Mentors and Recombinators:  
Multi-Dimensional Social Learning*

Srinivas Arigapudi†  Omer Edhan†  Yuval Heller§  
Ziv Hellman¶

Preliminary Partial Draft, April 2022

Abstract

We study games in which the set of strategies is multi-dimensional, and new agents might learn various strategic dimensions from different 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 fully characterise stationary states and stable states under these dynamics, and we show that they predict novel behaviour in various applications.

Keywords: revision protocols, imitation, non-monotone dynamics, evolutionary stability.  JEL codes: C73, D83

---

*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 behavioral 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: 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 helpful for fully analysing trajectories and asymptotic convergence results in re-combinator 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 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 formalised in our second 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.

It is well-known (see Claims 2–3) that a stationary state is asymptotically stable under the standard replicator dynamics essentially if and only if it satisfies:

1. *internal stability*: the payoff matrix restricted to the incumbent types is negative-definite, and
2. **external stability** the payoffs of external types is lower than the incumbents.

Our main results (Theorems 1–2) extend this characterisation to the recombinator dynamics. Specifically, we show that a stationary state is asymptotically stable under the combinator dynamics essentially\(^1\) if and only if

1. **internal stability**: the \( r \)-Jacobian matrix (which replaces the standard payoffs with the \( r \)-payoffs described above) is negative-definite,

2. **external stability against traits**: the payoffs of external traits are lower than those of the incumbents; we show that each small invasion of an external trait induces a unique distribution of types who carry this trait (partners), and this distribution is used to calculate the trait’s payoff.

3. **external stability against types**: the payoffs of external types is lower than the incumbents’ payoff multiplied by \( \frac{1}{1-r} > 1 \), where this factor \( \frac{1}{1-r} > 1 \) reflects the fact that stability against external types become easier the higher the recombination rate \( r \) (because traits of a successful external type recombine with the incumbent traits into hybrid types, which might have lower payoffs).

We apply our dynamics in two simple applications. The first application is a prisoner’s dilemma with partially enforceable contracts in which each player chooses (1) making either a costless enforceable promise or a costly non-enforceable promise to cooperate, and (2) either cooperating or defecting. All recombination rates admit an inefficient stable state in which all players make costly non-enforceable promises and defect. By contrast, if the recombination rate is not too low, the game also admits a Pareto-optimal stable state in which all agents make enforceable promises and cooperate. Our second application (Example 4) is a hawk-dove game which is enriched by adding three feasible approaches: (1-2) two specialised traits: being emotional (rational), which is somewhat helpful for a hawkish (dovish) player, while being harmful for a dovish (hawkish) player, and (3) being versatile (no effect on the game payoffs).

\(^1\) Asymptotic stability implies the weak variants of these 3 conditions, and it is implied by their strict counterparts.
We show that there is a recombination threshold, such that for any recombination rate below (above) this threshold, the unique stable state involves specialised (versatile) 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 in a large variety of applications.

2 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).2

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 the family of payoff 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. 3

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 behavior 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.4 Arad and Rubinstein (2012) and Arad and Penczynski (2018) present experimental evidence that perceive games with large set of strategies as multi-dimensional, and that they think how to play in each strategic dimension, independently of the

---

2 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).

3 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 (Hart, 2002; Arieli and Young, 2016), and best reply dynamics (Hwang and Newton, 2017; Babichenko, 2018; Sawa and Zusai, 2019). The few models that violate Weak payoff positivity include action-sampling dynamics (Sandholm, 2001; Oyama et al., 2015; Arigapudi et al., 2022) and payoff-sampling dynamics (Sethi, 2000; Sandholm et al., 2020; Arigapudi et al., 2021).

4 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.
other dimensions. Motivated by this, Arad and Rubinstein (2019) present a static (non-evolutionary) solution concept that is related to the case of $r = 1$ under the assumption of each player perceiving the distribution of traits as uniform.

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 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.\footnote{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).}

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) points 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 behavior in the population. The strategic aspect of the payoff structure (i.e., the fact that an agent’s payoff crucially depends on the behavior of other agents in the population) is a key factor in our model. This dependency of the agent’s payoff on the behavior of others yields qualitatively different results.
3 Model

3.1 Basic Setup

Let $G = (A, u)$ be a two-player symmetric normal-form game, where $A$ is a (finite) set of actions and $u : A^2 \to \mathbb{R}^{++}$ 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 $\alpha$ is

\begin{equation}
\tag{1}
   u_\alpha(a) := \sum_{a' \in A} \alpha(a') u(a, a'),
\end{equation}

and the average (mean) payoff in state $\alpha$ is

\begin{equation}
\tag{2}
   u_\alpha := \sum_{a \in A} \alpha(a) u_\alpha(a) = \sum_{a, a' \in A} \alpha(a) \alpha(a') u(a, a').
\end{equation}

We define a state $\alpha$ to be a Nash equilibrium if $u_\alpha(a) \geq u_\alpha(a)$ 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 $\alpha^0$ and the state at time $t$ along this trajectory by $\alpha^t$. 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).
3.2 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 action is imitated. However, we introduce here an extension to the replicator dynamics, by positing that there is a set $D = \{1, \ldots, |D|\}$ 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 type $a \in A$ is now defined to be a $D$-tuple, i.e., $a = (a_d)_{d \in D} = (a_1, \ldots, a_{|D|})$.

