



# An approximation of populations on a habitat with large carrying capacity

Naor Bauman<sup>1</sup> · Pavel Chigansky<sup>1</sup> · Fima Klebaner<sup>2</sup>

Received: 27 March 2023 / Revised: 10 October 2023 / Accepted: 18 February 2024 /

Published online: 18 March 2024

© The Author(s) 2024

## Abstract

We consider stochastic dynamics of a population which starts from a small colony on a habitat with large but limited carrying capacity. A common heuristics suggests that such population grows initially as a Galton–Watson branching process and then its size follows an almost deterministic path until reaching its maximum, sustainable by the habitat. In this paper we put forward an alternative and, in fact, more accurate approximation which suggests that the population size behaves as a special nonlinear transformation of the Galton–Watson process from the very beginning.

**Keywords** Population dynamics · Branching processes · Limit theorems · Approximation

**Mathematics Subject Classification** 92D25 · 60J80

## 1 Introduction

### 1.1 The model

A large population often starts from a few individuals who colonize a new habitat. Initially, in abundance of resources and lack of competition it grows rapidly until

---

✉ Pavel Chigansky  
Pavel.Chigansky@mail.huji.ac.il

Naor Bauman  
Naor.Bauman@mail.huji.ac.il

Fima Klebaner  
fima.klebaner@monash.edu

<sup>1</sup> Department of Statistics, The Hebrew University of Jerusalem, Mount Scopus Campus, 9190501 Jerusalem, Israel

<sup>2</sup> School of Mathematics, Monash University, Clayton, VIC 3058, Australia

reaching the carrying capacity. Then the population fluctuates around the carrying capacity for a very long period of time, until, by chance, it eventually dies out, see, e.g., Haccou et al. (2007); Hamza et al. (2016).

This cycle is captured by a stochastic model of density dependent branching process  $Z = (Z_n, n \in \mathbb{Z}_+)$  generated by the recursion

$$Z_n = \sum_{j=1}^{Z_{n-1}} \xi_{n,j}, \quad n \in \mathbb{N}, \tag{1}$$

started at an initial colony size  $Z_0$ . The random variables  $\xi_{n,j}$  take integer values and, for each  $n \in \mathbb{N}$ , are conditionally i.i.d. given all previous generations

$$\mathcal{F}_{n-1} = \sigma\{\xi_{m,j} : m < n, j \in \mathbb{N}\}.$$

The object of our study is the *density* process of the population  $\bar{Z}_n := Z_n/K$  relative to the carrying capacity parameter  $K > 0$ . The common distribution of the random variables  $\xi_{n,j}$  is assumed to depend on the density  $\bar{Z}_{n-1}$ :

$$P(\xi_{n,1} = \ell | \mathcal{F}_{n-1}) = p_\ell(\bar{Z}_{n-1}), \quad \ell \in \mathbb{Z}_+, \tag{2}$$

and is determined by the functions  $p_\ell : \mathbb{R}_+ \mapsto [0, 1]$ .

Both processes  $Z$  and  $\bar{Z}$  are indexed by  $K$ , but this dependence is suppressed in the notation. The mean and the variance of offspring distribution when the density process has value  $x$  are denoted by

$$m(x) = \sum_{k=0}^{\infty} k p_k(x) \quad \text{and} \quad \sigma^2(x) = \sum_{k=0}^{\infty} (k - m(x))^2 p_k(x), \quad x \in \mathbb{R}_+, \tag{3}$$

assumed to exist. Consequently,

$$E(\xi_{n,1} | \mathcal{F}_{n-1}) = m(\bar{Z}_{n-1}) \quad \text{and} \quad \text{Var}(\xi_{n,1} | \mathcal{F}_{n-1}) = \sigma^2(\bar{Z}_{n-1}).$$

If the offspring mean function satisfies

$$m(x) \begin{cases} > 1, & x < 1 \\ = 1, & x = 1 \\ < 1, & x > 1 \end{cases} \tag{4}$$

the process  $Z$  has a supercritical reproduction below the capacity  $K$ , critical reproduction at  $K$  and a subcritical reproduction above  $K$ . Thus a typical trajectory of  $Z$  grows rapidly until it reaches the vicinity of  $K$ . It then stays there fluctuating for a very long period of time and gets extinct eventually if  $p_0(x) > 0$  for all  $x \in \mathbb{R}_+$ . Thus the lifespan of such population roughly divides between the emergence stage, at

which the population establishes itself, the quasi-stationary stage around the carrying capacity and the decline stage which ends up with inevitable extinction.

**Remark 1** While (4) is typical for populations with quasi stable equilibrium at the capacity, it is not needed in the proofs and will not be assumed in what follows.

**1.2 Large initial colony**

A more quantitative picture can be obtained by considering the dynamics for the density process derived from (1) by setting  $f(x) := xm(x)$ , dividing by  $K$  and rearranging:

$$\bar{Z}_n = f(\bar{Z}_{n-1}) + \frac{1}{K} \sum_{j=1}^{Z_{n-1}} (\xi_{n,j} - m(\bar{Z}_{n-1})). \tag{5}$$

The second term on the right has zero mean and conditional variance

$$\text{Var} \left( \frac{1}{K} \sum_{j=1}^{Z_{n-1}} (\xi_{n,j} - m(\bar{Z}_{n-1})) \middle| \mathcal{F}_{n-1} \right) = K^{-1} \bar{Z}_{n-1} \sigma^2(\bar{Z}_{n-1}).$$

Consequently (5) can be viewed as a deterministic dynamical system perturbed by small noise of order<sup>1</sup>  $O_p(K^{-1/2})$ . If the initial colony size is relatively large, i.e., proportional to the carrying capacity:

$$\bar{Z}_0 = Z_0/K \xrightarrow{K \rightarrow \infty} x_0 > 0,$$

then  $\bar{Z}_n \xrightarrow{K \rightarrow \infty} x_n$  where  $x_n$  follows the unperturbed deterministic dynamics

$$x_n = f(x_{n-1}), \quad n \in \mathbb{N}, \tag{6}$$

started at  $x_0$ . If (4) is assumed,  $x = 1$  is the stable fixed point of  $f$  and if, in addition,  $f$  is an increasing function, then the sequence  $x_n$  increases to 1 with  $n$  when  $x_0 < 1$ . This limit also implies that the probability of early extinction tends to zero as  $K \rightarrow \infty$ .

Moreover, the stochastic fluctuations about the deterministic limit converge to a Gaussian process  $V = (V_n, n \in \mathbb{Z}_+)$  in distribution:

$$\sqrt{K}(\bar{Z}_n - x_n) \xrightarrow{K \rightarrow \infty} V_n$$

where  $V_n$  satisfies the recursion, Klebaner and Nerman (1994),

$$V_n = f'(x_{n-1})V_{n-1} + \sqrt{x_{n-1}\sigma^2(x_{n-1})}W_n, \quad n \in \mathbb{N},$$

<sup>1</sup> The usual notations for probabilistic orders is used throughout. In particular, for a sequence of random variables  $\zeta(K)$  and a sequence of numbers  $\alpha(K) \searrow 0$  as  $K \rightarrow \infty$ , the notation  $\zeta(K) = O_p(\alpha(K))$  means that the sequence  $\alpha(K)^{-1}\zeta(K)$  is bounded in probability.

with  $N(0, 1)$  i.i.d. random variables  $W_n$ 's.

