*Submitted to the Bernoulli*

# Small Value Probabilities for Supercritical Branching Processes with Immigration

WEIJUAN CHU[1,2] , WENBO V. LI[3] and YAN-XIA REN[1,4]

[1] *LMAM School of Mathematical Sciences, Peking University, Beijing 100871, P. R. China.*

[2] *Department of Mathematics, Nanjing University, Nanjing 210093, P. R. China.*
*E-mail:* chuwj@nju.edu.cn

[3] *Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, USA.*
*E-mail:* wli@math.udel.edu

[4] *Corresponding author; Center for Statistical Science, Peking University, Beijing 100871, P. R. China. E-mail:* yxren@math.pku.edu.cn

We consider a supercritical Galton-Watson branching process with immigration. It is well known that under suitable conditions on the offspring and immigration distributions, there is a finite, strictly positive limit $\mathcal{W}$ for the normalized population size. Small value probabilities for $\mathcal{W}$ are obtained. Precise effects of the balance between offspring and immigration distributions are characterized.

*Keywords:* Supercritical Galton-Watson branching process, small value probability, immigration.

## 1. Introduction and main results

Small value probability for a *positive* random variable $V$ studies the rate of decay of the so called left tail probability $\mathbb{P}(V \le \varepsilon)$ as $\varepsilon \to 0^+$. When $V$ is the norm of a random element in a Banach space, one is dealing with small ball probability, see [LS01] for a survey of Gaussian measures. When $V$ is the maximum of a continuous random process starting at zero, one is estimating lower tail probability which is closely related to studies of boundary crossing probabilities or the first exit time associated with a general domain, see [L03] and [LS04] for Gaussian processes. A comprehensive study of small value probability is emerging and available in various talks and lecture notes in [L12+], see also the literature compilation [Lif11].

In this paper, we further study the most natural aspect of the branching tree approach originated in [MO08] on the martingale limit of a supercritical Galton-Watson process. The problem has been solved initially in [D71a], [D71b], see Theorem 1. The main goal is developing additional tools to treat small value probabilities for the martingale limit of a supercritical Galton-Watson process with immigration. The interplay between the offspring and the immigration distribution can be seen clearly from our main result Theorem 2. We next provide a more detailed and precise discussion by introducing additional notations, surveying relevant results and stating our results.

Let $(Z_n, n \geq 0)$ be a supercritical Galton-Watson branching process with $Z_0$=1, offspring distribution $p_k = \mathbb{P}(X = k), k \geq 0$, and mean $m = \mathbb{E}\,X \in (1, \infty)$. To avoid non-branching case, we suppose $p_k < 1$ for all $k$ throughout this paper. Under the natural condition $\mathbb{E}\,[X \log^+ X] < \infty$, the positive martingale $Z_n m^{-n}$ converges to a nontrivial random variable $W < \infty$ in the sense (see Kesten and Stigum [KS66])

$$Z_n m^{-n} \longrightarrow W \qquad \text{a.s. \& } L^1 \text{ as } n \to \infty.$$

Here and throughout this paper, $\log^+ x = \log \max(x, 1) \geq 0$. The distribution of the limit $W$ is of great interests in various applications. However, except for some very special cases, the explicit distribution of $W$ is not available, see, for example, Harris [H48], Hambly [H95] and Williams [W08] Section 0.9. In general, it is known that $W$ has a continuous positive density on $(0, \infty)$ satisfying a Lipschitz condition, see Athreya and Ney [AN72], Ch. II, p.84 Lemma 2. However it is not clear what type of densities can arise in this way. This lack of complete information on the distribution of $W$ prompts a search for asymptotic information such as the behavior of the left tail, or the small value probabilities of $W$ and its density.

In [D71b], the following results were given with assumption $p_0 = 0$ which holds without loss of generality after the standard Harris-Sevastyanov transformation, see [H48], p.478 Theorem 3.2 or [B88] p.216. Here and throughout this paper we use $g_1(x) \asymp g_2(x)$ as $x \to 0^+$ $(\infty)$ to represent $c \leq g_1(x)/g_2(x) \leq C$ as $x \to 0^+$ $(\infty)$ for two constants $C > c > 0$ and $g_1(x) \sim g_2(x)$ as $x \to 0^+$ $(\infty)$ to represent $g_1(x)/g_2(x) \to 1$ as $x \to 0^+$ $(\infty)$.

**Theorem 1** *(Dubuc 1971(b))*

*(a) If $p_1 > 0$, then*

$$\mathbb{P}(W \leq \varepsilon) \asymp \varepsilon^{|\log p_1|/\log m} \qquad \text{as } \varepsilon \to 0^+.$$

*(b) If $p_1 = 0$, then*

$$-\log \mathbb{P}(W \leq \varepsilon) \asymp \varepsilon^{-\beta/(1-\beta)} \qquad \text{as } \varepsilon \to 0^+,$$

*with $\beta := \log \gamma / \log m$ and $\gamma := \inf\{n : p_n > 0\} \geq 2$.*

Note that the rough asymptotic $\asymp$ in Theorem 1 can not be improved into more precise asymptotic $\sim$ and the oscillation is very small. This is the so called near-constancy phenomenon that were described and studied theoretically or numerically in [D82], [B88], [BP88] and [BB91]. In fact, it is still an open conjecture that the Laplace transform of $W$ being non-oscillating near $\infty$ (and hence the small value probability of $W$ being non-oscillating near $0^+$) is only specific to the case $p_1 > 0$ in [KM68a] p.127. General estimates, near-constancy phenomena, specific examples, and various implications have been studied to various degree of accuracy in Harris [H48], Karlin and McGregor [KM68a] [KM68b], Dubuc [D71a] , [D71b] and [D82], Barlow and Perkins [BP88], Goldstein [G87], Kusuoka [K87], Bingham [B88], Biggins and Bingham [BB91] and [BB93], Biggins and Nadarajah [BN93], Fleischman and Wachtel [FW07] and [FW09]. Recently, Berestycki, Gantert, Mörters and Sidorova [BGMS12] studied limit behaviors of the Galton-Watson tree conditioned on $W < \varepsilon$ as $\varepsilon \downarrow 0$.

