

NEW FRONTIERS IN APPLIED PROBABILITY

A Festschrift for SØREN ASMUSSEN
Edited by P. GLYNN, T. MIKOSCH and T. ROLSKI

Part 5. Stochastic growth and branching

POPULATION-SIZE-DEPENDENT, AGE-STRUCTURED BRANCHING PROCESSES LINGER AROUND THEIR CARRYING CAPACITY

PETER JAGERS, *Chalmers University of Technology and University of Gothenburg*
Mathematical Sciences, Chalmers University of Technology, Chalmers, SE-412 96 Gothenburg,
Sweden. Email address: jagers@chalmers.se

FIMA C. KLEBANER, *Monash University*
School of Mathematical Sciences, Monash University, Clayton, Victoria 3800, Australia.



APPLIED PROBABILITY TRUST
AUGUST 2011

POPULATION-SIZE-DEPENDENT, AGE-STRUCTURED BRANCHING PROCESSES LINGER AROUND THEIR CARRYING CAPACITY

BY PETER JAGERS AND FIMA C. KLEBANER

Abstract

Dependence of individual reproduction upon the size of the whole population is studied in a general branching process context. The particular feature under scrutiny is that of reproduction changing from supercritical in small populations to subcritical in large populations. The transition occurs when the population size passes a critical threshold, known in ecology as the carrying capacity. We show that populations either die out directly, never coming close to the carrying capacity, or grow quickly towards the carrying capacity, subsequently lingering around it for a time that is expected to be exponentially long in terms of a carrying capacity tending to infinity.

Keywords: Age structure; branching process; carrying capacity

2010 Mathematics Subject Classification: Primary 60J80
Secondary 92D25

1. Introduction

The notion of a *carrying capacity* is fundamental in biological population dynamics. It is vague but common knowledge that a virgin population either dies out while small or else it grows roughly exponentially until coming close to the capacity of its habitat. Then it levels away, and indeed is subcritical above the carrying capacity. Such behaviour underlies many phenomena in evolutionary biology and ecological theory, in particular in adaptive dynamics (see [3], [8], [15], [18], and [19], just to mention a few titles in the vast literature). Our purpose is to render this knowledge more precise.

In an earlier paper we studied the pattern in terms of a simple but nevertheless illustrative Galton–Watson-type process with binary splitting, i.e. individuals obtaining zero or two children after living one season, [14]. Let K denote the carrying capacity, and assume that the probability of an individual splitting into two is $p = K/(K+z)$, if the season population size is z . Clearly, the mean reproduction $2K/(K+z)$ decreases in z and passes 1 precisely at the carrying capacity; the process is supercritical below and subcritical above K .

This toy model can easily be rendered slightly more serious, or at least general, by replacing binary splitting by general, population-size-dependent reproduction with a mean m_z^K , larger than 1 only when $z < K$. From there, the step is short to age- and size-dependent Bellman–Harris branching processes where individuals do not necessarily have the same life span but live for random times independently, producing a random number of offspring in a population-size-dependent manner at death.

Such processes can be further generalised by allowing life spans to also be influenced by population size, through a hazard rate $h_z^K(a)$ of an a -aged individual in a population of size z .

(Not to complicate matters, we assume that life spans are continuous.) Pushing generality one step further, we may let the number of offspring at death depend upon the mother's age, a , thus obtaining a function $m_z^K(a)$ for the mean offspring. In a final step we may follow classical demographic theory and also allow (female) individuals to give birth according to age-specific birth rates, now to be taken as size dependent. What unites these types of process and renders them amenable to analysis is that they are *Markovian in the age structure* [10, p. 208].

The broadest possible framework would be completely general branching processes, supercritical below and subcritical above the carrying capacity, though some care has to be exercised in such descriptions, since the age distribution plays a role in the fertility of the population. Indeed, denote the expected reproduction of an individual of age a in a population of size z by $\mu_z(da)$. The most natural criticality concept might then be that of *annealed criticality*: a population is annealed critical at size z if and only if $\mu_z(\infty) = 1$. Being critical in this sense would, however, not exclude the population tending to increase or decrease for certain age configurations. That goal is attained if the population is *strictly critical*, meaning that $\mu_z(da) = L_z(da)$ for all a , where L_z is the individual life span distribution in a population of size z . Here and in the sequel, dependence upon K is often not spelled out.

We give a rigorous formulation of age- and population-size-dependent processes which are Markovian in the age structure and have well-defined intensities of birth and death. We intend to treat the general case elsewhere. However, the reader might note that the present framework encompasses virtually all classical population dynamics, such as age-dependent branching processes as well as birth-and-death processes with age-dependent intensities, and various deterministic frameworks, such as age-structured population dynamics. In the case of Bellman–Harris processes, where reproduction distributions are unaffected by the mother's age at death (but not of population size), conditions simplify substantially.

