Mutational meltdown in asexual populations doomed to extinction

Peter Olofsson[1,2], Logan Chipkin[3], Ryan C. Daileda[2] and Ricardo B. R. Azevedo[3*]

[1]Department of Mathematics, Trinity University, San Antonio, Texas 78212, U.S.A.
[2]Department of Mathematics, Physics and Chemical Engineering, Jönköping University, 551 11, Jönköping, Sweden.
[3]Department of Biology & Biochemistry, University of Houston, Houston, Texas 77204, U.S.A.

*Corresponding author(s). E-mail(s): razevedo@uh.edu;

Abstract

Asexual populations are expected to accumulate deleterious mutations through a process known as Muller's ratchet. Lynch and colleagues proposed that the ratchet eventually results in a vicious cycle of mutation accumulation and population decline that drives populations to extinction. They called this phenomenon mutational meltdown. Here, we analyze mutational meltdown using a multi-type branching process model where, in the presence of mutation, populations are doomed to extinction. We analyse the change in size and composition of the population and the time of extinction under this model.

Keywords: Mutational meltdown, Muller's ratchet, lethal mutagenesis, extinction, branching processes, evolutionary rescue.

Statements and Declarations

- **Funding:** This work was supported by National Science Foundation grants DEB-1354952 and DEB-2014566 (awarded to R.B.R.A.) and National Institutes of Health grant R15GM093957 (awarded to P.O.).
- **Competing interests:** The authors have no competing interests to declare that are relevant to the content of this article.
- **Ethics approval:** Not applicable.
- **Consent to participate:** Not applicable.
- **Consent for publication:** Not applicable.
- **Availability of data and materials:** Not applicable.
- **Code availability:** Numerical calculations and stochastic simulations of the branching process models were done using software written in Python 3.7 and are available at https://github.com/rbazev/doomed.
- **Authors' contributions:** Not applicable.

1 Introduction

"All populations are doomed to eventual extinction." (Lynch and Gabriel, 1990)

In the absence of back mutations, an asexual individual cannot produce offspring carrying fewer deleterious mutations than itself. Indeed, it is always possible that individual offspring will accrue additional deleterious mutations. As a result, the class of individuals with the fewest deleterious mutations may, by chance, disappear irreversibly from the population, a process known as Muller's ratchet (Muller, 1964; Felsenstein, 1974; Haigh, 1978). Successive "clicks" of the ratchet will cause the fitness of asexual populations to decline. Muller's ratchet has been invoked to explain the evolution of sex (Muller, 1964; Felsenstein, 1974; Gordo and Campos, 2008), the extinction of small populations (Lynch and Gabriel, 1990; Lynch et al, 1993; Gabriel et al, 1993), the accelerated rate of evolution of endosymbiotic bacteria (Moran, 1996), the degeneration of Y-chromosomes (Charlesworth, 1978; Gordo and Charlesworth, 2000a), and cancer progression (McFarland et al, 2013, 2014).

Haigh (1978) argued that the ratchet should click at a rate inversely proportional to the size of the least loaded class in a population. If k is the lowest number of deleterious mutations present in an individual in the population, the size of the least loaded class at mutation-selection-drift equilibrium is

$$\hat{n}_k = Ne^{-U/s} \tag{1}$$

where N is the size of the population, U is the expected number of new deleterious mutations per genome per generation, and s is the deleterious effect of a mutation. Haigh argued that genetic drift causes the actual value of n_k to deviate stochastically from $\hat{n}_k$. The smaller the value of $\hat{n}_k$, the greater the probability that n_k will hit zero, causing the ratchet to click. If $\hat{n}_k > 1$,

then after a click of the ratchet, the size of the new least loaded class will go to a new equilibrium, $\hat{n}_{k+1}$, equal to $\hat{n}_k$ in equation (1). Haigh concluded that Muller's ratchet should click faster in small populations, experiencing a high deleterious mutation rate, and mutations with milder deleterious effects (low s). Subsequent work has derived more accurate estimates of the rate of clicking of the ratchet, both when $\hat{n}_k > 1$ (Stephan et al, 1993; Gordo and Charlesworth, 2000a,b; Jain, 2008; Neher and Shraiman, 2012; Metzger and Eule, 2013) and when $\hat{n}_k < 1$ (Gessler, 1995; Rouzine et al, 2003, 2008).

Beginning with Haigh's foundational study, most research on Muller's ratchet has assumed that the size of a population remains constant as deleterious mutations accumulate (e.g., Gessler, 1995; Gordo and Charlesworth, 2000a,b; Rouzine et al, 2003; Jain, 2008; Metzger and Eule, 2013). This assumption is biologically unrealistic—if true, fitness would decline continuously but the population would be immortal (Lynch and Gabriel, 1990; Melzer and Koeslag, 1991). Lynch and colleagues studied more realistic models where the fitness of an individual influences its fertility and populations experience density-dependent regulation (Lynch and Gabriel, 1990; Lynch et al, 1993; Gabriel et al, 1993; Lynch et al, 1995). They concluded that Muller's ratchet causes population size to decline, which accelerates the ratchet, which further reduces population size. This positive feedback results in a "mutational meltdown" that drives the population to extinction (Lynch and Gabriel, 1990; Lynch et al, 1993; Gabriel et al, 1993).

In one model, Lynch et al (1993) considered a population of asexual organisms subject to density-dependent regulation. Each individual produces R offspring. The number of mutations is Poisson distributed with rate U. The

offspring then undergo viability selection with a probability of survival

$$w_k = (1 - s)^k \tag{2}$$

where $k \geq 0$ is the number of deleterious mutations in the individual offspring, and $0 < s < 1$ is the deleterious effect of each mutation. If the number of offspring surviving viability selection N' exceeds the carrying capacity K then only K individuals survive and $N' - K$ individuals die, independently of their fitness; if $N' \leqslant K$, all N' individuals survive. Reproduction occurs after viability selection and density-dependent regulation. Assuming that initially all individuals in the population are mutation-free and that $NR > K$, Muller's ratchet proceeds in three phases in this model. First, mutations enter the population and accumulate rapidly. As the distribution of mutation numbers approaches mutation-selection-drift equilibrium mutation accumulation slows down. Second, the rate of mutation accumulation settles into a steady rate. This phase proceeds as in the classic constant population size model of Muller's ratchet (Haigh, 1978) and lasts while $NR\overline{w} > K$. Third, when mean viability falls below $\overline{w} = 1/R$ (i.e., when $NR\overline{w} = K$) the population size begins to decline, triggering mutational meltdown. During this phase the population is doomed to extinction.

Lynch et al (1993) derived some analytical expressions to describe the dynamics of mutation accumulation during the first two phases and the times at which these two phases end. However, they did not present any analytical theory on the dynamics or duration of the mutational meltdown phase itself (see also Gabriel et al, 1993; Lynch et al, 1995). Recently, Lansch-Justen et al (2022) derived analytical expressions to describe the mean number of deleterious mutations, population size, and extinction time during mutational

meltdown. Here we model mutational meltdown using a multi-type branching process. We analyse the change in size and composition of the population and the time of extinction under this model.

2 Mutational Parameters

We conducted a survey of the literature looking for empirical estimates of the deleterious mutation rate, u, and the mean deleterious effect of a mutation, s. The results are summarized in Figure 1. We adopted the median values of the mutational parameters ($u = 0.023$ and $s = 0.075$) as the default in this paper.

