

## ASYMPTOTICS OF ITERATED BRANCHING PROCESSES

DIDIER PIAU,\* *Université Lyon 1*

### Abstract

Gawęł and Kimmel (1996) introduced and studied iterated Galton–Watson processes,  $(X_n)_{n \geq 0}$ , possibly with thinning, as models of the number of repeats of DNA triplets during some genetic disorders. Our main results are the following. If the process indeed involves some thinning then extinction,  $\{X_n \rightarrow 0\}$ , and explosion,  $\{X_n \rightarrow \infty\}$ , can have positive probability simultaneously. If the underlying (simple) Galton–Watson process is nondecreasing with mean  $m$  then, conditionally on explosion, the ratios  $(\log X_{n+1})/X_n$  converge to  $\log m$  almost surely. This simplifies the arguments of Gawęł and Kimmel, and confirms and extends a conjecture of Pakes (2003).

*Keywords:* Branching process; trinucleotide repeat expansion; genetic disease

2000 Mathematics Subject Classification: Primary 60J80  
Secondary 92D10

### 1. Introduction

Iterated Galton–Watson (IGW) processes are Markov chains  $(X_n)_{n \geq 0}$  with nonnegative integer values, whose distribution depends on two parameters: a real number  $u$  in the interval  $(0, 1]$ , called the thinning parameter, and a probability measure  $\mathbf{p} := (p_k)_{k \geq 0}$  on the nonnegative integers, called the reproduction law. The stochastic evolution of  $(X_n)_{n \geq 0}$  is as follows.

Let  $(Z^{(n)})_{n \geq 0}$  denote an auxiliary independent and identically distributed (i.i.d.) collection of Galton–Watson processes with common reproducing distribution  $\mathbf{p}$  and starting from  $Z_0^{(n)} := 1$ . Let  $(B_k^{(n)})_{n \geq 0, k \geq 1}$  denote an auxiliary i.i.d. collection of Bernoulli random variables with common distribution

$$P(B_k^{(n)} = 1) := u, \quad P(B_k^{(n)} = 0) := 1 - u.$$

Assume that  $X_0$ ,  $(Z^{(n)})_n$ , and  $(B_k^{(n)})_{n,k}$  are independent. For every positive  $x$ , let  $S_x^{(n)}$  denote the population produced by the  $n$ th Galton–Watson process  $Z^{(n)}$  up to level  $x$ , that is,

$$S_x^{(n)} := \sum_{y=1}^x Z_y^{(n)}.$$

If  $X_n = 0$  then  $X_{n+1} := 0$ . If  $X_n = x \geq 1$  then

$$X_{n+1} := B_1^{(n)} + \cdots + B_y^{(n)}, \quad \text{where } y := S_x^{(n)}.$$

In other words, the distribution of  $X_{n+1}$  conditionally on  $\{X_n = x\}$  and on  $S_x^{(n)}$  is binomial with parameters  $(S_x^{(n)}, u)$ .

Received 28 May 2004; revision received 11 July 2008.

\* Current address: Institut Fourier UMR 5582, Université Joseph Fourier Grenoble 1, 100 rue des Maths, BP 74, 38402 Saint Martin d’Hères, France. Email address: didier.piau@ujf-grenoble.fr

Gaweł and Kimmel [1] introduced IGW processes to model the explosive growth of the number of repeats of DNA triplets in specific regions of the genome during heritable disorders such as fragile X-syndrome. Here,  $1 + X_n$  models the length of a linear chain of DNA repeats after  $n$  replications, and what we know of the molecular mechanism of replication suggests describing the evolution of  $(X_n)$  as above. (We use notation that is mostly compatible with that of Gaweł and Kimmel, with the exception that these authors defined the IGW process as  $(X'_n)_{n \geq 0}$  with  $X'_n := 1 + X_n$ .) Indeed, simulations in Gaweł and Kimmel when  $p_0 = 0$  and  $u < 1$  suggested that IGW processes either die out or grow extremely fast after a period of relative quiescence. This is precisely the behaviour of the number of repeats of DNA triplets in successive generations of carriers of these genetic disorders. See also the book by Kimmel and Axelrod [2, Chapter 3.7, pp. 51–56], which repeats the analysis of Gaweł and Kimmel.

