Journal of Theoretical Biology 605 (2025) 112089

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/jtbi

Many-strategy games in groups with relatives and the evolution of coordinated cooperation

Nadiah P. Kristensen [a,*], Ryan A. Chisholm [a], Hisashi Ohtsuki [b]

[a] Department of Biological Sciences, National University of Singapore, 16 Science Drive 4, Singapore, 117558, Singapore
[b] Research Center for Integrative Evolutionary Science, SOKENDAI (The Graduate University for Advanced Studies), Shonan Village, Hayama, Kanagawa, 240-0193, Japan

ARTICLE INFO

Keywords:
Evolutionary game dynamics
Kin selection
Relatedness
Replicator equation
Transformed payoff
Coordinated cooperation

ABSTRACT

Humans often cooperate in groups with friends and family members with varying degrees of genetic relatedness. Past kin selection can also be relevant to interactions between strangers, explaining how the cooperation first arose in the ancestral population. However, modelling the effects of relatedness is difficult when the benefits of cooperation scale nonlinearly with the number of cooperators (e.g., economies of scale). Here, we present a direct fitness method for rigorously accounting for kin selection in n-player interactions with m discrete strategies, where a genetically homophilic group-formation model is used to calculate the necessary higher-order relatedness coefficients. Our approach allows us to properly account for non-additive fitness effects between relatives (synergy). Analytical expressions for dynamics are obtained, and they can be solved numerically for modestly sized groups and numbers of strategies. We illustrate with an example where group members can verbally agree (cheap talk) to contribute to a public good with a sigmoidal benefit function, and we find that such coordinated cooperation is favoured by kin selection. As interactions switched from family to strangers, in order for coordinated cooperation to persist and for the population to resist invasion by liars, either some level of homophily must be maintained or following through on the agreement must be in the self-interests of contributors. Our approach is useful for scenarios where fitness effects are non-additive and the strategies are best modelled in a discrete way, such as behaviours that require a cognitive 'leap' of insight into the situation (e.g., shared intentionality, punishment).

1. Introduction

The theory of kin selection (Hamilton, 1964; Taylor and Frank, 1996; Lehmann and Rousset, 2010) not only explains cooperation between family members (Burnstein, 2015) but can also provide hypotheses for the origin of cooperation between nonkin. Cooperation between nonkin and strangers (Ledyard, 1995; Raihani and Bshary, 2015) may have originated when human ancestors interacted in small groups of mostly kin, thus evolving cooperative psychological impulses that were subsequently misapplied to nonkin today, i.e., the maladaptation hypothesis (Burnham and Johnson, 2005; Hagen and Hammerstein, 2006; El Mouden et al., 2012). Alternatively, a past kin-assortative environment may have facilitated the invasion of a cooperative behaviour that, once established in a population, is adaptive in interactions with cooperative nonkin (e.g., Bach et al., 2006; Boyd et al., 2010; Takezawa and Price, 2010; Cornforth et al., 2012; Boyd et al., 2014; Schonmann and Boyd, 2016; Kristensen et al., 2022). For example, reciprocal strategies like tit-for-tat in the iterated Prisoner's Dilemma are mutually beneficial among strangers, but in finitely repeated games in very large populations, they cannot invade a population of defectors (Axelrod and Hamilton, 1981); one possibility is that they first arose in the past, when dispersal was low and individuals typically interacted with family, and their invasion was facilitated by kin selection (Axelrod and Hamilton, 1981; Carter, 2021; Kristensen et al., 2022). Given the multiple ways kin selection likely influenced the evolution of human cooperation, we are interested in modelling techniques that combine kin assortativity with evolutionary game theory.

One way kin selection can operate is if individuals have a preference to interact with family members, and anthropological evidence suggests this is common (Alvard, 2009; Jaeggi and Gurven, 2013). For example, reciprocal food sharing networks provide a buffer against uncertainty, but individuals often choose reciprocal partners from among their kin (Allen-Arave et al., 2008; Nolin, 2010; Koster and Leckie, 2014). As another example, collective livestock-herding provides mutual

* Corresponding author.
E-mail addresses: nadiah@nadiah.org (N.P. Kristensen), ryan.chis@gmail.com (R.A. Chisholm), ohtsuki_hisashi@soken.ac.jp (H. Ohtsuki).
URL: https://nadiah.org/ (N.P. Kristensen)

https://doi.org/10.1016/j.jtbi.2025.112089
Received 20 November 2024; Received in revised form 3 March 2025; Accepted 5 March 2025
Available online 28 March 2025

benefits through economies of scale (Næss et al., 2010), and members of nomadic herding groups are typically—though not always—relatives (Næss, 2021). Such groups may form temporarily, e.g., changing composition between seasons (Næss, 2019). Therefore, we want to model ephemeral groups (*sensu* Godfrey-Smith and Kerr, 2009) that contain a mix of kin and nonkin.

However, it can be difficult to model kin selection when cooperation has non-additive benefits, i.e., when the benefit from many cooperators working together differs from the sum of the benefits that would have accrued if they had worked individually. Non-additivity is ubiquitous in biology (Archetti and Scheuring, 2012; Vásárhelyi and Scheuring, 2013; Archetti et al., 2020), including the evolutionarily important task of acquiring resources; confrontational scavenging (Bickerton and Szathmáry, 2011), hunting (Alvard and Nolin, 2002; Boza and Számadó, 2010), and agricultural food preparation (Hames and McCabe, 2007) all involve economies of scale. Non-additivity is also important for contemporary cooperation problems; conservation areas can possess ecological tipping points (Nobre and De Simone, 2009), and coordination in international climate agreements can induce nonlinearities in the benefits returned (Barrett, 2016). Moreover, models must account for non-additivity because it can significantly alter the evolutionary dynamics of a system (Hauert et al., 2006), e.g., it can permit a stable coexistence between cooperative and non-cooperative types where coexistence would not otherwise be possible (Archetti and Scheuring, 2011).

If more than two individuals are involved in a non-additive interaction, a model must account for all possible combinations of types within the group (e.g., Eq. (6), Allen and Nowak, 2016), which means accounting for the probability of all kin + nonkin combinations, i.e., the higher-order genetic associations (Ohtsuki, 2010) or other collective relatedness measures (Allen et al., 2024). One can avoid complicated genetic accounting with the help of certain assumptions (Van Veelen, 2009; Van Cleve, 2015) when modelling continuous traits, e.g., the amount contributed to a public good (Coder Gylling and Brännström, 2018) or the probability of cooperating (Peña et al., 2015). However, many human cooperative strategies involve a conceptual 'leap'—such as inventing a punishment institution (Sigmund et al., 2010), or conditioning one's cooperation on the presence of such an institution (Garcia and De Monte, 2013) or the cooperation of others (Takezawa and Price, 2010)—that is better modelled in a discrete way.

Here, we detail an evolutionary game-theoretic approach that can be applied to discrete strategies in non-additive group games played between relatives. Our approach builds on the higher-order genetic association approach of Ohtsuki (2014), which we here extend to scenarios with more than two strategies. We present our formulae in three ways. The first form is comparable to simple replicator dynamics (where group members are chosen at random), and we find that, when the group size and number of strategies is small, we can partition the dynamics into a kin-interaction and random-interaction component. The second form transforms the homophilic scenario into an equivalent well-mixed scenario, which can be analysed using the same methods as replicator dynamics. The third form is less analytically interpretable but more efficient for computation. We provide code and illustrate use of the code in a worked example.

We apply our methods to the example of coordinated cooperation in a public goods game with a sigmoidal benefit function. We find that coordinated cooperation can be favoured by kin selection, which allows it to invade, and it may subsequently persist if homophily declines and interactions shift from family to strangers. Thus, kin selection provides a stepping-stone towards coordinated cooperation between nonkin. However, if the benefits function is not nonlinear enough, then the population may be vulnerable to invasion by a 'lying' strategy that presents as a coordinating cooperator but subsequently defects. In that case, some level of homophily must be maintained in order for coordinated cooperation to persist evolutionarily. Our approach can be used for any similar example with a matrix-payoff game and kin selection to explore different narratives about the origin and persistence of cooperation.

2. Model

The higher-order genetic associations approach extends the replicator dynamics approach (Hofbauer and Sigmund, 1998) to include kin selection. Kin selection is modelled by accounting for the relatedness between group members due to genetically homophilic group formation.

2.1. Notation

Throughout the paper, we will use round braces for vectors, curly braces for sets, and square braces for multisets. Multisets are sets that allow the same element to appear more than once. For example, $(3, 2, 1, 1)$ is a four-dimensional vector, $\{3, 2, 1\}$ is a set with three elements, 3, 2 and 1, and $[3, 2, 1, 1]$ is a multiset that includes 3 once, 2 once, and 1 twice. Most importantly, $(3, 2, 1, 1)$ and $(1, 1, 2, 3)$ are different as vectors, but $[3, 2, 1, 1]$ and $[1, 1, 2, 3]$ are the same as multisets. Symbols that are frequently used are summarised in Table 1.

2.2. Life-cycle assumptions

We assume an infinitely large, clonally reproducing population of haploid individuals. Each individual plays one of m genetically determined pure strategies, $s_1, \ldots, s_m$ (e.g., Cooperate, Defect, etc.) in a social interaction or 'game' (e.g., a public goods game). Every timestep, individuals play games in groups of size n, which are formed independently according to a group-formation model (examples detailed below). Each individual is on average involved in one game per timestep. The payoffs each individual receives from the games determine the number of clonal offspring it has. A proportion s of adults survive each timestep, and offspring compete on an equal basis for the vacancies created by adult deaths. Because groups form and dissolve for each interaction, offspring compete in a global pool, and there is no local, within-group competition. This means that we have kin interaction but no kin competition in our model.

2.3. Homophilic group-formation model

Under genetic homophily, individuals preferentially group with family members. In our model, 'family' means the members have inherited their strategy from a shared common ancestor, i.e., they are identical by descent (IBD). Two individuals with different strategies are never family. However, two individuals with the same strategy may be family or they may not be — the latter case occurs when they have inherited the same strategy from different mutational events.

A homophilic group-formation model allows us to calculate the probabilities that groups will form with different family structures. We assume each individual has the same degree of family preference regardless of their strategy, and consequently the relatedness distribution in a group is independent of the strategy distribution. We assume that families are large enough that every individual can potentially form a group with n family members; however, we also assume that families are not so large that the strategy distribution among any set of non-family deviates from the frequency in the population (i.e., $n <$ family size $\ll$ population size$(= \infty)$). For example, we would not allow a single family to take up half the population, because then recruitments of non-family members will have a different strategy distribution to the frequency distribution of strategies in the population.

Let $\boldsymbol{q}$ be a multiset of positive integers that sum to n (i.e., an integer partition of n) that represents the family partition structure of a group. For example, $\boldsymbol{q} = [2, 2, 1]$ describes a group of 5 individuals with 2 individuals from one family, 2 individuals from a second family, and 1 individual from a third. A homophilic group formation model allows us to calculate the probability $F_{\boldsymbol{q}}$ that a randomly sampled group has family partition structure $\boldsymbol{q}$. The family structure can also be described by a group's family-size distribution $\boldsymbol{y}$, which is a vector of length n whose

Table 1
Notation. "r.v." stands for random variable.

Symbol	Attribute	Description
n	number	number of players
m	number	number of strategies
$\boldsymbol{q}$	multiset	family structure of a group, given as a partition of n
$\boldsymbol{y}$	vector	family structure of a group, given as a family size distribution
$F_{\boldsymbol{q}}$, (or $F_{\boldsymbol{y}}$)	number	probability that the family structure is $\boldsymbol{q}$, (or $\boldsymbol{y}$)
$\boldsymbol{z} = (\boldsymbol{z}_1, \dots, \boldsymbol{z}_m)$	vector of vectors	detailed family structure in a group
$\boldsymbol{z}_i = (z_{i,1}, \dots, z_{i,n})$	vector	detailed family structure of s_i strategists in a group, given as a family size distribution
r_2	number	dyadic relatedness
r_3	number	triadic relatedness
$\boldsymbol{g}_0 = (g_{0,1}, \dots, g_{0,m})$	vector	indicator of strategy of the focal player
$\boldsymbol{g}_j = (g_{j,1}, \dots, g_{j,m})$	vector	indicator of strategy of jth player
$\boldsymbol{g}_{\mathrm{nf}} = (g_{\mathrm{nf},1}, \dots, g_{\mathrm{nf},m})$	vector	indicator of strategies of the nonfocal players in a group
$\boldsymbol{g}_{\mathrm{a}} = (g_{\mathrm{a},1}, \dots, g_{\mathrm{a},m})$	vector	indicator of strategies of all players in a group
π_0	number	payoff of the focal player
$\overline{\pi}_x$	number	average payoff of players pursuing strategy x
$\overline{\pi}$	number	average payoff in the whole population
w_0	number	fitness of the focal player
$\overline{w}$	number	average fitness in the whole population (= 1)
p_x	number	frequency of players pursuing strategy x
$\boldsymbol{G}_0 = (G_{0,1}, \dots, G_{0,m})$	vector of r.v.	random variable for $\boldsymbol{g}$
$\boldsymbol{G}_j = (G_{j,1}, \dots, G_{j,m})$	vector of r.v.	random variable for $\boldsymbol{g}_j$
$\boldsymbol{G}_{\mathrm{nf}} = (G_{\mathrm{nf},1}, \dots, G_{\mathrm{nf},m})$	vector of r.v.	random variable for $\boldsymbol{g}_{\mathrm{nf}}$
$\boldsymbol{G}_{\mathrm{a}} = (G_{\mathrm{a},1}, \dots, G_{\mathrm{a},m})$	vector of r.v.	random variable for $\boldsymbol{g}_{\mathrm{a}}$
Π_0	random variable	random variable for π_0
W_0	random variable	random variable for w_0

kth element counts how many families in the group have k members. For example, if $\boldsymbol{q} = [2, 2, 1]$, then $\boldsymbol{y} = (1, 2, 0, 0, 0)$ because there is 1 family of size one, 2 families of size two, 0 families of size three/four/five. Therefore, a homophilic group-formation model also calculates the probability $F_{\boldsymbol{y}}$ that a randomly sampled group has family size distribution $\boldsymbol{y}$.

Our approach can be used with any group-formation model that provides probabilities $F_{\boldsymbol{q}}$ for all $\boldsymbol{q}$ (or equivalently, $F_{\boldsymbol{y}}$ for all $\boldsymbol{y}$), and three examples are detailed in Kristensen et al. (2022):

1. Leader driven: The leader is chosen at random from the population, and recruits/attracts kin with probability h and nonkin with probability $1 - h$. The only possible family partition structures have one family containing ℓ individuals and $n - \ell$ families containing a solitary individual, represented by the multiset $\boldsymbol{q} = [\ell, 1, \dots, 1]$. The reason why members who are not IBD with the leader have no other family members in the group is because we assume the population is infinitely large but families are not large. The nonzero family partition probabilities are

$$F_{[\ell,1,\dots,1]} = \binom{n-1}{\ell-1} h^{\ell-1}(1-h)^{n-\ell}. \tag{1}$$

2. Members recruit: The initial member is chosen at random. Current group members have an equal chance to recruit the next member, and recruit kin with probability h and nonkin with probability $1 - h$. An algorithm for computing $F_{\boldsymbol{q}}$ is provided in Kristensen et al. (2022).

3. Members attract: The initial member is chosen at random. Current group members have equal weighting 1 of attracting a new member who is kin, but nonkin members are also attracted to the group itself with collective weighting $\alpha \in [0, \infty)$. Therefore, the probability that the jth recruit (or $j + 1$th group member) is from a new family is

$$\mathbb{P}[j \text{ is from a new family}] = \frac{\alpha}{\alpha + j},$$

or from family a having n_a members already in the group is

$$\mathbb{P}[j \text{ is from family } a] = \frac{n_a}{\alpha + j}.$$

Then, the family partition probabilities are described by Ewens' formula (Ewens, 1972)

$$F_{\boldsymbol{y}} = \frac{n!\,\alpha^{\|\boldsymbol{y}\|}}{\prod_{k=1}^{n} k^{y_k} y_k!(\alpha + k - 1)}, \tag{2}$$

where y_k is the number of families with k members and $\|\boldsymbol{y}\| = \sum_{k=1}^{n} y_k$ is the number of families in the group.

2.4. Payoff and fitness

The payoff an individual receives from the game depends on its own strategy and the strategies of the other group members. We index the individuals in the group $j = 0, \dots, n-1$, reserving the index 0 to denote the focal individual. We define the strategy indicator for an individual j as $\boldsymbol{g}_j = (g_{j,1}, \dots, g_{j,m})$ such that $g_{j,x} = 1$ if individual j pursues strategy s_x and $g_{j,x} = 0$ otherwise. Thus, if individual j pursues strategy s_x, then $\boldsymbol{g}_j = \boldsymbol{e}_x$ where $\boldsymbol{e}_x$ is an m-dimensional vector with a 1 in the xth position and 0 in all others. The whole-group strategy composition is $\boldsymbol{g}_{\mathrm{a}} = \boldsymbol{g}_0 + \boldsymbol{g}_1 + \dots + \boldsymbol{g}_{n-1}$ (a is for "all") such that the ith component counts the number of group members who pursue strategy s_i. Similarly, the nonfocal strategy composition is $\boldsymbol{g}_{\mathrm{nf}} = \boldsymbol{g}_1 + \dots + \boldsymbol{g}_{n-1}$ (nf is for "non focal").

We assume that the payoff to the focal individual from the game can depend on the strategy of the focal individual and the nonfocal strategy composition, but that it is invariant against any permutation of labelling of the nonfocal individuals, $j = 1, \dots, n-1$ (cf. Gokhale and Traulsen, 2010). Thus, the payoff can be written as

$$\pi_0 := \pi(\boldsymbol{g}_0, \boldsymbol{g}_{\mathrm{nf}}) = \sum_{i=1}^{m} g_{0,i}\,\pi(\boldsymbol{e}_i, \boldsymbol{g}_{\mathrm{nf}}). \tag{3}$$

The payoff can also be written in terms of the whole-group strategy composition as

$$\pi_0 := \hat{\pi}(\boldsymbol{g}_0, \boldsymbol{g}_{\mathrm{a}}) = \sum_{i=1}^{m} g_{0,i}\,\hat{\pi}(\boldsymbol{e}_i, \boldsymbol{g}_{\mathrm{a}}). \tag{4}$$

These two definitions lead to the two formulations presented below, the accounting based on other-member versus whole-group composition, respectively.

Fitness in our model is defined as the number of offspring an individual has in the next generation, including the individual itself if it survives. Given the lifecycle assumptions, the expected fitness of the focal individual is the probability it survives plus the expected number of its offspring that are recruited to the population

$$w_0 = s + (1 - s)\frac{1 + \delta\pi_0}{1 + \delta\overline{\pi}}, \tag{5}$$

where $\overline{\pi}$ is the average payoff in the whole population, and the selection strength δ scales how important the social interaction is relative to the baseline reproduction with value 1. It is assumed that δ is small enough to ensure $1+\delta\pi_0 > 0$ and $1+\delta\overline{\pi} > 0$ (i.e., we assume w-weak selection *sensu* Wild and Traulsen, 2007).

2.5. Evolutionary dynamics

The change in the proportions of different strategies in the population is derived from first principles from the Price equation (Price, 1970). Let $\boldsymbol{p} = (p_1, \dots, p_m)$ be the vector of proportions p_x of the population pursuing strategy s_x. We would like to derive Δp_x, which is the change in the proportion of s_x strategists in the population. For that purpose, we define several random variables. In what follows, capital letters are often used for random variables, while corresponding small letters are often used for their values; see Table 1.

We randomly sample a group from the population, which we call the "focal group", and we denote its strategy composition by $\boldsymbol{G}_{\mathrm{a}} = (G_{\mathrm{a},1}, \dots, G_{\mathrm{a},m})$, where its ith component $G_{\mathrm{a},i}$ represents the number of players pursuing strategy s_i in the group, which is a random variable. Next, we randomly allocate labelling $j = 0, 1, \dots, n-1$ to the members of the group. The player with label 0 is designated the "focal player". We denote the strategy of the jth player by $\boldsymbol{G}_j = (G_{j,1}, \dots, G_{j,m})$, where its ith component $G_{j,i}$, which is a random variable, is 1 if this player pursues strategy s_i and 0 otherwise. The strategy composition among the $n-1$ non-focal players is denoted by $\boldsymbol{G}_{\mathrm{nf}}$ and calculated as $\boldsymbol{G}_{\mathrm{nf}} \equiv \boldsymbol{G}_1 + \dots, \boldsymbol{G}_{n-1}$. We also denote the payoff and the fitness of the focal player by Π_0 and W_0, respectively. They are also random variables.

With these notations, the evolutionary dynamics is given by the Price equation (Price, 1970)

$$\Delta p_x = \frac{\mathbb{C}\mathrm{ov}[G_{0,x}, W_0]}{\overline{w}}, \tag{6}$$

where $\mathbb{C}\mathrm{ov}$ represents covariance, and $\overline{w}$ is the average fitness in the population, which under the assumption of constant population size takes value $\overline{w} = 1$. By repeatedly applying two covariance identities to Eq. (6) (details in SI A.2), we obtain

$$\Delta p_x = \frac{\delta(1-s)}{1+\delta\overline{\pi}}\mathbb{C}\mathrm{ov}[G_{0,x}, \Pi_0] \propto \mathbb{C}\mathrm{ov}[G_{0,x}, \Pi_0]. \tag{7}$$

3. Analytical results

The purpose of this section is two-fold. In Section 3.1, we derive evolutionary game dynamics based on other-member-accounting payoff function π in Eq. (3). We find that we can derive difference equations or differential equations that are in the form of replicator equations, but expected payoffs are calculated with various relatedness coefficients. The theory we derive gives us analytical insights into how kin-interactions affect evolutionary dynamics. In Section 3.2, on the other hand, we derive evolutionary game dynamics based on whole-group-accounting payoff function $\hat{\pi}$ in Eq. (4). The derived final expressions are difficult to intuitively understand, but the main purpose is to give a computationally tractable formula, which gives us a practical method for numerically computing evolutionary dynamics of games with any given number of players, n, and with any given number of strategies, m. In other words, the goal of Section 3.1 is to give an intuitive formula, whereas the goal of Section 3.2 is to give a computational method. They look very different but they are mathematically equivalent, and we believe providing both types of formulae will help various types of readers with different interests.

3.1. Accounting based on other-member composition

Substituting the payoff to the focal individual from Eq. (3) into Eq. (7), we obtain the dynamics

$$\Delta p_x \propto \mathbb{E}[G_{0,x}\pi(\boldsymbol{e}_x, \boldsymbol{G}_{\mathrm{nf}})] - p_x \sum_{i=1}^{m} \mathbb{E}[G_{0,i}\pi(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{nf}})], \tag{8}$$

where $\mathbb{E}$ is expectation (see SI A.3). The random variable $\boldsymbol{G}_{\mathrm{nf}}$ can take values from the set

$$\mathcal{G}_{\mathrm{nf}} = \left\{ \boldsymbol{g}_{\mathrm{nf}} \in \mathbb{N}_{\geq 0}^{m} \;\middle|\; \sum_{i=1}^{m} g_{\mathrm{nf},i} = n-1 \right\}, \tag{9}$$

where $\mathbb{N}_{\geq 0} = \{0, 1, 2, \dots\}$ represents the set of natural numbers. Define $\overline{\pi}_i$ as the expected payoff to s_i strategists

$$\begin{aligned} \overline{\pi}_i &:= \mathbb{E}[\pi(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{nf}}) \mid \boldsymbol{G}_0 = \boldsymbol{e}_i] \\ &= \sum_{\boldsymbol{g}_{\mathrm{nf}} \in \mathcal{G}_{\mathrm{nf}}} \pi(\boldsymbol{e}_i, \boldsymbol{g}_{\mathrm{nf}}) \mathbb{P}[\boldsymbol{G}_{\mathrm{nf}} = \boldsymbol{g}_{\mathrm{nf}} \mid \boldsymbol{G}_0 = \boldsymbol{e}_i], \end{aligned} \tag{10}$$

which is the sum over all possible $\boldsymbol{g}_{\mathrm{nf}}$ of the product of the payoff to an s_i strategist when the others are $\boldsymbol{g}_{\mathrm{nf}}$ and the probability that an s_i strategist is grouped with others that are $\boldsymbol{g}_{\mathrm{nf}}$. Expanding Eq. (8) and substituting in the definition in Eq. (10), it can be shown (details in SI A.3) that the change in the proportion of s_x in the population is proportional to

$$\Delta p_x \propto p_x\left(\overline{\pi}_x - \sum_{i=1}^{m} p_i \overline{\pi}_i\right) = p_x\left(\overline{\pi}_x - \overline{\pi}\right), \tag{11}$$

which is a discrete-time replicator equation (Hofbauer and Sigmund, 1998). By taking a proper limit (for example, letting $\delta(1-s) \to 0$ and taking a proper time scale) in Eq. (7), we can obtain a continuous-time replicator equation (Hofbauer and Sigmund, 1998), as

$$\dot{p}_x = p_x\left(\overline{\pi}_x - \overline{\pi}\right). \tag{12}$$

Here and hereafter $\dot{p}_x$ represents a time-derivative of p_x.

The term $\mathbb{P}[\boldsymbol{G}_{\mathrm{nf}} = \boldsymbol{g}_{\mathrm{nf}} \mid \boldsymbol{G}_0 = \boldsymbol{e}_i]$ in Eq. (10) depends on the group-formation model. In a well-mixed population, where group formation is random uniform, we have $\mathbb{P}[\boldsymbol{G}_{\mathrm{nf}} = \boldsymbol{g}_{\mathrm{nf}} \mid \boldsymbol{G}_0 = \boldsymbol{e}_i] = \mathbb{P}[\boldsymbol{G}_{\mathrm{nf}} = \boldsymbol{g}_{\mathrm{nf}}]$, meaning that other-member composition is independent of the focal player, and the nonfocal strategy composition is determined by $n-1$ independent samples from the population, which follows a multinomial distribution. However, if group formation is genetically homophilic, the group is biased towards having more individuals pursuing the same strategy.

3.1.1. Replicator dynamics written in terms of relatedness coefficients

To obtain an intuition for Eq. (11), let us first consider a game with $m = 3$ strategies played between $n = 2$ players. The probabilities of each nonfocal strategy conditional on the focal individual being an s_1 strategist are

$$\begin{aligned} &\mathbb{P}[\boldsymbol{G}_{\mathrm{nf}} = \boldsymbol{e}_1 \mid \boldsymbol{G}_0 = \boldsymbol{e}_1] = F_{[2]} + F_{[1,1]}p_1, \\ &\mathbb{P}[\boldsymbol{G}_{\mathrm{nf}} = \boldsymbol{e}_2 \mid \boldsymbol{G}_0 = \boldsymbol{e}_1] = F_{[1,1]}p_2, \\ &\mathbb{P}[\boldsymbol{G}_{\mathrm{nf}} = \boldsymbol{e}_3 \mid \boldsymbol{G}_0 = \boldsymbol{e}_1] = F_{[1,1]}p_3. \end{aligned} \tag{13}$$

Therefore, the expected payoff to s_1-strategists is

$$\begin{aligned} \overline{\pi}_1 = &(F_{[2]} + F_{[1,1]}p_1)\pi(\boldsymbol{e}_1, \boldsymbol{e}_1) + (F_{[1,1]}p_2)\pi(\boldsymbol{e}_1, \boldsymbol{e}_2) \\ &+ (F_{[1,1]}p_3)\pi(\boldsymbol{e}_1, \boldsymbol{e}_3). \end{aligned} \tag{14}$$

The expected payoff to the s_1 strategist $\overline{\pi}_1$ can also be written in terms of dyadic relatedness. The probability that both individuals are IBD is $r_2 = F_{[2]}$ (i.e., the dyadic relatedness), and therefore we have $1 - r_2 = F_{[1,1]}$, resulting in

$$\overline{\pi}_1 = r_2\pi(\boldsymbol{e}_1, \boldsymbol{e}_1) + (1-r_2)\underbrace{\left(p_1\pi(\boldsymbol{e}_1, \boldsymbol{e}_1) + p_2\pi(\boldsymbol{e}_1, \boldsymbol{e}_2) + p_3\pi(\boldsymbol{e}_1, \boldsymbol{e}_3)\right)}_{:=\overline{\pi}_1^{\mathrm{mix}}} \tag{15}$$

where $\overline{\pi}_1^{\text{mix}}$ is the average payoff of strategy s_1 in a well-mixed population.

In general, for a 2-player game with m strategies, we have

$$\overline{\pi}_1 = r_2\pi(\boldsymbol{e}_1,\boldsymbol{e}_1) + (1-r_2)\underbrace{\sum_{i=1}^{m} p_i\pi(\boldsymbol{e}_1,\boldsymbol{e}_i)}_{\overline{\pi}_1^{\text{mix}}}. \tag{16}$$

At one extreme with perfect homophily, we have $r_2 = 1$ and therefore $\overline{\pi}_1 = \pi(\boldsymbol{e}_1,\boldsymbol{e}_1)$. At the other extreme with $r_2 = 0$, the average payoff is $\overline{\pi}_1 = \overline{\pi}_1^{\text{mix}}$, which recovers the replicator equation for a well-mixed population. Thus, the general effect of genetic homophily on fitness is to produce expected payoffs between these two extremes.

Now let us consider games played between $n = 3$ players. The expected payoff to an s_1 strategist is

$$\begin{aligned}\overline{\pi}_1 = {} & F_{[3]}\pi(\boldsymbol{e}_1,2\boldsymbol{e}_1) + F_{[2,1]}\sum_{i=1}^{m} p_i\left[\frac{1}{3}\pi(\boldsymbol{e}_1,2\boldsymbol{e}_i) + \frac{2}{3}\pi(\boldsymbol{e}_1,\boldsymbol{e}_1+\boldsymbol{e}_i)\right] \\ & + F_{[1,1,1]}\sum_{j=1}^{m}\sum_{k=1}^{m} p_j p_k \pi(\boldsymbol{e}_1,\boldsymbol{e}_j+\boldsymbol{e}_k),\end{aligned} \tag{17}$$

which can be derived as follows. $F_{[3]}$ is the probability that an s_1 strategist finds itself grouped with two others with whom it is IBD, and therefore those others are also s_1 strategists. $F_{[2,1]}$ is the probability the family partition structure is $[2,1]$. With probability $1/3$, the focal s_1 strategist is a member of the size-1 family, in which case the other two are IBD to each other, and they both are s_i strategists with probability p_i. With probability $2/3$, the focal s_1 strategist is a member of the size-2 family, in which case one group member is IBD to the focal and is an s_1 strategist, and the other group member is an s_i strategist with probability p_i. $F_{[1,1,1]}$ is the probability that none are IBD, and therefore nonfocal group members are random samples from the population.

The expected payoff to the s_1 strategist $\overline{\pi}_1$ can also be written in terms of relatedness. The triadic relatedness r_3 is the probability that 3 individuals drawn from the group will be IBD, which is $r_3 = F_{[3]}$. The dyadic relatedness r_2 is the probability that 2 individuals drawn from the group without replacement will be IBD, which is $r_2 = F_{[3]} + (F_{[2,1]}/3)$. These give $F_{[3]} = r_3$, $F_{[2,1]} = 3(r_2 - r_3)$, and $F_{[1,1,1]} = 1 - 3r_2 + 2r_3$. Therefore, the expected payoff to an s_1 strategist can be written in terms of the triadic and dyadic relatedness as

$$\begin{aligned}\overline{\pi}_1 = {} & r_3\pi(\boldsymbol{e}_1,2\boldsymbol{e}_1) + (r_2 - r_3)\sum_{i=1}^{m} p_i\left[\pi(\boldsymbol{e}_1,2\boldsymbol{e}_i) + 2\pi(\boldsymbol{e}_1,\boldsymbol{e}_1+\boldsymbol{e}_i)\right] \\ & + (1 - 3r_2 + 2r_3)\sum_{j=1}^{m}\sum_{k=1}^{m} p_j p_k \pi(\boldsymbol{e}_1,\boldsymbol{e}_j+\boldsymbol{e}_k).\end{aligned} \tag{18}$$

Similar to before, if $r_2 = r_3 = 0$, then $\overline{\pi}_1 = \overline{\pi}_1^{\text{mix}}$; and if $r_2 = r_3 = 1$, then $\overline{\pi}_1 = \pi(\boldsymbol{e}_1,2\boldsymbol{e}_1)$.

In general, for games with n players, the probability distribution of group compositions, and thus the expected payoffs, are written in terms of the set of family composition probabilities $\{F_{\boldsymbol{q}}\}$, where $\boldsymbol{q}$ ranges over all possible partitions of integer n by positive integers. The "partition function" P_n counts the number of possible partitions of n. For example, the first six are $P_1 = 1$, $P_2 = 2$, $P_3 = 3$, $P_4 = 5$, $P_5 = 7$ and $P_6 = 11$. The sum of probabilities $F_{\boldsymbol{q}}$ over all $\boldsymbol{q}$ is one, so we need $P_n - 1$ free parameters to describe the dynamics.

For $n > 3$ players, more complicated relatedness terms than dyadic, triadic, tetradic, etc., are needed. For $n = 2$, $P_2 - 1 = 1$ suggests that r_2 is enough; for $n = 3$, $P_3 - 1 = 2$ suggests that r_2, r_3 are enough; however, for $n = 4$, $P_4 - 1 = 4$ suggests that r_2, r_3, r_4 are not enough and an additional term is needed. In SI B, we characterise these additional relatedness terms, and we sketch a scheme for choosing which among the possible $P_n - 1$ relatedness terms to include to fully parameterise the system. In SI C, we show how the form of the payoff function may allow one to parameterise the model with fewer than $P_n - 1$ relatedness coefficients (e.g., a linear public goods game can be parameterised with only one relatedness coefficient, r_2).

3.1.2. Payoff transformation

Because, in our framework, the relatedness coefficients are independent of the strategy frequencies, the dynamics under homophily are equivalent to the dynamics in a transformed game in a well-mixed population where the transformation modifies the payoffs in a way that accounts for relatedness (Grafen, 1979). A 2-player game can be expressed in a matrix form, and thus the transformation of the payoffs can be expressed in terms of a transformed payoff matrix (e.g., Van Veelen, 2011; García et al., 2014). Analogously, an n-player game can be represented by a multidimensional matrix of dimension n, so we can obtain a transformed multidimensional matrix. Payoff-matrix transformations can be used to explore different mechanisms of inducing population structure (e.g., games played on graphs (Ohtsuki and Nowak, 2006)) and different cooperation mechanisms apart from kin selection (Taylor and Nowak, 2007) (e.g., the tit-for-tat strategy transforms the Prisoner's Dilemma payoff matrix into a Stag Hunt matrix Bowles and Gintis, 1998). The advantage of using a payoff-matrix transformation is that all of the techniques developed to study well-mixed populations can now be used study the more complex game. For example, by using a payoff-matrix transformation, Peña et al. (2015) was able to harness the theory of Bernstein polynomials (Peña et al., 2014) to study n-player, continuous 2-strategy games with genetic assortativity.

To gain the intuition, consider the generic 3-player m-strategy game above (Eq. (17)).

We seek a transformed payoff function π' that satisfies

$$\overline{\pi}_i = \sum_{j=1}^{m}\sum_{k=1}^{m} p_j p_k \pi'(\boldsymbol{e}_i,\boldsymbol{e}_j,\boldsymbol{e}_k), \tag{19}$$

which is the expression for the expected payoff to s_i-strategists in a well-mixed population. We also wish to preserve the symmetry of the original π function (e.g., $\pi(\boldsymbol{e}_i,\boldsymbol{e}_j+\boldsymbol{e}_k) = \pi(\boldsymbol{e}_i,\boldsymbol{e}_k+\boldsymbol{e}_j)$), so we require that the transformed payoff matrix also have the symmetry (e.g., $\pi'(\boldsymbol{e}_i,\boldsymbol{e}_j,\boldsymbol{e}_k) = \pi'(\boldsymbol{e}_i,\boldsymbol{e}_k,\boldsymbol{e}_j)$). We find that the following transformation satisfies our requirements

$$\begin{aligned}\pi'(\boldsymbol{e}_i,\boldsymbol{e}_j,\boldsymbol{e}_k) = {} & F_{[3]}\pi(\boldsymbol{e}_i,2\boldsymbol{e}_i) \\ & + F_{[2,1]}\left[\frac{1}{3}\left(\frac{\pi(\boldsymbol{e}_i,2\boldsymbol{e}_j)+\pi(\boldsymbol{e}_i,2\boldsymbol{e}_k)}{2}\right)\right. \\ & \left. + \frac{2}{3}\left(\frac{\pi(\boldsymbol{e}_i,\boldsymbol{e}_i+\boldsymbol{e}_j)+\pi(\boldsymbol{e}_i,\boldsymbol{e}_i+\boldsymbol{e}_k)}{2}\right)\right] \\ & + F_{[1,1,1]}\pi(\boldsymbol{e}_i,\boldsymbol{e}_j+\boldsymbol{e}_k).\end{aligned} \tag{20}$$

This approach can be generalised to the n-player, m-strategy game, and the methods for obtaining the transformed payoff matrix, describing the dynamics, assessing stability, and testing invasion fitness, are described in more detail in SI D. We have also made a detailed comparison between the payoff transformation proposed here and relevant results in kin selection literature in SI E.

3.2. Accounting based on whole-group composition

Substituting the payoff to the focal individual from Eq. (4) into Eq. (7), we obtain the dynamics

$$\Delta p_x \propto \mathbb{E}[G_{0,x}\hat{\pi}(\boldsymbol{e}_x,\boldsymbol{G}_{\text{a}})] - p_x\sum_{i=1}^{m}\mathbb{E}[G_{0,i}\hat{\pi}(\boldsymbol{e}_i,\boldsymbol{G}_{\text{a}})]. \tag{21}$$

Note that previously (Eq. (8)), we calculated payoff as a function of focal and non-focal members' strategies $\pi(\boldsymbol{e}_i,\boldsymbol{G}_{\text{nf}})$, whereas in Eq. (21), we calculate payoff as a function of focal and all members' strategies $\hat{\pi}(\boldsymbol{e}_i,\boldsymbol{G}_{\text{a}})$. The random variable $\boldsymbol{G}_{\text{a}}$ can take values from the set

$$\mathcal{G}_{\text{a}} = \left\{\boldsymbol{g}_{\text{a}} \in \mathbb{N}_{\geq 0}^{m} \;\middle|\; \sum_{i=1}^{m} g_{\text{a},i} = n\right\}. \tag{22}$$

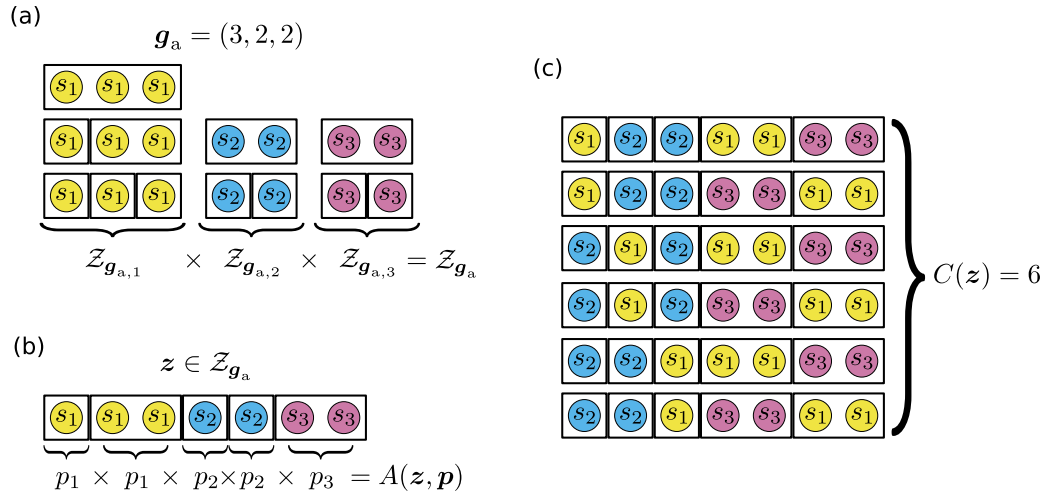


Fig. 1. An example calculating elements that contribute to $\mathbb{P}[\boldsymbol{G}_{\mathrm{a}} = \boldsymbol{g}_{\mathrm{a}}]$ where circles represent individuals in the group, different colours indicate different strategies pursued, and boxes delineate family members. (a) The set of all strategywise family-size distributions consistent with the strategy composition $\boldsymbol{g}_{\mathrm{a}} = (3,2,2)$, $\mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}}$, is the Cartesian product of distributions $\mathcal{Z}_{\boldsymbol{g}_{\mathrm{a},i}}$ that are consistent with the number of individuals pursuing each strategy i (see also Example 3 in SI A.4). (b) For a given strategywise family-size distribution $\boldsymbol{z}$, the probability of drawing those strategies for each family, $A(\boldsymbol{z}, \boldsymbol{p})$, is a product of strategy frequencies in the population. (c) The number of ways to allocate strategies to families, $C(\boldsymbol{z})$, is the count of multiset permutations of strategies across family-group sizes.

Eq. (21) can be written as (SI A.4)

$$\begin{aligned}\Delta p_x &\propto \mathbb{E}\left[\frac{G_{\mathrm{a},x}}{n}\hat{\pi}(\boldsymbol{e}_x, \boldsymbol{G}_{\mathrm{a}})\right] - p_x \sum_{i=1}^{m} \mathbb{E}\left[\frac{G_{\mathrm{a},i}}{n}\hat{\pi}(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{a}})\right] \\ &= \sum_{\boldsymbol{g}_{\mathrm{a}} \in \mathcal{G}_{\mathrm{a}}} \left(\frac{g_{\mathrm{a},x}}{n}\hat{\pi}(\boldsymbol{e}_x, \boldsymbol{g}_{\mathrm{a}}) - p_x \sum_{i=1}^{m} \frac{g_{\mathrm{a},i}}{n}\hat{\pi}(\boldsymbol{e}_i, \boldsymbol{g}_{\mathrm{a}})\right)\mathbb{P}[\boldsymbol{G}_{\mathrm{a}} = \boldsymbol{g}_{\mathrm{a}}].\end{aligned} \tag{23}$$

To derive $\mathbb{P}[\boldsymbol{G}_{\mathrm{a}} = \boldsymbol{g}_{\mathrm{a}}]$, which is the probability that the group composition is $\boldsymbol{g}_{\mathrm{a}}$, denote the family-size distribution of the $g_{\mathrm{a},i}$ individuals in the group who are s_i strategists as $\boldsymbol{z}_i = (z_{i,1}, \ldots, z_{i,n})$, which is a vector of length n whose kth element counts the number of families of size k among group members pursuing strategy s_i (see SI A for examples). Define the whole group's strategywise family-size distribution as the vector of each strategy's family-size distribution $\boldsymbol{z} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m)$, which is a collection of m vectors. Then the set of all family-size distributions consistent with $g_{\mathrm{a},i}$ is

$$\mathcal{Z}_{g_{\mathrm{a},i}} = \left\{\boldsymbol{z}_i \;\middle|\; \sum_{j=1}^{n} j\, z_{i,j} = g_{\mathrm{a},i}\right\}, \tag{24}$$

and the set of all strategywise family-size distributions consistent with $\boldsymbol{g}_{\mathrm{a}}$ is the Cartesian product

$$\mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}} = \bigtimes_{i=1}^{m} \mathcal{Z}_{g_{\mathrm{a},i}}. \tag{25}$$

The whole group's family-size distribution is $\boldsymbol{y} = \mathrm{sum}(\boldsymbol{z}) \left(:= \sum_{i=1}^{m} \boldsymbol{z}_i\right)$.

With these, the probability that the strategy distribution $\boldsymbol{g}_{\mathrm{a}}$ is realised through a particular group's strategy-wise family distribution $\boldsymbol{z} \in \mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}}$ is the probability $F_{\boldsymbol{y}}$ that a group has family-size distribution $\boldsymbol{y} = \mathrm{sum}(\boldsymbol{z})$, times the number of ways to allocate strategies to families, which is given by the product of multinomial coefficients

$$C(\boldsymbol{z}) = \prod_{k=1}^{n} \frac{\left(\sum_{i=1}^{m} z_{i,k}\right)!}{\prod_{i=1}^{m} z_{i,k}!}, \tag{26}$$

times the probability of drawing the family strategies properly to each family, which is given by

$$A(\boldsymbol{z}, \boldsymbol{p}) = \prod_{i=1}^{m} p_i^{\|\boldsymbol{z}_i\|}, \tag{27}$$

where $\|\boldsymbol{z}_i\|$ is the number of families in the group pursuing strategy s_i, i.e., $\|\boldsymbol{z}_i\| = \sum_{k=1}^{n} z_{i,k}$ (Fig. 1).

Therefore, we obtain (see SI A.4 for detailed derivation)

$$\mathbb{P}[\boldsymbol{G}_{\mathrm{a}} = \boldsymbol{g}_{\mathrm{a}}] = \sum_{\boldsymbol{z} \in \mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}}} C(\boldsymbol{z})\, A(\boldsymbol{z}, \boldsymbol{p})\, F_{\mathrm{sum}(\boldsymbol{z})}. \tag{28}$$

As a result, the change in the proportion of s_x in the population is

$$\begin{aligned}\Delta p_x \propto &\sum_{\boldsymbol{g}_{\mathrm{a}} \in \mathcal{G}_{\mathrm{a}}} \left(\frac{g_{\mathrm{a},x}}{n}\hat{\pi}(\boldsymbol{e}_x, \boldsymbol{g}_{\mathrm{a}}) - p_x \sum_{i=1}^{m} \frac{g_{\mathrm{a},i}}{n}\hat{\pi}(\boldsymbol{e}_i, \boldsymbol{g}_{\mathrm{a}})\right) \\ &\left(\sum_{\boldsymbol{z} \in \mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}}} C(\boldsymbol{z})\, A(\boldsymbol{z}, \boldsymbol{p})\, F_{\mathrm{sum}(\boldsymbol{z})}\right).\end{aligned} \tag{29}$$

Similarly to Section 3.1, by taking a proper limit (for example, letting $\delta(1-s) \to 0$ and taking a proper time scale) in Eq. (7), the above equation becomes a continuous-time differential equation, as

$$\begin{aligned}\dot{p}_x = &\sum_{\boldsymbol{g}_{\mathrm{a}} \in \mathcal{G}_{\mathrm{a}}} \left(\frac{g_{\mathrm{a},x}}{n}\hat{\pi}(\boldsymbol{e}_x, \boldsymbol{g}_{\mathrm{a}}) - p_x \sum_{i=1}^{m} \frac{g_{\mathrm{a},i}}{n}\hat{\pi}(\boldsymbol{e}_i, \boldsymbol{g}_{\mathrm{a}})\right) \\ &\left(\sum_{\boldsymbol{z} \in \mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}}} C(\boldsymbol{z})\, A(\boldsymbol{z}, \boldsymbol{p})\, F_{\mathrm{sum}(\boldsymbol{z})}\right).\end{aligned} \tag{30}$$

Our approach assumes that the level of genetic homophily in group formation is independent of the strategies in the group. Therefore, to reduce the computational effort required to numerically analyse Δp_x or $\dot{p}_x$, we precalculate the values $C(\boldsymbol{z})$ and powers $\|\boldsymbol{z}_i\|$ in the expression for $A(\boldsymbol{z}, \boldsymbol{p})$ for an arbitrary ordering of strategies and given (n, m) combination. These values are stored and can be referenced at each necessary stage of the calculations (see SI F for details). In the following section, we illustrate the approach's use with an example.

4. Example: coordinated cooperation in a public goods game with sigmoidal benefits

The aim of this section is to demonstrate the use of our theoretical framework. In particular, we study a public goods game in a scenario where interactions were between relatives in the ancestral past but are between non-kin today.

4.1. Background

Shared intentionality is a key feature distinguishing human cooperation from that of other apes, and involves individuals forming a collective 'we' with a jointly optimised goal and coordinating their actions

towards achieving it (Genty et al., 2020; Tomasello, 2020). This capacity, theorised to have evolved in coordination problems with interdependent participants and mutually beneficial outcomes, underpins humans' ability to cooperate with nonkin (Tomasello et al., 2012). It is often observed in experimental threshold public goods games (Palfrey and Rosenthal, 1984; Cadsby and Maynes, 1999; Archetti and Scheuring, 2012), where groups designate a threshold number of members to contribute, and those designated follow through on their commitments (Van de Kragt et al., 1983; Mak et al., 2015; Palfrey et al., 2017). This behaviour aligns with intuition: when 8 people face a task requiring 4, it is logical to attempt to designate 4 members to the task, and failing that, to abstain to avoid wasting effort.

Newton (2017b) demonstrated that shared intentionality evolves under fairly general conditions in a public goods game (PGG), specifically proving that shared intentionalists can both invade a population of defectors and persist. In this context, a PGG was defined as a game where one player's contribution never negatively affects another player. Shared intentionality was modelled as agents coordinating their actions to achieve higher payoffs than the Nash equilibrium (typically universal defection). For example, in a discrete public goods game, shared intentionalists might agree to cooperate only if a threshold number of others also commit to cooperation.

Despite these promising findings, two caveats to the evolution of shared intentionality remain. First, if contribution conflicts with individual rationality (e.g., as will occur in a linear PGG), shared intentionalists may be vulnerable to 'liars' who signal willingness to contribute but do not follow through (Newton, 2017b). Second, if coordination capacity entails a cost (e.g., cognitive abilities, communication), shared intentionalists may struggle to invade a population of defectors who do not pay these costs.

We investigated these caveats, beginning with the premise that, over the course of human lineage, there has been a shift from kin to nonkin interactions (Kuhn et al., 2001; Gamble et al., 2011; Brooks et al., 2018; Sehasseh et al., 2021; Ringbauer et al., 2021). Consequently, we modelled kin selection to facilitate the initial invasion of Coordinated Cooperation (cf., Boyd et al., 2010; Takezawa and Price, 2010; Boyd et al., 2014; Schonmann and Boyd, 2016; Kristensen et al., 2022). We also examined the strategy's robustness to declining homophily over time, particularly its ability to resist invasion by liars. Previous research has also shown that even unconditional cooperation can persist evolutionarily in well-mixed populations if the public good's benefit has a nonlinear relationship with the number of cooperators (Peña et al., 2014). This persistence occurs in threshold PGGs (Palfrey and Rosenthal, 1984; Cadsby and Maynes, 1999; Archetti and Scheuring, 2012) and sigmoid-shaped games (Bach et al., 2006; Boza and Számadó, 2010; Archetti and Scheuring, 2011; Peña et al., 2014; Archetti, 2018). Therefore, we modelled a PGG with a flexible sigmoid-shaped benefits function. To model coordination, we modelled agents who use a random lottery to designate contributors, a method that is often observed in experimental threshold games (Van de Kragt et al., 1983) and has been used historically to assign tasks and allocate burdens (Elster, 1988).

4.2. Model

We model a nonlinear PGG where the benefit returned to *each* group member $B(k)$ has a sigmoid relationship with the number of individuals k who contribute. Contributors pay a cost c, and every group member receives benefit $B(k)$ regardless of whether or not they contributed. We constrain the parameter values so that contributing will present a social dilemma to a lone contributor ($B(1) < c$).

