

Rate of extinction of the fittest genotype in the slow ratchet regime. Stochastic fluctuations of the fittest class can approximately be described by a diffusion equation:

$$\partial_t p(x, t) = -\partial_x [D_1(x)p(x, t)] + \partial_x^2 [D_2(x)p(x, t)] \quad (1)$$

where $p(x)$ is the probability distribution of the nose occupancy, $x = N^{-1}n_{k_*}$ being defined as the fraction of the population in the "top bin" k_* . Diffusive "drift" is represented by $D_2(x) = x(1-x)/2N$ and effect of selection acting in the top bin is represented by $D_1(x) = \hat{s}x(1-x/x_*)$, where $x_* = e^{-U/s}$ is the equilibrium nose occupancy, where $\hat{s} = \gamma s$ and γ is a "phenomenological" parameter introduced by HAIGH (1978). Both $x = 1$ (fixation) and $x = 0$ (extinction) are absorbing boundary conditions (JAIN, 2008; STEPHAN and KIM, 2002). The density function for fixation/extinction time $\varphi(t; x)$ satisfies the backward Kolmogorov equation

$$\partial_t \varphi(t; x) = D_1(x)\partial_x \varphi(t; x) + D_2(x)\partial_x^2 \varphi(t; x) \quad (2)$$

where t is the time interval between initial state x and fixation or extinction. The mean time to fixation/extinction starting at $x = y$ at $t = 0$ is given by $\bar{t}(y) = \int_0^\infty t\varphi(t; y)dt$, which satisfies

$$-1 = D_1(y)\partial_y \bar{t}(y) + D_2(y)\partial_y^2 \bar{t}(y). \quad (3)$$

Using the integrating factor $\varphi(x) = e^{\int_0^x dz \frac{D_1(z)}{D_2(z)}}$, yields for the mean time to extinction of the fittest class:

$$\bar{t}(x_*) = \int_0^{x_*} dy \frac{1}{\varphi(y)} \int_y^1 d\zeta \frac{\varphi(\zeta)}{D_2(\zeta)}, \quad (4)$$

where

$$\varphi(x) = (1 - \zeta)^{\frac{2N\hat{s}(1-x_*)}{x_*}} e^{\frac{2N\hat{s}\zeta}{x_*}}, \quad (5)$$

The second integral in the limit $Ns \gg 1$ can be approximated as

$$2N \int_y^1 d\zeta \frac{(1 - \zeta)^{\frac{2N\hat{s}(1-x_*)}{x_*}} e^{\frac{2N\hat{s}\zeta}{x_*}}}{\zeta(1 - \zeta)} \approx 2N \int_y^1 d\zeta \frac{e^{\frac{2N\hat{s}x_*}{x_*} \left(\frac{2\zeta}{x_*} - \frac{\zeta^2}{x_*^2} \right)}}{\zeta} \quad (6)$$

and for small x , $\varphi(x)$ approximates to

$$\varphi(x) \approx e^{2N\hat{s}x_* \left(\frac{2\zeta}{x_*} - \frac{\zeta^2}{x_*^2} \right)} \quad (7)$$

Using the above approximations and $\alpha = N\hat{s}x_*$, $n_* = Nx_*$ and changing the variables, $\eta = y/x_* - 1$, $z = \zeta/x_* - 1$ yields:

$$\bar{t}(x_*) \approx 2n_{k_*} \int_{-1}^0 d\eta e^{\alpha\eta^2} \int_\eta^{1/x_*-1} dz \frac{e^{-\alpha z^2}}{1+z} \quad (8)$$