In the present paper, we consider the supercritical branching process with immigration denoted by $(\mathcal{Z}_n, n \geq 0)$, and follow the definition in [AN72], Ch. VI, Section 7.1, p.263. To be more precise, we have

$$\mathcal{Z}_0 = Y_0, \quad \mathcal{Z}_{n+1} = X_1^n + X_2^n + \cdots + X_{\mathcal{Z}_n}^n + Y_{n+1}, \quad n \geq 0,$$

where $X_1^n, X_2^n, \cdots$ are i.i.d. with the same offspring distribution, $Y_0, Y_1, \cdots$ are i.i.d. with the same immigration distribution $\{q_k, k \geq 0\}$, and $X's$ and $Y's$ are independent. Recall that the offspring number $X$ has distribution $p_k = \mathbb{P}(X = k), k \geq 0$ and mean $m = \mathbb{E}\, X$. Suppose $Y$ has distribution $\{q_k, k \geq 0\}$. We use $f(s)$ and $h(s)$ to denote the generating function of $X$ and $Y$ respectively, i.e.

$$f(s) = \mathbb{E}\, s^X = \sum_{k=0}^{\infty} p_k s^k \quad \text{and} \quad h(s) = \mathbb{E}\, s^Y = \sum_{k=0}^{\infty} q_k s^k, \quad 0 < s < 1. \tag{1.1}$$

It is a classical result, see Seneta [S70] for example, that

$$\lim_{n \to \infty} \mathcal{Z}_n / m^n = \mathcal{W} \tag{1.2}$$

exists and is finite a.s. if and only if

$$\mathbb{E}\, \log^+ Y < \infty \quad \text{and} \quad \mathbb{E}\,(X \log^+ X) < \infty. \tag{1.3}$$

Our main result of this paper is the following small value probabilities for $\mathcal{W}$, which can be expressed as weighted summation of an infinite independent sequence of $W$'s, see (2.2).

**Theorem 2** *Assume the condition (1.3) holds.*
*(a) If $p_0 = 0$ and $0 < q_0 < 1$, then*

$$\mathbb{P}(\mathcal{W} \leq \varepsilon) \asymp \varepsilon^{|\log q_0|/\log m} \qquad as\ \varepsilon \to 0^+. \tag{1.4}$$

*(b) If $p_0 = 0$, $q_0 = 0$ and $p_1 > 0$, then*

$$\log \mathbb{P}(\mathcal{W} \leq \varepsilon) \sim -\frac{K|\log p_1|}{2(\log m)^2} \cdot |\log \varepsilon|^2 \qquad as\ \varepsilon \to 0^+, \tag{1.5}$$

*with $K = \inf\{n : q_n > 0\}$.*
*(c) If $p_0 = 0$, $q_0 = 0$ and $p_1 = 0$, then*

$$\log \mathbb{P}(\mathcal{W} \leq \varepsilon) \asymp -\varepsilon^{-\beta/(1-\beta)} \qquad as\ \varepsilon \to 0^+,$$

*with $\beta$ being defined as in Theorem 1(b).*
*(d) If $p_0 > 0$, then*

$$\mathbb{P}(\mathcal{W} \leq \varepsilon) \asymp \varepsilon^{|\log h(\rho)|/\log m} \qquad as\ \varepsilon \to 0^+, \tag{1.6}$$

*where $\rho$ is the solution of $f(s) = s$ between $(0,1)$, $f$ and $h$ are defined in* (1.1).

Note that there are additional phase transitions appearing in the case with immigration, in particular between the case where the immigration distribution has a positive mass at 0 and where there is no mass at 0. In the $p_0 > 0$ case, the extinction probability of the branching process $(Z_n, n \geq 0)$ (without immigration) is strictly positive, and plays the dominating role in the small value probability of $\mathcal{W}$. Our approach is outlined in Section 2 and detailed proof of Theorem 2 is give in sections 3, 4 and 5.

## 2. Our approach

Our proof of Theorem 2 is based on Dubuc's result stated in Theorem 1. In [D71b], an integral composition transform is used together with some nontrivial complex analysis, which is powerful but inflexible and un-intuitive. It seems impossible to extend the involved analytic method to the branching process with immigration. On the other hand, Mörters and Ortgiese [MO08] provided a very useful probabilistic approach for Theorem 1, called the "branching tree heuristic" method. Our approach is built on the top of their powerful arguments, and overcomes additional difficulties of immigration effects. More specifically, we start with a fundamental decomposition for $\mathcal{W}$ given in (2.2). Then a suitable truncation is used in order to handle the infinite series. To estimate the lower bound of $\mathbb{P}(\mathcal{W} \leq \varepsilon)$, we investigate when the least population size happens. For the upper bound, we use the exponential Chebyshev's inequality and estimate the Laplace transform of $\mathcal{W}$. The property of $\mathbb{P}(\mathcal{W} \leq \varepsilon)$ is then obtained through Tauberian type theorems.

Now we consider recursive distribution identities for $(\mathcal{Z}_n, n \geq 0)$ satisfying $\mathcal{Z}_0 = Y_0$. For fixed integers $r \geq 0$ and $l \geq 0$, let $\xi_r(1), \cdots, \xi_r(\mathcal{Z}_r)$ be the individuals in generation $r$, and $\eta_l(j), j = 1, \cdots, Y_l$ be the individuals of immigration in generation $l$. Then for any $r \geq 0$ and $n \geq r+1$,

$$\mathcal{Z}_n = \sum_{i=1}^{\mathcal{Z}_r} Z_{n-r}(\xi_r(i)) + \sum_{l=r+1}^{n} \sum_{j=1}^{Y_l} Z_{n-l}(\eta_l(j)).$$

Here $(Z_n(v), n \geq 0)$ is a supercritical G-W branching process initiated with one individual $v$ and $W(v)$ is the limit of the positive martingale $m^{-n} Z_n(v)$.

Dividing both sides of the above equality by $m^n$, then letting $n \to \infty$, we get

$$\mathcal{W} = m^{-r} \sum_{i=1}^{\mathcal{Z}_r} W(\xi_r(i)) + \sum_{l=r+1}^{\infty} m^{-l} \sum_{j=1}^{Y_l} W(\eta_l(j)). \tag{2.1}$$

For simplicity, we rewrite (2.1) as

$$\mathcal{W} = m^{-r} \sum_{i=1}^{\mathcal{Z}_r} W_i + \sum_{l=r+1}^{\infty} m^{-l} \sum_{j=1}^{Y_l} W_l^j. \tag{2.2}$$

Here all the $W_i, W_l^j, i = 1, \cdots, \mathcal{Z}_r, l = r+1, \cdots, n, j = 1, \cdots, Y_l$ are independent and identically distributed as $W$. The relation (2.2) is the fundamental distribution identity of $\mathcal{W}$ and it is used repeatedly in our approach.

