



**HAL**  
open science

# BERRY-ESSEEN'S BOUND AND CRAMÉR'S LARGE DEVIATION EXPANSION FOR A SUPERCRITICAL BRANCHING PROCESS IN A RANDOM ENVIRONMENT

Ion Grama, Quansheng Liu, Eric Miqueu

► **To cite this version:**

Ion Grama, Quansheng Liu, Eric Miqueu. BERRY-ESSEEN'S BOUND AND CRAMÉR'S LARGE DEVIATION EXPANSION FOR A SUPERCRITICAL BRANCHING PROCESS IN A RANDOM ENVIRONMENT. *Stochastic Processes and their Applications*, 2017, 127, pp.1255-1281. 10.1016/j.spa.2016.07.014 . hal-01270162

**HAL Id: hal-01270162**

**<https://hal.science/hal-01270162>**

Submitted on 5 Feb 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# BERRY-ESSEEN'S BOUND AND CRAMÉR'S LARGE DEVIATION EXPANSION FOR A SUPERCRITICAL BRANCHING PROCESS IN A RANDOM ENVIRONMENT

ION GRAMA, QUANSHENG LIU, AND ERIC MIQUEU

ABSTRACT. Let  $(Z_n)$  be a supercritical branching process in a random environment  $\xi = (\xi_n)$ . We establish a Berry-Esseen bound and a Cramér's type large deviation expansion for  $\log Z_n$  under the annealed law  $\mathbb{P}$ . We also improve some earlier results about the harmonic moments of the limit variable  $W = \lim_{n \rightarrow \infty} W_n$ , where  $W_n = Z_n / \mathbb{E}_\xi Z_n$  is the normalized population size.

## 1. INTRODUCTION AND MAIN RESULTS

A branching process in a random environment (BPRE) is a natural and important generalisation of the Galton-Watson process, where the reproduction law varies according to a random environment indexed by time. It was introduced for the first time in Smith and Wilkinson [24] to modelize the growth of a population submitted to an environment. For background concepts and basic results concerning a BPRE we refer to Athreya and Karlin [4, 3]. In the critical and subcritical regime the process goes out and the research interest is concentrated mostly on the survival probability and conditional limit theorems for the branching process, see e.g. Afanasyev, Böinghoff, Kersting and Vatutin [1, 2], Vatutin [26], Vatutin and Zheng [27], and the references therein. In the supercritical case, a great deal of current research has been focused on large deviation principle, see Bansaye and Berestycki [5], Böinghoff and Kersting [12], Bansaye and Böinghoff [6, 7, 8], Huang and Liu [17]. In the particular case when the offspring distribution is geometric, precise asymptotics can be found in Kozlov [19], Böinghoff [11], Nakashima [21]. In this article, we complete on these results by giving the Berry-Esseen bound and asymptotics of large deviations of Cramér's type for a supercritical BPRE.

A BPRE can be described as follows. The random environment is represented by a sequence  $\xi = (\xi_0, \xi_1, \dots)$  of independent and identically distributed random variables (i.i.d. r.v.'s); each realization of  $\xi_n$  corresponds to a probability law  $\{p_i(\xi_n) : i \in \mathbb{N}\}$  on  $\mathbb{N} = \{0, 1, 2, \dots\}$ , whose probability generating function is

$$(1.1) \quad f_{\xi_n}(s) = f_n(s) = \sum_{i=0}^{\infty} p_i(\xi_n) s^i, \quad s \in [0, 1], \quad p_i(\xi_n) \geq 0, \quad \sum_{i=0}^{\infty} p_i(\xi_n) = 1.$$

---

*Date:* February 5, 2016.

*2000 Mathematics Subject Classification.* Primary 60F17, 60J05, 60J10. Secondary 37C30 .

*Key words and phrases.* Branching processes, random environment, harmonic moments, Stein's method, Berry-Esseen bound, change of measure.

Define the process  $(Z_n)_{n \geq 0}$  by the relations

$$(1.2) \quad Z_0 = 1, \quad Z_{n+1} = \sum_{i=1}^{Z_n} N_{n,i}, \quad \text{for } n \geq 0,$$

where  $N_{n,i}$  is the number of children of the  $i$ -th individual of the generation  $n$ . Conditionally on the environment  $\xi$ , the r.v.'s  $N_{n,i}$  ( $i = 1, 2, \dots$ ) are independent of each other with common probability generating function  $f_n$ , and also independent of  $Z_n$ .

In the sequel we denote by  $\mathbb{P}_\xi$  the *quenched law*, i.e. the conditional probability when the environment  $\xi$  is given, and by  $\tau$  the law of the environment  $\xi$ . Then  $\mathbb{P}(dx, d\xi) = \mathbb{P}_\xi(dx)\tau(d\xi)$  is the total law of the process, called *annealed law*. The corresponding quenched and annealed expectations are denoted respectively by  $\mathbb{E}_\xi$  and  $\mathbb{E}$ . We also define, for  $n \geq 0$ ,

$$m_n = m_n(\xi) = \sum_{i=0}^{\infty} i p_i(\xi_n) \quad \text{and} \quad \Pi_n = \mathbb{E}_\xi Z_n = m_0 \dots m_{n-1},$$

where  $m_n$  represents the average number of children of an individual of generation  $n$  when the environment  $\xi$  is given. Let

$$(1.3) \quad W_n = \frac{Z_n}{\Pi_n}, \quad n \geq 0,$$

be the normalized population size. It is well known that under  $\mathbb{P}_\xi$ ,  $(W_n)_{n \geq 0}$  is a non-negative martingale with respect to the filtration

$$\mathcal{F}_n = \sigma(\xi, N_{k,i}, 0 \leq k \leq n-1, i = 1, 2, \dots),$$

where by convention  $\mathcal{F}_0 = \sigma(\xi)$ . Then the limit  $W = \lim W_n$  exists  $\mathbb{P}$ -a.s. and  $\mathbb{E}W \leq 1$ .

An important tool in the study of a BPRE is the associated random walk

$$S_n = \log \Pi_n = \sum_{i=1}^n X_i, \quad n \geq 1,$$

where the r.v.'s  $X_i = \log m_{i-1}$  ( $i \geq 1$ ) are i.i.d. depending only on the environment  $\xi$ . It turns out that the behavior of the process  $(Z_n)$  is mainly determined by the associated random walk which is seen from the decomposition

$$(1.4) \quad \log Z_n = S_n + \log W_n.$$

For the sake of brevity set  $X = \log m_0$ ,

$$\mu = \mathbb{E}X \quad \text{and} \quad \sigma^2 = \mathbb{E}(X - \mu)^2.$$

We shall assume that the BPRE is supercritical, with  $\mu \in (0, \infty)$ ; together with  $\mathbb{E}|\log(1 - p_0(\xi_0))| < \infty$  this implies that the population size tends to infinity with positive probability (see [4]). We also assume that the random walk  $(S_n)$  is non-degenerate with  $0 < \sigma^2 < \infty$ ; in particular this implies that

$$(1.5) \quad \mathbb{P}(Z_1 = 1) = \mathbb{E}p_1(\xi_0) < 1.$$

Throughout the paper, we assume the following condition:

$$(1.6) \quad \mathbb{E} \frac{Z_1 \log^+ Z_1}{m_0} < \infty,$$

which implies that the martingale  $W_n$  converges to  $W$  in  $L^1(\mathbb{P})$  (see e.g. [25]) and

$$\mathbb{P}(W > 0) = \mathbb{P}(Z_n \rightarrow \infty) = \lim_{n \rightarrow \infty} \mathbb{P}(Z_n > 0) > 0.$$

Furthermore, we assume in the sequel that each individual has at least one child, which means that

$$(1.7) \quad p_0 = 0 \quad \mathbb{P} - \text{a.s.}$$

In particular this implies that the associated random walk has positive increments,  $Z_n \rightarrow \infty$  and  $W > 0$   $\mathbb{P}$  - a.s. Throughout the paper, we denote by  $C$  an absolute constant whose value may differ from line to line.

Our first result is a Berry-Esseen type bound for  $\log Z_n$ , which holds under the following additional assumptions:

**A1.** There exists a constant  $\varepsilon > 0$  such that

$$(1.8) \quad \mathbb{E} X^{3+\varepsilon} < \infty.$$

**A2.** There exists a constant  $p > 1$  such that

$$(1.9) \quad \mathbb{E} \left( \frac{Z_1}{m_0} \right)^p < \infty.$$

**Theorem 1.1.** *Under conditions **A1** and **A2**, we have*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left( \frac{\log Z_n - n\mu}{\sigma\sqrt{n}} \leq x \right) - \Phi(x) \right| \leq \frac{C}{\sqrt{n}},$$

where  $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$  is the standard normal distribution function.

Theorem 1.1 completes the results of [17] by giving the rate of convergence in the central limit theorem for  $\log Z_n$ . The proof of this theorem is based on Stein's method and is deferred to Section 2.

Our next result concerns the asymptotic behavior of the left-tail of the r.v.  $W$ . For the Galton-Watson process this problem is well studied, see e.g. [14] and the references therein. For a BPRES, some interesting results have been obtained in [15] and [17]. In particular, for the annealed law, Huang and Liu ([17], Theorem 1.4) have found a necessary and sufficient condition for the existence of harmonic moments of  $W$ , under the following hypothesis:

$$(H) \quad \exists \delta > 0 \text{ and } A > A_1 > 1 \text{ such that } A_1 \leq m_0 \text{ and } \sum_{i=1}^{\infty} i^{1+\delta} p_i(\xi_0) \leq A^{1+\delta} \text{ a.s.}$$

However, this hypothesis is very restrictive; it implies in particular that  $1 < A_1 \leq m_0 \leq A$ . We will show (see Theorem 1.2 below) the existence of harmonic moments under the following significantly less restrictive assumption:

**A3.** The r.v.  $X = \log m_0$  has an exponential moment, i.e. there exists a constant  $\lambda_0 > 0$  such that

$$(1.10) \quad \mathbb{E}e^{\lambda_0 X} = \mathbb{E}m_0^{\lambda_0} < \infty.$$

Under this hypothesis, since  $X$  is a positive random variable, the function  $\lambda \mapsto \mathbb{E}e^{\lambda X}$  is finite for all  $\lambda \in (-\infty, \lambda_0]$  and is increasing.

**Theorem 1.2.** *Assume condition A3. Let*

$$(1.11) \quad a_0 = \begin{cases} \frac{\lambda_0}{1 - \log \mathbb{E}m_0^{\lambda_0} / \log \mathbb{E}p_1} & \text{if } \mathbb{P}(p_1 > 0) > 0, \\ \lambda_0 & \text{otherwise.} \end{cases}$$

Then, for all  $a \in (0, a_0)$ ,

$$\mathbb{E}W^{-a} < \infty.$$

Yet, a necessary and sufficient condition for the existence of harmonic moments of order  $a > 0$  under condition **A3** is still an open question.

The previous theorem allows us to obtain a Cramér type large deviation expansion for a BPRE. To state the corresponding result we need more notations. Let  $L$  and  $\psi$  be respectively the moment and cumulant generating function of the random variable  $X$ :

$$(1.12) \quad L(\lambda) = \mathbb{E}e^{\lambda X} = \mathbb{E}(m_0^\lambda),$$

$$(1.13) \quad \psi(\lambda) = \log L(\lambda).$$

Then  $\psi$  is analytical for  $\lambda \leq \lambda_0$  and we have  $\psi(\lambda) = \sum_{k=1}^{\infty} \frac{\gamma_k}{k!} \lambda^k$ , where  $\gamma_k = \frac{d^k \psi}{d\lambda^k}(0)$  is the cumulant of order  $k$  of the random variable  $X$ . In particular for  $k = 1, 2$ , we have  $\gamma_1 = \mu$  and  $\gamma_2 = \sigma^2$ . We shall use the Cramér's series of the associated random walk  $(S_n)_{n \geq 0}$  defined by

$$(1.14) \quad \mathcal{L}(t) = \frac{\gamma_3}{6\gamma_2^{3/2}} + \frac{\gamma_4\gamma_2 - 3\gamma_3^2}{24\gamma_2^3}t + \frac{\gamma_5\gamma_2^2 - 10\gamma_4\gamma_3\gamma_2 + 15\gamma_3^3}{120\gamma_2^{9/2}}t^2 + \dots$$

(see Petrov [22]) which converges for  $|t|$  small enough.

