MAKING THE MOST OF POTENTIAL: POTENTIAL GAMES AND GENOTYPIC CONVERGENCE

OMER EDHAN, ZIV HELLMAN & ILAN NEHAMA

ABSTRACT. We consider genotypic convergence of populations and show that under fixed fitness asexual and haploid sexual populations attain monomorphic convergence (even under genetic linkage between loci) to basins of attraction with locally exponential convergence rates; the same convergence obtains in single locus diploid sexual reproduction but to polymorphic populations. Furthermore, we show that there is a unified underlying theory underlying these convergences: all of them can be interpreted as instantiations of players in a potential game implementing a multiplicative weights updating algorithm to converge to equilibrium, making use of the Baum–Eagon Theorem. To analyse varying environments, we introduce the concept of 'virtual convergence', under which, even if fixation is not attained, the population nevertheless achieves the fitness growth rate it would have had under convergence to an optimal genotype. Virtual convergence is attained by asexual, haploid sexual, and multi-locus diploid reproducing populations, even if environments vary arbitrarily. We also study conditions for true monomorphic convergence in asexually reproducing populations in varying environments.

1. INTRODUCTION

One of the central questions of evolutionary theory has long been identifying conditions for asymptotic convergence to fixation on a monomorphic population. The classical example of such a result is the simplest case of asexual reproduction without mutation (e.g., bacteria reproducing in a petri dish) in which a version of the fundamental theorem of natural selection obtains: the mean fitness of the

The authors thank Daniel B. Weissman for suggesting corrections to an earlier version of this paper. Ziv Hellman and Ilan Nehama acknowledge research support by Israel Science Foundation grant 1626/18.

population, which follows the dynamic of the replicator equation, increases monotonically, leading to asymptotic fixation to a monomorphic population consisting of an optimal genotype with respect to the fitness environment.

Even this strong result, however, fails to hold once one considers arbitrarily varying fitness environments over time, even in asexually reproducing populations; in sexually reproducing populations the matter is more complicated still. In this paper we consider the general question of genotypic convergence of populations implementing various reproductive strategies under conditions of both fixed and varying environments. To this end we also introduce a concept that we term 'virtual convergence', applying ideas originally developed for the study of algorithms.

In greater detail, we consider here three discrete time population reproductive strategies: asexual, haploid sexual, and diploid sexual. The relevant state spaces for all of these is a polytope Θ. In the asexual case $\Theta = \Delta(\Gamma)$, the space of probability distributions over the set Γ of possible genotypes; what is of interest is tracing over time the relative frequency of the genotypes. In the sexual cases the focus instead is on the relative frequency in the population of alleles at each locus; if there are m loci with $k+1$ alleles per locus, the polytope of interest is $\Theta = \Delta_1^k \times \ldots \times \Delta_m^k$.

The dynamics considered will in general be describable by a transformation $T : \Theta \to \Theta$. That is, if the population is at state $\theta \in \Theta$ at time t, under the model it will be in state $T(\theta)$ at time $t+1$. The main matter studied is then the asymptotics of the trajectory defined by $T^n(\theta)$, starting from any θ, as n increases. If for each initial point $x \in \Theta$ there is a point $y \in \Theta$ such that $\lim_{n\to\infty} T^n(x) = y$ then the dynamic converges polymorphically; if y is a point distribution in Δ_i^k for each $1 \leq i \leq m$ then the convergence is monomorphic.

1.1. **Fixed Environments and Convergence in Potential Games.** The first question we consider asks which of these dynamics is guaranteed to converge, either monmorphically or polymorphically, when environments are fixed and unchanging over time, and we show that the asexual replicator dynamic, the sexual haploid dynamic – even under genetic inheritance linkage between loci – and the single-locus diploid dynamic all converge.

Furthermore, we provide a unified explanation for the convergence of all of these dynamics in the discrete time setting: all of them may be considered to be manifestations of potential games in

which the players monotonically increase the mean potential payoff by application of a multiplicative weights updating algorithm and exploiting the Baum–Eagon Theorem. When analysing discrete time dynamics, the standard tools of continuous time gradient climbing, which depend on partial derivatives, are not available. The Baum–Eagon inequality (see Section A), which was originally intended for application to the study of hidden Markov models, provides an alternative tool that has proven to be extremely relevant to analysing discrete time evolutionary models.

The Baum–Eagon Theorem is also an essential element in our proof of a general theorem that players in a potential game independently implementing polynomial multiplicative weights updating algorithms will asymptotically converge to a fixed point that is a Nash equilibrium. The surprising element of this theorem is that monotonic climbing of the mean potential payoff is attained even though there is no coordinating element to the updating of the players, each of whom updates based on the private information of the stage payoff received without explicitly taking into account the payoffs and updated distributions of the other players.

This is especially pertinent to our study of convergence in the sexual haploid model, where the dynamic can be described as an identical interests game being played by the loci, with the objective being identifying an optimal genotype; the theorem shows that the replicator dynamic conducted independently amongst the alleles at each locus, which is the essence of the sexual reproduction model, is guaranteed to converge. The exception to all this is the multi-locus diploid model under linkage disequilibrium, where the disequilibrium term prevents application of the Baum–Eagon inequality, and in fact it has long been known that convergence under that model is not guaranteed.

A further advantage of undergirding the fixed environment theorems by appeal to the Baum–Eagon Theorem is that it enables us to make use of theorems from [Baum and Sell, 1968] to obtain finer resolution insights into the dynamic paths followed by population along the way towards convergence. This includes the fact that surrounding each pure Nash equilibrium there exists a basin of attraction, and even more strongly a basin of attraction that is exponentially stable. This implies that an observer following a path through the state space (including that of any potential game in which the

players are implementing the polynomial multiplicative weights updating algorithm) will for a long time register relatively small increased in mean payoff until the path enters the exponental basin of attraction, at which point an acceleration will be noted with exponentially fast convergence to a fixed point of local maximal mean payoff.

1.2. **Varying Environments and Virtual Convergence.** When environments vary, sufficiently wildly varying environments from one time period to the next can make it impossible for the dynamic to converge to any single population state in Θ. To contend with this we introduce here a new concept of 'virtual convergence'. This is defined using tools borrowed from computer science and introduced to the population genetics literature in the past decade, which involve regret minimisation algorithms. The metaphor often used to describe this approach is that of selecting an action with respect to varying payoff functions in subsequent time periods based on advices offered by a collection of experts. The objective is attaining asymptotically the payoff that would have been achieved had one followed from the start the advice of the best expert in hindsight in every time period; a no regret algorithm achieves this objective.

In the evolutionary setting, the analogues of the experts of the previous paragraph are genotypes and the payoffs are fitness values. The question then becomes: is it the case that, no matter what sequence of environments and hence fitness values is realised, the reproducing population asymptotically attains the mean fitness that is the growth rate that it would have achieved had it been comprised from the start monomorphically by the optimal-in-hindsight genotype? If yes, then we say that virtual convergence is attained.

With these definitions, we study here modes of convergence for the asexual, haploid sexual, and diploid sexual reproduction, variously under independence of inheritance between loci as well as genetic linkage, fixed fitness and varying fitness conditions.[1] A summary of some of the results appears in Figure 1.

As can be seen in the summary, all of the reproduction models studied here attain virtual convergence, no matter how wildly environments vary. They attain this by exploiting the regret minimisation aspect of the multplicative weights updating algorithm.

[1] In all models in this paper generations are discrete and non-overlapping, populations are infinite, and no mutation, migration, or genetic drift is included in the models.

Reproduction	*Fixed fitness convergence*	*Varying fitness convergence*
Asexual	Monomorphic	Virtual
Haploid	Monomorphic	Virtual
Single Locus Diploid	Polymorphic	Virtual
Multi-Locus Diploid	Virtual	Virtual

Figure 1. Summary of modes of convergence.

In a sense, it can be said that the reproductive models are opportunistic: when environments are sufficiently well-behaved and they can take advantage of the payoff climbing afforded by the Baum–Eagon inequality, they will do so to converge to Nash equilibrium. If that is not available, then even under the worst possible conditions they will at least attain virtual convergence, minimising regret in hindsight.

1.3. **Non-arbitrarily Varying Environments.** The gap between fixed environments and entirely arbitrarily varying environments is large. The subject of convergence when environments vary in a structural way is explored here only with respect to the asexual replicator model, where we show that convergence to a monomorphic population is guaranteed under ergodically varying environments and under a broader property we introduce that we call one-step-ahead superiority.

There is much scope remaining to be researched in this topic, which we leave to future papers.

2. Basic Models and Notation

2.1. **Simplices.** For an integer m, Δ^m denotes the standard finite dimensional simplex over $m+1$ points. For a finite set Γ, $\Delta(\Gamma)$ denotes the collection of probability mass functions over the elements of Γ. We will denote the subset of $\Delta(\Gamma)$ consisting of distributions with support on one element of G alone by $\Delta_1(\Gamma)$, and the element of $\Delta_1(\Gamma)$ placing all support on $g \in \Gamma$ will be denoted by 1_g.

2.2. **Potential Games.** Let I be a finite set of m players. Associate with each player i a finite set of actions A_i. Denote $A = A_1 \times \ldots \times A_m$, and the cross product of all action sets except from i by A_{-i}. A game is defined by a payoff function $u : A \to \mathbb{R}^m$. The projection of the payoff function to the payoff of player i is denoted $u_i(a_i, a_{-i})$.

Payoff functions extend in the obvious multi-linear manner to payoff functions of mixed strategies.

An identical interests game is a game satisfying the property that $u_i(a) = u_j(a)$ for each $a \in A$ and each $i, j \in I$. A potential game is game with a potential function $\Phi : A \to \mathbb{R}$ satisfying for all $a_{-i} \in A_{-i}$ and all $a_i', a_i'' \in A_i$,

$$\Phi(a_i', a_{-i}) - \Phi(a_i'', a_{-i}) = u_i(a_i', a_{-i}) - u_i(a_i'', a_{-i}).$$

An ordinal potential game is game with a potential function $\Phi : A \to \mathbb{R}$ satisfying for all $a_{-i} \in A_{-i}$ and all $a_i', a_i'' \in A_i$,

$$\Phi(a_i', a_{-i}) - \Phi(a_i'', a_{-i}) > 0 \iff u_i(a_i', a_{-i}) - u_i(a_i'', a_{-i}) > 0.$$

Every identical interests game is a potential game and every potential game is an ordinal potential game.

2.3. **The Discrete Replicator Equation.** Much of the background material for the population genetics models here is from [Bürger, 2000] and [Edwards, 2000].

Time is discrete and denoted by positive integers t. Let $f^t : \Delta^m \to \mathbb{R}^m$ be given for each time t, with $f_i^t : \Delta^m \to \mathbb{R}$ for each $1 \leq i \leq m$ being the standard coordinate projection of f^t.

The mean value function associated with f^t, denoted $\overline{f}^t$, maps $\theta \in \Delta^m$ to $\mathbb{R}$ by

$$\overline{f}^t(\theta) := \sum_{1 \leq i \leq m} \theta_i f_i^t(\theta).$$

The discrete replicator equation is then the recursive mapping from Δ^m to Δ^m defined by

$$\theta_i^{t+1} := \theta_i^t \frac{f_i^t(\theta)}{\overline{f}^t(\theta)}. \tag{1}$$

2.4. **Alleles and Genotypes.** The model assumptions which will be maintained throughout are that populations are infinite (i.e., only proportions of genotypes and alleles in the population are of interest, not absolute numbers), that generations are discrete and non-overlapping, that selection occurs but not mutation or migration, and that stochastic genetic drift over time does not occur.

We assume that each gentoype is composed of m genetic loci. Each locus i is associated with a set of alleles $\mathcal{A}_i$ composed of $k+1$ alleles. A *genotype* is then formally a string $g = a_1 a_2 \ldots a_m$, such that $a_i \in \mathcal{A}_i$

for each $1 \leq i \leq m$. Denote the collection of all possible genotypes by Γ.

At each time t there is an adult population Π^t composed of individuals, each of which bears a genotype $g \in G$. The sub-population of individuals bearing genotype g at time t is denoted Π^t_g.

The adults in the population at time t in reproduce (asexually, haploid sexually, or diploid sexually, depending on the particular model being studied). After the adults in population Π^t reproduce, an offspring population Ω^t comes into existence. The sub-population of individuals bearing genotype g at time t is denoted Ω^t_g.

Denote by $d^t_g \in \Delta(\Gamma)$ the weight or relative proportion of genotype g at time t in the offspring population, i.e. the proportion of the set Ω^t_g in Ω^t. At the beginning of period $t+1$, the adult population Π^t dies, and as the individuals in Ω^t attain maturity they form the adult population Π^{t+1}.

A selection fitness value $w^t_g \in [0,1]$ is associated with each genotype g at each time t. This is interpreted as the probability that an offspring individual bearing genotype g in population Ω^t will survive and attain reproductive maturity as an adult in population Π^{t+1}.

2.5. **Asexual (Clonal) Reproduction Model.** In this model, at each time t each individual in Π^t_g produces ζ offspring (where ζ is a positive integer), each of whom bears the same genotype g as its parent. The offspring thus produced in population Ω^t then mature into the adults in population Π^{t+1}, subject to selection as determined by $\{\omega^t_g\}_{\omega \in \Gamma}$.

The relevant state space of the dynamic is the simplex $\Delta(\Gamma)$. The *mean fitness* at time t is

$$\overline{w}^t := \sum_{g \in \Gamma} d^t_g w^t_g. \tag{2}$$

In models in which w^t_g is constant over time we may suppress the time denotation and simply write w_g, and hence $\overline{w}^t = \sum_{g \in \Gamma} d^t_g w_g$.