Roughly speaking, this implies that when  $K$  is large,  $Z_n$  grows towards the capacity  $K$  along the deterministic path  $Kx_n$  and its fluctuations are of order  $O_p(K^{1/2})$ :

$$Z_n = x_n K + V_n K^{1/2} + o_p(K^{1/2}), \quad n \in \mathbb{N}. \tag{7}$$

If  $p_0(x) > 0$  for all  $x \in \mathbb{R}_+$  and (4) is imposed, zero is an absorbing state and hence the population gets extinct eventually. Large deviations analysis, see for example Klebaner and Zeitouni (1994); Klebaner et al. (1998), and Jung (2013); Högnäs (2019), shows that the mean of the time to extinction  $\tau_e = \inf\{n \geq 0 : Z_n = 0\}$  grows exponentially with  $K$ . In this paper we are concerned with the establishment stage of the population, which occurs well before the ultimate extinction, on the time scale of  $\log K$ .

### 1.3 Small initial colony

When  $Z_0$  is a fixed integer, say  $Z_0 = 1$ , then  $Z_0/K \rightarrow x_0 = 0$  and, since  $f(0) = 0$ , the solution to (6) is  $x_n = 0$  for all  $n \in \mathbb{N}$ . In this case the approximation (7) ceases to provide useful information. An alternative way to describe the stochastic dynamics in this setting was suggested recently in Barbour et al. (2016); Chigansky et al. (2018, 2019). It is based on the long known heuristics (Kendall 1956; Whittle 1955; Metz 1978), according to which such a population behaves initially as the Galton–Watson branching process and, if it manages to avoid extinction at this early stage, it continues to grow towards the carrying capacity following an almost deterministic curve.

This heuristics is made precise in Chigansky et al. (2019) as follows. We couple  $Z$  to a supercritical Galton–Watson branching process  $Y = (Y_n, n \in \mathbb{Z}_+)$  started at  $Y_0 = Z_0 = 1$ ,

$$Y_n = \sum_{j=1}^{Y_{n-1}} \eta_{n,j} \tag{8}$$

with the offspring distribution identical to that of  $Z$  at zero density size

$$P(\eta_{1,1} = \ell) = p_\ell(0), \quad \ell \in \mathbb{Z}_+.$$

This coupling is defined under assumption (a1.) below in Sect. 3.2.

Denote by  $\rho := m(0) > 1$ , define<sup>2</sup>  $n_c := n_c(K) = \lceil \log_\rho K^c \rceil$  for some  $c \in (\frac{1}{2}, 1)$  and let  $\bar{Y}_n := Y_n/K$  be the density of  $Y$ . Then  $\bar{Z}_n = Z_n/K$  is approximated in Chigansky et al. (2019) by

$$\begin{cases} \bar{Y}_n, & n \leq n_c, \\ f^{n-n_c}(\bar{Y}_{n_c}), & n > n_c, \end{cases}$$

<sup>2</sup>  $[x]$  and  $\{x\} = x - [x]$  denote the integer and fractional part of  $x \in \mathbb{R}_+$ .

where  $f^k$  stands for the  $k$ -th iterate of  $f$ . As is well known (Athreya and Ney 1972)

$$\rho^{-n} Y_n \xrightarrow[n \rightarrow \infty]{\text{P-a.s.}} W,$$

where  $W$  is an a.s. finite random variable. Moreover, under certain technical conditions on  $f$ , the limit

$$H(x) := \lim_{n \rightarrow \infty} f^n(x/\rho^n), \quad x \in \mathbb{R}_+ \tag{9}$$

can be shown to exist and define a continuous function.

**Theorem 1** (Chigansky et al. 2019) *Let  $n_1 := n_1(K) = \lceil \log_\rho K \rceil$  then*

$$\bar{Z}_{n_1} - H\left(W \rho^{-\lceil \log_\rho K \rceil}\right) \xrightarrow[K \rightarrow \infty]{\text{P}} 0. \tag{10}$$

In particular, this result implies that when  $K$  is a large integer power of  $\rho$  the distribution of  $\bar{Z}_{n_1}$  is close to that of  $H(W)$ . Moreover,

$$\bar{Z}_{n_1+n} \xrightarrow[K \rightarrow \infty]{\text{P}} x_n, \quad n \in \mathbb{N},$$

where  $x_n$  solves (6) started from the *random* initial condition  $H(W)$ . This approximation also captures the early extinction event since  $H(0) = 0$  and  $\text{P}(W = 0) = \text{P}(\lim_n Y_n = 0)$ , the extinction probability of the Galton–Watson process  $Y$ .

### 1.4 This paper’s contribution

In this paper we address the question of the rate of convergence in (10). Note that if the probabilities in (2) are constant with respect to  $x$  then  $f(x) = \rho x$ , consequently  $H(x) = x$ , and the processes  $Z$  and  $Y$  coincide. In this case

$$\sqrt{K}(\bar{Y}_{n_1} - W \rho^{-\lceil \log_\rho K \rceil}) = \rho^{-\frac{1}{2} \lceil \log_\rho K \rceil} \sqrt{\rho^{n_1}}(\rho^{-n_1} Y_{n_1} - W) = O_{\text{P}}(1) \tag{11}$$

where the order of convergence is implied by the CLT for the Galton–Watson process (Heyde 1970) by which  $\sqrt{\rho^{n_1}}(\rho^{-n_1} Y_{n_1} - W)$  converges in distribution to a mixed normal law as  $n \rightarrow \infty$ . Thus it can be expected that at best the sequence in (10) is of order  $O_{\text{P}}(K^{-1/2})$  as  $K \rightarrow \infty$ . However, the best rate of convergence in the approximation in Theorem 1 described above, is achieved with  $c = \frac{5}{8}$  and it is only  $O_{\text{P}}(K^{-1/8} \log K)$ . This can be seen from a close examination of the proof in Chigansky et al. (2019).

The goal of this paper is to put forward a different approximation with much faster rate of convergence of order  $O_{\text{P}}(K^{-1/2} \log K)$ . This is still slower than the rate achievable in the density independent case, but only by a logarithmic factor. The new proof highlights a better understanding of population dynamics at the emergence stage, which shows that, in fact, a sharper approximation is given by the Galton–Watson process transformed by a nonlinear function  $H$  arising in deterministic dynamics (9).

It is not clear at the moment whether the  $\log K$  factor is avoidable and whether a central limit type theorem holds. These questions are left for further research.

## 2 The main result

We will make the following assumptions.

- a1. The offspring distribution  $F_x(t) = \sum_{\ell \leq t} p_\ell(x)$  is stochastically decreasing with respect to the population density: for any  $y \geq x$ ,

$$F_y(t) \geq F_x(t), \quad \forall t \in \mathbb{R}_+.$$

- a2. The second moment of the offspring distribution, cf. (3),

$$m_2(x) = \sigma^2(x) + m(x)^2$$

is  $L$ -Lipschitz for some  $L > 0$ .