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. With slight abuse of notation we write $a_d \in a$ when trait $a_d$ is one of the components in the $D$-tuple $a$.

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, $s$) or an ambiguous non-enforceable promise (abbreviated, $a$) to cooperate, and

2. cooperating (abbr., $c$) or defecting (abbreviated, $d$) in a prisoner’s dilemma.

The set of actions is two-dimensional and includes $4 = 2^2$ actions: $A = (sc, sd, ac, ad)$, 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.
Table 1: Payoff Matrix of a Partially-Enforceable Prisoner’s Dilemma

|     | sc  | sd  | ac  | ad  |
|-----|-----|-----|-----|-----|
| sc  | 6,6 | 6,5 | 6,5 | 2,7 |
| sd  | 5,6 | 5,5 | 5,5 | 1,7 |
| ac  | 5,6 | 5,5 | 5,5 | 1,7 |
| ad  | 7,2 | 7,1 | 7,1 | 3,3 |

Note that action ad strictly dominates action sc, which, in turn, strictly dominates the two remaining actions sd and ac. ♦

3.3 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 \( u_\alpha \) from Equations (1) and (2). To this we add new expressions. Let the (marginal) frequency of trait \( a_d \in A_d \) (resp., trait profile \( a_{-d} \in A_{-d} \)) in state \( \alpha \) be denoted by

\[
\alpha(a_d) := \sum_{a_{-d} \in A_{-d}} \alpha(a_d, a_{-d}), \quad \alpha(a_{-d}) := \sum_{a_d \in A_d} \alpha(a_d, a_{-d}).
\]

Let \( \text{supp}(\alpha) \) (resp., \( \text{supp}_d(\alpha) \), \( \text{supp}_{-d}(\alpha) \)) denote the set of actions (resp., traits, trait-profiles) with positive frequency in state \( \alpha \), that is,

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

\[
\text{supp}_{-d}(\alpha) := \{a_{-d} \in A_{-d} \mid \alpha(a_{-d}) > 0\}.
\]

Let \( \text{Int}(\Delta(A)) \) denote the set of interior (full-support) states, i.e., \( \text{Int}(\Delta(A)) = \{(\alpha \in \Delta(A) \mid \text{supp}(\alpha) = A)\} \).

Definition 1. For a given state \( \alpha \), we call the collection of actions such that each of the traits in each action has positive frequency in state \( \alpha \) the rectangular closure of the support of \( \alpha \) and denote it by \( \overline{\text{supp}}(\alpha) \). In detail,

\[
\overline{\text{supp}}(\alpha) := \{a \in A \mid \alpha(a_d) > 0 \text{ for all } a_d \in a\}.
\]
It is immediate that \( \text{supp}(\alpha) \subseteq \overline{\text{supp}}(\alpha) \). We say that \( \alpha \) 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 \( \alpha_1 \) be the state that places weight \( \frac{1}{2} \) on each of the types \( \text{sc} \) and \( \text{ad} \). Observe that state \( \alpha_1 \) does not have rectangular support: \( \text{supp}(\alpha_1) = \{\text{sc}, \text{ad}\} \neq \overline{\text{supp}}(\alpha_1) = \Delta(A) \). Let \( \alpha_2 \) be the state that places weight \( \frac{1}{2} \) on each of the types \( \text{sc} \) and \( \text{sd} \). Then state \( \alpha_2 \) has rectangular support: \( \text{supp}(\alpha_2) = \overline{\text{supp}}(\alpha_2) = \{\text{sc}, \text{sd}\} \).

For each \( a_d \in \text{supp}_d(\alpha) \) define \( u_{\alpha}(a_d) \) to be the mean (marginal) 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 \( \alpha \) place weight 0.2 on type \( \text{sc} \), 0.2 on \( \text{sd} \), 0.1 on \( \text{ac} \), and 0.5 on \( \text{ad} \). These weights imply that the marginal frequencies are: \( \alpha(s) = 0.2 + 0.2 = 0.4 \), \( \alpha(a) = 1 - \alpha(s) = 0.6 \), \( \alpha(c) = 0.3 \), \( \alpha(d) = 0.7 \). A simple calculation shows that the types’ payoffs are: \( u_{\alpha}(\text{sc}) = 4 \), \( u_{\alpha}(\text{sd}) = 3 \), \( u_{\alpha}(\text{ac}) = 3 \), \( u_{\alpha}(\text{sd}) = 5 \), and that the mean payoff in the population is \( u_{\alpha} = 4.1 \). Applying (5) implies that trait payoffs are: \( u_{\alpha}(s) = \frac{0.2 \cdot 4 + 0.2 \cdot 3}{0.4} = 3.5 \), \( u_{\alpha}(a) = \frac{0.1 \cdot 3 + 0.5 \cdot 5}{0.6} = 4.67 \), \( u_{\alpha}(c) = 3.67 \), \( u_{\alpha}(d) = 4.42 \). Next, we observe that the average marginal payoff in each dimension is equal to the average payoff of the population \( u_{\alpha} \). This is so because:

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

### 3.4 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 $\alpha$. 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 $\alpha$, 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 $\alpha$, 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 recombinitr dynamics is

\begin{equation}
\dot{\alpha}(a) = (1 - r) \frac{\alpha(a) u_\alpha(a)}{u_\alpha} + r \prod_{a_d \in a} \alpha(a_d) u_\alpha(a_d) u_\alpha - \alpha(a),
\end{equation}

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.

