

Mentors and Recombinators: Multi-Dimensional Social Learning*

Srinivas Arigapudi[†] Omer Edhan[‡] Yuval Heller[§] Ziv Hellman[¶]

April 12, 2022

Abstract

We study games in which the set of strategies is multi-dimensional, and new agents might learn from multiple mentors. We introduce a new family of dynamics, the recombinator dynamics, which is characterized by a single parameter, the recombination rate $r \in [0, 1]$. The case of $r = 0$ coincides with the standard replicator dynamics. The opposite case of $r = 1$ corresponds to a setup in which each new agent learns each new strategic dimension from a different mentor, and combines these dimensions into her adopted strategy. We present two complete characterisations of the stationary states under these dynamics, and we demonstrate the applicability of the new dynamics to the study of strategic interactions.

*SA is supported in part at the Technion by a Fine Fellowship. YH gratefully acknowledge research support from the European Research Council (#677057), the US-Israel Binational Science Foundation (#2020022) and the Israel Science Foundation (#2443/19). ZH acknowledges research support from the Israel Science Foundation (#1626/18).

[†]Industrial Engineering and Management, Technion, Israel, email:arigapudi@campus.technion.ac.il.

[‡]Department of Economics, University of Manchester, UK, email:omeredhan.idan@manchester.ac.uk

[§]Department of Economics, Bar-Ilan University, Israel, email:yuval.heller@biu.ac.il.

[¶]Department of Economics, Bar-Ilan University, Israel, email:ziv.hellman@biu.ac.il

1 INTRODUCTION

Ben Zoma said: Who is wise? One who learns from every one, as it is written
(Psalms 119:99): 'From all of my teachers I have gained understanding'.

Ethics of the Fathers, Chapter 4

The theory of the dynamics of behavioural traits in populations, developed in the context of imitative games, has been one of the most successful sub-fields of evolutionary game theory. A rich set of results and insights regarding the convergence and the stability of equilibria of population-level traits forms the core of the subject, and have been applied to models of social interactive situations, economic models, and biological models of evolution and ecology.

Our contribution here begins with the observation that in many of the replicator population game models in the literature to date the traits that are the main objects of study are dimensionless, and player types are usually entirely identified with single traits. This is far from realistic; each of us is composed of an ensemble of traits and behaviours, from the way we speak and dress to the subjects that we study and the professions that we choose. In a broad sense, each of us may be identified with those trait ensembles, and social and professional success often depends on the entire suite of those traits.

The standard replicator paradigm of evolutionary game theory typically posits a population with a set of traits, while each agent in the population (which may be an individual, a firm, or a social or economic unit) bears only a single trait from that set. A two-player symmetric normal-form game exists, and a trait is associated with a population-dependent fitness value, representing the payoff a player with that trait would receive on average playing against the current distribution of traits in the population.

Here we extend this model to what we term the recombinator model by positing that there is a set of dimensions of cardinality $|D|$, where each dimension is itself a set of traits. An agent is now associated with a $|D|$ -tuple of traits, one from each dimension. Such a tuple of traits is a type, and each type is assigned a fitness payoff.

A newly born agent can either sample a single mentor, as in the standard replicator model, and then imitate the entirety of that mentor's type for his own type, or alternatively be what we term a combinator, who independently samples $|D|$ incumbent agents, each of whom is a mentor. For each $1 \leq d \leq |D|$, the combinator agent imitates the d -th trait of the d -th mentor.

As an example, consider a newly formed commercial enterprise. It may entirely emulate another

successful company, copying every aspect of the mentor company's corporate structure and strategic practices; this would be replicator imitation. Alternatively, in combinator imitation, the newly formed company would copy the corporate governance of one mentor company, the brand management of a second company, and the marketing and customer services approach of a third company.

For each $r \in [0, 1]$, which we term the recombination rate, we obtain a recombinator model, in which a newly born agent is a combinator with probability r and a replicator with probability $1 - r$. In this way one obtains a recombinator law of motion equation; when $r = 0$ the recombinator equation reduces to the standard replicator equation, and when $r = 1$ a pure combinator model is attained. From here, we may inquire about trajectories, convergence, stability, and similar properties.

The tuple of traits comprising the type of an agent and the recombinator model were inspired by considering models of DNA and genetic recombination in biological evolution. Interestingly, one of the central insights of the past century in biological evolution, the gene-centred view in which the interactions of genes are paramount over those of individuals for tracing long-term trajectories, is paralleled in the social learning setting of this paper, where the analogy is a traits-centred view.

We can identify two games occurring in parallel in our model: at a visible level, agents interact in the two-player normal form game G_P , which determines their population-dependent payoffs. More subtly, there is also a game being played between the traits: if we regard each dimension as a player whose action set is the set of traits of that dimension, then a multi-population common interests game emerges. In this context one can interpret each trait as competing against the other traits within its dimension, analogous to the competition between genetic alleles.

This view of parallel types and traits games is not simply a modelling convenience; it is essential for fully analysing trajectories and asymptotic convergence results in recombinator models. For one thing, when $r > 0$ the recombinator dynamics may violate payoff monotonicity, a property that is necessary for many of the standard results in imitative games. This can lead to stable states in which all agents have strictly dominated types.

To remedy this we introduce what we term the r -payoff function, which combines the effects of the combinator and replicator components of the dynamics into a single vector field along which trajectories in the recombinator model flow. Our first main result (Proposition 1 shows that the dynamics are monotone with respect to these r -payoffs, and uses this to present our first "if and only if" characterisation of the set of stationary states.

The traits-centric perspective also proves to be very fruitful and insightful. In this perspective, it is not the simplex of the relative proportions of the types that is of interest but rather the polytope of

the cross product of the simplices of the relative proportions of the traits within each dimension.

Despite what might appear at first sight to be a significant complication, this actually leads to cleaner results. For example at the types level, at the stationary state of a convergent trajectory of the recombinator dynamic the surviving types may exhibit different payoffs. In contrast, at the traits level, the dynamics reliably select for traits with higher payoffs (in ensemble with other traits), with lower performing traits eventually becoming extinct. At stationary state convergence, all surviving traits have exactly the same payoff. This is formalized in our second main result (Proposition 2), which presents a second “if and only if” characterisation of stationary states in terms of the traits payoffs and the correlation in the population between different traits.

Real-life social interactions, whether between individuals or economic units, are typically multi-dimensional, as opposed to the single-dimensional quality of many of the existing replicator dynamics in the literature. We therefore expect that the recombinator model and the solution concept of recombinator equilibrium can serve in future research efforts as a tool for understanding observed behaviour.

1.1 Related Literature

Our research belongs to the evolutionary game theory literature (pioneered in the seminal paper of [Maynard-Smith and Price, 1973](#), with an earlier brief discussion in John Nash’s unpublished dissertation, see [Weibull, 1994](#)). This literature considers a game that is played over and over again by biologically or socially conditioned players who are randomly drawn from large populations. Occasionally, new agents join the population (or incumbents revise their behaviour), and they learn how to play based on observing the (possibly noisy) behaviour and payoffs of some of the incumbent agents (see, [Weibull, 1997](#) and [Sandholm, 2010](#) for a textbook introduction, and [Newton, 2018](#) for a comprehensive recent survey of the literature).

