



Supporting Information

for *Adv. Sci.*, DOI 10.1002/adv.202414297

A Model of Epigenetic Inheritance Accounts for Unexpected Adaptation to Unforeseen Challenges

Dino Osmanović, Yitzhak Rabin and Yoav Soen*

Supplementary Information to a model of epigenetic inheritance accounts for unexpectedly rapid adaptations of real populations

Dino Osmanović*

*Department of Mechanical and Aerospace Engineering,
University of California, Los Angeles, Los Angeles 90095, CA, USA*

Yitzhak Rabin

Department of Physics, Bar-Ilan University, Ramat Gan 5290002, Israel

Yoav Soen

*Department of Bio-molecular Sciences,
Weizmann Institute of Science, Rehovot, Israel*

(Dated: October 29, 2024)

CONTENTS

I. Derivation of the extended McKendrick–von Foerster equation	2
II. Symbiosis leading to effective drift	4
III. Solution for $R = 0, I = \delta(\chi - \chi')$ model	6
IV. Stationary Distribution	9
V. Adaptation to a Sudden change in the environment	10
VI. Continual variation of the environment	13
VII. Conditions on survival	14
VIII. Fertility, Inheritance and their effect on adaptation to environment	16
IX. χ -dependent replication	21

* osmanovic.dino@gmail.com

X. More realistic birth rates	23
References	24

I. DERIVATION OF THE EXTENDED MCKENDRICK–VON FOERSTER EQUATION

We shall extend the McKendrick–von Foerster equation to include the effect of changes, both pre-programmed and random, that take place during the lifetime of individuals and affect their survival probabilities (epigenetics, microbiome, etc.)

The equations that govern the time (t) evolution of the distribution of ages T in a population $n(T, t)$ are given by:

$$\frac{\partial n(T, t)}{\partial t} = -\frac{\partial n(T, t)}{\partial T} - P_D(T, t) n(T, t) \quad (1)$$

$$n(0, t) = \int P_R(T', t) n(T', t) dT' \quad (2)$$

$$n(\infty, t) = 0, \quad (3)$$

which can be derived by considering the fundamental processes involved. Here $P_D(T, t)$ and $P_R(T', t)$ are the death and replication rates, respectively. Note that death is incorporated as a sink term in Eq. 1 and replication enters as an initial condition (at age $T = 0$), in Eq. 2.

We shall now derive the extended equation by including an additional degree of freedom in this equation. To begin, we consider an extended distribution, which includes both age T and some other parameter χ .

$$n(T, \chi, t) = \sum_{i=1}^N \delta(T - T_i(t)) \delta(\chi - \chi_i(t)) \quad (4)$$

where after a period of time Δt , we have:

$$n(T, \chi, t + \Delta t) = \sum_{i=1}^N \delta(T - T_i(t + \Delta t)) \delta(\chi - \chi_i(t + \Delta t)) \quad (5)$$

In order to convert this into a differential equation, we need to specify update rules for the parameters T and χ . For each individual i , the update rule for age in time is simply:

$$T_i(t + \Delta t) = T_i(t) + \Delta t \quad (6)$$

For the update of the additional parameter x_i we choose a drift-diffusion scheme:

$$\chi_i(t + \Delta t) = \chi_i(t) + f(\chi_i, t)\Delta t + \xi_i(t) \quad (7)$$

where ξ_i is a random number of mean 0 with correlation function $\langle \xi_i(t)\xi_j(t') \rangle = 2D\delta_{i,j}\delta(t-t')$, and $f(\chi_i, t)$ is a deterministic change in χ_i per unit time (here and in the following we assume that f does not depend on the age of the individual, T_i). Note that the parameter χ_i changes *during the lifetime* of the individual. In the following we assume that both the form of the deterministic function f and the statistical properties (e.g., variance) of the random variable ξ , are the same for all individuals in the population.

When combined, these update rules correspond to the following Fokker-Planck like equation that we will refer to as the extended McKendrick–von Foerster equation:

$$\frac{\partial n(T, \chi, t)}{\partial t} = -\frac{\partial n(T, \chi, t)}{\partial T} - \frac{\partial}{\partial \chi} (f(\chi)n(T, \chi, t)) + D\frac{\partial^2 n(T, \chi, t)}{\partial \chi^2} - P_D(T, \chi, t)n(T, \chi, t) \quad (8)$$

This equation corresponds to individuals aging in time, with some other parameter χ of the individuals evolving according to the drift-diffusion equation, with a drift term $f(\chi, t)$ that depends itself on the parameter χ . This state parameter could correspond to any facet of the individual. Note that the death rate $P_D(T, \chi, t)$ is, in general, a function of both the age and state parameters T and χ , respectively, and of time t .

The replication rule is more complicated. It is given by:

$$n(0, \chi, t) = \int P_R(T', \chi, t)I(\chi - \chi')n(T', \chi', t)d\chi'dT' \quad (9)$$

This equation is similar to Eq. 1, except now we have an additional term accounting for the inheritance of the state parameter χ , given by the function $I(\chi - \chi')$. For example, if the inheritance function is given by $I(\chi - \chi') = \delta(\chi - \chi')$ this corresponds to exact inheritance of the parameter χ , which changes during the parent's lifetime. The possibility of imperfect inheritance can be modelled by a Gaussian inheritance function.

We therefore arrive at the modified McKendrick–von Foerster equation (with its boundary

conditions at $T = 0$ and $T = \infty$):

$$\frac{\partial n(T, \chi, t)}{\partial t} = -\frac{\partial n(T, \chi, t)}{\partial T} - \frac{\partial}{\partial \chi} (f(\chi)n(T, \chi, t)) + D \frac{\partial^2 n(T, \chi, t)}{\partial \chi^2} - P_D(T, \chi)n(T, \chi, t) \quad (10)$$

$$n(0, \chi, t) = \int P_R(T', \chi', t) I(\chi - \chi') n(T', \chi', t) d\chi' dT' \quad (11)$$

$$n(\infty, \chi, t) = 0 \quad (12)$$

II. SYMBIOSIS LEADING TO EFFECTIVE DRIFT

We shall in the following motivate particular choices of the drift function $f(\chi)$ appearing in the main text by considering a system of two populations co-evolving under the following assumptions:

- One population breeds much faster than the other. We call the fast breeders “bacteria” and the slow breeders “hosts”
- A population of the bacteria is associated with an individual host. In a population of hosts, we have a population of populations of bacteria, each one corresponding to a particular host, i.e. a microbiome.
- The probability of survival of each host depends, in some sense, on the properties of its bacterial population. For example, if we take some set of genomic properties of the fast population, the survival (or death) probability of the slow breeder will depend on these properties.

One particularly attractive example would correspond to a system where the death probability of the host depends on the mean $\langle \chi \rangle$ of some set of bacterial properties. We can use equation 10 with $f(\chi) = 0$ to construct a Darwinian model of the bacteria under some environment (defined through some parameters that characterize the dependence of the death probability P_D on χ).

If the mean of the bacterial population $\langle \chi \rangle$ is away from the minimum of the death rate $P_D(\chi) \sim \chi^2$, the mean will evolve in time until it reaches the minimum. When the dynamics of the mean are observed in this simple scenario, they are seen to be exponential (see fig 1B). The importance of this observation is that if we look at the process on the timescale

of the host, it would appear as if it is responding to its environment by changing its value of $\langle\chi\rangle$ by introducing a drift function of the following form:

$$f(\chi) = v \frac{\partial P_D(\chi)}{\partial \chi} \quad (13)$$

In other words, an organism with parameter $\langle\chi\rangle$ responds to environment, by moving away from regions of large death rate for negative v and by moving towards these regions for positive v . We stress that this arises only by considering coupled co-evolution over two divergent time scales. When $v < 0$, the evolutionary objectives of both host and bacteria are aligned and the bacterial adaptation proceeds in a direction which lowers the death rate of the host. However, the opposite can be the case where $v > 0$. In this scenario, the direction of bacterial adaptation leads to greater death rates of the host, as would be the case for pathogenic bacteria.