- a3. The function  $f(x) = xm(x)$  has two continuous bounded derivatives and <sup>3</sup>

$$\|f'\|_\infty = f'(0) = \rho.$$

**Remark 2** Assumption (a1.) means that the reproduction drops with population density. In particular, it implies that  $x \mapsto m(x)$  is a decreasing function and hence,

$$f'(x) = m(x) + xm'(x) \leq m(x) \leq \rho, \quad \forall x \in \mathbb{R}_+,$$

which is only slightly weaker than (a3.). The assumption (a2.) is technical.

**Remark 3** The distribution of the process  $\bar{Z}$  does not depend on the values of  $\{p_\ell(0), \ell \in \mathbb{Z}_+\}$  for any  $K$ , while the distribution of  $W$  and, therefore, of  $H(W)$  does. This is not a contradiction since our assumptions imply continuity of  $x \mapsto p_\ell(x)$  at  $x = 0$  for all  $\ell \in \mathbb{Z}_+$ . Indeed,  $m(x) = \int_0^\infty (1 - F_x(t))dt$  and therefore

$$\int_0^\infty (F_x(t) - F_0(t))dt = m(0) - m(x) \xrightarrow{x \rightarrow 0} 0$$

where the convergence holds since  $m(x)$  is differentiable and a fortiori continuous at  $x = 0$ . By the stochastic order assumption (a1.),  $F_x(t) - F_0(t) \geq 0$  for any  $t \geq 0$ . Since both  $F_x$  and  $F_0$  are discrete with jumps at integers, for any  $s \geq 0$ ,

$$F_x(s) - F_0(s) = \int_{[s]}^{[s]+1} (F_x(t) - F_0(t))dt \leq \int_0^\infty (F_x(t) - F_0(t))dt \xrightarrow{x \rightarrow 0} 0.$$

This in turn implies that  $p_\ell(x) \rightarrow p_\ell(0)$  as  $x \rightarrow 0$  for all  $\ell$ .

<sup>3</sup>  $\|f\|_\infty = \sup_x |f(x)|$

**Theorem 2** Under assumptions (a1.)–(a3.)

$$\bar{Z}_{n_1} - H\left(W\rho^{-\{\log_\rho K\}}\right) = O_P(K^{-1/2} \log K), \text{ as } K \rightarrow \infty.$$

**Example 1** The binary splitting model from Chigansky et al. (2018) satisfies the above assumptions. Another example is Geometric offspring distribution

$$p_\ell(x) = q(x)^\ell(1 - q(x)), \quad \ell \in \mathbb{Z}_+$$

where  $q : \mathbb{R}_+ \mapsto [0, 1]$  is a decreasing function. This distribution satisfies the stochastic order condition (a1.). The normalization  $m(0) = \rho$  and  $m(1) = 1$  implies that  $q(0) = \rho/(1 + \rho)$  and  $q(1) = 1/2$ . A direct calculation shows that, e.g.,

$$q(x) = \frac{\rho}{1 + \rho} \exp\left(-x \log \frac{2\rho}{1 + \rho}\right), \quad x \geq 0$$

satisfies both (a2.) and (a3.).

**Example 2** Stochastic Ricker model (Högnäs 1997) is given by a density dependent branching process with the offspring distribution

$$p_\ell(x) = q_\ell e^{-\gamma x},$$

where  $\gamma > 0$  is a constant,  $q_\ell, \ell \geq 1$  is a given probability distribution, and no offspring are produced with probability  $1 - e^{-\gamma x}$ . This model satisfies the stochastic ordering assumption (a1.). The mean value of the distribution  $q_\ell$  is denoted by  $e^r$ , to emphasize the relation to the deterministic Ricker model. With such notation,

$$m(x) = e^{r-\gamma x}, \quad f(x) = x e^{r-\gamma x}.$$

Under normalization  $m(0) = \rho$  and  $m(1) = 1$  this becomes

$$m(x) = \rho^{1-x}, \quad f(x) = x\rho^{1-x}.$$

A direct calculation verifies the assumptions (a2.) and (a3.).

### 3 Proof of Theorem 2

We will construct the process  $Z$  defined in (1) and the Galton–Watson process  $Y$  from (8) on a suitable probability space so that  $Y_n \geq Z_n$  for all  $n \in \mathbb{N}$  and the trajectories of these processes remain sufficiently close at least for relatively small  $n$ 's (Sect. 3.2). We will then show that  $H$  is twice continuously differentiable (Sect. 3.1) and use Taylor's approximation to argue (Sect. 3.3) that

$$\bar{Z}_{n_1} - H(\bar{Y}_{n_1}) = O_P(K^{-1/2} \log K), \text{ as } K \rightarrow \infty.$$

This convergence combined with (11) implies the result. Below we will write  $C$  for a generic constant whose value may change from line to line.

### 3.1 Properties of $H$

In this section we establish existence of the limit (9) under the standing assumptions and verify its smoothness. The proof of existence relies on a result on functional iteration, shown in Baker et al. (2020).

**Lemma 3** (Baker et al. 2020, Lemma 1) *Let  $x_{m,n}$  be the sequence generated by the recursion*

$$x_{m,n} = \rho x_{m-1,n}(1 + Cx_{m-1,n}), \quad m = 1, \dots, n$$

*subject to the initial condition  $x_{0,n} = x/\rho^n > 0$ , where  $\rho > 1$  and  $C \geq 0$  are constants. There exists a locally bounded function  $\psi : \mathbb{R}_+ \mapsto \mathbb{R}_+$  such that for any  $n \in \mathbb{N}$*

$$x_{m,n} \leq \psi(x)\rho^{m-n}, \quad m = 1, \dots, n. \tag{12}$$

Throughout we will use the notation  $H_n(x) := f^n(x/\rho^n)$ .

**Lemma 4** *Under assumption (a3.) there exists a continuous function  $H : \mathbb{R}_+ \mapsto \mathbb{R}_+$  and a locally bounded function  $g : \mathbb{R}_+ \mapsto \mathbb{R}_+$  such that*

$$|H_n(x) - H(x)| \leq g(x)\rho^{-n}, \quad n \in \mathbb{N}.$$

**Proof** By assumption (a3.)

$$f(x) = \rho x + \int_0^x \int_0^t f''(s) ds dt \tag{13}$$

and hence for any  $x, y \in \mathbb{R}_+$

$$|f(y) - f(x)| \leq \rho|y - x| + \frac{1}{2} \|f''\|_\infty |y^2 - x^2| \leq \rho(1 + C|y| \vee |x|)|y - x| \tag{14}$$

with  $C = \|f''\|_\infty/\rho$ . Thus the sequence  $x_{m,n} := f^m(x/\rho^n)$  satisfies

$$x_{m,n} = f(x_{m-1,n}) \leq \rho(1 + Cx_{m-1,n})x_{m-1,n}$$

and  $x_{0,n} = x/\rho^n$ . By Lemma 3 there exists a locally bounded function  $\psi$  such that for any  $n \in \mathbb{N}$

$$|f^m(x/\rho^n)| \leq \psi(x)\rho^{m-n}, \quad m = 1, \dots, n. \tag{15}$$



The bound (14) also implies

$$\begin{aligned}
 |f^{m+1}(x/\rho^{n+1}) - f^m(x/\rho^n)| &= |f \circ f^m(x/\rho^{n+1}) - f \circ f^{m-1}(x/\rho^n)| \\
 &\leq \rho(1 + CF_{m,n})|f^m(x/\rho^{n+1}) - f^{m-1}(x/\rho^n)|
 \end{aligned}
 \tag{16}$$

where, in view of (15),

$$F_{m,n} := f^m(x/\rho^{n+1}) \vee f^{m-1}(x/\rho^n) \leq \psi(x)\rho^{m-1-n}.$$