A commonly applied dynamic to capture how the aggregate behaviour gradually changes in such a learning process is the *replicator dynamic* ([Taylor and Jonker, 1978](#)) in which the relative (per capita) change in the proportion of agents playing each action a (henceforth, a -agents) is proportional to the average payoff of the a -agents. Although the replicator dynamic was originally developed to describe natural selection in a genetic evolution, it has been successfully applied to many situations of social learning (examples for various applications of the replicator dynamic and its extensions include [Börgers and Sarin, 1997](#); [Hopkins, 2002](#); [Skyrms, 2004](#); [Cressman and Tao, 2014](#); [Sawa and Zusai, 2014](#); [Mertikopoulos and Sandholm, 2018](#)). In particular, the replicator dynamic models imitative processes

in which new agents imitate the behaviour of successful incumbents (mentors), with the probability that a specific mentor is chosen to be imitated proportional to that mentor's payoff (Björnerstedt and Weibull, 1994).¹

The replicator dynamic is part of a broad family of dynamics that satisfy *payoff monotonicity*: the relative growth of action a is larger than the relative growth of action a' if and only if a receives higher payoff than a' . The so-called Folk Theorem of evolutionary game theory (see, e.g., Nachbar, 1990; Hofbauer and Sigmund, 2003) states that there are close relations between stable states in this broad family of monotone dynamics and Nash equilibria, namely, that stable stationary points are Nash equilibria of the game, interior trajectories converge to Nash equilibria, and strict Nash equilibria are asymptotically stable.²

The vast majority of the evolutionary game theory literature assumes that the learning process is one-dimensional. In particular, most of the literature on imitative processes assumes that a new agent mimics the behaviour of only a single mentor. In what follows we describe the relatively small literature that deals with multi-dimensional learning, in which new agents may combine the learning of various traits from different mentors.³

The combination of different traits in a new agent in this social learning situation resembles the way that genetic inheritance is passed through generations in sexual inheritance: namely, just as a newly formed individual inherits DNA that combines genes from both of her parents (in contrast to the standard replicator dynamic which resembles asexual inheritance), a new agent in our social learning model combines traits from several mentors. The local stability of phenotypic behaviour that is determined by the combination of genes (at different loci in the DNA) has long been studied in the biological literature, see, e.g., Karlin, 1975; Eshel and Feldman, 1984; Matessi and Di Pasquale, 1996).

Waldman (1994) studies a setup in which the action of each agent is two-dimensional, where each dimension reflects a finite choice regarding the level of a different bias; for example, the first dimension may reflect the amount of overconfidence, and the second dimension the amount of disutility

¹ Experimental evidence for the predictions induced by imitative dynamics (similar to the replicator dynamic) is presented in Oprea et al. (2011); Cason et al. (2014); Hoffman et al. (2015); Benndorf et al. (2016, 2021).

² The folk theorem holds under a weaker property than monotonicity, namely, weak payoff positivity ((Weibull, 1994, Proposition 4.11)). Weak payoff positivity is the condition that if there exist some actions that yield strictly higher payoffs than the average payoff in the population, then at least one of these actions has a positive growth rate. This weak assumption holds in many dynamics, such as better reply dynamics (Arieli and Young, 2016), and best reply dynamics (Hwang and Newton, 2017; Babichenko, 2018; Sawa and Zusai, 2019).

³ Our notion of multi-dimensional learning should not be confused with Arieli and Mueller-Frank's (2019) different use of the same phrase. In their setup the phrase describes agents who take actions sequentially, and the order in which actions are taken is determined by a multi-dimensional integer lattice rather than a line as in the standard model of herding.

from work. [Waldman](#) shows that a pair of biases can be evolutionarily stable under sexual inheritance if the level of each bias is optimal when taking the level of the other bias as fixed ([Waldman](#) calls such pairs “second-best adaptations”). [Frenkel et al. \(2018\)](#), extended this analysis to a setup in which the level of each bias is a continuum, and shows that in that case although second-best adaptations do not exist, biases, which approximately compensate for the errors that any one of them would give rise to in isolation, may persist for relatively long periods.⁴

In the last decade some research papers have studied the relation between evolutionary dynamics and learning algorithms (see, e.g., [Livnat et al., 2008](#); [Chastain et al., 2013, 2014](#); [Barton et al., 2014](#); [Meir and Parkes, 2015](#)). These papers showed that sexual inheritance achieves ‘regret minimisation’ learning and convergence to Nash equilibria, and that it is helpful to regard genetic alleles in separate loci as playing a common interest game.

[Edhan et al. 2017](#), point out an important advantage of sexual inheritance over asexual inheritance in setups in which the set of possible genetic combinations is large. In such setups, individuals in any generation can only bear a tiny sample of the large genotype space. An asexual population samples once and finds a local maximum within that sample. In contrast, by continuously re-sampling, the sexual population more reliably attains an asymptotically globally superior action. As [Edhan et al. \(2017\)](#) put it, the strong random element in genetic recombination reproduction is ‘a feature, not a bug’, enabling unbiased sampling while at the same time the population climbs a mean fitness payoff gradient.

[Palaiopanos et al. \(2017\)](#) studied the learning behaviour of the polynomial multiplicative weights update (MWU) algorithm. In that paper they showed that interior trajectories always converge to pure Nash equilibria in congestion games in which each player separately applies an MWU algorithm. [Edhan et al. \(2021\)](#), noting that the replicator is a special case of the MWU algorithm and that genetic recombination reproduction can be cast as a potential game between genetic loci separately implementing a replicator dynamic, build on a similar result to show that haploid sexually reproducing populations exhibit monotonic increase in mean payoff and converge to pure Nash equilibria.

A key difference between our proposed research and the existing literature is that the latter focuses on situations in which an agent’s payoff essentially depends only on her own action, independently of the aggregate behaviour in the population. The strategic aspect of the payoff structure (i.e., the fact that an agent’s payoff crucially depends on the behaviour of other agents in the population)

⁴ Other related models that explain how pairs of biases, which approximately compensate for each other, can be stable are [Herold and Netzer \(2011\)](#); [Steiner and Stewart \(2016\)](#); [Netzer et al. \(2021\)](#).

is a key factor in our model. This dependency of the agent's payoff on the behaviour of others yields qualitatively different results.

2 MODEL

2.1 Basic Setup

Let $G_P = (A, u)$ be a two-player symmetric normal-form game, where A is a (finite) set of actions and $u : A^2 \rightarrow \mathbb{R}^{++}$ is a payoff function. We interpret $u(a, a') \in \mathbb{R}^+$ as the payoff of an agent playing action a against one playing action a' .

A continuum of agents of mass one is presumed. Each agent is associated with an action $a \in A$; this identification is called the type of the agent. The state space is the simplex $\Delta(A)$, where we interpret a population state (abbr., *state*) $\alpha \in \Delta(A)$ as a distribution of types with $\alpha(a)$ expressing the frequency of agents in the population playing a . The payoff (fitness) of an agent of type a in state α is

$$(1) \quad u_\alpha(a) := \sum_{a' \in A} \alpha(a') u(a, a'),$$

and the average (mean) payoff in state α is

$$(2) \quad u_\alpha := \sum_{a \in A} \alpha(a) u_\alpha(a) = \sum_{a, a' \in A} \alpha(a) \alpha(a') u(a, a').$$

We define a state α to be a *Nash equilibrium* if $u_\alpha \geq u_\alpha(a)$ for each $a \in A$; i.e, in a Nash equilibrium no action can induce a payoff that is greater than the average payoff.