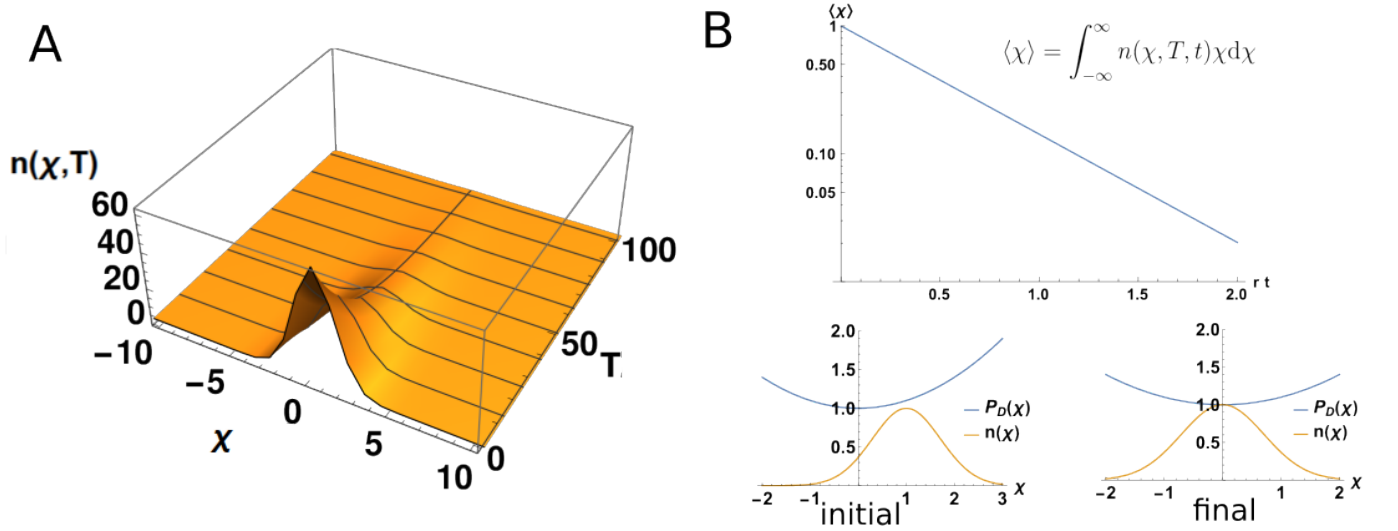


FIG. S1. Fundamental elements of the mathematical model. We construct expressions for how the distribution $n(\chi, T, t)$ changes in time. In panel A) we show an example distribution for an environment in which the death rate scales as $\sim \chi^2$, with $D = 1, v = 0, r = 2, a = 1, b = 1$. This leads to a steady state population distribution that is exponentially declining with age T and is Gaussian in the state parameter χ . B) The change in time of the mean of an initially misaligned initial population, subject to only diffusive dynamics. The mean relaxes exponentially, which can be interpreted as drift, on a longer timescale. ($a=1, b=1, D=1, v=0, r=1$)

III. SOLUTION FOR $R = 0, I = \delta(\chi - \chi')$ MODEL

We will in the following focus on an analytical example; this requires introducing Fourier transforms in χ and Laplace transforms in T . We will use the form of the death rate that we have introduced in the main text:

$$P_D(\chi) = a \int d\chi' dT' n(T', \chi', t) + b\chi^2 \quad (14)$$

This form of death rate has a term proportional to the total amount of individuals (global feedback due to population) and a term that depends quadratically on χ and ensures that the death rate has a minimum at $\chi = 0$ (“harmonic landscape”). Here and in the following we assume that the death rate does not depend on the age parameter and is a function of the state parameter χ only. Notice that while this simplifying assumption appears to be unphysical, even in the case of a constant death rate the population decreases exponentially with age. Therefore, the expected exponential increase of the death rate with T for ages exceeding some value T^* would act as a cutoff of the population distribution at T^* and will have a negligible effect on our results.

The additional assumption that was mentioned in the main text is that of perfect replication, $I(\chi - \chi') = \delta(\chi - \chi')$. This corresponds to offspring having exactly the same value of χ as their parents. The replication rate is given by:

$$P_R(T) = r \exp(-RT) \quad (15)$$

where r is the maximal replication rate and $1/R$ defines a time scale on which the ability to replicate decreases with age. Here and in the following we assume that the replication rate does not depend on the state parameter χ and is time-independent (i.e., does not depend on the changing environment).

As our starting point, we apply a Laplace transform in age T , and a Fourier transform in variable χ to the population distribution $n(T, \chi, t)$:

$$\hat{n}(S, k, t) = \int_{-\infty}^{\infty} d\chi \exp(-ik\chi) \int_0^{\infty} dT \exp(-ST) n(T, \chi, t) \quad (16)$$

We note here that the total number of individuals is given by $\hat{n}(0, 0, t)$. Applying the transforms to Eq. 10 yields:

$$\begin{aligned} \frac{\partial \hat{n}(S, k, t)}{\partial t} = & -S\hat{n}(S, k, t) + r\hat{n}(R, k, t) + 2vbk \frac{\partial \hat{n}(S, k, t)}{\partial k} - Dk^2 \hat{n}(S, k, t) \\ & - a\hat{n}(S, k, t)\hat{n}(0, 0, t) + b \frac{\partial^2 \hat{n}(S, k, t)}{\partial k^2} \end{aligned} \quad (17)$$

For now, we ignore the fact that the total population $\hat{n}(0, 0, t)$ is a functional of $\hat{n}(S, k, t)$, and treat it as some given function of time (we shall later demand that the solution is self-consistent in order to reintroduce this feedback). Furthermore, we ignore the age dependence of the replication time of the individual, by setting $R = 0$. In this case the entire behavior of the system is driven by the $S = 0$ mode in Laplace space, that obeys the following equation:

$$\begin{aligned} \frac{\partial \hat{n}(0, k, t)}{\partial t} = & r\hat{n}(0, k, t) + 2vbk \frac{\partial \hat{n}(0, k, t)}{\partial k} - Dk^2 \hat{n}(0, k, t) \\ & - a\hat{n}(0, k, t)N(t) + b \frac{\partial^2 \hat{n}(0, k, t)}{\partial k^2} \end{aligned} \quad (18)$$

Let us separate $\hat{n}(0, k, t)$ as $K(k)L(t)$, this yields (after dividing through by \hat{n}):

$$\frac{1}{L(t)} \frac{\partial L(t)}{\partial t} + aN(t) - r = +2vbk \frac{1}{K(k)} \frac{\partial K(k)}{\partial k} - Dk^2 + b \frac{1}{K(k)} \frac{\partial^2 K(k)}{\partial k^2} \quad (19)$$

Having successfully separated the variables here, we can solve each side being equal to some constant $-\mu$. This yields:

$$L(t) = A \exp \left(\int_0^t dt' (r - aN(t') - \mu) \right) \quad (20)$$

$$K(k) = B e^{-\frac{1}{2}k^2 \sqrt{\frac{bv^2 + \mathcal{D}}{b}} - \frac{k^2 v}{2}} H_{\frac{-vb - \sqrt{\frac{bv^2 + \mathcal{D}}{b}} b + \mu}{2b \sqrt{\frac{bv^2 + \mathcal{D}}{b}}}} \left(k \sqrt[4]{\frac{bv^2 + \mathcal{D}}{b}} \right) \quad (21)$$

where $H_m(x)$ is a Hermite polynomial of rank m and A and B are some constants. As these polynomials form of complete set, we set the degree of the polynomial equal to an integer m and solve for μ , giving $\mu = b \left((2m+1) \sqrt{\frac{\mathcal{D}}{b} + v^2} + v \right)$, and we can write down the complete solution:

$$\hat{n}(0, k, t) = \sum_m c_m e^{\left(\int_0^t dt' (r - aN(t') - b((2m+1)\sqrt{\frac{\mathcal{D}}{b} + v^2} + v)) \right)} e^{-\frac{1}{2}k^2 (\sqrt{\frac{\mathcal{D}}{b} + v^2} + v)} H_m \left(k \sqrt[4]{v^2 + \frac{\mathcal{D}}{b}} \right) \quad (22)$$

We see a common factor appearing, $\sqrt{\frac{\mathcal{D}}{b} + v^2}$, which we define as

$$\lambda^2 = \sqrt{\frac{\mathcal{D}}{b} + v^2} \quad (23)$$

for brevity.