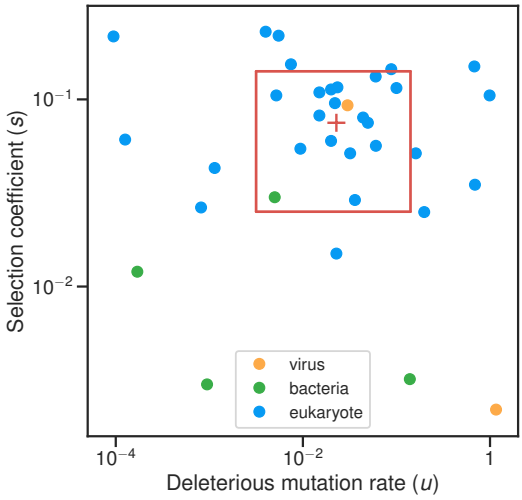

Fig. 1 Empirical estimates of mutational parameters. We obtained 35 estimates of the deleterious mutation rate (u) and mean deleterious effect of a mutation (s) from 25 studies on 16 species. When a single study included multiple estimates for a single species based on independent data sets, we used the median of those estimates. The red cross shows the median values: $u = 0.023$ and $s = 0.075$. Half of the estimates ($17/35 = 49\%$) are contained within the red box.

3 Model

3.1 Overview

We model a population of asexually reproducing organisms. Every generation, each individual has a probability u of acquiring a deleterious mutation. The mutation rate in our model is defined differently from that in standard models of Muller's ratchet (e.g., Haigh, 1978); however, the two formulations are similar if u is small. All mutations have the same deleterious effect (s), do not interact epistatically, and are irreversible.

Generations are non-overlapping. The absolute fitness of an individual with k deleterious mutations is defined by equation (2) and does not depend on the size of the population. We model reproduction and death using a branching process (see next section). Every generation, individuals reproduce or die independently of each other and experience hard selection (Wallace, 1975).

Initially, the population consists of N_0 individuals. The size of the population can change over time so the population can go extinct. Mutations doom the population to eventual extinction.

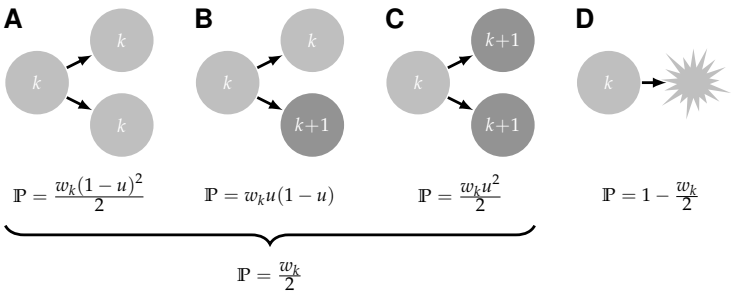

Fig. 2 Multi-type branching process. At each time step, an individual of type k—i.e, with k deleterious mutations (light gray)—can have one of four fates (A–D) with different probabilities, $\mathbb{P}$; w_k is the absolute fitness of an individual of type k defined in equation (2). It can either split into two daughters (A–C) or die (D). The daughters inherit the k mutations from their mother. A daughter can acquire one additional mutation and become an individual of type $k+1$ (dark gray) (B–C).

3.2 Branching process model

We model reproduction and death using a discrete-time, multi-type branching process (Haccou et al, 2005). We refer to individuals with k deleterious mutations as belonging to type k. Each generation, an individual of type k reproduces by splitting into two daughters with probability $\mathbb{P} = w_k/2$ (Fig. 2A–C) and dies with probability $\mathbb{P} = 1 - w_k/2$ (Fig. 2D), where w_k is the expected number of offspring of an individual with $k \geqslant 0$ deleterious mutations defined in equation (2) and $0 < s \leqslant 1$ is the deleterious effect of each mutation.

Individual offspring may independently acquire one deleterious mutation with probability u. The number of mutant offspring of a surviving individual of any type is, therefore, binomially distributed with parameters 2 and u (Fig. 2A–C).

As the mean number of offspring per individual is less than 1, the process is subcritical and will go extinct with probability 1 (Haccou et al, 2005).

3.3 Size and composition of the population

Initially (i.e., at generation $t = 0$), a population is described by the vector

$$\mathbf{Z}_0 = \begin{pmatrix} Z_{0,0} & Z_{0,1} & \cdots & Z_{0,g} \end{pmatrix} \tag{3}$$

where $Z_{0,k}$ is the number of individuals with k deleterious mutations and g is the greatest initial load (i.e., the maximum number of mutations carried by an individual; $Z_{0,k} = 0$ for $k > g$). The lowest initial load is ℓ (i.e., the minimum number of mutations carried by an individual; $Z_{0,k} = 0$ for $0 \leqslant k < \ell$).

At generation t, a population is described by the vector

$$\mathbf{Z}_t = \begin{pmatrix} Z_{t,0} & Z_{t,1} & \cdots & Z_{t,t+g} \end{pmatrix} \tag{4}$$

where $Z_{t,k}$ is the number of individuals with k deleterious mutations. Note that $Z_{t,k} = 0$ for $k > t + g$ because individuals can acquire no more than one mutation per generation.

In the following generation, the expected composition of the population is:

$$\mathbb{E}[\mathbf{Z}_{t+1}] = \mathbb{E}[\mathbf{Z}_t]\,\mathbf{M} \tag{5}$$

where $\mathbf{M}$ is the mean reproduction matrix with entries

$$\begin{cases} m_{k,k} &= w_k(1-u) \\ m_{k,k+1} &= w_k u \end{cases} \tag{6}$$

where $m_{i,j}$ is the expected number of offspring of type j generated by an individual of type i; all other entries of $\mathbf{M}$ are 0. The expected total number of offspring of an individual of type k is the absolute fitness $w_k = m_{k,k} + m_{k,k+1}$ defined in equation (2).

Iterating (5) we get

$$\mathbb{E}[\mathbf{Z}_t] = \mathbf{Z}_0\,\mathbf{M}^t \tag{7}$$

where $\mathbf{Z}_0$ is the (constant) initial state of the population (3). For any t, $\mathbf{M}^t$, the t-th power of $\mathbf{M}$, is upper triangular (i.e., all its entries below the diagonal are 0). Assuming fitness function (2), we can get an explicit form for the entries of $\mathbf{M}^t$, the expected number of descendants of an individual of type k carrying j *additional* mutations after t generations:

$$m^{(t)}_{k,k+j} = \begin{cases} u^j(1-u)^{t-j}(1-s)^{kt+\frac{j(j-1)}{2}} \displaystyle\prod_{i=1}^{j} \frac{1-(1-s)^{t+1-i}}{1-(1-s)^i} & , 0 \leqslant j \leqslant t \\ 0 & , \text{otherwise} \end{cases} \tag{8}$$

For a proof of equation (8) see Appendix A.

The elements of $\mathbb{E}[\mathbf{Z}_t]$ in equation (7) can be rewritten as

$$\begin{aligned}\mathbb{E}[Z_{t,k}] &= \sum_{i=0}^{g} Z_{0,i}\, m_{i,k}^{(t)} \\ &= (1-u)^t \sum_{i=0}^{g} Z_{0,i}\, (1-s)^{it}\, h_{t,k-i}\end{aligned} \tag{9}$$

where

$$h_{t,j} = \frac{u^j\,(1-s)^{\frac{j(j-1)}{2}}}{(1-u)^j} \prod_{i=1}^{j} \frac{1-(1-s)^{t+1-i}}{1-(1-s)^i}$$

As $t \to \infty$, we thus have $h_{t,j} \to h_j$ where

$$h_j = \frac{u^j\,(1-s)^{\frac{j(j-1)}{2}}}{(1-u)^j} \prod_{i=1}^{j} \frac{1}{1-(1-s)^i}$$

and equation (9) yields

$$\frac{\mathbb{E}[Z_{t,k}]}{(1-u)^t} \to Z_{0,0}\, h_k \tag{10}$$

as $t \to \infty$.