Moving from static descriptions to a dynamic law of motion, the time $t \geq 0$ is continuous. We denote an initial state of a trajectory by α^0 and the state at time t along this trajectory by α^t . The derivative with respect to time is denoted by $\dot{\alpha} := \frac{d\alpha}{dt}$.

At each time t there is a flow one of agents who die regardless of their type. Each dying agent is replaced by a new agent (or, equivalently, one can think of this flow as capturing agents who occasionally revise their strategies).

2.2 Strategic Dimensions

If we were describing the standard imitation interpretation of the replicator dynamics (Weibull 1997, Section 4.4.3), we would at this point imagine that a newly born agent selects a single 'mentor' whose

Table 1: Payoff Matrix of a Partially-Enforceable Prisoner's Dilemma

	<i>sc</i>	<i>sd</i>	<i>ac</i>	<i>ad</i>
<i>sc</i>	6, 6	6, 5	6, 5	2, 7
<i>sd</i>	5, 6	5, 5	5, 5	1, 7
<i>ac</i>	5, 6	5, 5	5, 5	1, 7
<i>ad</i>	7, 2	7, 1	7, 1	3, 3

action is imitated. However, we introduce here an extension to the replicator dynamics, by positing that there is a set $D = \{1, \dots, |D|\}$ of dimensions of behaviours, with $|D| \geq 2$. A typical dimension will be denoted by d . Each $d \in D$ is associated with a finite set A_d of traits, and a typical trait in dimension d will be denoted $a_d \in A_d$. We also write $A_{-d} := \prod_{d' \neq d} A_{d'}$.

The set of actions A , which in most of the evolutionary game theory literature is just a collection of elements lacking internal features, is defined here as follows: each element $a \in A$ is now defined to be a D -tuple, i.e., $a = (a_d)_{d \in D} = (a_1, \dots, a_{|D|})$. The interpretation is that the set of traits within the set A_d of dimension d are mutually exclusive; an agent can exhibit only one trait $a_d \in A_d$. In contrast, traits in different dimensions are complementary. A type a is now identified with a full tuple $a = (a_d)_{d \in D}$.

We demonstrate the multi-dimensionality of types in the following example.

Example 1 (The Partially-Enforceable Prisoner's Dilemma). Consider an interaction in which each player simultaneously makes two choices:

1. making either a simple contract-enforceable promise (abbreviated, \mathfrak{s}) or an ambiguous non-enforceable promise (abbreviated, \mathfrak{a}) to cooperate, and
2. cooperating (abbr., \mathfrak{c}) or defecting (abbreviated, \mathfrak{d}) in a prisoner's dilemma.

The set of actions is two-dimensional and includes $4 = 2^2$ actions: $A = (\mathfrak{sc}, \mathfrak{sd}, \mathfrak{ac}, \mathfrak{ad})$, where the first dimension describes the type of promise, and the second dimension describes the behaviour in the prisoner's dilemma. An ambiguous promise induces a cost of 1. When both players cooperate, they obtain a payoff of 6. When a player's promise is ambiguous, a player can gain 2 by defecting, and in this case her opponent loses 4. By contrast, when a player's promise is simple, she loses 1 from defecting (without affecting her opponent's payoff.) The payoff matrix is summarised in Table 1. Note that action \mathfrak{ad} strictly dominates action \mathfrak{sc} , which, in turn, strictly dominates the two remaining actions \mathfrak{sd} and \mathfrak{ac} . ◆

2.3 Frequencies of Traits and Payoffs

The payoff function $u : A \times A \rightarrow \mathbb{R}^+$ remains as before, as do the notations $u_\alpha(a)$ and u_α from Equations (1) and (2). To this we add new expressions. Let the (marginal) frequency of trait a_d in state α be denoted by

$$\alpha(a_d) := \sum_{a_{-d} \in A_{-d}} \alpha(a_d, a_{-d}).$$

Let $\text{supp}(\alpha)$ (resp., $\text{supp}_d(\alpha)$) denote the set of actions (resp., traits) with positive frequency in state α , that is,

$$(3) \quad \text{supp}(\alpha) := \{a \in A \mid \alpha(a) > 0\}, \quad \text{supp}_d(\alpha) := \{a_d \in A_d \mid \alpha(a_d) > 0\}.$$

Let $\text{Int}(A)$ denote the set of interior (full-support) states, i.e., $\text{Int}(\Delta(A)) = \{\alpha \in \Delta(A) \mid \text{supp}(\alpha) = A\}$. Let $\overline{\text{supp}}(\alpha)$, which we call the *rectangular closure of the support* of state α , denote the set of actions such that each of the traits in each action has positive frequency, i.e.,

$$(4) \quad \overline{\text{supp}}(\alpha) := \{a \in A \mid \alpha(a_d) > 0 \text{ for all } d \in D\}.$$

It is immediate that $\text{supp}(\alpha) \subseteq \overline{\text{supp}}(\alpha)$. We say that α has *rectangular support* if $\text{supp}(\alpha) = \overline{\text{supp}}(\alpha)$. Let $\text{Rec}(\Delta(A))$ denote the set of states with rectangular support. Note that any interior state has rectangular support, i.e., $\text{Int}(\Delta(A)) \subseteq \text{Rec}(\Delta(A))$. These notions of support are illustrated by revisiting Example 1.

Example 1 (continued). Let α_1 be the state that places weight $\frac{1}{2}$ on each of the types \mathfrak{sc} and \mathfrak{ad} . Observe that state α_1 does not have rectangular support: $\text{supp}(\alpha_1) = \{\mathfrak{sc}, \mathfrak{ad}\} \neq \overline{\text{supp}}(\alpha_1) = A$. Let α_2 be the state that places weight $\frac{1}{2}$ on each of the types \mathfrak{sc} and \mathfrak{sd} . Then state α_2 has rectangular support: $\text{supp}(\alpha_2) = \overline{\text{supp}}(\alpha_2) = \{\mathfrak{sc}, \mathfrak{sd}\}$. ◆

For each $a_d \in \text{supp}_d(\alpha)$ define $u_\alpha(a_d)$ to be the mean payoff of agents with trait a_d :

$$(5) \quad u_\alpha(a_d) := \frac{1}{\alpha(a_d)} \sum_{a_{-d} \in A_{-d}} \alpha(a_d, a_{-d}) \cdot u_\alpha(a_d, a_{-d}).$$

The definitions of the trait frequencies and payoffs are illustrated as follows.