The dynamics of asexual reproduction are governed by the asexual replicator equation as the equation of motion,

$$d^{t+1}_g = d^t_g \frac{w^t_g}{\overline{w}^t}. \tag{3}$$

This follows the schema of Equation (1), with the fitness w^t_g in the role of f_i and $\overline{w}$ corresponding to $\overline{f}$.

It will sometimes be convenient to express Equation (3) generically as

$$d_g^+ = d_g \frac{w_g}{\overline{w}}, \tag{4}$$

suppressing reference to t when its value is clear from context.

2.6. **Haploid Sexual Reproduction Model.** This model will be central to much of the paper, hence we present its assumptions here in some detail. We suppose a monoecious sexually reproducing haploid population, with panmictic mating occuring in pairs. Initially it will be supposed that there is no linkage between loci, i.e., each offspring at each locus bears the allele of one of the parents at the corresponding locus with equal probability. This assumption will be relaxed subsequently.

We will sometimes denote allele j at locus i by a_{ij},where $1 \leq i \leq m$ and $1 \leq j \leq k+1$, when convenient without confusion by context. When it is important to distinguish the j-th allele in locus i from the j-th allele at locus i', we will explicitly write a_{ij_i}. Let $C_{a_{ij_i}}$ denote the collection of all possible genotypes that contain a_{ij_i} at the slot for locus i. Write $\Pi^t_{a_{ij_i}} := \bigcup_{g \in C_{a_{ij_i}}} \Pi^t_g$ and $\Omega^t_{a_{ij_i}} := \bigcup_{g \in C_{a_{ij_i}}} \Omega^t_g$.

Denote by $q^t_{ij_i}$ the *allelic frequency* of allele a_{ij_i} at locus i at time t in population Ω^t, i.e., $q^t_{ij_i}$ is the proportion of $\Omega^t_{a_{ij_i}}$ in Ω^t. Call $q^t_i = \{q^t_{ij}\}_{1 \leq j \leq k+1}$ the *allelic frequency distribution* of locus i at t; this is an element of a k-simplex, which we will denote Δ^k_i. The relevant phase space for studying the evolutionary dynamic is then a polytope composed of an m-cross product of simplices:

$$\Theta := \Delta^k_1 \times \ldots \times \Delta^k_m. \tag{5}$$

The topology for studying convergence is the product topology of the simplices regarded as manifolds.

Going from $\Delta(\Gamma)$ to Θ is always possible, since we defined $q^t_{ij_i}$ as the proportion of $\Omega^t_{a_{ij_i}}$ in Ω^t for each a_{ij_i}. Denote the mapping thus defined by $\rho : \Delta(\Gamma) \to \Theta$.

For $g = a_{1j_1}, a_{2j_2}, , \ldots, a_{mj_m}$ and an allelic frequency distribution $q^t \in \Theta$ denote

$$q^t_g = q^t_{1j_1} q^t_{2j_2} \ldots q^t_{mj_m}. \tag{6}$$

If $d^t_g = q^t_g$ for all g, a population is said to be in linkage equilibrium. When linkage equilibrium obtains, the inverse mapping $\rho^{-1} : \Theta \to \Delta(\Gamma)$ is well-defined by applying Equation (6). When we make use

of this inverse mapping, given $x \in \Theta$ we will write $[\rho^{-1}(x)]_g$ to stand for the g-th component of $\rho^{-1}(x) \in \Delta(\Gamma)$.

The *marginal fitness* of allele a_{ij} at time t is defined as

$$w_{ij}^t := \sum_{g \in C_{a_{ij}}} w_g^t \frac{d_g^t}{\sum_{g' \in C_{a_{ij_i}}} d_{g'}^t}. \tag{7}$$

From the collection $\{w_{ij}^t\}_{1 \leq j \leq k+1}$ we furthermore can calculate the *mean payoff* for locus i, which is $\overline{w}_i^t := \sum_{j=1}^{k+1} q_{ij}^t w_{ij}^t$. But this yields nothing new, because $\overline{w}_i^t = \overline{w}^t$ of Equation (2) for all loci i.

The dynamic in this model is the *haploid sexual replicator* which can be shown to be

$$q_{ij}^{t+1} = q_{ij}^t \frac{w_{ij}^t}{\overline{w}^t}, \tag{8}$$

and applies at every allele j of every locus i. This clearly follows the schema of Equation (1) with w_i^t here in the role of f^t in Equation (1) and $\overline{w}_i^t$ as $\overline{f}^t$. As before, it will sometimes be convenient to express Equation (8) generically as

$$q_{ij}^+ = q_{ij} \frac{w_{ij}}{\overline{w}}, \tag{9}$$

suppressing reference to t when its value is clear from context.

The haploid sexual replicator dynamic maps points in Θ to points in Θ, and hence also maps points in Δ_i^k to points in Δ_i^k under the projection from Θ to Δ_i^k.

2.6.1. *Haploid Reproduction as an Identical Interests Game.* In the fixed fitness case, the collection of fitness values $\{w_g\}_{g \in \Gamma}$ can be regarded as defining an identical interests game between the loci. Here we essentially rewrite some of the previous sections in notation that is more familiar from the game theory literature, and unite the analysis of haploid dynamics with that of the dynamics of player strategies in repeated identical interests games when the players implement the haploid sexual replicator equation, Equation (8), in updating their strategies.

From the fixed fitness haploid model with polytope Θ of alleles at the various loci, define an identical interest game W_Θ as follows. Each locus i becomes a player i. The set of alleles of locus i becomes the set of pure actions A_i of player i. For each profile of pure actions $(a_1, \ldots, a_m) \in A_1 \times \ldots \times A_m$, the payoff $w_i(a_1, \ldots, a_m) = w(a_1, \ldots, a_m)$

to each player i is identical and defined to be

$$w(a_1, \ldots, a_m) := w_g,$$

where $g = a_1 \ldots a_m$ is the genotype defined by $a_i \in \mathcal{A}_i$ for each i and w_g is fitness payoff to genotyper g.

Here $q_i \in \Delta_i^k$, which previously denoted the distribution of alleles in locus i, is interpreted as a mixed strategy. The mean fitness $\overline{w}$ is interpreted as $\overline{w} = w(q_1, \ldots, q_m)$, the expected payoff (to each player in the game) when each player/locus i plays mixed strategy q_i. The expected payoff/mean fitness $\overline{w}$ plays the role of the potential function in the identical interests game W_Θ.

Every potential game (and hence every identical interests game) admits at least one pure strategy Nash equilibrium, namely the pure strategy profile yielding the highest potential payoff. The set of all pure Nash equilibria is the set of local maxima of the potential. Denote this set of pure Nash equilibria of W_Θ by N_{W_Θ}.

Each $\nu \in N_{W_\Theta}$ is by definition a profile of alleles $(a_1, \ldots, a_m)$, one from each locus. Hence it is naturally associated with a particular genotype that we will denote $g_\nu \in \Gamma$.

Note that if the set of mixed strategy profiles is restricted to a subset $\Theta' \subset \Theta$, a different identical interests game $W_{\Theta'}$ is induced. The set of pure Nash equilibria of $W_{\Theta'}$ may differ from the set of pure Nash equilibria of W_Θ.

We may write $q_{i_{a_j}}$ as a synonym for q_{ij} when q_i is the mixed strategy of i. We can write $w(p; q_{-i})$ for the expected payoff when locus i plays mixed strategy p while all the other loci play mixed strategy q_{-i}. In a special case this notation becomes $w(a_j; q_{-i})$, standing for the expected payoff when pure action/allele $a_j \in \mathcal{A}_i$ is chosen at locus i while all the other loci play mixed strategy q_{-i}; this is none other than the game interpretation of the marginal fitness of allele $a_j \in \mathcal{A}_i$, which was above written as w_{ij}. Then for each i,

$$w(q_1, \ldots, q_m) = \sum_{a_\ell \in \mathcal{A}_i} q_{i_{a_\ell}} w(a_\ell; q_{-i}).$$

2.7. **Diploid Sexual Reproduction Model.**

2.7.1. *One Locus.* In the single locus diploid model, with a set of alleles $\mathcal{A}$, one needs to keep track of pairs of alleles, $a_i a_j \in \mathcal{A}$, which constitute the genotypes. We suppose no position effects and hence do not distinguish between $a_i a_j$ and $a_j a_i$. Random mating is also assumed, hence Hardy–Weinberg ratios hold during the mating phase

(with selection then moving the adult population away from the Hardy–Weinberg ratios).

Label the frequency of allele a_i at time t by p_i^t and the frequency of genotype $a_i a_j$ by $P_{ij}^t = P_{ji}^t = p_i^t p_j^t$. Denote the fitness of genotype $a_i a_j$ by $W_{ij}^t = W_{ji}^t$, and the population mean fitness by

$$\overline{W}^t = \sum_{i,j} P_{ij}^t W_{ij}^t = \sum_i (p_i^t p_i^t W_{ii}^t + \sum_{j \neq i} p_i^t p_j^t W_{ij}^t). \tag{10}$$

Define the marginal fitness of allele a_i as

$$W_i = \sum_j p_i W_{ij} = p_i W_{ii} + \sum_{j \neq i} p_j W_{ij}. \tag{11}$$

With all the preliminaries in place, the dynamic is once again defined by a straight-forward replicator as the equation of motion

$$p_i^+ = p_i \frac{W_i}{\overline{W}}, \tag{12}$$

for each allele.

2.7.2. *Multiple Loci.* The multi-locus diploid model is complicated to describe; we omit most details and present only the minimal notation needed for our purposes here.

As before we suppose that there are m loci with $k+1$ alleles per locus. The state space is $\Delta_1^k \times \ldots \times \Delta_m^k$ and trajectories are elements $(p_1, \ldots, p_m) \in \Delta_1^k \times \ldots \times \Delta_m^k$.

Within each locus i, as in the single locus model, the alleles are between themselves playing at each time period a symmetric potential game with a fitness $W_{i_k i_l}^t$ assigned to each pairing $a_{i_k} a_{i_l}$, where $a_{i_k}, a_{i_l} \in A_i$. However, $W_{i_k i_l}^t$ is now a function not only of $a_{i_k} a_{i_l}$ but of the entire profile p_{-i} of the allelic distributions of the other loci.

The standard analysis in the literature tracks the distribution of gametes (where each gamete is one possible haploid half of a diploid genotype). Each gamete g can be assigned a marginal fitness W_g as a function of the fitnesses of the pairings at each locus and the allelic frequency, and from this the mean fitness $\overline{W}$ of the population is calculated. Denoting the frequency of gamete g by r_g, one can derive a recursion formula that is reminiscent of but not identical to the replicator equation

$$r_g^+ = r_g \frac{W_g}{\overline{W}} - D_g \tag{13}$$

where D_g is the linkage disequilibrium for g. The existence of the disequilibrium term D_g means that the diploid multi-locus dynamic is not a replicator dynamic, making the analysis of this dynamic different from all the other models studied in this paper.

3. Multiplicative Weights, Potential Games, and Virtual Convergence

3.1. **Regret Minimisation.** The objective of many of the on-line learning algorithms developed in the literature in recent years is the attainment of regret minimisation. Let $\mathcal{K} \subset \mathbb{R}^k$ be non-empty, bounded, compact, and convex. At each iteration time t algorithm $\mathcal{A}$ selects an element $x_t \in \mathcal{K}$, while a concave function $\ell_t : \mathcal{K} \to \mathbb{R}$ is revealed.

The goal of the algorithm is to minimise the average regret over any n rounds, defined as

$$R_n(\mathcal{A}) := \sum_{t=1}^{n} \ell_t(x^*) - \sum_{t=1}^{n} \ell_t(x_t),$$

where $x^* \in \arg\max_{x \in \mathcal{K}} \sum_{t=1}^{n} \ell_t(x)$. In other words, the objective is to have minimal regret relative to having selected the best possible $x^* \in \mathcal{K}$ from the start and playing x^* in a fixed manner at every time period.

An algorithm implements asymptotic regret minimisation if its regret is sub-linear, i.e., $R_n(\mathcal{A}) = o(n)$ as $n \to \infty$. When this holds

$$\limsup_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \ell_t(x^*) - \frac{1}{n} \sum_{t=1}^{n} \ell_t(x_t) \leq 0 \tag{14}$$

where x^* is the element of $\mathcal{K}$ with the optimal average payoff.[2] In other words, the average regret converges to zero in the limit and the payoff of the algorithm approaches that of having selected the optimal in hindsight x^* from the start and monotonically selecting only that point at every iteration.

3.2. **Multiplicative Weights Update Algorithm.** The multiplicative weights update algorithm comes in two flavours: a polynomial and exponential version. In the polynomial version, $d \in \Delta^k$ is mapped

[2] Strictly speaking we need to consider the lim sup in Equation (14) because the limiting average payoff value might not be well defined.

to $d^+ \in \Delta^k$, conditional on receipt of a given tuple of real numbers $(\ell_1, \ldots, \ell_k)$, by

$$d_i^+ = \frac{d_i(1+\eta\ell_i)}{\sum_j d_j(1+\eta\ell_j)} \tag{15}$$

for some $\eta > 0$. Dividing the numerator and denominator of Equation (15) by η changes nothing, hence Equation (15) can be equivalently expressed as

$$d_i^+ = \frac{d_i(\frac{1}{\eta}+\ell_i)}{\sum_j d_j(\frac{1}{\eta}+\ell_j)}. \tag{16}$$

In the special case in which $\eta \to \infty$, sometimes called the parameter-free version of the algorithm (cf. [Meir and Parkes, 2015]), Equation (16) becomes

$$d_i^+ = \frac{d_i\ell_i}{\sum_j d_j\ell_j} = d_i\frac{\ell_i}{\overline{\ell}}, \tag{17}$$

which is exactly the replicator equation.