If we assume weak mutation (i.e., low u) then $(1-u)^t \approx e^{-ut}$, and we can write equation (10) as

$$\mathbb{E}[Z_{t,k}] \sim e^{-ut} Z_{0,0}\, h_k \tag{11}$$

Therefore, in the long run, the numbers of individuals of all types are expected to decline exponentially at a rate equal to the mutation rate, u.

The total size of the population at time t is defined as $N_t = Z_{t,0} + Z_{t,1} + Z_{t,2} + \ldots + Z_{t,t+g}$; equation (9) indicates that its expected value can be computed as

$$\mathbb{E}[N_t] = \sum_{k=0}^{g} \left(Z_{0,k} \sum_{j=0}^{t} m_{k,k+j}^{(t)} \right)$$

which turns out to be equivalent to

$$\mathbb{E}[N_t] = n_t \prod_{i=0}^{t-1} \left(1 - u + u(1-s)^i\right) \tag{12}$$

where

$$n_t = \sum_{k=0}^{g} Z_{0,k}(1-s)^{kt}$$

is the expected population size in the absence of mutation. For a proof of equation (12) see Appendix B.

Under weak mutation, equation (12) can be approximated asymptotically as $t \to \infty$ as:

$$\mathbb{E}[N_t] \sim e^{-ut} Z_{0,0}\, e^{u/s} \tag{13}$$

which is proved in Appendix C. Fig. 3A illustrates this approximation for default mutational parameters.

By equations (11) and (13), we thus get the ratio

$$\frac{\mathbb{E}[Z_{t,k}]}{\mathbb{E}[N_t]} \sim h_k\, e^{-u/s} \tag{14}$$

which can be viewed as a quasi-stable distribution over the classes $k = 0, 1, \ldots,$ as $t \to \infty$ ("quasi" because all classes are bound to die out eventually). In particular, as $h_0 = 1$, the mutation-free class has relative size $e^{-u/s}$, the same proportion as predicted under the Haigh model. Fig. 3B illustrates this approximation for default mutational parameters.

The computations above tacitly assume that initially the population includes some mutation-free individuals (i.e., $Z_{0,0} > 0$). However, if the least loaded type (i.e., the lowest number of mutations carried by an individual in the population) is not 0 but some number $\ell > 0$, the ratio in equation (14) still holds with h_k replaced by h_{k-l} for $k \geq l$. This is easily realized by replacing

$(1-u)^t$ by the product $(1-u)^t(1-s)^{lt}$ in the proof in Appendix C. Thus, if the ratchet clicks, the distribution will simply be displaced by one mutation (e.g., from ℓ to $\ell+1$). Haigh (1978) described a similar phenomenon in his model of Muller's ratchet under constant population size.

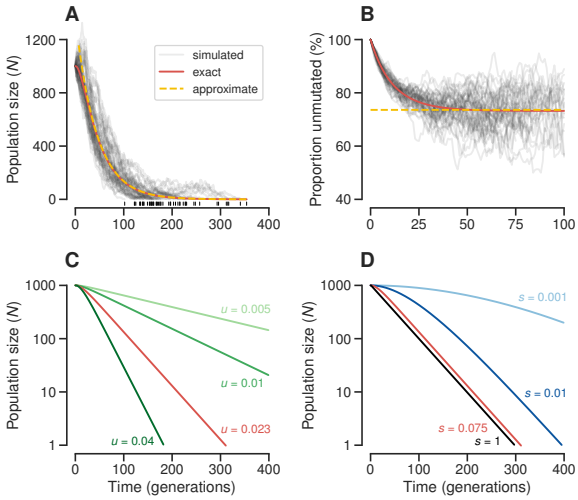

Fig. 3 Dynamics of total population size, N_t, of doomed populations. **(A)** Dynamics of N_t in 50 populations founded by $N_0 = Z_{0,0} = 1000$ mutation-free individuals and subject to mutations with deleterious effect $s = 0.075$ and rate $u = 0.023$. Gray lines show N_t of individual simulated populations. The red line shows the exact expected N_t from equation (12). The yellow dashed line shows the long-term weak mutation approximation from equation (13). Black vertical lines above the time axis indicate extinction times. **(B)** Dynamics of the proportion of mutation-free individuals ($Z_{t,0}/N_t$) in the simulations shown in (A). The red line shows the ratio of the exact expected values given in equations (9) and (12). The yellow dashed line shows the long-term weak mutation approximation from equation (14). **(C)** Effect of u on the dynamics of N_t ($N_0 = Z_{0,0} = 1000$, $s = 0.075$). **(D)** Effect of s on the dynamics of N_t ($N_0 = Z_{0,0} = 1000$, $u = 0.023$). Note that the N_t axes in (C) and (D) are log-transformed. The red lines in all plots correspond to the default mutational parameters ($u = 0.023$, $s = 0.075$).

3.4 Time to extinction

By well-known results from the theory of branching processes (Haccou et al, 2005), if the mean number of offspring per individual is strictly less than 1, the extinction time of the population has a finite expected value, $\mathbb{E}[T]$. By equation (2), an unmutated individual has $w_0 = 1$ offspring on average.

Equation (6), however, shows that $m_{k,k} < 1$ for any $k \geqslant 0$, that is, individuals of any type have on average less than one offspring of their own type, provided that mutations occur and are deleterious ($u > 0$ and $s > 0$). Therefore, the population has a finite expected extinction time.

An individual of type k may have offspring of types k or $k+1$ according to the joint probability generating function (p.g.f.)

$$\Phi_k(\mathbf{x}) = 1 - \frac{w_k}{2} + \frac{w_k}{2}\Big((1-u)^2 x_k^2 + 2u(1-u)x_k x_{k+1} + u^2 x_{k+1}^2\Big) \tag{15}$$

with $k = 0, 1, \ldots$ where $\mathbf{x} = (x_0, x_1, \ldots)$. Note that although $\mathbf{x}$ is of infinite length, each Φ_k depends on only finitely many coordinates. Denote the joint p.g.f. of $\mathbf{Z}_1$ by

$$\mathbf{\Phi}(\mathbf{x}) = \Big(\Phi_0(\mathbf{x})\ \Phi_1(\mathbf{x})\ \cdots\Big)$$

By standard branching process theory, the joint p.g.f. of $\mathbf{Z}_t$ is given by the t-fold composition of $\mathbf{\Phi}$ with itself

$$\mathbf{\Phi}^{(t)}(\mathbf{x}) = \mathbf{\Phi}\left(\mathbf{\Phi}^{(t-1)}(\mathbf{x})\right) \tag{16}$$

with $t = 1, 2, \ldots$, $\mathbf{\Phi}^{(0)}(\mathbf{x}) = \mathbf{x}$, and $\mathbf{\Phi}^{(1)}(\mathbf{x}) = \mathbf{\Phi}(\mathbf{x})$.

More standard theory yields that the probability that the population is extinct *by* generation t is

$$\mathbb{P}[N_t = 0] = \prod_{k=0}^{\infty}\left(\Phi_k\left(\mathbf{\Phi}^{(t-1)}(\mathbf{0})\right)\right)^{Z_{0,k}} \tag{17}$$

where $\mathbf{0} = (0, 0, 0, \ldots)$, and $\Phi_k(\mathbf{x})$ and $\mathbf{\Phi}^{(t-1)}(\mathbf{0})$ are computed by equations (15) and (16), respectively (Mode, 1971). The probability that the population goes extinct *in* generation t is $\mathbb{P}[N_t = 0] - \mathbb{P}[N_{t-1} = 0]$. Fig. 4 illustrates

the use of equation (17) to obtain the distribution of extinction times under default mutational parameters.