Example 1 (continued). Let α place weight 0.2 on type \mathfrak{sc} , 0.2 on \mathfrak{sd} , 0.1 on \mathfrak{ac} , and 0.5 on \mathfrak{ad} . These weights imply that the marginal frequencies are: $\alpha(\mathfrak{s}) = 0.2 + 0.2 = 0.4$, $\alpha(\mathfrak{a}) = 1 - \alpha(\mathfrak{s}) = 0.6$, $\alpha(\mathfrak{c}) = 0.3$, $\alpha(\mathfrak{d}) = 0.7$. A simple calculation shows that the types' payoffs are: $u_\alpha(\mathfrak{sc}) = 4$, $u_\alpha(\mathfrak{sd}) = 3$, $u_\alpha(\mathfrak{ac}) =$

3, $u_\alpha(\mathfrak{s}\mathfrak{d}) = 5$, and that the mean payoff in the population is $u_\alpha = 4.1$. Applying (5) implies that trait payoffs are: $u_\alpha(\mathfrak{s}) = \frac{0.2 \cdot 4 + 0.2 \cdot 3}{0.4} = 3.5$, $u_\alpha(\mathfrak{a}) = \frac{0.1 \cdot 3 + 0.5 \cdot 5}{0.6} = 4.67$, $u_\alpha(\mathfrak{c}) = 3.67$, $u_\alpha(\mathfrak{d}) = 4.42$. \blacklozenge

Next, we observe that the average marginal payoff in each dimension is equal to the average payoff of the population u_α . This is so because:

$$(6) \quad \sum_{a_d \in A_d} \alpha(a_d) u_\alpha(a_d) = \sum_{a_d \in A_d} \sum_{a_{-d} \in A_{-d}} \alpha(a_d, a_{-d}) u_\alpha(a_d, a_{-d}) = \sum_{a \in A} \alpha(a) u_\alpha(a) = u_\alpha.$$

2.4 Recombinator Dynamics

In our model, a new agent may either with probability $1 - r$ select a single incumbent (mentor) and directly imitate all the traits of that mentor, or with probability r sample $|D|$ mentors, one for each dimension, and imitate a trait from each of those mentors. In greater detail, if the new agent has a single mentor, then that mentor is sampled from the current state α . If instead there are multiple mentors, the new agent independently samples $|D|$ different mentors and for each d imitates the d -th trait of the d -th mentor. Importantly, the sampling of mentors is not uniform: the more successful a mentor of type a is at state α , as measured by that mentor's payoff relative to the population average, the more mentees he or she attracts. At the same time, the greater $\alpha(a)$, that is, the greater the proportion of action a in population α , the greater the likelihood that one of the mentors of type a will be selected.

Calling $r \in [0, 1]$ the *recombination rate*, the resulting *recombinator dynamics* is

$$(7) \quad \dot{\alpha}(a) = (1 - r) \frac{\alpha(a) u_\alpha(a)}{u_\alpha} + r \prod_{a_d \in a} \frac{\alpha(a_d) u_\alpha(a_d)}{u_\alpha} - \alpha(a),$$

where the first component is the inflow of new agents who imitate a single mentor, the second component is the inflow of new agents who combine learning from $|D|$ mentors, and the last term is the outflow of dying agents.

When $r = 0$, Equation (7) reduces to:

$$(8) \quad \dot{\alpha}(a) = \frac{\alpha(a) u_\alpha(a)}{u_\alpha} - \alpha(a) = \frac{1}{u_\alpha} \alpha(a) (u_\alpha(a) - u_\alpha),$$

which is replicator dynamics (up to a payoff-dependent rescaling of time; see, [Weibull, 1997](#), Section 4.4.3).

When $r = 1$, we have a model of pure combination of traits at each time by each agent that we call

the *combinator dynamics*, characterised by the equation of motion:

$$(9) \quad \dot{\alpha}(a) = \prod_{a_d \in a} \frac{\alpha(a_d) u_{\alpha}(a_d)}{u_{\alpha}} - \alpha(a).$$

Remark 1. Weibull (1997, Section 4.4.3) presents a more general imitation dynamic (for the case of $r = 0$) in which $u_{\alpha}(a)$ (resp., u_{α}) in Equation (8) is replaced by $w(u_{\alpha}(a))$ (resp., $w(u_{\alpha})$), where $w : \mathbb{R} \rightarrow \mathbb{R}^{++}$ is a strictly monotone function. Our dynamics can capture this general version by a normalisation of the payoff function. That is, if the original payoff function is denoted $\pi : A \rightarrow \mathbb{R}$ (which might be measured in dollars), then $u \equiv w(\pi)$ is the normalised payoff following a monotone transformation to cardinal units, measuring the probability of being chosen as a mentor that is induced by the dollar payoff. \blacklozenge

2.5 Forward Invariance

It is well known that under the replicator dynamics the support of any state remains identical along trajectories at all finite times $t \geq 0$; this property is called *forward invariance*. The support may decrease (but not increase) as $t \rightarrow \infty$ (i.e., if $r = 0$, then $\text{supp}(\alpha^t) = \text{supp}(\alpha^0) \forall t > 0$, and $\lim_{t \rightarrow \infty} \text{supp}(\alpha^t) \subseteq \text{supp}(\alpha^0)$). Moreover, these properties hold in the broader class of imitative dynamics (Sandholm 2010, Section 5.4).

A related property holds for the recombinator dynamics with $r > 0$ with one key difference: the support, if it is not rectangular at time zero, immediately increases to its rectangular closure for any $t > 0$ (as demonstrated in the example below). That is, for any $r > 0$ and any trajectory starting at α^0 :

$$(10) \quad \text{supp}(\alpha^t) = \overline{\text{supp}}(\alpha^0) \text{ for all } t > 0, \quad \text{and} \quad \lim_{t \rightarrow \infty} \text{supp}(\alpha^t) \subseteq \overline{\text{supp}}(\alpha^0).$$

Example 1 (continued). Let the initial state α^0 place positive weights on types \mathfrak{sc} and \mathfrak{ad} (i.e., $\text{supp}(\alpha^0) = \{\mathfrak{sc}, \mathfrak{ad}\}$). Observe that $\text{supp}(\alpha^t) = \overline{\text{supp}}(\alpha^0) = A$ for all $t > 0$. This is because every time a new agent is born there is a positive probability that a mentor pair $(\mathfrak{sc}, \mathfrak{ad})$ will be sampled, leading to the creation of type \mathfrak{sd} , and similarly a positive probability that mentor pair $(\mathfrak{ad}, \mathfrak{sc})$ will be sampled (the ordering makes a difference), leading to the creation of type \mathfrak{ac} . Note that this always holds, even though actions \mathfrak{sd} and \mathfrak{ac} induce strictly dominated payoffs. \blacklozenge

The recombinator dynamics exhibit the same continuous (and forward invariant) behaviour as the imitative dynamics (Sandholm 2010, Section 5.4) at all positive times $t > 0$. Instantaneous discon-

tinuities in the recombinator dynamics can only occur at time zero, and only when the support of the initial state is not rectangular.

3 RESULTS

In this section, we explore the monotonicity properties of the recombinator dynamics, and present two characterizations for its stationary states. We first demonstrate that the recombinator dynamics violates payoff monotonicity. We then define a new payoff function, called r -payoff. We show that the recombinator dynamics is monotone with respect to the r -payoffs, and this induces our first characterization of stationary states. Finally, we show that the induced dynamics on the traits (rather than on the types) does satisfy payoff monotonicity (with respect to the original payoff function u), and use this to derive a second characterization of stationary states in terms of the trait dynamics.

3.1 Non-Monotonicity of the Recombinator Dynamics

Dynamics are payoff monotone (see, e.g., [Weibull 1997](#), Definition 4.2) if a type with a higher payoff grows at a higher rate.