Using the orthogonality of the Hermite polynomials allows us to write down the constants c_m :

$$c_m = \frac{\lambda}{\sqrt{\pi} 2^m m!} \int dk \hat{n}(0, k, 0) e^{\frac{1}{2}vk^2 - \frac{1}{2}\lambda^2 k^2} H_m(k\lambda) \quad (24)$$

Under the scenario where the initial population is a delta function (for example), $\hat{n}(0, k, 0) = 1$ and the integral can be performed, obtaining (for the even values):

$$c_{2m} = \frac{N_0 \lambda}{\sqrt{\pi} 2^{2m} 2m!} (-1)^m 2^{2m + \frac{1}{2}} \Gamma\left(m + \frac{1}{2}\right) \sqrt{\frac{1}{\lambda^2 - v}} \left(\frac{v + \lambda^2}{v - \lambda^2}\right)^m \quad (25)$$

where N_0 is the total number of individuals initially.

Some simplifications can be made to our solution:

$$\hat{n}(0, k, t) = e^{rt - bvt} e^{\left(\int_0^t dt' (-aN(t'))\right)} \sum_m c_{2m} e^{-b((4m+1)\lambda^2)t} e^{-\frac{1}{2}k^2(\lambda^2 + v)} H_{2m}(k\lambda) \quad (26)$$

The total number of individuals is given by this equation with $k = 0$. The Hermite polynomial of rank $2m$ at $k = 0$ is equal to $\frac{(-1)^m (2m)!}{m!}$. We can substitute our expression for the value of c_{2m} into the expression as well:

$$\hat{n}(0, 0, t) = e^{rt - bvt} e^{\left(\int_0^t dt' (-aN(t'))\right)} \sum_m c_{2m} e^{-b((4m+1)\lambda^2)t} \frac{(-1)^m (2m)!}{m!} \quad (27)$$

Fortunately, the sum has an analytical solution, yielding:

$$\hat{n}(0, 0, t) = e^{rt - bvt} e^{\left(\int_0^t dt' (-aN(t'))\right)} \frac{\sqrt{2}\lambda N_0 \sqrt{\frac{1}{\lambda^2 - v}} e^{-b\lambda^2 t}}{\sqrt{\frac{(\lambda^2 + v)e^{-4b\lambda^2 t}}{\lambda^2 - v}} + 1} \quad (28)$$

From here we can close the loop by recognizing that $n(0, 0, t)$ is equal to $N(t)$ (as we mentioned at the start); additionally we introduce the function:

$$B(t) = \left(e^{rt - bvt} \frac{\sqrt{2}\lambda N_0 \sqrt{\frac{1}{\lambda^2 - v}} e^{-b\lambda^2 t}}{\sqrt{\frac{(\lambda^2 + v)e^{-4b\lambda^2 t}}{\lambda^2 - v}} + 1} \right) \quad (29)$$

By taking logarithms of eq. 28 and then taking the derivative, we can derive a differential equation for $N(t)$:

$$\frac{dN(t)}{dt} = \frac{B'(t)}{B(t)} N(t) - aN(t)^2 \quad (30)$$

In words, the entire process described above enables us to write down a modified logistic growth equation with a time-dependent birth rate $d \ln B/dt$. It should be noted that this function $B(t)$ will be dependent on the initial conditions specified in the process, leading to non-universal forms. Once $N(t)$ is solved, we can return to the solution in order to derive how the distribution itself (in Fourier space) is changing in time.

Therefore, the generic time evolution of a system with the harmonic death rate is given by:

$$N(t) = \frac{N_0 \exp \left(\int_0^t G(t') dt' \right)}{N_0 \int_0^t a \exp \left(\int_0^{t'} G(t'') dt'' \right) dt' + 1} \quad (31)$$

where in this example:

$$G(t) = \frac{B'(t)}{B(t)} = r + \frac{b(v^2 - \lambda^4) \left(e^{4b\lambda^2 t} - 1 \right)}{(\lambda^2 - v) e^{4b\lambda^2 t} + \lambda^2 + v} \quad (32)$$

This is the landscape-dependent growth rate. In the limit that $t \rightarrow \infty$ this expression becomes a constant, which is a good proxy to the total population as at this point the total births and deaths must cancel:

$$G_\infty = r - b(v + \lambda^2) \quad (33)$$

The total population will be this expression divided by a . As b is a constant greater than zero, we can see here that it's preferable to have v be as negative as possible (sign convention) and λ as small as possible. Reintroducing λ :

$$G_\infty = r - b \left(v + \sqrt{\frac{\mathcal{D}}{b} + v^2} \right) \quad (34)$$

Despite the fact that technically G_∞ does increase for stronger drifts, there are diminishing returns to v becoming increasingly more negative due to the presence of the term in the square root.

IV. STATIONARY DISTRIBUTION

Having handled the non-linearity, we can return to how we expect the distributions to appear. We can recall our definition of the distribution:

$$\hat{n}(0, k, t) = e^{rt - bvt} e^{\left(\int_0^t dt' (-aN(t')) \right)} \sum_m c_{2m} e^{-b((4m+1)\lambda^2)t} e^{-\frac{1}{2}k^2(\lambda^2+v)} H_{2m}(k\lambda) \quad (35)$$

For a given $N(t)$, this is the distribution in Fourier space. We can observe that there are many different modes relaxing independently. It can also be seen that the higher modes relax down to 0 much faster than the 0 mode, as the relaxation time for the mode is decreasing in the mode number. In this case, it would appear that a good approximation as to the form of distribution in k space as $t \rightarrow \infty$ will be given by:

$$\hat{n}(0, k, t \rightarrow \infty) \sim \exp \left(-\frac{1}{2} k^2 (\lambda^2 + v) \right) \quad (36)$$

Therefore, in real space the distribution will have the following form:

$$\hat{n}(0, \chi, t \rightarrow \infty) \sim \frac{1}{\sqrt{\lambda^2 + v}} e^{-\frac{\chi^2}{2(\lambda^2 + v)}} \quad (37)$$

where we can see that negative v serves to make the distribution narrower and that larger \mathcal{D} broadens the distribution.

V. ADAPTATION TO A SUDDEN CHANGE IN THE ENVIRONMENT

In order now to analyze how the adaptation to a sudden change in the environment can proceed, we can imagine the following scenario:

- an original population is left to relax in an environment $b(\chi - \chi_0)^2$. This will form an distribution with the Gaussian character shown in the previous section
- at time $t = 0$, we shift the environment to be $b\chi^2$. We can also shift the magnitude of the death rate by changing a .
- we observe how the population relaxed to the previous environment adapts to the new environment as a function of the magnitude of the shift.

Using the results of the previous section, we begin with a population of distribution:

$$n(\chi, t) = \frac{G_\infty}{a_{old} \sqrt{2\pi} \sqrt{\lambda^2 + v}} e^{-\frac{\chi^2}{2(\lambda^2 + v)}} \quad (38)$$

This equation is the stationary (equilibrium) state of the original equation, as can be checked by substitution. We can make this equation correspond to a different landscape simply by shifting the χ variable to $\chi - \chi_0$. Taking the Fourier transform of this equation yields:

$$\hat{n}(0, k, 0) = \frac{G_\infty}{a_{old}} \exp \left(-\frac{1}{2} (v + \lambda^2) k^2 \right) \exp (ik\chi_0) \quad (39)$$

Recalling our definition of the Hermite series coefficients:

$$c_m = \frac{\lambda}{\sqrt{\pi} 2^m m!} \int dk \hat{n}(0, k, 0) e^{\frac{1}{2}vk^2 - \frac{1}{2}\lambda^2 k^2} H_m(k\lambda) \quad (40)$$

where the generic function $B(t)$ in terms of these functions is given by:

$$B(t) = e^{rt-bvt} \sum_{m=0}^{\infty} c_{2m} e^{-b((4m+1)\lambda^2)t} \frac{(-1)^m (2m)!}{(m)!} \quad (41)$$