We model $m = 4$ strategies split into two types: unconditional and communicative. Unconditional types act independently and always play the same strategy. Unconditional Cooperators (U) always contribute and Unconditional Defectors (D) always defect. They do not discuss their plans with other group members, and it is assumed they are incapable of doing so.

Communicative players use a lottery (e.g., drawing straws) to determine who among them will contribute to the PGG. We make the simplifying assumption that they have enough experience or insight into the game to aim for τ contributors, which corresponds to the inflection point of the sigmoid benefits function, and is the contributor-group size that maximises the incentive for each contributor to remain a contributor (i.e., maximises $B(\tau) - c - B(\tau - 1)$). We assume that these cognitive and communicative abilities entail a small cost ϵ that is paid whether or not the lottery takes place. Importantly, we assume that the agreement reached via the lottery is only verbal. Verbal agreement allows for a 'lying' strategy that participates in the lottery but defects if chosen, thus inducing contributions from others and then free-riding on those contributions. Thus, we model two communicative types: Coordinating Cooperators (C), who participate in the lottery and follow-through if chosen; and Liars (L), who participate in the lottery, but will not contribute even if they are chosen.

The game is played in two stages. In the first stage, if the number of communicative group members (C + L) meets or exceeds the quorum τ, then the lottery takes place, and τ participants are randomly chosen to be contributors. If the quorum is not met, the lottery does not take place. In the second stage, group members independently decide whether or not to contribute to the public good: U always contribute; D and L always defect; and C contribute if a lottery took place and they were chosen, otherwise they defect.

We model the shift in human social structures—from interactions primarily among kin to those among nonkin—as a trend of decreasing homophily in time. Specifically, we adopt the leader-driven group-formation model introduced in Section 2.3, where h is the parameter representing the degree of homophily (Eq. (1)). We obtain analytical results for two extreme cases in a small-group setting: an ancestral state with perfect homophily ($h = 1$), where groups consist exclusively of family members; and a contemporary state with zero homophily ($h = 0$), where the population mixes freely. We also study how the change in homophily level, especially from larger values to smaller ones, affect the result. We then extend our investigation to larger groups using numerical methods.

4.3. Results

4.3.1. Analytic results from a three-player game

To build our intuition before exploring a larger numerical example (next section), we performed an analysis of a three-player game ($n = 3$) with threshold $\tau = 2$. We defined a sigmoid-shaped benefits function

$$B(k) = \begin{cases} 0 & \text{if } k = 0, \\ \beta & \text{if } k = 1, \\ b - \beta & \text{if } k = 2, \\ b & \text{if } k = 3. \end{cases} \tag{31}$$

where β is a parameter that we can vary to explore the effect of nonlinearity on the dynamics, and we imposed the constraint $\beta < c$ to ensure that contributing presented a social dilemma to a lone contributor.

We explored games between the two extremes of nonlinearity, from the threshold PGG when $\beta = 0$ to the linear PGG when $\beta = b/3$ (Fig. 2). We also compared between the two extremes of homophily—perfect homophily ($h = 1$) representing an ancestral state versus zero homophily ($h = 0$) representing a contemporary society—such that our analytical results are independent of the group-formation model. To perform the analysis under homophily, we derived an expression for the transformed payoff matrix, which allowed us to treat the dynamics as though they were in a well-mixed population. We analysed the dynamics of pairs of strategies using the techniques of Peña et al. (2015), and we investigated whether or not new strategies can invade coexisting pairs. See SI G for the full analysis.

As a result, we found that in a well-mixed population without any family structure (contemporary scenario, $h = 0$), Coordinating Cooper-

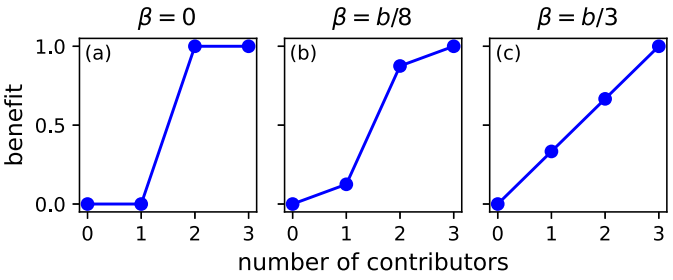


Fig. 2. The benefits returned from the public good as a function of the number of contributors for (a) $\beta = 0$, threshold game; (b) $\beta = b/8$, sigmoid game; and (c) $\beta = b/3$, linear game. The maximum benefit has been normalised to $b = 1$.

ators can persist in stable coexistence with Defectors (Fig. 3a). Coexistence can occur even in situations where Unconditional Cooperators cannot persist, including when the benefits function is linear (necessary condition in Eq. G.18). However, because of the cognitive cost ε, Coordinating Cooperators cannot invade from rarity into a population of all Defectors, which raises the question of how Coordinated Cooperation first arose.

We found that Coordinated Cooperation could have arisen in the ancestral past if there was some degree of homophily (cf., Kristensen et al., 2022). Homophily facilitates the invasion of C from rarity into a population of all-D (SI G.4.1). If the ancestral state had perfect homophily ($h = 1$), then either a population of all Coordinating Cooperators was stable (as shown in Fig. 3a) or a population of all Unconditional Cooperators was stable (Fig. G.3), depending on the parameter values (SI G.4.5). If the ancestral state was all-U, then as homophily declined over time, the population would eventually become invadable by Coordinating Cooperators (SI G.4.5).

Assuming that Coordinating Cooperators evolved in one of the ways described above, we must still explain how it persists today, particularly how the population resists invasion by liars (Fig. 3b). We found that, when homophily is zero, Liars can invade the C + D coexistence if $b - 2\beta < c$ (SI G.5.2). This occurs when the contribution cost c is high, and it is guaranteed when the PGG is linear ($\beta = b/3$) provided contribution presents a social dilemma to a lone contributor ($\beta < c$). However, if

$$b - 2\beta > c, \tag{32}$$

then Liars can never invade a C + D coexistence regardless of the homophily level (SI G.5.2). Eq. (32) is also a necessary condition for the strategy profile with 2 contributors to be a Nash equilibrium in the well-mixed game with untransformed payoffs (Eq. H.15). For later use, we note that Eq. (32) can be re-written $B(2) - B(1) > c$.

4.3.2. Numerical results from a many-player game

We generalised the sigmoid benefits function for a many-player game (Archetti, 2018)

$$B(k) = \frac{L(k) - L(0)}{L(n) - L(0)}; \qquad L(k) = \frac{1}{1 + e^{\sigma(\tau - 0.5 - k)/n}}, \tag{33}$$

where n is the number of players, k is the number of contributors, $\tau - 0.5$ is the midpoint of the sigmoid benefits function, and σ is the steepness parameter. When $\sigma = 0$, the game is a linear PGG, and as $\sigma \to \infty$, the game approaches the threshold game.

Similar to the analytical results from the 3-player game above, we found that the evolutionary dynamics in a many-player game can be divided into two main regimes depending on the degree of nonlinearity in benefit function (SI H.2.1). If the switching gain from non-contribution to contribution at the coordination point is greater the cost of contribution

$$B(\tau) - B(\tau - 1) > c, \tag{34}$$

which occurs when contribution cost is low and the benefit function is more nonlinear, then in a well-mixed population, the C + D coexistence can resist invasion by Liars. Note that Eq. (34) corresponds to Eq. (32) for a 3-player game. Eq. (34) is also the condition for an all-C population to resist invasion by Liars (SI H.2.4); a necessary condition for the strategy profile with τ contributors to be a Nash equilibrium in the well-mixed game (SI H.2.2); and the condition under which Unconditional Cooperators will have a positive switching gain against Unconditional Defectors, which is a necessary but not sufficient condition for a coexistence between Unconditional Cooperators and Defectors.

Next, we studied a numerical example for each of the two regimes, and we explored how the dynamics changed as homophily declined from an ancestral state of perfect homophily to a state of zero homophily today.

For a scenario where Eq. (34) is satisfied, as anticipated by the 3-player example, we found an evolutionary trajectory that ended with a stable coexistence between Coordinated Cooperation and Unconditional Defection (Fig. 4a; details in I.1). When ancestral homophily was high, Unconditional Cooperation was the ancestral state. As homophily declined, the all-U population became invadable by Coordinated Cooperators, resulting in a U + C coexistence. A further decline in homophily allowed Unconditional Defectors to invade, resulting in a C + D coexistence that persisted until zero homophily. Thus, the evolution of Coordinated Cooperation between strangers is achieved.

However, the trajectory above assumes that, as homophily declines, each evolutionary steady state is invaded soon after it becomes invadable by a new strategy. When this assumption is relaxed, other trajectories are possible that end with an all-Defector population instead of C + D. For example, if the initial all-U population persists for long enough as homophily declines, then it may be invaded by D first instead of C, resulting in a U + D coexistence (see the arrow from "U invadable" to the yellow "U + D" bar in Fig. 4a). If homophily declines further and Coordination is not 'invented' soon enough, then the population will become 'trapped' in the U + D coexistence (see the uninvadable part of the yellow "U + D" bar that is marked with borders in Fig. 4a). As anticipated by the 3-player model (SI G.5.3), trapping occurs at the same homophily level that the all-D population becomes uninvadable to C. Specifically, a separatrix appears near the U-D axis (shown as a red dotted line in Fig. 5), and separates the C + D steady state from both the all-D and U + D states trapping populations there. As homophily declines further, the separatrix moves right while the stable-unstable U + D steady-state pair move towards each other. Eventually, the pair collides, the U + D coexistence disappears, and the population will end in an all-D state (see the arrow from "U + D" to the red "D" bar in Fig. 4a).

For our second model scenario, where Eq. (34) is not satisfied, the nonlinearity level σ tends to be smaller that the first scenario. In this scenario, the strategy profile with τ contributors is no longer a Nash equilibrium in the well-mixed game, and the only Nash equilibrium is non-contribution.

We found that some level of homophily must be maintained to prevent cooperation from being entirely lost from the population (Fig. 4b, details in I.2). Initially, as homophily decreased from the ancestral state, the evolutionary dynamics changed in similar ways to the first example, with an evolutionary route from all-U through U + C coexistence to C + D coexistence (see the arrows from "U" to "U + C" and from "U + C" to "C + D" in Fig. 4b). However, as homophily declined further, the C + D coexistence became invadable by Liars, which were subsequently invaded by D, leading to an all-Defector evolutionary endpoint (see the arrows from "C + D" to "C + D + L" and from "C + D + L" to "D" in Fig. 4b).

4.3.3. Results about computational costs

As the group size and number of strategies increases, the computation time required to evaluate $\dot{\boldsymbol{p}}$ increases combinatorially (Fig. 6a). This occurs because the total number of possible group strategy compositions increases combinatorially (cf. Allen et al., 2024), and the number of family-partition structures also increases at an increasing though lesser rate (SI J). In the Coordinated Cooperation example above, we

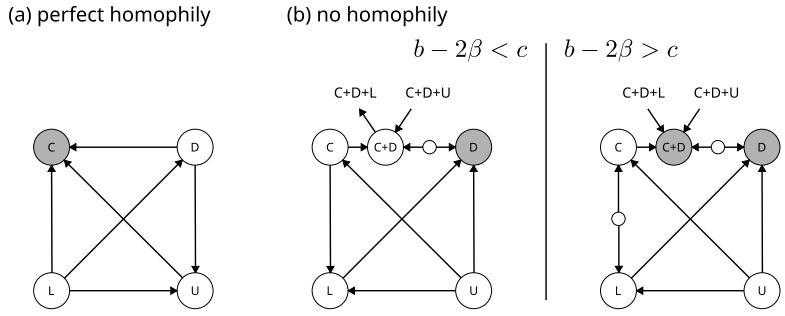


Fig. 3. An example of the evolutionary dynamics between the four strategies, where arrows represent the evolutionary trajectories between population states, and shaded nodes indicate uninvadable evolutionary end-states. In this example, (a) under perfect homophily, Coordinated Cooperation is the global evolutionary end-state (see Fig. G.3 for the alternative case where Unconditional Cooperation is the end-state). (b) When the population is well-mixed, Unconditional Defectors are always one possible end-state. If the condition $b - 2\beta > c$ (Eq. (32)) is satisfied, then a coexistence between Coordinating Cooperators and Defectors (C + D) is also a possible end-state; however, if the condition is not satisfied, then C + D can be invaded by Liars.

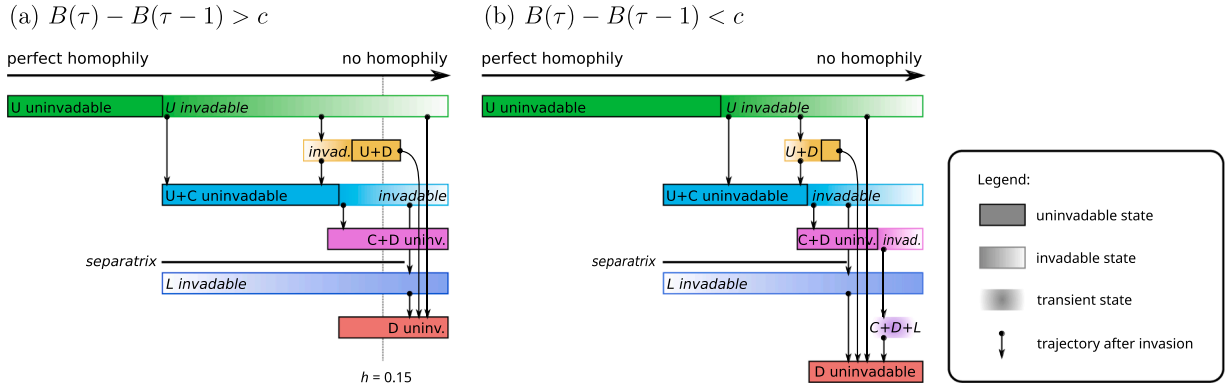


Fig. 4. A qualitative summary of the evolutionary dynamics as homophily declines over time for scenarios where the strategy profile with τ contributors is (a) a Nash equilibrium in the well-mixed game ($\sigma = 10$), or (b) not a Nash equilibrium ($\sigma = 6$). Horizontal bars represent population states and arrows indicate the evolutionary trajectories after an invasion. To simplify the summary, we assume that new strategies invade one at a time; that the evolutionary dynamics have enough time to stabilise before the next invasion, i.e., separation of timescales; and that Liars cannot be invaded before Coordinating Cooperators, i.e., it does not make sense to lie about an activity that has not yet been invented (see SI for full analysis). Default parameter values: $n = 8$, $\tau = 5$, $c = -0.25$, and $\varepsilon = 0.02$. The evolutionary dynamics when homophily $h = 0.15$ are plotted in Fig. 5.

were able to achieve numerical tractability because we considered only 4 strategies. In general, numerical tractability is restricted to problems with a modest group size and number of strategies.

Whether it is faster to obtain solutions using the whole-group accounting (Eq. (29)) or other-member accounting (e.g., Eq. (20)) depends on the particulars of the problem. For example, when numerically solving for the steady states, $\dot{\boldsymbol{p}}$ must be evaluated many times. When using the other-member accounting, the transformed payoffs can be stored in a matrix form (details in SI D), and then each evaluation of $\dot{\boldsymbol{p}}$ is obtained by repeated matrix multiplications. Calculating the transformed payoffs entails a high overhead cost; however, the overhead can be worth the trade-off in efficiency gained for each evaluation of $\dot{\boldsymbol{p}}$ (e.g., Fig. 6b), which is efficient because because matrix multiplication is an operation for which numerical software is typically optimised.

5. Discussion

We have provided an analytical description for the evolutionary dynamics of strategies in games with kin assortativity. Our approach involves the use of genetically homophilic group-formation models, which provide the higher-order relatedness coefficients needed to describe the dynamics, and we have demonstrated how the dynamics can be solved with examples. We studied a public goods game with 4 strategies, where both the Coordinated Cooperation and Liar strategies required insight into the game and communication abilities. We discuss each in turn below.

5.1. General approach

We have provided an analytical description of the evolutionary dynamics of discrete strategies in group games where group members are a mix of kin and nonkin and game payoffs are not necessarily additive. When games are played between 2 or 3 individuals, the dynamics can be described analytically in terms of dyadic and triadic relatedness (Eqs. (16) and (17)). However, as the group size and number of strategies increases, the number of relatedness terms needed increases combinatorially (cf. Allen et al., 2024). Nevertheless, we found that it is still practical to solve the dynamics numerically for modestly sized scenarios of theoretical interest.

Our central results are as follows. In Section 3.1 we have derived replicator equations, Eqs. (11), (12), for family structured populations. We have found that they are described by family partition probabilities $\{F_q\}$ or by a proper set of relatedness coefficients (detailed in Section 3.1.1; see also SI B and SI C). We have also found a formula to transform the payoff functions. These transformed payoff functions allow us to study evolutionary game dynamics in a family-structured population as if it were a well-mixed population (detailed in Section 3.1.2; see also SI D). Our method can be used to rederive previous results from specific models of synergistic interactions, and we do so for a selection of models (Queller, 1985; Gardner et al., 2011; Taylor and Maciejewski, 2012; Taylor, 2017) in SI E. Moreover, we have derived another game-dynamics equations based on the whole-group accounting, Eqs. (29), (30), which are mathematically equivalent to the replicator equations we have derived in Eqs. (11), (12). Although those equations look less

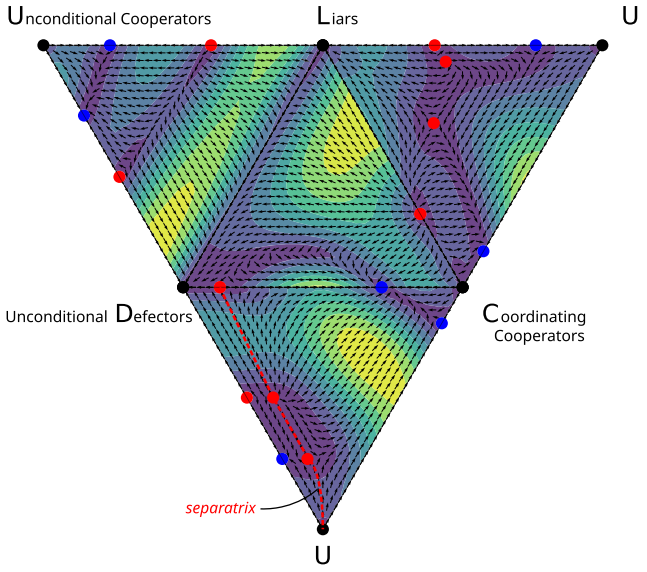


Fig. 5. Evolutionary dynamics and steady states on the faces of the tetrahedral strategy space for the scenario where Eq. (34) is satisfied (Fig. 4 a, τ contributors is a Nash equilibrium in the well-mixed game) when homophily is low ($h = 0.15$). Blue dots represent states of stable coexistence between the strategies present in the population that may nevertheless be invadable by other strategies, and red dots represent unstable equilibria (2 unstable interior equilibria not shown because they lie inside the tetrahedron). A population at the U + D coexistence cannot be invaded by Coordinating Cooperators due to the separatrix (red dashed) on the (D, U, C) face (the bottom triangle). As homophily declines further, the stable-unstable equilibria pair on the D-U axis will collide and disappear and the population will evolve to an all-D state. Therefore, if Coordinated Cooperation has not been established in the population before the collision, then cooperation will be lost from the population.

intuitive, they are computationally more tractable, and those expressions greatly help us implement the code for numerical calculations (detailed in Section 3.2; see also SI F).

Our approach is particularly aimed at scenarios where (1) group members are potentially related, (2) fitness effects are not necessarily additive, and (3) strategies are best modelled as discrete. Historically, there has been a split (discussed in Ohtsuki, 2014) between kin-selection models that address point (1) (e.g., Taylor and Frank, 1996) and evolutionary game-theoretic models that address points (2) and (3) (e.g. Hauert et al., 2006). Nevertheless, all three can co-occur. For example, the benefits returned from punishment likely have a nonlinear relationship with the number of punishers (Raihani and Bshary, 2011; Roberts, 2013; Raihani and Bshary, 2015), and food sharing between relatives (Gurven et al., 2000; Schweinfurth and Call, 2019; Jaeggi and Gurven, 2013) combines inclusive fitness benefits with reciprocity (Allen-Arave et al., 2008) in potentially synergistic ways (Jones, 2000; Van Cleve and Akçay, 2014).

Our assumption that groups form and dissolve with each interaction (Godfrey-Smith and Kerr, 2009) implies that offspring compete globally, which is an assumption that favours the evolution of cooperation under kin assortativity (West et al., 2006). This can be contrasted with limited dispersal models where the inclusive fitness benefits of increasing the fecundity of one's kin can be cancelled out by the concordant increase in competition between kin (Taylor, 1992). It can also be contrasted with social learning models where high-payoff strategies are copied from group members instead of the population at large, and where kin assortment can also hinder the evolution of cooperation (Martin and Lessard, 2023).

We have assumed that groups' family partition-structure probabilities are independent of strategy composition, which is a very strong assumption. For example, in the Wright's infinite-islands model, where an infinite population is split into an infinite number of subpopulations ('islands') with occasional dispersal between them (Wright, 1931; Lehmann and Rousset, 2010; Rousset, 2013), this assumption can be easily violated, because under natural selection different strategies can have different genealogical relationships. Moreover, under natural selection, strategy frequencies in the current population are generally different from those frequencies in the past, which is another obstacle in formulating the dynamics. In our approach, dynamical sufficiency is obtained by assuming that family partition probabilities are independent of strategy composition and that the strategy that each family adopts is determined by the current frequencies of strategies in the population. Those ideal assumptions can hold in a broader class of group-formation models under neutrality, where there are no difference between "strategies", so our approach remains valid as an approximation if one assumes that strength of selection, δ, is weak enough (see SI K). We highlight this as an area for future work.

Our group-formation models may be directly reinterpreted in the style of the 'matching rule' models commonly used in economics, which we hope will be useful to workers in that field. Economists typically model homophilic attraction on the basis of similarity in strategy rather than ancestry, and the resulting matching rule is a function that maps from the population strategy frequencies ($\boldsymbol{p}$) to the probability distribution of different group strategy compositions ($\mathbb{P}[\boldsymbol{G} = \boldsymbol{g}]$) (e.g., Bergstrom, 2003). Matching-rule models have been used to study a diverse range of topics including Kantian morality (Alger and Weibull, 2013), supply-chain dynamics (Chai et al., 2021), cross-cultural cooperation (Bilancini et al., 2018), and feedback from group-level changes in assortativity through democratic (Nax and Rigos, 2016; Wu et al., 2016) or top-down (Wu, 2019) processes. However, although the matching-rule formalism (Jensen and Rigos, 2018) can be applied to any n-player m-type scenario, most work has focused on 2 players and 2 types. Applying matching rules in many-player many-strategy scenarios is difficult for the same reason as in kin-assortative models: one must account for the statistical dependency between the focal and nonfocal members' strategies (Alger and Weibull, 2014, 2019). However, one can account for that dependency using our models by reinterpreting our genetically homophilic process in terms of strategy matching: replace 'recruitment of a family member' with 'recruitment of the same type', and replace 'recruitment of a non-family member' with a 'mistake' that results in a random draw from the population.

We wish to highlight some related modelling approaches that circumvent some of the assumptions made in our own work. First, we have assumed a fixed group size; however, Garcia and De Monte (2013) modelled a variable-sized public goods game where group formation was by players (microbes) physically adhering to one another. Second, we have assumed that the order in which individuals are recruited to the game does not influence their fitness effects on other group members; however, Gokhale and Traulsen (2010) described a model where different orderings resulted in different payoffs. This may be particularly relevant to our leader-driven group-formation model; we assumed the leader plays the same role in the collective action as any other group member, but real-life leadership typically involves additional costs and benefits to the role (e.g., Lévi-Strauss, 1945). Third, we have assumed the same level of genetic homophily for all individuals regardless of their strategies (cf. Bergstrom, 2013). This precludes investigation of the evolution of assortativity (e.g., Newton, 2017a) and may be relevant when a shortage of viable group members can prevent the game from occurring. Indeed, we assumed an infinite population where there are always recruitment candidates available to form a group; however, in finite populations, there may be a trade-off between restricting membership to kin and the size of the group that can be recruited towards a collective task (Avilés et al., 2004; Mehdiabadi et al., 2006). Finally, we wish to draw the reader's attention to Allen et al. (2024), who similarly considered nonlinear interactions with assortment but in a finite population, and derived an expression for 'collective inclusive fitness' that centred on collectives as actors instead of individuals.

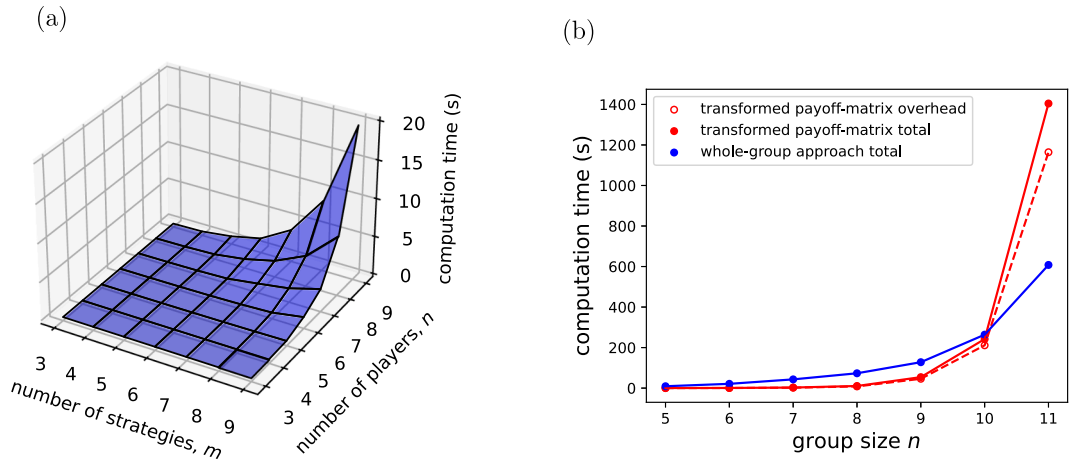


Fig. 6. How computation time varies with the parameters of the problem and method used. (a) The computational time required to calculate $\dot{p}_x$ as a function of group size and number of strategies using the whole-group approach (Eq. (29)). (b) Comparing the computational time required to identify the evolutionary steady states using the whole-group approach versus the transformed payoff-matrix approach (SI D) using the example from Fig. 4 a with homophily $h = 0.15$ and $\tau = \text{floor}(5n/8)$.

We offer some general advice for those wishing to use our methods. If the key attributes of the scenario can be captured with just a few players, then it may be possible to obtain analytic insights into the dynamics under homophily by analysing the transformed payoff matrix (see, for example, Peña et al. (2014) and Peña et al. (2015) for the analysis of 2-strategy games). For example, we began our analysis of Coordinated Cooperation by analysing 3-player games (SI G), and our code, which uses the SymPy symbolic mathematics library, may assist readers with their own analysis. As the number of players and strategies increases, a combinatorial explosion in the number of terms in Eq. (29) necessarily occurs. Nevertheless, for modestly sized groups and numbers of strategies, it still practical to obtain numerical solutions (Fig. 6a). Whether one should use the other-member or whole-group accounting depends on the particulars of the problem (e.g., Fig. 6b). The whole-group accounting (i.e., Eq. (29)) is more efficient when group size is large, but in situations where $\Delta \boldsymbol{p}$ must be evaluated many times, e.g., when numerically solving for the steady states, the overhead cost of evaluating the transformed payoff matrix may be worth the trade-off (Fig. 6b). Python code to perform all calculations has been uploaded to the Github collaborative platform where tutorials can also be found (SI L).

5.2. Coordinated cooperation example

We used our new methods to investigate how Coordinated Cooperation might evolve in a public goods game with economies of scale (sigmoid payoff). Coordination is significant both because it relies on distinctively human faculties and because it can allow cooperation to persist between nonkin in situations where unconditional cooperation cannot. We considered a scenario where genetic homophily was higher in the past and declined over time, which reflects the general trend that has occurred over the course of the human lineage (Gamble et al., 2011; Marwick, 2003; David-Barrett, 2020; Ringbauer et al., 2021). We found that high ancestral homophily facilitated the initial invasion of Coordinated Cooperation; however, in order for it to persist, either some level of homophily must be maintained, or contributing at the coordination point must be in players' self-interests (Fig. 4). We discuss each of these findings in turn.

Our first finding, that kin selection facilitated the invasion of Coordinated Cooperation, means that ancestral homophily can provide a stepping stone for the evolution of cooperation between nonkin (cf., Kristensen et al., 2022). In our model, the resource costs of cognition and communication (ϵ) prevented Coordinated Cooperation from evolving between nonkin (cf., Newton, 2017b). However, if Coordinated Cooperation invades first by kin selection, then it may persist even if homophily later declines (see also Dos Santos and West, 2018). Ancestral kin selection has been proposed to explain how other cooperative behaviours first evolved, including: conditional cooperation in iterated games (Takezawa and Price, 2010; Schonmann and Boyd, 2016), reciprocity (Axelrod and Hamilton, 1981; Carter, 2021), punishment (Boyd et al., 2010, 2014), and unconditional cooperation in threshold games (Bach et al., 2006; Kristensen et al., 2022) (but see: Lehmann et al., 2007; Martin and Lessard, 2023, 2024). In these examples, past kin selection facilitated the initial invasion of the strategy, which was then able to persist even as interactions shifted from kin to strangers. This dynamical role for kin selection can be contrasted with the cooptation and preadaptation hypotheses, where kin psychology provides a framework for coalitional structure (Smith, 2003; Read, 2010; Moffett, 2013) and a substrate to build more inclusive group-membership rules to solve more complex coordination problems (Alvard, 2009). Cooptation hypotheses are not mutually exclusive with dynamical facilitation, and all mechanisms likely played a role.

Our second finding was that, once Coordinated Cooperation has invaded by kin selection, in order to persist as homophily declines, it must either be in coordinators' self-interests to follow through on the agreement (e.g., Fig. 4a) or some degree of homophily must be maintained in the population (e.g., Fig. 4b). Self-interest is satisfied when costs are low and benefits are sufficiently nonlinear. Specifically, when there is a positive switching gain from non-contribution to contribution at the coordination point (Eq. (32); Eq. (34)), then Coordinated Cooperation can persist in coexistence with Defection and the population can resist invasion by Liars even if homophily declines to zero. However, if the switching-gain condition is not satisfied, then in order for the population to resist invasion by Liars, some degree of homophily must be maintained. Homophily helps by increasing the probability that individuals pursuing the same strategy are in the same group, including the probability that Liars will be grouped with Liars. The degree of homophily required to maintain Coordinated Cooperation is lower than that needed for Coordinating Cooperators to invade a population of Defectors in the first place (e.g., in Fig. 4b, as homophily declines, the uninvadable D-state begins before uninvadable C + D state ends). In general, coordination is known to interact synergistically with relatedness (Jones, 2000).

We have made the simplifying assumption that Coordinating Cooperators choose the number of contributors that matches the number that is most conducive to the maintenance of cooperation (i.e., τ); however, the problem of choosing the coordination point can be non-trivial. For example, in an experimental common-pool resource game, Ostrom et al. (1992) observed groups who were either unable to identify the group optimum or devised ambiguous or unnecessarily complicated rules that hampered their ability to achieve the social optimum. However, we also note that 'knowledge' of the best coordination point does not have to be

mechanistic (Henrich, 2021; O'Madagain and Tomasello, 2022), and future work may involve modelling the coordination point as a cultural trait.

One alternative explanation for the emergence of Coordinated Cooperation is that populations are sufficiently small for mutant invaders to overcome the separatrix through chance. The separatrix on the D-C axis (i.e., the unstable steady-state (red dot) in Fig. 5) puts a small subpopulation of Coordinating Cooperators at a selective disadvantage. Because our model assumes a very large (infinite) population with deterministic dynamics, such a subpopulation is inevitably driven to extinction. However, in finite populations, stochastic forces come into play, potentially driving the frequency of a selectively disadvantaged subpopulation past the separatrix through drift. For example, Nowak (Sec. 7.4, 2006) demonstrated this principle with tit-for-tat (TFT) strategies, which cannot invade a population of all-defectors in infinite populations because TFT receives the sucker's payoff in the first round (Axelrod and Hamilton, 1981). Yet in finite populations, TFT can both invade and reach fixation, and is in fact favoured by selection because its fixation probability exceeds that of defectors invading an all-TFT population. In a similar way, Coordinating Cooperators may invade a population of all Defectors, particularly when the cognitive cost is small, which positions the separatrix close to the all-D state. Moreover, they may stochastically increase beyond the deterministic coexistence frequencies (i.e., blue dot on D-C axis in Fig. 5) and temporarily achieve fixation.

Another alternative explanation for the emergence of Coordinated Cooperation is that the probability that players explore alternative strategies is sufficiently high to overcome the separatrix. For example, Imhof et al. (2005) investigated the deterministic dynamics in an infinite population of TFT, always defect (ALLD), and always cooperate (ALLC) players. Importantly, their dynamics were governed by a 'replicator-mutator equation', which explicitly includes the mutation rate rather than modelling mutations as very rare events like our model. They found that, when the separatrix between ALLD and TFT was sufficiently close to ALLD, and when the mutation rate was sufficiently high, ALLD loses stability, and TFT (and ALLC) can invade (see also greyed areas in Fig. 1 of Traulsen et al. (2009)). When strategies are genetically encoded, it is less justifiable to model a mutation process with high rates that causes large changes in strategy phenotype. However, such assumptions may be appropriate in social learning or cultural evolution models, where 'mutation' represents players exploring alternative strategies (Traulsen et al., 2009).

Our model is consistent with some experimental results but not others. Behavioural experiments find that cheap talk enhances cooperation in threshold PGGs (Van de Kragt et al., 1983; Mak et al., 2015; Palfrey et al., 2017), and that participants will believe others' announcements about whether or not they intend to contribute (Palfrey and Rosenthal, 1991). Our model explains that these commitments are believable because they are in the commitment-makers' self-interest and therefore may persist evolutionarily among strangers. Such self-enforcing commitments hold practical importance, e.g., for designing effective climate-change agreements (Barrett, 2016). Understanding the origins of self-enforcing commitments also bridges the gap between cooperative and non-cooperative game theory (Newton, 2017b).

A more challenging puzzle, however, is how to explain commitments among non-kin that are not self-enforcing. This raises questions about the commitment norm, i.e., that one should do what one has promised (Kerr and Kaufman-Gilliland, 1994), and why some participants follow through on their commitments even if doing so goes against their self-interest (Balliet, 2010). Commitment behaviour develops early in humans (Kachel et al., 2018; Kachel and Tomasello, 2019; Chalik and Rhodes, 2020), and joint commitment is one of the key attributes that distinguishes human cooperation from that of other apes (Genty et al., 2020; Tomasello, 2020). For example, in a collaborative situation where one individual receives their reward early, 3.5-year-old children will continue contributing until their partner also receives their reward (Hamann et al., 2012), whereas chimpanzees do not distinguish between continuing to help in an existing collaboration versus starting a new one (Greenberg et al., 2010). It has been hypothesised that joint commitment evolved in concert with shared intentionality in the context of interdependent collaboration (Tomasello, 2020); however, our model shows that interdependence is not enough for joint commitments to persist. If shared intentionality did indeed evolve in this context, then commitment beyond self-interest must be interpreted as a 'mistake' similar to altruistic behaviour in general, i.e., either as the result of ambiguous cues and an adaptive bias towards less costly errors (Haselton et al., 2015) or as a maladaptive response to a novel social environment (Burnham and Johnson, 2005; Hagen and Hammerstein, 2006; El Mouden et al., 2012). Alternatively, such commitments may persist through an additional mechanism, such as indirect reciprocity, where players selectively cooperate based on their partner's reputation from previous interactions (reviewed in Okada, 2020). This amounts to a strategy-based rather than familial homophily. Then such commitments may not be maladaptive unless they involve a 'mistake' regarding the true reputational costs of reneging. Understanding how cooperation under reputation evolves requires investigating reputational rules: how players in 'good' or 'bad' standing should treat others based on their respective standings (Ohtsuki and Iwasa, 2006; Fujimoto and Ohtsuki, 2023). The model of Krellner and Han (2025) is particularly relevant here, as it examines scenarios where players can make a public commitment to cooperate before playing a Prisoner's Dilemma, with standing determined by commitment adherence alone. They demonstrated that cooperation can be sustained when defection is judged as 'bad' but only when a commitment was broken, which aligns with the typical formulation of commitment norms (Kerr and Kaufman-Gilliland, 1994), the additional clout promises gain from being made explicit (Kachel and Tomasello, 2019), and the normative judgements made against those who renege (e.g., tattling, Kachel et al., 2018).

As a conclusion of this coordination example part, our analysis reveals how the distinctively human capacity for coordinated cooperation may have threaded a narrow evolutionary pathway: first emerging in ancestral environments rich with kinship ties, then persisting through a combination of residual levels of homophily and mutual benefits. This model can help resolve apparent paradoxes in human social behaviour, where self-interest and sophisticated cooperation coexist. It also suggests that successful coordination in contemporary settings may depend less on appealing to altruism or moral obligations, and more on creating conditions where coordination agreements naturally align with individual interests.

CRediT authorship contribution statement

Nadiah P. Kristensen: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Funding acquisition, Formal analysis; **Ryan A. Chisholm:** Writing – review & editing, Supervision, Project administration; **Hisashi Ohtsuki:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Formal analysis, Conceptualization, Funding acquisition, Formal analysis, Conceptualization

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Ryan A. Chisholm reports financial support was provided by Government of Singapore Ministry of Education. Nadiah P. Kristensen reports financial support was provided by Government of Singapore Ministry of Education. Hisashi Ohtsuki reports financial support was provided by Japan Society for the Promotion of Science. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

NPK and RAC were funded by the Ministry of Education, Singapore, under a Tier 1 Grant (WBS A-0004766-00-00). HO acknowledges the support by JSPS KAKENHI Grant Number JP20K06812.

Supplementary material

Supplementary material associated with this article can be found in the online version at 10.1016/j.jtbi.2025.112089.

References

Alger, I., Weibull, J.W., 2013. Homo moralis—preference evolution under incomplete information and assortative matching. Econometrica 81 (6), 2269–2302.
Alger, I., Weibull, J.W., 2014. Evolutionarily stable strategies, preferences and moral values, in n-player interactions. LERNA Working Paper.
Alger, I., Weibull, J.W., 2019. Evolutionary models of preference formation. Annu. Rev. Econ. 11, 329–354.
Allen, B., Khwaja, A.-R., Donahue, J.L., Kelly, T.J., Hyacinthe, S.R., Proulx, J., Lattanzio, C., Dementieva, Y.A., Sample, C., 2024. Nonlinear social evolution and the emergence of collective action. PNAS Nexus 3 (4), 131.
Allen, B., Nowak, M.A., 2016. There is no inclusive fitness at the level of the individual. Curr. Opin. Behav. Sci. 12, 122–128.
Allen-Arave, W., Gurven, M., Hill, K., 2008. Reciprocal altruism, rather than kin selection, maintains nepotistic food transfers on an ache reservation. Evol. Hum. Behav. 29 (5), 305–318.
Alvard, M., 2009. Kinship and cooperation. Hum. Nat. 20 (4), 394–416.
Alvard, M.S., Nolin, D.A., 2002. Rousseau's whale hunt? Coordination among big-game hunters. Curr. Anthropol. 43 (4), 533–559.
Archetti, M., 2018. How to analyze models of nonlinear public goods. Games 9 (2), 17.
Archetti, M., Scheuring, I., 2011. Coexistence of cooperation and defection in public goods games. Evolution 65 (4), 1140–1148.
Archetti, M., Scheuring, I., 2012. Game theory of public goods in one-shot social dilemmas without assortment. J. Theor. Biol. 299, 9–20.
Archetti, M., Scheuring, I., Yu, D., 2020. The non-tragedy of the non-linear commons. Preprints 2020040226.
Avilés, L., Fletcher, J.A., Cutter, A.D., 2004. The kin composition of social groups: trading group size for degree of altruism. Am. Nat. 164 (2), 132–144.
Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. Science 211 (4489), 1390–1396.
Bach, L.A., Helvik, T., Christiansen, F.B., 2006. The evolution of n-player cooperation-threshold games and ESS bifurcations. J. Theor. Biol. 238 (2), 426–434.
Balliet, D., 2010. Communication and cooperation in social dilemmas: a meta-analytic review. J. Confl. Resolut. 54 (1), 39–57.
Barrett, S., 2016. Coordination vs. voluntarism and enforcement in sustaining international environmental cooperation. Proc. Natl. Acad. Sci. 113 (51), 14515–14522.
Bergstrom, T.C., 2003. The algebra of assortative encounters and the evolution of cooperation. Int. Game Theory Rev. 5 (03), 211–228.
Bergstrom, T.C., 2013. Measures of assortativity. Biol. Theory 8 (2), 133–141.
Bickerton, D., Szathmáry, E., 2011. Confrontational scavenging as a possible source for language and cooperation. BMC Evol. Biol. 11 (1), 1–7.
Bilancini, E., Boncinelli, L., Wu, J., 2018. The interplay of cultural intolerance and action-assortativity for the emergence of cooperation and homophily. Eur. Econ. Rev. 102, 1–18.
Bowles, S., Gintis, H., 1998. The moral economy of communities: structured populations and the evolution of pro-social norms. Evol. Hum. Behav. 19 (1), 3–25.
Boyd, R., Gintis, H., Bowles, S., 2010. Coordinated punishment of defectors sustains cooperation and can proliferate when rare. Science 328 (5978), 617–620.
Boyd, R., Schonmann, R.H., Vicente, R., 2014. Hunter–gatherer population structure and the evolution of contingent cooperation. Evol. Hum. Behav. 35 (3), 219–227.
Boza, G., Számadó, S., 2010. Beneficial laggards: multilevel selection, cooperative polymorphism and division of labour in threshold public good games. BMC Evol. Biol. 10 (1), 1–12.
Brooks, A.S., Yellen, J.E., Potts, R., Behrensmeyer, A.K., Deino, A.L., Leslie, D.E., Ambrose, S.H., Ferguson, J.R., d'Errico, F., Zipkin, A.M., Whittaker, S., Post, J., Veatch, E.G., Foecke, K., Clark, J.B., 2018. Long-distance stone transport and pigment use in the earliest middle stone age. Science 360 (6384), 90–94.
Burnham, T.C., Johnson, D. D.P., 2005. The biological and evolutionary logic of human cooperation. Anal. Krit. 27 (1), 113–135.
Burnstein, E., 2015. Altruism and genetic relatedness. In: Buss, D.M. (Ed.), The handbook of evolutionary psychology. Wiley Online Library, pp. 528–551.
Cadsby, C.B., Maynes, E., 1999. Voluntary provision of threshold public goods with continuous contributions: experimental evidence. J. Public Econ. 71 (1), 53–73.
Carter, G.G., 2021. Co-option and the evolution of food sharing in vampire bats. Ethology 127 (10), 837–849.
Chai, C., Francis, E., Xiao, T., 2021. Supply chain dynamics with assortative matching. J. Evolut. Econ. 31 (1), 179–206.
Chalik, L., Rhodes, M., 2020. Groups as moral boundaries: a developmental perspective. In: Benson, J. (Ed.), Advances in Child Development and Behavior. Elsevier. Vol. 58, pp. 63–93.
Coder Gylling, K., Brännström, Å., 2018. Effects of relatedness on the evolution of cooperation in nonlinear public goods games. Games 9 (4), 87.
Cornforth, D.M., Sumpter, D. J.T., Brown, S.P., Brännström, Å., 2012. Synergy and group size in microbial cooperation. Am. Nat. 180 (3), 296–305.
David-Barrett, T., 2020. Herding friends in similarity-based architecture of social networks. Sci. Rep. 10 (1), 1–6.
Dos Santos, M., West, S.A., 2018. The coevolution of cooperation and cognition in humans. Proc. R. Soc. B 285 (1879), 20180723.
El Mouden, C., Burton-Chellew, M., Gardner, A., West, S.A., 2012. What do humans maximize? In: Okashi, S., Binmore, K.G. (Eds.), Evolution and Rationality: Decisions, Cooperation and Strategic Behaviour. Cambridge University Press, Cambridge, pp. 23–49. https://doi.org/10.1017/CBO9780511792601.003
Elster, J., 1988. Taming chance: randomization in individual and social decisions. Tanner Lect. Hum. Values 9, 107–179.
Ewens, W.J., 1972. The sampling theory of selectively neutral alleles. Theor. Popul. Biol. 3 (1), 87–112.
Fujimoto, Y., Ohtsuki, H., 2023. Evolutionary stability of cooperation in indirect reciprocity under noisy and private assessment. Proc. Natl. Acad. Sci. 120 (20), e2300544120.
Gamble, C., Gowlett, J., Dunbar, R., 2011. The social brain and the shape of the palaeolithic. Camb. Archaeol. J. 21 (1), 115–136.
García, J., van Veelen, M., Traulsen, A., 2014. Evil green beards: tag recognition can also be used to withhold cooperation in structured populations. J. Theor. Biol. 360, 181–186.
Garcia, T., De Monte, S., 2013. Group formation and the evolution of sociality. Evolution 67 (1), 131–141.
Gardner, A., West, S.A., Wild, G., 2011. The genetical theory of kin selection. J. Evol. Biol. 24 (5), 1020–1043.
Genty, E., Heesen, R., Guéry, J.-P., Rossano, F., Zuberbühler, K., Bangerter, A., 2020. How apes get into and out of joint actions: shared intentionality as an interactional achievement. Interact. Stud. 21 (3), 353–386.
Godfrey-Smith, P., Kerr, B., 2009. Selection in ephemeral networks. Am. Nat. 174 (6), 906–911.
Gokhale, C.S., Traulsen, A., 2010. Evolutionary games in the multiverse. Proc. Natl. Acad. Sci. 107 (12), 5500–5504.
Grafen, A., 1979. The hawk-dove game played between relatives. Anim. Behav. 27, 905–907.
Greenberg, J.R., Hamann, K., Warneken, F., Tomasello, M., 2010. Chimpanzee helping in collaborative and noncollaborative contexts. Anim. Behav. 80 (5), 873–880.
Gurven, M., Hill, K., Kaplan, H., Hurtado, A., Lyles, R., 2000. Food transfers among Hiwi foragers of Venezuela: tests of reciprocity. Hum. Ecol. 28 (2), 171–218.
Hagen, E.H., Hammerstein, P., 2006. Game theory and human evolution: a critique of some recent interpretations of experimental games. Theor. Popul. Biol. 69 (3), 339–348.
Hamann, K., Warneken, F., Tomasello, M., 2012. Children's developing commitments to joint goals. Child Dev. 83 (1), 137–145.
Hames, R., McCabe, C., 2007. Meal sharing among the Ye'kwana. Hum. Nat. 18 (1), 1–21.
Hamilton, W.D., 1964. The genetical theory of social behaviour. I, II. J. Theor. Biol. 7, 1–52.
Haselton, M.G., Nettle, D., Murray, D.R., 2015. The evolution of cognitive bias. In: Buss, D.M. (Ed.), The Handbook of Evolutionary Psychology. Wiley Online Library, Hoboken, NJ, pp. 1–20. https://onlinelibrary.wiley.com/doi/full/10.1002/9781119125563.evpsych241.
Hauert, C., Michor, F., Nowak, M.A., Doebeli, M., 2006. Synergy and discounting of cooperation in social dilemmas. J. Theor. Biol. 239 (2), 195–202.
Henrich, J., 2021. Cultural evolution: is causal inference the secret of our success? Curr. Biol. 31 (8), R381–R383.
Hofbauer, J., Sigmund, K., 1998. Evolutionary Games and Population Dynamics. Cambridge University Press, Cambridge.
Imhof, L.A., Fudenberg, D., Nowak, M.A., 2005. Evolutionary cycles of cooperation and defection. Proc. Natl. Acad. Sci. 102 (31), 10797–10800.
Jaeggi, A.V., Gurven, M., 2013. Reciprocity explains food sharing in humans and other primates independent of kin selection and tolerated scrounging: a phylogenetic meta-analysis. Proc. R. Soc. B 280 (1768), 20131615.
Jensen, M.K., Rigos, A., 2018. Evolutionary games and matching rules. Int. J. Game Theory 47 (3), 707–735.
Jones, D., 2000. Group nepotism and human kinship. Curr. Anthropol. 41 (5), 779–809.
Kachel, U., Svetlova, M., Tomasello, M., 2018. Three-year-olds' reactions to a partner's failure to perform her role in a joint commitment. Child Dev. 89 (5), 1691–1703.
Kachel, U., Tomasello, M., 2019. 3-and 5-year-old children's adherence to explicit and implicit joint commitments. Dev. Psychol. 55 (1), 80.
Kerr, N.L., Kaufman-Gilliland, C.M., 1994. Communication, commitment, and cooperation in social dilemma. J. Pers. Soc. Psychol. 66 (3), 513.
Koster, J.M., Leckie, G., 2014. Food sharing networks in lowland Nicaragua: an application of the social relations model to count data. Soc. Netw. 38, 100–110.
Van de Kragt, A. J.C., Orbell, J.M., Dawes, R.M., 1983. The minimal contributing set as a solution to public goods problems. Am. Polit. Sci. Rev. 77 (1), 112–122.
Krellner, M., Han, T.A., 2025. Words are not wind–how joint commitment and reputation solve social dilemmas, without repeated interactions or enforcement by third parties. arXiv:2307.06898.
Kristensen, N.P., Ohtsuki, H., Chisholm, R.A., 2022. Ancestral social environments plus nonlinear benefits can explain cooperation in human societies. Sci. Rep. 12, 20252.
Kuhn, S.L., Stiner, M.C., Reese, D.S., Güleç, E., 2001. Ornaments of the earliest upper paleolithic: new insights from the levant. Proc. Natl. Acad. Sci. 98 (13), 7641–7646.
Ledyard, J.O., 1995. Public goods: a survey of experimental research. In: Kagel, J., Roth, A. (Eds.), Handbook of Experimental Economics. Princeton University Press, Princeton, pp. 111–193.