We define the explosion  $F$  and death  $D$  of the IGW process as the events

$$F := \{X_n \rightarrow \infty\}, \quad D := \{X_n \rightarrow 0\} = \{X_n = 0 \text{ for large enough } n\}.$$

Thus,  $F$  and  $D$  are mutually exclusive. (In the context of genetic disorders, the death  $D$  of the process corresponds to the extinction of the diseased gene lineage, and the explosion  $F$  corresponds to unlimited expansion of DNA triplet repeats, interrupted by the death of the mutation carrier.) Gaweł and Kimmel showed that  $P(D) = 1$  as soon as  $p_0 > 0$ , and that  $P(F) = 1$  when  $p_0 = 0$ ,  $p_1 < 1$ , and  $u = 1$ . Later on, their arguments were simplified by Pakes [4]. Pakes also conjectured that, when  $p_0 = 0$ ,  $p_1 < 1$ , and  $u = 1$ ,  $(\log X_{n+1})/X_n$  converges almost surely to  $\log m(\mathbf{p})$ , where  $m(\mathbf{p}) > 1$  denotes the mean of  $\mathbf{p}$ , that is,

$$m(\mathbf{p}) := \sum_{k \geq 0} kp_k.$$

In this paper we determine the asymptotic behaviour of every IGW process. We confirm the conjecture of Pakes and extend it to IGW processes with thinning, that is, to the case where  $u < 1$ , and we refine partial results of Gaweł and Kimmel which are not recalled above.

The two propositions below collect some simple facts about the mean behaviour of IGW processes and about their almost-sure behaviour in some degenerate cases. These are mainly due to Gaweł and Kimmel, or to Pakes. Proposition 2 for instance is due to Pakes when  $m(\mathbf{p}) > 1$  (note that if  $p_0 = 0$  and  $p_1 < 1$ , then  $m(\mathbf{p}) > 1$ ). We write  $P_x$  and  $E_x$  for the probability  $P$  and the expectation  $E$ , conditionally on  $X_0 = x$ .

**Proposition 1.** *The mean behaviour of the IGW process is as follows.*

- (i) *If  $m(\mathbf{p}) > 1$  then  $E_x(X_n) \rightarrow \infty$  for every initial population  $x \geq 1$ .*
- (ii) *If  $m(\mathbf{p}) < 1$ , or if  $m(\mathbf{p}) = 1$  and  $u < 1$ , then  $E_x(X_n) \rightarrow 0$  for every initial population  $x \geq 0$ .*
- (iii) *If  $m(\mathbf{p}) = u = 1$  then  $E_x(X_n) = x$  for every initial population  $x \geq 0$ .*

**Proposition 2.** *The almost-sure behaviour of the IGW process in some degenerate cases is as follows.*

- (i) *If  $p_0 > 0$  then  $P_x(D) = 1$  for every initial population  $x \geq 0$ .*
- (ii) *If  $p_0 = 0$ ,  $p_1 < 1$ , and  $u = 1$ , then  $P_x(F) = 1$  for every initial population  $x \geq 1$ .*

We stress that the hypotheses of Propositions 1(i) and 2(i) are compatible; hence, we can have  $E_x(X_n) \rightarrow \infty$  and  $P_x(D) = 1$  simultaneously for every initial population  $x \geq 1$ . This contrasts with the behaviour of usual Galton–Watson processes.

We now turn to the nondegenerate case, which Gawel and Kimmel, and Pakes omitted.

**Theorem 1.** *Assume that  $p_0 = 0$  and  $p_1 < 1$ .*

- (i) *For every initial population  $x \geq 0$ ,  $P_x(D) + P_x(F) = 1$ .*
- (ii) *Assume further that  $u < 1$ . Then, for every initial population  $x \geq 1$ ,  $P_x(D)$  and  $P_x(F)$  are both positive, and  $P_x(D) \leq P_1(D)^x$ . Thus,  $P_x(F) \rightarrow 1$  when  $x \rightarrow \infty$ .*

We stress that Theorem 1(ii) does not state that  $P_x(D)$  decreases geometrically when the initial population  $x \rightarrow +\infty$ . In fact, we can show in this case that, when  $x \rightarrow +\infty$ ,

$$\log \frac{1}{P_x(D)} \gg x.$$

We now state our main result.