Consider the following assumption:

**A4.** There exists a constant  $p > 1$  such that

$$(1.15) \quad \mathbb{E} \frac{Z_1^p}{m_0} < \infty.$$

Note that under (1.7) condition **A4** implies **A2**. The intuitive meaning of these conditions is that the process  $(Z_n)$  cannot deviate too much from its mean  $\Pi_n$ .

The following theorem gives a Cramér's type large deviation expansion of a BPRE.

**Theorem 1.3.** *Assume conditions A3 and A4. Then, for  $0 \leq x = o(\sqrt{n})$ , we have, as  $n \rightarrow \infty$ ,*

$$(1.16) \quad \frac{\mathbb{P}\left(\frac{\log Z_n - n\mu}{\sigma\sqrt{n}} > x\right)}{1 - \Phi(x)} = \exp\left\{\frac{x^3}{\sqrt{n}} \mathcal{L}\left(\frac{x}{\sqrt{n}}\right)\right\} \left[1 + O\left(\frac{1+x}{\sqrt{n}}\right)\right]$$

and

$$(1.17) \quad \frac{\mathbb{P}\left(\frac{\log Z_n - n\mu}{\sigma\sqrt{n}} < -x\right)}{\Phi(-x)} = \exp\left\{-\frac{x^3}{\sqrt{n}} \mathcal{L}\left(-\frac{x}{\sqrt{n}}\right)\right\} \left[1 + O\left(\frac{1+x}{\sqrt{n}}\right)\right].$$

As a consequence of this result we obtain a large deviation approximation by the normal law in the normal zone  $x = o(n^{1/6})$  :

**Corollary 1.4.** *Under the assumptions of Theorem 1.3, we have for  $0 \leq x = o(n^{1/6})$ , as  $n \rightarrow \infty$ ,*

$$(1.18) \quad \frac{\mathbb{P}\left(\frac{\log Z_n - n\mu}{\sigma\sqrt{n}} > x\right)}{1 - \Phi(x)} = 1 + O\left(\frac{x^3}{\sqrt{n}}\right)$$

and

$$(1.19) \quad \frac{\mathbb{P}\left(\frac{\log Z_n - n\mu}{\sigma\sqrt{n}} < -x\right)}{\Phi(-x)} = 1 + O\left(\frac{x^3}{\sqrt{n}}\right).$$

Note that Theorem 1.3 is more precise than the moderate deviation principle established in [17], and, moreover, is stated under weaker assumptions. Indeed, let  $a_n$  be a sequence of positive numbers satisfying  $\frac{a_n}{n} \rightarrow 0$  and  $\frac{a_n}{\sqrt{n}} \rightarrow \infty$ . Then by Theorem 1.6 of [17], under hypothesis (H), we have, for  $x_n = \frac{xa_n}{\sigma\sqrt{n}}$  with fixed  $x \in \mathbb{R}$ ,

$$(1.20) \quad \log \mathbb{P}\left(\frac{\log Z_n - n\mu}{a_n} > x\right) \sim -\frac{x_n^2}{2}.$$

Using the weaker condition **A3** (instead of condition (H)) Theorem 1.3 implies that

$$(1.21) \quad \mathbb{P}\left(\frac{\log Z_n - n\mu}{\sigma\sqrt{n}} > x_n\right) = (1 - \Phi(x_n)) \exp\left(\frac{x_n^3}{\sqrt{n}} \mathcal{L}\left(\frac{x_n}{\sqrt{n}}\right)\right) \left(1 + O\left(\frac{1+x_n}{\sqrt{n}}\right)\right),$$

which sharpens (1.20) without the log-scaling.

The rest of the paper is organized as follows. In Section 2, we prove Theorem 1.1. In Section 3, we study the existence of harmonic moments of  $W$  and give a proof of Theorem 1.2. Section 4 is devoted to the proof of Theorem 1.3.

## 2. THE BERRY-ESSEN BOUND FOR $\log Z_n$

In this section we establish a Berry-Esseen bound for the normalized branching process

$$\frac{\log Z_n - n\mu}{\sigma\sqrt{n}},$$

based on Stein's method. In Section 2.1, we recall briefly the main idea of Stein's method. Section 2.2 contains some auxiliary results to be used latter in the proofs. In Section 2.3, we give a proof of Theorem 1.1.

**2.1. Stein's method.** Let us recall briefly some facts on the Stein method to be used in the proofs. For more details, the reader can consult the excellent reviews [10, 23] or the more complete book [9]. The main idea is to describe the closeness of the law of a r.v.  $X$  to the standard normal law using *Stein's operator*

$$(2.1) \quad \mathcal{A}f(w) = f'(w) - wf(w),$$

which can be seen as a substitute of the classical Fourier-transform tool. For any  $x \in \mathbb{R}$  let  $f_x$  be a solution of *Stein's equation* :

$$(2.2) \quad \mathbf{1}(w \leq x) - \Phi(x) = f'_x(w) - wf_x(w),$$

for all  $w \in \mathbb{R}$ . The Kolmogorov distance between the law of the random variable  $X$  and the normal law  $\mathcal{N}(0, 1)$  can be expressed in term of *Stein's expectation*  $\mathbb{E}\mathcal{A}f_x(X)$ . Indeed, substituting  $w$  by  $X$  in (2.2), taking expectation and the supremum over  $x \in \mathbb{R}$ , we obtain

$$(2.3) \quad \sup_{x \in \mathbb{R}} |\mathbb{P}(X \leq x) - \Phi(x)| = \sup_{x \in \mathbb{R}} |\mathbb{E}(f_x(X) - Xf_x(X))| = \mathbb{E}\mathcal{A}f_x(X).$$

The key point is that Stein's operator  $\mathcal{A}$  characterizes the standard normal law, as shown by the following Lemma.

**Lemma 2.1** (Characterization of the normal law). *A random variable  $Z$  is of normal law  $\mathcal{N}(0, 1)$  if and only if  $\mathbb{E}\mathcal{A}f(Z) = 0$  for all absolutely continuous function  $f$  such that  $\mathbb{E}|f'(Z)| < \infty$ .*

By Lemma 2.1, it is expected that if the distribution of  $X$  is close to the normal law  $\mathcal{N}(0, 1)$  in the sense of Kolmogorov's distance, then  $\mathbb{E}\mathcal{A}f(X)$  is close to 0 for a large class of functions  $f$  including the solutions  $f_x$  of Stein's equation (2.2). This permits to study the convergence of  $X$  to the normal law by using only the structure of  $X$  and the qualitative properties of  $f_x$ . We will use the following result, where we use the notation  $\|\cdot\|$  for the infinity norm.

**Lemma 2.2.** *For each  $x \in \mathbb{R}$ , Stein's equation (2.2) has a unique bounded solution (see [16], Lemma 1.1) given by*

$$(2.4) \quad \begin{aligned} f_x(w) &= e^{w^2/2} \int_w^\infty e^{-t^2/2} (\Phi(x) - \mathbf{1}(t \leq x)) dt \\ &= \begin{cases} \sqrt{2\pi} e^{w^2/2} \Phi(w) [1 - \Phi(x)] & \text{if } w \leq x, \\ \sqrt{2\pi} e^{w^2/2} \Phi(x) [1 - \Phi(w)] & \text{if } w > x. \end{cases} \end{aligned}$$

Moreover, we have for all real  $x$ ,

$$(2.5) \quad \|f_x\| \leq 1, \quad \|f'_x\| \leq 1,$$

and for all real  $w, s$  and  $t$  (see [16], Lemma 1.3),

$$(2.6) \quad \begin{aligned} |f'_x(w+s) - f'_x(w+t)| &\leq (|t| + |s|)(|w| + 1) + \mathbf{1}(x-t \leq w \leq x-s) \mathbf{1}(s \leq t) \\ &\quad + \mathbf{1}(x-s \leq w \leq x-t) \mathbf{1}(s > t). \end{aligned}$$

The next result gives a bound of order  $n^{-1/2}$  of Stein's expectation of a sum of i.i.d. r.v.'s.

**Lemma 2.3.** *Let  $X_1, \dots, X_n$  be a sequence of i.i.d. r.v.'s with  $\mu = \mathbb{E}X_1 \in \mathbb{R}$ ,  $\sigma^2 = \mathbb{E}[X_1 - \mu]^2 < \infty$  and  $\rho = \mathbb{E}|X_1|^3 < \infty$ . Define  $Y_n = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^n (X_k - \mu)$ . For each  $x \in \mathbb{R}$ , the unique bounded solution  $f_x$  of Stein's equation (2.2) satisfies*

$$(2.7) \quad |\mathbb{E}[f'_x(Y_n) - Y_n f_x(Y_n)]| \leq C\rho/\sqrt{n},$$

where  $C$  is an absolute constant.

Note that from (2.3) and (2.7) one gets the classical Berry-Esseen theorem. The proof of Lemma 2.3 can be found in [16].

**2.2. Auxiliary results.** In the proof of Theorem 1.1 we make use of the following two assertions. The first one is a consequence of the Marcinkiewicz-Zygmund inequality (see [20], Lemma 1.4), which will be used several times.

**Lemma 2.4** ([20], Lemma 1.4). *Let  $(X_i)_{i \geq 1}$  be a sequence of i.i.d. centered r.v.'s. Then we have for  $p \in (1, \infty)$ ,*

$$(2.8) \quad \mathbb{E} \left| \sum_{i=1}^n X_i \right|^p \leq \begin{cases} (B_p)^p \mathbb{E}(|X_i|^p) n, & \text{if } 1 < p \leq 2, \\ (B_p)^p \mathbb{E}(|X_i|^p) n^{p/2}, & \text{if } p > 2, \end{cases}$$

where  $B_p = 2 \min \{k^{1/2} : k \in \mathbb{N}, k \geq p/2\}$  is a constant depending only on  $p$  (so that  $B_p = 2$  if  $1 < p \leq 2$ ).

The second one is a result concerning the exponential rate of convergence of  $W_n$  to  $W$  in  $L^p(\mathbb{P})$  from [18], Theorem 1.5.

**Lemma 2.5.** *Under **A2**, there exist two constants  $C > 0$  and  $\delta \in (0, 1)$  such that*

$$(\mathbb{E}|W_n - W|^p)^{1/p} \leq C\delta^n.$$

The next result concerns the existence of positive moments of the r.v.  $\log W$ .

**Lemma 2.6.** *Assume that  $\mathbb{E}|\log m_0|^{2p} < \infty$ , for some  $p > 1$ . Then we have, for all  $q \in (0, p)$ ,*

$$\mathbb{E}|\log W|^q < \infty \quad \text{and} \quad \sup_{n \in \mathbb{N}} \mathbb{E}|\log W_n|^q < \infty.$$

We prove Lemma 2.6 by studying the asymptotic behavior of the Laplace transform of  $W$ . Define the quenched and annealed Laplace transform of  $W$  by

$$\phi_\xi(t) = \mathbb{E}_\xi e^{-tW} \quad \text{and} \quad \phi(t) = \mathbb{E}\phi_\xi(t) = \mathbb{E}e^{-tW},$$

where  $t \geq 0$ . Then by Markov's inequality, we have for  $t > 0$ ,

$$(2.9) \quad \mathbb{P}(W < t^{-1}) \leq e \mathbb{E}e^{-tW} = e \phi(t).$$

*Proof of Lemma 2.6.* By Hölder's inequality, it is enough to prove the assertion of the lemma for  $q \in (1, p)$ . It is obvious that there exists a constant  $C > 0$  such that  $\mathbb{E}|\log W|^q \mathbb{1}(W \geq 1) \leq C\mathbb{E}W < \infty$ . So it remains to show that  $\mathbb{E}|\log W|^q \mathbb{1}(W \leq 1) < \infty$ . By (2.9) and the fact that

$$(2.10) \quad \mathbb{E}|\log W|^q \mathbb{1}(W \leq 1) = q \int_1^{+\infty} \frac{1}{t} (\log t)^{q-1} \mathbb{P}(W \leq t^{-1}) dt,$$



it is enough to show that, as  $t \rightarrow \infty$ ,

$$\phi(t) = O(\log t)^{-p}.$$

It is well-known that  $\phi_\xi(t)$  satisfies the functional relation

$$(2.11) \quad \phi_\xi(t) = f_0 \left( \phi_{T\xi} \left( \frac{t}{m_0} \right) \right),$$

where  $f_0$  is the generating function defined by (1.1) and  $T^n$  is the shift operator defined by  $T^n(\xi_0, \xi_1, \dots) = (\xi_n, \xi_{n+1}, \dots)$  for  $n \geq 1$ . Using (2.11) and the fact that  $\phi_{T\xi}^k \left( \frac{t}{m_0} \right) \leq \phi_{T\xi}^2 \left( \frac{t}{m_0} \right)$  for all  $k \geq 2$ , we obtain

$$(2.12) \quad \begin{aligned} \phi_\xi(t) &\leq p_1(\xi_0) \phi_{T\xi} \left( \frac{t}{m_0} \right) + (1 - p_1(\xi_0)) \phi_{T\xi}^2 \left( \frac{t}{m_0} \right) \\ &= \phi_{T\xi} \left( \frac{t}{m_0} \right) \left( p_1(\xi_0) + (1 - p_1(\xi_0)) \phi_{T\xi} \left( \frac{t}{m_0} \right) \right). \end{aligned}$$