Lehmann, L., Rousset, F., 2010. How life history and demography promote or inhibit the evolution of helping behaviours. Philos. Trans. R. Soc. B 365 (1553), 2599–2617.
Lehmann, L., Rousset, F., Roze, D., Keller, L., 2007. Strong reciprocity or strong ferocity? A population genetic view of the evolution of altruistic punishment. Am. Nat. 170 (1), 21–36.
Lévi-Strauss, C., 1945. The social and psychological aspects of chieftainship in a primitive tribe: the Nambikuara of Northwestern Mato Grosso. Trans. N. Y. Acad. Sci. 2 (7/1), 16–32.
Mak, V., Zwick, R., Rao, A.R., Pattaratanakun, J.A., 2015. "Pay what you want" as threshold public good provision. Organ. Behav. Hum. Decis. Process. 127, 30–43.
Martin, É., Lessard, S., 2023. Assortment by group founders always promotes the evolution of cooperation under global selection but can oppose it under local selection. Dyn. Games Appl., 13, 1194–1218.
Martin, É., Lessard, S., 2024. Evolution of cooperation in social dilemmas with assortment in finite populations. J. Theor. Biol. 592, 111891.
Marwick, B., 2003. Pleistocene exchange networks as evidence for the evolution of language. Camb. Archaeol. J. 13 (1), 67–81.
Mehdiabadi, N.J., Jack, C.N., Farnham, T.T., Platt, T.G., Kalla, S.E., Shaulsky, G., Queller, D.C., Strassmann, J.E., 2006. Kin preference in a social microbe. Nature 442 (7105), 881–882.
Moffett, M.W., 2013. Human identity and the evolution of societies. Hum. Nat. 24 (3), 219–267.
Næss, M.W., 2019. From hunter-gatherers to nomadic pastoralists: forager bands do not tell the whole story of the evolution of human cooperation. SocArXiv. https://doi.org/10.31235/osf.io/9c8bm
Næss, M.W., 2021. Collaborative foundations of herding: the formation of cooperative groups among Tibetan pastoralists. J. Arid Environ. 186, 104407.
Næss, M.W., Bårdsen, B.-J., Fauchald, P., Tveraa, T., 2010. Cooperative pastoral production—the importance of kinship. Evol. Hum. Behav. 31 (4), 246–258.
Nax, H.H., Rigos, A., 2016. Assortativity evolving from social dilemmas. J. Theor. Biol. 395, 194–203.
Newton, J., 2017a. The preferences of homo moralis are unstable under evolving assortativity. Int. J. Game Theory 46 (2), 583–589.
Newton, J., 2017b. Shared intentions: the evolution of collaboration. Games Econ. Behav. 104, 517–534.
Nobre, C.A., De Simone, B.L., 2009. 'Tipping points' for the Amazon forest. Curr. Opin. Environ. Sustain. 1 (1), 28–36.
Nolin, D.A., 2010. Food-sharing networks in Lamalera, Indonesia. Hum. Nat. 21 (3), 243–268.
Nowak, M.A., 2006. Evolutionary Dynamics: Exploring the Equations of Life. Harvard University Press, Cambridge, MA.
Ohtsuki, H., 2010. Evolutionary games in Wright's island model: kin selection meets evolutionary game theory. Evolution 64 (12), 3344–3353.
Ohtsuki, H., 2014. Evolutionary dynamics of n-player games played by relatives. Philos. Trans. R. Soc. B 369 (1642), 20130359.
Ohtsuki, H., Iwasa, Y., 2006. The leading eight: social norms that can maintain cooperation by indirect reciprocity. J. Theor. Biol. 239 (4), 435–444.
Ohtsuki, H., Nowak, M.A., 2006. The replicator equation on graphs. J. Theor. Biol. 243 (1), 86–97.
Okada, I., 2020. A review of theoretical studies on indirect reciprocity. Games 11 (3), 27.
O'Madagain, C., Tomasello, M., 2022. Shared intentionality, reason-giving and the evolution of human culture. Philos. Trans. R. Soc. B 377 (1843), 20200320.
Ostrom, E., Walker, J., Gardner, R., 1992. Covenants with and without a sword: self-governance is possible. Am. Polit. Sci. Rev. 86 (2), 404–417.
Palfrey, T., Rosenthal, H., Roy, N., 2017. How cheap talk enhances efficiency in threshold public goods games. Games Econ. Behav. 101, 234–259.
Palfrey, T.R., Rosenthal, H., 1984. Participation and the provision of discrete public goods: a strategic analysis. J. Public Econ. 24 (2), 171–193.
Palfrey, T.R., Rosenthal, H., 1991. Testing for effects of cheap talk in a public goods game with private information. Games Econ. Behav. 3 (2), 183–220.
Peña, J., Lehmann, L., Nöldeke, G., 2014. Gains from switching and evolutionary stability in multi-player matrix games. J. Theor. Biol. 346, 23–33.
Peña, J., Nöldeke, G., Lehmann, L., 2015. Evolutionary dynamics of collective action in spatially structured populations. J. Theor. Biol. 382, 122–136.
Price, G.R., 1970. Selection and covariance. Nature 227 (5257), 520–521.
Queller, D.C., 1985. Kinship, reciprocity and synergism in the evolution of social behaviour. Nature 318 (6044), 366–367.
Raihani, N.J., Bshary, R., 2011. The evolution of punishment in n-player public goods games: a volunteer's dilemma. Evolution 65 (10), 2725–2728.
Raihani, N.J., Bshary, R., 2015. Why humans might help strangers. Front. Behav. Neurosci. 9, 39. https://doi.org/10.3389/fnbeh.2015.00039
Read, D., 2010. From experiential-based to relational-based forms of social organization: a major transition in the evolution of *Homo sapiens*. Proc. Br. Acad. 158, 199–229.
Ringbauer, H., Novembre, J., Steinrücken, M., 2021. Parental relatedness through time revealed by runs of homozygosity in ancient DNA. Nat. Commun. 12 (1), 1–11.
Roberts, G., 2013. When punishment pays. PLoS One 8 (3), e57378.
Rousset, F., 2013. Genetic Structure and Selection in Subdivided Populations. Vol. 40. Princeton University Press, Princeton.
Schonmann, R.H., Boyd, R., 2016. A simple rule for the evolution of contingent cooperation in large groups. Philos. Trans. R. Soc. B 371 (1687), 20150099.
Schweinfurth, M.K., Call, J., 2019. Revisiting the possibility of reciprocal help in non-human primates. Neurosci. Biobehav. Rev. 104, 73–86.
Sehasseh, E.M., Fernandez, P., Kuhn, S., Stiner, M., Mentzer, S., Colarossi, D., Clark, A., Lanoe, F., Pailes, M., Hoffmann, D., et al., 2021. Early middle stone age personal ornaments from Bizmoune Cave, Essaouira, Morocco. Sci. Adv. 7 (39), eabi8620.
Sigmund, K., De Silva, H., Traulsen, A., Hauert, C., 2010. Social learning promotes institutions for governing the commons. Nature 466 (7308), 861.
Smith, E.A., 2003. Human cooperation: Perspectives from behavioral ecology. In: Hammerstein, P. (Ed.), Genetic and Cultural Evolution of Cooperation. MIT Press, Cambridge, Massachusetts, pp. 401–427.
Takezawa, M., Price, M.E., 2010. Revisiting "the evolution of reciprocity in sizable groups" continuous reciprocity in the repeated n-person prisoner's dilemma. J. Theor. Biol. 264 (2), 188–196.
Taylor, C., Nowak, M.A., 2007. Transforming the dilemma. Evolution 61 (10), 2281–2292.
Taylor, P., 2017. Inclusive fitness in finite populations—effects of heterogeneity and synergy. Evolution 71 (3), 508–525.
Taylor, P., Maciejewski, W., 2012. An inclusive fitness analysis of synergistic interactions in structured populations. Proc. R. Soc. B 279 (1747), 4596–4603.
Taylor, P.D., 1992. Altruism in viscous populations—an inclusive fitness model. Evol. Ecol. 6 (4), 352–356.
Taylor, P.D., Frank, S.A., 1996. How to make a kin selection model. J. Theor. Biol. 180 (1), 27–37.
Tomasello, M., 2020. The moral psychology of obligation. Behav. Brain Sci. 43, e56.
Tomasello, M., Melis, A.P., Tennie, C., Wyman, E., Herrmann, E., 2012. Two key steps in the evolution of human cooperation: the interdependence hypothesis. Curr. Anthropol. 53 (6), 673–692.
Traulsen, A., Hauert, C., De Silva, H., Nowak, M.A., Sigmund, K., 2009. Exploration dynamics in evolutionary games. Proc. Natl. Acad. Sci. 106 (3), 709–712.
Van Cleve, J., 2015. Social evolution and genetic interactions in the short and long term. Theor. Popul. Biol. 103, 2–26.
Van Cleve, J., Akçay, E., 2014. Pathways to social evolution: reciprocity, relatedness, and synergy. Evolution 68 (8), 2245–2258.
Van Veelen, M., 2009. Group selection, kin selection, altruism and cooperation: when inclusive fitness is right and when it can be wrong. J. Theor. Biol. 259 (3), 589–600.
Van Veelen, M., 2011. The replicator dynamics with n players and population structure. J. Theor. Biol. 276 (1), 78–85.
Vásárhelyi, Z., Scheuring, I., 2013. Invasion of cooperators in lattice populations: linear and non-linear public good games. BioSystems 113 (2), 81–90.
West, S.A., Gardner, A., Shuker, D.M., Reynolds, T., Burton-Chellow, M., Sykes, E.M., Guinnee, M.A., Griffin, A.S., 2006. Cooperation and the scale of competition in humans. Curr. Biol. 16 (11), 1103–1106.
Wild, G., Traulsen, A., 2007. The different limits of weak selection and the evolutionary dynamics of finite populations. J. Theor. Biol. 247 (2), 382–390.
Wright, S., 1931. Evolution in Mendelian populations. Genetics 16 (2), 97.
Wu, J., 2019. Social connections and cultural heterogeneity. J. Evolut. Econ. 29 (2), 779–798.
Wu, J., et al., 2016. Evolving assortativity and social conventions. Econ. Bull. 36 (2), 936–941.

Many-strategy games in groups with relatives and the evolution of coordinated cooperation

Supplementary Information

Nadiah P. KRISTENSEN, Ryan A. CHISHOLM, Hisashi OHTSUKI

A Analytic results

A.1 Strategy indicator variables

We consider games with m possible strategies, $s_1, \ldots, s_m$, played in groups of size n. We index the individuals in the group $j = 0, \ldots, n-1$, and reserve index 0 for the focal individual.

Let $\boldsymbol{e}_x$ be an m-dimensional vector with a 1 in the x-th position and 0 in all others

$$\boldsymbol{e}_x = (0, \ldots, 0, \underbrace{1}_{x-\text{th}}, 0, \ldots, 0). \tag{A.1}$$

The strategy indicator for an individual j that pursues strategy s_x is

$$\boldsymbol{g}_j = \boldsymbol{e}_x. \tag{A.2}$$

The whole-group strategy composition is defined as the sum of individual strategy indicators

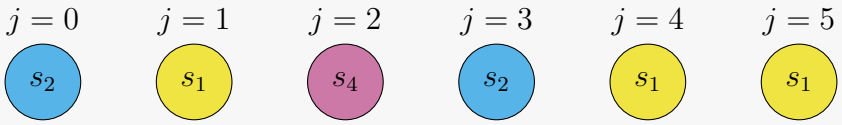
$$\boldsymbol{g}_{\text{a}} = \boldsymbol{g}_0 + \boldsymbol{g}_1 + \ldots + \boldsymbol{g}_{n-1}, \tag{A.3}$$

such that the i-th component counts the number of group members who pursue strategy s_i, and similarly the nonfocal strategy composition is defined as

$$\boldsymbol{g}_{\text{nf}} = \boldsymbol{g}_1 + \ldots + \boldsymbol{g}_{n-1}, \tag{A.4}$$

which counts the number of individuals in the group apart from the focal individual who pursue each strategy s_i.

Example 1. Consider the group illustrated below, which is playing a game with $m = 4$ possible strategies, where each coloured circle represents an individual j, and the colours indicate the individual's strategy, which is also written within the circle



The strategy indicators for each individual are

$$\begin{aligned} \boldsymbol{g}_1 = \boldsymbol{g}_4 = \boldsymbol{g}_5 &= (1,0,0,0), \\ \boldsymbol{g}_0 = \boldsymbol{g}_3 &= (0,1,0,0), \\ \boldsymbol{g}_2 &= (0,0,0,1). \end{aligned}$$

The whole-group strategy composition is

$$\boldsymbol{g}_{\text{a}} = (3,2,0,1),$$

and the nonfocal strategy composition is

$$\boldsymbol{g}_{\text{nf}} = (3,1,0,1). \qquad \triangleleft$$

A.2 Evolutionary dynamics from Price equation

Let p_x be the proportion of individuals in the population pursuing strategy s_x. From the Price equation (Price, 1970), the change in the proportion of s_x strategists in the population is given by

$$\overline{w}\,\Delta p_x = \mathbb{C}\mathrm{ov}[G_{0,x}, W_0], \tag{A.5}$$

where $\overline{w}$ is the average fitness in the population ($\overline{w} = 1$ when population size constant), $\mathbb{C}\mathrm{ov}$ is covariance, and $G_{0,x}$ is the x-th element of the focal individual's strategy indicator $\boldsymbol{G}_0$ and an indicator variable

$$G_{0,x} = \begin{cases} 1 & \text{if the focal individual pursues strategy } s_x, \\ 0 & \text{otherwise.} \end{cases} \tag{A.6}$$

The expected fitness W_0 of the focal individual is

$$W_0 = s + (1-s)\frac{1+\delta\Pi_0}{1+\delta\overline{\pi}}, \tag{A.7}$$

where Π_0 is the payoff of the focal individual and $\overline{\pi}$ is the average payoff in the population. The survival probability s, and selection strength δ are constants.

To obtain Eq. 7 in the main text, we substitute Eq. A.7 into Eq. A.5 and apply the covariance identity

$$\mathbb{C}\mathrm{ov}[X, aY + b] = a\mathbb{C}\mathrm{ov}[X, Y], \tag{A.8}$$

where a, b are constants. Specifically, we have

$$\begin{aligned} \Delta p_x &= \mathbb{C}\mathrm{ov}\left[G_{0,x}, s + (1-s)\frac{1+\delta\Pi_0}{1+\delta\overline{\pi}}\right], \\ &= \frac{\delta(1-s)}{1+\delta\overline{\pi}}\mathbb{C}\mathrm{ov}\left[G_{0,x}, \Pi_0\right], \end{aligned} \tag{A.9}$$

and therefore

$$\Delta p_x \propto \mathbb{C}\mathrm{ov}\left[G_{0,x}, \Pi_0\right]. \tag{A.10}$$

A.3 Accounting based on other-member composition

In the light of Eq. 3, the payoff to the focal individual from the game can be written in terms of the nonfocal strategy composition

$$\Pi_0 = \sum_{i=1}^{m} G_{0,i}\pi(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{nf}}). \tag{A.11}$$

Substituting Eq. A.11 into Eq. A.10 and applying the identity

$$\mathbb{C}\mathrm{ov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y], \tag{A.12}$$

we obtain

$$\begin{aligned}\Delta p_x &\propto \mathbb{C}\mathrm{ov}\left[G_{0,x},\ \sum_{i=1}^{m} G_{0,i}\pi(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{nf}})\right], \\ &= \mathbb{E}\left[\sum_{i=1}^{m} G_{0,x}G_{0,i}\pi(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{nf}})\right] - \mathbb{E}\left[G_{0,x}\right]\ \mathbb{E}\left[\sum_{i=1}^{m} G_{0,i}\pi(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{nf}})\right], \\ &= \mathbb{E}\left[G_{0,x}\pi(\boldsymbol{e}_x, \boldsymbol{G}_{\mathrm{nf}})\right] - p_x \sum_{i=1}^{m} \mathbb{E}\left[G_{0,i}\pi(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{nf}})\right], \end{aligned} \tag{A.13}$$

where we used $\mathbb{E}\left[G_{0,x}\right] = p_x$.

Let $\boldsymbol{g}_{\mathrm{nf}}$ be the value of $\boldsymbol{G}_{\mathrm{nf}}$. It can take values from the set

$$\mathcal{G}_{\mathrm{nf}} = \left\{ \boldsymbol{g}_{\mathrm{nf}} \in \mathbb{N}_{\geq 0}^{m} \,\middle|\, \sum_{i=1}^{m} g_{\mathrm{nf},i} = n - 1 \right\}. \tag{A.14}$$

The expectations in Eq. A.13 is thus

$$\begin{aligned}\mathbb{E}\left[G_{0,i}\pi(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{nf}})\right] &= \sum_{\boldsymbol{g}_{\mathrm{nf}} \in \mathcal{G}_{\mathrm{nf}}} \pi(\boldsymbol{e}_i, \boldsymbol{g}_{\mathrm{nf}}) \mathbb{P}[\boldsymbol{G}_0 = \boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{nf}} = \boldsymbol{g}_{\mathrm{nf}}], \\ &= \sum_{\boldsymbol{g}_{\mathrm{nf}} \in \mathcal{G}_{\mathrm{nf}}} \pi(\boldsymbol{e}_i, \boldsymbol{g}_{\mathrm{nf}}) \underbrace{\mathbb{P}[\boldsymbol{G}_0 = \boldsymbol{e}_i]}_{p_i} \mathbb{P}[\boldsymbol{G}_{\mathrm{nf}} = \boldsymbol{g}_{\mathrm{nf}} \mid \boldsymbol{G}_0 = \boldsymbol{e}_i], \\ &= p_i \underbrace{\sum_{\boldsymbol{g}_{\mathrm{nf}} \in \mathcal{G}_{\mathrm{nf}}} \pi(\boldsymbol{e}_i, \boldsymbol{g}_{\mathrm{nf}}) \mathbb{P}[\boldsymbol{G}_{\mathrm{nf}} = \boldsymbol{g}_{\mathrm{nf}} \mid \boldsymbol{G}_0 = \boldsymbol{e}_i]}_{\overline{\pi}_i}, \end{aligned} \tag{A.15}$$

where $\overline{\pi}_i$ is the expected payoff to s_i strategists.

Substituting the expectations (Eq. A.15) into Eq. A.13, we obtain Eq. 11 in the main text:

$$\Delta p_x \propto p_x \left(\overline{\pi}_x - \sum_{i=1}^{m} p_i \overline{\pi}_i \right) = p_x \left(\overline{\pi}_x - \overline{\pi} \right). \tag{A.16}$$

By taking a proper limit and a proper time scale, we obtain Eq. 12 in the main text:

$$\dot{p}_x = p_x \left(\overline{\pi}_x - \overline{\pi} \right). \tag{A.17}$$

A.4 Accounting based on whole-group composition

In the light of Eq. 4, the payoff to the focal individual from the game can be written in terms of the whole group's strategy composition

$$\Pi_0 = \sum_{i=1}^{m} G_{0,i}\hat{\pi}(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{a}}). \tag{A.18}$$

By substituting Eq. A.18 into Eq. A.10 and applying the identity in Eq. A.12, we obtain

$$\begin{aligned}\Delta p_x &\propto \mathbb{C}\mathrm{ov}\left[G_{0,x}, \sum_{i=1}^{m} G_{0,i}\hat{\pi}(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{a}})\right], \\ &= \mathbb{E}\left[G_{0,x}\hat{\pi}(\boldsymbol{e}_x, \boldsymbol{G}_{\mathrm{a}})\right] - p_x \sum_{i=1}^{m} \mathbb{E}\left[G_{0,i}\hat{\pi}(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{a}})\right]. \end{aligned} \tag{A.19}$$

Let $\boldsymbol{g}_{\mathrm{a}}$ be the value of $\boldsymbol{G}_{\mathrm{a}}$. It can take values from the set

$$\mathcal{G}_{\mathrm{a}} = \left\{ \boldsymbol{g}_{\mathrm{a}} \in \mathbb{N}^m_{\geq 0} \,\middle|\, \sum_{i=1}^{m} g_{\mathrm{a},i} = n \right\}. \tag{A.20}$$

The expectations in Eq. A.19 is thus

$$\begin{aligned}\mathbb{E}\left[G_{0,i}\hat{\pi}(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{a}})\right] &= \sum_{\boldsymbol{g}_{\mathrm{a}} \in \mathcal{G}_{\mathrm{a}}} \hat{\pi}(\boldsymbol{e}_i, \boldsymbol{g}_{\mathrm{a}}) \mathbb{P}[\boldsymbol{G}_0 = \boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{a}} = \boldsymbol{g}_{\mathrm{a}}], \\ &= \sum_{\boldsymbol{g}_{\mathrm{a}} \in \mathcal{G}_{\mathrm{a}}} \hat{\pi}(\boldsymbol{e}_i, \boldsymbol{g}_{\mathrm{a}}) \underbrace{\mathbb{P}[\boldsymbol{G}_0 = \boldsymbol{e}_i \mid \boldsymbol{G}_{\mathrm{a}} = \boldsymbol{g}_{\mathrm{a}}]}_{\frac{g_{\mathrm{a},i}}{n}} \mathbb{P}[\boldsymbol{G}_{\mathrm{a}} = \boldsymbol{g}_{\mathrm{a}}], \\ &= \mathbb{E}\left[\frac{G_{\mathrm{a},i}}{n}\hat{\pi}(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{a}})\right]. \end{aligned} \tag{A.21}$$

Substituting the expectations (Eq. A.21) into Eq. A.19, we obtain

$$\begin{aligned}\Delta p_x &\propto \mathbb{E}\left[\frac{G_{\mathrm{a},x}}{n}\hat{\pi}(\boldsymbol{e}_x, \boldsymbol{G}_{\mathrm{a}})\right] - p_x \sum_{i=1}^{m} \mathbb{E}\left[\frac{G_{\mathrm{a},i}}{n}\hat{\pi}(\boldsymbol{e}_i, \boldsymbol{G}_{\mathrm{a}})\right] \\ &= \sum_{\boldsymbol{g}_{\mathrm{a}} \in \mathcal{G}_{\mathrm{a}}} \left(\frac{g_{\mathrm{a},x}}{n}\hat{\pi}(\boldsymbol{e}_x, \boldsymbol{g}_{\mathrm{a}}) - p_x \sum_{i=1}^{m} \frac{g_{\mathrm{a},i}}{n}\hat{\pi}(\boldsymbol{e}_i, \boldsymbol{g}_{\mathrm{a}}) \right) \mathbb{P}[\boldsymbol{G}_{\mathrm{a}} = \boldsymbol{g}_{\mathrm{a}}]. \end{aligned} \tag{A.22}$$

The probability $\mathbb{P}[\boldsymbol{G}_{\mathrm{a}} = \boldsymbol{g}_{\mathrm{a}}]$ depends on the family partition structure of the group, and we have the probability distribution of family partition structures from the homophilic group-formation model. However, the family structures that are possible for a given strategy composition $\boldsymbol{g}_{\mathrm{a}}$ are constrained: individuals with the same strategy can be from different families, but individuals with different strategies cannot be from the same family. Therefore, to identify the family partition structures consistent with $\boldsymbol{g}_{\mathrm{a}}$, we identify the possible partition structures for each strategy that are consistent with in $\boldsymbol{g}_{\mathrm{a}}$.

We start with the family-size distribution of s_i strategists in a group. Let

$$\boldsymbol{z}_i = (z_{i,1}, z_{i,2}, \ldots, z_{i,n}). \tag{A.23}$$

represents the the family-size distribution of s_i strategists, where its k-th component $z_{i,k}$ represents the number of size-k families among s_i strategists. The whole group's strategywise family-size distribution is defined as a vector of each strategy's family-size distribution

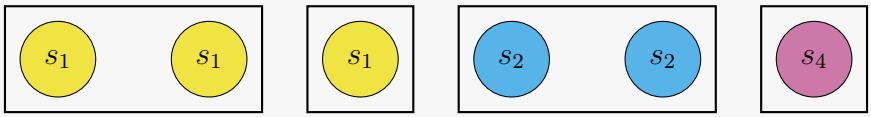
$$\boldsymbol{z} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m), \tag{A.24}$$

which is an m dimensional vector, each of whose component is an n dimensional vector. The

group's family-size distribution can be calculated from the strategywise family-size distribution as

$$\boldsymbol{y} = \mathrm{sum}(\boldsymbol{z}) := \sum_{i=1}^{m} \boldsymbol{z}_i.$$

Example 2. Consider the group illustrated below, where the boxes group together individuals who are family members



Each strategy's family-size distribution is

$$\begin{aligned}
\boldsymbol{z}_1 &= (1,1,0,0,0,0),\\
\boldsymbol{z}_2 &= (0,1,0,0,0,0),\\
\boldsymbol{z}_3 &= (0,0,0,0,0,0),\\
\boldsymbol{z}_4 &= (1,0,0,0,0,0),
\end{aligned}$$

and the group's strategywise family-size distribution is

$$\boldsymbol{z} = \big((1,1,0,0,0,0),\ (0,1,0,0,0,0),\ (0,0,0,0,0,0),\ (1,0,0,0,0,0)\big),$$

The group's family-size distribution is

$$\boldsymbol{y} = \mathrm{sum}(\boldsymbol{z}) = (2,2,0,0,0,0).$$ ◁

Given the strategy composition $\boldsymbol{g}_\mathrm{a} \in \mathcal{G}_\mathrm{a}$ in a group, we ask which strategywise family-size distribution $\boldsymbol{z}$ of the group is consistent with $\boldsymbol{g}_\mathrm{a}$. According to $\boldsymbol{g}_\mathrm{a}$, there are individuals with different strategies that must be from different families, and there are individuals with the same strategy that may be from the same family or different families, so $\boldsymbol{z}_i$ must be consistent with this. In fact, $\boldsymbol{z}_i$ is consistent with $g_{\mathrm{a},i}$ if and only if

$$\sum_{k=1}^{n} k\ z_{i,k} = g_{\mathrm{a},i} \tag{A.25}$$

holds. We write the set of all family-size distributions that are consistent with $g_{\mathrm{a},i}$, as

$$\mathcal{Z}_{g_{\mathrm{a},i}} = \left\{ \boldsymbol{z}_i \,\middle|\, \sum_{k=1}^{n} k\ z_{i,k} = g_{\mathrm{a},i} \right\}. \tag{A.26}$$

Therefore, the set of all strategywise family-size distributions $\boldsymbol{z}$ that are consistent with $\boldsymbol{g}$ is

$$\mathcal{Z}_{\boldsymbol{g}_\mathrm{a}} = \left\{ \boldsymbol{z} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m) \,\middle|\, \boldsymbol{z}_1 \in \mathcal{Z}_{g_{\mathrm{a},1}}, \ldots, \boldsymbol{z}_m \in \mathcal{Z}_{g_{\mathrm{a},m}} \right\}, \tag{A.27}$$

which is the Cartesian product of $g_{\mathrm{a},i}$'s (see Example 3 for an illustrative example)

$$\mathcal{Z}_{\boldsymbol{g}_\mathrm{a}} = \mathcal{Z}_{g_{\mathrm{a},1}} \times \ldots \times \mathcal{Z}_{g_{\mathrm{a},m}}. \tag{A.28}$$

Example 3. Consider a group with strategy composition

$$\boldsymbol{g}_\mathrm{a} = (3, 2, 0, 1).$$

This group has 3 members pursuing strategy s_1, 2 members pursuing strategy s_2, 1 members pursuing strategy s_4.

The members who are pursuing different strategies (e.g., s_1 and s_2) must be in separate families from each other. However, the members who are pursuing the same strategy (e.g., the 2 members pursuing strategy s_2) can either be in the same family or in separate families.

For the 3 members pursuing strategy s_1, the possible family-size distributions ($\mathcal{Z}_3$) can shown graphically as below, where boxes indicate the family boundaries

$$\boxed{s_1 \;\; s_1 \;\; s_1}$$

$$\boxed{s_1 \;\; s_1}\boxed{s_1}$$

$$\boxed{s_1}\boxed{s_1}\boxed{s_1}$$

For the 2 members pursuing strategy s_2, the possible family-size distributions ($\mathcal{Z}_2$) are

$$\boxed{s_2 \;\; s_2}$$

$$\boxed{s_2}\boxed{s_2}$$

For the single member pursuing strategy s_4, there is only 1 possible family-size distribution ($\mathcal{Z}_1$)

$$\boxed{s_4}$$

Therefore, the set of all possible strategywise family-size distributions is the Cartesian product of each strategy's possible family-size distributions above

$$\boxed{s_1 \;\; s_1 \;\; s_1}\boxed{s_2 \;\; s_2}\boxed{s_4}$$

$$\boxed{s_1 \;\; s_1 \;\; s_1}\boxed{s_2}\boxed{s_2}\boxed{s_4}$$

$$\boxed{s_1 \;\; s_1}\boxed{s_1}\boxed{s_2 \;\; s_2}\boxed{s_4}$$

$$\boxed{s_1 \;\; s_1}\boxed{s_1}\boxed{s_2}\boxed{s_2}\boxed{s_4}$$

$$\boxed{s_1}\boxed{s_1}\boxed{s_1}\boxed{s_2 \;\; s_2}\boxed{s_4}$$

$$\boxed{s_1}\boxed{s_1}\boxed{s_1}\boxed{s_2}\boxed{s_2}\boxed{s_4}$$

In math, the set of all family-size distributions consistent with the $g_{\mathrm{a},1} = 3$ s_1-strategists in the group is

$$\mathcal{Z}_3 = \{(0,0,1,0,0,0), (1,1,0,0,0,0), (3,0,0,0,0,0)\},$$

consistent with the $g_{\mathrm{a},2} = 2$ s_2-strategists in the group is

$$\mathcal{Z}_2 = \{(0,1,0,0,0,0), (2,0,0,0,0,0)\},$$

consistent with the $g_{\mathrm{a},3} = 0$ s_3-strategists in the group is

$$\mathcal{Z}_0 = \{(0,0,0,0,0,0)\},$$

and consistent with the $g_{\mathrm{a},4} = 1$ s_4-strategists in the group is

$$\mathcal{Z}_1 = \{(1,0,0,0,0,0)\}.$$

The set of all possible strategywise family-size distributions consistent with strategy $(3,2,0,1)$ is the Cartesian product of each strategy's possible family-size distributions

$$\begin{aligned}
\mathcal{Z}_{(3,2,0,1)} = \;& \mathcal{Z}_3 \times \mathcal{Z}_2 \times \mathcal{Z}_0 \times \mathcal{Z}_1, \\
= \Big\{ & \Big((0,0,1,0,0,0),\ (0,1,0,0,0,0),\ (0,0,0,0,0,0),\ (1,0,0,0,0,0)\Big), \\
& \Big((0,0,1,0,0,0),\ (2,0,0,0,0,0),\ (0,0,0,0,0,0),\ (1,0,0,0,0,0)\Big), \\
& \Big((1,1,0,0,0,0),\ (0,1,0,0,0,0),\ (0,0,0,0,0,0),\ (1,0,0,0,0,0)\Big), \\
& \Big((1,1,0,0,0,0),\ (2,0,0,0,0,0),\ (0,0,0,0,0,0),\ (1,0,0,0,0,0)\Big), \\
& \Big((3,0,0,0,0,0),\ (0,1,0,0,0,0),\ (0,0,0,0,0,0),\ (1,0,0,0,0,0)\Big), \\
& \Big((3,0,0,0,0,0),\ (2,0,0,0,0,0),\ (0,0,0,0,0,0),\ (1,0,0,0,0,0)\Big) \Big\}.
\end{aligned}$$

◁

Let $\boldsymbol{Z}$ be the random variable denoting the group's strategywise family-size distribution, which takes values $\boldsymbol{z}$. We write $\boldsymbol{Y} := \mathrm{sum}(\boldsymbol{Z}) = \sum_{i=1}^{m} \boldsymbol{Z}_i$ and $\boldsymbol{y} := \mathrm{sum}(\boldsymbol{z}) = \sum_{i=1}^{m} \boldsymbol{z}_i$ below. Then

$$\begin{aligned}
\mathbb{P}[\boldsymbol{G}_{\mathrm{a}} = \boldsymbol{g}_{\mathrm{a}}] &= \sum_{\boldsymbol{z} \in \mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}}} \mathbb{P}[\boldsymbol{Z} = \boldsymbol{z}], \\
&= \sum_{\boldsymbol{z} \in \mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}}} \mathbb{P}[\boldsymbol{Y} = \boldsymbol{y}]\ \mathbb{P}[\boldsymbol{Z} = \boldsymbol{z} \mid \boldsymbol{Y} = \boldsymbol{y}], \\
&= \sum_{\boldsymbol{z} \in \mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}}} F_{\boldsymbol{y}}\ \mathbb{P}[\boldsymbol{Z} = \boldsymbol{z} \mid \boldsymbol{Y} = \boldsymbol{y}], \qquad \text{(A.29)}
\end{aligned}$$

where the quantity $F_{\boldsymbol{y}}$ is calculated using the homophilic group-formation model and $\mathbb{P}[\boldsymbol{Z} = \boldsymbol{z} \mid \boldsymbol{Y} = \boldsymbol{y}]$ remains to be specified.

Given the group family structure $\boldsymbol{y}$, the number of ways to order strategies in $\boldsymbol{z}$ across families is a count of the multiset permutations

$$C(\boldsymbol{z}) = \prod_{j=1}^{n} \frac{(\sum_{i=1}^{m} z_{i,j})!}{\prod_{i=1}^{m} z_{i,j}!}, \qquad \text{(A.30)}$$

and the probability of drawing the family strategies given a particular ordering is

$$A(\boldsymbol{z}, \boldsymbol{p}) = \prod_{i=1}^{m} p_i^{\|\boldsymbol{z}_i\|}, \tag{A.31}$$

where $\|\boldsymbol{z}_i\|$ counts the number of families in a group pursuing strategy s_i

$$\|\boldsymbol{z}_i\| = \sum_{j=1}^{n} z_{i,j}. \tag{A.32}$$

Therefore

$$\mathbb{P}[\boldsymbol{G}_{\mathrm{a}} = \boldsymbol{g}_{\mathrm{a}}] = \sum_{\boldsymbol{z} \in \mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}}} F_{\mathrm{sum}(\boldsymbol{z})}\, C(\boldsymbol{z})\, A(\boldsymbol{z}, \boldsymbol{p}). \tag{A.33}$$

Substituting Eq. A.33 into Eq. A.22 gives Eq. 29 in the main text:

$$\Delta p_x \propto \sum_{\boldsymbol{g}_{\mathrm{a}} \in \mathcal{G}_{\mathrm{a}}} \left(\frac{g_{\mathrm{a},x}}{n} \hat{\pi}(\boldsymbol{e}_x, \boldsymbol{g}_{\mathrm{a}}) - p_x \sum_{i=1}^{m} \frac{g_{\mathrm{a},i}}{n} \hat{\pi}(\boldsymbol{e}_i, \boldsymbol{g}_{\mathrm{a}}) \right) \left(\sum_{\boldsymbol{z} \in \mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}}} C(\boldsymbol{z})\, A(\boldsymbol{z}, \boldsymbol{p})\, F_{\mathrm{sum}(\boldsymbol{z})} \right). \tag{A.34}$$

By taking a proper limit and a proper time scale, we obtain Eq. 30 in the main text:

$$\dot{p}_x = \sum_{\boldsymbol{g}_{\mathrm{a}} \in \mathcal{G}_{\mathrm{a}}} \left(\frac{g_{\mathrm{a},x}}{n} \hat{\pi}(\boldsymbol{e}_x, \boldsymbol{g}_{\mathrm{a}}) - p_x \sum_{i=1}^{m} \frac{g_{\mathrm{a},i}}{n} \hat{\pi}(\boldsymbol{e}_i, \boldsymbol{g}_{\mathrm{a}}) \right) \left(\sum_{\boldsymbol{z} \in \mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}}} C(\boldsymbol{z})\, A(\boldsymbol{z}, \boldsymbol{p})\, F_{\mathrm{sum}(\boldsymbol{z})} \right). \tag{A.35}$$

Example 4. In Example 3 above, take

$$\boldsymbol{z} = \Big((1,1,0,0,0,0),\ (2,0,0,0,0,0),\ (0,0,0,0,0,0),\ (1,0,0,0,0,0) \Big) \in \mathcal{Z}_{(3,2,0,1)}$$

as an example, which corresponds to the following strategywise family-size distribution:

s_1	s_1	s_1	s_2	s_2	s_4

For this $\boldsymbol{z}$, we have

$$C(\boldsymbol{z}) = \underbrace{\frac{(1+2+0+1)!}{1!\,2!\,0!\,1!}}_{j=1} \cdot \underbrace{\frac{(1+0+0+0)!}{1!\,0!\,0!\,0!}}_{j=2} = 12$$

and

$$A(\boldsymbol{z}, \boldsymbol{p}) = p_1^{1+1}\, p_2^2\, p_3^0\, p_4^1 = p_1^2\, p_2^2\, p_4.$$

Since

$$\boldsymbol{y} = \mathrm{sum}(\boldsymbol{z}) = (4,1,0,0,0,0),$$

it follows that

$$C(\boldsymbol{z})\, A(\boldsymbol{z}, \boldsymbol{p})\, F_{\mathrm{sum}(\boldsymbol{z})} = 12 p_1^2\, p_2^2\, p_4\, F_{(4,1,0,0,0,0)}.$$

◁

B Relationships between relatedness coefficients and group family-partition probabilities

Our homophilic group-formation models parameterise the evolutionary dynamics in terms of group family partition-structure probabilities $F_{\boldsymbol{q}}$; however, kin-selection models have traditionally been parameterised using relatedness coefficients like dyadic relatedness r_2. In this Supplement, we provide a method of converting between partition-structure probabilities and relatedness coefficients. Relatedness coefficients may provide a more biologically intuitive parameterisation in keeping with the traditional approach, and can reduce the number of parameters needed to describe the evolutionary dynamics if a game can be decomposed into a sum of smaller-group-size games (details in next Supplement C).

In the main text, we showed that for games of arbitrary degree, relatedness coefficients up to r_2 and r_3 are sufficient to parameterise the strategy dynamics for group sizes $n = 2$ and $n = 3$, respectively. However, for $n \geq 4$, coefficients up to r_n are no longer sufficient to parameterise the dynamics. The purpose of this Supplement is to generalise the concept of the relatedness coefficient to $r_{\boldsymbol{\rho}}$ (e.g., $r_{[2,2]}$ for $n = 4$ below), where $\boldsymbol{\rho}$ is a multiset, which allows us to include additional relatedness coefficients to parameterise for large n.

B.1 How to calculate $r_{\boldsymbol{\rho}}$ given the family partition structure probabilities $F_{\boldsymbol{q}}$

We generalise the concept of the relatedness coefficient from r_k to $r_{\boldsymbol{\rho}}$, where $\boldsymbol{\rho} = \left[\rho_1, \rho_2, \ldots, \rho_{|\boldsymbol{\rho}|}\right]$ is a multiset that describes a way of sampling without replacement members from a group into sets of size ρ_i. Here, with some abuse of notation, $|\boldsymbol{\rho}|$ represents the number of components in multiset $\boldsymbol{\rho}$. In particular, $\boldsymbol{\rho}$ specifies that we sample ρ_1 group members and place them in one set, ρ_2 group members and place them in a second set, and so on for $|\boldsymbol{\rho}|$ sets. We will call the outcome of drawing a sample according to $\boldsymbol{\rho}$ an outcome that '*satisfies* $\boldsymbol{\rho}$' if, for every sampled set, all individuals within that set are family members. The relatedness coefficient $r_{\boldsymbol{\rho}}$ is the probability that a randomly drawn group sampled according to $\boldsymbol{\rho}$ will satisfy $\boldsymbol{\rho}$. For example, $r_{[2,2]}$ represents the probability that, when we sample without replacement $2 + 2 = 4$ members from a randomly drawn group from the population, 1st and 2nd members belong to the same family and 3rd and 4th members belong to the same family. Note that those two families can be either the same family or different families.

Given the family partition structure probabilities $\{F_{\boldsymbol{q}}\}$, where each $\boldsymbol{q}$ is a multiset (see Table 1 in the main text), $r_{\boldsymbol{\rho}}$ can be calculated as follows. The first step is to identify all possible ways to allocate sets structured according to $\boldsymbol{\rho}$ into a group with family partition structure $\boldsymbol{q}$ such that the allocation satisfies $\boldsymbol{\rho}$. Recall that $\boldsymbol{q} = \left[q_1, q_2, \ldots, q_{|\boldsymbol{q}|}\right]$ is the family partition structure of the group, which specifies that q_1 group members are members of one family, q_2 group members are members of a second family, and so on for all $|\boldsymbol{q}|$ family partitions, and $\boldsymbol{q}$ satisfies $\sum_{j=1}^{|\boldsymbol{q}|} q_j = n$. Let $\boldsymbol{\alpha} = \left(\alpha_1, \alpha_2, \ldots, \alpha_{|\boldsymbol{q}|}\right)$ be an allocation of sets structured according to $\boldsymbol{\rho}$ into $\boldsymbol{q}$ that satisfies $\boldsymbol{\rho}$, where each element α_j is the set of indices i of $\boldsymbol{\rho}$ that have been allocated to partition q_j (see Example 5 below). It is possible for more than one set to be allocated to the same family partition. All partitions must be large enough to contain the sets allocated to it. Therefore, the set of all possible allocations that satisfy $\boldsymbol{\rho}$ is

$$\mathcal{A}_{\boldsymbol{\rho},\boldsymbol{q}} = \left\{ \boldsymbol{\alpha} \;\middle|\; \sum_{i \in \alpha_j} \rho_i \leq q_j \;\forall j \right\}. \tag{B.1}$$

If a group with family partition structure $\boldsymbol{q}$ is sampled according to $\boldsymbol{\rho}$, then the probability

that the allocation outcome $\boldsymbol{a}$ will be equal to a particular allocation $\boldsymbol{\alpha} \in \mathcal{A}_{\boldsymbol{\rho},\boldsymbol{q}}$ is

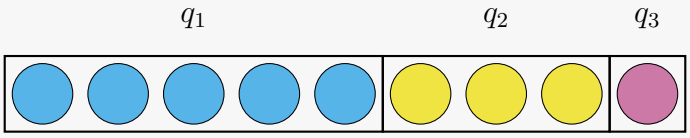
$$\mathbb{P}[\boldsymbol{a} = \boldsymbol{\alpha} \mid \boldsymbol{\rho}, \boldsymbol{q}] = \frac{\left(n - \sum_{k=1}^{|\boldsymbol{\rho}|} \rho_k\right)!}{n!} \prod_{j=1}^{|\boldsymbol{q}|} \frac{q_j!}{\left(q_j - \sum_{i \in \alpha_j} \rho_i\right)!}. \tag{B.2}$$

The first term represents the decreasing number of individuals remaining in the group as the sampling without replacement progresses. The second term represents the decreasing number of individuals remaining in partition q_j as the individuals allocated to q_j are sampled. Therefore, the probability that a set sampled according $\boldsymbol{\rho}$ from a group with family partition structure $\boldsymbol{q}$ satisfies $\boldsymbol{\rho}$ is

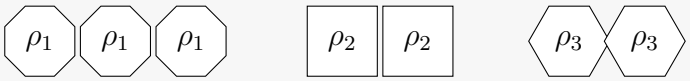
$$r_{\boldsymbol{\rho},\boldsymbol{q}} \coloneqq \mathbb{P}[\boldsymbol{a} \in \mathcal{A}_{\boldsymbol{\rho},\boldsymbol{q}} \mid \boldsymbol{\rho}, \boldsymbol{q}] = \sum_{\boldsymbol{\alpha} \in \mathcal{A}_{\boldsymbol{\rho},\boldsymbol{q}}} \mathbb{P}[\boldsymbol{a} = \boldsymbol{\alpha} \mid \boldsymbol{\rho}, \boldsymbol{q}]. \tag{B.3}$$

Example 5. Calculate the probability $r_{\boldsymbol{\rho},\boldsymbol{q}}$ that a sample according to $\boldsymbol{\rho} = [3, 2, 2]$ drawn from a group with family partition structure $\boldsymbol{q} = [5, 3, 1]$ will satisfy $\boldsymbol{\rho}$.

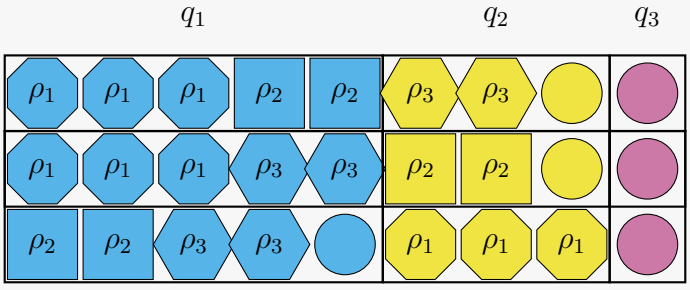
We illustrate the group with $\boldsymbol{q} = [5, 3, 1]$ below, where the boxes separate the families and shared family membership is indicated by colours



We illustrate our sampled sets ρ_i by three shapes:



In order to satisfy $\boldsymbol{\rho}$, individuals in the sampled sets must share the same family membership. There are three possible allocations that satisfy $\boldsymbol{\rho}$, which are illustrated



The allocations that satisfy $\boldsymbol{\rho}$ are written

$$\mathcal{A}_{\boldsymbol{\rho},\boldsymbol{q}} = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\},$$

where

$$\begin{aligned}
\boldsymbol{\alpha}_1 &= (\{1, 2\}, \{3\}, \{\}), \\
\boldsymbol{\alpha}_2 &= (\{1, 3\}, \{2\}, \{\}), \\
\boldsymbol{\alpha}_3 &= (\{2, 3\}, \{1\}, \{\}).
\end{aligned}$$

The probability of drawing $\boldsymbol{\alpha}_1$ is (Eq. B.2)

$$\mathbb{P}[\boldsymbol{a} = \boldsymbol{\alpha}_1 \mid \boldsymbol{\rho}, \boldsymbol{q}] = \frac{(9-7)!}{9!} \cdot \frac{5!}{(5-3-2)!} \cdot \frac{3!}{(3-2)!} \cdot \frac{1!}{1!} = \underbrace{\frac{5}{9} \cdot \frac{4}{8} \cdot \frac{3}{7}}_{\rho_1 \to q_1} \cdot \underbrace{\frac{2}{6} \cdot \frac{1}{5}}_{\rho_2 \to q_1} \cdot \underbrace{\frac{3}{4} \cdot \frac{2}{3}}_{\rho_3 \to q_2} = \frac{1}{252},$$

where the expanded form has been annotated to indicate which components of the equation refer to which allocations. In this example, the probability of drawing $\boldsymbol{\alpha}_2$ and $\boldsymbol{\alpha}_3$ is also $\frac{1}{252}$. Thus, the total probability is the sum of allocation probabilities

$$r_{\boldsymbol{\rho},\boldsymbol{q}} = \frac{1}{252} + \frac{1}{252} + \frac{1}{252} = \frac{1}{84},$$

which corresponds to Eq. B.3. ◁

In a given population, the probability that a randomly sampled group will have partition structure $\boldsymbol{q}$ is $F_{\boldsymbol{q}}$. Therefore, the probability that a randomly drawn group sampled according to $\boldsymbol{\rho}$ will satisfy $\boldsymbol{\rho}$ is

$$r_{\boldsymbol{\rho}} = \sum_{\boldsymbol{q} \vdash n} r_{\boldsymbol{\rho},\boldsymbol{q}} F_{\boldsymbol{q}}, \tag{B.4}$$

where the sum over $\boldsymbol{q} \vdash n$ is the sum over $\boldsymbol{q}$ that are integer partitions of n.

B.2 Our proposal for how to choose the relatedness-coefficient parameterisation

For a given n, there are more $r_{\boldsymbol{\rho}}$ coefficients than are needed to describe the dynamics. Therefore, in this subsection, we propose a scheme for choosing which $r_{\boldsymbol{\rho}}$ to include in the parameterisation.

We typically need $P_n - 1$ coefficients to obtain dynamical sufficiency, where P_n is the partition function number counting the total number of possible integer partitions of the group size n. The reason why we need $P_n - 1$ is because there are P_n partition-structure probabilities $F_{\boldsymbol{q}_k}$, but also $\sum_{k=1}^{P_n} F_{\boldsymbol{q}_k} = 1$; therefore, only $P_n - 1$ partition probabilities $F_{\boldsymbol{q}_k}$ are needed. Therefore, if we are parameterising instead using $r_{\boldsymbol{\rho}}$, we expect we will need $P_n - 1$ coefficients (see proof below). However, there are $\sum_{k=1}^{n} P_k$ relatedness coefficients to choose from.

For any relatedness coefficient $r_{\boldsymbol{\rho}}$ with $\rho \in \boldsymbol{\rho}$ such that $\rho = 1$, that size-1 set counts the probability that an individual is always in the same family as itself, which is always simply 1. Intuitively, these sets seem redundant. Therefore, we propose that such coefficients should be excluded, and models should be parameterised only with the set of relatedness coefficients that do not concern size-1 sets, i.e.,

$$\mathcal{R}_n = \{r_{\boldsymbol{\rho}} \,|\, \boldsymbol{\rho} \vdash k,\ 2 \leq k \leq n,\ \rho_i > 1\ \forall i\}\,, \tag{B.5}$$

where $\boldsymbol{\rho} \vdash k$ means that multiset $\boldsymbol{\rho}$ is a partition of integer k. In words, $r_{\boldsymbol{\rho}} \in \mathcal{R}_n$ is a relatedness coefficient based on multiset $\boldsymbol{\rho}$ which is a partition of some k ($2 \leq k \leq n$) and which does not contain 1 in its element. We can verify that the size of this set is $|\mathcal{R}_n| = P_n - 1$ as needed. Let E_k be the number of partitions of k that contain at least one element equal to 1. For group sizes 1 and 2, we have: $P_1 = 1$, $E_1 = 1$, $|\mathcal{R}_1| = 0$; and $P_2 = 2$, $E_2 = 1$, $|\mathcal{R}_2| = 1$. In general, for $n \geq 2$, $|\mathcal{R}_n| = \sum_{k=2}^{n}(P_k - E_k)$. The partitions of k that contain at least one element equal to 1 are obtained by adding the element 1 to each partition of $k-1$; therefore, $E_k = P_{k-1}$. Therefore, $|\mathcal{R}_n| = \sum_{k=2}^{n}(P_k - P_{k-1}) = P_n - P_1 = P_n - 1$.

Example 6. For $n = 6$, $P_n - 1 = 10$, the set of relatedness coefficients we suggest to parameterise the model is

$$\mathcal{R}_n = \left\{ r_{[2]}, r_{[3]}, r_{[2,2]}, r_{[4]}, r_{[3,2]}, r_{[5]}, r_{[4,2]}, r_{[3,3]}, r_{[2,2,2]}, r_{[6]} \right\}, \qquad \triangleleft$$

which are all the relatedness coefficients except those that involve sampling a set of size 1.

$\mathcal{R}_n$ in Eq. B.5 is equivalent to

$$\mathcal{R}_n = \{ r_{\boldsymbol{\rho}} \,|\, \boldsymbol{\rho} = [q \in \boldsymbol{q} \mid q \neq 1], \ \boldsymbol{\rho} \neq \varnothing, \ \forall \boldsymbol{q} \vdash n \}. \tag{B.6}$$

The multiset $\boldsymbol{\rho} = [q \in \boldsymbol{q} \mid q \neq 1]$ is only empty for one partition of n, which is $\boldsymbol{q} = [1, 1, \ldots, 1]$; therefore, again, the size of the set of relatedness coefficients $|\mathcal{R}_n| = P_n - 1$.

Example 7. If $\boldsymbol{q} = [4, 3, 3, 1, 1]$, its corresponding $\boldsymbol{\rho}$ in Eq. B.6 is that same multiset with all the ones removed, i.e., $\boldsymbol{\rho} = [4, 3, 3]$. $\triangleleft$

In particular, $\{\mathcal{R}_n\}$ is an increasing sequence of sets, as $\mathcal{R}_2 \subsetneq \mathcal{R}_3 \subsetneq \mathcal{R}_4 \subsetneq \cdots$. One benefit of choosing relatedness coefficients according to Eq. B.5 is that, if the extent of nonlinearity of the game is low, then the subset of $\mathcal{R}_n$ needed to to describe the dynamics can be easily identified. For 2-strategy games, Ohtsuki (2014) wrote the payoffs to a focal A- and B-strategist as polynomial functions of the number k of A-strategists among the $n-1$ other players and defined the *degree* of the game as the maximum degree of the polynomials. Ohtsuki (2014) proved that dynamical sufficiency of an n-player game of degree d can be obtained with relatedness coefficients up to $d+1$, i.e., $\mathcal{R}_{d+1}$. For games with more than 2 strategies, we hypothesise that analogous subsets can be identified (SI C).

