The Annals of Probability
2004, Vol. 32, No. 1A, 337–364

IMMORTAL BRANCHING MARKOV PROCESSES: AVERAGING PROPERTIES AND PCR APPLICATIONS

BY DIDIER PIAU

Université Lyon-I

The immortal branching Markov process (iBMP) is a modification of the usual branching model, in which each particle of generation n is counted, in addition to its offspring, as a member of generation $n+1$, its state being unchanged. When the number of offspring is Bernoulli, iBMP accounts, for instance, for the variability of the biological sequences that are produced by polymerase chain reactions (PCRs). This variability is due to the mutations and to the incomplete replications that affect the PCR. Estimators of PCR mutation rate and efficiency have been proposed that are based, in particular, on the mean empirical law η_n of the mutations of a sequence. Unfortunately, η_n is not analytically tractable. However, the infinite-population limit η_n^* of η_n is easily characterized in the two following, biologically relevant, cases. The Markovian kernel describes a homogeneous random walk, either on the integers or on some finite Cartesian product of a finite set. In the PCR context, this corresponds to infinite or finite targets, respectively. In this paper, we provide bounds of the discrepancy between η_n and η_n^* in these two cases. As a consequence, iBMP exhibits a strong averaging effect, even for surprisingly small starting populations. The bounds are explicit functions of the offspring law, the Markovian kernel, the number of steps n, the size of the initial population and, in the finite-target case, the size of the target. They concern every moment and, what might be less expected, the histogram itself. In the finite-target case, some of the bounds undergo a phase transition at an explicit value of the mutation rate per site and per cycle. We use precise estimates of the harmonic means of classical nondecreasing branching processes, whose proofs are included in the Appendix.

1. Introduction.

1.1. *iBMP.* We consider the following random process. At time n, there exists a finite number of particles x. The state of particle x is $s(x)$. Each particle x gives birth to a random number Z_x of offspring y by the usual i.i.d. procedure of branching processes. Hence, each state $s(y)$ depends on $s(x)$ along Markovian transitions. The population at time $n+1$ is made of these offspring y and of the population of particles x at time n. We call this an immortal branching Markov process (iBMP), since a particle lives and gives birth at every epoch posterior to its

---

Received June 2001; revised February 2002.

AMS 2000 subject classifications. Primary 60J80, 92D20; secondary 60K37, 60J85.

Key words and phrases. Markov branching process, Galton–Watson process, mean field approximation, polymerase chain reaction (PCR), error-prone PCR, mutation rate, DNA amplification, estimation.

creation. This paper is devoted to the study of this new type of process. The iBMP model has proved useful for the study of a biochemical reaction called PCR, see Section 1.3, and we believe it will do so for other types of studies as well.

The overall population S_n at time n of the iBMP performs a classical branching process of reproducing mechanism $L_x := Z_x + 1 \geq 1$. Thus, $(S_n)_{n \geq 0}$ is a nondecreasing supercritical branching process, except in the trivial case $Z_x = 0$ a.s., which we exclude. Thus, iBMP may be viewed as a branching Markov process (BMP), in which, for each particle, the state of one offspring does not behave like those of the others. We call independent BMP the model where the states of all the offspring of x are i.i.d. conditional on $s(x)$.

We study the mean empirical law η_n of an iBMP at time n. If the states $s(x)$ are nonnegative real numbers, the (Laplace) transform of η_n is, for $|u| \leq 1$,

$$L_n(u) := \mathbb{E}\left(S_n^{-1} \sum_{x=1}^{S_n} u^{s(x)} \right).$$

Here, the sum from $x = 1$ to S_n is a shorthand for the sum along all the particles of the population at time n. For independent BMP, the description of L_n is trivial. To see this in a simple case, assume that the random variables $s(y) - s(x)$, for $x \to y$, are i.i.d. copies of ξ. Denote by $g(u) := \mathbb{E}(u^{\xi})$ the transform of ξ, and by Θ the expectation of the reproducing law. Then

$$\Theta^{-n}\, g(u)^{-n} \sum_{x=1}^{S_n} u^{s(x)}$$

is a well-known martingale, which is a basic tool for the study of the limit of S_n/Θ^n; see, for instance, Biggins (1977) or Lyons (1997). Hence, the expectation of this martingale at time n is entirely determined by the initial population and states. Joffe (1993) introduced the modification

$$S_n^{-1}\, g(u)^{-n} \sum_{x=1}^{S_n} u^{s(x)},$$

and he showed that this is also a martingale. Thus, for independent BMP,

$$L_n(u) = g(u)^n\, S_0^{-1} \sum_{x=1}^{S_0} u^{s(x)}.$$

In other words, the states of the particles and the size of the population are uncorrelated. For iBMP, the first martingale above is still available, but not the variant by Joffe.

1.2. *Mean field limits.* The $S_0 \to \infty$ limit η_n^* of the mean empirical law η_n of an iBMP is easy to describe when the population at time 0 is homogeneous, that is, when $s(x) = s_0$ for all x at time 0.

DEFINITION 1. Let σ_k denote the result of k steps of the Markovian kernel that describes the evolution of the states, starting from $\sigma_0 := s_0$. The type $t(x)$ of a particle x is $t(x) := k$ if x is obtained through exactly k offspring steps. Let $S_n(k)$ denote the number of particles at time n of type k.

The random variables $S_n(k)$ are related by branching relations of the form

$$S_{n+1}(k+1) = S_n(k+1) + \sum_{x=1}^{S_n(k)} Z_x. \tag{1}$$

Here, the sum from $x = 1$ to $x = S_n(k)$ is really the sum along the particles at time n of type k. Furthermore,

$$L_n(u) = \sum_{k=0}^{n} \zeta_n(k)\mathbb{E}(u^{\sigma_k}), \qquad \zeta_n(k) := \mathbb{E}\left(\frac{S_n(k)}{S_n}\right). \tag{2}$$

This indicates once again why $L_n(u)$ is not tractable in general, since the computation of the coefficients $\zeta_n(k)$ in (2), starting from (1), is not an easy task. However, when $S_0 \to \infty$, the overall population is the sum of the independent populations that are generated by each starting particle. By the standard law of large numbers, S_n/S_0 converges almost surely to

$$\mathbb{E}_1(S_n) = (1+\lambda)^n, \qquad \lambda := \mathbb{E}(Z_x),$$

when $S_0 \to \infty$. Here, $\mathbb{E}_1$ is the expectation when $S_0 = 1$. From the same law of large numbers, $S_n(k)/S_0$ converges almost surely to

$$\mathbb{E}_1(S_n(k)) = \binom{n}{k}\lambda^k.$$

Thus, when $S_0 \to \infty$, L_n converges to its mean field version L_n^*, with

$$L_n^*(u) = \sum_{k=0}^{n} \binom{n}{k}\lambda^k (1+\lambda)^{-n}\mathbb{E}(u^{\sigma_k}).$$

In other words, the empirical law η_n^* of the mean field iBMP is the law of the Markov chain σ_τ, when considered at a random time τ which is independent of the Markov chain $(\sigma_k)_k$ and whose law is binomial of parameters (n, q), with

$$q := \frac{\lambda}{1+\lambda}, \qquad p := 1 - q, \qquad \lambda := \mathbb{E}(Z). \tag{3}$$

Let t_n denote the mean empirical type of a particle at time n, and t_n^* the mean field version of t_n. Let ζ_n and ζ_n^* denote the laws of t_n and t_n^*. Thus, ζ_n^* is the binomial law of parameters (n, q), and

$$\zeta_n(k) := \mathbb{P}(t_n = k) := \mathbb{E}\left(\frac{S_n(k)}{S_n}\right). \tag{4}$$

Below, the random variables M_n and M_n^* follow the laws η_n and η_n^*. For any bounded function F,

$$\mathbb{E}(F(M_n)) = \mathbb{E}\left(S_n^{-1}\sum_{x=1}^{S_n} F(s(x))\right).$$

1.3. *PCR randomness.* The iBMP can model any reproduction process where particles give birth at every epoch after their creation, without aging. When Z_x is Bernoulli, that is, almost surely 0 or 1, iBMP models the variability of genetic sequences produced by polymerase chain reactions (PCRs). Recall that the PCR is a powerful method of production of large amounts of DNA from a small number of template molecules. Quoting from Weiss and von Haeseler (1997), it "has revolutionized work with genetic material." Apart from the sequencing of DNA, its applications include cancer research, inherited disease diagnosis, forensic medicine and ancient DNA. Like any biochemical mechanism, the PCR is imperfect, in the sense that some replications are incomplete and that the replicated molecules present some mutations. These errors can have noticeable consequences; see Jacobs, Tscholl, Sek, Pfreundschuh, Daus and Trümper (1999) for an example related to the early diagnosis of cancer.

