Suppression of beneficial mutations in dynamic microbial populations
Supplemental Information

Philip Bittihn,[1,*] Jeff Hasty,[1,2,3,†] and Lev S. Tsimring[1,†]

[1] *BioCircuits Institute, University of California San Diego, La Jolla, California 92093, USA*
[2] *Department of Bioengineering, University of California San Diego, La Jolla, California 92093, USA*
[3] *Molecular Biology Section, Division of Biological Science, University of California San Diego, La Jolla, California 92093, USA*

(Dated: October 6, 2016)

* pbittihn@ucsd.edu
† Co-senior authors

Figures and equations in the main text are referenced using a superscript letter "M". For example, Eq. (1) in the main text would be referenced as Eq. $(1)^M$.

I. COARSE-GRAINED MUTATION DYNAMICS

Here, we will give a formal argument to show that, for sufficiently small μ, mutations in the model

$$X \xrightarrow{\alpha(1-\mu)} 2X \tag{1a}$$

$$X \xrightarrow{\alpha\mu} X + Y \tag{1b}$$

$$Y \xrightarrow{(1+s)\alpha} 2Y \tag{1c}$$

occur according to a Poisson process with rate $\alpha\mu\overline{n_X}$, where $\overline{n_X}$ is the time-averaged number of wild-type cells for $\mu = 0$. The argument is based on the separation of time scales between the fast fluctuations in population size and rare mutations, similar to the derivation of the slow-scale stochastic simulation algorithm in Ref. [1]. However, in our case, it is necessary to consider the properties of the population dynamics not described by the reactions of Eq. (1) themselves. For a constant population, the statement above is clearly true since $n_X = \overline{n_X}$ for all times and therefore the rate of Eq. (1b) is $\alpha\mu\overline{n_X}$ by definition. However, as we will show below, the same formula is also a good approximation for the rate of mutations in the periodic dilution scenario if μ is small.

For simplicity, the number of wild-type individuals n_X will be denoted as x in this section.

Assume that there are currently no mutant cells in the population. What is the probability density $p_{\text{next}}(t)$ for the next mutation to occur at some time $t > 0$? If we split up the interval from 0 to t into $M \in \mathbb{N}$ equally spaced intervals $\Delta t = t/M$, the probability of no mutation occuring in the ith subinterval is $1 - \alpha\mu x(i\Delta t)\Delta t$. As the probability for a mutation occuring between t and $t + \mathrm{d}t$ is $\alpha\mu x(t)\,\mathrm{d}t$,

$$p_{\text{next}}(t)\,\mathrm{d}t = \alpha\mu x(t)\,\mathrm{d}t \cdot \lim_{M\to\infty} \prod_{i=1}^{M} \left[1 - \alpha\mu x\left(i\frac{t}{M}\right)\frac{t}{M}\right]. \tag{2}$$

Taking the log of the second factor allows us to convert it to the integral

$$\begin{aligned}
\lim_{M\to\infty} \log \prod_{i=1}^{M} \left[1 - \alpha\mu x\left(i\frac{t}{M}\right)\frac{t}{M}\right] &= \lim_{M\to\infty} \sum_{i=1}^{M} \log\left[1 - \alpha\mu x\left(i\frac{t}{M}\right)\frac{t}{M}\right] \\
&= -\alpha\mu \lim_{M\to\infty} \sum_{i=1}^{M} x\left(i\frac{t}{M}\right)\frac{t}{M} \\
&= -\alpha\mu \int_0^t x(\tau)\,\mathrm{d}\tau.
\end{aligned}$$

Substituting this into Eq. (2) yields

$$p_{\text{next}}(t) = \alpha\mu x(t) \cdot \exp\left(-\alpha\mu \int_0^t x(\tau)\,\mathrm{d}\tau\right) \tag{3}$$

This is the exact probability density for the time of the next mutation. The following

argument holds if the dynamics of x is made up of short time intervals (called "cycles" hereafter according to the model in the main text, but the logic applies more generally) that are all statistically identical, which is the case for the periodic dilution scenario. Then, let Δt_{meso} be some mesoscopic time period, where mesoscopic means that it is large compared to the length of the cycles and small compared to the time scale on which the second (exponential) factor of Eq. (3) changes. Choosing Δt_{meso} is always possible if μ is sufficiently small. This choice implies that x averaged over Δt_{meso} is close to $\overline{n_X}$ and that there are close to $\Delta t_{\text{meso}}/T$ cycles within that time period, where T is the average length of the cycles (in the case of periodic dilutions, the length of all cycles is exactly T). The probability that the next mutation will occur anywhere within $[t, t + \Delta t_{\text{meso}}]$ is

$$\tilde{p}_{\text{next}}(t)\Delta t_{\text{meso}} = \int_t^{t+\Delta t_{\text{meso}}} \alpha\mu x(t') \cdot \exp\left(-\alpha\mu \int_0^{t'} x(\tau)\,\mathrm{d}\tau\right) \mathrm{d}t' \tag{4}$$

$$= \left[-\exp\left(-\alpha\mu \int_0^{t'} x(\tau)\,\mathrm{d}\tau\right)\right]_t^{t+\Delta t_{\text{meso}}} \tag{5}$$

$$= \exp\left(-\alpha\mu \int_0^{t} x(\tau)\,\mathrm{d}\tau\right)\left[1 - \exp\left(-\alpha\mu \int_t^{t+\Delta t_{\text{meso}}} x(\tau)\,\mathrm{d}\tau\right)\right] \tag{6}$$

$$\approx \alpha\mu\overline{n_X} \exp\left(-\alpha\mu \int_0^{t} x(\tau)\,\mathrm{d}\tau\right)\Delta t_{\text{meso}}, \tag{7}$$

where the last step was possible due to the mesoscopic scale of Δt_{meso}. Because $\alpha\mu$ is assumed to be sufficiently small, the integral will already be close to $\overline{n_X}t$ when the argument to the exponential deviates substantially from zero. Replacing this term yields:

$$\tilde{p}_{\text{next}}(t)\Delta t_{\text{meso}} = \alpha\mu\overline{n_X} \exp(-\alpha\mu\overline{n_X}t)\Delta t_{\text{meso}}. \tag{8}$$