For $\alpha \gg 1$ the integral the second integral can be readily approximated yielding

$$\bar{t}(x_*) \approx 2n_{k_*} \sqrt{\frac{\pi}{4\alpha}} \int_{-1}^0 d\eta e^{\alpha\eta^2} [erf(\sqrt{\alpha}\beta) - erf(\sqrt{\alpha}\eta)] \quad (9)$$

where $\beta = \frac{1}{k_*} - 1$. Using $\eta^2 = 1 - \theta^2/\alpha$ allows the evaluation of the integral in the $\alpha \gg 1$ limit and yields

$$\bar{t}(x_*) \approx n_{k_*} \sqrt{\pi\alpha}^{-3/2} e^\alpha. \quad (10)$$

Assuming an exponential distribution for time of extinction of the fittest class, the rate of extinction is given by $r_- = 1/\bar{t}(x_*)$. Replacing $\alpha = \gamma sn_{k_*}$ yields:

$$r_- \approx e^{-\gamma sn_{k_*}} \gamma s \sqrt{\gamma sn_{k_*} / \pi} \quad (11)$$

which we use to arrive at Eq. (11) of the main text.

This standard calculation of extinction probability has glossed over the non-trivial "many-body" element of the problem: the fact that the strength of selection acting on the fittest genotype is defined relative to the population mean, which depends on the deviations from the steady state for all other, less-fit, genotypes. The latter in turn depends on the history of the dynamics in the top fitness "bin". As in all of the previous work, we deal with this problem here by introducing a phenomenological parameter γ , known as the "Haigh factor" (GORDO and CHARLESWORTH, 2000a,b; HAIGH, 1978; STEPHAN and KIM, 2002). However, unlike earlier work we shall not assume this factor to be a constant and determine its λ -dependence numerically (see discussion of the slow ratchet regime in the main text). An analytic method for describing the effect of history dependence on the fluctuations of the top bin goes beyond the scope of the present work and will be described in a separate publication (Neher and Shraiman, submitted).

Asymptotic expressions for ϵ_c The matching condition used to determine ϵ_c in the fast-ratchet regime can be rewritten as follows.

$$\frac{1}{\sigma \sqrt{\epsilon k_*}} = N e^{-\lambda(1-2\lambda) + \frac{k_*}{2} \log \frac{1-\epsilon}{\epsilon}} J_{k_*}(2\alpha). \quad (12)$$

In the limit of $\alpha \gg 1$, the zero of the Bessel function is approximately at $k_* + 1 \approx \alpha$ and the Bessel function at k_* evaluates to roughly $\sim k_*^{2/3}$ (Abramowitz 9.3.33 and 9.1.27). Hence we have

$$\lambda(1-2\epsilon) - \lambda \sqrt{\epsilon(1-\epsilon)} \log \frac{1-\epsilon}{\epsilon} = \log Ns \quad (13)$$

where we have neglected powers to the $\frac{1}{6}$ and $\mathcal{O}(1)$ factors inside the logarithm. In the limit $\epsilon \ll 1$ (but $\lambda^2 \epsilon \gg 1$), the matching simplifies to

$$1 - 2\epsilon + \sqrt{\epsilon} \log \epsilon = \lambda^{-1} \log Ns \quad (14)$$

which simplifies further to

$$\sqrt{\epsilon} \log \epsilon \approx \lambda^{-1} \log(Ns) - 1 = -2z \quad (15)$$

where we defined z for convenience. This can be solved for ϵ :

$$\epsilon_c = \frac{z^2}{W(-z)^2} \approx \frac{z^2}{(\log(z) - \log(-\log(z)))^2} \quad (16)$$

where $W(x)$ is the -1 branch of Lambert's W-function, i.e. the solutions of $W(x)e^{W(x)} = x$. The linear correction in Eq. 14 can be incorporated iteratively.

$$\epsilon_{i+1} = \frac{(z + \epsilon_i)^2}{W(-z + \epsilon_i)^2}. \quad (17)$$

This iteration converges for small C and ϵ . At larger C , the branch of the $W(x)$ function is lost. The result ϵ_2 obtained after the first iteration, starting with Eq. 16, is also shown in Fig. S1 as dashed red line.