**Theorem 2.** *Assume that  $p_0 = 0$ . For every initial population  $x \geq 1$ , conditionally on the explosion  $F$ , the random variable  $(\log X_{n+1})/X_n$  converges almost surely to  $\log m(\mathbf{p})$ . In particular, conditionally on  $F$ ,  $X_{n+1}/X_n$  converges almost surely to  $\infty$ .*

As regards the process conditional on the death  $D$ , we could try to show that, if suitably renormalized, the death time, that is, the first hitting time of 0, converges in distribution, conditionally on  $D$  and when the starting point of the process goes to  $\infty$ . As noted by a referee, even a conjecture about this presumed limit distribution would be useful, and could be compared to some simulations. We do not pursue this avenue here.

The rest of the paper is organised as follows. In Section 2 we (re)prove Propositions 1 and 2. In Section 3 we deal with the easy parts of Theorem 1, that is, part (i) and the fact that the probability of death is positive and at most geometric. In Section 4 we provide explicit upper bounds of the probability of death in some specific cases. Section 5 exposes a strategy of proof for the study of the explosion case. In Section 6 we apply this strategy, first to the proof of Theorem 1(ii), that is, of the fact that the probability of explosion is positive, and then to the proof of Theorem 2.

## 2. Proofs of Propositions 1 and 2

*Proof of Proposition 1.* From the construction of the IGW process, we see that, for every initial population  $x \geq 0$ ,  $E_x(X_1) = \chi(x)$ , where the sequence  $\chi$  is defined by  $\chi(0) := 0$  and, for every integer  $x \geq 1$ ,

$$\chi(x) := u(m + \dots + m^x).$$

When  $m > 1$ ,  $\chi$  can be extended to a convex function on  $[0, +\infty)$ , which we still call  $\chi$ . (This step of the proof would be false for  $m < 1$ .) Thus, for every initial population  $x \geq 0$  and every time  $n \geq 0$ ,

$$E_x(X_{n+1}) \geq \chi(E_x(X_n)).$$

Choose  $x_0 \geq 1$  large enough such that  $\chi(x_0) \geq 2x_0$ . Then,  $E_{x_0}(X_n) \geq x_0 2^n$  for every  $n \geq 0$ . (This step of the proof uses the fact that  $\chi$  is nondecreasing.) For smaller values of  $x$ , for instance for  $x = 1$ , consider the event that  $X_n \geq x_0$ . For large enough  $n$ ,  $n = n_0$  say, this event

has positive probability, say  $r$ , with respect to  $P_1$ . Finally, (i) holds since, for every  $n \geq n_0$  and every  $x \geq 1$ ,

$$E_x(X_n) \geq E_1(X_n) \geq P(X_{n_0} \geq x_0) E_{x_0}(X_{n-n_0}) \geq rx_0 2^{n-n_0}.$$

When  $m = 1$ ,  $E_x(X_1) = ux$ ; thus, the  $m = 1$  cases in (ii) and (iii) are obvious. Finally, when  $m < 1$ ,  $E_x(X_1) \leq umx$  and  $um < 1$ ; thus,  $E_x(X_n) \rightarrow 0$ . This completes the proof of (ii).

*Proof of Proposition 2.* The proof of (ii) is direct since in this case  $X_{n+1} \geq X_n$  almost surely. As regards (i), if  $p_0 > 0$ , for each  $n \geq 0$ , the event that  $Z_1^{(n)} = 0$  has positive probability  $p_0$  and these events are independent. One of these events is almost surely realized, say  $Z_1^{(n)} = 0$ . Then  $S_x^{(n)} = 0$  for every integer  $x \geq 1$  and  $X_{n+1} = 0$ , which proves (i). The argument also shows that

$$P_x(X_n \geq 1) \leq (1 - p_0)^n$$

for every time  $n$  and initial population  $x$ . In other words, the time to absorption is uniformly stochastically dominated by the geometric distribution of parameter  $p_0$ .

### 3. Elementary parts of Theorem 1

#### 3.1. Almost surely, death or explosion

This is straightforward, and analogous to the usual Galton–Watson case. We have to check that 0 is the only nontransient state of the Markov chain  $(X_n)$ . If  $p_0 = 0$  and  $u = 1$ , this is true; see the proof of Proposition 1(i). If  $p_0 = 0$  and  $u < 1$ , the return to  $x \geq 1$ , starting from  $x$ , assumes that the first step is not to 0. As such, the probability of this event is at most  $1 - P_x(X_1 = 0) < 1$ ; see the next subsection.

#### 3.2. Probability of death, not zero

When  $X_0 = x$ ,  $X_1 = 0$  if and only if the thinning phase kills each and every  $S_x^{(0)}$  individual. Hence, for every  $x \geq 0$ ,

$$P_x(D) \geq P_x(X_1 = 0) = E((1 - u)^{S_x}),$$

which is positive.