During each PCR cycle, each molecule is replicated with probability $\lambda \in (0; 1)$, and if the replication is successful, some mutations occur that segregate the copy from the original. Thus, the whole PCR can be modeled by an iBMP where Z_x follows the Bernoulli law

$$\lambda\delta_1 + (1-\lambda)\delta_0.$$

Two Markovian kernels are relevant. First, in the additive model, $s(x)$ is the total number of mutations that occurred on x and on its ancestors. Thus, $(\sigma_n)_{n\geq 0}$ in Definition 1 is a random walk on the integer line with i.i.d. nonnegative increments. This describes well the state of x insofar as the number G of sites in the replicated molecules is large and the mutation rate per site μ is so low that mutations rarely occur twice at a same site. The law of the increments of σ_n is then approximately Poisson of parameter

$$\widetilde{\mu} := G\mu.$$

The full model, however, is a product model that takes into account the facts that G is finite and that several mutations can occur at same site. Then, x is represented by an element of $\mathcal{A}^G$, where $\mathcal{A}$ is the encoding alphabet of DNA or, more generally, a finite set of size $N \geq 2$. That is, x is identified by

$$\{\ell_i(x); 1 \leq i \leq G\}, \qquad \ell_i(x) \in \mathcal{A}.$$

The sites evolve independently and the mutation rate from ℓ to ℓ' at site i is independent of the site i and of the letters (ℓ, ℓ'). We are well aware that these two assumptions may be too simplifying in actual biological settings. When the duplication $x \to y$ occurs, each letter $\ell_i(x)$ is modified to $\ell_i(y)$, independently of the other letters, according to a Markov chain of mutation rate $\mu \in (0; 1)$. That is, the letter ℓ remains ℓ itself with probability $1 - \mu$, and ℓ becomes ℓ' with probability $\mu/(N-1)$ for each letter $\ell' \neq \ell$. Assuming, without loss of generality, that every letter of every initial sequence is 0, the state of x is the number of "false" letters in x, that is,

$$s(x) := \sum_{i=1}^{G} \mathbf{1}_{\{\ell_i(x) \neq 0\}}.$$

1.4. *Mean field PCR.* Sun (1995) and Weiss and von Haeseler (1995), building on earlier works that include Krawczak, Reiss, Schmidtke and Rosler (1989), Reiss, Krawczak, Schlösser, Wagner and Cooper (1990), Hayashi (1990) and Maruyama (1990), use the Poisson additive model to study statistical properties of PCR sequences. Their approach is adapted to the product model by Wang, Zhao, Cheng and Sun (2000). In practice, these authors solve the mean field cases, as explained in Section 1.2, and they propose estimators of the parameters λ and $\widetilde{\mu}$ or μ. In particular, Sun (1995) bases an estimator of $\widetilde{\mu}$ on η_n and he proves that the estimator is unbiased in the mean field case. Similar results are obtained for the mean empirical law of the Hamming distance between two sequences.

In Piau (2001), we provide explicit bounds of the discrepancy between η_n and η_n^*, as regards the mean and the histogram, for the Poisson additive model. The same program is realized for the product model in Piau (2002), as regards every moment and the histogram. Consequences of these results are, first, that Sun's estimator systematically underestimates $\widetilde{\mu}$ when the initial population is finite, a fact that seems to have been overlooked; see, for instance, Brunnert, Müller and Urfer (2000) and Wang, Zhao, Cheng and Sun (2000). More important, the mean field approximation is excellent, even for relatively small initial populations S_0. For instance, a condition mentioned by Sun (1995), namely that

$$S_0 \gg (1 + \lambda)^{2n},$$

is simply not relevant. This validates, in a way, the whole approach since, as mentioned by Sun himself after Saiki et al. (1988), the efficiency λ can only be considered as constant when S_0 is moderate, that is, in the so-called exponential region. For greater S_0 values, λ progressively decreases to 0. For more on PCR, on the biological relevance of the model, on the values of the parameters, on error-prone PCR and on the relevance of the additive Poisson model versus the product model, we refer to the above-mentioned papers and to the references in Piau (2001, 2002).

1.5. *Results.* We concentrate below on the accuracy of the approximation of η_n by η_n^*. In contrast with asymptotic mathematical results, the main task here is to obtain rigorous nonasymptotic bounds, which are often the only relevant ones for practical purposes. For instance, the considered objects have quite simple limits when $n \to \infty$.

We study the additive model and the product model. Basically, we prove that a strong averaging effect occurs, as regards the difference of the moments and the distance in total variation. To get tractable bounds of the various discrepancies between η_n and η_n^*, our main strategy is to look for exact recursion relations between η_n and η_{n+1}, such that η_n^* solves the main term of the recursion and such that the error terms can be explicitly bounded. Sun (1995) computes various distributions related to the most recent common ancestor of two particles in the mean field case. Weiss and von Haeseler (1995) use Sun's results; Weiss and von Haeseler (1997) model the genealogy of PCR samples by a modified coalescent process. Thus, our method is entirely different.

In the additive model, we prove approximations of the mean and variance of M_n of order $1/S_0$, the mean field values being of order n. This result generalizes to every offspring law and every law of the increments of $s(x)$ the bound on the mean of M_n that was obtained in Piau (2001) for Bernoulli offspring and Poisson increments, as well as the fact, already noted in Piau (2001), that the mean field expectation of M_n^* overestimates the expectation of M_n. One could mimic our methods to get similar approximations of the higher moments of M_n.

In the product model, we generalize the bounds of Piau (2002) to every offspring law. Namely, we prove an approximation of every factorial moment of M_n of order k, at the order $G^k \alpha^n / S_0$, for any $\mu \in (0; \beta')$, where $\beta' \in (0; 1)$ and $\alpha \in (0; 1)$ are explicit. For higher values of μ, this bound has to be replaced by $G^k \widehat{\alpha}_k^n / S_0$, where the explicit value of $\widehat{\alpha}_k$ depends on k and μ.

In both models, η_n and η_n^* are within $1/S_0$ distance in total variation, for any mutation rate. The bound on the histogram of the additive model in Piau (2001) is only in $1/\sqrt{S_0}$ and is restricted to small values of the mutation rate. The bound on the histogram of the additive model in Piau (2002) is only in $1/\sqrt{S_0}$.

A surprising (to us) feature of all these results is that the bounds are uniform in n. Thus, η_n does not diverge from η_n^*, when $n \to \infty$. Finally, we mention that one can apply the same methods to study the mean empirical law of a uniform sample of size greater than 1 from the population at time n. For instance, if H_n denotes the empirical Hamming distance between two sequences in the additive model and H_n^* the mean field limit, one can show that

$$\mathbb{E}(H_n) = \mathbb{E}(H_n^*) + O(1)/S_0, \qquad \mathbb{E}(H_n^*) = 2q\mathbb{E}(\xi)n.$$

1.6. *Summary of notation.* Our results, stated in the next section, use notation that can be summarized as follows. Let Z_x denote the number of particles that are duplicated from x. In both models, we assume that Z is square integrable when necessary. We say $p_k := \mathbb{P}(Z = k)$ for $k \geq 0$, $q_0 := 1 - p_0$, $\sigma^2 := \operatorname{var}(Z)$, and

$$\lambda := \mathbb{E}(Z), \qquad p := 1/(1+\lambda), \qquad q := 1 - p.$$

For any duplication $x \to y$, let $\xi_y := s(y) - s(x)$. In the additive model, the random variables ξ_y are i.i.d. copies of ξ, which is integrable enough. Let $\mu := \mathbb{E}(\xi)$ (instead of $\widetilde{\mu}$ in previous sections) and assume, for the sake of simplicity, that $\mu \geq 0$. In the product model, recall that N is the size of the encoding alphabet of the sequences, G is the number of sites of the sequences and μ is the probability that the letter at a given site of an offspring y of x is different in y and in x. Introduce the parameters

$$\beta := 1 - 1/N, \qquad \mu' := 1 - \mu/\beta, \qquad \beta' := \beta \min\{1, q_0/q\},$$

and, for any $k \geq 0$,

$$\alpha_k := p + q(\mu')^k. \tag{5}$$

Thus, $\alpha := \alpha_1 = 1 - q\,\mu/\beta$ and $q_0\,\beta \leq \beta' \leq \beta$.

1.7. *Mathematical statements.*

THEOREM 1 (Moments of the additive model).

(i) *There exist* Θ^+ *and* Θ^- *in* $(0; +\infty)$, *which depend only on the offspring law*, *such that*, *for any* $n \geq 1$,

$$\mathbb{E}(M_n^*) - \mu\Theta^+/S_0 \leq \mathbb{E}(M_n) \leq \mathbb{E}(M_n^*) - \mu\Theta^-/(S_0 + 1).$$

If $p_1 \neq 0$, *this holds with*

$$\Theta^+ := 2\sigma^2/(p_0^3 q_0), \qquad \Theta^- := p_0 p_1 p^3.$$

If $p_1 = 0$ *but* $p_k \neq 0$, *this holds with the same value of* Θ^+ *and with*

$$\Theta^- := k p_0 p_k p^3.$$

(ii) *There exists* Γ *in* $(0; +\infty)$, *independent of* n *and* S_0, *such that*, *for any* $n \geq 0$,

$$|\operatorname{var}(M_n) - \operatorname{var}(M_n^*)| \leq \Gamma/S_0.$$

(iii) *The mean field values are*

$$\mathbb{E}(M_n^*) = n\mu q,$$
$$\operatorname{var}(M_n^*) = n\big(\operatorname{var}(\xi)q + \mu^2 pq\big).$$

In contrast to the additive model, in the product model, M_n^* cannot be the sum of n copies of a given random variable, simply because M_n is always in $\{0, 1, \dots, G\}$. For $k \le G$, let $e_k(t) := t(t-1)\cdots(t-k+1)$ denote the factorial polynomial of degree k. We prove that there exist $m_n^{(k)}$ and $m_{k,n}^*$ independent of G, such that

$$\mathbb{E}(e_k(M_n)) =: e_k(G)m_n^{(k)}, \qquad \mathbb{E}(e_k(M_n^*)) =: e_k(G)m_{k,n}^*.$$

THEOREM 2 (Moments of the product model). *For any $\mu \in (0; \beta')$ and $k \ge 1$, there exists Υ_k in $(0; +\infty)$, independent of n, such that*

$$|m_n^{(k)} - m_{k,n}^*| \le \Upsilon_k \frac{\alpha^n}{S_0}.$$

The cases $k = 1$ and $k = 2$ read as follows. Set $m_n := m_n^{(1)}$ and $m_n^* := m_{1,n}^*$.