Definition 1. Dynamics $\dot{\alpha}$ are *payoff monotone* if $u_{\alpha}(a) > u_{\alpha}(a') \Leftrightarrow \frac{\dot{\alpha}(a)}{\alpha(a)} > \frac{\dot{\alpha}(a')}{\alpha(a')}$ for each state $\alpha \in \Delta(A)$ and for each pair of types $a, a' \in \text{supp}(\alpha)$.

It is well known that the replicator dynamic satisfies payoff monotonicity. In fact, the property of payoff monotonicity turns out to be a central feature of the replicator dynamic: dynamics that satisfy payoff monotonicity nearly always satisfy theorems essentially identical to those of the replicator dynamic with respect to convergence to equilibria and various modes of stability (see Chapter 4 of [Weibull 1997](#)).

Given this, it is noteworthy that our next example demonstrates that the recombinator dynamics may violate payoff monotonicity when the recombination rate is positive, and that this can allow strictly dominated types to be asymptotically stable.

Example 2 (Non-monotonicity in the Partially-Enforceable Prisoner's Dilemma). Fix a recombination rate $r > \frac{1}{7}$, and a sufficiently small $\varepsilon \ll 1$. Consider an initial state α that puts weight $1 - \varepsilon$ on type \mathfrak{sc} and weight ε on type \mathfrak{ad} . The payoff matrix (Table 1) implies that $u_{\alpha}(\mathfrak{sc}) = 6 - 4\varepsilon < u_{\alpha}(\mathfrak{ad}) = 7 - 4\varepsilon$. In what follows we show that $\frac{\dot{\alpha}(\mathfrak{sc})}{\alpha(\mathfrak{sc})} > \frac{\dot{\alpha}(\mathfrak{ad})}{\alpha(\mathfrak{ad})}$, which violates payoff monotonicity. Observe that $u_{\alpha}(\mathfrak{s}) = u_{\alpha}(\mathfrak{c}) = 6 - 4\varepsilon$, $u_{\alpha}(\mathfrak{a}) = u_{\alpha}(\mathfrak{d}) = 6 - 4\varepsilon$, and $u_{\alpha} = 6 - 3\varepsilon$. Substituting these values in the recombinator

dynamics Equation (7) yields:

$$\dot{\alpha}(sc) = (1-r) \frac{(1-\varepsilon)(6-4\varepsilon)}{6-3\varepsilon} + r \left(\frac{(1-\varepsilon)(6-4\varepsilon)}{6-3\varepsilon} \right)^2 - (1-\varepsilon) = O(\varepsilon) \Rightarrow \frac{\dot{\alpha}(sc)}{\alpha(sc)} = O(\varepsilon),$$

$$\dot{\alpha}(ad) = (1-r) \frac{\varepsilon(7-4\varepsilon)}{6-3\varepsilon} + r \left(\frac{\varepsilon(7-4\varepsilon)}{6-3\varepsilon} \right)^2 - \varepsilon = \varepsilon \frac{7(1-r)}{6} - \varepsilon + O(\varepsilon^2) \Rightarrow \frac{\dot{\alpha}(ad)}{\alpha(ad)} = \frac{1-7r}{6} + O(\varepsilon).$$

Observe that for a sufficiently small ε , $\frac{\dot{\alpha}(sc)}{\alpha(sc)} > \frac{\dot{\alpha}(ad)}{\alpha(ad)}$ iff $r > \frac{1}{7}$. One can show that the strictly dominated type ac is asymptotically stable⁵ if and only if $r > \frac{1}{7}$ (and the only other asymptotically stable type is sd). Figure 1 illustrates the basin of attractions of ac and sd for two values of r : 0.5 and 1.

3.2 r -Payoffs

In this section, we define a new payoff function $z_\alpha^r(a)$, and show that the recombinator dynamics is monotone with respect to $z_\alpha^r(a)$.

The recombinator dynamics (7) can be rewritten as:

$$(11) \quad \dot{\alpha}(a) = \alpha(a)(1-r) \frac{u_\alpha(a)}{u_\alpha} + r \prod_{a_d \in a} \alpha(a_d) \prod_{a_d \in a} \frac{u_\alpha(a_d)}{u_\alpha} - \alpha(a).$$

This implies that for any $a \in \text{supp}(\alpha)$:

$$(12) \quad \frac{\dot{\alpha}(a)}{\alpha(a)} = (1-r) \frac{u_\alpha(a)}{u_\alpha} + r \frac{\prod_{a_d \in a} \alpha(a_d)}{\alpha(a)} \prod_{a_d \in a} \frac{u_\alpha(a_d)}{u_\alpha} - 1.$$

Let us now define for any $a \in \text{supp}(\alpha)$ the *trait-to-type ratio* $m_\alpha(a)$

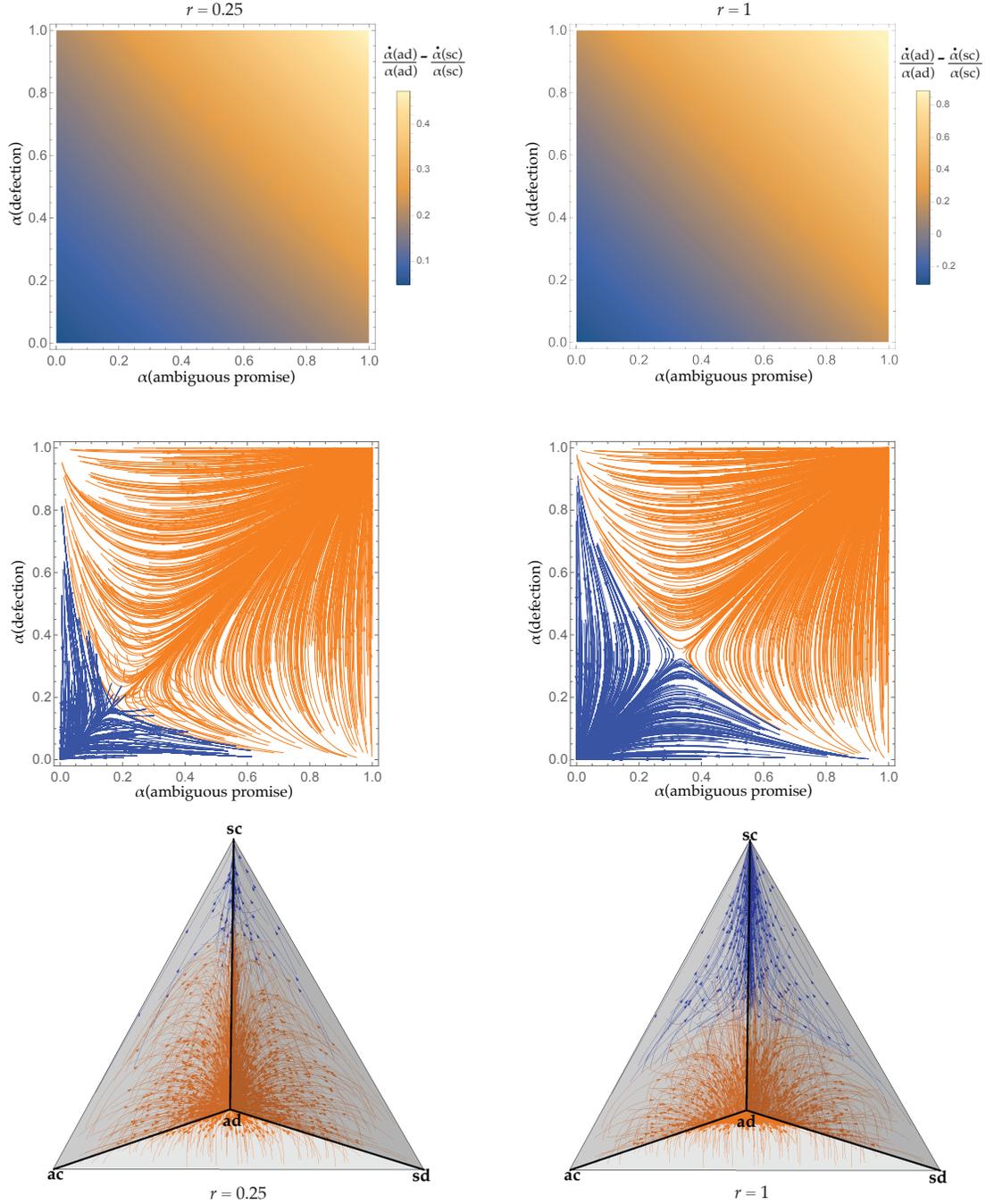
$$m_\alpha(a) := \frac{\prod_{a_d \in a} \alpha(a_d)}{\alpha(a)}.$$