Recall that the dynamics of the population subsequent to the perturbation is given by Eq. 31. The integral over the Hermite polynomials can be performed, leading to:

$$c_m = \frac{G_{\infty}}{a_{old} 2^m m!} \frac{\chi_0^m e^{-\frac{\chi_0^2}{4\lambda^2}}}{(-i)^m \lambda^m} \quad (42)$$

Inputting this expression into the formula for $B(t)$ again yields an analytic expression after the summation is performed:

$$B(t) = e^{rt-bvt} \frac{\sqrt{\pi} G_{\infty}}{a_{old}} e^{-\frac{4b\lambda^4 t + \chi_0^2}{4\lambda^2}} \cosh\left(\frac{\chi_0 e^{-2b\lambda^2 t}}{2\lambda}\right) \quad (43)$$

This leads to the growth function $G(t)$ of the form:

$$G(t) = r - b(\lambda^2 + v) - b\lambda\chi_0 e^{-2b\lambda^2 t} \tanh\left(\frac{\chi_0 e^{-2b\lambda^2 t}}{2\lambda}\right) \quad (44)$$

where it is easy to observe that over long timescales this growth function returns back to our previous form. However, we are now interested in the dynamics of adaption to this new environment. To that end we need to take the integral of $G(t)$ between 0 and t .

$$\int_0^t dt' G(t') = -bt(\lambda^2 + v) + \log\left(\cosh\left(\frac{\chi_0 e^{-2b\lambda^2 t}}{2\lambda}\right)\right) + rt - \log\left(\cosh\left(\frac{\chi_0}{2\lambda}\right)\right) \quad (45)$$

Although the expression in the denominator of $N(t)$ has no analytic form, it can be calculated numerically. This again shows that populations with large v adapt better to a sudden change. Plots of various scenarios are displayed over the next few pages.

The relaxation time in the main text is calculated by applying a shift χ_0 to a stationary population and measuring how long it takes for the population to recover to 99% of its equilibrium value.

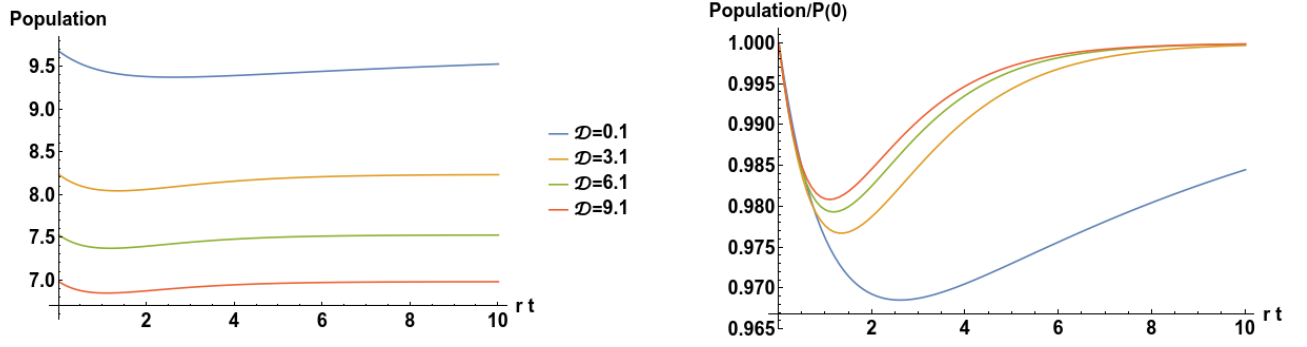


FIG. S2. Recovery of a population with $v = 0$ to a perturbation to χ for different D . Populations with a larger D have a smaller initial population, but a faster recovery time. ($a=1, b=1$)

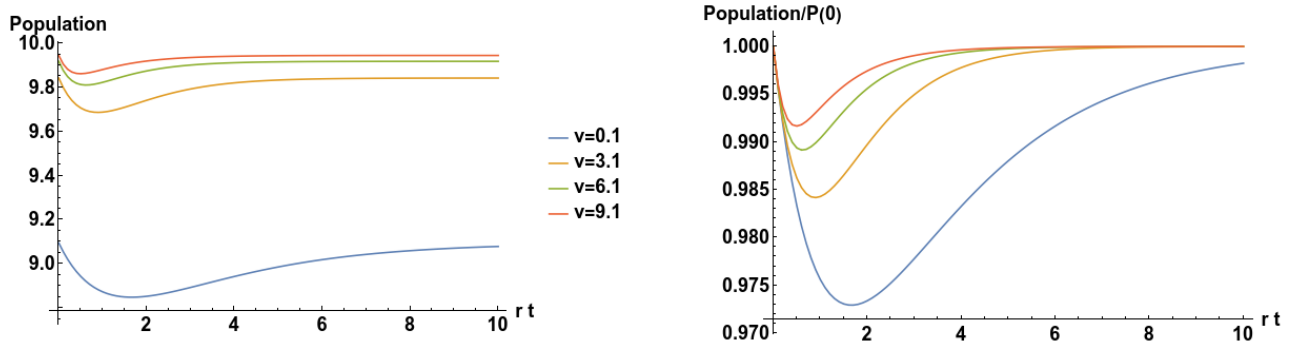


FIG. S3. Recovery of a population with $D = 1$ to a perturbation to χ for different values of v . Unlike the prior figure, populations with larger v have a larger initial population, but also a faster recovery time. ($a=1, b=1$)

For completeness, an analytic expression for the time evolution of a initially equilibrium population N_{eq} to a shift χ_0 is given by:

$$N(t) = \frac{N_{eq} \exp \left(-b\lambda^2 t - b\lambda v + \frac{\chi_0^2 e^{-4b\lambda^2 t}}{4\lambda^2} + rt - \frac{\chi_0^2}{4\lambda^2} \right)}{a N_{eq} e^{-bt(\lambda^2 + v) - \frac{\chi_0^2}{4\lambda^2}} \left(e^{rt} E_{\frac{1}{4}} \left(\frac{r - bv}{b\lambda^2} + 3 \right) \left(-\frac{e^{-4bt\lambda^2} \chi_0^2}{4\lambda^2} \right) - e^{bt(\lambda^2 + v)} E_{\frac{1}{4}} \left(\frac{r - bv}{b\lambda^2} + 3 \right) \left(-\frac{\chi_0^2}{4\lambda^2} \right) \right)} + 1 \quad (46)$$

where E is the exponential integral function.

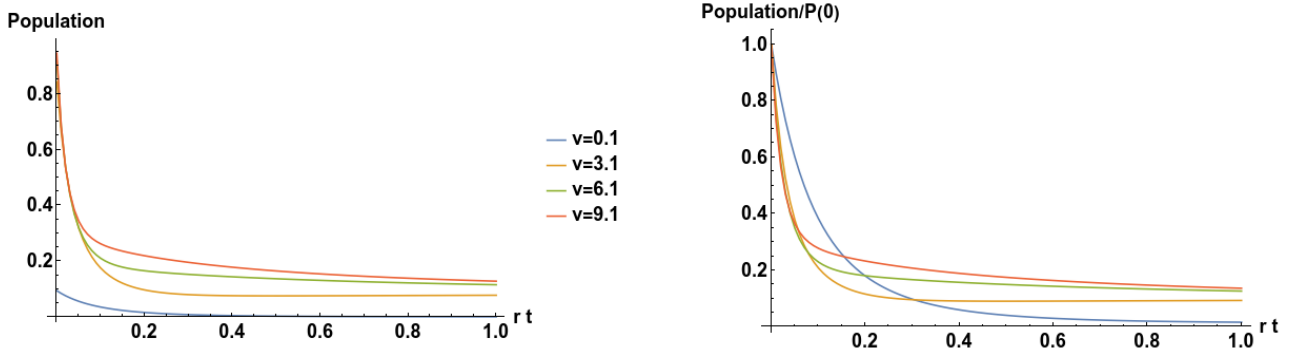


FIG. S4. Recovery of a population with $D = 1$ to a perturbation to χ and an additional increase in the death rate due to the change of environment, from $a = 1$ to $a = 10$, for different values of v . Populations with larger v have a larger initial population, but the effects of an environmental change such as this (reduction in niche size) are more catastrophic for the fast adaptors than the slow adaptors. ($b=1$)

