Abstract. The Bare Bones model provides a simple approach to the adaptive population dynamics of binary splitting cells, allowing replication to depend on population size, cf. [10]. Here we use it to analyse how a population stemming from one single mutant, appearing in a large wild-type population, can establish itself. By “large” we then mean that the wild-type population has a size of the order of the carrying capacity $K$. Under appropriate conditions, the mutant may then establish itself after a time of order $\log K$, as $K \to \infty$. The densities (i.e. sizes divided by the carrying capacity) of the wild-type and mutant populations can be approximated by a deterministic curve with a random initial condition. We give a precise description of this random initial condition in terms of the scaling limit $H$ of the corresponding non-linear two-dimensional deterministic dynamics and a random variable $W$ appearing as an almost sure limit of the initial binary splitting process of the mutant. This complements the “random shift” approximation in [3], and gives a first result for the simultaneous development of two competing subpopulations.

1. Introduction

1.1. Evolutionary model. There has been much work in stochastic adaptive dynamics and evolutionary branching, see [4], [5], [13] to mention just a few. Here we consider a specific question, the establishment of a mutant subpopulation in a large wild-type population, using the Bare Bones model, introduced in [10]. This is a density-dependent binary splitting, in which each subpopulation, the mutant and the original or wild-type, has its own carrying capacity. We assume that these are large and of comparable size, thus writing $a_1 K$ for the first, and $a_2 K$ for the second population, where the parameter $K$ is large, and $a_1, a_2$ are positive constants.

We assume that the established original population is in what could be called macroscopic equilibrium around its carrying capacity $a_1 K$. This means that it starts at the value, $Z_0^{(1)} = a_1 K$, where reproduction is assumed critical. Then one new cell, the mutant, appears and initiates its own population, $Z_0^{(2)} = 1$. Each population develops by binary splitting with probabilities dependent on the numbers of cells, with transitions from generation $n$ to $n + 1$ described by the recursion

$$Z_{n+1} = \left( Z_{n+1}^{(1)}, Z_{n+1}^{(2)} \right) = \left( \sum_{k=1}^{Z_n^{(1)}} \xi^{(1)}_{nk}, \sum_{k=1}^{Z_n^{(2)}} \xi^{(2)}_{nk} \right)$$

(1.1)
where \( \xi^{(1)}_{nk}, k = 1, 2, \ldots \) are conditionally independent random variables, given the preceding generation \( Z_n \), with probabilities

\[
\begin{align*}
\mathbb{P}\left( \xi^{(1)}_{nk} = 2 | Z_n \right) &= 1 - \mathbb{P}\left( \xi^{(1)}_{nk} = 0 | Z_n \right) = \frac{a_1K}{a_1K + Z_n^{(1)} + \gamma Z_n^{(2)}}, \\
\mathbb{P}\left( \xi^{(2)}_{nk} = 2 | Z_n \right) &= 1 - \mathbb{P}\left( \xi^{(2)}_{nk} = 0 | Z_n \right) = \frac{a_2K}{a_2K + \gamma Z_n^{(1)} + Z_n^{(2)}}.
\end{align*}
\]  

In the absence of mutants at the initial point the established population thus follows critical reproduction, as advertised, whereas the mutant population starts supercritically provided \( a_2 > \gamma a_1 \). This is assumed throughout the paper.

This branching process has two competing types, but it is not genuinely multi-type, because none of them can produce the other trait. The interaction coefficient \( \gamma \) is assumed to satisfy \( 0 < \gamma < 1 \). This means that cells of one type encroach less upon the reproduction of the other cell type than do cells of the same type. That \( \gamma \) is the same in both definitions means that influence is symmetric between the cell types. We could have worked with different interaction coefficients, \( \gamma_1, \gamma_2 \), but refrain from this in order to exhibit the basic pattern undisturbed. Biologically, this amounts to symmetry in the relation between cells of the two different types.

1.2. Randomly perturbed dynamics. A better insight into the invasion stage is provided by the density process, \([8, 9]\)

\[
X_n = (X_n^{(1)}, X_n^{(2)}) = (Z_n^{(1)}/K, Z_n^{(2)}/K),
\]

that is, the population numbers relative to \( K \). Note that the offspring distributions in (1.2) are in fact functions of the density; at \( x = (x_1, x_2) \)

\[
\begin{align*}
\mathbb{P}\left( \xi^{(1)}_{nk} = 2 | X_n = x \right) &= \frac{a_1}{a_1 + x_1 + \gamma x_2}, \\
\mathbb{P}\left( \xi^{(2)}_{nk} = 2 | X_n = x \right) &= \frac{a_2}{a_2 + \gamma x_1 + x_2}.
\end{align*}
\]

We denote the offspring mean at \( x \) by \( m(x) = (m_1(x), m_2(x)) \),

\[
\begin{align*}
m_1(x) &= \mathbb{E}\left( \xi^{(1)}_{nk} | X_n = x \right) = \frac{2a_1}{a_1 + x_1 + \gamma x_2}, \\
m_2(x) &= \mathbb{E}\left( \xi^{(2)}_{nk} | X_n = x \right) = \frac{2a_2}{a_2 + \gamma x_1 + x_2}.
\end{align*}
\]

In terms of these, the density process can be represented as randomly perturbed deterministic dynamics:

\[
\begin{align*}
X_{n+1}^{(1)} &= X_n^{(1)} + m_1(X_n) + \frac{1}{K} \sum_{j=1}^{K} \left( \xi^{(1)}_{nj} - m_1(X_n) \right), \\
X_{n+1}^{(2)} &= X_n^{(2)} + m_2(X_n) + \frac{1}{K} \sum_{j=1}^{K} \left( \xi^{(2)}_{nj} - m_2(X_n) \right),
\end{align*}
\]

(1.3)

the deterministic dynamics given by \( f(x) = (f_1(x), f_2(x)) \) with \( f_i(x) = x m_i(x), i = 1, 2, \) so that

\[
\begin{align*}
f_1(x) &= \frac{2x_1a_1}{a_1 + x_1 + \gamma x_2}, \\
f_2(x) &= \frac{2x_2a_2}{a_2 + \gamma x_1 + x_2},
\end{align*}
\]
and

\[ x_{n+1} = f(x_n), \quad (1.4) \]

subject to the initial condition \( x_0 = (a_1, 1/K) \).