Since  $f$  has bounded second derivative and  $f'(0) = \rho$ , cf. (13),

$$|f(x/\rho^{n+1}) - x/\rho^n| \leq \frac{1}{2}\|f''\|_\infty(x/\rho)^2\rho^{-2n}.$$

Plugging this bound into (16) and iterating  $n$  times we obtain

$$\begin{aligned}
 |f^{n+1}(x/\rho^{n+1}) - f^n(x/\rho^n)| &\leq |f(x/\rho^{n+1}) - x/\rho^n| \rho^n \prod_{m=1}^n (1 + CF_{m,n}) \\
 &\leq \frac{1}{2}\|f''\|_\infty(x/\rho)^2\rho^{-2n} \rho^n \prod_{m=1}^n (1 + C\psi(x)\rho^{m-1-n}) \leq \tilde{g}(x)\rho^{-n}
 \end{aligned}$$

where we defined

$$\tilde{g}(x) := \frac{1}{2}\|f''\|_\infty(x/\rho)^2 \prod_{k=1}^\infty (1 + C\psi(x)\rho^{-k}).$$

Thus the limit  $H(x) = \lim_{n \rightarrow \infty} f^n(x/\rho^n)$  exists and satisfies the claimed bound with  $g(x) = \tilde{g}(x)/(1 - \rho^{-1})$ . Continuity of  $H$  follows since  $H_n$  are continuous for each  $n$  and the convergence is uniform on compacts. □

**Corollary 5**  *$f$  is topologically semiconjugate to its linearization at the origin:*

$$H(x) = f \circ H(x/\rho), \quad \forall x \in \mathbb{R}_+.$$

**Proof** Since  $f$  is continuous

$$H(x) = \lim_{n \rightarrow \infty} f^{n+1}(x/\rho^{n+1}) = \lim_{n \rightarrow \infty} f \circ f^n((x/\rho)\rho^{-n}) = f \circ H(x/\rho).$$

□

The next lemma shows that  $H$  is strictly increasing in a vicinity of the origin and is therefore a local conjugacy.

**Lemma 6** *There exists an  $a > 0$  such that  $H$  is strictly increasing on  $[0, a]$  and*

$$f(x) = H(\rho H^{-1}(x)), \quad x \in [0, H(a)]. \tag{17}$$

**Proof** Let  $c := \|f''\|_\infty$  and  $r := \rho/c$  then

$$f'(x) \geq \rho - cx > 0, \quad \forall x \in [0, r].$$

Since  $f$  is  $\rho$ -Lipschitz and  $f(0) = 0$ , for any  $j = 1, \dots, n$  and  $x \in [0, r]$ ,

$$f^{n-j}(x/\rho^n) \leq x/\rho^j \in [0, r]$$

and hence for all  $x \in [0, r]$

$$\begin{aligned} H'_n(x) &= \prod_{j=1}^n \frac{1}{\rho} f'(f^{n-j}(x/\rho^n)) \geq \prod_{j=1}^n \left(1 - \frac{c}{\rho} f^{n-j}(x/\rho^n)\right) \\ &\geq \prod_{j=1}^n \left(1 - \frac{c}{\rho} x \rho^{-j}\right) \geq 1 - \frac{c}{\rho} x \sum_{j=1}^n \rho^{-j} \geq 1 - \frac{c}{\rho - 1} x, \end{aligned}$$

where we used the Bernoulli inequality. Thus we can choose a number  $a \in (0, r)$  such that  $H'_n(x) \geq 1/2$  for all  $x \in [0, a]$ . It then follows that for any  $y > x$  in the interval  $[0, a]$

$$H_n(y) - H_n(x) = \int_x^y H'_n(t) dt \geq \frac{1}{2}(y - x) > 0.$$

Taking the limit  $n \rightarrow \infty$  implies that  $H$  is strictly increasing on  $[0, a]$ . Being continuous,  $H$  is invertible and (17) holds by Corollary 5. □

**Remark 4** Under additional assumption that  $f$  is strictly increasing on the whole  $\mathbb{R}_+$ , the function  $H$  is furthermore a global conjugacy, i.e. (17) holds on  $\mathbb{R}_+$ .

The next lemma establishes differentiability of  $H$ .

**Lemma 7**  *$H$  has continuous derivative*

$$H'(x) = \prod_{j=1}^\infty \frac{1}{\rho} f'(H(x\rho^{-j})), \quad \forall x \in \mathbb{R}_+ \tag{18}$$

where the series converges uniformly on compacts.

**Proof** Step 1. Let us first argue that the infinite product in (18)

$$G(x) := \prod_{j=1}^\infty \frac{1}{\rho} f'(H(x\rho^{-j})) \tag{19}$$

is well defined. By assumption (a3.),  $f$  is  $\rho$ -Lipschitz and hence  $f^n$  is  $\rho^n$ -Lipschitz. Consequently,  $H_n$  is 1-Lipschitz for all  $n \in \mathbb{N}$  and so is  $H$ . This will be used in the proof on several occasions. Let  $c := \|f''\|_\infty$  and  $r := \frac{1}{2}\rho/c$ , then

$$f'(x) \geq \rho - cx > 0, \quad \forall x \in [0, r]. \tag{20}$$

For  $x > 0$  define the function  $j(x) := \lceil \log_\rho(x/r) \rceil$ . Then for any  $j > j(x)$ ,

$$\begin{aligned} \left| \log \frac{1}{\rho} f'(H(x\rho^{-j})) \right| &= -\log \frac{1}{\rho} f'(H(x\rho^{-j})) \leq -\log \left( 1 - \frac{c}{\rho} H(x\rho^{-j}) \right) \\ &\stackrel{\dagger}{\leq} 2\frac{c}{\rho} H(x\rho^{-j}) \leq 2\frac{c}{\rho} x\rho^{-j} =: Cx\rho^{-j}, \end{aligned} \tag{21}$$

where  $\dagger$  holds since  $-\log(1 - u) \leq 2u$  for all  $u \in [0, \frac{1}{2}]$ . The partial products in (19) can be written as

$$\begin{aligned} G_n(x) &:= \prod_{j=1}^n \frac{1}{\rho} f'(H(x\rho^{-j})) \\ &= \left( \prod_{j=1}^{j(x)} \frac{1}{\rho} f'(H(x\rho^{-j})) \right) \exp \left( \sum_{j=j(x)+1}^n \log \frac{1}{\rho} f'(H(x\rho^{-j})) \right) =: T(x) \exp(L_n(x)). \end{aligned}$$

In view of the estimate (21),  $G_n(x)$  converges to  $G(x) := T(x) \exp(L(x))$  for any  $x \in \mathbb{R}_+$  where  $L(x) = \lim_n L_n(x)$ . Furthermore,

$$\begin{aligned} |G_n(x) - G(x)| &= |T(x)| \left| \exp(L_n(x)) - \exp(L(x)) \right| \\ &\leq \exp(L(x) \vee L_n(x)) |L(x) - L_n(x)| \end{aligned} \tag{22}$$

where we used the bound  $|T(x)| \leq 1$ . For any  $R > 0$  and all  $x \in [0, R]$  the estimate (21) implies

$$|L(x) - L_n(x)| = \sum_{j=n+1}^\infty \left| \log \frac{1}{\rho} f'(H(x\rho^{-j})) \right| \leq \sum_{j=n+1}^\infty Cx\rho^{-j} \leq CR \frac{\rho^{-n}}{\rho - 1},$$

and thus, in view of the bound (22), we obtain

$$\sup_{x \leq R} |G_n(x) - G(x)| \rightarrow 0. \tag{23}$$

Since  $G_n$  is continuous for any  $n$ , this uniform convergence implies that  $G$  is continuous as well.