In its exponential version the multiplicative weights update algorithm, also known as the Hedge algorithm ([Freund and Schapire, 1997])), maps $d \in \Delta^k$ to $d^+ \in \Delta^k$, conditional on receipt of a given tuple of real numbers $(\ell_1, \ldots, \ell_k)$, by

$$d_i^+ = \frac{d_i \cdot \mathrm{e}^{\eta\ell_i}}{\sum_j d_j \cdot \mathrm{e}^{\eta\ell_j}} \tag{18}$$

for some $\eta > 0$.

The replicator equation can also be shown to be a special case of the exponential algorithm ([Edhan et al., 2017]) as expressed in Equation (1). The key is to register not the fitness payoffs at each time period but the logarithms of the fitnesses: given a fitness tuple $f = (f_1, \ldots, f_m)$, form the tuple $(\ell_1, \ldots, \ell_m)$ by setting $\ell_i = \frac{1}{\eta}\ln(f_i)$. Then apply Equation (18):

$$d_i^+ = \frac{d_i \cdot \mathrm{e}^{\eta\ell_i}}{\sum_j d_j \cdot \mathrm{e}^{\eta\ell_j}} = \frac{d_i \cdot \mathrm{e}^{\eta(\frac{1}{\eta}\ln(f_i))}}{\sum_j d_j \cdot \mathrm{e}^{\eta(\frac{1}{\eta}\ln(f_j))}} = d_i\frac{f_i}{\overline{f}}.$$

It is well known in the literature that the multiplicative weights update algorithm attains regret minimisation. In the genetic context studied here, this translates into attaining asymptotic average growth rates equal to that of having selected the optimal-in-hindsight genotype g^* from the start and hypothetically running history again with a population consisting solely of g^* at every time period.

Furthermore, since the haploid sexual reproductive strategy can be interpreted as an implementation of the replicator independently in each locus, the interpretation of the replicator as an instantiation of the multiplicative weights updating algorithm is applicable in several of the models in this paper, beyond the asexual model.

Several papers studying the applicability of multiplicative weights updating algorithms to evolutionary models have been published in recent years. A brief list of such papers includes [Livnat et al., 2008, Chastain et al., 2013, Chastain et al., 2014, Meir and Parkes, 2015].

3.3. **Multiplicative Weights and Baum–Eagon.** It is instructive to compare the multiplicative weights updating algorithm, especially in its parameter-free version:

$$q_i^+ = q_i \frac{\ell_i}{\bar{\ell}},$$

and Baum–Eagon updating (as in Equation (24))

$$q_{ij}^+ := q_{ij} \frac{\frac{\partial U}{\partial q_{ij}}}{\sum_s q_{is} \frac{\partial U}{\partial q_{is}}}.$$

From the perspective of each Δ^i, the Baum–Eagon updating is a special case of the multiplicative weights updating algorithm in which the payoff ℓ_{ij} is given as the partial derivative of a potential function U with respect to q_{ij}. This perspective will have a significant role here, as many of the dynamics that will be studied benefit both from the monotonic potential increase afforded by the Baum–Eagon Theorem and the regret-minimisation given by the multiplicative weights algorithm aspect.

3.4. **Convergence in Potential Games.** The content of the following theorem is technically equivalent to a theorem in [Palaiopanos et al., 2017] (see also [Panageas et al., 2019]), which is expressed and proved there in the context of congestion games. We present it here with a full proof for two reasons: a) an independent proof for potential games is of value; b) the proof here can readily be understood in the context of reproductive strategies, such as haploid sexual reproduction, given the interpretation of such strategies as implementing the multiplicative weights updating algorithm, as described in Section 3.2, in the context of a potential game between loci, with alleles in the role of pure actions, as described in Section 2.6.1.

Theorem 1. *Suppose that each of a finite set of players playing a potential game implements the polynomial multiplicative weights update algorithm at discrete time periods to update his mixed strategy, starting from a mixed strategy of full support.*

Then the strategy profile of the players will converge to a fixed point that is a Nash equilibrium.

Theorem 1 is a stronger result than may appear at first glance, because there is no explicit coordinating element between the players that is assumed. To see why this may be surprising, consider the following extremely simple 2×2 game, which is an identical interests game (and hence a potential game):

	S	B
S	$(2,2)$	$(0,0)$
B	$(0,0)$	$(1,1)$

One may interpret this as a coordination game between a couple, who wish to meet. If they are both at the symphony hall (action profile (S,S)) they each receive a payoff of 2; if they are both at the beach (action profile (B,B)) they each receive a payoff of 1; otherwise they fail to meet and receive zero payoff. Suppose that both players simultaneously implement a simple-minded best reply strategy, beginning at action profile (S,B). Then in the next time period, the action profile will be (B,S), followed by (S,B) etc. Lacking a coordinating element, no convergence to a fixed point is attained.

In contrast, Theorem 1 does guarantee convergence under the multiplicative weights update algorithm, even though there is no coordination between the players and each player updates his or her mixed strategy from one time period to the next entirely independently of the other players. It is as if coordination is attained 'for free'. This result is attained by virtue of the Baum–Eagon Theorem, which underlies the proof of the theorem and guarantees that, despite the lack of coordination, a monotonic climb up the potential of the game ensues at each time period.

3.5. **Virtual Convergence.** Let $\Theta := \Delta_1^k \times \ldots \times \Delta_m^k$ be a polytope, with $T : \Theta \to \Theta$ a transformation.

We will say that the dynamic defined by T *converges polymorphically* if for each initial point $x \in \Theta$ there is a point $y \in \Theta$ such that $\lim_{n\to\infty} T^n(x) = y$. In the special case that for each x the limit $y = \lim_{n\to\infty} T^n(x) = (q_1, \ldots, q_m)$ satisfies the condition that q_i is a

point distribution in Δ_i^k for each $1 \leq i \leq m$, the dynamic *converges monomorphically*.

Suppose now that a linear fitness function $\ell_t : \Theta \to \mathbb{R}$ is revealed for each time t. For an initial point $x \in \Theta$, denote $x_n := T^n(x)$, with $T^0(x) = x$. We will say that the dynamic defined by T *virtually converges polymorphically* if for any sequence $\ell_1, \ell_2, \ldots$ of payoffs and any initial point $x \in \Theta$, there is a point $y^* \in \Theta$ such that

$$\left| \limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \ell_t(y^*) - \frac{1}{n} \sum_{t=1}^{n} \ell_t(x_t) \right| = 0.$$

In the special case that virtual convergence is to a y^* that is a point distribution in Δ_i^k for each $1 \leq i \leq m$, we say that *virtual monomorphic convergence* obtains.

4. Asexual (Clonal) Reproduction

The dynamics of frequency independent asexual reproduction without mutation is perhaps the simplest of evolutionary dynamics – essentially 'bacteria in a petri dish'. Despite the apparent simplicity, there is much to be said here that will also have implications for the analysis presented in later sections.

4.1. **Fixed Fitness.** We suppose here a fixed fitness value w_g for each genotype at each time period, generically with a genotype $g^* \in \Gamma$ whose fitness w_{g^*} is maximal amongst the genotypes. There are several ways to analyse this; in the spirit of this paper, we may regard this dynamic as a single-player potential game. In this interpretation, there is one player whose mixed strategy at time t is a probability measure $d^t \in \Delta(\Gamma)$ over the genotypes in Γ. The expected payoff is $\sum_{g\in\Gamma} d_g^t w_g$. Theorem 1 then implies convergence to a fixed point in $\Delta(\Gamma)$.

Alternatively, we may directly apply the Baum–Eagon Theorem. The dynamics are governed by the asexual replicator equation,

$$d_g^+ = d_g \frac{w_g}{\overline{w}}. \tag{19}$$

Since $\overline{w} = \sum_{g\in\Gamma} d_g w_g$, it follows that $\frac{\partial \overline{w}}{\partial d_g} = w_g$, hence Equation (19) is an application of the Baum–Eagon transformation as expressed in Equation (24).

Denote by $T_0 : \Delta(\Gamma) \to \Delta(\Gamma)$ the transformation that defines $d^+ = T_0(d)$ by mapping d_g to d_g^+ for each g according to Equation (19). Since the population mean fitness is increasing montonically, $\lim_{n\to\infty} T_0^n(d)$

for any starting distribution $d \in \Delta(\Gamma)$ converges to a point in $\Delta_1(\Gamma)$, i.e., a fixed point that is a point distribution, since the only fixed points of Equation (19) are point distributions. All the weight is asymptotically on 1_{g^*}, where g^* is the genotype of maximal fitness.

This implies that the interior of the simplex $\Delta(\Gamma)$ forms a global exponentially stable basin of attraction. If, in contrast, the initial point lies within a strict subface $F \subset \Delta(\Gamma)$, then the convergence will again be to a monomorphic population whose genotype is the genotype of maximal fitness within F. This will clearly be sub-optimal if $g^* \notin F$.

4.2. **Temporally Varying Fitness.** The fixed fitness setting of asexual reproduction is the simplest evolutionary model, yielding perhaps the strongest result that can be expected, of monotonic and rapid fitness increase towards convergent fixation to the globally optimal genotype. This satisfactory result, however, may not necessarily obtain if fitnesses are no longer fixed in time.

In a temporally varying fitness model, we suppose that there is a collection of possible environments Ω, such that each $\omega \in \Omega$ determines a fitness landscape such that each genotype g is assigned a fitness value $w_g(\omega)$ under ω. At each time t one environment ω from Ω is selected, with the payoff to the genotypes registered in accordance to the fitness landscape of that environment.

A simple hill climbing dynamic cannot be applicable here because there is a different 'hill' (i.e., fitness gradient derived from the environment) at each time period; the trajectory under the transformation T_0 will no longer be monotonically increasing in mean fitness. Despite this, the replicator algorithm does an excellent job at learning, even under conditions of temporally varying fitness. This can be seen in several ways.

Consider first a discrete i.i.d. model in which there is a probability measure μ over Ω determining the selection of the environment at each time period, repeated indefinitely. This determines for each genotype g an expected fitness payoff under μ. An optimal population will (generically) be composed of the genotype with maximal expected fitness payoff, and the replicator reliably identifies this genotype. More generally:

Proposition 1. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space over a collection Ω of environments. For each genotype $g \in \Gamma$, define a random variable $w_g(\omega) \in [0,1]$, interpreted as the fitness of g under environment $\omega \in \Omega$, from which the expected fitness is given as $E(w_g \mid \mu) = \int_\Omega w_g(\omega)\, d\mu(\omega)$.*

Let $S : \Omega \to \Omega$ be a stationary and ergodic transformation defining a stochastic process for each g by $w_g^t(\omega) = w_g(S^t(\omega))$. Then under the asexual replicator dynamic, with probability one the population asymptotically converges to a monomorphic population consisting of the genotype with maximal expected fitness.

Proposition 1 indicates that when there is sufficient structure to the stochastic process of the varying environments, at least as expressed in stationary ergodicity (which include i.i.d. as a special case), the replicator dynamic will be able to extract the information inherent in the process to identify the optimal genotype and converge to that genotype, from any initial population state (that at least minimally includes the optimal genotype).

From here one can ask what happens when the stochastic process of varying environments can be any process at all. It is not difficult to conjure examples of temporally varying environments that do not admit convergence to a single genotype. For example, let

$$w_g^t = \begin{cases} e^{\frac{1}{2}} & \text{for } t \leq 100 \mod 200 \\ e^{\frac{1}{3}} & \text{for } t > 100 \mod 200 \end{cases}$$

$$w_h^t = \begin{cases} e^{\frac{1}{3}} & \text{for } t \leq 100 \mod 200 \\ e^{\frac{1}{2}} & \text{for } t > 100 \mod 200 \end{cases}$$

Then clearly both $\liminf \frac{1}{T}\sum_{t=1}^{T} \ln w_g^t < \limsup \frac{1}{T}\sum_{t=1}^{T} \ln w_h^t$ and $\liminf \frac{1}{T}\sum_{t=1}^{T} \ln w_h^t < \limsup \frac{1}{T}\sum_{t=1}^{T} \ln w_g^t$. When one genotype is strong the other is weak, each temporarily overtaking the other only to fall back later.

Nevertheless, it is possible to extend Proposition 1 to much more general environments using the notion of *one-step-ahead expected log-fitness*. The one-step ahead expected log-fitness is the expected log-fitness of a generation conditional on the past generations.

Definition 2. *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space over a collection Ω of environments and let $(\psi_t)_{t\geq 1}$ be a stochastic process of environments relative to $(\Omega, \mathcal{B}, \mu)$. For each genotype $g \in \Gamma$, define a process by $w_g^{\psi_t} = w_g(\psi_t)$, interpreted as the fitness of g under environmental process ψ_t. Let $\rho_t^g = \ln w_g^{\psi_t}$ denote the log-fitness; assume that ρ_t^g is always bounded.*

We will call $\widehat{\rho}_t^g = \frac{1}{t}\sum_{s=1}^{t} E[\rho_{s+1}^g | \rho_s^g, ..., \rho_1^g]$ the average one-step-ahead expected log-fitness of g at t.

Definition 3. *A gentoype g is* asymptotically one-step-ahead superior on average *if* $\liminf_{t\to\infty} \widehat{\rho}_t^g > \limsup_{t\to\infty} \widehat{\rho}_t^h$ *with probability one for all genotypes $h \in \Gamma$ with $h \neq g$.*

Theorem 2. *If a genotype $g \in \Gamma$ is asymptotically one-step-ahead superior on average then, under the asexual replicator dynamic, with probability one the population asymptotically converges to a monomorphic population consisting of the genotype g.*

It is worthwhile noting here that in the case of an ergodic environment, $\liminf\limits_{t\to\infty} \widehat{\rho}_t^g$ and $\limsup\limits_{t\to\infty} \widehat{\rho}_t^g$ are one and the same and equal to the constant $E[\ln w_g]$ almost surely. Thus the sufficient condition established in Theorem 2, namely asymptotic one-step-ahead superiority, is reduced to $E[\ln w_g] > E[\ln w_h]$.