It is easy to see that the noise term in the stochastic system (1.3) is of order \( 1/\sqrt{K} \). Indeed, given \( X_n = x \),

\[ \eta^{(i)}_{n+1}(x) = \frac{1}{\sqrt{K}} \sum_{j=1}^{Kx_i} (\xi^{(i)}_{nj} - m_i(x)), \quad i = 1, 2 \]

has zero mean and a conditional variance which is bounded in \( K \). Hence we can write (1.3) as the system

\[ X_{n+1} = f(X_n) + \frac{1}{\sqrt{K}} \eta^{(i)}_{n+1}(X_n), \quad (1.5) \]

with the initial condition \( X_0 = (a_1, 1/K) \). Note that the stochastic dynamics, as well as the initial condition, depends on \( K \).

The function \( f \) in (1.4), that generates the deterministic dynamical system, has four fixed points, i.e. solutions to \( f(x) = x \):

\[ x^{(0)} = (0, 0), \quad x^{(1)} = (a_1, 0), \quad x^{(2)} = (0, a_2), \quad x^{(*)} = \left( \frac{a_1 - \gamma a_2}{1 - \gamma^2}, \frac{a_2 - \gamma a_1}{1 - \gamma^2} \right), \quad (1.6) \]

where both coordinates of \( x^{(*)} \) are positive, if the following coexistence condition holds:

\[ a_1 - \gamma a_2 > 0, \quad a_2 - \gamma a_1 > 0. \quad (C) \]

An elementary analysis of the Jacobian matrix \( \nabla f(x) \) shows that \( x^{(*)} \) and \( x^{(0)} \) are stable and unstable fixed points respectively of the linearized dynamics, and \( x^{(1)} \) and \( x^{(2)} \) its saddle points.

1.3. The large capacity limit. The classical result in perturbation theory of dynamical systems, see e.g. \([7]\) and \([11]\), asserts that the trajectory \( X^K_n \) of the stochastic dynamics (1.5) converges to that of the deterministic system (1.4), started from the initial condition \( x_0 = \lim_{K \to \infty} X^K_0 \):

\[ \max_{n \leq N} \left| X^K_n - x_n \right| \overset{p}{\to} 0, \quad N \in \mathbb{N}. \quad (1.7) \]

In our setup, \( x_0 = x^{(1)} = (a_1, 0) \) is a fixed point of \( f \) and therefore the corresponding limit trajectory is constant, \( x_n \equiv x_0 \) for all \( n \geq 0 \). Consequently, the limit (1.7) fails to provide any information on the transition to a new coexistence equilibrium: if such a transition occurs, it becomes visible much later, at a time depending on \( K \).

A new type of limit theorems, capable of capturing this transition was recently discovered in \([3]\), \([2]\), \([6]\), \([1]\). They point at a time shift which grows logarithmically in \( K \). In \([3]\) this shift is random and the process \( X^K_n \) is approximated by the trajectory of the deterministic system (1.4). In \([2]\), \([6]\), \([1]\), the shift is deterministic, but \( X^K_n \) converges to a trajectory of (1.4), started from a random initial condition.

While the two approaches are related, they are not equivalent. The main building block in the random initial condition theory is a certain scaling limit of the deterministic flow, which does not appear in the random shift theory. Existence of this limit was so far established only in the one dimensional case. This work is the first such result in two
dimensions. Having established it, we can complement the “random shift” in [3] with the “random initial condition” for the Bare Bones model.

1.4. The main result. As mentioned, this builds upon two approximations:

- Keeping the stochasticity, but making a local approximation of the population dynamics, by assuming the surrounding environment constant, viz. \((a_1, 0)\), “linearisation”, and

- Disregarding stochasticity, but respecting the dependence structure (1.4) with the initial value \((a_1, 0)\) perturbed.

For the latter, note that the Jacobian matrix of \(f(\cdot)\) at the unstable fixed point \(x^{(1)} = (a_1, 0)\) is given by

\[
A := Df(x^{(1)}) = \left( \begin{array}{cc}
\frac{1}{2} & -\frac{\gamma}{2} \\
0 & \rho
\end{array} \right), \quad \text{where } \rho := \frac{2a_2}{a_2 + \gamma a_1},
\]  

(1.8)

Here \(\rho = m_2(x^{(1)})\) and \(1 < \rho < \frac{2}{1+\gamma}\) under the coexistence condition \((C)\).

The stochastic approximation of the mutant at the point \((a_1, 0)\) is given by the supercritical Galton-Watson binary splitting process \(Y = (Y_n, n \in \mathbb{Z}^+)\) with

\[
Y_{n+1} = \sum_{j=1}^{Y_n} \zeta_{nj}
\]

where \(\zeta_{nj} \in \{0, 2\}\) are i.i.d. random variables with \(P(\zeta_{nj} = 2) = \rho/2\). As is well known, the scaled process \(W_n := \rho^{-n}Y_n\) is a martingale with the a.s. and \(L_2\) limit

\[
W = \lim_{n \to \infty} \rho^{-n}Y_n.
\]

As opposed to this, the wild-type population with conditions frozen at \((a_1, 0)\) is critical. Below \(f^n(\cdot)\) denotes the \(n\)-fold iterate of \(f(\cdot)\), and thus mirrors the deterministic growth of our process.

