A Branching Process Model of Evolutionary Rescue

Ricardo B. R. Azevedo* and Peter Olofsson†

*Department of Biology and Biochemistry, University of Houston, Houston, Texas, U.S.A.
†Department of Mathematics, Physics and Chemical Engineering, Jönköping University, Sweden.

June 4, 2021

Abstract

Evolutionary rescue is the process whereby a declining population may start growing again, thus avoiding extinction, via an increase in the frequency of beneficial genotypes. These genotypes may either already be present in the population in small numbers, or arise by mutation as the population declines. We present a simple two-type discrete-time branching process model and use it to obtain results such as the probability of rescue, the shape of the population growth curve of a rescued population, and the time until the first rescuing mutation occurs. Comparisons are made to existing results in the literature in cases where both the mutation rate and the selective advantage of the beneficial mutations are small.

Keywords: evolutionary rescue, adaptation, extinction, beneficial mutations, branching process, population dynamics

1 Introduction

A declining population may be saved from otherwise inevitable extinction by the establishment of one or more beneficial mutations, a phenomenon known as evolutionary rescue [3, 5, 9, 13]. Such rescue may be a desirable outcome, such as recovery of an endangered species, or an undesirable one, such as cancer becoming resistant to chemotherapy. A typical feature of evolutionary rescue is a U-shaped population growth curve corresponding to initial decline, stabilization, and ultimate recovery. This behavior has been established in mathematical models [9, 13, 19] and observed in experiments on real populations [1, 4, 24].

Orr and Unckless wrote two seminal papers on the mathematics of evolutionary rescue [18, 19]. They considered a population that has experienced sudden environmental deterioration such that it can no longer sustain itself. They then analyzed the possibility that this doomed population is able to adapt sufficiently rapidly to avoid extinction. They took into account the contributions of both standing genetic variation and new mutations to adaptation. Using branching processes methods in combination with Haldane's [10] classic approximation of the probability of fixation, Orr and Unckless obtained several highly interesting formulas for, among other things, the probability that a population is rescued by a new mutation and the expected size of a rescued population at a given time.

Crucially, Orr and Unckless assumed that selection is weak: the change in the environment reduces the mean absolute fitness of the population slightly below 1, and beneficial mutations raise it slightly above 1; or, in branching process parlance, they assumed that the doomed population is just barely subcritical, and the rescued population is just barely supercritical. They further assumed that mutation is weak relative to selection.

In the present work, we make no assumptions about the strengths of selection or mutation, and establish general results that coincide with those of Orr and Unckless when selection and mutation are weak, but can differ significantly in other scenarios. Strong selection is frequently observed in nature, specially as a result of human activity [22]. Furthermore, mutation rates are widely variable among organisms [17]. Thus, our results are of more than purely theoretical interest.

2 Model and Preliminaries

2.1 Branching Process

The population is modeled by a discrete-time branching process consisting of two types of individuals: wildtype individuals and individuals carrying a beneficial mutation. For simplicity, we refer to the latter metonymically as "beneficial individuals." The reproduction law is such that wildtype individuals can have offspring of both types, whereas beneficial individuals can only have offspring of their own type.

Denote by X the total number of offspring of a wildtype parent, where X is a nonnegative integer-valued random variable with mean $E[X] = 1 - r$, where r is the degree of maladaptation of a wildtype individual. The absolute fitness of a wildtype individual is, therefore, $1 - r$. We assume throughout that $0 < 1 - r < 1$, that is, $0 < r < 1$. Thus, a population composed entirely of wildtype individuals is subcritical and doomed to extinction[1]. The maladaptation is assumed to arise from an abrupt change in the environment [9, 13, 18, 19].

Each offspring of a wildtype individual may be beneficial with probability u (the beneficial mutation rate) or wildtype with probability $1 - u$, independently of other offspring. We denote the number of wildtype and beneficial offspring by W and B, respectively, so that we have $X = W + B$. Conditioned on X, W and B have binomial distributions with success probabilities $1 - u$ and u, respectively:

$$P(W = j, B = n - j \mid X = n) = \binom{n}{j}(1 - u)^j u^{n-j}$$

Beneficial individuals may exist in the population as standing genetic variation or arise *de novo* through mutation. We assume a multiplicative fitness model throughout, that is, beneficial individuals have absolute fitness $(1 - r)(1 + s)$, where s is the beneficial effect of the mutation they carry. Beneficial individuals cannot accumulate additional beneficial mutations. We assume that s is large enough to make $(1-r)(1+s) > 1$. Thus, the process of beneficial individuals is supercritical and has the potential to rescue the population from extinction.

[1] Many of the results in the paper are valid also in the critical case $1 - r = 1$, that is, $r = 0$.

Let $\varphi(t)$ be the probability generating function (pgf) of X

$$\varphi(t) = E\left[t^X\right] = \sum_{j=0}^{\infty} t^j P(X = j) \tag{1}$$

A complete specification of the pgf in Equation (1) requires that the distribution of X be known. To avoid uninteresting special cases, we make the natural assumption that $\varphi(1) = P(X < \infty) = 1$. Also, as $E[X] < 1$, we must have $\varphi(0) = P(X = 0) > 0$.

With W and B as above, let $P(j,k) = P(W = j, B = k)$ and $P(k) = P(B = k)$ be the offspring distributions for a wildtype and a beneficial parent, respectively, and define the pgfs

$$F(v,t) = \sum_{j,k\geq 0} v^j t^k P(W = j, B = k) \tag{2}$$

and

$$G(t) = \sum_{k\geq 0} t^k P(B = k) \tag{3}$$

Obviously $\varphi(t) = F(t,t)$ and for computations, the following result is of interest.

Lemma 2.1.

$$F(v,t) = \varphi\big((1-u)v + ut\big)$$

Proof.

$$\begin{aligned}
F(v,t) &= \sum_{i,j} v^i t^j P(W = i, B = j) \\
&= \sum_{n=0}^{\infty} \sum_{i=0}^{n} v^i t^{n-i} P(W = i, B = j \,|\, X = n) P(X = n) \\
&= \sum_{n=0}^{\infty} \sum_{i=0}^{n} v^i t^{n-i} \binom{n}{i} (1-u)^i u^{n-i} P(X = n) \\
&= \sum_{n=0}^{\infty} ((1-u)v + ut)^n P(X = n) \\
&= \varphi\big((1-u)v + ut\big)
\end{aligned}$$

by the binomial theorem. ∎

2.2 Extinction

Let W_n and B_n be the number of wildtype and beneficial individuals, respectively, in generation n. The total population size is $Z_n = W_n + B_n$. Let further $\mathcal{D}$ denote the event of extinction (death) of the population, that is

$$\mathcal{D} = \bigcup_{n=0}^{\infty} \{Z_n = 0\}$$

Denote by q_w and q_b the probabilities of extinction of the population when starting from one wildtype individual and one beneficial individual, respectively:

$$q_w = P(\mathcal{D} \,|\, W_0 = 1, B_0 = 0)$$

$$q_b = P(\mathcal{D} \,|\, W_0 = 0, B_0 = 1)$$

For convenience, we denote the corresponding survival probabilities by

$$\begin{aligned} p_w &= 1 - q_w \\ p_b &= 1 - q_b \end{aligned} \tag{4}$$

As the beneficial type is supercritical we have $q_b < 1$, and by standard branching process theory, q_b can be found as the smallest solution in $[0, 1]$ to the equation $t = G(t)$ (recalling that the process of beneficial individuals is a single-type process)

Next we must find the extinction probability q_w of a population that starts from one wildtype individual. As there is a chance of a beneficial mutation and beneficials have a chance of avoiding extinction, we realize that $q_w < 1$. It can further be computed explicitly according to the following result.