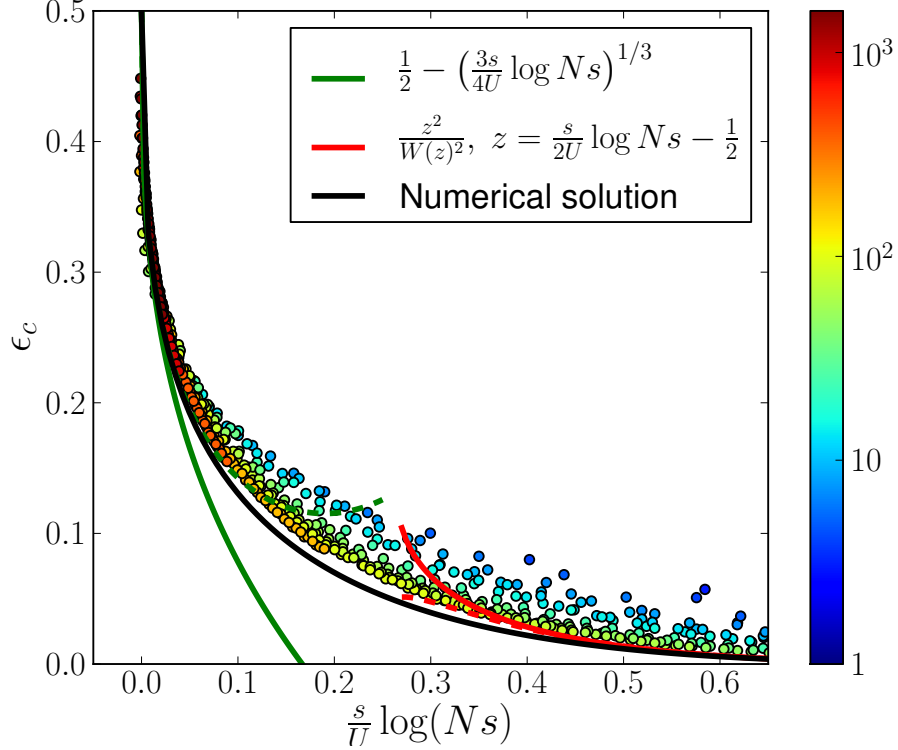


Figure S1. Dynamic mutation-selection balance. The figure shows ϵ_c for many combinations of U , s and N as a function of $\frac{s}{U} \log Ns$, while the color codes for $\epsilon_c U^2 s^{-2}$. If $\epsilon_c U^2 s^{-2} \gg 1$, ϵ_c is solely a function of $\frac{s}{U} \log Ns$ and is well described by the numerical solution of Eq. 6 in Main text, shown as a black line. The asymptotic approximations for large ϵ_c (Eq. 19) and small ϵ_c (Eq. 16) are shown as green and red lines, respectively. The dashed lines correspond to the more accurate version mentioned in the main text.

The other limit that is amenable to analytic calculations is the limit $\epsilon \rightarrow 1/2$. To this end, we define $\delta = \frac{1}{2} - \epsilon$ and expand the right hand side of Eq. 14

$$1 - 2\epsilon - \sqrt{\epsilon(1-\epsilon)} \log \frac{1-\epsilon}{\epsilon} = \frac{4}{3}\delta^3 + \frac{44}{15}\delta^5 + \mathcal{O}(\delta^7) = \lambda^{-1} \log Ns. \quad (18)$$

From this, we find

$$\epsilon \approx \frac{1}{2} - \left(\frac{3}{4\lambda} \log(Ns) \right)^{1/3}. \quad (19)$$

This expression is compared to simulation results in Fig. S1. The term proportional to δ^5 can again be included by iteration and the result is shown as the dashed green line in Fig. S1.