**Theorem 1.1.** Under the basic assumptions stated, the limit

\[
H(x) := \lim_n f^n(x^{(1)} + x/\rho^n), \quad x \in \mathbb{R} \times \mathbb{R}_+
\]

exists and the convergence is uniform over compacts.

This auxiliary theorem leads on to the main result:

**Theorem 1.2.** Let \((X_n^K, n \in \mathbb{Z}_+)\) be generated by (1.5) (or equivalently by (1.1)) subject to initial condition \(X_0^K = (a_1, 1/K)\). Assume that the coexistence condition \((C)\) holds and let \(n_1(K) := \log_{\rho} K\). Then

\[
X_{n_1(K)+n}^{K} \xrightarrow{d} \frac{d}{K \to \infty} f^n \left( H\left( (0, W) \right) \right),
\]

(1.10)

along any sequence of integers \(n_1(K)\).
Remark 1.3.

(1) Numerical calculations indicate that $H(x)$ is constant with respect to perturbations $x_1$, the first entry of $x$, in the wild-type population density, Figure 1. This is consistent with the criticality of that population at the density $a_1$.

(2) Since $H((0,W)) = x^{(1)}$ on the event $\{W = 0\}$, the limit in (1.10) also equals $x^{(1)}$ on this event. This corresponds to early extinction of the mutant. In other words, with probability $\mathbb{P}(W = 0) = 2/\rho - 1$ the mutant fails to establish itself alongside the large original population. This interesting fact allows for approximate calculation of survival probability when the carrying capacities are large.

(3) Recently heuristics for similar random initial conditions for selective sweeps in large populations in one dimension were given in [12].

2. Proofs

The proof consists of two major parts, first proving that the limit $H(x)$ in Theorem 1.1 exists, then constructing the approximation, which implies the main Theorem 1.2.

2.1. The limit $H(x)$.

2.1.1. An auxiliary recursion in dimension one. Let us start with an auxiliary one dimensional quadratic recursion

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

subject to initial condition $x_{0,n} = x/\rho^n$ with $x > 0$, where $C \geq 0$ and $\rho > 1$ are constant coefficients. In what follows we will need the following estimate on its solution:

**Lemma 2.1.** There exists a finite function $\psi : \mathbb{R}_+ \mapsto \mathbb{R}_+$, such that

$$x_{m,n} \leq \psi(x)\rho^{n-m}, \quad m = 1, \ldots, n.$$  (2.2)
Proof. If we multiply both sides of (2.1) by $C$, a recursion for $\tilde{x}_{m,n} := Cx_{m,n}$ is obtained and hence $C = 1$ can and will be assumed without loss of generality:

$$x_{m,n} = p(x_{m-1,n}), \quad m = 1, \ldots, n$$

with $p(x) = \rho x(1 + x)$. Since $x_{m,n} \geq \rho x_{m-1,n}$, proving the desired bound is equivalent to showing that

$$\sup_n x_{n,n} < \infty, \quad \forall x \geq 0.$$  

To this end, consider the Schröder functional equation

$$\phi(f(x)) = s\phi(x), \quad x \in [0, \infty) \tag{2.4}$$

where $s = \frac{1}{\rho} \in (0, 1)$ and $f(x) = \sqrt{\frac{1}{4} + sx - \frac{1}{2}}$ is the inverse of the parabola $p(\cdot)$ on $\mathbb{R}_+$. Function $f(x)$ satisfies the following conditions

1. $f$ is continuous and strictly increasing on $[0, \infty)$
2. $f(0) = 0$ and $0 < f(x) < x$ for $0 < x < \infty$
3. $f(x)/x \to s$ as $x \to 0+$
4. $f(x)$ is concave (and therefore $f(x)/x$ is decreasing on $\mathbb{R}_+$)
5. for all $\delta > 0$

$$\int_0^\delta \frac{|f(x) - sx|}{x^2} dx < \infty.$$ 

Under these conditions [14] shows that the limit

$$\phi(x) := \lim_n \frac{f^n(x)}{s^n}, \quad x \in [0, \infty)$$

exists, solves (2.4) and satisfies the following properties

a) $0 < \phi(x) < \infty$ in $(0, \infty)$ (nontrivial limit)

b) $\phi(x)/x$ is monotone on $(0, \infty)$

c) $\phi'(0+) = 1$.

d) $\phi(x)$ is invertible

We can write (2.4) as

$$\phi(y) = s\phi(p(y)), \quad y \in \mathbb{R}_+$$

and, inverting, we obtain the conjugacy

$$p(y) = \phi^{-1}(\rho \phi(y)).$$

Hence

$$x_{n,n} = p^n(x/\rho^n) = \phi^{-1}(\rho^n\phi(x/\rho^n)) \xrightarrow{n \to \infty} \phi^{-1}(x\phi'(0+)) = \phi^{-1}(x).$$

In particular, (2.3) and, therefore also (2.2), hold.

\[ \square \]

\[ \text{1} \text{see the remark in the paragraph following (3.1) in [14]} \]
2.1.2. *Growth estimates*. Let us summarize some relevant properties of the function \( f \) around the fixed point \( x^{(1)} = (a_1, 0) \). Define \( g(x) := f(x^{(1)} + x) - x^{(1)} \). Then \( g(0) = 0 \) and

\[
 f^n(x^{(1)} + x) = g^n(x) + x^{(1)}. \tag{2.5}
\]