THEOREM 3 (First moment of the product model). *The mean field value is*

$$m_n^* = \beta(1 - \alpha^n).$$

There exists κ in $(0; +\infty)$, which depends only on the offspring law, such that, for any $\mu \in (0; \beta')$,

$$|m_n - m_n^*| \le \kappa \frac{\beta^2 \mu}{(\beta' - \mu)^2} \frac{\alpha^n}{S_0}.$$

For smaller values of μ, more explicit bounds are obtained. If $\mu < \beta q_0/(4q)$ (the condition $\mu < \frac{1}{4} q_0 \beta$ is enough to ensure this), one has

$$|m_n - m_n^*| \le \frac{4\sigma^2}{q_0 p_0^3} \mu \frac{\alpha^n}{S_0}.$$

THEOREM 4 (Variance of the product model). *For $\mu \in (0; \beta')$,*

$$\begin{aligned}\operatorname{var}(M_n) = {} & G\beta(1-2\beta)(1-\alpha^n) + G\beta^2(1-\alpha_2^n) \\ & + G^2\beta^2(\alpha_2^n - \alpha^{2n}) + G^2 O\left(\frac{\alpha^n}{S_0}\right).\end{aligned}$$

By convention, $O(\cdot)$ is uniform, except with respect to β and to the laws of Z and ξ, and it can be explicitly bounded. Canceling the $O(\cdot)$ term, one gets the mean field value $\operatorname{var}(M_n^*)$.

REMARK 2. As was expected, $\operatorname{var}(M_n) \to G\beta(1-\beta)$ when $n \to \infty$. However, for finite n, two phase transitions occur. First, the error term dominates the α^{2n} term when $S_0 \ll \alpha^{-n}$. Second, the error term dominates the α^{2n} and α_2^n terms when $S_0 \ll (\alpha/\alpha_2)^n$. Recall that $(\alpha/\alpha_2)^n \ll \alpha^{-n}$.

REMARK 3. Some terms are missing in the expression of $\operatorname{var}(M_n^*)$ that one would deduce from Theorem 2 of Wang, Zhao, Cheng and Sun (2000), if one replaces the assertion "the lim sup is at most" in item (ii) of this theorem by "the limit is equal to." The missing terms being $(-G m_{2,n}^*)$, that is, negative, the assertion "the lim sup is at most" in this theorem holds as it is stated, and it can be strengthened to "the limit is $-\infty$." [The notation of the papers is mostly compatible, except that $e^{-\mu}$ in Wang, Zhao, Cheng and Sun (2000) is our $1-\mu$, and their a is our $1 - \frac{4}{3}\mu = 1 - \mu/\beta$.] The problem might come from the fact that these authors, "for simplicity, approximate the distribution of the number of base changes... by a Poisson random variable...."

REMARK 4. Theorems 2–4 can be extended to every value of the parameters as follows. For any $k \geq 1$, there exist Υ_k in $(0; +\infty)$, independent of n, and $\widehat{\alpha}_k \in (0; 1)$, such that

$$|m_n^{(k)} - m_{k,n}^*| \leq \Upsilon_k \frac{\widehat{\alpha}_k^n}{S_0}.$$

For $k = 1$, this holds with

$$\widehat{\alpha}_1 := \max\{p_0, p, \alpha\}.$$

For $k \geq 2$, this holds with $\widehat{\alpha}_k := \widehat{\alpha}_2$, where

$$\widehat{\alpha}_2 := \max\{p_0, p, \alpha, \alpha_2\}.$$

Thus, $\widehat{\alpha}_1$ is as follows. If $q_0 < q$ (i.e., if $\beta' < \beta$), $\widehat{\alpha}_1 = \alpha$ for $\mu \in (0; \beta')$, and $\widehat{\alpha}_1 = p_0$ for $\mu \in (\beta'; 1)$. If $q_0 \geq q$ (i.e., if $\beta' = \beta$), $\widehat{\alpha}_1 = \alpha$ for $\mu \in (0; \beta)$, and $\widehat{\alpha}_1 = p$ for $\mu \in (\beta; 1)$. The values of $\widehat{\alpha}_2$ are as follows. If $\mu \in (0; \beta)$, $\widehat{\alpha}_2 = \widehat{\alpha}_1$. If $\mu \in (\beta; 1)$, $\widehat{\alpha}_2 = \max\{p_0, \alpha_2\} = \max\{p_0, p + q(\mu')^2\}$.

We omit the proofs, which are similar to the proofs of Theorems 2–4. Furthermore, for instance, for $k = 1$, if $\mu > \beta'$, $\widehat{\alpha}_1^n \gg \alpha^n$, although m_n converges to its $n \to \infty$ limit at rate α^n. In our opinion, this makes the cases that our Theorems 2–4 leave out somewhat less interesting.

As regards the distance in total variation, recall that one sets, for instance,

$$\|\zeta_n - \zeta_n^*\|_{\mathrm{TV}} := \frac{1}{2} \sum_{k \geq 0} |\zeta_n(k) - \zeta_n^*(k)|.$$

Furthermore, the type t_n coincides with the number of mutations in the additive model with $\xi \equiv 1$. Part (i) of Theorem 5 below asserts that this case implies the general case.

THEOREM 5 (Histograms of both models). (i) *In both models*,

$$\|\eta_n - \eta_n^*\|_{\mathrm{TV}} \le \|\zeta_n - \zeta_n^*\|_{\mathrm{TV}}.$$

(ii) *There exists* Λ *in* $(0;+\infty)$, *which depends only on the offspring law, such that*

$$\|\zeta_n - \zeta_n^*\|_{\mathrm{TV}} \le \Lambda/S_0.$$

This holds with

$$\Lambda := 2\sigma^2/(p_0^3 q_0).$$

(iii) *The mean field law* ζ_n^* *dominates* ζ_n *stochastically.*

REMARK 5. We mention, see Piau (2002) for the proof, that, in part (ii) of Theorem 5, the upper bound Λ/S_0 can be replaced by $\Lambda'/\sqrt{S_0}$, with

$$\Lambda' := 2\sigma p/q_0.$$

For small values of S_0, this bound can be a sharper alternative.

REMARK 6. Part (iii) of Theorem 5 implies that, in the additive model of Theorem 1, $\mathbb{E}(M_n^*)$ is indeed greater than $\mathbb{E}(M_n)$ when $\mathbb{E}(\xi)$ is positive.

1.8. *Organization.* As Sun already observed for the additive model, η_n can be decomposed along the law of the types. This is explained in Section 2 in more detail than in Section 1.2. Section 3 provides a description of the mean field limit η_n^*. Section 4 collects some partly new estimates of branching processes. The proofs of the theorems are deferred to Sections 5–8. The Appendix, which can be read separately from the rest of the paper, states and proves precise estimates of the harmonic means of branching processes used in the proofs.

2. Decomposition by types.

The law of the mean empirical type t_n determines the law of M_n. Recall that the type $t(x)$ of a sequence x is the number of replications that produced x. In other words, if x exists before the first cycle, $t(x) := 0$, and if $x \to y$ is a replication, $t(y) := t(x) + 1$.

In the additive model, conditionally on $\{t(x) = k\}$, $s(x)$ is the sum σ_k of k i.i.d. copies of ξ. In the product model, for any replication $x \to y$, conditionally on $\{s(x) = k\}$, $s(y) - s(x)$ is the sum of G independent random variables. Among them, k follow the law η_-, which indicates the possible replacement of a false letter by the correct letter 0, and the $(G - k)$ others follow the law η_+, which indicates the possible creation of a new false letter. Then

$$\eta_- := \big(1 - \mu/(N-1)\big)\delta_0 + \big(\mu/(N-1)\big)\delta_{-1},$$
$$\eta_+ := (1-\mu)\delta_0 + \mu\delta_1.$$

If $t(x) = n$, $s(x)$ is the result of the mutations that occurred during n replications. Introduce a Markov chain $(\rho_n)_{n\geq 0}$ on $\{0, 1\}$, of transition matrix

$$\mathbb{P}(\rho_{n+1} = 1|\rho_n = 0) = \mu,$$
$$\mathbb{P}(\rho_{n+1} = 0|\rho_n = 1) = \mu/(N-1),$$

and starting from $\rho_0 = 0$. The law of ρ_n is $\mu_n\delta_1 + (1-\mu_n)\delta_0$, where

$$\mu_n := \beta\bigl(1 - (1-\mu/\beta)^n\bigr).$$

Thus, $\mu_0 = 0$, $\mu_1 = \mu$, $\mu_2 = 2\mu - \mu^2/\beta$ and $\mu_n \to \beta$ when $n \to \infty$. Finally, conditionally on $\{t(x) = n\}$, $s(x)$ is the sum of G i.i.d. copies of ρ_n. The stationary law of ρ_n is

$$(1-\beta)\delta_0 + \beta\delta_1.$$

Hence, after many replications, each letter of a sequence is, independently of the others, false with probability β.

3. Mean field description.

In the mean field limit $S_0 \to \infty$, t_n converges in law to t_n^* of binomial (n, p) law, and M_n^*, conditional on t_n^*, follows the law of M_n, conditional on t_n. For $|u| \leq 1$, let

$$L_n^*(u) := \mathbb{E}\bigl(u^{M_n^*}\bigr).$$

THEOREM 6. *In the additive model*,

$$L_n^*(u) = \bigl(p + q\mathbb{E}(u^\xi)\bigr)^n =: L_*(u)^n.$$

THEOREM 7. *In the product model*,

$$\begin{aligned} m_{k,n}^* &= \sum_{i=0}^{n} \binom{n}{i} q^i p^{n-i} \mu_i^k \\ &= \beta^k \sum_{i=0}^{k} \binom{k}{i} (-1)^i \alpha_i^n, \\ L_n^*(u) &= \sum_{k=0}^{n} \binom{n}{k} q^k p^{n-k} \bigl(1 - \mu_k(1-u)\bigr)^G \\ &= \sum_{k=0}^{G} \binom{G}{k} \bigl(1 - \beta(1-u)\bigr)^{G-k} \bigl(\beta(1-u)\bigr)^k \alpha_k^n. \end{aligned}$$

Thus, with $\alpha := \alpha_1$, one has

$$m_n^* := m_{1,n}^* = \beta(1-\alpha^n), \qquad m_{2,n}^* = \beta^2(1 - 2\alpha^n + \alpha_2^n).$$

REMARK 7. The $G \to \infty$ Poisson limit of M_n^* in the product model is the $S_0 \to \infty$ limit M_n^{**} of M_n in the additive Poisson model. That is, assuming that $G\mu \to \widetilde{\mu}$,

$$\mathbb{E}(M_n^{**}) = n\widetilde{\mu}q,$$

$$\operatorname{var}(M_n^{**}) = n(\widetilde{\mu}^2 pq + \widetilde{\mu}q).$$

This follows directly from the fact that, in the mean field limit of the additive model, M_n^{**} follows the law of the sum of a binomial (n, q) number of i.i.d. Poisson random variables of parameter $\widetilde{\mu}$.