By iteration, this leads to

$$(2.13) \quad \phi_\xi(t) \leq \phi_{T^n \xi} \left( \frac{t}{\Pi_n} \right) \prod_{j=0}^{n-1} \left( p_1(\xi_j) + (1 - p_1(\xi_j)) \phi_{T^n \xi} \left( \frac{t}{\Pi_n} \right) \right).$$

Taking expectation and using the fact that  $\phi_{T^n \xi}(t) \leq 1$ , we get

$$\phi(t) \leq \mathbb{E} \left[ \prod_{j=0}^{n-1} \left( p_1(\xi_j) + (1 - p_1(\xi_j)) \phi_{T^n \xi} \left( \frac{t}{\Pi_n} \right) \right) \right].$$

Using a simple truncation and the fact that  $\phi_\xi(\cdot)$  is non-increasing, we have, for all  $A > 1$ ,

$$\begin{aligned} \phi(t) &\leq \mathbb{E} \left[ \prod_{j=0}^{n-1} \left( p_1(\xi_j) + (1 - p_1(\xi_j)) \phi_{T^n \xi} \left( \frac{t}{A^n} \right) \right) \mathbf{1}(\Pi_n \leq A^n) \right] + \mathbb{P}(\Pi_n \geq A^n) \\ &\leq \mathbb{E} \left[ \prod_{j=0}^{n-1} \left( p_1(\xi_j) + (1 - p_1(\xi_j)) \phi_{T^n \xi} \left( \frac{t}{A^n} \right) \right) \right] + \mathbb{P}(\Pi_n \geq A^n). \end{aligned}$$

Since  $T^n \xi$  is independent of  $\sigma(\xi_0, \dots, \xi_{n-1})$ , and the r.v.'s  $p_1(\xi_i)$  ( $i \geq 0$ ) are i.i.d., we have

$$\phi(t) \leq \left[ \mathbb{E} p_1(\xi_0) + (1 - \mathbb{E} p_1(\xi_0)) \phi \left( \frac{t}{A^n} \right) \right]^n + \mathbb{P}(\Pi_n \geq A^n).$$

By the dominated convergence theorem, we have  $\lim_{t \rightarrow \infty} \phi(t) = 0$ . Thus, for any  $\gamma \in (0, 1)$ , there exists a constant  $K > 0$  such that, for all  $t \geq K$ , we have  $\phi(t) \leq \gamma$ . Then for all  $t \geq K A^n$ , we have  $\phi \left( \frac{t}{A^n} \right) \leq \gamma$ . Consequently, for  $t \geq K A^n$ ,

$$(2.14) \quad \phi(t) \leq \alpha^n + \mathbb{P}(\Pi_n \geq A^n),$$

where, by (1.5),

$$(2.15) \quad \alpha = \mathbb{E} p_1(\xi_0) + (1 - \mathbb{E} p_1(\xi_0)) \gamma \in (0, 1).$$

Recall that  $\mu = \mathbb{E}X$  and  $S_n = \log \Pi_n = \sum_{i=1}^n X_i$ . Choose  $A$  such that  $\log A > \mu$  and let  $\delta = \log A - \mu > 0$ . By Markov's inequality and Lemma 2.4, there exists a constant  $C > 0$  such that, for  $n \in \mathbb{N}$ ,

$$\begin{aligned} \mathbb{P}(\Pi_n \geq A^n) &\leq \mathbb{P}(|S_n - n\mu| \geq n\delta) \\ &\leq \frac{\mathbb{E} |\sum_{i=1}^n (X_i - \mu)|^{2p}}{n^{2p}\delta^{2p}} \\ &\leq \frac{C}{n^p}. \end{aligned}$$

Then, by (2.14), we get, for  $n$  large enough and  $t \geq KA^n$ ,

$$(2.16) \quad \phi(t) \leq \frac{C}{n^p}.$$

For  $t \geq K$ , define  $n_0 = n_0(t) = \left\lfloor \frac{\log(t/K)}{\log(A)} \right\rfloor \geq 0$ , where  $[x]$  stands for the integer part of  $x$ , so that

$$\frac{\log(t/K)}{\log(A)} - 1 \leq n_0 \leq \frac{\log(t/K)}{\log(A)} \quad \text{and} \quad t \geq KA^{n_0}.$$

Coming back to (2.16), with  $n = n_0$ , we get for  $t \geq K$ ,

$$\phi(t) \leq \frac{C(\log A)^p}{(\log(t/K))^p} \leq C(\log t)^{-p},$$

which proves that  $\mathbb{E}|\log W|^q < \infty$  for all  $q \in (1, p)$ , (see (2.10)). Furthermore, since  $x \mapsto |\log^q(x)|\mathbf{1}(x \leq 1)$  is a non-negative and convex function for  $q \in (1, p)$ , by Lemma 2.1 of [17] we have

$$\sup_{n \in \mathbb{N}} \mathbb{E} |\log W_n|^q \mathbf{1}(W_n \leq 1) = \mathbb{E} |\log W|^q \mathbf{1}(W \leq 1).$$

By a standard truncation we obtain

$$(2.17) \quad \sup_{n \in \mathbb{N}} \mathbb{E} |\log W_n|^q \leq C\mathbb{E}W + \mathbb{E} |\log W|^q \mathbf{1}(W \leq 1) < \infty,$$

which ends the proof of the lemma.  $\square$

The next result concerns the exponential speed of convergence of  $\log W_n$  to  $\log W$ .

**Lemma 2.7.** *Assume **A2** and there exists a constant  $q > 2$  such that  $\mathbb{E}|\log m_0|^q < \infty$ . Then there exist two constants  $C > 0$  and  $\delta \in (0, 1)$  such that for all  $n \geq 0$ ,*

$$(2.18) \quad \mathbb{E} |\log W_n - \log W| \leq C\delta^n.$$

*Proof.* From (1.2) and (1.3) we get the following useful decomposition:

$$(2.19) \quad W_{n+1} - W_n = \frac{1}{\Pi_n} \sum_{i=1}^{Z_n} \left( \frac{N_{n,i}}{m_n} - 1 \right),$$

which reads also

$$(2.20) \quad \frac{W_{n+1}}{W_n} - 1 = \frac{1}{Z_n} \sum_{i=1}^{Z_n} \left( \frac{N_{n,i}}{m_n} - 1 \right).$$

By (2.20) we have the decomposition

$$(2.21) \quad \log W_{n+1} - \log W_n = \log(1 + \eta_n),$$

with

$$(2.22) \quad \eta_n = \frac{W_{n+1}}{W_n} - 1 = \frac{1}{Z_n} \sum_{i=1}^{Z_n} \left( \frac{N_{n,i}}{m_n} - 1 \right).$$

Under  $\mathbb{P}_\xi$  the r.v.'s  $\frac{N_{n,i}}{m_n} - 1$  ( $i \geq 1$ ) are i.i.d., centered and independent of  $Z_n$ .

Choose  $p \in (1, 2]$  such that **A2** holds. We first show that

$$(2.23) \quad (\mathbb{E}|\eta_n|^p)^{1/p} \leq C\delta^n,$$

for some constants  $C > 0$  and  $\delta \in (0, 1)$ . Applying Lemma 2.4 under  $\mathbb{P}_\xi$  and using the independence between the r.v.'s  $\frac{N_{n,i}}{m_n}$  ( $i \geq 1$ ) and  $Z_n$ , we get

$$\mathbb{E}_\xi |\eta_n|^p \leq 2^p \mathbb{E}_\xi [Z_n^{1-p}] \mathbb{E}_\xi \left| \frac{N_{n,1}}{m_n} - 1 \right|^p.$$

By **A2** and the fact that under the probability  $\mathbb{P}$  the random variable  $\frac{N_{n,1}}{m_n}$  has the same law as  $\frac{Z_1}{m_0}$ , we obtain

$$(2.24) \quad \mathbb{E} |\eta_n|^p \leq 2^p \mathbb{E} \left| \frac{Z_1}{m_0} - 1 \right|^p \mathbb{E} [Z_n^{1-p}].$$

We shall give a bound of the harmonic moment  $\mathbb{E}Z_n^{1-p}$ . By (1.2), using the convexity of the function  $x \mapsto x^{1-p}$  and the independence between the r.v.'s  $Z_n$  and  $N_{n,i}$  ( $i \geq 1$ ), we get

$$\begin{aligned} \mathbb{E} [Z_{n+1}^{1-p}] &= \mathbb{E} \left[ \left( \sum_{i=1}^{Z_n} N_{n,i} \right)^{1-p} \right] \\ &\leq \mathbb{E} \left[ Z_n^{1-p} \frac{1}{Z_n} \left( \sum_{i=1}^{Z_n} N_{n,i}^{1-p} \right) \right] \\ &\leq \mathbb{E} \left[ \mathbb{E} \left( Z_n^{1-p} \frac{1}{Z_n} \left( \sum_{i=1}^{Z_n} N_{n,i}^{1-p} \right) \middle| Z_n \right) \right] \\ &= \mathbb{E} [Z_n^{1-p}] \mathbb{E} [N_{n,1}^{1-p}] \\ &= \mathbb{E} [Z_n^{1-p}] \mathbb{E} [Z_1^{1-p}]. \end{aligned}$$

By induction, we obtain

$$\mathbb{E} [Z_{n+1}^{1-p}] \leq (\mathbb{E}Z_1^{1-p})^{n+1}.$$

By (1.7), we have  $\mathbb{E}Z_1^{1-p} < 1$ . So the above inequality (2.24) gives (2.23) with  $C = 2 \left( \mathbb{E} \left| \frac{Z_1}{m_0} - 1 \right|^p \right)^{1/p} < \infty$  and  $\delta = (\mathbb{E}Z_1^{1-p})^{1/p} < 1$ .

Now we prove (2.18). Let  $K \in (0, 1)$ . Using the decomposition (2.21) and a standard truncation, we have

$$\begin{aligned} \mathbb{E} |\log W_{n+1} - \log W_n| &= \mathbb{E} |\log(1 + \eta_n)| \mathbf{1}(\eta_n \geq -K) + \mathbb{E} |\log(1 + \eta_n)| \mathbf{1}(\eta_n < -K) \\ (2.25) \qquad \qquad \qquad &= A_n + B_n. \end{aligned}$$

We first find a bound for  $A_n$ . It is obvious that there exists a constant  $C > 0$  such that for all  $x > -K$ ,  $|\ln(1+x)| \leq C|x|$ . By (2.23), we get

$$(2.26) \qquad A_n \leq C \mathbb{E} |\eta_n| \leq C (\mathbb{E} |\eta_n|^p)^{1/p} \leq C \delta^n.$$

Now we find a bound for  $B_n$ . Note that by (2.21) and Lemma 2.6, we have, for any  $r \in (0, q/2)$ ,

$$(2.27) \qquad \sup_{n \in \mathbb{N}} \mathbb{E} |\log(1 + \eta_n)|^r < \infty.$$

Let  $r, s > 1$  be such that  $\frac{1}{s} + \frac{1}{r} = 1$  and  $r < q/2$ . By Hölder's inequality, (2.27), Markov's inequality and (2.23), we have

$$\begin{aligned} B_n &\leq (\mathbb{E} |\log(1 + \eta_n)|^r)^{1/r} \mathbb{P}(\eta_n < -K)^{1/s} \\ &\leq C \mathbb{P}(|\eta_n| > K)^{1/s} \\ &\leq C (\mathbb{E} |\eta_n|^p)^{1/s} \\ (2.28) \qquad \qquad \qquad &\leq C \delta^n. \end{aligned}$$

Thus by (2.25), (2.26) and (2.28), there exist two constants  $C > 0$  and  $\delta \in (0, 1)$  such that

$$(2.29) \qquad \mathbb{E} |\log W_{n+1} - \log W_n| \leq C \delta^n.$$

Using the triangular inequality, we have for all  $k \in \mathbb{N}$ ,

$$\begin{aligned} \mathbb{E} |\log W_{n+k} - \log W_n| &\leq C (\delta^n + \dots + \delta^{n+k-1}) \\ &\leq \frac{C}{1 - \delta} \delta^n. \end{aligned}$$

Letting  $k \rightarrow \infty$ , we get

$$\mathbb{E} |\log W - \log W_n| \leq \frac{C}{1 - \delta} \delta^n,$$

which proves Lemma 2.7. □

We now prove a concentration inequality for the joint law of  $(S_n, \log Z_n)$ .