The configuration of fixed points (1.6) imply that the subset

\[
 E = \left\{ x \in \mathbb{R}^2 : x_1 \geq x^{(*)}_1 - x^{(1)}_1, 0 \leq x_2 \leq x^{(*)}_2 \right\}
\]

is forward invariant under \( g \). In what follows, \( \| \cdot \| \) stands for the \( \ell_\infty \) norm for vectors and the corresponding operator norm for matrices. In particular, the matrix \( A \) defined in (1.8) satisfies \( \| A \| = \rho > 1 \). The linear subspace \( E_0 = \{ x \in \mathbb{R}^2 : x_2 = 0 \} \) is invariant under \( A \) and

\[
 \sup_{x \in E_0} \| Ax \| = \frac{1}{2} \tag{2.6}
\]

Below \( C, C_1, \) etc. stand for constants, which depend only on \( a_1, a_2 \) and \( \gamma \) and whose values may change from line to line.

The first coordinate of \( g \) can be written as

\[
 g_1(x) = \frac{2a_1(a_1 + x_1)}{2a_1 + x_1 + \gamma x_2} - a_1 = \frac{a_1 x_1}{2a_1 + x_1 + \gamma x_2} - \frac{a_1 \gamma x_2}{2a_1 + x_1 + \gamma x_2} = \frac{1}{2} x_1 \left( 1 - \frac{x_1 + \gamma x_2}{2a_1 + x_1 + \gamma x_2} \right) - \frac{\gamma}{2} x_2 \left( 1 - \frac{x_1 + \gamma x_2}{2a_1 + x_1 + \gamma x_2} \right),
\]

and, similarly,

\[
 g_2(x) = \rho x_2 \left( 1 - \frac{x_2 + \gamma x_1}{a_2 + \gamma a_1 + x_2 + \gamma x_1} \right).
\]

Hence \( g(x) \) can be written as

\[
 g(x) = (I - B(x)) Ax \tag{2.7}
\]

where matrix \( B(x) \) satisfies the bound

\[
 \| B(x) \| \leq C \| x \|, \quad x \in E \tag{2.8}
\]

with a constant \( C \). Similar calculation also shows that for \( x, y \in E \)

\[
 g(x) - g(y) = (A + F(x, y))(x - y) \tag{2.9}
\]

where matrix \( F(x, y) \) satisfies

\[
 \| F(x, y) \| \leq C (\| x \| \vee \| y \|). \tag{2.10}
\]

Using these formulas and Lemma 2.1, the following growth estimate is obtained:

**Lemma 2.2.** For any \( x \in \mathbb{R} \times \mathbb{R}_+ \) and all \( n \) large enough,

\[
 \| g^n(x/\rho^n) \| \leq \psi(\| x \|) \rho^{m-n} \tag{2.11}
\]

with a finite function \( \psi(r), \ r \geq 0 \).
Proof. For any \( x \in \mathbb{R} \times \mathbb{R}_+ \) and all \( n \) large enough \( x/\rho^n \in E \) and, since \( E \) is invariant, \( g^m(x/\rho^n) \in E \) for all \( m \). Hence by (2.7) the sequence \( x_j := g^j(x) \) satisfies
\[
\|x_j\| = \|g(x_{j-1})\| = \|(I - B(x_{j-1}))Ax_{j-1}\| \leq \|I - B(x_{j-1})\|\|A\|\|x_{j-1}\| \leq \rho\|x_{j-1}\|(1 + C\|x_{j-1}\|).
\]
The claim now follows from Lemma 2.1. \( \square \)

2.1.3. Proof of Theorem 1.1.

Proof. We will argue that the increments of \( g^n(x/\rho^n) \) are absolutely summable, uniformly over compacts in \( \mathbb{R} \times \mathbb{R}_+ \). Let \( n \) be large enough so that \( x/\rho^n \in E \) and therefore, by invariance, \( g^m(x/\rho^n) \in E \) for all \( m \geq 1 \). Consider the array
\[
g^m(x/\rho^{n+1}) - g^{m-1}(x/\rho^n), \quad m = 1, \ldots, n.
\]
By (2.7) for \( m = 1 \)
\[
g(x/\rho^{n+1}) - x/\rho^n = Ax/\rho^{n+1} - x/\rho^n - B(x/\rho^{n+1})Ax/\rho^{n+1} = \\
\rho^{-n}(A/\rho - I)x + \rho^{-2n}v_n =: \rho^{-n}u + \rho^{-2n}v_n
\]
where \( u \in E_0 \) and, by (2.8), \( v_n \) is a sequence of vectors with norm, uniformly bounded in \( n \). Both \( u \) and \( v_n \) depend continuously on \( x \), which is omitted from the notations. For \( m \geq 1 \), (2.9) implies
\[
g^{m+1}(x/\rho^{n+1}) - g^m(x/\rho^n) = g \circ g^m(x/\rho^{n+1}) - g \circ g^{m-1}(x/\rho^n) = \\
(A + F(g^m(x/\rho^{n+1}), g^{m-1}(x/\rho^n)))(g^m(x/\rho^{n+1}) - g^{m-1}(x/\rho^n))
\]
and, letting \( F_{m,n} = F(g^m(x/\rho^{n+1}), g^{m-1}(x/\rho^n)) \) for brevity, we get
\[
g^{n+1}(x/\rho^{n+1}) - g^n(x/\rho^n) = \left\{ \prod_{m=1}^n (A + F_{m,n}) \right\} (\rho^{-n}u + \rho^{-2n}v_n), \quad (2.12)
\]
where the product is understood as \( \prod_{j=1}^n M_j := M_n \ldots M_1 \).