The statement of Theorem 2 supposes that genotype g is asymptotically one-step-ahead superior on average with probability one with respect to all environments. Suppose instead that a genotype g is is asymptotically one-step-ahead superior on average only with respect to a subset A_g^∞ of the collection of environments. Then we obtain the following corollary.

Corollary 4. *If $Pr(A_g^\infty) > 0$, where A_g^∞ is the set of environments in which genotype g is is asymptotically one-step-ahead superior on average, then under the asexual replicator dynamic, with probability one in A_g^∞ the population asymptotically converges to a monomorphic population consisting of the genotype g.*

Algorithms such as the multiplicative weights and mirror ascent algorithms have been developed in the computer science literature in recent years for the sake of optimisation under conditions of no statistical structure. The replicator dynamic, it turns out, exploits the results afforded by these algorithms.

Theorem 3. *Under the replicator dynamic, for any arbitrary temporally varying fitness there is an optimal-in-hindsight genotype g^* such that for any initial point in the interior of the simplex, asexual reproduction virtually converges monmorphically to g^*.*

In summary, we interpret the results of this section from a learning perspective: the objective is to learn which genotype is best fit for the environment process, via the algorithmic tool of the replicator.

When the environment is fixed, the replicator homes in on the objectively fittest genotype. When the environment process is sufficiently structured, as in a stationary ergodic process, the replicator

makes use of time averaging to identify a winning genotype. Failing that, in the worst case in which there is insufficient structure for predictive learning, the replicator still manages to extract information, by application of regret minimisation via the multiplicative weights updating algorithm; virtual convergence occurs in the sense that one can imagine a population which from the start consisted of only the optimal-in-hindsight genotype and attaining the same asymptotic average growth rate as actually attained.

5. Haploid Sexual Reproduction

5.1. **Fixed Fitness.** In this section the population will be presumed to reproduce via haploid sexual reproduction under a fitness landscape $\{w_g\}_{g\in\Gamma}$ that is fixed throughout time.

5.1.1. *Under Linkage Equilibrium.* Under linkage equilibrium, in population Ω_g^t the equation $d_g^t = q_{1j_1}^t q_{2j_2}^t \dots q_{mj_m}^t$ holds for each genotype $g = a_{1j_1}, a_{2j_2}, , \dots, a_{mj_m}$. As Ω_g^t matures into Π_g^{t+1}, selection applies such that linkage equilibrium does *not* hold for Π_g^{t+1}; however, by assumption random mating between the reproducing adults in Π_g^{t+1} immediately restores linkage equilibrium in the next offspring generation Ω_g^{t+1}.

One advantage of working with an assumption of linkage equilibrium is that we may identify in a bijective manner a point in the allelic frequency space Θ and a corresponding point in the genotypic frequency space $\Delta(\Gamma)$. We shall freely do so in this section as follows.

Recalling the haploid sexual replicator,

$$q_{ij_i}^+ = q_{ij_i}\frac{w_{ij_i}}{\overline{w}}, \tag{20}$$

define $\tau : \Theta \to \Theta$ to be the transformation given by the mapping of q_{ij} to q_{ij}^+ for each locus i and allele j in i. Exploiting the linkage equilibrium assumption, define a transformation $T_1 : \Delta(\Gamma) \to \Delta(\Gamma)$ by

$$T_1(d) = \rho^{-1} \circ \tau \circ \rho(d). \tag{21}$$

Abusing terminology, we will call both τ and T_1 haploid sexual replicator transformations. This enables us to analyse the dynamics equally well under either T_1 or τ; both define discrete dynamical systems determining trajectory paths in $\Delta(\Gamma)$ or in Θ, respectively.

As in the asexual case, the Baum–Eagon inequality applies (see a similar argument in [Novak and Barton, 2017]). The domain is the polytope Θ as defined in Equation (5).

Lemma 5. *The haploid sexual replicator transformation (under linkage equilibrium and without genetic linkage between loci) satisfies the Baum–Eagon inequality, with mean fitness $\overline{w}$ as a Lyapunov function.*

It follows that the population will asymptotically converge to a fixed point of the dynamics defined by the transformation T_1 along paths of monotonically increasing mean fitness.

From here it will be convenient to continue the analysis from the equivalent perspective of the identical interests game interpretation.

Theorem 4. *Under haploid recombinative sexual reproduction (under linkage equilibrium and without genetic linkage between loci), trajectories almost always increase mean fitness monotonically.*

Beginning from almost any interior point of $\Delta(\Gamma)$ the haploid sexual replicator dynamic converges asymptotically to a monomorphic population in which each individual bears a genotype g_ν from the set N_{W_Θ} of pure Nash equilibria of the associated potential game W_Θ.

Corollary 6. *If the initial point of the allelic frequency of the population lies in any face Θ' of Θ then the dynamic converges asymptotically to a monomorphic population consisting of genotypes from the set of pure Nash equilibria of the associated potential game W'_Θ.*

There are immediate interesting implications of Theorem 4. One of these is that $\Delta(\Gamma)$ is entirely partitioned into asymptotically stable basins of attraction (deterministically in this model).

Theorem 5. *For each pure Nash equilibrium $\nu \in N_{W_\Theta}$, there exists $B_\nu \subset \Delta(\Gamma)$ containing $1_{g_\nu} \in \Delta_1(\Gamma)$ such that starting from any initial point in B_ν the population under the dynamic will converge to a monomorphic population consisting solely of genotype g_ν, i.e., $T_1^n(x) \longrightarrow 1_{g_\nu}$ for every $x \in B_\nu$. Apart from separatrices between these basins of attraction, which are of negligible measure, the sets in the collection $\{B_\nu\}$ form a partition of $\Delta(\Gamma)$.*

Even more than that can be said here. By Theorem 6 of [Baum and Sell, 1968], any transformation of the form defined in Equation (24) increases U-homotopically, from which it follows that the haploid sexual replicator transformation $T_1 : \Delta(\Gamma) \to \Delta(\Gamma)$ increases $\overline{w}$-homotopically.

Proposition 7. *Let $S_t(x) = tT_1(x) + (1-t)x$. For each pure Nash equilibrium $\nu \in N_{W_\Theta}$, there exists a neighbourhood $H_\nu \subset \Delta(\Gamma)$ of 1_{g_ν} such that $S_t(H_\nu) \subset H_\nu$ for $0 < t \leq 1$, and for every $x \in H_\nu$, $T_1^n(x) \longrightarrow 1_{g_\nu}$. Furthermore, each H_ν has the homotopy type of a disk.*

The significance of the 'basin of homotopic attraction' H_ν of Proposition 7 is that not only does every point $x \in H_\nu$ converge to g_ν under the dynamic, also a small perturbation of around x perserves this property. In contrast, around any pure strategy point that is neither a local maximum or a local minimum there are points such that a small perturbation can lead to asymptotic convergence to different fixed points.

Finally:

Proposition 8. *For each pure Nash equilibrium $\nu \in N_{W_\Theta}$, there exists a neighbourhood $E_\nu \subset \Delta(\Gamma)$ of 1_{g_ν} that is an exponentially stable basin of attraction.*

Within the exponentially stable basin of attraction around a Nash equilibrium, the haploid sexual replicator dynamics resembles the asexual replicator dynamics, with exponential convergence to an equilbrium point.

The containment relations are $1_{g_\nu} \in E_\nu \subseteq H_\nu \subseteq B_\nu$. This implies that an observer following a trajectory starting in B_ν far from g_ν will likely initially see a slow and moderate increase in mean fitness, with broad polymorphism, for a long time, but once the trajectory enters E_ν suddenly an extremely fast rise in mean fitness will be registered along with rapid convergence to a monomorphic population.

5.1.2. *Under Genetic Linkage Between Loci.* Genetic linkage in this section means physical linkage between loci: in the context of haploid reproduction, this means that an offspring zygote might inherit a pair (or more) of alleles from one of the parents as a package, in contrast to an assumption of independent inheritance with probability 0.5 from each parent at each locus.

The term linkage disequilibrium refers to a population genotype distribution that does not equal the cross product of the marginal distribution as reflected in the allelic distribution. We note that in the game theoretic terms that we have been applying throughout to the study of genetic reproduction, linkage equilibrium corresponds to independent strategy selection of each player, while linkage disequilibrium corresponds to dependencies in the selection of strategies.

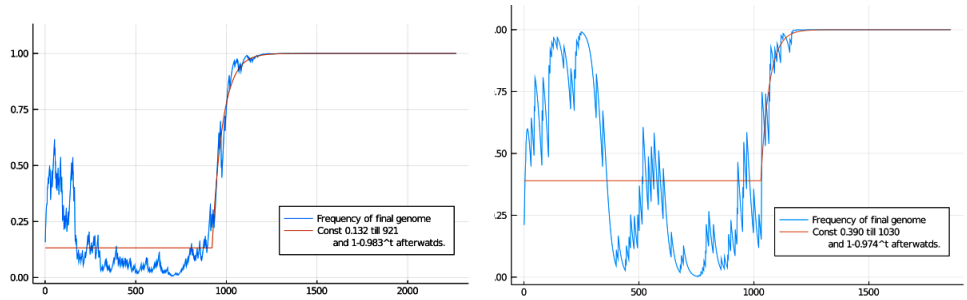

Figure 2. An illustration of the convergence to a monomorphic populations consisting of one genotype under haploid sexual reproduction. In both simulations, the frequency of the final genome in the population over generations is tracked. This frequency appears to drift most of time, then converges at an exponential rate to fixation towards the end, in accordance with Proposition 8.

We initially present an analysis of the haploid sexual replicator under genetic linkage between loci in the case of two loci, for clarity of exposition.

Let $r \in [0, 1]$ be the recombination rate. Suppose one starts with a point $d \in \Delta(\Gamma)$ representing the population distribution. Project d to Θ via $\theta = \rho(d)$. Under the asexual replicator d is mapped to $T_0(d)$, and under the haploid replicator θ is mapped to $\tau(\theta)$. Then the replicator equation under recombination rate r is

$$d^+ = r[\rho^{-1}(\tau(\theta))] + (1 - r)\alpha(d)$$

or, using the transformation T_1 defined in Equation (21),

$$d^+ = rT_1(d) + (1 - r)T_0(d). \tag{22}$$

We may denote by T_r the transformation defined by Equation (22) which is consistent with our labelling of T_1 and T_0. When $r \neq 1$, genetic linkage between the loci occurs.

The recombination rate r is intended to describe a situation in which each offspring is produced by sexual recombination with probability r and is produced by asexual cloning with probability $1 - r$. This results in an offspring population, such that within that population, a weight r of the offspring is descended from a sexual reproduction event and weight $1 - r$ is descended from an asexual reproduction event.

We may instead consider the following situation, which is mathematically equivalent and more convenient for our purposes: Create two separate copies Π_0 and Π_1 of the reproducing population Π, maintaining the genotype frequencies of the original population in each copy, with relative population size proportions $\frac{|\Pi_1|}{|\Pi_0|} = \frac{r}{1-r}$. Let Π_0 reproduce asexually to produce offspring population Ω_0 and Π_1 reproduce haploid sexually to produce offspring population Ω_1, finally combining them into $\Omega = \Omega_0 \cup \Omega_1$ and regarding the genotypic frequency of Ω.

Slightly more generally, select fraction r of the population at random to reproduce by the haploid sexual transformation, with the remaining $1 - r$ of the population reproducing by the asexual transformation. All of these alternatives result in an offspring population with weight r descending from a sexual reproduction event and weight $1 - r$ descending from an asexual reproduction event, which is what is relevant.

Proposition 9. *In a population starting at an initial point in linkage equilibrium, under two-locus haploid recombinative sexual reproduction with recombination rate r, trajectories always increase mean fitness monotonically.*

Beginning from any such interior point of $\Delta(\Gamma)$ the haploid sexual replicator dynamic converges asymptotically to a monomorphic population in which each individual bears a genotype g_ν from the set N_{W_Θ} of pure Nash equilibria of the associated potential game W_Θ.

In greater generality, suppose that there are m loci. Let λ be a partition of $\{1, \ldots, m\}$ into $\ell \leq m$ partition elements. An individual will be of λ-type if, when reproducing, the genes of that individual undergo physical genetic linkage according to λ. In other words, if two λ-type individuals I_1 and I_2 mate and produce an offspring O, then for each partition element $\overline{\lambda}$ of λ, all the alleles in the loci included in $\overline{\lambda}$ in the genotype of O will be identical to either the alleles of $\overline{\lambda}$ in the genotype of I_1 or the alleles of $\overline{\lambda}$ in the genotype of I_2, with equal probability. If the entire population reproduces in this way, denote the resulting transformation from $\Delta(\Gamma)$ to $\Delta(\Gamma)$ by T_λ.

If λ is the coarsest partition, consisting of only one partition element, this describes asexual reproduction. For any other partition, λ-type reproduction with $1 < \ell \leq m$ partition elements reduces to haploid sexual reproduction: simply regard the ℓ partition elements as ℓ independent loci. If λ is the finest partition, in which each locus

is its own partition element, this describes haploid sexual reproduction under independence of inheritance at each locus.

Let Λ be the set of all partitions of $\{1, \ldots, m\}$. For each $\lambda \in \Lambda$ let $r_\lambda \in [0, 1]$, such that $\sum_{\lambda \in \Lambda} r_\lambda = 1$. Interpret r_λ as the probability that an offspring is produced by physical genetic linkage in accordance with partition λ. Mathematically this is equivalent to selecting at random at each generation, for each $\lambda \in \Lambda$, a fraction r_λ of the population which reproduces by λ-type reproduction.