VI. CONTINUAL VARIATION OF THE ENVIRONMENT

In order to analyze the effect of fast environmental switching, we shall first have to expound the nature of our solution under an arbitrary initial condition. One way that the equations so far can simulate dynamics is as follows:

- Select a starting initial condition $\hat{n}(0, k, 0)$ under some death probability
- Solve for

$$c_m = \frac{\lambda}{\sqrt{\pi} 2^m m!} \int dk \hat{n}(0, k, 0) e^{\frac{1}{2} v k^2 - \frac{1}{2} \lambda^2 k^2} H_m(k\lambda)$$

- Solve for

$$B(t) = \sum_m c_{2m} e^{-b((4m+1)\lambda^2)t} \frac{(-1)^m (2m)!}{m!}$$

- Solve for

$$N(t) = \frac{e^{rt-bvt} B(t) N(0)}{B(0) + aN(0) \int_0^t dt' e^{rt'-bvt'} B(t')}$$

- Solve for the final distribution

$$\hat{n}(0, k, t) = e^{rt-bvt} e^{\left(\int_0^t dt' (-aN(t'))\right)} \sum_m c_m e^{-b((2m+1)\lambda^2)t} e^{-\frac{1}{2} k^2 (\lambda^2 + v)} H_m(k\lambda)$$

- Shift the distribution by multiplying by $e^{ik\chi_0}$ and then use that distribution as an input and repeat

We will here distinguish between two types of shift strategies. When we multiply the distribution by $e^{ik\chi_0}$, we are in fact establishing that the population prior to the shift in the death probability was relaxing to the death rate $(\chi - \chi_0)^2$. We can either choose that the χ_0 we shift to are alternating between the same two points, hence allowing for lack of adaption to be good once the system returns back to its “initial” environment. Alternatively, we can choose χ_0 to be constantly moving in the same direction, such that the system is trying to relax to a completely new environment each time. In this case, lack of adaption is even more detrimental. We might anticipate that these two scenarios entail different dynamics.

We plot an example of how the distributions $n(\chi, T, t)$ shift for χ_0 that alternates between two points in figure S5

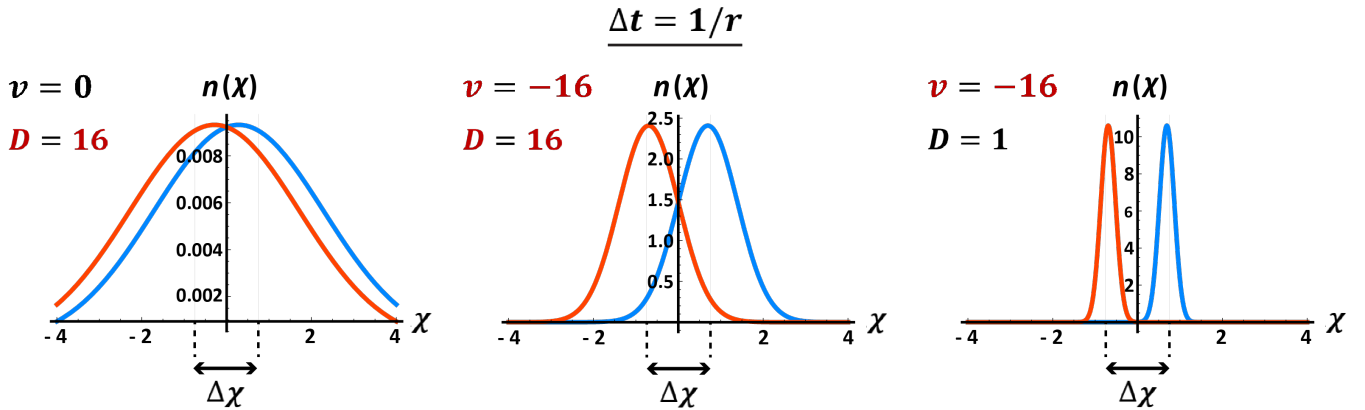


FIG. S5. Effect of negative v on the shape of the distribution $n(\chi)$ immediately prior to each switch. Shown for $v = -16$ with high and low D (middle and right panels) vs. $v = 0$ and high D (left).

VII. CONDITIONS ON SURVIVAL

In order to derive conditions on survival and extinction of populations, we consider the behavior of populations when subjected to a single shift. From this, we can infer the properties of repeated applications of environmental shifts on the population.

We begin with the time evolution equation of each mode $c_m(t)$, which is given by:

$$\frac{dc_m(t)}{dt} = (N_{eq}a - 2mb\lambda^2 - aN(t))c_m(t) \quad (47)$$

In order to estimate relaxation times, we focus on the zeroth mode of the system, which is also the largest mode:

$$\frac{dc_0(t)}{dt} = (N_{eq}a - aN(t))c_0(t) \quad (48)$$

Our estimate of the timescale shall proceed from the following facts: at equilibrium, only the zeroth mode of the system is non-zero, $c_m = 0, m > 0$, and we want the population to be the same, $N(\Delta t) = N(0)$. This corresponds to full relaxation of the zeroth mode after a shift. Equation (48) is solved by:

$$c_0(\Delta t) = c_0(0) \exp \left(N_{eq}a\Delta t - a \int_0^{\Delta t} N(t)dt \right) \quad (49)$$

The value $c_0(0)$ can be calculated by considering the equilibrium state and applying a shift of $\Delta\chi$ to it, then recalculating the value of $c_0(0)$. This is given by:

$$c_0(0) = N(0) \exp(-\Delta\chi^2/4\lambda^2) \quad (50)$$

We can also use the above expression to see that the relaxation to the new equilibrium is complete when $c_0(\Delta t) = N(0)$. We can therefore rearrange equation (50) to yield the following condition:

$$N(0) = N(0) \exp(-\Delta\chi^2/4\lambda^2) \left(N_{eq}a\Delta t - a \int_0^{\Delta t} N(t)dt \right) \quad (51)$$

where $N(0)$, the initial population number, cancels out. Equation (51) defines the condition for full relaxation of the first mode after a shift in the environment is applied. As $aN(t)$ is always positive, we can derive a rather simple condition from equation (51), which is that, minimally:

$$-\Delta\chi^2/4\lambda^2 + N_{eq}a\Delta t > 0 \quad (52)$$

in order for full relaxation of the zeroth mode to be completed in a time Δt after a shift $\Delta\chi$

This is then a condition on the survival of a population. Substituting for N_{eq} in eq. 52 gives:

$$\frac{\Delta P_D(\Delta\chi)}{4b\lambda^2} \leq [r - b(v + \lambda^2)] \Delta t \quad (53)$$

where $\Delta P_D(\Delta\chi) = b\Delta\chi^2$ corresponds to the initial jump in the rate of death due to the sudden displacement by $\Delta\chi$ from the state that minimized the death rate prior to the jump. Recalling that $r\Delta t$ is the number of births within a time interval Δt yields:

$$\tau\Delta P_D(\Delta\chi) + \frac{\Delta t}{4\tau} + bv\Delta t \leq N_{births}(\Delta t) \quad (54)$$

where $\tau = 1/4b\lambda^2$ is the characteristic timescale for broadening of the distribution $n(\chi)$. To achieve complete recovery within Δt , the number of births must compensate for the number of deaths in this time interval. The latter is displayed on the left-hand-side of eq. 54 as a sum of 3 components (left to right): (i) number of “excess” deaths (due to the shift by $\Delta\chi_0$ during the characteristic time for alleviating the increase in death rate by broadening the distribution toward the new minimum, (ii) number of deaths in time Δt due to limited carrying capacity, and (iii) the reduction (for $v < 0$) or increase (for $v > 0$) in the number of deaths by the directed change (drift) in χ over time Δt .

Note that the above analysis neglects the contributions of higher modes and therefore provides only an approximation to the exact necessary conditions for survival. Since higher modes decay faster than the lowest mode, the lowest mode approximation is expected to be better for higher values of Δt . This is clearly observed in figures 3C and 4B of the main text. The approximation breaks down for small Δt in the case periodic switching (figure 3C), concurring with the expectation that for sufficiently rapid periodic changes of the environment, the population can survive even without adapting to the higher death rate in the new environment.