Next we turn to consider a slightly different type of supercritical branching process with immigration, which is denoted by $(\widetilde{\mathcal{Z}}_n, n \geq 0)$. The only difference is to assume $\widetilde{\mathcal{Z}}_0 = 1$. The corresponding limit of $\widetilde{\mathcal{Z}}_n/m^n$ is denoted by $\widetilde{\mathcal{W}}$. Then by simple computation we get that

$$\widetilde{\mathcal{W}} =^d W + \frac{\mathcal{W}}{m} \tag{2.3}$$

in distribution, denoted by $=^d$ throughout this paper, where $W$ and $\mathcal{W}$ are independent. Then owing to (2.3) and the fact that

$$\begin{aligned} \mathbb{P}(W + \mathcal{W}/m \leq \varepsilon) &\geq \mathbb{P}(W \leq \varepsilon/2) \cdot \mathbb{P}(\mathcal{W}/m \leq \varepsilon/2), \\ \mathbb{P}(W + \mathcal{W}/m \leq \varepsilon) &\leq \mathbb{P}(W \leq \varepsilon) \cdot \mathbb{P}(\mathcal{W}/m \leq \varepsilon), \end{aligned} \tag{2.4}$$

we obtain the following result as a consequence of combining Theorem 1 and Theorem 2.

**Theorem 3** *Assume the condition (1.3) holds.*
*(a) If $p_0 = 0$, $p_1 > 0$ and $q_0 > 0$, then*

$$\mathbb{P}(\widetilde{\mathcal{W}} \leq \varepsilon) \asymp \varepsilon^{|\log(p_1 q_0)|/\log m} \qquad as\ \varepsilon \to 0^+.$$

*(b) If $p_0 = 0$, $p_1 > 0$ and $q_0 = 0$, then*

$$\log \mathbb{P}(\widetilde{\mathcal{W}} \leq \varepsilon) \sim -\frac{K|\log p_1|}{2(\log m)^2}|\log \varepsilon|^2 \qquad as\ \varepsilon \to 0^+,$$

*with $K$ being defined as in Theorem 2(b).*
*(c) If $p_0 = 0$ and $p_1 = 0$, then*

$$\log \mathbb{P}(\widetilde{\mathcal{W}} \leq \varepsilon) \asymp -\varepsilon^{-\beta/(1-\beta)} \qquad as\ \varepsilon \to 0^+,$$

*with $\beta$ being defined as in Theorem 1(b).*
*(d) If $p_0 > 0$, then*

$$\mathbb{P}(\widetilde{\mathcal{W}} \leq \varepsilon) \asymp \varepsilon^{|\log h(\rho)|/\log m} \qquad as\ \varepsilon \to 0^+.$$

Note that when $q_0 = 1$, i.e., without immigration, Theorem 3 recovers Theorem 1.

# 3. Proof of Theorem 2: Lower bound

We start with a simple but crucial probability estimate that is a consequence of the condition $\mathbb{E} \log^+ Y < \infty$ in (1.3).

**Lemma 1** *Under the condition that* $\mathbb{E}\log^+ Y < \infty$ *in* (1.3), *for any fixed constant* $\delta > 0$, *there exists an integer* $l$ *such that*

$$\mathbb{P}(\max_{i\geq l+1} Y_i e^{-\delta i} \leq 1) \quad \geq \quad e^{-1}. \tag{3.1}$$

**Proof.** For any given $\delta > 0$,

$$\begin{aligned}\sum_{i=1}^{\infty}\mathbb{P}(\log^+ Y \geq \delta i) &= \sum_{i=1}^{\infty}\sum_{k=i}^{\infty}\mathbb{P}(k \leq \delta^{-1}\log^+ Y < k+1)\\ &= \sum_{k=1}^{\infty} k\mathbb{E}\,\mathbb{I}(k \leq \delta^{-1}\log^+ Y < k+1)\\ &\leq \delta^{-1}\mathbb{E}\log^+ Y < \infty.\end{aligned}$$

Let $Y_i$ and $Y$ be our independent and identically distributed immigration random variables. Then for any large integer $l$ such that

$$\sum_{i=l+1}^{\infty}\mathbb{P}(\log^+ Y \geq \delta i) \leq 1/2 \tag{3.2}$$

we have

$$\begin{aligned}\mathbb{P}(\max_{i\geq l+1} Y_i e^{-\delta i} \leq 1) &\geq \prod_{i=l+1}^{\infty}\big(1-\mathbb{P}(\log^+ Y \geq \delta i)\big)\\ &\geq \exp\left(-2\sum_{i=l+1}^{\infty}\mathbb{P}(\log^+ Y \geq \delta i)\right)\\ &\geq e^{-1},\end{aligned}$$

here we used the fact that $(1-x)e^{2x}$ is increasing for $0 \leq x < 1/2$. This finishes our proof of the lemma.

**Proof of (a) and (b).** For any $\varepsilon > 0$, let $k = k_\varepsilon$ be the integer such that

$$m^{-k} \leq \varepsilon < m^{-k+1}, \tag{3.3}$$

which is equivalent to saying

$$k-1 < |\log\varepsilon|/\log m \leq k, \quad \text{or} \quad k = \lceil |\log\varepsilon|/\log m\rceil. \tag{3.4}$$

Using the fundamental distribution identity (2.2) with $r = 0$, we have for a fixed integer $l$ to be chosen later,

$$\begin{aligned}\mathbb{P}(\mathcal{W} \leq \varepsilon) &= \mathbb{P}\bigg(\sum_{i=0}^{\infty} m^{-i}\sum_{j=1}^{Y_i} W_i^j \leq \varepsilon\bigg)\\ &\geq \mathbb{P}\bigg(\sum_{i=0}^{k+l} m^{-i}\sum_{j=1}^{Y_i} W_i^j \leq \frac{\varepsilon}{2}\bigg)\cdot\mathbb{P}\bigg(\sum_{i=k+l+1}^{\infty} m^{-i}\sum_{j=1}^{Y_i} W_i^j \leq \frac{\varepsilon}{2}\bigg).\end{aligned} \tag{3.5}$$

The second term in (3.5) can be estimated by using $\varepsilon \ge m^{-k}$ in (3.3) as below

$$\begin{aligned}\mathbb{P}\Big(\sum_{i=k+l+1}^{\infty} m^{-i}\sum_{j=1}^{Y_i} W_i^j \le \frac{\varepsilon}{2}\Big) &\ge \mathbb{P}\Big(\sum_{i=k+l+1}^{\infty} m^{-i}\sum_{j=1}^{Y_i} W_i^j \le \frac{m^{-k}}{2}\Big)\\ &= \mathbb{P}\Big(\sum_{i=l+1}^{\infty} m^{-i}\sum_{j=1}^{Y_i} W_i^j \le \frac{1}{2}\Big). \qquad (3.6)\end{aligned}$$

