} \, m^{\alpha n} \, \mathbf{P}(R_n \geq \varepsilon_n) \tag{17a}$$

$$\leq \limsup_{n\uparrow\infty} \ \varepsilon_n^{2\alpha} \, m^{\alpha n} \, \mathbf{P}(R_n \geq \varepsilon_n) \ \leq \ V^* \Gamma_\alpha \ < \ \infty, \tag{17b}$$

where

$$V_* \ := \ \liminf_{u\downarrow 0} \ u^{1-\alpha} w(u), \quad V^* \ := \ \limsup_{u\downarrow 0} \ u^{1-\alpha} w(u) \tag{18}$$

and

$$\Gamma_\alpha \ := \ \frac{2^{\alpha-1}\,\Gamma(\alpha+1/2)}{\alpha\sqrt{\pi}}\,\sigma^{2\alpha}. \tag{19}$$

(b) *If* X_1^+ *has a tail of index* $\theta \in (2, 1+\alpha)$ *and* $\varepsilon_n m^{\varkappa n} \to \infty$, *then*

$$\lim_{n\uparrow\infty} \varepsilon_n^\theta \, m^{(\theta-1)n} \, \mathbf{P}(R_n \geq \varepsilon_n) \ = \ a \, I_\theta \, , \tag{20}$$

where

$$I_\theta \ := \ \frac{1}{\Gamma(\theta-1)} \int_0^\infty \varphi(v)\, v^{\theta-2}\, \mathrm{d}v. \tag{21}$$

(c) *If* X_1^+ *has a tail of index* $\theta \in (2, 1+\alpha)$ *and* $\varepsilon_n m^{\varkappa n} \to$ *some* $\tau^{-1} \in (0, \infty)$, *then*

$$\begin{aligned} \tau^{2\alpha} V_* \Gamma_\alpha + \tau^\theta a \, I_\theta \ &\leq \liminf_{n\uparrow\infty} \ m^{\alpha(\theta-2)n/(2\alpha-\theta)} \, \mathbf{P}(R_n \geq \varepsilon_n) \\ &\leq \limsup_{n\uparrow\infty} m^{\alpha(\theta-2)n/(2\alpha-\theta)} \, \mathbf{P}(R_n \geq \varepsilon_n) \\ &\leq \tau^{2\alpha} V^* \Gamma_\alpha + \tau^\theta a \, I_\theta \, . \end{aligned} \tag{22}$$

Of course, here $\Gamma(\cdot)$ refers to the Gamma function.

Remark 6 (*Case* $\alpha \ \leq \ 1$). Because of our general assumption $0 \ < \ \mathbf{E}X_1^2 \ < \ \infty$, condition (16) can be dropped in the case $\alpha \leq 1$. For these values of the Schröder constant, part (a) describes all possible large deviation probabilities, i.e. for any choice of ε_n and X_1 satisfying our general assumptions. On the other hand, for $\alpha > 1$ the rates of large deviations may depend on the tail of X_1^+ and on the speed of ε_n .

Remark 7 (*Critical value of* θ). If $1 < \alpha < \infty$, Theorem 5 leaves open the case that X_1^+ has a tail of index $\theta = \alpha + 1$. Our methods allow to prove that part (a) holds, if $\varepsilon_n n^{1/(\alpha-1)} \to 0$. On the other hand, if $\varepsilon_n n^{1/(\alpha-1)} \to \infty$, then

$$\lim_{n\uparrow\infty} n^{-1}\, \varepsilon_n^{1+\alpha}\, m^{\alpha n}\, \mathbf{P}(R_n \geq \varepsilon_n) \;=\; a\, J_\alpha \tag{23}$$

where

$$J_\alpha \;:=\; \frac{1}{\Gamma(\alpha)} \int_1^m \mathsf{S}\big(\varphi(v)\big)\, v^{\alpha-1}\, \mathrm{d}v. \tag{24}$$

Finally, if $\varepsilon_n n^{1/(\alpha-1)} \to \tau^{-1} \in (0,\infty)$ holds, then a similar statement as in (c) is true.

Under the assumptions in part (a) (of Theorem 5), the sum at the right hand side of (12) is determined by those values of k which are of order ε_n^{-2}. As we already mentioned, this corresponds to lower deviations of Z and normal deviations of S_k . Large deviations as in part (b) have a different structure: the main contribution comes from k of order m^n, which corresponds to normal deviations of Z_n and large deviations of S_k . In part (c) we have a combination of regimes appearing in (a) and (b): the values of k of orders ε_n^{-2} and m^n contribute at the same level.

In the proof of Theorem 5 (in Sect. 3.1), we shall split the sum in decomposition (12) according to the structure of large deviations as just described:

part (a): $k \in (0,\ \delta/\varepsilon_n^2],\ k \in (\delta/\varepsilon_n^2,\ A/\varepsilon_n^2],\ \ k \in (A/\varepsilon_n^2,\ \infty)$;
part (b): $k \in (0,\ \delta m^n],\ \ k \in (\delta m^n,\ \infty)$;
part (c): $k \in (0,\ \delta/\varepsilon_n^2],\ k \in (\delta/\varepsilon_n^2,\ A/\varepsilon_n^2],\ k \in (A/\varepsilon_n^2,\ \delta m^n],\ \ k \in (\delta m^n,\ \infty)$;

where $\delta \in (0,1)$ and $A \geq 1$ are constants. The interplay between Z_n and S_k in each of these cases will be considered in Sect. 2.2.

Next we recall some known facts on the asymptotic behavior of supercritical Galton–Watson processes in the Schröder case. With q and γ introduced in the beginning of Sect. 1.2 and with f_n denoting the iterates of f, the following limit exists:

$$\lim_{n\uparrow\infty} \frac{f_n(s)-q}{\gamma^n} \;=:\; \mathsf{S}(s) \;=:\; \sum_{j=0}^{\infty} \nu_j s^j, \quad 0 \leq s < 1. \tag{25}$$

Hence,

$$\lim_{n\uparrow\infty} \gamma^{-n}\, \mathbf{P}(Z_n = k) \;=\; \nu_k\,, \quad k \geq 1. \tag{26}$$

The Schröder constant $\alpha < \infty$ describes the behavior of the density function $w(u)$ as $u \downarrow 0$. In fact, according to Biggins and Bingham [5], there is a continuous, positive multiplicatively periodic function V such that

$$u^{1-\alpha} w(u) \;=\; V(u) + o(1) \quad \text{as} \quad u \downarrow 0. \tag{27}$$

The function V in (27) can be replaced by a (positive) constant V_0 if and only if

$$\mathsf{S}\big(\varphi(h)\big) = \; V_0 h^{-\alpha}, \quad h \geq 0, \tag{28}$$

where φ denotes the Laplace transform of the limit random variable W (cf. Asmussen and Hering [1, p. 96]. In this case, $V^* = V_* = V_0$ in Theorem 5, that is, we get the following conclusion.

Corollary 8 (Schröder under an additional regularity of Z). *Suppose that* (28) *holds. Then, under the assumptions of Theorem* 5(a)*,*

$$\lim_{n\uparrow\infty} \varepsilon_n^{2\alpha} m^{\alpha n} \mathbf{P}(R_n \geq \varepsilon_n) \;=\; V_0 \, \Gamma_\alpha \tag{29}$$

[*with* Γ_α *from* (19)]. *Moreover, under the assumptions of Theorem* 5(c)*,*

$$\lim_{n\uparrow\infty} m^{\alpha(\theta-2)n/(2\alpha-\theta)} \mathbf{P}(R_n \geq \varepsilon_n) \;=\; \tau^{2\alpha} V_0 \, \Gamma_\alpha + \tau^\theta a \, I_\theta \tag{30}$$

[*with* I_θ *from* (21)].

1.4 Large deviations in the Böttcher case

As well-known, in the Böttcher case the following limit

$$\lim_{n\uparrow\infty} (f_n(s))^{(\mu^{-n})} \;=:\; \mathsf{B}(s), \quad 0 \leq s \leq 1, \tag{31}$$

exists, is positive and continuous [with $\mu \geq 2$ from Definition 2]. From this it follows that in general $f_n(s)$ does not converge as $n \uparrow \infty$. But taking logarithms, we have

$$\lim_{n\uparrow\infty} \mu^{-n} \log f_n(s) \;=\; \log \mathsf{B}(s). \tag{32}$$

On the other hand, our result on lower deviations in the Böttcher case (Proposition 4) is also only for log-scaled probabilities. These two facts explain the use of a logarithmic scaling in our following theorem.