#### 3.3. Probability of death, at most geometric

Assume that  $p_0 = 0$ , and fix  $x \geq 1$  and  $y \geq 1$ . The fundamental branching property of the Galton–Watson process  $Z^{(n)}$  means that  $Z_{x+y}^{(n)}$  is the sum of  $Z_x^{(n)} \geq 1$  random variables distributed like  $Z_y^{(n)}$ . Hence,  $S_{x+y}^{(n)}$  is stochastically greater than the sum of  $S_x^{(n)}$  and of an independent copy of  $S_y^{(n)}$ . After thinning, this shows that  $X_1$  under  $P_{x+y}$  is stochastically greater than the sum of  $X_1$  under  $P_x$  and of an independent copy of  $X_1$  under  $P_y$ . By recursion over  $n \geq 1$ , the same assertion holds when we replace  $X_1$  by  $X_n$ . Hence,

$$P_{x+y}(X_n = 0) \leq P_x(X_n = 0) P_y(X_n = 0).$$

This implies that  $P_x(D) \leq P_1(D)^x$  for every  $x \geq 0$ , an assertion of Theorem 1(ii).

Note that the important step here is to prove that  $P_1(D) < 1$ . We do this in Section 4 in some specific cases, the general case is given in Section 6.1.

### 4. Upper bounds of the probability of death

When  $p_0 = 0$ , in the special case  $mu > 1$ , we can explicitly bound  $P_1(D)$  by some  $q < 1$ , using the generating function of the distribution  $(p_k)_k$ . Thanks to Section 3.3, this proves that  $P_x(D) \leq q^x$ .

#### 4.1. Special case

We can often obtain an upper bound of  $q := P_1(D)$  at small cost, as follows. Since the sequence  $(P_x(D))_{x \geq 0}$  is submultiplicative,

$$q = E_1(P_{X_1}(D)) \leq E_1(q^{X_1}).$$

The generating function of  $X_1$  is  $g(s) := E_1(s^{X_1}) = f(1 - u + us)$ , with

$$f(s) := \sum_{k \geq 1} p_k s^k.$$

Thus,  $q \leq g(q)$ , and it is not hard to see that  $q$  is at most the smallest root of the equation  $s = g(s)$ . In the supercritical case,  $g'(1^-) = mu > 1$ ,  $s \leq g(s)$  for  $s = 1$  and  $s \leq q_0$ , where  $q_0$  in  $(0, 1)$  is defined by  $q_0 = g(q_0)$ . Finally,  $q \leq q_0$ . In particular,  $P_x(D) < 1$ .

#### 4.2. Binary case

Assume that the Galton–Watson process describes a binary replication with efficiency  $\lambda$ . Thus,  $f(s) := (1 - \lambda)s + \lambda s^2$  and  $m = 1 + \lambda$ . Elementary computations then yield the following. If  $u > 1/m$ ,  $q \leq q(\lambda, u)$  with

$$q(\lambda, u) := \frac{(1 - u)(1 - \lambda u)}{\lambda u^2}.$$

Note that  $0 < q(\lambda, u) < 1$ , except when  $u = 1$ , and then  $q = q(\lambda, 1) = 0$ , and when  $u = 1/m$ , and then  $q(\lambda, u) = 1$  but  $q < 1$ .

### 5. A general strategy

Assume that  $p_0 = 0$ . We first explain our strategy to study the probability of explosion. We fix an integer-valued sequence  $(\varphi(x))_x$  such that  $\varphi(x) \geq x + 1$  for every  $x$ , and an integer-valued sequence  $(\psi(x))_x$ .

**Definition 1.** For every  $x \geq 1$ , let  $\eta_x$  denote a random variable of binomial distribution of parameters  $(\psi(x), u)$ , and let  $S_x$  be a random variable distributed like every  $S_x^{(n)}$ . Introduce the probabilities  $A(x)$  and  $B(x)$ , defined as

$$A(x) := P(S_x \leq \psi(x)), \quad B(x) := P(\eta_x < \varphi(x)),$$

and the event  $C$ , defined as

$$C := \{\text{for all } n \geq 0, X_{n+1} \geq \varphi(X_n)\}.$$

To show that  $P_x(C) > 0$  for every  $x \geq 1$ , we start from the inequality

$$P_x(X_1 < \varphi(x)) \leq A(x) + B(x).$$

The Markov inequality and the fact that  $S_x \geq Z_x$  imply that

$$A(x) \leq P(Z_x \leq \psi(x)) \leq \psi(x) E\left(\frac{1}{Z_x}\right) \leq \psi(x)e^{-\kappa x},$$

with a positive  $\kappa$  that depends only on  $p$ ; see, for instance, Proposition A.2 of [5].