**Lemma 2.8.** *Assume **A1** and **A2**. Then for all  $x \in \mathbb{R}$ , we have*

$$(2.30) \qquad \mathbb{P} \left( \frac{\log Z_n - n\mu}{\sigma\sqrt{n}} \leq x, \frac{S_n - n\mu}{\sigma\sqrt{n}} \geq x \right) \leq \frac{C}{\sqrt{n}}$$

and

$$(2.31) \qquad \mathbb{P} \left( \frac{\log Z_n - n\mu}{\sigma\sqrt{n}} \geq x, \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x \right) \leq \frac{C}{\sqrt{n}}.$$

Before giving the proof of Lemma 2.8, let us give some heuristics of the proof, following Kozlov [19]. By (2.20), we can write

$$(2.32) \quad W_{n+1} = W_n \times \left( Z_n^{-1} \sum_{i=1}^{Z_n} \frac{N_{n,i}}{m_n} \right).$$

Since  $Z_n \rightarrow \infty$ , by the law of large numbers,  $Z_n^{-1} \sum_{i=1}^{Z_n} \frac{N_{n,i}}{m_n}$  is close to 1, and then  $W_{n+1}/W_n$  is also close to 1 when  $n$  is large enough. Therefore we can hope to replace  $\log W_n$  by  $\log W_m$  without losing too much, when  $m = m(n)$  is an increasing subsequence of integers such that  $m/n \rightarrow 0$ . Denote

$$(2.33) \quad Y_{m,n} = \sum_{i=m+1}^n \frac{X_i - \mu}{\sigma\sqrt{n}}, \quad Y_n = Y_{0,n} \quad \text{and} \quad V_m = \frac{\log W_m}{\sigma\sqrt{n}}.$$

Then, the independence between  $Y_{m,n}$  and  $(Y_m, V_m)$  allows us to use a Berry-Esseen approximation on  $Y_{m,n}$  to get the result.

*Proof of Lemma 2.8.* We first prove (2.30). Let  $\alpha_n = \frac{1}{\sqrt{n}}$  and  $m = m(n) = \lfloor n^{1/2} \rfloor$ , where  $\lfloor x \rfloor$  stands for the integer part of  $x$ . Let  $D_m = V_n - V_m$ . By a standard truncation, using Markov's inequality and Lemma 2.7, there exists  $\delta \in (0, 1)$  such that

$$(2.34) \quad \begin{aligned} \mathbb{P}(Y_n + V_n \leq x, Y_n \geq x) &\leq \mathbb{P}(Y_n + V_m \leq x + \alpha_n, Y_n \geq x) + \mathbb{P}(|D_m| > \alpha_n) \\ &\leq \mathbb{P}(Y_n + V_m \leq x + \alpha_n, Y_n \geq x) + \delta^m. \end{aligned}$$

Now we find a bound for the right-hand side of (2.34). Obviously we have the decomposition

$$(2.35) \quad Y_n = Y_m + Y_{m,n}.$$

For  $x \in \mathbb{R}$ , let  $G_{m,n}(x) = \mathbb{P}(Y_{m,n} \leq x)$  and  $G_n(x) = G_{0,n}(x)$ . Denote by  $\nu_m(ds, dt) = \mathbb{P}(Y_m \in ds, V_m \in dt)$  the joint law of  $(Y_m, V_m)$ . By conditioning and using the independence between  $Y_{m,n}$  and  $(Y_m, V_m)$ , we have

$$(2.36) \quad \begin{aligned} &\mathbb{P}(Y_n + V_m \leq x + \alpha_n, Y_n \geq x) \\ &= \mathbb{P}(Y_{m,n} + Y_m + V_m \leq x + \alpha_n, Y_{m,n} + Y_m \geq x) \\ &= \int \mathbb{P}(Y_{m,n} + s + t \leq x + \alpha_n, Y_{m,n} + s \geq x) \nu_m(ds, dt) \\ &= \int \mathbf{1}(t \leq \alpha_n) (G_{m,n}(x - s - t + \alpha_n) - G_{m,n}(x - s)) \nu_m(ds, dt). \end{aligned}$$

For the terms  $G_{m,n}(x - s - t + \alpha_n)$  and  $G_{m,n}(x - s)$  we are going to use the normal approximation using the Berry-Esseen theorem. Since  $(1-x)^{-1/2} = 1 + \frac{x}{2} + o(x)$  ( $x \rightarrow 0$ ), we have  $\frac{\sqrt{n}}{\sqrt{n-m}} = (1 - \frac{m}{n})^{-1/2} = 1 + R_n$ , where  $0 \leq R_n \leq C/\sqrt{n}$  and  $n \geq 2$ . Therefore, we obtain

$$G_{m,n}(x) = \mathbb{P}\left(\sum_{i=m+1}^n \frac{X_i - \mu}{\sigma\sqrt{n}} \leq x\right) = G_{n-m}\left(\frac{x\sqrt{n}}{\sqrt{n-m}}\right) = G_{n-m}(x(1 + R_n)).$$

Furthermore, by the mean value theorem, we have

$$(2.37) \quad |\Phi(x(1 + R_n)) - \Phi(x)| \leq R_n |x\Phi'(x)| \leq \frac{R_n e^{-1/2}}{2\pi} \leq \frac{C}{\sqrt{n}},$$

where we have used the fact that the function  $x \mapsto x\Phi'(x) = xe^{-x^2/2}/\sqrt{2\pi}$  attains its maximum at  $x = \pm 1$ . Therefore, by the Berry-Esseen theorem, we have for all  $x \in \mathbb{R}$ ,

$$(2.38) \quad \begin{aligned} |G_{m,n}(x) - \Phi(x)| &\leq |G_{n-m}(x(1 + R_n)) - \Phi(x(1 + R_n))| + |\Phi(x(1 + R_n)) - \Phi(x)| \\ &\leq \frac{C}{\sqrt{n}}. \end{aligned}$$

From this and (2.36), we get

$$(2.39) \quad \begin{aligned} &\mathbb{P}(Y_n + V_m \leq x + \alpha_n, Y_n \geq x) \\ &\leq \int \mathbf{1}(t \leq \alpha_n) |\Phi(x - s - t + \alpha_n) - \Phi(x - s)| \nu_m(ds, dt) + \frac{C}{\sqrt{n}}. \end{aligned}$$

Using again the mean value theorem and the fact that  $|\Phi'(x)| \leq 1$ , we obtain

$$(2.40) \quad |\Phi(x - s - t + \alpha_n) - \Phi(x - s)| = |-t + \alpha_n| \leq |t| + \frac{1}{\sqrt{n}}.$$

Moreover, by Lemma 2.6 and the definition of  $\nu_m$ , we have

$$(2.41) \quad \int |t| \nu_m(ds, dt) = \frac{\mathbb{E}|\log W_m|}{\sigma\sqrt{n}} \leq \frac{C}{\sqrt{n}}.$$

Hence, from (2.39) and (2.40), we get

$$\mathbb{P}(Y_n + V_m \leq x + \alpha_n, Y_n \geq x) \leq \frac{C}{\sqrt{n}}.$$

Implementing this bound into (2.34) gives (2.30). The inequality (2.31) is obtained in the same way.  $\square$

**2.3. Proof of Theorem 1.1.** In this section we prove a Berry-Esseen bound for  $\log Z_n$  using Stein's method. In order to simplify the notational burden, let

$$Y_n = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n (X_i - \mu), \quad V_n = \frac{\log W_n}{\sigma\sqrt{n}} \quad \text{and} \quad \tilde{Y}_n = \frac{\log Z_n - n\mu}{\sigma\sqrt{n}} = Y_n + V_n.$$

By (2.3), it is enough to find a suitable bound of Stein's expectation

$$(2.42) \quad |\mathbb{E}[f'_x(\tilde{Y}_n) - \tilde{Y}_n f_x(\tilde{Y}_n)]|,$$

where  $x \in \mathbb{R}$  and  $f_x$  is the unique bounded solution of Stein's equation (2.2). For simplicity, in the following we write  $f$  for  $f_x$ . By the triangular inequality, we have

$$(2.43) \quad \begin{aligned} |\mathbb{E}[f'(\tilde{Y}_n) - \tilde{Y}_n f(\tilde{Y}_n)]| &\leq |\mathbb{E}[f'(\tilde{Y}_n) - Y_n f(Y_n)]| \\ &+ |\mathbb{E}[Y_n f(Y_n) - Y_n f(\tilde{Y}_n)]| + |\mathbb{E}[V_n f(\tilde{Y}_n)]|. \end{aligned}$$

By **A1** and Lemma 2.6, we have  $\sup_n \mathbb{E} |\log W_n|^{3/2} < \infty$ . Therefore, by the definition of  $V_n$ , we have

$$(2.44) \quad |\mathbb{E}[V_n f(\tilde{Y}_n)]| \leq \frac{\|f\|}{\sqrt{n}} \sup_n \mathbb{E} |\log W_n| \leq \frac{C}{\sqrt{n}}.$$

Moreover, using the fact that  $f$  is a Lipschitz function with  $\|f'\| \leq 1$ , together with Hölder's inequality and Lemma 2.4, we get

$$(2.45) \quad \begin{aligned} |\mathbb{E}[Y_n f(Y_n) - Y_n f(\tilde{Y}_n)]| &\leq \mathbb{E}[|Y_n| |f(\tilde{Y}_n) - f(Y_n)|] \\ &\leq \|f'\| \mathbb{E}[|Y_n| |V_n|] \\ &\leq \frac{1}{n} \left[ \mathbb{E} \left| \sum_{i=1}^n (X_i - \mu) \right|^3 \right]^{1/3} \left[ \mathbb{E} |\log W_n|^{3/2} \right]^{2/3} \\ &\leq \frac{C}{n} \left( B_3^3 \mathbb{E} |X_1 - \mu|^3 n^{3/2} \right)^{1/3} \\ &\leq \frac{C}{\sqrt{n}}. \end{aligned}$$

Again, by the triangular inequality, we have

$$(2.46) \quad \begin{aligned} |\mathbb{E}[f'(\tilde{Y}_n) - Y_n f(Y_n)]| &\leq |\mathbb{E}[f'(Y_n + V_n) - f'(Y_n)]| \\ &\quad + |\mathbb{E}[f'(Y_n) - Y_n f(Y_n)]|. \end{aligned}$$

Applying (2.6) for  $w = Y_n$ ,  $s = V_n$  and  $t = 0$ , we get

$$\begin{aligned} |\mathbb{E}[f'(Y_n + V_n) - f'(Y_n)]| &\leq \mathbb{E}(|Y_n| |V_n|) + \mathbb{E}|V_n| + \mathbb{P}(Y_n + V_n \leq x, Y_n \geq x) \\ &\quad + \mathbb{P}(Y_n + V_n \geq x, Y_n \leq x). \end{aligned}$$

As for (2.44) and (2.45), we have  $\mathbb{E}|V_n| \leq \frac{C}{\sqrt{n}}$  and  $\mathbb{E}(|Y_n| |V_n|) \leq \frac{C}{\sqrt{n}}$ . From these bounds and the concentration inequalities of Lemma 2.8, we have

$$(2.47) \quad |\mathbb{E}[f'(Y_n + V_n) - f'(Y_n)]| \leq \frac{C}{\sqrt{n}}.$$

Furthermore, since  $Y_n$  is a sum of i.i.d. random variables, by Lemma 2.3, it follows that

$$(2.48) \quad |\mathbb{E}[f'(Y_n) - Y_n f(Y_n)]| \leq \frac{C}{\sqrt{n}}.$$

Thus, coming back to (2.43) and using the bounds (2.44), (2.45), (2.46), (2.47) and (2.48), we get

$$|\mathbb{E}[f'(\tilde{Y}_n) - \tilde{Y}_n f(\tilde{Y}_n)]| \leq \frac{C}{\sqrt{n}},$$

which ends the proof of Theorem 1.1.