Example 8. For a 6-player, 2-strategy game with degree $d = 4$, only relatedness coefficients up to $d + 1 = 5$ are needed to describe the dynamics: $\mathcal{R}_5 = \left\{ r_{[2]}, r_{[3]}, r_{[2,2]}, r_{[4]}, r_{[3,2]}, r_{[5]} \right\}$. The elements in $\mathcal{R}_6 \setminus \mathcal{R}_5 = \left\{ r_{[4,2]}, r_{[3,3]}, r_{[2,2,2]}, r_{[6]} \right\}$ are not needed. $\triangleleft$

B.2.1 Proof that $\mathcal{R}_n$ is sufficient to describe the dynamics

To prove that the set of parameters $\mathcal{R}_n$ defined in Eq. B.6 is sufficient to describe the dynamics, we will show that it has a 1-to-1 relationship with $P_n - 1$ of the probabilities $F_{\boldsymbol{q}}$. To show this, we first specify vectors with particular ordering for the elements of $\mathcal{R}_n$ and for corresponding $F_{\boldsymbol{q}}$ probabilities, and then we describe the relationship between the two vectors by a matrix M. We then prove that M is upper triangular, which implies that it is an invertible matrix and therefore the relationship is 1-to-1.

Let $\vec{\boldsymbol{\rho}}$ denote a particular ordering of sampling schemes $\boldsymbol{\rho}$, and let $\vec{\boldsymbol{\rho}}_{(i)}$ denote the i-th sampling scheme in that ordering. Define $\vec{\boldsymbol{r}}$ as the vector of all relatedness coefficients $r_{\boldsymbol{\rho}} \in \mathcal{R}_n$ such that their order in the vector $\vec{\boldsymbol{r}}$ corresponds to the ordering of $\vec{\boldsymbol{\rho}}$; that is, $\vec{\boldsymbol{r}} = \left(r_{\vec{\boldsymbol{\rho}}_{(1)}}, r_{\vec{\boldsymbol{\rho}}_{(2)}}, \ldots, r_{\vec{\boldsymbol{\rho}}_{(P_n-1)}} \right)$.

Let $\vec{\boldsymbol{q}}$ denote a particular ordering of family partition structures $\boldsymbol{q}$, and let $\vec{\boldsymbol{q}}_{(i)}$ denote the i-th partition structure in that ordering. Denote the set of all family partition-structure probabilities apart from the structure where no group members are from the same family by

$$\mathcal{F}_n = \{F_{\boldsymbol{q}} \,|\, \boldsymbol{q} \vdash n, \boldsymbol{q} \neq [1, 1, \ldots, 1]\}\,.$$

$\mathcal{F}_n$ has size $|\mathcal{F}_n| = P_n - 1$. Define $\vec{\boldsymbol{F}}$ as the vector of all partition-structure probabilities $F_{\boldsymbol{q}} \in \mathcal{F}_n$ such that their order in the vector $\vec{\boldsymbol{F}}$ corresponds to the ordering of $\vec{\boldsymbol{q}}$; that is, $\vec{\boldsymbol{F}} = \left(F_{\vec{\boldsymbol{q}}_{(1)}}, \ldots, F_{\vec{\boldsymbol{q}}_{(P_n-1)}}\right)$.

Denote the size of the multiset $\vec{\boldsymbol{\rho}}_{(i)}$ by $|\vec{\boldsymbol{\rho}}_{(i)}|$, and denote the sum of its elements by

$$\|\vec{\boldsymbol{\rho}}_{(i)}\| = \sum_{k=1}^{|\vec{\boldsymbol{\rho}}_{(i)}|} (\vec{\boldsymbol{\rho}}_{(i)})_k.$$

We can now specify an ordering of $\vec{\boldsymbol{r}}$ and $\vec{\boldsymbol{F}}$ by specifying an ordering of $\vec{\boldsymbol{\rho}}$ and $\vec{\boldsymbol{q}}$. We specify that $\vec{\boldsymbol{\rho}}$ and $\vec{\boldsymbol{q}}$ must satisfy the following three conditions.

$$\text{Condition 1:} \qquad \vec{\boldsymbol{q}}_{(i)} = \vec{\boldsymbol{\rho}}_{(i)} \cup \underbrace{[1, 1, \ldots, 1]}_{n - \|\vec{\boldsymbol{\rho}}_{(i)}\|} \tag{B.7}$$

where $\cup$ is a multiset union. Condition 1 reflects the correspondence described in Eq. B.6 between the relatedness coefficients we have chosen and $P_n - 1$ of the family partition structures. Condition 1 means that the ordering of $\vec{\boldsymbol{q}}$ is defined by the ordering of $\vec{\boldsymbol{\rho}}$; therefore, the next two conditions specify constraints on the ordering of $\vec{\boldsymbol{\rho}}$ only.

$$\text{Condition 2:} \qquad \|\vec{\boldsymbol{\rho}}_{(i)}\| < \|\vec{\boldsymbol{\rho}}_{(j)}\| \implies i < j \tag{B.8}$$

which means that $\boldsymbol{\rho}$ with smaller sums come earlier in the indexing.

$$\text{Condition 3:} \qquad \text{if } \|\vec{\boldsymbol{\rho}}_{(i)}\| = \|\vec{\boldsymbol{\rho}}_{(j)}\|, \quad \text{then } |\vec{\boldsymbol{\rho}}_{(i)}| > |\vec{\boldsymbol{\rho}}_{(j)}| \implies i < j. \tag{B.9}$$

which means that, within sequences of $\boldsymbol{\rho}$ whose sums are equal, the $\boldsymbol{\rho}$ with greater length come earlier in the indexing.

Example 9. For $n = 5$, we have $P_5 - 1 = 7 - 1 = 6$, and $\mathcal{R}_5$ and $\mathcal{F}_5$ are

$$\mathcal{R}_5 = \left\{r_{[2]}, r_{[3]}, r_{[2,2]}, r_{[4]}, r_{[3,2]}, r_{[5]}\right\},$$
$$\mathcal{F}_5 = \left\{F_{[5]}, F_{[4,1]}, F_{[3,2]}, F_{[3,1,1]}, F_{[2,2,1]}, F_{[2,1,1,1]}\right\}.$$

The ordering satisfying Conditions 1-3 above is unique, and it is

$$\vec{\boldsymbol{\rho}} = (\,[2]\,, \quad [3]\,, \quad [2,2]\,, \quad [4]\,, \quad [3,2]\,, [5]),$$
$$\vec{\boldsymbol{q}} = (\,[2,1,1,1]\,, [3,1,1]\,, [2,2,1]\,, [4,1]\,, [3,2]\,, [5]),$$

so we have

$$\vec{\boldsymbol{r}} = (r_{[2]}, \quad r_{[3]}, \quad r_{[2,2]}, \quad r_{[4]}, \quad r_{[3,2]}, r_{[5]}),$$
$$\vec{\boldsymbol{F}} = (F_{[2,1,1,1]}, F_{[3,1,1]}, F_{[2,2,1]}, F_{[4,1]}, F_{[3,2]}, F_{[5]}).$$

That is, for $\vec{\boldsymbol{r}}$, a multiset with a smaller sum comes first, and a larger-sized multiset comes first if sums are equal, and to obtain $\vec{\boldsymbol{F}}$, fill the corresponding multiset in $\vec{\boldsymbol{r}}$ with 1's to make the sum equal to n.

In contrast, for $n = 6$, the ordering satisfying Conditions 1-3 is not unique, and both

$$\vec{\boldsymbol{r}} = (r_{[2]}, \quad r_{[3]}, \quad r_{[2,2]}, \quad r_{[4]}, \quad r_{[3,2]}, \quad r_{[5]}, \quad r_{[2,2,2]}, r_{[4,2]}, r_{[3,3]}, r_{[6]}),$$
$$\vec{\boldsymbol{F}} = (F_{[2,1,1,1,1]}, F_{[3,1,1,1]}, F_{[2,2,1,1]}, F_{[4,1,1]}, F_{[3,2,1]}, F_{[5,1]}, F_{[2,2,2]}, F_{[4,2]}, F_{[3,3]}, F_{[6]}),$$

and

$$\vec{\boldsymbol{r}} = (r_{[2]}, \quad r_{[3]}, \quad r_{[2,2]}, \quad r_{[4]}, \quad r_{[3,2]}, \quad r_{[5]}, \quad r_{[2,2,2]}, r_{[3,3]}, r_{[4,2]}, r_{[6]}),$$
$$\vec{\boldsymbol{F}} = (F_{[2,1,1,1,1]}, F_{[3,1,1,1]}, F_{[2,2,1,1]}, F_{[4,1,1]}, F_{[3,2,1]}, F_{[5,1]}, F_{[2,2,2]}, F_{[3,3]}, F_{[4,2]}, F_{[6]}),$$

satisfy the required criteria. Specifically, multisets $[3, 3]$ and $[4, 2]$ have the same sum and the same size, so either can come first. ◁

To prove that the mapping from $\mathcal{R}_n$ to $\mathcal{F}_n$ is 1-to-1, we show that the matrix M that describes the relationship $\vec{\boldsymbol{r}} = M\vec{\boldsymbol{F}}$ has all non-zero elements on the diagonal and all zero elements below the diagonal. Examples of the M matrices up to size $n = 6$, from which one may directly observe that the matrices are upper-triangular, are given in the next subsection (Section B.3).

Here we provide a proof. The element $m_{i,j}$ of M is the probability $r_{\vec{\boldsymbol{\rho}}_{(i)},\vec{\boldsymbol{q}}_{(j)}}$ (Eq. B.4). From Eq. B.2 and Eq. B.3, $m_{i,j} = 0$ if and only if $\mathcal{A}_{\vec{\boldsymbol{\rho}}_{(i)},\vec{\boldsymbol{q}}_{(j)}} = \varnothing$.

First, let us first consider the diagonal elements $m_{i,i}$ of M. From Condition 1, we know that $\vec{\boldsymbol{q}}_{(i)} = \vec{\boldsymbol{\rho}}_{(i)} \cup [1, 1, \ldots, 1]$, so there is at least one allocation of $\vec{\boldsymbol{q}}_{(i)}$ into $\vec{\boldsymbol{\rho}}_{(i)}$ that satisfies $\vec{\boldsymbol{\rho}}_{(i)}$, namely $\boldsymbol{\alpha} = \big(\{1\}, \{2\}, \ldots, \{|\vec{\boldsymbol{\rho}}_{(i)}|\}\big)$. Therefore, $\mathcal{A}_{\vec{\boldsymbol{\rho}}_{(i)},\vec{\boldsymbol{q}}_{(i)}} \neq \varnothing$, and therefore $m_{i,i} \neq 0$.

Second, let us consider lower-triangular region of $m_{i,j}$ where $i > j$. Taking the contrapositive of Condition 2, there are three possible cases:

1. $\|\vec{\boldsymbol{\rho}}_{(i)}\| > \|\vec{\boldsymbol{\rho}}_{(j)}\|$: It is not possible to sample more individuals than there are available, therefore $\mathcal{A}_{\vec{\boldsymbol{\rho}}_{(i)},\vec{\boldsymbol{q}}_{(j)}} = \varnothing$ and therefore $m_{i,j} = 0$.
2. $\|\vec{\boldsymbol{\rho}}_{(i)}\| = \|\vec{\boldsymbol{\rho}}_{(j)}\|$ and $|\vec{\boldsymbol{\rho}}_{(i)}| < |\vec{\boldsymbol{\rho}}_{(j)}|$: At least one set in $\vec{\boldsymbol{\rho}}_{(i)}$ must be split between multiple family partitions in $\vec{\boldsymbol{q}}_{(j)}$, therefore $\mathcal{A}_{\vec{\boldsymbol{\rho}}_{(i)},\vec{\boldsymbol{q}}_{(j)}} = \varnothing$ and therefore $m_{i,j} = 0$.
3. $\|\vec{\boldsymbol{\rho}}_{(i)}\| = \|\vec{\boldsymbol{\rho}}_{(j)}\|$ and $|\vec{\boldsymbol{\rho}}_{(i)}| = |\vec{\boldsymbol{\rho}}_{(j)}|$: In this case, $\mathcal{A}_{\vec{\boldsymbol{\rho}}_{(i)},\vec{\boldsymbol{q}}_{(j)}} \neq \varnothing$ if and only if $\vec{\boldsymbol{\rho}}_{(i)} = \vec{\boldsymbol{\rho}}_{(j)}$. But each element of $\vec{\boldsymbol{\rho}}$ is unique and $i \neq j$ and therefore $\vec{\boldsymbol{\rho}}_{(i)} \neq \vec{\boldsymbol{\rho}}_{(j)}$. Therefore $\mathcal{A}_{\vec{\boldsymbol{\rho}}_{(i)},\vec{\boldsymbol{q}}_{(j)}} = \varnothing$ and therefore $m_{i,j} = 0$.

In all cases when $i > j$, $m_{i,j} = 0$. This ends the proof.

Example 10. Let us consider $n = 6$, and let us consider the following ordering

$$\vec{\boldsymbol{\rho}} = (\,[2]\,, \quad [3]\,, \quad [2,2]\,, \quad [4]\,, \quad [3,2]\,, \quad [5]\,, \quad [2,2,2]\,, [4,2]\,, [3,3]\,, [6]),$$
$$\vec{\boldsymbol{q}} = (\,[2,1,1,1,1]\,, [3,1,1,1]\,, [2,2,1,1]\,, [4,1,1]\,, [3,2,1]\,, [5,1]\,, [2,2,2]\,, [4,2]\,, [3,3]\,, [6]),$$

and

$$\vec{\boldsymbol{r}} = (r_{[2]}, \quad r_{[3]}, \quad r_{[2,2]}, \quad r_{[4]}, \quad r_{[3,2]}, \quad r_{[5]}, \quad r_{[2,2,2]}, r_{[4,2]}, r_{[3,3]}, r_{[6]}),$$
$$\vec{\boldsymbol{F}} = (F_{[2,1,1,1,1]}, F_{[3,1,1,1]}, F_{[2,2,1,1]}, F_{[4,1,1]}, F_{[3,2,1]}, F_{[5,1]}, F_{[2,2,2]}, F_{[4,2]}, F_{[3,3]}, F_{[6]}).$$

The first part of the proof above says $m_{5,5} \neq 0$, for example, because sampling scheme $\vec{\boldsymbol{\rho}}_{(5)} = [3, 2]$ (i.e, taking 3 individuals first, and then 2 individuals second) can be compatible with family partition structure $\vec{\boldsymbol{q}}_{(5)} = [3, 2, 1]$ (i.e. if the first 3 individuals are taken from the family of size 3 in $[3, 2, 1]$ and second 2 individuals are taken from the family of size 2 in $[3, 2, 1]$, then the first 3 individuals belong to one family and the second 2 individuals belong to another family).

For the second part of the proof, case 1 above says $m_{5,4} = 0$, for example, because sampling scheme $\vec{\boldsymbol{\rho}}_{(5)} = [3, 2]$ can never be compatible with family partition structure $\vec{\boldsymbol{q}}_{(4)} = [4, 1, 1]$. The reason for this is because $3 + 2 > 4$ and therefore it is impossible that the first 3 individuals belong to one family and at the same time the second 2 individuals belong to another family under $\vec{\boldsymbol{q}}_{(4)}$.

Case 2 above says $m_{6,5} = 0$, for example, because sampling scheme $\vec{\boldsymbol{\rho}}_{(6)} = [5]$ can never be compatible with family partition structure $\vec{\boldsymbol{q}}_{(5)} = [3, 2, 1]$. The reason for this is because $\vec{\boldsymbol{\rho}}_{(6)} = [5]$ is a smaller-sized partition of 5 than $\vec{\boldsymbol{\rho}}_{(5)} = [3, 2]$ is and therefore it is impossible that 5 individuals belong to the same family under $\vec{\boldsymbol{q}}_{(5)}$.

Case 3 above says $m_{9,8} = 0$, for example, because sampling scheme $\vec{\boldsymbol{\rho}}_{(9)} = [3, 3]$ can never be compatible with family partition structure $\vec{\boldsymbol{q}}_{(8)} = [4, 2]$. The reason for this is because both $\vec{\boldsymbol{\rho}}_{(9)} = [3, 3]$ and $\vec{\boldsymbol{\rho}}_{(8)} = [4, 2]$ are partitions of 6 and their sizes are the same (=2) but they are different partitions, and therefore it is impossible that the first 3 individuals belong to one family and at the same time the second 3 individuals belong to another family under $\vec{\boldsymbol{q}}_{(8)}$. ◁

B.3 Coefficients for converting between $r_{\boldsymbol{\rho}}$ and $F_{\boldsymbol{q}}$ up to $n = 6$

Code to calculate the coefficients needed to convert between $r_{\boldsymbol{\rho}}$ and $F_{\boldsymbol{q}}$ can be found in the Github repository: `scripts/related2partprob/calc_matrices.py`. In this subsection, we tabulate the coefficients up to $n = 6$ and illustrate their use.

In the main text, for $n = 3$, we calculated

$$\begin{aligned} r_{[2]} &= \frac{F_{[2,1]}}{3} + F_{[3]}, \\ r_{[3]} &= F_{[3]}, \end{aligned} \tag{B.10}$$

which gives

$$\begin{aligned} F_{[2,1]} &= 3(r_{[2]} - r_{[3]}), \\ F_{[3]} &= r_{[3]}. \end{aligned} \tag{B.11}$$

(In this case, $F_{[1,1,1]} = 1 - 3r_{[2]} + 2r_{[3]}$). Tables B.1 and B.2 present the coefficients computed for converting from $F_{\boldsymbol{q}}$ to $r_{\boldsymbol{\rho}}$ and visa versa, respectively, which correspond to the coefficients in the equations presented in the main text.

Table B.1: Coefficients to convert from family partition probabilities to relatedness coefficients for $n = 3$.

	factor	$F_{[2,1]}$	$F_{[3]}$
$r_{[2]}$	1/6	2	6
$r_{[3]}$	1/6	0	6

Table B.2: Coefficients to convert from relatedness coefficients to family partition probabilities for $n = 3$.

	factor	$r_{[2]}$	$r_{[3]}$
$F_{[2,1]}$	3	1	-1
$F_{[3]}$	1	0	1

Table B.3: Coefficients to convert from family partition probabilities to relatedness coefficients for $n = 4$.

	factor	$F_{[2,1,1]}$	$F_{[3,1]}$	$F_{[2,2]}$	$F_{[4]}$
$r_{[2]}$	1/12	2	6	4	12
$r_{[3]}$	1/24	0	6	0	24
$r_{[2,2]}$	1/24	0	0	8	24
$r_{[4]}$	1/24	0	0	0	24

For $n = 4$, we computed the coefficients presented in Tables B.3 and B.4.
The coefficients in Table B.3 give relationships

$$\begin{aligned} r_{[2]} &= \frac{F_{[2,1,1]}}{6} + \frac{F_{[3,1]}}{2} + \frac{F_{[2,2]}}{3} + F_{[4]}, \\ r_{[3]} &= \frac{F_{[3,1]}}{4} + F_{[4]}, \\ r_{[2,2]} &= \frac{F_{[2,2]}}{3} + F_{[4]}, \\ r_{[4]} &= F_{[4]}. \end{aligned} \tag{B.12}$$

Table B.4: Coefficients to convert from relatedness coefficients to family partition probabilities for $n = 4$.

	factor	$r_{[2]}$	$r_{[3]}$	$r_{[2,2]}$	$r_{[4]}$
$F_{[2,1,1]}$	6	1	-2	-1	2
$F_{[3,1]}$	4	0	1	0	-1
$F_{[2,2]}$	3	0	0	1	-1
$F_{[4]}$	1	0	0	0	1

The coefficients in Table B.4 give relationships

$$\begin{aligned} F_{[2,1,1]} &= 6\left(r_{[2]} - 2r_{[3]} - r_{[2,2]} + 2r_{[4]}\right), \\ F_{[3,1]} &= 4\left(r_{[3]} - r_{[4]}\right), \\ F_{[2,2]} &= 3\left(r_{[2,2]} - r_{[4]}\right), \\ F_{[4]} &= r_{[4]}. \end{aligned} \tag{B.13}$$

Coefficients for $n = 5$ and $n = 6$ are presented in Tables B.5 to B.8. Relationships between $r_{\boldsymbol{\rho}}$ and $F_{\boldsymbol{q}}$ can derived in an analogous way to the examples above for $n = 3$ and $n = 4$.

Table B.5: Coefficients to convert from family partition probabilities to relatedness coefficients for $n = 5$.

	factor	$F_{[2,1,1,1]}$	$F_{[3,1,1]}$	$F_{[2,2,1]}$	$F_{[4,1]}$	$F_{[3,2]}$	$F_{[5]}$
$r_{[2]}$	1/20	2	6	4	12	8	20
$r_{[3]}$	1/60	0	6	0	24	6	60
$r_{[2,2]}$	1/120	0	0	8	24	24	120
$r_{[4]}$	1/120	0	0	0	24	0	120
$r_{[3,2]}$	1/120	0	0	0	0	12	120
$r_{[5]}$	1/120	0	0	0	0	0	120

Table B.6: Coefficients to convert from relatedness coefficients to family partition probabilities for $n = 5$.

	factor	$r_{[2]}$	$r_{[3]}$	$r_{[2,2]}$	$r_{[4]}$	$r_{[3,2]}$	$r_{[5]}$
$F_{[2,1,1,1]}$	10	1	-3	-3	6	5	-6
$F_{[3,1,1]}$	10	0	1	0	-2	-1	2
$F_{[2,2,1]}$	15	0	0	1	-1	-2	2
$F_{[4,1]}$	5	0	0	0	1	0	-1
$F_{[3,2]}$	10	0	0	0	0	1	-1
$F_{[5]}$	1	0	0	0	0	0	1

Table B.7: Coefficients to convert from family partition probabilities to relatedness coefficients for $n = 6$.

	factor	$F_{[2,1,1,1,1]}$	$F_{[3,1,1,1]}$	$F_{[2,2,1,1]}$	$F_{[4,1,1]}$	$F_{[3,2,1]}$	$F_{[5,1]}$	$F_{[2,2,2]}$	$F_{[4,2]}$	$F_{[3,3]}$	$F_{[6]}$
$r_{[2]}$	1/30	2	6	4	12	8	20	6	14	12	30
$r_{[3]}$	1/120	0	6	0	24	6	60	0	24	12	120
$r_{[2,2]}$	1/360	0	0	8	24	24	120	24	72	72	360
$r_{[4]}$	1/360	0	0	0	24	0	120	0	24	0	360
$r_{[3,2]}$	1/720	0	0	0	0	12	120	0	48	72	720
$r_{[5]}$	1/720	0	0	0	0	0	120	0	0	0	720
$r_{[2,2,2]}$	1/720	0	0	0	0	0	0	48	144	0	720
$r_{[4,2]}$	1/720	0	0	0	0	0	0	0	48	0	720
$r_{[3,3]}$	1/720	0	0	0	0	0	0	0	0	72	720
$r_{[6]}$	1/720	0	0	0	0	0	0	0	0	0	720

Table B.8: Coefficients to convert from relatedness coefficients to family partition probabilities for $n = 6$.

	factor	$r_{[2]}$	$r_{[3]}$	$r_{[2,2]}$	$r_{[4]}$	$r_{[3,2]}$	$r_{[5]}$	$r_{[2,2,2]}$	$r_{[4,2]}$	$r_{[3,3]}$	$r_{[6]}$
$F_{[2,1,1,1,1]}$	15	1	-4	-6	12	20	-24	3	-18	-8	24
$F_{[3,1,1,1]}$	20	0	1	0	-3	-3	6	0	3	2	-6
$F_{[2,2,1,1]}$	45	0	0	1	-1	-4	4	-1	5	2	-6
$F_{[4,1,1]}$	15	0	0	0	1	0	-2	0	-1	0	2
$F_{[3,2,1]}$	60	0	0	0	0	1	-1	0	-1	-1	2
$F_{[5,1]}$	6	0	0	0	0	0	1	0	0	0	-1
$F_{[2,2,2]}$	15	0	0	0	0	0	0	1	-3	0	2
$F_{[4,2]}$	15	0	0	0	0	0	0	0	1	0	-1
$F_{[3,3]}$	10	0	0	0	0	0	0	0	0	1	-1
$F_{[6]}$	1	0	0	0	0	0	0	0	0	0	1

C Game degree determines the maximum relatedness order needed to parameterise the model

In previous work on n-player 2-strategy (strategies A and B) group games, Ohtsuki (2014) defined the degree of a game d as the maximum order of polynomial needed to write the payoffs to a focal player as a function of the number k of A-strategists among the $n-1$ nonfocal players. Ohtsuki (2014) proved that, if a game has degree d, then dynamical sufficiency can be obtained using only relatedness coefficients up to order $d+1$. For example, the payoffs in a linear PGG can be written as polynomials of degree $d=1$, and a linear PGG can be parameterised with only one relatedness coefficient, r_2, regardless of how large the group is.

In this Supplement, we sketch a generalisation of this principle to m-strategy games using worked examples. The core concept is that, if an n-player game has degree $d\,(<n-1)$, then the n-player payoffs can be written as a sum of $(d+1)$-player payoffs, and only relatedness coefficients up to order $(d+1)$, by which we mean relatedness coefficients that appear in $\mathcal{R}_{d+1}$ (see Eq. B.6), are needed to describe the dynamics.

C.1 Overview

In a game with n players and m strategies, there are $T=\frac{(n+m-2)!}{(n-1)!(m-1)!}$ possible nonfocal strategy distributions, and therefore there are up to T possible unique payoffs that a focal s_x strategist may receive. Therefore, in general and for arbitrary payoffs, the payoff function can be written as a polynomial of degree $n-1$ with T coefficients

$$\begin{aligned}&\pi_n\left(\boldsymbol{e}_x,\left(g_{\mathrm{nf},1},g_{\mathrm{nf},2},\ldots,n-1-\sum_{i=1}^{m-1}g_{\mathrm{nf},i}\right)\right)\\&\quad=c_1+c_2g_{\mathrm{nf},1}+c_3g_{\mathrm{nf},2}+\ldots+c_mg_{\mathrm{nf},m-1}+c_mg_{\mathrm{nf},1}^2+c_{m+1}g_{\mathrm{nf},1}g_{\mathrm{nf},2}+\ldots+c_Tg_{\mathrm{nf},m-1}^{n-1},\end{aligned}\tag{C.1}$$

where the subscript n on π_n emphasises that this is the payoff for an n-player game.

We observe that, if the maximum degree across focal strategies of the polynomial π_n is $d\,(<n-1)$ (we propose to call this d "degree of the game" in analogy with Ohtsuki (2014)), then the full game payoff can be written as an aggregation of games with $d+1$ players

$$\pi_n(\boldsymbol{e}_x,\boldsymbol{g}_{\mathrm{nf}})=\sum_{\boldsymbol{\gamma}_{\mathrm{nf}}\in\Gamma_{\boldsymbol{g}_{\mathrm{nf}}}}\pi_{d+1}(\boldsymbol{e}_x,\boldsymbol{\gamma}_{\mathrm{nf}}),\tag{C.2}$$

where $\Gamma_{\boldsymbol{g}_{\mathrm{nf}}}$ is the set of all d-player nonfocal strategy compositions that can be obtained by selecting d out of $n-1$ individuals from the full nonfocal strategy composition $\boldsymbol{g}_{\mathrm{nf}}$. Specifically,

$$\Gamma_{\boldsymbol{g}_{\mathrm{nf}}}=\left\{\boldsymbol{\gamma}_{\mathrm{nf}}\in\mathbb{N}_{\geq 0}^m\,\middle|\,\boldsymbol{\gamma}_{\mathrm{nf}}=\sum_{j\in\mathcal{J}}\boldsymbol{g}_j,\ \boldsymbol{g}_{\mathrm{nf}}=\sum_{j=1}^{n-1}\boldsymbol{g}_j,\ \mathcal{J}\in S_d(\{1,\ldots,n-1\})\right\},\tag{C.3}$$

where $\boldsymbol{g}_j$ are the strategy indicator functions for individual j indexed in arbitrary order, and $S_d(\mathcal{A})$ is the set of all d-sized subsets of the set $\mathcal{A}$. The size of the set is $|\Gamma_{\boldsymbol{g}_{\mathrm{nf}}}|=\binom{n-1}{d}$.

Given that the game can be reinterpreted as an aggregation of $d+1$ player games, then it is obvious that the evolutionary dynamics can be described with relatedness coefficients up to order $(d+1)$, i.e., $\mathcal{R}_{d+1}$ in Eq. B.6.

C.2 Examples

C.2.1 A three-player game that is a sum of two-player games

It is known that any n-player linear public goods game with two strategies, cooperate and defect, can be parameterised with the dyadic relatedness coefficient $r_{[2]}$ alone. In this subsection, we detail an example where a three-player can be written as a sum of 2-player games and therefore parameterised with $r_{[2]}$ alone.

Consider a 3-player linear public goods game with contribution cost c and maximum benefit $b = 1$. Let s_1 be Cooperators and s_2 be Defectors. The payoffs to Cooperators are

$$\pi_3(\boldsymbol{e}_1, 2\boldsymbol{e}_1) = 1 - c, \qquad \pi_3(\boldsymbol{e}_1, \boldsymbol{e}_1 + \boldsymbol{e}_2) = 2/3 - c, \qquad \pi_3(\boldsymbol{e}_1, 2\boldsymbol{e}_2) = 1/3 - c,$$

and the payoffs to Defectors are

$$\pi_3(\boldsymbol{e}_2, 2\boldsymbol{e}_1) = 2/3, \qquad \pi_3(\boldsymbol{e}_2, \boldsymbol{e}_1 + \boldsymbol{e}_2) = 1/3, \qquad \pi_3(\boldsymbol{e}_2, 2\boldsymbol{e}_2) = 0.$$

Both payoff functions can be written as polynomials of degree $d = 1$. For Cooperators

$$\pi_3(\boldsymbol{e}_1, (g_{\mathrm{nf},1}, 2 - g_{\mathrm{nf},1})) = \frac{1}{3} - c + \frac{g_{\mathrm{nf},1}}{3},$$

and for Defectors

$$\pi_3(\boldsymbol{e}_2, (g_{\mathrm{nf},1}, 2 - g_{\mathrm{nf},1})) = \frac{g_{\mathrm{nf},1}}{3}.$$

The degree of the game, d, is the largest order of those two polynomials above, and therefore $d = 1$. We find that the three-player games above can also be written as the sum of 2 two-player games. For both Cooperators and Defectors, we can write

$$\begin{aligned}
\pi_3(\boldsymbol{e}_x, 2\boldsymbol{e}_1) &= 2\pi_2(\boldsymbol{e}_x, \boldsymbol{e}_1), \\
\pi_3(\boldsymbol{e}_x, \boldsymbol{e}_1 + \boldsymbol{e}_2) &= \pi_2(\boldsymbol{e}_x, \boldsymbol{e}_1) + \pi_2(\boldsymbol{e}_1, \boldsymbol{e}_2), \\
\pi_3(\boldsymbol{e}_x, 2\boldsymbol{e}_2) &= 2\pi_2(\boldsymbol{e}_x, \boldsymbol{e}_2),
\end{aligned}$$

where both π_2 are polynomials of degree 1. For Cooperators

$$\pi_2(\boldsymbol{e}_1, (g_{\mathrm{nf},1}, 1 - g_{\mathrm{nf},1})) = \frac{1}{6} - \frac{c}{2} + \frac{g_{\mathrm{nf},1}}{3},$$

and for Defectors

$$\pi_2(\boldsymbol{e}_2, (g_{\mathrm{nf},1}, 1 - g_{\mathrm{nf},1})) = \frac{g_{\mathrm{nf},1}}{3}.$$

C.2.2 A four-player game that is a sum of three-player games

In this subsection, we detail an example of a 4-player game whose payoffs can be expressed as polynomials of degree $d = 2$ and as an aggregation of 3-player games. Therefore, this 4-player game's dynamics can be parameterised using only relatedness coefficients in the set $\mathcal{R}_3 = \{r_{[2]}, r_{[3]}\}$.

We consider a game with $m = 3$ strategies played between $n = 4$ players. In general, the payoff function is a polynomial of degree $d = 3$ with 10 coefficients

$$\begin{aligned}
\pi_4\big(\boldsymbol{e}_x, (g_{\mathrm{nf},1}, g_{\mathrm{nf},2}, 3 - g_{\mathrm{nf},1} - g_{\mathrm{nf},2})\big) = {} & c_1 + c_2 g_{\mathrm{nf},1} + c_3 g_{\mathrm{nf},2} + c_4 g_{\mathrm{nf},1}^2 + c_5 g_{\mathrm{nf},1} g_{\mathrm{nf},2} + c_6 g_{\mathrm{nf},2}^2 \\
& + c_7 g_{\mathrm{nf},1}^3 + c_8 g_{\mathrm{nf},1}^2 g_{\mathrm{nf},2} + c_9 g_{\mathrm{nf},1} g_{\mathrm{nf},2}^2 + c_{10} g_{\mathrm{nf},2}^3. \quad \text{(C.4)}
\end{aligned}$$

Here, we consider a special case where the payoff function is a polynomial of degree $d = 2$. We consider the payoff function with 6 coefficients

$$\pi_4\big(\boldsymbol{e}_x, (g_{\mathrm{nf},1}, g_{\mathrm{nf},2}, 3 - g_{\mathrm{nf},1} - g_{\mathrm{nf},2})\big) = 9 - 3g_{\mathrm{nf},1} + 3g_{\mathrm{nf},2} + 3g_{\mathrm{nf},1}^2 + 4g_{\mathrm{nf},1}g_{\mathrm{nf},2} + 5g_{\mathrm{nf},2}^2, \quad \text{(C.5)}$$

where these coefficients have been chosen arbitrarily but to produce integer payoffs. We believe the 4-player payoffs can be written as

$$\pi_4(\boldsymbol{e}_x, \boldsymbol{g}_{\mathrm{nf}}) = \sum_{\boldsymbol{\gamma}_{\mathrm{nf}} \in \Gamma_{\boldsymbol{g}_{\mathrm{nf}}}} \pi_3(\boldsymbol{e}_x, \boldsymbol{\gamma}_{\mathrm{nf}}),$$

and we wish to determine the value of each $\pi_3(\boldsymbol{e}_x, \boldsymbol{\gamma}_{\mathrm{nf}})$.

Table C.9 lists all possible nonfocal distributions $\boldsymbol{g}_{\mathrm{nf}}$ in the four-player game and the payoffs π_4 to the focal s_x strategist calculated according to Eq. C.5. For each four-player nonfocal distribution, three corresponding three-player nonfocal distributions $\boldsymbol{\gamma}_{\mathrm{nf}}$ are identified according to Eq. C.3. The relationship between three- and four-player game payoffs is written as an equation in the final column.

Table C.9: A list of all four-player games and the corresponding aggregation of three-player games.

	four-player game		3 three-player games			
#	$\boldsymbol{g}_{\mathrm{nf}}$	$\pi_4(\boldsymbol{e}_x, \boldsymbol{g}_{\mathrm{nf}})$	$\boldsymbol{\gamma}_{\mathrm{nf}}$	$\boldsymbol{\gamma}_{\mathrm{nf}}$	$\boldsymbol{\gamma}_{\mathrm{nf}}$	equation relating 3- and 4-player payoffs
1	$(3,0,0)$	27	$(2,0,0)$	$(2,0,0)$	$(2,0,0)$	$\pi_4 = 3\pi_3(\boldsymbol{e}_x, (2,0,0))$
2	$(0,3,0)$	63	$(0,2,0)$	$(0,2,0)$	$(0,2,0)$	$\pi_4 = 3\pi_3(\boldsymbol{e}_x, (0,2,0))$
3	$(0,0,3)$	9	$(0,0,2)$	$(0,0,2)$	$(0,0,2)$	$\pi_4 = 3\pi_3(\boldsymbol{e}_x, (0,0,2))$
4	$(2,1,0)$	31	$(2,0,0)$	$(1,1,0)$	$(1,1,0)$	$\pi_4 = \pi_3(\boldsymbol{e}_x, (2,0,0)) + 2\pi_3(\boldsymbol{e}_x, (1,1,0))$
5	$(2,0,1)$	15	$(2,0,0)$	$(1,0,1)$	$(1,0,1)$	$\pi_4 = \pi_3(\boldsymbol{e}_x, (2,0,0)) + 2\pi_3(\boldsymbol{e}_x, (1,0,1))$
6	$(1,2,0)$	43	$(0,2,0)$	$(1,1,0)$	$(1,1,0)$	$\pi_4 = \pi_3(\boldsymbol{e}_x, (0,2,0)) + 2\pi_3(\boldsymbol{e}_x, (1,1,0))$
7	$(0,2,1)$	35	$(0,2,0)$	$(0,1,1)$	$(0,1,1)$	$\pi_4 = \pi_3(\boldsymbol{e}_x, (0,2,0)) + 2\pi_3(\boldsymbol{e}_x, (0,1,1))$
8	$(1,0,2)$	9	$(0,0,2)$	$(1,0,1)$	$(1,0,1)$	$\pi_4 = \pi_3(\boldsymbol{e}_x, (0,0,2)) + 2\pi_3(\boldsymbol{e}_x, (1,0,1))$
9	$(0,1,2)$	17	$(0,0,2)$	$(0,1,1)$	$(0,1,1)$	$\pi_4 = \pi_3(\boldsymbol{e}_x, (0,0,2)) + 2\pi_3(\boldsymbol{e}_x, (0,1,1))$
10	$(1,1,1)$	21	$(1,1,0)$	$(1,0,1)$	$(0,1,1)$	$\pi_4 = \pi_3(\boldsymbol{e}_x, (1,1,0)) + \pi_3(\boldsymbol{e}_x, (1,0,1)) + \pi_3(\boldsymbol{e}_x, (0,0,1))$

To rewrite the four-player game as an aggregation of three-player games, we must solve for the three-player payoffs. For scenarios where the two nonfocals pursue the same strategy, the three-player payoffs can be solved independently from Rows 1, 2, and 3 of Table C.9

$$\pi_3(\boldsymbol{e}_x, (2,0,0)) = \frac{\pi_4(\boldsymbol{e}_x, (3,0,0))}{3} = 9,$$
$$\pi_3(\boldsymbol{e}_x, (0,2,0)) = \frac{\pi_4(\boldsymbol{e}_x, (0,3,0))}{3} = 21,$$
$$\pi_3(\boldsymbol{e}_x, (0,0,2)) = \frac{\pi_4(\boldsymbol{e}_x, (0,0,3))}{3} = 3.$$

For scenarios where the two nonfocals pursue different strategies, we can choose 3 rows from

Rows 4-9. For example, when we choose Rows 4, 5, and 7

$$\pi_3(\boldsymbol{e}_x,(1,1,0)) = \frac{\pi_4(\boldsymbol{e}_x,(2,1,0)) - \pi_3(\boldsymbol{e}_x,(2,0,0))}{2} = 11,$$
$$\pi_3(\boldsymbol{e}_x,(1,0,1)) = \frac{\pi_4(\boldsymbol{e}_x,(2,0,1)) - \pi_3(\boldsymbol{e}_x,(2,0,0))}{2} = 3,$$
$$\pi_3(\boldsymbol{e}_x,(0,1,1)) = \frac{\pi_4(\boldsymbol{e}_x,(0,2,1)) - \pi_3(\boldsymbol{e}_x,(0,2,0))}{2} = 7.$$

We can verify that the π_3 solutions above are consistent with the other rows. For example, Row 10 asserts

$$\pi_4(\boldsymbol{e}_x,(1,1,1)) = \pi_3(\boldsymbol{e}_x,(1,1,0)) + \pi_3(\boldsymbol{e}_x,(1,0,1)) + \pi_3(\boldsymbol{e}_x,(0,0,1)) = 11+3+7 = 21,$$

which is correct.

In general, if π_4 is degree $d = 2$ (i.e. $c_7 = c_8 = c_9 = c_{10} = 0$ in Eq. C.4), then the polynomial describing payoffs in three-player games is shown to be

$$\pi_3(\boldsymbol{e}_x,\boldsymbol{\gamma}_{\text{nf}}) = \frac{c_1}{3} + \frac{c_2 - c_4}{2}\gamma_{\text{nf},1} + \frac{c_3 - c_6}{2}\gamma_{\text{nf},2} + c_4\gamma_{\text{nf},1}^2 + c_5\gamma_{\text{nf},1}\gamma_{\text{nf},2} + c_6\gamma_{\text{nf},2}^2.$$

D How we implemented the other-member accounting: the transformed payoff approach

In this supplement, we provide the general method for obtaining the transformed payoff matrix (or payoff tensor), and we describe how to test stability and determine invasion fitness.

The methods described here are implemented in the `TransmatBase` class in the code repository `functions/transmat_base.py`. Examples of their analytic use in a small example can be found in SI G, and helpful functions for performing symbolic operations in `SymPy` can be found in the module `functions/symbolic_transformed.py`.

D.1 Background and small-n examples

When the relatedness coefficients are independent of strategy frequencies in the population, as they are in our approach, then the dynamics under homophily are equivalent to the dynamics in a well-mixed population of a transformed game, where the transformation modifies the payoffs in a way that accounts for relatedness (Grafen, 1979). When the payoffs are expressed in a matrix form, homophilic dynamics can be expressed in terms of operations on a transformed payoff matrix (e.g., Van Veelen, 2011; García et al., 2014).

For example, consider a 2-player, m-strategy game in a well-mixed population. We can construct a payoff matrix A such that each element $a_{i,j}$ is the payoff an i strategist receives when matched with a j strategist

$$a_{i,j} = \pi(\boldsymbol{e}_i, \boldsymbol{e}_j). \tag{D.1}$$

Then the replicator dynamics can be written

$$\dot{p}_i = p_i(\overline{\pi}_i - \overline{\pi}) = p_i((A\boldsymbol{p})_i - \boldsymbol{p}^T A\boldsymbol{p}), \tag{D.2}$$

where $\overline{\boldsymbol{\pi}} = A\boldsymbol{p}$ is a vector of expected payoffs to each focal strategist, and $\overline{\pi} = \boldsymbol{p}^T A\boldsymbol{p}$ is the expected payoff in the population as a whole (Hofbauer and Sigmund, 1998). For example, we can write out $\overline{\boldsymbol{\pi}} = A\boldsymbol{p}$ in detail

$$\begin{pmatrix} \overline{\pi}_1 \\ \vdots \\ \overline{\pi}_i \\ \vdots \\ \overline{\pi}_m \end{pmatrix} = \begin{pmatrix} a_{1,1} & \dots & a_{1,m} \\ \vdots & & \vdots \\ a_{i,1} & \dots & a_{i,m} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,m} \end{pmatrix} \begin{pmatrix} p_1 \\ \vdots \\ p_i \\ \vdots \\ p_m \end{pmatrix} = \begin{pmatrix} a_{1,1}p_1 + \ldots + a_{1,m}p_m \\ \vdots \\ a_{i,1}p_1 + \ldots + a_{i,m}p_m \\ \vdots \\ a_{m,1}p_1 + \ldots + a_{m,m}p_m \end{pmatrix}. \tag{D.3}$$

Now consider a situation with genetic homophily, with dyadic relatedness r_2. This means that an i-strategist will be matched with an i-strategist with probability r_2, and i-strategist will be matched with a strategy randomly drawn from the population with probability $1 - r_2$ (see Eq. 16). Then we can create a transformed payoff matrix B

$$B = r_2 \begin{pmatrix} a_{1,1} & \dots & a_{1,1} \\ \vdots & & \vdots \\ a_{i,i} & \dots & a_{i,i} \\ \vdots & & \vdots \\ a_{m,m} & \dots & a_{m,m} \end{pmatrix} + (1 - r_2) \begin{pmatrix} a_{1,1} & \dots & a_{1,m} \\ \vdots & & \vdots \\ a_{i,1} & \dots & a_{i,m} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,m} \end{pmatrix}, \tag{D.4}$$

such that the evolutionary dynamics of a game with payoffs B played in a well-mixed population

$$\dot{p}_i = p_i((B\boldsymbol{p})_i - \boldsymbol{p}^T B\boldsymbol{p}), \tag{D.5}$$

are equivalent to the dynamics of game with payoffs A played in the homophilic population with relatedness r_2.

3-player game. In the main text, we found the transformed payoff function π' for the 3-player m-strategy game that preserved the symmetry of the original π function and also satisfied

$$\overline{\pi}_i = \sum_{j=1}^{m} \sum_{k=1}^{m} p_j p_k \pi'(\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k). \tag{D.6}$$

The transformation was

$$\begin{aligned} \pi'(\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k) &= F_{[1,1,1]} \pi(\boldsymbol{e}_i, \boldsymbol{e}_j + \boldsymbol{e}_k) + F_{[3]} \pi(\boldsymbol{e}_i, 2\boldsymbol{e}_i) \\ &+ F_{[2,1]} \left[\frac{1}{3} \left(\frac{\pi(\boldsymbol{e}_i, 2\boldsymbol{e}_j) + \pi(\boldsymbol{e}_i, 2\boldsymbol{e}_k)}{2} \right) + \frac{2}{3} \left(\frac{\pi(\boldsymbol{e}_i, \boldsymbol{e}_i + \boldsymbol{e}_j) + \pi(\boldsymbol{e}_i, \boldsymbol{e}_i + \boldsymbol{e}_k)}{2} \right) \right]. \end{aligned} \tag{D.7}$$

The transformed "payoff matrix" is not a matrix anymore, but it may be called "payoff tensor" B, and it has size m^3 with elements

$$b_{i,j,k} = \pi'(\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k), \tag{D.8}$$

illustrated below (Fig. D.1).

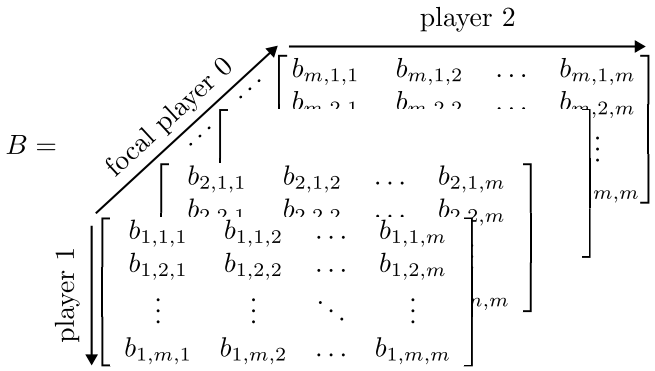


Figure D.1: The 3-dimensional payoff tensor for the 3-player game. Each dimension corresponds to a player, the strategy of the focal player is indicated by the first index, and m entries correspond to the m strategies.

Example 11. A 3-dimensional tensor $B = \{b_{i,j,k}\}$ is naturally identifiable with a function $B : \mathbb{R}^m \otimes \mathbb{R}^m \otimes \mathbb{R}^m \to \mathbb{R}$ such that for $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \boldsymbol{x}^{(3)} \in \mathbb{R}^m$ it returns a value,

$$B(\boldsymbol{x}^{(1)} \otimes \boldsymbol{x}^{(2)} \otimes \boldsymbol{x}^{(3)}) := \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{k=1}^{m} x_i^{(1)} x_j^{(2)} x_k^{(3)} b_{i,j,k}.$$

In what follows we will use the following short-hand notations,

$$\begin{aligned} B_i \boldsymbol{p}^{\otimes 2} &:= B(\boldsymbol{e}_i \otimes \boldsymbol{p} \otimes \boldsymbol{p}) \\ B \boldsymbol{p}^{\otimes 3} &:= B(\boldsymbol{p} \otimes \boldsymbol{p} \otimes \boldsymbol{p}) \end{aligned}, \tag{D.9}$$

where the first expression represents the average payoff of s_i-strategists in a well-mixed population where game B is played, and the second expression is the average payoff in the whole

population.

