Sex with no regrets: How sexual reproduction uses a no regret learning algorithm for evolutionary advantage

DOI:
10.1016/j.jtbi.2017.05.018

Document Version
Accepted author manuscript

Link to publication record in Manchester Research Explorer

Citation for published version (APA):

Published in:
Journal of Theoretical Biology

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
SEX WITH NO REGRETS: HOW SEXUAL REPRODUCTION USES A NO REGRET LEARNING ALGORITHM FOR EVOLUTIONARY ADVANTAGE

OMER EDHAN¹, ZIV HELLMAN², AND DANA SHERILL-ROFE³,⁴

ABSTRACT. The question of ‘why sex’ has long been a puzzle. The randomness of recombination, which potentially produces low fitness progeny, contradicts notions of fitness landscape hill climbing. We use the concept of evolution as an algorithm for learning unpredictable environments to provide a possible answer. While sex and asex both implement similar machine learning no-regret algorithms in the context of random samples that are small relative to a vast genotype space, the algorithm of sex constitutes a more efficient goal-directed walk through this space. Simulations indicate this gives sex an evolutionary advantage, even in stable, unchanging environments. Asexual populations rapidly reach a fitness plateau, but the learning aspect of the no-regret algorithm most often eventually boosts the fitness of sexual populations past the maximal viability of corresponding asexual populations. In this light, the randomness of sexual recombination is not a hindrance but a crucial component of the ‘sampling for learning’ algorithm of sexual reproduction.

1. INTRODUCTION

The question of ‘why sex’ has long been a puzzle of evolution. Several advantages of sexual reproduction have been proffered over the years, including increased population genotypic variability and novel genotypes created by recombination, Muller’s ratchet (the suggestion that the genomes of an asexual population accumulate deleterious mutations in an irreversible manner), the Red Queen hypothesis according to which sexual reproduction grants an advantage in combatting parasites, and Kondrashov’s deterministic mutation hypothesis, under which negative epistasis leads to mutations having increasingly negative fitness effects in asexual reproduction. All of these explanations, to one extent or another, have mostly been deemed insufficient (Otto and Gerstein (2006); Otto (2009); Barton (2009)).

¹. School of Social Sciences, University of Manchester, UK.
². Department of Economics, Bar Ilan University, Israel.
³. Department of Developmental Biology and Cancer Research at Hadassah Medical School and the Federmann Center for the Study of Rationality, Hebrew University, Israel.
⁴. Department of Management, Bar Ilan University, Israel.
As against these there is a long list of costs of sex, amongst which is the issue that sex tends to disassociate beneficial gene combinations by random shuffling.\(^1\) The very randomness of recombination shuffling, which potentially produces low fitness progeny from high fitness parents (Eshel and Feldman (1970); Lewotin (1971)), seems to contradict the foundational Fundamental Theorem of Natural Selection, which states that the mean fitness of a population monotonically increases in each generation until the population attains a fixed optimum (Akin (1979)).

In fact, it has long been known that in multi-locus sexually reproducing populations with epistasis the mean fitness can decrease over time and there is even no hope of finding any function, fitness or otherwise, that is non-decreasing and climbs a gradient (Akin (1983); Ewens (2008); Hastings (1981)). This sharpens the questions. Which trait is optimised by sexual reproduction? How could sex out-compete pure asexual reproduction when the latter does monotonically increase mean fitness in each generation in a rapid climb to optimal fitness? What is the point of the random shuffling of genes inherent in sexual recombination, which at first glance would appear to have a neutral effect on population variance in fitness?

Our answer is that sex is an instantiation of a machine learning algorithm (in particular, a type of algorithm called ‘no regret learning’ that has recently been intensely studied in the computer science literature) in the context of random sampling of the genotype space, with populations that are small in comparison to the space of all genotypes serving as the samples of the algorithm.

This is sub-optimal for fitness but is sufficient to give sex an edge over asex; in most cases the boost in fitness beyond the initial sample afforded by the learning algorithm is enough to outpace the best that asex can do.

The sampling aspect is key to our results and is a perspective that is often missing in models of the evolution of sexual reproduction in the literature. The space of all possible genotypic combinations for any species is astronomically large. Each generation is only a sparse sample of the vast potential genotype space. The randomness of sexual recombination in this view is not an anomaly but inherently necessary to ensure that the sampling in the generations yields unbiased estimators of various parameters, as detailed in the body of the paper.

In outline, our results are as follows. Theorem 1 and Theorem 2 establish, by careful consideration of the replicator dynamics of sexual and asexual populations and comparison with the Hedge algorithm (a well-known machine learning algorithm), that in infinite

\[^1\] We deliberately avoid discussing the two-fold cost of sex, by utilising a model of monoecious organisms in which every individual is capable of bearing offspring. We do this for the sake of isolating effects, since a monoecious species avoids a two-fold cost but does bear the potential disadvantages of genetic shuffling. In other words, the two-fold cost of sex is a cost that is borne over and above the more basic apparent cost of genetic shuffling.
population models the end goals of both reproductive strategies are the implementation of no regret learning strategies. Corollary 1 and the example following it show that these strategies actually give asexual reproduction a theoretical advantage over sexual reproduction.

In the finite population model with generational sampling from the full potential genotype space, the situation is quite different. Proposition 1 and Theorem 3 show that under sampling the sexually reproducing population always attains the same viability outcomes as it would in the infinite population case, a result that is not true for asexually reproducing populations; indeed, the larger the genotype space, the less likely that a random sampling of a finite asexual population will contain an optimal genotype.

In Section 5 our central claim, that the learning strategy of sexually reproducing populations, which are small compared to the space of all genotypes, eventually enables them to overtake competing asexual populations, is tested in simulations. We report that our simulations indicate that the learning capacity of sexually reproducing populations implementing learning algorithms is impressively consistent and reliable under many different conditions. In a majority of our simulation trials, the sexual population is initially at a disadvantage but as the learning aspect of the no-regret algorithm proceeds, the mean values steadily climb, most often past the maximal viability of the asexual population.

2. INFINITE POPULATION MODEL

2.1. Background. The Fundamental Theorem of Natural Selection, as presented in standard textbooks (see, for example, Chapter 3 of Gillespie (2004)), states that the change in the mean fitness of a population undergoing natural selection is proportional to the additive genetic variation in fitness. As variances are always positive, the mean fitness will always increase when natural selection changes the allele frequency. This was given a precise formulation by Siavash Shahshahni, who showed that the dynamics of a population with additive fitness under pure selection follows a vector field that is a gradient with respect to a Riemannian metric introduced by Shahshani himself (Shahshahni (1979); Akin (1979)). In plainer language, a population under pure selection reliably climbs a gradient to attain an optimal point in the fitness landscape.

In the presence of epistasis, however, this result does not hold beyond one locus (Crow (2002); Ewens (1989); Nagylaki (1991)). As noted in Ewens (1989), “It is well known that in discrete-time models, when fitness depends on the genes at more than one locus, the total change in mean fitness can be negative, even under random mating, essentially because of recombination effects.”

The fact that even in the simplest possible model, the two-locus-two-allele setting, it is possible to construct a fitness landscape in which the end result of the dynamics is suboptimal was noted as early as Moran (1964). Further work by several authors (Akin (1982,
Hastings (1981)) established in much more general terms that in all multi-locus and multi-allele settings there is no way to define any population-wide gradient that all systems will climb. In general, equilibrium paths in such systems will fail to attain stable optimal results, and even if the system is temporarily in an optimal state the dynamics may push the system back ‘down the slope’.

Organisms typically have thousands of loci in their genotypes, with many epistatic interactions. The dynamics of asexual reproduction, in all situations, are easily shown reliably always to climb a gradient and home in on the genotypes of optimal fitness. If in contrast the dynamics of sexual reproduction can demonstrably often be shown to lead to sub-optimal solutions, how could sexual reproduction compete against asexual reproduction?

Furthermore, it has long been recognised in the literature that sex breaks down favourable combinations of genes (Eshel and Feldman (1970); Lewontin (1971)), posing a challenge to why sex has been an evolutionarily successful reproduction strategy. As Otto (2009) states, ‘by breaking down the genetic combinations built by past selection, sex almost always creates offspring of lower average fitness than asexual reproduction’.

If we regard sexual reproduction as a random walk through genotype space, the randomness that is a central element of its mechanism might seem to imply that such a walk would produce neutral results. Barton (2009) writes: ‘we still do not have a compelling general explanation for why so many organisms go to so much trouble to mix their genomes with others’, and goes on to note that ‘crucially, recombination does not necessarily either increase or decrease the variance in fitness. If genes are already randomly associated in the population, shuffling them will make no difference.’

A long-prevailing suggested answer to this question claimed that since genetic shuffling will tend to reduce fitness if there are positive associations between favourable genes, there ought to be a preponderance of negative interactive effects (i.e., negative epistasis) to explain the advantage of sex. There is, however, little evidence found to date to support this hypothesis (Bonhoeffer et al (2004); de Visser and Elena (2007); Elena and Lenski (1997); Kouyos et al (2007)). The prominent role that random gene shuffling plays in sexual reproduction has continued to be baffling because it would seem actively to work against any possibility of attaining optimal population fitness.

2.2. Infinite Sexually Reproducing Population Model. To contend with the issues mentioned above, we work with a model of a monoecious sexually reproducing haploid population with arbitrary numbers of loci and alleles, random mating, discrete time, and non-overlapping adult generations. A classic example of such an organism is the yeast *Saccharomyces cerevisiae*, which has the ability to change its mating type from $a$ to $\alpha$ and from $\alpha$ to $a$, lives in a natural state of haploidy, and the reproduces in a semelparous manner.