Note that the last equality follows from the independence and identical distribution of all $W_i^j$'s and $Y_i$'s. Next we have by controlling the size of $Y_i$, $i \ge l+1$, given in Lemma 1,

$$\begin{aligned}&\mathbb{P}\Big(\sum_{i=l+1}^{\infty} m^{-i}\sum_{j=1}^{Y_i} W_i^j \le \frac{1}{2}\Big)\\ \ge\ &\mathbb{P}\Big(\sum_{i=l+1}^{\infty} m^{-i}\sum_{j=1}^{Y_i} W_i^j \le \frac{1}{2}, \max_{i\ge l+1} Y_i e^{-\delta i} \le 1\Big)\\ \ge\ &\mathbb{P}\Big(\sum_{i=l+1}^{\infty} m^{-i}\sum_{j=1}^{\lceil \exp(\delta i)\rceil} W_i^j \le \frac{1}{2}\Big)\cdot \mathbb{P}\Big(\max_{i\ge l+1} Y_i e^{-\delta i} \le 1\Big). \qquad (3.7)\end{aligned}$$

Using Chebyshev's inequality for the first part of (3.7), we get

$$\begin{aligned}\mathbb{P}\Big(\sum_{i=l+1}^{\infty} m^{-i}\sum_{j=1}^{\lceil \exp(\delta i)\rceil} W_i^j \le \frac{1}{2}\Big) &\ge 1-2\mathbb{E}\sum_{i=l+1}^{\infty} m^{-i}\sum_{j=1}^{\lceil \exp(\delta i)\rceil} W_i^j\\ &\ge 1-2\sum_{i=l+1}^{\infty} m^{-i}(e^{\delta i}+1). \qquad (3.8)\end{aligned}$$

For $\delta$ satisfying $e^{\delta} < m$, we have $\sum_{i=l+1}^{\infty} m^{-i}(e^{\delta i}+1) < \infty$. Then we choose $\delta$ small enough and integer $l$ large enough so that

$$2\sum_{i=l+1}^{\infty} m^{-i}(e^{\delta i}+1) < \frac{1}{2}. \qquad (3.9)$$

Combining (3.6)–(3.9) and Lemma 1, we obtain

$$\mathbb{P}\Big(\sum_{i=k+l+1}^{\infty} m^{-i}\sum_{j=1}^{Y_i} W_i^j \le \frac{\varepsilon}{2}\Big) \ge \mathbb{P}\Big(\sum_{i=l+1}^{\infty} m^{-i}\sum_{j=1}^{Y_i} W_i^j \le \frac{1}{2}\Big) \ge \frac{1}{2e}. \qquad (3.10)$$

Now back to the first part of (3.5), we have to handle it under conditions (a) and (b) separately. In the case (a) with $q_0 > 0$, we have the simple estimate

$$\mathbb{P}\Big(\sum_{i=0}^{k+l} m^{-i}\sum_{j=1}^{Y_i} W_i^j \le \frac{\varepsilon}{2}\Big) \ge \mathbb{P}\big(Y_0 = \cdots = Y_{k+l} = 0\big) = q_0^{k+l+1}. \qquad (3.11)$$

Using $k-1 < |\log \varepsilon|/\log m$ in (3.4), it's easy to deduce that

$$q_0^k \quad \geq \quad q_0 \cdot q_0^{|\log \varepsilon|/\log m} = q_0 \varepsilon^{|\log q_0|/\log m}. \tag{3.12}$$

Combining (3.5) and (3.10)–(3.12) we have shown the lower bound in Theorem 2(a).

For the case (b) with $q_0 = 0$, we have, recalling the definition of $K = \inf\{n : q_n > 0\}$,

$$\begin{aligned} \mathbb{P}\Big(\sum_{i=0}^{k+l} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \frac{\varepsilon}{2}\Big) \quad &\geq \quad \mathbb{P}\Big(\sum_{i=0}^{k+l} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \frac{\varepsilon}{2}, Y_0 = \cdots = Y_{k+l} = K\Big) \\ &= \quad \mathbb{P}\Big(\sum_{i=0}^{k+l} m^{-i} \sum_{j=1}^{K} W_i^j \leq \frac{\varepsilon}{2}\Big) \cdot q_K^{k+l+1}. \end{aligned} \tag{3.13}$$

The above probability of sums can be bounded termwise, and thus

$$\begin{aligned} \mathbb{P}\Big(\sum_{i=0}^{k+l} m^{-i} \sum_{j=1}^{K} W_i^j \leq \frac{\varepsilon}{2}\Big) \quad &\geq \quad \mathbb{P}\Big(\max_{0\leq i\leq k+l}\max_{1\leq j\leq K} m^{-i} W_i^j \leq \frac{\varepsilon/2}{K(k+l+1)}\Big) \\ &= \quad \prod_{i=0}^{k+l} \mathbb{P}^K\Big(m^{-i} W \leq \frac{\varepsilon/2}{K(k+l+1)}\Big) \\ &\geq \quad \prod_{i=0}^{k+l} \mathbb{P}^K\Big(W \leq \frac{m^{i-k}/2}{K(k+l+1)}\Big), \end{aligned} \tag{3.14}$$

where we use the independent and identically distributed property of all $W_i^j$'s in the last equality and $\varepsilon \geq m^{-k}$ from (3.3) in the last inequality.

From Theorem 1(a) there exists a constant $c > 0$ such that, for $i = 0, 1, \cdots, k+l$,

$$\mathbb{P}\Big(W \leq \frac{m^{i-k}/2}{K(k+l+1)}\Big) \quad \geq \quad c\Big(\frac{m^{i-k}/2}{K(k+l+1)}\Big)^{|\log p_1|/\log m}. \tag{3.15}$$

Combining (3.5), (3.10) and (3.13)–(3.15) together, and taking summation over $0 \leq i \leq k+l$ after taking logarithm, we have

$$\begin{aligned} \log \mathbb{P}(\mathcal{W} \leq \varepsilon) \quad &\geq \quad -\frac{K|\log p_1|}{2} k^2 - O(k \log k) \\ &\geq \quad -\frac{K|\log p_1|}{2(\log m)^2} |\log \varepsilon|^2 - O(\log \varepsilon^{-1} \log\log \varepsilon^{-1}), \end{aligned}$$

which follows easily from $k < 1 + |\log \varepsilon|/\log m$ in (3.4).