4. Branching estimates.

4.1. *Lemmas.* Let $\mathcal{F}_n$ denote the σ-algebra generated by the n first cycles, that is, by all the replications and mutations during the n first generations of the iBMP. Let f denote the generating function of $L := Z + 1$, that is, for $|u| \le 1$,

$$f(u) := u\mathbb{E}(u^Z) = p_0 u + p_1 u^2 + \cdots .$$

Lemma 8 is in Joffe (1993). Lemma 10 provides estimations of the error term in Lemma 9. Lemma 11 is a direct consequence of Proposition A.5 in the Appendix.

LEMMA 8. *For any x in generation n,*

$$\mathbb{E}\biggl(\frac{1 + Z_x}{S_{n+1}}\bigg|\mathcal{F}_n\biggr) = \frac{1}{S_n}.$$

LEMMA 9. *One has*

$$\mathbb{E}\biggl(\frac{S_n}{S_{n+1}}\bigg|\mathcal{F}_n\biggr) = p + A(S_n),$$

where $A(S_n)$ is nonnegative and defined by

$$A(S) := \int_0^1 h(u) f'(u) f(u)^S \, du, \qquad h(u) := -\frac{1}{f'(u)}\biggl(\frac{f(u)}{uf'(u)}\biggr)'.$$

LEMMA 10. *If $p_1 \neq 0$, there exist h^- and h^+ in $(0; +\infty)$ such that, for any $u \in (0; 1)$, $h^- \le h(u) \le h^+$. One can choose*

$$h^+ := \sigma^2 / p_0^3, \qquad h^- := p_0 p_1 p^3.$$

A consequence of Lemmas 9 and 10 is

$$h^-/(S + 1) \le A(S) \le h^+/(S + 1). \tag{6}$$

We show, in the proof of Lemma 10, that (6) still holds when $p_1 = 0$, with a different value of h^-. Corollary 12 below stems from Lemma 11 and (6).

LEMMA 11.

$$\sum_{k\geq 0}\mathbb{E}\big((S_k+1)^{-1}\big)\leq \big(1-\mathbb{E}((1+Z)^{-1})\big)^{-1}S_0^{-1}\leq 2/(q_0\,S_0).$$

COROLLARY 12. *For any* $n\geq 0$,

$$h^-/(S_0+1)\leq\sum_{k=0}^{n}\mathbb{E}(A(S_k))\leq 2h^+/(q_0\,S_0).$$

4.2. *Technicalities.*

PROOF OF LEMMA 9. One writes

$$\mathbb{E}\bigg(\frac{S_n}{S_{n+1}}\bigg|\mathcal{F}_n\bigg)=S_n\int_0^1 f(u)^{S_n}\frac{du}{u}=\int_0^1 (f(u)^{S_n})'\bigg(\frac{f(u)}{uf'(u)}\bigg)\,du.$$

Integration by parts of the last expression yields p as the integrated term and the integral of $hf'f^{S_n}$ as the term to integrate. □

PROOF OF LEMMA 10. A slight rewriting of h yields $h=g/k^3$, with

$$\begin{aligned}g(u)&:=\sum_n(n+1)^2p_nu^{n+1}\sum_n p_nu^{n+1}-\bigg(\sum_n(n+1)p_nu^{n+1}\bigg)^2\\&=\sum_{n,n'}\tfrac{1}{2}(n-n')^2p_np_{n'}u^{n+n'+2},\\k(u)&:=\sum_n(n+1)p_nu^{n+1}.\end{aligned}$$

For $u\in(0;1)$, $g(u)\leq g(1)u^3=\sigma^2u^3$ and $k(u)\geq p_0u$. On the other hand, $g(u)\geq p_0p_1u^3$ and $k(u)\leq k(1)u=u/p$. This yields h^+ and h^- when $p_1\neq 0$.

If $p_1=0$, the value of h^- is different. To see this, assume that $p_k\neq 0$ with $k\geq 2$. Since $g(u)\geq p_0p_kk^2u^{k+2}$ and $k(u)\leq u/p$, for any $v\in(0;1)$ and $u\in(v;1)$, one has

$$h(u)\geq q^3p_0p_kk^2u^{k-1}\geq q^3p_0p_kk^2v^{k-1}.$$

On the other hand,

$$\int_v^1(f(u)^S)'\,du=\frac{1-f(v)^{S+1}}{S+1}\geq\frac{1-v^2}{S+1}.$$

Thus, for any v, one can choose

$$h^-:=p^3p_0p_kk^2(v^{k-1}-v^{k+1}).$$

Optimization over v and simple estimates finally yield

$$h^-:=\frac{4}{3\sqrt{3}}p^3p_0p_kk<p^3p_0p_kk.$$

□

5. Moments of the additive model. Let T_n denote the sum from $x = 1$ to S_n of $s(x)$. Then

$$T_{n+1} = \sum_{x=1}^{S_n}\left(s(x)(1+Z_x) + \sum_{y=1}^{Z_x}\xi_y\right).$$

From Lemmas 8 and 9,

$$\begin{aligned}\mathbb{E}(M_{n+1}|\mathcal{F}_n) &= \mathbb{E}\left(\frac{T_{n+1}}{S_{n+1}}\Big|\mathcal{F}_n\right)\\ &= M_n + \mu\sum_{x=1}^{S_n}\mathbb{E}\left(\frac{Z_x}{S_{n+1}}\Big|\mathcal{F}_n\right)\\ &= M_n + \mu q - \mu A(S_n).\end{aligned}$$

An integration of this and the subsequent iteration yield

$$\begin{aligned}\mathbb{E}(M_n) &= n\mu q - \mu\sum_{k=0}^{n-1}\mathbb{E}(A(S_k))\\ &= \mathbb{E}(M_n^*) - \mu\sum_{k=0}^{n-1}\mathbb{E}(A(S_k)).\end{aligned}\tag{7}$$

Thus, Corollary 12 implies the theorem with

$$\Theta^+ := 2h^+/q_0, \qquad \Theta^- := h^-.$$

Turning to the evaluation of the second moment and following the path that leads to (7), one gets

$$\begin{aligned}\mathbb{E}(M_{n+1}^2) = \mathbb{E}(M_n^2) &+ 2\mu q\mathbb{E}(M_n) + \mathbb{E}(\xi^2)q\\ &- 2\mu\mathbb{E}\big(M_nA(S_n)\big) - \mathbb{E}(\xi^2)\mathbb{E}(A(S_n)),\end{aligned}\tag{8}$$

where the only additional ingredient is the fact that

$$\begin{aligned}\mathbb{E}\left(\frac{T_n}{S_{n+1}}\Big|\mathcal{F}_n\right) &= \mathbb{E}\left(M_n\mathbb{E}\left(\frac{S_n}{S_{n+1}}\Big|\mathcal{F}_n\right)\right)\\ &= p\mathbb{E}(M_n) + \mathbb{E}\big(M_nA(S_n)\big).\end{aligned}$$

From (7) and (8),

$$\mathrm{var}(M_{n+1}) = \mathrm{var}(M_n) + \mathbb{E}(\xi^2)q - \mu^2q^2 + B_n,$$

where the error term B_n is

$$\begin{aligned}B_n = 2\mu\mathbb{E}(M_n)\mathbb{E}(A(S_n)) &- 2\mu\mathbb{E}\big(M_nA(S_n)\big) - \mathbb{E}(\xi^2)\mathbb{E}(A(S_n))\\ &- \mu^2\mathbb{E}(A(S_n))^2 + 2\mu^2q\mathbb{E}(A(S_n)).\end{aligned}\tag{9}$$

It remains to prove that $|B_n|$ is summable. Since $\mathbb{E}(M_n)$ increases like n, and since $\mathbb{E}(A(S_n))$ decreases at least as fast as $\mathbb{E}(S_n^{-1})$, that is, geometrically, the only problematic term in (9) is $\mathbb{E}(M_n A(S_n))$. Starting from

$$\frac{1}{S(1+S)} = \int_0^1 t^S (1-t) \frac{dt}{t}$$

and using standard manipulations, one gets

$$\mathbb{E}\Big(\frac{M_{n+1}}{S_{n+1}+1}\Big|\mathcal{F}_n\Big) = (M_n+\mu)\int_0^1 S_n f'(t) f(t)^{S_n-1}(1-t)\,dt \\ -\mu\int_0^1 S_n f(t)^{S_n}(1-t)\,\frac{dt}{t}.$$

We are interested only in crude estimates of the left-hand side. We use

$$p_0(1-t) \le 1 - f(t)$$

in the first integral of the right-hand side, we perform this integral and we cancel the second integral of the right-hand side. This yields

$$\mathbb{E}\Big(\frac{M_{n+1}}{S_{n+1}+1}\Big|\mathcal{F}_n\Big) \le p_0^{-1}\frac{M_n+\mu}{S_n+1},$$

that is,

$$\mathbb{E}\Big(\frac{M_{n+1}}{S_{n+1}+1}\Big) \le p_0^{-1}\mathbb{E}\Big(\frac{M_n}{S_n+1}\Big) + \mu p_0^{-1}\mathbb{E}\Big(\frac{1}{S_n+1}\Big).$$

This, along with the estimates of the Appendix, is enough to show that $|B_n|$ is summable and that the sum of the series is at most a multiple of $1/S_0$.