Consider a collection of individuals with ages $(a^1, \dots, a^z) = A$, where $z = |A|$ is the number of elements in A . It is convenient to regard the collection of ages A as a measure

$$A = \sum_{i=1}^z \delta_{a_i},$$

where δ_a denotes the point measure at a . As usual, the following notation is used for a function f on \mathbb{R} and measure A :

$$(f, A) = \int f(x)A(dx) = \sum_{i=1}^z f(a^i).$$

The expression on the right-hand side follows of course only if the measure is purely atomic.

In such a z -sized population, an individual of age a has a random life span with hazard rate $h_z(a)$. During life she gives birth to single children with intensity $b_z(a)$ at age a . If she dies, she splits into a random number of children which follows a distribution that may depend upon z . Its expectation is denoted by $m_z(a)$ and the second moment by $v_z(a)$, if the mother's age at death was a .

Let us emphasize that we take the basic parameters b_z, h_z , and m_z to be dependent upon the population size z , rather than upon the whole array A , because we wish to focus only upon essentials. The following general theory would work equally well if the parameters could be influenced by the age distribution at hand as well as by the population size. We shall also allow ourselves a slight inconsistency in notation, suffixing parameters sometimes by the population

size, sometimes by the whole age distribution A , and sometimes by the population ‘density’ $x = z/K$.

If there are no deaths and no births, the population changes only by ageing. When an individual dies, its point mass disappears and an offspring number of point masses appear at 0. Similarly, when she gives birth during life, a point mass appears at the origin. Thus, population evolution is given by a measure-valued process $\{A_t^K; t \geq 0\}$. Since process parameters depend upon K , there is a family of such processes indexed by K .

In terms of the intensities and the mean number of children at birth, the population is strictly critical at size z if and only if $b_z(a) + h_z(a)(m_z(a) - 1) = 0$ for all a . If $b_z = 0$ and m_z is independent of age, the process reduces to a Bellman–Harris age-dependent branching process with population-size dependence. Then strict and annealed criticality coincide, and reduce to the classical condition $m_z = 1$. Generally, strict criticality implies annealed. We note that there is a third concept of criticality: a population is *critical with respect to the age composition* A if and only if

$$(b_{|A|} + h_{|A|}(m_{|A|} - 1), A) = 0.$$

Clearly, strict criticality implies not only annealed but also composition criticality. In the next section we shall see that there is no trend to population change, neither to growth nor to decline, when the population is composition critical. Then change is random in the sense that it has a martingale character.

Under fairly general assumptions, we prove first that a small population either dies out directly, without approaching the carrying capacity, or else comes close to K , i.e. reaches any band $[(1 - \varepsilon)K, (1 + \varepsilon)K]$, $0 < \varepsilon < 1$, in a time of order $\log K$. Once the population size has reached such a level, it stays there for an exponentially long time, i.e. its expected persistence time is $O(e^{cK})$ for some $c > 0$. Usually, such results are derived from a large deviation principle, yielding the time to exit from a domain of attraction of a fixed point; cf. [7]. Here, we give a new proof using an exponential martingale inequality.

In the past, populations have been studied using measure-valued Markov processes with various setups; see, e.g. [2], [3], [5], [6, Section 9.4], [16], [17], and [18]. We take the state space to be the finite positive Borel measures on \mathbb{R}^+ with the topology of weak convergence, i.e. $\lim_{n \rightarrow \infty} \mu_n = \mu$ if and only if $\lim_{n \rightarrow \infty} \int f d\mu_n = \int f d\mu$ for any bounded and continuous function f on \mathbb{R}^+ . Métivier [16] and Borde-Boussion [2] embedded the space of measures into a weighted Sobolev space. Oelschläger [17] used yet another topology for the state space of the set of signed measures. Our model is closest to Oelschläger’s, and the French school around Méléard (cf. [3], [18], and other papers), but formulated in terms of branching rather than birth-and-death processes. As in this paper, Tran [18] added an age structure to the birth-and-death framework. He considered large populations, scaled them, and studied the limit behaviour. In the case of a size-independent birth rate, no splitting, and a logistic death rate (i.e. one of the form $h_z(a) = d(a) + \eta z$ for a fixed base-line death rate d plus a component proportional to size), he obtained results on large deviations from a limiting process.

In some literature branching populations are also allowed to diffuse according to Brownian motion [1]. Such ideas have had a substantial mathematical impact, giving rise to super Brownian motion and other super processes in the limit. From that point of view, our case is degenerate in the sense that no Laplacian is present in the process generator. Therefore, the linear operators that enter the description (they are not generators) are of first order only.

While we could have borrowed certain results from the abovementioned papers, we find it more transparent to give a self-contained treatment, especially since our model is given in a simpler setup and has an immediate biological background.