The resulting $\{r_\lambda\}_{\lambda \in \Lambda}$-tuple replicator equation is

$$d^+ = \sum_{\lambda \in \Lambda} r_\lambda T_\lambda(d). \tag{23}$$

Theorem 6. *In a population starting at an initial point in linkage equilibrium, under m-locus haploid recombinative sexual reproduction with recombination tuple $\{r_\lambda\}_{\lambda \in \Lambda}$, trajectories always increase mean fitness monotonically.*

Beginning from any such interior point of $\Delta(\Gamma)$ the haploid sexual replicator dynamic converges asymptotically to a monomorphic population in which each individual bears a genotype g_ν from the set N_{W_Θ} of pure Nash equilibria of the associated potential game W_Θ.

The conclusion is that whether or not there is genetic linkage, in haploid sexual reproduction mean fitness increases monotonically and the population always converges to a monomorphic population corresponding to a pure Nash equilibrium of the potential game (this statement also holds true for asexual reproduction, since the equilibrium of maximal mean fitness is itself a pure Nash equilibrium of the potential game). Results similar to those in Theorem 5 and Proposition 7 also attain whether or not there is genetic linkage, with sensitivity to intial conditions as before.

However, although under any genetic linkage structure convergence to some pure Nash equilibrium occurs, the probability of converging to any particular pure Nash equilibrium, starting from the same initial allelic distribution, differs from one linkage structure to another. The same initial point can converge to different equilibria points depending on the linkage structure (as can be seen for example in the extreme case of no recombination, under which convergence will always be to the asexual globally optimal fitness equilibrium.)

5.1.3. *External and Internal Environments.* Let $\mathcal{E}$ be a collection of possible environments. Each $e \in \mathcal{E}$ determines a fitness landscape in the

sense that e is identified with an identical interests game W_e played by the loci, relative to a fixed allelic frequency space Θ.

In this section we will make a distinction between what we term the 'external environment' and the 'internal environment'. The external environment is the realisation $e^1, e^2, \ldots, e^t, \ldots$ and the corresponding $W_{e^1}, W_{e^2}, \ldots, W_{e^t}, \ldots$. Parallel to this, we may take the perspective of any particular locus i. From this perspective, the alleles in locus i are amongst themselves implementing an asexual replicator dynamic as follows. Let $(q_1^t, \ldots, q_m^t)$ denote the profile of allelic frequencies over time. At time t, the identical interests game is W_{e^t}, and we may write the time t growth rate as the stage t game payoff $w_{e^t}(q_1, \ldots, q_m)$. Let $q_{-i}^t = (q_1^t, \ldots, q_{i-1}^t, q_{i+1}^t, \ldots, q_m^t)$ denote the profile of the $m-1$ loci apart from i. Call the sequence $(q_{-i}^1, q_{-i}^2, \ldots)$ the *internal environment* from the perspective of locus i.

Define the *total environment* from the perspective of locus i at time t to be $w_{e^t}(\cdot; q_{-i}^t)$, meaning that each choice of $q_i \in \Delta(A_i)$ yields the payoff $w_{e^t}(q_i; q_{-i})$. In this way we may reduce the dynamic of each locus to the asexual replicator, with the alleles in locus i implementing the asexual dynamic with respect to the total environment from their perspective.

	a_{21}	a_{22}	a_{23}
a_{11}	$w_{a_{11},a_{21}} = .40$	$w_{a_{11},a_{22}} = .60$	$w_{a_{11},a_{23}} = .80$
a_{12}	$w_{a_{12},a_{21}} = .48$	$w_{a_{12},a_{22}} = .55$	$w_{a_{12},a_{23}} = .75$
a_{13}	$w_{a_{13},a_{21}} = .20$	$w_{a_{13},a_{22}} = .51$	$w_{a_{13},a_{23}} = .70$

	a_{21}	a_{22}	a_{23}
a_{11}	$w_{a_{11},a_{21}} = .40$	$w_{a_{11},a_{22}} = .48$	$w_{a_{11},a_{23}} = .20$
a_{12}	$w_{a_{12},a_{21}} = .60$	$w_{a_{12},a_{22}} = .55$	$w_{a_{12},a_{23}} = .51$
a_{13}	$w_{a_{13},a_{21}} = .80$	$w_{a_{13},a_{22}} = .75$	$w_{a_{13},a_{23}} = .70$

Table 1. Two examples of fitness matrices for haploid reproduction with two loci and three alleles per locus

5.1.4. *Virtual Convergence.* Table 1 presents examples of two fitness matrices with two loci and three alleles per locus. One may interpret these as representing a population that is exposed to two possible environments, one per matrix, where the top is is interpreted as a 'rainy year' environment and the bottom one is a 'drought year' environment.

The top matrix has one pure Nash equilibrium, at (a_{11}, a_{23}), and the bottom matrix similarly one at (a_{13}, a_{21}). It is possible, however, for an environment realisation to create a trajectory that almost always remains extremely close to points that are not Nash equilibria – even though from the perspective of both matrices there is a repulsion from those points. For example, there can be a trajectory that begins close to (a_{13}, a_{21}) and remains there under an alternating realisation, rainy one year and drought the next. This is because although at each time period the single period dynamic pushes away from (a_{13}, a_{21}), the directions of push away from that point are nearly opposite, hence the trajectory never wanders far.

Alternatively, it is possible to imagine environment realisations in which very long stretches of one environment bring the population very nearly to convergence to (a_{11}, a_{23}), followed by equally long stretches subsequently driving the population very nearly to convergence to (a_{13}, a_{21}), repeated periodically in such a way that there is no candidate even for 'near convergence'. These simple examples indicate that the behaviour under time varying fitness can be complex and highly dependent on initial conditions and random realisations. Nevertheless, we can state the following theorem on virtual convergence. The intuition behind it is that even when the external and internal environments change upredictably, from the perspective of each individual locus the internal alleles are implementing the simple replicator equation, hence each locus experiences virtual convergence.

Theorem 7. *The hapoid sexually reproducing dynamic virtually converges monomorphically under any environment realisation.*

Note that although the result of Theorem 7 guarantees virtual monomorphic convergence under every environment realisation, different realisations can lead to different virtual growth rates. If the external environment process is sufficiently regular, however, then every realisation will lead to the same virtual convergence, even though the corresponding internal environment process may not follow parallel regularity.

Proposition 10. *If the external environment follows a stationary and ergodic stochastic process then every environment realisation leads to the same virtual monomorphic growth rate.*

6. Diploid Sexual Reproduction

6.1. **Single Locus Model.**

6.1.1. *Fixed Fitness.* The diploid sexual reproduction in the single locus model is extremely similar to the haploid two locus model, which enables many of the results from haploid sexual reproduction to be carried over almost entirely (and arguably relatively simply, since the parallel is to two loci and not m loci) but for one very significant difference: where in the haploid two locus model fitness is represented by a potential matrix with a separate set of alleles for the row player and the column player (corresponding to different alleles in the different loci), in the single locus diploid model the same alleles appear as both row players and column players in the matrix.

The state space is the allelic frequency space $\Theta = \Delta(A)$, with the trajectories in $\Delta(A)$ recursively following the replicator equation. The fitness landscape[3] determined by fitness W_{ij} corresponding to gamete $a_i a_j$, denoted here as before by W_Θ is a symmetric matrix.

Dynamics with respect to symmetric matrices have long been studied in the literature of evolutionary game theory. The parallels are clear: both the diploid single locus and the population dynamic cases can be thought of as a single player game, in which the player selects a mixed strategy (e.g., the ratio of hawks to doves in the population, or the ratio of allele A to allele B), receives an expected payoff, and in the next time period updates the mixed strategy in accordance with a replicator equation. In the evolutionary game theory literature, it is well known that such a dynamic leads to convergence to evolutionarily stable Nash equilibria, which may be either pure or mixed Nash equilibria.

In comparison with the haploid sexual model, the significant new element is, of course, the possibility of convergence to mixed Nash equilibria. One possible evolutionary advantage of maintaining mixed equilibria versus convergence to pure equilibria is that mixed equilibria may be similar to constantly re-balanced portfolios in investment theory; it is well known that re-balanced portfolios (analogous to mixed equilibria) can significantly outperform single stock portfolios (analogous to pure equilibria).

There is one major difference between the dynamics of the diploid sexual reproduction and the population dynamics of a typical evolutionary game theory model. In the fixed fitness/matrix setting, evolutionary game theory models can frequently exhibit cyclic or chaotic trajectories that never converge. In contrast, diploid single locus dynamic trajectories are generically monotonically increase in

[3] We assume here that there are no position effects.

mean fitness and converge, to either a stable monomorphic population (pure Nash equilibrium) or a stable polymorphic population (mixed Nash equilibrium).

This is because the diploid single locus dynamic follows the Baum–Eagon inequality, montonically increasing in mean fitness. Theorem 1 is not directly applicable here, because of the non-linearity in the payoffs to the players with respect to the payoff matrix. One can instead show directly from the equations of motion that the Baum–Eagon inequality holds. This result is known in the literature (see for example [Edwards, 2000]); we reproduce it here for completeness.

Recall that by Equation (11), $W_i = \sum_j p_i W_{ij}$ and that by Equation (10), $\overline{W} = \sum_i (p_i^2 W_{ii} + \sum_{j \neq i} p_i p_j W_{ij})$. Calculating the partial derivative, $\frac{\partial \overline{W}}{\partial p_i} = 2p_i W_{ii} + \sum_{j \neq i} 2p_j W_{ij} = 2(p_i W_{ii} + \sum_{j \neq i} p_j W_{ij})$ (taking into account the fact that i will appear both in W_{ij} and W_{ji} for each j).

It follows that $\sum_j p_i \frac{\partial W_i}{\partial p_i} = 2(\sum_i (p_i^2 W_{ii} + \sum_{j \neq i} p_i^t p_j^t W_{ij})) = 2\overline{W}$. Hence

$$\begin{aligned} p_i \frac{\frac{\partial \overline{W}}{\partial p_i}}{\sum_j p_j \frac{\partial \overline{W}}{\partial p_j}} &= p_i \frac{2(p_i W_{ii} + \sum_{j \neq i} p_j W_{ij})}{2\overline{W}} \\ &= p_i \frac{2W_i}{2\overline{W}} = p_i \frac{W_i}{\overline{W}} \\ &= p_i^+, \end{aligned}$$

with the last equality following Equation (12). Hence the Baum–Eagon Theorem applies to the dynamic in $\Delta(A)$ and one concludes that in the single locus diploid sexual reproduction dynamic fitness monotonically increases until a local maximum is attained.

The upshot is that, apart from the possible convergence to stable polymorphism when mixed strategies are the end result of the dynamic, the diploid single locus model parallels the haploid two-locus model in the crucial aspects of monotonic Baum–Eagon mean fitness increase while following a replicator recursion. This enables us to adapt many of the results from the haploid analysis to the diploid model.

Theorem 8. *In the diploid single-locus model, for each ESS equilibrium ν, there exists $B_\nu \subset \Delta(\Gamma)$ containing 1_{g_ν} such that starting from any initial point in B_ν the population under the dynamic will converge to a monomorphic population consisting solely of genotype g_ν, i.e., $T_1^n(x) \longrightarrow 1_{g_\nu}$ for every $x \in B_\nu$. Apart from separatrices between these basins of*

attraction, which are of negligible measure, the sets in the collection $\{B_\nu\}$ *form a partition of* $\Delta(\Gamma)$.

Proposition 11. *Let* $S_t(x) = tT_1(x) + (1-t)x$. *In the diploid single-locus model, for each ESS equilibrium* ν, *there exists a neighbourhood* $H_\nu \subset \Delta(\Gamma)$ *of* 1_{g_ν} *such that* $S_t(H_\nu) \subset H_\nu$ *for* $0 < t \leq 1$, *and for every* $x \in H_\nu$, $T_1^n(x) \longrightarrow g_\nu$. *Furthermore, each* H_ν *has the homotopy type of a disk.*

6.1.2. *Temporally Variable Fitness.* The time varying fitness results of the haploid two-locus model similarly carry over to the diploid single-locus model.

Theorem 9. *The single-locus diploid sexually reproducing dynamic under varying fitness environments virtually converges polymorphically under any environment realisation.*

6.2. **Multiple Locus Model.** The diploid multi-locus model under linkage equilibrium re-capitulates the diploid single locus model at every locus. Hence the results of the previous section apply without changes.

Many of the tools of the previous sections fail to apply in the diploid multi-locus model under linkage disequilibrium. The main equation of motion, Equation (13), is similar to but not quite a replicator equation. More to the point, the presence of the disequilibrium term causes the dynamic to fail to conform to the Baum–Eagon conditions. Hence the Baum–Eagon theorem, even under fixed fitness, cannot be used to conclude that monotonic fitness increase occurs; indeed, it has long been known that there are examples of fitness landscapes in which diploid populations can exhibit reductions in mean fitness over stretches of time and even periodically cycling trajectories. This leaves no possibility for a general theorem on either monomorphic or polymorphic convergence.

However, virtual convergence (possibly polymorphic), which does not depend on monotonic fitness increase, does obtain here under both fixed and temporally varying fitness. The proof is essentially the same as the proof of haploid virtual convergence (Theorem 7) under temporally varying fitness. As before, we suppose a choice of environment realisation $e^1, e^2, \ldots$ from a set of possible environments, and distinguish between the external environment, represented by such a realisation, and the internal environment perceived by a locus i, which is the allelic frequencies of the other loci at any time t.