**Proof of (c).** First observe that, in this setting with $\gamma = \inf\{n : p_n > 0\} \geq 2$, $K = \inf\{n : q_n > 0\} \geq 1$, the smallest number of particles in generation $n$ ($n \geq 1$) is

$$b(n) := K(\gamma^n + \gamma^{n-1} + \cdot\cdot +1) = K(\gamma^{n+1} - 1)/(\gamma - 1). \tag{3.16}$$

It is also easy to see that the chance this occurs is

$$\mathbb{P}(\mathcal{Z}_n = b(n)) = p_\gamma^{b(n-1)+\cdots+b(0)} q_K^{n+1} := p_\gamma^{B(n)} q_K^{n+1}, \tag{3.17}$$

where

$$B(0) = 0, \qquad B(n) = b(n-1) + \cdots + b(0) = \frac{K\left(\gamma^{n+1} - (n+1)\gamma + n\right)}{(\gamma-1)^2}, \ n \geq 1. \tag{3.18}$$

Given $\varepsilon > 0$, we can choose $k = k_\varepsilon$ such that

$$\frac{\gamma^k}{m^k} \leq \varepsilon < \frac{\gamma^{k-1}}{m^{k-1}}, \tag{3.19}$$

which is equivalent to saying

$$k - 1 < |\log \varepsilon| / \log(m/\gamma) \leq k, \quad \text{or} \quad k = \lceil |\log \varepsilon| / \log(m/\gamma) \rceil. \tag{3.20}$$

Next let $l$ be an integer that will be determined later. Using the fundamental distribution identity (2.2) with $r = k + l$ and (3.17), we have

$$\begin{aligned}
& \mathbb{P}(\mathcal{W} \leq \varepsilon) \\
\geq\ & \mathbb{P}\big(\mathcal{W} \leq (\gamma/m)^k | \mathcal{Z}_{k+l} = b(k+l)\big)\mathbb{P}(\mathcal{Z}_{k+l} = b(k+l)) \\
=\ & \mathbb{P}\bigg(m^{-k-l} \sum_{i=1}^{b(k+l)} W_i + \sum_{i=k+l+1}^{\infty} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq (\gamma/m)^k\bigg) p_\gamma^{B(k+l)} q_K^{k+l+1} \\
\geq\ & \mathbb{P}\bigg(\sum_{i=1}^{b(k+l)} W_i \leq \frac{m^l\gamma^k}{2}\bigg)\mathbb{P}\bigg(\sum_{i=1}^{\infty} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \frac{m^l\gamma^k}{2}\bigg) p_\gamma^{B(k+l)} q_K^{k+l+1}. \qquad (3.21)
\end{aligned}$$

For the first term in (3.21) we have by Chebyshev's inequality and choosing suitable $l$

$$\begin{aligned}
\mathbb{P}\bigg(\sum_{i=1}^{b(k+l)} W_i \leq m^l\gamma^k/2\bigg) \quad &\geq \quad 1 - \frac{2}{m^l\gamma^k}\mathbb{E} \sum_{i=1}^{b(k+l)} W_i = 1 - \frac{2b(k+l)}{m^l\gamma^k} \\
&\geq \quad 1 - \frac{2K\gamma}{\gamma-1}(\gamma/m)^l \geq 1/2, \qquad (3.22)
\end{aligned}$$

where $\mathbb{E}\, W = 1$ and $b(n) \leq K(\gamma-1)^{-1}\gamma^{n+1}$ from (3.16) are used.

For the second part of (3.21), we have

$$\begin{aligned}
\mathbb{P}\bigg(\sum_{i=1}^{\infty} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \frac{m^l\gamma^k}{2}\bigg) \quad &= \quad \mathbb{P}\bigg(\sum_{i=l+1}^{\infty} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \frac{\gamma^k}{2}\bigg) \\
&\geq \quad \mathbb{P}\bigg(\sum_{i=l+1}^{\infty} m^{-i} \sum_{j=1}^{Y_i} W_i^j \leq \frac{1}{2}\bigg) \geq e^{-1}/2, \qquad (3.23)
\end{aligned}$$

where the last inequality follows from (3.10).

Combing (3.21)–(3.23), we get

$$\mathbb{P}(\mathcal{W} \le \varepsilon) \ge p_\gamma^{B(k+l)} q_K^{k+l+1} e^{-1}/4. \tag{3.24}$$

Recalling the definition of $B(k+l)$ in (3.18) and $k-1 < |\log \varepsilon|/\log(m/\gamma)$ in (3.20), we see

$$B(k+l) \le \frac{K}{(\gamma-1)^2}\gamma^{k+l+1} \le C\gamma^{|\log\varepsilon|/\log(m/\gamma)} = C\varepsilon^{-\beta/(1-\beta)},$$

where $\beta$ is defined as in Theorem 1(b) and $C$ is a positive constant. Therefore from (3.24) we obtain

$$\log \mathbb{P}(\mathcal{W} \le \varepsilon) \ge -C\varepsilon^{-\beta/(1-\beta)}$$

for some constant $C > 0$.

# 4. Proof of Theorem 2: Upper bound

As we can see from the arguments in Section 3, only the finite terms in (2.2) are contributing to the small value probabilities of $\mathcal{W}$. Hence we take only $r = 0$ in (2.2), choose a suitable cut off $k$, and focus on properties of $\sum_{l=0}^k m^{-l} \sum_{j=1}^{Y_l} W_l^j$.

**Proof of (a).** Let $k = k_\varepsilon$ be the integer defined as in (3.3). Using the fundamental distribution identity (2.2) with $r = 0$ and exponential Chebyshev's inequality, we have

$$\begin{aligned}\mathbb{P}(\mathcal{W} \le \varepsilon) &\le \mathbb{P}\Big(\sum_{i=0}^k m^{-i} \sum_{j=1}^{Y_i} W_i^j \le \varepsilon\Big) \\ &\le e^{\lambda\varepsilon} \cdot \mathbb{E}\exp\Big(-\lambda \sum_{i=0}^k m^{-i} \sum_{j=1}^{Y_i} W_i^j\Big) \qquad \text{for any } \lambda > 0. \end{aligned} \tag{4.1}$$

Noticing that all the $(W_i^j, i = 0, \cdots, k, j = 1, \cdots, Y_i)$ are independent, we have

$$\mathbb{E}\exp\Big(-\lambda\sum_{i=0}^k m^{-i}\sum_{j=1}^{Y_i} W_i^j\Big) = \prod_{i=0}^k \mathbb{E}\exp\Big(-\lambda m^{-i}\sum_{j=1}^{Y_i} W_i^j\Big). \tag{4.2}$$