2. Age- and population-dependent processes

In this section we give definitions and basic properties to be used in the subsequent analysis. They consist essentially of the generator of our measure-valued Markov process and an integral representation known as Dynkin’s formula. The subscript A in P_A and E_A indicates that the population started from $z = (1, A)$ individuals, of ages $A = (a^1, \dots, a^z)$, at time $t = 0$. No subscript means that the population started from some implicit age configuration. The generator of a Markov measure-valued population-age-dependent branching process was given in [12]. As mentioned, we specialise to population-size-dependent parameters b_z, h_z, m_z , and v_z^2 , the latter two being defined as follows. Let $Y(a)$ denote the number of offspring an individual dying at age a splits into. Its expectation is $m_z(a)$ if the death and simultaneous splitting occur in a population of size z . The second entity denotes the corresponding variance.

Theorem 1. ([12].) *For a bounded differentiable function F on \mathbb{R}^+ and a continuously differentiable function f on \mathbb{R}^+ , the following limit exists:*

$$\lim_{t \rightarrow 0} \frac{1}{t} E_A \{ F((f, A_t)) - F((f, A)) \} = \mathfrak{G}F((f, A)).$$

Here

$$\begin{aligned} \mathfrak{G}F((f, A)) &= F'((f, A))(f', A) + \sum_{j=1}^z b_z(a^j) \{ F(f(0) + (f, A)) - F((f, A)) \} \\ &\quad + \sum_{j=1}^z h_z(a^j) [E_A \{ F(Y(a^j)f(0) + (f, A) - f(a^j)) \} - F((f, A))]. \end{aligned}$$

Consequently, Dynkin’s formula holds: for a bounded C^1 function F on \mathbb{R} and a C^1 function f on \mathbb{R}^+ ,

$$F((f, A_t)) = F((f, A_0)) + \int_0^t \mathfrak{G}F((f, A_s)) ds + M_t^{F,f}, \tag{1}$$

where $M_t^{F,f}$ is a local martingale with predictable quadratic variation

$$\langle M^{F,f}, M^{F,f} \rangle_t = \int_0^t \mathfrak{G}F^2((f, A_s)) ds - 2 \int_0^t F((f, A_s)) \mathfrak{G}F((f, A_s)) ds.$$

As a corollary, the following representation was also obtained; see [12].

Theorem 2. *For a C^1 function f on \mathbb{R}^+ ,*

$$(f, A_t) = (f, A_0) + \int_0^t (L_{Z_s} f, A_s) ds + M_t^f,$$

where $Z_s = (1, A_s) = |A|$ is the population size at time s , the linear operators L_z are defined by

$$L_z f = f' - h_z f + f(0)(b_z + h_z m_z), \tag{2}$$

and M_t^f is a local square integrable martingale with the sharp bracket given by

$$\langle M^f, M^f \rangle_t = \int_0^t (f^2(0)b_{Z_s} + f^2(0)v_{Z_s}^2 h_{Z_s} + h_{Z_s} f^2 - 2f(0)m_{Z_s} h_{Z_s} f, A_s) ds.$$

Corollary 1. *The special choice $f = 1$ yields*

$$Z_t = Z_0 + \int_0^t (b_{Z_s} + h_{Z_s}(m_{Z_s} - 1), A_s) ds + M_t^1.$$

Furthermore, from Theorem 2.3 of [12], it follows that, if $f \geq 0$ satisfies the (linear growth) condition, i.e.

$$(H1) \quad |(L_z f, A)| \leq C(1 + (f, A)) \text{ for some } C > 0 \text{ and any } A,$$

and if (f, A_0) is integrable, then so is (f, A_t) . Its expectation is bounded by

$$E(f, A_t) \leq (E(f, A_0) + Ct) \left(1 + \frac{e^{Ct}}{C} \right). \tag{3}$$

We use expectation without an index to indicate that the starting age configuration A_0 may well be random. Here and in the sequel, C denotes a constant, not necessarily the same in different contexts.

We call a family of functions f_z uniformly bounded if $\sup_{z,a} |f_z(a)| < \infty$. The following corollary is easy to check.

Corollary 2. *Suppose that the functions b_z , m_z , and h_z are uniformly bounded. If f and f' are bounded, then the growth condition (H1) is satisfied and conclusion (3) holds.*

In particular, the function $f(x) = 1$ satisfies condition (H1) and so

$$EZ_t \leq (EZ_0 + Ct) \left(1 + \frac{e^{Ct}}{C} \right).$$

Furthermore, if the functions v_z are uniformly bounded as well then M_t^f is a square integrable martingale with the quadratic variation

$$\langle M^f, M^f \rangle_t \leq C \int_0^t Z_s ds. \tag{4}$$

3. The time of ascent

Now consider a general population-size-dependent branching process, as described, starting at time $t = 0$ from z individuals. To ease notation, we take all individuals to be newborns. Such a population must die out eventually [11]. What are the chances that the population will reach a size in the vicinity of the carrying capacity, before ultimately dying out? We write T for the time to extinction and T_d for the time it takes the population to first attain a size greater than or equal to dK , $0 < d < 1$; the number of ancestors $z < dK$. Clearly,