6. First moment of the product model.

6.1. *Exact recursions.* Setting $\ell(x) := \ell_1(x)$, one has $\mathbb{E}(M_n) = G m_n$ with

$$m_n := \mathbb{E}\Big(\frac{T_n}{S_n}\Big), \qquad T_n := \sum_{x=1}^{S_n} \mathbf{1}_{\{\ell(x)\ne 0\}}.$$

If the duplication $x \to y$ occurs,

$$\mathbb{P}\big(\ell(y)=0|\ell(x)\ne 0\big) = \mu/(N-1), \qquad \mathbb{P}\big(\ell(y)\ne 0|\ell(x)=0\big) = \mu.$$

These two relations can be written as

$$\mathbb{P}\big(\ell(y)\ne 0|\ell(x)\big) = \mu + \mu'\mathbf{1}_{\{\ell(x)\ne 0\}}, \qquad \mu' := 1 - \mu/\beta.$$

Using Lemma 8, and

$$T_{n+1} = \sum_{x=1}^{S_n}\Big(\mathbf{1}_{\{\ell(x)\ne 0\}} + \sum_{y=1}^{Z_x}\mathbf{1}_{\{\ell(y)\ne 0\}}\Big),$$

one gets

$$
\begin{aligned}
y_n &:= \mathbb{E}\Big(\frac{T_{n+1}}{S_{n+1}}\Big|\mathcal{F}_n\Big)\\
&= \sum_{x=1}^{S_n} \mathbf{1}_{\{\ell(x)\neq 0\}}\mathbb{E}\Big(\frac{1+\mu' Z_x}{S_{n+1}}\Big|\mathcal{F}_n\Big) + \mu\mathbb{E}\Big(\frac{Z_x}{S_{n+1}}\Big|\mathcal{F}_n\Big)\\
&= \alpha\frac{T_n}{S_n} + \beta(1-\alpha) + \mu A(S_n)\Big(\frac{T_n}{\beta S_n} - 1\Big),
\end{aligned}
$$

where $\alpha = \alpha_1 \in (0;1)$ is defined by (5). Iterating this recursion and using $m_{n+1} = \mathbb{E}(y_n)$, one gets

$$m_n = \beta(1-\alpha^n) + \mu C_n,$$

where

$$C_n := \sum_{k=0}^{n-1} \alpha^{n-1-k} c_k, \qquad c_k := \mathbb{E}\Big(A(S_k)\Big(\frac{T_k}{\beta S_k} - 1\Big)\Big).$$

6.2. *Estimations of the error terms.* It seems difficult to obtain further exact expressions of m_n. Turning to approximations of C_n, we first note that

$$|c_k| \le \mathbb{E}(A(S_k)),$$

because $0 \le T_k \le S_k$ almost surely. From Lemma 10 and Proposition A.2,

$$|c_k| \le h^+\mathbb{E}((1+S_k)^{-1}) \le h^+\kappa k^* r^k/S_0,$$

where $r := \max\{p_0, p\}$. Since $\alpha = 1 - q\mu/\beta$, the condition $\alpha > r$ is equivalent to

$$\mu < \beta' := \beta \min\{1, q_0/q\}.$$

Assuming this, one gets by summation

$$|C_n| \le h^+\kappa\alpha\beta^2 q^{-2}(\beta' - \mu)^{-2}\alpha^n/S_0,$$

if $n^* = n+1$ in Proposition A.2, and, in the other case,

$$|C_n| \le h^+\kappa\beta q^{-1}(\beta' - \mu)^{-1}\alpha^n/S_0.$$

Further simplifications lead, for $\mu \in (0;\beta')$, to

$$|m_n - m_n^*| \le h^+\kappa q^{-2}\frac{\beta^2\mu}{(\beta'-\mu)^2}\frac{\alpha^n}{S_0}. \tag{10}$$

REMARK 13. The upper bound is small in two respects. First, S_0 is usually large. Second, the ratio

$$\frac{\beta^2\mu}{(\beta'-\mu)^2}$$

is comparable to μ as long as μ is not too close to β'. In the PCR context, the mutation rate μ per cycle and per site is often quite small. We refer to the discussion in Piau (2001) for the values of μ in actual PCR.

REMARK 14. For smaller values of μ, Proposition A.5 allows to cancel the unspecified factor κ. If $\mu < \beta q_0/(2q) =: \beta''$, the right-hand side of (10) can be replaced by

$$h^+ q^{-1} \frac{\beta\mu}{\beta'' - \mu} \frac{\alpha^n}{S_0}.$$

If $\mu < \beta''/2$, (10) holds with the simpler $(4h^+/q_0)\mu\alpha^n/S_0$.

7. Higher moments of the product model. The method of Section 6 yields every moment of M_n. Recall that M_n is $\{0, 1, \dots, G\}$ valued. For any $k \le G$,

$$\mathbb{E}(e_k(M_n)) = e_k(G) m_n^{(k)}, \qquad m_n^{(k)} := \mathbb{E}\biggl(\frac{T_n^{(k)}}{S_n}\biggr),$$

where $e_k(t) := t(t-1)\cdots(t-k+1)$ denotes the factorial polynomial of degree k and where $T_n^{(k)}$ is an analogue of T_n, defined by

$$T_n^{(k)} := \sum_{x=1}^{S_n} U_k(x), \qquad U_k(x) := \prod_{i=1}^{k} \mathbf{1}_{\{\ell_i(x) \ne 0\}}.$$

Conditionally on $\mathcal{F}_n$,

$$\mathbb{E}(T_{n+1}^{(k)} | \mathcal{F}_n) = \sum_{x=1}^{S_n} \Biggl(U_k(x) + \sum_{y=1}^{Z_x} \mathbb{E}\bigl(U_k(y) | (\ell_i(x))_{1 \le i \le k}\bigr) \Biggr).$$

Since each letter evolves independently of the others, the conditioning is

$$\mathbb{E}\bigl(U_k(y) | (\ell_i(x))_{1 \le i \le k}\bigr) = \prod_{i=1}^{k} \bigl(\mu + \mu' \mathbf{1}_{\{\ell_i(x) \ne 0\}}\bigr) =: V_k(x).$$

This makes possible our intermediate step, which is to compute

$$y_n^{(k)} := \mathbb{E}\biggl(\frac{T_{n+1}^{(k)}}{S_{n+1}} \bigg| \mathcal{F}_n\biggr).$$

With the notation of Section 6, one finds

$$y_n^{(k)} = \frac{T_n^{(k)}}{S_n} D_n + \sum_{i=0}^{k} \binom{k}{i} \mu^{k-i} (\mu')^i \frac{T_n^{(i)}}{S_n} D_n',$$

where, for any member x of generation n,

$$D_n := \mathbb{E}\biggl(\frac{S_n}{S_{n+1}} \bigg| \mathcal{F}_n\biggr), \qquad D_n' := \mathbb{E}\biggl(\frac{Z_x S_n}{S_{n+1}} \bigg| \mathcal{F}_n\biggr).$$

From Lemmas 8 and 9,

$$D_n = p + A(S_n), \qquad D'_n = q - A(S_n).$$

From now on, we assume, for simplicity, that $k = 2$:

$$y_n^{(2)} = \alpha_2 \frac{T_n^{(2)}}{S_n} + q\left(\mu^2 + 2\mu\mu' \frac{T_n}{S_n}\right) + \text{some terms in } A(S_n).$$

This means that $m_n^{(2)}$ satisfies the recursion

$$m_{n+1}^{(2)} = \alpha_2 m_n^{(2)} + \alpha_{2,1} m_n + \alpha_{2,0} + c_n^{(2)},$$

where α_2, $\alpha_{2,1}$ and $\alpha_{2,0}$ are constant and $c_n^{(2)}$ is an error term. More precisely, α_2 is defined by (5), and

$$\begin{aligned} \alpha_{2,1} &:= 2\mu\mu' q, \\ \alpha_{2,0} &:= \mu^2 q, \\ c_n^{(2)} &:= \mathbb{E}\left(A(S_n)\left((1 - (\mu')^2)\frac{T_n^{(2)}}{S_n} - 2\mu\mu' \frac{T_n}{S_n} - \mu^2\right)\right). \end{aligned}$$

The term in the inner parentheses in $c_n^{(2)}$ is $O(1)$, $A(S_n)$ is nonnegative and $\mathbb{E}(A(S_n))$ is of the order of $\mathbb{E}(S_n^{-1})$, that is, from the results of the Appendix, at most of the order of nr^n/S_0 with

$$r := \max\{p_0, p\}.$$

Iterating this, $m_n^{(2)}$ can be written as the sum of $m_{2,n}^*$ and of two error terms. The first one is

$$\alpha_{2,1}\mu \sum_{i=0}^{n-1} C_i \alpha_2^{n-1-i}.$$

Since C_i is at most a multiple of α^i/S_0 and since $\alpha_2 < \alpha$ for $\mu < \beta'$, this sum is at most a multiple of α^n/S_0. The second error term is

$$\sum_{i=0}^{n-1} c_i^{(2)} \alpha_2^{n-1-i}.$$

Since $c_i^{(2)}$ is at most a multiple of $(i+1)r^i/S_0$ and since $\alpha_2 < r$, this sum is at most a multiple of nr^n/S_0, which is negligible with respect to α^n/S_0 since $r < \alpha$. This yields the $k = 2$ case of the theorem when $\mu < \beta'$. The proof for $k \geq 3$ is similar and omitted.

8. Total variation bounds. Part (i) of Theorem 5 is a straightforward consequence of the fact that the law of M_n conditioned by t_n is equal to the law of M_n^* conditioned by t_n^*, and of the following characterization of the distance in total variation: $\|\eta_n - \eta_n^*\|_{\mathrm{TV}}$ is the infimum of $\mathbb{P}(M \neq M^*)$ over the couples (M, M^*) of random variables, such that the law of M is η_n and the law of M^* is η_n^*.