This confirms the intuitive fact that, on a coarse-grained time scale (where the fluctuations of the population size are averaged out), mutations occur according to a Poisson process with rate $\alpha\mu\overline{n_X}$ and, therefore, the average period of mutations is $T_{\text{mut}} = (\alpha\mu\overline{n_X})^{-1}$.

After a mutation has occurred, the population consists of both Xs and Ys, but only for a transitional homogenization period τ_{hom}, until the mutation has either been fixed or eliminated. For each single mutation τ_{hom} will depend on the state the population was in when the mutation was introduced. However, if the typical τ_{hom} is much smaller than the typical period τ_{mut} at which mutations occur (i.e. for small enough μ), then the transient coexistence of wild-type and mutant individuals will not significantly alter the statistics and the fixation time is

$$\tau = \frac{\tau_{\text{mut}}}{p} + \tau_{\text{hom}}^{+} \approx \frac{\tau_{\text{mut}}}{p} = (\alpha\mu\overline{n_X}p)^{-1}, \tag{9}$$

where p is the probability that a mutation occurring at unknown time and population state will eventually become fixed. τ_{hom}^{+} is the homogenization time conditioned on the fixation of the mutation. However, the contribution to τ is small if τ_{mut} is large and hence we ignore it after the first step in Eq. (9).

II. PROBABILITY OF MUTATION WITHIN A CYCLE

In general, the fixation probability of a mutation can depend on the time it is introduced and the state of the population at that time. To calculate the unconditional fixation probability p for an arbitrary mutation, it is therefore necessary to know the probability distribution of the introduction of a mutation across different times and states.

Assume that no mutation has occurred up to time t. As shown above, for sufficiently small μ, there exists a time scale Δt_{meso} spanning several cycles, on which the exponential factor in Eq. (3) does not change significantly. Therefore, the probability distribution for the time to the next mutation (i.e. when reaction (1b) fires next) is proportional to $\alpha\mu x(t)$ for several cycles, without the dampening factor. Because $T < \Delta t_{\text{meso}}$, the same is certainly true for the next complete cycle. In addition, all the cycles are assumed to be statistically identical. Therefore, the timing of the next mutation within the cycle in which it occurs does not depend on t (the time before which no mutation has occurred). This means that there is a universal probability $p_{\text{mut}}(x^*, \theta)$ that the next mutation happens at time θ into a cycle (in our case, the time since the last dilution event), when there are x^* wild-type individuals. Since $p_{\text{mut}}(x^*, \theta) \propto \alpha\mu x^*$, the probability is given by

$$p_{\text{mut}}(x^*, \theta)\,\mathrm{d}\theta = \frac{p_{\text{x}}(x^*, \theta) \cdot x^*\,\mathrm{d}\theta}{\int_0^\infty \sum_{k=0}^\infty p_{\text{x}}(k, \theta') \cdot k\,\mathrm{d}\theta'}, \tag{10}$$

where $p_x(x^*, \theta)$ is the probability that the population has size x^* at time θ into the cycle. For our specific case in the main text, only $p_{\text{mut}}(\theta)$ is required, i.e. the probability that the next mutation occurs at time θ after the last dilution event. We obtain this probability by summing over all possible x^*:

$$p_{\text{mut}}(\theta) = \sum_{k=0}^\infty p_{\text{mut}}(x^*, \theta) = \frac{\sum_{k=0}^\infty p_{\text{x}}(k, \theta) \cdot k\,\mathrm{d}\theta}{\int_0^\infty \sum_{k=0}^\infty p_{\text{x}}(k, \theta') \cdot k\,\mathrm{d}\theta'} = \frac{m_1(N_s; \alpha; \theta)}{\int_0^\infty m_1(N_s; \alpha; \theta')\,\mathrm{d}\theta'}, \tag{11}$$

where m_1 is the first moment of the population size. Since every cycle starts with N_s individuals which are constantly dividing with rate α, the first moment is simply $m_1(N_s; \alpha; \theta) = N_s \exp(\alpha t)$, leading to

$$p_{\text{mut}}(\theta) = \frac{\exp(\alpha\theta)}{\int_0^T \exp(\alpha\theta')\,\mathrm{d}\theta'}. \tag{12}$$