We suppose as a primitive the existence of a finite set of alleles $A$ (which may be thought of as the basic ‘gene pool’). We furthermore suppose the existence of a fundamental
partition $\mathcal{L} := \{L_1, \ldots, L_\ell\}$ of $\mathcal{A}$, i.e., $\bigcup_j L_j = \mathcal{A}$ and $L_i \cap L_j = \emptyset$ for every pair $L_i, L_j \in \mathcal{L}$ such that $i \neq j$. Each partition element $L_j$ is a locus and $\mathcal{L}$ is the set of loci, of cardinality $\ell$. For ease of notation, denote $I = (1, \ldots, \ell)$.

In principle the analysis here could be conducted assuming populations of any $n$-ploidy for $n \geq 1$. For concreteness and mathematical simplicity we focus on the haploid (i.e., $n = 1$) case. A (haploid) genotype $g$ is a selector function $g : \mathcal{L} \to \mathcal{A}$; i.e., $g$ satisfies the condition that $g(L_j) \in L_j$ for each $j \in I$.

More simply, a generic genotype $g$ can be depicted as a string of alleles,

$$a_1a_2 \ldots a_{\ell-1}a_\ell,$$

such that $a_j \in L_j$ for all $j$.

By the space of all genotypes we mean the set of all such $\ell$ long possible combinations $g$ of alleles with $a_j \in L_j$. We denote this genotype space by $\Gamma$.

The models we construct here exclude the possibility of mutations, in the sense that the set of alleles $\mathcal{A}$ is fixed throughout; against this background, mutation would mean expanding $\mathcal{A}$. We concentrate on natural selection learning within a fixed genotype space, which roughly corresponds to the period of reproductive generations between the introduction of beneficial mutations.

We initially suppose an infinite founding population with at least one individual bearing each possible genotype. There is no physical linkage between alleles at different loci in the model (stochastic linkage, however, may arise through selection favouring specific allele combinations).

Individuals in the adult population at time $t$ with genotypic frequency $\{p^t_g\} g \in \Gamma$ are paired, reproduce, and then die. Each reproducing pair conceives $2\zeta$ meiotic products of zygotes, where $\zeta$ is an integer. The haploid genotype of each meiotic product of the zygote is formed by recombination, by which is meant independently picking at each locus an allele with equal probability from one of the two parent genotypes. This produces an offspring population with genotypic frequency $\{\tilde{p}^t_g\} g \in \Gamma$ that forms the basis for the adult population at time $t + 1$.

Not all of the haploid individuals arising from zygote meiosis conceived during reproduction at time $t$ survive to adulthood. The viability $v^t_g$ is the probability that an individual bearing $g$ conceived at $t$ survives to adulthood at $t + 1$. The collection $\{v^t_g\} g \in \Gamma$ forms a tensor at time $t$. Arbitrary epistatic relations between loci may be expressed in the viability variables. Note that the term epistasis is used here in a broad sense: the measure $v^t_g$ can be an arbitrary function from the genotype $g$ to a real number and may depend on the entire genotypic string. In particular, there is no assumption that epistasis necessarily measures some deviation from a linear combination of individual allelic contributions to viability, as
we take the view that there is no such thing as a naked gene without some broader genomic context.

For each $g$, we may regard $v^t_g$ as a time-dependent random variable that cannot deterministically be foreseen. One may consider the underlying randomness to be due to unpredictable environmental variables. There is no assumption of normality of distributions or even independent identically distributed random variables. We will, however, adopt here the following assumptions:

1. The values of $v^t_g$ range as $0 \leq v^t_g \leq 1$;
2. For all $t$, for each allele $a$ there is at least one $g \in \Gamma$ containing $a$ such that $v^t_g > 0$.
3. There is a minimal $v^* > 0$ such that for all $t$ and all $g$, if $v^t_g \neq 0$ then $v^t_g \geq v^*$.

The first assumption allows for some genetic combinations to be absolutely lethal: for example, it may happen that certain combinations lead to such gross embryonic developmental defects that they necessarily cause foetal or post-natal mortality. On the other hand, it is not the aim of this work to depict the process of extinction of species, hence we require that at each time $t$ the value of $v^t_g > 0$ for at least one genotype $g$, to rule out a sudden extinction event. But we also need a stronger condition to rule out a ‘sieve’ effect that would kill off each allele separately over time until none are left. The second assumption suffices to avoid these extreme cases.

The third assumption is mainly needed for technical reasons but it is amenable to a natural interpretation. A genotype with $v^t_g = 0$ has no hope at all to attain maturity. It is essentially conceived with a flaw that guarantees early death; one might as well assume pre-natal mortality. Since there are a finite number of possible genotypes, there is some $v^{*t} > 0$ such that for all $g$ at time $t$, if $v^t_g \neq 0$ then $v^t_g \geq v^{*t}$, i.e., conditional on surviving gestation and being born at time $t$, each individual has at least some minimal probability $v^{*t}$ of surviving to adulthood. Assumption (3) then states that although the viability landscape can change considerably over time, there is some minimal $v^* > 0$ such that conditional on being born each individual has at least the probability $v^*$ (which can be an extremely small but non-zero value) of surviving to adulthood, independently of which time period is under consideration.

From here we will denote a generic allele as $a_{j,i}$, interpreted as allele $i$ in locus $j$. The mean population viability at time $t$ is $\bar{v}^t = \sum_{g \in G} p^t_g v^t_g$. As reproduction produces $\zeta$ meiotic products of zygotes per capita, the population growth rate at time $t$ is $\rho^t = \zeta \bar{v}^t$. Alleles at each locus compete over which allele will have a greater long run population share. Denote by $q^t_{j,i}$ the fraction of the adult population at time $t$ bearing allele $a_{j,i}$, and denote $\tilde{q}^t_{j,i}$ the corresponding fraction of the offspring population. These values are given
SEX WITH NO REGRETS: HOW SEXUAL REPRODUCTION USES A NO REGRET LEARNING ALGORITHM FOR EVOLUTIONARY ADVANTAGE

as

\[
q_{j,i}^t := \sum_{g \in G_{j,i}} p_g^t .
\]

and

\[
\tilde{q}_{j,i}^t := \sum_{g \in G_{j,i}} \tilde{p}_g^t .
\]

The dynamics of the adult population from generation \( t \) to \( t + 1 \) is given by

\[
p_{g}^{t+1} = v_g^t \tilde{q}_{j,i}^t / \tilde{q}_{j,i}^t ,
\]

Let \( \tilde{p}_g^t \mid_{a_{j,i}} = \tilde{p}_g^t / q_{j,i}^t \) be the population weight of individuals of type \( g \) at time \( t \) relative to the population restricted to those bearing allele \( a_{j,i} \). The mean viability of haploid meiotic products of zygotes bearing \( a_{j,i} \) at time \( t \) is \( \varphi_{j,i}^t = \sum_{g \in G} v_g^t \tilde{p}_g^t \mid_{a_{j,i}} \). Again, multiplying by \( \zeta \) gives the mean expected number of meiotic products of zygotes per capita bearing \( a_{j,i} \) surviving to maturity, \( \rho_{j,i}^t = \zeta \varphi_{j,i}^t \), that we term the allele’s portfolio growth at time \( t \).

Next, note that

\[
\tilde{q}_{j,i}^t = q_{j,i}^t
\]

for all \( t \) and every allele \( a_{j,i} \). To see this, consider an allele \( a_{j,i} \) and a reproducing individual bearing \( a_{j,i} \). In expectation, \( \zeta \) of that individual’s offspring will also bear \( a_{j,i} \). As this is true of every such individual, and \( q_{j,i}^t \) is the fraction of the population bearing \( a_{j,i} \), it follows that the fraction of offspring bearing \( a_{j,i} \) will also be \( q_{j,i}^t \), i.e., that \( \tilde{q}_{j,i}^t = q_{j,i}^t \). In other words, alleles may be shuffled around into different genetic combinations during reproduction, but they maintain their frequencies relative to each other at each locus.

This sets the stage for deriving the the per locus dynamic, which is

\[
q_{j,i}^{t+1} = q_{j,i}^t \rho_{j,i}^t / \tilde{q}_{j,i}^t .
\]

For the proof, note that

\[
\varphi_{j,i}^t = \sum_{g \in \Gamma} v_g^t \tilde{q}_g^t \mid_{a_{j,i}} = \sum_{g \in G_{j,i}} v_g^t \tilde{p}_g^t / q_{j,i}^t .
\]

From this deduce

\[
q_{j,i}^t \varphi_{j,i}^t = \sum_{g \in G_{j,i}} v_g^t \tilde{p}_g^t .
\]
hence

\[
q_{j,i}^t \rho_{j,i}^t = q_{j,i}^t \frac{\gamma_{j,i}^t \zeta}{\nu^t \zeta} = \sum_{g \in G_{j,i}} v_{g}^{t} \rho_{g}^{t} \frac{p_{g}}{p_{j,i}^t} = \sum_{g \in G_{j,i}} p_{g}^{t+1} = q_{j,i}^{t+1}
\]

2.3. **Infinite Asexually Reproducing Population Model.** The dynamics of asexual reproduction are simpler than those of sexual reproduction. In asexual reproduction, the main variable to trace is the population frequency of the genotypes, \(p^t\).

In detail, in the model an individual bearing genotype \(g \in \Gamma\) at time \(t\) conceives \(\zeta\) offspring, each of which also bears the same genotype \(g\). The probability of survival to adulthood of each of these haploid meiotic products of the reproduction is given stochastically by \(v_{g}^{t}\). As before, the generation at time \(t\) dies shortly after reproduction and the offspring born at time \(t\) that survive to adulthood form the generation at time \(t + 1\).

Denoting again the population frequency of genotype \(g\) by \(p_{g}^t\), the genotypic growth at time \(t\) of genotype \(g\) is

\[
\rho_{g}^t = \zeta v_{g}^{t}.
\]

The population growth rate is then

\[
\rho^t = \sum_{g \in \Gamma} p_{g}^t \rho_{g}^t.
\]

The population frequency dynamic is a straight-forward replicator:

\[
p_{g}^{t+1} = p_{g}^t \frac{\rho_{g}^t}{\rho^t} = p_{g}^t \frac{v_{g}^{t}}{\nu^t}.
\]

3. **Reproduction and the No Regret Algorithm**

Equation (5) is a typical single-locus haploid model replicator equation under conditions of recombination prior to selection. Equation (8) is a standard asexual population replicator. They have the same form; replacing \(q_{j,i}^{t+1}\) and \(\rho_{j,i}^t\) by \(p_{g}^{t+1}\) and \(\rho_{g}^t\), respectively, converts Equation (5) into Equation (8).