The trait-to-type ratio $m_\alpha(a)$ is the ratio between the product of the weights of the traits in α to the weight of α itself. In states in which the event of a randomly chosen agent having a trait in one dimension (say, being cooperative in Example 1) is independent of that agent having a trait in another dimension (say, giving a simple promise) $m_\alpha(a) \equiv 1$. We call such states *trait independent*.

Definition 2. A state α is *trait independent* if $\alpha(a) = \prod_{a_d \in a} \alpha(a_d)$ for each $a \in A$.

Thus, the trait-to-type ratio $m_\alpha(a)$ captures the distance of $\alpha(a)$ from trait independence. Values

⁵A stationary state is asymptotically stable if any nearby initial state converges to the stationary state.



This figure illustrates the relative growth rates and the basins of attraction in the partially-enforceable Prisoner's Dilemma. In all of the panels, the left side describes recombination rate of $r = 0.4$ and the right side describes $r = 1$. The upper panel illustrates the difference between the relative growth rates of the types ad and sc (i.e., $\frac{\dot{\alpha}(ad)}{\alpha(ad)} - \frac{\dot{\alpha}(sc)}{\alpha(sc)}$) under the recombinator dynamics given trait-independent states. The x (y) axis describes the frequency of trait a (d). The middle panel illustrates the evolution of the projection of the trajectories on the plane defined by the frequencies of the traits a and d . The bottom panel illustrates the evolution of the trajectories (i.e., the phase plot) in the full three dimensional space of $\Delta(A)$. Trajectories that converge to sc (everyone giving a simple promise and cooperating) appear in blue, and those that converge to ad (everyone giving an ambiguous promise and defecting) appear in orange.

Figure 1: Relative Growth Rates and Phase Portraits in Example 1.

of $m_\alpha(a) < 1$ (resp., $m_\alpha(a) > 1$) represent positive correlation between the traits composing type a , which yield that the probability of a randomly chosen agent having type a is larger (resp., smaller), then the product of the probabilities of $|D|$ randomly chosen agents each having one of the traits a_d in a . Define

$$(13) \quad z_\alpha^r(a) := (1-r) \frac{u_\alpha(a)}{u_\alpha} + r m_\alpha(a) \prod_{a_d \in a} \frac{u_\alpha(a_d)}{u_\alpha}.$$

We call $z_\alpha^r(a)$ the r -payoff of action $a \in \text{supp}(\alpha)$ at state α .

Observe that without recombination ($r = 0$) the r -payoff coincides with the standard payoff function $u_\alpha(a)$ (up to a normalisation attained by dividing by the mean payoff u_α): $z_\alpha^0(a) = \frac{u_\alpha(a)}{u_\alpha}$. Further observe that in the opposite case of full recombination ($r = 1$) the r -payoff of a strategy a depends only on the payoffs of its traits (and not of its own payoff $u_\alpha(a)$): $z_\alpha^1(a) = m_\alpha(a) \prod_{a_d \in a} \frac{u_\alpha(a_d)}{u_\alpha}$. In the general case of $r \in (0, 1)$ the r -payoff is a convex combination of these two terms: $z_\alpha^r(a) = (1-r)z_\alpha^0(a) + rz_\alpha^1(a)$.

Substituting z_α^0 and $m_\alpha(a)$ in Equation (12) yields:

$$(14) \quad \frac{\dot{\alpha}(a)}{\alpha(a)} = z_\alpha^r(a) - 1.$$

Equation (14) implies that the recombinator dynamics are monotone with respect to $z_\alpha^r(a)$, and that a state α is stationary if and only if all incumbent types have r -payoff of 1.

Proposition 1. *The recombinator dynamics (8) satisfy the following:*

1. *r -payoff monotonicity: $z_\alpha^r(a) > z_\alpha^r(a') \Leftrightarrow \frac{\dot{\alpha}(a)}{\alpha(a)} > \frac{\dot{\alpha}(a')}{\alpha(a')}$ for each state $\alpha \in \Delta(A)$ and for each pair of types $a, a' \in \text{supp}(\alpha)$.*
2. *A state α is stationary if and only if $z_\alpha^r(a) = 1$ for all $a \in \text{supp}(\alpha)$.*

Proof. Part (1) is immediate from Equation (14). We prove part (2). Suppose that $z_\alpha^r(a) = 1$ for all $a \in \text{supp}(\alpha)$. Then $\dot{\alpha}(a) = 0$ for $a \in \text{supp}(\alpha)$, by Equation (14). Hence each $a \in \text{supp}(\alpha)$ satisfies the property that $\alpha^t(a) = \alpha^0(a)$ for all $t \geq 0$. For the other direction of part (2), suppose that $z_\alpha^r(a) \neq 1$ for some $a \in \text{supp}(\alpha)$. Then by Equation (14), $\dot{\alpha}(a) \neq 0$ and hence stationarity cannot obtain. If α has not got rectangular support, then $\text{supp}(\alpha^0) \subsetneq \overline{\text{supp}(\alpha^0)} = \text{supp}(\alpha^t)$ for all $t > 0$. Once again, this immediately implies that stationarity does not obtain. \square

3.3 Payoff Monotonicity of the Trait-Centric Dynamics

The recombinator dynamics, which is defined over the set of types, induces dynamics over the set of traits. These induced dynamics can be interpreted as a game between the traits (which lies behind the original game between the types G). The gene-centered view of genetic evolution (Williams; Dawkins 1976) highlights the way in which biological natural selection chooses the fitness-maximising genes, rather than choosing fitness-maximising individuals. Similarly, in what follows, we show that the social learning process that is captured by the recombinator dynamics leads to the survival of the payoff-maximising traits (rather than payoff-maximising types), where each trait $a_d \in A_d$ is essentially competing against the other traits in A_d .