Applying this to ζ_n and ζ_n^*, one gets random variables τ and τ^* of law ζ_n and ζ_n^* such that $\mathbb{P}(\tau \neq \tau^*)$ is close to $\|\zeta_n - \zeta_n^*\|_{\mathrm{TV}}$. Using τ and τ^* to construct M and M^* yields an admissible couple (M, M^*) such that $\{M \neq M^*\} \subset \{\tau \neq \tau^*\}$. This proves part (i).

As regards part (ii), recall that $S_n(k)$ denotes the number of particles of type k at time n and that

$$S_{n+1}(k) = S_n(k) + \sum_{x=1}^{S_n(k-1)} Z_x.$$

Thus, $\mathbb{P}(t_{n+1} = k)$ involves $S_n(k)$, $S_n(k-1)$ and S_{n+1}. Replacing S_{n+1} by S_n/q in this recursion relation yields, owing to the lemmas of Section 4, error terms which involve $A(S_n)$ and which can be controlled. More precisely,

$$\mathbb{E}\biggl(\frac{S_{n+1}(k)}{S_{n+1}}\bigg|\mathcal{F}_n\biggr) = \bigl(p + A(S_n)\bigr)\frac{S_n(k)}{S_n} + \bigl(q - A(S_n)\bigr)\frac{S_n(k-1)}{S_n}.$$

Integrating this, one gets

$$\mathbb{P}(t_{n+1} = k) = p\mathbb{P}(t_n = k) + q\mathbb{P}(t_n = k-1) + \mathbb{E}(V_n(k)), \tag{11}$$

where $V_n(k)$ is an error term defined by

$$S_n V_n(k) := A(S_n)\bigl(S_n(k) - S_n(k-1)\bigr).$$

The mean field version t_n^* solves

$$\mathbb{P}(t_{n+1}^* = k) = p\mathbb{P}(t_n^* = k) + q\mathbb{P}(t_n^* = k-1).$$

Finally, note that

$$\sum_{k \geq 0} |S_n(k) - S_n(k-1)| \leq 2S_n.$$

This implies that

$$\|\zeta_{n+1} - \zeta_{n+1}^*\|_{\mathrm{TV}} \leq \|\zeta_n - \zeta_n^*\|_{\mathrm{TV}} + \mathbb{E}(A(S_n)).$$

This proves part (ii) with $\Lambda := 2h^+/q_0$.

Turning to part (iii), pick a nondecreasing function φ. Then (11) yields the recursion

$$\begin{aligned} &\mathbb{E}\bigl(\varphi(t_{n+1}^*) - \varphi(t_{n+1})\bigr) \\ &\quad = \mathbb{E}(W_n(\varphi)) + p\mathbb{E}\bigl(\varphi(t_n^*) - \varphi(t_n)\bigr) + q\mathbb{E}\bigl(\varphi(t_n^* + 1) - \varphi(t_n + 1)\bigr), \end{aligned} \tag{12}$$

with

$$W_n(\varphi) := -\sum_{k\geq 0}\varphi(k)V_n(k).$$

Abel's transform yields

$$S_n W_n(\varphi) = A(S_n)\sum_{k\geq 0}\big(\varphi(k+1)-\varphi(k)\big)S_n(k)\geq 0.$$

Finally, if t_n^* dominates t_n stochastically, all the terms of the right-hand side of (12) are nonnegative. Hence,

$$\mathbb{E}(\varphi(t_{n+1}^*))\geq \mathbb{E}(\varphi(t_{n+1})).$$

This proves that t_{n+1}^* dominates t_{n+1} stochastically.

APPENDIX

On the harmonic means of branching processes. We study the behavior of $\mathbb{E}(S_n^{-1})$, where $(S_n)_{n\geq 0}$ denotes the homogeneous branching process of reproducing law $(p_k)_{k\geq 1}$, starting from a deterministic $S_0\geq 1$. We assume that $p_1\neq 0$ and $p_1\neq 1$, and we denote by $L\geq 1$ an integer-valued random variable of law $(p_k)_{k\geq 1}$. For $|t|\leq 1$ and $n\geq 0$, let

$$f(t) := \mathbb{E}(t^L), \qquad f_n(t) := \mathbb{E}\big(t^{S_n}\big).$$

Let $m := \mathbb{E}(L)$, with the convention that $m^{-1} := 0$ is L is not integrable. Introduce

$$r := \max\{p_1, m^{-1}\}.$$

Finally, let $\mathbb{P}_\gamma$ denote the law of $(S_n)_{n\geq 0}$ when $p_k := \gamma(1-\gamma)^{k-1}$, and $\mathbb{P}^\varepsilon$ the law of $(S_n)_{n\geq 0}$ when $p_1 := 1-\varepsilon$ and $p_2 := \varepsilon$.

REMARK A.1. The random variable L plays the role of $1+Z$ in the iBMP context. Thus, m corresponds to $1+\lambda = 1/p$, and p_1 in the Appendix is p_0 in the rest of the paper and so on.

A.1. *Results.* Joffe (1993) mentions, as an unpublished result, the fact that, for any $a>0$,

$$\mathbb{E}(S_n^{-a})\leq cu^n \qquad \text{for any } u > \max\{p_1, m^{-a}\},$$

where c is independent of n. In the same vein, Athreya (1994) shows that, if $p_1 m^a > 1$, then

$$p_1^{-n}\mathbb{E}(S_n^{-a}|S_0 = 1)$$

is a nondecreasing sequence that converges to a finite limit

$$\int_0^1 g(t)|\log t|^{a-1}\frac{dt}{t}.$$

Here, g is the unique solution on $[0; 1)$ of the functional equation

$$g(f(t)) = p_1 g(t), \qquad g(0) = 0, g'(0^+) = 1, g(t) \neq 0 \text{ for } t \neq 0.$$

We refine the Athreya–Joffe bounds, mainly for the first moment of S_n^{-1}. The emphasis is put on explicit and nonasymptotic bounds, as well as on the dependence of the bounds on S_0. Athreya and other authors deal with the case $S_0 = 1$; see Remark A.4.

PROPOSITION A.2. *Let* $n^* := n + 1$ *if* $p_1 = m^{-1}$ *and* $S_0 = 1$, *and* $n^* := 1$ *otherwise.*

(i) $\mathbb{E}(S_n^{-1}) \geq m^{-n}/S_0$.
(ii) *There exists* κ, *independent of n and* S_0, *such that*

$$\mathbb{E}(S_n^{-1}) \leq \kappa n^* r^n / S_0.$$

REMARK A.3. The apparition of n in the upper bound when $p_1 = m^{-1}$ and $S_0 = 1$ is not due to an artifact of our proof, as the example of $\mathbb{P}_\gamma$ shows; see Proposition A.7 and Remark A.4.

REMARK A.4. After the completion of this paper, we discovered that the harmonic moments of supercritical branching processes are not an entirely new subject, owing to the historical notes in a recent preprint that we ran into by Ney and Vidyashankar (2001).

According to these authors, Heyde and Brown (1971) conjectured that $\mathbb{E}(S_n^{-1})$ is equivalent to a multiple of m^{-n} in some cases, and they found a case where $\mathbb{E}(S_n^{-1})$ is equivalent to a multiple of nm^{-n} (apparently our $\mathbb{P}_\gamma$). Pakes (1975) established that $\mathbb{E}(S_n^{-1})$ is equivalent to a multiple of r^{-n} when $p_1 \neq m^{-1}$ and $L \log L$ is integrable, and he conjectured the case $p_1 = m^{-1}$. Ney and Vidyashankar settled the case $p_1 = m^{-1}$ and they recovered the case $p_1 \neq m^{-1}$, both results under the weaker assumption that L is integrable. Additionally, they found the exact equivalent of $\mathbb{E}(S_n^{-i})$ for any $i > 0$.

All these authors assume that $S_0 = 1$. As a consequence, they miss the fact that the degeneracy of the case $p_1 = m^{-1}$ and $S_0 = 1$, namely that the exact equivalent of $\mathbb{E}(S_n^{-1})$ is a multiple of nr^n and not a multiple of r^n, disappears when $S_0 \geq 2$. For the higher moments of S_n^{-1}, see Section A.2.

Below, we trade the mostly unspecified constant κ of Proposition A.2 against nonoptimal geometric terms. For instance, from Proposition A.5,

$$(1 + \varepsilon)^{-n}/S_0 \leq \mathbb{E}^\varepsilon(S_n^{-1}) \leq (1 - \varepsilon/2)^n/S_0.$$

PROPOSITION A.5. $\mathbb{E}(S_n^{-1}) \leq \mathbb{E}(L^{-1})^n/S_0$.

The comparison of $\mathbb{P}$ with $\mathbb{P}_\gamma$ yields additional upper bounds. Proposition A.6 compares $\mathbb{P}$ to $\mathbb{P}^\varepsilon$ to $\mathbb{P}_\gamma$ and Proposition A.7 solves the case $\mathbb{P}_\gamma$.

PROPOSITION A.6. *One has $f \le f^\varepsilon \le f_\gamma$ for $\varepsilon = 1 - p_1$ and $\gamma = 1/(1+\varepsilon)$. Hence, starting from the same S_0,*

$$\mathbb{E}(S_n^{-1}) \le \mathbb{E}^\varepsilon(S_n^{-1}) \le \mathbb{E}_\gamma(S_n^{-1}).$$

PROPOSITION A.7. *The sequence of general term $b_\gamma^n := \mathbb{E}_\gamma(S_n^{-1})/\gamma^n$ is increasing and starts from $b_\gamma^0 = 1/S_0$. If $S_0 \ge 2$, $b_\gamma^n \to 1/(S_0 - 1)$. If $S_0 = 1$, $b_\gamma^n \sim n \log(1/\gamma)$, since*

$$\mathbb{E}_\gamma(S_n^{-1}) = \log(1/\gamma^n)\gamma^n/(1-\gamma^n).$$

COROLLARY A.8. *If $S_0 \ge 2$,*

$$\mathbb{E}(S_n^{-1}) \le (2 - p_1)^{-n}/(S_0 - 1).$$

If $S_0 = 1$,

$$\mathbb{E}(S_n^{-1}) \le (2 - p_1)^{-n}(2 - p_1)/(1 - p_1).$$

In Corollary A.8, the order of magnitude of the upper bound when $n \to \infty$ is not correct when $(2 - p_1)^{-1} > r$. The exception is $\mathbb{P}^\varepsilon$, for which we get surprisingly tight estimates.