We propose here that this similarity between the two population replicators is due to both equations being instantiations of a machine learning algorithm called a no-regret algorithm (Aurora, Hazan, and Kale (2012); Auer et al (1995)). There is no possibility to
know ahead of time, even in probabilistic terms, what future environment (i.e., viability function) will be encountered. A reproductive strategy that implements a simple machine learning (also called on-line learning) algorithm that learns the environment as it is encountered grants evolutionary advantage over non-learning strategies.

Metaphorically, each locus (or asexually reproducing population) is like an investor who does not know which expert (alleles in the sexual case, genotypes for the asexual population) will optimally increase his wealth. He divides his wealth amongst the experts and collects periodic reviews. Experts with above average returns get a proportionally greater share of the investment and those below average less in the next period.

This fits the dynamic of Equations (5) and (8) and is, as we now proceed to explain, an instantiation of a machine learning algorithm called a no-regret algorithm (Aurora, Hazan, and Kale (2012); Auer et al (1995)). The metaphor of an investor who divides his wealth amongst experts and at each time period ‘boosts’ the fraction given to the experts whose portfolios achieve above-average returns, is often used as motivation for the Hedge algorithm. (The Hedge algorithm was originally due to Freund and Schapire (1997). It has since been significantly generalised and is now a major element in the study of on-line learning algorithms. An extensive recent survey of the subject appears in Aurora, Hazan, and Kale (2012).) We present a variant of the Hedge algorithm in detail here as Algorithm 1.

The cumulative gain of this algorithm over time periods from 1 to $T$ is

$$\Phi_H^T := \sum_{t=1}^{T} y_t = \sum_{t=1}^{T} \sum_{i=1}^{k} q_t^i r_t^i.$$ 

One can ask what would have been the cumulative gain if instead of following the algorithm the investor had in every time period given all his wealth to expert $i$ (i.e., always set the weight $w_t^i = 1$ in the algorithm) and nothing to the other experts, in which case the cumulative gain would have been

$$\Phi_i^T := \sum_{t=1}^{T} r_t^i.$$ 

The regret at not having selected *ex ante* the expert who was revealed *ex post* to be the optimal expert is the difference

$$\text{Reg}_T = \max_i \Phi_i^T - \Phi_H^T.$$ 

One can also ask what will be the asymptotic average regret, $\lim_T \text{Reg}_T / T$, as $T$ increases.

Lemma 3.1 and Theorem 3.2 of Auer et al (1995) show that, depending on the specifics of the problem studied, $\eta$ can be tuned so as to guarantee that the average regret of the
Algorithm 1: Exponential Hedge Algorithm

**Parameters:**
- number of experts $k$,
- maximal single-period gain $M$,
- initial weight vector $q^1 \in \Delta^k$,
- a real number $\eta > 0$,
- number of trials $T$

**for** $t = 1, 2, \ldots, T$ **do**
1. Receive reward vector $r^t \in [0, M]^k$
2. Score the weighted mean value $y^t = \sum_{i=1}^{k} q^t_i r^t_i$
3. Update the weight vector by
   $$\hat{q}^{t+1}_i = q^t_i \cdot e^{\eta r^t_i}$$
4. Normalize $\hat{q}^{t+1}$ into an allocation $q^{t+1} \in \Delta^k$ by setting
   $$q^{t+1}_i = \frac{\hat{q}^{t+1}_i}{\sum_j \hat{q}^{t+1}_j}$$
   **for each** $i = 1, \ldots, k$

**end for**

Hedge algorithm decreases at the rate $O(\sqrt{\ln k/T})$, and hence that it asymptotically vanishes. We denote the optimal $\eta$ attaining this goal by $\eta^*$.

We now recast the allelic frequency dynamics of a single locus in a sexually reproducing population, as expressed in Equation (5), as an implementation of a Hedge-styled algorithm. The details are presented in Algorithm 2, which is composed in a way that makes direct comparison to Algorithm 1 as transparent as possible.

The vector dynamics determined by Algorithm 2 come down to $q^{t+1}_i = q^t_i \rho^t_i$, exactly mimicking the dynamics derived above for the alleles. The gain scored in each generation for each allele $i$ is $\ln(\rho^t_i / \rho^t_j)$, i.e., the log of the quotient of that allele’s portfolio growth rate and the worst performing portfolio (this ensures that the quotient is always greater than or equal to 1 and therefore that the log gain is non-negative). The overall gain $y^t$ scored in each generation is the mean log gain, $\sum_i q^t_i \ln(\rho^t_i / \rho^t_j)$, scaled by $\frac{1}{\eta}$ (note that, similarly to $\rho^t$, the mean log gain is a population-level value even though it is calculated at a single locus).

Ostensibly, tracking the cumulative mean log gain may look wrong, as compared to the more standard log mean return, i.e., $\ln(\sum_i q^t_i \rho^t_i)$. There are two justifications for the mean log gain. Firstly, the mean log gain may be considered to be a proxy for measuring
Algorithm 2 Allelic Frequency Algorithm (at a single locus)

Parameters:
number of alleles $k$,
initial allelic frequency vector $q^1 \in \Delta^k$,
number of generations $T$

for $t = 1, 2, \ldots, T$ do
1. Receive the revealed viability landscape $V^t$
2. Calculate, for each allele $i$, the portfolio growth $\rho^t_i$
3. Calculate the average growth rate $\bar{\rho}^t = \sum_{i=1}^k q^t_i \rho^t_i$ and the minimal non-zero growth rate $\hat{\rho}^t$
4. Score a gain $y^t := \frac{1}{\eta^*} \sum_{i=1}^k q^t_i \ln(\rho^t_i / \rho^t)$, with the understanding that the sum does not include indices $i$ such that $\rho^t_i = 0$
5. Update the frequency vector by
   $$\hat{q}^t_{i+1} = q^t_i \cdot e^{\eta^* (1/\eta^* \ln(\rho^t_i / \rho^t))} = q^t_i (\rho^t_i / \hat{\rho}^t),$$
   with the understanding that $\hat{q}^t_{i+1} = 0$ if $\rho^t_i = 0$
6. Normalize $\hat{q}^{t+1}$ into an allocation $q^{t+1} \in \Delta^k$ by setting
   $$q^{t+1}_i = \frac{\hat{q}^{t+1}_i}{\sum_j \hat{q}^{t+1}_j} = q^t_i \frac{\rho^t_i}{\hat{\rho}^t}$$
   for each $i \in 1, \ldots, k$
end for

growth rapidity at each time period, in the sense that, all other things being equal, a population with a higher mean log return will experience greater absolute size growth than a comparable population with a lower mean log return.

Secondly and perhaps more convincingly, the result of the Hedge algorithm itself provides a justification ‘in hindsight’. In detail, for a positive integer $T$ define the cumulative mean log gain as

$$\Psi^T_S := \eta^* \sum_{t=1}^T y^t = \sum_{t=1}^T \sum_{i=1}^k q^t_i \ln \rho^t_i$$

and the corresponding cumulative gain in hindsight from sticking to one allele throughout (i.e., always setting the weight $q^t_i = 1$ in the algorithm for some particular allele $i$)

$$\Psi^T_i := \sum_{t=1}^T \ln \rho^t_i.$$
Then, denoting $\Phi^T_S := \frac{1}{\eta} (\Psi_S - \sum_{t=1}^T \ln \rho^t)$ and $\Phi^T_i := \frac{1}{\eta} (\Psi_i - \sum_{t=1}^T \ln \rho^t)$ we can, as before, inquire about the asymptotic regret $Reg_T = \max_i \Phi^T_i - \Phi^T_S$. Making use again of standard results with respect to regret under the Hedge algorithm, one concludes that $Reg_T/T$ goes to zero as $T$ increases without bound.

In the multi-locus setting, this dynamic is implemented at every locus simultaneously. We then understand $\Psi_S$ to refer to the cumulative log growth that results from this dynamic at every locus, and $\Psi_{a_{j,i}}$ to refer to the cumulative log growth that results when only allele $a_{j,i}$ is present at locus $j$ in the population at every generation.

In a similar vein, $\Psi_{a_{j,i},a_{k,m}}$ will mean the cumulative log growth when alleles $a_{j,i}$ and $a_{k,m}$ are respectively fixed at loci $j$ and $k$, while the other loci follow the Algorithm 2, and so on, until we get to the cumulative log growth attained by supposing that the entire population consists solely of individuals bearing a single genotype $g$.

Furthermore, define an optimal-in-hindsight $a^*_j$ allele at locus $j$ to be an allele that attains the highest asymptotic growth rate, relative to all other alleles at locus $j$, as measured by $\Psi_{a_{j,i}}$ as $T$ increases without bound. (Clearly, an optimal allele cannot be one that ever has a portfolio growth rate of zero, even once, because then this last asymptotic value goes to zero. It is possible that two or more alleles are optimal in hindsight; in that case we may choose $i^*$ arbitrarily from amongst them.)

An optimal-in-hindsight sexual genotype is a genotype $g^S = a^*_1a^*_2\ldots a^*_{\ell-1}a^*_\ell$ such that there exists a dynamic that leads, for each locus $j$, to the allele $a^*_j$ to be optimal-in-hindsight with respect to the alleles in all the other loci. (There may be several optimal-in-hindsight sexual genotypes, not necessarily a unique one. This may be understood by considering the loci to be playing a coordination game and choosing strategies (i.e., alleles) by way of a no-regret algorithm, which may lead to any of several pure Nash equilibria. In this regard, see Kleinberg et al. (2009); Reshef and Parkes (2015)).