The second term in (2.12) satisfies the following bound
\[
\left\| \left\{ \prod_{m=1}^n (A + F_{m,n}) \right\} \rho^{-2n}v_n \right\| \leq \rho^{-2n}\|v_n\| \prod_{m=1}^n (\|A\| + \|F_{m,n}\|) \leq \\
\rho^{-2n}C_1 \prod_{m=1}^n (\rho + C(\|g^m(x/\rho^{n+1})\| + \|g^{m-1}(x/\rho^n)\|)) \leq \\
\rho^{-2n}C_1 \prod_{m=1}^n (\rho + C_2\rho^{m-n}) \leq \rho^{-n}C_1 \prod_{m=1}^n (1 + C_2\rho^{m-n}) \leq C_3\rho^{-n}, \quad (2.13)
\]
where we used (2.11), (2.10) and \( \|A\| = \rho \). Constant \( C_3 \) here depends continuously on \( \|x\| \).
Let us now bound the first term in (2.12). To this end, observe that

\[
\prod_{m=1}^{n} (A + F_{m,n}) = \prod_{m=2}^{n} (A + F_{m,n}) F_{1,n} + \\
\prod_{m=3}^{n} (A + F_{m,n}) F_{2,n} A + \\
\prod_{m=4}^{n} (A + F_{m,n}) F_{3,n} A^2 + \cdots + \\
(A + F_{n,n}) F_{n-1,n} A^{n-2} + F_{n,n} A^{n-1} + A^n
\]

Since \( u \in E_0 \) and \( E_0 \) is invariant under \( A \), we have \( \|A^k u\| \leq (1/2)^k \|u\| \) by (2.6). Therefore for all \( k = 0, \ldots, n - 2 \)

\[
\| \prod_{m=k+2}^{n} (A + F_{m,n}) F_{k+1,n} A^k u \| \leq (1/2)^k \|u\| \| \prod_{m=k+2}^{n} (A + F_{m,n}) F_{k+1,n} \| \leq C_1(1/2)^k \|F_{k+1,n}\| \prod_{m=k+2}^{n} (\|A\| + \|F_{m,n}\|) \leq C_1(1/2)^k \|x\| \rho^{k-n} \prod_{m=k+2}^{n} (\rho + C \|x\| \rho^{m-n}) \leq C_2(1/2)^k \prod_{m=k+2}^{n} (1 + C_3 \rho^{m-n}) \leq C_2(1/2)^k \exp \left( C_3 \sum_{m=k+2}^{n} \rho^{m-n} \right) \leq C_4(1/2)^k,
\]

where all constants \( C_j \) depend on \( x \). Consequently

\[
\| \prod_{m=1}^{n} (A + F_{m,n}) \rho^{-n} u \| \leq C_4 \rho^{-n}
\]

Plugging this and (2.13) into (2.12) yields

\[
\| g^{n+1}(x/\rho^{n+1}) - g^n(x/\rho^n) \| \leq C_5 \rho^{-n},
\]

and, in turn, the claimed uniform convergence of \( g^n(x/\rho^n) \). Using a telescoping sum and (2.5) existence of the limit \( H \) in (1.9) now follows and Theorem 1.1 is proved. \( \square \)

2.2. The main approximation. We construct the random variable \( W \) on the same probability space and show that convergence (1.10) holds in probability. To this end, let \( U_{nj} \) and \( V_{nj} \) be i.i.d. random variables distributed uniformly over the unit interval \([0, 1]\) and define \( \xi_{nj}^{(1)} \) and \( \xi_{nj}^{(2)} \) in (1.1) as follows:

\[
\xi_{nj}^{(1)} = 2 \cdot 1 \left\{ U_{nj} \leq \frac{a_1 K}{a_1 K + Z_n^{(1)} + \gamma Z_n^{(2)}} \right\} \\
\xi_{nj}^{(2)} = 2 \cdot 1 \left\{ V_{nj} \leq \frac{a_2 K}{a_2 K + \gamma Z_n^{(1)} + Z_n^{(2)}} \right\}
\]
Define Galton-Watson branching processes

\[ Y_{n+1}^{(1)} = \sum_{j=1}^{Y_{n}^{(1)}} 2 \cdot 1 \left\{ U_{nj} \leq \frac{1}{2} \right\} \quad \text{and} \quad Y_{n+1}^{(2)} = \sum_{j=1}^{Y_{n}^{(2)}} 2 \cdot 1 \left\{ V_{nj} \leq \frac{1}{2}\rho \right\}, \]

subject to \( Y_{0}^{(1)} = a_{1}K \) and \( Y_{0}^{(2)} = 1 \), and the corresponding densities \( \bar{Y}_{n}^{(1)} := K^{-1}Y_{n}^{(1)} \) and \( \bar{Y}_{n}^{(2)} := K^{-1}Y_{n}^{(2)} \). Finally let \( n_{c}(K) := \log_{\rho} K^{c} \) with a constant \( c \in (\frac{1}{2}, 1) \) and define the approximating process

\[ \bar{Z}_{n} = \begin{cases} \bar{Y}_{n} & n \leq n_{c} \\ f^{n-n_{c}}(\bar{Y}_{n_{c}}) & n > n_{c} \end{cases} \]

By continuity, the assertion of Theorem 1.2 holds for any \( n \geq 0 \), if it holds for \( n = 0 \), which, in turn, holds if

\[ \bar{Z}_{n_{1}} = f^{n_{1}-n_{c}}(\bar{Y}_{n_{c}}) \xrightarrow{p} H((0, W)) \quad (2.15) \]

and

\[ \bar{Z}_{n_{1}} - \bar{Z}_{n_{1}} \xrightarrow{p} 0, \quad (2.16) \]

where \( \bar{Z}_{n} := K^{-1}Z_{n} = X_{n} \).