Similarly, for an n-player game, an n-dimensional tensor $B = \{b_{i_1,\dots,i_n}\}$ defines a game, where and it is identified with a function $B : (\mathbb{R}^m)^{\otimes n} \to \mathbb{R}$ such that for $\boldsymbol{x}^{(1)}, \dots, \boldsymbol{x}^{(n)} \in \mathbb{R}^m$ it returns a value,

$$B(\boldsymbol{x}^{(1)} \otimes \cdots \otimes \boldsymbol{x}^{(n)}) := \sum_{i_1=1}^{m} \cdots \sum_{i_n=1}^{m} x_{i_1}^{(1)} \cdots x_{i_n}^{(n)} b_{i_1,\dots,i_n}.$$

We will use the following short-hand notations,

$$\begin{aligned} B_i \boldsymbol{p}^{\otimes(n-1)} &:= B(\boldsymbol{e}_i \otimes \underbrace{\boldsymbol{p} \otimes \cdots \otimes \boldsymbol{p}}_{(n-1) \text{ times}}) \\ B\boldsymbol{p}^{\otimes n} &:= B(\underbrace{\boldsymbol{p} \otimes \boldsymbol{p} \otimes \cdots \otimes \boldsymbol{p}}_{n \text{ times}}). \end{aligned} \tag{D.10}$$

Their meanings are similar to above. ◁

By using notations that are suggested in Example 11, the evolutionary dynamics can be written analogous to the 2-dimensional case (Eq. D.5)

$$\dot{p}_i = p_i \left(\underbrace{B_i \boldsymbol{p}^{\otimes 2}}_{\text{payoff of } s_i} - \underbrace{B\boldsymbol{p}^{\otimes 3}}_{\text{population average}} \right). \tag{D.11}$$

4-player game. For the 4-player m-strategy game, we seek the transformed payoff function π' that satisfies

$$\overline{\pi}_i = \sum_{j=1}^{m} \sum_{k=1}^{m} \sum_{l=1}^{m} p_j p_k p_l \pi'(\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k, \boldsymbol{e}_l), \tag{D.12}$$

and that transformation is

$$\begin{aligned} &\pi'(\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k, \boldsymbol{e}_l) \\ &= F_{[1,1,1,1]} \pi(\boldsymbol{e}_i, \boldsymbol{e}_j + \boldsymbol{e}_k + \boldsymbol{e}_l) \\ &\quad + F_{[2,1,1]} \left[\frac{1}{2} \left(\frac{\pi(\boldsymbol{e}_i, \boldsymbol{e}_i + \boldsymbol{e}_j + \boldsymbol{e}_k) + \pi(\boldsymbol{e}_i, \boldsymbol{e}_i + \boldsymbol{e}_j + \boldsymbol{e}_l) + \pi(\boldsymbol{e}_i, \boldsymbol{e}_i + \boldsymbol{e}_k + \boldsymbol{e}_l)}{3} \right) \right. \\ &\qquad \left. + \frac{1}{2} \left(\frac{\begin{array}{l} \pi(\boldsymbol{e}_i, 2\boldsymbol{e}_j + \boldsymbol{e}_k) + \pi(\boldsymbol{e}_i, 2\boldsymbol{e}_j + \boldsymbol{e}_l) + \pi(\boldsymbol{e}_i, 2\boldsymbol{e}_k + \boldsymbol{e}_j) \\ \quad + \pi(\boldsymbol{e}_i, 2\boldsymbol{e}_k + \boldsymbol{e}_l) + \pi(\boldsymbol{e}_i, 2\boldsymbol{e}_l + \boldsymbol{e}_j) + \pi(\boldsymbol{e}_i, 2\boldsymbol{e}_l + \boldsymbol{e}_k) \end{array}}{6} \right) \right] \\ &\quad + F_{[3,1]} \left[\frac{3}{4} \left(\frac{\pi(\boldsymbol{e}_i, 2\boldsymbol{e}_i + \boldsymbol{e}_j) + \pi(\boldsymbol{e}_i, 2\boldsymbol{e}_i + \boldsymbol{e}_k) + \pi(\boldsymbol{e}_i, 2\boldsymbol{e}_i + \boldsymbol{e}_l)}{3} \right) \right. \\ &\qquad \left. + \frac{1}{4} \left(\frac{\pi(\boldsymbol{e}_i, 3\boldsymbol{e}_j) + \pi(\boldsymbol{e}_i, 3\boldsymbol{e}_k) + \pi(\boldsymbol{e}_i, 3\boldsymbol{e}_l)}{3} \right) \right] \\ &\quad + F_{[2,2]} \left[\frac{\pi(\boldsymbol{e}_i, \boldsymbol{e}_i + 2\boldsymbol{e}_j) + \pi(\boldsymbol{e}_i, \boldsymbol{e}_i + 2\boldsymbol{e}_k) + \pi(\boldsymbol{e}_i, \boldsymbol{e}_i + 2\boldsymbol{e}_l)}{3} \right] \\ &\quad + F_{[4]} \pi(\boldsymbol{e}_i, 3\boldsymbol{e}_i). \end{aligned} \tag{D.13}$$

The transformed payoff tensor B has size m^4 with elements

$$b_{i,j,k,l} = \pi'(\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k, \boldsymbol{e}_l). \tag{D.14}$$

D.2 General n-player solution

In general, the payoffs in a well-mixed population can be stored in an n-dimensional $m \times \cdots \times m$ tensor A where each dimension corresponds to an individual in the group and the index is the strategy it plays. We organise A so that the leading index refers to the focal player. The elements $a_{\boldsymbol{u}}$ of A are indexed by a vector $\boldsymbol{u} = (u_0, \ldots, u_{n-1})$ of length n (group size) where u_0 is the strategy of the focal player and elements $1 \leq u_j \leq m$ are the strategies of the nonfocal players. Therefore, each element $a_{\boldsymbol{u}}$ of A is the payoff that a u_0-strategist receives when grouped with other individuals playing strategies u_1 to u_{n-1}

$$a_{\boldsymbol{u}} = \pi\left(\boldsymbol{e}_{u_0}, \sum_{j=1}^{n-1} \boldsymbol{e}_{u_j}\right). \tag{D.15}$$

Eq. D.15 indicates that the payoff tensor A is symmetric in the sense that for any permutation $\sigma : \{1, \ldots, n-1\} \to \{1, \ldots, n-1\}$ of nonfocal-players' labels the payoff remains the same:

$$a_{\boldsymbol{u}} = a_{\boldsymbol{u}'} \quad \text{where} \quad \boldsymbol{u} = (u_0, u_1, \ldots, u_{n-1}) \text{ and } \boldsymbol{u}' = (u_0, u_{\sigma(1)}, \ldots, u_{\sigma(n-1)}). \tag{D.16}$$

For example,

$$a_{1,2,3,4} = a_{1,2,4,3} = a_{1,3,2,4} = a_{1,3,4,2} = a_{1,4,2,3} = a_{1,4,3,2}.$$

The evolutionary dynamics in a well-mixed population are

$$\dot{p}_i = p_i \left(A_i \boldsymbol{p}^{\otimes(n-1)} - A\boldsymbol{p}^{\otimes n}\right). \tag{D.17}$$

To obtain the transformed payoff tensor, we seek an n-dimensional $m \times \cdots \times m$ tensor B that is (i) a function of A, that is (ii) symmetric in the sense that for any permutation $\sigma : \{1, \ldots, n-1\} \to \{1, \ldots, n-1\}$ of nonfocal-players' labels the payoff remains the same:

$$b_{\boldsymbol{u}} = b_{\boldsymbol{u}'} \quad \text{where} \quad \boldsymbol{u} = (u_0, u_1, \ldots, u_{n-1}) \text{ and } \boldsymbol{u}' = (u_0, u_{\sigma(1)}, \ldots, u_{\sigma(n-1)}), \tag{D.18}$$

and (iii) by which the evolutionary dynamics in the homophilic game are described by

$$\dot{p}_i = p_i \left(B_i \boldsymbol{p}^{\otimes(n-1)} - B\boldsymbol{p}^{\otimes n}\right). \tag{D.19}$$

Each element of B corresponds to the transformed payoff

$$b_{\boldsymbol{u}} = \pi'\left(\boldsymbol{e}_{u_0}, \boldsymbol{e}_{u_1}, \ldots, \boldsymbol{e}_{u_{n-1}}\right). \tag{D.20}$$

In general, the elements $b_{\boldsymbol{u}}$ of B involve a combination of elements $a_{\boldsymbol{v}}$ of A whose indices $\boldsymbol{v}$ are combinations with replacement of elements of $\boldsymbol{u}$ (see examples in previous subsections). To help us specify the relationship between $\boldsymbol{v}$ and $\boldsymbol{u}$, we define a vector of indices of $\boldsymbol{u}$ called $\boldsymbol{j} = (j_0, \ldots, j_{n-1})$ with elements $j_k \in \{0, \ldots, n-1\}$, and we define $\boldsymbol{u}_{\boldsymbol{j}}$ is the vector obtained when the indices $\boldsymbol{j}$ are applied to $\boldsymbol{u}$ such that the k-th element of $\boldsymbol{u}_{\boldsymbol{j}}$ is

$$(\boldsymbol{u}_{\boldsymbol{j}})_k = u_{j_k}. \tag{D.21}$$

Example 12. If $\boldsymbol{u} = (1, 3, 5, 2)$ and $\boldsymbol{j} = (0, 0, 2, 2)$, then $\boldsymbol{u}_{\boldsymbol{j}} = (1, 1, 5, 5)$. ◁

By inspection of the examples for $n = 3$ and $n = 4$ in the previous subsection, we can see that the transformation is a sum over all possible partitions of $\boldsymbol{q}$ of n, and the elements $a_{\boldsymbol{u}_{\boldsymbol{j}}}$ involved in each term correspond to the partition structure $\boldsymbol{q}$

$$b_{\boldsymbol{u}} = \sum_{\boldsymbol{q} \vdash n} F_{\boldsymbol{q}} \left(\sum_{q_0 \in \boldsymbol{q}} \frac{q_0}{n|\mathcal{J}_{q_0,\boldsymbol{q}}|} \left(\sum_{\boldsymbol{j} \in \mathcal{J}_{q_0,\boldsymbol{q}}} a_{\boldsymbol{u}_{\boldsymbol{j}}} \right) \right). \tag{D.22}$$

We identify $\mathcal{J}_{q_0,\boldsymbol{q}}$ by inspection. The first element of $\boldsymbol{u}_{\boldsymbol{j}}$ is always equal to the first element of $\boldsymbol{u}$, i.e., $j_0 = 0$, which preserves the strategy pursued by the focal player. The value 0 is repeated q_0 times, which is the size of the focal's family. Subsequent entries of $\boldsymbol{j}$ are repetitions of unique integers between 1 and $n - 1$, where the number of repetitions corresponds to the size of a nonfocal family.

We define some notation to help us describe $\mathcal{J}_{q_0,\boldsymbol{q}}$. Let $\boldsymbol{a} = (\boldsymbol{b}, \boldsymbol{c})$ denote the operation that concatenates together vectors $\boldsymbol{b}$ and $\boldsymbol{c}$ into a single vector, i.e., $\boldsymbol{a} = (b_1, \ldots, b_{|\boldsymbol{b}|}, c_1, \ldots, c_{|\boldsymbol{c}|})$. Then the set of indices in the sum in Eq. D.22 is

$$\mathcal{J}_{q_0,\boldsymbol{q}} = \left\{ \boldsymbol{j} \in \{0, \ldots, n-1\}^n \;\middle|\; \begin{array}{l} \boldsymbol{j} = (\boldsymbol{i}_0, \boldsymbol{i}_1, \boldsymbol{i}_2, \ldots, \boldsymbol{i}_{|\boldsymbol{q}|-1}), \\ \quad \boldsymbol{i}_0 = (0, \ldots, 0), \;\; |\boldsymbol{i}_0| = q_0, \\ \quad \boldsymbol{i}_k = (i_k, \ldots, i_k), \;\; |\boldsymbol{i}_k| = q_k, \;\; \text{for} \;\; 1 \le k \le |\boldsymbol{q}| - 1, \\ \quad i_k \in \{1, \ldots, n-1\}, \;\; i_k < i_l \;\; \forall k < l, \\ \quad \left[q_0, q_1, \ldots, q_{|\boldsymbol{q}|-1}\right] = \boldsymbol{q} \text{ (equal as multisets)} \end{array} \right\}. \tag{D.23}$$

The method for obtaining the transformed payoff tensor described above is implemented in the method `create_transformed_payoff_matrix()` of the `TransmatBase` class in the code repository
functions/transmat_base.py.

Example 13. Calculate the set of indices $\mathcal{J}_{q_0,\boldsymbol{q}}$ for $\boldsymbol{q} = [2, 1, 1]$ and $q_0 = 1$, and show how it relates to the relevant term in the transformed payoff payoff function $\pi'(\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k, \boldsymbol{e}_l)$ for the 4-player game (Eq. D.13).

We have $|\boldsymbol{q}| = 3$, so we choose $q_1, q_2 \in \mathbb{N}$ such that $[q_0, q_1, q_2] = [1, q_1, q_2]$ equals $\boldsymbol{q} = [2, 1, 1]$ as multisets. $\boldsymbol{j}$ has elements

$$\boldsymbol{j} = (\boldsymbol{i}_0, \boldsymbol{i}_1, \boldsymbol{i}_2) = (\underbrace{i_0, \ldots, i_0}_{q_0 \text{ times}}, \underbrace{i_1, \ldots, i_1}_{q_1 \text{ times}}, \underbrace{i_2, \ldots, i_2}_{q_2 \text{ times}}) = (0, \underbrace{i_1, \ldots, i_1}_{q_1 \text{ times}}, \underbrace{i_2, \ldots, i_2}_{q_2 \text{ times}}).$$

because $i_0 = 0$ by definition and because $q_0 = 1$ by our assmpution.

Possible values of q's and i's that satisfy all the constraints in Eq. D.23 are $(q_1, q_2) = (2, 1), (1, 2)$ and $(i_1, i_2) = (1, 2), (1, 3), (2, 3)$. Therefore, the set of indices is

$$\mathcal{J}_{1,[2,1,1]} = \{(0, 1, 1, 2), (0, 1, 1, 3), (0, 2, 2, 3), (0, 1, 2, 2), (0, 1, 3, 3), (0, 2, 3, 3)\}.$$

When the indices $\boldsymbol{j} \in \mathcal{J}_{1,[2,1,1]}$ are applied to the indices $\boldsymbol{u} = (i, j, k, l)$ indicating elements of tensor A, the sum in Eq. D.22 becomes

$$\sum_{\boldsymbol{j} \in \mathcal{J}_{q_0, \boldsymbol{q}}} a_{\boldsymbol{u}_{\boldsymbol{j}}} = a_{i,j,j,k} + a_{i,j,j,l} + a_{i,k,k,l} + a_{i,j,k,k} + a_{i,j,l,l} + a_{i,k,l,l}.$$

This sum can be written in terms of the untransformed payoffs

$$\pi(\boldsymbol{e}_i, \boldsymbol{e}_j + \boldsymbol{e}_j + \boldsymbol{e}_k) + \pi(\boldsymbol{e}_i, \boldsymbol{e}_j + \boldsymbol{e}_j + \boldsymbol{e}_l) + \ldots + \pi(\boldsymbol{e}_i, \boldsymbol{e}_k + \boldsymbol{e}_l + \boldsymbol{e}_l),$$

which is equal to the numerator of the second coefficient of $F_{[2,1,1]}$ in the transformed payoff function in Eq. D.13. ◁

D.3 The Jacobian matrix

Let $\boldsymbol{p}^*$ be a steady-state distribution of $M \leq m$ coexisting strategies in a population. The Jacobian matrix can be used to establish the stability of $\boldsymbol{p}^*$. If all the eigenvalues of the Jacobian matrix have negative real parts, then the coexistence is stable, whereas if any eigenvalue has a positive real part, the coexistence is unstable.

Below we discuss specific local stability. Suppose that $\boldsymbol{p}^*$ consists only of the first $M(\leq m)$ strategies, $s_1, \ldots, s_M$, which means that $p_i^* > 0$ for $1 \leq i \leq M$ and $p_i^* = 0$ for $M+1 \leq i \leq m$. Then we can define a set

$$\mathrm{sub}(\boldsymbol{p}^*) := \{(p_1, \ldots, p_M, 0, \ldots, 0) \in \mathbb{R}^m \mid p_i \geq 0, \sum_{j=1}^{M} p_j = 1\}, \tag{D.24}$$

which is an $(M-1)$-dimensional sub-simplex of the $(m-1)$-dimensional simplex, $S_m := \{(p_1, \ldots, p_m) \in \mathbb{R}^m \mid p_i \geq 0, \sum_{j=1}^{m} p_j = 1\}$. Local asymptotic stability of $\boldsymbol{p}^*$ on $\mathrm{sub}(\boldsymbol{p}^*)$ is different from local asymptotic stability of $\boldsymbol{p}^*$ on the whole simplex S_m. The difference is as follows. The stability in the former sense means that any small perturbation of $\boldsymbol{p}^*$ on $\mathrm{sub}(\boldsymbol{p}^*)$, without the possibility that a new strategy $p_j > 0\,(M+1 \leq j \leq m)$ is introduced, will be eventually cancelled out by the dynamics, whereas the stability in the latter sense means that any small perturbation of $\boldsymbol{p}^*$ on S_m, including the possibility that a new strategy $s_j > 0\,(M+1 \leq j \leq m)$ is introduced, will be eventually cancelled by the dynamics. Below, we will discuss the former stability; that is, stability of $\boldsymbol{p}^*$ on $\mathrm{sub}(\boldsymbol{p}^*)$. The possibility of whether a new strategy $s_j\,(M+1 \leq j \leq m)$ can invade $\boldsymbol{p}^*$ or not is separately discussed in section D.4.

Because $\sum_{i=1}^{M} p_i = 1$, one of the dimensions of the system is redundant, and therefore the Jacobian matrix is then an $(M-1) \times (M-1)$ matrix with elements

$$J_{i,j} = \left.\frac{\partial \dot{p}_i}{\partial p_j}\right|_{\boldsymbol{p}^*} \qquad (1 \leq i, j \leq M-1).$$

We rewrite the M dynamical equations as an $M-1$ dimensional system by substituting $p_M = 1 - \sum_{i=1}^{M-1} p_i$

$$\dot{p}_i = p_i \left(\overline{\pi}_i - \sum_{k=1}^{M} p_k \overline{\pi}_k \right) = p_i \left(\overline{\pi}_i - \overline{\pi}_M - \sum_{k=1}^{M-1} p_k (\overline{\pi}_k - \overline{\pi}_M) \right). \tag{D.25}$$

The derivative of the expression in Eq. D.25 is,

$$\frac{\partial \dot{p}_i}{\partial p_i} = \left(\overline{\pi}_i - \overline{\pi}_M - \sum_{k=1}^{M-1} p_k (\overline{\pi}_k - \overline{\pi}_M) \right) + p_i \left(\frac{\partial \overline{\pi}_i}{\partial p_i} - \frac{\partial \overline{\pi}_M}{\partial p_i} - (\overline{\pi}_i - \overline{\pi}_M) - \sum_{k=1}^{M-1} p_k \left(\frac{\partial \overline{\pi}_k}{\partial p_i} - \frac{\partial \overline{\pi}_M}{\partial p_i} \right) \right),$$

and for $i \neq j$,

$$\frac{\partial \dot{p}_i}{\partial p_j} = p_i \left(\frac{\partial \overline{\pi}_i}{\partial p_j} - \frac{\partial \overline{\pi}_M}{\partial p_j} - (\overline{\pi}_j - \overline{\pi}_M) - \sum_{k=1}^{M-1} p_k \left(\frac{\partial \overline{\pi}_k}{\partial p_j} - \frac{\partial \overline{\pi}_M}{\partial p_j} \right) \right).$$

At the steady state $\boldsymbol{p}^*$, the expected fitnesses of all strategies are equal, and thus $\overline{\pi}_k - \overline{\pi}_M = 0$ for all k. Therefore, each element of the Jacobian matrix is

$$J_{i,j} = \left. \frac{\partial \dot{p}_i}{\partial p_j} \right|_{\boldsymbol{p}^*} = p_i^* \left[\left. \frac{\partial \overline{\pi}_i}{\partial p_j} \right|_{\boldsymbol{p}^*} - \sum_{k=1}^{M-1} p_k^* \left. \frac{\partial \overline{\pi}_k}{\partial p_j} \right|_{\boldsymbol{p}^*} - \left(1 - \sum_{k=1}^{M-1} p_k^* \right) \frac{\partial \overline{\pi}_M}{\partial p_j} \right] = p_i^* \left(\left. \frac{\partial \overline{\pi}_i}{\partial p_j} \right|_{\boldsymbol{p}^*} - \sum_{k=1}^{M} p_k^* \left. \frac{\partial \overline{\pi}_k}{\partial p_j} \right|_{\boldsymbol{p}^*} \right). \tag{D.26}$$

Therefore, to find the Jacobian matrix, we must obtain a general expression for the partial derivatives terms in Eq. D.26. For example, let us derive the derivatives for a 3-player game. The expected payoff is

$$\overline{\pi}_i = \sum_{j=1}^{M} \sum_{k=1}^{M} p_j p_k \, b_{i,j,k},$$

which is a function of $p_1, \ldots, p_{M-1}$ because $p_M = 1 - \sum_{j=1}^{M-1} p_j$. Writing the partial differential operator $\partial / \partial p_x \, (1 \leq x \leq M-1)$ simply as ∂_x, we have

$$\partial_x p_i = \begin{cases} 1 & (\text{if } i = x) \\ -1 & (\text{if } i = M) \\ 0 & (\text{otherwise}), \end{cases}$$

which is simply written, by using the Kronecker delta, as

$$\partial_x p_i = \delta_{ix} - \delta_{iM}, \tag{D.27}$$

where

$$\delta_{ij} = \begin{cases} 1 & (\text{if } i = j) \\ 0 & (\text{if } i \neq j). \end{cases}$$

By the chain rule, therefore, we have

$$\begin{aligned}
\partial_x \overline{\pi}_i &= \sum_{j=1}^{M}\sum_{k=1}^{M} \partial_x(p_j p_k)\, b_{i,j,k} \\
&= \sum_{j=1}^{M}\sum_{k=1}^{M} (\partial_x p_j \cdot p_k + p_j \cdot \partial_x p_k)\, b_{i,j,k} \\
&= \sum_{j=1}^{M}\sum_{k=1}^{M} \{(\delta_{jx} - \delta_{jM})p_k + p_j(\delta_{kx} - \delta_{kM})\}\, b_{i,j,k} \\
&= \sum_{k=1}^{M} p_k(b_{i,x,k} - b_{i,M,k}) + \sum_{j=1}^{M} p_j(b_{i,j,x} - b_{i,j,M}) \\
&= 2\sum_{j=1}^{M} p_j(b_{i,j,x} - b_{i,j,M}),
\end{aligned}$$

where at the last line we have used the symmetry, $b_{i,j,k} = b_{i,k,j}$.

Similarly, for a 4-player game, the expected payoff is

$$\overline{\pi}_i = \sum_{j=1}^{M}\sum_{k=1}^{M}\sum_{l=1}^{M} p_j p_k p_l\, b_{i,j,k,l}.$$

Applying the chain rule to $\partial_x(p_j p_k p_l)$ and using a similar reasoning to above lead to

$$\partial_x \overline{\pi}_i = 3\sum_{j=1}^{M}\sum_{k=1}^{M} p_j p_k (b_{i,j,k,x} - b_{i,j,k,M}).$$

In general, for n-person game the expected payoff is

$$\overline{\pi}_i = \sum_{j_1=1}^{M} \cdots \sum_{j_{n-1}=1}^{M} \left(\prod_{\kappa=1}^{n-1} p_{j_\kappa}\right) b_{i,j_1,\ldots,j_{n-1}},$$

and it is easy to see that

$$\partial_x \overline{\pi}_i = (n-1)\sum_{j_1=1}^{M} \cdots \sum_{j_{n-2}=1}^{M} \left(\prod_{\kappa=1}^{n-2} p_{j_\kappa}\right) (b_{i,j_1,\ldots,j_{n-2},x} - b_{i,j_1,\ldots,j_{n-2},M})$$

holds.

The Jacobian is calculated in the method `calc_jacobian()` of the `TransmatBase` class in the code repository `functions/transmat_base.py`.

D.4 Invasion fitness

Consider a stable coexistence between M strategies at strategy frequencies $p_1^*, p_2^*, \ldots, p_M^*$ (> 0). In order for a new strategy s_{M+1} to successfully invade the population, the fitness of s_{M+1}-strategists when rare must be higher than the average fitness of the other strategies. Let $\boldsymbol{p}^* = (p_1^*, p_2^*, \ldots, p_M^*, 0, \ldots, 0)$, where the first 0 represents the rare s_{M+1}-strategists. Then, in

order for s_{M+1} to successfully invade, its per capita growth rate must be positive:

$$\frac{1}{p_{M+1}}\dot{p}_{M+1}\bigg|_{\boldsymbol{p}^*} = \overline{\pi}_{M+1}|_{\boldsymbol{p}^*} - \sum_{j=1}^{M} p_j^* \, \overline{\pi}_j|_{\boldsymbol{p}^*} > 0.$$

At the steady state, the expected payoff of each strategy that is present in $\boldsymbol{p}^*$ is equal, and therefore it is also equal to the expected payoff in the whole population:

$$\sum_{j=1}^{M} p_j^* \, \overline{\pi}_j|_{\boldsymbol{p}^*} = \overline{\pi}_k|_{\boldsymbol{p}^*} \quad \forall k \in \{1, \ldots, M\}.$$

Therefore, the condition for successful invasion of s_{M+1} is

$$\overline{\pi}_{M+1}|_{\boldsymbol{p}^*} - \overline{\pi}_1|_{\boldsymbol{p}^*} > 0. \tag{D.28}$$

Note that $\overline{\pi}_1$ above can be any $\overline{\pi}_j$ for $j \in \{1, \ldots, M\}$, because they are all equal.

Once the transformed payoff tensor B is obtained, the expected payoffs $\overline{\pi}_j$ can be evaluated by

$$\overline{\pi}_j|_{\boldsymbol{p}^*} = B_j(\boldsymbol{p}^*)^{\otimes(n-1)}, \tag{D.29}$$

where we have used the notation of Eq. D.10, described in Example 11.

The invasion fitness is calculated in the method `calc_invasion_fitness()` of the `TransmatBase` class in the code repository `functions/transmat_base.py`.

E Comparison with kin selection literature

There has been considerable effort to understand synergistic interactions between individuals who are potentially genetically related. The most popular example in the literature about this is the two-player game with two strategies, whose payoff matrix is given by

$$M = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} B-C+D & -C \\ B & 0 \end{pmatrix}, \tag{E.1}$$

where strategy 1 is cooperation and strategy 2 is non-cooperation (Queller, 1985). Assuming an infinitely large population with global competition, and also assuming that the current frequency of cooperators in the population is given by p, our Eq. D.4 suggests that the proper transformation of this payoff matrix is

$$\begin{aligned} M' &= \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \\ &= r_2 \begin{pmatrix} B-C+D & B-C+D \\ 0 & 0 \end{pmatrix} + (1-r_2) \begin{pmatrix} B-C+D & -C \\ B & 0 \end{pmatrix} \\ &= \begin{pmatrix} B-C+D & -C+r_2(B+D) \\ (1-r_2)B & 0 \end{pmatrix}, \end{aligned} \tag{E.2}$$

and that the resulting replicator equation with this transformed payoff matrix is

$$\dot{p} = p(1-p)\left[r_2 B - C + \{r_2 + (1-r_2)p\}\, D\right] \tag{E.3}$$

(see Eq. D.5). Therefore, we conclude the sign of the square-bracketed term above,

$$\Delta \equiv r_2 B - C + \{r_2 + (1-r_2)p\}\, D, \tag{E.4}$$

determines the direction of evolution.

We, however, would like to stress that the result above is perfectly consistent with results derived by using inclusive fitness theory. In fact, Queller (1985) was the first to show that $\Delta > 0$ is the condition for evolution of cooperation. Gardner et al. (2011) have further shown that the condition, $\Delta > 0$, has a proper interpretation as a Hamilton's rule as

$$\Delta = r_2 b - c > 0 \tag{E.5}$$

where the "appropriate" b and c terms were shown to be

$$\begin{aligned} b &= B + \frac{1}{1+r_2}\{r_2 + (1-r_2)p\}\, D, \\ c &= C - \frac{1}{1+r_2}\{r_2 + (1-r_2)p\}\, D. \end{aligned} \tag{E.6}$$

Moreover, Eq. E.4 can be rewritten in a different way as

$$\Delta = r_2 B - C + \left[\frac{r_2+1}{2} + \left(p - \frac{1}{2}\right)(1-r_2)\right] D, \tag{E.7}$$

which agrees with the format of inclusive fitness effect W_{IF} of the allele that encodes strategy

1, which was shown in Eq. (3.5) in Taylor and Maciejewski (2012):

$$W_{\mathrm{IF}} = \beta B - \gamma C + \left[\frac{\beta+\gamma}{2} + \left(p - \frac{1}{2}\right)\alpha\right] D \tag{E.8}$$

with $\alpha = 1 - r_2, \beta = r_2$ and $\gamma = 1$. Furthermore, at $p = 1/2$ we have

$$\Delta|_{p=1/2} = r_2 \underbrace{\left(B + \frac{D}{2}\right)}_{\equiv \mathcal{B}} - \underbrace{\left(C - \frac{D}{2}\right)}_{\equiv \mathcal{C}}, \tag{E.9}$$

and this corresponds to the payoff transformation to an "equivalent" one for a finite population model which was suggested by Taylor (2017):

$$\mathcal{M} = \begin{pmatrix} \mathcal{B} - \mathcal{C} & -\mathcal{C} \\ \mathcal{B} & 0 \end{pmatrix}. \tag{E.10}$$

These observations confirm that our general mathematical framework for (multidimensional) matrix form games in the current paper includes many of previous arguments developed in kin-selection literature as special cases. One gap in our approach is that we did not consider any types of spatial structure of the population or local competition therein. For the corresponding analyses in kin-selection literature, rich results are already known, and we refer to Lehmann and Keller (2006), Lehmann et al. (2007), Gardner and West (2010), Ohtsuki (2010), Taylor (2013), Taylor (2016), and Taylor (2017) as examples.

F How we implemented the whole-group accounting

To implement the whole-group accounting, we used our expression for $\dot{\boldsymbol{p}}$ to find steady states numerically (SI F.1), and then for each steady state, we checked whether or not it was locally stable (SI F.2), and we determined if any of the other strategies not present in the population could invade (SI F.3). The methods described here are implemented in the `ModelBase` class in the code repository `functions/model_base.py`. An example is given in SI I.

F.1 Calculating $\dot{\boldsymbol{p}}$

In the code, we reorganised the calculation of Eq. 30 so that the combinatorial terms (components of $C(\boldsymbol{z})$ and $A(\boldsymbol{z},\boldsymbol{p})$) could be precalculated and stored in minimally sized files. In this subsection, we sketch the approach we used.

In the main text, the dynamics are written in Eq. 30 compactly as

$$\dot{p}_x = \sum_{\boldsymbol{g}_\mathrm{a}\in\mathcal{G}_\mathrm{a}} \left(\frac{g_{\mathrm{a},x}}{n}\hat{\pi}(\boldsymbol{e}_x,\boldsymbol{g}_\mathrm{a}) - p_x \sum_{i=1}^{m} \frac{g_{\mathrm{a},i}}{n}\hat{\pi}(\boldsymbol{e}_i,\boldsymbol{g}_\mathrm{a}) \right) \left(\sum_{\boldsymbol{z}\in\mathcal{Z}_{\boldsymbol{g}_\mathrm{a}}} C(\boldsymbol{z})\, A(\boldsymbol{z},\boldsymbol{p})\, F_{\mathrm{sum}(\boldsymbol{z})} \right).$$

We can split Eq. 30 as

$$\begin{aligned} \dot{p}_x = &\left(\sum_{\boldsymbol{g}_\mathrm{a}\in\mathcal{G}_\mathrm{a}^{(x)}} \hat{\pi}(\boldsymbol{e}_x,\boldsymbol{g}_\mathrm{a}) \underbrace{\sum_{\boldsymbol{z}\in\mathcal{Z}_{\boldsymbol{g}_\mathrm{a}}} \frac{g_{\mathrm{a},x}}{n} C(\boldsymbol{z})\, A(\boldsymbol{z},\boldsymbol{p})\, F_{\mathrm{sum}(\boldsymbol{z})}}_{H_x(\boldsymbol{g}_\mathrm{a},\boldsymbol{p})} \right) \\ &- \left(p_x \sum_{i=1}^{m} \sum_{\boldsymbol{g}_\mathrm{a}\in\mathcal{G}_\mathrm{a}^{(i)}} \hat{\pi}(\boldsymbol{e}_i,\boldsymbol{g}_\mathrm{a}) \underbrace{\sum_{\boldsymbol{z}\in\mathcal{Z}_{\boldsymbol{g}_\mathrm{a}}} \frac{g_{\mathrm{a},i}}{n} C(\boldsymbol{z})\, A(\boldsymbol{z},\boldsymbol{p})\, F_{\mathrm{sum}(\boldsymbol{z})}}_{H_i(\boldsymbol{g}_\mathrm{a},\boldsymbol{p})} \right), \end{aligned} \tag{F.1}$$

where the sum over all strategy compositions $\mathcal{G}_\mathrm{a}$ has been replaced with a sum over only those compositions where the corresponding strategy is present, i.e.,

$$\mathcal{G}_\mathrm{a}^{(i)} = \left\{ \boldsymbol{g}_\mathrm{a} \in \mathbb{N}_{\geq 0}^{m} \,\middle|\, \sum_{k=1}^{m} g_{\mathrm{a},k} = n \text{ and } g_{\mathrm{a},i} > 0 \right\}.$$

This reduces the number of combinatorial outcomes that must be considered.

The calculation of each $H_i(\boldsymbol{g}_\mathrm{a},\boldsymbol{p})$ has an identical structure, which we took advantage of in the code by re-indexing $\boldsymbol{p}$. For example, when calculating H_x, we re-ordered $\boldsymbol{p}$ so the focal frequency p_x was the first element in the vector and instead calculated H_1. By re-indexing $\boldsymbol{p}$, the same coefficients $C(\boldsymbol{z})$ and power terms $\|\boldsymbol{z}_i\|$ from $A(\boldsymbol{z},\boldsymbol{p})$ can be used to calculate each H_i.

We stored the coefficients and power terms corresponding to the calculation of $H_1(\boldsymbol{g}_\mathrm{a},\boldsymbol{p})$ in matrices with $|\mathcal{G}_\mathrm{a}^{(1)}|$ rows and P_n columns. Precalculated matrices can be found in the repository in `results/partn2prob/`, and their interpretation is illustrated with an example below.

Example 14. For groups of size $n = 4$ playing a game with $m = 2$ possible strategies, vector $\boldsymbol{H}_1$ that lists up $H_1(\boldsymbol{g}_\mathrm{a}, \boldsymbol{p})$ for all $\boldsymbol{g}_\mathrm{a} \in \mathcal{G}_\mathrm{a}^{(1)}$ can be written by using a matrix that stores all the pre-calculated information and a vector of the family-size distribution probabilities, named $\boldsymbol{F}$, as

$$\underbrace{\begin{bmatrix} H_1((1,3),\boldsymbol{p}) \\ H_1((2,2),\boldsymbol{p}) \\ H_1((3,1),\boldsymbol{p}) \\ H_1((4,0),\boldsymbol{p}) \end{bmatrix}}_{\boldsymbol{H}_1} = \underbrace{\frac{1}{4}}_{1/n} \underbrace{\begin{bmatrix} 4p_1\,p_2^3 & 2\,p_1\,p_2^2 & p_1\,p_2 & 0 & 0 \\ 12\,p_1^2\,p_2^2 & 2\,p_1^2\,p_2 + 2\,p_1\,p_2^2 & 0 & 4p_1\,p_2 & 0 \\ 12\,p_1^3\,p_2 & 6\,p_1^2\,p_2 & 3\,p_1\,p_2 & 0 & 0 \\ 4p_1^4 & 4p_1^3 & 4p_1^2 & 4p_1^2 & 4p_1 \end{bmatrix}}_{\text{matrix}} \underbrace{\begin{bmatrix} F_{(4,0,0,0)} \\ F_{(2,1,0,0)} \\ F_{(1,0,1,0)} \\ F_{(0,2,0,0)} \\ F_{(0,0,0,1)} \end{bmatrix}}_{\boldsymbol{F}}.$$

The coefficients of the elements in the matrix can be stored in matrix form as

```
coef = [[[4],       [2],   [1],    [],    []],
        [[12],   [2, 2],    [],   [4],    []],
        [[12],      [6],   [3],    [],    []],
        [[4],       [4],   [4],   [4],   [4]]],
```

and the powers as

```
pwrs = [[[1, 3]],          [[1, 2]],   [[1, 1]],         [],         []
        [[2, 2]],   [[2, 1], [1, 2]],         [],   [[1, 1]],         [],
        [[3, 1]],          [[2, 1]],   [[1, 1]],         [],         [],
        [[4, 0]],          [[3, 0]],   [[2, 0]],   [[2, 0]],   [[1, 0]]].
```

For example, for the sum $2p_1^2p_2 + 2p_1p_2^2$, the entry `[2, 2]` in the second row and second column of `coef` provides the coefficient for each product of powers, and `[[2, 1], [1, 2]]` in the second row and second column of `pwrs` provides the powers. The .csv file of these matrices is stored at `results/partn2prob/CW_groupsize4_nustrategies2.csv`. ◁

F.2 The Jacobian Matrix

Let $\boldsymbol{p}^*$ be a steady-state distribution of M ($\leq m$) coexisting strategies in a population. The Jacobian matrix can be used to establish the stability of $\boldsymbol{p}^*$. If all the eigenvalues of the Jacobian matrix have negative real parts, then the coexistence is stable, whereas if any eigenvalue has a positive real part, the coexistence is unstable. Note that similarly to Section D.3, here we discuss asymptotic stability of $\boldsymbol{p}^*$ on $\mathrm{sub}(\boldsymbol{p}^*)$, which is a different concept of asymptotic stability of $\boldsymbol{p}^*$ on the whole simplex, S_m. See Section D.3 for their difference.

The $(M-1) \times (M-1)$ Jacobian matrix J has elements

$$J_{i,j} = \left.\frac{\partial \dot{p}_i}{\partial p_j}\right|_{\boldsymbol{p}^*} \quad (1 \leq i, j \leq M-1),$$

where we treat p_M not as an independent variable because $p_M = 1 - \sum_{i=1}^{M-1} p_i$. If all the eigenvalues of J have negative real parts, then the steady state is stable on $\mathrm{sub}(\boldsymbol{p}^*)$.

From Eq. F.1, the dynamics of these strategies are described by

$$\dot{p}_x = \left(\sum_{\boldsymbol{g}_\mathrm{a} \in \mathcal{G}_\mathrm{a}^{(x)}} \hat{\pi}(\boldsymbol{e}_x, \boldsymbol{g}_\mathrm{a}) H_x(\boldsymbol{g}_\mathrm{a}, \boldsymbol{p}) \right) - \left(p_x \sum_{i=1}^{M} \sum_{\boldsymbol{g}_\mathrm{a} \in \mathcal{G}_\mathrm{a}^{(i)}} \hat{\pi}(\boldsymbol{e}_i, \boldsymbol{g}_\mathrm{a}) H_i(\boldsymbol{g}_\mathrm{a}, \boldsymbol{p}) \right), \tag{F.2}$$

where $\mathcal{G}_\mathrm{a}^{(i)}$ is the set of possible group strategy compositions composed of the M strategies with $g_{\mathrm{a},i} > 0$.

The diagonal elements of the Jacobian are, for $1 \leq x \leq M-1$,

$$\begin{aligned} J_{x,x} &= \left. \frac{\partial \dot{p}_x}{\partial p_x} \right|_{\boldsymbol{p}^*} \\ &= \sum_{\boldsymbol{g}_\mathrm{a} \in \mathcal{G}_\mathrm{a}^{(x)}} \hat{\pi}(\boldsymbol{e}_x, \boldsymbol{g}_\mathrm{a}) \left. \frac{\partial H_x(\boldsymbol{g}_\mathrm{a}, \boldsymbol{p})}{\partial p_x} \right|_{\boldsymbol{p}^*} - \sum_{i=1}^{M} \sum_{\boldsymbol{g}_\mathrm{a} \in \mathcal{G}_\mathrm{a}^{(i)}} \hat{\pi}(\boldsymbol{e}_i, \boldsymbol{g}_\mathrm{a}) \left(p_x \left. \frac{\partial H_i(\boldsymbol{g}_\mathrm{a}, \boldsymbol{p})}{\partial p_x} \right|_{\boldsymbol{p}^*} + H_i(\boldsymbol{g}_\mathrm{a}, \boldsymbol{p}^*) \right), \end{aligned} \tag{F.3}$$

and the off-diagonal elements are, for $x \neq y$ with $1 \leq x, y \leq M-1$,

$$\begin{aligned} J_{x,y} &= \left. \frac{\partial \dot{p}_x}{\partial p_y} \right|_{\boldsymbol{p}^*} \\ &= \sum_{\boldsymbol{g}_\mathrm{a} \in \mathcal{G}_\mathrm{a}^{(x)}} \hat{\pi}(\boldsymbol{e}_x, \boldsymbol{g}_\mathrm{a}) \left. \frac{\partial H_x(\boldsymbol{g}_\mathrm{a}, \boldsymbol{p})}{\partial p_y} \right|_{\boldsymbol{p}^*} - p_x \sum_{i=1}^{M} \sum_{\boldsymbol{g}_\mathrm{a} \in \mathcal{G}_\mathrm{a}^{(i)}} \hat{\pi}(\boldsymbol{e}_i, \boldsymbol{g}_\mathrm{a}) \left. \frac{\partial H_i(\boldsymbol{g}_\mathrm{a}, \boldsymbol{p})}{\partial p_y} \right|_{\boldsymbol{p}^*}, \end{aligned} \tag{F.4}$$

where partial derivatives of H's that appear in Eqs. F.3 and F.4 are taken by regarding H's as $(M-1)$-variable functions of $p_1, \ldots, p_{M-1}$. More explicitly, from $p_M = 1 - \sum_{i=1}^{M-1} p_i$ we have

$$H_x(\boldsymbol{g}_\mathrm{a}, \boldsymbol{p}) = \frac{g_{\mathrm{a},x}}{n} \sum_{\boldsymbol{z} \in \mathcal{Z}_{\boldsymbol{g}_\mathrm{a}}} C(\boldsymbol{z}) \underbrace{\left(\prod_{j=1}^{M-1} p_j^{\|\boldsymbol{z}_j\|} \right) \left(1 - \sum_{k=1}^{M-1} p_k \right)^{\|\boldsymbol{z}_M\|}}_{=A(\boldsymbol{z}, \boldsymbol{p})} F_{\mathrm{sum}(\boldsymbol{z})}, \tag{F.5}$$

and therefore, for all $1 \leq x, y \leq M-1$,

$$\left. \frac{\partial H_x(\boldsymbol{g}_\mathrm{a})}{\partial p_y} \right|_{\boldsymbol{p}^*} = \frac{g_{\mathrm{a},x}}{n} \sum_{\boldsymbol{z} \in \mathcal{Z}_{\boldsymbol{g}_\mathrm{a}}} C(\boldsymbol{z}) \left(\frac{\|\boldsymbol{z}_y\|}{p_y^*} - \frac{\|\boldsymbol{z}_M\|}{p_M^*} \right) \underbrace{\left(\prod_{k=1}^{M} (p_k^*)^{\|\boldsymbol{z}_k\|} \right)}_{=A(\boldsymbol{z}, \boldsymbol{p}^*)} F_{\mathrm{sum}(\boldsymbol{z})}. \tag{F.6}$$

The calculation of the Jacobian matrix is implemented in the `calc_jacobian()` method of `ModelBase`.

F.3 Invasion fitness

Consider a stable coexistence between $M(< m)$ strategies with population frequencies $p_1^*, p_2^*, \ldots, p_M^*$. In order for a new strategy s_{M+1} to successfully invade the population, the growth rate of s_{M+1}-strategists when rare must be positive. Let $\boldsymbol{p}^* = (p_1^*, p_2^*, \ldots, p_M^*, 0, \ldots, 0)$, where the first 0 represents the rare s_{M+1}-strategists. Then, in order for s_{M+1} to successfully

invade

$$\lim_{\boldsymbol{p}\to\boldsymbol{p}^*}\frac{1}{p_{M+1}}\dot{p}_{M+1}$$
$$=\underbrace{\lim_{\boldsymbol{p}\to\boldsymbol{p}^*}\left(\sum_{\boldsymbol{g}_{\mathrm{a}}\in\mathcal{G}_{\mathrm{a}}^{(M+1)}}\hat{\pi}(\boldsymbol{e}_{M+1},\boldsymbol{g}_{\mathrm{a}})\frac{H_{M+1}(\boldsymbol{g}_{\mathrm{a}},\boldsymbol{p})}{p_{M+1}}\right)}_{\text{LHT}}-\underbrace{\lim_{\boldsymbol{p}\to\boldsymbol{p}}\left(\sum_{i=1}^{M+1}\sum_{\boldsymbol{g}_{\mathrm{a}}\in\mathcal{G}_{\mathrm{a}}^{(i)}}\hat{\pi}(\boldsymbol{e}_i,\boldsymbol{g}_{\mathrm{a}})H_i(\boldsymbol{g}_{\mathrm{a}},\boldsymbol{p})\right)}_{\text{RHT}}>0$$

where the limit is taken from the positive side, $p_{M+1}>0$, while keeping $p_{M+2}=p_{M+3}=\cdots=0$, where $\mathcal{G}_{\mathrm{a}}^{(i)}$ is the set of possible group strategy compositions with $g_{\mathrm{a},i}>0$, and where

$$H_i(\boldsymbol{g}_{\mathrm{a}},\boldsymbol{p})=\sum_{\boldsymbol{z}\in\mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}}}\frac{g_{\mathrm{a},i}}{n}C(\boldsymbol{z})\,F_{\mathrm{sum}(\boldsymbol{z})}\prod_{k=1}^{M+1}p_k^{\|\boldsymbol{z}_k\|}.$$

Consider the right-hand term (RHT) first, and consider its $i=M+1$ term in the summation. Since $\boldsymbol{g}_{\mathrm{a}}\in\mathcal{G}_{\mathrm{a}}^{(M+1)}$, we know $g_{\mathrm{a},M+1}>0$, which means that any strategywise family-size distribution $\boldsymbol{z}\in\mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}}$ contains at least one family of s_{M+1}-strategists, and hence it should satisfy $\|\boldsymbol{z}_{M+1}\|\geq 1$. This means that the $(p_{M+1}^*)^{\|\boldsymbol{z}_{M+1}\|}$ term that appear in $H_{M+1}(\boldsymbol{g}_{\mathrm{a}},\boldsymbol{p}^*)$ is always zero. Therefore, this right-hand term is equivalent to the same RHT in the absence of the invader,

$$\text{RHT}=\sum_{i=1}^{M}\sum_{\boldsymbol{g}_{\mathrm{a}}\in\mathcal{G}_{\mathrm{a}}^{(i)}}\hat{\pi}(\boldsymbol{e}_i,\boldsymbol{g}_{\mathrm{a}})H_i(\boldsymbol{g}_{\mathrm{a}},\boldsymbol{p}^*),\tag{F.7}$$

which is the average payoff to an individual in the population when the invader is nonexistent.

Now consider the left-hand term (LHT). The fraction in the LHT is calculated as

$$\lim_{\boldsymbol{p}\to\boldsymbol{p}^*}\frac{H_{M+1}(\boldsymbol{g}_{\mathrm{a}},\boldsymbol{p})}{p_{M+1}}=\sum_{\boldsymbol{z}\in\mathcal{Z}_{\boldsymbol{g}_{\mathrm{a}}}}\frac{g_{\mathrm{a},M+1}}{n}C(\boldsymbol{z})\,F_{\mathrm{sum}(\boldsymbol{z})}\left[\lim_{\boldsymbol{p}\to\boldsymbol{p}^*}\left(\prod_{k=1}^{M}(p_k)^{\|\boldsymbol{z}_k\|}\right)(p_{M+1})^{\|\boldsymbol{z}_{M+1}\|-1}\right].$$

Since $\|\boldsymbol{z}_{M+1}\|\geq 1$ is always satisfied for the same reason as above, the limit inside the square brackets above always exists and it is

$$\lim_{\boldsymbol{p}\to\boldsymbol{p}^*}\left(\prod_{k=1}^{M}(p_k)^{\|\boldsymbol{z}_k\|}\right)(p_{M+1})^{\|\boldsymbol{z}_{M+1}\|-1}=\begin{cases}0 & \text{if } \|\boldsymbol{z}_{M+1}\|>1,\\ \prod_{k=1}^{M}(p_k^*)^{\|\boldsymbol{z}_k\|} & \text{if } \|\boldsymbol{z}_{M+1}\|=1.\end{cases}$$

Recall that $\|\boldsymbol{z}_{M+1}\|$ counts the number of families in the group pursuing strategy s_{M+1}. Therefore, the LHT calculates the expected payoff to an invader when it is a member of the only family in the group pursuing the invading strategy.

Our interpretation of LHT above can be compared with the more familiar situation of replicator dynamics in a well-mixed population. In a well-mixed population, the analogous left-hand term calculates the invader's expected payoff when it is the lone *individual* in the group pursuing the invading strategy instead of a member of the lone *family*. This explains why homophily can facilitate the invasion of cooperative strategies. Under homophily, cooperative individuals are typically accompanied by family members they recruited/attracted, so they fare better than in a well-mixed scenario, where they will be the lone cooperator in a group of defectors.

The calculation of invasion fitness is implemented in the `calc_invasion_fitness()` method of `ModelBase`.

G Analytic results for 3-player example

In this Supplement, we detail the analysis of the 3-player version of the sigmoid public goods game that is presented in the main text. The purpose of this Supplement is to provide an example of how the transformed payoffs can be used to analyse the dynamics under homophily.

Code to assist analytic work can be found in the `functions/symbolic_transformed.py` module in the code repository, and a quick-start tutorial using examples from this appendix can be found in `tutorials/`.

G.1 Payoff matrices

We explore the evolutionary dynamics of the same four strategies as the main text, which we index as follows:

$$
\begin{aligned}
s_1 &= \text{Unconditional Defectors 'D'},\\
s_2 &= \text{Coordinating Cooperators 'C'},\\
s_3 &= \text{Liars 'L'},\\
s_4 &= \text{Unconditional Cooperators 'U'}.
\end{aligned}
$$

The untransformed payoff tensor A is a $4 \times 4 \times 4$ tensor, where each of the 3 dimensions corresponds to each of the $n = 3$ players, and the entries 1 to 4 correspond to the strategies s_1 to s_4 (Fig. G.2).

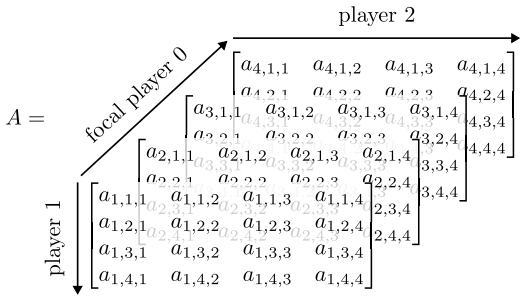


Figure G.2: The 3-dimensional payoff tensor for the 3-player game. Each dimension corresponds to a player, the strategy of the focal player is indicated by the first index, and 4 entries correspond to the four strategies.

Consider the first layer, which specifies the payoffs the focal player (player 0) will receive when it is a D-strategist. The row index corresponds to the strategy of player 1 and the column to the strategy of player 2. For example, element $a_{1,2,2} = b - \beta$ is the payoff the focal Unconditional Defector (s_1) receives when grouped with two Coordinating Cooperators (s_2), element $a_{1,4,4} = b - \beta$ is the payoff when grouped with two Unconditional Cooperators (s_4), and element $a_{1,2,3} = \beta$ is the payoff when grouped with one Coordinating Cooperator (s_2), and one Liar (s_3). In matrix form

$$
A_{1,:} = \begin{bmatrix} 0 & 0 & 0 & \beta \\ 0 & b-\beta & \beta & \beta \\ 0 & \beta & 0 & \beta \\ \beta & \beta & \beta & b-\beta \end{bmatrix}. \tag{G.1}
$$

The second layer, which specifies the payoffs to a focal C-strategist, is

$$A_{2,:} = \begin{bmatrix} -\varepsilon & b-\beta-c-\varepsilon & \beta-c-\varepsilon & \beta-\varepsilon \\ b-\beta-c-\varepsilon & b-\beta-\frac{2c}{3}-\varepsilon & \frac{b}{3}+\frac{\beta}{3}-\frac{2c}{3}-\varepsilon & b-c-\varepsilon \\ \beta-c-\varepsilon & \frac{b}{3}+\frac{\beta}{3}-\frac{2c}{3}-\varepsilon & \frac{2\beta}{3}-\frac{2c}{3}-\varepsilon & b-\beta-c-\varepsilon \\ \beta-\varepsilon & b-c-\varepsilon & b-\beta-c-\varepsilon & b-\beta-\varepsilon \end{bmatrix}. \tag{G.2}$$

The third layer, which specifies the payoffs to a focal L-strategist, is

$$A_{3,:} = \begin{bmatrix} -\varepsilon & \beta-\varepsilon & -\varepsilon & \beta-\varepsilon \\ \beta-\varepsilon & \frac{b}{3}+\frac{\beta}{3}-\varepsilon & \frac{2\beta}{3}-\varepsilon & b-\beta-\varepsilon \\ -\varepsilon & \frac{2\beta}{3}-\varepsilon & -\varepsilon & \beta-\varepsilon \\ \beta-\varepsilon & b-\beta-\varepsilon & \beta-\varepsilon & b-\beta-\varepsilon \end{bmatrix}. \tag{G.3}$$

And finally, the fourth layer, which specifies the payoffs to a focal U-strategist, is

$$A_{4,:} = \begin{bmatrix} \beta-c & \beta-c & \beta-c & b-\beta-c \\ \beta-c & b-c & b-\beta-c & b-\beta-c \\ \beta-c & b-\beta-c & \beta-c & b-\beta-c \\ b-\beta-c & b-\beta-c & b-\beta-c & b-c \end{bmatrix}. \tag{G.4}$$

To analyse the evolutionary dynamics under homophily, we analyse the dynamics in a well-mixed population with payoffs given by the transformed payoff tensor B (SI D). B is calculated using Eqs. D.22 and D.23; however, we have also written a function that automates the algebraic manipulations using the `SymPy` library: `create_transformed_payoff_matrix()`, which can be found in the `functions/symbolic_transformed.py` module in the code repository. The replicator dynamics with this transformed payoff tensor is given by Eq. D.11.