Theorem 10. *The multi-locus diploid sexually reproducing dynamic virtually converges polymorphically under any environment realisation.*

APPENDIX A. BAUM–EAGON DYNAMICS

We make use extensive use here of the Baum–Eagon inequality (originally developed for the study of hidden Markov models by use of the Baum–Welch algorithm). The concepts and results in this section are from [Baum and Eagon, 1967] and [Baum and Sell, 1968]. A brief exposition on the Baum–Eagon inequality with applications to population genetics appears in [Edwards, 2000].

Baum–Eagon Inequality. Let Θ be a polytope given by a cross product of simplices,[4] i.e., $\Theta := \Delta_1^k \times \ldots \times \Delta_m^k$. Denote the j-th element in the i-th simplex by x_{ij}.

Let $U(x_{ij})$ be a real-valued polynomial function with non-negative coefficients over the variables $\{x_{ij}\}_{i,j}$. Let $\mathbf{x}$ be a point in the domain Θ. Let $T(\mathbf{x})$ denote the point of Θ whose i,j-th coordinate is given by

$$T(\mathbf{x})_{ij} := x_{ij} \frac{\left.\frac{\partial U}{\partial x_{ij}}\right|_{\mathbf{x}}}{\sum_s x_{is} \left.\frac{\partial U}{\partial x_{is}}\right|_{\mathbf{x}}}, \tag{24}$$

where the denominator is a normalising element.

Then $U(T(\mathbf{x})) > U(\mathbf{x})$ unless $T(\mathbf{x}) = \mathbf{x}$. ∎

It is possible to give the Baum–Eagon inequality an interesting gradient interpretation. Fix i, i.e., concentrate on the i-th simplex, with each element of Δ_i^k denoted as a tuple $\mathbf{x}^i := (x_{i1}, x_{i2}, \ldots, x_{ik})$. From the perspective of Δ_i^k, U may be considered to be a 'potential' function, involving $(x_{i1}, x_{i2}, \ldots, x_{ik})$ and other parameters.

Consider a Euclidean gradient vector derived from the potential in this perspective, that is, $\nabla U(\mathbf{x}^i) = (\frac{\partial U}{\partial x_{i1}}, \frac{\partial U}{\partial x_{i2}}, \ldots, \frac{\partial U}{\partial x_{i1}})$. Then the transformation of Equation (24) can be considered as mapping $\mathbf{x}^i$ to $\Delta U(\mathbf{x}^i) \cdot \mathbf{x}^i$ for each i separately, followed by projection to the simplex by way of the normalisation. In a sense, the Baum–Eagon dynamic is an application of a form of 'gradient hill climbing', locally within each simplex of the polytope Θ, that taken together ensures a global climb.

[4] As before, in greater generality it is possible to allow each simplex to be of different dimension and attain the same results. For simplicity of exposition, we restrict here to the special case in which all the simplices are of the same dimension.

APPENDIX B. PROOFS

Proof of Proposition 1. This is a straight-forward application of Birkhoff's ergodic theorem. Again we register the log fitness. By the ergodic theorem, for each g, $\lim_{t\to\infty} \frac{1}{t}\ln(w_g^t(\omega)) = E(\ln w_g \mid \mu)$ with probability one. Hence the genotype g^* with the greatest expected log fitness (which is also the one with the greatest expected fitness) dominates, as it grows at the fastest average rate. ∎

Proof of Theorem 2. Consider the field $F_t^g = \sigma(\rho_1^g, ..., \rho_t^g)$ and the random variable

$$Z_t^g := \sum_{s=1}^{t} \rho_s^g - \sum_{s=1}^{t} E[\rho_{s+1}^g | F_t^g]. \tag{25}$$

This is a martingale since

$$E[Z_{t+1}^g - Z_t^g | F_t^g] = E\left[\rho_{t+1}^g | F_t^g\right] - E\left[\rho_{t+1}^g | F_t^g\right] = 0, \tag{26}$$

and it is bounded by assumption. Thus, by the Azuma-Hoeffding inequality:

$$Pr(|t^{-1}Z_t^g| \geq ct^{-1/4}) \leq \exp\left(-\frac{c^2 t^{3/2}}{c^2 t}\right) = \exp(-\sqrt{t}). \tag{27}$$

Let E_t be the event that $|t^{-1}Z_t^g| \geq ct^{-1/4}$. Then we have shown that

$$\sum_{t=1}^{\infty} Pr(E_t) \leq \sum_{t=1}^{\infty} \exp(-\sqrt{t}) < \infty. \tag{28}$$

Using the Borel-Cantelli Lemma we deduce

$$Pr\left(\bigcap_{s\geq 1}\bigcup_{t\geq s} E_t\right) = 0. \tag{29}$$

Notice that

$$\bigcap_{s\geq 1}\bigcup_{t\geq s} E_t = \{\left|t^{-1}Z_t^g\right| \geq ct^{-1/4} \ \ \forall t \geq 1\}, \tag{30}$$

implying that almost surely $\frac{1}{t}Z_t^g \to 0$ as $t \to \infty$. By the definition of Z_t^g we deduce that almost surely

$$\lim_{t\to\infty}\left(\frac{1}{t}\sum_{s=1}^{t}\rho_s^g - \frac{1}{t}\sum_{s=1}^{t} E[\rho_{s+1}^g | F_s^g]\right) = 0. \tag{31}$$

Thus, by adding $\limsup_{t\to\infty} \frac{1}{t}\sum_{s=1}^{t} E[\rho^g_{s+1}|F^g_s] = \limsup_{t\to\infty} \widehat{\rho}^g_t$ to both sides of the equals sign in Equation (31), we obtain that almost surely

$$\limsup_{t\to\infty} \frac{1}{t}\sum_{s=1}^{t} \rho^g_s \leq \limsup_{t\to\infty} \widehat{\rho}^g_t, \tag{32}$$

and by an entirely similar argument

$$\liminf_{t\to\infty} \frac{1}{t}\sum_{s=1}^{t} \rho^g_s \geq \liminf_{t\to\infty} \widehat{\rho}^g_t. \tag{33}$$

Next, recall that by assumption g is asymptotically one-step-ahead superior on average, meaning that by definition $\liminf \widehat{\rho}^g_t > \limsup \widehat{\rho}^h_t$ for all genotypes $h \neq g$. Combining this with the inequalities in (32) and (33), which holds for every genotype, one obtains that for every genotype h

$$\liminf_{t\to\infty} \frac{1}{t}\sum_{s=1}^{t} \rho^g_s > \limsup_{t\to\infty} \frac{1}{t}\sum_{s=1}^{t} \rho^h_s. \tag{34}$$

This is sufficient to deduce the statement of the theorem. ∎

Proof of Corollary 4. As $Pr(A^\infty_g) > 0$ by assumption, we can consider the process ψ^g_t which is the restriction of ψ_t to A^∞_g. The corollary then follows by applying Theorem 2 to the process ψ^g_t. ∎

Proof of Theorem 3. As shown in [Edhan et al., 2017], the replicator is an instantiation of Hedge, the exponential version of the multiplicative weights update algorithm. It follows that the replicator attain asymptotic zero regret.

Translating this mathematical result back to the evolutionary setting, this is equivalent to stating that asexual reproduction virtually converges monomorphically to an optimal-in-hindsight genotype g^*. ∎

Proof of Lemma 5. Focus on a particular allele a_{ij_i} and its attendant q_{ij_i}. Recall that by Equation (2), $\overline{w} = \sum_{g\in\Gamma} d_g w_g$, that by Equation (7), $w_{ij_i} = \sum_{g\in C_{a_{ij_i}}} w_g d_{g,i_{j_i}}$, while by Equation (20) the haploid sexual replicator is $q^+_{ij_i} = q_{ij_i}\frac{w_{ij_i}}{\overline{w}}$.

For $g \notin C_{a_{ij_i}}$, one has $\frac{\partial(d_g w_g)}{\partial q_{ij_i}} = 0$. For $g \in C_{a_{ij_i}}$, using $d_g = q_{1j_1}q_{2j_2}\dots q_{mj_m}$ (by linkage equilibrium) yields

$$q_{ij_i}\frac{\partial(d_g w_g)}{\partial q_{ij_i}} = q_{1j_1}q_{2j_2}\dots q_{mj_m} w_g = d_g w_g.$$

Hence $\frac{\partial \overline{w}}{\partial q_{ij_i}} = \sum_{g \in C_{a_{ij_i}}} w_g d_{g,ij_i} = w_{ij_i}$. It follows that $q^+_{ij_i} = q_{ij_i} \frac{\partial \overline{w}}{\partial q_{ij_i}} / \overline{w}$. This is the schema for applying the Baum–Eagon theorem of Equation (24), with T_1 as the transformation and $\overline{w}$ the Lyapunov function. ∎

Proof of Theorem 4. By Lemma 5, monotonic mean fitness increase along trajectories is immediate from the Baum–Eagon Theorem.

Any pure strategy profile of the game W_Γ (corresponding to a point distribution concentrated on a single allele at each locus) constitutes a fixed point of Equation (8), and in fact the only fixed points of this dynamic are pure strategy profiles. However, if g is a pure strategy profile of w that is not a Nash equilibrium then it is not a stable point of the dynamic; using standard dynamics arguments involving nullclines and separatrices, there exists around g a neighbourhood such that any $\overline{w}$-increasing trajectory with initial point in the interior of that neighbourhood eventually leaves that neighborhood.

In other words there is a basin of repulsion around every such point; hence the dynamic cannot converge to non-Nash equilibria points. Convergence will therefore always be to a pure strategy Nash equilibrium point, i.e., a local maximum of the potential $\overline{w}$. ∎

Proof of Theorem 5 This is more or less a corollary of Theorem 4. For each pure Nash equilibrium $\nu \in N_{W_\Theta}$, let B_ν be the set of elements in $\Delta(\Gamma)$ that asymptotically converge to 1_{g_ν}. By definition, this forms an asymptotically stable basin of attraction. ∎

Proof of Proposition 7. This is an almost direct application of Theorem 3 from [Baum and Sell, 1968]. We can identify H_ν as follows: for any $\eta > 0$, let V_η be the connected component in $\Delta(\Gamma)$ of $\{x \in \Delta(\Gamma) \mid w(x) > w(1_{g_\nu}) - \eta\}$ that contains 1_{g_ν}.

Let $\eta_0 > 0$ be the smallest real number such that $\overline{V}_{\eta_0}$, the closure of V_{η_0}, contains another critical point of w in addition to 1_{g_ν}. Set $H_\nu = \bigcup_{\eta < \eta_0} V_\eta$. Theorem 3 of [Baum and Sell, 1968] now applies to H_ν to attain the conclusion. ∎

Proof of Proposition 9. Suppose that at time t the population has mean fitness $\overline{w}$. A sexually reproducing sub-population of weight r is selected, whose mean fitness is $\overline{w}_r$, and the complementary asexual sub-population of weight $1-r$ therefore has mean fitness $\overline{w}_{1-r}$, such that $\overline{w} = r\overline{w}_r + (1-r)\overline{w}_{1-r}$. At time $t+1$, the offspring population of the sexual reproducers has mean fitness $\overline{w}^+_r$, and the offspring population of the asexual reproducers has mean fitness $\overline{w}^+_{1-r}$.

Since both the T_0 and the T_1 transformations increase mean fitness (except at fixed points), it follows that $\overline{w}_r^+ > \overline{w}_r$ and $\overline{w}_{1-r}^+ > \overline{w}_{1-r}$. But $\overline{w}^+ = r\overline{w}_r^+ + (1-r)\overline{w}_{1-r}^+$, hence $\overline{w}^+ > \overline{w}$.

This argument relies on the T_1 increasing mean fitness monotonically, which rests on Theorem 4, which in turn ultimately relies on Theorem 1. In game theoretic terms, Theorem 1 presumes independent choices of strategies on the parts of the players (as opposed to correlated strategies), hence an initial point of linkage equilibrium neeeds to be assumed. Starting from any such interior point, the trajectory will follow increasing mean fitness until it arrives at a local maximum, which will be a pure Nash equilibrium point. ∎

Proof of Theorem 6. Suppose that at time t the population has mean fitness $\overline{w}$. For each partition $\lambda \in \Lambda$, a sub-population of weight r_λ of reproduction type λ is selected, whose mean fitness is $\overline{w}_\lambda$ such that $\overline{w} = \sum_{\lambda\in\Lambda} r_\lambda \overline{w}_\lambda$. At time $t+1$, the offspring population of the λ-type reproducers has mean fitness $\overline{w}_\lambda^+$, with population mean fitness $\overline{w}^+ = \sum_{\lambda\in\Lambda} r_\lambda \overline{w}_\lambda^+$.

Using similar argumentation as in the proof of Proposition 9, since the T_λ transformations increase mean fitness (except at fixed points) for all partitions λ, it follows that $\overline{w}_\lambda^+ > \overline{w}_\lambda$ for all partitions λ. Hence $\overline{w}^+ > \overline{w}$.

Starting from any linkage equilibrium point, the trajectory will follow increasing mean fitness until it arrives at a local maximum, which will be a pure Nash equilibrium point. ∎

Proof of Theorem 7. In this proof we ask a different question from the usual questions of regret minimisation: instead of asking whether an algorithm attains the same asymptotic rate as the best expert, we suppose that the algorithm converges to the rate of the best expert and ask whether that rate is equal to some exogenous rate.

Let $e^1, e^2, \ldots$ be any environment realisation. Let i be a locus. By Equation (8), the alleles in locus i are each implementing a replicator equation (with respect to the identical interests game they are playing against the alleles in the other loci). This implies that the reproductive dynamic internal to the locus follows a multiplicative weights updating algorithm with respect to the total environment payoffs, taking into account both external and internal environments. Hence, regret minimisation applies to this dynamic and in the limit the locus attains the average growth rate it would have attained had it implemented the optimal fixed strategy-in-hindsight within $\Delta(A_i)$

at all times. In other words, the individual locus attains virtual convergence.