Proposition 2.2. *Let $\psi(t) = \varphi\big((1-u)t + uq_b\big)$. The extinction probability q_w is the unique solution in $[0, 1]$ to the equation $t = \psi(t)$.*

Proof. If $u = 0$, the equation reduces to $t = \varphi(t)$, the usual equation for the extinction probability, and as $1 - r < 1$, the unique solution is $t = 1$. Assume $u > 0$ and let $\mathcal{D}$ denote the event of extinction. Condition on the total number of offspring X and use the binomial theorem to obtain

$$
\begin{aligned}
q_w &= \sum_{n\geqslant 0} P(\mathcal{D}\,|\,X=n)P(X=n) \\
&= \sum_{n\geqslant 0}\sum_{j=0}^{n} P(\mathcal{D}\,|\,W=j, B=n-j)P(W=j, B=n-j\,|\,X=n)P(X=n) \\
&= \sum_{n\geqslant 0}\sum_{j=0}^{n} q_w^j q_b^{n-j}\binom{n}{j}(1-u)^j u^{n-j}(X=n) \\
&= \sum_{n=0}^{\infty}(q_w(1-u)+uq_b)^n P(X=n) \\
&= \varphi\left(q_w(1-u)+uq_b\right)
\end{aligned}
$$

Thus q_w solves the equation $t = \psi(t)$. We next demonstrate that there is exactly one solution in $[0,1]$ so the equation does indeed uniquely determine q_w. The claim follows if we can show that (a) ψ is strictly increasing and strictly convex, and has (b) $\psi(0) > 0$ and (c) $\psi(1) < 1$; the graph of $y = \psi(t)$ must then intersect the line $y = t$ exactly once in $[0,1]$. Here, (a) follows since $\psi'(t) > 0$ and $\psi''(t) > 0$ are readily seen by differentiation. For (b), note that

$$\psi(0) = \varphi(uq_b) > \varphi(0) > 0$$

and for (c),

$$\psi(1) = \varphi(1-u+uq_b) < \varphi(1) = 1$$

The proof is complete. ■

2.3 Rescue

If a population does not go extinct, we consider it rescued, and define the event $\mathcal{R} = \mathcal{D}^c$, so that the probability of rescue is simply one minus the probability of extinction: $P(\mathcal{R}) = 1 - P(\mathcal{D})$. Hence, a population starting from W_0 wildtype and B_0 beneficial individuals goes extinct if all the $W_0 + B_0$ subpopulations go extinct, and the probability of rescue is

$$P(\mathcal{R}) = 1 - q_w^{W_0} q_b^{B_0} \tag{5}$$

If initially the population is composed entirely of wildtype individuals ($B_0 = 0$), rescue can only occur through new mutations. The probability that this kind of rescue occurs is given by (see Equation (5))

$$P_{\text{new}}(\mathcal{R}) = 1 - q_w^{W_0} \tag{6}$$

Alternatively, rescue can occur from the B_0 beneficial ancestors. This scenario is modeled by assuming there is no mutation, that is, $u = 0$. In this case, $q_w = 1$ because the wildtype population is subcritical, and Equation (5) becomes

$$P_{\text{sv}}(\mathcal{R}) = 1 - q_b^{B_0} \tag{7}$$

The results from this section enable us to compute the probability of rescue for any offspring distribution, any degree of wildtype maladaptation (r), any selective benefit of a mutation (s), any beneficial mutation rate (u), and any initial composition (W_0, B_0). Figure 1 illustrates the effect of some of these parameters on the probability of rescue from new mutations: the P_{new} increases with increasing s (Figure 1A), decreasing r (Figure 1B), and increasing u (Figure 1C and 1D).

In the next section we will consider the case of weak selection/weak mutation, that is, when u, s, and r are all small. The formulas given by Orr and Unckless [18, 19] arise as approximations in our general framework.

2.4 Weak Selection/Weak Mutation Approximation

Assuming weak selection, that is, sufficiently small r and s, we can get approximations of Equations (6) and (7) that express the survival probabilities in terms of u, r, s, W_0, and B_0. We also assume that u is small, by orders of magnitude, compared to r and s (i.e., weak mutation).

In this case, the process of beneficials has mean $(1-r)(1+s) \approx 1 + (s-r)$ and Haldane's [10] approximation $p_b \approx 2(s-r)$ holds. Thus, in the standing variation case with $u = 0$, the probability of rescue in (7) becomes

$$\begin{aligned} P_{\text{sv}}(\mathcal{R}) &\approx 1 - (1 - 2(s-r))^{B_0} \\ &\approx 1 - \exp(-2B_0(s-r)) \end{aligned}$$

which is "P_{stand}" in Equation (2) in Orr and Unckless [19].

For the new mutation case, a standard first-order Taylor expansion of the natural logarithm of $\varphi(t)$ around $t = 1$ gives

$$\log \varphi(t) \approx \log \varphi(1) + \frac{\varphi'(1)}{\varphi(1)}(t-1)$$

$$= (1-r)(t-1) \tag{8}$$

because $\varphi'(1) = 1 - r$ and $\varphi(1) = 1$. (Note that for the Poisson offspring distribution, this Taylor approximation is exact.) Now recall

$$\psi(t) = \varphi\left((1-u)t + uq_b\right)$$

from Proposition 2.2. By Equation (8):

$$\log \psi(t) \approx (1-r)\left((1-u)t + uq_b - 1\right) \tag{9}$$

and the equation $t = \psi(t)$ becomes

$$\log t \approx (1-r)\left((1-u)t + uq_b - 1\right) \tag{10}$$

Using $q_b \approx 1 - 2(s-r)$ and the second-order Taylor expansion $\log t \approx t - 1$, around $t = 1$, Equation (10) becomes

$$t - 1 \approx (1-r)\left((1-u)t + u(1-2(s-r)) - 1\right)$$

the solution of which is

$$q_w \approx \frac{2r^2u + 2su - ru - 2rsu - u - r}{-r - u + ru}$$

which gives

$$\begin{aligned} p_w &= 1 - q_w \\ &\approx \frac{2ru - 2su - 2r^2u - 2rsu}{-r - u + ru} \\ &\approx \frac{2u(s-r)}{r} \end{aligned} \tag{11}$$

where the last approximation neglects the third-order terms $2r^2u$ and $2rsu$ in the numerator, and u and ru in the denominator. Equation (6) becomes

$$\begin{aligned} P_{\text{new}}(\mathcal{R}) &\approx 1 - \exp(-p_w W_0) \\ &\approx 1 - \exp\left(-\frac{2W_0 u(s-r)}{r}\right) \end{aligned} \tag{12}$$

which is "P_{new}" in Equation (2) in Orr and Unckless [18].

The weak selection/weak mutation approximation in Equation (12) performs well when the assumptions are met (Figure 1A, green arrow). However, it performs less well when selection is relatively strong (Figure 1A and 1B). In some scenarios, the approximation breaks down completely, even when the mutation rate is weak relative to selection (Figure 1D, high u).

3 Population Size

We next turn to the question of how the mean population size changes over time. Let the vector (W_n, B_n) be the number of wildtype and beneficial individuals in generation n, and consider its expected value, the vector $(E[W_n], E[B_n])$. In standard branching process notation, let m_{ij} denote the number of type-j offspring of a mother of type i. For simplicity, let the types be denoted by w and b, for wildtype and beneficial, respectively. The mean reproduction matrix is

$$M = \begin{pmatrix} m_{ww} & m_{wb} \\ 0 & m_{bb} \end{pmatrix} \tag{13}$$

In our specific model, recall $E[X] = 1 - r$, the total number of offspring of a wildtype individual to get $m_{ww} = (1-r)(1-u)$, $m_{wb} = (1-r)u$, and $m_{bb} = (1-r)(1+s)$ so that

$$M = (1-r)\begin{pmatrix} 1-u & u \\ 0 & 1+s \end{pmatrix}$$

and it is readily seen by induction that its nth power is

$$M^n = (1-r)^n \begin{pmatrix} (1-u)^n & u\dfrac{(1+s)^n - (1-u)^n}{s+u} \\ 0 & (1+s)^n \end{pmatrix} \tag{14}$$

which grows $\sim (1-r)^n(1+s)^n$ as $n \to \infty$, and the beneficials ultimately dominate the population. Specifically, as $n \to \infty$,