The transformed payoff tensor B is a $4 \times 4 \times 4$ matrix with 3 layers as follows. The first layer, which specifies the payoffs to a focal D-strategist, is

$$B_{1,:} = \begin{bmatrix} 0 & \frac{F_{[2,1]}(b-\beta)}{6} & 0 & F_{[1,1,1]}\beta + F_{[2,1]}\left(\frac{b}{6}+\frac{\beta}{6}\right) \\ \frac{F_{[2,1]}(b-\beta)}{6} & F_{[1,1,1]}\left(b-\beta\right) + F_{[2,1]}\left(\frac{b}{3}-\frac{\beta}{3}\right) & F_{[1,1,1]}\beta + F_{[2,1]}\left(\frac{b}{6}-\frac{\beta}{6}\right) & F_{[1,1,1]}\beta + \frac{F_{[2,1]}b}{3} \\ 0 & F_{[1,1,1]}\beta + F_{[2,1]}\left(\frac{b}{6}-\frac{\beta}{6}\right) & 0 & F_{[1,1,1]}\beta + F_{[2,1]}\left(\frac{b}{6}+\frac{\beta}{6}\right) \\ F_{[1,1,1]}\beta + F_{[2,1]}\left(\frac{b}{6}+\frac{\beta}{6}\right) & F_{[1,1,1]}\beta + \frac{F_{[2,1]}b}{3} & F_{[1,1,1]}\beta + F_{[2,1]}\left(\frac{b}{6}+\frac{\beta}{6}\right) & F_{[1,1,1]}\left(b-\beta\right) + F_{[2,1]}\left(\frac{b}{3}+\frac{\beta}{3}\right) \end{bmatrix}. \tag{G.5}$$

The second layer, which specifies the payoffs to a focal C-strategist, is

$$B_{2,:} = \begin{bmatrix} F_{[2,1]}\left(\frac{2b-2\beta-2c}{3}\right) + X_2 & F_{[1,1,1]}\left(b-\beta-c\right) + F_{[2,1]}\left(\frac{5b-5\beta-4c}{6}\right) + X_2 & F_{[1,1,1]}\left(\beta-c\right) + F_{[2,1]}\left(\frac{4b}{9}-\frac{\beta}{9}-\frac{2c}{3}\right) + X_2 & F_{[1,1,1]}\beta + F_{[2,1]}\left(\frac{5b}{6}-\frac{\beta}{2}-\frac{2c}{3}\right) + X_2 \\ F_{[1,1,1]}\left(b-\beta-c\right) + F_{[2,1]}\left(\frac{5b-5\beta-4c}{6}\right) + X_2 & F_{[1,1,1]}\left(b-\beta-\frac{2c}{3}\right) + F_{[2,1]}\left(b-\beta-\frac{2c}{3}\right) + X_2 & F_{[1,1,1]}\left(\frac{b+\beta-2c}{3}\right) + F_{[2,1]}\left(\frac{11b-5\beta-12c}{18}\right) + X_2 & F_{[1,1,1]}\left(b-c\right) + F_{[2,1]}\left(b-\frac{2\beta}{3}-\frac{2c}{3}\right) + X_2 \\ F_{[1,1,1]}\left(\beta-c\right) + F_{[2,1]}\left(\frac{4b}{9}-\frac{\beta}{9}-\frac{2c}{3}\right) + X_2 & F_{[1,1,1]}\left(\frac{b+\beta-2c}{3}\right) + F_{[2,1]}\left(\frac{11b-5\beta-12c}{18}\right) + X_2 & F_{[1,1,1]}\left(\frac{2\beta}{3}-\frac{2c}{3}\right) + F_{[2,1]}\left(\frac{2b}{9}+\frac{4\beta}{9}-\frac{2c}{3}\right) + X_2 & F_{[1,1,1]}\left(b-\beta-c\right) + F_{[2,1]}\left(\frac{11b+\beta-12c}{18}\right) + X_2 \\ F_{[1,1,1]}\beta + F_{[2,1]}\left(\frac{5b}{6}-\frac{\beta}{2}-\frac{2c}{3}\right) + X_2 & F_{[1,1,1]}\left(b-c\right) + F_{[2,1]}\left(b-\frac{2\beta}{3}-\frac{2c}{3}\right) + X_2 & F_{[1,1,1]}\left(b-\beta-c\right) + F_{[2,1]}\left(\frac{11b+\beta-12c}{18}\right) + X_2 & F_{[1,1,1]}\left(b-\beta\right) + F_{[2,1]}\left(b-\frac{\beta}{3}-\frac{2c}{3}\right) + X_2 \end{bmatrix}, \tag{G.6}$$

where $X_2 = F_{[3]}\left(b-\beta-\frac{2c}{3}\right) - \varepsilon$.

The third layer, which specifies the payoffs to a focal L-strategist, is

$$B_{3,:} = \begin{bmatrix} -\varepsilon & F_{[1,1,1]}\beta + F_{[2,1]}\left(\frac{b}{18}+\frac{5\beta}{18}\right) - \varepsilon & -\varepsilon & F_{[1,1,1]}\beta + F_{[2,1]}\left(\frac{b}{6}+\frac{\beta}{6}\right) - \varepsilon \\ F_{[1,1,1]}\beta + F_{[2,1]}\left(\frac{b}{18}+\frac{5\beta}{18}\right) - \varepsilon & F_{[1,1,1]}\left(\frac{b}{3}+\frac{\beta}{3}\right) + F_{[2,1]}\left(\frac{b}{9}+\frac{5\beta}{9}\right) - \varepsilon & F_{[1,1,1]}\left(\frac{2\beta}{3}\right) + F_{[2,1]}\left(\frac{b}{18}+\frac{5\beta}{18}\right) - \varepsilon & F_{[1,1,1]}\left(b-\beta\right) + F_{[2,1]}\left(\frac{2b}{9}+\frac{4\beta}{9}\right) - \varepsilon \\ -\varepsilon & F_{[1,1,1]}\left(\frac{2\beta}{3}\right) + F_{[2,1]}\left(\frac{b}{18}+\frac{5\beta}{18}\right) - \varepsilon & -\varepsilon & F_{[1,1,1]}\beta + F_{[2,1]}\left(\frac{b}{6}+\frac{\beta}{6}\right) - \varepsilon \\ F_{[1,1,1]}\beta + F_{[2,1]}\left(\frac{b}{6}+\frac{\beta}{6}\right) - \varepsilon & F_{[1,1,1]}\left(b-\beta\right) + F_{[2,1]}\left(\frac{2b}{9}+\frac{4\beta}{9}\right) - \varepsilon & F_{[1,1,1]}\beta + F_{[2,1]}\left(\frac{b}{6}+\frac{\beta}{6}\right) - \varepsilon & F_{[1,1,1]}\left(b-\beta\right) + F_{[2,1]}\left(\frac{b}{3}+\frac{\beta}{3}\right) - \varepsilon \end{bmatrix}. \tag{G.7}$$

And finally, the fourth layer, which specifies the payoffs to a focal U-strategist, is

$$B_{4,:} = \begin{bmatrix} F_{[1,1,1]}\left(\beta-c\right) + F_{[2,1]}\left(\frac{2b}{3}-\frac{\beta}{3}-c\right) + X_4 & F_{[1,1,1]}\left(\beta-c\right) + F_{[2,1]}\left(\frac{5b}{6}-\frac{\beta}{2}-c\right) + X_4 & F_{[1,1,1]}\left(\beta-c\right) + F_{[2,1]}\left(\frac{2b}{3}-\frac{\beta}{3}-c\right) + X_4 & F_{[1,1,1]}\left(b-\beta-c\right) + F_{[2,1]}\left(\frac{5b}{6}-\frac{\beta}{6}-c\right) + X_4 \\ F_{[1,1,1]}\left(\beta-c\right) + F_{[2,1]}\left(\frac{5b}{6}-\frac{\beta}{2}-c\right) + X_4 & F_{[1,1,1]}\left(b-c\right) + F_{[2,1]}\left(b-\frac{2\beta}{3}-c\right) + X_4 & F_{[1,1,1]}\left(b-\beta-c\right) + F_{[2,1]}\left(\frac{5b}{6}-\frac{\beta}{2}-c\right) + X_4 & F_{[1,1,1]}\left(b-\beta-c\right) + F_{[2,1]}\left(b-\frac{\beta}{3}-c\right) + X_4 \\ F_{[1,1,1]}\left(\beta-c\right) + F_{[2,1]}\left(\frac{2b}{3}-\frac{\beta}{3}-c\right) + X_4 & F_{[1,1,1]}\left(b-\beta-c\right) + F_{[2,1]}\left(\frac{5b}{6}-\frac{\beta}{2}-c\right) + X_4 & F_{[1,1,1]}\left(\beta-c\right) + F_{[2,1]}\left(\frac{2b}{3}-\frac{\beta}{3}-c\right) + X_4 & F_{[1,1,1]}\left(b-\beta-c\right) + F_{[2,1]}\left(\frac{5b}{6}-\frac{\beta}{6}-c\right) + X_4 \\ F_{[1,1,1]}\left(b-\beta-c\right) + F_{[2,1]}\left(\frac{5b}{6}-\frac{\beta}{6}-c\right) + X_4 & F_{[1,1,1]}\left(b-\beta-c\right) + F_{[2,1]}\left(b-\frac{\beta}{3}-c\right) + X_4 & F_{[1,1,1]}\left(b-\beta-c\right) + F_{[2,1]}\left(\frac{5b}{6}-\frac{\beta}{6}-c\right) + X_4 & F_{[1,1,1]}\left(b-c\right) + F_{[2,1]}\left(b-c\right) + X_4 \end{bmatrix}, \tag{G.8}$$

where $X_4 = F_{[3]}(b-c)$.

G.2 Analytic method

G.2.1 Qualitative dynamics between pairs of strategies

To gain insights into the homophilic replicator dynamics, we follow Peña et al. (2015) and apply the theory of Bernstein polynomials to the transformed game. Peña et al.'s analysis is performed between pairs of strategies A and B. They define the switching gain d_k as the payoff gain an individual would receive if they switched from being a B- to A-strategist when grouped with k A-strategists and $n-1-k$ B-strategists, which is

$$d_k = \pi(\boldsymbol{e}_A, k\boldsymbol{e}_A + (n-1-k)\boldsymbol{e}_B). \tag{G.9}$$

Then, as detailed in Peña et al. (2014), insights into the replicator dynamics between A- and B-strategists can be obtained from the sign pattern of the gain sequence $\boldsymbol{d} = [d_0, \ldots, d_{n-1}]$ (see also: Peña and Nöldeke, 2016; Nöldeke and Peña, 2016; Archetti, 2018; Peña and Nöldeke, 2018; Nöldeke and Peña, 2020; Peña et al., 2022).

G.2.2 Qualitative analysis of invasibility to a dimorphic population of coexistence pairs

We investigated the invasibility to a dimorphic population of coexisting pairs under homophily by investigating the invasibility in the transformed game in a well-mixed population (see D.4 for theory).

Consider a stable coexistence between two strategies s_1 and s_2, with strategy frequencies p_1^* and p_2^*, being invaded by a third strategy s_3. In order for s_3 to successfully invade the population, the fitness of rare s_3-strategists must higher than the average fitness in the population. Let $\boldsymbol{p}^* = (p_1^*, p_2^*, 0, 0)$, where the first zero represents s_3, and the second zero represent the other strategy s_4 that is not considered here. Then, the condition for s_3 to successfully invade is

$$\frac{1}{p_3}\dot{p}_3\bigg|_{\boldsymbol{p}^*} = \overline{\pi}_3|_{\boldsymbol{p}^*} - \sum_{j=1}^{2} p_j^*\, \overline{\pi}_j|_{\boldsymbol{p}^*} > 0.$$

At the steady state, the expected payoffs to s_1-strategists and s_2-strategists are equal and also equal to the average payoff. Therefore, the condition for s_3 to invade is

$$\overline{\pi}_3|_{\boldsymbol{p}^*} - \overline{\pi}_j|_{\boldsymbol{p}^*} > 0, \tag{G.10}$$

which can be evaluated for either $j=1$ or $j=2$. Evaluating for $j=1$ or $j=2$ are equivalent; however, as we will see below, because we have not specified $\boldsymbol{p}^*$, one or the other may give clearer analytical insights into the dynamics.

Let us choose $j=1$ for the fitness comparison. Let k be the number of s_1-strategists among the $n-1$ nonfocal players. Writing in terms of the transformed payoffs π', the expected payoff to s_1-strategists at the s_1+s_2 coexistence is

$$\overline{\pi}_1|_{\boldsymbol{p}^*} = \sum_{k=0}^{n-1} \binom{n-1}{k} (p_1^*)^k (p_2^*)^{n-1-k}\, \pi'(\boldsymbol{e}_1, \underbrace{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_1}_{k \text{ times}}, \underbrace{\boldsymbol{e}_2, \ldots, \boldsymbol{e}_2}_{(n-1-k) \text{ times}}),$$

and the expected payoff to rare invading s_3-strategists is

$$\overline{\pi}_3|_{\boldsymbol{p}^*} = \sum_{k=0}^{n-1} \binom{n-1}{k} (p_1^*)^k (p_2^*)^{n-1-k}\, \pi'(\boldsymbol{e}_3, \underbrace{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_1}_{k \text{ times}}, \underbrace{\boldsymbol{e}_2, \ldots, \boldsymbol{e}_2}_{(n-1-k) \text{ times}}),$$

Therefore, we are able to simplify the condition for s_3 to invade (Eq. G.10) as

$$\sum_{k=0}^{n-1} \binom{n-1}{k} (p_1^*)^k (p_2^*)^{n-1-k} \left\{ \pi'(\boldsymbol{e}_3, \underbrace{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_1}_{k \text{ times}}, \underbrace{\boldsymbol{e}_2, \ldots, \boldsymbol{e}_2}_{(n-1-k) \text{ times}}) - \pi'(\boldsymbol{e}_1, \underbrace{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_1}_{k \text{ times}}, \underbrace{\boldsymbol{e}_2, \ldots, \boldsymbol{e}_2}_{(n-1-k) \text{ times}}) \right\} > 0 \tag{G.11}$$

For brevity, define

$$h_k \coloneqq \pi'(\boldsymbol{e}_3, \underbrace{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_1}_{k \text{ times}}, \underbrace{\boldsymbol{e}_2, \ldots, \boldsymbol{e}_2}_{(n-1-k) \text{ times}}) - \pi'(\boldsymbol{e}_1, \underbrace{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_1}_{k \text{ times}}, \underbrace{\boldsymbol{e}_2, \ldots, \boldsymbol{e}_2}_{(n-1-k) \text{ times}}). \tag{G.12}$$

Therefore, a sufficient (but not necessary) condition for s_3 to successfully invade the s_1+s_2 coexistence is that every h_k term is positive

$$h_k > 0 \quad \forall k \in \{0, \ldots, n-1\}. \tag{G.13}$$

G.3 Main analytic results

In the process of performing the analysis, we uncovered two conditions that we deemed useful to impose as additional assumptions on the rest of the analysis. First, we found that $b - \beta - c - \varepsilon > 0$ is a necessary condition for the persistence in a well-mixed population of Coordinating Cooperators (detailed in G.4.1). We are specifically interested in the evolution of coordination, so we impose this necessary condition as an assumption. Second, the assumption $c - \beta > \varepsilon$ simplified the C vs. U pairwise analysis (G.4.5). This second assumption can be compared to the more natural assumption $c - \beta > 0$, which ensures that a lone contributor is presented with a social dilemma. The assumption $c - \beta > \varepsilon$ can be interpreted as strengthening this social-dilemma requirement.

We are particularly interested in the scenario where the level of homophily gradually changes from the ancestral state of perfect homophily, $F_{[1,1,1]} = F_{[2,1]} = 0, F_{[3]} = 1$, to the state of no homophily, $F_{[1,1,1]} = 1, F_{[2,1]} = F_{[3]} = 0$, during which $F_{[3]}$ can possibly decrease and $F_{[1,1,1]}$ can possibly increase. We do not particularly specify how $F_{[2,1]}$ changes.

We found that the evolutionary dynamics are divided into different qualitative regimes depending on two main parameter conditions (Fig. G.3). First, when the condition

$$\frac{c}{3} - \beta - \varepsilon > 0 \tag{G.14}$$

is satisfied, the ancestral state under perfect homophily is an all-C population; whereas when $\frac{c}{3} - \beta - \varepsilon < 0$, the ancestral state is all-U. In general, declining homophily facilitates the invasion of D and L (Table G.10). If the ancestral state is all-U, then under zero homophily a C+U coexistence is possible.

The second condition determined whether or not Coordinating Cooperators can resist invasion by Liars. If

$$b - 2\beta > c, \tag{G.15}$$

then Liars can never invade a C+D coexistence regardless of the homophily level. The condition in Eq. G.15 also ensures that Liars can never invade a population of all-C, and it is a necessary condition for coexistences D+U and L+U (provided ε is small) in a well-mixed population.

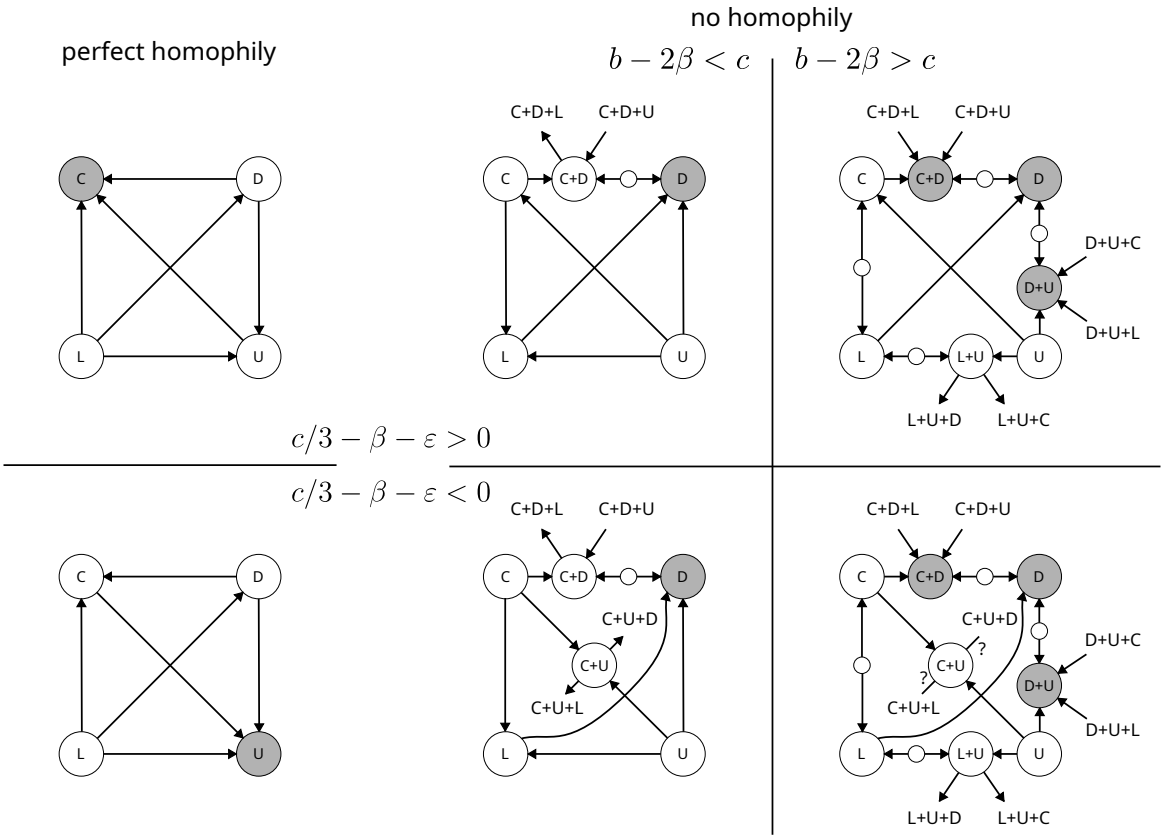


Figure G.3: A summary of the evolutionary dynamics between the four strategies, where arrows represent the evolutionary trajectories between population states, and shaded nodes indicate uninvadable evolutionary end states, subject to the assumptions described in the main text. Edges marked with a question mark indicate invasion scenarios for which we could not find a simple analytic condition. Note that under no homophily with $b - 2\beta > c$, the coexistences D+U and L+U depicted are possible but not necessarily present, depending on the parameter values. They are included here to provide a fuller accounting of the dynamics.

Table G.10: Summary of the pairwise analysis. The conditions marked with subscripts were imposed as assumptions: (1) allows Coordinated Cooperation to persist in a well-mixed population in coexistence with Defectors, and (2) ensures the social dilemma faced by the lone contributor is not arbitrarily weak.

Pair	Additional constraints	No homophily	Perfect homophily	Effect of homophily	Effect of nonlinearity
D vs. C		D can invade all-C. C cannot invade all-D.	C dominates.	D invasion impeded. C invasion facilitated.	Under no homophily, coexistence facilitated.
	$^{(1)}$ $b-\beta-c-\varepsilon>0$	Coexistence possible.			
D vs. L		D can invade all-L. L cannot invade all-D. D dominates.	D dominates.	None.	None.
D vs. U		D can invade all-U. U cannot invade all-D.	U dominates.	D invasion impeded. U invasion facilitated.	Under no homophily, coexistence facilitated.
	$b-2\beta-c>0$	Coexistence possible.			
	$b-2\beta-c<0$	D dominates.			
C vs. L		C cannot invade all-L.	C dominates.	C invasion facilitated. L invasion impeded.	Under no homophily, shifts from L-dominates to bistability regime.
	$b-2\beta-c>0$	L cannot invade all-C. Bistability.			
	$b-2\beta-c<0$	L can invade all-C. L dominates.			
C vs. U	$^{(2)}$ $c-\beta>\varepsilon$	C can invade all-U U cannot invade all-C		C invasion impeded.	Shift from $c/3-\beta-\varepsilon<0$ to $c/3-\beta-\varepsilon>0$ regime.
	$c/3-\beta-\varepsilon>0$	C dominates.	C dominates.		
	$c/3-\beta-\varepsilon<0$	Coexistence.	U dominates.		
L vs. U		L can invade all-U. U cannot invade all-L.	U dominates.	L invasion impeded. U invasion facilitated.	Under no homophily, coexistence facilitated.
	$b-2\beta-c>0$	Coexistence possible.			
	$b-2\beta-c<0$	L dominates.			

The condition in Eq. G.15 also has an intuitive interpretation: it is the condition under which contributing to the public good is in the self-interests of those who have been chosen by the lottery as contributors when there is already $\tau - 1 = 1$ contributor, because $b - 2\beta = (b - \beta) - \beta = B(2) - B(1)$ represents the marginal benefit of contribution, whereas c is its cost. In other words, the condition where lying is against one's self-interests. The condition can also be interpreted geometrically using Peña et al.'s framework: Eq. G.15 is the requirement that contributors have a positive switching gain against defectors when there are $\tau - 1$ other contributors in the group (Fig. G.4a).

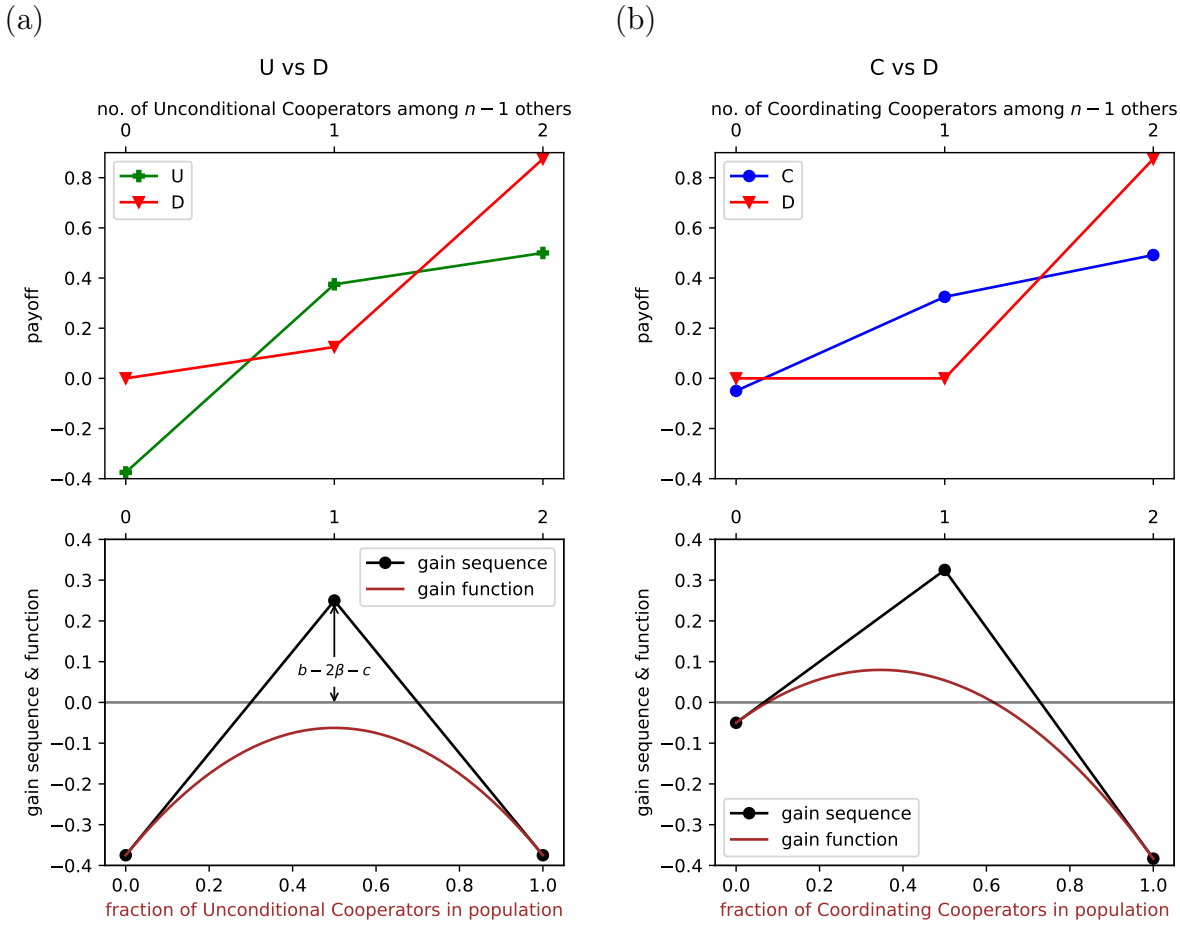


Figure G.4: An example comparing payoffs, gain sequences (see Eq. G.9), and gain functions (see Peña et al. (2014); it is defined as the difference in average payoffs between two focal strategies at given frequencies of strategies; U minus D is shown in (a), and C minus D is shown in (b)) between strategy pairs (a) Unconditional Cooperators versus Unconditional Defectors, and (b) Coordinating Cooperators versus Unconditional Defectors. Although Unconditional Cooperators are dominated by Unconditional Defectors, Coordinating Cooperators can coexist with Unconditional Defectors, and the positive switching gain $b - 2\beta - c > 0$ ensures that Liars cannot invade (see text). Parameter values: $n = 3$, $\tau = 2$, $\beta = 1/8$, $c = 0.5$, $\varepsilon = 0.05$.

The invasibility analysis of the coexistence pairs (Table G.11) also reveals some relationships that are mirrored in the results from the larger numerical example. For example, in the 3-player game, we found analytically that C can invade D+U if and only if C can invade all-D. This is observed numerically in the 8-player game, where the separatrix preventing the invasion of C into D+U appears at the same time as the separatrix on the D-C (Fig. 5).

Table G.11: Summary of the invasion analysis into coexisting pairs.

Pair	Invader	No-homophily results	Other results
D+U	C	C cannot invade D+U.	Increasing homophily facilitates C invasion. C can invade D+U iff C can invade all-D.
	L	L cannot invade D+U.	Neither nonlinearity nor homophily affects L invasion (invasion fitness is a constant $-\varepsilon$).
C+D	L	L can invade C+D iff $b-2\beta-c<0$.	If C+D can resist L invasion at zero homophily, C+D can resist L invasion at any homophily level.
	U	U cannot invade C+D.	Increasing homophily facilitates U invasion. U can invade C+D iff U can invade all-D.
C+U	D	If $b-2\beta-c<0$, D can invade C+U.	If D can invade C+U, D can invade all-C.
	L	If $b-2\beta-c<0$, L can invade C+U.	
L+U	D	D can invade L+U.	Neither nonlinearity nor homophily affects D invasion (invasion fitness is a constant ε).
	C	Provided $c-\beta>0$ and $\varepsilon\to+0$, then C can invade L+U.	

G.4 Detailed results: Qualitative dynamics between pairs of strategies

G.4.1 Unconditional Defectors versus Coordinating Cooperators

We consider the switching gains d_k from C to D, where k is the number of Defectors among the $n-1$ other group members.

Under perfect homophily ($F_{[1,1,1]}=F_{[2,1]}=0, F_{[3]}=1$)

$$d_0=d_1=d_2=-b+\beta+\frac{2c}{3}+\varepsilon<0, \tag{G.16}$$

which indicates that C dominates (Result 3.1.a Peña et al., 2014).

Under zero homophily ($F_{[1,1,1]}=1, F_{[2,1]}=F_{[3]}=0$)

$$\boldsymbol{d}=(\underbrace{\varepsilon+\frac{2c}{3}}_{(+)},\underbrace{-b+\beta+c+\varepsilon}_{(x)},\underbrace{\varepsilon}_{(+)}). \tag{G.17}$$

Split into two cases:

1. If $(x)=-b+\beta+c+\varepsilon>0$, then $\boldsymbol{d}=(+,+,+)$, which indicates that D dominates (Result 3.1.b, Peña et al., 2014).

2. If $(x) = -b + \beta + c + \varepsilon < 0$, then $\boldsymbol{d} = (+, -, +)$, which indicates that C+D coexistence is possible depending on the specific parameter values (Result 4, Peña et al., 2014).

Because we are interested in the scenario where Coordinating Cooperators can persist in a well-mixed population with Defectors, we will therefore impose $(x) < 0$ as an assumption for the remainder of the analysis:

$$\text{Assume:} \quad -b + \beta + c + \varepsilon < 0. \tag{G.18}$$

Under zero homophily, as nonlinearity increases (i.e., β decreases from $b/3$ to 0), the negative d_1 term becomes more negative while d_0 and d_2 are unchanged. Therefore, nonlinearity facilitates coexistence.

Under intermediate homophily, after applying the simplification $F_{[2,1]} = 1 - F_{[1,1,1]} - F_{[3]}$, we obtain

$$d_0 = F_{[1,1,1]} \underbrace{\left(\frac{2b}{3} - \frac{2\beta}{3}\right)}_{(+)} + F_{[3]} \underbrace{\left(-\frac{b}{3} + \frac{\beta}{3}\right)}_{(-)} - \frac{2b}{3} + \frac{2\beta}{3} + \frac{2c}{3} + \varepsilon \tag{G.19}$$

$$d_1 = F_{[1,1,1]} \underbrace{\left(-\frac{b}{3} + \frac{\beta}{3} + \frac{c}{3}\right)}_{(-)} + F_{[3]} \underbrace{\left(-\frac{b}{3} + \frac{\beta}{3}\right)}_{(-)} - \frac{2b}{3} + \frac{2\beta}{3} + \frac{2c}{3} + \varepsilon \tag{G.20}$$

$$d_2 = F_{[1,1,1]} \underbrace{\left(\frac{2b}{3} - \frac{2\beta}{3} - \frac{2c}{3}\right)}_{(+)} + F_{[3]} \underbrace{\left(-\frac{b}{3} + \frac{\beta}{3}\right)}_{(-)} - \frac{2b}{3} + \frac{2\beta}{3} + \frac{2c}{3} + \varepsilon \tag{G.21}$$

Increasing homophily decreases $F_{[1,1,1]}$ and/or increases $F_{[3]}$, so the effect of increasing homophily is to:

- decrease d_0, which implies the invasion of D into all-C is impeded, and
- decrease d_2, which implies the invasion of C into all-D is facilitated.

G.4.2 Unconditional Defectors versus Liars

The switching gains d_k from L to D, where k is the number of Defectors among the $n-1$ other group members, are all equal

$$d_0 = d_1 = d_2 = \varepsilon > 0. \tag{G.22}$$

The switching gain sign pattern $\boldsymbol{d} = (+, +, +)$ indicates that D dominates (Result 3.1.b Peña et al., 2014). The switching gains are independent of the family partition structure probabilities $F_{\boldsymbol{q}}$; therefore, the dynamics are unaffected by homophily. The switching gains are also independent of β; therefore, the dynamics are unaffected by the degree of nonlinearity of the benefits function.

G.4.3 Unconditional Defectors versus Unconditional Cooperators

We consider the switching gains d_k from U to D, where k is the number of Defectors among the $n-1$ other group members.

Under perfect homophily ($F_{[1,1,1]} = F_{[2,1]} = 0, F_{[3]} = 1$)

$$d_0 = -b + c < 0, \tag{G.23}$$

which indicates that U dominates (Result 3.1.a Peña et al., 2014).

Under zero homophily ($F_{[1,1,1]} = 1, F_{[2,1]} = F_{[3]} = 0$)

$$\boldsymbol{d} = (\underbrace{-\beta + c}_{(+)}, \underbrace{-b + 2\beta + c}_{-(g)}, \underbrace{-\beta + c}_{(+)}). \tag{G.24}$$

Two cases depending on the sign of (g):

1. If $b - 2\beta - c < 0$, the sign pattern is $\boldsymbol{d} = (+, +, +)$, and D dominates (Result 3.1.b, Peña et al., 2014).

2. If $b - 2\beta - c > 0$, the sign pattern is $\boldsymbol{d} = (+, -, +)$, and coexistence U+D is possible depending on the specific parameter values (Result 4, Peña et al., 2014).

Under zero homophily, as nonlinearity increases (i.e., β decreases from $b/3$ to 0), coexistence is facilitated. In a linear game, $b = 3\beta$, then $d_0 = d_1 = d_2 = -\beta + c > 0$, which means D dominates. However, in a threshold game, $\beta = 0$, then $\boldsymbol{d} = (+, -, +)$, which means U+D coexistence is possible.

Under intermediate homophily, after applying the simplification $F_{[2,1]} = 1 - F_{[1,1,1]} - F_{[3]}$, we obtain

$$d_0 = F_{[1,1,1]} \underbrace{\left(\frac{2b}{3} - \frac{4\beta}{3}\right)}_{(+)} + F_{[3]} \underbrace{\left(-\frac{b}{3} - \frac{\beta}{3}\right)}_{(-)} - \frac{2b}{3} + \frac{\beta}{3} + c, \tag{G.25}$$

$$d_1 = F_{[1,1,1]} \underbrace{\left(-\frac{b}{3} + \frac{5\beta}{3}\right)}_{(+/-)} + F_{[3]} \underbrace{\left(-\frac{b}{3} - \frac{\beta}{3}\right)}_{(-)} - \frac{2b}{3} + \frac{\beta}{3} + c, \tag{G.26}$$

$$d_2 = F_{[1,1,1]} \underbrace{\left(\frac{2b}{3} - \frac{4\beta}{3}\right)}_{(+)} + F_{[3]} \underbrace{\left(-\frac{b}{3} - \frac{\beta}{3}\right)}_{(-)} - \frac{2b}{3} + \frac{\beta}{3} + c. \tag{G.27}$$

Increasing homophily decreases $F_{[1,1,1]}$ and/or increases $F_{[3]}$, so the effect of increasing homophily is to:

- decrease d_0, which implies the invasion of D into all-U is impeded, and
- decrease d_2, which implies the invasion of U into all-D is facilitated.

Under intermediate homophily in a linear game ($\beta = b/3$), all switching gains are equal. Increasing homophily uniformly decreases their value. Therefore, there is a sudden transition from a D-dominates to U-dominates regime.

Under intermediate homophily in a threshold game ($\beta = 0$),

$$d_1 = F_{[1,1,1]} \underbrace{\left(\frac{2b}{3}\right)}_{(+)} + F_{[3]} \underbrace{\left(-\frac{b}{3}\right)}_{(-)} - \frac{2b}{3} + c, \tag{G.28}$$

$$d_2 = F_{[1,1,1]} \underbrace{\left(-\frac{b}{3}\right)}_{(-)} + F_{[3]} \underbrace{\left(-\frac{b}{3}\right)}_{(-)} - \frac{2b}{3} + c, \tag{G.29}$$

$$d_3 = F_{[1,1,1]} \underbrace{\left(\frac{2b}{3}\right)}_{(+)} + F_{[3]} \underbrace{\left(-\frac{b}{3}\right)}_{(-)} - \frac{2b}{3} + c. \tag{G.30}$$

Increasing homophily decreases $F_{[1,1,1]}$ and/or increases $F_{[3]}$, so the effect of increasing homophily is to:

- decrease d_0, which implies the invasion of D into all-U is impeded, and
- decrease d_2, which implies the invasion of U into all-D is facilitated.

G.4.4 Coordinating Cooperators versus Liars

We consider the switching gains d_k from L to C, where k is the number of Coordinating Cooperators among the $n-1$ other group members.

Under perfect homophily ($F_{[1,1,1]} = F_{[2,1]} = 0, F_{[3]} = 1$)

$$d_0 = d_1 = d_2 = b - \beta - \frac{2c}{3} > 0, \tag{G.31}$$

which indicates that C dominates (Result 3.1.b, Peña et al., 2014).

Under zero homophily ($F_{[1,1,1]} = 1, F_{[2,1]} = F_{[3]} = 0$)

$$\boldsymbol{d} = (1/3)(\underbrace{2\beta - 2c}_{(-)}, \underbrace{b - \beta - 2c}_{(f)}, \underbrace{2b - 4\beta - 2c}_{2\times(g)}), \tag{G.32}$$

where $(g) > 0$ is a necessary condition for the coexistence U+D (G.4.3 above).

Note that $(g) = b - \beta + (f)$ where $b - \beta > 0$; therefore, $(g) > (f)$. Therefore, we can split into two cases depending on the sign of (g):

1. If $b - 2\beta - c < 0$, the switching gain sign pattern is or $\boldsymbol{d} = (-,-,-)$, which indicates that C cannot invade all-L, L can invade all-C, and L dominates (Result 3.1.a, Peña et al., 2014).

2. If $b - 2\beta - c > 0$, the switching gain sign pattern is either $\boldsymbol{d} = (-,-,+)$ or $\boldsymbol{d} = (-,+,+)$, which indicates that C cannot invade all-L, L cannot invade all-C, and there is bistability between them (Result 3.2.a, Peña et al., 2014).

Under zero homophily, as nonlinearity increases (i.e., β decreases from $b/3$ to 0), the term marked (g) switches from a negative to positive value. Therefore, the effect of increasing nonlinearity in the benefits function is to transition the dynamics from the L-dominates regime to the bistability regime.

Under intermediate homophily, after applying the simplification $F_{[2,1]} = 1 - F_{[1,1,1]} - F_{[3]}$, we obtain

$$d_0 = F_{[1,1,1]} \underbrace{\left(-\frac{2b}{9} + \frac{2\beta}{9}\right)}_{(-)} + F_{[3]} \underbrace{\left(\frac{7b}{9} - \frac{13\beta}{9}\right)}_{(+) \text{ because } \beta \leq b/3} + \frac{2b}{9} + \frac{4\beta}{9} - \frac{2c}{3}, \tag{G.33}$$

$$d_1 = F_{[1,1,1]} \underbrace{\left(-\frac{2b}{9} + \frac{2\beta}{9}\right)}_{(-)} + F_{[3]} \underbrace{\left(\frac{4b}{9} - \frac{4\beta}{9}\right)}_{(+)} + \frac{5b}{9} - \frac{5\beta}{9} - \frac{2c}{3}, \tag{G.34}$$

$$d_2 = F_{[1,1,1]} \underbrace{\left(-\frac{2b}{9} + \frac{2\beta}{9}\right)}_{(-)} + F_{[3]} \underbrace{\left(\frac{b}{9} + \frac{5\beta}{9}\right)}_{(+)} + \frac{8b}{9} - \frac{14\beta}{9} - \frac{2c}{3}. \tag{G.35}$$

Increasing homophily decreases $F_{[1,1,1]}$ and/or increases $F_{[3]}$, so the effect of increasing homophily is to:

- increase d_0, which implies the invasion of C into all-L is facilitated,
- increase d_2, which implies the invasion of L into all-C is impeded.

Under intermediate homophily in a linear game ($\beta = b/3$), all switching gain terms are equal

$$d_i = -\frac{4F_{[1,1,1]}\beta}{9} + \frac{8F_{[3]}\beta}{9} + \frac{10\beta}{9} - \frac{2c}{3}. \tag{G.36}$$

Increasing homophily decreases $F_{[1,1,1]}$ and/or increases $F_{[3]}$, so the effect of increasing homophily is to decrease all d_i, which implies a sudden switch from an L-dominates to C-dominates regime.

Under intermediate homophily in a threshold game ($\beta = 0$), the signs of the coefficients in front of $F_{[1,1,1]}$ and $F_{[3]}$ are unchanged from Eqs. G.33 to G.35; therefore, the qualitative effects of increasing homophily are the same:

- the invasion of C into all-L is facilitated, and
- the invasion of L into all-C is impeded.

G.4.5 Coordinating Cooperators versus Unconditional Cooperators

We consider the switching gains d_k from U to C, where k is the number of Coordinating Cooperators among the $n-1$ other group members.

Under perfect homophily ($F_{[1,1,1]} = F_{[2,1]} = 0, F_{[3]} = 1$)

$$d_0 = d_1 = d_2 = \frac{c}{3} - \beta - \varepsilon, \tag{G.37}$$

which splits into two cases:

1. if $\frac{c}{3} - \beta - \varepsilon > 0$, C dominates; else
2. if $\frac{c}{3} - \beta - \varepsilon < 0$, U dominates.

Under perfect homophily, as nonlinearity increases (i.e., β decreases from $b/3$ to 0), the system switches from a U-dominates to C-dominates regime.

Under zero homophily ($F_{[1,1,1]} = 1, F_{[2,1]} = F_{[3]} = 0$)

$$\boldsymbol{d} = (-\beta + c - \varepsilon, \beta - \varepsilon, \underbrace{\frac{c}{3} - \beta - \varepsilon}_{(h)}). \tag{G.38}$$

Let us impose the condition $c - \beta > \varepsilon$. We previously assumed $c > \beta$, which means that a lone contributor is faced with a social dilemma. The condition $c - \beta > \varepsilon$ strengthens this requirement so that the social dilemma is not arbitrarily weak.

First, consider the case where $\beta > \varepsilon$, i.e., a benefits function that is not arbitrarily close to a threshold game. Then we can split the dynamics into two cases:

1. if $\frac{c}{3} - \beta - \varepsilon > 0$, $\boldsymbol{d} = (+, +, +)$, C dominates (Result 3.1b, Peña et al., 2014); else,
2. if $\frac{c}{3} - \beta - \varepsilon < 0$, $\boldsymbol{d} = (+, +, -)$, C+U coexist (Result 3.2b, Peña et al., 2014).

Now let us consider the threshold game ($\beta = 0$). Then $\boldsymbol{d} = (+, -, +)$. Technically, this sign pattern allows for the possibility of two interior equilibria (Result 5, Peña et al., 2014). However, because ε is very small, the gain function will not cross the zero axis, and therefore C will dominate.

From the analysis above, **under zero homophily, as nonlinearity increases** (i.e., β decreases from $b/3$ to 0), the dynamics switch from a C+U coexistence to C-dominates regime.

Under intermediate homophily, after applying the simplification $F_{[2,1]} = 1 - F_{[1,1,1]} - F_{[3]}$, we obtain

$$d_0 = F_{[1,1,1]} \underbrace{\left(-\frac{2\beta}{3} + \frac{2c}{3}\right)}_{(+)} + F_{[3]} \underbrace{\left(-\frac{2\beta}{3}\right)}_{(-)} \underbrace{-\frac{\beta}{3} + \frac{c}{3} - \varepsilon}_{(+)}, \tag{G.39}$$

$$d_1 = F_{[1,1,1]} \underbrace{\left(\frac{4\beta}{3} - \frac{c}{3}\right)}_{(+) \text{ if } h<0; \text{ else } (+/-)} + F_{[3]} \underbrace{\left(-\frac{2\beta}{3}\right)}_{(-)} \underbrace{-\frac{\beta}{3} + \frac{c}{3} - \varepsilon}_{(+)}, \tag{G.40}$$

$$d_2 = \underbrace{(F_{[1,1,1]} + F_{[3]})}_{(1-F_{[2,1]})} \left(-\frac{2\beta}{3}\right) \underbrace{-\frac{\beta}{3} + \frac{c}{3} - \varepsilon}_{(+)}. \tag{G.41}$$

Increasing homophily decreases $F_{[1,1,1]}$ and/or increases $F_{[3]}$, so the effect of increasing homophily is to decrease d_0, which implies that the invasion of C into all-U is impeded.

For the three group-formation models that we consider (see Sec. 2.3 in the main text), $F_{[2,1]}$ has a hump-shaped relationship with our (non-)homophily parameter, h or α. Therefore, d_2 has a v-shaped relationship with homophily, which implies that, as homophily increases, the invasion of U into all-C is initially facilitated but then impeded.

Under intermediate homophily in a threshold game ($\beta = 0$), $d_0 > 0$ and $d_2 = c/3 - \varepsilon > 0$; therefore, C can invade all-U and U cannot invade all-C.

G.4.6 Liars versus Unconditional Cooperators

We consider the switching gains d_k from U to L, where k is the number of Liars among the $n - 1$ other group members.

Under perfect homophily ($F_{[1,1,1]} = F_{[2,1]} = 0, F_{[3]} = 1$)

$$d_0 = d_1 = d_2 = -b + c - \varepsilon < 0, \tag{G.42}$$

which indicates that U dominates (Result 3.1.a, Peña et al., 2014).

Under zero homophily ($F_{[1,1,1]} = 1, F_{[2,1]} = F_{[3]} = 0$), assuming that $c - \beta > \varepsilon$ (see G.4.5), then

$$\boldsymbol{d} = (\underbrace{-\beta + c - \varepsilon}_{(+)}, \underbrace{-b + 2\beta + c}_{-(g)} - \varepsilon, \underbrace{-\beta + c - \varepsilon}_{(+)}), \tag{G.43}$$

which splits into two cases:

1. If $(g) > 0$, $\boldsymbol{d} = (+, -, +)$, the switching gain sign pattern is $\boldsymbol{d} = (+, -, +)$, which indicates that L can invade all-U, U cannot invade all-L, and U+L coexistence is possible (Result 4.2, Peña et al., 2014).

2. If $(g) < 0$, $\boldsymbol{d} = (+, +, +)$, the switching gain sign pattern is $\boldsymbol{d} = (+, +, +)$, which indicates L dominates (Result 3.1.a, Peña et al., 2014).

Note that the term marked (g) above is also involved in the dynamics of U vs. D and C vs. L. Specifically, under zero homophily, $(g) > 0$ is: a necessary condition for the coexistence U+D (G.5.3); and a necessary and sufficient condition for all-C to resist invasion by L (G.4.4).

Under zero homophily, as nonlinearity increases (i.e., β decreases from $b/3$ to 0), the dynamics can potentially switch from an L-dominates to a U+L coexistence regime.

Under intermediate homophily, after applying the simplification $F_{[2,1]} = 1 - F_{[1,1,1]} - F_{[3]}$, we obtain

$$d_0 = F_{[1,1,1]} \underbrace{\left(\frac{2b}{3} - \frac{4\beta}{3}\right)}_{(+)} + F_{[3]} \underbrace{\left(-\frac{b}{3} - \frac{\beta}{3}\right)}_{(-)} - \frac{2b}{3} + \frac{\beta}{3} + c - \varepsilon, \tag{G.44}$$

$$d_1 = F_{[1,1,1]} \underbrace{\left(-\frac{b}{3} + \frac{5\beta}{3}\right)}_{(+/-)} + F_{[3]} \underbrace{\left(-\frac{b}{3} - \frac{\beta}{3}\right)}_{(-)} - \frac{2b}{3} + \frac{\beta}{3} + c - \varepsilon, \tag{G.45}$$

$$d_2 = F_{[1,1,1]} \underbrace{\left(\frac{2b}{3} - \frac{4\beta}{3}\right)}_{(+)} + F_{[3]} \underbrace{\left(-\frac{b}{3} - \frac{\beta}{3}\right)}_{(-)} - \frac{2b}{3} + \frac{\beta}{3} + c - \varepsilon. \tag{G.46}$$

Increasing homophily decreases $F_{[1,1,1]}$ and/or increases $F_{[3]}$, so the effect of increasing homophily is to

- decrease d_0, which implies the invasion of L into all-U is impeded,
- decrease d_1, which implies the invasion of U into all-L is facilitated.

Under intermediate homophily in a linear game ($\beta = b/3$), all switching gains are equal. Increasing homophily uniformly decreases their value. Therefore, there is a sudden transition from an L-dominates regime to a U-dominates regime.

Under intermediate homophily in a threshold game ($\beta = 0$),

$$d_1 = F_{[1,1,1]} \underbrace{\left(\frac{2b}{3}\right)}_{(+)} + F_{[3]} \underbrace{\left(-\frac{b}{3}\right)}_{(-)} - \frac{2b}{3} + c - \varepsilon, \tag{G.47}$$

$$d_2 = F_{[1,1,1]} \underbrace{\left(-\frac{b}{3}\right)}_{(-)} + F_{[3]} \underbrace{\left(-\frac{b}{3}\right)}_{(-)} - \frac{2b}{3} + c - \varepsilon, \tag{G.48}$$

$$d_3 = F_{[1,1,1]} \underbrace{\left(\frac{2b}{3}\right)}_{(+)} + F_{[3]} \underbrace{\left(-\frac{b}{3}\right)}_{(-)} - \frac{2b}{3} + c - \varepsilon. \tag{G.49}$$

Increasing homophily decreases $F_{[1,1,1]}$ and/or increases $F_{[3]}$, so the effect of increasing homophily is to:

- decrease d_0, which implies the invasion of L into all-U is impeded, and
- decrease d_2, which implies the invasion of U into all-L is facilitated.

G.5 Detailed results: Qualitative analysis of invasibility of coexistence pairs

In the previous subsection, we found that stable coexistence was possible between some pairs of strategies. In this section, we investigate whether or not pairs of strategies in a stable coexistence can be invaded by the other strategies.