Time is continuous and is denoted by the superscript $t \geq 0$. The initial state of
a trajectory is denoted by $\alpha^0$, and the dynamics of a trajectory is determined by the
equation of motion given by Equation (7), with $\alpha^t$ substituting for $\alpha$ and \( \alpha^t \equiv \frac{d\alpha^t}{dt} \)
denoting the time derivative of the state.

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

\begin{equation}
\alpha(a) = \frac{\alpha(a) u_\alpha(a)}{u_\alpha} - \alpha(a) = \frac{1}{u_\alpha} \alpha(a)(u_\alpha(a) - u_a),
\end{equation}

which is the 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:

\begin{equation}
\dot{\alpha}(a) = \prod_{a_d \in a} \frac{\alpha(a_d) u_\alpha(a_d)}{u_\alpha} - \alpha(a).
\end{equation}
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., \( u_\alpha \)) in Equation (8) is replaced by \( w(u_\alpha(a)) \) (resp., \( w(u_\alpha) \)), where \( w : \mathbb{R} \to \mathbb{R}^{++} \) 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 \( \pi : A \to \mathbb{R} \) (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.

3.5 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 \( t \to \infty \) (i.e., if \( r = 0 \), then \( \text{supp}(\alpha_t) = \text{supp}(\alpha_0) \forall t > 0 \), and \( \lim_{t \to \infty} \text{supp}(\alpha_t) \subseteq \text{supp}(\alpha_0) \)). 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 \( \alpha^0 \):

\[
\text{supp}(\alpha_t) = \text{supp}(\alpha^0) \text{ for all } t > 0, \quad \text{and} \quad \lim_{t \to \infty} \text{supp}(\alpha_t) \subseteq \text{supp}(\alpha^0).
\]

Example 1 (continued). Let the initial state \( \alpha^0 \) place positive weights on types \( sc \) and \( ad \) (i.e., \( \text{supp}(\alpha^0) = \{sc, ad\} \)). Observe that \( \text{supp}(\alpha_t) = \text{supp}(\alpha^0) = A \) for all \( t > 0 \). This is because every time a new agent is born there is a positive probability that a mentor pair \( (sc, ad) \) will be sampled, leading to the creation of type \( sd \), and similarly a positive probability that mentor pair \( (ad, sc) \) will be sampled (the ordering makes a difference), leading to the creation of type \( ac \). Note that this always holds, even though actions \( sd \) and \( ac \) induce strictly dominated payoffs.
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 discontinuities in the recombinator dynamics can only occur at time zero, and only when the support of the initial state is not rectangular.

3.6 Stability and Convergence

We conclude this section with a few standard definitions of dynamic stability and convergence. A state is stationary if it is a fixed point of the dynamics.

Definition 2. A state $\alpha \in \Delta(A)$ is stationary if $\dot{\alpha}(a) = 0$ for each $a \in A$.

It is well-known that any convergent limit state of a trajectory $\lim_{t \to \infty} \alpha^t$ must be a stationary state (Proposition 6.3 of Weibull, 1997).

A state is attracting if starting from every sufficiently close initial condition, the trajectories converge to that state. More precisely,

Definition 3. A state $\alpha^* \in \Delta(A)$ is attracting if there is an open neighbourhood $U$ of $\alpha^*$ such that all trajectories initially in $U$ converge to $\alpha^*$, i.e., $\alpha^0 \in U \Rightarrow \lim_{t \to \infty} \alpha^t = \alpha^*$.

A state is Lyapunov stable if the trajectory of a population starting out near that state always remains close, and it is asymptotically stable if, in addition, the trajectory eventually converges to the equilibrium state. A state is unstable if it is not Lyapunov stable. It is well known (see, e.g., Weibull, 1997, Section 6.4) that every Lyapunov stable state is stationary.

Definition 4. A state $\alpha \in \Delta(A)$ is Lyapunov stable if for every neighbourhood $U$ of $\alpha^*$ there is a neighbourhood $V \subseteq U$ of $\alpha^*$ such that if the initial state of a trajectory satisfies $\alpha_0 \in V$ then $\alpha^t \in U$ for all $t > 0$. A state is unstable if it is not Lyapunov stable.

Definition 5. A state $\alpha^* \in \Delta(A)$ is asymptotically stable if it is Lyapunov stable and attracting.
The basin of attraction of a state $\alpha^*$, which is denoted by $BA(\alpha^*)$ is the set of initial states that converge to this state, i.e.,

\[ BA(\alpha^*) = \{ \alpha \in A | \alpha^0 = \alpha \Rightarrow \lim_{t \to \infty} \alpha^t = \alpha^* \}. \]

Finally, we say that state $\alpha^*$ is globally stable if its basin of attraction includes all interior states, i.e., $\text{Int}(\Delta(A)) \subseteq BA(\alpha^*)$.