COROLLARY A.9. *Let $d_\varepsilon^n := (1+\varepsilon)^n \mathbb{E}^\varepsilon(S_n^{-1})$. If $S_0 \ge 2$,*

$$1/S_0 \le d_\varepsilon^n \le 1/(S_0 - 1) \le 2/S_0.$$

If $S_0 = 1$,

$$1 \le d_\varepsilon^n \le 1 + \varepsilon^{-1},$$

and the limit of d_ε^n when $n \to \infty$ is greater than $\frac{1}{2}(1 + \varepsilon^{-1})$. Thus, for any $S_0 \ge 1$,

$$(1+\varepsilon)^{-n}/S_0 \le \mathbb{E}^\varepsilon(S_n^{-1}) \le (1 + \varepsilon^{-1})(1+\varepsilon)^{-n}/S_0.$$

REMARK A.10. For general laws, there is a gap between our upper and lower bounds when $p_1 > m^{-1}$. In that case, Athreya's result shows that $\mathbb{E}(S_n^{-1})/p_1^n$ converges to a finite limit when $S_0 = 1$. However, the refinement of (i) into

$$\mathbb{E}(S_n^{-1}) \ge c p_1^n / S_0$$

cannot hold with any absolute constant c, as the example of $\mathbb{P} = \frac{1}{2}(\delta_1 + \delta_{N^2})$ and $S_0 = N$ shows, when $N \to \infty$ (we leave the details to the reader). We mention without proof the following results. One always has

$$\mathbb{E}(S_n^{-1}) \ge p_1^{nS_0}, \qquad \mathbb{E}(S_n^{-1}) \ge m^{-n}/S_0.$$

On the other hand, if $p_1^{S_0} > m^{-1}$, $\mathbb{E}(S_n^{-1}) \le c c_1^{S_0} p_1^{nS_0}$, and, if $p_1^{S_0} < m^{-1}$, $\mathbb{E}(S_n^{-1}) \le c' m^{-n}$. Thus, as $n \to \infty$,

$$n^{-1} \log \mathbb{E}(S_n^{-1}) \to r(S_0) := \max\{p_1^{S_0}, m^{-1}\}.$$

A.2. *Higher moments.* We mention without proof that the same techniques apply to moments of higher orders. Using the concavity of the function of $t > 0$ defined by

$$\prod_{i=1}^{k} (t^{-1/k} + c_i)^{-1}$$

for any nonnegative c_i, one can prove that

$$\mathbb{E}(S_n^{-k} | S_0) \le S_0^{-k} \mathbb{E}(S_n^{-k} | S_0 = 1),$$
$$\mathbb{E}\big(e_k(S_n)^{-1} | S_0\big) \le e_k(S_0)^{-1} \mathbb{E}(S_n^{-k} | S_0 = 1).$$

The comparison with $\mathbb{P}_\gamma$ uses the moments of $e_k(S_n)^{-1}$ for $S_0 = 1$. These are such that, for any $k \ge 2$,

$$\mathbb{E}_\gamma\big(e_k(S_n)^{-1} | S_0 = 1\big) \sim \gamma^n/(k-2)!.$$

For $k = 2$, the "$\sim$" sign can be replaced by a "$\le$" sign. Thus, for instance,

$$\mathbb{E}\big((S_n(S_n+1))^{-1}\big) \le (2 - p_1)^{-n}/S_0^2.$$

A.3. *PCR.* Some general comments are in order, as regards the applications of these bounds in the PCR or iBMP context. First, the PCR reaction involves $\mathbb{P}^\varepsilon$. For a general $\mathbb{P}$, the iterations of our proof of iBMP bounds always converge. They yield bounds on the order of α^n if and only if $\alpha > r$, where

$$\alpha := 1 - (1 - m^{-1})\mu/\beta,$$

and where μ and β in (0; 1) are parameters of the model that are independent of the branching process generated by the offspring. See Piau (2001) for the infinite-target limit and Piau (2002) for the finite-target case when $\mathbb{P} = \mathbb{P}^\varepsilon$, and the rest of the present paper for a general $\mathbb{P}$. Thus, for $\mathbb{P}^\varepsilon$, the domain of validity of the α^n bounds is exactly $\mu < \beta$. For a general $\mathbb{P}$, if one uses a bound of the form

$$\mathbb{E}(S_n^{-1}) \le \kappa \rho^n / S_0,$$

the domain of validity of our PCR-like bounds on the order of α^n is $\alpha > \rho$. Thus, from Proposition A.2, a necessary condition is $\alpha > r$, that is,

$$\mu < \beta(1 - r)/(1 - m^{-1}).$$

This condition is met if $\mu < (1 - p_1)\beta$.

On the other hand, Proposition A.5 and Corollary A.8 allow for explicit values of κ. Since the iBMP bounds involve κ, one would prefer to use these. Starting from Corollary A.8, the condition becomes

$$\mu < \beta(1-(2-p_1)^{-1})/(1-m^{-1}),$$

which is satisfied if $\mu < \frac{1}{2}(1-p_1)\beta$. This is the meaning of Remark 14 in Section 6.

A.4. *Technicalities.* We start with simple convexity considerations. Proposition A.5 is part (ii) of Lemma A.11. Then we define the orbit of a given point under the action of f. Lemma A.13 provides estimates of this orbit.

LEMMA A.11. (i) $\mathbb{E}(L)^{-1}S_n^{-1} \le \mathbb{E}(S_{n+1}^{-1}|S_n) \le \mathbb{E}(L^{-1})S_n^{-1}$.

(ii) $m^{-n}S_0^{-1} \le \mathbb{E}(S_n^{-1}) \le \mathbb{E}(L^{-1})^n S_0^{-1}$.

(iii) $\mathbb{E}(S_n^{-1}|S_0) \le S_0^{-1}\mathbb{E}(S_n^{-1}|S_0=1)$.

DEFINITION A.12. Let $a_0 \in (0;1)$ and let $(a_k)_{k\in\mathbb{Z}}$ denote the orbit of a_0 with respect to the action of f, that is,

$$a_k := f^{(k)}(a_0),$$

where $f^{(k)}$ denotes the composition of f with itself k times if $k \ge 1$, and the composition of the inverse of f with itself $(-k)$ times if $k < 0$.

Since $f(t) < t$ for $t \in (0;1)$, $(a_k)_k$ is decreasing, $a_k \to 0$ when $k \to +\infty$ and $a_k \to 1$ when $k \to -\infty$.

LEMMA A.13. (i) *For any* $k \ge 0$, $a_k \ge a_0 p_1^k$ *and* $1-a_{-k} \ge (1-a_0)m^{-k}$.

(ii) *For any* $k \ge 0$, $a_k \le c_1 p_1^k$, *where* c_1 *depends only on* p_1.

(iii) *For any* $q > m^{-1}$, $1-a_{-k} \le q^k$ *for* k *large enough.*

(iv) *If* L *is square integrable, for any* $k \ge 0$, $1-a_{-k} \le c_2 m^{-k}$, *where* c_2 *depends only on* m *and* $\mathbb{E}(L^2)$.

(v) *If* $L\log L$ *is integrable, for any* $k \ge 0$, $1-a_{-k} \le c_3 m^{-k}$, *where* c_3 *depends only on the law of* L.

Careful computations give $c_1 \le e^{2/p_1}$ in (i) of Lemma A.13. Part (v) is given in Athreya and Ney (1972) and we omit the proof.

A.5. *Proofs.*

PROOF OF LEMMA A.11. Let L_i denote i.i.d. copies of L. Conditionally on $\{S_n = S\}$, S_{n+1} follows the law of $L_1 + \cdots + L_S$. Since t^{-1} is a convex function of $t > 0$,

$$\mathbb{E}(S_{n+1}^{-1}|S_n = S) \ge \mathbb{E}(S_{n+1}|S_n = S)^{-1} = S^{-1}\mathbb{E}(L)^{-1}.$$

On the other hand, $(t^{-1}+c)^{-1}$ is a concave function of $t>0$, for any $c>0$. For any i.i.d. random variables Y_i and Y, applying this to $t=Y_i^{-1}$ successively from $i=1$ to $i=S$, one gets

$$\mathbb{E}\big((Y_1+\cdots+Y_S)^{-1}\big)\le\mathbb{E}(Y^{-1})S^{-1}. \tag{A.1}$$

For $Y_i:=L_i$, (A.1) implies (i), hence (ii). Let S^i denote i.i.d. copies of S_n when $S_0=1$. Conditionally on $\{S_0=S\}$, S_n follows the law of $S^1+\cdots+S^S$. For $Y_i:=S^i$, (A.1) implies (iii). $\square$

PROOF OF LEMMA A.13. Part (i) stems from

$$f(t)\ge p_1t, \qquad 1-f(t)\le m(1-t).$$

As regards (ii), we first prove that, for any $q>p_1$, $a_k\le q^k$ for k large enough. To see this, note that $f(t)\sim p_1t$ when $t\to0$ implies that, for any $p_1<q'<q$, $f(t)\le q't$ for s small enough and that $a_{k+1}\le q'a_k$ for k large enough since $a_k\to0$.

Using this first step and the inequality $f(t)\le p_1t+(1-p_1)t^2$, one gets, for any $q>p_1$,

$$a_{k+1}\le p_1a_k+(1-p_1)a_k^2\le p_1a_k+(1-p_1)q^{2k}$$

for k large enough. Iterating this,

$$a_k\le p_1^{k-k_0}a_{k_0}+(1-p_1)\sum_{i=k_0}^{k-1}p_1^{k-1-i}q^{2i}.$$

If $q<\sqrt{p_1}$, the sum is bounded by a multiple of p_1^k; hence, (i) holds for k large enough. Changing the value of c_1 yields (i) for any $k\ge0$.