Conditioning on $Y_i = 0$ or $Y_i \ge 1$, we have

$$\mathbb{E}\exp\Big(-\lambda m^{-i}\sum_{j=1}^{Y_i} W_i^j\Big) \le q_0 + (1-q_0)\mathbb{E}\exp\big(-\lambda m^{-i} W_i^1\big) \le q_0(1+\delta_i), \tag{4.3}$$

where

$$\delta_i = q_0^{-1}\mathbb{E}\,\exp\big(-\lambda m^{-i}W_i^1\big) = q_0^{-1}\mathbb{E}\,\exp\big(-\lambda m^{-i}W\big),\ i=0,\cdots,k. \tag{4.4}$$

Substituting (4.3) into (4.1) and letting $\lambda = \varepsilon^{-1}$, we obtain

$$\mathbb{P}(\mathcal{W}\le\varepsilon)\le e q_0^{k+1}\prod_{i=0}^{k}(1+\delta_i).$$

Since $k\ge|\log\varepsilon|/\log m$ in (3.4), we have

$$q_0^k\le\varepsilon^{|\log q_0|/\log m}.$$

So we finish the proof by showing

$$\sum_{i=0}^{k}\log(1+\delta_i)\le\sum_{i=0}^{k}\delta_i\le M, \tag{4.5}$$

where $M>0$ is a constant independent of $\varepsilon$ (noticing that the $k$ depends on $\varepsilon$).

In order to estimate $\delta_i$, we need the following fact given in Li [L12+].

**Lemma 2** *(i) Assume $V$ is a positive random variable and $\alpha>0$ is a constant. Then*

$$\mathbb{P}(V\le t)\le C_1t^{\alpha}\quad \textit{for some constant } C_1>0 \textit{ and all } t>0$$

*is equivalent to*

$$\mathbb{E}\,e^{-\lambda V}\le C_2\lambda^{-\alpha}\quad \textit{for some constant } C_2>0 \textit{ and all } \lambda>0.$$

*(ii) Assume $V$ is a positive random variable and $\alpha>0, \theta\in\mathbb{R}$, or $\alpha=0,\theta>0$ are constants. Then we have*

$$\log\mathbb{P}(V\le t)\le -C_1t^{-\alpha}|\log t|^{\theta}\quad \textit{for some constant } C_1>0 \textit{ and all } t>0$$

*is equivalent to*

$$\log\mathbb{E}\,e^{-\lambda V}\le -C_2\lambda^{\alpha/(1+\alpha)}(\log\lambda)^{\theta/(1+\alpha)}\quad \textit{for some constant } C_2>0 \textit{ and all } \lambda>0.$$

To show (4.5), we have to argue separately according to $p_1>0$ or $p_1=0$. When $p_1>0$, by Theorem 1(a) and Lemma 2(i), there exists a constant $C>0$ satisfying that

$$\mathbb{E}\,e^{-\lambda W}\le C\lambda^{-|\log p_1|/\log m},\qquad \lambda>0. \tag{4.6}$$

Combining (4.4) with $\lambda = \varepsilon^{-1}$, then using (4.6), we have

$$\begin{aligned}\sum_{i=0}^{k} \delta_i &= q_0^{-1} \sum_{i=0}^{k} \mathbb{E} \exp(-\varepsilon^{-1} m^{-i} W) \\ &\le q_0^{-1} C \sum_{i=0}^{k} (\varepsilon m^i)^{|\log p_1|/\log m} \\ &= C q_0^{-1} \varepsilon^{|\log p_1|/\log m} \sum_{i=0}^{k} p_1^{-i} \\ &\le C' \varepsilon^{|\log p_1|/\log m} \cdot p_1^{-k} \le C' p_1^{-1},\end{aligned}$$

where $C'$ is a constant and the last inequality follows from (3.4).

When $p_1 = 0$, using Theorem 1(b) and Lemma 2(ii) with $\alpha = \beta/(1-\beta)$ and $\theta = 0$, we have for some constant $b > 0$,

$$\log \mathbb{E}\, e^{-\lambda W} \le -b\lambda^{\beta}, \qquad \lambda > 0, \tag{4.7}$$

from which it's similar to show that (4.5) holds. Indeed, setting $\lambda = \varepsilon^{-1}$ in (4.4), and then using (4.7) and $\varepsilon < m^{-k+1}$ from (3.3), we obtain

$$\begin{aligned}\sum_{i=0}^{k} \delta_i &= q_0^{-1} \sum_{i=0}^{k} \mathbb{E} \exp(-\varepsilon^{-1} m^{-i} W) \\ &\le q_0^{-1} \sum_{i=0}^{k} \exp(-b\varepsilon^{-\beta} m^{-i\beta}) \\ &\le q_0^{-1} \sum_{i=0}^{k} \exp(-b m^{(k-i-1)\beta}) \\ &\le q_0^{-1} \sum_{i=0}^{\infty} \exp(-b m^{(i-1)\beta}) < \infty.\end{aligned}$$

**Proof of (b).** Let $k$ be defined as in (3.3). Using (4.1) and $Y_i \ge K$ for any $i \ge 0$,

$$\mathbb{P}(\mathcal{W} \le \varepsilon) \le e^{\lambda\varepsilon} \prod_{i=0}^{k} \prod_{j=1}^{K} \mathbb{E} \exp\left(-\lambda m^{-i} W_i^j\right), \qquad \lambda > 0. \tag{4.8}$$

In the case (b) with $p_1 > 0$, substituting (4.6) into (4.8) with $\lambda = \varepsilon^{-1}$, we obtain

$$\mathbb{P}(\mathcal{W} \le \varepsilon) \le e \prod_{i=0}^{k} \prod_{j=1}^{K} C(\varepsilon m^i)^{|\log p_1|/\log m}.$$

Taking the logarithm we obtain

$$\begin{aligned}
\log \mathbb{P}(\mathcal{W} \le \varepsilon) &\le 1 + K(k+1)(\log C - |\log \varepsilon| \cdot |\log p_1|/\log m) + k(k+1) \cdot K|\log p_1|/2 \\
&= -k \cdot |\log \varepsilon| \cdot K|\log p_1|/\log m + (k-1)^2 \cdot K|\log p_1|/2 + O(k) \\
&\le -\frac{K|\log p_1|}{2(\log m)^2}|\log \varepsilon|^2 + O(|\log \varepsilon|),
\end{aligned}$$

where the last inequality follows from $k - 1 < |\log \varepsilon|/\log m \le k$, which is given in (3.4).