4 Payoff Monotonicity and Stationary States

In this section we explore the monotonicity properties of the recombinator dynamics. We first demonstrate that the recombinator dynamics violates payoff monotonicity. We then define a new payoff function, called $r$-payoff, and we show that the recombinator dynamics is monotone with respect to the $r$-payoffs. 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 $u$).

4.1 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 6.** Dynamics $\dot{\alpha}$ are payoff monotone if $u_\alpha(a) > u_\alpha(a') \iff \frac{\dot{\alpha}(a)}{\alpha(a)} > \frac{\dot{\alpha}(a')}{\alpha(a')}$ 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, which implies that its stationary states are those that satisfy the property that all incumbent types have the same payoff. Formally

**Claim 1.** (Weibull, 1997, Proposition 5.9)

1. The replicator dynamic is payoff monotone.
2. A state $\alpha$ is stationary if and only if $u_\alpha(a) = u_\alpha(a')$ for all $a, a' \in \text{supp}(\alpha)$.

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 $\varepsilon << 1$. Consider an initial state $\alpha$ that puts weight $1 - \varepsilon$ on type $sc$ and weight $\varepsilon$ on type $ad$. The payoff matrix (Table 1) implies that $u_\alpha(sc) = 6 - 4\varepsilon < u_\alpha(ad) = 7 - 4\varepsilon$. In what follows we show that $\dot{\alpha}(sc) > \dot{\alpha}(ad)$, which violates payoff monotonicity. Observe that $u_\alpha(s) = u_\alpha(c) = 6 - 4\varepsilon$, $u_\alpha(a) = u_\alpha(d) = 6 - 4\varepsilon$, 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 $\varepsilon$, $\frac{\alpha(sc)}{\alpha(ad)} > \frac{\alpha(ad)}{\alpha(ad)}$ iff $r > \frac{1}{7}$. We later show that the strictly dominated type $sc$ is asymptotically stable if and only if $r > \frac{1}{7}$ (and the only other asymptotically stable type is $ad$). Figure 1 illustrates the basin of attractions of $ac$ and $sd$ for two values of $r$: 0.25 and 1.

### 4.2 $r$-Payoffs

In this section, we define a new payoff function $z'_a(a)$, and show that the recombinator dynamics is monotone with respect to $z'_a(a)$.

The recombinator dynamics (7) can be rewritten as:

$$\dot{\alpha}(a) = \alpha(a)(1 - r)\frac{u_\alpha(a)}{u_\alpha} + r \prod_{a_d \in a} \alpha(a_d) \prod_{a_d \in a} \frac{u_\alpha(a_d)}{u_\alpha} - \alpha(a).$$
This 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 $r = 0.25$ 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., $\frac{\alpha(ad)}{\alpha(ad)} - \frac{\alpha(sc)}{\alpha(sc)}$) 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.
This implies that for any \( a \in \text{supp}(\alpha) \):

\[
\frac{\dot{\alpha}(a)}{\alpha(a)} = (1 - r) \frac{u_\alpha(a)}{u_\alpha} + r \prod_{a_d \in a} \frac{\alpha(a_d)}{\alpha(a)} \prod_{a_d \in a} \frac{u_\alpha(a_d)}{u_\alpha} - 1.
\]

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 \( \alpha \) to the weight of \( \alpha \) 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 7.** A state \( \alpha \) 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 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

\[
z_\alpha^r(a) := (1 - r) \frac{u_\alpha(a)}{u_\alpha} + rm_\alpha(a) \prod_{a_d \in a} \frac{u_\alpha(a_d)}{u_\alpha}.
\]

We call \( z_\alpha^r(a) \) the \( r \)-payoff of action \( a \in \text{supp}(\alpha) \) at state \( \alpha \). We extend the definition to \( a \notin \text{supp}(\alpha) \) by defining \( z_\alpha^r(a) = 0 \) in such case.

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 \( u_\alpha \)): \( z_\alpha^0(a) = \frac{u_\alpha(a)}{u_\alpha} \). 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 \( m_α(a) \) in Equation (13) yields:

\[
\dot{\alpha}(a) = z_α^r(a) - 1.
\]

Equation (15) implies that the recombinator dynamic is monotone with respect to \( z_α^r(a) \), and that a state \( \alpha \) is stationary if and only if all incumbent types have \( r \)-payoff of 1. This is formalised in Proposition 1, which generalises Claim 1 to recombinator dynamics, where the game payoff function \( u \) is replaced by the \( r \)-payoff function \( z^r \).

**Proposition 1.** The recombinator dynamic of Equation (8) satisfies:

1. \( r \)-payoff monotonicity: \( z_α^r(a) > z_α^r(a') \iff \frac{\dot{\alpha}(a)}{\alpha(a)} > \frac{\dot{\alpha}(a')}{\alpha(a')} \) for each state \( \alpha \in \Delta(A) \) and for each pair of types \( a, a' \in \text{supp}(\alpha) \).