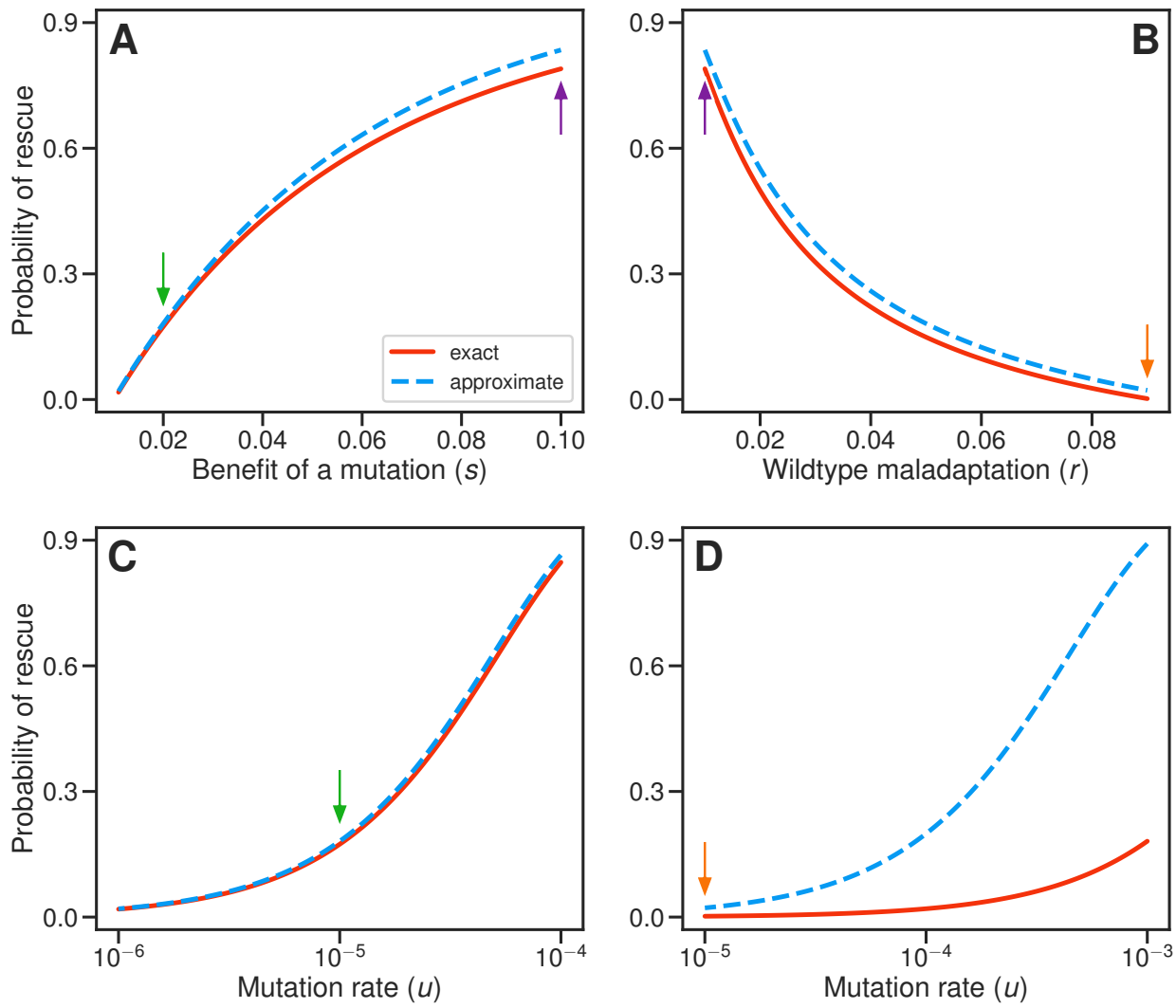

Figure 1: Effect of different parameters on the probability of rescue from new mutations, $P_{\text{new}}(\mathcal{R})$. In all cases we assumed an initial population size of $W_0 = 10^4$ wildtype individuals ($B_0 = 0$ beneficial individuals). (A) Selection coefficient of a beneficial mutation, s. Other parameters: $r = 0.01$ and $u = 10^{-5}$. (B) Degree of maladaptation of wildtype individuals, r. Other parameters: $s = 0.1$ and $u = 10^{-5}$. (C–D) Beneficial mutation rate, u. Other parameters: (C) $r = 0.01$ and $s = 0.02$; (D) $r = 0.09$ and $s = 0.1$. Exact probabilities (red, continuous) were calculated using Equation (6), assuming a Poisson offspring distribution. Approximate probabilities (blue, dashed) assuming weak selection/weak mutation were calculated using Equation (12). Arrows of particular colors indicate identical combinations of parameters.

$$\frac{1}{(1-r)^n(1+s)^n} M^n \to \begin{pmatrix} 0 & \dfrac{u}{s+u} \\ 0 & 1 \end{pmatrix}$$

Starting from fixed numbers (W_0, B_0), we have

$$E[(W_n, B_n)] = (W_0, B_0)\, M^n$$

The total population size is $Z_n = W_n + B_n$. Thus, the expected total population size, $E[Z_n] = E[W_n + B_n]$, starting from (W_0, B_0) ancestors is

$$E[Z_n] = W_0\,(1-r)^n \left(\frac{u}{s+u}(1+s)^n + \frac{s}{s+u}(1-u)^n \right) + B_0\,(1-r)^n (1+s)^n \tag{15}$$

and asymptotically

$$E[Z_n] \sim (1-r)^n (1+s)^n \left(W_0 \frac{u}{s+u} + B_0 \right)$$

Equation (15) takes into account all populations, including those that go extinct. In section 5.1 we derive the expected population size of rescued populations.

4 Mutations

4.1 Number of Beneficial Mutations

Let $\mathcal{A}$ be the event that at least one beneficial mutation has arisen in a population started from one wildtype individual, and let $\eta = P(\mathcal{A})$. As above, let φ be the pgf of the number of offspring X (Equation (1)).

Proposition 4.1. *The probability η is the unique solution to the equation*

$$t = 1 - \varphi\big((1-t)(1-u)\big)$$

Proof. The RHS function is increasing in t. By putting in $t = 0$ and $t = 1$, respectively, we get the endpoint values $1 - \varphi(1-u)$ and $1 - \varphi(0)$, both of which are in $(0,1)$, and hence the RHS function intersects the line $y = t$ exactly once in $[0,1]$. Next, condition on the number of offspring X of the ancestor, and the number Y of those offspring that are mutants:

$$\begin{aligned}
\eta = P(\mathcal{A}) &= \sum_{k=1}^{\infty} P(\mathcal{A} \,|\, X = k) P(X = k) \\
&= \sum_{k=1}^{\infty} \sum_{j=0}^{k} P(\mathcal{A} \,|\, X = k, Y = j) P(Y = j \,|\, X = k) P(X = k) \\
&= \sum_{k=1}^{\infty} \sum_{j=0}^{k} P(\mathcal{A} \,|\, X = k, Y = j) \binom{k}{j} u^j (1-u)^{k-j} P(X = k)
\end{aligned}$$

the sum starting at $k = 1$ because $P(\mathcal{A} \mid X = 0) = 0$. Next, note that we have $P(\mathcal{A} \mid X = k, Y = j) = 1$ for $j \geqslant 1$, and for $j = 0$:

$$P(\mathcal{A} \mid X = k, Y = 0) = 1 - (1 - \eta)^k$$

The term with $j = 0$ equals

$$\sum_{k=1}^{\infty} \left(1 - (1 - \eta)^k\right) (1 - u)^k P(X = k)$$

and the rest of the sum is

$$\sum_{k=1}^{\infty} \sum_{j=1}^{k} \binom{k}{j} u^j (1 - u)^{k-j} P(X = k) = \sum_{k=1}^{\infty} \left(1 - (1 - u)^k\right) P(X = k)$$

Add these together to get $1 - \varphi\left((1 - \eta)(1 - u)\right)$ and the proof is complete. ■

Note that, as there has to be at least one beneficial mutation for there to be rescue, we obviously have $\eta > p_w$. Also note that with $v = 1 - t$, the proposition may be rewritten as $v = \varphi(v(1 - u))$ where, by Lemma 2.1, the RHS equals the pgf of W integrated over the event that $B = 0$:

$$\begin{aligned} \varphi(v(1 - u)) &= F(v, 0) \\ &= \sum_{j=0}^{\infty} v^j P(W = j, B = 0) \\ &= E\left[v^W; B = 0\right] \end{aligned}$$

Next let K be the total number of mutations ever arising in a population, noting that $P(K > 0) = \eta$ above. Its expected value is

Proposition 4.2.

$$E[K] = W_0 \frac{uE[X]}{1 - (1 - u)E[X]}$$

Proof. First let $W_0 = 1$ and let $\nu = E[K]$. We first rule out the possibility that $\nu = \infty$. To this end, let W be the total number of individuals ever born in a branching process with mean $E[X] = 1 - r < 1$. Then

$$\nu < E[W] = E\left[\sum_{n=0}^{\infty} W_n\right] = \sum_{n=0}^{\infty}(1-r)^n = \frac{1}{r} < \infty$$