2.2.1. Proof of (2.15). Observe that

\[
\begin{align*}
 f^{n_{1}-n_{c}}(\bar{Y}_{n_{c}}) &= f^{n_{1}-n_{c}}\left( (a_{1} + (\bar{Y}_{n_{c}}^{(1)} - a_{1}), \bar{Y}_{n_{c}}^{(2)}) \right) \\
 &= f^{n_{1}-n_{c}}\left( (a_{1} + \rho^{-(n_{1}-n_{c})}K^{1-c}(\bar{Y}_{n_{c}}^{(1)} - a_{1}), \rho^{-(n_{1}-n_{c})}\rho^{n_{c}}Y_{n_{c}}^{(2)}) \right).
\end{align*}
\]

Since \( EY_{n}^{(1)} = a_{1}K \) and \( \text{Var}(Y_{n}^{(1)}) \leq na_{1}K \) we have

\[ E(\bar{Y}_{n_{c}}^{(1)} - a_{1})^{2} \leq a_{1}K^{-1} \log_{\rho} K^{c} \quad (2.17) \]

and hence for \( c \in (\frac{1}{2}, 1) \),

\[ K^{1-c}(\bar{Y}_{n_{c}}^{(1)} - a_{1}) \xrightarrow{p} 0. \]

Convergence (2.15) now follows by Theorem 1.1.

2.2.2. Proof of (2.16). Since

\[ \|\bar{Z}_{n_{1}} - \bar{Z}_{n_{1}}\| \leq \|\bar{Z}_{n_{1}} - f^{n_{1}-n_{c}}(\bar{Z}_{n_{c}})\| + \|f^{n_{1}-n_{c}}(\bar{Z}_{n_{c}}) - f^{n_{1}-n_{c}}(\bar{Y}_{n_{c}})\| \]

it suffices to prove that

\[ \|\bar{Z}_{n_{1}} - f^{n_{1}-n_{c}}(\bar{Z}_{n_{c}})\| \xrightarrow{p} 0 \quad (2.18) \]

and

\[ \|f^{n_{1}-n_{c}}(\bar{Z}_{n_{c}}) - f^{n_{1}-n_{c}}(\bar{Y}_{n_{c}})\| \xrightarrow{p} 0. \quad (2.19) \]

Let us first prove (2.18). Recall that the density process \( X_{n} = Z_{n} \) solves (1.5):

\[ \bar{Z}_{n} = f(\bar{Z}_{n-1}) + \frac{1}{\sqrt{K}}\eta_{n} \]
and hence the difference \( \delta_n := Z_n - f^{n-n_c}(Z_{n_c}) \) satisfies
\[
\delta_n = f(Z_{n-1}) - f(Z_{n-1} - \delta_{n-1}) + \frac{1}{\sqrt{K}}\eta_n, \quad n > n_c
\]
subject to \( \delta_{n_c} = 0 \). The Jacobian of \( f(x) \) is bounded, \( \tilde{\rho} := \sup_{x \in \mathbb{R}^2} \|Df(x)\| \leq 2 \) and hence is \( \tilde{\rho} \)-Lipschitz on \( \mathbb{R}^2_+ \) with respect to \( \ell_\infty \) norm. Hence
\[
\|\delta_n\| \leq \tilde{\rho}\|\delta_{n-1}\| + \frac{1}{\sqrt{K}}\|\eta_n\|
\]
and consequently
\[
E\|\delta_{n_1}\| \leq \frac{1}{\sqrt{K}} \sum_{j=n_c}^{n_1} \tilde{\rho}^{n_1-j}E\|\eta_j\| \leq \frac{1}{\sqrt{K}}(n_1 - n_c)\tilde{\rho}^{n_1-n_c} \sup_{j \leq n_1} E\|\eta_j\| \leq CK^{(1-c)\beta - \frac{1}{2}} \log \tilde{\rho} K^{1-c} \to 0,
\]
where \( \beta := \log_{\tilde{\rho}} \rho \) and convergence (2.18) holds if \( c \) is chosen close enough to 1.

To check (2.19), write
\[
\|f^{n_1-n_c}(Z_{n_c}) - f^{n_1-n_c}(Y_{n_c})\| = \left\|f^{n_1-n_c}(x^{(1)} + \rho^{-(n_1-n_c)}(Z_{n_c} - Kx^{(1)})) - f^{n_1-n_c}(x^{(1)} + \rho^{-(n_1-n_c)}(Y_{n_c} - Kx^{(1)}))\right\|
\]

Since, by (2.17), the vector \( \rho^{-n_c}(Y_{n_c} - Kx^{(1)}) \) converges to \( (0, W) \) in probability and by Theorem 1.1 functions \( f^n(x^{(1)} + x/\rho^n) \) converge uniformly on compacts to \( H(x) \), it suffices to show that \( \rho^{-n_c}\|Z_{n_c} - Y_{n_c}\| \xrightarrow{P} 0 \), that is
\[
K^{-c}|Z_{n_c}^{(j)} - Y_{n_c}^{(j)}| \to 0, \quad j = 1, 2,
\]
where \( c < 1 \) has been already fixed in the previous calculations.