From here the proof proceeds inductively. Suppose that under the true dynamic each locus i exhibits the mixed strategy sequence $\mu^i = (\mu_i^1, \mu_i^2, \ldots)$, where $\mu_i^t \in \Delta(A_i)$ for each t. Further, denote by L the lim sup average growth rate payoff that is attained under the profile $(\mu^1, \mu^2, \ldots, \mu^m)$ of these strategy sequences (which is equal for each locus).

Let a_1^* represent the pure strategy of locus 1 whose lim inf attains asymptotically zero regret, as in Equation (14), i.e, that asymptotically does as well as L. Locus 2 can then take the perspective of facing an environment consisting of the external environment along with internal environment $(a_1^*, \mu_i^3, \ldots)$, and virtually attain the same payoff with fixed optimal-in-hindsight a_2^*, i.e., $(a_1^*, a_2^*, \mu_i^3, \ldots)$ virtually attains L.

By induction, the sequence $(a_1^*, a_2^*, \ldots, a_{j-1}^*, \mu_i^j, \ldots)$ attains L. At locus j, implement the optimal-in-hindsight a_j^* against $(a_1^*, a_2^*, \ldots, a_{j-1}^*, \mu_i^{j+1}, \ldots)$ to attain payoff L under $(a_1^*, a_2^*, \ldots, a_{j-1}^*, a_j^*, \mu_i^{j+1}, \ldots)$.

Continuing by induction, in this way eventually one concludes that population consisting entirely of the genotype $g^* = a_1^* a_2^* \ldots a_m^*$ is the optimal-in-hindsight genotype. ∎

Proof of Proposition 10. Following the same reasoning as in previous proofs, consider the perspective of locus i. The payoff received by locus i is equal to what it would gain if all the other loci were to play the pure strategy profile $q_{-i} = (a_1^*, \ldots, a_{i-1}^*, a_{i+1}^*, \ldots, a_m^*)$.

By assumption the external environment process selecting the realisations $W_{e^1}, W_{e^2}, \ldots$ is stationary and ergodic. Since the q_{-i} profile is virtually pure and fixed throughout time, the internal environment $w_{e^t}(\cdot; q_{-1})$ reflects the external environment and is similarly stationary and ergodic. Hence the virtual convergence of locus i, which is equivalent to that of an asexually reproducing population under the conditions of a stationary and ergodic environment process, is always to the same payoff. This holds equally true for all loci. ∎

Proof of Theorem 8. The proof is the same as the proof of Theorem 5. ∎

Proof of Proposition 11. The proof is the same as the proof of Proposition 7. ∎

Proof of Theorem 9. By Equation (12), the alleles in the locus are implementing a replicator equation (with respect to the symmetric potential game they are playing amongst themselves). This implies that the reproductive dynamic in the locus follows a multiplicative weights updating algorithm with respect to the total environment payoffs, taking into account both external and internal environments. Hence, regret minimisation applies to this dynamic and in the limit the locus attains the average growth rate it would have attained had it implemented the optimal fixed strategy-in-hindsight within $\Delta(A)$ at all times. ∎

Proof of Theorem 10. Let $e^1, e^2, \ldots$ be any environment realisation. Let i be a locus. By Equation (12), the alleles in locus i are each implementing a replicator equation (with respect to the symmetric potential game they are playing amongst themselves). This implies that the reproductive dynamic internal to the locus follows a multiplicative weights updating algorithm with respect to the total environment payoffs, taking into account both external and internal environments. Hence, regret minimisation applies to this dynamic and in the limit the locus attains the average growth rate it would have attained had it implemented the optimal fixed strategy-in-hindsight within $\Delta(A_i)$ at all times. In other words, the individual locus attains virtual (possibly polymorphic) convergence.

From here the proof proceeds inductively as in the proof of Theorem 7. Suppose that under the true dynamic each locus i exhibits the mixed strategy sequence $\mu^i = (\mu_i^1, \mu_i^2, \ldots)$, where $\mu_i^t \in \Delta(A_i)$ for each t. Further, denote by L the lim sup average growth rate payoff that is attained under the profile $(\mu^1, \mu^2, \ldots, \mu^m)$ of these strategy sequences (which is equal for each locus).

Letting s_1^* represent the fixed (possibly mixed) strategy of locus 1 whose lim inf attains asymptotically zero regret, as in Equation (14), i.e, that asymptotically does as well as L. Locus 2 can then take the perspective of facing an environment consisting of the external environment along with internal environment $(s_1^*, \mu_i^3, \ldots)$, and virtually attain the same payoff with fixed optimal-in-hindsight s_2^*, i.e., $(s_1^*, s_2^*, \mu_i^3, \ldots)$ virtually attains L.

Continuing argument by induction, in this way eventually one concludes that population consisting entirely of the genotype $g^* = s_1^* s_2^* \ldots s_m^*$ is the optimal-in-hindsight genotype. ∎

APPENDIX C. PROOF OF CONVERGENCE IN POTENTIAL GAMES UNDER THE POLYNOMIAL MULTIPLICATIVE WEIGHTS UPDATE ALGORITHM

C.1. **Preliminary Setup.** Let (A, u, ϕ) be a potential game, where $A = A_1 \times \ldots \times A_m$ is the set of action profiles, $u : A \to \mathbb{R}^m$ the payoff function, and $\Phi : A \to \mathbb{R}$ the potential. For $x \in A$, we use x_i to denote the action in x of the player i and x_{-i} to denote the actions in x of the players apart from player i. For simplicity we will assume that $|A_i| = k$ uniformly for all players; the extension to the more general case is straightforward. Enumerating the elements of A_i, the j-th action of player i is a_{i_j}.

A mixed strategy of player i will be denoted $q_i = (q_{i_1}, \ldots, q_{i_k})$, and a profile of strategies $q = (q_1, \ldots, q_m) \in \Delta(A_1) \times \ldots \times \Delta(A_m)$. The application of a profile q yields an expected payoff for player i that we will denote $u_i(q)$. Given an action profile $x = (x_{1j_1}, x_{2j_2}, \ldots, x_{mj_m}) \in A$ and a profile of strategies q, denote

$$q_x = \prod_{1 \leq \nu \leq m} q_{\nu_{j_\nu}} \tag{35}$$

and

$$q_{x_{-i}} = \prod_{1 \leq \nu \leq m; \nu \neq i} q_{\nu_{j_\nu}}. \tag{36}$$

Suppose that each player is applying the multiplicative updates algorithm to update the mixed strategy he uses from one time period to the next. To interpret what is meant by this, we need to specify the payoff player i receives for placing weight q_{i_j} on action a_{i_j} when his overall expected payoff is $u_i(q)$.

Suggestively borrowing notation introduced earlier here in the context of alleles, denote $C_{a_{ij}} = \{x \in A \mid x_i = a_{ij}\}$, i.e., the set of action profiles with the action of player i fixed at a_{i_j}. Next suppose that player i fixes action a_{i_j} while the other players choose mixed strategies q_{-i}. In this case, denote the expected payoff for player i by $u_{i_j}(q_{-i})$, which is

$$u_{i_j}(q_{-i}) := \sum_{x \in C_{a_{ij}}} q_{x_{-i}} u_i(x). \tag{37}$$

Denote by $u_{i_j}(q_{i_j}, q_{-i})$ the payoff player i receives for placing weight q_{i_j} on action a_{i_j} when the other players choose q_{-i}. Using Equation

(37), this is

$$u_{i_j}(q_{i_j}, q_{-i}) := q_{i_j} u_{i_j}(q_{-i}). \tag{38}$$

With that we can specify what it means for each player to apply the multiplicative weights updates algorithm for $\eta > 0$. When q is the profile of mixed strategies, player i views the tuple $(u_{i_1}(q_{i_1}, q_{-i}), \ldots, u_{i_k}(q_{i_k}, q_{-i}))$. In response, the mixed strategy that player i chooses in the next time period is given by

$$q_{i_j}^+ = q_{i_j} \frac{1 + \eta u_{i_j}(q_{-i})}{\sum_h q_{i_h}(1 + \eta u_{i_h}(q_{-i}))} \tag{39}$$

For ease of reading, we will from here express Equation (39) more simply as

$$q_{i_j}^+ \propto q_{i_j}(1 + \eta u_{i_j}(q_{-i})), \tag{40}$$

supressing the denominator whose entire purpose is only to ensure that the result is a probability distribution.

C.2. Proof of Theorem 1. We prove this first in the special case that the potential game is an identical interests game and that the updating rule is the parameter-free case, i.e., for each i, j, $u_i(x) = u_j(x) = \Phi(x)$, so that each player gets the same payoff (the potential) for each profile of actions, and Equation (40) becomes

$$q_{i_j}^+ \propto q_{i_j} u_{i_j}(q_{-i}). \tag{41}$$

From here most of the work is unravelling of definitions. From Equations (37) and (38) and $u_i(x) = \Phi(x)$ we obtain

$$u_{i_j}(q_{-i}) = \sum_{x \in C_{a_{ij}}} q_{x_{-i}} \Phi(x). \tag{42}$$

At the same time, the expected payoff of player i under q is $\Phi(q) = \sum_{x \in A} q_x \Phi(x)$. It follows that for each available action a_{i_j},

$$\frac{\partial \Phi}{\partial q_{i_j}} = \sum_{x \in C_{a_{ij}}} q_{x_{-i}} \Phi(x). \tag{43}$$

Putting it all together yields

$$q_{i_j}^+ \propto q_{i_j} \frac{\partial \Phi}{\partial q_{i_j}},$$

which is exactly what is needed for application of the Baum–Eagon Theorem (since $\Phi(q)$ is a polynomial function of the various probability weights of q). It follows that under this dynamic the value

of Φ increases monotonically from one time period to the next, with convergence to a fixed point that is a Nash equilibrium.

Moving on from the parameter-free case, consider next the more general polynomial update rule

$$q_{i_j}^+ \propto q_{i_j}(1 + \eta u_{i_j}(q_{-i})) = q_{i_j} + q_{i_j}(\eta u_{i_j}(q_{-i})),$$

but still maintain the assumption of an identical interest game, i.e., $u_i(x) = u_j(x) = \Phi(x)$ for all i, j, so that Equation (42) still holds.

Define

$$\Psi(q) = \sum_{1 \le i \le m} \sum_{1 \le j \le k} q_{i_j} + \eta \Phi(q).$$

Then

$$\frac{\partial \Psi}{\partial q_{i_j}} = 1 + \eta \frac{\partial \Phi}{\partial q_{i_j}}, \tag{44}$$

while Equation (43) holds as before, i.e., $\frac{\partial \Phi}{\partial q_{i_j}} = \sum_{x \in C_{a_{ij}}} q_{x_{-i}} \Phi(x)$. Putting it all together yields

$$q_{i_j}^+ \propto q_{i_j} \frac{\partial \Psi}{\partial q_{i_j}} = q_{i_j} + q_{i_j}(\eta u_{i_j}(q_{-i})),$$

which is sufficient for obtaining the result we seek by appeal to the Baum–Eagon Theorem.

Finally, in greatest generality suppose that the game is fully a potential game as opposed to an identical interests game. In this case, it is well known that the potential game can be decomposed into an identical interests game and a dummy game. That is, $u_i(q) = \Phi(q) + D_i(q_{-i})$, where Φ is an identical interests game and D_i the payoff to i from the dummy game D, depends solely on q_{-i} but does not change at all with changes in q_i.

It follows that $\frac{\partial u_i}{\partial q_{i_j}} = \frac{\partial \Phi}{\partial q_{i_j}}$: in other words, this case essentially reduces to the case of an identical interests game from the perspective of partial differentiation with respect to the weights of the actions of player i. Hence, with minor modifications the same proof as applied earlier applies to the general case; we omit the obvious details. ∎

APPENDIX D. BASINS OF ATTRACTION AND REPULSION

As proved in [Novak and Barton, 2017], using the Baum–Eagon inequality, the haploid sexual replicator dynamic, under fixed fitness conditions, converges to fixed points of the dynamic.[5]

It is clear from inspection of Equation (8) that any pure strategy profile is a fixed point of the dynamic (and conversely the only fixed points are pure strategy profiles); the set of pure Nash equilibria is a proper subset of the set of pure strategy profiles. Details regarding basins of attraction and repulsion, however, need also to be shown.

We will make use of a collection of relative distribution weights (relative to an allele a_{ij}) $\{e^t_{g,ij}\}_{g\in C_{a_{ij}}}$. To define this, let $e^t_{g,ij} = 0$ if $g \notin C_{a_{ij_i}}$. Otherwise, define

$$e^t_{g,ij} := \frac{d^t_g}{\sum_{g'\in C_{a_{ij_i}}} d^t_{g'}}. \tag{45}$$

In words, $e^t_{g,ij}$ is the weight of gentotype g amongst the genotypes containing a_{ij}. Clearly, $\sum_{g\in C_{a_{ij}}} e^t_{g,ij} = 1$.

When linkage equilibrium holds, we can re-express $e^t_{g,ij}$ associated with allele a_{ij_i} when $g = a_{1j_1}a_{2j_2}\dots a_{mj_m}$ as

$$e^t_{g,ij} = \frac{d^t_g}{\sum_{g'\in C_{a_{ij_i}}} d^t_{g'}} = \frac{q^t_{1j_1}q^t_{2j_2}\dots q^t_{mj_m}}{\sum_{g'\in C_{a_{ij_i}}} d^t_{g'}}. \tag{46}$$

Equation (7), defining the marginal fitness of allele a_{ij} at time t is rewritten in these terms as

$$w^t_{ij} := \sum_{g\in C_{a_{ij}}} w^t_g e^t_{g,ij}. \tag{47}$$

It will be useful to express Equation (47) solely in terms of $\{q_i\}$ and $\{w_g\}$. To that end, for a fixed $a_{ij_i} \in \mathcal{A}_i$, enumerate the elements of $C_{a_{ij_i}}$ as $(g_1, \dots, g_l)$. Each such $g \in C_{a_{ij_i}}$ is by definition a string of alleles $a_{1j_1}a_{2j_2}\dots a_{ij_i}\dots a_{mj_m}$ where a_{ij_i} is the same for each $g \in C_{a_{ij_i}}$ but the other alleles vary from one such genotype to the other.