Next, condition on the number X of offspring of the ancestor, and the number Y of those that are mutants. Noting that $E[K \mid X = 0] = 0$ we get

$$\begin{aligned}
\nu = E[K] &= \sum_{k=1}^{\infty}\sum_{j=0}^{k} E[K \mid X = k, Y = j]P(Y = j \mid X = k)P(X = k) \\
&= \sum_{k=1}^{\infty}\sum_{j=0}^{k}(j + (k-j)\nu)\binom{k}{j}u^j(1-u)^{k-j}P(X = k) \\
&= (1-\nu)\sum_{k=1}^{\infty}\sum_{j=0}^{k} j\binom{k}{j}u^j(1-u)^{k-j}P(X = k) \\
&\quad + \nu\sum_{k=1}^{\infty} k\sum_{j=0}^{k}\binom{k}{j}u^j(1-u)^{k-j}P(X = k) \\
&= (1-\nu)uE[X] + \nu E[X]
\end{aligned}$$

where we used that, for $j \geqslant 0$,

$$j\binom{k}{j} = k\binom{k-1}{j-1}$$

Solving $\nu = (1-\nu)uE[X] + \nu E[X]$ and multiplying by W_0 concludes the proof. ■

Note that the proposition also follows from the informal argument that the expected number of beneficial offspring of the ancestor is $uE[X]$ and the populations starting from the remaining $(1-u)E[X]$ offspring produce on average ν new mutations each. Hence $\nu = uE[X] + (1-u)E[X]\nu$, which also gives the result.

For example, with the parameters used in Figures 2A and 2B ($W_0 = 10^4$, $r = 0.01$, and $u = 10^{-5}$), the expected number of independent beneficial mutations occurring during the course of evolution is $E[K] = 9.9$. The corresponding number for Figures 2C and 2D ($W_0 = 10^4$, $r = 0.01$, and $u = 10^{-4}$) is $E[K] = 98.0$. These numbers of mutations include mutations that occur in populations that ultimately go extinct and mutations that occur in rescued populations but disappear and do not contribute to the rescue.

By Equation (4), a given mutation is successful in rescuing the population with probability p_b. With K_S denoting the number of successful mutations, we thus get

$$E[K_S] = W_0 \frac{p_b u E[X]}{1-(1-u)E[X]} \tag{16}$$

Equation (16) takes into account all populations, including those that go extinct ($K_S = 0$). In section 5.3 we derive the $E[K_S]$ of rescued populations.

4.2 Waiting Time for a Beneficial Mutation

We next turn our attention to the time (generation) T of the first appearance of a beneficial mutation, aiming for the probabilities $P(T > n)$, $n = 0, 1, 2, ...$, in a population starting from wildtype individuals only. From these probabilities we can obtain the distribution function

$$F(n) = 1 - P(T > n)$$

and the probability mass function

$$p(n) = F(n) - F(n-1)$$

Note that $P(T = \infty) > 0$, because the population may die out without experiencing any mutations (specifically, $P(T = \infty) = 1 - \eta$, by Proposition 4.1). Thus, it also follows that $E[T] = \infty$.

We will establish an expression for $P(T > n)$. To that end, first recall the pgf $\varphi(t)$ of the number of offspring X of a wildtype parent (Equation (1)). Next, define a sequence of functions $H_0, H_1, \ldots$ recursively through

$$H_{k+1}(t) = \varphi(t\, H_k(t)) \tag{17}$$

where $H_0(t) \equiv 1$. Thus $H_1(t) = \varphi(t), H_2(t) = \varphi(t\,\varphi(t))$, and so on. Consider a population that starts from W_0 wildtype individuals and no beneficial individuals, $B_0 = 0$.

Proposition 4.3.

$$P(T > n) = \big(H_n(1-u)\big)^{W_0}$$

Proof. As $P(T > n \,|\, W_0) = (P(T > n \,|\, W_0 = 1))^{W_0}$, we can assume $W_0 = 1$. Recall W_n and B_n, the number of wildtype and beneficial individuals in generation n, respectively. Having $T > n$ means that there are no mutation events at or before the nth generation, that is

$$P(T > n) = P(B_1 = 0, B_2 = 0, \ldots, B_n = 0)$$

For any i, j, k, let

$$p(i, j) = P(W_{k+1} + B_{k+1} = j \,|\, W_k = i, B_k = 0)$$

and note that

$$P(W_{k+1} = j, B_{k+1} = 0 \,|\, W_k = i, B_k = 0) = (1-u)^j p(i,j)$$

and

$$\sum_{j=0}^{\infty} t^j p(i,j) = \big(\varphi(t)\big)^i$$

Repeated use of the above identities yields

$$\begin{aligned}
P(T > n) &= P(B_1 = 0, B_2 = 0, \ldots, B_n = 0) \\
&= \sum_{k_1,k_2,\ldots,k_n} P(W_1 = k_1, B_1 = 0, W_2 = k_2, B_2 = 0, \ldots, W_n = k_n, B_n = 0) \\
&= \sum_{k_1,k_2,\ldots,k_n} (1-u)^{k_1} p(1,k_1) \cdots (1-u)^{k_{n-2}} p(k_{n-2},k_{n-1}) (1-u)^{k_n} p(k_{n-1},k_n) \\
&= \sum_{k_1,k_2,\ldots,k_{n-1}} (1-u)^{k_1} p(1,k_1) \cdots (1-u)^{k_{n-1}} p(k_{n-2},k_{n-1}) \big(\varphi(1-u)\big)^{k_{n-1}} \\
&= \sum_{k_1,k_2,\ldots,k_{n-2}} (1-u)^{k_1} p(1,k_1) \cdots \big((1-u)\varphi(1-u)\big)^{k_{n-2}} \\
&= \cdots = H_n(1-u)
\end{aligned}$$

■

Note that, by the definition of H_0, we have $P(T > 0) = 1$. Also, Proposition 4.1 and Proposition 4.3 with $W_0 = 1$ yield

$$H_n(1-u) \to 1 - \eta$$

which is easily corroborated numerically.

Recall that any given beneficial mutation is successful in rescuing the population with probability p_b, and let T_S be the time until such a successful mutation. The probabilities $P(T_S > n)$ can be found by applying Proposition 4.3 to the function $\widetilde{H}_n$, defined recursively through Equation (17) but instead using the pgf

$$\widetilde{\varphi}(t) = \varphi\big((1-u+up_b)t + uq_b\big)$$

With B_S and B_F denoting the numbers of successful and failed mutants, respectively, $\tilde{\varphi}$ is the pgf of $W + B_S$ (recall Lemma 2.1). The total number of offspring is $X = W + B_S + B_F$, and conditioned on X, the three types of offspring follow a multinomial distribution with probabilities $1 - u, up_b$, and uq_b. Thus, the conditional distribution of B_S given $W + B_S$ is binomial with success probability

$$\widetilde{u} = \frac{up_b}{1 - u + up_b}$$

and we get

Corollary 4.3.1.

$$P(T_S > n) = \left(\widetilde{H}_n(1 - \widetilde{u})\right)^{W_0}$$

5 Rescued Populations

5.1 Population Size

For populations that survive, the mean population size (Equation (15)) is a poor measure as it takes into account all the populations that go extinct. For example, with an extinction probability of 99%, the 1% of populations that survive will have sizes far above the mean population size. Thus, we consider instead the mean population size conditioned on rescue. The analysis hinges upon use of the elementary result that for any integrable random variable Y, we have

$$E[Y \mid \mathcal{R}] = \frac{E[Y] - E[Y \mid \mathcal{D}]P(\mathcal{D})}{P(\mathcal{R})} \tag{18}$$

by the Law of Total Expectation.