III. DIFFUSION APPROXIMATION FOR CONSTANT POPULATION

Although the fixation probability for a constant population was successfully derived in Eq. $(5)^M$ as a limit of the dynamic case, a more accurate approximation—without taking into account the effect of ξ^2—can be obtained directly by treating each division event as a "generation" in the context of Kimura's diffusion theory:
The probability that a non-mutant will divide next is $N_c(1-y)\alpha/[N_c(1-y)\alpha + N_c y\alpha(1+s)] = (1-y)/[(1-y)+y(1+s)]$, where y is the fraction of mutants in the population. Similarly,

the probability of a mutant dividing next is $y(1+s)/[(1-y)+y(1+s)]$. For the number of mutants in the population to change, either a mutant has to divide and then a non-mutant is chosen to be removed, or vice versa. The probability for the removal of a non-mutant after a mutant has divided is $(1-y)/(1+1/N_c)$, whereas the probability for the removal of a mutant after a non-mutant has divided is $y/(1+1/N_c)$. Therefore, the mean change in one cell division is:

$$M_{\delta y,c}(y) = \frac{1}{N_c}\left[\frac{y(1+s)}{1-y+y(1+s)}\cdot\frac{1-y}{1+1/N_c} - \frac{1-y}{1-y+y(1+s)}\cdot\frac{y}{1+1/N_c}\right] \tag{13}$$

$$= \frac{sy(1-y)}{(1+sy)(N_c+1)} \tag{14}$$

Similarly, the second moment is $\frac{(s+2)y(1-y)}{N_c(1+sy)(N_c+1)}$ and therefore, the variance evaluates to:

$$V_{\delta y,c}(y) = \frac{(s+2)y(1-y)}{N_c(1+sy)(N_c+1)} - \left(\frac{sy(1-y)}{(1+sy)(N_c+1)}\right)^2 \tag{15}$$

$$= \frac{(s+2)y(1-y)(1+sy)(N_c+1) - N_c s^2 y^2 (1-y)^2}{N_c(1+sy)^2(N_c+1)^2}. \tag{16}$$

Dividing the two equations yields

$$\frac{M_{\delta y,c}(y)}{V_{\delta y,c}(y)} = \frac{sN_c(1+sy)(N_c+1)}{(s+2)(1+sy)(N_c+1) - N_c s^2 y(1-y)}. \tag{17}$$

Expanding about $s=0$ shows that to second order, $\frac{M_{\delta y,c}(y)}{V_{\delta y,c}(y)}$ is independent of y:

$$\frac{M_{\delta y,c}(y)}{V_{\delta y,c}(y)} = \frac{N_c}{4}s(2-s) + \mathcal{O}(s^3) \tag{18}$$

We can therefore use $\Lambda_c = N_c s(2-s)/4$ in Eq. $(2)^M$ to calculate $u_c(y)$. In this case, every mutation starts with a single mutant cell in a population of N_c+1 cells. To be fixed, the mutation has to survive the removal step of the Moran process and then become fixed starting from a frequency $1/N_c$, finally leading to

$$p_c = \frac{N_c}{N_c+1}u_c(1/N_c) = \frac{N_c}{N_c+1}\cdot\frac{1-\exp(-\frac{1}{2}s(2-s))}{1-\exp(-\frac{N_c}{2}s(2-s))}. \tag{19}$$

We can read off Eq. (19) that $\bar{y}_c = 1/(N_c+1)$ if we wanted to use $p_c \approx u_c(\bar{y}_c)$ in analogy with the dynamic case. Note that Eq. $(5)^M$ (for $\xi^2 = 1$) is just a less accurate version of Eq. (19), which assumes $N_c/(N_c+1) \approx 1$ and uses only a first-order approximation for Λ_c. The two formulas converge for large N_c and small s, and therefore, Eq. (19) would still lead to the same result for the asymptotic fixation time ratio Δ in Eq. $(7)^M$.

IV. DIFFUSION APPROXIMATION FOR PERIODIC DILUTION: Λ_d

To calculate Λ_d, we note that the first and second moment a population of cells that starts out with N_0 cells and then grows by dividing with a rate λ are given by

$$m_1(N_0,\lambda,t) = N_0\exp(\lambda t) \tag{20a}$$

$$m_2(N_0,\lambda,t) = N_0\exp(\lambda t)[(\xi^2+N_0)\exp(\lambda t) - \xi^2] \tag{20b}$$

for $\xi^2 = 1$. These moments can be evaluated at $t = T$ to obtain the mean and variance contributed by the growth cycle for each sub-population, i.e. non-mutants and mutants, which start out with $(1-y)N_s$ and yN_s cells. In general, for a population starting with N_0 individuals and dividing with rate λ, the mean $M_\lambda(N_0)$ and the variance $V_\lambda(N_0)$ evaluate to:

$$M_\lambda(N_0) = N_0 \exp(\lambda T) \tag{21a}$$

$$\begin{aligned} V_\lambda(N_0) &= N_0 \exp(\lambda T)[(\xi^2 + N_0)\exp(\lambda T) - \xi^2] - N_0^2 \exp^2(\lambda T) \\ &= \xi^2 N_0 \exp(\lambda T)[\exp(\lambda T) - 1] \end{aligned} \tag{21b}$$

Note that ξ^2 in Eq. (20b) effectively scales the variance here. Assume the fraction of mutants at the beginning of a cycle is y. What we are first interested in are the mean and variance of the fraction of mutants at the end of the growth phase as a function of y. Since $y = n_Y/(n_X + n_Y)$, we use second and first-order approximations for the mean and variance of a ratio, respectively:

$$\left\langle \frac{a}{b} \right\rangle \approx \frac{\langle a \rangle}{\langle b \rangle} - \frac{\mathrm{cov}(a,b)}{\langle b \rangle^2} + \frac{\mathrm{var}(b)\langle a \rangle}{\langle b \rangle^3} \tag{22a}$$

$$\mathrm{var}\left(\frac{a}{b}\right) \approx \frac{\langle a \rangle^2}{\langle b \rangle^2}\left[\frac{\mathrm{var}(a)}{\langle a \rangle^2} - 2\frac{\mathrm{cov}(a,b)}{\langle a \rangle \langle b \rangle} + \frac{\mathrm{var}(b)}{\langle b \rangle^2}\right] \tag{22b}$$