$$T < T_d \implies Z_t < dK \text{ for all } t.$$

Now assume that reproduction decreases with increasing population. A tilded variable will pertain to a population-size-independent branching process with the fixed parameters b_{dK} , h_{dK} , and m_{dK} . Then

$$P(T < T_d) \leq P(\tilde{T} < \infty) = \tilde{q}^z,$$

where \tilde{q} is the extinction probability of the population-size-independent branching process. If $\tilde{m}_d > 1$ and $\tilde{\sigma}_d^2$ denote the mean and variance of the total reproduction of an individual in

this latter process, we find, by Haldane’s inequality (see [9, p. 125]) that the probability of the original population never reaching dK is

$$P(T < T_d) \leq \left(1 - \frac{2(\tilde{m}_d - 1)}{\tilde{\sigma}_d^2 + \tilde{m}_d(\tilde{m}_d - 1)}\right)^z.$$

With a positive chance, the population will thus reach a size of order K . Since it grows quicker than the process \tilde{Z}_t while under the level dK , and the latter process grows exponentially, we can conclude that dK will be attained after a time of order $\log K$.

Theorem 3. *If reproduction decreases with population size and the population is annealed critical at K , then any population size dK , $0 < d < 1$, is attained with positive probability within a time $T_d = O(\log K)$ as $K \rightarrow \infty$.*

4. The time of pseudo-stability

We proceed to show that once the population size has reached the interval $[K - \varepsilon K, K + \varepsilon K]$ it remains there an exponentially long time. In other words, in terms of the process scaled by K , it takes exponentially long to exit from $[1 - \varepsilon, 1 + \varepsilon]$. The property ensuring this lingering around the carrying capacity is that the population reproduces subcritically above level K , supercritically below level k , and critically at K . First we shall see that this property forces the scaled population size to converge to 1.

Criticality is understood in the strict sense. Define the criticality function in terms of the operator L defined in (2):

$$\chi_A = b_A + h_A(m_A - 1) = L_A 1.$$

Then criticality means that $\chi_A(a) = 0$ for all a , as soon as $|A| = K$. Indeed, we assume that dependence on the population is through the scaled population size $x = z/K$. Then any A with total mass K , or scaled mass 1, is a criticality point and we write, abusing notation,

$$\chi_A = \chi_K = \chi_1 = 0.$$

The dependence of parameters and variables on the carrying capacity is usually not spelled out, sometimes indicated by a superscript K .

Finally, assume that χ satisfies a Lipschitz condition in the neighbourhood of 1.

Assumption 1. *There is a constant C such that*

$$|\chi_x| = |\chi_x - \chi_1| \leq C|x - 1| \tag{5}$$

for x close to 1.

Theorem 4. *Suppose that $X_0^K \rightarrow 1$ in probability as $K \rightarrow \infty$, and that Assumption 1 holds. Then $X_t^K = Z_t^K / K$, i.e. the total population size scaled by the carrying capacity, converges in probability to 1, uniformly on any time interval $[0, T]$, $T > 0$. In other words, for any $\eta > 0$,*

$$\lim_{K \rightarrow \infty} P\left(\sup_{t \leq T} |X_t^K - 1| > \eta\right) = 0.$$

Proof. By Corollary 1,

$$X_t^K = X_0^K + \int_0^t \left(\chi_{X_s^K}, \frac{1}{K} A_s^K \right) ds + \frac{1}{K} M_t^{1,K}. \tag{6}$$

In terms of the reproduction variance $\sigma_x^2 = v_x - m_x^2$, the martingale quadratic variation reduces to

$$\left\langle \frac{1}{K} M^{1,K}, \frac{1}{K} M^{1,K} \right\rangle_t = \frac{1}{K} \int_0^t \left((b_{X_s^K} + \sigma_{X_s^K}^2 + (m_{X_s^K} - 1)^2) h_{X_s^K}, \frac{1}{K} A_s^K \right) ds.$$

First we show that

$$\sup_{t \leq T} X_t^K \leq \left(X_0^K + \sup_{t \leq T} \frac{1}{K} |M_t^{1,K}| \right) e^{CT}. \tag{7}$$

Indeed, since the criticality function is bounded by some $C > 0$, and $(1, A_s^K/K) = X_s^K \leq \sup_{u \leq s} X_u^K$,

$$\begin{aligned} \sup_{t \leq T} X_t^K &\leq X_0^K + \frac{1}{K} \int_0^T |(\chi_{X_s^K} 1, A_s^K)| ds + \sup_{t \leq T} \frac{1}{K} |M_t^{1,K}| \\ &\leq X_0^K + \frac{1}{K} \int_0^T C(1, A_s^K) ds + \sup_{t \leq T} \frac{1}{K} |M_t^{1,K}| \\ &\leq X_0^K + \sup_{t \leq T} \frac{1}{K} |M_t^{1,K}| + C \int_0^T \sup_{u \leq s} X_u^K ds. \end{aligned}$$

Gronwall’s inequality in the form given in [13, p. 21] yields (7).