Starting from $1-f(t)\sim m(1-t)$ when $t\to1$, the proof of (iii) is similar to the first step of (ii). The proof of (iv) is similar to the second step of (ii), as follows. From $f''\le f''(1)=\mathbb{E}(L(L-1))=:2v^2$ and $f'(1)=m$, one gets

$$1-f(t)\ge m(1-t)-v^2(1-t)^2.$$

Setting $b_k:=1-a_{-k}$, this implies

$$b_{k-1}=f(b_k)\ge mb_k-v^2b_k^2.$$

Hence, for any $q>m^{-1}$,

$$b_k\le m^{-1}b_{k-1}+v^2m^{-1}b_k^2\le m^{-1}b_{k-1}+v^2m^{-1}q^{2k}$$

for k large enough. Iterating this and choosing $q<1/\sqrt{m}$ yields (iv). $\square$

PROOF OF PROPOSITION A.2. Part (i) stems from (ii) of Lemma A.11. As regards (ii), by a truncation argument, one can assume that $L\log L$ is integrable. From (iii) of Lemma A.11, it is enough to consider the case $S_0=1$. Then

$$\mathbb{E}(S_n^{-1})=\int_0^1 f_n(t)\frac{dt}{t}=\sum_{k\in\mathbb{Z}}A_k,$$

where

$$A_k := \int_{a_{k+1}}^{a_k} f_n(t)\frac{dt}{t}.$$

We denote by B_+, B_0 and B_- the sums of A_k for $k \geq 0$, $-n \leq k \leq 0$ and $k \leq -n$, respectively. Since $f_n(t)/t$ is nondecreasing, for any k,

$$A_k \leq f_n(a_k)a_k^{-1}(a_k - a_{k+1}).$$

Since $S_0 = 1$, $f_n(a_k) = f_0(f^{(n)}(a_k)) = a_{k+n}$. Hence,

$$A_k \leq a_{k+n}a_k^{-1}(a_k - a_{k+1}).$$

If $k \geq 0$, from (ii) of Lemma A.13 and from the fact that $a_{k+1} \geq p_1 a_k$,

$$B_+ \leq \sum_{k \geq 0} c_1(1 - p_1)p_1^{k+n} = c_1 p_1^n.$$

If $k \leq 0$, $a_k^{-1} \leq a_0^{-1}$ and $a_k - a_{k+1} \leq 1 - a_{k+1} \leq c_3 m^k$ from (v) of Lemma A.13. If $k \leq -n$, $a_{k+n} \leq 1$, then

$$B_- \leq \sum_{k \leq -n} a_0^{-1} c_3 m^k = cm^{-n}.$$

If $k \geq -n$, $a_{k+n} \leq c_1 p_1^{k+n}$. Thus,

$$B_0 \leq \sum_{-n \leq k \leq 0} a_0^{-1} c_3 c_1 p_1^{k+n} m^k.$$

If $p_1 \neq m^{-1}$, B_0 is bounded above by a multiple of the largest of p_1^n and m^{-n}. If $p_1 = m^{-1} = r$, B_0 is bounded above by a multiple of nr^n. This proves (ii) when $S_0 = 1$.

Finally, if $p_1 = m^{-1} = r$ and $S_0 \geq 2$, from $f_n(a_k) \leq a_{k+n}^2$ and the fact that $f_n(t)/t^2$ is nondecreasing,

$$A_k \leq a_{k+n}^2 a_k^{-2}(a_k - a_{k+1}).$$

The method above then shows that B_+, B_0 and B_- are bounded by multiples of r^{2n}, r^{2n} and r^n, respectively. Thus, the result holds for $S_0 = 2$. For any $S_0 \geq 2$, S_n is stochastically greater than the sum of $S_0/2$ i.i.d. copies of S_n conditioned by $S_0 = 2$. The result obtains from the analogue of (iii) in Lemma A.11. □

PROOF OF PROPOSITIONS A.6 AND A.7. One uses the following relations:

$$f^{\varepsilon}(t) = (1 - \varepsilon)t + \varepsilon t^2, \qquad f_\gamma \circ f_{\gamma'} = f_{\gamma\gamma'},$$

$$f_\gamma(t) = \frac{\gamma t}{1 - (1 - \gamma)t}, \qquad \mathbb{E}(S_n^{-1}) = \int_0^1 f_n(t)\frac{dt}{t}.$$

Since $f_\gamma^{(n)} = f_{\gamma^n}$, the change of variable $t \to u := f_\gamma^{(n)}(t)$ yields

$$\mathbb{E}_\gamma(S_n^{-1}) = \int_{t=0}^{t=1} u^{S_0} \frac{dt}{t} = \gamma^n \int_{u=0}^{u=1} u^{S_0 - 1} \frac{du}{w(u)},$$

where $w(u) := \gamma^n + (1 - \gamma^n)u$. An exact computation then yields the case $S_0 = 1$ of Proposition A.7, and the bound $w(u) \geq u$ yields the case $S_0 \geq 2$. □

PROOF OF COROLLARY A.8. The $S_0 \geq 2$ part stems from Propositions A.6 and A.7. Writing $\mathbb{E}_S$ for $\mathbb{E}(\cdot | S_0 = S)$, the $S_0 = 1$ part stems from

$$\mathbb{E}_1(S_{n+1}^{-1}) = \sum_{k \geq 1} p_k \mathbb{E}_k(S_n^{-1}) \leq p_1 \mathbb{E}_1(S_n^{-1}) + (1 - p_1)\mathbb{E}_2(S_n^{-1}),$$

since $\mathbb{E}_k(S_n^{-1}) \leq \mathbb{E}_2(S_n^{-1})$ for any $k \geq 2$, and from $\mathbb{E}_2(S_n^{-1}) \leq (2 - p_1)^{-n}$. □

REFERENCES

ATHREYA, K. B. (1994). Large deviation rates for branching processes I. Single type case. *Ann. Appl. Probab.* **4** 779–790.

ATHREYA, K. B. and NEY, P. E. (1972). *Branching Processes*. Springer, New York.

BIGGINS, J. (1977). Martingale convergence in the branching random walk. *J. Appl. Probab.* **14** 25–37.

BRUNNERT, M., MÜLLER, O. and URFER, W. (2000). Genetical and statistical aspects of polymerase chain reactions. Technical Report 6/2000, Univ. Dortmund.

GREY, D. (1980). A new look at convergence of branching processes. *Ann. Probab.* **8** 377–380.

HAYASHI, K. (1990). Mutations induced during the polymerase chain reaction. *Technique* **2** 216–217.

HEYDE, C. C. and BROWN, B. M. (1971). An invariance principle and some convergence rate results for branching processes. *Z. Wahrsch. Verw. Gebiete* **20** 271–278.

JACOBS, G., TSCHOLL, E., SEK, A., PFREUNDSCHUH, M., DAUS, H. and TRÜMPER, L. (1999). Enrichment polymerase chain reaction for the detection of Ki-ras mutations: Relevance of Taq polymerase error rate, initial DNA copy number, and reaction conditions on the emergence of false-positive mutant bands. *J. Cancer Res. Clin. Oncol.* **125** 395–401.

JOFFE, A. (1993). A new martingale in branching random walk. *Ann. Appl. Probab.* **3** 1145–1150.

KRAWCZAK, M., REISS, J., SCHMIDTKE, J. and ROSLER, U. (1989). Polymerase chain reaction: Replication errors and reliability of gene diagnosis. *Nucleic Acids Res.* **17** 2197–2201.

LYONS, R. (1997). A simple path to Biggins' martingale convergence for branching random walk. In *Classical and Modern Branching Processes* (K. B. Athreya and P. Jagers, eds.) 217–221. Springer, New York.

MARUYAMA, I. N. (1990). Estimation of errors in the polymerase chain reaction. *Technique* **2** 302–303.

NEY, P. E. and VIDYASHANKAR, A. N. (2001). Harmonic moments and large deviation rates for supercritical branching processes. Technical Report Stat2001-5, Univ. Georgia.

PAKES, A. G. (1975). Non-parametric estimation in the Galton–Watson process. *Math. Biosci.* **26** 1–18.

PIAU, D. (2001). Processus de branchement en champ moyen et réaction PCR. *Adv. in Appl. Probab.* **33** 391–403.

PIAU, D. (2002). Mutation–replication statistics for polymerase chain reactions. *J. Comput. Biol.* **9** 831–847.

REISS, J., KRAWCZAK, M., SCHLÖSSER, M., WAGNER, M. and COOPER, D. N. (1990). The effect of replication errors on the mismatch analysis of PCR-amplified DNA. *Nucleic Acids Res.* **18** 973–978.

SAIKI, R. K., GELFAND, D. H., STOFFEL, S., SCHARF, S. J., HIGUCHI, R., HORN, G. T., MULLIS, K. B. and ERLICH, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239** 487–491.

SUN, F. (1995). The polymerase chain reaction and branching processes. *J. Comput. Biol.* **2** 63–86.

WANG, D., ZHAO, C., CHENG, R. and SUN, F. (2000). Estimation of the mutation rate during error-prone polymerase chain reaction. *J. Comput. Biol.* **7** 143–158.

WEISS, G. and VON HAESELER, A. (1995). Modeling the polymerase chain reaction. *J. Comput. Biol.* **2** 49–61.

WEISS, G. and VON HAESELER, A. (1997). A coalescent approach to the polymerase chain reaction. *Nucleic Acids Res.* **25** 3082–3087.

LAPCS—LABORATOIRE DE PROBABILITÉS,
COMBINATOIRE ET STATISTIQUE
UNIVERSITÉ CLAUDE BERNARD LYON-I
DOMAINE DE GERLAND
50, AVENUE TONY-GARNIER
69366 LYON CEDEX 7
FRANCE
E-MAIL: Didier.Piau@univ-lyon1.fr