3. HARMONIC MOMENTS OF  $W$ 

In this section, we study the existence of harmonic moments of the random variable  $W$ . Section 3.1 is devoted to the proof of Theorem 1.2. For the needs of Cramér's type large deviations, in Section 3.2 we shall prove the existence of the harmonic moments of  $W$  under the changed probability measure, which generalizes the result of Theorem 1.2.

**3.1. Existence of harmonic moments under  $\mathbb{P}$ .** Following the line of Lemma 2.6 we prove Theorem 1.2 by studying the asymptotic behavior of the Laplace transform of  $W$ . Actually Theorem 1.2 is a simple consequence of Theorem 3.1 below. Recall that

$$\phi_\xi(t) = \mathbb{E}_\xi e^{-tW} \quad \text{and} \quad \phi(t) = \mathbb{E}\phi_\xi(t) = \mathbb{E}e^{-tW}, \quad t \geq 0.$$

**Theorem 3.1.** *Assume condition **A3**. Let  $a_0 > 0$  be defined by (1.11). Then for any  $a \in (0, a_0)$ , there exists a constant  $C > 0$  such that for all  $t > 0$ ,*

$$\phi(t) \leq Ct^{-a}.$$

*In particular  $\mathbb{E}W^{-a} < \infty$  for all  $a \in (0, a_0)$ .*

*Proof.* By (2.14), we have for  $A > 1$  and  $t \geq KA^n$ ,

$$(3.1) \quad \phi(t) \leq \alpha^n + \mathbb{P}(\Pi_n \geq A^n),$$

where

$$(3.2) \quad \alpha = \mathbb{E}p_1(\xi) + (1 - \mathbb{E}p_1(\xi))\gamma \in (0, 1).$$

Using Markov's inequality and condition **A3**, there exists  $\lambda_0 > 0$  such that

$$\mathbb{P}(\Pi_n \geq A^n) \leq \frac{\mathbb{E}\Pi_n^{\lambda_0}}{A^{n\lambda_0}} = \left( \frac{\mathbb{E}m_0^{\lambda_0}}{A^{\lambda_0}} \right)^n.$$

Setting  $A = \left( \frac{\mathbb{E}m_0^{\lambda_0}}{\alpha} \right)^{1/\lambda_0} > 1$ , we get for any  $n \in \mathbb{N}$  and  $t \geq KA^n$ ,

$$(3.3) \quad \phi(t) \leq 2\alpha^n.$$

Now, for any  $t \geq K$ , define  $n_0 = n_0(t) = \left\lfloor \frac{\log(t/K)}{\log A} \right\rfloor \geq 0$ , where  $[x]$  stands for the integer part of  $x$ , so that

$$\frac{\log(t/K)}{\log A} - 1 \leq n_0 \leq \frac{\log(t/K)}{\log A} \quad \text{and} \quad t \geq KA^{n_0}.$$

Then, for  $t \geq K$ ,

$$\phi(t) \leq 2\alpha^{n_0} \leq 2\alpha^{-1}(t/K)^{\frac{\log \alpha}{\log A}} = C_0 t^{-a},$$

with  $C_0 = 2\alpha^{-1}K^a$  and  $a = -\frac{\log \alpha}{\log A} > 0$ . Thus we can choose a constant  $C > 0$  large enough such that for all  $t > 0$ ,

$$(3.4) \quad \phi(t) \leq Ct^{-a}.$$



This proves the first inequality of Theorem 3.1. The existence of harmonic moments of  $W$  of order  $s \in (0, a)$  is deduced from (3.4) and the fact that

$$\mathbb{E}W^{-s} = \frac{1}{\Gamma(s)} \int_0^{+\infty} \phi(t)t^{s-1}dt,$$

where  $\Gamma$  is the Gamma function.

Now we prove (ii). By the definition of  $a$ ,  $A$  and  $\alpha$ , we have

$$\begin{aligned} a &= -\lambda_0 \frac{\log \alpha}{\log \mathbb{E}m_0^{\lambda_0} - \log \alpha} \\ &= -\lambda_0 \frac{\log(\mathbb{E}p_1 + (1 - \mathbb{E}p_1)\gamma)}{\log \mathbb{E}m_0^{\lambda_0} - \log(\mathbb{E}p_1 + (1 - \mathbb{E}p_1)\gamma)}, \end{aligned}$$

where  $\gamma \in (0, 1)$  is an arbitrary constant. Since  $a \rightarrow a_0$  as  $\gamma \rightarrow 0$ , this concludes the proof of Theorem 3.1.  $\square$

**3.2. Existence of harmonic moments under  $\mathbb{P}_\lambda$ .** In this section, we establish a uniform bound for the harmonic moments of  $W$  under the probability measures  $\mathbb{P}_\lambda$ , uniformly in  $\lambda \in [0, \lambda_0]$ .

Let  $m(x) = \mathbb{E}[Z_1 | \xi_0 = x] = \sum_{k=1}^{\infty} kp_k(x)$ . By **A3**, for all  $\lambda \leq \lambda_0$ , we can define the conjugate distribution function  $\tau_{0,\lambda}$  as

$$(3.5) \quad \tau_{0,\lambda}(dx) = \frac{m(x)^\lambda}{L(\lambda)} \tau_0(dx).$$

Note that (3.5) is just Cramér's change of measure for the associated random walk  $(X_n)_{n \geq 1}$ . Consider the new branching process in a random environment whose environment distribution is  $\tau_\lambda = \tau_{0,\lambda}^{\otimes \mathbb{N}}$ . The corresponding annealed probability and expectation are denoted by

$$(3.6) \quad \mathbb{P}_\lambda(dx, d\xi) = \mathbb{P}_\xi(dx) \tau_\lambda(d\xi)$$

and  $\mathbb{E}_\lambda$  respectively. Note that, for any  $\mathcal{F}_n$ -measurable random variable  $T$ , we have

$$(3.7) \quad \mathbb{E}_\lambda T = \frac{\mathbb{E}e^{\lambda S_n} T}{L(\lambda)^n}.$$

It is easily seen that under  $\mathbb{P}_\lambda$ , the process  $(Z_n)$  is still a supercritical branching process in a random environment, which verifies the condition (1.7), and that  $(W_n)_{n \in \mathbb{N}}$  is still a non-negative martingale which converges a.s. to  $W$ . We shall show under the additional assumption **A4** that there exists a constant  $a > 0$  such that for all  $b \in (0, a)$ ,

$$\sup_{0 \leq \lambda \leq \lambda_0} \mathbb{E}_\lambda W^{-b} < \infty.$$

Denote the Laplace transforms of  $W$  under  $\mathbb{P}_\lambda$  by

$$\phi_\lambda(t) = \mathbb{E}_\lambda \phi_\xi(t) = \mathbb{E}_\lambda e^{-tW},$$

where  $t \geq 0$  and  $\lambda \leq \lambda_0$ . The following theorem gives a bound on  $\phi_\lambda(t)$  and  $\mathbb{E}_\lambda W^{-a}$  uniformly in  $\lambda \in [0, \lambda_0]$ .

**Theorem 3.2.** *Assume conditions **A3** and **A4**. Then there exist constants  $a > 0$  and  $C > 0$  such that for all  $t > 0$ ,*

$$\sup_{0 \leq \lambda \leq \lambda_0} \phi_\lambda(t) \leq Ct^{-a}.$$

*In particular, we have  $\sup_{0 \leq \lambda \leq \lambda_0} \mathbb{E}_\lambda W^{-b} < \infty$  for all  $b \in (0, a)$ .*

For the proof of the previous theorem we need to control the exponential speed of convergence in  $L^p$  of  $W_n$  to  $W$ , uniformly under the class of probability measures  $(\mathbb{P}_\lambda)_{0 \leq \lambda \leq \lambda_0}$ .

**Lemma 3.3.** *Assume that **A3** holds for some  $\lambda_0 > 0$ , and **A4** holds for some  $p \in (1, 2]$ . Then for  $\lambda_0 > 0$  small enough, there exist constants  $C > 0$  and  $\delta_0 \in (0, 1)$  such that, for all  $n \geq 1$ ,*

$$\sup_{0 \leq \lambda \leq \lambda_0} (\mathbb{E}_\lambda |W_n - W|^p)^{1/p} \leq C\delta_0^n.$$

*Proof.* Applying Lemma 2.4 under  $\mathbb{E}_\xi$  to the decomposition (2.19) and using the independence between  $Z_n$  and  $\frac{N_{n,i}}{m_n}$  ( $i \geq 1$ ), we get

$$\begin{aligned} \mathbb{E}_\xi |W_{n+1} - W_n|^p &\leq 2^p \Pi_n^{-p} \mathbb{E}_\xi Z_n \mathbb{E}_\xi \left| \frac{N_n}{m_n} - 1 \right|^p \\ &= 2^p \Pi_n^{1-p} \mathbb{E}_\xi \left| \frac{N_n}{m_n} - 1 \right|^p. \end{aligned}$$

Note that under  $\mathbb{P}_\lambda$ , the r.v.'s  $m_0, \dots, m_{n-1}$  are i.i.d., independent of  $\frac{N_n}{m_n}$ , and  $\frac{N_n}{m_n}$  has the same law as  $\frac{Z_1}{m_0}$ . Thus, taking expectation  $\mathbb{E}_\lambda$ , we get

$$(3.8) \quad \mathbb{E}_\lambda |W_{n+1} - W_n|^p \leq 2^p (\mathbb{E}_\lambda m_0^{1-p})^n \mathbb{E}_\lambda \left| \frac{Z_1}{m_0} - 1 \right|^p.$$

Recall that  $m_0 > 1$ . Choose  $\lambda_0 > 0$  small enough such that  $p - \lambda_0 > 1$ . By condition **A4**, for all  $0 \leq \lambda \leq \lambda_0$ ,

$$\mathbb{E}_\lambda \left( \frac{Z_1}{m_0} \right)^p = \frac{1}{\mathbb{E} m_0^\lambda} \mathbb{E} \left( \frac{Z_1^p}{m_0^{p-\lambda}} \right) \leq \mathbb{E} \left( \frac{Z_1^p}{m_0} \right) < +\infty.$$

Since  $1 - p + \lambda_0 < 0$ , we have, for all  $0 \leq \lambda \leq \lambda_0$ ,  $\mathbb{E}_\lambda m_0^{1-p} = \frac{1}{\mathbb{E} m_0^\lambda} \mathbb{E} m_0^{1-p+\lambda} \leq \mathbb{E} m_0^{1-p+\lambda_0} < 1$ . Hence by (3.8), for  $\delta_0 = (\mathbb{E} m_0^{1-p+\lambda_0})^{1/p} < 1$  and  $C = 2 \left( \mathbb{E} \left( \frac{Z_1^p}{m_0} \right)^{1/p} + 1 \right) < \infty$ , we have

$$(3.9) \quad \sup_{0 \leq \lambda \leq \lambda_0} (\mathbb{E}_\lambda |W_{n+1} - W_n|^p)^{1/p} \leq C\delta_0^n.$$

Using the triangular inequality, for all  $k \in \mathbb{N}$ ,

$$\begin{aligned} \sup_{0 \leq \lambda \leq \lambda_0} (\mathbb{E}_\lambda |W_{n+k} - W_n|^p)^{1/p} &\leq C \left( \delta_0^n + \dots + \delta_0^{n+k-1} \right) \\ &\leq \frac{C}{1 - \delta_0} \delta_0^n. \end{aligned}$$

Letting  $k \rightarrow \infty$ , we get

$$(3.10) \quad \sup_{0 \leq \lambda \leq \lambda_0} (\mathbb{E}_\lambda |W - W_n|^p)^{1/p} \leq \frac{C}{1 - \delta_0} \delta_0^n,$$

which concludes the proof of Lemma 3.3.  $\square$

Now we proceed to prove Theorem 3.2.