In our case, $a = n_Y$ and $b = n_X + n_Y$, where n_X and n_Y are uncorrelated and so $\mathrm{cov}(a,b) = \mathrm{var}(n_Y)$. Therefore, all means and variances necessary for the above approximation can be calculated exactly from Eq. (21), giving approximations for the mean fraction of mutants $M_{\mathrm{end}}(y)$ at the end of the growth cycle and its variance $V_{\mathrm{end}}(y)$ introduced by the growth process:

$$\begin{aligned} M_{\mathrm{end}}(y) = {} & \frac{M_{\alpha(1+s)}(yN_s)}{M_{\alpha(1+s)}(yN_s) + M_\alpha((1-y)N_s)} - \frac{V_{\alpha(1+s)}(yN_s)}{[M_{\alpha(1+s)}(yN_s) + M_\alpha((1-y)N_s)]^2} \\ & + \frac{[V_\alpha((1-y)N_s) + V_{\alpha(1+s)}(yN_s)]M_{\alpha(1+s)}(yN_s)}{[M_{\alpha(1+s)}(yN_s) + M_\alpha((1-y)N_s)]^3} \\ V_{\mathrm{end}}(y) = {} & \left(\frac{M_{\alpha(1+s)}(yN_s)}{M_{\alpha(1+s)}(yN_s) + M_\alpha((1-y)N_s)}\right)^2 \cdot \\ & \left[\frac{V_{\alpha(1+s)}(yN_s)}{[M_{\alpha(1+s)}(yN_s)]^2} - 2\frac{V_{\alpha(1+s)}(yN_s)}{[M_{\alpha(1+s)}(yN_s) + M_\alpha((1-y)N_s)] \cdot M_{\alpha(1+s)}(yN_s)}\right. \\ & \left. + \frac{V_\alpha((1-y)N_s) + V_{\alpha(1+s)}(yN_s)}{[M_{\alpha(1+s)}(yN_s) + M_\alpha((1-y)N_s)]^2}\right] \end{aligned}$$

The dilution itself is the selection of N_s cells according to a hypergeometric distribution. For a mean growth cycle with a fraction of $z = M_{\mathrm{end}}(y)$ mutants at the end, this selection process will have a mean $N_s z$ and variance $N_s z(1-z)\frac{N_{\mathrm{end}}^{\mathrm{total}} - N_s}{N_{\mathrm{end}}^{\mathrm{total}} - 1}$, where $N_{\mathrm{end}}^{\mathrm{total}} = M_{\alpha(1+s)}(yN_s) + M_\alpha((1-y)N_s)$ is the total number of cells at the end of the cycle, before the dilution happens. Dividing by N_s and N_s^2, respectively, yields the mean fraction of mutants after selection and the variance due to the selection process alone:

$$M_{\mathrm{select}}(y) = M_{\mathrm{end}}(y) \tag{23}$$

$$V_{\mathrm{select}}(y) = \frac{1}{N_s} M_{\mathrm{end}}(y)(1 - M_{\mathrm{end}}(y)) \frac{M_{\alpha(1+s)}(yN_s) + M_\alpha((1-y)N_s) - N_s}{M_{\alpha(1+s)}(yN_s) + M_\alpha((1-y)N_s) - 1} \tag{24}$$

However, in reality, z is itself a random variable with variance $V_{\text{end}}(y)$. If this variance is small, such that the variance of the selection process is approximately constant for different possible z around $z = M_{\text{end}}(y)$, then the variances just add up. Thus, we have for the periodic dilution case (subscript d)

$$M_{\delta y,d}(y) = M_{\text{end}}(y) - y \tag{25}$$

$$V_{\delta y,d}(y) = V_{\text{end}}(y) + V_{\text{select}}(y). \tag{26}$$

Expanding $M_{\delta y,d}/V_{\delta y,d}$ around $s = 0$ reveals that it is independent of y up to first order:

$$\frac{M_{\delta y,d}}{V_{\delta y,d}} = \frac{(N_s - \xi^2)(fN_s - 1)f\log(f)}{(f-1)(fN_s(1+\xi^2) - \xi^2)}s + \mathcal{O}(s^2), \tag{27}$$

where $f = \exp(\alpha T)$ as in the main text. Thus we obtain

$$\Lambda_d = \frac{(N_s - \xi^2)(fN_s - 1)f\log(f)}{(f-1)(fN_s(1+\xi^2) - \xi^2)}s \tag{28}$$

$$\approx \frac{(N_s - \xi^2)f\log(f)}{(f-1)(1+\xi^2)}s. \tag{29}$$

In the last step, we neglected the small additive corrections to fN_s terms in the enumerator and the denominator.

V. DIFFUSION APPROXIMATION FOR PERIODIC DILUTION: $\bar{y}_d$

To estimate the average initial fraction of mutants $\bar{y}_d$, we consider the growth cycle in which the mutation first occurs: The probability of a mutation due to reaction (1b) in a time interval $\mathrm{d}\theta$ is $\mathrm{d}\theta \cdot p_x(x,\theta)\cdot x / \int_0^T \sum_{x=N_s}^{\infty} p_x(x,\theta')\cdot x\,\mathrm{d}\theta'$ according to Eq. (10) (for small μ). If the mutation happens at time θ into the cycle, the expected number of mutants at the end of the cycle is given by $m_1(m_s;\alpha(1+s);T-\theta) = m_s\exp[(1+s)\alpha(T-\theta)]$, where $m_1(N_0;\lambda;t)$ is defined as in Eq. (20) and m_s is the initial size of the mutant subpopulation. Therefore, the first moment for the number of *m*utants at the end of the *i*nitial cycle (subscript mi) can be calculated as