Let us prove (2.20) for \( j = 2 \), omitting the similar proof for the case \( j = 1 \). To this end, choose constants \( \alpha_{1\ell}, \alpha_{1u} \) and \( \alpha_2 \) so that
\[
c < \alpha_2 < \alpha_{1u} < \alpha_{1\ell} < 1
\]
and define two auxiliary Galton-Watson branching processes \( L_n \) and \( U_n \) as follows:
\[
L_n^{(1)} = \sum_{j=1}^{L_n^{(1)}-1} 2 \cdot 1 \left\{ U_{nj} \leq r_K^- \right\}, \quad L_0^{(1)} = a_1 K
\]
\[
U_n^{(1)} = \sum_{j=1}^{U_n^{(1)}-1} 2 \cdot 1 \left\{ U_{nj} \leq r_K^+ \right\}, \quad U_0^{(1)} = a_1 K
\]
with
\[
r_K^- := \frac{a_1}{a_1 + a_1(1 + K^{\alpha_{1u}^{-1}}) + \gamma K^{\alpha_2-1}} \quad \text{and} \quad r_K^+ := \frac{a_1}{a_1 + a_1(1 - K^{\alpha_{1\ell}^{-1}})}
\]
and

\[ L_n^{(2)} = \sum_{j=1}^{I_n^{(2)}} 2 \cdot 1 \left\{ V_{nj} \leq \frac{1}{2} \rho_K^- \right\}, \quad L_0^{(2)} = 1 \]

\[ U_n^{(2)} = \sum_{j=1}^{I_n^{(2)}} 2 \cdot 1 \left\{ V_{nj} \leq \frac{1}{2} \rho_K^+ \right\}, \quad U_0^{(2)} = 1 \]

where

\[ \rho_K^+ = \frac{2a_2}{a_2 + \gamma a_1 (1 + K^{\alpha_1 u - 1}) + K^{\alpha_2 - 1}} \quad \text{and} \quad \rho_K^- = \frac{2a_2}{a_2 + \gamma a_1 (1 - K^{\alpha_1 t - 1})}, \]

and \( U_{nj} \) and \( V_{nj} \) are the random variables in (2.14). Define the stopping times

\[ \tau^{1t} = \min \{ n : Z_n^{(1)} \leq a_1 (K - K^{\alpha_1 t}) \} \]
\[ \tau^{1u} = \min \{ n : Z_n^{(1)} \geq a_1 (K + K^{\alpha_1 u}) \} \]
\[ \tau^2 = \min \{ n : Z_n^{(2)} \geq K^{\alpha_2} \} \]

and set \( \tau = \tau^{1t} \land \tau^{1u} \land \tau^2 \). Then by construction, \( \mathbb{P}(L_n^{(2)} \leq Y_n^{(2)} \leq U_n^{(2)}) = 1 \) and

\[ \{ \tau \geq n \} \subseteq \{ L_n^{(2)} \leq Z_n^{(2)} \leq U_n^{(2)} \}. \]

Hence

\[ |Z_n^{(2)} - Y_n^{(2)}| \leq (U_n^{(2)} - L_n^{(2)}) 1_{\{\tau \geq n\}} + |Z_n^{(2)} - Y_n^{(2)}| 1_{\{\tau < n\}} \leq (U_n^{(2)} - L_n^{(2)}) + Z_n^{(2)} 1_{\{\tau < n\}} + Y_n^{(2)} 1_{\{\tau < n\}}. \]

The first term is positive and

\[ K^{-c} \mathbb{E}(U_n^{(2)} - L_n^{(2)}) = \left( (\rho_K^+ / \rho)^{\log_{\rho} K^c} - (\rho_K^- / \rho)^{\log_{\rho} K^c} \right) \leq C \log_{\rho} K^c \rho_K^+ - \rho_K^- \left( \rho_K^+ / \rho \right)^{\log_{\rho} K^c} \leq C \log_{\rho} K^c \left| K^{\alpha_1 u - 1} + K^{\alpha_1 t - 1} \right| (1 + O(K^{\alpha_1 t - 1})) \log_{\rho} K^c \xrightarrow{K \to \infty} 0. \]

Since \( \mathbb{E} Z_n^{(2)} \leq C \rho^n \) and \( \rho^{-n} Y_n^{(2)} \to W \), to bound the other two terms it suffices to check

\[ \mathbb{P}(\tau < n_c) \xrightarrow{K \to \infty} 0. \quad (2.21) \]

To this end, define

\[ \sigma^{1t} = \min \{ n : L_n^{(1)} \leq a_1 (K - K^{\alpha_1 t}) \} \]
\[ \sigma^{1u} = \min \{ n : U_n^{(1)} \geq a_1 (K + K^{\alpha_1 u}) \} \]
\[ \sigma^2 = \min \{ n : U_n^{(2)} \geq K^{\alpha_2} \} \]

and set \( \sigma = \sigma^{1u} \land \sigma^{1t} \land \sigma^2 \). Then

\[ \mathbb{P}(\tau < n_c) \leq \mathbb{P}(\sigma < n_c) \leq \mathbb{P}(\sigma^{1t} < n_c) + \mathbb{P}(\sigma^{1u} < n_c) + \mathbb{P}(\sigma^2 < n_c). \]
The last term satisfies the bound
\[ P(\sigma_2 < n_c) = P \left( \max_{n \leq n_c} U_n^{(2)} \geq K^{\alpha_2} \right) = P \left( (\rho_K^+) - n_c \max_{n \leq n_c} U_n^{(2)} \geq (\rho_K^+) - n_c K_2^\alpha \right) \leq \]
\[ \mathbb{P} \left( \max_{n \leq n_c} (\rho_K^+) - n U_n^{(2)} \geq (\rho_K^+) - n_c K^{\alpha_2} \right) \leq (\rho_K^+) n_c K^{\alpha_2} \leq C K^{\gamma - \alpha_2} \xrightarrow{K \to \infty} 0. \]