G.5.1 Coordinating Cooperators + Unconditional Cooperators coexistence

Invasion of Defectors into C+U. Under zero homophily, choosing U for the fitness comparison,

$$h_0 = c - \beta > 0, \tag{G.50}$$

$$h_1 = -b + 2\beta + c = -(g), \tag{G.51}$$

$$h_2 = c - \beta > 0. \tag{G.52}$$

(see Eq. G.12 for the definition of h_k). Remember that $(g) > 0$ is, as derived above, a necessary and sufficient condition for all-C to resist invasion by L (G.4.4). Therefore, if L can invade all-C, then $(g) < 0$, then $h_0, h_1, h_2 > 0$, then D can invade C+U.

Under intermediate homophily, choosing C for the fitness comparison

$$h_0 = F_{[1,1,1]} \underbrace{\varepsilon}_{(+)} + F_{[2,1]} \underbrace{\left(-\frac{2b}{3} + \frac{2\beta}{3} + \frac{2c}{3} + \varepsilon\right)}_{(-)} + F_{[3]} \underbrace{\left(-b + \beta + \frac{2c}{3} + \varepsilon\right)}_{(-)}, \tag{G.53}$$

$$h_1 = F_{[1,1,1]} \underbrace{(-b + \beta + c + \varepsilon)}_{(-)} + F_{[2,1]} \underbrace{\left(-\frac{2b}{3} + \frac{2\beta}{3} + \frac{2c}{3} + \varepsilon\right)}_{(-)} + F_{[3]} \underbrace{\left(-b + \beta + \frac{2c}{3} + \varepsilon\right)}_{(-)} \tag{G.54}$$

$$h_2 = F_{[1,1,1]} \underbrace{\left(\frac{2c}{3} + \varepsilon\right)}_{(+)}, + F_{[2,1]} \underbrace{\left(-\frac{2b}{3} + \frac{2\beta}{3} + \frac{2c}{3} + \varepsilon\right)}_{(-)} + F_{[3]} \underbrace{\left(-b + \beta + \frac{2c}{3} + \varepsilon\right)}_{(-)}. \tag{G.55}$$

When no groups contain no family members (i.e., $F_{[1,1,1]} = 0, F_{[2,1]} + F_{[3]} = 1$), D cannot invade C+U. As $F_{[1,1,1]}$ increases, h_2 switches from negative to positive before h_0. Therefore, all-C becomes invadable to D before C+U. Therefore, a necessary condition for D to invade C+U is D can invade all-C.

Invasion of Liars into C+U. Under zero homophily, choosing C for the fitness comparison

$$h_0 = 0, \tag{G.56}$$
$$h_1 = c - \beta > 0, \tag{G.57}$$
$$h_2 = \frac{-2b + 4\beta + 2c}{3} = -2\frac{(g)}{3}. \tag{G.58}$$

Remember that $(g) > 0$ is, as derived above, a necessary and sufficient condition for all-C to resist invasion by L (G.4.4). Therefore, if L can invade all-C, then $(g) < 0$, then $h_0, h_1, h_2 > 0$, then L can invade C+U.

Choosing U for the fitness comparison

$$h_0 = -\frac{2b}{3} + \frac{\beta}{3} + c - \varepsilon, \tag{G.59}$$
$$h_1 = c - \varepsilon > 0, \tag{G.60}$$
$$h_2 = -\beta + c - \varepsilon > 0. \tag{G.61}$$

Therefore, a sufficient condition for L to invade C+U is

$$-2b + \beta + 3c - 3\epsilon > 0. \tag{G.62}$$

Under intermediate homophily, no obvious insights were found.

G.5.2 Coordinating Cooperators + Unconditional Defectors coexistence

Invasion of Liars into C+D. Under zero homophily, choosing C for the fitness comparison, the invasion fitness coefficients for rare L invading the C+D coexistence are

$$h_0 = 0, \tag{G.63}$$
$$h_1 = \underbrace{-b + 2\beta + c}_{-(g)}, \tag{G.64}$$
$$h_2 = \underbrace{-\frac{2b}{3} + \frac{4\beta}{3} + \frac{2c}{3}}_{-2(g)/3}, \tag{G.65}$$

where $(g) > 0$ is, as derived above, a necessary and sufficient condition for all-C to resist invasion by L (G.4.4). Therefore, under zero homophily, C+D can resist invasion by Liars if and only if all-C can resist invasion by Liars.

Under intermediate homophily

$$h_0 = F_{[2,1]} \underbrace{\left(-\frac{2b}{3} + \frac{2\beta}{3} + \frac{2c}{3}\right)}_{(-)} + F_{[3]} \underbrace{\left(-b + \beta + \frac{2c}{3}\right)}_{(-)}, \tag{G.66}$$

$$h_1 = F_{[2,1]} \underbrace{\left(\frac{2b}{9} - \frac{8\beta}{9} - \frac{c}{3}\right)}_{(x)} + \underbrace{F_{[3]} \left(-\beta - \frac{c}{3}\right)}_{(-)} + \underbrace{-b + 2\beta + c}_{-(g)}, \tag{G.67}$$

$$h_2 = F_{[2,1]} \underbrace{\left(-\frac{2b}{9} + \frac{2\beta}{9}\right)}_{(-)} + F_{[3]} \underbrace{\left(-\frac{b}{3} - \frac{\beta}{3}\right)}_{(-)} + \underbrace{-\frac{2b}{3} + \frac{4\beta}{3} + \frac{2c}{3}}_{-(g)} \tag{G.68}$$

If $(g) > 0$, then h_0 and h_2 are negative. To show h_1 is negative, assume (x) is positive, and find the maximum possible value of h_1, which is when $F_{[2,1]} = 1$ and $F_{[1,1,1]} = F_{[3]} = 0$. Then

$$9h_1 = -5 \underbrace{(b - 2\beta - c)}_{(g)} + \underbrace{c - 2b}_{(-)} < 0. \tag{G.69}$$

Therefore, if C+D can resist invasion by Liars under zero homophily, then C+D can resist invasion by Liars at any level of homophily.

Invasion of Unconditional Cooperators into C+D. Under zero homophily, choosing D for the fitness comparison.

$$h_0 = h_1 = h_2 = \beta - c < 0, \tag{G.70}$$

therefore U cannot invade the C+D coexistence.

Under intermediate homophily, choosing D for the fitness comparison, all terms are equal:

$$h_i = F_{[1,1,1]} \underbrace{\left(-\frac{2b}{3} + \frac{4\beta}{3}\right)}_{(-)} + \underbrace{F_{[3]} \left(\frac{b}{3} + \frac{\beta}{3}\right)}_{(+)} + \frac{2b}{3} - \frac{\beta}{3} - c \tag{G.71}$$

The effect of increasing homophily is to decrease $F_{[1,1,1]}$ and/or increase $F_{[3]}$, therefore the effect of increasing homophily is to increase all h_i by the same amount, which is to facilitate the the invasion of U into C+D.

If $h_2 > 0$, then U can invade all-D. Therefore, U can invade C+D iff U can invade all-D.

G.5.3 Unconditional Defectors coexistence + Unconditional Cooperators coexistence

Invasion of Coordinating Cooperators into D+U. Under zero homophily, choosing D for the fitness comparison

$$h_0 = h_1 = h_2 = -\varepsilon < 0, \tag{G.72}$$

therefore C cannot invade the D+U coexistence.

Under intermediate homophily, choosing D for the fitness comparison, all terms are equal:

$$h_i = F_{[1,1,1]} \underbrace{\left(-\frac{2b}{3} + \frac{2\beta}{3} + \frac{2c}{3}\right)}_{(-)} + F_{[3]} \underbrace{\left(\frac{b}{3} - \frac{\beta}{3}\right)}_{(+)} + \frac{2b}{3} - \frac{2\beta}{3} - \frac{2c}{3} - \varepsilon \tag{G.73}$$

Increasing homophily decreases $F_{[1,1,1]}$ and/or increases $F_{[3]}$; therefore, increasing homophily

facilitates the invasion of C into D+U.

If $h_2 > 0$, then C can invade all-D. Therefore, C can invade D+U if and only if C can invade all-D.

Invasion of Liars into D+U. Under intermediate homophily, choosing D for the fitness comparison, all terms are equal

$$h_i = -\varepsilon, \tag{G.74}$$

therefore L cannot invade D+U, and neither nonlinearity nor homophily has an effect on its invasion.

G.5.4 Liars + Unconditional Cooperators coexistence

Invasion of Unconditional Defectors into L+U. Choosing L for the fitness comparison, the invasion fitness coefficients for rare D invading the L+U coexistence are all constant and equal:

$$h_0 = h_1 = h_2 = \varepsilon > 0. \tag{G.75}$$

Therefore, D can always invade L+U, and neither nonlinearity nor homophily has an effect on invasion.

Invasion of Coordinating Cooperators into L+U. Under zero homophily, choosing U for the fitness comparison, we have

$$h_0 = -\frac{\beta}{3} + \frac{c}{3} - \varepsilon, \tag{G.76}$$

$$h_1 = -\varepsilon, \tag{G.77}$$

$$h_2 = -\beta + c - \varepsilon. \tag{G.78}$$

Consider the case of $c > \beta$. If $\varepsilon = 0$, we find that $h_0 > 0, h_1 = 0$ and $h_2 > 0$. This means the invasion condition, Eq. G.11, which requires a weighted sum of h_i terms to be positive, is satisfied and therefore C invades L+U. Since the left hand side of Eq. G.11 is obviously a continuous function of ε, by the continuity argument C can still invade L+U as long as $\varepsilon \to 0$ is small enough. Therefore we can conclude that $c > \beta$ and $\varepsilon \to 0$ is a sufficient condition for C to invade L+U.

Under intermediate homophily, on the other hand, the h_i have mixed signs and no obvious insights found.

H Analytic results for 8-player Coordinated Cooperation example

H.1 Payoffs

We calculated the expected payoffs in the sigmoid PGG as follows. For clarity, let us index the variables with the strategy names instead of the strategy indices. Let g_s be the number of s-strategists in the whole group ($s \in \{D, C, L, U\}$), and let $\boldsymbol{g}_{\mathrm{a}} = (g_D, g_C, g_L, g_U)$ If the lottery meets quorum, i.e., $g_C + g_L \geq \tau$, then the probability that j Coordinating Cooperators and $\tau - j$ Liars will be designated as contributors is

$$P_j = \begin{cases} \frac{\binom{g_C}{j}\binom{g_L}{\tau-j}}{\binom{g_C+g_L}{\tau}} & \text{if } \max(0, \tau - g_L) \leq j \leq \min(g_C, \tau), \\ 0 & \text{otherwise.} \end{cases} \tag{H.1}$$

Therefore, the expected payoff to an individual pursuing each of the strategies are as follows: to Unconditional Defectors

$$\hat{\pi}(\boldsymbol{e}_D, \boldsymbol{g}_{\mathrm{a}}) = \begin{cases} B(g_U) & \text{if } g_C + g_L < \tau, \\ \sum_{j=0}^{\tau} P_j B(j + g_U) & \text{if } g_C + g_L \geq \tau, \end{cases} \tag{H.2}$$

to Coordinating Cooperators

$$\hat{\pi}(\boldsymbol{e}_C, \boldsymbol{g}_{\mathrm{a}}) = \begin{cases} B(g_U) - \varepsilon & \text{if } g_C + g_L < \tau, \\ \left(\sum_{j=0}^{\tau} P_j \left(B(j + g_U) - \frac{j}{g_C} c\right)\right) - \varepsilon & \text{if } g_C + g_L \geq \tau, \end{cases} \tag{H.3}$$

to Liars

$$\hat{\pi}(\boldsymbol{e}_L, \boldsymbol{g}_{\mathrm{a}}) = \begin{cases} B(g_U) - \varepsilon & \text{if } g_C + g_L < \tau, \\ \left(\sum_{j=0}^{\tau} P_j B(j + g_U)\right) - \varepsilon & \text{if } g_C + g_L \geq \tau, \end{cases} \tag{H.4}$$

and to Unconditional Cooperators

$$\hat{\pi}(\boldsymbol{e}_U, \boldsymbol{g}_{\mathrm{a}}) = \begin{cases} B(g_U) - c & \text{if } g_C + g_L < \tau, \\ \left(\sum_{j=0}^{\tau} P_j B(j + g_U)\right) - c & \text{if } g_C + g_L \geq \tau. \end{cases} \tag{H.5}$$

H.2 Condition for Coordinated Cooperation to persist and resist invasion by Liars

In the main text, we stated that a C+D coexistence in a well-mixed population (no homophily) can resist invasion by Liars if the following condition is satisfied (Eq. 34)

$$B(\tau) - B(\tau - 1) > c.$$

In this section, we will prove the statement from the main text (H.2.1), and we will detail its other interpretations. We will show that Eq. 34 also has the following meanings:

- it is a necessary condition for the strategy profile with τ contributors to be a Nash equilibrium in the well-mixed game with untransformed payoffs (H.2.2);
- it means that Unconditional Cooperators have a positive switching gain against Unconditional Defectors when there are $\tau - 1$ other Unconditional Cooperators, which is a necessary but not sufficient condition for coexistence between Unconditional Cooperators and Defectors (H.2.3); and
- it is the condition under which an all-C population can resist invasion by Liars (H.2.4).

H.2.1 Condition for a C+D coexistence to resist invasion by Liars

In this subsection, we will consider a well-mixed population (no homophily) with a pre-existing stable coexistence between Coordinating Cooperators and Defectors, and we will show that the condition given in Eq. 34 in the main text is the condition under which the population can resist invasion by Liars.

Define $\boldsymbol{p}^* = (p_C^*, p_D^*, 0)$ as the interior steady state of the C + D system, with the final 0 representing a 0 proportion of Liar types in the population at invasion. In order for the coexistence between C and D types to resist invasion by Liars, the invasion fitness of Liars must be negative

$$\frac{1}{p_L}\dot{p}_L\bigg|_{\boldsymbol{p}^*} < 0. \tag{H.6}$$

The replicator dynamics for the Liar type

$$\dot{p}_L = p_L\left(\overline{\pi}_L - \sum_{i\in\{C,D,L\}} p_i\overline{\pi}_i\right). \tag{H.7}$$

Substituting Eq. H.7 into Eq. H.6

$$\frac{1}{p_L}\dot{p}_L\bigg|_{\boldsymbol{p}^*} = \overline{\pi}_L|_{\boldsymbol{p}^*} - \left(\underbrace{p_C^*\,\overline{\pi}_C|_{\boldsymbol{p}^*} + p_D^*\,\overline{\pi}_D|_{\boldsymbol{p}^*}}_{=\overline{\pi}_C|_{\boldsymbol{p}^*}=\overline{\pi}_D|_{\boldsymbol{p}^*}} + \underbrace{p_L^*}_{=0}\,\overline{\pi}_L|_{\boldsymbol{p}^*}\right),$$

and therefore the condition for the population to resist invasion by Liars (Eq. H.6) becomes

$$\overline{\pi}_L|_{\boldsymbol{p}^*} - \overline{\pi}_C|_{\boldsymbol{p}^*} < 0. \tag{H.8}$$

Let K be the random variable denoting the number of Unconditional Cooperators k among the $(n-1)$ nonfocal group members. In a well-mixed population (no homophily), $\mathbb{P}[K = k \mid \boldsymbol{p}^*]$

is binomially distributed and independent of the strategy of the focal group member. Therefore, the expected payoff to Coordinating Cooperators is

$$\overline{\pi}_C|_{\boldsymbol{p}^*} = -\varepsilon + \sum_{k=\tau-1}^{n-1} \mathbb{P}[K = k \mid \boldsymbol{p}^*] \left(B(\tau) - \frac{\tau}{k+1} c \right). \tag{H.9}$$

Because the invading L type is rare and the population is well-mixed, there will only ever be 1 Liar in the group. Recall that the Liar contributes 1 towards the quorum τ for the lottery to take place, but the Liar does not contribute if chosen by the lottery. If the lottery takes place, the probability that the Liar is chosen is $\tau/(k+1)$, resulting in a public good benefit of $B(\tau-1)$; and the probability that the Liar is not chosen is $(k+1-\tau)/(k+1)$, resuling in a public good benefit of $B(\tau)$. Therefore, the expected payoff to the invading Liar is

$$\overline{\pi}_L|_{\boldsymbol{p}^*} = -\varepsilon + \sum_{k=\tau-1}^{n-1} \mathbb{P}[K = k \mid \boldsymbol{p}^*] \left(\frac{\tau B(\tau-1) + (k+1-\tau) B(\tau)}{k+1} \right). \tag{H.10}$$

Substituting $\overline{\pi}_C|_{\boldsymbol{p}^*}$ (Eq. H.9) and $\overline{\pi}_L|_{\boldsymbol{p}^*}$ (Eq. H.10) into the invasion-resistance condition in Eq. H.8, the invasion-resistance condition becomes

$$\tau \Big(B(\tau-1) - (B(\tau) - c) \Big) \left(\sum_{k=\tau-1}^{n-1} \frac{\mathbb{P}[K = k \mid \boldsymbol{p}^*]}{k+1} \right) < 0. \tag{H.11}$$

The summation term in Eq. H.11 is positive; therefore, the condition for invasion resistance becomes

$$B(\tau) - B(\tau-1) > c, \tag{H.12}$$

which is Eq. 34 in the main text.

H.2.2 Condition for τ contributors and $n-\tau$ defectors to be a Nash equilibrium

In this subsection, we derive the conditions under which the strategy profile with τ contributors and $n-\tau$ non-contributors is a Nash equilibrium in the well-mixed game (untransformed payoffs) when we consider only two strategies, Unconditional Cooperator and Unconditional Defector, and we show that Eq. 34 in the main text is a necessary condition for this strategy profile to be a Nash equilibrium.

Let S_j be the set of all possible strategies for individual j. Let $u_j(s_j, s_{-j})$ be the payoff to individual j when individual j pursues strategy s_j and all other players apart from j pursue strategy profile s_{-j}. Let $s^* = (s_j^*, s_{-j}^*)$ be a strategy profile. Then s^* is a strict Nash equilibrium if

$$u_i(s_j^*, s_{-j}^*) > u_j(s_j, s_{-j}^*) \quad \forall s_j (\neq s_j^*) \in S_j. \tag{H.13}$$

The possible strategies are $S_j = \{D, U\}$. In the situation where the focal j in s^* is an Unconditional Defector

$$\begin{aligned} s_j^* &= D, \\ s_{-j}^* &= \{\underbrace{U, \ldots, U}_{\tau}, \underbrace{D, \ldots, D}_{n-\tau-1}, \} \end{aligned}$$

and Eq. H.13 gives

$$B(\tau) > B(\tau+1) - c. \tag{H.14}$$

In the situation where the focal j in s^* is an Unconditional Cooperator

$$\begin{aligned} s_j^* &= U, \\ s_{-j}^* &= \{\underbrace{U, \ldots, U}_{\tau-1}, \underbrace{D, \ldots, D}_{n-\tau}, \} \end{aligned}$$

and Eq. H.13 gives

$$B(\tau) - c > B(\tau - 1). \tag{H.15}$$

Eq. H.15 is equivalent to Eq. 34 from the main text. Therefore, Eq. H.15 is a necessary condition for the strategy profile with τ contributors to be a Nash equilibrium.

H.2.3 Switching gain for Unconditional Cooperators against Unconditional Defectors

In this subsection, we find the switching gains for Unconditional Cooperators against Unconditional Defectors. We show that, in a group with $\tau - 1$ Unconditional Cooperators, the switching gain is positive when Eq. 34 from the main text is satisfied.

The switching gains for Unconditional Cooperators against Defectors, d_k, are the payoff gains an individual would receive when they switch from being a D- to U-strategist when grouped with k U-strategists and $(n - 1 - k)$ D-strategists. The payoff to an Unconditional Cooperator is

$$\pi_U(k) = B(k + 1) - c, \tag{H.16}$$

and the payoff to an Unconditional Defector is

$$\pi_D(k) = B(k), \tag{H.17}$$

therefore, the switching gain is

$$d_k = \pi_U(k) - \pi_D(k) = B(k + 1) - B(k) - c. \tag{H.18}$$

When $k = \tau - 1$, the condition for a positive switching gain, $d_{\tau-1} > 0$, becomes

$$B(\tau) - B(\tau - 1) > c, \tag{H.19}$$

which is equivalent to Eq. 34.

For the sigmoid game we consider, d_0 and d_{n-1} are negative, d_k is increasing for $k < \tau - 1$, reaches its maximum at $k = \tau - 1$, and is decreasing for $k > \tau - 1$. Therefore, by Results 3.1 and 4.1 in Peña et al. (2014), Eq. H.19 is a necessary (but not sufficient) condition for stable coexistence between U and D-strategists.

H.2.4 Condition under which Coordinating Cooperators can resist invasion by Liars

In this subsection, we derive the condition under which a well-mixed population (no homophily) of all Coordinating Cooperators can resist invasion by Liars and show that it is equivalent to Eq. 34 in the main text.

Define $\boldsymbol{p}^* = (p_C^*, 0) = (1, 0)$ as the strategy distribution in the all-C population where the final 0 represents the 0 proportion of Liar types in the population at invasion. In order for the all-C population to resist invasion by Liars, the invasion fitness of Liars must be negative

$$\frac{1}{p_L}\dot{p}_L\bigg|_{\boldsymbol{p}^*} < 0. \tag{H.20}$$

The replicator dynamics for the Liar type

$$\dot{p}_L = p_L \left(\overline{\pi}_L - \overline{\pi}_C \right), \tag{H.21}$$

and therefore the condition for the all-C population to resist invasion by Liars becomes

$$\overline{\pi}_L|_{\boldsymbol{p}^*} - \overline{\pi}_C|_{\boldsymbol{p}^*} < 0. \tag{H.22}$$

In the all-C population, the lottery is always held with all n members participating, τ contributors are always chosen, and the probability that an individual will be chosen as a contributor is τ/n. Therefore, the expected payoff to a Coordinating Cooperator is

$$\overline{\pi}_C|_{\boldsymbol{p}^*} = B(\tau) - \frac{\tau}{n} c - \varepsilon. \tag{H.23}$$

Because the invading L type is rare and the population is well-mixed, there will only ever be 1 Liar in the group. The lottery always takes place. The probability that the Liar is chosen is τ/n, resulting in a public good benefit of $B(\tau - 1)$; and the probability that the Liar is not chosen is $(n - \tau)/n$, resulting in a public good benefit of $B(\tau)$. Therefore, the expected payoff to the invading Liar is

$$\overline{\pi}_L|_{\boldsymbol{p}^*} = \frac{\tau B(\tau - 1) + (n - \tau) B(\tau)}{n} - \varepsilon. \tag{H.24}$$

Substituting the expected payoffs for C-strategists (Eq. H.23) and rare L-strategists (Eq. H.24) into the invasion-resistance condition in Eq. H.22, the condition for invasion resistance becomes

$$B(\tau) - B(\tau - 1) - c > 0, \tag{H.25}$$

which is equivalent to Eq. 34.

H.3 Coexistence of Coordinating Cooperators and Defectors

The purpose of this subsection is to show that Coordinating Cooperators can persist in a well-mixed population (no homophily) in evolutionary coexistence with Unconditional Defectors provided that (1) τ Coordinating Cooperators will receive a positive payoff from the game (when we neglect cognitive cost, ε), and (2) the cognitive cost of being a Coordinating Cooperator ε is small enough.

We assume that τ Coordinating Cooperators will receive a positive payoff from the game when we neglect cognitive cost, ε:

$$\text{Assumption 1:} \qquad B(\tau) - c > 0. \tag{H.26}$$

First, let us consider a scenario where there is no cognitive cost to being a Coordinating Cooperator, $\varepsilon = 0$ (e.g., Fig. H.5a). Let k be the number of Coordinating Cooperators among the $n-1$ nonfocal members of the group. Then the payoffs to a focal Coordinating Cooperator are

$$\pi_C(k) = \begin{cases} 0 & \text{if } k < \tau - 1, \\ B(\tau) - \frac{\tau}{k+1} c & \text{if } k \geq \tau - 1, \end{cases} \tag{H.27}$$

the payoffs to a focal Unconditional Defector are

$$\pi_D(k) = \begin{cases} 0 & \text{if } k < \tau, \\ B(\tau) & \text{if } k \geq \tau, \end{cases} \tag{H.28}$$

and therefore the switching gains from D to C are

$$d_k = \begin{cases} 0, & \text{if } k < \tau - 1, \\ B(\tau) - c & \text{if } k = \tau - 1, \\ -\frac{\tau}{k+1} c & \text{if } k > \tau - 1. \end{cases} \tag{H.29}$$

The gain sequence $\boldsymbol{d}$ has a single sign change, its first non-zero entry has a positive sign, and its last non-zero entry has a negative sign; therefore, by Result 3.2.b of Peña et al. (2014), an all-C population is unstable, an all-D population is also unstable, and there is an interior steady state representing the stable coexistence of C and D. The result also implies that if we let $g(p)$ be the gain function, which is defined as the average payoff of Coordinating Cooperators minus the average payoff of Unconditional Defectors, calculated at the population frequency of Coordinated Cooperators being p (see Peña et al. (2014)), its maximal value $\overline{g}$ satisfies

$$\overline{g} > 0. \tag{H.30}$$

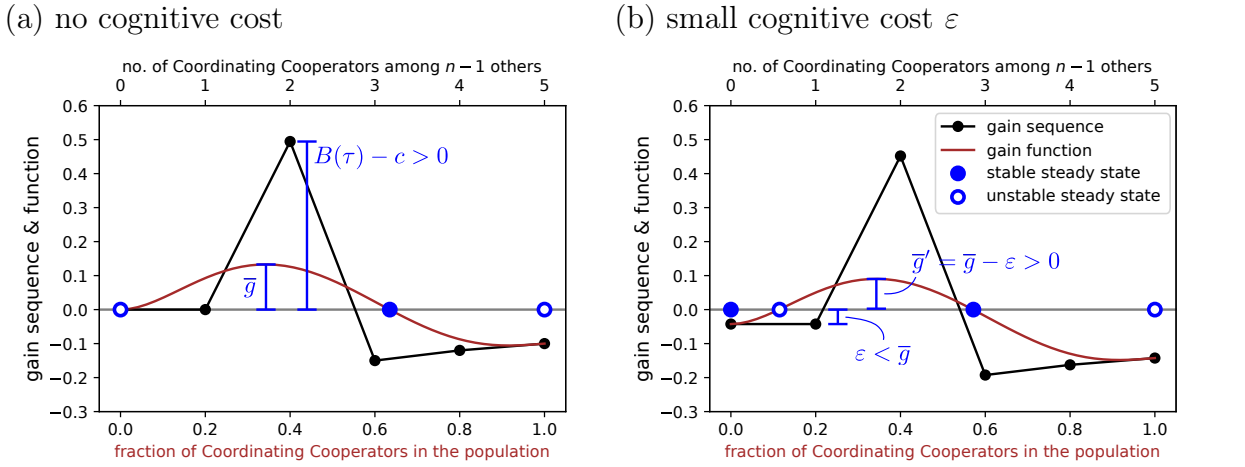


Figure H.5: An example of a gain sequence, $d_0, \ldots, d_{n-1}$ and gaint function, $g(p)$ for Coordinating Cooperators versus Unconditional Defectors that illustrates the main principle behind the proof: introducing a cognitive cost ε shifts both functions down by ε. Consequently, provided $B(\tau) - c > 0$ and $\varepsilon < \overline{g}$, coexistence between C and D is possible. (a) When there are no cognitive costs to being a Coordinating Cooperator, then provided $B(\tau) - c > 0$, there will be one interior steady state that is stable (filled blue circle). (b) When a cognitive cost is introduced that satisfies $\varepsilon < \overline{g}$, then there will be two interior steady states, one stable and one unstable (empty blue circle). Parameter values used: $n = 6$, $\tau = 3$, $\sigma = 10$, $c = 0.2$.

Now let us consider a scenario where there is a cognitive cost to being a Coordinating Cooperator (e.g., Fig. H.5b). We assume that the cost is not too high; specifically,

$$\text{Assumption 2:} \qquad 0 < \varepsilon < \overline{g}. \tag{H.31}$$

Compared to the $\varepsilon = 0$ scenario above, The effect of imposing a cognitive cost is to shift the entire gain sequence uniformly down by ε (e.g., Fig. H.5), i.e., the gain sequence becomes

$$d'_k = \begin{cases} -\varepsilon & \text{if } k < \tau - 1, \\ B(\tau) - c - \varepsilon & \text{if } k = \tau - 1, \\ -\frac{c\tau}{k+1} - \varepsilon & \text{if } k > \tau - 1. \end{cases} \tag{H.32}$$

The gain sequence $\boldsymbol{d}'$ now has two sign changes and its first non-zero entry has a negative sign. By Result 4.1 of Peña et al. (2014), the number of steady states and their signs depends on the the maximal value of the new gain function, $\overline{g}'$. Shifting the entire gain sequence down by ε also shifts the gain function g' uniformly down, and the new maximal value of the gain function is

$$\overline{g}' = \overline{g} - \varepsilon > 0. \tag{H.33}$$

By Result 4.1.c of Peña et al. (2014): the all-C population is unstable, the all-D population is stable, and there is one stable steady state p_C^* and one unstable steady state p_C^{**} satisfying $0 < p_C^{**} < p_C^* < 1$. Therefore, a stable C+D coexistence is possible.

I Numerical results for 8-player Coordinated Cooperation example

In this Supplement, we explore the evolutionary dynamics of the two example scenarios in detail, and we explore how the dynamics change as homophily decreases from a high ancestral level to a low contemporary level. We use the whole-group accounting approach for Example 1, and the payoff-matrix approach for Example 2. For both examples, we use the leader-driven group-formation model to calculate parameter F (Eq. 1). The code used to produce these results is available in the online repository, which we hope will provide a practical help for future workers.

Throughout this Supplement, the evolutionary dynamics and steady states on the faces of the tetrahedral strategy space are illustrated in the same way as in Fig. 5 in the main text. Specifically, red dots represent unstable equilibria, and blue dots represent stable equilibria in the subspace sub($\boldsymbol{p}^*$). Stability on sub($\boldsymbol{p}^*$) differs from local asymptotic stability of $\boldsymbol{p}^*$ on the whole simplex S_m in that these steady states may be invadable by new strategies not in the population (see Supplement D.3 for details of the difference).

I.1 Example scenario 1

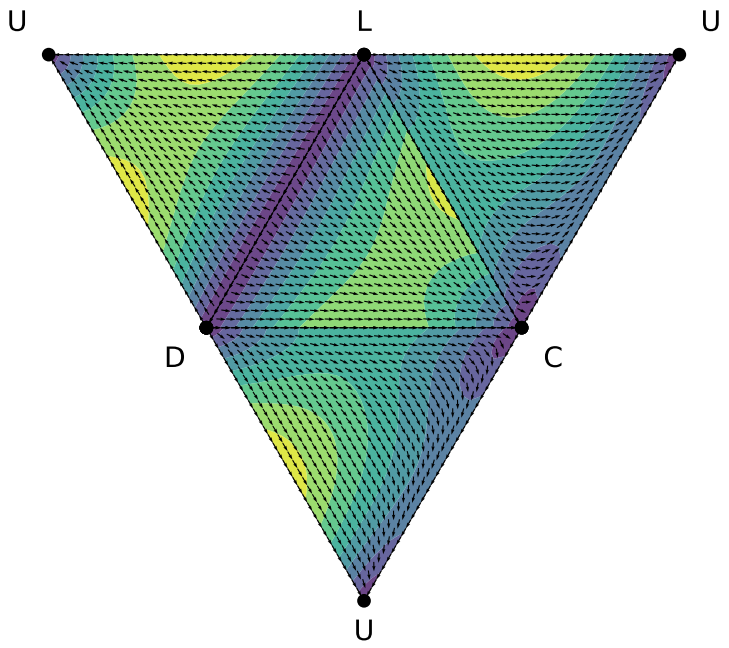
In example scenario 1, the parameter values are chosen such that the strategy profile with τ contributors and $n-\tau$ non-contributors is a Nash equilibrium (Table I.12). Code to reproduce these results can be found in `scripts/sigmoid_UDCL/` and additional results can be found in `results/sigmoid_UDCL/`.

Table I.12: Parameter values used for example scenario 1.

parameter value	description
$n = 8$	Number of group members playing the game.
$\tau = 5$	Lottery quorum and midpoint of benefits function minus 0.5.
$\sigma = 10$	Steepness of the sigmoid benefits function.
$c = 0.25$	Cost of contributing to the public good.
$\varepsilon = 0.02$	Cognitive cost of being a communicative player (C or L).

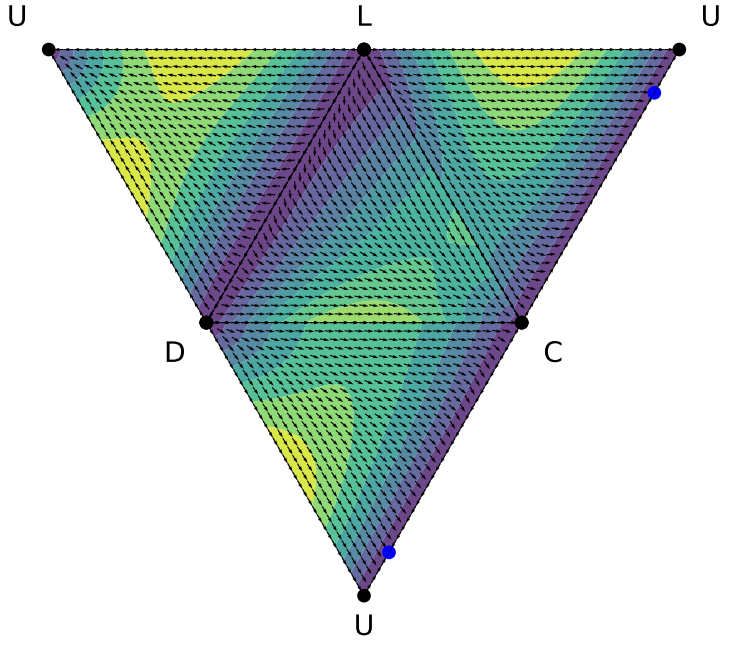
At maximum homophily, with parameter $h = 1$ in the leader-driven group-formation model, the globally evolutionarily stable strategy is Unconditional Cooperation.

$h = 1$

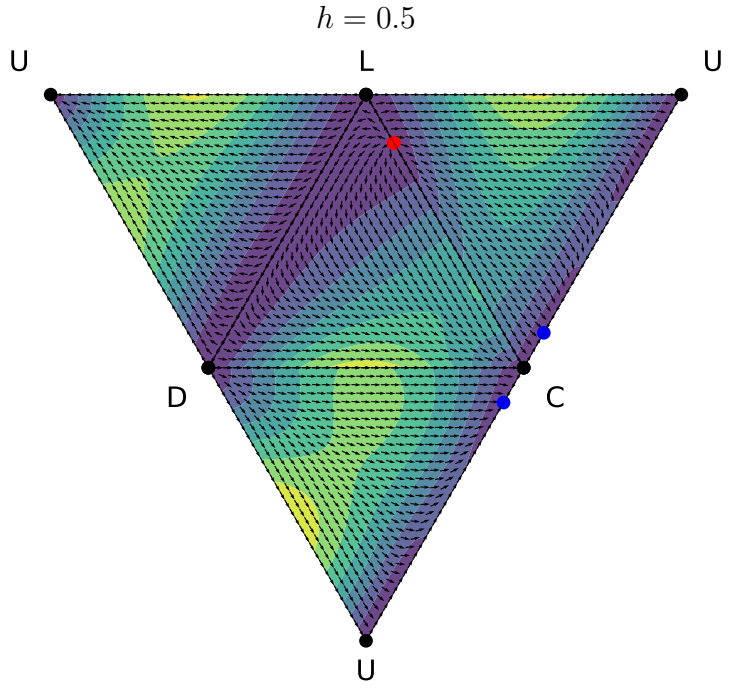


As homophily decreases, the first event (occurring between $h = 0.65$ and $q = 0.645$) is the appearance of a coexistence between U and C, which is a global attractor. It emerges at $p_U = 1$ and moves away as homophily decreases.

$h = 0.645$



The next event is the appearance of a separatrix between L and C (occurring before $h = 0.5$), which emerges near $p_L = 1$ and moves away as homophily decreases.



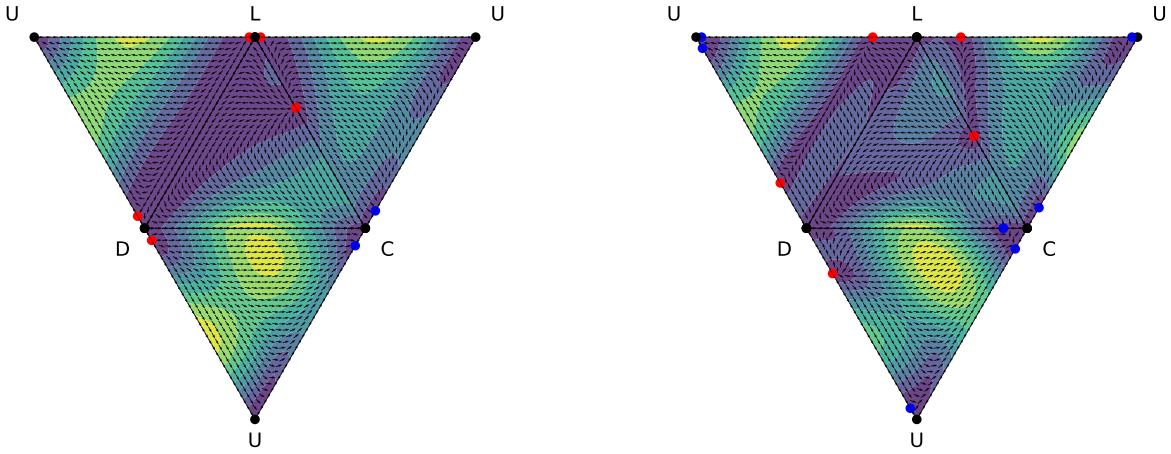
At $h = 0.5$, the U+C coexistence is uninvadable to all other strategies.

Table I.13: Steady states from `fixedpts_stability_sigmoidUDCL_v1_leader_driven_ngrid_9_q_5000.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
0	0	0	1		-0.196584	0.014948	-0.216584	
0	0	1	0		0.020000	-0.040833		0.105206
0	1	0	0		-0.143905		-0.258182	0.050131
1	0	0	0			0.132027	-0.020000	0.085206
0	0.872272	0	0.127728	stable	-0.212479		-0.266321	
0	0.176902	0.823098	0	unstable				

The next events (occurring before $h = 0.4$) are the appearances of unstable steady states between D and U (emerging at $p_D = 1$) and between L and U (emerging at $p_L = 1$). As homophily decreases further, these unstable states each become paired with a stable steady state, both emerging at $p_U = 1$ (occurs before $h = 0.3$). Within the same interval, a stable steady state appears between C and D (emerging at $p_C = 1$), which will also become paired with an unstable steady state as homophily decreases further.

$h = 0.4$ $h = 0.3$



Thus, by $h = 0.3$, there are 4 stable steady states in the system: U+C, U+L, C+D, and U+D. The U+C coexistence remains uninvadable to all other strategies; the U+L coexistence is invadable by both D and C; the U+D coexistence is invadable by invadable to C; and the C+D coexistence is invadable by U. Thus, sequential invasions will terminate in U+C, which remains the evolutionary endpoint.

Table I.14: Steady states from `fixedpts_stability_sigmoidUDCL_v1_leader_driven_ngrid_9_q_7000.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
0	0	0	1		0.035176	0.068581	0.015176	
0	0	1	0		0.020000	-0.116350		-0.106830
0	1	0	0		0.047010		-0.156719	0.092966
1	0	0	0			0.014863	-0.020000	-0.126830
0	0	0.800290	0.199710	unstable				
0	0	0.024850	0.975150	stable	0.020000	0.059705		
0	0.892269	0	0.107731	stable	-0.032696		-0.138100	
0	0.518106	0.481894	0	unstable				
0.762992	0	0	0.237008	unstable				
0.058471	0	0	0.941529	stable		0.033806	-0.020000	
0.107715	0.892285	0	0	stable			-0.160745	0.021924

The next event (occurring before $h = 0.27$) is the appearance of an unstable steady state on the (D, C, U) face. This unstable steady state emerges out of the C+D coexistence, which switches the C+D coexistence from being invadable by U to being uninvadable by all other strategies.

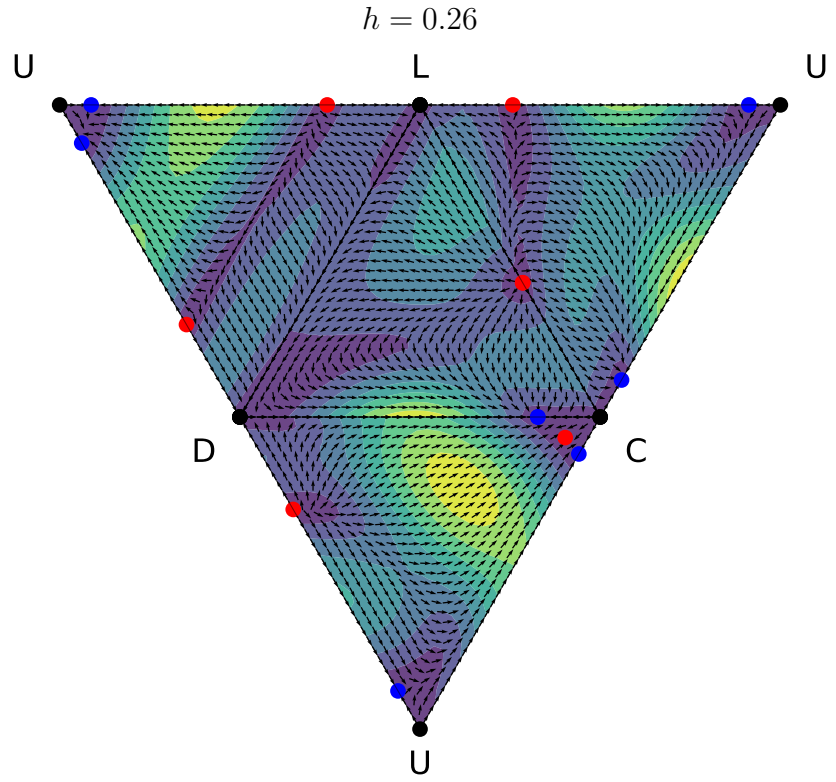


Table I.15: Steady states from `fixedpts_stability_sigmoidUDCL_v1_leader_driven_ngrid_9_q_7400.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
0	0	0	1		0.074126	0.087590	0.054126	
0	0	1	0		0.020000	-0.125229		-0.136153
0	1	0	0		0.081145		-0.137569	0.109483
1	0	0	0			0.001863	-0.020000	-0.156153
0	0	0.742465	0.257535	unstable				
0	0	0.087433	0.912567	stable	0.020000	0.051911		
0	0.881502	0	0.118498	stable	-0.005512		-0.110307	
0	0.569733	0.430267	0	unstable				
0.999831	0	0	0			0.001870	-0.019982	-0.155995
0.703764	0	0	0.296236	unstable				
0.122346	0	0	0.877654	stable		0.013946	-0.020000	
0.173223	0.826777	0	0	stable			-0.143651	-0.014060
0.064089	0.869981	0	0.065929	unstable				

As homophily decreases further, the unstable steady state on the (D, C, U) face moves towards and collides with the U+C coexistence (occurs before $h = 0.25$), which switches the U+C coexistence from being uninvadable by all other strategies to being invadable by D. Simultaneously, an unstable steady state appears between C and D (emerging at $p_D = 1$), which now renders the all-D population uninvadable by all other strategies. Thus, by $h = 0.24$, the potential evolutionary endpoints are the all-D population and D+C coexistence.

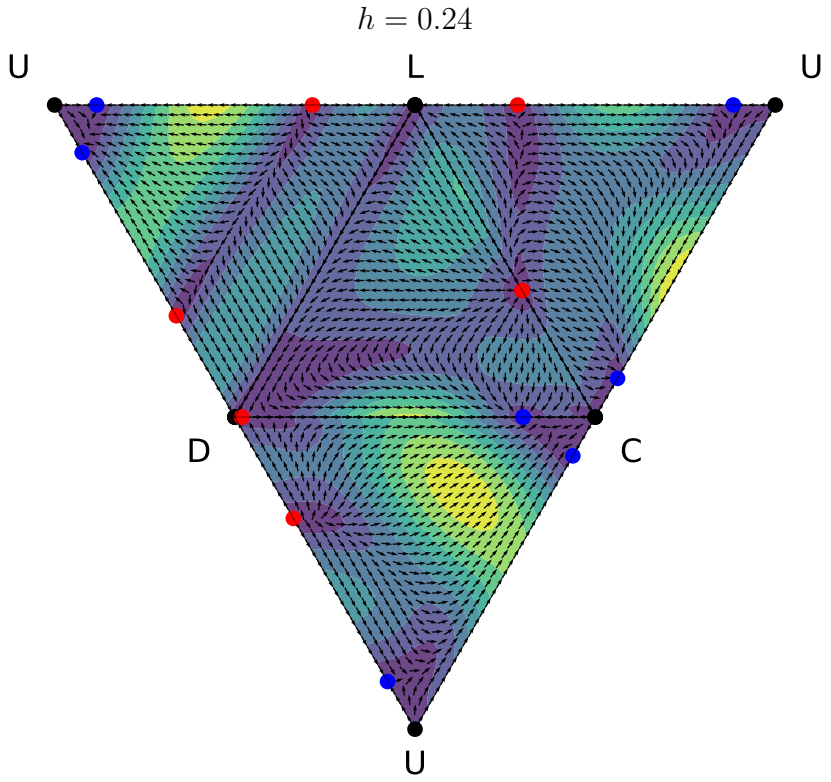
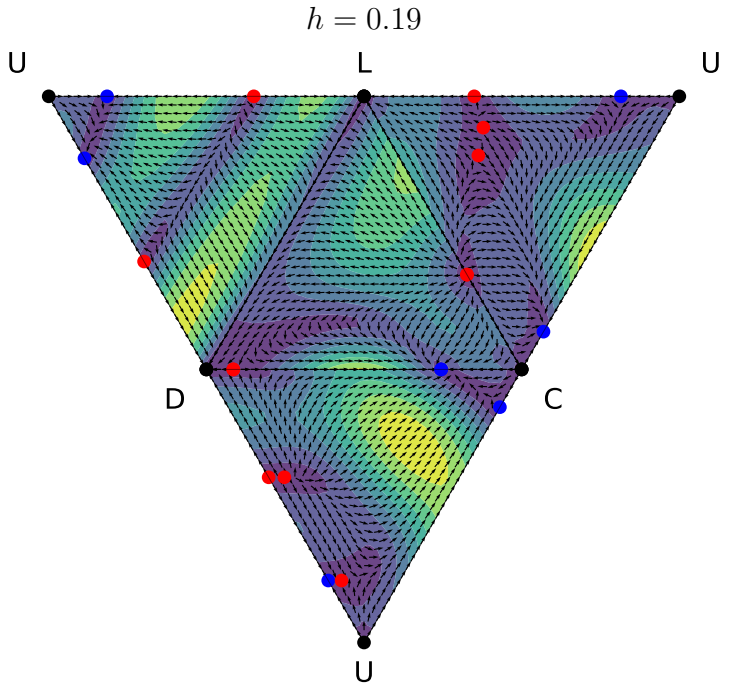


Table I.16: Steady states from `fixedpts_stability_sigmoidUDCL_v1_leader_driven_ngrid_9_q_7600.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
0	0	0	1		0.092228	0.097804	0.072228	
0	0	1	0		0.020000	-0.129042		-0.149034
0	1	0	0		0.096760		-0.128235	0.117591
1	0	0	0			-0.003282	-0.020000	-0.169034
0	0	0.715246	0.284754	unstable				
0	0	0.116556	0.883444	stable	0.020000	0.047555		
0	0.875868	0	0.124132	stable	0.007392		-0.096166	
0	0.594324	0.405676	0	unstable				
0.675320	0	0	0.324680	unstable				
0.152688	0	0	0.847312	stable		0.006012	-0.020000	
0.981066	0.018934	0	0	unstable				
0.200252	0.799748	0	0	stable			-0.134922	-0.030236

Soon after the emergence of the D+C unstable steady state, two more unstable steady states appear on the (U, D, C) face, emerging out of the U+D unstable-stable steady-state pair (occurs before $h = 0.2$). The significant effect of the unstable steady states on the (U, D, C) face is that, whereas previously the population at U+D was invadable by C providing a potential route to C+D coexistence, now any population at U+D is 'trapped' in that state. Thus, by $h = 0.19$, the potential evolutionary endpoints of an invasion sequence are: all-D, C+D, and U+D.



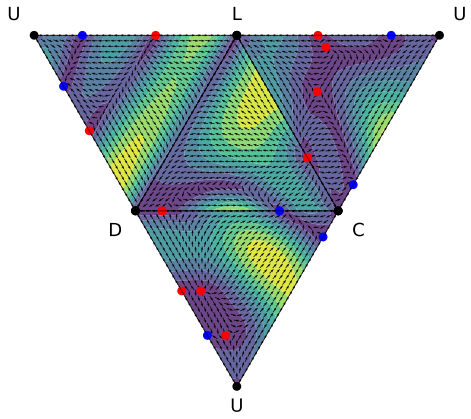
Two unstable steady states also appear in the interior of the strategy space.

Table I.17: Steady states from `fixedpts_stability_sigmoidUDCL_v1_leader_driven_ngrid_9_q_8100.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
0	0	0	1		0.132996	0.124381	0.112996	
0	0	1	0		0.020000	-0.136928		-0.176097
0	1	0	0		0.130418		-0.105742	0.135739
1	0	0	0			-0.012546	-0.020000	-0.196097
0	0	0.650451	0.349549	unstable				
0	0	0.185160	0.814840	stable	0.020000	0.035670		
0	0.861799	0	0.138201	stable	0.037740		-0.060665	
0	0.652873	0.347127	0	unstable				
0	0.114652	0.564229	0.321119	unstable				
0	0.216628	0.528963	0.254409	unstable				
0.604903	0	0	0.395097	unstable				
0.226984	0	0	0.773016	stable		-0.008372	-0.020000	
0.914641	0.085359	0	0	unstable				
0.255317	0.744683	0	0	stable			-0.112862	-0.065413
0.555233	0.049910	0	0.394857	unstable				
0.184680	0.042306	0	0.773014	unstable				
0.016939	0.259279	0.497743	0.226039	unstable				
0.030128	0.090624	0.541191	0.338056	unstable				

As homophily decreases further, the U+D steady-state pair move towards each other.