$$m_{1,mi} = \int_0^T \sum_{x=N_s}^{\infty} m_1(m_s;\alpha(1+s);T-\theta)\frac{p_x(x,\theta)\cdot x}{\int_0^T m_1(N_s;\alpha;\theta')\,\mathrm{d}\theta'}\,\mathrm{d}\theta \tag{30}$$

$$= \frac{m_s f(f^s - 1)}{s(f-1)} \tag{31}$$

Similarly, the expected number of *w*ildtype individuals at the end of the *i*nitial cycle (subscript wi) is $m_1(x,\alpha;T-\theta) = x\exp[\alpha(T-\theta)]$, yielding

$$m_{1,\text{wi}} = \int_0^T \sum_{x=N_s}^{\infty} x\exp[\alpha(T-\theta)]\frac{p_x(x,\theta)\cdot x}{\int_0^T m_1(N_s;\alpha;\theta')\,\mathrm{d}\theta'}\,\mathrm{d}\theta$$

$$= f\left[N_s + \xi^2\left(1 - \frac{\log(f)}{f-1}\right)\right] \tag{32}$$

by realizing that the sum over all the x-dependent terms under the integral simply evaluates to $m_2(N_s;\alpha;\theta)$. Setting $\xi^2 = 1$ and $m_s = 1$, it is noteworthy that, even in the limit of $s \to 0$ (i.e. mutants having the same growth rate as non-mutants), the sum of the two terms (31) and (32) becomes $(1+N_s)f$, which means that the population on average has a

larger size at the end of a cycle, if a mutation is introduced, than it has otherwise. In fact, even the number of non-mutants is larger than usual at the end of these cycles, because $\log(f)/(f-1)$ is smaller than 1 for all $f > 1$ in Eq. (32). This is due to the fact that the cases we select from the set of *all* cycles to calculate this average are conditioned on the introduction of a mutation, which is more likely to happen at larger population sizes and therefore introduces a bias. Consistently, Eq. (32) converges to exactly fN_s if there are no population size fluctuations during growth, corresponding to $\xi^2 = 0$.
Eqs. (31) and (32) lead to the estimate

$$\bar{y}_d \approx \frac{m_{1,\mathrm{mi}}}{m_{1,\mathrm{wi}} + m_{1,\mathrm{mi}}} \approx \frac{(f^s - 1)m_s}{s(f-1)N_s}, \tag{33}$$

where the last step is valid for large N_s. Since $\bar{y}_d$ is in fact a ratio of two random quantities, we also calculated higher moments and the covariance of the two subpopulations in the initial cycle to obtain a more accurate estimate of $\bar{y}_d$ with the help of Eq. (22), as done in section IV for Λ_d. However, we found that the more accurate value of $\bar{y}_d$ does not lead to appreciable changes in the predictions, in particular in Fig. 2^M a and b. Moreover, $\bar{y}_d$ still becomes independent of ξ^2 for large N_s. Therefore, we chose the approximation with the simpler formula, which is Eq. (33).

VI. RANGE OF s-DEPENDENCE FOR SMALL POPULATION SIZES

Equation $(7)^M$ specifies the the asymptotic ratio of fixation probabilities Δ for large populations and small s. Combined with the slope at $s = 0$ from Eq. $(6)^M$, this allows for an order-of-magnitude estimation of the range δs of (small) selective advantages where the fixation probability ratio actually depends on s. By assuming that the linear relationship of Eq. $(6)^M$ holds until p_c/p_d reaches $\Delta(f)$, we get

$$\delta s \sim \frac{(f-1)(1+\xi^2)}{fN_s \log(f)}. \tag{34}$$

While this only gives a rough estimate, it can be useful in determining whether mutations in a given range of selective advantages are in danger of unequal suppression by the experimental protocol, which might be undesirable for experimental evolution experiments. In accordance with Fig. 2^M b, Eq. (34) shows that δs is reduced for increasing population sizes.

VII. BRANCHING PROCESS APPROXIMATION FOR PERIODIC DILUTIONS

Conceptually, the calculation of p_d in the branching process limit simply requires substituting the rates of Eq. $(10)^M$ into Eq. $(8)^M$, taking the limit $\sigma \to 0$ and finally weighting the result appropriately for different θ. In practice, however, the calculation is rather non-trivial, so it is carried out explicitly in this section.
First, consider the integral

$$I = \int_0^\infty (\lambda + \delta)(t) \exp\left(- \int_0^t (\lambda - \delta)(\tau)\, \mathrm{d}\tau \right) \mathrm{d}t, \tag{35}$$

which can be split up into several integrals I_i, each from $t = t_i = i \cdot (\log f)/\alpha - \theta$ to $t = t_{i+1}$. We take I_0 to mean the first integral from $t = 0$ to $t = t_1$, which is the only integral that does not contain a period σ of pruning and has a length of $T - \theta$ instead of T. For $i = 0$,

the integrands therefore reduce to

$$(\lambda + \delta)(t) = \alpha(1+s) \tag{36}$$

$$-\int_0^t (\lambda - \delta)(\tau)\,\mathrm{d}\tau = -\alpha(1+s)t. \tag{37}$$

In contrast, for $i = 1, 2, \ldots$, they evaluate to

$$(\lambda + \delta)(t) = \alpha(1+s) + \log f \cdot \begin{cases} \frac{1}{\sigma} & \text{for } 0 < t - t_i < \sigma \\ 0 & \text{otherwise} \end{cases} \tag{38}$$

$$-\int_0^t (\lambda - \delta)(\tau)\,\mathrm{d}\tau = -\alpha(1+s)t + \log f \cdot \begin{cases} i - 1 + \frac{t-t_i}{\sigma} & \text{for } 0 < t - t_i < \sigma \\ i & \text{otherwise} \end{cases} \tag{39}$$