*Proof of Theorem 3.2.* Let  $\varepsilon \in (0, 1)$ . By a truncation argument, we have for all  $\lambda \in [0, \lambda_0]$ , and  $n \in \mathbb{N}$ ,

$$(3.11) \quad \begin{aligned} \phi_\lambda(t) &= \mathbb{E}_\lambda e^{-tW} [\mathbf{1}(|W_n - W| \leq \varepsilon^n) + \mathbf{1}(|W_n - W| > \varepsilon^n)] \\ &\leq e^{t\varepsilon^n} \mathbb{E}_\lambda e^{-tW_n} + \mathbb{P}_\lambda(|W_n - W| > \varepsilon^n). \end{aligned}$$

Using Markov's inequality and Lemma 3.3, there exists  $\delta_0 \in (0, 1)$  such that

$$(3.12) \quad \sup_{0 \leq \lambda \leq \lambda_0} \mathbb{P}_\lambda(|W_n - W| > \varepsilon^n) \leq C\beta_1^n,$$

where  $\beta_1 = \delta_0/\varepsilon < 1$  for  $\varepsilon > \delta_0$ .

Now we proceed to bound the first term in the right-hand side of (3.11). Recall that  $L(\cdot)$  is increasing. Furthermore, since  $x \mapsto e^{-tx}$  is a non-negative and convex function, we have (see Lemma 2.1 of [17]) that  $\sup_{n \in \mathbb{N}} \mathbb{E} e^{-tW_n} = \mathbb{E} e^{-tW} = \phi(t)$ . Then, again using truncation, we have for all  $\lambda \in [0, \lambda_0]$ ,  $n \in \mathbb{N}$  and  $c > \mu$ ,

$$(3.13) \quad \begin{aligned} \mathbb{E}_\lambda e^{-tW_n} &= \mathbb{E}_\lambda e^{-tW_n} [\mathbf{1}(S_n \leq cn) + \mathbf{1}(S_n > cn)] \\ &\leq e^{\lambda_0 cn} \phi(t) + \mathbb{P}_\lambda(S_n > cn). \end{aligned}$$

By the exponential Markov's inequality, we have for  $\lambda \leq \lambda_0/2$ ,

$$\begin{aligned} \mathbb{P}_\lambda(S_n > cn) &\leq \left( \mathbb{E}_\lambda e^{\lambda X} \right)^n e^{-\lambda cn} \\ &= e^{n(\psi(2\lambda) - \psi(\lambda) - \lambda c)}, \end{aligned}$$

where  $\psi(\lambda) = \log \mathbb{E} e^{\lambda X}$  and  $\psi(2\lambda) - \psi(\lambda) - \lambda c = \lambda\mu - \lambda c + o(\lambda)$  as  $\lambda \rightarrow 0$ . Since  $c > \mu$  we can choose  $\lambda_0 > 0$  small enough, such that for all  $0 \leq \lambda \leq \lambda_0$ ,  $\psi(2\lambda) - \psi(\lambda) - \lambda c \leq \lambda(\mu - c)/2 < 0$ . Thus we have

$$(3.14) \quad \sup_{0 \leq \lambda \leq \lambda_0} \mathbb{P}_\lambda(S_n > cn) \leq \beta_2^n,$$

where  $\beta_2 = e^{\lambda_0(\mu - c)/2} < 1$ . Furthermore by Theorem 3.1, for all  $a \in (0, a_0)$ , there exists  $C > 0$  such that  $\phi(t) \leq Ct^{-a}$  for all  $t > 0$ . Thus implementing (3.12), (3.13) and (3.14) into (3.11) leads to

$$(3.15) \quad \sup_{0 \leq \lambda \leq \lambda_0} \phi_\lambda(t) \leq e^{t\varepsilon^n} \left( e^{\lambda_0 cn} Ct^{-a} + \beta_2^n \right) + C\beta_1^n.$$

Since  $\phi_\lambda(t)$  is decreasing in  $t$ , we have for any  $t \geq t_n = \varepsilon^{-n}$ ,

$$(3.16) \quad \sup_{0 \leq \lambda \leq \lambda_0} \phi_\lambda(t) \leq \sup_{0 \leq \lambda \leq \lambda_0} \phi_\lambda(t_n) \leq e \left( e^{\lambda_0 cn} C\varepsilon^{an} + \beta_2^n \right) + C\beta_1^n.$$

Choosing  $\lambda_0 > 0$  small enough such that  $\beta_3 = e^{\lambda_0 c} \varepsilon^a < 1$ , we find that there exists a constant  $C > 0$  and  $\beta = \max \{\beta_1, \beta_2, \beta_3\} \in (0, 1)$  such that, for any  $t \geq \varepsilon^{-n}$ ,

$$(3.17) \quad \sup_{0 \leq \lambda \leq \lambda_0} \phi_\lambda(t) \leq C\beta^n.$$

The rest of the proof is similar to that of Theorem 1.2, starting from (3.3).  $\square$

#### 4. PROOF OF CRAMÉR'S LARGE DEVIATION EXPANSION

In this section, we prove Theorem 1.3. The starting point is the decomposition (1.4). We will show that the Cramér-type large deviation expansion of  $\log Z_n$  is determined by that of the associated random walk  $(S_n)$ . Our proof is based on Cramér's change of measure  $\mathbb{P}_\lambda$  defined by (3.6). An important step in the approach is to have a good control of the joint law of the couple  $(S_n, \log Z_n)$  under the changed measure  $\mathbb{P}_\lambda$  uniformly in  $\lambda \in [0, \lambda_0]$ , for some small  $\lambda_0$ , which is done in Section 4.1. The proof of Theorem 1.3 is deferred to Section 4.2.

In the sequel we shall use the first three moments of the r.v.  $X = \log m_0$  under the changed probability measure  $\mathbb{P}_\lambda$ :

$$(4.1) \quad \mu_\lambda = \mathbb{E}_\lambda X = \psi'(\lambda) = \sum_{k=1}^{\infty} \frac{\gamma_k}{(k-1)!} \lambda^{k-1},$$

$$(4.2) \quad \sigma_\lambda = \mathbb{E}_\lambda (X - \mu_\lambda)^2 = \psi''(\lambda) = \sum_{k=2}^{\infty} \frac{\gamma_k}{(k-2)!} \lambda^{k-2},$$

$$(4.3) \quad \rho_\lambda = \mathbb{E}_\lambda |X - \mu_\lambda|^3,$$

with  $\psi$  defined in (1.13).

**4.1. Auxiliary results.** In this section we prove a uniform concentration inequality bound for the class of probability measures  $(\mathbb{P}_\lambda)_{0 \leq \lambda \leq \lambda_0}$ . First we give uniform bounds for the first three moments of  $X$  under  $\mathbb{P}_\lambda$ . It is well known that, for  $\lambda_0$  small enough and for any  $\lambda \in [0, \lambda_0]$ ,

$$(4.4) \quad |\mu_\lambda - \mu| \leq C_1 \lambda, \quad |\sigma_\lambda - \sigma| \leq C_2 \lambda, \quad |\rho_\lambda - \rho| \leq C_3 \lambda,$$

where  $C_1, C_2, C_3$  are absolute constants. These bounds allow us to obtain an uniform rate of convergence for the process  $(\log W_n)$  under  $\mathbb{P}_\lambda$ .

**Lemma 4.1.** *Assume **A3** and **A4**. Then there exists  $\delta_0 \in (0, 1)$  such that*

$$(4.5) \quad \sup_{0 \leq \lambda \leq \lambda_0} \mathbb{E}_\lambda |\log W_n - \log W| \leq \delta_0^n.$$

*Proof.* The proof is similar to that in Lemma 2.7: it is enough to replace  $\mathbb{E}$  by  $\mathbb{E}_\lambda$  and to ensure that all the bounds in that proof still hold uniformly in  $\lambda \in [0, \lambda_0]$ , for  $\lambda_0 > 0$  small enough.

We first prove that for some constants  $\lambda_0 > 0$ ,  $\delta \in (0, 1)$  and  $C > 0$ ,

$$(4.6) \quad \sup_{0 \leq \lambda \leq \lambda_0} (\mathbb{E}_\lambda |\eta_n|^p)^{1/p} \leq C\delta^n,$$

where  $\eta_n$  is defined (2.20). In fact, we have, for  $p \in (1, 2)$ ,

$$\mathbb{E}_\lambda |\eta_n|^p \leq 2^p \mathbb{E}_\lambda \left| \frac{Z_1}{m_0} - 1 \right|^p \left( \mathbb{E}_\lambda [Z_1^{1-p}] \right)^n.$$

By the dominated convergence theorem and the fact that  $m_0 > 1$ , we have  $\mathbb{E}_\lambda Z_1^{1-p} \leq \mathbb{E} Z_1^{1-p} m_0^\lambda \xrightarrow{\lambda \rightarrow 0} \mathbb{E} Z_1^{1-p} < 1$ . Thus there exists a  $\lambda_0 > 0$  small enough such that

$$\mathbb{E}_\lambda Z_1^{1-p} \leq \mathbb{E} Z_1^{1-p} m_0^{\lambda_0} < 1.$$

By **A3** and **A4**, for some small enough  $\lambda_0 \in (0, p-1]$  and all  $\lambda \in [0, \lambda_0]$  we have,

$$\mathbb{E}_\lambda \left( \frac{Z_1}{m_0} \right)^p = \left( \mathbb{E} m_0^\lambda \right)^{-1} \mathbb{E} \frac{Z_1^p}{m_0^{p-\lambda_0}} \leq \mathbb{E} \frac{Z_1^p}{m_0^{p-\lambda_0}} \leq \mathbb{E} \frac{Z_1^p}{m_0} < \infty.$$

Therefore, (4.6) holds with  $C \leq 2 \left[ \left( \mathbb{E} \frac{Z_1^p}{m_0} \right)^{1/p} + 1 \right] < \infty$  and  $\delta \leq \left( \mathbb{E} Z_1^{1-p} m_0^{\lambda_0} \right)^{1/p} < 1$ .

Next we show that

$$(4.7) \quad \sup_{n \in \mathbb{N}} \sup_{0 \leq \lambda \leq \lambda_0} \mathbb{E}_\lambda |\log(1 + \eta_n)|^r < \infty,$$

for all  $r > 0$ . It is easily seen that there exists a constant  $C_r > 0$  such that  $\mathbb{E}_\lambda |\log W|^r \leq C_r (\mathbb{E}_\lambda W^{-\alpha} + \mathbb{E}_\lambda W) \leq C_r (\mathbb{E}_\lambda W^{-\alpha} + 1)$ . Then, by **A3** and Theorem 3.2, for all  $r > 0$ , we have

$$(4.8) \quad \sup_{0 \leq \lambda \leq \lambda_0} \mathbb{E}_\lambda |\log W|^r < \infty.$$

Thus by (2.17) and (2.21) we get (4.7).

We finally end the proof in the same way as in Lemma 2.7, using the uniform bounds (4.7) and (4.6).  $\square$

Now we give a control of the joint law of  $(S_n, \log Z_n)$  for the convergence to the distribution function  $\Phi([0, x]) \mathbf{1}(x \geq 0)$ ,  $x \in \mathbb{R}$ , uniformly in  $\lambda \in [0, \lambda_0]$ , where  $\Phi([0, x]) = \Phi(x) - \Phi(0)$  (recall that  $\Phi$  is the distribution function of the standard normal law).