We now estimate  $B(x)$ . The Chebyshev inequality for  $\eta_x$  and the fact that  $\eta_x$  is distributed like the sum of  $\psi(x)$  i.i.d. copies of a Bernoulli random variable  $B$  of parameter  $u$  imply that

$$B(x) \leq e^{\varphi(x)} E(e^{-B})^{\psi(x)} = e^{\varphi(x)} \left(1 - u + \frac{u}{e}\right)^{\psi(x)}.$$

Computing the first and second derivatives of the function

$$u \mapsto e^{-u^2} - 1 + u - \frac{u}{e},$$

we see that it is positive on  $(0, 1)$ ; hence,  $1 - u + u/e \leq e^{-u^2}$  and

$$B(x) \leq \exp(\varphi(x) - u^2\psi(x)).$$

**Definition 2.** We say that  $(\gamma(x))_{x \geq 1}$  is admissible with respect to  $(\varphi(x))_x$  and  $(\psi(x))_x$  if the sequence  $(\gamma(x))_{x \geq 1}$  is nonincreasing, summable, and such that, for every  $x \geq 1$ ,

$$\psi(x)e^{-\kappa x} + \exp(\varphi(x) - u^2\psi(x)) \leq \gamma(x) < 1.$$

Let  $(\gamma(x))_x$  denote an admissible sequence. For  $x \geq 1$ , recursively define  $(\gamma_k(x))_{k \geq 0}$  by

$$\gamma_0(x) := \gamma(x), \quad \gamma_{k+1}(x) := \gamma_k(\varphi(x)) \quad (k \geq 0).$$

Then  $P_x(X_1 < \varphi(x)) \leq \gamma(x)$  for every  $x \geq 1$ . Conditioning successively on the values of  $X_n$  and iterating the above yields

$$P_x(C) \geq \prod_{k \geq 0} (1 - \gamma_k(x)).$$

The iteration uses both the fact that  $\varphi(x) \geq x$  and the fact that  $(\gamma(x))_x$  is nonincreasing.

Since  $\varphi(x) \geq x + 1$  and  $(\gamma(x))_x$  is nonincreasing,  $\gamma_k(x) \leq \gamma(k + x)$ . Hence,

$$P_x(C) \geq \alpha(x), \quad \alpha(x) := \prod_{y \geq x} (1 - \gamma(y)).$$

If an admissible sequence exists,  $\alpha(x) > 0$ . In particular,  $P_x(C) > 0$  for every  $x \geq 1$ . Furthermore,  $C \subset F$ ; hence,  $P_x(F) \geq P_x(C)$  and  $P_x(F) > 0$  as well. Finally, the sequence  $(\alpha(x))_x$  increases to 1 when  $x \rightarrow +\infty$ ; hence,  $P_x(F)$  and  $P_x(C)$  both converge to 1 when  $x \rightarrow +\infty$ .

We need a slight refinement of this, stated as Lemma 1, below.

**Lemma 1.** *If an admissible sequence exists then  $\liminf X_n/n \geq 1$  almost surely, conditionally on  $F$ .*

*Proof.* Consider the event  $H := \{\liminf X_n/n \geq 1\}$ . Then  $C \subset H = H_k$  for every nonnegative integer  $k$ , with

$$H_k := \left\{ \liminf \frac{X_{n+k}}{n} \geq 1 \right\}.$$

Furthermore, for every  $x \geq 1$ , on the event  $\{X_k \geq 1\}$ ,

$$P_x(H_k \mid X_k) = P_{X_k}(H) \geq P_{X_k}(C) \geq \alpha(X_k).$$

Since  $F \subset \{X_k \geq 1\}$ , this shows that

$$P_x(H) \geq E_x(\alpha(X_k) \mathbf{1}_F).$$

When  $k \rightarrow +\infty$ ,  $X_k \rightarrow +\infty$ ; hence,  $\alpha(X_k) \rightarrow 1$  on  $F$ , and  $P_x(H) \geq P_x(F)$ . This completes the proof.