2. A state \( \alpha \) is stationary if and only if \( z_α^r(a) = 1 \) for all \( a \in \text{supp}(\alpha) \).

**Proof.** Part (1) is immediate from Equation (15). We prove part (2). Suppose that \( z_α^r(a) = 1 \) for all \( a \in \text{supp}(\alpha) \). Then \( \dot{\alpha}(a) = 0 \) for \( a \in \text{supp}(\alpha) \), by Equation (15). 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 \( z_α^r(a) \neq 1 \) for some \( a \in \text{supp}(\alpha) \). Then by Equation (15), \( \dot{\alpha}(a) \neq 0 \) and hence stationarity cannot obtain. If \( \alpha \) has not got rectangular support, then \( \text{supp}(\alpha^0) \subsetneq \text{supp}(\alpha') = \text{supp}(\alpha^t) \) for all \( t > 0 \). Once again, this immediately implies that stationarity does not obtain. \( \square \)

### 4.3 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, 1966; Dawkins, 1976) highlights the ways in which biological natural selection chooses 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 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 \(\alpha\) (i.e., \(\alpha(a_d) > 0\)). Let us slightly rewrite Equation (12) from the perspective of a particular \(a_d \in a\):

\[
\dot{\alpha}(a_d, a_{-d}) = (1 - r) \frac{\alpha(a_d, a_{-d})}{u_{\alpha}} u_{\alpha}(a_d, a_{-d}) + r \frac{\alpha(a_d)}{u_{\alpha}} u_{\alpha}(a_{d'}) \prod_{a_{d'} \in a_{-d}} \alpha(a_{d'}) u_{\alpha}(a_{d'}) - \alpha(a_d, a_{-d}).
\]

Dividing Equation (16) by \(\alpha(a_d)\) and summing over all \(a'_{-d} \in A_{-d}\) yields the following trait-centric recombinator dynamics:

\[
\frac{\dot{\alpha}(a_d)}{\alpha(a_d)} = \sum_{a'_{-d} \in A_{-d}} \frac{\alpha(a_d, a'_{-d})}{\alpha(a_d)} = \frac{(1 - r)}{u_{\alpha}} u_{\alpha}(a_d) + \frac{r}{u_{\alpha}} u_{\alpha}(a_d) \sum_{a'_{-d} \in A_{-d}} \prod_{a'_{d'} \in a'_{-d}} \alpha(a'_{d'}) u_{\alpha}(a'_{d'}) - 1
\]

\[
= u_{\alpha}(a_d) \left( \frac{(1 - r)}{u_{\alpha}} + \frac{r}{u_{\alpha}} \sum_{a'_{-d} \in A_{-d}} \prod_{a'_{d'} \in a'_{-d}} \alpha(a'_{d'}) u_{\alpha}(a'_{d'}) \right) - 1.
\]

Note that the right-hand side of Equation (17) 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 (17) satisfy the following properties:

1. Payoff monotonicity: \(u_{\alpha}(a_d) > u_{\alpha}(a'_{d}) \Leftrightarrow \frac{\alpha(a_d)}{\alpha(a_{d'})} > \frac{\alpha(a'_{d})}{\alpha(a_{d'})}\) 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 \(\alpha\) 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) $\left(1 - r\right) \frac{u_d(a)}{u_a} + rm_\alpha(a) = 1$ 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 (17) 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 $\alpha$ is stationary. For the other direction, suppose that $\alpha$ is stationary. Proposition 1 implies that $z_{\alpha}^r(a) = 1$ for all $a \in \text{supp}(\alpha)$. By Part (1), the stationarity of $\alpha$ 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 $z_{\alpha}^r(a)$ (14) implies that $(1 - r) \frac{u_d(a)}{u_a} + rm_\alpha(a) = 1$ for any $a \in \text{supp}(\alpha)$.

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 to 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 $\alpha^*$ (as described in Table 2) in which 53% of the agents have type $sc$, 11% have type $ad$, and 18% have each of the remaining types ($ac$ and $sd$). The payoff matrix (Table 1) implies that $ad$ has the highest payoff of 6.56, $sc$ 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
Table 2: Heterogeneous Stationary State $\alpha^*$ in the Contractable Prisoner’s Dilemma

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

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

(satisfying condition (2-a) on Proposition 2. Observe that the average of each type’s relative payoff $\frac{u_\alpha (a)}{u_a}$ and its trait-to-type ratio $m_\alpha (a)$ is equal to one, which satisfies condition (2-b). For example, the highest payoff 1.24 of a$d$ is compensated by having the lowest trait-to-type ratio of 0.76. Thus, Proposition 2 implies that $\alpha^*$ is stationary.

5 Characterisation of Stable States

In this section, we characterise asymptotic stability in our setup.

5.1 Benchmark Result for the Replicator Dynamics

It is well known that asymptotic stability of stationary states under the replicator dynamic is characterised by three conditions:

1. *Internal stability:* the payoff matrix restricted to the incumbent types is (semi-) negative-definite (i.e., all of its eigenvalues are negative (non-positive)).