**Theorem 1.** *In the infinite population model with full initial genotypic support, the cumulative growth of a sexually reproducing population converges asymptotically to the cumulative growth of a population consisting solely of individuals bearing one of the optimal-in-hindsight sexual genotypes.*

The frequency dynamics of an asexually reproducing population, as expressed in Equation (8), are also an implementation of a Hedge-styled algorithm. The details are presented in Algorithm 3. There are significant differences with respect to a sexually reproducing population. In the asexual setting, not alleles but genotypes directly compete with each other. Furthermore, there is no multi-locus focus in the asexual context, as now the entire population implements the Hedge algorithm.

The vector dynamics determined by Algorithm 3 come down to $p^{t+1}_g = q_g \rho^t$, mimicking those of Equation (8). The gain $y^t$ scored in each generation is again the mean log return.
Algorithm 3 Asexual Reproduction Algorithm

**Parameters:**
- number of genotypes $m$ in genotype space $\Gamma$,
- initial genotypic frequency vector $p^1 \in \Delta^m$,
- number of generations $T$

for $t = 1, 2, \ldots, T$ do

1. Receive the revealed viability landscape $V^t$
2. Calculate, for each genotype $g$, the genotypic growth $\rho^t_g$
3. Calculate the average growth rate $\bar{\rho}^t = \sum_{g \in \Gamma} p^t_g \rho^t_g$ and the minimal non-zero growth rate $\rho^t$
4. Score a gain $y^t := \frac{1}{\eta^*} \sum_{g \in \Gamma} p^t_g \ln(\rho^t_g/\bar{\rho}^t)$, with the understanding that the sum does not include indices $i$ such that $\rho^t_i = 0$
5. Update the genotypic frequency vector by
\[
\hat{p}^{t+1}_g = p^t_g \cdot e^{\eta^* \left( \frac{1}{\rho^t_g} \ln(\rho^t_g/\bar{\rho}^t) \right)} = p^t_g \rho^t_g
\]
with the understanding that $\hat{p}^{t+1}_g = 0$ if $\rho^t_g = 0$
6. Normalize $\hat{p}^{t+1}$ into an allocation $p^{t+1} \in \Delta^n$ by setting
\[
p^{t+1}_g = \frac{\hat{p}^{t+1}_g}{\sum_{g'} \hat{p}^{t+1}_{g'}} = p^t_g \frac{\rho^t_g}{\bar{\rho}^t}
\]
for each $g \in \Gamma$

end for

Define the cumulative mean log growth gain as

\[
(11) \quad \Psi_A := \eta^* \sum_{t=1}^T y^t = \sum_{t=1}^T \sum_{g \in \Gamma} p^t_g \ln \rho^t_g
\]

and the corresponding cumulative gain in hindsight from sticking to one genotype throughout (i.e., always set the weight $p^t_g = 1$ in the algorithm for one particular $g \in \Gamma$)

\[
(12) \quad \Psi_g := \sum_{t=1}^T \ln \rho^t_g.
\]

Then, denoting $\Phi_A := \frac{1}{\eta^*} \Psi_A - \sum_{t=1}^T \ln \rho^t_g$ and $\Phi_g := \frac{1}{\eta^*} \Psi_g - \sum_{t=1}^T \ln \rho^t_g$ we can again inquire about the asymptotic regret between $\Phi_A$ and $\max_g \Phi_g$. 
Definition 1. *The optimal-in-hindsight asexual genotype is*

\[ g^A := \arg \max_{g \in \Gamma} \lim_{T \to \infty} \frac{1}{T} \Psi_g. \]

♦

Theorem 2. *In the infinite population model, the cumulative growth of an asexually reproducing population converges asymptotically to the cumulative growth of a population consisting solely of individuals bearing the optimal-in-hindsight asexual genotype \( g^A \).*

Corollary 1. *In the infinite population model, asymptotic cumulative asexual population growth is greater than or equal to asymptotic cumulative sexual population growth.*

That it is possible for the inequality of Corollary 1 to be strict, i.e., \( \Phi_{g^A} > \Phi_{g^S} \), is established by the following example (inspired by an example in Livnat et al (2008)).

Example. We work with a haploid two-locus model with three alleles per locus. The alleles in locus 1 are denoted by \( a_1, a_2, a_3 \) and those in locus 2 by \( b_1, b_2, b_3 \).

The viability landscape is constant for all time. Denoting by \( v_{kl} \) the viability of genotype \( a_k b_l \), let

\[
\begin{array}{ccc}
  v_{11} = 0.420 & v_{12} = 0.523 & v_{13} = 0.550 \\
  v_{21} = 0.520 & v_{22} = 0.530 & v_{23} = 0.500 \\
  v_{31} = 0.510 & v_{32} = 0.525 & v_{33} = 0.410 \\
\end{array}
\]

Suppose that initially there is a uniform distribution of genotypes. Under asexual reproduction the optimal-in-hindsight asexual genotype \( g^A = a_1 b_3 \), which consistently has the greatest viability, 0.55.

Under sexual reproduction, consider an initial population with each genotype represented by an equal number of individuals. The alleles with the highest mixability in this example are \( a_2 \) and \( b_2 \), and their frequencies are correspondingly boosted in the next generation above those of the other alleles. This results in the increasing of the mixability of these two alleles in the succeeding generation, further boosting their frequencies and so on in each generation. The asymptotic dynamic yields a population growth rate of the optimal-in-hindsight sexual genotype \( g^S = a_2 b_2 \), whose 0.53 viability is outpaced by that of \( g^A \). ♦

The Fundamental Theorem of Natural Selection is a statement about the change in mean fitness in each generation; in contrast, Theorem 1 involves only asymptotic values. The dynamic of the fundamental theorem leads to an equilibrium population composition while Theorem 1 states only that the expected cumulative growth rate converges to what would have been the growth rate had the population hypothetically consisted only of individuals bearing genotype \( g^S \) at each time period, allowing possible cycling through different
population compositions (see Fig. 3). Furthermore, attaining optimal mean viability is not always the end result of the dynamic of Theorem 1, as shown by the example above.

Theorem 1 is similar in content to results in Chastain et al (2013) and Livnat et al (2008), which also use a no-regret approach. Chastain et al (2013) assume weak selection, which is not assumed here (in independent work, Reshef and Parkes (2015) have also managed to obtain similar results without assuming weak selection), and the dynamic in Chastain et al (2013) is deterministic in contrast to the stochastic viability landscapes here. Livnat et al (2008) also include computer simulations of competition between sexual and asexual populations, confirming the objective of the asexual reproductive algorithm is attaining optimal fitness in contrast to the objective of maximal allele mean viability in the sexual reproductive algorithm. In virtually all of those simulations, the asexual mean fitness outcompetes the sexual population.

4. Sampling and the Finite Population Model

If sexual reproduction leads to the dynamic of Theorem 1, which as shown in the example in previous section may potentially be sub-optimal, how could sex compete? Sexual reproduction as a rule induces significant costs relative to asexual reproduction. If, in addition, sexual reproduction is also less efficient in long-term fitness, one would suppose that a majority of biological species would be using the asexual reproductive strategy, and yet exactly the opposite is true.

Corollary 1, however, relates to a model that includes the two assumptions of an infinite population and initial full genotypic population support. We show in this section that removing these assumptions significantly changes the picture.

Even if one interprets the infinite population model as referring to a ‘sufficiently large’ finite population, in most cases only a ‘relatively small’ finite population model can be realistic. For most species, |Γ| is a far bigger number than the number of organisms that have ever existed in the species. Every generation can only contain a tiny fraction of the immense collection of all possible genotypes.

If the allelic frequency dynamic cannot be implemented over the entire genotype space Γ, the next best thing is to do what election polling firms do when they cannot poll an entire population, namely, use of a relatively small random sample. In this view, the genotypic population $X^t$ in each generation of a finite sexually reproducing species is in effect a random sample of the entire genotype space.\(^2\)

\(^2\) That is, small relative to the full genotype space.

Some care should be taken with regard to this metaphor. In election polls each random sample ideally satisfies the independent identical distributions conditions. In contrast, the genotype space samplings of sexual reproduction is not independently distributed since the probability of the formation of a particular genotype at time $t$ does depend on the genotypic frequency at time $t-1$.\(^3\)

\(^3\) That is, small relative to the full genotype space.
**Figure 1. Learning From a Sparse Population Sample.**

- **a.** Within sparse samples of genotype space, asex attains the local maximum viability, represented by a red circle. In contrast, sexual reproduction reliably attains the viability of an optimal sexual genotype (represented by a green circle), no matter where the initial sample lies within the space.

- **b.** A histogram of a random sample of the initial population’s viability from one of our simulations. The initial population represents only a small fraction of all possible genotypes, hence not all possible viabilities are represented. The asexual population will eventually consist nearly exclusively of individuals whose viability is maximal out of the initial sample. In the sexual population, genetic mixture will produce new levels of viability, quickly covering a large portion of the viability spectrum, as can be seen in the attached films of the simulations.

Even when the population size is rather large in terms of population sizes encountered in nature, it will be extremely small compared to the vastly immense space of all genotypes. The size of the population can, however, have an effect on the dynamic; all other things being equal, a smaller population will likely exhibit worse approximation to the ideal infinite population values in any given generation than a larger population.\(^4\)

In the infinite population model we supposed a founding population with at least one individual bearing each possible genotype. Since we want the finite population to represent only a small sample of the genotype space, the founding finite population cannot meet that standard. Instead we require only that the potential of eventually creating every possible genotype by recombination to be inherent in the founding population. Formally,

\(^4\) Compare with Barton and Paixão (2016), where population size has effects on population traits.
this translates into requiring that for each distinct allele there be at least one individual in the finite founding population with a genotype bearing that allele. This dictates a certain minimal population size: denoting the number of distinct alleles in locus \( j \) by \( |L_j| \), the population size must be at least \( \max(\{|L_1|, |L_2|, \ldots, |L_\ell|\}) \) (the proof of this can be found in Lemma 1 in the appendix).