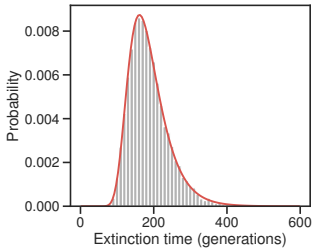

Fig. 4 Distribution of extinction times, T. Histogram based on 10^4 stochastic simulations with the same parameters as those used in Fig. 3A ($N_0 = Z_{0,0} = 1000$, $u = 0.023$, $s = 0.075$). The red line shows exact probabilities computed using equation (17).

The expected extinction time T of the population can be computed by the formula

$$\mathbb{E}[T] = \sum_{t=0}^{\infty} \mathbb{P}[T > t] \tag{18}$$

which holds for non-negative, integer-valued random variables. Moreover, as $\mathbb{P}[T > t] = \mathbb{P}[N_t > 0]$, the expression becomes

$$\begin{aligned} \mathbb{E}[T] &= \mathbb{P}[T > 0] + \sum_{t=1}^{\infty} \Big(1 - \mathbb{P}[N_t = 0]\Big) \\ &= 1 + \sum_{t=1}^{\infty} \Big(1 - \mathbb{P}[N_t = 0]\Big) \end{aligned} \tag{19}$$

where $\mathbb{P}(N_t = 0)$ is given by equation (17), and $\mathbb{P}[T > 0] = 1$ because $N_t > 0$ at time $t = 0$.

Finally, the variance of T is given by,

$$\mathrm{Var}[T] = \mathbb{E}[T(T-1)] + \mathbb{E}[T] - \mathbb{E}^2[T] \tag{20}$$

where

$$\mathbb{E}[T(T-1)] = 2\sum_{t=1}^{\infty} t\mathbb{P}(T > t)$$
$$= 2\sum_{t=1}^{\infty} t\Big(1 - \mathbb{P}[N_t = 0]\Big) \quad .$$

Fig. 5 uses equations (19) and (20) to compute the expected value and variance of extinction time T under a wide range of values of the mutational parameters. High values of the deleterious mutation rate, u, and selection coefficient of the mutations, s, both accelerate extinction (Fig. 5A). However, over the range explored u has a much larger effect on $\mathbb{E}[T]$ than s. For example, at the default selection coefficient $s = 0.075$, reducing the mutation rate by two orders of magnitude (from $u = 0.25$ to 0.0025) causes extinction times to increase 27-fold. In contrast, at the default mutation rate $u = 0.023$, reducing the selection coefficient by two orders of magnitude (from $s = 1$ to 0.01), causes extinction times to increase by only 41%. Low values of u and high values of s make extinction times more variable (Fig. 5B). Again, u has a stronger effect on the variability of T than s.

3.5 The first click of the ratchet

Let ℓ be the least loaded class at time $t = 0$ so that initially the population includes a fixed number $Z_{0,\ell} > 0$ of least loaded individuals (ancestors). Because of the irreversible nature of mutations in this model, once the least loaded class goes extinct, it can never reappear.

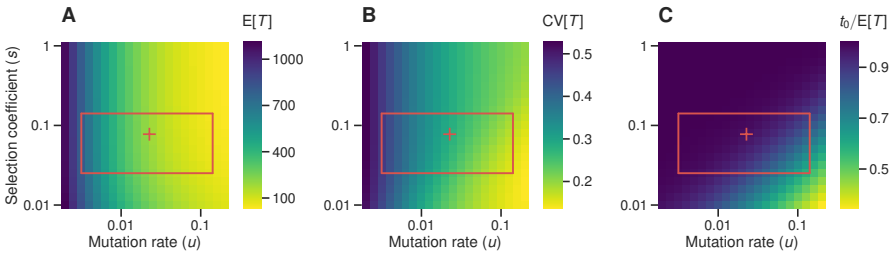

Fig. 5 High deleterious mutation rate accelerates extinction in doomed populations. **(A)** Expected extinction time, $\mathbb{E}[T]$, of populations founded by $N_0 = Z_{0,0} = 1000$ mutation-free individuals and subject to mutations with deleterious effect s and rate u. $\mathbb{E}[T]$ was calculated using equation (19) for $21 \times 21 = 441$ combinations of values of s and u evenly spaced on a log-scale spanning 2 orders of magnitude. **(B)** Coefficient of variation of extinction time, $\mathrm{CV}[T] = \sqrt{\mathrm{Var[T]}}/\mathbb{E}[T]$ for the populations shown in (A). $\mathrm{Var}[T]$ was calculated using equation (20). **(C)** Expected extinction time of the mutation-free class, t_0, as a proportion of $\mathbb{E}[T]$ for the populations shown in (A). t_0 was calculated using equation (21). The red box and cross are described in Fig. 1.

Let τ_ℓ denote the time of the first click of the ratchet during mutational meltdown (i.e., the extinction of type-ℓ individuals), that is

$$\tau_\ell = \min\{t \geqslant 0 : Z_{t,\ell} = 0\}$$

By the standard results from probability theory introduced in equations (18) and (19)

$$\begin{aligned} t_\ell = \mathbb{E}[\tau_\ell] &= \sum_{t=0}^{\infty} \mathbb{P}[\tau_\ell > t] \\ &= 1 + \sum_{t=1}^{\infty} \Big(1 - \mathbb{P}[Z_{t,\ell} = 0]\Big) \end{aligned} \tag{21}$$

where $\mathbb{P}[\tau_\ell > 0] = 1$ because least loaded individuals are present at time $t = 0$. Fig. 5C shows that the time of first click t_0 makes up a substantial proportion of the total extinction time except when u/s is large. For example, at the default mutational parameters, the expected time of first click is $t_0 = 187.2$, that is, 97% of the total time to extinction.

The time of extinction of the entire type-ℓ subpopulation is the time of extinction of the $Z_{0,\ell}$ independent subpopulations started from the ancestors. Thus, let $z_{t,\ell}^{(j)}$ denote the number of type-ℓ individuals in generation t stemming from the jth ancestor and $\tau_\ell^{(j)}$ be the time of extinction of the subpopulation started from the jth type-ℓ ancestor, $j = 1, 2, ..., Z_{0,\ell}$. We then have

$$Z_{t,\ell} = \sum_{j=1}^{Z_{0,\ell}} z_{t,\ell}^{(j)}$$

and

$$\tau_\ell = \max\left\{\tau_\ell^{(1)}, \tau_\ell^{(2)}, \ldots, \tau_\ell^{(Z_{0,\ell})}\right\}$$

and the equivalence

$$\tau_\ell^{(j)} \leqslant t \;\Leftrightarrow\; z_{t,\ell}^{(j)} = 0$$

An individual of type k has offspring of type k according to the the p.g.f.

$$\varphi_k(x) = 1 - \frac{w_k}{2} + \frac{w_k}{2}\Big((1-u)^2x^2 + 2u(1-u)x + u^2\Big) \tag{22}$$

Note that $\varphi_k(x)$ is obtained by applying $\Phi_k(\mathbf{x})$ to x_k in equation (15) and ignoring x_{k+1}. By another standard result (Jagers, 1994), the p.g.f. of $Z_{t,k}$ is given by the t-fold composition of φ_k with itself, denoted by $\varphi_k^{(t)}$.

For $t > 0$ we get the probability in equation (21)

$$\begin{aligned}\mathbb{P}[Z_{t,\ell} = 0] &= \prod_{j=1}^{Z_{0,\ell}} \mathbb{P}[z_{t,\ell}^{(j)} = 0] \\ &= \left(\varphi_\ell^{(t)}(0)\right)^{Z_{0,\ell}}\end{aligned}$$

Let $X_{\ell+1}$ be the number of individuals of type $\ell+1$ present after the first click of the ratchet (i.e., the time τ_ℓ of extinction of type ℓ). $X_{\ell+1}$ is a random

variable on $\{0, 1, 2, \dots\}$, noting in particular that $X_{\ell+1}$ may be 0, in which case both types ℓ and $\ell+1$ are extinct. The expected value of $X_{\ell+1}$ is given by

$$\begin{aligned} x_{\ell+1} = \mathbb{E}[X_{\ell+1}] &= Z_{0,\ell}\mathbb{E}\left[m_{\ell,\ell+1}^{(\tau_\ell)}\right] \\ &\approx Z_{0,\ell} m_{\ell,\ell+1}^{(t_\ell)} \end{aligned} \tag{23}$$

Note that equation (23) is a first-order Taylor approximation.