The integrals can then be calculated as

$$I_0 = 1 - \exp(\alpha(1+s)\theta) f^{-(1+s)} \tag{40}$$

$$I_i = f^{-is} \cdot \frac{f^{-(s+2)} \exp(\alpha(1+s)\theta)[(f^{1+s} - f)\alpha(1+s)\sigma + (f + f^{1+s} - 2\exp(-\alpha(1+s)\sigma) f^{2+s}) \log f]}{\alpha(1+s)\sigma - \log f}. \tag{41}$$

The only i-dependent term in I_i is f^{-is}, so $I = \sum_{i=0}^{\infty} I_i$ can be calculated using the geometric series. Using Eq. $(8)^M$, the fixation probability for finite σ is then given by

$$p_d(\theta, \sigma) = \frac{2}{1+I} = \left[1 + \frac{\exp(\alpha(1+s)\theta)(1 - \exp(-\alpha(1+s)\sigma) f) \log f}{(f^{1+s} - f)(\alpha(1+s)\sigma - \log f)}\right]^{-1} \tag{42}$$

We obtain the fixation probability for instantaneous selection and mutations occurring at time θ into a cycle by taking the limit $\sigma \to 0$:

$$p_d(\theta) = \left[1 + \frac{\exp(\alpha(1+s)\theta)(f-1)}{(f^{1+s} - f)}\right]^{-1} \tag{43}$$

With the probability that a mutation actually occurs at time θ into the cycle from Eq. (12), the unconditional fixation probability p_d therefore is

$$p_d = \frac{\int_0^T \exp(\alpha\theta) p_d(\theta)\,\mathrm{d}\theta}{\int_0^T \exp(\alpha\theta)\,\mathrm{d}\theta} = \frac{\alpha}{f-1} \int_0^T \exp(\alpha\theta) \left[1 + \frac{\exp(\alpha(1+s)\theta)(f-1)}{(f^{1+s} - f)}\right]^{-1} \mathrm{d}\theta, \tag{44}$$

where $T = (\log f)/\alpha$. The integrand is of the form $A^x(a + bB^x)^{-m}$, which has the indefinite integral

$$\int A^x (a + bB^x)^{-m}\,\mathrm{d}x = (\log A)^{-1} a^{-m} A^x {}_2F_1\left(m, \frac{\log A}{\log B}; \frac{\log A}{\log B} + 1; -\frac{bB^x}{a}\right) \tag{45}$$

In our case, $A = \exp(\alpha)$, $B = \exp(\alpha(1+s))$, $a = 1$, $b = \frac{f-1}{f^{1+s} - f}$, and $m = 1$, so the indefinite integral of the integrand in Eq. (44) is

$$\alpha^{-1} \exp(\alpha\theta) {}_2F_1\left(1, 1/(1+s), 1 + 1/(1+s), -\frac{f-1}{f^{1+s} - f} \exp(\alpha(1+s)\theta)\right).$$

Subtracting the values at the integral limits 0 and $T = (\log f)/\alpha$ then yields the result shown in Eq. $(11)^M$:

$$p_d = \frac{1}{f-1}\left[f\,{}_2F_1\left(1, \frac{1}{1+s}; 1+\frac{1}{1+s}; -\frac{f-1}{f^{1+s}-f}f^{1+s}\right) - {}_2F_1\left(1, \frac{1}{1+s}; 1+\frac{1}{1+s}; -\frac{f-1}{f^{1+s}-f}\right)\right] \tag{46}$$

VIII. RATIO OF FIXATION PROBABILITIES IN THE BRANCHING PROCESS APPROXIMATION

The fixation probability $p_c = s/(1+s)$ can easily be calculated directly from Eq. $(8)^M$ as described in the main text. However, it is worth noting that, as for the diffusion theory, p_d can be used for the constant population case by letting f tend to 1. In this case, it is particularly simple, since the limit $f \to 1$ also implies that $\theta \to 0$, because the length of growth phase tends to zero as $f \to 1$. Therefore, we can use the intermediate result of Eq. (43) to obtain

$$p_c = \lim_{f\to 1}\lim_{\theta\to 0} p_d(\theta) = \lim_{f\to 1}\left[1 + \frac{(f-1)}{(f^{1+s}-f)}\right]^{-1} = \left[1+\frac{1}{s}\right]^{-1} \tag{47}$$

$$= \frac{s}{1+s} \tag{48}$$

To determine the asymptotic ratio p_c/p_d for $s \to 0$ from the branching process approximation, it is actually easier to use again the intermediate result of Eq. (43) instead of carrying out the limit directly based on Eq. (46). As p_c does not depend on θ, we can first determine

$$\lim_{s\to 0}\frac{p_d(\theta)}{p_c} = \exp(-\alpha\theta)\frac{f\log f}{f-1}. \tag{49}$$

Integration with the probability $p_{\mathrm{mut}}(\theta)$ for the occurrence of a mutation at time θ into the cycle from Eq. (12) then leads to

$$\lim_{s\to 0}\frac{p_d}{p_c} = \int_0^T p_{\mathrm{mut}}(\theta)\lim_{s\to 0}\frac{p_d(\theta)}{p_c}\,\mathrm{d}\theta = \frac{\int_0^T \frac{f\log f}{f-1}\,\mathrm{d}\theta}{\int_0^T \exp(\alpha\theta)\,\mathrm{d}\theta} \tag{50}$$

$$= f\left(\frac{\log f}{f-1}\right)^2, \tag{51}$$

which is the inverse of the desired quantity Δ.