Step 2. To show that  $H(x)$  is differentiable and to verify the claimed formula for the derivative, it remains to show that the sequence of derivatives

$$H'_n(x) = \prod_{j=1}^n \frac{1}{\rho} f'(f^{n-j}(x/\rho^n))$$

converges to  $G$  uniformly on compacts. Fix an  $R > 0$ , define  $J(R) = \lceil \log_\rho(R/r) \rceil$  and, for  $n > J(R)$ , let

$$\tilde{P}_n(x) := \prod_{j=1}^{J(R)} \frac{1}{\rho} f'(f^{n-j}(x/\rho^n)), \quad P_n(x) := \prod_{j=J(R)+1}^n \frac{1}{\rho} f'(f^{n-j}(x/\rho^n))$$

and

$$\tilde{Q}_n(x) := \prod_{j=1}^{J(R)} \frac{1}{\rho} f'(H(x\rho^{-j})), \quad Q_n(x) := \prod_{j=J(R)+1}^n \frac{1}{\rho} f'(H(x\rho^{-j})).$$

Since  $\|f'\|_\infty = \rho$  all these functions are bounded by 1 and

$$\begin{aligned} |H'_n(x) - G(x)| &\leq |H'_n(x) - G_n(x)| + |G_n(x) - G(x)| \\ &= |\tilde{P}_n(x)P_n(x) - \tilde{Q}_n(x)Q_n(x)| + |G_n(x) - G(x)| \\ &\leq |P_n(x) - Q_n(x)| + |\tilde{P}_n(x) - \tilde{Q}_n(x)| + |G_n(x) - G(x)|. \end{aligned}$$

Since  $f'$  is continuous and the convergence  $H_n \rightarrow H$  is uniform on compacts, it follows that

$$\sup_{x \leq R} |\tilde{P}_n(x) - \tilde{Q}_n(x)| = \sup_{x \leq R} \left| \prod_{j=1}^{J(R)} \frac{1}{\rho} f'(H_{n-j}(x\rho^{-j})) - \prod_{j=1}^{J(R)} \frac{1}{\rho} f'(H(x\rho^{-j})) \right| \xrightarrow{n \rightarrow \infty} 0,$$

and hence, to complete the proof, we need to show that

$$\sup_{x \leq R} |P_n(x) - Q_n(x)| \xrightarrow{n \rightarrow \infty} 0. \tag{24}$$

To this end, in view of Corollary 5,

$$\begin{aligned} H(x\rho^{-j}) &= f \circ H(x\rho^{-j-1}) = f^2 \circ H(x\rho^{-j-2}) = \dots \\ &= f^{n-j} \circ H(x\rho^{-j-(n-j)}) = f^{n-j} \circ H(x\rho^{-n}) \end{aligned}$$

and hence

$$P_n(x) - Q_n(x) = \prod_{j=J(R)+1}^n \frac{1}{\rho} f'(f^{n-j}(x\rho^{-n})) - \prod_{j=J(R)+1}^n \frac{1}{\rho} f'(f^{n-j}(H(x\rho^{-n}))).$$

Consequently, for all  $x \in (0, R]$ ,

$$\begin{aligned}
 & \left| \log P_n(x) - \log Q_n(x) \right| \\
 & \leq \sum_{j=J(R)+1}^n \left| \log \frac{1}{\rho} f'(f^{n-j}(x\rho^{-n})) - \log \frac{1}{\rho} f'(f^{n-j}(H(x\rho^{-n}))) \right| \\
 & \stackrel{\dagger}{\leq} \frac{1}{\rho - cr} \|f''\|_\infty \sum_{j=J(R)+1}^n |f^{n-j}(x\rho^{-n}) - f^{n-j}(H(x\rho^{-n}))| \tag{25} \\
 & \leq \frac{1}{\rho - cr} \|f''\|_\infty \sum_{j=1}^n \rho^{n-j} |x\rho^{-n} - H(x\rho^{-n})| \leq C\rho^n |\rho^{-n}x - H(x\rho^{-n})| \\
 & = C\rho^n |H \circ H^{-1}(x\rho^{-n}) - H(x\rho^{-n})| \stackrel{\ddagger}{\leq} C|\rho^n H^{-1}(x\rho^{-n}) - x|.
 \end{aligned}$$

Here the bound  $\dagger$  holds since for  $j > J(R)$  both arguments of  $f'$  are smaller than  $r$  and thus (20) applies. The inequality  $\ddagger$  is true since  $H$  is 1-Lipschitz. The inverses in the last line of (25) are well defined for  $n \geq k := \lceil \log_\rho(R/H(a)) \rceil + 1$  where  $a$  is the constant guaranteed by Lemma 6. Moreover, for all such  $n$

$$\begin{aligned}
 |\rho^n H^{-1}(x\rho^{-n}) - x| &= \rho^k |\rho^{n-k} H^{-1}(x\rho^{-k}\rho^{-(n-k)}) - x\rho^{-k}| \\
 &= \rho^k |H^{-1} \circ f^{n-k}(x\rho^{-k}\rho^{-(n-k)}) - x\rho^{-k}| = \rho^k |H^{-1} \circ H_{n-k}(x\rho^{-k}) - x\rho^{-k}| \xrightarrow{n \rightarrow \infty} 0.
 \end{aligned} \tag{26}$$

Moreover, the sequence of functions  $D_n(x) := \rho^n H^{-1}(x\rho^{-n})$  is decreasing on  $[0, R]$  for all  $n$  large enough:

$$D_{n+1}(x) = \rho^n \rho H^{-1}(x\rho^{-n-1}) = \rho^n H^{-1} \circ f(x\rho^{-n-1}) \leq \rho^n H^{-1}(x\rho^{-n}) = D_n(x),$$

where the inequality holds since  $H^{-1}$  is increasing near the origin. It follows now from Dini’s theorem that the convergence in (26) is uniform:

$$\sup_{x \leq R} |\rho^n H^{-1}(x\rho^{-n}) - x| \xrightarrow{n \rightarrow \infty} 0.$$

The convergence in (24) holds since both  $Q_n$  and  $P_n$  are bounded by 1 and

$$\sup_{x \leq R} |P_n(x) - Q_n(x)| \leq \sup_{x \leq R} |P_n(x) \vee Q_n(x)| \sup_{x \leq R} |\log P_n(x) - \log Q_n(x)| \xrightarrow{n \rightarrow \infty} 0.$$