Fix any trait in the support of α (i.e., $\alpha(a_d) > 0$). Let us slightly rewrite Equation (11) from the perspective of a particular $a_d \in a$:

$$(15) \quad \dot{\alpha}(a_d, a_{-d}) = (1-r) \frac{\alpha(a_d, a_{-d}) u_\alpha(a_d, a_{-d})}{u_\alpha} + r \frac{\alpha(a_d) u_\alpha(a_d)}{u_\alpha^{|D|}} \prod_{a_{d'} \in a_{-d}} \alpha(a_{d'}) u_\alpha(a_{d'}) - \alpha(a_d, a_{-d}).$$

Dividing Equation (15) by $\alpha(a_d)$ and summing over all $a'_{-d} \in A_{-d}$ yields the following *trait-centric recombinator dynamics*:

$$(16) \quad \frac{\dot{\alpha}(a_d)}{\alpha(a_d)} = \sum_{a'_{-d} \in A_{-d}} \frac{\dot{\alpha}(a_d, a'_{-d})}{\alpha(a_d)} = \frac{(1-r)}{u_\alpha} u_\alpha(a_d) + \frac{r}{u_\alpha^{|D|}} u_\alpha(a_d) \sum_{a'_{-d} \in A_{-d}} \prod_{a_{d'} \in a'_{-d}} \alpha(a_{d'}) u_\alpha(a_{d'}) - 1$$

$$= u_\alpha(a_d) \left(\frac{(1-r)}{u_\alpha} + \frac{r}{u_\alpha^{|D|}} \sum_{a'_{-d} \in A_{-d}} \prod_{a_{d'} \in a'_{-d}} \alpha(a_{d'}) u_\alpha(a_{d'}) \right) - 1$$

Note that the right-hand side of Equation (16) can be decomposed into an expression involving $u_\alpha(a_d)$ and a sum involving only elements of a_{-d} . This implies that the trait-centric recombinator dynamics is monotone in the trait payoffs. This yields a simple trait-centred characterisation of stationary states: (a) all traits obtain the same payoff, and (b) all types obtain the same r -weighted average of the *relative type's payoff* (i.e., the ratio between the type's payoff and the average payoff, $\frac{u_\alpha(a)}{u_\alpha}$) and the trait-to-type ratio, $m_\alpha(a)$.

Proposition 2. *The trait-centric recombinator dynamics (16) satisfy the following properties:*

1. *Payoff monotonicity: $u_\alpha(a_d) > u_\alpha(a'_d) \Leftrightarrow \frac{\dot{\alpha}(a_d)}{\alpha(a_d)} > \frac{\dot{\alpha}(a'_d)}{\alpha(a'_d)}$ for each state $\alpha \in \Delta(A)$ and for each pair of traits $a_d, a'_d \in \text{supp}_d(\alpha)$.*

2. A state α is stationary if and only if

- (a) $u_\alpha(a_d) = u_\alpha$ for any dimension $d \in D$ and any $a_d \in \text{supp}_d(\alpha)$.
- (b) $(1 - r) \frac{u_\alpha(a)}{u_\alpha} + rm_\alpha(a) = 1$ for any $a \in \text{supp}(\alpha)$.

Proof.

1. Part (1) is implied by the fact that replacing a_d with a'_d in the right-hand side of Equation (16) leaves the expression in the bracket unchanged. Appealing to Equation (6), we further conclude that $u_{\alpha^*}(a_d) = u_{\alpha^*}(a'_d) = u_{\alpha^*}$.
2. Observe that (a) and (b) jointly imply that $z_\alpha^r(a) = 1$ for all $a \in \text{supp}(\alpha)$, which implies, by Proposition 1, that α is stationary. For the other direction, suppose that α is stationary. Proposition 1 implies that $z_\alpha^r(a) = 1$ for all $a \in \text{supp}(\alpha)$. By Part (1), the stationarity of α implies that $u_\alpha(a_d) = u_\alpha$ for every d and $a_d \in \text{supp}_d(\alpha)$. Substituting this equality in the definition of $z_\alpha^r(a)$ (13) implies that $(1 - r) \frac{u_\alpha(a)}{u_\alpha} + rm_\alpha(a) = 1$ for any $a \in \text{supp}(\alpha)$. \square

Part (1) of Proposition 2 implies that the recombinator dynamics selects for traits with higher (state-dependent) payoffs, and that traits that are consistently having lower payoffs would become extinct. When the population converges into a stationary state, it must be the case that all surviving traits have exactly the same payoff (note that traits within a dimension are directly competing with each other; traits in separate dimensions are not competing and may be regarded as complementing each other since the underlying game between the dimensions is a common interests game). In contrast, the surviving types in a stationary state may have different payoffs, under the constraint that types with higher payoffs have lower trait-to-type ratios. This is illustrated in the following example.

Example 3 (Heterogeneous stationary state). Consider the partially-enforceable prisoner's dilemma with recombination rate $r = 0.5$. Consider the state α^* (as described in Table 2) in which 53% of the agents have type \mathfrak{sc} , 11% have type \mathfrak{ad} , and 18% have each of the remaining types (\mathfrak{ac} and \mathfrak{sd}). The payoff matrix (Table 1) implies that \mathfrak{ad} has the highest payoff of 6.56, \mathfrak{sc} has payoff of 5.56, and the remaining two types have the lowest payoff of 5.56. Calculating the payoff of each trait (as the weighted mean of the payoffs of the types that has this trait) shows that each trait has the same payoff of 5.31 (satisfying condition (2-a) on Proposition 2. Observe that the average of each type's relative payoff $\frac{u_\alpha(a)}{u_\alpha}$ and its trait-to-type ratio $m_\alpha(a)$ is equal to one, which satisfies condition (2-b).

Table 2: Heterogeneous Stationary State α^* in the Contactable Prisoner's Dilemma

Types in the stationary state α^*					Traits in the stationary state α^*		
a	$\alpha(a)$	$u_\alpha(a)$	$\frac{u_\alpha(a)}{u_\alpha}$	$m_\alpha(a)$	a_d	$\alpha(a_d)$	$u_\alpha(a_d)$
\mathfrak{sc}	53%	5.56	1.05	0.95	\mathfrak{c}	53%+18%=71%	$\frac{0.53}{0.71} \cdot 5.56 + \frac{0.18}{0.71} \cdot 4.56 = 5.31$
\mathfrak{ac}	18%	4.56	0.86	1.14	\mathfrak{d}	18%+11%=29%	$\frac{0.18}{0.29} \cdot 4.56 + \frac{0.11}{0.29} \cdot 6.56 = 5.31$
$\mathfrak{s}\mathfrak{d}$	18%	4.56	0.86	1.14	\mathfrak{s}	18%+53%=71%	$\frac{0.53}{0.71} \cdot 5.56 + \frac{0.18}{0.71} \cdot 4.56 = 5.31$
$\mathfrak{a}\mathfrak{d}$	11%	6.56	1.24	0.76	\mathfrak{a}	11%+18%=29%	$\frac{0.18}{0.29} \cdot 4.56 + \frac{0.11}{0.29} \cdot 6.56 = 5.31$

The left table describes the frequencies, payoffs and trait-to-type ratios of the types in state α^* . The right table describes the frequencies and payoffs of the traits in α^* .