By standard branching process results [2, 15], a supercritical branching process conditioned on extinction is a subcritical branching process, whose reproduction law can be given explicitly. To that end, recall the notation $P(j,k) = P(W = j, B = k)$ for a wildtype parent and $P(k) = P(B = k)$ for a beneficial parent, and the corresponding pgfs $F(v,t)$ and $G(t)$ (Equations (2) and (3)). Recall also the extinction probabilities q_w and q_b, starting from one wildtype and one beneficial individual, respectively. Following Jagers and Lagerås [15], the process conditioned on extinction has reproduction law given by

$$\widehat{P}(j,k) = q_w^{j-1} q_b^k P(j,k) \tag{19}$$

$$\widehat{P}(k) = q_b^{k-1} P(k)$$

with the corresponding pgfs

$$\widehat{F}(v,t) = \frac{1}{q_w} F(q_w v, q_b t)$$

$$\widehat{G}(t) = \frac{1}{q_b} G(q_b t)$$

By standard properties of pgfs, we can compute the entries of the mean reproduction matrix $\widehat{M}$, in analogy with Equation (13), as

$$\widehat{m}_{ww} = \widehat{F}'_v(1,1) = F'_v(q_w, q_b)$$

$$\widehat{m}_{wb} = \widehat{F}'_t(1,1) = \frac{q_b}{q_w} F'_t(q_w, q_b)$$

$$\widehat{m}_{bb} = \widehat{G}'(1) = G'(q_b)$$

so that

$$\widehat{M} = \begin{pmatrix} F'_v(q_w, q_b) & \frac{q_b}{q_w} F'_t(q_w, q_b) \\ 0 & G'(q_b) \end{pmatrix}$$

the nth power of which is

$$\widehat{M}^n = \begin{pmatrix} F'_v(q_w, q_b)^n & \frac{q_b}{q_w} F'_t(q_w, q_b) \frac{G'(q_b)^n - F'_v(q_w, q_b)^n}{G'(q_b) - F'_v(q_w, q_b)} \\ 0 & G'(q_b)^n \end{pmatrix} \tag{20}$$

From Equations (18) and (20), we can now compute the nth generation mean population sizes conditioned on rescue, starting from W_0 wildtype and B_0 beneficials:

Proposition 5.1.

$$E[(W_n, B_n) \,|\, \mathcal{R}] = (W_0, B_0) \, \frac{M^n - q_w^{W_0} q_b^{B_0} \widehat{M}^n}{1 - q_w^{W_0} q_b^{B_0}}$$

The total population size is $Z_n = W_n + B_n$ and we get

$$E[Z_n \,|\, \mathcal{R}] = (W_0, B_0) \, \frac{M^n - q_w^{W_0} q_b^{B_0} \widehat{M}^n}{1 - q_w^{W_0} q_b^{B_0}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{21}$$

Figure 2 shows an example of the application of Equation (21).

5.2 Weak Selection/Weak Mutation Approximation

We again consider the case of weak selection/weak mutation where u, r, and s are all small. We first consider the standing variation case where rescue is due to pre-existing beneficial individuals and not to new mutations. Let the population start from (W_0, B_0) individuals where B_0 is small; to ensure no mutations we set $u = 0$. We then have

$$E[Z_n|\mathcal{D}] = W_0(1-r)^n + E[B_n\,|\,\mathcal{R}]$$

where, by Equation (3)

$$E[B_n\,|\,\mathcal{D}] = B_0 G'(q_b)^n \tag{22}$$

As

$$E[B_n] = B_0 m_{bb}^n$$

Equations (18) and (22) yield

$$E[B_n\,|\,\mathcal{R}] = B_0 \frac{(1-r)^n(1+s)^n - q_b^{B_0} G'(q_b)^n}{P_{\mathrm{sv}}(\mathcal{R})}$$

If q_b is close to 1 and B_0 small, we get the approximation

$$P_{\mathrm{sv}}(\mathcal{R}) = 1 - (1-p_b)^{B_0} \approx p_b B_0$$

which yields

$$E[B_n\,|\,\mathcal{R}] \approx \frac{m_{bb}^n - q_b^{B_0} G'(q_b)^n}{p_b}$$

With Haldane's approximation $p_b \approx 2(s-r)$, we get

$$E[Z_n\,|\,\mathcal{R}] \approx W_0(1-r)^n + \frac{(1-r)^n(1+s)^n - (1-2(s-r))^{B_0}\left(G'(1-2(s-r))\right)^n}{2(s-r)}$$

which is essentially Equation (10) in Orr and Unckless [19], noting that they use the approximation $(1-r)^n(1+s)^n \approx e^{(s-r)n}$, and neglect the term $E[B_n\,|\,\mathcal{D}]P(\mathcal{D})$ in Equation (18).

For the new mutation case, the population is rescued by mutation and not by pre-existing beneficial individuals. We thus let $B_0 = 0$ and start from W_0 wildtype individuals. Again, neglect the term $E[B_n \,|\, \mathcal{D}]P(\mathcal{D})$ in Equation (18) to obtain

$$E[Z_n \,|\, \mathcal{R}] \approx W_0(1-r)^n(1-u)^n + \frac{E[B_n]}{P_{\text{new}}(\mathcal{R})}$$

and Equation (14) gives, for the second term

$$\frac{E[B_n]}{P_{\text{new}}(\mathcal{R})} = \frac{W_0 u(1-r)^n(1+s)^n - (1-u)^n}{(s+u)P_{\text{new}}(\mathcal{R})}$$

where

$$P_{\text{new}}(\mathcal{R}) = 1 - (1-p_w)^{W_0} \approx p_w W_0$$

Neglecting the term $(1-u)^n$ and using Equation (11) to approximate p_w, we now get

$$\begin{aligned}\frac{E[B_n]}{P_{\text{new}}(\mathcal{R})} &\approx \frac{rW_0 u(1-r)^n(1+s)^n}{(s+u)2u(s-r)W_0} \\ &\approx \frac{r}{s}\frac{(1-r)^n(1+s)^n}{2(s-r)}\end{aligned}$$

And, therefore,

$$E[Z_n \,|\, \mathcal{R}] \approx W_0(1-r)^n(1-u)^n + \frac{r}{s}\frac{(1-r)^n(1+s)^n}{2(s-r)} \tag{23}$$

in agreement with Equation (19) in Orr and Unckless [19].

Figure 2 shows an example of the application of Equation (23). The weak selection/weak mutation approximation can perform reasonably well when selection is strong (Figure 2B), but can break down even when the mutation rate is weak relative to selection (Figure 2D). Paradoxically, the weak selection/weak mutation approximation tends to underestimate population size (Figure 2) while overestimating the probability of rescue (Figure 1).