For the finite population case, we model a finite adult population at time \( t \), \( \hat{X}^t \), as a sampling of \( \Gamma \). The individuals in the population mate to produce offspring population \( \tilde{X}^t \). Denote by \( \hat{q}^t_i \) the fraction of the offspring at \( t \) bearing allele \( i \), by \( \hat{\phi}^t_i \) the mean viability and \( \hat{\rho}^t_i \) the portfolio growth. The dynamics proceed as before, but in interpretation \( \hat{\phi}^t_i \) and \( \hat{\rho}^t_i \) are estimators of \( \phi^t_i \) and \( \rho^t_i \) that would have been attained if we were following an infinite population in which every genotype is represented.

The sampling metaphor is key to our results here. It is the randomness of genetic shuffling that enables \( \hat{\rho}^t_{j,i} \) to function as an unbiased estimator, i.e., \( E(\hat{\rho}^t_{j,i}) = \rho^t_{j,i} \).

**Proposition 1.** If the Hedge algorithm (Algorithm 1 above) is implemented by sampling in each time period, with the unbiased estimator of the reward vector \( \hat{r}^t \) replacing the true reward vector \( r^t \) in each time period beginning with an initial sample of size \( k_1 \), then the long-term weight vector \( \hat{q}^T \) resulting from using the estimator closely approximates the true long-term weight vector \( q^T \), with the approximation rapidly improving with increasing initial sample sizes.

An outline of the proof of Proposition 1 is as follows. In Algorithm 1, the main dynamic variable that is being tracked is the weight vector \( q^t \), with that dynamic ultimately determined by the value of the reward vector \( r^t \) at every time \( t \). The reward vector \( r^t \) can be thought of as the return on investment in the ‘portfolio’ of relative holdings over the elements of the genotype space.

Note that even in the infinite population case, both \( q^t \) and \( r^t \) are finite \( k \)-length vectors, not infinite vectors (in fact, all the algorithms presented in this paper make use throughout only of finite quantities). The only difference between the infinite and finite population situations is that in the infinite population case the portfolio held at each time \( t \) will have full support over the genotype space; in the finite case the portfolio may be very ‘sparse’, with only a fraction of the genotype space having non-zero representation in the portfolio. If despite this the finite case reward vector \( \hat{r}^t \) is an unbiased estimator of \( r^t \) (e.g., by virtue of random sampling), then \( \hat{r}^t \) will fairly faithfully follow \( r^t \) over time. One can then apply concentration of measure methods to show that the dynamic of \( \hat{q}^T \) closely approximates the true \( q^T \) as the size of the initial sample increases.

The full details appear in the appendix in Lemma 2. There it is shown that in general, whatever the size of the initial population, the subsequent process will tend to two possible types of dynamics: one in which the population size is forever small and bounded and one
in which it grows exponentially. The probability of the population becoming trapped in
size limiting dynamic, however, is very small; the probability of the alternative, of pop-
ulation size growing rapidly, is large and increases with the size of the initial population.
Once the population size becomes sufficiently great at some point in time, the effect of the
initial population size on the outcome of the process becomes negligible.

**Theorem 3.** *In the finite population model considered as sampling from gentoype space,
the cumulative growth of a sexually reproductive population asymptotically approximates
the cumulative growth of a population consisting solely of individuals bearing one of the
optimal-in-hindsight sexual genotypes.*

It ought to be kept in mind that Theorem 3 holds true under ideal conditions in which
there is no genetic drift caused by the irrevocable loss of an allele from some generation
onwards due to random probabilistic effects (an event that would then prevent some pos-
sible genotypes from ever being sampled again). The chances of this happening can be
significantly mitigated by supposing that \( v^t_g \) is positive for all \( g \) and all \( t \), which is not a
very strong assumption, but even this is not sufficient to guarantee that genetic drift will
never occur.

The statement of Theorem 3 holds the key to our explanation for why sex prevails as a
reproductive strategy. Consider a founding population containing a small sample \( \Omega_0 \subset \Gamma \)
of distinct genotypes (e.g., a founding population with five distinct genotypes colonising
a new area, with genotype viabilities 0.5071, 0.5361, 0.5529, 0.5816, 0.6702). If this pop-
ulation reproduces asexually, the asymptotic viability it will attain is that of the genotype
\( g_0 \in \Omega_0 \) of maximal viability (e.g., 0.6702). Until a beneficial mutation occurs, such
a population’s viability is bounded from above by the viability of \( g_0 \). Its walk through
genotype space is limited by the bounds of the original sampling \( \Omega_0 \).

In contrast, if that same population, founded by \( \Omega_0 \), reproduces sexually it will execute
an algorithmically goal-directed walk through \( \Gamma \), asymptotically attaining the viability of
\( g^S \), an optimal-in-hindsight sexual genotype. If the asymptotic viability of \( g^S \) more often
than not is greater than that of \( g_0 \), then sex has an evolutionary advantage over asex. In
technical terms, asex converges to the global optimum of a small sub-tensor of the viability
tensor, while sex converges to one of the local optima of the entire tensor spanned by the
support of all alleles, and the latter is usually much better (see Fig. 1).

In our model, selection and randomness in sexual reproduction do not work at cross-
purposes but reinforce each other. Directionality in the search through genotype space
is due to the implementation of a multi-locus no regret algorithm. At the same time,
randomness is needed to ensure an unbiased estimation of the ‘ideal’ values by way of
sampling, yielding not an entirely random walk but a directed walk through genotype
space underpinned by randomness.
5. Simulations

To test the claims of this paper, we conducted two different types of simulations: simulation of competition between sexually and asexually reproducing populations, and simulation comparing the growth rates of a finite sexually reproducing population starting with small samples with the comparable rate of an idealised infinite population.

5.1. Sexually Reproducing versus Asexually Reproducing Populations. The simulations were distinguished in numbers of loci, numbers of alleles per locus and epistatic interactions (see appendix). They all, however, followed the format of two founding populations with the same initial genotypes, one population sexual and the other asexual.

These populations were tracked through some maximal number of generations, with the mean population viabilities in each generation recorded. At the end of each such ‘trial’, the population with higher mean viability was declared the winner of the trial. In each simulation 1000 trials were conducted. Under all conditions, sexual populations exhibited...
Figure 3. Viability Histograms. a. A histogram of the distribution of viabilities in a simulated sexual population after several hundred generations. The histogram approximates a normal curve with a mean viability that is higher than that of the parallel asexual population. The dashed red line is the maximal viability level within the initial sample. b. A histogram of the distribution of viabilities in a simulated asexual population after several hundred generations. Virtually all of the population bears genotypes with the same level of viability, coinciding with the maximal viability level within the initial sample given by the red dashed line.

clear advantages over asexual populations (see Fig. 2). The percentage of trials in which sex outstripped asex in viability ranged from 71.7% to 92.3%, giving sex an overwhelming edge.
One of the significant aspects of the simulations was the consistency and reliability exhibited by the sexual no-regret learning algorithm. Figure 4 presents representative data from trials held under different conditions. The graph signatures are consistently similar. They virtually all feature a rapid asexual mean viability plateau. The sexual population is initially at a disadvantage but as the learning aspect of the no-regret algorithm proceeds, the mean values steadily climb, most often past the maximal viability of the asexual population.

We consider the fact that our simulations yielded clear results in favour of the sexual reproductive strategy even with relatively small population sizes to be indicative of the robustness of our mathematical model—one would expect the sexual population to follow the ideal theoretical estimator even more closely with larger population samples.

5.2. Finite versus Infinite Sexual Populations. In these simulations, two sexually reproducing populations consisting of 10 loci and 2 alleles per locus were compared in identical environments. Simulating an actual infinite population was, for fairly obvious reasons, not possible for us. Instead we compared a population with a very small founding population (16 or 32 individuals) with a population starting out with 5000 individuals, with each genotype represented within that large founding population. The goal was to compare the long-term mean viability rates attained by very small initial samples with the corresponding rates attained by a population that initially does not sample the genotype space (as it covers the entire genotype space at the beginning).

In total, 100 simulation trials in which the founding population consisted of 16 individuals were conducted, alongside 100 simulation trials of founding populations of 32 individuals. In each simulation run, the mean viability attained by a parallel population of 5000 individuals served as a baseline against which to compare the performance of the population beginning with far smaller samples of the genotype space.

The simulations indicated that a larger initial population yielded more accurate results, as would be expected (see Figure 5). The populations founded by 32 individuals attained outcomes that were worse than the 5000-founded populations by 2% or more in 18% of the simulation trials, while populations founded by 16 individuals were off by 2% or more 24% of the time.

Interestingly, the small sample populations in the simulations also sometimes exhibited better long-term mean viability results than the large sample populations. This might be explained in several ways. There could be a founder effect being exhibited, in which an initial unequal weighting of alleles in the founding population has long-term effects on allelic frequency. In addition, genetic drift, in which some alleles are lost forever due to random stochastic events, or more general drift effects, may lead to to different long-term outcomes.
SEX WITH NO REGRETS: HOW SEXUAL REPRODUCTION USES A NO REGRET LEARNING ALGORITHM FOR EVOLUTIONARY ADVANTAGE

(A) Simple average epistasis, 200 Loci

(B) Simple average epistasis, 1200 Loci

(c) CES epistasis, 200 loci

**Figure 4.** Representative trial runs under different epistatic assumptions. The dotted red lines represent the mean asexual population viability average, with the continuous blue lines representing the mean sexual viability average. The graph signatures of the trial runs are extremely consistent. The asexual population rapidly reaches a plateau; the sexual population is initially at a disadvantage but at a certain point begins to ‘learn the landscape’, increasing its mean viability and overtaking the asexual population.
Overall, however, the simulation results here were robust: in a clear majority of simulation runs (75% or more) the small sample founding populations attained long-term viability rates that were either equal to or better than those of a large founding population, or worse by at most 2%.