3.6 Large least loaded class

The expected time to extinction of the least loaded class, t_ℓ, is given by equation (21). Following Jagers et al (2007), there exists a sequence $c\left(Z_{0,\ell}\right) \to c$ as $Z_{0,\ell} \to \infty$ such that

$$t_\ell = -\frac{\ln Z_{0,\ell} + c\left(Z_{0,\ell}\right)}{\ln m_{\ell,\ell}} \tag{24}$$

where $m_{\ell,\ell} < 1$ is the expected number of least loaded offspring per least loaded individual (equation (6)) and $Z_{0,\ell}$ is the initial size of the least loaded class. Note that the value of c depends on the parameters.

Equation (24) shows that t_ℓ grows logarithmically with $Z_{0,\ell}$ with a slope of $-1/\ln m_{\ell,\ell}$. If mutation and selection are both weak, the slope becomes $\approx 1/(u + \ell s)$. Thus, increasing the initial size of the least loaded class delays its extinction more when mutation and selection are weak than when they are strong.

We now investigate the limiting behavior of $x_{\ell+1}$ as $Z_{0,\ell} \to \infty$. By equations (8), (23) and (24) we get

$$x_{\ell+1} \approx Z_{0,\ell} m_{\ell,\ell+1}^{(t_\ell)}$$

$$= \frac{C(Z_{0,\ell})u}{s(1-u)}\left(1-(1-s)^{t_\ell}\right)$$
$$\to \frac{Cu}{s(1-u)} \tag{25}$$

as $Z_{0,\ell} \to \infty$, where $C(Z_{0,\ell}) = e^{-c(Z_{0,\ell})}$ and $C = e^{-c}$. For a proof, see Appendix D. Interestingly, equation (25) shows that $x_{\ell+1}$ approaches a constant as $Z_{0,\ell}$ increases.

If the least loaded class is mutation-free ($\ell = 0$), the value of t_0 is not affected by the effects of mutations, s (equations (21) and (24)), because the rate at which individuals "leave" the mutation-free class is independent of s. The selection coefficient does, however, affect the size of the new least loaded class, x_1 (equations (23) and (25)), and therefore the total time to extinction.

4 Discussion

Most models of Muller's ratchet have assumed that populations maintain a constant size as deleterious mutations accumulate (e.g., Haigh, 1978; Gessler, 1995; Gordo and Charlesworth, 2000a,b; Rouzine et al, 2003; Jain, 2008; Metzger and Eule, 2013). This is typically justified as resulting from density-dependent regulation of population size. However, the assumption is unrealistic because it prevents populations from ever going extinct (Lynch and Gabriel, 1990; Melzer and Koeslag, 1991). In a series of studies relaxing the assumption of constant population size, Lynch and colleagues argued that Muller's ratchet eventually generates a positive feedback where the ratchet clicks, which causes population size to decline, which strengthens genetic drift relative to natural selection, which in turn accelerates the ratchet (Lynch and Gabriel, 1990; Lynch et al, 1993; Gabriel et al, 1993). They called this vicious cycle mutational meltdown and concluded that it drives populations to extinction.

Until recently, there was no quantitative theory of the mutational meltdown phase. We believe this theoretical neglect is unwarranted. Mutational meltdown offers a last chance for the population to be rescued by beneficial mutations and avoid extinction. Thus, the dynamics and duration of the mutational meltdown phase are expected to be important determinants of the probability of evolutionary rescue (Orr and Unckless, 2008; Martin et al, 2013; Orr and Unckless, 2014; Azevedo and Olofsson, 2021).

Lansch-Justen et al (2022) derived analytical expressions to describe the mutational meltdown phase in a model similar to that originally analyzed by Lynch et al (1993). Briefly, Lansch-Justen et al began by deriving the mean number of deleterious mutations in a population assuming that natural selection is no longer effective. They then used this number to derive the expected population size through time, N_t—analogous to our equation (12)—and then estimated time to extinction as the t at which N_t drops below 1. Lansch-Justen et al (2022) found that their approximation for extinction time matched simulation results well.

The approach described by Lansch-Justen et al (2022), however, does not work well in our model. For example, the (exact) expected time to extinction under default mutational parameters is $\mathbb{E}[T] = 192.8$ generations according to equation (19) (Fig. 4). Under the same parameters, the time required for the expected population size to drop below $\mathbb{E}[N_t] = 1$ is $t = 311$ generations according to equation (12), a time at which 96.6% of populations have gone extinct (Fig. 4). Furthermore, the expected population size at $\mathbb{E}[T]$ is over 15 individuals.

The extent to which real populations undergo mutational meltdown is unclear. A population must first enter the doomed regime. Models of Muller's ratchet in populations of constant size have identified three major risk factors

that can drive populations into the doomed regime: long-term reductions in population size, increases in mutation rate, and intermediate deleterious effects of mutations. Next, we consider each risk factor in turn.

Population size can decline as a result of changes in the environment, such as, climate change, decreased food availability, emergence of infectious diseases, and habitat loss or fragmentation. For example, the emergence of Devil Facial Tumor Disease, a transmissible cancer, has caused the size of the Tasmanian devil population to decline by ∼77% within 5 years (Hawkins et al, 2006; Lazenby et al, 2018). As a result, the devils are under risk of extinction (McCallum et al, 2009). Our results indicate that increasing population size causes relatively small delays in extinction in doomed populations. $E[T]$ is approximately proportional to the logarithm of initial population size (equation (24)). Similar results have been obtained in other stochastic models of population dynamics (Lande, 1993; Jagers et al, 2007).

Increased mutation rates can drive even very large populations into the doomed regime—a phenomenon known as lethal mutagenesis (Bull et al, 2007). Increases in mutation rate have been observed directly in experimental populations. For example, a population of *Escherichia coli* adapting to a constant environment evolved a mutator mutation after ∼25,000 generations that increased mutation rate by ∼150-fold (Barrick et al, 2009; Wielgoss et al, 2013). Evolution experiments have revealed that real populations can, indeed, experience increased extinction risk when the mutation rate is high. Zeyl et al (2001) allowed 12 populations of the yeast *Saccharomyces cerevisiae* with genetically elevated mutation rate to evolve and found that two of them went extinct within 2,900 generations. One of these populations went extinct shortly after a large decline in fitness. Bank et al (2016) subjected two populations of influenza A virus to gradually increasing concentrations of favipiravir,

a drug that increases the mutation rate of the virus, and observed that both populations accumulated mutations rapidly and went extinct. The results from both of these studies are broadly consistent with the occurrence of a mutational meltdown.

The results described in the previous paragraph indicate that mutational meltdown might have clinical applications. Mutagenic agents are being explored as antiviral drugs (Loeb et al, 1999; Crotty et al, 2001; Pariente et al, 2001; Bank et al, 2016). Increased mutation rate in tumor cells has been found to correlate with improved outcomes for some cancers (Silva et al, 2000; Birkbak et al, 2011; Andor et al, 2016). Several inhibitors of key components of the DNA-repair and DNA damage-response machinery (e.g., PARP inhibitors, Lord et al, 2015), are currently being used to treat cancer, or are under preclinical or clinical development (Brown et al, 2017).