□

**Lemma 8** *H has continuous second derivative*

$$H''(x) = H'(x) \sum_{i=1}^{\infty} \frac{f''(H(x\rho^{-i}))}{f'(H(x\rho^{-i}))} H'(x\rho^{-i})\rho^{-i}. \tag{27}$$

**Proof** The partial products in (18)

$$G_n(x) := \prod_{j=1}^n \frac{1}{\rho} f'(H(x\rho^{-j}))$$

satisfy

$$\begin{aligned} G'_n(x) &= \sum_{i=1}^n \left( \prod_{j=1, j \neq i}^n \frac{1}{\rho} f'(H(x\rho^{-j})) \right) \frac{1}{\rho} f''(H(x\rho^{-i})) H'(x\rho^{-i}) \rho^{-i} \\ &= G_n(x) \sum_{i=1}^n \frac{f''(H(x\rho^{-i}))}{f'(H(x\rho^{-i}))} H'(x\rho^{-i}) \rho^{-i}, \end{aligned}$$

where the convention  $0/0 = 0$  is used. By assumption (a3.),  $f''/f'$  is bounded uniformly on a vicinity of the origin.  $H'$  is continuous by Lemma 7 and therefore is bounded on compacts. Hence the series is compactly convergent. By Lemma 7, so is  $G_n$ . Thus  $G'_n(x)$  converges compactly, its limit is continuous and coincides with  $H''(x)$ .  $\square$

### 3.2 The auxiliary Galton–Watson process

Let  $(U_{n,j} : n \in \mathbb{N}, j \in \mathbb{Z}_+)$  be an array of i.i.d.  $U([0, 1])$  random variables and define

$$\xi_{n,j}(x) = F_x^{-1}(U_{n,j}) := \min \{t \geq 0 : F_x(t) \geq U_{n,j}\},$$

where  $F_x(t)$  is the offspring distribution function when the population density is  $x$ , cf. assumption (a1.). Then  $P(\xi_{n,j}(x) = k) = p_k(x)$  for all  $k \in \mathbb{Z}_+$ . Let  $\eta_{n,j} := \xi_{n,j}(0)$ . By assumption (a1.)

$$\xi_{n,j}(x) \leq \eta_{n,j} \quad \forall x \in \mathbb{R}_+, n, j \in \mathbb{N}. \tag{28}$$

Let  $Z = (Z_n, n \in \mathbb{Z}_+)$  and  $Y = (Y_n, n \in \mathbb{Z}_+)$  be processes generated by the recursions

$$Z_n = \sum_{j=1}^{Z_{n-1}} \xi_{n,j}(\bar{Z}_{n-1}) \quad \text{and} \quad Y_n = \sum_{j=1}^{Y_{n-1}} \eta_{n,j}$$

started from the same initial conditions  $Z_0 = Y_0 = 1$ . By construction these processes coincide in distribution with (1) and (8) respectively. Moreover, in view of (28), by induction

$$Z_n \leq Y_n, \quad \forall n \in \mathbb{Z}_+. \tag{29}$$

### 3.3 The approximation

In view of (11),

$$\bar{Y}_{n_1} - W\rho^{-\{\log_\rho K\}} = \rho^{-\{\log_\rho K\}}(\rho^{-n_1}Y_{n_1} - W) = O_P(\rho^{-n_1/2}) = O_P(K^{-1/2}).$$

Since  $H$  has continuous derivative it follows that

$$H(\bar{Y}_{n_1}) - H(W\rho^{-\{\log_\rho K\}}) = O_P(K^{-1/2}).$$

Thus to prove the assertion of Theorem 2 it remains to show that

$$\bar{Z}_{n_1} - H(\bar{Y}_{n_1}) = O_P(K^{-1/2} \log K), \quad K \rightarrow \infty.$$

The process  $\bar{Y}_n = K^{-1}Y_n$  satisfies

$$\bar{Y}_n = \rho\bar{Y}_{n-1} + \frac{1}{K} \sum_{j=1}^{Y_{n-1}} (\eta_{n,j} - \rho).$$

By Taylor’s approximation and in view of Corollary 5

$$\begin{aligned} H(\bar{Y}_n) &= H(\rho\bar{Y}_{n-1}) + H'(\rho\bar{Y}_{n-1}) \frac{1}{K} \sum_{j=1}^{Y_{n-1}} (\eta_{n,j} - \rho) + R_n(K) \\ &= f(H(\bar{Y}_{n-1})) + H'(\rho\bar{Y}_{n-1}) \frac{1}{K} \sum_{j=1}^{Y_{n-1}} (\eta_{n,j} - \rho) + R_n(K) \end{aligned} \tag{30}$$

where

$$R_n(K) := \frac{1}{2} H''(\theta_{n-1}(K)) \left( \frac{1}{K} \sum_{j=1}^{Y_{n-1}} (\eta_{n,j} - \rho) \right)^2 \tag{31}$$

with  $\theta_{n-1}(K) \geq 0$  satisfying

$$|\theta_{n-1}(K) - \rho\bar{Y}_{n-1}| \leq \left| \frac{1}{K} \sum_{j=1}^{Y_{n-1}} (\eta_{n,j} - \rho) \right|. \tag{32}$$

Since  $\|f'\|_\infty = \rho$  is assumed,  $f$  is  $\rho$ -Lipschitz. By subtracting equation (5) from (30) we obtain the bound for  $\delta_n := |H(\bar{Y}_n) - \bar{Z}_n|$ :

$$\delta_n \leq \rho\delta_{n-1} + |\varepsilon_n^{(1)}| + |\varepsilon_n^{(2)}| + |\varepsilon_n^{(3)}| + |R_n(K)| \tag{33}$$

subject to  $\delta_0 = |H(1/K) - 1/K|$ , where we defined

$$\begin{aligned} \varepsilon_n^{(1)} &= \left( H'(\rho \bar{Y}_{n-1}) - 1 \right) \frac{1}{K} \sum_{j=1}^{Y_{n-1}} (\eta_{n,j} - \rho), \\ \varepsilon_n^{(2)} &= \frac{1}{K} \sum_{j=1}^{Z_{n-1}} \left( (\eta_{n,j} - \rho) - (\xi_{n,j}(\bar{Z}_{n-1}) - m(\bar{Z}_{n-1})) \right), \\ \varepsilon_n^{(3)} &= \frac{1}{K} \sum_{j=Z_{n-1}+1}^{Y_{n-1}} (\eta_{n,j} - \rho). \end{aligned}$$

Consequently,

$$\delta_n \leq \rho^n \delta_0 + \sum_{j=1}^n \rho^{n-j} \left( |\varepsilon_j^{(1)}| + |\varepsilon_j^{(2)}| + |\varepsilon_j^{(3)}| + |R_j(K)| \right)$$

and it is left to show that the contribution of each term at time  $n_1 = \lceil \log_\rho K \rceil$  is of order  $O_P(K^{-1/2} \log K)$  as  $K \rightarrow \infty$ .

### 3.3.1 Contribution of the initial condition

Since  $H(0) = 0$  and, by (18),  $H'(0) = 1$ , Taylor’s approximation implies that for all  $K$  large enough

$$\delta_0 = |H(1/K) - 1/K| \leq \frac{1}{2} \sup_{x \leq 1} |H''(x)| K^{-2} = CK^{-2}$$

and, consequently,  $|\rho^{n_1} \delta_0| \leq CK^{-1}$ .