### 6. Some applications of the strategy

#### 6.1. Probability of explosion, not zero

We first apply the strategy explained in Section 5 with  $\varphi(x) := x + 1$  and  $\psi(x) := x^2$ . We can choose  $\gamma(x) < 1$  for every  $x \geq 1$  and such that  $\gamma(x) \leq e^{-cx}$  when  $x \rightarrow \infty$ , with a positive  $c$ ; hence, admissible sequences exist and  $P_x(F) > 0$  for every  $x \geq 1$ . This is the missing part of Theorem 1(ii), completing the proof of Theorem 1.

#### 6.2. Lower bound in Theorem 2

Our second application of the strategy explained in Section 5 is more involved. We choose  $\mu$  in the interval  $(1, m)$  and an integer sequence  $(\varphi(x))_x$  such that  $\mu^{-x}\varphi(x) \rightarrow 1$  when  $x \rightarrow +\infty$ . Then we choose  $\nu$  in the interval  $(\mu, m)$  and an integer sequence  $(\psi(x))_x$  such that  $\nu^{-x}\psi(x) \rightarrow 1$  when  $x \rightarrow +\infty$ . Since  $\nu > \mu$ , the contributions  $(B(x))_x$  are summable.

As regards the contributions  $(A(x))_x$ , we use standard estimates of the harmonic moments of  $Z_x$ . Choose a positive real number  $r$  such that  $p_1 m^r < 1$ , this is possible as soon as  $m$  is finite and  $p_1 < 1$ . Results in [3] allow us to control the behaviour of  $E(1/(Z_x)^r)$  as follows.

To keep things simple, we choose a real number  $\varrho$  in  $(\nu, m)$  and we use the following easy consequence of the results of [3]. There exists a finite constant  $c_0$  such that  $E(1/(Z_x)^r) \leq c_0 \varrho^{-rx}$  for every  $x \geq 1$ .

The Markov inequality for  $(Z_x)^r$  yields, for large enough  $x$ ,

$$A(x) \leq P(Z_x \leq \psi(x)) \leq \psi(x)^r E\left(\frac{1}{(Z_x)^r}\right) \leq c_0 \left(\frac{\nu}{\varrho}\right)^{rx}.$$

Finally, the sequence  $(\gamma(x))_x$  is allowed to decrease geometrically, and, hence, to be summable.

Since admissible sequences exist,  $P_x(C) > 0$ . Furthermore, introduce

$$G := \{\liminf Y_n \geq \log \mu\}, \quad Y_n := \frac{\log X_{n+1}}{X_n}.$$

Since  $C \subset G$  and  $G$  is asymptotic,  $G$  is almost sure on  $F$ . Finally, conditionally on  $F$ ,

$$\liminf Y_n \geq \log m \quad \text{almost surely.}$$

### 6.3. Upper bound in Theorem 2

As regards the other side of the equality in Theorem 2, fix  $\mu > m$ . We use the simple fact that if  $Y_n \geq \log \mu$  then, conditionally on  $X_n = x$ , we have  $S_x^{(n)} \geq X_{n+1} \geq \mu^x$ . Furthermore,

$$P(S_x \geq \mu^x) \leq \mu^{-x} E(S_x) \leq \frac{(m/\mu)^x m}{m-1},$$

which is summable. By Lemma 1,  $\liminf X_n/n \geq 1$  almost surely, conditionally on  $F$ . This implies that the events  $\{Y_n \geq \log \mu\}$  are realized at most for a finite number of values of  $n$ , conditionally on  $F$  or not. Thus,  $\limsup Y_n \leq \log \mu$  almost surely. This concludes the proof of Theorem 2.

### References

- [1] GAWEL, B. AND KIMMEL, M. (1996). The iterated Galton–Watson process. *J. Appl. Prob.* **33**, 949–959.
- [2] KIMMEL, M. AND AXELROD, D. E. (2002). *Branching Processes in Biology* (Interdisciplinary Appl. Math. **19**). Springer, New York.
- [3] NEY, P. E. AND VIDYASHANKAR, A. N. (2003). Harmonic moments and large deviation rates for supercritical branching processes. *Ann. Appl. Prob.* **13**, 475–489.
- [4] PAKES, A. G. (2003). Biological applications of branching processes. In *Stochastic Processes: Modelling and Simulation* (Handbook Statist. **21**), North-Holland, Amsterdam, pp. 693–773.
- [5] PIAU, D. (2004). Immortal branching Markov processes: averaging properties and PCR applications. *Ann. Prob.* **32**, 337–364.