Next, by Doob’s inequality,

$$\begin{aligned} \mathbb{P} \left(\sup_{t \leq T} \frac{1}{K} |M_t^{1,K}| > a \right) &\leq \frac{1}{a^2} \mathbb{E} \left\langle \frac{1}{K} M^{1,K}, \frac{1}{K} M^{1,K} \right\rangle_T \\ &\leq \frac{C}{a^2 K} \int_0^T \mathbb{E} X_t^K dt \\ &\leq \frac{C_1 \mathbb{E} X_0^K}{a^2 K} \\ &\rightarrow 0, \end{aligned}$$

where the bound from (4) was used. Hence,

$$\sup_{t \leq T} \frac{1}{K} M_t^{1,K} \xrightarrow{P} 0 \quad \text{as } K \rightarrow \infty. \tag{8}$$

We have, from Corollary 1,

$$\begin{aligned} \sup_{t \leq T} |X_t^K - 1| &\leq |X_0^K - 1| + \int_0^T \left| \left(\chi_{X_s^K}, \frac{1}{K} A_s^K \right) \right| ds + \sup_{t \leq T} \frac{1}{K} |M_t^{1,K}| \\ &\leq |X_0^K - 1| + \sup_{t \leq T} \frac{1}{K} |M_t^{1,K}| + C \sup_{t \leq T} X_t^K \int_0^T \sup_{u \leq s} |X_s^K - 1| ds, \end{aligned}$$

thanks to the Lipschitz condition (5). Using Gronwall’s inequality, we again conclude that

$$\begin{aligned} \sup_{t \leq T} |X_t^K - 1| &\leq \left(|X_0^K - 1| + \sup_{t \leq T} \frac{1}{K} |M_t^{1,K}| \right) \exp \left[C \sup_{t \leq T} X_t^K \right] \\ &\leq \left(|X_0^K - 1| + \sup_{t \leq T} \frac{1}{K} |M_t^{1,K}| \right) \exp \left[C \left(X_0^K + \sup_{t \leq T} \frac{1}{K} |M_t^{1,K}| \right) e^{CT} \right], \end{aligned}$$

where we have used (7) to obtain the second inequality. Since $X_0^K - 1$ converges in probability to 0 by assumption and $\sup_{t \leq T} |M_t^{1,K}|/K$ converges in probability to 0 by (8), the exponential term converges to 1, and the statement follows.

An exponential bound on the exit time from the vicinity of K requires exponential moments of the process. Hence, we assume bounded exponential moments of offspring distributions. Then the process $Z_t^K = (1, A_t^K)$ has exponential moments as well. At this junction, let $\phi_A(t)(a) = E_A\{e^{tY(a)}\}$ denote the conditional moment generating function given A of the number $Y(a)$ of offspring at the death of an a -aged individual splitting in a population with age composition A . Similarly, in this proof P_A denotes offspring probabilities in a population of size $|A|$ and composition A .

The following condition (see [18, Assumption 3]) may seem strange at first sight, but it serves to give the subcriticality above the carrying capacity a strict form. We let $\phi_A(t)$ denote the function $\phi_A(t)(\cdot)$.

Assumption 2. For any K , there exists a population size $V_K > K$ such that

$$(e^{1/K} - 1)b_A + \left(\phi_A \left(\frac{1}{K} \right) e^{-1/K} - 1 \right) h_A \leq 0 \quad \text{whenever } (1, A) > V_K,$$

and V_K/K is bounded for large K .

Since the reproduction is subcritical for population sizes larger than K , such a number exists. Indeed, for large K and any $a \geq 0$,

$$(e^{1/K} - 1)b_A(a) + \left(\phi_A \left(\frac{1}{K} \right) (a) e^{-1/K} - 1 \right) h_A(a) \sim \frac{1}{K} (b_A(a) + (m_A(a) - 1)h_A(a)),$$

which is negative for large $(1, A)$. The assumption needed is that the expression turns negative not too far away from K , when K is also large. An example is provided by the binary splitting with $b = 0$ and $Y = Y(a)$ independently of the splitting age, mentioned in the introduction and further explored in [14]: $P_A(Y = 2) = K/(K + z)$, $z = (1, A)$. Then V_K is determined from

$$\frac{z}{K + z} e^{-1/K} + \frac{K}{K + z} e^{1/K} = E_z e^{(Y-1)/K} = 1.$$

Solving in z gives $V_K = e^{1/K} K$.