### 3.3.2 Contribution of $R_n(K)$

To estimate the residual, defined in (31), let us show first that the family of random variables

$$\max_{m \leq n_1} \left| H''(\theta_m(K)) \right| \tag{34}$$

is bounded in probability as  $K \rightarrow \infty$ . By equation (32),

$$\begin{aligned} E\theta_{n-1}(K) &\leq \rho E\bar{Y}_{n-1} + E \left| \frac{1}{K} \sum_{j=1}^{Y_{n-1}} (\eta_{n,j} - \rho) \right| \\ &\leq \rho E\bar{Y}_{n-1} + \frac{1}{K} \sqrt{EY_{n-1}\sigma^2(0)} \leq \frac{1}{K} \rho^n + \frac{1}{K} \sqrt{\rho^n \sigma^2(0)} \\ &\leq K^{-1} \rho^n + CK^{-1} \rho^{n/2} \leq 2CK^{-1} \rho^n. \end{aligned}$$



If  $H''$  is bounded then (34) is obviously bounded. Let us proceed assuming that  $H''$  is unbounded. Define  $\psi(M) := \max_{x \leq M} |H''(x)|$ . By continuity,  $\psi(M)$  is finite, continuous and increases to  $\infty$ . Let  $\psi^{-1}$  be its generalized inverse

$$\psi^{-1}(t) = \inf\{x \geq 0 : \psi(x) \geq t\}.$$

Since  $\psi$  is continuous and unbounded,  $\psi^{-1}$  is nondecreasing (not necessarily continuous) and  $\psi^{-1}(t) \rightarrow \infty$  as  $t \rightarrow \infty$ . Then for any  $R \geq 0$ , by the union bound,

$$\begin{aligned} \mathbb{P}\left(\max_{m \leq n_1} |H''(\theta_m(K))| \geq R\right) &\leq \mathbb{P}\left(\max_{m \leq n_1} \psi(\theta_m(K)) \geq R\right) \\ &\leq \sum_{m=1}^{n_1} \mathbb{P}\left(\psi(\theta_m(K)) \geq R\right) \leq \sum_{m=1}^{n_1} \mathbb{P}\left(\theta_m(K) \geq \psi^{-1}(R)\right) \\ &\leq \sum_{m=1}^{n_1} \frac{\mathbb{E}\theta_m(K)}{\psi^{-1}(R)} \leq \frac{1}{\psi^{-1}(R)} \sum_{m=1}^{n_1} 2CK^{-1}\rho^m \leq \frac{\rho}{\rho-1} \frac{2C}{\psi^{-1}(R)} \xrightarrow{R \rightarrow \infty} 0. \end{aligned}$$

This proves that (34) is bounded in probability. The contribution of  $R_n(K)$  in (33) can now be bounded as

$$\left| \sum_{m=1}^{n_1} \rho^{n_1-m} R_m(K) \right| \leq \max_{j \leq n_1} |H''(\theta_j(K))| \sum_{m=1}^{n_1} \rho^{n_1-m} \left( \frac{1}{K} \sum_{j=1}^{Y_{m-1}} (\eta_{m,j} - \rho) \right)^2$$

where

$$\begin{aligned} \mathbb{E} \sum_{m=1}^{n_1} \rho^{n_1-m} \left( \frac{1}{K} \sum_{j=1}^{Y_{m-1}} (\eta_{m,j} - \rho) \right)^2 \\ = \sum_{m=1}^{n_1} \rho^{n_1-m} \frac{1}{K^2} \mathbb{E} Y_{m-1} \sigma^2(0) \leq \sum_{m=1}^{n_1} \rho^{n_1-m} \frac{1}{K^2} \rho^m \sigma^2(0) \leq CK^{-1} \log K. \end{aligned}$$

Hence

$$\left| \sum_{m=1}^{n_1} \rho^{n_1-m} R_m(K) \right| = O_{\mathbb{P}}(1) O_{\mathbb{P}}(K^{-1} \log K) = O_{\mathbb{P}}(K^{-1} \log K).$$

### 3.3.3 Contribution of $\varepsilon^{(3)}$

By conditional independence of  $\eta_{n,j}$ 's

$$\mathbb{E}(\varepsilon_m^{(3)})^2 = \frac{\sigma^2(0)}{K} \mathbb{E}(\bar{Y}_{m-1} - \bar{Z}_{m-1}).$$

In view of (29), the sequence  $D_m := \bar{Y}_m - \bar{Z}_m \geq 0$  satisfies

$$\begin{aligned}
 \mathbb{E}D_m &= \frac{1}{K} \mathbb{E} \left( \sum_{j=1}^{Y_{m-1}} \eta_{m,j} - \sum_{j=1}^{Z_{m-1}} \xi_{m,j}(\bar{Z}_{m-1}) \right) \\
 &= \frac{1}{K} \mathbb{E} \sum_{j=Z_{m-1}+1}^{Y_{m-1}} \eta_{m,j} + \frac{1}{K} \mathbb{E} \sum_{j=1}^{Z_{m-1}} (\eta_{m,j} - \xi_{m,j}(\bar{Z}_{m-1})) \\
 &= \rho \mathbb{E}D_{m-1} + \frac{1}{K} \mathbb{E} \sum_{j=1}^{Z_{m-1}} (\rho - m(\bar{Z}_{m-1})) \\
 &= \rho \mathbb{E}D_{m-1} + \mathbb{E}\bar{Z}_{m-1}(\rho - m(\bar{Z}_{m-1})) \\
 &= \rho \mathbb{E}D_{m-1} + \mathbb{E}(\rho \bar{Z}_{m-1} - f(\bar{Z}_{m-1})) \\
 &\leq \rho \mathbb{E}D_{m-1} + \frac{1}{2} \|f''\|_\infty \mathbb{E}\bar{Z}_{m-1}^2 \\
 &\leq \rho \mathbb{E}D_{m-1} + CK^{-2} \rho^{2m},
 \end{aligned}$$

where the last bound holds in view of (29) and the well known formula for the second moment of the Galton–Watson process. Since  $D_0 = 0$  it follows that

$$\mathbb{E}D_m \leq \sum_{\ell=1}^m \rho^{m-\ell} CK^{-2} \rho^{2\ell} \leq CK^{-2} \rho^{2m}.$$

Hence the contribution of  $\varepsilon^{(3)}$  in (33) is bounded by

$$\begin{aligned}
 \mathbb{E} \left| \sum_{m=1}^{n_1} \rho^{n_1-m} \varepsilon_m^{(3)} \right| &\leq C \sum_{m=1}^{n_1} \rho^{n_1-m} K^{-1/2} \sqrt{\mathbb{E}D_m} \\
 &\leq C \sum_{m=1}^{n_1} \rho^{n_1-m} K^{-1/2} K^{-1} \rho^m \leq CK^{-1/2} \log K.
 \end{aligned}$$