Models of Muller's ratchet have shown that the rate at which the mean fitness of a population, $\bar{w}$, declines before entering the doomed phase is maximal at intermediate selection coefficients (Gabriel et al, 1993; Lynch et al, 1993; Gessler, 1995; Gordo and Charlesworth, 2000a,b; Lansch-Justen et al, 2022). For example, under the default mutation rate of $u = 0.023$, the decline in the $\bar{w}$ of a population of $N = 1000$ individuals is predicted to be fastest when deleterious mutations have selection coefficients of $s = 0.0063$ (decline in $\bar{w}$ per generation was measured as $s/\Delta t$, where Δt is the time between clicks of the ratchet calculated using the method of Gordo and Charlesworth, 2000a,b). This pattern arises because when deleterious mutations have small effects selection is less efficient at purging them so the ratchet clicks faster; however, because s is small those clicks cause $\bar{w}$ to decline slowly. Conversely, when deleterious mutations have large effects selection is efficient at purging them so the ratchet clicks slowly which also results in a slow decline of $\bar{w}$. Faster declines of $\bar{w}$ are achieved for intermediate values of s because selection

is inefficient enough at purging deleterious mutations that the ratchet clicks relatively quickly *and* each click has a significant impact on $\bar{w}$.

As we pointed out above, mutational meltdown offers a last chance for a population in the throes of Muller's ratchet or (Lynch and Gabriel, 1990; Lynch et al, 1993; Gabriel et al, 1993) or lethal mutagenesis (Bull et al, 2007; Matuszewski et al, 2017) to undergo evolutionary rescue (Orr and Unckless, 2008; Martin et al, 2013; Orr and Unckless, 2014; Azevedo and Olofsson, 2021). A standard model of evolutionary rescue (Orr and Unckless, 2008, 2014; Azevedo and Olofsson, 2021) considers a population that is somehow maladapted (perhaps due to sudden environmental deterioration) such that it is doomed to extinction. In the absence of beneficial mutations, this population can be modeled by a single type branching process. Models of evolutionary rescue typically ignore deleterious mutations. Interestingly, our results indicate that that simplifying assumption may be reasonable in the range of mutational parameters explored in this paper. Fig. 5A shows that for values of u and s spanning two orders of magnitude of values, the total extinction time is reasonably approximated by the case of $s = 1$, which in practive is a single type branching process.

Acknowledgments. We thank Alex Stewart, Ata Kalirad, Herbert Levine, and Erin Kelleher for helpful discussions and Richard Neher for comments on an earlier version of the manuscript. The National Science Foundation (grants DEB-1354952 and DEB-2014566 awarded to R.B.R.A.) and National Institutes of Health (grant R01GM101352 awarded to R.B.R.A. and grant R15GM093957 awarded to P.O.) funded this work.

Appendix A Proof of equation (8)

Given real numbers b and x and $n \in \mathbb{N}$ (the nonnegative integers), let $\mathbf{A}$ denote the "almost diagonal" $n \times n$ matrix

$$\mathbf{A} = \begin{pmatrix} 1 & b & & & \\ & x & bx & & \\ & & x^2 & bx^2 & \\ & & & \ddots & \ddots \\ & & & & x^{n-1} \end{pmatrix}$$

whose ith row is simply x^{i-1} multiplied by $\begin{pmatrix} 0 & 0 & \cdots & 0 & 1 & b & 0 & 0 & 0 & \cdots \end{pmatrix}$, the 1 occuring in the ith position. In equation (5), $\mathbf{M} = (1-u)\mathbf{A}$ with $x = 1-s$ and $b = \dfrac{u}{1-u}$.

Since $\mathbf{A}$ is upper triangular, so is its kth power (for $k \in \mathbb{N}$), with diagonal entries

$$a_{i,i}^{(k)} = x^{k(i-1)}$$

The superdiagonal entries are more complicated, but can also be expressed explicitly in terms of x, k and b.

Lemma 1 *For $n, k \in \mathbb{N}$, $1 \leqslant i \leqslant n-1$ and $1 \leqslant j \leqslant n-i$ one has*

$$a_{i,i+j}^{(k)} = b^j x^{\frac{j(j-1)}{2} + k(i-1)} \prod_{\ell=1}^{j} \frac{x^{k+1-\ell} - 1}{x^\ell - 1} \tag{A1}$$

provided that we declare $x^0 = 1$.

Proof We induct on k. When $k = 1$, for $j = 1$ and any i equation (A1) becomes

$$bx^{i-1} \frac{x-1}{x-1} = bx^{i-1} = a_{i,i+1}$$

When $j \geqslant 2$, then the $\ell = 2$ factor in the product in equation (A1) is $\frac{x^{2-2}-1}{x^2-1} = 0$, so that regardless of i the entire expression becomes 0, which again equals $a_{i,i+j}$. We conclude that equation (A1) holds for $k = 1$.

Now assume the result is true for some $k \geqslant 1$. Since $\mathbf{A}^{k+1} = \mathbf{A} \cdot \mathbf{A}^k$ and $a_{i,\ell} = 0$ unless $\ell \in \{i, i+1\}$,

$$\begin{aligned} a_{i,i+j}^{(k+1)} &= \sum_{\ell=1}^{n} a_{i,\ell} a_{\ell,i+j}^{(k)} \\ &= x^{i-1} a_{i,i+j}^{(k)} + b x^{i-1} a_{i+1,i+j}^{(k)} \\ &= x^{i-1} \left(a_{i,i+j}^{(k)} + b a_{i+1,(i+1)+(j-1)}^{(k)} \right) \end{aligned}$$

Using the inductive hypothesis[1] we obtain

$$\begin{aligned} a_{i,i+j}^{(k+1)} &= b^j x^{i-1+\frac{(j-1)(j-2)}{2}+ki} \left(\prod_{\ell=1}^{j-1} \frac{x^{k+1-\ell}-1}{x^\ell - 1} \right) \left(x^{j-k-1} \frac{x^{k+1-j}-1}{x^j-1} + 1 \right) \\ &= b^j x^{i-1+\frac{(j-1)(j-2)}{2}+ki} \left(\prod_{\ell=1}^{j-1} \frac{x^{k+1-\ell}-1}{x^\ell - 1} \right) \left(\frac{1 - x^{j-k-1} + x^j - 1}{x^j - 1} \right) \\ &= b^j x^{i-1+\frac{(j-1)(j-2)}{2}+ki+j-k-1} \left(\prod_{\ell=1}^{j-1} \frac{x^{k+1-\ell}-1}{x^\ell - 1} \right) \left(\frac{x^{k+1}-1}{x^j-1} \right) \\ &= b^j x^{\frac{j(j-1)}{2}+(k+1)(i-1)} \left(\prod_{\ell=1}^{j-1} \frac{x^{k+1-\ell}-1}{x^\ell - 1} \right) \left(\frac{x^{k+1}-1}{x^j-1} \right) \\ &= b^j x^{\frac{j(j-1)}{2}+(k+1)(i-1)} \left(\prod_{\ell=1}^{j} \frac{1}{x^\ell - 1} \right) \left(\prod_{\ell=0}^{j-1} (x^{k+1-\ell} - 1) \right) \\ &= b^j x^{\frac{j(j-1)}{2}+(k+1)(i-1)} \left(\prod_{\ell=1}^{j} \frac{x^{(k+1)+1-\ell}-1}{x^\ell - 1} \right) \end{aligned}$$

which shows that equation (A1) holds for the exponent $k+1$. This concludes the proof. □

[1] Strictly speaking, the inductive hypothesis will only apply to the term $a_{i+1,(i+1)+(j-1)}^{(k)}$ when $j \geqslant 2$. However, if we adopt the convention that any empty product is equal to one, the expression stated in the result agrees with $a_{i+1,(i+1)+(j-1)}^{(k)}$ when $j = 1$ as well.