6. Conclusions

The concept of evolution as fundamentally being a process of learning is a relatively new idea that is entering the field in parallel to an increasing awareness of the centrality of learning algorithms in the study of computer science, game theory, economics, finance, and many other disciplines. Learning complements and extends previous paradigms that were focussed on optimisation. It takes an inherently more dynamic view, not necessarily assuming fixed probability distributions as the backdrop against which fixed optima can be identified and attained. Instead, learning models seek dynamic processes that have the flexibility to contend with any stochastic environment they encounter, including inherently unpredictable environments, by taking simple reactive steps that cumulatively add up to desirable outcomes.

This approach seems especially well fitted for the study of evolution (Barton, Novak, and Paixão (2014)). Fitness can only coherently be defined with respect to an environment. One aspect of the environment is its changeability. Biological entities generally lack the capacity to predict ahead of time which environments they will encounter, or how those environments will change over time (which may be entirely unpredictable by any methods). Being able to react in simple steps that nevertheless can bring positive results under nearly all conditions should in theory yield evolutionary advantages. Several previously suggested models have connected sexual reproduction to unstable environments but provided only partial explanations, as they could not explain asexual reproduction in unstable environments or sexual reproduction in stable ones.

Important contributions to the literature on learning in evolutionary processes, especially with regards to the evolution of sex, appear in several recent papers (Livnat et al (2008); Chastain et al (2014)). Those papers, as this one, consider the question of why sex breaks down favorable combination of genes (Eshel and Feldman (1970); Lewotin (1971)), which is a problem if one assumes that sex should facilitate an increase in population mean fitness (Karlin (1972, 1974)). The answer to this that was proposed by Livnat et al (2008) is that ‘the role of sex is to enable selection not for highly favorable specific combinations of genes but for genes that are favorable in many different combinations’, which in turn, as suggested by Chastain et al (2014), points to a trade-off between increasing cumulative fitness and entropy that may helps us understand the role of sex in evolution.

In response to Chastain et al (2014), Barton, Novak, and Paixão (2014) point out that the same learning algorithms can apply also to asexual populations, rendering it unclear
SEX WITH NO REGRETS: HOW SEXUAL REPRODUCTION USES A NO REGRET LEARNING ALGORITHM FOR EVOLUTIONARY ADVANTAGE

Figure 5. Results of 100 simulation trial runs for a founding population of 16 individuals (A) and a founding population of 32 (B), comparing long-term mean viability against the baseline of a population initially containing 5000 individuals, in a model of genotypes consisting of 10 loci and 2 alleles per locus.
as to how learning can be understood as granting an advantage to sex over asex. That is exactly what we address here. Our answer is, in effect, that the randomness of sexual reproduction that some have considered a vexing puzzle (Barton (2009)) is ‘not a bug but a feature’. That is, the randomness is an integral element of the learning by sampling conducted by sexual populations. While both asexual and sexual populations implement learning algorithms, the learning by sampling eventually enables the sexual population to attain greater growth rates.

Our simulations bore this out with a surprising reliability. In a variety of randomly selected environments and a range of epistatic conditions, the asexual populations typically homed in on the local maximum extremely quickly. The sexual populations had a more difficult up-hill climb with their slower learning method, but their directed search through genotype space eventually paid off, in fitness terms, very well.

As pointed out in Barton, Novak, and Paixão (2014), studying evolutionary and natural selection by way of learning algorithms raises the question of how rapidly evolutionary learning can occur under various conditions. How does the run-time scale with the dimensionality of the environment, genome size, population size, and so on? Our simulations indicate that in general, the larger the genome size, the more time is required for efficient learning, which is not surprising as increasing the genome translates into a larger genotype space that is to be searched. There is still much to be studied here, in both quantitative and qualitative terms.

Future studies are likely to contribute to a better understanding of the results of the model presented here. One possible future direction is to consider the changes in lifetime reproductive output, marked as ζ in this work. Adding stochasticity to ζ or to the effect of genotypes on reproductive rates, in addition to varying differential rates of ζ in both sexual and asexual populations over time, may yield new insights on changes in the population fitness.

Finally, we emphasise once again that our model in this paper does not include mutations, despite the paramount importance of mutations in evolution. In the context of the model here, mutations expand the genotype space (by adding novel alleles to the set $A$). Here, learning occurs in the period of reproductive generations between the emergence of beneficial mutations. Introducing mutations into the present model is an important future research direction.

7. Appendix: Methods and Proofs

7.1. Methods. Three different models were used for the simulations, which were conducted in Matlab code.
7.1.1. Simulation 1 - Simple Averaging with 200 Loci. In this simulation, we worked with a genotype space consisting of genotypes with 200 loci and 4 alleles per locus.

For each iteration, which may be termed a ‘trial’, a selection of five distinct genotypes from this space were selected at random. From this, two founding populations, one asexual and the other sexual, were formed. Each founding population consisted of 100 individuals, with each individual bearing one of the five previously selected genotypes such that 20 individuals bore each genotype. Mating between pairs of individuals in each generation in the sexual population was by random matching.

Every possible genotype in the genotype space was assigned a viability value between 0 and 1 in the following manner: the 200 loci were partitioned into 20 ‘epistatic units’ of 10 loci each. Each epistatic unit was associated with a randomly selected viability function that assigned a value between 0 and 1 to each of the $4^{10}$ possible combinations of the four alleles at each of the ten locus locations comprising the unit (this is the most general form of epistatic interaction between those loci). The viability value of the full 200-locus genotype was then calculated as the simple average of the viabilities of each of the 20 epistatic units.

Each trial was run up to a maximal number of 250 generations. The maximal permitted population size was 5,000 individuals; if a population exceeded this size in a particular generation it was reduced by way of half of the individuals in the population being killed off randomly. The half that survived this procedure then proceeded to mate and produce the next generation.

In each generation, the mean viability of each population was recorded. At the end of each trial, the population with higher mean viability was declared the winner of the trial (a trial was halted if the sexual population attained higher mean viability than the asexual population at any point in time, to conserve computational resources). The simulation was comprised of 1000 trials.

7.1.2. Simulation 2 - Simple Averaging with 1200 Loci. The details of this simulation model are nearly identical to those of the previous model, except that: genotypes here contained 1200 loci, the number of alleles per locus was 3, the maximal number of generations per simulation run was 500, the maximal population size was 20,000, and the number of epistatic units was 120. The simulation was again comprised of 1000 trials.

7.1.3. Simulation 3 - CES with 200 Loci. In this simulation model, genotypes contained 200 loci, the number of alleles per locus was 3, the maximal number of generations per simulation run was 350, the maximal population size was 15,000, and the number of epistatic units was 20.

The major difference between the simulation model here and the previous ones is in the calculation of the viability of a genotype. As before, the viability value of an epistatic unit
is given by a randomly chosen function assigning a value between 0 and 1 to strings of alleles of length 10. However, the total viability of a genotype is now determined by a formula aggregating the viability values using a constant elasticity of substitution formula (a concept borrowed from economics). In application, a power $0 < r < 1$ and a set of coefficients $\{a_1, a_2, \ldots, a_{20}\}$ (one per epistatic unit) satisfying $\sum_{i=1}^{20} a_i = 1$ are chosen at random, and the viability of a genotype is given by $\sum_{i=1}^{20} (a_i v_i)^{1/r}$, where $v_i$ is the viability value of the $i$-th epistatic unit.

The simulation was comprised of 1000 trials. In each trial a different viability function for the epistatic units was randomly selected. The constant elasticity of substitution power and coefficients were held constant for each block of five trials, and then selected randomly for the next block of five.

7.1.4. Simulation 4 - Finite versus Infinite Populations. The underlying code used for this simulation was nearly identical to the code used in the previously listed simulations. In the simulation model here, genotypes contained 10 loci, the number of alleles per locus was 2, and the maximal number of generations was 250 per simulation run.

The simulation was conducted twice, each time with 100 trials. In the first simulation, in each trial, the long-term mean viability of a population with a founding population of 5000 individuals, covering the entire genotype space, was recorded, to serve as a baseline of comparison. This was then compared against the long-term mean viability of a population founded by a small sample of 16 individuals, under identical conditions.

In the second simulation, in each trial, the long-term mean viability of a population founded by a small sample of 32 individuals was similarly compared against the baseline of a founding population of 5000 individuals.

7.2. Proofs of Propositions and Theorems.

**Lemma 1.** Let $A$ be a set of alleles, partitioned into loci $\{L_1, \ldots, L_\ell\}$. Denoting the number of distinct alleles in locus $j$ by $|L_j|$, the minimal size of a population satisfying the condition that for each allele $\bar{a}$ there exists at least one individual in the population bearing a genotype $g$ such that $g$ contains allele $\bar{a}$ is $m := \max(|L_1|, |L_2|, \ldots, |L_\ell|)$.

**Proof.** Suppose first that the number of alleles per locus is uniformly $m$.

Consider the $\ell \times m$ matrix $A := (a_{ji})$, in which matrix element $a_{ji}$ represents the $i$-th possible allele in locus $j$. For each $1 \leq i \leq m$, use column $i$ of $A$ to form an individual $i$ bearing genotype $g_i = a_{1i} a_{2i} \ldots a_{\ell i}$. There are $m$ individuals in the resulting population and each distinct allele is borne by an individual in this population. Clearly, any population of size smaller than $m$ will necessarily omit at least one allele.

If the number of alleles per locus is not uniform but bound above by $m$, form a similar matrix, with repetitions of elements now allowed across rows. Then once again, the $m$
columns of the matrix enable one to define a population of \( m \) individuals such that each distinct allele is borne by at least one individual in the population.

**Proof of Theorem 1.** As Algorithm 2, an instantiation of the Hedge algorithm, is being implemented at each locus, the asymptotic average regret between the cumulative logarithmic population growth achieved under \( \Phi_S \) and the corresponding growth attained in the counter-factual situation in which at locus 1 the optimal-in-hindsight allele \( a_1^* \) (with respect to the dynamic at the other loci) alone is fixed while Algorithm 2 is implemented at all other loci, \( \lim_{T \to \infty} \frac{1}{T} (\Phi_S - \Phi_{a_1^*}) \), vanishes.