**Proof of (c).** It is clear that

$$\mathbb{P}(\mathcal{W} \le \varepsilon) \le \mathbb{P}(W \le \varepsilon), \tag{4.9}$$

and therefore we finish the proof of $(c)$ by using estimate in Theorem 1(b).

## 5. Proof of Theorem 2(d)

If $p_0 > 0$, then $f(s) = s$ has a unique solution $\rho \in (0, 1)$ and $\mathbb{P}(W = 0) = \rho$. By means of the Harris-Sevastyanov transformation

$$\widetilde{f}(s) := \frac{f((1-\rho)s + \rho) - \rho}{(1-\rho)},$$

$\widetilde{f}$ defines a new branching mechanism with $\widetilde{p}_0 = 0$ and $\widetilde{f}'(1) = m$. We use $(\widetilde{Z}_n, n \ge 0)$ to denote the corresponding branching process and $\widetilde{W}$ to denote the limit of $m^{-n}\widetilde{Z}_n$. By Theorem 3.2 in [H48],

$$W =^d W_0 \cdot \widetilde{W}, \tag{5.1}$$

where $W_0$ is independent of $\widetilde{W}$ and takes the values 0 and $1/(1-\rho)$ with probabilities $\rho$ and $1-\rho$ respectively. Notice that the small value probability of $\widetilde{W}$ has the asymptotic behavior described in Theorem 1(a) with $\widetilde{p}_1 = \widetilde{f}'(0) = f'(\rho) > 0$, and $\tau = |\log \widetilde{p}_1|/\log m$, that is

$$\mathbb{P}(\widetilde{W} \le \varepsilon) \asymp \varepsilon^\tau. \tag{5.2}$$

Now we start to prove the lower bound. For any $\varepsilon > 0$, let $k = k_\varepsilon$ be the integer defined in (3.3). Then using (3.5) and (3.10), we only need to estimate the first part of (3.5):

$$\begin{aligned}
\mathbb{P}\bigg(\sum_{i=0}^{k+l} m^{-i} \sum_{j=1}^{Y_i} W_i^j \le \frac{\varepsilon}{2}\bigg) &\ge \prod_{i=0}^{k+l} \mathbb{P}\big(\sum_{j=1}^{Y_i} W_i^j = 0\big) \\
&= \prod_{i=0}^{k+l} \left(\sum_{n=0}^{\infty} q_n \mathbb{P}^n(W=0)\right) = h(\rho)^{k+l+1},
\end{aligned} \tag{5.3}$$

where $h$ is the generating function of immigration $Y$. Using $k-1<|\log\varepsilon|/\log m$ given in (3.4), it's easy to deduce that

$$h(\rho)^k \quad\geq\quad h(\rho)\cdot h(\rho)^{|\log\varepsilon|/\log m}=h(\rho)\cdot\varepsilon^{|\log h(\rho)|/\log m}. \tag{5.4}$$

Combining (3.5), (3.10), (5.3) and (5.4) we obtain the lower bound of (d).

Next we show the upper bound. Using (5.1), we have

$$\mathbb{E}\,e^{-\lambda W}=\rho+\mathbb{E}\,e^{-\lambda W}\mathbb{I}_{\{W>0\}}:=\rho+\delta(\lambda),\qquad \lambda>0. \tag{5.5}$$

Using (4.1), (4.2) and the independent and identically distributed property of all the $(W_i^j, i=0,\cdots,k, j=1,\cdots,Y_i)$, we have

$$\begin{aligned}\mathbb{P}(\mathcal{W}\leq\varepsilon) &\leq e^{\lambda\varepsilon}\prod_{i=0}^{k}h\left(\rho+\delta(\lambda m^{-i})\right)\\ &= (h(\rho))^{k+1}\exp\left(\lambda\varepsilon+\sum_{i=0}^{k}\log\left(h\left(\rho+\delta(\lambda m^{-i})\right)/h(\rho)\right)\right),\end{aligned} \tag{5.6}$$

where $\lambda=\lambda_k$ depends on $k(=k_\varepsilon)$ and is given later. Since $k\geq|\log\varepsilon|/\log m$ from (3.4), we have

$$(h(\rho))^k\leq\varepsilon^{|\log h(\rho)|/\log m}. \tag{5.7}$$

Next we show there is a constant $M>0$, which does not depend on $\varepsilon$, such that

$$\begin{aligned}&\lambda\varepsilon+\sum_{i=0}^{k}\log\left(h\left(\rho+\delta(\lambda m^{-i})\right)/h(\rho)\right)\\ \leq\ &\lambda m^{-k+1}+h(\rho)^{-1}\sum_{i=0}^{k}\left(h\left(\rho+\delta(\lambda m^{-i})\right)-h(\rho)\right)\leq M.\end{aligned} \tag{5.8}$$

Since $\delta(\lambda m^{-x})$ is increasing with respect to $x$, we have

$$\sum_{i=0}^{k}\left(h\left(\rho+\delta(\lambda m^{-i})\right)-h(\rho)\right)\leq\int_0^{k+1}\left(h\left(\rho+\delta(\lambda m^{-x})\right)-h(\rho)\right)\mathrm{d}x. \tag{5.9}$$

Note that $\delta(\lambda)=(1-\rho)\mathbb{E}\,e^{-(\lambda/(1-\rho))\widetilde{W}}$. By (5.2) and Lemma 2(i), there exists a constant $C>0$ such that

$$\delta(\lambda m^{-x})\leq C(\lambda m^{-x})^{-\tau} \tag{5.10}$$

with $\tau = |\log f'(\rho)|/\log m$. Thus we have

$$\begin{aligned}
& \sum_{i=0}^{k} \left(h\left(\rho + \delta(\lambda m^{-i})\right) - h(\rho)\right) \\
\leq \;& \int_0^{k+1} \left(h\left(\rho + C(\lambda m^{-x})^{-\tau}\right) - h(\rho)\right) \mathrm{d}x \\
= \;& 1/(\tau \log m) \cdot \int_{\lambda^{-\tau}}^{\lambda^{-\tau} m^{(k+1)\tau}} 1/y \cdot \left(h\left(\rho + Cy\right) - h(\rho)\right) \mathrm{d}y \\
\leq \;& 1/(\tau \log m) \cdot \int_{0}^{\lambda^{-\tau} m^{(k+1)\tau}} 1/y \cdot \left(h\left(\rho + Cy\right) - h(\rho)\right) \mathrm{d}y. \qquad (5.11)
\end{aligned}$$