**Lemma 4.2.** *Assume **A3** and **A4**. There exist positive constants  $C$ ,  $\beta_1$ ,  $\beta_2$  and  $\delta \in (0, 1)$  such that for any  $x > 0$ ,*

$$(4.9) \quad \sup_{0 \leq \lambda \leq \lambda_0} \left| \mathbb{P}_\lambda \left( \frac{S_n - n\mu}{\sigma_\lambda \sqrt{n}} \leq x, \frac{\log Z_n - n\mu}{\sigma_\lambda \sqrt{n}} \geq 0 \right) - \Phi([0, x]) \right| \leq \frac{C}{\sqrt{n}},$$

and

$$(4.10) \quad \begin{aligned} & \sup_{0 \leq \lambda \leq \lambda_0} \mathbb{P}_\lambda \left( \frac{S_n - n\mu}{\sigma_\lambda \sqrt{n}} \leq -x, \frac{\log Z_n - n\mu}{\sigma_\lambda \sqrt{n}} \geq 0 \right) \\ & \leq C \left( x + \frac{1}{\sqrt{n}} \right) e^{-\beta_1 x \sqrt{n}} + \min \left( e^{-\beta_2 x \sqrt{n}}, \delta \sqrt{n} x^{-1/2} n^{-1/4} \right). \end{aligned}$$

*Proof.* Let  $m = m(n) = \lfloor n^{1/2} \rfloor$ , with  $\lfloor x \rfloor$  denoting the integer part of  $x$ , and

$$Y_{m,n}^\lambda = \sum_{i=m+1}^n \frac{X_i - \mu_\lambda}{\sigma_\lambda \sqrt{n}}, \quad Y_n^\lambda = Y_{0,n}^\lambda \quad \text{and} \quad V_m^\lambda = \frac{\log W_m}{\sigma_\lambda \sqrt{n}}.$$

The proof of (4.9) is similar to that of Lemma 2.8 with  $\mathbb{P}$  replaced by  $\mathbb{P}_\lambda$ . The only difference is that the bounds (2.38) and (2.41) have to be uniform in  $\lambda \in [0, \lambda_0]$ . The uniformity in (2.38) is ensured by the Berry-Esseen theorem and (4.4) which imply that

$$(4.11) \quad \sup_{\lambda \in [0, \lambda_0]} \left| G_{m,n}^\lambda(x) - \Phi(x) \right| \leq \frac{C}{\sqrt{n}},$$

where  $G_{m,n}^\lambda(x) = \mathbb{P}_\lambda(Y_{m,n}^\lambda \leq x)$ . The uniformity in (2.41) is a consequence of Lemma 4.1. Further details of the proof are left to the reader.

Now we prove (4.10). Let  $D_m^\lambda = V_n^\lambda - V_m^\lambda$ . By considering the events  $\{|D_m^\lambda| \leq \frac{x}{2}\}$  and  $\{|D_m^\lambda| > \frac{x}{2}\}$  we have

$$(4.12) \quad \begin{aligned} \mathbb{P}_\lambda \left( Y_n^\lambda \leq -x, Y_n^\lambda + V_n^\lambda \geq 0 \right) &\leq \mathbb{P}_\lambda \left( Y_n^\lambda \leq -x, Y_n^\lambda + V_m^\lambda \geq -\frac{x}{2} \right) \\ &+ \mathbb{P}_\lambda \left( |D_m^\lambda| > \frac{x}{2} \right). \end{aligned}$$

We first find a suitable bound of the first term of the right-hand side of (4.12). Again by decomposing  $Y_n^\lambda = Y_{m,n}^\lambda + Y_m^\lambda$ , using (4.11) and the fact that  $\Phi([a, b]) \leq b - a$ , we have

$$\begin{aligned} &\mathbb{P}_\lambda \left( Y_n^\lambda \leq -x, Y_n^\lambda + V_m^\lambda \geq -\frac{x}{2} \right) \\ &= \int \mathbf{1} \left( t > \frac{x}{2} \right) \mathbb{P}_\lambda \left( Y_{m,n}^\lambda \in \left[ -\frac{x}{2} - s - t, -x - s \right] \right) \nu_m^\lambda(ds, dt) \\ &\leq \int \mathbf{1} \left( t > \frac{x}{2} \right) \left[ \Phi \left( \left[ -\frac{x}{2} - s - t, -x - s \right] \right) + \frac{C}{\sqrt{n}} \right] \nu_m^\lambda(ds, dt) \\ &\leq \int \mathbf{1} \left( t > \frac{x}{2} \right) \left[ \left( t - \frac{x}{2} \right) + \frac{C}{\sqrt{n}} \right] \nu_m^\lambda(ds, dt) \\ &\leq \mathbb{E}_\lambda \left[ V_m^\lambda \mathbf{1} \left( V_m^\lambda \geq \frac{x}{2} \right) \right] + \left[ \frac{x}{2} + \frac{C}{\sqrt{n}} \right] \mathbb{P}_\lambda \left( V_m^\lambda > \frac{x}{2} \right). \end{aligned}$$

By Markov's inequality, we have  $\mathbb{P}_\lambda \left( V_m^\lambda > \frac{x}{2} \right) \leq e^{-\frac{x}{2} \sigma_\lambda \sqrt{n}}$ . Moreover, using Hölder's and Markov's inequalities, we get by (4.8) and the definition of  $V_m$  that

$$\mathbb{E}_\lambda \left[ V_m^\lambda \mathbf{1} \left( V_m^\lambda \geq \frac{x}{2} \right) \right] \leq \left( \mathbb{E}_\lambda |V_m^\lambda|^2 \right)^{1/2} \mathbb{P}_\lambda \left( V_m^\lambda \geq \frac{x}{2} \right)^{1/2} \leq \frac{C}{\sigma_\lambda \sqrt{n}} e^{-\frac{x}{4} \sigma_\lambda \sqrt{n}}.$$

Since, by (4.4),  $\sigma_\lambda$  is bounded uniformly in  $\lambda \in [0, \lambda_0]$ , there exists  $\beta_1 > 0$  such that for any  $\lambda \in [0, \lambda_0]$ ,

$$(4.13) \quad \mathbb{P}_\lambda \left( Y_n^\lambda \leq -x, Y_n^\lambda + V_m^\lambda \geq -\frac{x}{2} \right) \leq C \left( x + \frac{1}{\sqrt{n}} \right) e^{-\beta_1 x \sqrt{n}}.$$

We now search for a suitable bound for the second term of the right-hand side of (4.12). By Hölder's inequality and Theorem 3.2, there exist some constants  $C > 0$ ,  $a > 0$  and  $0 < \alpha < \min(1/2, a/2)$  such that, for all  $\lambda \in [0, \lambda_0]$  and  $n \in \mathbb{N}$ ,

$$\mathbb{E}_\lambda \left( \frac{W_n}{W_m} \right)^\alpha \leq \left( \mathbb{E}_\lambda W_n^{2\alpha} \right)^{1/2} \left( \mathbb{E}_\lambda W_m^{-2\alpha} \right)^{1/2} \leq \left( \mathbb{E}_\lambda W^{2\alpha} \right)^{1/2} \left( \mathbb{E}_\lambda W^{-2\alpha} \right)^{1/2} \leq C.$$

Thus, by Markov's inequality and (4.4), there exists a constant  $\beta_2 > 0$  (independent of  $(\lambda, n, x)$ ) such that, for all  $\lambda \in [0, \lambda_0]$ ,

$$(4.14) \quad \begin{aligned} \mathbb{P}_\lambda \left( |D_m^\lambda| > \frac{x}{2} \right) &\leq \mathbb{P}_\lambda \left( \left( \frac{W_n}{W_m} \right)^\alpha > e^{\alpha \sigma_\lambda \sqrt{n} \frac{x}{2}} \right) + \mathbb{P}_\lambda \left( \left( \frac{W_m}{W_n} \right)^\alpha > e^{\alpha \sigma_\lambda \sqrt{n} \frac{x}{2}} \right) \\ &\leq C e^{-\beta_2 x \sqrt{n}}. \end{aligned}$$

Moreover, by Markov and Jensen's inequalities and Lemma 4.1, there exists  $\delta_0 \in (0, 1)$  such that for  $\lambda \in [0, \lambda_0]$ ,

$$(4.15) \quad \begin{aligned} \mathbb{P}_\lambda \left( |D_m^\lambda| > \frac{x}{2} \right) &\leq \mathbb{P}_\lambda \left( |\log W_n - \log W_m|^{1/2} > \frac{x^{1/2} n^{1/4}}{\sqrt{2}} \right) \\ &\leq C \delta_0^{m/2} x^{-1/2} n^{-1/4}. \end{aligned}$$

From (4.14) and (4.15) we have, for any  $\lambda \in [0, \lambda_0]$ ,

$$(4.16) \quad \mathbb{P}_\lambda \left( |D_m^\lambda| > \frac{x}{2} \right) \leq C \min \left( e^{-\beta_2 x \sqrt{n}}, \delta_0^{\sqrt{n}/2} x^{-1/2} n^{-1/4} \right).$$

Using (4.12), (4.13) and (4.16), we get (4.10) with  $\delta = \delta_0^{1/2}$ . This ends the proof of the lemma.  $\square$

**4.2. Proof of Theorem 1.3.** We shall prove only the first assertion, the second one being proved in the same way.

For  $0 \leq x \leq 1$ , the theorem follows from the Berry-Esseen estimate in Theorem 1.1. So we assume that  $1 \leq x = o(\sqrt{n})$ . Using the change of measure (3.7), for any  $\lambda \in [0, \lambda_0]$ , we have

$$\mathbb{P} \left( \frac{\log Z_n - n\mu}{\sigma \sqrt{n}} > x \right) = L(\lambda)^n \mathbb{E}_\lambda \left[ e^{-\lambda S_n} \mathbf{1}(\log Z_n - n\mu > x \sigma \sqrt{n}) \right].$$

Denote

$$(4.17) \quad Y_n^\lambda = \frac{S_n - n\mu_\lambda}{\sigma_\lambda \sqrt{n}} \quad \text{and} \quad V_n^\lambda = \frac{\log W_n}{\sigma_\lambda \sqrt{n}}.$$

Using the decomposition (1.4), centering and reducing  $S_n$  under  $\mathbb{P}_\lambda$ , we get

$$\begin{aligned} & \mathbb{P}\left(\frac{\log Z_n - n\mu}{\sigma\sqrt{n}} > x\right) \\ &= \exp(n\psi(\lambda) - n\mu_\lambda) \mathbb{E}_\lambda \left[ e^{-\lambda\sigma_\lambda\sqrt{n}Y_n^\lambda} \mathbf{1}\left(Y_n^\lambda + V_n^\lambda > \frac{x\sigma\sqrt{n} - n(\mu_\lambda - \mu)}{\sigma_\lambda\sqrt{n}}\right) \right], \end{aligned}$$

with  $\psi$  defined in (1.13). It is well known that for  $x = o(\sqrt{n})$  as  $n \rightarrow \infty$ , the equation

$$(4.18) \quad x\sigma\sqrt{n} = n(\mu_\lambda - \mu),$$

has a unique solution  $\lambda(x)$  which can be expressed as the power series

$$(4.19) \quad \lambda(x) = \frac{t}{\sqrt{\gamma_2}} - \frac{\gamma_3}{2\gamma_2^2}t^2 - \frac{\gamma_4\gamma_2 - 3\gamma_3^2}{6\gamma_2^{7/2}}t^3 + \dots$$

with  $t = \frac{x}{\sqrt{n}}$  (see [22] for details). Choosing  $\lambda = \lambda(x)$ , it follows that

$$(4.20) \quad \begin{aligned} \mathbb{P}\left(\frac{\log Z_n - n\mu}{\sigma\sqrt{n}} > x\right) &= \exp(n\psi(\lambda) - n\mu_\lambda) \mathbb{E}_\lambda \left[ e^{-\lambda\sigma_\lambda\sqrt{n}Y_n^\lambda} \mathbf{1}(Y_n^\lambda + V_n^\lambda > 0) \right] \\ &= \exp(n\psi(\lambda) - n\mu_\lambda) I, \end{aligned}$$

where

$$(4.21) \quad \begin{aligned} I &= \mathbb{E}_\lambda \left[ e^{-\lambda\sigma_\lambda\sqrt{n}Y_n^\lambda} \mathbf{1}(Y_n^\lambda + V_n^\lambda > 0) \right] \\ &= \int e^{-\lambda\sigma_\lambda\sqrt{n}Y_n^\lambda} \mathbf{1}(Y_n^\lambda + V_n^\lambda > 0) d\mathbb{P}_\lambda. \end{aligned}$$