2. *External stability:* The payoffs of types outside the support is lower than the incumbents’ payoff.
The strict variants of these conditions (i.e., negative-definiteness and quasi-strictness) imply asymptotic stability, and their weak counterparts (semi-negative-definiteness and being a Nash equilibrium) are implied by asymptotic stability. This is formalised as follows. Given a domain supp(\(\alpha\)) we define \(U\) as a square matrix of \(|\text{supp}(\alpha)|\), where the \(a_i,a_j\) element of \(U\) is equal to \(u(a_i,a_j)\).

**Claim 2.** A stationary state \(\alpha\) is asymptotically stable under the replicator dynamics if it satisfies

1. Internal stability: the payoff matrix \(U\) restricted to \(\text{supp}(\alpha)\) is negative definite.

2. External stability: \(U_\alpha > U_\alpha(a)\) for each \(a \notin \text{supp}(\alpha)\).

**Claim 3.** If a state \(\alpha\) is asymptotically stable under the replicator dynamics, then it satisfies

1. Weak internal stability: the payoff matrix \(U\) restricted to \(\text{supp}(\alpha)\) is semi-negative definite.

2. Weak external stability: \(U_\alpha \geq U_\alpha(a)\) for each \(a \notin \text{supp}(\alpha)\).

We omit the proofs of these two well-known results (which are implied by combining Theorem 9.2.7, Corollary 9.4.2, Theorem 9.4.4, and Theorem 9.4.8 in Van Damme, 1991).

The main results of this section generalise these claims, and characterise asymptotic stability under the recombinator dynamics. In order to do so, we first need to present the notion of \(r\)-Jacobian matrix (which generalises the payoff matrix in condition (1) of internal stability), and the notion of a payoff of an external trait, which will be used when generalising condition (2).

### 5.2 \(r\)-Jacobian Matrix

For each state \(\alpha\), let \(J'_\alpha\) denote the Jacobian matrix at \(\alpha\) with respect to the \(r\)-payoff function \(z';\) that is, \(J'_\alpha\) is a square matrix of size \(|\text{supp}(\alpha)|\), where the \(ij\)-th element in
the matrix is the partial derivative of $z_\alpha(a_i)$ with respect to $\alpha(a_j)$. Formally:

$$J_{\alpha}^{(r)}(a_i, a_j) = \frac{\partial z_\alpha(a_i)}{\partial \alpha(a_j)} \quad \text{for all } a_i, a_j \in \text{supp}(\alpha).$$

Observe that when $r = 0$ (the replicator dynamic) the $r$-Jacobian coincides with the payoff matrix (restricted to $\text{supp}(\alpha)$) up to multiplication by the constant $u_\alpha$, i.e.:

$$(J_\alpha^0)_{a_i, a_j} = \frac{\partial z_\alpha^0(a_i)}{\partial \alpha(a_j)} = \frac{\partial(u_\alpha(a_i)/u_\alpha)}{\partial \alpha(a_j)} = \frac{u(a_i, a_j)}{u_\alpha} \quad \text{for all } a_i, a_j \in \text{supp}(\alpha).$$

The main results of this section show that one can replace the payoff matrix in condition (1) of Claims 2–3 by the $r$-Jacobian when characterising asymptotic stability under the recombinator dynamics.

### 5.3 Stable Partner Distribution and Invading Trait Payoff

Consider an invasion of a population $\alpha$ by mutants who carry trait $a_d \notin \text{supp}(\alpha)$. In this scenario, there is a qualitative difference between the replicator dynamics ($r = 0$) and the recombinator dynamics with a positive recombination rate ($r > 0$). Under the replicator dynamics, the mutant type that carries trait $a_d$, say type $a = (a_d, a_{-d})$ (where we refer to the trait profile $a_{-d}$ as the partner of trait $a_d$), remains constant, and, thus, the success of trait $a_d$ in invading the population depends solely on the payoff of the mutant type combining the trait $a_d$ and its partner $a_{-d}$.

In contrast, under recombination the distribution of partners of an invading trait $a_d$ typically changes after a mutant carrying $a_d$ is introduced into a stationary population $\alpha$. In what follows we show that the distribution of partners converges towards a unique stable distribution of partners, which is independent of the initial mutant type that introduces $a_d$ to the population. We denote this stable distribution of partners of an invading trait $a_d$ by $\eta^a_\alpha$. Distribution $\eta^a_\alpha$ induces each type $(a_d, a_{-d})$ (where $a_{-d} \in \text{supp}(\alpha)$) with a frequency that is an $r$-weighted average of two elements: (1) its own payoff times its own frequency, and (2) the product of the frequencies of traits in
$a_{-d}$. Formally,

**Definition 8.** Fix $r > 0$, a stationary state $\alpha$, and a trait $a_d \notin \text{supp}(\alpha)$ Then $\eta^a_d \in \Delta(\text{supp}_{-d}(\alpha))$, which we call the *stable partner distribution* of trait $a_d$, is the unique solution to the following set of $|\text{supp}_{-d}(\alpha)|$ equations:

\[
\eta(a_{-d}) = (1 - r) \frac{\eta(a_{-d}) u_{\alpha}(a_d, a_{-d})}{\sum_{a_{-d} \in \text{supp}_{-d}(\alpha)} \eta(a_{-d}) u_{\alpha}(a_d, a_{-d})} + r \prod_{a_{d'} \in a_{-d}} \alpha(a_{d'})
\]

We begin by showing that the stable partner distribution is well-defined.