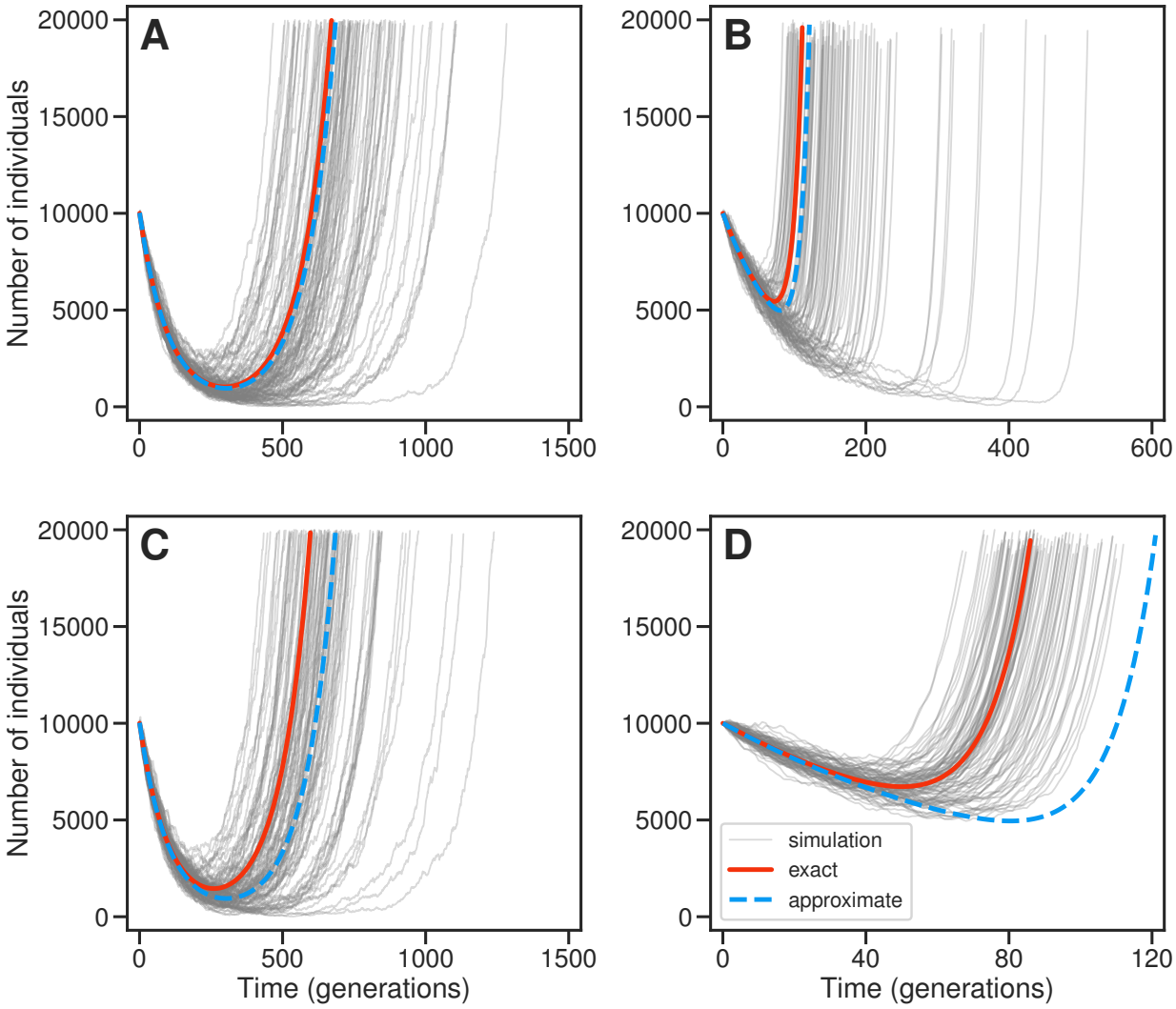

Figure 2: Total population size of rescued populations. We considered evolutionary rescue based on new mutations. In all cases we assumed an initial population size of $W_0 = 10^4$ wildtype individuals ($B_0 = 0$ beneficial individuals), a degree of wildtype maladaptation of $r = 0.01$, and a Poisson offspring distribution. (A) Weak selection/weak mutation: $s = 0.02$ and $u = 10^{-5}$ (green arrow, Figure 1A). (B) Stronger selection: $s = 0.1$ and $u = 10^{-5}$ (purple arrow, Figure 1A). (C) Stronger mutation: $s = 0.02$ and $u = 10^{-4}$ (far right, Figure 1C). (D) Stronger selection and stronger mutation: $s = 0.1$ and $u = 10^{-4}$. Gray lines show total sizes (Z_n) of individual simulated populations experiencing rescue (defined as reaching $Z_n > 2 \times 10^4$). A sample of 100 individual trajectories is shown for each parameter combination. Exact population sizes (red, continuous) were calculated using Equation (21). Approximate population sizes (blue, dashed) assuming weak selection/weak mutation were calculated using Equation (23)

5.3 Number of Rescuing Mutations

In Equation (16), we established an expression for the expected number of successful mutations, $E[K_S]$, in a population. By noting that

$$E[K_S] = E[K_S \,|\, \mathcal{R}]P(\mathcal{R}) + E[K_S \,|\, \mathcal{D}]P(\mathcal{D})$$

where, obviously, $E[K_S \,|\, \mathcal{D}] = 0$, we get the following expression for the expected number of rescuing mutations in a rescued population starting from W_0 wildtype individuals:

$$E[K_S \,|\, \mathcal{R}] = \frac{W_0 p_b u E[X]}{(1-(1-u)E[X])\left(1-q_w^{W_0}\right)} \tag{24}$$

As there must be at least one rescuing beneficial mutation in a rescued population, we realize that $E[K_S \,|\, \mathcal{R}] > 1$.

For example, with the parameters used in Figure 2A ($W_0 = 10^4$, $r = 0.01$, $s = 0.02$, and $u = 10^{-5}$), the expected number of independent beneficial mutations occurring during the course of evolution is $E[K_S \,|\, \mathcal{R}] = 1.1$ (Equation (24)). The corresponding numbers for the other parts of Figure 2 are: 2.0 (B), 2.2 (C), and 15.6 (D).

5.4 Waiting Time for a Rescuing Mutation

We next turn our attention to the times T and T_S of the first appearances of a beneficial mutation and a successful mutation, respectively, in a rescued population. Proposition 4.3 and Corollary 4.3.1 give the unconditional probabilities $P(T > n)$ and $P(T_S > n)$, and the Law of Total Probability gives

$$P(T > n \,|\, \mathcal{R}) = \frac{P(T > n) - P(T > n \,|\, D)P(\mathcal{D})}{P(\mathcal{R})}$$

with a similar expression for T_S. Note that conditioning on rescue implies that there is at least one mutation event in finite time, that is, we have $P(T = \infty \,|\, \mathcal{R}) = 0$. In contrast, both $P(T = \infty \,|\, \mathcal{D}) > 0$ and $P(T = \infty) > 0$, because the population may die out without experiencing any mutations. Therefore, $E[T] = \infty$ and $E[T \,|\, \mathcal{D}] = \infty$, whereas $E[T \,|\, \mathcal{R}]$ may still be finite (which also means that we cannot apply Equation (18) to find $E[T \,|\, \mathcal{R}]$).

Proposition 4.3 shows how to compute the probability $P(T > n)$. To get to $P(T > n \,|\, \mathcal{R})$ we need to compute $P(T > n \,|\, \mathcal{D})$ by use of the conditional offspring law $\widehat{P}$ of Equation (19), but alas, there is no immediate analog of Proposition 4.3. This proposition relies upon the conditional distribution of mutants being binomial when conditioned on the total number of offspring; this property does not transfer to $\widehat{P}$ (which can be realized by simple examples, for example binary splitting). However, the case of a Poisson offspring law does lend itself to such analysis, which we address in the next section.

Things are easier when it comes to T_S, which is also the more interesting variable, answering the question when the (first) rescuing mutation arises. As extinction obviously means that there are no successful mutations, we have $P(T_S > n \,|\, \mathcal{D}) \equiv 1$ and hence

$$P(T_S > n \,|\, \mathcal{R}) = \frac{P(T_S > n) - P(\mathcal{D})}{P(\mathcal{R})}$$

which gives

Corollary 5.1.1.

$$P(T_S > n \,|\, \mathcal{R}) = \frac{\left(\widetilde{H}_n(1-\widetilde{u})\right)^{W_0} - q_w^{W_0}}{1 - q_w^{W_0}}$$

by Corollary 4.3.1. We can now obtain the probability mass function

$$P(T_S = n \,|\, \mathcal{R}) = P(T_S > n-1 \,|\, \mathcal{R}) - P(T_S > n \,|\, \mathcal{R}),\ n = 1, 2, \ldots$$

and the expected value

$$E[T_S \,|\, \mathcal{R}] = \sum_{n=0}^{\infty} P(T_S > n \,|\, \mathcal{R})$$

Orr and Unckless [19] provided the following weak selection/weak mutation approximation (Supporting Information Text S2)

$$P(T_S = n \,|\, \mathcal{R}) \approx r e^{-rn} \frac{C e^{-C(1-e^{-rn})}}{1 - e^{-C}} \tag{25}$$

where $C = 2W_0 u(s-r)/r$. Figure 3 shows that this approximation works well, even in conditions where the approximation for population size breaks down (compare Figures 2D and 3D).