Convergence to the dynamic mutation-selection balance state. In the main text we discuss the possibility that ϵ itself depends on the absolute fitness. As the simplest model of this dependence here we consider the case of $\epsilon(k) = k/L$ which may be thought of as a "genome" of length L , that for

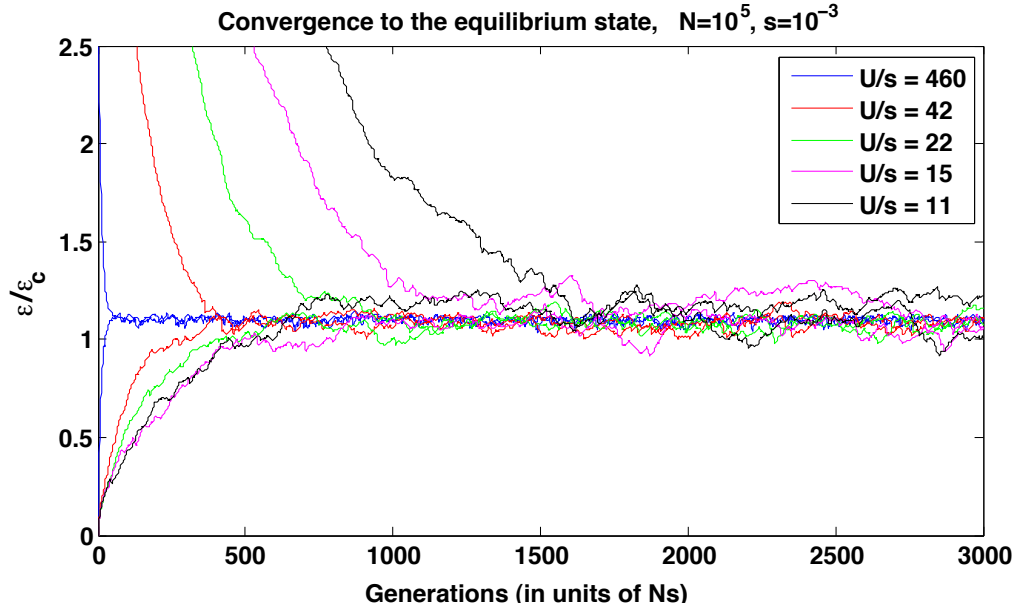


Figure S2. Evolution of the beneficial mutation fraction in the model with $\epsilon(k) = k/L$. The figure shows time evolution of ϵ towards e_c for populations which start above or below the absolute fitness (i.e. average number of mutations \bar{k} per genome) corresponding to the dynamic balance state which is realized when $\epsilon(k)$ at the nose of the distribution is at ϵ_c . The rate of convergence depends on the population parameters Ns, λ . Fluctuations around ϵ_c arise from the stochastic nature of the dynamic balance state.

$k = 0$ cannot be improved in fitness, but allows for a k/L fraction of compensatory mutations, once $k > 0$. Dependence of beneficial mutation fraction on absolute fitness is parameterized by L here set to 10^3 , so that differences in ϵ within any population are relatively small as $\lambda/L \ll 1$. Figure S2 presents the result of numerical simulations for this model, for different values of λ and different population sizes (see Methods section of the main text for the description of the numerical method). We ran simulations starting with populations in the high ($k = 0$) and low ($k = L$) fitness states and observed the dynamics of population averaged fitness \bar{k} decrease or increase with time eventually approaching a statistically steady state. This asymptotic state is characterized by average ϵ close to ϵ_c appropriate for the given population parameters. This convergence is evident in Figure S2 which displays the time dependence of population averaged ϵ normalized to $\epsilon_c(N, \lambda)$. Asymptotic value of the $\epsilon/\epsilon_c(N, \lambda)$ deviates slightly from unity because of the residual variation of ϵ within the population in the equilibrium state. The rate of evolution towards dynamic balance slows down as the population approaches its ϵ_c , which is expected as the $v \rightarrow 0$ as $\epsilon \rightarrow \epsilon_c$.