It follows that the asymptotic mean growth rate of \( \Psi_S^T \) converges to that of \( \Psi_{a_1^*} \). Next we suppose that \( a_1^* \) is fixed at locus 1 while the other loci follow Algorithm 2. By the same reasoning, this time concentrating on fixing the optimal allele at locus 2 w.r.t. to the dynamic at the other loci, we can conclude that the asymptotic mean growth rate of \( \Psi_{a_1^*a_2^*}^T \) converges to that of \( \Psi_{a_1^*} \), and hence also to that of \( \Psi_S^T \).

Continuing inductively to all loci, the conclusion is that the expected asymptotic mean growth rate of \( \Psi_S^T \) converges to that of \( \Psi_{a_1^*a_2^*...a_\ell^*} \), which is the same as \( \Psi_{g^*} \).

**Proof of Theorem 2.** The proof follows the same reasoning as that of Theorem 1, namely applying the no-regret result of the Hedge algorithm represented by Algorithm 3, but it is even simpler here since there is no need to keep track of multiple loci. The result is straightforward by construction of the algorithm.

**Proof of Corollary 1.** Given Theorems 1 and 2, the statement here is tantamount to showing that

\[
\Phi_{g^A} \geq \Phi_{g^S},
\]

where \( g^A \) is the optimal-in-hindsight asexual genotype and \( g^S \) is the optimal-in-hindsight sexual genotype.

The weak inequality in Equation (13) follows trivially from the fact that \( g^A \) is by definition an optimal growth genotype in \( \Gamma \), hence its cumulative growth must be greater than or equal to that of \( g^S \).

**Lemma 2.** If the Hedge algorithm (Algorithm 1 above) is implemented by sampling in each time period, with the unbiased estimator of the reward vector \( \hat{r}_t \) replacing the true reward vector \( r_t \) in each time period beginning with an initial sample of size \( k_1 \), then the long-term weight vector \( \hat{q}^T \) resulting from using the estimator converges almost surely to the true long-term weight vector \( q^T \) as \( T \to \infty \), with the convergence uniform in the initial population size \( k_1 \).

**Proof of Lemma 2.** The presentation of Algorithm 1, for expositional clarity, divides the updating of the weight vector \( q^T \) into two steps, first calculating \( \hat{q}^{t+1}_i = q^t_i \cdot e^{\eta r^t_i} \), and then normalising into a distribution by setting \( q^{t+1}_i = \frac{\hat{q}^{t+1}_i}{\sum_j \hat{q}^{t+1}_j} \). This can be reduced to
one step by multiplying each $r_t^i$ by a factor $\beta_t$ that ensures that $q_t^{i+1} = q_t^i e^{\eta T \beta_t r_t^i}$. We will suppose here w.l.o.g. that $r_t^i$ already incorporates this normalisation factor and therefore that $q_t^{i+1} = q_t^i e^{\eta T r_t^i}$ for each $t$.

It follows that the long-term weight vector element $q_i^T$ is determined as

$$q_i^T = q_1^i \cdot e^{\eta T \sum_{t=1}^T r_t^i}.$$  \hspace{1cm} (14)

For a sample of size $k_t$ in generation $t$, we implement the algorithm with the reward $\circ r_t^i$, which is given as the average of $k_t$ independent random variables $z_{t,1}^j, \ldots, z_{t,k_t}^j$ with expectation $r_t^i$, namely $\circ r_t^i = \frac{1}{k_t} \sum_{j=1}^{k_t} z_{t,j}^j$. The resulting random variable $\circ r_t^i$ is an unbiased estimator of $r_t^i$, and the long-term weight vector element $\circ q_i^T$ is then determined as

$$\circ q_i^T = q_1^i \cdot e^{\eta T \sum_{t=1}^T \circ r_t^i}.$$  \hspace{1cm} (15)

Dividing Equation (14) by Equation (15), the long-term ratio between $q_i^T$ and $\circ q_i^T$ as $T$ increases comes down to the difference $\sum_{t=1}^T (r_t^i - \circ r_t^i)$.

Let

$$X_t^i = \circ r_t^i - r_t^i = \frac{1}{k_t} \sum_{j=1}^{k_t} (z_{t,j}^j - r_t^i).$$  \hspace{1cm} (16)

Let $\tau(n) = \max\{\tau : \sum_{t=1}^\tau k_t < n, \tau \geq 1\}$, and $S(n) = \sum_{t=1}^{\tau(n)} k_t$, and define a sequence of random variables $\{Y^n\}_{n \geq 1}$ by

$$Y^n = \sum_{t=\lceil \sqrt{T} \rceil}^{\tau(n)} \sum_{j=1}^{k_t} \frac{z_{t,j}^j - r_t^i}{k_t} + \sum_{j=1}^{n-S(n)} \frac{z_{\tau(n)+1}^j - r_{\tau(n)+1}^i}{k_{\tau(n)+1}}.$$  \hspace{1cm} (17)

Notice that as $z_{t,j}^j$ is an unbiased estimator of $r_t^i$ for each $j$ and $t$, it follows that $E(z_{t,j}^j) = r_t^i$ and hence $E(Y^n_{\circ \tau(n)}, z_{1,\circ \tau(n)}, \ldots, z_{k_t,\circ \tau(n)})$ is a martingale with bounded differences.

In addition, for each $T$ define

$$Y_T^0 = \sum_{t=1}^{\lfloor \sqrt{T} \rfloor-1} \sum_{j=1}^{k_t} \frac{z_{t,j}^j - r_t^i}{k_t} + \sum_{j=1}^{n-S(n)} \frac{z_{\tau(n)+1}^j - r_{\tau(n)+1}^i}{k_{\tau(n)+1}}.$$  \hspace{1cm} (18)
Next, denote by $m > 1$ the average growth rate of the population. Let $s$ be arbitrarily chosen for now. For each $s$, $\varepsilon > 0$, and $\delta > 0$ with $\frac{m}{1+\delta} > 1$ we have

$$P \left( e^{s|Y_p^0 + Y^{S(T)}|} > e^{s\varepsilon} \right)$$

$$= P \left( k_{\sqrt{T}} < \left( \frac{m}{1+\delta} \right)^{\sqrt{T}} k_1 \right) P \left( e^{s|Y_p^0 + Y^{S(T)}|} > e^{s\varepsilon} \left| k_{\sqrt{T}} \right| < \left( \frac{m}{1+\delta} \right)^{\sqrt{T}} k_1 \right)$$

$$+ P \left( k_{\sqrt{T}} \geq \left( \frac{m}{1+\delta} \right)^{\sqrt{T}} k_1 \right) P \left( e^{s|Y_p^0 + Y^{S(T)}|} > e^{s\varepsilon} \left| k_{\sqrt{T}} \right| \geq \left( \frac{m}{1+\delta} \right)^{\sqrt{T}} k_1 \right)$$

(19)

$$\leq P \left( k_{\sqrt{T}} < \left( \frac{m}{1+\delta} \right)^{\sqrt{T}} k_1 \right) + e^{-s\varepsilon} E \left( e^{s|Y_p^0 + Y^{S(T)}|} \left| k_{\sqrt{T}} \geq \left( \frac{m}{1+\delta} \right)^{\sqrt{T}} k_1 \right| \right)$$

Now, Equation (19) is the same as

$$P \left( k_{\sqrt{T}} < \left( \frac{m}{1+\delta} \right)^{\sqrt{T}} k_1 \right) +$$

$$e^{-s\varepsilon} E \left[ E \left( e^{s|Y_p^0 + Y^{S(T)}| Y^0_T, Y^{S(T)-1}_T, k_{T-1}, z_{T-1}^1, \ldots, z_{T-1}^{k_{T-1}}, k_{\sqrt{T}} \geq \left( \frac{m}{1+\delta} \right)^{\sqrt{T}} k_1 \right) \right]$$

$$\leq P \left( k_{\sqrt{T}} < \left( \frac{m}{1+\delta} \right)^{\sqrt{T}} k_1 \right) +$$

$$e^{-s\varepsilon} E \left[ E \left( e^{s \frac{|z_{T-1}^1 - T_{k_{T-1}}|}{k_T}} Y^0_T, Y^{S(T)-1}_T, k_{T-1}, z_{T-1}^1, \ldots, z_{T-1}^{k_{T-1}}, Y^0_T, k_{\sqrt{T}} \geq \left( \frac{m}{1+\delta} \right)^{\sqrt{T}} k_1 \right) \right]$$

$$= P \left( k_{\sqrt{T}} < \left( \frac{m}{1+\delta} \right)^{\sqrt{T}} k_1 \right) +$$

$$e^{-s\varepsilon} E \left[ e^{s|Y_p^0 + Y^{S(T)-1}|} E \left( e^{s \frac{|z_{T-1}^1 - T_{k_{T-1}}|}{k_T}} Y^0_T, Y^{S(T)-1}_T, k_{T-1}, z_{T-1}^1, \ldots, z_{T-1}^{k_{T-1}}, Y^0_T, k_{\sqrt{T}} \geq \left( \frac{m}{1+\delta} \right)^{\sqrt{T}} k_1 \right) \right]$$

$$\leq P \left( k_{\sqrt{T}} < \left( \frac{m}{1+\delta} \right)^{\sqrt{T}} k_1 \right) + e^{-s\varepsilon} E \left( e^{s|Y_p^0 + Y^{S(T)-1}|} e^{s \frac{|z_{T-1}^1 - T_{k_{T-1}}|}{k_T}} k_{\sqrt{T}} \geq \left( \frac{m}{1+\delta} \right)^{\sqrt{T}} k_1 \right).$$
Repeating recursively we obtain

\[
P \left( e^{s|Y^{S(T)}|} > e^{s\epsilon} \right) \leq \frac{e^{-s\epsilon}}{P \left( k_{\sqrt{T}} < \left( \frac{m}{1 + \delta} \right)^{\sqrt{T}} k_1 \right)} + e^{-s\epsilon} \mathbb{E} \left( e^{s|Y^{0}| + \sum_{t=\sqrt{T}+1}^{T} \sum_{j=1}^{k_{T}} \frac{|z_{j}^t - r_{j}^t|^2}{k_{T}^2} } | k_{\sqrt{T}} \geq \left( \frac{m}{1 + \delta} \right)^{\sqrt{T}} k_1 \right).
\]