Theorem 9 (Böttcher under light tails concerning X_1). *Assume the Böttcher case and that* $\mathbf{E}e^{h|X_1|}$ *is finite for some* $h > 0$. *Then*

$$\mu \log \mathsf{B}\Big(\varphi(1/2\sigma^2)\Big) \le \liminf_{n\uparrow\infty} \varepsilon_n^{-2\beta} m^{-\beta n} \log \mathbf{P}(R_n \ge \varepsilon_n) \tag{33a}$$

$$\le \limsup_{n\uparrow\infty} \varepsilon_n^{-2\beta} m^{-\beta n} \log \mathbf{P}(R_n \ge \varepsilon_n) \le \mu^{-1} \log \mathsf{B}\Big(\varphi(1/2\sigma^2)\Big). \tag{33b}$$

If, additionally, $\varepsilon_n = m^{-\lambda_n/2}$ *for integers* $\lambda_n \to \infty$ *with* $\lambda_n = o(n)$ *as* $n \uparrow \infty$, *then*

$$\lim_{n\uparrow\infty} \varepsilon_n^{-2\beta} m^{-\beta n} \log \mathbf{P}(R_n \ge \varepsilon_n) \;=\; \log \mathsf{B}\Big(\varphi(1/2\sigma^2)\Big). \tag{34}$$

According to this theorem, the main contribution to $\mathbf{P}(R_n \ge \varepsilon_n)$ comes from lower deviations of Z_n and large deviations of S_n. In order to explain this heuristically, we note that by Proposition 4 there exist (positive and finite) constants $c_1 \ge c_2$ such that

$$\exp\Big[-c_1\,(k/m^n)^{-\beta/(1-\beta)}\Big] \;\le\; m^n\, \mathbf{P}(Z_n = k) \;\le\; \exp\Big[-c_2\,(k/m^n)^{-\beta/(1-\beta)}\Big]. \tag{35}$$

On the other hand (for details see the proof of Theorem 9 in Sect. 3.2 below),

$$\exp[-c_3\,\varepsilon_n^2 k] \;\le\; \mathbf{P}(S_k \ge \varepsilon_n k) \;\le\; \exp[-c_4\,\varepsilon_n^2 k] \tag{36}$$

for some $c_3 \ge c_4$. Then, roughly speaking,

$$\mathbf{P}(R_n \ge \varepsilon_n) \;\sim\; m^{-n} \sum_{k=\mu^n}^{\infty} \exp\Big[-a\,(k/m^n)^{-\beta/(1-\beta)} - b\,\varepsilon_n^2 k\Big] \tag{37}$$

with $a, b > 0$. Obviously, the value of this sum is determined, in a sense, by the maximal summand. It can now easily be seen, that the function

$$g(u) \;:=\; a\,(u/m^n)^{-\beta/(1-\beta)} + b\,\varepsilon_n^2 u, \quad u > 0, \tag{38}$$

achieves its minimum at $u_* := c\,\varepsilon_n^{-2(1-\beta)} m^{n\beta}$ [with c we always denote a constant which might change its value from place to place], and consequently,

$$g(u_*) \;=\; c\,\varepsilon_n^{2\beta} m^{n\beta}. \tag{39}$$

This is in line with the normalizing sequence in Theorem 9 (except a constant factor). Evidently, the values k of order $\varepsilon_n^{-2(1-\beta)} m^{n\beta}$ correspond to lower deviations of Z_n and large deviations of S_k.

If we put formally $\alpha = \infty$ in the conditions in Theorem 5 (b) (passing to the Böttcher case), then (20) should hold under the condition $\varepsilon_n m^{n/2} \to \infty$, since $\varkappa \to 1/2$ as $\alpha \uparrow \infty$. But we prove it only under a slightly stronger condition on ε_n:

Theorem 10 (Böttcher under heavier tails concerning X_1^+). *Suppose the Böttcher case and that X_1^+ has a tail of index $\theta > 2$. If $\varepsilon_n m^{n/2} n^{-1/2\beta} \to \infty$, then (20) is true.*

There is the same "philosophy" behind Theorem 10 as it is behind Theorem 5(b). The main influence of normal deviations of Z_n explains also the independence of (20) of the parameters α and β. Note also that in the special case $\varepsilon_n \equiv \varepsilon$, Theorem 5(b) was proved in [12].

We stress the fact, that our results in the Böttcher case are weaker than those in the Schröder case. In fact, in the case of light tails of X_1^+, we found only log-scaled asymptotics for large deviation probabilities. Moreover, in the case of regularly varying tails, we have additional restrictions on ε_n. Finally, there is a gap between the tail conditions in Theorems 9 and 10.

Remark 11 (*Possible generalizations*). Many conditions in our results are too restrictive, but allow us to make proofs slightly shorter and clearer. Here we mention some (almost evident) generalizations of our theorems.

(a) It is possible to prove versions of Theorem 5 for X_1 from the domain of attraction of a stable law of any index.
(b) Theorems 5 and 10 can be generalized to the case $\mathbf{P}(X_1 \geq x) = L(x)x^{-\theta}$ with some L slowly varying at infinity.
(c) We conjecture that condition $\mathbf{E}Z_1 \log Z_1 < \infty$ can be dropped in all of our theorems. In fact, we need it only for inequality (42) below, taken from Theorem II.4.2 of Athreya and Ney [2]. But it should be possible to prove this bound for all supercritical Galton–Watson processes.
(d) In [13], $\mathbf{P}\left(Z_n \geq \varepsilon_n\,;\ Z_n \geq v_n\right)$ is considered with $v_n \to \infty$ and $\varepsilon_n \equiv \varepsilon$. Our methods allow to deal with the case $v_n = o(m^n)$ and $\varepsilon_n \to 0$.

Remark 12 (*On critical Galton–Watson processes*). For the moment, suppose that the Galton–Watson process Z is critical, that is, $m = 1$. Furthermore, assume that $\varsigma^2 := \mathbf{V}\mathrm{ar} Z_1 \in (0, \infty)$. Then, analogously to (5),

$$\lim_{n\uparrow\infty} \mathbf{P}\left(n^{1/2} R_n < x \,\middle|\, Z_n > 0\right) = \frac{2}{\varsigma^2} \int_0^\infty \Phi\left(\tfrac{x u^{1/2}}{\sigma}\right) \mathrm{e}^{-2u/\varsigma^2}\, \mathrm{d}u. \tag{40}$$

For the proof of this convergence in the two special cases of X_1 as mentioned in Sect. 1.1, see [10] and [15], respectively. From (40) we find that for critical processes the domain of large deviations is defined by the relation $\varepsilon_n^2 n \to \infty$ as $n \uparrow \infty$. The special case $\varepsilon_n \equiv \varepsilon$ was treated by Athreya and Vidyashankar [4]. If now $\varepsilon_n \to 0$ and $\varepsilon_n^2 n \to \infty$, then

$$\lim_{n\uparrow\infty} \varepsilon_n^2 n\, \mathbf{P}\left(R_n \geq \varepsilon_n \,\middle|\, Z_n > 0\right) = \frac{\sigma^2}{\varsigma^2}\,. \tag{41}$$

Actually, (41) is similar to the statement of Theorem 5(a) in the case $\alpha = 1$ and if m^n is replaced by the order n of $\mathbf{E}\{Z_n \,|\, Z_n > 0\}$. Also, the proof of (41) is close to the proof of Theorem 5(a) in the case $\alpha = 1$. There are only two differences. First, instead of (42) below, we have to use $\mathbf{P}\left(Z_n = k \,\middle|\, Z_n > 0\right) \le c\,n^{-1}$, which is derived in Nagaev and Vakhtel [14]. Second, we have to apply the local limit theorem for critical Galton–Watson processes instead of Proposition 3. For the proof of this local limit theorem under a second moment assumption, see [14].