**Proposition 3.** Equation (19) admits a unique solution in $\Delta(\text{supp}_{-d}(\alpha))$ for any $r > 0$, stationary state $\alpha$, and trait $a_d \notin \text{supp}_{-d}(\alpha)$.

*Proof.* TBD (in the next version)

The fact that the partner distribution of an invading trait converges to a stable distribution (as shown in the proof of Theorem 1) allows us to define the payoff of an invading trait as the weighted average of the types carrying this mutant trait, where the weights are according to the stable distribution of partners. Formally,

**Definition 9.** Fix a stationary state $\alpha$ and recombination rate $r$. The *payoff of invading trait* $a_d \notin \text{supp}_{-d}(\alpha)$ is defined as follows:

\[
u^r_{\alpha}(a_d) = \sum_{a_{-d} \in \text{supp}(\alpha)} \eta^a_d (a_{-d}) \cdot u_{\alpha}(a_{-d}, a_d).
\]

The stable partner distribution is degenerate if the stationary state is homogeneous (pure). For example, the stable partner distribution of trait $d$ with respect to the stationary state $sc$ in the partially enforceable prisoner’s dilemma assigns mass one to $s$. The following example demonstrates the stable partner distribution for a heterogeneous stationary state in a new game, which we call the emotional hawk dove game.

**Example 4 (Stable partner distribution in the emotional hawk dove game).** Consider a bargaining interaction in which each player simultaneously makes two choices:
1. being a “hawk” (trait $h$) or a “dove” (trait $d$) in the bargaining.

2. being “emotional” (trait $e$), “rational” (trait $r$) or “versatile” (trait $v$) during the bargaining process (where the latter trait allows transitions from emotional to rational phases within the bargaining process).

The two dimensional set of types includes $6 = 2 \cdot 3$ elements $A = \{hr, hv, he, hr, hv, he\}$. The basic game is hawk-dove: two players have to divide a surplus worth 100 between them. Two doves divide it equally (50 each). A hawk obtain a large share of 70 against a dovish opponent. Finally, when two hawks are matched bargaining often fails, and hence each player obtains a low payoff of 10.

The payoffs of the basic hawk-dove game are modified by the choice of each player’s emotional approach. Being emotional helps a hawkish player and adds two to her share of the surplus, while it harms a dovish player and reduces four units from her share. Similarly, being rational helps a dovish play (adds two units to her share) but harms a hawkish player (reduces four units from her share). Finally, being versatile does not affect a player’s payoff.

Observe that the heterogeneous state $\alpha$ that assigns mass 50% to type $hv$ and the remaining mass of 50% to type $hv$ is stationary. Consider an invasion of this population by a rare mutant type with the emotional trait $e$. Substituting the example’s parameters in Eq. (19) yields the following value of $\eta^e_\alpha(h)$ as a function of $r$:

$$\eta(h) = (1-r)\frac{42\eta(h)}{42\eta(h) + 36(1-\eta(h))} + \frac{r}{2} \Rightarrow \eta^e_\alpha(h)(r) = \frac{\sqrt{169r^2 - 4r + 4 - 13r + 2}}{4}.$$

Figure 2 illustrates the stable share of hawkish partners of the invading trait $e$ as a function of the recombination rate $r$. When the recombination rate is close to zero, almost all the partners of $e$ are hawkish. This share is decreases in $r$ and it converges to $50/50$ as the recombination rate converges to 1. The payoff of the invading trait is

$$u_\alpha(e) = \eta^e_\alpha(r)(h) \cdot u_{\alpha}(he) + \eta^e_\alpha(r)(d) \cdot u_{\alpha}(he) = \eta^e_\alpha(r)(h)\frac{72+12}{2} + \eta^e_\alpha(r)(d)\frac{46+26}{2},$$

28
which is larger than \( u_\alpha = 40 \) iff \( \eta_\alpha^e(r)(h) > \frac{2}{3} \iff r < \frac{1}{6} \). Thus the stationary state \( \alpha \) is stable against an invasion of the trait \( e \) iff \( r > \frac{1}{6} \) (one can show that this condition implies that \( \alpha \) is asymptotically stable).