Appendix B Proof of equation (12)

Equation (12) provides an exact calculation of the expected population size at generation t given the initial state $\mathbf{Z}_0$:

$$E[N_t] = n_t \prod_{i=1}^{t-1} \left(1 - u + u(1-s)^i\right) \tag{B1}$$

where

$$n_t = \sum_{k=0}^{\ell} Z_{0,k}(1-s)^{kt}$$

and ℓ is the maximum number of mutations carried by individuals present at time $t = 0$.

Proof Let $z_{t,k} = E[Z_{t,k}]$ when starting from one individual of type k, so that

$$E[N_t] = \sum_{k=0}^{\ell} Z_{0,k} z_{t,k}$$

We also have

$$z_{t,k} = \sum_{j=0}^{t} m_{k,k+j}^{(t)} \tag{B2}$$

where $m_{k,k+j}^{(t)}$ is defined in equation (8).

Equation (B1) will follow if we can prove that

$$z_{t,k} = (1-s)^{kt} \prod_{i=1}^{t-1} \left(1 - u + u(1-s)^i\right) \tag{B3}$$

We will use induction on t. From equation (B2) we get $z_{0,k} = m_{k,k}^{(0)} = 1$ and $z_{1,k} = m_{k,k}^{(1)} + m_{k,k+1}^{(1)} = (1-s)^k$, which agree with equation (B3). Now condition on the first generation. The expected number of offspring of the type-k ancestor is $(1-u)(1-s)^k$ of type k and $u(1-s)^k$ of type $k+1$ (equation (6)), and we get

$$z_{t,k} = (1-u)(1-s)^k z_{t-1,k} + u(1-s)^k z_{t-1,k+1}$$

As

$$m_{k+1,k+1+j}^{(t)} = (1-s)^t m_{k,k+j}^{(t)}$$

we get

$$z_{t-1,k+1} = (1-s)^{t-1} z_{t-1,k}$$

and hence the recurrence relation

$$z_{t,k} = z_{t-1,k}(1-s)^k \left(1 - u + u(1-s)^{t-1}\right) \tag{B4}$$

The induction hypothesis is that equation (B3) holds for $t-1$, whence it immediately follows that if holds also for t. □

Appendix C Proof of equation (13)

Equation (13) provides the following asymptotic formula under the assumption of weak mutation (i.e., small u):

$$\mathbb{E}[N_t] \sim e^{-ut} Z_{0,0}\, e^{\frac{u}{s}}$$

Proof We have

$$\mathbb{E}[N_t] = n_t \prod_{i=0}^{t-1} \left(1 - u + u(1-s)^i\right)$$

where

$$n_t = \sum_{k=0}^{\ell} Z_{0,k}(1-s)^{kt}$$

First note that because $0 < 1 - s < 1$, we get $n_t \to Z_{0,0}$ as $t \to \infty$. Next, let

$$Q_t = \prod_{i=0}^{t-1} \left(1 + \frac{u}{1-u}(1-s)^i\right)$$

so that $\mathbb{E}[N_t] = n_t(1-u)^t Q_t$. For small u, we use the approximation $1+x \approx e^x$ to get

$$\begin{aligned} Q_t &\approx \prod_{i=0}^{t-1} \exp\left(\frac{u}{1-u}(1-s)^i\right) \\ &= \exp\left(\frac{u}{1-u}\sum_{i=0}^{t-1}(1-s)^i\right) \\ &= \exp\left(\frac{u}{1-u}\cdot\frac{1-(1-s)^t}{s}\right) \\ &\to \exp\left(\frac{u}{(1-u)s}\right) \approx e^{u/s} \end{aligned}$$

proving equation (13). □

Appendix D Proof of equation (25)

We have

$$x_{k+1} \approx Z_{0,k}\, m_{k,k+1}^{(t_k)} \tag{C1}$$

where by equation (8)

$$\begin{aligned} m_{k,k+1}^{(t_k)} &= u(1-u)^{t_k-1}(1-s)^{kt_k} \cdot \frac{1-(1-s)^{t_k}}{s} \\ &= \Big((1-u)(1-s)^k\Big)^{t_k} \cdot \frac{u}{s(1-u)} \cdot \Big(1-(1-s)^{t_k}\Big) \\ &= \big(m_{k,k}\big)^{t_k} \cdot \frac{u}{s(1-u)} \cdot \Big(1-(1-s)^{t_k}\Big) \end{aligned}$$

By equation (24),

$$\begin{aligned} \big(m_{k,k}\big)^{t_k} &= \left(\big(m_{k,k}\big)^{-\frac{1}{\ln m_{k,k}}}\right)^{\ln Z_{0,k} + c(Z_{0,k})} \\ &= \frac{1}{Z_{0,k}} \cdot e^{-c(Z_{0,k})} \end{aligned}$$

noting that, for any $a > 0$,

$$a^{-\frac{1}{\ln a}} = e^{-1}$$

Thus, equation (C1) becomes

$$x_{k+1} \approx \frac{C(Z_{0,k})u}{s(1-u)}(1-(1-s)^{t_k})$$

$$\to \frac{Cu}{s(1-u)}$$

as $Z_{0,k} \to \infty$, which is equation (25).

References

Andor N, Graham TA, Jansen M, et al (2016) Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. Nat Med 22(1):105–113. https://doi.org/10.1038/nm.3984

Azevedo RBR, Olofsson P (2021) A branching process model of evolutionary rescue. Math Biosci 341:108,708. https://doi.org/10.1016/j.mbs.2021.108708

Bank C, Renzette N, Liu P, et al (2016) An experimental evaluation of drug-induced mutational meltdown as an antiviral treatment strategy. Evolution 70(11):2470–2484. https://doi.org/10.1111/evo.13041

Barrick JE, Yu DS, Yoon SH, et al (2009) Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. Nature 461(7268):1243–7. https://doi.org/10.1038/nature08480

Birkbak NJ, Eklund AC, Li Q, et al (2011) Paradoxical relationship between chromosomal instability and survival outcome in cancer. Cancer Res 71(10):3447–52. https://doi.org/10.1158/0008-5472.CAN-10-3667

Brown JS, O'Carrigan B, Jackson SP, et al (2017) Targeting DNA repair in cancer: beyond PARP inhibitors. Cancer Discov 7(1):20–37. https://doi.org/10.1158/2159-8290.CD-16-0860

Bull JJ, Sanjuán R, Wilke CO (2007) Theory of lethal mutagenesis for viruses. J Virol 81(6):2930–2939. https://doi.org/10.1128/JVI.01624-06

Charlesworth B (1978) Model for evolution of Y chromosomes and dosage compensation. Proc Natl Acad Sci U S A 75(11):5618–5622. https://doi.org/10.1073/pnas.75.11.5618

Crotty S, Cameron CE, Andino R (2001) Rna virus error catastrophe: direct molecular test by using ribavirin. Proc Natl Acad Sci U S A 98(12):6895–900. https://doi.org/10.1073/pnas.111085598

Felsenstein J (1974) The evolutionary advantage of recombination. Genetics 78(2):737–756. https://doi.org/10.1093/genetics/78.2.737

Gabriel W, Lynch M, Bürger R (1993) Muller's ratchet and mutational meltdowns. Evolution 47(6):1744–1757. https://doi.org/10.1111/j.1558-5646.1993.tb01266.x

Gessler DDG (1995) The constraints of finite size in asexual populations and the rate of the ratchet. Genet Res 66(03):241–253. https://doi.org/10.1017/S0016672300034686

Gordo I, Campos PRA (2008) Sex and deleterious mutations. Genetics 179(1):621–626. https://doi.org/10.1534/genetics.108.086637