Theorem 5. Let X_t^K be the population size scaled by the carrying capacity K . Suppose that all exponential moments of the offspring number at splitting exist and that Assumptions 1 and 2 are in force. Then, there is a constant C , independent of K , such that, for any t ,

$$E e^{X_t^K} \leq E e^{X_0^K} e^{Ct}. \tag{9}$$

Furthermore,

$$E \exp \left[\int_0^1 X_s^K ds \right] \leq C E e^{X_0^K}. \tag{10}$$

Proof. We consider a process for fixed K , and in order to ease notation, we suppress K in the proof of this result. The statement follows by taking F in (1) as the exponential function, or rather, to be precise, letting it equal smooth bounded functions that agree with the exponential on bounded intervals and making use of a localising sequence $T_n = \inf\{t : Z_t > n\}$.

With $x = z/K$ we have

$$\begin{aligned} g(F(1, A)) &= F'((f, A))(f', A) + \sum_{j=1}^z b_A(a^j)[F(f(0) + (f, A)) - F((f, A))] \\ &\quad + \sum_{j=1}^z h_A(a^j)[E\{F(Y_A f(0) + (f, A)) - f(a^j)\} - F((f, A))] \\ &= (e^{x+1/K} - e^x) \sum_{j=1}^z b_A(a^j) + \sum_{j=1}^z h_A(a^j)(E\{e^{x+(Y_A-1)/K}\} - e^x) \\ &= e^x((e^{1/K} - 1)(b_A, A) + (E\{e^{(Y_A-1)/K}\} - 1)(h_A, A)). \end{aligned}$$

Hence, we obtain, by (1),

$$e^{X_t} = e^{X_0} + \int_0^t e^{X_s} ((e^{1/K} - 1)b_{A_s} + (E\{e^{(Y_{A_s}-1)/K}\} - 1)h_{A_s}, A_s) ds + M_t^{\text{exp}},$$

where M_t^{exp} is a local martingale. Localising and taking the expectation,

$$E e^{X_{t \wedge T_n}} = E e^{X_0} + E \int_0^{t \wedge T_n} e^{X_s} ((e^{1/K} - 1)b_{A_s} + (E\{e^{(Y_{A_s}-1)/K}\} - 1)h_{A_s}, A_s) ds.$$

Now we use the facts that the reproduction is subcritical above K , that the parameters b_A and h_A are bounded, and that the function under the integral is negative for values of $Z_s > V_K$ or $X_s > V_K/K$. For $Z_s < V_K$, the inequalities $e^{1/K} - 1 \leq C/K$ and $|\phi_A^K(1/K)e^{-1/K} - 1| \leq C/K$ show that the integrand does not exceed $C V_K/K$. We have

$$\begin{aligned} E e^{X_{t \wedge T_n}} &\leq E e^{X_0} + C \frac{V_K}{K} E \int_0^t e^{X_{s \wedge T_n}} \mathbf{1}(Z_s \leq V_K) ds \\ &\leq E e^{X_0} + CE \int_0^t e^{X_{s \wedge T_n}} ds, \end{aligned}$$

where C is a constant independent of K , since V_K/K is assumed to be bounded. Gronwall's inequality yields

$$E e^{X_{t \wedge T_n}} \leq E e^{X_0} e^{Ct}, \tag{11}$$

where C does not depend on K . Letting $n \rightarrow \infty$, we obtain (9).

To prove (10), we use Jensen's inequality for the uniform distribution on $[0, 1]$ combined with the exponential function: for any integrable function g on $[0, 1]$, $\int_0^1 e^{g(s)} ds \geq \exp[\int_0^1 g(s) ds]$. Applying this inequality with $g(s) = X_s^K$, and using (11), we obtain $E \exp[\int_0^1 X_s^K ds] \leq E \int_0^1 e^{X_s^K} ds \leq CE e^{X_0^K}$.

Theorem 6. Assume that $X_0^K \rightarrow 1$ in probability. For any $\varepsilon > 0$, let

$$\tau^K = \inf\{t : |X_t^K - 1| > \varepsilon\}.$$

Suppose that the previous assumptions hold and also that the number of offspring through splitting at death is bounded by some constant. Then $E\{\tau^K\}$ is exponentially large in K , i.e. for some positive constants C and c ,

$$E\{\tau^K\} > Ce^{cK}.$$

Proof. First we establish an exponential bound for the probability of exit up to time 1, when the normed population started at $x \in (1 - \eta, 1 + \eta)$ for $\eta < \varepsilon/6$:

$$P_x(\tau^K \leq 1) = P_x\left(\sup_{t \leq 1} |X_t^K - 1| > \varepsilon\right) \leq e^{-CK}. \tag{12}$$

Write the dynamics of X_t^K as in (6), i.e.