### 3.3.4 Contribution of $\varepsilon^{(2)}$

By assumption (a2.),

$$\begin{aligned}
 \mathbb{E}(\varepsilon_m^{(2)})^2 &= K^{-2} \mathbb{E} \sum_{j=1}^{Z_{m-1}} (\eta_{m,j} - \xi_{m,j}(\bar{Z}_{m-1}) - (\rho - m(\bar{Z}_{m-1})))^2 \\
 &\leq K^{-2} \mathbb{E} \sum_{j=1}^{Z_{m-1}} (\eta_{m,j} - \xi_{m,j}(\bar{Z}_{m-1}))^2 \stackrel{\dagger}{\leq} K^{-2} \mathbb{E} \sum_{j=1}^{Z_{m-1}} (m_2(0) - m_2(\bar{Z}_{m-1})) \\
 &\leq K^{-2} \mathbb{E} \sum_{j=1}^{Z_{m-1}} L \bar{Z}_{m-1} \leq CK^{-3} \rho^{2m}
 \end{aligned}$$

where † holds by (28). Hence  $\varepsilon^{(2)}$  contributes

$$\mathbb{E} \left| \sum_{m=1}^{n_1} \rho^{n_1-m} \varepsilon_m^{(2)} \right| \leq C \sum_{m=1}^{n_1} \rho^{n_1-m} K^{-3/2} \rho^m \leq C K^{-1/2} \log K.$$

### 3.3.5 Contribution of $\varepsilon^{(1)}$

The function  $g(x) := H'(x) - 1$  is continuously differentiable with  $g(0) = 0$  and thus Taylor’s approximation gives

$$\varepsilon_n^{(1)} = g(\rho \bar{Y}_{n-1}) \frac{1}{K} \sum_{j=1}^{Y_{n-1}} (\eta_{n,j} - \rho) = g'(\zeta_{n-1}(K)) \rho \bar{Y}_{n-1} \frac{1}{K} \sum_{j=1}^{Y_{n-1}} (\eta_{n,j} - \rho)$$

where  $\zeta_{n-1}(K)$  satisfies  $0 \leq \zeta_{n-1}(K) \leq \rho \bar{Y}_{n-1}$ . Here

$$\begin{aligned} & \mathbb{E} \left| \bar{Y}_{n-1} \frac{1}{K} \sum_{j=1}^{Y_{n-1}} (\eta_{n,j} - \rho) \right| \\ & \leq \left( \mathbb{E}(\bar{Y}_{n-1})^2 \right)^{1/2} \left( \mathbb{E} \left( \frac{1}{K} \sum_{j=1}^{Y_{n-1}} (\eta_{n,j} - \rho) \right)^2 \right)^{1/2} \\ & \leq (K^{-2} \rho^{2n})^{1/2} \left( K^{-2} \mathbb{E} Y_{n-1} \sigma^2(0) \right)^{1/2} \leq C K^{-2} \rho^{3/2n}. \end{aligned}$$

It follows that

$$\mathbb{E} \sum_{m=1}^{n_1} \rho^{n_1-m} \left| \bar{Y}_{m-1} \frac{1}{K} \sum_{j=1}^{Y_{m-1}} (\eta_{m,j} - \rho) \right| \leq \sum_{m=1}^{n_1} \rho^{n_1-m} C K^{-2} \rho^{3/2m} \leq C K^{-1/2}.$$

It is then argued as in Sect. 3.3.2 that

$$\sum_{m=1}^{n_1} \rho^{n_1-m} \varepsilon_m^{(1)} = O_{\mathbb{P}}(1) O_{\mathbb{P}}(K^{-1/2}) = O_{\mathbb{P}}(K^{-1/2}).$$

**Acknowledgements** The research was supported by ARC grant DP220100973.

**Funding** Open access funding provided by Hebrew University of Jerusalem.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Athreya KB, Ney PE (1972) Branching processes. Die Grundlehren der mathematischen Wissenschaften, Band 196, p 287. Springer, New York
- Baker J, Chigansky P, Jagers P, Klebaner FC (2020) On the establishment of a mutant. *J Math Biol* 80(6):1733–1757. <https://doi.org/10.1007/s00285-020-01478-x>
- Barbour AD, Chigansky P, Klebaner FC (2016) On the emergence of random initial conditions in fluid limits. *J Appl Probab* 53(4):1193–1205. <https://doi.org/10.1017/jpr.2016.74>
- Chigansky P, Jagers P, Klebaner FC (2018) What can be observed in real time PCR and when does it show? *J Math Biol* 76(3):679–695. <https://doi.org/10.1007/s00285-017-1154-1>
- Chigansky P, Jagers P, Klebaner FC (2019) Populations with interaction and environmental dependence: from few, (almost) independent, members into deterministic evolution of high densities. *Stoch Models* 35(2):108–118. <https://doi.org/10.1080/15326349.2019.1575755>
- Haccou P, Jagers P, Vatutin VA (2007) Branching processes: variation, growth, and extinction of populations. Cambridge studies in adaptive dynamics, vol. 5, p 316. Cambridge University Press, Cambridge, IIASA, Laxenburg
- Hamza K, Jagers P, Klebaner FC (2016) On the establishment, persistence, and inevitable extinction of populations. *J Math Biol* 72(4):797–820. <https://doi.org/10.1007/s00285-015-0903-2>
- Heyde CC (1970) A rate of convergence result for the super-critical Galton-Watson process. *J Appl Probab* 7(2):451–454. <https://doi.org/10.2307/3211980>
- Högnäs G (1997) On the quasi-stationary distribution of a stochastic Ricker model. *Stoch Process Appl* 70(2):243–263. [https://doi.org/10.1016/S0304-4149\(97\)00064-1](https://doi.org/10.1016/S0304-4149(97)00064-1)
- Högnäs G (2019) On the lifetime of a size-dependent branching process. *Stoch Models* 35(2):119–131. <https://doi.org/10.1080/15326349.2019.1578241>
- Jung B (2013) Exit times for multivariate autoregressive processes. *Stoch Process Appl* 123(8):3052–3063. <https://doi.org/10.1016/j.spa.2013.03.003>
- Kendall DG (1956) Deterministic and stochastic epidemics in closed populations. In: Proceedings of the third Berkeley symposium on mathematical statistics and probability, 1954–1955, vol. IV, pp 149–165. University of California Press, Berkeley-Los Angeles, CA
- Klebaner FC, Nerman O (1994) Autoregressive approximation in branching processes with a threshold. *Stoch Process Appl* 51(1):1–7. [https://doi.org/10.1016/0304-4149\(93\)00000-6](https://doi.org/10.1016/0304-4149(93)00000-6)
- Klebaner FC, Zeitouni O (1994) The exit problem for a class of density-dependent branching systems. *Ann Appl Probab* 4(4):1188–1205
- Klebaner FC, Lazar J, Zeitouni O (1998) On the quasi-stationary distribution for some randomly perturbed transformations of an interval. *Ann Appl Probab* 8(1):300–315. <https://doi.org/10.1214/aoap/1027961045>
- Metz JAJ (1978) The epidemic in a closed population with all susceptibles equally vulnerable; some results for large susceptible populations and small initial infections. *Acta Biotheor* 27(1):75–123. <https://doi.org/10.1007/BF00048405>
- Whittle P (1955) The outcome of a stochastic epidemic: a note on Bailey’s paper. *Biometrika* 42(1–2):116–122. <https://doi.org/10.1093/biomet/42.1-2.116>

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.