Table 3: Payoff Matrix for Emotional Hawk-Dove Game

| \( d_r \) | \( d_v \) | \( d_e \) | \( h_r \) | \( h_v \) | \( h_e \) |
|---------|---------|---------|---------|---------|---------|
| 50, 50  | 52, 48  | 56, 44  | 36, 64  | 32, 68  | 30, 70  |
| 48, 52  | 50, 50  | 54, 46  | 34, 66  | 30, 70  | 28, 72  |
| 44, 56  | 50, 50  | 54, 46  | 30, 70  | 26, 74  | 24, 76  |
| 64, 36  | 66, 34  | 70, 30  | 10, 10  | 6, 14   | 4, 16   |
| 68, 32  | 70, 30  | 74, 26  | 14, 6   | 10, 10  | 8, 12   |
| 70, 30  | 72, 28  | 76, 24  | 16, 4   | 12, 8   | 10, 10  |

Figure 2: Stable Share of Hawkish Partners \( \eta_\alpha^e(h)(r) \) in Example 4

5.4 Main Results

In what follows we characterise the set of asymptotically stable states. The characterisation extends Claims 2–3 by showing that the asymptotic stability of a stationary state is characterised by the following three conditions:

1. **Internal stability:** The \( r \)-Jacobian matrix is (semi-)negative-definite (i.e., all of its eigenvalues are negative (non-positive)).
2. **External stability against traits**: The payoffs of traits outside the support are lower than the incumbent traits’ payoffs.

3. **External stability against types**: The payoffs of types outside the support are at most $1 - r$ times higher than the average incumbents’ payoff.

Our main results are, essentially, if and only if characterisation of asymptotic stability. Specifically, the strict variants of these conditions (i.e., negative-definiteness and strictly lower payoffs of external traits/types) imply asymptotic stability, and their weak counterparts (semi-negative-definiteness and weakly lower payoffs of external traits/types) are implied by asymptotic stability. Formally,

**Theorem 1.** A stationary state $\alpha$ is asymptotically stable under the recombinator dynamics with recombination rate $r$ if it satisfies

1. **Internal stability**: the $r$-Jacobian matrix $J_\alpha^r$ is negative definite.

2. **External stability against traits**: $u_\alpha > u_\alpha^r(a_d)$ for each $a_d \notin \text{supp}_d(\alpha)$.

3. **External stability against types**: $u_\alpha > (1 - r)u_\alpha(a)$ for each $a \notin \text{supp}(\alpha)$.

**Theorem 2.** If a state $\alpha$ is asymptotically stable under the recombinator dynamics, then it satisfies

1. **Internal stability**: the $r$-Jacobian matrix $J_\alpha^r$ is semi-negative definite.

2. **External stability against traits**: $u_\alpha \geq u_\alpha^r(a_d)$ for each $a_d \notin \text{supp}_d(\alpha)$.

3. **External stability against types**: $u_\alpha \geq (1 - r)u_\alpha(a)$ for each $a \notin \text{supp}(\alpha)$.

The intuition for the internal stability condition is the same as in the standard case of the replicator dynamics. The intuition for the second condition (external stability against traits) was explained in the previous subsection. In what follows we explain the added factor of $1 - r$ (which multiplies the payoff of the external types). If the invading type’s payoff is more than $\frac{1}{1-r}$ larger than the incumbents’ payoff, then the share of $1 - r$ of the new agents who imitate all traits from the same mentor is sufficient for...
the frequency of this successful invading trait to increase, regardless of what the new agents who combine multiple traits do. In contrast, if the invading payoff is less than \( \frac{1}{1-r} \) times the incumbents’ payoff, then the ability to invade the population crucially depends on whether the external traits that compose the invading type can invade the population. If none of these traits can invade the population, the invading type will become extinct.

*Proof.* TBD in the next version

### References

Arad, A. and S. Penczynski (2018). Multi-dimensional reasoning in competitive resource allocation games: Evidence from intra-team communication. Technical report, Working paper.

Arad, A. and A. Rubinstein (2012). Multi-dimensional iterative reasoning in action: The case of the colonel blotto game. *Journal of Economic Behavior & Organization 84*(2), 571–585.

Arad, A. and A. Rubinstein (2019). Multidimensional reasoning in games: framework, equilibrium, and applications. *American Economic Journal: Microeconomics 11*(3), 285–318.

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.

Arigapudi, S., Y. Heller, and I. Milchtaich (2021). Instability of defection in the prisoner’s dilemma under best experienced payoff dynamics. *Journal of Economic Theory 197*, 105174.
Arigapudi, S., Y. Heller, and A. Schreiber (2022). Sampling dynamics and stable mixing in hawk-dove games. *mimeo.*

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.

Hart, S. (2002). Evolutionary dynamics and backward induction. *Games and Economic Behavior* 41(2), 227–264.

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.

Oyama, D., W. H. Sandholm, and O. Tercieux (2015). Sampling best response dynamics and deterministic equilibrium selection. *Theoretical Economics* 10(1), 243–281.
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. (2001). Almost global convergence to $p$-dominant equilibrium. International Journal of Game Theory 30(1), 107–116.

Sandholm, W. H. (2010). Population games and evolutionary dynamics. MIT press.

Sandholm, W. H., S. S. Izquierdo, and L. R. Izquierdo (2020). Stability for best experienced payoff dynamics. Journal of Economic Theory 185, 104957.

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.

Sethi, R. (2000). Stability of equilibria in games with procedurally rational players. Games and Economic Behavior 32(1), 85–104.

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.

Van Damme, E. (1991). Stability and perfection of Nash equilibria, Volume 339. Springer.

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. (1966). *Adaptation and natural selection*. Princeton, NJ: Princeton university press.