VIII. FERTILITY, INHERITANCE AND THEIR EFFECT ON ADAPTATION TO ENVIRONMENT

We proceed to extend our baseline model to more complicated scenarios by relaxing both the $I = \delta(\chi - \chi')$ condition, as well as the $R = 0$ assumption. The first order effect of imperfect inheritance can be captured through resetting of the diffusion coefficient, as we shall demonstrate for a Gaussian inheritance function

$$I(\chi - \chi') = (1/\sqrt{2\pi\sigma^2}) \exp(-(\chi - \chi')^2/(2\sigma^2)). \quad (55)$$

We will use the convolution theorem to Fourier transform the first term in the second of Eqs. 1 into k -space:

$$\int_{-\infty}^{\infty} d\chi \exp(-ik\chi) \int P_R(T') I(\chi - \chi') n(T', \chi', t) d\chi' dT' = r \exp\left(-\frac{1}{2}k^2\sigma^2\right) \hat{n}(R, k, t) \quad (56)$$

Taylor expansion of the exponential yields the following:

$$r \left(1 - \frac{1}{2}k^2\sigma^2 + \frac{1}{4}k^4\sigma^4 - \dots\right) \hat{n}(R, k, t) \quad (57)$$

For a not too large value of σ , we ignore the corrections going as $\mathcal{O}(k^4)$ and higher. This might yield inaccurate solutions at large values of k , however, given how we designed the death rate, the number of individuals at large values of k is very small. To first order then we can see that the effect of imperfect inheritance can be given simply by:

$$D \rightarrow D + \frac{1}{2}r\sigma^2 \quad (58)$$

The other extension, accounting for declining fertility ($R > 0$) means that our transformed equations now exist in two modes, 0 and R :

$$\frac{\partial \hat{n}(R, k, t)}{\partial t} = (r - R)\hat{n}(R, k, t) + 2vbk \frac{\partial \hat{n}(R, k, t)}{\partial k} - Dk^2 \hat{n}(R, k, t) - a\hat{n}(R, k, t)N(t) + b \frac{\partial^2 \hat{n}(R, k, t)}{\partial k^2} \quad (59)$$

$$\frac{\partial \hat{n}(0, k, t)}{\partial t} = r\hat{n}(R, k, t) + 2vbk \frac{\partial \hat{n}(0, k, t)}{\partial k} - Dk^2 \hat{n}(0, k, t) - a\hat{n}(0, k, t)N(t) + b \frac{\partial^2 \hat{n}(0, k, t)}{\partial k^2} \quad (60)$$

This set of equations appears decidedly more complex than the previous ones, however, they have a similar stationary state, which is given by:

$$\hat{n}_S(0, k) = \frac{1}{a} \left(r - R - b \left(v + \sqrt{\frac{\mathcal{D}}{b} + v^2} \right) \right) \exp \left(-\frac{1}{2} \left(v + \sqrt{\frac{\mathcal{D}}{b} + v^2} \right) k^2 \right) \quad (61)$$

$$\hat{n}_S(R, k) = \frac{r - R}{ra} \left(r - R - b \left(v + \sqrt{\frac{\mathcal{D}}{b} + v^2} \right) \right) \exp \left(-\frac{1}{2} \left(v + \sqrt{\frac{\mathcal{D}}{b} + v^2} \right) k^2 \right) \quad (62)$$

The mode R , which tells us how much the species is reproducing, is equal to the mode 0 multiplied by the factor $(r - R)/r$. Moreover, it is possible to prove that once the system has reached a point where $n(R, k, t) = ((r - R)/r)n(0, k, t)$ no further dynamics in the ratio of $n(R, k, t)/n(0, k, t)$ occur, no matter the form of environmental change. This means that

the effective dynamics reduce again down to a single mode even with declining fertility. We can prove this by analyzing the following derivative:

$$\frac{\partial}{\partial t} \left(\frac{\hat{n}(R, k, t)}{\hat{n}(0, k, t)} \right) = \frac{\hat{n}(0, k, t)\hat{n}'(R, k, t) - \hat{n}(R, k, t)\hat{n}'(0, k, t)}{\hat{n}(0, k, t)^2} \quad (63)$$

where the equation has been derived using the quotient rule (n' refers here to derivative with respect to t). The differential equations given in 59 and 60 can be substituted into equation 63.

We can rewrite equations 59 and 60 as:

$$\frac{\partial \hat{n}(R, k, t)}{\partial t} = \hat{O}\hat{n}(R, k, t) + (r - R)\hat{n}(R, k, t) \quad (64)$$

$$\frac{\partial \hat{n}(0, k, t)}{\partial t} = \hat{O}\hat{n}(0, k, t) + r\hat{n}(0, k, t) \quad (65)$$

where \hat{O} is a linear operator. Substituting into 63 yields:

$$\frac{\partial}{\partial t} \left(\frac{\hat{n}(R, k, t)}{\hat{n}(0, k, t)} \right) = \frac{\hat{n}(0, k, t)(\hat{O}\hat{n}(R, k, t) + (r - R)\hat{n}(R, k, t)) - \hat{n}(R, k, t)(\hat{O}\hat{n}(0, k, t) + r\hat{n}(0, k, t))}{\hat{n}(0, k, t)^2} \quad (66)$$

Now we imagine that one of the modes is given as a multiple of another, for example $\hat{n}(R, k, t) = c\hat{n}(0, k, t)$. We could, for instance, instantiate the system in this configuration if we wished. This expression can be cleaned up, yielding:

$$\frac{\partial}{\partial t} \left(\frac{\hat{n}(R, k, t)}{\hat{n}(0, k, t)} \right) = c(r - R) - rc^2 \quad (67)$$

We make the observation that the dynamics of the quotient between modes R and 0 depends only on a constant if the two modes are initially a multiplicative constant of one another. When $c = (r - R)/r$, the right hand side of expression 63 is zero, i.e. the modes no longer change their relative magnitude compared to one another.

This proof doesn't demonstrate that the system will always be found in this state, but it does show that under a special condition we can ignore the dynamics of one of the modes. Moreover, analysis of equations 59 and 60 shows that both equations have the same dependence on k . In other words, they should approach the same distribution in k as time increases, but, as we already demonstrated, if the two modes are merely multiples of one another (which occurs when they have the same distribution in k) we can dispense with one of the modes. Therefore the model with declining fertility, despite appearing more complicated, can be mapped onto a single mode equation through the substitution

$\hat{n}(R, k, t) \rightarrow ((r - R)/r)\hat{n}(0, k, t)$ and we can be assured nothing funny will occur in the other modes.

Taking as our key mode $S = 0$, we have the following equation for a system with declining fertility and imperfect inheritance:

$$\frac{\partial \hat{n}(0, k, t)}{\partial t} = (r - R)\hat{n}(0, k, t) + 2vbk \frac{\partial \hat{n}(0, k, t)}{\partial k} - (D + \frac{1}{2}r\sigma^2)k^2 \hat{n}(0, k, t) - a\hat{n}(0, k, t)N(t) + b \frac{\partial^2 \hat{n}(0, k, t)}{\partial k^2} \quad (68)$$

whereupon we see that if we define $\tilde{r} = r - R$ and $\tilde{D} = (D + \frac{1}{2}r\sigma^2)$ we recover the perfect inheritance, no declining fertility model with new values of the replication and diffusion.