For example, the highest payoff 1.24 of $\mathfrak{a}\mathfrak{d}$ is compensated by having the lowest trait-to-type ratio of 0.76. Thus, Proposition 2 implies that α^* is stationary.

4 CONCLUSION

In this paper we have introduced a new concept of “recombinator dynamics”. These population learning dynamics involve games in which the set of actions is multi-dimensional, and new agents who join the population might learn different strategic dimensions from different mentors (i.e., successful incumbents). This family of dynamics has a single parameter r , where the case of $r = 0$ coincides with the classical replicator dynamics, and the case of $r = 1$ is the case in which each new agent learns each new property from a different mentor.

We have presented two “if and only if” characterisations for stationary states in this setup, and we have applied the analysis to an interesting example of a prisoner’s dilemma game with partially-enforceable contracts. We plan in a work in progress to characterise the subset of these stationary states that are dynamically stable (both locally and globally), and to apply the recombinator dynamics to additional applications.

References

Arieli, I. and M. Mueller-Frank (2019). Multidimensional social learning. *The Review of Economic Studies* 86(3), 913–940.

- Arieli, I. and H. P. Young (2016). Stochastic learning dynamics and speed of convergence in population games. *Econometrica* 84(2), 627–676.
- Babichenko, Y. (2018). Fast convergence of best-reply dynamics in aggregative games. *Mathematics of Operations Research* 43(1), 333–346.
- Barton, N. H., S. Novak, and T. Paixão (2014). Diverse forms of selection in evolution and computer science. *Proceedings of the National Academy of Sciences* 111(29), 10398–10399.
- Benndorf, V., I. Martínez-Martínez, and H.-T. Normann (2016). Equilibrium selection with coupled populations in hawk–dove games: Theory and experiment in continuous time. *Journal of Economic Theory* 165, 472–486.
- Benndorf, V., I. Martínez-Martínez, and H.-T. Normann (2021). Games with coupled populations: An experiment in continuous time. *Journal of Economic Theory*, 105281.
- Björnerstedt, J. and J. W. Weibull (1994). Nash equilibrium and evolution by imitation. Technical report, IUI Working Paper.
- Börgers, T. and R. Sarin (1997). Learning through reinforcement and replicator dynamics. *Journal of economic theory* 77(1), 1–14.
- Cason, T. N., D. Friedman, and E. Hopkins (2014). Cycles and instability in a rock–paper–scissors population game: A continuous time experiment. *Review of Economic Studies* 81(1), 112–136.
- Chastain, E., A. Livnat, C. Papadimitriou, and U. Vazirani (2013). Multiplicative updates in coordination games and the theory of evolution. In *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, pp. 57–58.
- Chastain, E., A. Livnat, C. Papadimitriou, and U. Vazirani (2014). Algorithms, games, and evolution. *Proceedings of the National Academy of Sciences* 111(29), 10620–10623.
- Cressman, R. and Y. Tao (2014). The replicator equation and other game dynamics. *Proceedings of the National Academy of Sciences* 111(Supplement 3), 10810–10817.
- Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press, Oxford, UK.
- Edhan, O., Z. Hellman, and I. Nehama (2021). Making the most of potential: potential games and genotypic convergence. *Royal Society Open Science*.

- Edhan, O., Z. Hellman, and D. Sherill-Rofe (2017). Sex with no regrets: How sexual reproduction uses a no regret learning algorithm for evolutionary advantage. *Journal of theoretical biology* 426, 67–81.
- Eshel, I. and M. W. Feldman (1984). Initial increase of new mutants and some continuity properties of ess in two-locus systems. *The American Naturalist* 124(5), 631–640.
- Frenkel, S., Y. Heller, and R. Teper (2018). The endowment effect as blessing. *International Economic Review* 59(3), 1159–1186.
- Herold, F. and N. Netzer (2011). Probability weighting as evolutionary second-best. *Working paper series/Socioeconomic Institute* (1005).
- Hofbauer, J. and K. Sigmund (2003). Evolutionary game dynamics. *Bulletin of the American mathematical society* 40(4), 479–519.
- Hoffman, M., S. Suetens, U. Gneezy, and M. A. Nowak (2015). An experimental investigation of evolutionary dynamics in the rock–paper–scissors game. *Scientific Reports* 5(1), 1–7.
- Hopkins, E. (2002). Two competing models of how people learn in games. *Econometrica* 70(6), 2141–2166.
- Hwang, S.-H. and J. Newton (2017). Payoff-dependent dynamics and coordination games. *Economic Theory* 64(3), 589–604.
- Karlin, S. (1975). General two-locus selection models: some objectives, results and interpretations. *Theoretical population biology* 7(3), 364–398.
- Livnat, A., C. Papadimitriou, J. Dushoff, and M. W. Feldman (2008). A mixability theory for the role of sex in evolution. *Proceedings of the National Academy of Sciences* 105(50), 19803–19808.
- Matessi, C. and C. Di Pasquale (1996). Long-term evolution of multilocus traits. *Journal of Mathematical Biology* 34(5), 613–653.
- Maynard-Smith, J. and G. R. Price (1973). The logic of animal conflict. *Nature* 246(5427), 15–18.
- Meir, R. and D. C. Parkes (2015). On sex, evolution, and the multiplicative weights update algorithm. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*.

- Mertikopoulos, P. and W. H. Sandholm (2018). Riemannian game dynamics. *Journal of Economic Theory* 177, 315–364.
- Nachbar, J. H. (1990). “evolutionary” selection dynamics in games: Convergence and limit properties. *International journal of game theory* 19(1), 59–89.
- Netzer, N., A. Robson, J. Steiner, and P. Kocourek (2021). *Endogenous Risk Attitudes*. Centre for Economic Policy Research.
- Newton, J. (2018). Evolutionary game theory: A renaissance. *Games* 9(2), 31.
- Oprea, R., K. Henwood, and D. Friedman (2011). Separating the hawks from the doves: Evidence from continuous time laboratory games. *Journal of Economic Theory* 146(6), 2206–2225.
- Palaiopanos, G., I. Panageas, and G. Piliouras (2017). Multiplicative weights update with constant step-size in congestion games: convergence, limit cycles and chaos. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 5874–5884.
- Sandholm, W. H. (2010). *Population games and evolutionary dynamics*. MIT press.
- Sawa, R. and D. Zusai (2014). Evolutionary imitative dynamics with population-varying aspiration levels. *Journal of Economic Theory* 154, 562–577.
- Sawa, R. and D. Zusai (2019). Evolutionary dynamics in multitasking environments. *Journal of Economic Behavior & Organization* 166, 288–308.
- Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge University Press.
- Steiner, J. and C. Stewart (2016). Perceiving prospects properly. *American Economic Review* 106(7), 1601–31.
- Taylor, P. D. and L. B. Jonker (1978). Evolutionary stable strategies and game dynamics. *Mathematical biosciences* 40(1-2), 145–156.
- Waldman, M. (1994). Systematic errors and the theory of natural selection. *The American Economic Review*, 482–497.
- Weibull, J. W. (1994). The mass-action interpretation of nash equilibrium. Technical report, IUI Working Paper.

Weibull, J. W. (1997). *Evolutionary Game Theory*. Cambridge, MA:MIT press.

Williams, G. C. Princeton, NJ: Princeton university press.