$$X_t^K = X_0^K + \int_0^t (\chi_{X_s^K}, A_s^K) ds + \frac{1}{K} M_t^{1,K} = x + I_t^K + \frac{1}{K} M_t^{1,K},$$

where $I_t^K = \int_0^t (\chi_{X_s^K}, A_s^K) ds$. Since $|X_t^K - 1| \leq |X_t^K - x| + |x - 1| \leq \eta + |X_t^K - x|$, we have

$$\begin{aligned} P_x\left(\sup_{t \leq 1} |X_t^K - 1| > \varepsilon\right) &\leq P_x\left(\sup_{t \leq 1} |I_t^K + \eta| > \frac{\varepsilon}{2}\right) + P_x\left(\sup_{t \leq 1} \left|\frac{1}{K} M_t^{1,K}\right| > \frac{\varepsilon}{2}\right) \\ &\leq P_x\left(\sup_{t \leq 1} |I_t^K| > \frac{2\varepsilon}{3}\right) + P_x\left(\sup_{t \leq 1} \left|\frac{1}{K} M_t^{1,K}\right| > \frac{\varepsilon}{2}\right). \end{aligned}$$

As $X_0^K \rightarrow 1, X_t^K \rightarrow 1$ by Theorem 4, and, consequently, $\chi_{X_t^K} \rightarrow \chi_1 = 0$, with all convergences taking place in probability. Hence, $(\chi_{X_s^K}, A_s^K) = o(K)$ in the same sense, and we have, by the exponential form of Chebyshev’s inequality,

$$\begin{aligned} P_x\left(\sup_{t \leq 1} |I_t^K| > \frac{2\varepsilon}{3}\right) &\leq P_x\left(\int_0^1 |(\chi_{X_s^K}, A_s^K)| ds > \frac{2\varepsilon K}{3}\right) \\ &\leq e^{-2\varepsilon K/3} E_x \exp\left[\int_0^1 |(\chi_{X_s^K}, A_s^K)| ds\right] \\ &= e^{-2\varepsilon K/3 + o(K)} \\ &\leq e^{-CK} \quad \text{for some } C. \end{aligned}$$

The second probability $P_x(\sup_{t \leq 1} |M_t^{1,K}| > K\varepsilon/2)$ is controlled by an exponential martingale inequality (see, e.g. [4, Lemma 4.2]), i.e.

$$P_x\left(\sup_{t \leq T} |M_t^{1,K}| > \varepsilon, \langle M^{1,K}, M^{1,K} \rangle_T \leq q\right) \leq 2e^{-\varepsilon^2/(B\varepsilon + q)},$$

where B is the maximal number of children at splitting, which is a bound on the jumps of $M^{1,K}$. Hence, replacing ε by $K\varepsilon/2$,

$$P_x\left(\sup_{t \leq 1} \frac{1}{K} |M_t^{1,K}| > \frac{\varepsilon}{2}\right) \leq 2e^{-(\varepsilon^2 K^2/2)/(BK\varepsilon + 2q)} + P_x\left(\left\langle \frac{1}{K} M^{1,K}, \frac{1}{K} M^{1,K} \right\rangle_1 > q\right). \tag{13}$$

Since the parameter functions are uniformly bounded, the quadratic variation is bounded by

$$\left\langle \frac{1}{K} M^{1,K}, \frac{1}{K} M^{1,K} \right\rangle_t \leq \frac{C}{K} \int_0^t X_s^K ds,$$

by Corollary 2. From Chebyshev’s inequality and bound (10),

$$\begin{aligned} P_x\left(\left\langle \frac{1}{K}M^{1,K}, \frac{1}{K}M^{1,K} \right\rangle_1 > q\right) &\leq P_x\left(\frac{C}{K} \int_0^1 X_s^K ds > q\right) \\ &\leq e^{-cK} E_x \exp\left[\int_0^1 X_s^K ds\right] \\ &\leq Ce^{-cK}. \end{aligned}$$

Hence, we obtain, from (13), $P_x(\sup_{t \leq 1} |M_t^{1,K}|/K > \varepsilon/2) \leq Ce^{-c(\varepsilon)K}$, and gathering terms gives (12). The final step is a recursive argument, formulated in terms of the filtration $\{\mathcal{F}_n := \sigma(\{A_t^K, t \leq n\})\}$:

$$\begin{aligned} P(\tau^K > n) &= P\left(\sup_{t \leq n} |X_t^K - 1| < \varepsilon\right) \\ &= P\left(\sup_{t \leq n-1} |X_t^K - 1| < \varepsilon, \sup_{n-1 \leq t \leq n} |X_t^K - 1| < \varepsilon\right) \\ &\geq P\left(\sup_{t \leq n-1} |X_t^K - 1| < \varepsilon, \sup_{n-1 \leq t \leq n} |X_t^K - 1| < \varepsilon, |X_{n-1}^K - 1| < \eta\right) \\ &= E\left\{P\left(\left(|X_{n-1}^K - 1| < \eta, \sup_{n-1 \leq t \leq n} |X_t^K - 1| < \varepsilon \mid \mathcal{F}_{n-1}\right)\right); \right. \\ &\quad \left. \sup_{t \leq n-1} |X_t^K - 1| < \varepsilon\right\} \\ &\geq \inf_{x \in (1-\eta, 1+\eta)} P_x\left(\sup_{t \leq 1} |X_t^K - 1| < \varepsilon\right)P(\tau > n - 1) \\ &\geq \left(\inf_{x \in (1-\eta, 1+\eta)} P_x\left(\sup_{t \leq 1} |X_t^K - 1| < \varepsilon\right)\right)^n. \end{aligned}$$