IX. COMPARISON WITH LIMITING CASES FOUND IN PREVIOUS STUDIES

To obtain analytical results, previous studies of the fixation probability in bottlenecked populations have focused mainly on the regime of large population sizes and small selective advantages [2, 3]. While our analytical approximations in the main text cover finite population sizes as well as larger selective advantages, they include the large population limit for small selective advantages, and we showed that both the diffusion approximation and the branching process approximation yield the same result in this particular limit, namely

Eq. (7)^M, according to which the fixation probability in this limit is reduced by a factor of

$$\frac{p_d}{p_c} = f\left(\frac{\log f}{f-1}\right)^2 \tag{52}$$

for serial passage with a dilution factor f. The result seems to contradict the fact that an optimal dilution factor was found in previous studies, because it converges monotonically to 1 for $f \to 1$.

First, in Ref. [2], the reduction factor for the fixation probability was calculated to be

$$D(\log D)^2 = (\log f)^2/f, \tag{53}$$

which has an optimum. This result is based on an earlier calculation by the same group according to which the fixation probability for a mutation occurring at time θ into a cycle for a cycle length T and wildtype growth rate α is (equation 12 in Ref. [4])

$$2s\frac{\alpha T}{\exp(\alpha\theta)}. \tag{54}$$

As we will show below, this approximation is only valid for large f (small D), in which case Eq. (52) and Eq. (53) indeed agree. Intuitively, the inapplicability of Eq. (54) for dilution factors close to 1 can be seen as follows: For a constant wild-type growth rate, we expect the fixation probability to converge to that of a constant population as $T \to 0$ (and $f \to 1$), because diluting very often by a factor close to 1 is experimentally indistinguishable from keeping the population constant. However, Eq. (54) converges to 0 for $T \to 0$. The reason for this discrepancy lies in authors' use of a binomial distribution to approximate the sampling process. The variance of the binomial distribution for selecting N_s individuals when the mutant ratio (i.e. the probability of selecting a mutant) is $y = n_Y/(n_X + n_Y)$ is

$$N_s y(1-y). \tag{55}$$

In reality, however, the sampling process is from a finite population, and so the hypergeometric distribution is appropriate. In this case, the variance for selecting N_s out of fN_s individuals when the mutant ratio is $y = n_Y/(n_X + n_Y)$ is

$$N_s y(1-y)\frac{fN_s - N_s}{fN_s - 1}, \tag{56}$$

which converges to Eq. (55) only for large f. For f closer to one, however, Eq. (55) greatly overestimates the variance, because Eq. (56) converges to zero in this limit. The gravity of this difference can be seen in the extreme case $f = 1$, when the whole population is selected for the next cycle. According to the binomial distribution, the number of mutants at the beginning of the next cycle would be a random quantity, whereas in reality (and according to the hypergeometric distribution), the number of mutants should be exactly the same as just before the dilution event. The overestimated stochasticity upon dilution leads to an underestimation of the fixation probability in Eq. (54) for f close to 1 (i.e. short T): $M_{\delta y}$ approaches 0 as $f \to 1$ (and therefore $T \to 0$), because there is no time for the fraction of mutants to change during the growth phase. If the binomial distribution is used, $V_{\delta y}$ stays finite according to Eq. (55) and therefore $M_{\delta y}/V_{\delta y} \to 0$ as $f \to 1$. In contrast, both $M_{\delta y}$ and $V_{\delta y}$ approach 0 as $f \to 1$ if the more realistic hypergeometric distribution is used, which leads to a consistent behavior of the fixation probability as $f \to 1$, as shown in the main text. Therefore, we conclude that Eq. (54) [4] and the factor (53) derived from it in Ref. [2] are only valid for large f and the optimum is an artifact stemming from the unphysical

behavior of the approximation towards $f = 1$.

The situation is different for the optimum found in Ref. [3]. The optimal dilution ratio for large populations and small s there is not found for the fixation *probability* (the quantity calculated in Ref. [2] and the present study), but for the *fixation rate per cycle*, derived as

$$\mu f N_s \frac{D(\log D)^2}{1-D} s = \mu N_s \frac{(\log f)^2}{1-1/f} s = \mu N_s \frac{f(\log f)^2}{f-1} s \tag{57}$$

in equation 12 of Ref. [3]. However, if the organism in question has a constant division rate and the only parameter that is varied experimentally is the dilution factor (and hence the dilution interval), then the fixation rate *per unit time* is a more appropriate measure of the adaptation rate of the population (the fixation rate per cycle might get smaller as the dilution factor approaches 1, but more cycles fit in a given time span). Using the authors' convention of a division rate equal to 1, the cycle length is $T = \log(f)$, and so Eq. (57) can be converted to a the *fixation rate per unit time* by dividing by T:

$$\mu N_s \frac{f \log f}{f-1} s \tag{58}$$

As expected, the above equation converges to $\mu N_s s$ for $f \to 1$, i.e. the fixation rate in a constant population of N_s individuals dividing with rate 1. Furthermore, it increases monotonically with f, such that there is no optimum, and therefore diluting the population more severely always leads to the loss of more beneficial mutations, if all other parameters are kept constant. In fact, our analysis yields the same fixation rate as in Eq. (58): The fixation rate per time is simply the number of mutations occurring per time multiplied by the fixation probability. The number of mutations per unit time for a growth rate of $\alpha = 1$ is $N_s(f-1)/\log(f)\mu$, and the fixation probability for large populations and small s is

$$f \left(\frac{\log f}{f-1} \right)^2 s \tag{59}$$

which is implied by Eq. $(7)^M$ or can be seen directly from Eq. $(4)^M$ by first taking the limit $N_s \to \infty$ and subsequently considering small s. Multiplying the two yields

$$N_s \mu f \frac{\log f}{f-1} s, \tag{60}$$

which is identical to Eq. (58). Our results are therefore consistent with Ref. [3]. However, the optimal dilution ratio found therein maximizes the fixation rate *per cycle*, which might be important for specific scenarios but is not the relevant quantity for the experimental situation we are considering: a population of cells dividing with fixed rate where only the dilution factor (and thus the dilution interval) can be chosen.