To realize the full potential of this model, we need to non-dimensionalize it. To this end we imagine the following substitutions:

$$\hat{n}(0, k, t) = \tilde{N}\tilde{n}(0, k, t) \quad (69)$$

$$k = k_c \tilde{k} \quad (70)$$

$$t = t_c \tau \quad (71)$$

$$\tilde{N} = \frac{r - R}{a} \quad (72)$$

This yields the following:

$$\begin{aligned} \frac{N}{t_c} \frac{\partial \tilde{n}(0, k, \tau)}{\partial \tau} = & (r - R)\tilde{N}\tilde{n}(0, k, \tau) + 2vb\tilde{N}\tilde{k} \frac{\partial \tilde{n}(0, \tilde{k}, \tau)}{\partial \tilde{k}} - \tilde{N}k_c^2 (D + \frac{1}{2}r\sigma^2)\tilde{k}^2 \tilde{n}(0, \tilde{k}, \tau) \quad (73) \\ & - a\tilde{N}^2 \tilde{n}(0, \tilde{k}, \tau) \tilde{n}(0, 0, \tau) + \frac{b\tilde{N}}{k_c^2} \frac{\partial^2 \tilde{n}(0, \tilde{k}, \tau)}{\partial \tilde{k}^2} \end{aligned}$$

dividing through by \tilde{N}/t_c gives:

$$\begin{aligned} \frac{\partial \tilde{n}(0, k, \tau)}{\partial \tau} = & (r - R)t_c \tilde{n}(0, k, \tau) + 2vbt_c \tilde{k} \frac{\partial \tilde{n}(0, \tilde{k}, \tau)}{\partial \tilde{k}} - t_c k_c^2 (D + \frac{1}{2}r\sigma^2)\tilde{k}^2 \tilde{n}(0, \tilde{k}, \tau) \quad (74) \\ & - aNt_c \tilde{n}(0, \tilde{k}, \tau) \tilde{n}(0, 0, \tau) + \frac{bt_c}{k_c^2} \frac{\partial^2 \tilde{n}(0, \tilde{k}, \tau)}{\partial \tilde{k}^2} \end{aligned}$$

where:

$$t_c = 1/(r - R) \quad (75)$$

$$k_c^2 = \frac{r - R}{D + \frac{1}{2}r\sigma^2} \quad (76)$$

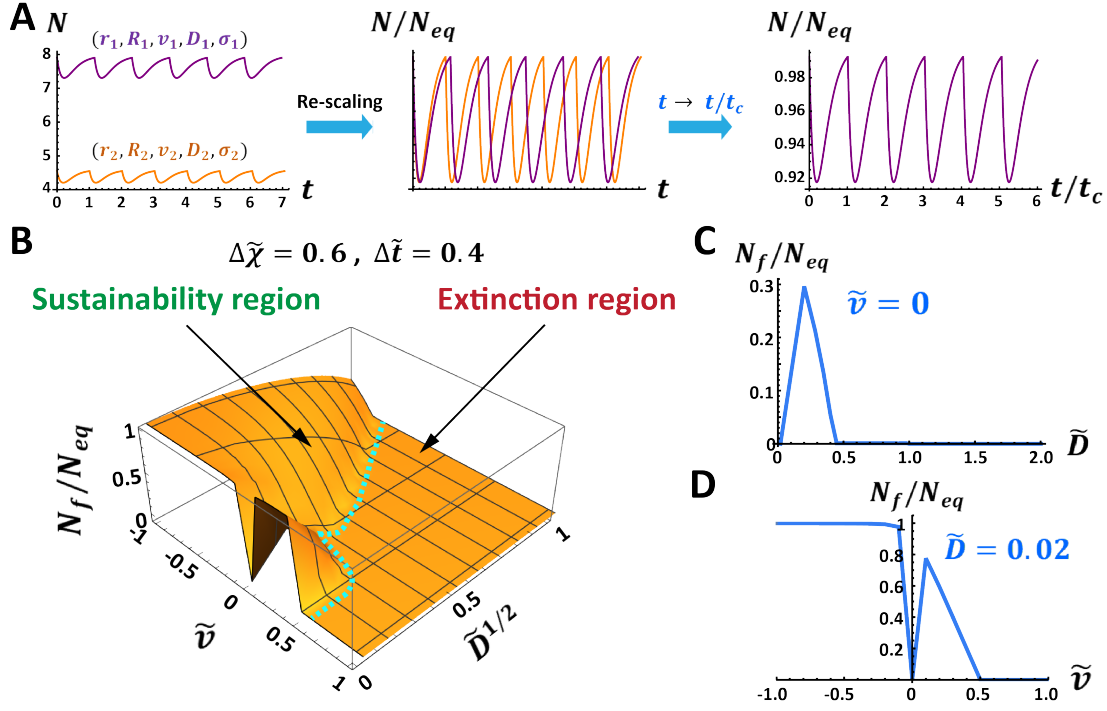


FIG. S6. Long-term solutions to successive shifts of the environment (Type II dynamics), taking into consideration the effects of imperfect fidelity of inheritance and age-dependent decline in fertility. Comparison between populations under the same effective challenge, is enabled by normalizing the shift magnitude $\Delta\chi$ and the duration between shifts Δt by χ_c and t_c . **(A)** Effects of scaling the population size by N_{eq} (middle) and measuring time in units of t_c (right). **(B)** Landscape of N_f/N_{eq} vs. \tilde{v} and $\tilde{D}^{1/2}$, for fixed $\tilde{\Delta\chi} = \Delta\chi/\chi_c = 0.6$ and $\tilde{\Delta t} = \Delta t/t_c = 0.4$. **(C,D)** Sections of the landscape in (B) for $\tilde{v} = 0$ (C) and $\tilde{D} = 0.02$ (D), demonstrating significant impacts of changes that are within a generation on long-term population survival and size

This equation reduces to:

$$\begin{aligned} \frac{\partial \tilde{n}(0, k, \tau)}{\partial \tau} = & \tilde{n}(0, k, \tau) + \frac{2vb}{r-R} \tilde{k} \frac{\partial \tilde{n}(0, \tilde{k}, \tau)}{\partial \tilde{k}} - \tilde{k}^2 \tilde{n}(0, \tilde{k}, \tau) \\ & - \tilde{n}(0, \tilde{k}, \tau) \tilde{n}(0, 0, \tau) + \frac{b(D + \frac{1}{2}r\sigma^2)}{(r-R)^2} \frac{\partial^2 \tilde{n}(0, \tilde{k}, \tau)}{\partial \tilde{k}^2} \end{aligned} \quad (77)$$

where there are only two dimensionless variables affecting the temporal change of the reduced

number of individuals in a population:

$$\tilde{v} = \frac{vb}{r-R} \text{ and } \tilde{D} = \frac{b(D + \frac{1}{2}r\sigma^2)}{(r-R)^2} \quad (78)$$

We conclude that, when measured in units of population of $(r-R)/a$ and units of time of $1/(r-R)$, two systems with the same value of \tilde{v} and of \tilde{D} will have exactly the same behavior.

The choice $k_c = \sqrt{b/r}\frac{1}{\chi}$ also works to reduce the possible number of parameters, and is the choice that we use in the main text.

One can then recast our previous results in terms of these new parameterizations. This is displayed in figure S6.

IX. χ -DEPENDENT REPLICATION

In the previous sections we considered the case in which the probability of death per unit time is affected by some state variable χ . We could imagine a different form of our model, one in which the probability to replicate depends on this variable instead. We may ask whether this model is isomorphic to the death rate model. To proceed, we write our equations in this updated form. For simplicity, we shall take the perfect inheritance version of our model $I(\chi - \chi') = \delta(\chi - \chi')$

We therefore have the following set of equations:

$$\frac{\partial n(T, \chi, t)}{\partial t} = -\frac{\partial n(T, \chi, t)}{\partial T} - \frac{\partial}{\partial \chi} (f(\chi)n(T, \chi, t)) + D\frac{\partial^2 n(T, \chi, t)}{\partial \chi^2} - P_D(T, \chi)n(T, \chi, t) \quad (79)$$

$$n(0, \chi, t) = \int P_R(T', \chi)n(T', \chi, t)dT' \quad (80)$$

$$n(\infty, x, t) = 0 \quad (81)$$

Another assumption that we make is that the replication function P_R is separable, $P_R(\chi, T) = \mathcal{X}(\chi)\mathcal{T}(T)$. The age-dependent effective replication rate will be given by:

$$\mathcal{T}(T) = r \exp(-RT) \quad (82)$$

We can take the age-related Laplace transform once more:

$$\frac{\partial \hat{n}(S, \chi, t)}{\partial t} = -S\hat{n}(S, \chi, t) + r\mathcal{X}(\chi)\hat{n}(R, \chi, t) - \frac{\partial}{\partial \chi} (f(\chi)\hat{n}(S, \chi, t)) + D\frac{\partial^2 \hat{n}(S, \chi, t)}{\partial \chi^2} - P_D(T, \chi)\hat{n}(S, \chi, t) \quad (83)$$

Let's consider the $R = 0$ case, leading to a model with a single mode:

$$\frac{\partial \hat{n}(0, \chi, t)}{\partial t} = r\mathcal{X}(\chi)\hat{n}(0, \chi, t) - \frac{\partial}{\partial \chi}(f(\chi)\hat{n}(0, \chi, t)) + D\frac{\partial^2 \hat{n}(0, \chi, t)}{\partial \chi^2} - P_D(T, \chi)\hat{n}(0, \chi, t) \quad (84)$$

This model is not identical to our baseline model. For one, we can see that the function $\mathcal{X}(\chi)$ must have a different form to the death rate probability we had previously. For if it were to go as $\sim \chi^2$ there would be regions where the probability to replicate would become very large as $|\chi|$ increases (N.B. this isn't a problem for the death rate). Similarly, a function of the form $-\chi^2$ would have the issue that the replication rate would become negative at certain points. Therefore, this function would have to be bounded to be realistic. It is not possible to straightforwardly map the results of our analytical model to this system.