2 Auxiliary results

In this section we prepare for the proofs of our theorems.

2.1 Separate considerations

As a first step, we state two bounds for local probabilities of our supercritical Galton–Watson process Z (satisfying $\mathbf{E}Z_1 \log Z_1 < \infty$).

Lemma 13 (Local probabilities of Z). *There is a constant c such that*

$$\mathbf{P}\left(Z_n = k \,\middle|\, Z_0 = \ell\right) \;\le\; c\,\frac{\ell}{k}, \quad k, \ell, n \ge 1. \tag{42}$$

Moreover, in the Schröder case, again for some constant c,

$$\mathbf{P}\left(Z_n = k \,\middle|\, Z_0 = 1\right) \;\le\; c\,\frac{k^{\alpha-1}}{m^{\alpha n}}, \quad k, n \ge 1. \tag{43}$$

Proof For aperiodic ($d = 1$) offspring laws, inequality (42) follows from the proof of Theorem II.4.2 in [2]. Indeed, from the last formula on p. 81 there, the inequality

$$2\pi\, k\,\mathbf{P}\left(Z_n = k \,\middle|\, Z_0 = \ell\right) \;\le\; \ell \int_{-\pi m^n}^{\pi m^n} m^{-n}\left|f_n'(\mathrm{e}^{iu/m^n})\right| \mathrm{d}u \tag{44}$$

follows, and the boundedness of this integral is shown in the end of that proof. The remaining case $d > 1$ can be dealt with in a similar way.

In proving (43) it is sufficient to assume that $k \le m^n$, otherwise (43) follows from (42). Under the present condition $\mathbf{E}Z_1 \log Z_1 < \infty$, formula (151) in [7] with $N = \ell_0 := 1 + [1/\alpha]$ and $j = n - a_k$ where $a_k := \min\{j \ge 1 : m^j \ge k\}$ gives

$$\sum_{\ell=\ell_0}^{\infty} \mathbf{P}(Z_{n-a_k} = \ell)\,\mathbf{P}\left(Z_{a_k} = k \,\middle|\, Z_0 = \ell\right) \;\le\; \frac{c}{m^{a_k}}\, f_{n-a_k}(\mathrm{e}^{-\delta}), \tag{45}$$

since $m^{a_k} \ge k$. It follows from (25) that the right hand side is bounded by $c\,m^{-a_k}\gamma^{n-a_k}$. Since

$$k \;\le\; m^{a_k} \;\le\; mk \quad \text{and} \quad \gamma = m^{-\alpha}, \tag{46}$$

we get the bound

$$\sum_{\ell=\ell_0}^{\infty} \mathbf{P}(Z_{n-a_k}=\ell)\,\mathbf{P}\left(Z_{a_k}=k \,\middle|\, Z_0=\ell\right) \;\le\; c\,\frac{k^{\alpha-1}}{m^{\alpha n}}\,. \tag{47}$$

If $\ell_0=1$, then the proof of (43) is complete, since the left hand side in (47) equals $\mathbf{P}(Z_n=k)$. Assume now that $\ell_0\ge 2$. From (42) it follows that

$$\sum_{\ell=1}^{\ell_0-1} \mathbf{P}(Z_{n-a_k}=\ell)\,\mathbf{P}\left(Z_{a_k}=k \,\middle|\, Z_0=\ell\right) \;\le\; c\,\frac{\ell_0}{k} \sum_{\ell=1}^{\ell_0-1} \mathbf{P}(Z_{n-a_k}=\ell). \tag{48}$$

By (26), $\lim_{n\uparrow\infty}\gamma^{-n}\,\mathbf{P}(Z_n=\ell)=\nu_\ell<\infty$, for every fixed ℓ. Hence,

$$\sum_{\ell=1}^{\ell_0-1} \mathbf{P}(Z_{n-a_k}=\ell) \;\le\; c\,\gamma^{n-a_k} \tag{49}$$

for all $n\ge 1$. Using again (46), we get

$$\sum_{\ell=1}^{\ell_0-1} \mathbf{P}(Z_{n-a_k}=\ell)\,\mathbf{P}\left(Z_{a_k}=k \,\middle|\, Z_0=\ell\right) \;\le\; c\,\frac{k^{\alpha-1}}{m^{\alpha n}}\,. \tag{50}$$

This completes the proof. $\square$

For easy citation purposes, we expose as a lemma the following two versions of the so-called Fuk–Nagaev inequality for tail probabilities of sums of i.i.d. variables, which is easily derived from Nagaev [11]. Recall that we assumed that X_1 is centered and has a positive finite variance σ^2.

Lemma 14 (Fuk–Nagaev inequality). *For* $k\ge 1$, $\varepsilon_n>0$, $n\ge 1$, $r>1$, *and* $t\ge 2$,

$$\mathbf{P}(S_k\ge\varepsilon_n k) \;\le\; k\,\mathbf{P}\left(X_1\ge r^{-1}\varepsilon_n k\right)+(\mathrm{e}\,r\sigma^2)^r\,\varepsilon_n^{-2r}\,k^{-r}, \tag{51}$$

and

$$\begin{aligned}\mathbf{P}(S_k\ge\varepsilon_n k) &\le k\,\mathbf{P}\left(X_1\ge r^{-1}\varepsilon_n k\right)+\exp\left[-\frac{2}{(t+2)^2\,\mathrm{e}^t\,\sigma^2}\,\varepsilon_n^2 k\right]\\ &\quad+\left(\frac{(t+2)\,r^{t-1}\,\mathbf{E}\left\{X_1^t;\;0\le X_1\le\varepsilon_n k\right\}}{t\,\varepsilon_n^t\,k^{t-1}}\right)^{tr/(t+2)}.\end{aligned} \tag{52}$$

Proof By (1.56) and (1.23) in [11], for all $u,v>0$,

$$\mathbf{P}(S_k\ge u) \;\le\; k\,\mathbf{P}(X_1\ge v)+\mathrm{e}^{u/v}\left(\frac{\sigma^2 k}{uv}\right)^{u/v} \tag{53}$$

and

$$\mathbf{P}(S_k \geq u) \leq k\,\mathbf{P}(X_1 \geq v) + \exp\Big[-\frac{2u^2}{(t+2)^2\,\mathrm{e}^t\,\sigma^2}\Big] + \left(\frac{(t+2)\,k\,\mathbf{E}\big\{X_1^t;\ 0 \leq X_1 \leq v\big\}}{t\,u\,v^{t-1}}\right)^{tu/(t+2)v}. \tag{54}$$

Putting here $u = \varepsilon_n k$ and $v = u/r$, we get (51) and (52), finishing the proof. □

Remark 15 (*On the case* $\varepsilon_n \equiv \varepsilon$). Here we prove a one-sided version of statement (7) concerning our general R_n, assuming the Schröder case and that $\mathbf{E}(X_1^+)^{1+\alpha} < \infty$. Take any $\varepsilon > 0$ and set $g_n(k) := m^{\alpha n}\,\mathbf{P}(Z_n = k)\,\mathbf{P}(S_k \geq \varepsilon k)$. From estimate (43) we get, for all $n, k \geq 1$, the inequality $g_n(k) \leq c\,\tilde{g}(k)$, where $\tilde{g}(k) := k^{\alpha-1}\,\mathbf{P}(S_k \geq \varepsilon k)$. Next we show that $\tilde{g}(k)$ is summable in k. Letting $\varepsilon_n = \varepsilon$ and $r = \alpha + 1$ in (51), we see that for all $k \geq 1$,

$$\tilde{g}(k) \;\leq\; k^{\alpha}\,\mathbf{P}\,(X_1 \geq \varepsilon k/(1+\alpha)) + c\,\varepsilon^{-2-2\alpha}k^{-2}. \tag{55}$$

But the summability of $k^{\alpha}\,\mathbf{P}(X_1 \geq ck)$ with some (hence all) positive c is equivalent to the finiteness of $\mathbf{E}(X_1^+)^{1+\alpha}$, and we get the claimed summability of $\tilde{g}(k)$.