Mutational Meltdown. While the focus of our analysis has been on understanding the properties of the dynamic balance state, we have in the main text of the paper also discussed the global condition for its stability: the monotonic increase in the beneficial mutation rate (i.e. ϵ) with decreasing absolute fitness. That discussion however assumed constant population size, which may not be a good assumption once the absolute fitness falls below certain threshold. Declining absolute population fitness can lead to a reduction in population size, speeding Muller’s ratchet and leading to further declines in absolute fitness

and eventual extinction, a process is known as “mutational meltdown.” We now consider the conditions under which the dynamic equilibrium state allows a population to avoid this mutational meltdown.

To address this question, we must first specify the relationship between absolute fitness and changes in population size. This depends on the details of the population dynamics, and a variety of different models are plausible. Here we use the framework proposed by LYNCH *et al.* (1993), in which the average number of offspring of an individual is Rw , where R is the average fecundity of a mutation-free individual and w is its absolute fitness. In our model an individual with ℓ deleterious mutations has $w = (1 - s)^\ell$. The population then melts down when the average individual produces less than one offspring, which occurs when

$$R(1 - s)^{\bar{\ell}} < 1. \quad (20)$$

To connect this condition to our analysis of the dynamic equilibrium state, we next need to specify the relationship between ϵ and absolute fitness. This depends on the specific model of back and compensatory mutations. Here we assume that ϵ increases linearly with the number of deleterious mutations,

$$\epsilon = \frac{c\ell}{L}, \quad (21)$$

where c represents the average number of back or compensatory mutations available per deleterious mutation and L represents the total number of potentially deleterious sites.

Given these assumptions, our dynamic equilibrium state must satisfy

$$\epsilon_c < \frac{c}{L} \frac{\ln R}{\ln(1 - s)^{-1}} \approx \frac{c \ln R}{sL} \quad (22)$$

in order to avoid mutational meltdown. We can simply apply our analysis as described above to compute ϵ_c ; if the resulting value is greater than this critical value given the relevant values of R , s , and L , the dynamic equilibrium will not be stable due to reductions in population size, and meltdown will occur. We note however that since ϵ_c is bounded above by $\frac{1}{2}$, meltdown can only occur for sufficiently strong selection, $sL > 2 \ln R$. At the same time, stronger selection tends to reduce ϵ_c . Thus meltdown will typically be a concern only for a specific range of intermediate selection strengths.

The assumptions we made about population dynamics and the linear relationship between ϵ and absolute fitness are of course arbitrary. Recent work by SILANDER *et al.* (2007) has argued that in fact ϵ increases faster than linearly with ℓ . In this case, mutational meltdown will be less likely than we have suggested above (i.e. larger ϵ_c is required for meltdown). The alternative scenario where ϵ increases less than linearly with ℓ is also possible in principle, in which case mutational meltdown becomes more probable, though this scenario seems implausible given that back-mutation rates should increase linearly with ℓ .

References

- GORDO, I., and B. CHARLESWORTH, 2000a The degeneration of asexual haploid populations and the speed of muller’s ratchet. *Genetics* **154**: 1379–1387.
- GORDO, I., and B. CHARLESWORTH, 2000b On the speed of muller’s ratchet. *Genetics* **156**: 2137–2140.
- HAIGH, J., 1978 The accumulation of deleterious genes in a population. *Theor. Popul. Biol.* **14**: 251–26.
- JAIN, K., 2008 Loss of least-loaded class in asexual populations due to drift and epistasis. *Genetics* **179**(4): 2125–34.
- LYNCH, M., R. BURGER, D. BUTCHER, and W. GABRIEL, 1993 The mutational meltdown in asexual populations. *J. Hered.* **84**: 339–344.

SILANDER, O., O. TENAILLON, and L. CHAO, 2007 Understanding the evolutionary fate of finite populations: the dynamics of mutational effects. *PLoS Biol.* **5**: e94.

STEPHAN, W., and Y. KIM, 2002 *Recent Applications of Diffusion Theory to Population Genetics*. Oxford University Press, Oxford, UK, 72–93.