As $\rho < 1$, we may choose $\delta_0 > 0$ such that $\rho + \delta_0 < 1$. Next we choose $\lambda = \left(C/\delta_0\right)^{1/\tau} m^{(k+1)}$ in order to assure $\rho + Cy < 1$ so that $h\left(\rho + Cy\right)$ is well defined. Indeed we have

$$\lambda m^{-k+1} = m^2 \left(C/\delta_0\right)^{1/\tau} := M_1, \qquad (5.12)$$

and

$$\rho + Cy \leq \rho + C\lambda^{-\tau} m^{(k+1)\tau} = \rho + \delta_0 < 1, \qquad y \leq \lambda^{-\tau} m^{(k+1)\tau}.$$

Then we follow (5.11) to get

$$\begin{aligned}
& \sum_{i=0}^{k} \left(h\left(\rho + \delta(\lambda m^{-i})\right) - h(\rho)\right) \\
\leq \;& 1/(\tau \log m) \cdot \int_0^{\delta_0/C} 1/y \cdot \left(h\left(\rho + Cy\right) - h(\rho)\right) \mathrm{d}y \\
:= \;& M_2 < \infty, \qquad (5.13)
\end{aligned}$$

where we used

$$\lim_{y \to 0} 1/y \cdot \left(h\left(\rho + Cy\right) - h(\rho)\right) = Ch'(\rho) < \infty.$$

From (5.8), (5.12) and (5.13) we obtain that (5.8) holds with $M = M_1 + M_2$, and finish the proof of Theorem 2(d).

## Acknowledgements

We would like to thank referees for insightful remarks and suggestions. The second author is supported in part by DMS–0805929, DMS–1106938, NSFC-10928103. The third author is supported in part by NSFC (Grant No. 10971003 and 11128101) and Specialized Research Fund for the Doctoral Program of Higher Education of China.

# References


[AN72] Athreya, K. B. and Ney, P. E. (1972). *Branching Processes.* New York: Springer-Verlag.

[BGMS12] Berestycki, N., Gantert, N., Mörters, P. and Sidorova, N. (2012). *Galton-Watson trees with vanishing martingale limit.* arXiv:1204.3080 [math.PR].

[B88] Bingham, N. H. (1988). On the limit of a supercritical branching process. *J. Appl. Probab.* **25A** 215-228.

[BB91] Biggins, J. D. and Bingham, N. H. (1991). Near-constancy of phenomena in branching processes. *Math. Proc. Cambridge Philos. Soc.* **110** 545-558.

[BB93] Biggins, J. D. and Bingham, N. H. (1993). Large deviation in the supercritical branching process. *Adv. Appl. Probab.* **25** 757-772.

[BN93] Biggins, J. D. and Nadarajah, S. (1993). Near-constancy of the Harris function in the simple branching process. *Comm. Statist. Stochastic Models* **n.3** 435-444.

[BP88] Barlow, M. T. and Perkins, E. A. (1988). Brownian Motion on the sierpinski gasket. *Probab. Theory Related Fields* **79** 543-623.

[D71a] Dubuc, S. (1971a). La Densité de la loi-limite *d'*un processus en cascade expansif. *Z. Wahrsch. Verw. Gebiete* **19** 281-290.

[D71b] Dubuc, S. (1971b). Problèmes relatifs *à l'*itération de fonctions suggé*ré*s par les processus en cascade. *Ann. Inst. Fourier(Grenoble)* **21** 171-251.

[D82] Dubuc, S. (1982). Étude théorique et numérique de la fonction de Karlin-McGregor. *J. Anal. Math.* **42** 15-37.

[FW07] Fleischmann, K. and Wachtel, V. (2007). Lower deviation probabilities for supercritical Galton-Watson processes. *Ann. Inst. H. Poincaré Probab. Stat.* **43** 233-255.

[FW09] Fleischmann, K. and Wachtel, V. (2009). On the left tail asymtotics for the limit law of supercritical Galton-Watson processes in the *Böttcher* case. *Ann. Inst. H. Poincaré Probab. Stat.* **45** 201-225.

[G87] Goldstein, S. (1987). Random walks and diffusions on fractals. *Percolation theory and ergodic theory of infinite particle systems (Minneapolis, Minn., 1984C1985), 121C129, IMA Vol. Math. Appl., 8, Springer, New York.*

[H95] Hambly, B. M. (1995). On constant tail behavior for the limiting random variable in a supercritical branching process. *J. Appl. Probab. Trust* **32** 267-273.

[H48] Harris, T. E. (1948). Branching processes. *Ann. Math. Statist.* **19** 474-494.

[KM68a] Karlin, S. and McGregor, J. (1968a). Embeddability of discrete time simple branching processes into continuous time branching processes. *Trans. Amer. Math. Soc.* **132** 115-136.

[KM68b] Karlin, S. and McGregor, J. (1968b). Embedding iterates of analytic functions with two fixed points into continuous groups. *Trans. Amer. Math. Soc.* **132** 137-145.

[KS66] Kesten, H. and Stigum, B. (1966). A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Statist.* **37** 1211-1223.

[K87] Kusuoka, S. (1987). A diffusion process on a fractal. *Probabilistic methods in mathematical physics (Katata/Kyoto, 1985)* 251-274.

[L03] Li, W. V. (2003). The first exit time of Brownian motion from unbounded convex domain, *Ann. Probab.* **31** 1078-1096.

[L12+] Li, W. V. (2012+). Small Value Probabilities: Techniques and Applications. Lecture notes. `http://www.math.udel.edu/~wli/svp.html`

[Lif11] Lifshits, M. A. (2011). Bibliography compilation on small deviation probabilities. `http://www.proba.jussieu.fr/pageperso/smalldev/biblio.pdf`, 2011

[LS01] Li, W. V. and Shao, Q.-M.(2001). Gaussian processes: inequalities, small ball probabilities and applications. *Stochastic processes: theory and methods, Handbook of Statist.* **19** 533–597.

[LS04] Li, W. V. and Shao, Q.-M. (2004). Lower tail probabilities for Gaussian processes. *Ann. Probab.* **32** 216-242.

[MO08] Mörters, P. and Ortgiese, M. (2008). Small value probabilities via the branching tree heuristic. *Bernoulli* **14(1)** 277-299.

[S70] Seneta, E. (1970). On the supercrirical branching process with immigration. *Math. Biosci.* **7** 9-14.

[W08] Williams, D. (2008). *Probability with Martingales.* Cambridge: Cambridge Mathematical textbooks.