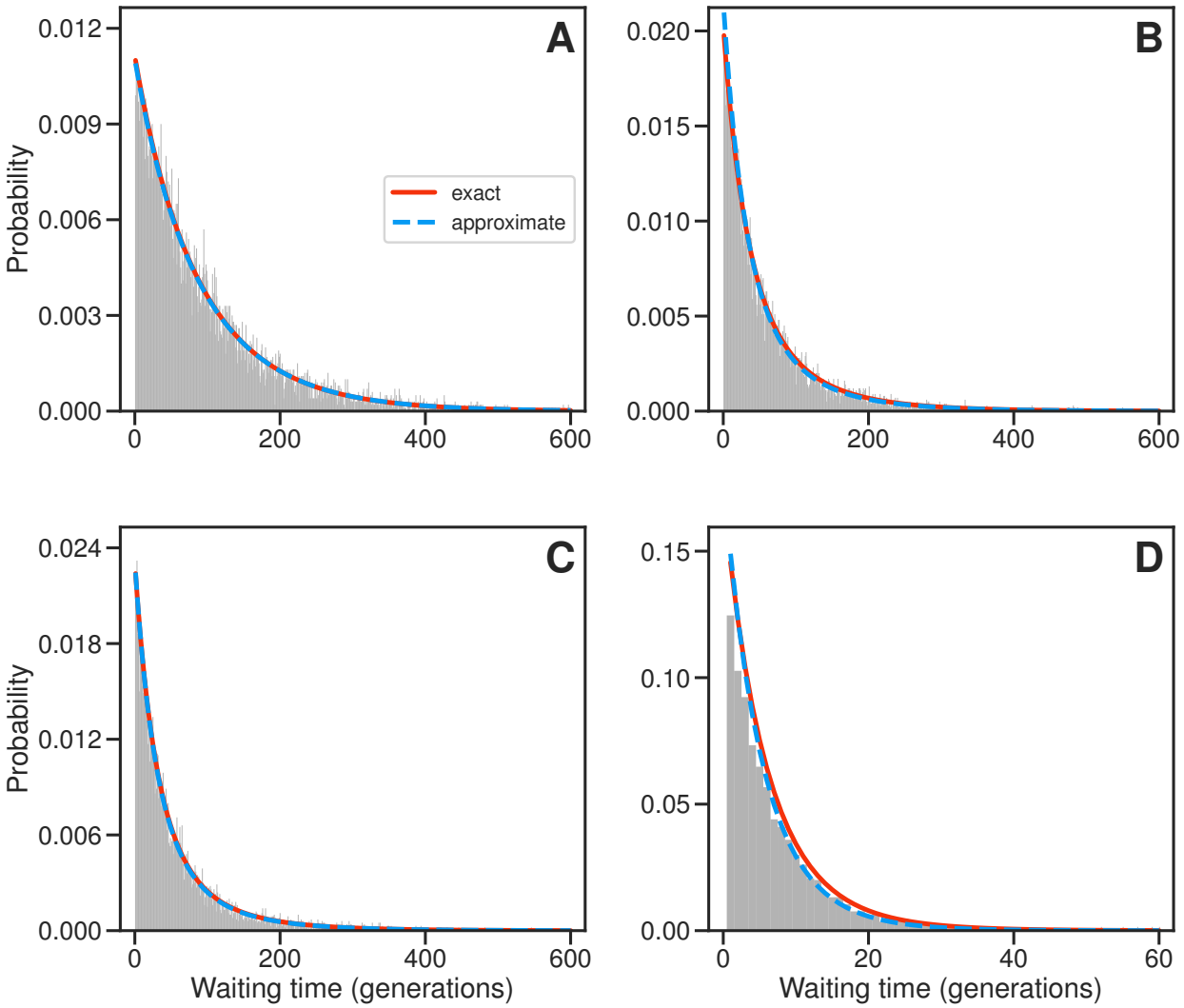

Figure 3: Waiting time for the first rescuing mutation. We considered evolutionary rescue based on new mutations. The parameter combinations in each part as the same as in Figure 2. In all cases we assumed an initial population size of $W_0 = 10^4$ wildtype individuals ($B_0 = 0$ beneficial individuals), a degree of wildtype maladaptation of $r = 0.01$, and a Poisson offspring distribution. (A) Weak selection/weak mutation: $s = 0.02$ and $u = 10^{-5}$ (green arrow, Figure 1A). (B) Stronger selection: $s = 0.1$ and $u = 10^{-5}$ (purple arrow, Figure 1A). (C) Stronger mutation: $s = 0.02$ and $u = 10^{-4}$ (far right, Figure 1C). (D) Stronger selection and stronger mutation: $s = 0.1$ and $u = 10^{-4}$. Histograms of waiting times are based on 10^4 stochastic simulations. The histograms are truncated for clarity, which leads to the exclusion of $< 3\%$ of values in each case. Rescue was defined as reaching $Z_n > 1.5 \times 10^4$. Exact probabilities (red, continuous) were calculated using Corollary 5.1.1. Approximate probabilities (blue, dashed) assuming weak mutation were calculated using Equation (25).

5.5 The Poisson Distribution

As the Poisson distribution is frequently used as offspring distribution (as in our simulations), we finish with a section about its properties in the context of evolutionary rescue. By standard properties of the Poisson distribution, the numbers of wildtype and beneficial offspring, W and B, are independent with $W \sim \mathrm{Poi}((1-r)(1-u))$ and $B \sim \mathrm{Poi}((1-r)\,u)$. As it turns out, both the Poisson distribution and the independence still hold when conditioning on the population going extinct. By Equation (19) and independence, the reproduction law conditioned on extinction is

$$\begin{aligned}\widehat{P}(j,k) &= q_w^{j-1} q_b^k P(j,k) \\ &= q_w^{j-1} q_b^k P_W(j) P_B(k) \\ &= \widehat{P}_W(j) \widehat{P}_B(k)\end{aligned}$$

The last equality follows from the multiplicative separation of j and k in the expression for $\widehat{P}(j,k)$; by this, the marginal distributions are independent. Hence $\widehat{P}_W(j) = c\, q_w^{j-1} P_W(j)$ where the normalizing constant c is obtained by summing over j:

$$\begin{aligned}\sum_{j=0}^{\infty} q_w^{j-1} P_W(j) &= \frac{1}{q_w} \varphi_W(q_w) \\ &= \frac{1}{q_w} e^{(1-r)(1-u)(q_w-1)}\end{aligned}$$

yielding

$$c = q_w e^{-(1-r)(1-u)(q_w-1)}$$

and thus

$$\begin{aligned}\widehat{P}_W(j) &= e^{-(1-r)(1-u)(q_w-1)} q_w^j P_W(j) \\ &= e^{-(1-r)(1-u)(q_w-1)} q_w^j e^{-(1-r)(1-u)} \frac{((1-r)(1-u))^j}{j!} \\ &= e^{-(1-r)(1-u)q_w} \frac{((1-r)(1-u)q_w)^j}{j!}\end{aligned}$$

which we recognize as a Poisson distribution with mean $(1-r)(1-u)q_w$ and may write as

$$\widehat{W} \sim \mathrm{Poi}\big((1-r)(1-u)q_w\big)$$

A similar calculation yields

$$\widehat{B} \sim \mathrm{Poi}\big((1-r)uq_b\big)$$

Put together, these also give the total number of offspring

$$\widehat{X} = \widehat{W} + \widehat{B} \sim \mathrm{Poi}\big((1-r)((1-u)q_w + uq_b)\big)$$

A beneficial individual has a number of (necessarily beneficial) offspring B which is $\mathrm{Poi}((1-r)(1+s))$. Conditioned on extinction, more similar calculations give

$$\widehat{B} \sim \mathrm{Poi}\big((1-r)(1+s)q_b\big)$$

for a beneficial parent. By standard properties of the Poisson distribution, we also get the conditional distributions of $\widehat{W}$ and $\widehat{B}$ given $\widehat{X}$ as binomial $(\widehat{X}, 1-\widehat{u})$ and $(\widehat{X}, \widehat{u})$, respectively, where

$$\widehat{u} = \frac{uq_b}{uq_b + (1-u)q_w}$$

As $q_w > q_b$, we have $\widehat{u} < u$ which makes sense.

The standard Poisson properties do not, however, carry over to the population conditioned on rescue. By the Law of Total Probability:

$$P(j,k \,|\, \mathcal{R}) = \frac{1 - q_w^j q_b^k}{1 - q_w} P_W(j) P_B(k)$$

showing that, conditioned on rescue, W and B are neither Poisson nor independent.

6 Discussion

We use a two-type branching process to model evolutionary rescue. The types are wildtype individuals and individuals carrying beneficial mutations. The parameters are: the degree of maladaptation of wildtype individuals (r), the selective benefit of a mutation (s), the beneficial mutation rate (u), and the initial size and composition of the population (W_0, B_0). Our model is simple, but provides a basis for an understanding of the interplay between fundamental parameters in evolutionary rescue. Our formulas for population growth, number of rescuing mutations, and time until rescue in

rescued populations boil down to those of Orr and Unckless [18, 19] in cases of weak selection/weak mutation, but the discrepancy can be large when selection is strong (Figures 1D and 2D).

Strong selection is frequently observed in nature, notably when humans are involved [22]. Examples include global climate change [8] and the use of antibiotics [21] and pesticides [11]. Strong levels of maladaptation (i.e., high r) are common. For example, shifting a population of the cowpea seed beetle, *Callosobruchus maculatus*, adapted to mung bean to lentil, caused approximately 99% mortality [25]. This extreme maladaptation caused multiple experimental populations to go extinct, but one population adapted quickly and reached 69% survival in 5 generations.