Gordo I, Charlesworth B (2000a) The degeneration of asexual haploid populations and the speed of Muller's ratchet. Genetics 154(3):1379–1387. https://doi.org/10.1093/genetics/154.3.1379

Gordo I, Charlesworth B (2000b) On the speed of Muller's ratchet. Genetics 156(4):2137–2140. https://doi.org/10.1093/genetics/156.4.2137

Haccou P, Jagers P, Vatutin VA (2005) Branching Processes: Variation, Growth, and Extinction of Populations. Cambridge Studies in Adaptive Dynamics, Cambridge University Press, https://doi.org/10.1017/CBO9780511629136

Haigh J (1978) The accumulation of deleterious genes in a population—Muller's ratchet. Theor Popul Biol 14:251–267. https://doi.org/10.1016/0040-5809(78)90027-8

Hawkins C, Baars C, Hesterman H, et al (2006) Emerging disease and population decline of an island endemic, the Tasmanian devil *Sarcophilus harrisii*. Biol Conservat 131(2):307–324. https://doi.org/10.1016/j.biocon.2006.04.010

Jagers P (1994) Branching processes with biological applications. University Microfilms, Ann Arbor, MI

Jagers P, Klebaner FC, Sagitov S (2007) On the path to extinction. Proc Natl Acad Sci U S A 104(15):6107–6111. https://doi.org/10.1073/pnas.0610816104

Jain K (2008) Loss of least-loaded class in asexual populations due to drift and epistasis. Genetics 179(4):2125–34. https://doi.org/10.1534/genetics.108.089136

Lande R (1993) Risks of population extinction from demographic and environmental stochasticity and random catastrophes. Am Nat 142(6):911–927. https://doi.org/10.1086/285580

Lansch-Justen L, Cusseddu D, Schmitz MA, et al (2022) The extinction time under mutational meltdown driven by high mutation rates. Ecol Evol 12(7). https://doi.org/10.1002/ece3.9046

Lazenby BT, Tobler MW, Brown WE, et al (2018) Density trends and demographic signals uncover the long-term impact of transmissible cancer in Tasmanian devils. J Appl Ecol 55(3):1368–1379. https://doi.org/10.1111/

1365-2664.13088

Loeb LA, Essigmann JM, Kazazi F, et al (1999) Lethal mutagenesis of hiv with mutagenic nucleoside analogs. Proc Natl Acad Sci U S A 96(4):1492–7. https://doi.org/10.1073/pnas.96.4.1492

Lord CJ, Tutt AN, Ashworth A (2015) Synthetic lethality and cancer therapy: lessons learned from the development of PARP inhibitors. Annu Rev Med 66(1):455–470. https://doi.org/10.1146/annurev-med-050913-022545

Lynch M, Gabriel W (1990) Mutation load and the survival of small populations. Evolution 44(7):1725–1737. https://doi.org/10.2307/2409502

Lynch M, Bürger R, Butcher D, et al (1993) The mutational meltdown in asexual populations. J Hered 84(5):339–344. https://doi.org/10.1093/oxfordjournals.jhered.a111354

Lynch M, Conery J, Bürger R (1995) Mutation accumulation and the extinction of small populations. Am Nat 146(4):489–518. https://doi.org/10.1086/285812

Martin G, Aguilée R, Ramsayer J, et al (2013) The probability of evolutionary rescue: towards a quantitative comparison between theory and evolution experiments. Phil Trans R Soc B 368(1610):20120,088. https://doi.org/10.1098/rstb.2012.0088

Matuszewski S, Ormond L, Bank C, et al (2017) Two sides of the same coin: A population genetics perspective on lethal mutagenesis and mutational meltdown. Virus Evolution 3(1). https://doi.org/10.1093/ve/vex004, URL http://dx.doi.org/10.1093/ve/vex004

McCallum H, Jones M, Hawkins C, et al (2009) Transmission dynamics of Tasmanian devil facial tumor disease may lead to disease-induced extinction. Ecology 90(12):3379–3392. https://doi.org/10.1890/08-1763.1

McFarland CD, Korolev KS, Kryukov GV, et al (2013) Impact of deleterious passenger mutations on cancer progression. Proc Natl Acad Sci U S A 110(8):2910–2915. https://doi.org/10.1073/pnas.1213968110

McFarland CD, Mirny LA, Korolev KS (2014) Tug-of-war between driver and passenger mutations in cancer and other adaptive processes. Proc Natl Acad Sci U S A 111(42):15,138–15,143. https://doi.org/10.1073/pnas.1404341111

Melzer AL, Koeslag JH (1991) Mutations do not accumulate in asexual isolates capable of growth and extinction—Muller's ratchet re-examined. Evolution 45(3):649–655. https://doi.org/10.1111/j.1558-5646.1991.tb04335.x

Metzger JJ, Eule S (2013) Distribution of the fittest individuals and the rate of Muller's ratchet in a model with overlapping generations. PLoS Comput Biol 9(11):e1003,303. https://doi.org/10.1371/journal.pcbi.1003303

Mode CJ (1971) Multitype Branching Processes: Theory and Applications. Modern Analytic and Computational Methods in Science and Mathematics, Elsevier, Oxford, U.K.

Moran NA (1996) Accelerated evolution and Muller's rachet in endosymbiotic bacteria. Proc Natl Acad Sci U S A 93(7):2873–2878. https://doi.org/10.1073/pnas.93.7.2873

Muller HJ (1964) The relation of recombination to mutational advance. Mutat Res 1(1):2–9. https://doi.org/10.1016/0027-5107(64)90047-8

Neher RA, Shraiman BI (2012) Fluctuations of fitness distributions and the rate of Muller's ratchet. Genetics 191(4):1283–1293. https://doi.org/10.1534/genetics.112.141325

Orr HA, Unckless RL (2008) Population extinction and the genetics of adaptation. Am Nat 172(2):160–169. https://doi.org/10.1086/589460

Orr HA, Unckless RL (2014) The population genetics of evolutionary rescue. PLoS Genet 10(8):e1004,551. https://doi.org/10.1371/journal.pgen.1004551

Pariente N, Sierra S, Lowenstein PR, et al (2001) Efficient virus extinction by combinations of a mutagen and antiviral inhibitors. J Virol 75(20):9723–30. https://doi.org/10.1128/JVI.75.20.9723-9730.2001

Rouzine IM, Wakeley J, Coffin JM (2003) The solitary wave of asexual evolution. Proc Natl Acad Sci U S A 100(2):587–592. https://doi.org/10.1073/pnas.242719299

Rouzine IM, Brunet É, Wilke CO (2008) The traveling-wave approach to asexual evolution: Muller's ratchet and speed of adaptation. Theor Popul Biol 73(1):24–46. https://doi.org/10.1016/j.tpb.2007.10.004

Silva JM, Garcia JM, Dominguez G, et al (2000) DNA damage after chemotherapy correlates with tumor response and survival in small cell lung cancer patients. Mutat Res 456(1-2):65–71. https://doi.org/10.1016/S0027-5107(00)00113-5

Stephan W, Chao L, Smale JG (1993) The advance of Muller's ratchet in a haploid asexual population: approximate solutions based on diffusion theory. Genet Res 61(3):225–231. https://doi.org/10.1017/S0016672300031384

Wallace B (1975) Hard and soft selection revisited. Evolution 29(3):465–473. https://doi.org/10.1111/j.1558-5646.1975.tb00836.x

Wielgoss S, Barrick JE, Tenaillon O, et al (2013) Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. Proc Natl Acad Sci U S A 110(1):222–227. https://doi.org/10.1073/pnas.1219574110

Zeyl C, Mizesko M, De Visser JAGM (2001) Mutational meltdown in laboratory yeast populations. Evolution 55(5):909–917. https://doi.org/10.1111/j.0014-3820.2001.tb00608.x