$h = 0.15$



$h = 0.11$

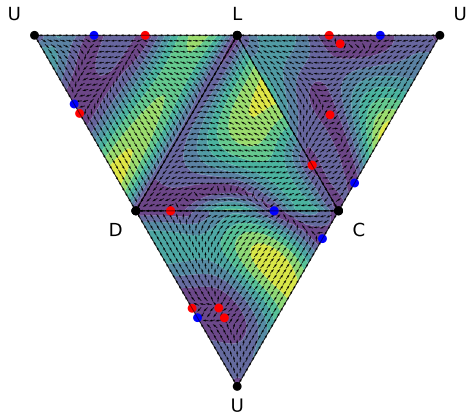
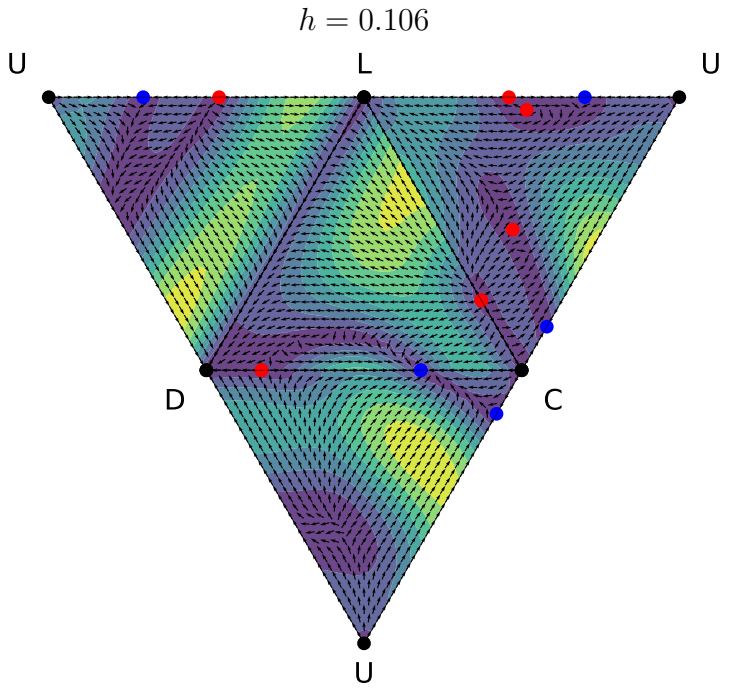


Table I.18: Steady states from `fixedpts_stability_sigmoidUDCL_v1_leader_driven_ngrid_9_q_8500.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p^*_D	p^*_C	p^*_L	p^*_U	on sub($\boldsymbol{p}^*$)	D	C	L	U
0	0	0	1		0.160569	0.145441	0.140569	
0	0	1	0		0.020000	-0.141723		-0.192644
0	1	0	0		0.150867		-0.088764	0.146490
1	0	0	0			-0.016805	-0.020000	-0.212644
0	0	0.599799	0.400201	unstable				
0	0	0.238158	0.761842	stable	0.020000	0.025269		
0	0.851124	0	0.148876	stable	0.060032		-0.032851	
0	0.697213	0.302787	0	unstable				
0	0.321853	0.442325	0.235822	unstable				
0	0.067884	0.526652	0.405463	unstable				
0.543884	0	0	0.456116	unstable				
0.290453	0	0	0.709547	stable		-0.015022	-0.020000	
0.868302	0.131698	0	0	unstable				
0.289673	0.710327	0	0	stable			-0.095350	-0.088335
0.450182	0.094193	0	0.455625	unstable				
0.200344	0.090079	0	0.709578	unstable				
0.036036	0.063158	0.488808	0.411998	unstable				
0.013588	0.344572	0.422645	0.219195	unstable				

When the U+D steady-state pair collide (occurs before $h = 0.106$), they annihilate one another, and any population that was previously at the U+D coexistence will evolve to an all-D population.



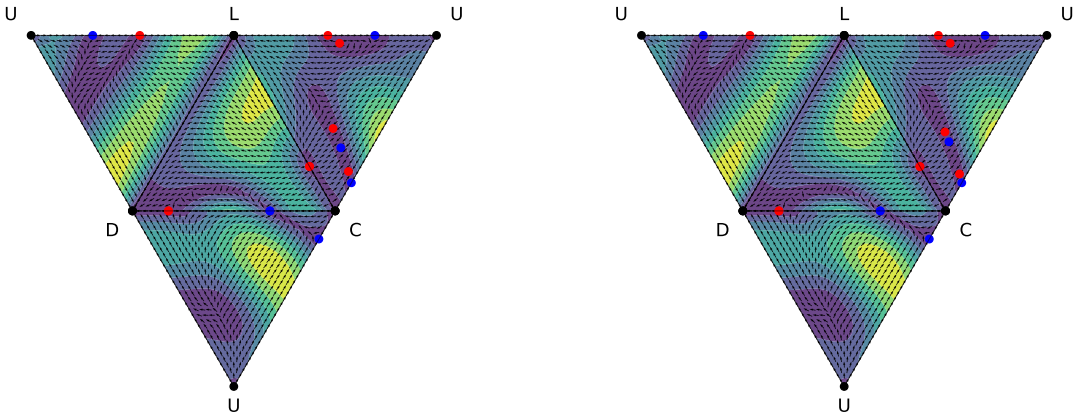
In summary, at $h = 0.106$, there are four stable steady states in the system: L+U, C+U, D, and C+D. The L+U state is invadable by both Defectors and Coordinating Cooperators; the C+U state is invadable by Defectors only; and both D and C+D are uninvadable evolutionary endpoints.

Table I.19: Steady states from `fixedpts_stability_sigmoidUDCL_v1_leader_driven_ngrid_9_q_8940.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
0	0	0	1		0.185309	0.166657	0.165309	
0	0	1	0		0.020000	-0.145661		-0.205990
0	1	0	0		0.166070		-0.071335	0.152934
1	0	0	0			-0.019115	-0.020000	-0.225990
0	0	0.540376	0.459624	unstable				
0	0	0.299660	0.700340	stable	0.020000	0.012549		
0	0.840523	0	0.159477	stable	0.082373		-0.003921	
0	0.743889	0.256111	0	unstable				
0	0.484845	0.285985	0.229171	unstable				
0	0.046665	0.460719	0.492616	unstable				
0.825029	0.174971	0	0	unstable				
0.320160	0.679840	0	0	stable			-0.076739	-0.108555

The next event is the appearance of an attractor-repellor pair on the (L, U, C) face (occurring before $h = 0.1031$). As homophily decreases, the pair move apart and eventually collide with another attractor-repellor pair.

$h = 0.1031$ $h = 0.1026$



Before the collision, the four stable steady states retain the same invasibility status as before: L+U is invadable by both D and C; C+U is invadable by D; and both D and C+D are uninvadable.

Table I.20: Steady states from `fixedpts_stability_sigmoidUDCL_v1_leader_driven_ngrid_9_q_8969.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
0	0	0	1		0.186727	0.167941	0.166727	
0	0	1	0		0.020000	-0.145876		-0.206703
0	1	0	0		0.166804		-0.070237	0.153140
1	0	0	0			-0.019202	-0.020000	-0.226703
0	0	0.536057	0.463943	unstable				
0	0	0.304100	0.695900	stable	0.020000	0.011627		
0	0.746896	0.253104	0	unstable				
0	0.839877	0	0.160123	stable	0.083760		-0.002102	
0	0.640449	0.152006	0.207545	stable	0.024955			
0	0.530664	0.244787	0.224550	unstable				
0	0.045265	0.455447	0.499289	unstable				
0	0.775281	0.046997	0.177722	unstable				
0.822468	0.177532	0	0	unstable				
0.321944	0.678056	0	0	stable			-0.075549	-0.109715

Once the collisions have occurred (occurs by $h = 0.08$), the C+U coexistence becomes invadable by Liars.

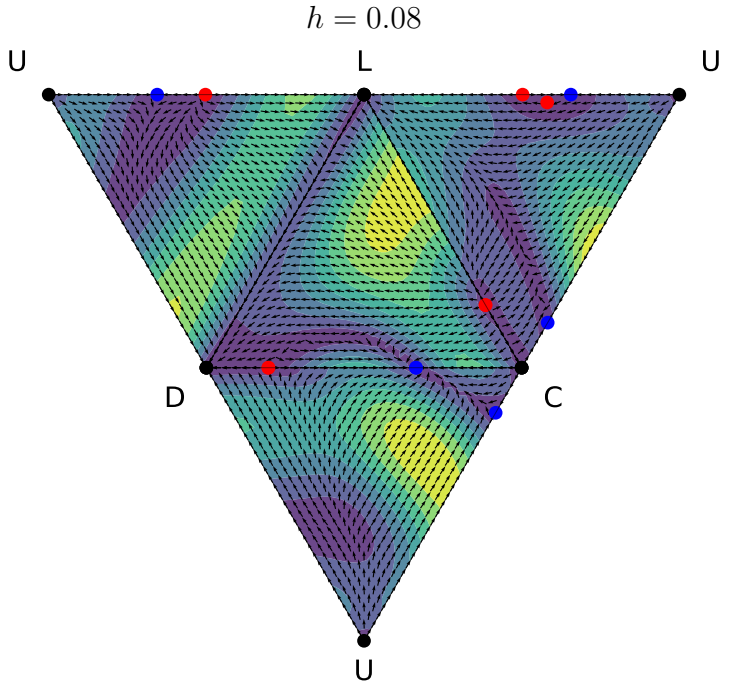


Table I.21: Steady states from `fixedpts_stability_sigmoidUDCL_v1_leader_driven_ngrid_9_q_9200.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
0	0	0	1		0.197070	0.177534	0.177070	
0	0	1	0		0.020000	-0.147413		-0.211705
0	1	0	0		0.171535		-0.061734	0.153818
1	0	0	0			-0.019695	-0.020000	-0.231705
0	0	0.496978	0.503022	unstable				
0	0	0.344087	0.655913	stable	0.020000	0.003452		
0	0.835001	0	0.164999	stable	0.094403		0.011896	
0	0.770548	0.229452	0	unstable				
0	0.028775	0.404749	0.566476	unstable				
0.803342	0.196658	0	0	unstable				
0.335266	0.664734	0	0	stable			-0.066275	-0.118239

The next events are the collision of the remaining repellor on the (L, U, C) face with the attractor of L+U coexistence, and the appearance of a pair of repellors on the (D, L, C) face (occurs by $h = 0.07$).

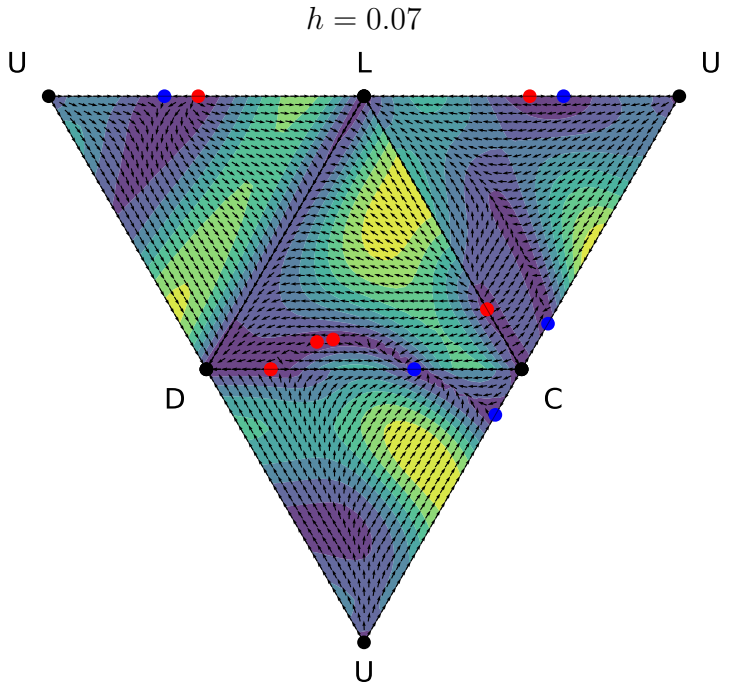


Table I.22: Steady states from `fixedpts_stability_sigmoidUDCL_v1_leader_driven_ngrid_9_q_9300.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
0	0	0	1		0.201021	0.181299	0.181021	
0	0	1	0		0.020000	-0.147985		-0.213512
0	1	0	0		0.173000		-0.058192	0.153594
1	0	0	0			-0.019817	-0.020000	-0.233512
0	0	0.474378	0.525622	unstable				
0	0	0.367048	0.632952	stable	0.020000	-0.001045		
0	0.780625	0.219375	0	unstable				
0	0.833047	0	0.166953	stable	0.098780		0.017660	
0.795757	0.204243	0	0	unstable				
0.340571	0.659429	0	0	stable			-0.062385	-0.121549
0.599144	0.300809	0.100047	0	unstable				
0.543608	0.347154	0.109238	0	unstable				

Finally, the attractor-repellor L+U pair collide. The evolutionary endpoints in a well-mixed population (no homophily) are all-D and a C+D coexistence.

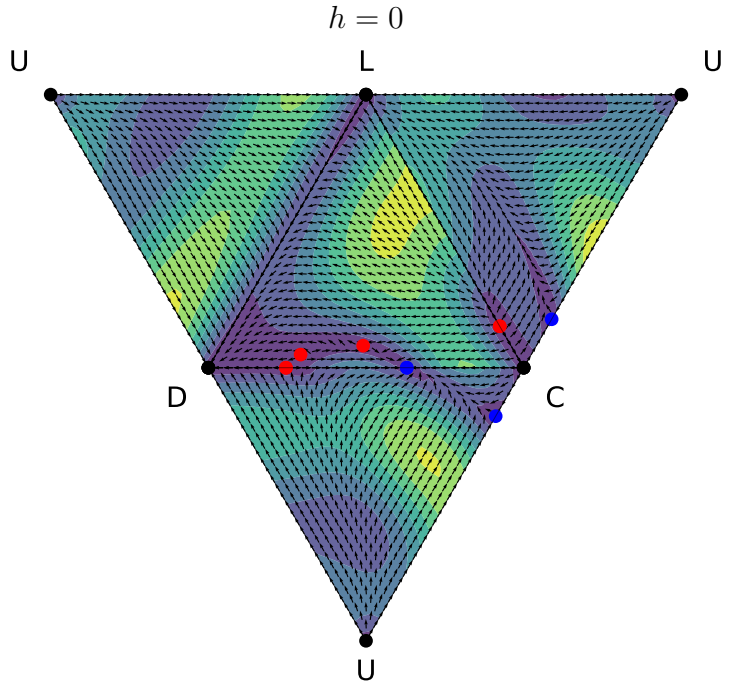


Table I.23: Steady states from `fixedpts_stability_sigmoidUDCL_v1_leader_driven_ngrid_9_q_10000.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
0	0	0	1		0.219861	0.199861	0.199861	
0	0	1	0		0.020000	-0.150636		-0.221018
0	1	0	0		0.176250		-0.036025	0.145444
1	0	0	0			-0.020000	-0.020000	-0.241018
0	0.822535	0	0.177465	stable	0.124994		0.051688	
0	0.848196	0.151804	0	unstable				
0.683425	0.268518	0.048057	0	unstable				
0.468771	0.450586	0.080642	0	unstable				
0.753897	0.246103	0	0	unstable				
0.370781	0.629219	0	0	stable			-0.037835	-0.139092

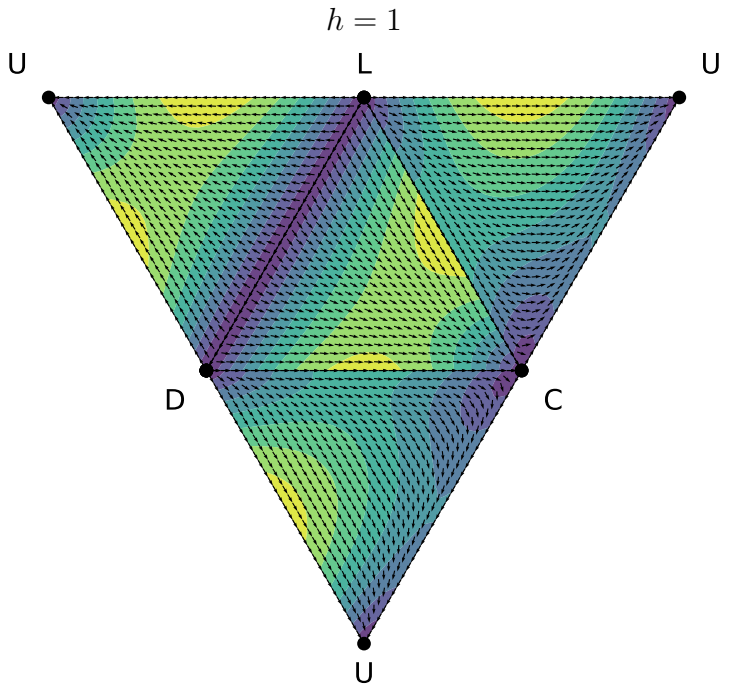
I.2 Example scenario 2

In example scenario 2, the parameter values are chosen such that the strategy profile with τ contributors and $n-\tau$ noncontributors is *not* a Nash equilibrium (Table I.24). This was achieved by reducing the steepness of the sigmoid benefits function (σ) in comparison to example scenario 2. Code to reproduce these results can be found in `scripts/transmat_sigmoid_UDCL/` and additional results can be found in `results/transmat_sigmoid_UDCL/`.

Table I.24: Parameter values used for example scenario 2.

parameter value	description
$n = 8$	Number of group members playing the game.
$\tau = 5$	Lottery quorum and midpoint of benefits function minus 0.5.
$\sigma = 6$	Steepness of the sigmoid benefits function.
$c = 0.25$	Cost of contributing to the public good.
$\varepsilon = 0.02$	Cognitive cost of being a communicative player (C or L).

At maximum homophily, with parameter $h = 1$ in the leader-driven group-formation model, the globally evolutionarily stable strategy is Unconditional Cooperation.



As homophily decreases, the first event (occurring before $h = 0.59$) is the appearance of a pair of steady states, one unstable and one stable, between U and C. The stable U+C coexistence is uninvadable by all other strategies, and there are no interior steady states.

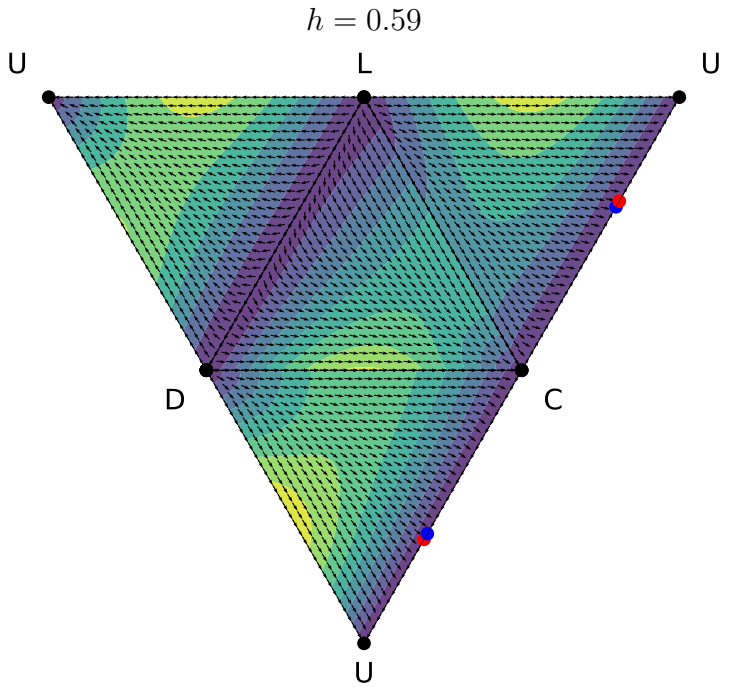
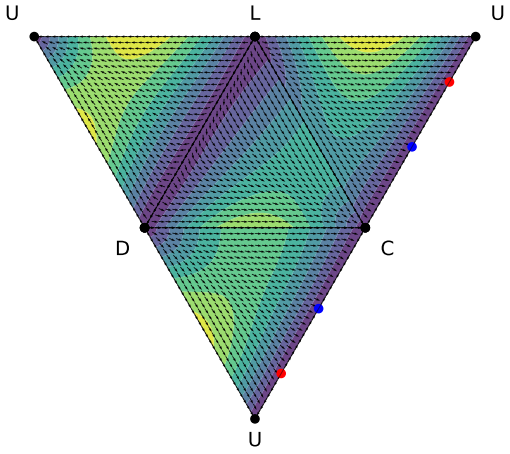


Table I.25: Steady states from `fixedpts_stability_transmat_sigmoidUDCL_v2_leader_driven_ngrid_9_q_4200.csv`

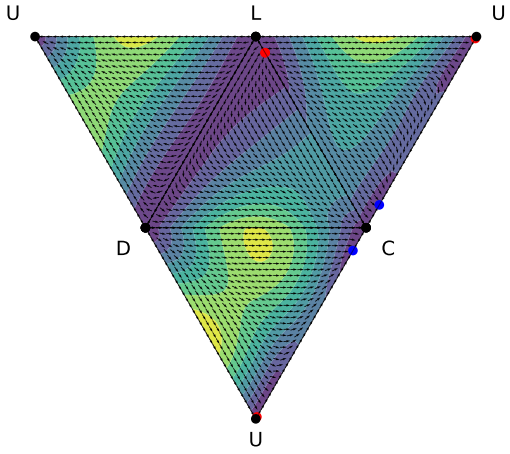
steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
1	0	0	0			0.177464	-0.020000	0.190443
0	1	0	0		-0.192100		-0.220019	0.056601
0	0	1	0		0.020000	0.055499		0.210443
0	0	0	1		-0.271503	-0.018685	-0.291503	
0	0.237578	0	0.762422	unstable				
0	0.576807	0	0.423193	stable	-0.294832		-0.262562	

As homophily decreases further, the two steady states separate, with the unstable steady-state moving towards U and the stable steady-state moving towards C. The unstable steady state collides with U and is annhilated whereas the stable steady-state persists. A separatrix appears between L and C, emerging near $p_L = 1$, and moves away from L and towards C as homophily declines.

$h = 0.58$

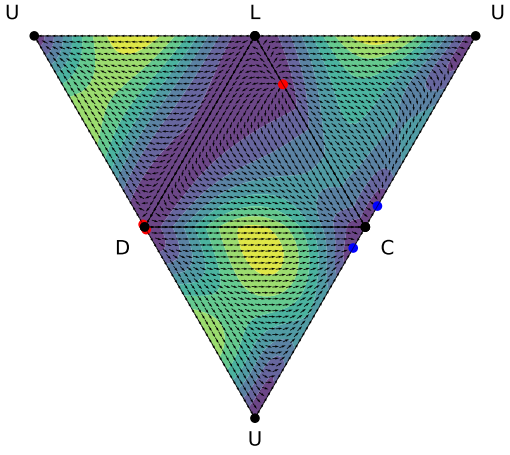


$h = 0.46$



The next events are the appearances of unstable states between D and U (emerging at $p_D = 1$ before $h = 0.4$) and between L and U (emerging at $p_L = 1$ before $h = 0.36$).

$h = 0.40$



$h = 0.36$

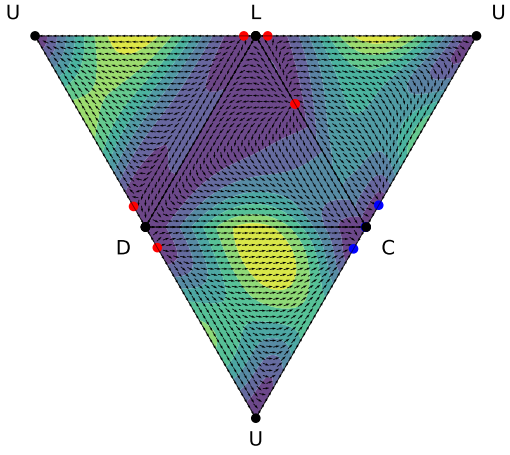


Table I.26: Steady states from `fixedpts_stability_transmat_sigmoidUDCL_v2_leader_driven_ngrid_9_q_6000.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
1	0	0	0			0.056491	-0.020000	-0.005170
0	1	0	0		-0.035833		-0.122982	0.054347
0	0	1	0		0.020000	-0.041738		0.014830
0	0	0	1		-0.076831	0.010250	-0.096831	
0.986388	0	0	0.013612	unstable				
0	0.253604	0.746396	0	unstable				
0	0.890236	0	0.109764	stable	-0.096036		-0.126183	

As homophily declines further, the next event is the appearance of a stable steady state between C and D (occurs before $h = 0.36$). This C+D coexistence is invadable by U but not by L.

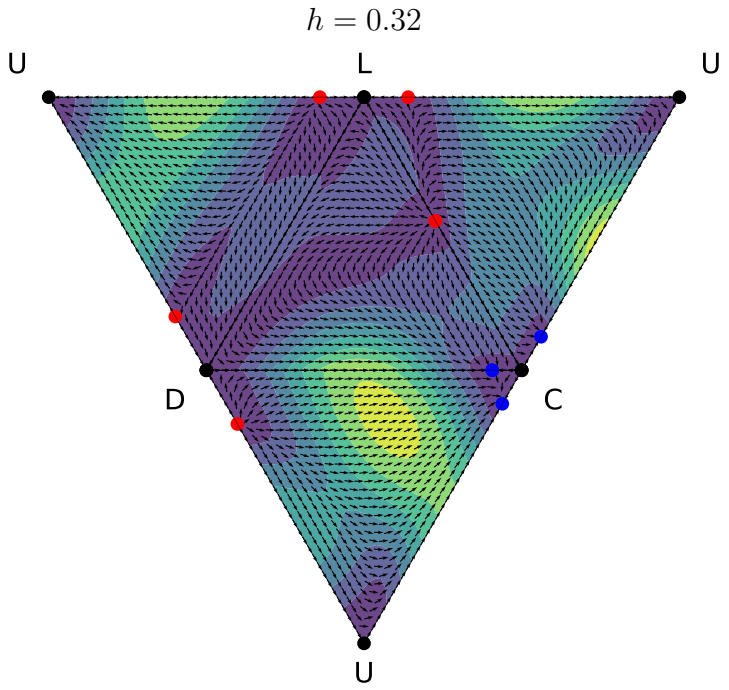


Table I.27: Steady states from `fixedpts_stability_transmat_sigmoidUDCL_v2_leader_driven_ngrid_9_q_6800.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p^*_D	p^*_C	p^*_L	p^*_U	on sub($\boldsymbol{p}^*$)	D	C	L	U
1	0	0	0			0.019101	-0.020000	-0.073095
0	1	0	0		0.036888		-0.084182	0.076929
0	0	1	0		0.020000	-0.072219		-0.053095
0	0	0	1		-0.002194	0.032532	-0.022194	
0.803292	0	0	0.196708	unstable				
0.093205	0.906795	0	0	stable			-0.086447	0.026316
0	0.453008	0.546992	0	unstable				
0	0.877209	0	0.122791	stable	-0.038003		-0.080534	
0	0	0.859934	0.140066	unstable				

As homophily declines further, a stable coexistence U+D emerges from $p_U = 1$ (occurs by $h = 0.315$), and this U+D coexistence is invadable by C.

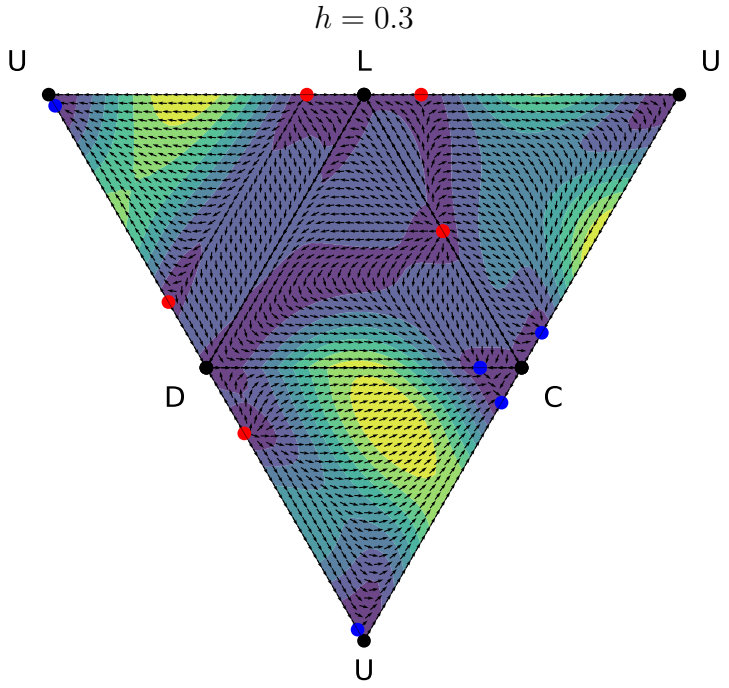
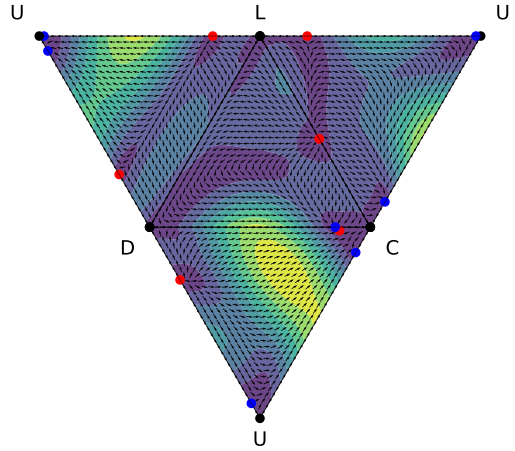


Table I.28: Steady states from `fixedpts_stability_transmat_sigmoidUDCL_v2_leader_driven_ngrid_9_q_7000.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
1	0	0	0			0.011870	-0.020000	-0.088008
0	1	0	0		0.054101		-0.075022	0.083492
0	0	1	0		0.020000	-0.078735		-0.068008
0	0	0	1		0.014919	0.039507	-0.005081	
0.759456	0	0	0.240544	unstable				
0.040844	0	0	0.959156	stable		0.024843	-0.020000	
0	0	0.818654	0.181346	unstable				
0.131519	0.868481	0	0	stable			-0.078071	0.008997
0	0.872054	0	0.127946	stable	-0.024848		-0.069229	
0	0.500143	0.499857	0	unstable				

As homophily declines further, between $h = 0.29$ and $h = 0.26$, a repellor appears on the (D,C,U) face near the C+D steady state and moves towards and collides with the C+U steady state. When the repellor appears, it renders the C+D steady state uninvadable, and when it collides with C+U, it renders the C+U steady state invadable by D.

$h = 0.284$



$h = 0.276$

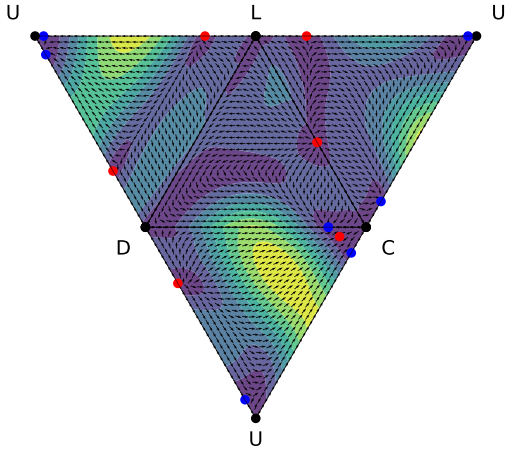


Table I.29: Steady states from `fixedpts_stability_transmat_sigmoidUDCL_v2_leader_driven_ngrid_9_q_7240.csv`

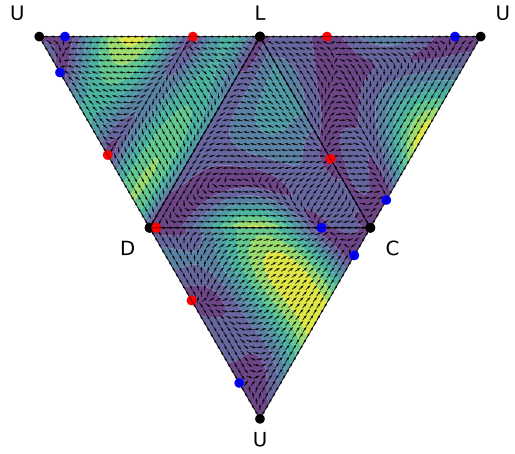
steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p^*_D	p^*_C	p^*_L	p^*_U	on sub($\boldsymbol{p}^*$)	D	C	L	U
1	0	0	0			0.004328	-0.020000	-0.104781
0	1	0	0		0.073856		-0.064344	0.091390
0	0	1	0		0.020000	-0.085999		-0.084781
0	0	0	1		0.034552	0.048569	0.014552	
0.706134	0	0	0.293866	unstable				
0.097363	0	0	0.902637	stable		0.014433	-0.020000	
0.171427	0.828573	0	0	stable			-0.068010	-0.010663
0	0	0.769833	0.230167	unstable				
0.095703	0.855105	0	0.049192	unstable				
0	0.865450	0	0.134550	stable	-0.009671		-0.055722	
0	0.556018	0.443982	0	unstable				
0	0	0.038640	0.961360	stable	0.020000	0.040190		

A separatrix appears between D and C making the all-D population an uninvadable state (occurs before $h = 0.26$).

Pairs of repellors appear on the (L, U, C) and (L, C, D) faces ($h = 0.232$).

Two repellors also appear on the (D, C, U) face (occurs just before $h = 0.232$). These repellors render the stable D+U coexistence uninvadable by C.

$h = 0.24$



$h = 0.232$

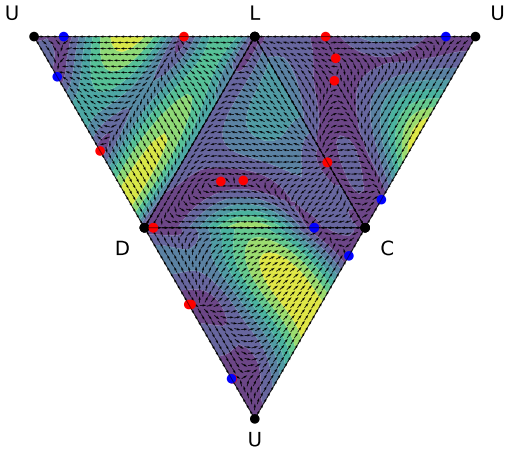
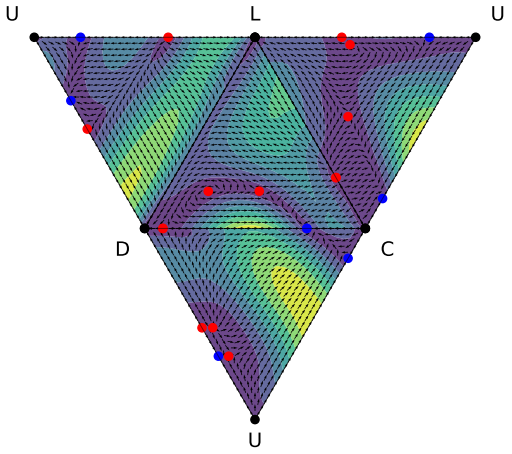


Table I.30: Steady states from `fixedpts_stability_transmat_sigmoidUDCL_v2_leader_driven_ngrid_9_q_7680.csv`

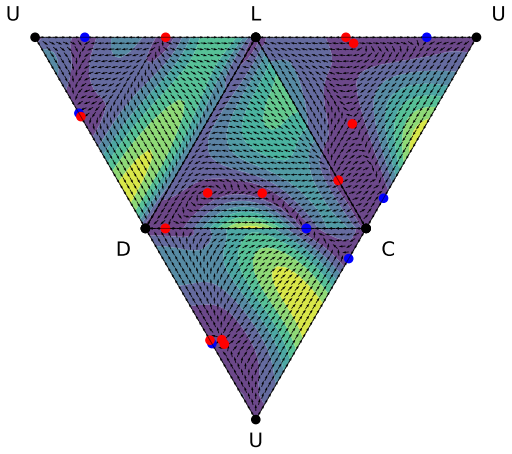
steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
1	0	0	0			-0.006389	-0.020000	-0.132326
0	1	0	0		0.106724		-0.045723	0.104933
0	0	1	0		0.020000	-0.097807		-0.112326
0	0	0	1		0.067836	0.066841	0.047836	
0.959620	0.040380	0	0	unstable				
0.230579	0.769421	0	0	stable			-0.049669	-0.043426
0.598422	0	0	0.401578	unstable				
0.210207	0	0	0.789793	stable		-0.000629	-0.020000	
0	0.658148	0.341852	0	unstable				
0	0.852864	0	0.147136	stable	0.016532		-0.031266	
0	0	0.679153	0.320847	unstable				
0	0	0.134227	0.865773	stable	0.020000	0.034903		
0.532484	0.225175	0.242341	0	unstable				
0.428340	0.324309	0.247351	0	unstable				
0	0.113620	0.576571	0.309809	unstable				
0	0.230979	0.523639	0.245382	unstable				
0.587834	0.010667	0	0.401499	unstable				
0.207131	0.003076	0	0.789793	unstable				

As homophily declines, the separatrix on the (D, C, U) face moves away from the U-D axis, and the pair of steady states between U and D move towards one another. When the steady states collide (occurs by $h = 0.193$), they annhilate one another, and a population that was at the U+D coexistence moves to an all-D state.

$h = 0.2$

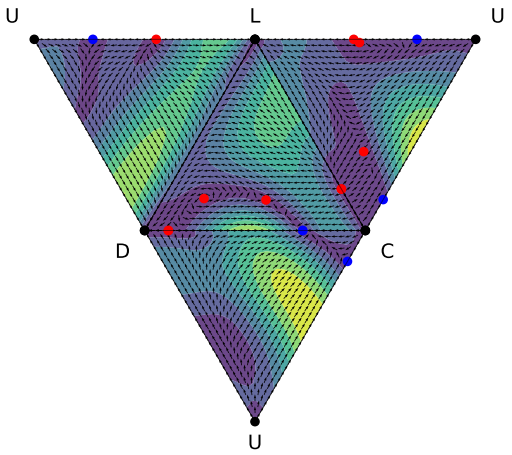


$h = 0.194$

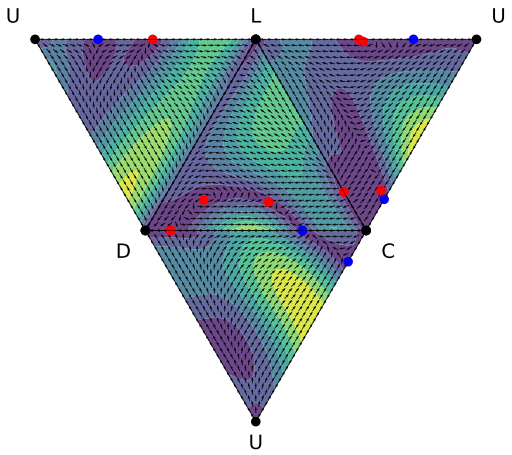


As homophily decreases further, the repellor on the (L, U, C) face moves towards the C+U coexistence. Once it collides (occurs before $h = 0.173$), it renders the C+U coexistence invadable by L.

$h = 0.18$

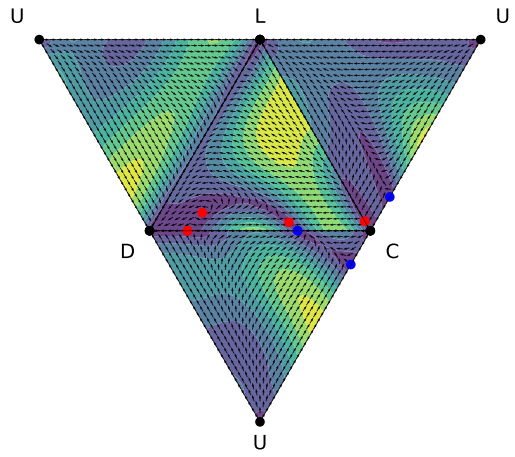


$h = 0.174$



As homophily declines further, the pair of steady states between L and U collide. The next major event occurs when the repellor between L and C collides with C and a repellor on the (L, D, C) face collides with the stable steady state C+D (occurs before $h = 0.1$). When the repellor collides with C, the all-C population becomes invadable by L, and when the repellor collides with C+D, the C+D coexistence becomes invadable by L.

$h = 0.12$



$h = 0.08$

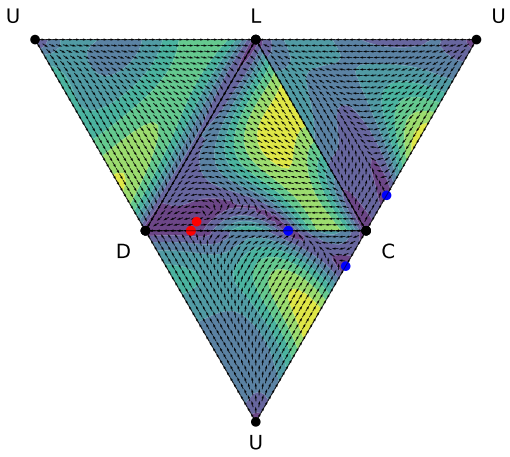
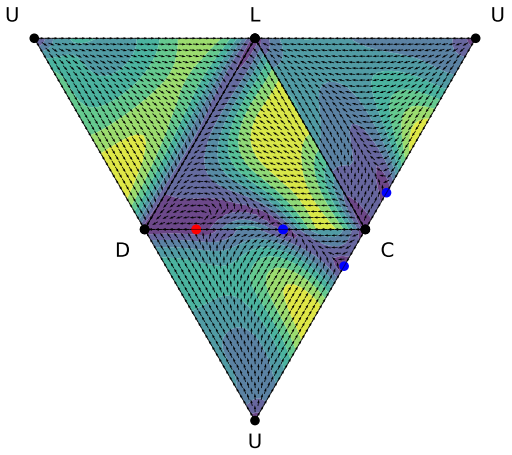


Table I.31: Steady states from `fixedpts_stability_transmat_sigmoidUDCL_v2_leader_driven_ngrid_9_q_9200.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p_D^*	p_C^*	p_L^*	p_U^*	on sub($\boldsymbol{p}^*$)	D	C	L	U
1	0	0	0			-0.019722	-0.020000	-0.196084
0	1	0	0		0.171794		0.007546	0.121187
0	0	1	0		0.020000	-0.125110		-0.176084
0	0	0	1		0.152746	0.133132	0.132746	
0.793923	0.206077	0	0	unstable				
0.353106	0.646894	0	0	stable			0.006873	-0.124494
0	0.814790	0	0.185210	stable	0.091035		0.044190	
0.744438	0.208122	0.047441	0	unstable				

The final even is when the repellor on the (L, U, C) collides with the unstable steady state between D and C, resulting in dynamics that are qualitatively similar to that of a well-mixed population (no homophily).

$h = 0.04$



$h = 0$

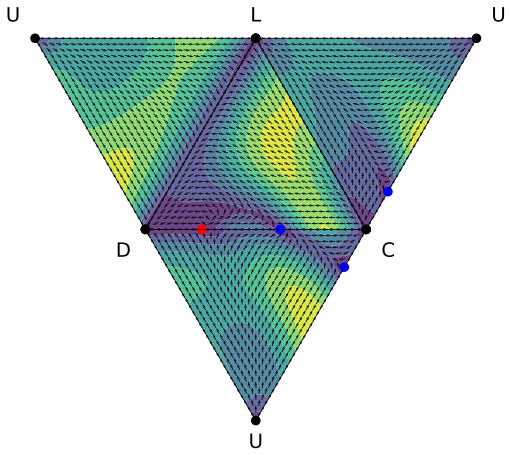


Table I.32: Steady states from `fixedpts_stability_transmat_sigmoidUDCL_v2_leader_driven_ngrid_9_q_10000.csv`

steady state $\boldsymbol{p}^*$				stability	invasion fitness of			
p^*_D	p^*_C	p^*_L	p^*_U	on sub($\boldsymbol{p}^*$)	D	C	L	U
1	0	0	0			-0.020000	-0.020000	-0.211683
0	1	0	0		0.176250		0.027456	0.106653
0	0	1	0		0.020000	-0.132302		-0.191683
0	0	0	1		0.177263	0.157263	0.157263	
0.744491	0.255509	0	0	unstable				
0.388576	0.611424	0	0	stable			0.027862	-0.148455
0	0.802645	0	0.197355	stable	0.118837		0.072508	

J Numerical tractability by reducing the number of strategies considered

As group size and number of strategies increases, there is a combinatorial increase in the number of group-composition probabilities that must be specified in order to parameterise the model. This is somewhat ameliorated by our approach, where we parameterise the problem in terms of the probabilities of family-partition outcomes rather than strategy-composition outcomes. However, the combinatorial nature of the problem is intrinsic, and ultimately constrains the size of the problem that can be solved.

First, let us compare the number of strategy compositions to family compositions. The total number of possible whole-group strategy compositions $\boldsymbol{g}_{\mathrm{a}}$ is (Jensen and Rigos, 2018)

$$\frac{(n+m-1)!}{n!(m-1)!}, \tag{J.1}$$

which grows rapidly with group size and number of strategies. Therefore, if a model is parameterised in terms of the probabilities of different group strategy compositions, the number of probabilities that must be specified also grows rapidly (dashed red line, Fig. J.6). In contrast, if strategy composition is determined genetic homophily only, then the model can be parameterised in terms of the probabilities of different family structures, which are fewer in number. The total number of possible family partition structures is equal to the partition number, which grows comparatively slowly (blue line, Fig. J.6).

However, to obtain the dynamics $\Delta\boldsymbol{p}$, one must still evaluate the payoffs for each possible group strategy composition given the family partition structures. This was achievable in our example because we modelled a scenario where the number of strategies was modest (i.e., up to $m = 4$ strategies used in our examples, solid red line, Fig. J.6). Consequently, the number of possible group strategy compositions had a similar magnitude to the partition number (true for group sizes up to approximately $n = 30$).

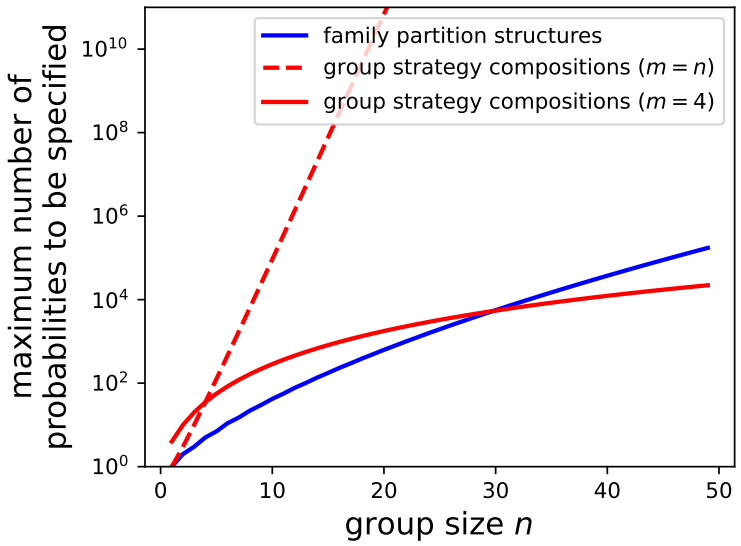


Figure J.6: An example of how the number of whole-group strategy compositions (red) and family partition structures (blue) increases as group size increases.

K General class of homophilic group-formation models under weak selection

The type of homophilic group-formation model we use (Kristensen et al., 2022) assumes that group formation depends only on family relationships and is independent of individuals' strategies. We assume that the probability distribution of family partition-structure probabilities $\{F_{\boldsymbol{q}}\}$ is independent of the strategy composition of the group. However, our approach can be applied to a broader class of homophilic group-formation models, where the family structure depends on the strategies, if one makes two additional assumptions. First, one assumes δ-weak selection, i.e., $\delta \to 0$. Second, one assumes that, under neutrality $\delta = 0$, the group formation probability agrees with what we have derived in our paper. Then, under weak selection, the dynamics can be approximated using our approach.

To demonstrate the idea, we first bring out selection-strength term δ from Eq. 7,

$$\Delta p_x \propto \delta \, \mathbb{C}\mathrm{ov}[G_{0,x}, \Pi_0], \tag{K.1}$$

which is of order δ. This renders Eq. 23, from the accounting based on whole-group composition, as

$$\Delta p_x \propto \delta \sum_{\boldsymbol{g}_\mathrm{a} \in \mathcal{G}_\mathrm{a}} \left(\frac{g_{\mathrm{a},x}}{n} \hat{\pi}(\boldsymbol{e}_x, \boldsymbol{g}_\mathrm{a}) - p_x \sum_{i=1}^{m} \frac{g_{\mathrm{a},i}}{n} \hat{\pi}(\boldsymbol{e}_i, \boldsymbol{g}_\mathrm{a}) \right) \mathbb{P}[\boldsymbol{G}_\mathrm{a} = \boldsymbol{g}_\mathrm{a}], \tag{K.2}$$

which is again of order δ. Here $\mathbb{P}[\boldsymbol{G}_\mathrm{a} = \boldsymbol{g}_\mathrm{a}]$ is the probability of the whole-group strategy composition is $\boldsymbol{g}_\mathrm{a}$. In a general class of models, $\mathbb{P}[\boldsymbol{G}_\mathrm{a} = \boldsymbol{g}_\mathrm{a}]$ may not agree with ours, but when we perform a Taylor expansion of $\mathbb{P}[\boldsymbol{Z} = \boldsymbol{z}]$ around $\delta = 0$, we have

$$\mathbb{P}[\boldsymbol{G}_\mathrm{a} = \boldsymbol{g}_\mathrm{a}] = \mathbb{P}[\boldsymbol{G}_\mathrm{a} = \boldsymbol{g}_\mathrm{a}]^{(0)} + \delta \, \mathbb{P}[\boldsymbol{G}_\mathrm{a} = \boldsymbol{g}_\mathrm{a}]^{(1)} + \mathcal{O}(\delta^2), \tag{K.3}$$

and from our assumption above, $\mathbb{P}[\boldsymbol{G}_\mathrm{a} = \boldsymbol{g}_\mathrm{a}]^{(0)}$ agrees with ours. Putting Eq. K.3 into Eq. K.2, we obtain

$$\Delta p_x \propto \delta \sum_{\boldsymbol{g}_\mathrm{a} \in \mathcal{G}_\mathrm{a}} \left(\frac{g_{\mathrm{a},x}}{n} \hat{\pi}(\boldsymbol{e}_x, \boldsymbol{g}_\mathrm{a}) - p_x \sum_{i=1}^{m} \frac{g_{\mathrm{a},i}}{n} \hat{\pi}(\boldsymbol{e}_i, \boldsymbol{g}_\mathrm{a}) \right) \mathbb{P}[\boldsymbol{G}_\mathrm{a} = \boldsymbol{g}_\mathrm{a}]^{(0)} + \mathcal{O}(\delta^2). \tag{K.4}$$

Note that terms of order δ and higher in Eq. K.3 are all absorbed in the $\mathcal{O}(\delta^2)$ term in Eq. K.4. Therefore, as long as we are interested in a weak selection regime, $|\delta| \ll 1$, one can still use our formula as an approximation, where errors are up to $\mathcal{O}(\delta^2)$.

L Overview of the code repository

Fig. L.7 provides an overview of the contents of the Github repository: github.com/nadiahpk/homophilic-many-strategy-PGG/. The repository has been archived with the Zenodo DOI: 10.5281/zenodo.14991697.

The best place to start is with the quickstart tutorials in the `tutorials` directory. From there, readers who are interested in using the whole-group accounting should refer to the worked examples in the `sigmoid_UDCL` directories, and readers who are interested in using the transformed payoff matrix approach should refer to the worked examples in the `transmat_sigmoid_UDCL` directories.

Figure L.7: Overview of the contents of the Github repository.