On the other hand, it follows from (26) that for every fixed k,

$$\lim_{n\uparrow\infty} g_n(k) \;=\; \nu_k\,\mathbf{P}(S_k \geq \varepsilon k). \tag{56}$$

Therefore, by dominated convergence,

$$\lim_{n\uparrow\infty} \sum_{k=1}^{\infty} g_n(k) \;=\; \sum_{k=1}^{\infty} \nu_k\,\mathbf{P}(S_k \geq \varepsilon k). \tag{57}$$

Recalling the definition of $g_n(k)$ and using (12), we obtain

$$\lim_{n\uparrow\infty} m^{\alpha n}\,\mathbf{P}(R_n \geq \varepsilon) \;=\; \sum_{k=1}^{\infty} \nu_k\,\mathbf{P}(S_k \geq \varepsilon k), \tag{58}$$

yielding the wanted one-sided version.

2.2 Interplay between the two competing forces

In the next five lemmas we prove estimates for different parts of the sum at the right hand side of decomposition (12), which are the crucial steps in the proof of Theorem 5.

Lemma 16 (A tail estimate). *Assume* X_1^+ *has a tail of index* $\theta > 2$. *Then*

$$\sum_{k \geq m^n} \mathbf{P}(Z_n = k)\, \mathbf{P}(S_k \geq \varepsilon_n k)$$
$$\leq c \left(\varepsilon_n^{-\theta}\, m^{-(\theta-1)n} + (\varepsilon_n^2 m^n)^{-1} \exp\left[-c\, \varepsilon_n^2\, m^n\right]\right), \quad \varepsilon_n > 0, \; n \geq 1. \tag{59}$$

Proof Letting $t = \theta + 1$ and $r = (t+2)/t$ in (52), and using that X_1^+ has a tail of index $\theta > 2$, we get the bound

$$\mathbf{P}(S_k \geq \varepsilon_n k) \leq c \left(\varepsilon_n^{-\theta}\, k^{-(\theta-1)} + \frac{\mathbf{E}\left\{X_1^{\theta+1};\; X_1 \in [0, \varepsilon_n k]\right\}}{\varepsilon_n^{\theta+1}\, k^{\theta}} \right) + \exp[-c\, \varepsilon_n^2 k]. \tag{60}$$

Clearly, under (14),

$$\mathbf{E}\left\{X_1^{\theta+1};\; X_1 \in [0, x]\right\} \sim a\, \theta x \quad \text{as} \quad x \uparrow \infty. \tag{61}$$

Thus,

$$\mathbf{E}\left\{X_1^{\theta+1};\; X_1 \in [0, x]\right\} \leq c\, x, \quad x \geq 1. \tag{62}$$

On the other hand, if $x \leq 1$,

$$\mathbf{E}\left\{X_1^{\theta+1};\; X_1 \in [0, x]\right\} \leq x^{\theta+1}\, \mathbf{P}(X_1 \in [0, x]) \leq x. \tag{63}$$

Therefore,

$$\mathbf{E}\left\{X_1^{\theta+1};\; X_1 \in [0, x]\right\} \leq c\, x, \quad x \geq 0. \tag{64}$$

Applying this to the expectation in (60), we get

$$\mathbf{P}(S_k \geq \varepsilon_n k) \leq c\, \varepsilon_n^{-\theta}\, k^{-(\theta-1)} + \exp[-c\, \varepsilon_n^2\, k]. \tag{65}$$

Moreover, combining this bound with (42) gives

$$\sum_{k \geq m^n} \mathbf{P}(Z_n = k)\, \mathbf{P}(S_k \geq \varepsilon_n k) \leq c\, \varepsilon_n^{-\theta} \sum_{k \geq m^n} k^{-\theta} + \sum_{k \geq m^n} k^{-1} \exp[-c\, \varepsilon_n^2 k]. \tag{66}$$

Obviously,

$$\sum_{k \geq m^n} k^{-\theta} \leq c\, m^{-(\theta-1)n}. \tag{67}$$

On the other hand,

$$\sum_{k\geq m^n} k^{-1}\exp[-c\,\varepsilon_n^2 k] \leq m^{-n}\sum_{k\geq m^n}\exp[-c\,\varepsilon_n^2 k]$$
$$\leq c\,(\varepsilon_n^2 m^n)^{-1}\exp\left[-c\,\varepsilon_n^2 m^n\right]. \tag{68}$$

Substituting (67) and (68) into (66) finishes the proof. □

Lemma 17 (Another tail estimate). *Assume that* X_1^+ *has a tail of index* $\theta\in(2,1+\alpha)$. *If* $\varepsilon_n\geq m^{-\varrho n}$ *for some* $\varrho\in(0,1/2)$, *then*

$$\limsup_{n\uparrow\infty}\left|\varepsilon_n^\theta\, m^{(\theta-1)n}\sum_{k>\delta m^n}\mathbf{P}(Z_n=k)\,\mathbf{P}(S_k\geq\varepsilon_n k)-a\,I_\theta\right| \leq c\,\delta^{1+\alpha-\theta}. \tag{69}$$

Proof It is known (see for example Borovkov [6]), that if $\mathbf{P}(X_1\geq x)$ is regularly varying as $x\uparrow\infty$ with index $\theta>2$, then for every sequence $a_k\to\infty$,

$$\lim_{k\uparrow\infty}\sup_{x:\ x\geq a_k(k\log k)^{1/2}}\left|\frac{\mathbf{P}(S_k\geq x)}{k\,\mathbf{P}(X_1\geq x)}-1\right| = 0. \tag{70}$$

Note that if $\delta>0$, $k\geq\delta m^n$, and $\varepsilon_n\geq m^{-\varrho n}$, then $\varepsilon_n\geq\delta^\varrho k^{-\varrho}$. Hence,

$$\frac{\varepsilon_n k}{(k\log k)^{1/2}} \geq \delta^\varrho\,\frac{k^{1/2-\varrho}}{(\log k)^{1/2}}. \tag{71}$$

Since $0<\varrho<1/2$, the right hand side goes to infinity as $k\uparrow\infty$, and we will take it as a_k. Thus, applying (70) gives, as $n\uparrow\infty$,

$$\sum_{k>\delta m^n}\mathbf{P}(Z_n=k)\,\mathbf{P}(S_k\geq\varepsilon_n k)=\bigl(1+o(1)\bigr)\sum_{k>\delta m^n}k\,\mathbf{P}(Z_n=k)\,\mathbf{P}(X_1\geq\varepsilon_n k)$$
$$=\bigl(1+o(1)\bigr)\,a\,\varepsilon_n^{-\theta}\sum_{k>\delta m^n}k^{-(\theta-1)}\,\mathbf{P}(Z_n=k), \tag{72}$$

where in the second step we used that X_1^+ has a tail of index $\theta\in(2,1+\alpha)$. By (43) we have

$$\sum_{1\leq k\leq\delta m^n}k^{-(\theta-1)}\,\mathbf{P}(Z_n=k) \leq c\,m^{-\alpha n}\sum_{1\leq k\leq\delta m^n}k^{\alpha-\theta} \leq c\,m^{-(\theta-1)n}\,\delta^{1+\alpha-\theta}.$$

By Theorem 1 of [12], for $\theta-1<\alpha$, we have $\mathbf{E}\left\{Z_n^{-(\theta-1)};\ Z_n>0\right\}\sim I_\theta\, m^{-(\theta-1)n}$ as $n\uparrow\infty$, with I_θ defined in (21). Hence, for all sufficiently large n,

$$\left|\sum_{k>\delta m^n}k^{-(\theta-1)}\,\mathbf{P}(Z_n=k)-I_\theta\, m^{-(\theta-1)n}\right| \leq c\,m^{-(\theta-1)n}\,\delta^{1+\alpha-\theta}. \tag{73}$$

Combining (72) and (73), the proof is finished. □

Recall our general assumption (13).