Using the fact that

$$e^{-\lambda\sigma_\lambda\sqrt{n}Y_n^\lambda} = \lambda\sigma_\lambda\sqrt{n} \int_{\mathbb{R}} \mathbf{1}(Y_n^\lambda < y) e^{-\lambda\sigma_\lambda\sqrt{n}y} dy$$

and Fubini's theorem, we obtain

$$I = \lambda\sigma_\lambda\sqrt{n} \int_{\mathbb{R}} e^{-\lambda\sigma_\lambda\sqrt{n}y} \mathbb{P}_\lambda(Y_n^\lambda < y, Y_n^\lambda + V_n^\lambda > 0) dy.$$

Obviously  $I = I_+ + I_-$ , with

$$I_+ = \lambda\sigma_\lambda\sqrt{n} \int_0^\infty e^{-\lambda\sigma_\lambda\sqrt{n}y} \mathbb{P}_\lambda(Y_n^\lambda < y, Y_n^\lambda + V_n^\lambda > 0) dy,$$

$$I_- = \lambda\sigma_\lambda\sqrt{n} \int_{-\infty}^0 e^{-\lambda\sigma_\lambda\sqrt{n}y} \mathbb{P}_\lambda(Y_n^\lambda < y, Y_n^\lambda + V_n^\lambda > 0) dy.$$

We shall show that

$$(4.22) \quad I = I_1(1 + O(\lambda)),$$

where

$$(4.23) \quad I_1 = \lambda\sigma_\lambda\sqrt{n} \int_0^\infty e^{-\lambda\sigma_\lambda\sqrt{n}y} \Phi([0, y]) dy.$$



By Lemma 4.2 (i) we get by a straightforward computation that

$$(4.24) \quad |I_+ - I_1| \leq \frac{C}{\sqrt{n}}.$$

By Lemma 4.2 (ii), we have

$$\begin{aligned} I_- \leq C \lambda \sigma_\lambda \sqrt{n} \int_{-\infty}^0 e^{\lambda \sigma_\lambda \sqrt{n} |y|} & \left[ \left( |y| + \frac{1}{\sqrt{n}} \right) e^{-\beta_1 |y| \sqrt{n}} \right. \\ & \left. + \min \left( e^{-\beta_2 |y| \sqrt{n}}, \delta \sqrt{n} |y|^{-1/2} n^{-1/4} \right) \right] dy. \end{aligned}$$

Recall that, by (4.4),  $\sigma_\lambda$  is bounded for  $\lambda$  small enough and, by (4.19), we have  $\lambda \rightarrow 0$  as  $n \rightarrow \infty$ . Then for  $0 < \varepsilon < \min(\beta_1, \beta_2)$ , we have  $\lambda \sigma_\lambda < \varepsilon$  for all  $n$  large enough. Thus, by a straightforward calculation and by choosing  $\varepsilon > 0$  small enough, it can be seen that

$$\begin{aligned} I_- & \leq C \lambda \sqrt{n} \int_{-\infty}^0 \left( |y| + \frac{1}{\sqrt{n}} \right) e^{-(\beta_1 - \varepsilon) |y| \sqrt{n}} dy \\ & \quad + C \lambda \sqrt{n} \int_{-\infty}^{-1} e^{-(\beta_2 - \varepsilon) |y| \sqrt{n}} dy \\ & \quad + C \lambda \sqrt{n} \int_{-1}^0 \delta \sqrt{n} |y|^{-1/2} n^{-1/4} e^{\varepsilon |y| \sqrt{n}} dy \\ & \leq \frac{C \lambda}{\sqrt{n}}. \end{aligned}$$

By (4.19), we get, as  $n \rightarrow \infty$ ,

$$(4.25) \quad I_- = o\left(\frac{1}{\sqrt{n}}\right).$$

From (4.24) and (4.25) it follows that

$$(4.26) \quad |I - I_1| \leq \frac{C}{\sqrt{n}}.$$

The integral  $I_1$  appears in the proof of the Cramér's large deviation expansion theorem for the i.i.d. case. For convenience, we state here some well known results concerning the asymptotic expansion of the cumulant generating function  $\psi(\lambda)$  and of the integral  $I_1$ . For details we refer the reader to [22].

**Lemma 4.3.** *Let  $X$  be a r.v. such that  $\mathbb{E}[e^{\lambda_0 |X|}] < \infty$  for some  $\lambda_0 > 0$ . For  $\lambda \in (-\lambda_0, \lambda_0)$ , let  $\psi(\lambda) = \log \mathbb{E}[e^{\lambda X}]$ ,  $\mu_\lambda = \psi'(\lambda)$  and  $\sigma_\lambda = \psi''(\lambda)$ . Set  $\mu = \mathbb{E}X$ . Then for  $1 \leq x = o(\sqrt{n})$ ,  $\lambda = \lambda(x)$  solution of (4.18) and  $n$  large enough, we have:*

(i) *the cumulant generating function  $\psi(\lambda) = \log \mathbb{E}[e^{\lambda X}]$  satisfies the identity*

$$(4.27) \quad \frac{x^2}{2} + n(\psi(\lambda) - \lambda \mu_\lambda) = \frac{x^3}{\sqrt{n}} \mathcal{L}\left(\frac{x}{\sqrt{n}}\right),$$

*where  $\mathcal{L}(t)$  is the Cramér's series defined by (1.14);*

(ii) the integral  $I_1$  defined by (4.23) satisfies the property that there exist some positive constants  $C_1, C_2 > 0$  such that

$$(4.28) \quad C_1 \leq \lambda \sigma_\lambda \sqrt{n} I_1 \leq C_2 ;$$

moreover, the integral  $I_1$  admits the following asymptotic expansion:

$$(4.29) \quad I_1 = \exp\left(\frac{x^2}{2}\right) [1 - \Phi(x)] \left(1 + O\left(\frac{x}{\sqrt{n}}\right)\right).$$

Now we can end the proof of Theorem 1.3. By (4.26), (4.28) and (4.19), we have

$$I = I_1 (1 + O(\lambda)) = I_1 \left(1 + O\left(\frac{x}{\sqrt{n}}\right)\right).$$

Coming back to (4.20) and using (4.29), we get

$$\mathbb{P}\left(\frac{\log Z_n - n\mu}{\sigma\sqrt{n}} > x\right) = \exp\left(\frac{x^2}{2} + n(\psi(\lambda) - \lambda\mu_\lambda)\right) (1 - \Phi(x)) \left(1 + O\left(\frac{x}{\sqrt{n}}\right)\right).$$

Then, by (4.27), we obtain the desired Cramér's large deviation expansion

$$(4.30) \quad \mathbb{P}\left(\frac{\log Z_n - n\mu}{\sigma\sqrt{n}} > x\right) = \exp\left(\frac{x^3}{\sqrt{n}} \mathcal{L}\left(\frac{x}{\sqrt{n}}\right)\right) (1 - \Phi(x)) \left(1 + O\left(\frac{x}{\sqrt{n}}\right)\right),$$

which ends the proof of the first assertion of Theorem 1.3.

## REFERENCES

- [1] V. I. Afanasyev, C. Böinghoff, G. Kersting, and V. A. Vatutin. Limit theorems for weakly subcritical branching processes in random environment. *J. Theoret. Probab.*, 25(3):703–732, 2012.
- [2] V. I. Afanasyev, C. Böinghoff, G. Kersting, and V. A. Vatutin. Conditional limit theorems for intermediately subcritical branching processes in random environment. *Ann. Inst. Henri Poincaré Probab. Stat.*, 50(2):602–627, 2014.
- [3] K. B. Athreya and S. Karlin. Branching processes with random environments: II: Limit theorems. *Ann. Math. Stat.*, 42(6):1843–1858, 1971.
- [4] K. B. Athreya and S. Karlin. On branching processes with random environments: I: Extinction probabilities. *Ann. Math. Stat.*, 42(5):1499–1520, 1971.
- [5] V. Bansaye and J. Berestycki. Large deviations for branching processes in random environment. *Markov Process. Related Fields*, 15(4):493–524, 2009.
- [6] V. Bansaye and C. Böinghoff. Upper large deviations for branching processes in random environment with heavy tails. *Electron. J. Probab.*, 16(69):1900–1933, 2011.
- [7] V. Bansaye and C. Böinghoff. Lower large deviations for supercritical branching processes in random environment. *Proc. Steklov Inst. Math.*, 282(1):15–34, 2013.
- [8] V. Bansaye and C. Böinghoff. Small positive values for supercritical branching processes in random environment. *Ann. Inst. Henri Poincaré Probab. Stat.*, 50(3):770–805, 2014.
- [9] A. D. Barbour and L. H. Y. Chen. *An introduction to Stein's method*, volume 4. World Scientific, 2005.
- [10] A. D. Barbour and L. H. Y. Chen. Stein's (magic) method. *arXiv:1411.1179 [math]*, 2014.
- [11] C. Böinghoff. Limit theorems for strongly and intermediately supercritical branching processes in random environment with linear fractional offspring distributions. *Stoch. Process. Appl.*, 124(11):3553–3577, 2014.

- [12] C. Böinghoff and G. Kersting. Upper large deviations of branching processes in a random environment - offspring distributions with geometrically bounded tails. *Stoch. Process. Appl.*, 120(10):2064–2077, 2010.
- [13] H. Cramér. Sur un nouveau théorème-limite de la théorie des probabilités. *Actualités scientifiques et industrielles*, 736(5-23):115, 1938.
- [14] K. Fleischmann and V. Wachtel. On the left tail asymptotics for the limit law of supercritical Galton–Watson processes in the Böttcher case. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(1):201–225, 2009.
- [15] B. Hambly. On the limiting distribution of a supercritical branching process in a random environment. *J. Appl. Probab.*, 29(3):499–518, 1992.
- [16] S. T. Ho and L. H. Y. Chen. An  $L_p$  bound for the remainder in a combinatorial central limit theorem. *Ann. Probab.*, 6(2):231–249, 1978.
- [17] C. Huang and Q. Liu. Moments, moderate and large deviations for a branching process in a random environment. *Stoch. Process. Appl.*, 122(2):522–545, 2012.
- [18] C. Huang and Q. Liu. Convergence in  $L_p$  and its exponential rate for a branching process in a random environment. *Electron. J. Probab.*, 19(104):1–22, 2014.
- [19] M. V. Kozlov. On large deviations of branching processes in a random environment: geometric distribution of descendants. *Discrete Math. Appl.*, 16(2):155–174, 2006.
- [20] Q. Liu. Local dimensions of the branching measure on a Galton–Watson tree. *Ann. Inst. Henri Poincaré Probab. Stat.*, 37(2):195–222, 2001.
- [21] M. Nakashima. Lower deviations of branching processes in random environment with geometrical offspring distributions. *Stoch. Process. Appl.*, 123(9):3560–3587, 2013.
- [22] V. V. Petrov. *Sums of independent random variables*. Berlin, 1975.
- [23] N. Ross. Fundamentals of Stein’s method. *Probab. Surv.*, 8:210–293, 2011.
- [24] W. L. Smith and W. E. Wilkinson. On branching processes in random environments. *Ann. Math. Stat.*, 40(3):814–827, 1969.
- [25] D. Tanny. A necessary and sufficient condition for a branching process in a random environment to grow like the product of its means. *Stoch. Process. Appl.*, 28(1):123–139, 1988.
- [26] V. A. Vatutin. A refinement of limit theorems for the critical branching processes in random environment. In *Workshop on Branching Processes and their Applications. Lect. Notes Stat. Proc.*, volume 197, pages 3–19. Springer, Berlin, 2010.
- [27] V. A. Vatutin and X. Zheng. Subcritical branching processes in a random environment without the Cramer condition. *Stoch. Process. Appl.*, 122(7):2594–2609, 2012.

*Current address*, Grama, I.: Université de Bretagne-Sud, LMBA, UMR CNRS 6205, Vannes, France

*E-mail address*: `ion.grama@univ-ubs.fr`

*Current address*, Liu, Q.: Université de Bretagne-Sud, LMBA, UMR CNRS 6205, Vannes, France

*E-mail address*: `quansheng.liu@univ-ubs.fr`

*Current address*, Miqueu, E.: Université de Bretagne-Sud, LMBA, UMR CNRS 6205, Vannes, France

*E-mail address*: `eric.miqueu@univ-ubs.fr`