Nevertheless, with these caveats included, a different interpretation of our model is possible, in the sense that we assign different meanings to terms. One can consider our normal model:

$$\begin{aligned} \frac{\partial \hat{n}(0, k, t)}{\partial t} = & r\hat{n}(0, k, t) + 2vbk\frac{\partial \hat{n}(0, k, t)}{\partial k} - Dk^2\hat{n}(0, k, t) \\ & - a\hat{n}(0, k, t)N(t) + b\frac{\partial^2 \hat{n}(0, k, t)}{\partial k^2} \end{aligned} \quad (85)$$

In our previous approach, we identified the term $r\hat{n}(0, k, t)$ as the term corresponding to reproduction, however one can consider the Fourier transform (with respect to k) of the combined term $r\hat{n}(0, k, t) + b\frac{\partial^2 \hat{n}(0, k, t)}{\partial k^2}$ as a χ dependent replication rate. One can bring these different interpretations to bear as the total growth of the population only depends on the relative balance between growth and death, and not their absolute values. One can consider the effect of χ as a reduction in birth rate up to the value $\chi = \sqrt{r/b}$ (where the total reproduction would be equal to zero). As we mentioned above, beyond this point this number becomes negative so can no longer be considered a "birth rate". Therefore this can also be considered an approximate, yet analytically tractable model of χ -dependent replication so long as the tails of the distribution don't contribute a lot to the dynamics. In such a view, the reproduction rate of an individual is reduced as they move in χ space, and individuals are removed from the population at a rate $a\hat{n}(0, k, t)N(t)$

X. MORE REALISTIC BIRTH RATES

We introduced a form of birth rate which is obviously inaccurate, that is that the birth rate probability is given by:

$$P_R(T) = r \exp(-RT) \quad (86)$$

which has the desired feature that fertility declines with age, but assumes unrealistically, that maximal fertility occurs at the moment of birth. We wish to conceive of a different scheme where fertility declines at large ages, but also starts at zero, meaning that organisms have some period of time before they begin to reproduce. In order to be able to preserve Laplace transforms for analytical simplicity, we introduce the following form of $P_R(T)$ that has the desired features:

$$P_R(T) = r (\exp(-R_1 T) - \exp(-R_2 T)) \quad (87)$$

So long as $R_2 > R_1$, we have a function that starts at zero at $T = 0$, and decays to zero as $T \rightarrow \infty$, but with finite reproduction at intermediate ages, with a maximum at $(\ln R_1/R_2)/(R_1 - R_2)$

We reuse the operator introduced in equations (64) to succinctly write down the time evolution equations for this more complicated model. In such a system we only have 3 important modes:

$$\frac{\partial \hat{n}(0, k, t)}{\partial t} = r(\hat{n}(R_1, k, t) - \hat{n}(R_2, k, t)) + \hat{O}\hat{n}(0, k, t) \quad (88)$$

$$\frac{\partial \hat{n}(R_1, k, t)}{\partial t} = (r - R_1)\hat{n}(R_1, k, t) - r\hat{n}(R_2, k, t) + \hat{O}\hat{n}(R_1, k, t) \quad (89)$$

$$\frac{\partial \hat{n}(R_2, k, t)}{\partial t} = r\hat{n}(R_1, k, t) - (r + R_2)\hat{n}(R_2, k, t) + \hat{O}\hat{n}(R_2, k, t) \quad (90)$$

Now we follow the familiar procedure of looking at how the ratios of these modes to one another appear. As, once again, the evolution in k -space is identical for every mode, these parts can be separated out and we can analyze the ratios of these modes to one another, i.e., we search for solutions where $\hat{n}(R_1, k, t) = c_1 \hat{n}(0, k, t)$ and $\hat{n}(R_2, k, t) = c_2 \hat{n}(0, k, t)$, we can employ the quotient rule and the equations (88)-(90) for this purpose:

$$\frac{\partial}{\partial t} \left(\frac{\hat{n}(R_1, k, t)}{\hat{n}(0, k, t)} \right) = (c_1 - 1)(c_1 - c_2)(-r) - c_1 R_1 \quad (91)$$

$$\frac{\partial}{\partial t} \left(\frac{\hat{n}(R_2, k, t)}{\hat{n}(0, k, t)} \right) = (c_2 - 1)(c_2 - c_1)r - c_2 R_2 \quad (92)$$

Once again, this suggests that there are particular values of the numbers c_1 and c_2 in which the quotients of the modes do not display any time evolution. In other words, the apparent fact that we have three modes can again be reduced to a single mode, despite the additional complexity. The equations above do in fact have the following solutions:

$$\begin{aligned}
c_1 &= 0 & c_2 &= 0 \\
c_1 &= -\frac{R_1(R_1-R_2+\sqrt{(R_1-R_2-4)(R_1-R_2)}-2)+2R_2}{2(R_1-R_2)} & c_2 &= \frac{1}{2} \left(-\frac{\sqrt{(R_1-R_2-4)(R_1-R_2)}R_2}{R_1-R_2} + R_2 + 2 \right) \\
c_1 &= \frac{R_1(-R_1+R_2+\sqrt{(R_1-R_2-4)(R_1-R_2)}+2)-2R_2}{2(R_1-R_2)} & c_2 &= \frac{1}{2} \left(\frac{\sqrt{(R_1-R_2-4)(R_1-R_2)}R_2}{R_1-R_2} + R_2 + 2 \right).
\end{aligned} \tag{93}$$

Of these solutions, only the second displays the property that $c_1 > c_2$ when $R_2 > R_1 > 0$, which is a relevant feature for real distributions as we expect the number of individuals to decline with age (we expect $\int_0^\infty dT n(T, x, t) \exp(-RT)$ to decrease as R increases).

Inserting these values into the equation for $\hat{n}(0, k, t)$ and simplifying leads to the following equation:

$$\frac{\partial \hat{n}(0, k, t)}{\partial t} = \left(\frac{1}{2} \left(\sqrt{(R_2 - R_1)(4r - R_1 + R_2)} - R_1 - R_2 \right) \right) \hat{n}(0, k, t) + 2vbk \frac{\partial \hat{n}(0, k, t)}{\partial k} - \tag{94}$$

$$Dk^2 \hat{n}(0, k, t) - a \hat{n}(0, k, t) N(t) + b \frac{\partial^2 \hat{n}(0, k, t)}{\partial k^2}$$

which, as we can see, is identical to the equations we have dealt with in the main text. The only remaining factor is to determine whether the fixed point is one to which the system naturally ends up in, which can be studied by ensuring that the flow of equations (91) and (92) is always towards the fixed point, even if it wasn't initialized at the fixed point. This can be checked graphically by looking at the flow plot against c_1, c_2 for a given value of r, R_1, R_2 . These graphs display a flow towards the fixed point value so long as there is a non-zero population that exists at this value. When considering long time phenomena, one can ignore initial transients associated with non-equilibrated values of c_1, c_2 and just observe what were to occur in such systems for the reduced model.