Lemma 18 (A further tail estimate). *Suppose the Schröder case and let* X_1^+ *satisfy moment condition* (16). *Then*

$$\limsup_{n\uparrow\infty} \varepsilon_n^{2\alpha} m^{\alpha n} \sum_{k\geq A/\varepsilon_n^2} \mathbf{P}(Z_n = k)\,\mathbf{P}(S_k \geq \varepsilon_n k) \;\leq\; \frac{c}{A}, \quad A \geq 1. \tag{74}$$

Proof Combining (43) and (51) with $r = \alpha + 1$ gives

$$\begin{aligned} &m^{\alpha n} \sum_{k\geq A/\varepsilon_n^2} \mathbf{P}(Z_n = k)\,\mathbf{P}(S_k \geq \varepsilon_n k) \\ &\quad\leq\; c \left(\sum_{k\geq A/\varepsilon_n^2} k^{\alpha}\,\mathbf{P}\left(X_1 \geq (\alpha+1)^{-1}\varepsilon_n k\right) + \varepsilon_n^{-2(\alpha+1)} \sum_{k\geq A/\varepsilon_n^2} k^{-2} \right). \end{aligned} \tag{75}$$

Note that

$$\varepsilon_n^{-2(\alpha+1)} \Bigl(\sum_{k\geq A/\varepsilon_n^2} k^{-2} \Bigr) \;\leq\; \frac{c}{A}\,\varepsilon_n^{-2\alpha}, \quad n > 0, \quad \varepsilon_n > 0, \quad A \geq 1. \tag{76}$$

On the other hand, to bound the first sum at the right hand side in (75), note first that

$$\int_{k-1}^{k} u^{\alpha}\,\mathbf{P}(X_1 \geq (\alpha+1)^{-1}\varepsilon_n u)\,\mathrm{d}u \;\geq\; (k-1)^{\alpha}\,\mathbf{P}\Bigl(X_1 \geq (\alpha+1)^{-1}\varepsilon_n k\Bigr), \quad k \geq 1.$$

This inequality can be continued by using $k - 1 \geq k/2$ for $k \geq 2$. Summing up gives for $\varepsilon_n^2 \leq 1/2$,

$$\begin{aligned} \sum_{k\geq A\varepsilon_n^2} k^{\alpha}\,\mathbf{P}\left(X_1 \geq (\alpha+1)^{-1}\varepsilon_n k\right) &\leq c \int_{A/\varepsilon_n^2 - 1}^{\infty} u^{\alpha}\,\mathbf{P}\Bigl(X_1 \geq (\alpha+1)^{-1}\varepsilon_n u\Bigr)\,\mathrm{d}u \\ &\leq c\,\varepsilon_n^{-\alpha-1} \int_{(A-\varepsilon_n^2)/(\alpha+1)\varepsilon_n}^{\infty} v^{\alpha}\mathbf{P}(X_1 \geq v)\,\mathrm{d}v. \end{aligned} \tag{77}$$

Recall that we assumed the moment condition (16) and that $\varepsilon_n \to 0$. Then the integral in (77) converges to zero as $n \uparrow \infty$, uniformly in $A \geq 1$. In particular, under $\alpha \geq 1$, (77) is of order $o(\varepsilon_n^{-2\alpha})$, uniformly in $A \geq 1$. On the other hand, if $\alpha < 1$

and since $\mathbf{E}X_1^2 < \infty$,

$$\int\limits_{(A-\varepsilon_n^2)/(\alpha+1)\varepsilon_n}^{\infty} v^{\alpha}\,\mathbf{P}(X_1 \geq v)\,\mathrm{d}v \;\leq\; c\,\frac{\varepsilon_n^{1-\alpha}}{(A-\varepsilon_n^2)^{1-\alpha}} \int\limits_{(A-\varepsilon_n^2)/(\alpha+1)\varepsilon_n}^{\infty} v\,\mathbf{P}(X_1 \geq v)\,\mathrm{d}v$$
$$= o(\varepsilon_n^{1-\alpha}) \;=\; o(\varepsilon_n^{-2\alpha}) \tag{78}$$

as $n \uparrow \infty$, uniformly in $A \geq 1$. Thus, for each $\alpha < \infty$ we have

$$\sup_{A\geq 1} \sum_{k\,\geq A/\varepsilon_n^2} k^{\alpha}\,\mathbf{P}\left(X_1 \geq (\alpha+1)^{-1}\varepsilon_n k\right) = o(\varepsilon_n^{-2\alpha}) \quad \text{as } n \uparrow \infty. \tag{79}$$

In particular,

$$\limsup_{n\uparrow\infty}\, \varepsilon_n^{2\alpha} \sum_{k\,\geq A/\varepsilon_n^2} k^{\alpha}\,\mathbf{P}\left(X_1 \geq (\alpha+1)^{-1}\varepsilon_n k\right) \;\leq\; \frac{c}{A}, \quad A \geq 1. \tag{80}$$

Combining (75), (76), and (80) gives the claim in the lemma. □

Lemma 19 (Initial part). *In the Schröder case,*

$$\sum_{1\,\leq k\,\leq\,\delta/\varepsilon_n^2} \mathbf{P}(Z_n = k)\,\mathbf{P}(S_k \geq \varepsilon_n k) \;\leq\; c\,\delta^{\alpha}\,\varepsilon_n^{-2\alpha}\,m^{-\alpha n}, \tag{81}$$

$\delta > 0,\;\; \varepsilon_n > 0,\;\; n \geq 1.$

Proof It follows from (43) that

$$\sum_{1\,\leq k\,\leq\,\delta/\varepsilon_n^2} \mathbf{P}(Z_n = k)\,\mathbf{P}(S_k \geq \varepsilon_n k) \;\leq\; \sum_{1\,\leq k\,\leq\,\delta/\varepsilon_n^2} \mathbf{P}(Z_n = k)$$
$$\leq\; \frac{c}{m^{\alpha n}} \sum_{1\,\leq k\,\leq\,\delta/\varepsilon_n^2} k^{\alpha-1} \;\leq\; c\,\delta^{\alpha}\,\varepsilon_n^{-2\alpha}\,m^{-\alpha n}, \tag{82}$$

finishing the proof. □

Lemma 20 (A central part and another initial part estimate). *Suppose* $1 < \alpha < \infty$ *and that* X_1^+ *has a tail of index* $\theta \in (2, 1+\alpha)$. *Then*

$$\sum_{A/\varepsilon_n^2\,\leq k\,\leq\,\delta m^n} \mathbf{P}(Z_n = k)\,\mathbf{P}(S_k \geq \varepsilon_n k)$$
$$\leq\; c\,\left(\delta^{1+\alpha-\theta}\,\varepsilon_n^{-\theta} m^{-(\theta-1)n} + A^{-1}\,\varepsilon_n^{-2\alpha}\,m^{-\alpha n}\right), \tag{83}$$

$A \geq 1, \ \ \delta > 0, \ \ \varepsilon_n > 0, \ \ n \geq 1$, *and*

$$\sum_{1 \leq k \leq \delta m^n} \mathbf{P}(Z_n = k)\, \mathbf{P}(S_k \geq \varepsilon_n k) \leq c \left(\delta^{1+\alpha-\theta} \varepsilon_n^{-\theta} m^{-(\theta-1)n} + \varepsilon_n^{-2\alpha} m^{-\alpha n}\right), \quad \delta > 0, \quad \varepsilon_n > 0, \quad n \geq 1. \tag{84}$$

Proof Combining (43) and (51) with $r = \alpha + 1$ gives

$$\sum_{A/\varepsilon_n^2 \leq k \leq \delta m^n} \mathbf{P}(Z_n = k)\, \mathbf{P}(S_k \geq \varepsilon_n k) \leq c\, m^{-\alpha n} \left(\sum_{A/\varepsilon_n^2 \leq k \leq \delta m^n} k^\alpha\, \mathbf{P}\left(X_1 \geq (\alpha+1)^{-1} \varepsilon_n k\right) + \varepsilon_n^{-2(\alpha+1)} \sum_{A/\varepsilon_n^2 \leq k \leq \delta m^n} k^{-2}\right). \tag{85}$$

From (76),

$$\varepsilon_n^{-2(\alpha+1)} \sum_{A/\varepsilon_n^2 \leq k \leq \delta m^n} k^{-2} \leq \frac{c}{A}\, \varepsilon_n^{-2\alpha}. \tag{86}$$

On the other hand, since X_1^+ has a tail of index $\theta \in (2, 1 + \alpha)$,

$$\sum_{A/\varepsilon_n^2 \leq k \leq \delta m^n} k^\alpha\, \mathbf{P}\left(X_1 \geq (\alpha + 1)^{-1} \varepsilon_n k\right) \leq c\, \varepsilon_n^{-\theta} \sum_{1 \leq k \leq \delta m^n} k^{\alpha-\theta} \leq c\, \varepsilon_n^{-\theta} \delta^{1+\alpha-\theta} m^{(1+\alpha-\theta)n}. \tag{87}$$

Combine (85)–(87) to get (83).