Application of our model to real populations is not straightforward because the mutational parameters (s and u) are difficult to estimate. The baseline values we adopted in our numerical exploration of the model ($s = 0.02$ and $u = 10^{-5}$; e.g., green arrow in Figure 1A) are consistent with estimates obtained in three independent experimental studies on populations of the bacterium *Escherichia coli* [12, 23, 27]. Stronger mutational effects and a lower mutation rate ($s \approx 0.1$ and $u \approx 10^{-7}$) were observed in a different study on the same species [28]; weaker mutational effects and a higher beneficial mutation rate ($s \approx 0.01$ and $u \approx 10^{-4}$) were observed in the yeast *Saccharomyces cerevisiae* [6]. Thus, the scenarios in the numerical illustrations in Figures 1–3 involve plausible values of the mutational parameters.

Our model makes the simplifying assumption that all beneficial mutations have the same effect. Studies in both *S. cerevisiae* [6] and the bacterium *Pseudomonas fluorescens* [14] found evidence that the beneficial effects of mutations are approximately exponentially distributed. A study in two different bacteriophages (ID11 and ϕ6) found support, instead, for a uniform distribution of mutational effects [26]. Our model could be extended to consider more than one type of beneficial individual.

Another central assumption of our model is that an individual can acquire at most one beneficial mutation. This assumption precludes rescue involving multiple mutations in our model. There is experimental evidence that evolutionary rescue through individuals carrying multiple mutations does occur in real populations. For example, populations of *S. cerevisiae* "rescued" from high copper concentrations often acquired multiple beneficial mutations [7]. Osmond et al. [20] showed that if beneficial mutations of small effect are more common than those of large effect then rescue by multiple mutations can be more likely than rescue by single mutations.

Most models of evolutionary rescue [9, 18, 19], including ours, ignore deleterious mutations. This modeling decision is questionable because deleterious mutations are expected to be orders of magnitude more likely to occur than beneficial ones. The accumulation of deleterious mutations is predicted to accelerate the decline in population size in a process known as mutational meltdown [16]. Thus, the probability of rescue from new mutations predicted in a model like ours is likely to be overly optimistic.

Our model makes several additional simplifying assumptions including, but not restricted to, constant environment, asexual reproduction, density- and frequency-independence, and hard selection. All of these assumptions have a potential to influence the probability, tempo, and mode of evolutionary rescue. Many of the assumptions mentioned above, specially those related to mutation, can be explored within the modeling framework we developed.

Acknowledgements

We thank Alex Stewart for helpful discussions and Logan Chipkin for his early contribution to this project. The National Science Foundation grant DEB-2014566 (awarded to R.B.R.A.) and National Institutes of Health grant R15GM093957 (awarded to P.O.) funded this work.

References

[1] D. Agashe, J. J. Falk, and D. I. Bolnick. Effects of founding genetic variation on adaptation to a novel resource. Evolution 65(9): 2481–2491, 2011.

[2] K. B. Athreya and P. E. Ney. Branching Processes. Springer Verlag, New York, 1972

[3] G. Bell. Evolutionary rescue. Annual Review of Ecology, Evolution and Systematics 48: 605–627, 2017.

[4] G. Bell and A. Gonzalez. Evolutionary rescue can prevent extinction following environmental change. Ecology Letters 12(9): 942–948, 2009.

[5] S. M. Carlson, C. J. Cunningham, P. A. H. Westley. Evolutionary rescue in a changing world. Trends in Ecology and Evolution 29(9): 521–530, 2014.

[6] E. M. Frenkel, B. H. Good, and M. M. Desai. The fates of mutant lineages and the distribution of fitness effects of beneficial mutations in laboratory budding yeast populations. Genetics 196(4): 1217–1226, 2014.

[7] A. C. Gerstein, J. Ono, D. S. Lo, M. L. Campbell, A. Kuzmin, and S. P. Otto. Too much of a good thing: The unique and repeated paths toward copper adaptation. Genetics 199(2): 555–571, 2015.

[8] P. Gienapp, M. Lof, T. E. Reed, J. McNamara, S. Verhulst, and M. E. Visser. Predicting demographically sustainable rates of adaptation: can great tit breeding time keep pace with climate change? Philosophical Transactions of the Royal Society of London B Biological Sciences 368(1610): 20120289, 2013.

[9] R. Gomulkiewicz and R. D. Holt. When does evolution by natural selection prevent extinction? Evolution 49(1): 201–207, 1995.

[10] J. B. S. Haldane. A mathematical theory of natural and artificial selection. Part V: Selection and mutation. Mathematical Proceedings of the Cambridge Philosophical Society 23(7): 838–844, 1927.

[11] N. J. Hawkins, C. Bass, A. Dixon, and P. Neve. The evolutionary origins of pesticide resistance. Biological Reviews 94(1): 135–155, 2018.

[12] M. Hegreness, N. Shoresh, D. Hartl, and R. Kishony. An equivalence principle for the incorporation of favorable mutations in asexual populations. Science 311(5767): 1615–1617, 2006.

[13] R. D. Holt and R. Gomulkiewicz. Conservation implication of niche conservatism and evolution in heterogeneous environments. In: R. Ferrière, U. Diekmann, and D. Couvert, editors. Evolutionary Conservation Biology, Cambridge University Press: 244–264, 2004.

[14] R. Kassen and T. Bataillon. Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. Nature Genetics 38(4): 484–488, 2006.

[15] P. Jagers and A. Lagerås. General branching processes conditioned on extinction are still branching processes. Electronic Communications in Probability 13: 540–547, 2008.

[16] M. Lynch, R. Bürger, D. Butcher, and W. Gabriel. The mutational meltdown in asexual populations. Journal of Heredity 84(5): 339–344, 1993.

[17] M. Lynch, M. S. Ackerman, J.-F. Gout, H. Long, W. Sung, W. K. Thomas, and P. L. Foster. Genetic drift, selection and the evolution of the mutation rate. Nature Reviews Genetetics 17(11): 704–714, 2016.

[18] H. Allen Orr and R. L. Unckless. Population extinction and the genetics of adaptation. American Naturalist 172(2): 160–169, 2008.

[19] H. Allen Orr and R. L. Unckless. The population genetics of evolutionary rescue. PLoS Genetics 10(8): 1–9, 2014.

[20] M. M. Osmond, S. P. Otto, and G. Martin. Genetic paths to evolutionary rescue and the distribution of fitness effects along them. Genetics 214(2): 493–510, 2020.

[21] T. Oz, A. Guvenek, S. Yildiz, E. Karaboga, Y. T. Tamer, N. Mumcuyan, V. B. Ozan, G. H. Senturk, M. Cokol, P. Yeh, and E. Toprak. Strength of selection pressure is an important parameter contributing to the complexity of antibiotic resistance evolution. Molecular Biology and Evolution 31(9): 2387–2401, 2014.

[22] S. R. Palumbi. Humans as the world's greatest evolutionary force. Science 293(5536): 1786–1790, 2001.

[23] L. Perfeito, L. Fernandes, C. Mota, and I. Gordo. Adaptive mutations in bacteria: high rate and small effects. Science 317(5839): 813–815, 2007.

[24] J. Ramsayer, O. Kaltz, and M. E. Hochberg. Evolutionary rescue in populations of *Pseudomonas fluorescens* across an antibiotic gradient. Evolutionary Applications 6(4): 608–616, 2013.

[25] A. Rêgo, F. J. Messina, Z. Gompert. Dynamics of genomic change during evolutionary rescue in the seed beetle *Callosobruchus maculatus*. Molecular Ecology 28(9): 2136–2154, 2019.

[26] D. R. Rokyta, C. J. Beisel, P. Joyce, M. T. Ferris, C. L. Burch, and H. A. Wichman. Beneficial fitness effects are not exponential for two viruses. Journal of Molecular Evolution 67(4): 368–76, 2008.

[27] A. Sousa, S. Magalhães, and I. Gordo. Cost of antibiotic resistance and the geometry of adaptation. Molecular Biology and Evolution 29(5): 1417–1428, 2012.

[28] W. Zhang, V. Sehgal, D. M. Dinh, R. B. R. Azevedo, T. F. Cooper, and R. Azencott. Estimation of the rate and effect of new beneficial mutations in asexual populations. Theoretical Population Biology 81(2): 168–178, 2012.