So

$$\begin{aligned} E\tau^K &> \sum_n P(\tau^K > n) \\ &> \sum_n \left(\inf_{x \in (1-\eta, 1+\eta)} P_x\left(\sup_{t \leq 1} |X_t^K - 1| < \varepsilon\right)\right)^n \\ &= \frac{1}{1 - \inf_{x \in (1-\eta, 1+\eta)} P_x(\sup_{t \leq 1} |X_t^K - 1| < \varepsilon)} \\ &= \frac{1}{\sup_{x \in (1-\eta, 1+\eta)} P_x(\sup_{t \leq 1} |X_t^K - 1| > \varepsilon)} \\ &> Ce^{cK}. \end{aligned}$$

Acknowledgements

This research was supported by the Swedish Research Council and the Australian Research Council under grant DP0881011. The authors are grateful to Serik Sagitov for penetrating discussions, to the anonymous referee for a number of improvements, and to Géza MeszÉna for suggesting that we should pursue this generalisation of our approach in [14].

References

- [1] ASMUSSEN, S. AND HERING, H. (1983). *Branching Processes*. Birkhäuser, Boston, MA.
- [2] BORDE-BOUSSION, A.-M. (1990). Stochastic demographic models: age of a population. *Stoch. Process. Appl.* **35**, 279–291.
- [3] CHAMPAGNAT, N., FERRIÈRE, R. AND MÉLÉARD, S. (2008). From individual stochastic processes to macroscopic models in adaptive evolution. *Stoch. Models* **24**, 2–44.
- [4] CHIGANSKY, P. AND LIPTSER, R. (2010). Moderate deviations for a diffusion-type process in a random environment. *Theory Prob. Appl.* **54**, 29–50.
- [5] DAWSON, D. A. (1993). Measure-valued Markov processes. In *École d'Eté de Probabilités de Saint-Flour XXI* (Lecture Notes Math. **1541**), Springer, Berlin, pp. 1–260.
- [6] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes*. John Wiley, New York.
- [7] FREIDLIN, M. I. AND WENTZELL, A. D. (1998). *Random Perturbations of Dynamical Systems*. Springer, New York.
- [8] GERITZ, S. A. H., KISDI, É., MESZÉNA, G. AND METZ, J. A. J. (1998). Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecology* **12**, 35–57.
- [9] HACCOU, P., JAGERS, P. AND VATUTIN, V. A. (2005). *Branching Processes: Variation, Growth, and Extinction of Populations*. Cambridge University Press.
- [10] JAGERS, P. (1975). *Branching Processes with Biological Applications*. Wiley-Interscience, London.
- [11] JAGERS, P. (1992). Stabilities and instabilities in population dynamics. *J. Appl. Prob.* **29**, 770–780.
- [12] JAGERS, P. AND KLEBANER, F. C. (2000). Population-size-dependent and age-dependent branching processes. *Stoch. Process. Appl.* **87**, 235–254.
- [13] KLEBANER, F. C. (2005). *Introduction to Stochastic Calculus with Applications*, 2nd. edn. Imperial College Press, London.
- [14] KLEBANER, F. C. *et al.* (2011). Stochasticity in the adaptive dynamics of evolution: the bare bones. *J. Biol. Dynamics* **5**, 174–162.
- [15] MÉLÉARD, S. AND TRAN, V. C. (2009). Trait substitution sequence process and canonical equation for age-structured populations. *J. Math. Biol.* **58**, 881–921.
- [16] MÉTIVIER, M. (1987). Weak convergence of measure valued processes using Sobolev imbedding techniques. In *Stochastic Partial Differential Equations and Applications* (Trento, 1985; Lecture Notes Math. **1236**), Springer, Berlin, pp. 172–183.
- [17] OELSCHLÄGER, K. (1990). Limit theorems for age-structured populations. *Ann. Prob.* **18**, 290–318.
- [18] TRAN, V. C. (2008). Large population limit and time behaviour of a stochastic particle model describing an age-structured population. *ESAIM Prob. Statist.* **12**, 345–386.
- [19] WAXMAN, D. AND GAVRILETS, S. (2005). 20 Questions on adaptive dynamics. *J. Evol. Biol.* **18**, 1139–1154.

PETER JAGERS, *Chalmers University of Technology and University of Gothenburg*

Mathematical Sciences, Chalmers University of Technology, Chalmers, SE-412 96 Gothenburg, Sweden.

Email address: jagers@chalmers.se

FIMA C. KLEBANER, *Monash University*

School of Mathematical Sciences, Monash University, Clayton, Victoria 3800, Australia.