Putting $A = 1$ in (83) and $\delta = 1$ in (81), we obtain (84), finishing the proof. □

Recall that (μ, d) refers to the type of the offspring law, $\alpha \in (0, \infty)$ to the Schröder constant, and that X_1 is assumed to have a finite variance σ^2. For $0 < \delta < 1 < A < \infty$, consider

$$\Sigma_n(\delta, A) := \sum_{\delta/\varepsilon_n^2 \leq k \leq A/\varepsilon_n^2} \mathbf{P}(Z_n = k)\, \mathbf{P}(S_k \geq \varepsilon_n k). \tag{88}$$

Lemma 21 (Another central part estimate). *Suppose to be in the Schröder case. Then for all* $0<\delta<1<A<\infty$,

$$V_*\int_\delta^A u^{\alpha-1}\,\overline{\Phi}(\sqrt{u}/\sigma)\,\mathrm{d}u \;\le \liminf_{n\uparrow\infty}\,\varepsilon_n^{2\alpha}\,m^{\alpha n}\,\Sigma_n(\delta,A)\le \limsup_{n\uparrow\infty}\,\varepsilon_n^{2\alpha}\,m^{\alpha n}\,\Sigma_n(\delta,A)$$
$$\le V^*\int_\delta^A u^{\alpha-1}\,\overline{\Phi}(\sqrt{u}/\sigma)\,\mathrm{d}u \tag{89}$$

with V_* *and* V^* *defined in* (18)*, and where* $\overline{\Phi}(x):=1-\Phi(x)$.

Proof In view of (9) in Proposition 3 with $k_n=\delta/\varepsilon_n^2$,

$$\Sigma_n(\delta,A)=\big(1+o(1)\big)\,d\sum_{k\in H(\delta,A)} m^{-n}\,w\left(\tfrac{k}{m^n}\right)\mathbf{P}(S_k\ge\varepsilon_n k)\quad\text{as } n\uparrow\infty \tag{90}$$

with $H(\delta,A):=\big\{k\in[\delta/\varepsilon_n^2,\ A/\varepsilon_n^2]:\ k\equiv\mu\ (\mathrm{mod}\,d)\big\}$. Clearly,

$$V_*(n)\sum_{k\in H(\delta,A)}\frac{k^{\alpha-1}}{m^{\alpha n}}\,\mathbf{P}(S_k\ge\varepsilon_n k)\le\sum_{k\in H(\delta,A)} m^{-n}\,w\left(\tfrac{k}{m^n}\right)\mathbf{P}(S_k\ge\varepsilon_n k)$$
$$\le V^*(n)\sum_{k\in H(\delta,A)}\frac{k^{\alpha-1}}{m^{\alpha n}}\,\mathbf{P}(S_k\ge\varepsilon_n k), \tag{91}$$

where we set

$$V_*(n)\ :=\ \inf_{u\le A/\varepsilon_n^2 m^n} u^{1-\alpha}w(u),\quad V^*(n)\ :=\ \sup_{u\le A/\varepsilon_n^2 m^n} u^{1-\alpha}w(u). \tag{92}$$

By the central limit theorem,

$$\sup_{k\in H(\delta,A)}\left|\mathbf{P}(S_k\ge\varepsilon_n k)-\overline{\Phi}\left(\sqrt{\varepsilon_n^2 k}/\sigma\right)\right|\to 0\quad\text{as } n\uparrow\infty. \tag{93}$$

Hence, as $n\uparrow\infty$,

$$\sum_{k\in H(\delta,A)} k^{\alpha-1}\,\mathbf{P}(S_k\ge\varepsilon_n k)=\big(1+o(1)\big)\sum_{k\in H(\delta,A)} k^{\alpha-1}\,\overline{\Phi}\left(\sqrt{\varepsilon_n^2 k}/\sigma\right)$$
$$=\varepsilon_n^{-2\alpha}\,\big(1+o(1)\big)\sum_{k\in H(\delta,A)}(\varepsilon_n^2 k)^{\alpha-1}\,\overline{\Phi}\left(\sqrt{\varepsilon_n^2 k}/\sigma\right)\varepsilon_n^2$$
$$=d^{-1}\varepsilon_n^{-2\alpha}\,\big(1+o(1)\big)\int_\delta^A u^{\alpha-1}\,\overline{\Phi}(\sqrt{u}/\sigma)\,\mathrm{d}u. \tag{94}$$

Substituting (94) into (91) and noting that we have $V_*(n) \to V_*$ and $V^*(n) \to V^*$ as $n \uparrow \infty$ by our velocity assumption (13) on ε_n, we obtain (89). □

Finally, we compute the limit, as $\delta \downarrow 0$ and $A \uparrow \infty$, of the integral from (89).

Lemma 22 (A moment formula for the Gaussian law). *For* $0 < \alpha < \infty$,

$$\int_0^\infty u^{\alpha-1}\, \overline{\Phi}(\sqrt{u}/\sigma)\, \mathrm{d}u = \frac{2^{\alpha-1}\,\Gamma(\alpha+1/2)}{\alpha\sqrt{\pi}}\,\sigma^{2\alpha} \;=\; \Gamma_\alpha\,. \tag{95}$$

Proof Substituting $v = \sqrt{u}/\sigma$, we have

$$\begin{aligned}
\int_0^\infty u^{\alpha-1}\, \overline{\Phi}(\sqrt{u}/\sigma)\, \mathrm{d}u &= 2\sigma^{2\alpha} \int_0^\infty v^{2\alpha-1}\, \overline{\Phi}(v)\, \mathrm{d}v \\
&= 2\sigma^{2\alpha} \int_0^\infty \mathrm{d}v\; v^{2\alpha-1} \int_v^\infty \mathrm{d}t\, \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-t^2/2} \\
&= 2\frac{\sigma^{2\alpha}}{\sqrt{2\pi}} \int_0^\infty \mathrm{d}t\; \mathrm{e}^{-t^2/2} \int_0^t \mathrm{d}v\; v^{2\alpha-1} \\
&= \frac{\sigma^{2\alpha}}{\alpha\sqrt{2\pi}} \int_0^\infty t^{2\alpha}\, \mathrm{e}^{-t^2/2}\, \mathrm{d}t.
\end{aligned} \tag{96}$$

Substituting now $v = t^2/2$, the chain of equalities can be continued with

$$= \frac{2^{\alpha-1}\sigma^{2\alpha}}{\alpha\sqrt{\pi}} \int_0^\infty v^{\alpha-1/2}\, \mathrm{e}^{-v}\, \mathrm{d}v \;=\; \frac{2^{\alpha-1}\,\Gamma(\alpha+1/2)}{\alpha\sqrt{\pi}}\,\sigma^{2\alpha}, \tag{97}$$

which equals Γ_α from (19). The proof is finished. □

3 Proof of the theorems

3.1 Schröder case: proof of Theorem 5

After all of the preparations in the previous section, the proof of Theorem 5 can easily be completed.