From Equation (6), under linkage equilibrium, for each such g we have $d^t_g = \prod_{\nu=1}^m q^t_{\nu j_\nu} = q^t_{1j_1}q^t_{2j_2}\dots q^t_{ij_i}\dots q^t_{mj_m}$. For future reference we

[5] This result is actually mentioned, without a detailed proof, all the way back in the original paper by Leonard E. Baum and J. A. Eagon [Baum and Eagon, 1967].

will want to use a 'reduced form' of this expression, defined as

$$d^t_{g,-i} = \prod_{1\leq \nu\leq m, \nu\neq i} q^t_{\nu j_\nu} = q^t_{1j_1} q^t_{2j_2} \cdots \cancel{q^t_{ij_i}} \cdots q^t_{mj_m} \tag{48}$$

by which we mean the $m-1$-fold product that does not include q^t_{ij} in it. Then we can re-write Equation (46) as

$$\begin{aligned} e^t_{g,ij_i} &= \frac{d^t_g}{\sum_{g'\in C_{a_{ij_i}}} d^t_{g'}} = \frac{q^t_{1j_1} q^t_{2j_2} \cdots q^t_{mj_m}}{\sum_{g'\in C_{a_{ij_i}}} d^t_{g'}} \\ &= q^t_{ij_i} \frac{d^t_{g,-i}}{\sum_{g'\in C_{a_{ij_i}}} d^t_{g'}} = \frac{q^t_{ij_i}}{q^t_{ij_i}} \frac{d^t_{g,-i}}{\sum_{g'\in C_{a_{ij_i}}} d^t_{g',-i}} \\ &= \frac{d^t_{g,-i}}{\sum_{g'\in C_{a_{ij_i}}} d^t_{g',-i}}. \end{aligned} \tag{49}$$

Note that this entirely removes dependence on $q^t_{ij_i}$; all dependence is on the allelic frequencies of loci *apart* from locus i. We can go even further. Denote in vector notation $\vec{d}^t_{ij_i,-i} := (d^t_{g_1,-i}, \ldots, d^t_{g_l,-i})$, with respect to the eumeration of $(g_1, \ldots, g_l)$ as the elements of $C_{a_{ij_i}}$. By tracing through the definitions, it becomes clear that the dependence on a_{ij_i} is superfluous. In other words, for $a_{ij_i}, a_{ij'_i} \in \mathcal{A}_i$, two alleles in locus i, one has $\vec{d}^t_{ij_i,-i} = \vec{d}^t_{ij'_i,-i}$. We can therefore denote this vector uniformly as $\vec{d}^t_{-i}$.

Similarly, $\sum_{g\in C_{a_{ij_i}}} d^t_{g,-i} = \sum_{g\in C_{a_{ij'_i}}} d^t_{g,-i}$ for $a_{ij_i}, a_{ij'_i} \in \mathcal{A}_i$. Hence we can choose any one of them and uniformly denote $N_{-i} := \sum_{g\in C_{a_{ij_i}}} d^t_{g,-i}$.

Continuing with the vector notation, denote $\vec{w}_{ij_i} := (w_{g_1}, \ldots, w_{g_l})$ with respect to the enumeration of $(g_1, \ldots, g_l)$ as the elements of $C_{a_{ij_i}}$. In this we cannot avoid dependence on the specific allele a_{ij_i}.

From here we can rewrite Equation (47) as

$$w^t_{ij_i} = \sum_{g\in C_{a_{ij_i}}} w^t_g e^t_{g,ij_i} = \frac{1}{N_{-i}} \langle \vec{d}^t_{-i}, \vec{w}_{ij_i} \rangle \tag{50}$$

using the vector dot product.

Finally, since the mean fitness $\overline{w}^t_i = \sum_{j=1}^k q^t_{ij} w^t_{ij}$, this becomes

$$\overline{w}^t_i = \frac{1}{N_{-i}} \sum_{j=1}^k q^t_{ij} \langle \vec{d}^t_{-i}, \vec{w}_{ij} \rangle. \tag{51}$$

D.1. Basins of Repulsion Around Non-Equilibrium Points. Suppose that fitness values w_g are fixed over time. Let $(a_1^*, a_2^*, \ldots, a_m^*)$ be a pure strategy profile of the game that is not a Nash equilibrium. In the biological interpretation, this corresponds to a population composed solely of genotypes $g = a_1^*, a_2^*, \ldots, a_m^*$, in which case $\overline{w} = w_g = w(a_1^*, a_2^*, \ldots, a_m^*)$.

By definition of a non-equilibrium, there is at least one player/locus $\bar{i}$ such that, if we write $a_{\bar{i}}^* = a_{\bar{i}j}$ then there is $a_{\bar{i}j'}$ such that $w(a_{\bar{i}j'}; a_{-\bar{i}}^*) > w(a_{\bar{i}}^*; a_{-\bar{i}}^*)$. For $\varepsilon > 0$, construct for each player i the mixed strategy q_i^ε that gives weight $1 - \varepsilon$ to pure action $a_i^* \in \mathcal{A}_i$ and spreads the rest of the weight ε amongst all the other actions in $\mathcal{A}_i$.

We claim that for sufficiently small ε, the haploid sexual replicator dynamic at the mixed strategy $(q_1^\varepsilon, \ldots, q_m^\varepsilon)$ draws away from $(a_1^*, a_2^*, \ldots, a_m^*)$.

To see this, note that for $(q_1^\varepsilon, \ldots, q_m^\varepsilon)$ for small ε, for any i and $ij_i \in \mathcal{A}_i$, it is the case that w_{ij_i} is very close to $w(a_{ij_i}; a_{-i}^*)$. This is because by Equation (50) ,

$$w_{ij_i} = \frac{1}{N_{-i}} \langle \vec{d}_{-i}^{\,t}, \vec{w}_{ij_i} \rangle,$$

where each element $d_{g,-i} \in \vec{d}_{-i}$, for $g = a_{1j_1} \ldots a_{ij_i} \ldots a_{mj_m}$, is defined by $d_{g,-i} = \prod_{1 \le \nu \le m, \nu \ne i} q_{\nu j_\nu}^\varepsilon$. By the definition of $(q_1^\varepsilon, \ldots, q_m^\varepsilon)$, for small ε the only relevant $d_{g,-i}$ that has appreciable weight is that associated with $g = a_1^* \ldots a_{ij_i} \ldots a_m^*$, where $d_{g,-i} = \prod_{1 \le \nu \le m, \nu \ne i} (1 - \varepsilon)$. Hence w_{ij_i} is very nearly $w(a_{ij_i}; a_{-i}^*)$.

In particular, $w_{\bar{i}j'} \approx w(a_{\bar{i}j'}; a_{-\bar{i}}^*) > w(a_{\bar{i}}^*; a_{-\bar{i}}^*) \approx \varphi_{\bar{i}^*}$. By continuity, for sufficiently small ε this strict inequality can be guaranteed to hold, $w_{\bar{i}j'} > w_{\bar{i}^*}$.

Since the players are implementing the haploid sexual replicator equation, Equation (8), in updating from one period to the next, when $w_{\bar{i}j'} > w_{\bar{i}^*}$ at time t it follows that at time $t+1$ in the distribution $q_{\bar{i}}^{t+1}$ relatively greater weight will be given to $a_{\bar{i}j'}$ and relatively less to $a_{\bar{i}}^*$. It follows that in the step from t to $t+1$ the dynamic moves away from $(a_1^*, a_2^*, \ldots, a_m^*)$.

This is sufficient to conclude that the dynamic must converge to a pure Nash equilibrium.

D.2. Proof of Proposition 8. Suppose that $a^* = (a_{1j_1^*} a_{2j_2^*} \ldots a_{mj_m^*})$ is a pure strategy Nash equilibrium profile of the game.

Let $B \subset D$ be defined as the collection of mixed strategy profiles $(q_1, \ldots, q_m) \in D$ such that for all i:

(1) $q_{ij_i^*} > q_{ij_i}$ for all $j_i \neq j_i^* \in \mathcal{A}_i$, and
(2) $w_{ij_i^*} > w_{ij_i}$ for all $j_i \neq j_i^* \in \mathcal{A}_i$.

In words, B is the set of mixed strategy profiles such that for each i, within the allelic frequency distribution q_i the greatest weight is placed on $a_{1j_1^*}$ and at the same time the marginal fitness of $a_{1j_1^*}$ is the highest against all its competing alleles in locus i. (Recall that by Equation (7) w_{ij_i} is a function of genotypic density and hence also a function of $(q_1, \ldots, q_m)$.)

We can now proceed to show that B is an exponentially stable basin of attraction around a^*. To see this, first note that B is not empty because the profile a^* is trivially an element of B. By continuity, B is then a neighbourhood of a^*.

Next, let $(q_1^t, \ldots, q_m^t) \in B$ at time t. For any i, since $w_{ij_i^*} > w_{ij_i}$ for all $j_i \neq j_i^*$, by Equation (8) it follows that at $t+1$ the weight of allele $a_{ij_i^*}$, i.e. $q_{ij_i^*}^{t+1}$, relative to $q_{ij_i}^{t+1}$ for any other allele, only increases. So the first condition for $q_{ij_i^*}^{t+1}$ is satisfied.

Next, since a^* is a Nash equilibrium, $w(a^*) > w(a_{ij_i}; a_{-i^*})$ for all $j_i \neq j_i^*$. Recall that $w_{ij_i}^{t+1} = \frac{1}{N_{-i}} \langle \vec{d}_{-i}^{t+1}, \vec{w}_{ij_i} \rangle$. Since the weight $q_{i'j_{i'}^*}^{t+1}$ increases relative to $q_{i'j_{i'}}^{t+1}$ for all i', and $w(a_{ij_i^*}; a_{-i^*}) > w(a_{ij_i}; a_{-i^*})$ for all $j_i \neq j_i^*$, it can only be the case that $\langle \vec{d}_{-i}^{t+1}, \vec{v}_{ij_i^*} \rangle > \langle \vec{d}_{-i}^{t+1}, \vec{v}_{ij_i} \rangle$, i.e., $w_{ij_i^*}^{t+1} > w_{ij_i}^{t+1}$. Hence the second condition is satisfied and it follows that $(q_1^{t+1}, \ldots, q_m^{t+1}) \in B$.

Finally, since for each i the weight $q_{ij_i^*}$ increases relative to any other q_{ij_i} monotonically (and even at an increasing rate), asymptotically starting from any $(q_1, \ldots, q_m) \in B$ the dynamic converges to a^*.

REFERENCES

[Baum and Eagon, 1967] Baum, L. E. and Eagon, J. A. (1967). An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model of ecology. *Bulletin of the American Mathematical Society*, 73(3):360–363.

[Baum and Sell, 1968] Baum, L. E. and Sell, G. R. (1968). Growth transformations for functions on manifolds. *Pacific Journal of Mathematics*, 27:221–227.

[Bürger, 2000] Bürger, R. (2000). *The Mathematical Theory of Selection, Recombination, and Mutation*. Wiley, New York.

[Chastain et al., 2013] Chastain, E., Livnat, A., Papadimitriou, C., and Vazirani, U. (2013). Multiplicative updates in coordination games and the theory of evolution. *Proceedings of ITCS 2013*.

[Chastain et al., 2014] Chastain, E., Livnat, A., Papadimitriou, C., and Vazirani, U. (2014). Algorithms, games, and evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 111(29):10620–10623.

[Edhan et al., 2017] Edhan, O., Hellman, Z., and Sherill-Rofe, D. (2017). Sex with no regrets: How sexual reproduction uses a no regret learning algorithm for evolutionary advantage. *Journal of Theoretical Biology*, 426:67–81.

[Edwards, 2000] Edwards, A. (2000). *Foundations of Mathematical Genetics*. Cambridge University Press, Cambridge.

[Freund and Schapire, 1997] Freund, Y. and Schapire, R. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55:119–139.

[Livnat et al., 2008] Livnat, A., Papadimitriou, C., Dushoff, J., and Feldman, M. (2008). A mixability theory for the role of sex in evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 105(50):19803–19808.

[Meir and Parkes, 2015] Meir, R. and Parkes, D. (2015). On sex, evolution, and the multiplicative weights update algorithm. *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 929–937.

[Novak and Barton, 2017] Novak, S. and Barton, N. H. (2017). When does frequency-independent selection maintain genetic variation? *Genetics*, 207:653–668.

[Palaiopanos et al., 2017] Palaiopanos, G., Panageas, I., and Piliouras, G. (2017). Multiplicative weights update with constant step-size in congestion games: Convergence, limit cycles and chaos. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pages 5874–5884.

[Panageas et al., 2019] Panageas, I., Piliouras, G., and Wang, X. (2019). Multiplicative weights updates as a distributed constrained optimization algorithm: Convergence to second-order stationary points almost always. *Proceedings of the 36th International Conference on Machine Learning*.

SCHOOL OF SOCIAL SCIENCES, UNIVERSITY OF MANCHESTER, MANCHESTER, UNITED KINGDOM.

DEPARTMENT OF ECONOMICS, BAR-ILAN UNIVERSITY, RAMAT GAN, ISRAEL.