Recalling that \( ||z_{j}^t - r_{j}^t||_{\infty} \leq 1 \), deduce that

\[
P \left( e^{s|Y^{0}_{T} + Y^{S(T)}|} > e^{s\epsilon} \right) \leq P(k_{\sqrt{T}} < \left( \frac{m}{1 + \delta} \right)^{\sqrt{T}} k_1) +
\]

\[
P \left( |Y^{0}_{T} + Y^{S(T)}| > \epsilon \right) \leq P \left( k_{\sqrt{T}} < \left( \frac{m}{1 + \delta} \right)^{\sqrt{T}} k_1 \right) +
\]

\[
E \left( \frac{e^{-\epsilon - |Y^{0}_{T}|^2 k_{\sqrt{T}}}}{2 \left( \frac{m}{1 + \delta} \right)^{\sqrt{T}} k_1} \prod_{t=\sqrt{T}+1}^{T} \frac{1}{\sum_{j=1}^{k_{T}} z_{j}^t r_{j}^t} \right) | k_{\sqrt{T}} \geq \left( \frac{m}{1 + \delta} \right)^{\sqrt{T}} k_1 \right) \leq
\]

\[
P \left( k_{\sqrt{T}} < \left( \frac{m}{1 + \delta} \right)^{\sqrt{T}} k_1 \right) +
\]

\[
P \left( \left( \frac{m}{1 + \delta} \right)^{\sqrt{T}} k_1 \right) +
\]

\[
E \left( \frac{e^{-\epsilon - |Y^{0}_{T}|^2}}{2 \left( \frac{m}{1 + \delta} \right)^{\sqrt{T}} k_1} \prod_{t=\sqrt{T}+1}^{T} \frac{1}{\sum_{j=1}^{k_{T}} z_{j}^t r_{j}^t} \right) | k_{\sqrt{T}} \geq \left( \frac{m}{1 + \delta} \right)^{\sqrt{T}} k_1 \right) \right).
\]
SEX WITH NO REGRETS: HOW SEXUAL REPRODUCTION USES A NO REGRET LEARNING ALGORITHM FOR EVOLUTIONARY ADVANTAGE

Let $g_\varepsilon = \inf_T (|Y^0_T| - \varepsilon)^2$. We can\footnote{First choose positive constants $c, \varepsilon''$ with $\varepsilon'' \pm \sqrt{c} \in (\varepsilon/2, \varepsilon)$. Let $\delta = \text{sign} \inf_T (Y^0_T - \varepsilon'')$. Then $g_{\varepsilon''+\sqrt{\delta}} = \inf_T ((|Y^0_T| - \varepsilon'')^2 + 2\sqrt{\delta}(|Y^0_T| - \varepsilon'') + c) \geq c + g_{\varepsilon''} + 2\sqrt{\delta} \inf_T (\delta(|Y^0_T| - \varepsilon'')) \geq g_{\varepsilon''} + c \geq c$. Now set $\varepsilon' = \varepsilon'' + \sqrt{\delta}.$} choose a random variable $\varepsilon'$ with values in $(\varepsilon/2, \varepsilon)$ such that $g_{\varepsilon'} \geq c$ almost surely, for a fixed constant $c > 0$. Thus

$$P \left( |Y^0_T + Y^S(T)| > \varepsilon \right) \leq P \left( |Y^0_T + Y^S(T)| > \varepsilon' \right)$$

$$\leq P \left( k_{[\sqrt{T}]} < \left( \frac{m}{1 + \delta} \right) \sqrt{T} k_1 \right) + E \left( e^{- \frac{g_{\varepsilon'} \left( \frac{m}{1 + \delta} \right)^2 k_1}{1 + \frac{\delta}{\sqrt{T}} T^{1/2} k_1}} \right).$$

Denote $g = g_{\varepsilon'}$ for convenience. Further denote by $A_{\varepsilon}$ the event that $|Y^S(T)| > \varepsilon$ infinitely often in $T$ for a given $k_1$, and by $A_{\varepsilon,T}$ the event $|Y^S(T)| > \varepsilon$ for a given $T$ and an initial sample size $k_1$. Note that $A_{\varepsilon} = \limsup_T A_{\varepsilon,T}$. Now, we wish to evaluate $\sum_{T=1}^{\infty} P(A_{\varepsilon,T})$. First notice that

$$E \left( \sum_{T=1}^{\infty} \left( e^{- \frac{g_{k_1}}{1 + \frac{\delta}{\sqrt{T}} T^{1/2}}} \right) \right).$$

$$= E \left( \sum_{T=1}^{\infty} \left( e^{- \frac{c \left( \frac{m}{1 + \delta} - 1 \right)^2 k_1}{1 + \frac{\delta}{\sqrt{T}} T^{1/2}}} \right) \right)^{T-1}$$

$$= E \left( \sum_{T=1}^{\infty} \left( e^{- \frac{c \left( \frac{m}{1 + \delta} - 1 \right)^2 k_1}{1 + \frac{\delta}{\sqrt{T}} T^{1/2}}} \right) \right)^{T-1}.$$
(25) \[ E \left( \sum_{T=1}^{\infty} \left( e^\left( \frac{c}{T+\frac{1}{T+\frac{1}{T+\cdots}}} \right) \right)^{T-1} \right), \]

with the equality in Equation (24) following from the stationarity of the process \( \frac{k_t}{k_1} \). As this is now a geometric progression with quotient < 1, we obtain

(26) \[ E \left( \sum_{T=1}^{\infty} \left( \frac{g_{k_1}}{2 \left( 1 + \frac{T}{\sum_{t=\lceil \sqrt{T} \rceil+1}^{\lceil \sqrt{T} \rceil} \frac{1}{k_t/k_1} \right)} \right) \right)^{\sqrt{T}}, \]

(27) \[ \leq E \left( \frac{1}{1 - e^\left( \frac{c}{T+\frac{1}{T+\frac{1}{T+\cdots}}} \right)} \right) \]

(28) \[ \leq E \left( \frac{1}{c^\left( \frac{m}{T+\frac{1}{T+\cdots}} - 1 \right) k_1} \right) = \frac{8}{c^\left( \frac{m}{k_t} - 1 \right) k_1} E \left( 1 + \sum_{t=2}^{\infty} \frac{1}{k_t/k_1} \right) < \infty. \]

For the second summand, as \( \ln \frac{k_t}{k_1} = \sum_{t=1}^{T} \ln \left( \frac{k_{t+1}}{k_t} \right) \) and as \( E \left( \ln \left( \frac{k_{t+1}}{k_t} \right) \right) \leq \ln E \left( \frac{k_{t+1}}{k_t} \right) = \ln m \), it follows that \( W^\tau = \sum_{t=1}^{T} \ln \left( \frac{k_{t+1}}{k_t} \right) \) is a super-martingale, with differences bounded by some number \( d \). Then by the Azuma-Hoeffding inequality

\[ P \left( W^{\lceil \sqrt{T} \rceil} < \lceil \sqrt{T} \rceil \delta \right) \]

\[ \leq e^{-\frac{\delta^2}{2d/\sqrt{T}}}, \]

\[ = \left( e^{-\frac{\delta^2}{2d}} \right)^{\sqrt{T}}. \]
Hence

\[
\sum_{T=1}^{\infty} P\left( W_{\lfloor \sqrt{T} \rfloor} \leq \lfloor \sqrt{T} \rfloor \delta \right) < \sum_{T=1}^{\infty} \left( e^{-\frac{\delta^2}{2T}} \right)^{\lfloor \sqrt{T} \rfloor} = \sum_{n=1}^{\infty} \left( 2n + 1 \right) \left( e^{-\frac{\delta^2}{2n}} \right)^n < \infty.
\]

We deduce that \( \sum_{T=1}^{\infty} P(A_{\epsilon,T}) < \infty \). The Borel-Cantelli Lemma now implies

(29) \[ P(\limsup_{T} A_{\epsilon,T}) = 0, \]

and therefore \( P(A_{\epsilon} = 0) \). As this is true for each \( \epsilon > 0 \), almost sure convergence as \( T \to \infty \) follows. As the convergences in Equations (22) and (25) can be made uniform in \( k_1 \), the almost sure convergence is uniform in \( k_1 \) as well.

**Proof of Proposition 1.** The Proposition follows almost immediately from the proof of Lemma 2, as \( k_1 \) grows without bound.

**Proof of Theorem 3.** By Theorem 1, the infinite population asymptotically attains the mean growth rate equal to that of \( g^S \), an optimal-in-hindsight sexual genotype, and does so by implementing the Hedge algorithm at each locus. The finite population also implements the Hedge algorithm, by using an unbiased estimator for the portfolio growth of each allele, facilitated by the randomness of recombination.

By Proposition 1, the finite population long-term allelic frequency that results from using estimators closely approximates that of the infinite population allelic frequency as the initial sample grows. The result follows.

**REFERENCES**


SEX WITH NO REGRETS: HOW SEXUAL REPRODUCTION USES A NO REGRET LEARNING ALGORITHM FOR EVOLUTIONARY ADVANTAGE


**Acknowledgements:** The authors thank Yannai Gonczarowski, Elad Hazan, Aviv Keren, Reshef Meir, Uzi Motro and the participants of the 2015 Manchester Workshop on Evolution, Economics and Computation, for helpful conversations and comments.