X. DENSITY-TRIGGERED DILUTION

In the main text, the population was pruned to N_s individuals periodically with period T. In this way, the population on average reaches a size of fN_s individuals before it is diluted, but the final size in individual cycles fluctuates due to the stochasticity of division times. We tested whether our results depend on the exact periodicity of the pruning and carried out numerical simulations where the population is reduced to N_s individuals once it reaches a fixed size fN_s. Now, instead of the final size, the period is a random quantity. Figures 1

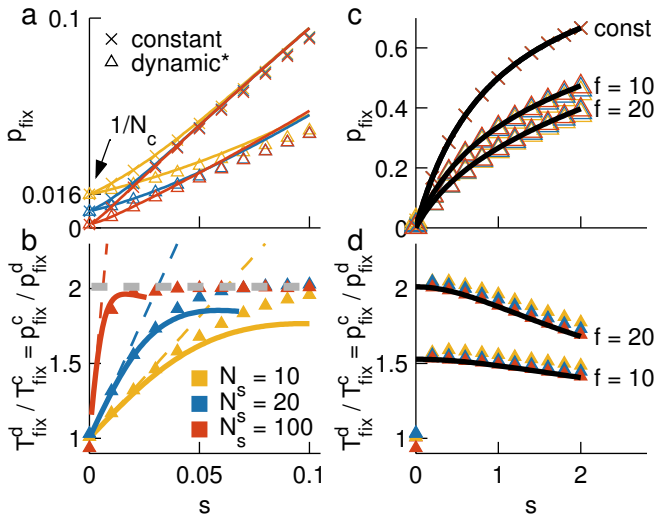

FIG. 1. Reproduction of Figure 2^M with numerical results for density-triggered instead of periodic dilutions. **(a)** Fixation probabilities from numerical simulations (symbols) compared to diffusion approximation, Eqs. $(5)^M$, $(4)^M$ (lines). **(b)** τ_d/τ_c from numerical simulations (symbols) and diffusion approximation (lines). Colored dashed lines: initial slope at $s = 0$ according to Eq. $(6)^M$; gray dashed line: asymptotic ratio, Eq. $(7)^M$. For all data in (a) and (b) $f = 20$, $\xi^2 = 1$, $m_s = 1$. **(c)**, **(d)** p and τ_d/τ_c from numerical simulations (symbols) compared to branching process approximation, Eqs. $(9)^M$, (46) (lines).

and 2 are identical to figures 2^M and 3^M, except that all data from numerical simulations was replaced with the alternative dynamic scenario, labeled "dynamic*". They show that the results are nearly identical, which might be important if the pruning is inherent to the cells, e.g., via an engineered density-triggered lysis mechanism.

[1] Y. Cao, D. T. Gillespie, and L. R. Petzold, The Journal of chemical physics **122**, 14116 (2005).
[2] L. M. Wahl, P. J. Gerrish, and I. Saika-Voivod, Genetics **162**, 961 (2002).
[3] J. E. Hubbarde and L. M. Wahl, Mathematical Biosciences **213**, 113 (2008).
[4] L. M. Wahl and P. J. Gerrish, Evolution **55**, 2606 (2001).

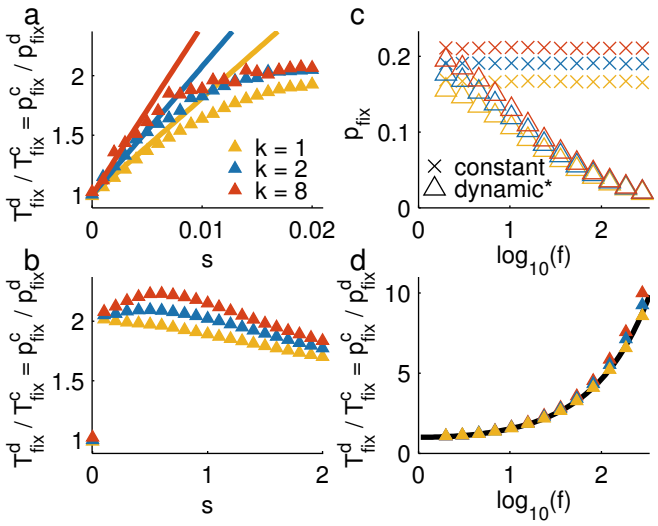

FIG. 2. Fixation probabilities and ratios in the multi-stage model. Reproduction of Figure 3^M with numerical results for density-triggered instead of periodic dilutions. **(a)** τ_d/τ_c for small s for different k in numerical simulations (symbols). Lines indicate the slope predicted by Eq. $(6)^M$ with $\xi^2 = \frac{2\log(2)^2}{k}$. **(b)** τ_d/τ_c from numerical simulations for larger s. **(c)** p as a function of the dilution factor f for a constant population and the two dynamic scenarios and different numbers of stages k. **(d)** τ_d/τ_c for the data shown in (c) compared to the analytical approximation, Eq. $(7)^M$. Parameters: $N_s = 50$, $f = 20$ (a,b) and $N_s = 20$, $s = 0.2$ (c,d).