(a) We start by showing that

$$\limsup_{n\uparrow\infty}\, \varepsilon_n^{2\alpha}\, m^{\alpha n} \sum_{k\,\geq A/\varepsilon_n^2} \mathbf{P}(Z_n = k)\, \mathbf{P}(S_k \geq \varepsilon_n k) \;\leq\; \frac{c}{A}, \quad A \geq 1. \tag{98}$$

In the case $\mathbf{E}(X_1^+)^{1+\alpha} < \infty$, this bound is already obtained in Lemma 18. Thus, we have to show (98) in the case if X_1^+ has a tail of index θ and $\varepsilon_n = o(m^{-\varkappa n})$. Combining (83) with $\delta = 1$ and Lemma 16, we get

$$\sum_{k \geq A/\varepsilon_n^2} \mathbf{P}(Z_n = k)\,\mathbf{P}(S_k \geq \varepsilon_n k) \\ \leq c\Big(A^{-1}\varepsilon_n^{-2\alpha}\, m^{-\alpha n} + \varepsilon_n^{-\theta} m^{-(\theta-1)n} + (\varepsilon_n^2 m^n)^{-1} \exp[-c\varepsilon_n^2 m^n]\Big). \tag{99}$$

Noting that

$$\varepsilon_n^{-\theta} m^{-(\theta-1)n} + (\varepsilon_n^2 m^n)^{-1} \exp[-c\varepsilon_n^2 m^n] = o(\varepsilon_n^{-2\alpha}\, m^{-\alpha n}) \tag{100}$$

under our assumptions $\varepsilon_n^2 m^n \to \infty$ and $\varepsilon_n = o(m^{-\varkappa n})$, the proof of (98) is finished.
Combining Lemmas 19, 21, and (98), and using that δ and A are arbitrary, we see that

$$\begin{aligned} V_* \int_0^\infty u^{\alpha-1}\, \overline{\Phi}(\sqrt{u}/\sigma)\, \mathrm{d}u &\leq \liminf_{n\uparrow\infty} \varepsilon_n^{2\alpha}\, m^{\alpha n} \sum_{k=1}^\infty \mathbf{P}(Z_n = k)\,\mathbf{P}(S_k \geq \varepsilon_n k) \\ &\leq \limsup_{n\uparrow\infty} \varepsilon_n^{2\alpha}\, m^{\alpha n} \sum_{k=1}^\infty \mathbf{P}(Z_n = k)\,\mathbf{P}(S_k \geq \varepsilon_n k) \\ &\leq V^* \int_0^\infty u^{\alpha-1}\, \overline{\Phi}(\sqrt{u}/\sigma)\, \mathrm{d}u. \end{aligned} \tag{101}$$

With Lemma 22 the proof of part (a) is completed.

(b) If $\varepsilon_n m^{\varkappa n} \to \infty$, then, obviously, $\varepsilon_n^{-2\alpha}\, m^{-\alpha n} = o(\varepsilon_n^{-\theta}\, m^{-(\theta-1)n})$. Therefore, by estimate (84),

$$\limsup_{n\uparrow\infty} \varepsilon_n^{\theta}\, m^{(\theta-1)n} \sum_{1 \leq k \leq \delta m^n} \mathbf{P}(Z_n = k)\,\mathbf{P}(S_k \geq \varepsilon_n k) \leq c\,\delta^{1+\alpha-\theta}. \tag{102}$$

Part (b) follows from Lemma 17 and (102) by letting $\delta \downarrow 0$.

(c) Finally, under $\varepsilon_n \sim \tau^{-1} m^{-\varkappa n}$, part (c) follows from (81), (89), (95), (83), and Lemma 17.

The proof is finished altogether. □

3.2 Böttcher under light tails concerning X_1 : proof of Theorem 9

It follows from the assumed finiteness of an exponential moment of X_1, see e.g. Lemma III.5 in Petrov [16], that for every $\delta \in (0, 1)$ there exists $h_\delta > 0$ such that

$$\mathbf{E}\mathrm{e}^{hX_1} \ \le\ \mathrm{e}^{\sigma^2(1+\delta)h^2/2}, \quad |h| \le h_\delta\,. \tag{103}$$

Thus, we may use the well-known Bernstein inequality, see Theorem III.15 in [16]. This gives, for all $k \ge 1$ and $\varepsilon_n \le h_\delta$,

$$\mathbf{P}(S_k \ge \varepsilon_n k) \ \le\ \exp\Big[-(1-\delta)\,\frac{\varepsilon_n^2 k}{2\sigma^2}\Big]. \tag{104}$$

Therefore,

$$\mathbf{P}(R_n \ge \varepsilon_n) \ \le\ f_n\Big(\exp\Big[-(1-\delta)\,\frac{\varepsilon_n^2}{2\sigma^2}\Big]\Big) \quad \text{if } \varepsilon_n \le h_\delta\,. \tag{105}$$

We may also assume that $\varepsilon_n \le 1/m$. Set $r_n := \max\{k \ge 1 : m^k \le \varepsilon_n^{-2}\}$. Then,

$$m^{-r_n-1} \ <\ \varepsilon_n^2 \ \le\ m^{-r_n}. \tag{106}$$

The left hand inequality together with the monotonicity of f_n gives

$$f_n\Big(\exp\Big[-(1-\delta)\,\frac{\varepsilon_n^2}{2\sigma^2}\Big]\Big) \ \le\ f_n\Big(\exp\Big[-(1-\delta)\,\frac{m^{-r_n-1}}{2\sigma^2}\Big]\Big). \tag{107}$$

Bounds (105), (107), and the right hand inequality in (106) imply

$$\varepsilon_n^{-2\beta} m^{-n\beta} \log \mathbf{P}(R_n \ge \varepsilon_n) \ \le\ \mu^{-n+r_n} \log f_n\Big(\exp\Big[-(1-\delta)\,\frac{m^{-r_n-1}}{2\sigma^2}\Big]\Big), \tag{108}$$

where we used $\mu = m^\beta$. Since $r_n \to \infty$, by the Kesten–Stigum theorem for supercritical Galton–Watson processes,

$$\lim_{n\uparrow\infty} f_{r_n+1}\Big(\exp\Big[-(1-\delta)\,\frac{m^{-r_n-1}}{2\sigma^2}\Big]\Big) \ =\ \varphi\left((1-\delta)/2\sigma^2\right). \tag{109}$$

On the other hand, from the assumption $\varepsilon_n^2 m^n \to \infty$ and the right hand inequality in (106) it follows that $n - r_n \to \infty$. Therefore, by (31) we have for $s \in [0, 1]$,

$$\lim_{n\uparrow\infty} \mu^{-n+r_n+1} \log f_{n-r_n-1}(s) \ =\ \log \mathsf{B}(s). \tag{110}$$

By the continuity of B, combining (109) and (110) we obtain

$$\lim_{n\uparrow\infty} \mu^{-n+r_n+1} \log f_n \left(\exp\left[-(1-\delta)\,\frac{m^{-r_n-1}}{2\sigma^2}\right]\right) = \log \mathsf{B}\left(\varphi\left((1-\delta)/2\sigma^2\right)\right). \tag{111}$$

Now (33b) follows from (108) and (111) letting $\delta \downarrow 0$.

In order to prove (33a) we will exploit the following version of Kolmogorov's inequality: for $0<\delta<1$ fixed, there exists a constant $D \in (0,\infty)$ such that

$$\mathbf{P}(S_k \geq \varepsilon_n k) \;\geq\; \exp\left[-(1+\delta)\,\frac{\varepsilon_n^2 k}{2\sigma^2}\right], \quad k > D/\varepsilon_n^2, \quad n \geq 1. \tag{112}$$

See Statulevicius [19]. Using (112) we obtain

$$\begin{aligned}\mathbf{P}(R_n \geq \varepsilon_n) &\geq \sum_{k>D/\varepsilon_n^2} \mathbf{P}(Z_n = k)\, \exp\left[-(1+\delta)\,\frac{\varepsilon_n^2 k}{2\sigma^2}\right] \\ &\geq f_n\left(\exp\left[-(1+\delta)\,\frac{\varepsilon_n^2}{2\sigma^2}\right]\right) - \mathbf{P}(Z_n \leq D/\varepsilon_n^2).\end{aligned} \tag{113}$$

Clearly, if $D/\varepsilon_n^2 < \mu^n$, then $\mathbf{P}(Z_n \leq D/\varepsilon_n^2) = 0$, and we pass directly to statement (117) below. Otherwise, it follows from Proposition 4 that

$$\mathbf{P}(Z_n \leq D/\varepsilon_n^2) \;\leq\; \exp\left[-c\, D^{-\beta/(1-\beta)}\, (\varepsilon_n^2 m^n)^{\beta/(1-\beta)}\right]. \tag{114}$$