The process \( a_1 K - L_n^{(1)} \) is a sub-martingale and hence
\[ P(\sigma_1 < n_c) = P \left( \min_{n \leq n_c} L_n^{(1)} \leq a_1 (K - K^{\alpha_1}) \right) = \]
\[ P \left( \max_{n \leq n_c} (a_1 K - L_n^{(1)}) \geq a_1 K^{\alpha_1} \right) \leq CK^{-\alpha_1 \epsilon} \mathbb{E} \left| a_1 K - L_n^{(1)} \right| \leq \]
\[ CK^{-\alpha_1 \epsilon} \left| a_1 K - \mathbb{E} L_n^{(1)} \right| + CK^{-\alpha_1 \epsilon} \sqrt{\text{Var}(L_n^{(1)})}. \]

The first term here satisfies the bound:
\[ K^{-\alpha_1 \epsilon} \left| a_1 K - \mathbb{E} L_n^{(1)} \right| \leq K^{-\alpha_1 \epsilon} \left| a_1 K - a_1 K (r_K^-)^{n_c} \right| \leq \]
\[ a_1 K^{1 - \alpha_1 \epsilon} \left| 1 - (1 - K^{\alpha_1 u - 1} - K^{\alpha_2 - 1})^{n_c} \right| \leq \]
\[ a_1 K^{1 - \alpha_1 \epsilon} (K^{\alpha_1 u - 1} + K^{\alpha_2 - 1})^{n_c} = a_1 (K^{\alpha_2 - \alpha_1 \epsilon} + K^{\alpha_1 u - \alpha_1 \epsilon}) \log K \xrightarrow{K \to \infty} 0, \]

and the second term satisfies
\[ K^{-\alpha_1 \epsilon} \sqrt{\text{Var}(L_n^{(1)})} \leq K^{-\alpha_1 \epsilon} \sqrt{a_1 K (r_K^-)^{2n_c}} \leq CK^{1 - \alpha_1 \epsilon} \xrightarrow{K \to \infty} 0. \]

Finally
\[ P(\sigma_1 < n_c) = P \left( \max_{n \leq n_c} U_n^{(1)} \geq a_1 (K + K^{\alpha_1 u}) \right) \leq \]
\[ P \left( \max_{n \leq n_c} (r_K^+) - n U_n^{(1)} - a_1 K \right) \geq a_1 K (r_K^+)^{1 - n_c - 1} + (r_K^+) - n_c K^{\alpha_1 u} \right) \leq \]
\[ CK^{-\alpha_1 u} \mathbb{E} \left( (r_K^+)^{1 - n_c - 1} U_n^{(1)} - a_1 K \right) \leq CK^{-\alpha_1 u} \sqrt{\text{Var}(U_n^{(1)})} \leq \]
\[ CK^{-\alpha_1 u} \sqrt{a_1 K (r_K^+)^{2n_c}} \leq CK^{2 - \alpha_1 u} \xrightarrow{K \to \infty} 0. \]

This verifies (2.21) and, in turn, of (2.19), completing the proof of the main Theorem 1.2.

REFERENCES
[1] J. Baker, P. Chigansky, K. Hamza, and F. C. Klebaner. Persistence of small noise and random initial conditions. \textit{Adv. in Appl. Probab.}, 50(suppl. A):67–81, 2018.
[2] A. D. Barbour, P. Chigansky, and F. C. Klebaner. On the emergence of random initial conditions in fluid limits. \textit{J. Appl. Probab.}, 53(4):1193–1205, 2016.
[3] A. D. Barbour, K. Hamza, Haya Kaspi, and F. C. Klebaner. Escape from the boundary in Markov population processes. \textit{Adv. in Appl. Probab.}, 47(4):1190–1211, 2015.
[4] N. Champagnat, R. Ferriare, and G. Ben Arous. The canonical equation of adaptive dynamics: A mathematical view. \textit{Selection}, 2(1-2):73–83, 2002.
[5] Nicolas Champagnat and Sylvie Mélédard. Polymorphic evolution sequence and evolutionary branching. \textit{Probab. Theory Related Fields}, 151(1-2):45–94, 2011.
What can be observed in real time PCR and when does it show? J. Math. Biol., 76(3):679–695, 2018.

M. I. Freidlin and A. D. Wentzell. Random perturbations of dynamical systems, volume 260 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, New York, 1984. Translated from the Russian by Joseph Szücs.

F. C. Klebaner. On population-size-dependent branching processes. Adv. in Appl. Probab., 16(1):30–55, 1984.

Fima C. Klebaner. Population-dependent branching processes with a threshold. Stochastic Process. Appl., 46(1):115–127, 1993.

Fima C. Klebaner, Serik Sagitov, Vladimir A. Vatutin, Patsy Haccou, and Peter Jagers. Stochasticity in the adaptive dynamics of evolution: the bare bones. J. Biol. Dyn., 5(2):147–162, 2011.

Thomas G. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. J. Appl. Probability, 7:49–58, 1970.

Guillaume Martin and Amaury Lambert. A simple, semi-deterministic approximation to the distribution of selective sweeps in large populations. Theoretical Population Biology, 101:40 – 46, 2015.

S. Sagitov, B. Mehlig, P. Jagers, and V. Vatutin. Evolutionary branching in a stochastic population model with discrete mutational steps. Theoretical Population Biology, 83:145 – 154, 2013.

E. Seneta. On Koenigs’ ratios for iterates of real functions. J. Austral. Math. soc., 10:207–213, 1969.

School of Mathematical Sciences, Monash University, Monash, VIC 3800, Australia
E-mail address: jeremy.baker@monash.edu

Department of Statistics, The Hebrew University, Mount Scopus, Jerusalem 91905, Israel
E-mail address: pchiga@mscc.huji.ac.il

Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE-412 96 Gothenburg, Sweden
E-mail address: jagers@chalmers.se

School of Mathematical Sciences, Monash University, Monash, VIC 3800, Australia
E-mail address: fima.klebaner@monash.edu