From (113), (114), and the left hand inequality in (106), we have

$$\mathbf{P}(R_n \geq \varepsilon_n) \;\geq\; f_n\left(\exp\left[-(1+\delta)\,\frac{m^{-r_n}}{2\sigma^2}\right]\right) - \exp\left[-c\,(\varepsilon_n^2 m^n)^{\beta/(1-\beta)}\right]. \tag{115}$$

Analogously to (111),

$$\lim_{n\uparrow\infty} \mu^{-n+r_n} \log f_n \left(\exp\left[-(1+\delta)\,\frac{m^{-r_n}}{2\sigma^2}\right]\right) \;=\; \log \mathsf{B}\left(\varphi\left((1+\delta)/2\sigma^2\right)\right). \tag{116}$$

By the left hand inequality of (106), $\mu^{n-r_n} \leq m^\beta (\varepsilon_n^2 m^n)^\beta$. Therefore, from the limit statement (116) we see that the second term at the right hand side of estimate (115) is negligible compared with the first term there, i.e.

$$\mathbf{P}(R_n \geq \varepsilon_n) \;\geq\; f_n\left(\exp\left[-(1+\delta)\,\frac{m^{-r_n}}{2\sigma^2}\right]\right)\big(1+o(1)\big). \tag{117}$$

Thus, using the left hand inequality in (106), we get the bound

$$\varepsilon_n^{-2\beta} m^{-n\beta} \log \mathbf{P}(R_n \geq \varepsilon_n) \geq \mu^{-n+r_n+1} \log f_n \left(\exp\left[-(1+\delta)\, \frac{m^{-r_n}}{2\sigma^2} \right] \right) + o(1). \tag{118}$$

Since δ is arbitrary, combining (118) and (116) completes the proof of (33a).

In the derivation of (117) from (113) we learned that the second term at the right hand side of (113) is small compared with the first term there. Thus, from (113) together with (105) we get

$$\begin{aligned} f_n\left(\exp\left[-(1+\delta)\, \frac{\varepsilon_n^2}{2\sigma^2} \right] \right) \big(1+o(1)\big) &\leq \mathbf{P}(R_n \geq \varepsilon_n) \\ &\leq f_n\left(\exp\left[-(1-\delta)\, \frac{\varepsilon_n^2}{2\sigma^2} \right] \right). \end{aligned} \tag{119}$$

Hence, if $\varepsilon_n^2 = m^{-\lambda_n}$ then (34) follows from these inequalities and (116) replacing there r_n by λ_n, and finally letting $\delta \downarrow 0$. Altogether, the proof of Theorem 9 is complete. □

3.3 Böttcher under heavier tails concerning X_1^+: proof of Theorem 10

With B_2 from Proposition 4, and $\theta > 2$ the tail index of X_1^+, define $k_n := m^n / \log^{(1-\beta)/\beta} m^{2n\theta/B_2}$. Then by Proposition 4, for all sufficiently large n,

$$\mathbf{P}(Z_n \leq k_n) \;\leq\; \exp\left[-(B_2/2)(k_n/m^n)^{-\beta/(1-\beta)} \right] \;=\; m^{-\theta n}. \tag{120}$$

Hence, for these n,

$$\sum_{k \leq k_n} \mathbf{P}(Z_n = k)\, \mathbf{P}(S_k \geq \varepsilon_n k) \;\leq\; \mathbf{P}(Z_n \leq k_n) \;\leq\; m^{-\theta n} \tag{121}$$

and

$$\sum_{k \leq k_n} k^{-(\theta-1)}\, \mathbf{P}(Z_n = k) \;\leq\; \mathbf{P}(Z_n \leq k_n) \;\leq\; m^{-\theta n}. \tag{122}$$

It is easy to verify that

$$\frac{\varepsilon_n k_n}{(k_n \log k_n)^{1/2}} \;=\; \big(c + o(1)\big)\, \varepsilon_n m^{n/2} n^{-1/2\beta} \quad \text{as } n \uparrow \infty. \tag{123}$$

By our assumption in the theorem, the right hand side converges to infinity. Then, we can use (70) with $a_k := \varepsilon_n(k/\log k)^{1/2}$ to obtain

$$\sum_{k>k_n} \mathbf{P}(Z_n = k)\,\mathbf{P}(S_k \geq \varepsilon_n k) = \big(1+o(1)\big) \sum_{k>k_n} k\,\mathbf{P}(Z_n = k)\,\mathbf{P}(X_1 \geq \varepsilon_n k)$$
$$= \big(1+o(1)\big)\, a\,\varepsilon_n^{-\theta} \sum_{k>k_n} k^{-(\theta-1)} \mathbf{P}(Z_n = k) \quad \text{as } n \uparrow \infty. \tag{124}$$

Theorem 1 of [12] and (122) yield

$$\sum_{k>k_n} k^{-(\theta-1)}\,\mathbf{P}(Z_n = k) \;=\; I_\theta\, m^{-(\theta-1)n} \big(1+o(1)\big) \quad \text{as } n \uparrow \infty. \tag{125}$$

Substituting this into (124) and combining with (121) completes the proof. □

Acknowledgments We thank Didier Piau for a suggestion to improve the representation of the proof of Lemma 13 in our original preprint. We are also grateful to the anonymous referees for careful reading and suggestions for an improvement of the representation.

References

1. Asmussen, A., Hering, H.: Branching Processes, Progress in Probability and Statistics, vol. 3. Birkhäuser Boston, Boston (1983)
2. Athreya, K.B., Ney, P.E.: Branching Processes. Springer, Berlin (1972)
3. Athreya, K.B.: Large deviation rates for branching processes I. Single type case. Ann. Appl. Probab. **4**, 779–790 (1994)
4. Athreya, K.B., Vidyashankar, A.N. : Large deviation rates for supercritical and critical branching processes. In: Classical and Modern Branching Processes, IMA Volumes in Mathematics and its Applications, vol. 84, pp. 1–18. Springer, Berlin (1997)
5. Biggins, J.D., Bingham, N.H.: Large deviations in the supercritical branching process. Adv. Appl. Probab. **25**(4), 757–772 (1993)
6. Borovkov, A.A.: Estimates for sums and maxima of sums of random variables when the Cramér condition is not satisfied. Sibirian Math. J. **41**, 811–848 (2000)
7. Fleischmann, K., Wachtel, V.: Lower deviation probabilities for supercritical Galton–Watson processes. Ann. Inst. Henri Poincaré Probab. Statist. **43**(2), 233–255 (2007)
8. Jacob, C., Peccoud, J.: Inference on the initial size of a supercritical branching process from migrating binomial observations. C. R. Acad. Sci. Paris Ser. I **322**, 875–888 (1996)
9. Jacob, C., Peccoud, J.: Estimation of the parameters of a branching process from migrating binomial observations. Adv. Appl. Probab. **30**, 948–967 (1998)
10. Nagaev, A.V.: On estimating the expected number of direct descendants of a particle in a branching process. Theor. Probab. Appl. **12**, 314–320 (1967)
11. Nagaev, S.V.: Large deviations of sums of independent random variables. Ann. Probab. **7**, 745–789 (1979)
12. Ney, P.E., Vidyashankar, A.N.: Harmonic moments and large deviation rates for supercritical branching processes. Ann. Appl. Probab. **13**, 475–489 (2003)
13. Ney, P.E., Vidyashankar, A.N.: Local limit theory and large deviations for supercritical branching processes. Ann. Appl. Probab. **14**, 1135–1166 (2004)
14. Nagaev, S.V., Vakhtel, V.I.: On the local limit theorem for critical Galton–Watson processes. Theor. Probab. Appl. **50**, 400–419 (2006)
15. Pakes, A.G.: Non-parametric estimation in the Galton–Watson processes. Math. Biosci. **26**, 1–18 (1975)

16. Petrov, V.V.: Sums of Independent Random Variables. Springer, Berlin (1975)
17. Piau, D.: Immortal branching Markov processes: averaging properties and PCR applications. Ann. Probab. **32**, 337–364 (2004)
18. Rouault, A.: Large deviations and branching processes. Pliska Stud. Math. Bulgar. **13**, 15–38 (2000)
19. Statulevicius, V.A.: On large deviations. Z. Wahrscheinlichkeitstheorie Verw. Geb. **6**, 133–144 (1966)