# Evolutionary dynamics on interdependent populations

Jesús Gómez-Gardeñes,[1,2] Carlos Gracia-Lázaro,[1,2] Luis Mario Floría,[1,2] and Yamir Moreno[2,3]

[1]*Departamento de Física de la Materia Condensada, University of Zaragoza, Zaragoza 50009, Spain*
[2]*Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Zaragoza 50018, Spain*
[3]*Departamento de Física Teórica, University of Zaragoza, Zaragoza 50009, Spain*

Although several mechanisms can promote cooperative behavior, there is no general consensus about why cooperation survives when the most profitable action for an individual is to defect, especially when the population is well mixed. Here we show that when a replicator such as evolutionary game dynamics takes place on interdependent networks, cooperative behavior is fixed on the system. Remarkably, we analytically and numerically show that this is even the case for well-mixed populations. Our results open the path to mechanisms able to sustain cooperation and can provide hints for controlling its rise and fall in a variety of biological and social systems.

PACS number(s): 89.75.Fb, 05.45.−a, 05.10.Gg

## I. INTRODUCTION

The onset of global cooperation in large populations of unrelated agents when defective actions provide the largest short-term benefits at the individual level constitutes one of the most amazing puzzles for evolutionary dynamics [1–4]. Recently, the structure of the interactions among individuals seems to have provided a way out for cooperation to survive in those scenarios, such as the Prisoner's Dilemma (PD) game, in which defective behaviors are evolutionarily favored under the well-mixed assumption [5–7]. Although recent results have shown that network reciprocity is not always a viable mechanism to explain cooperation among humans [8], larger cooperative levels are achieved if an evolutionary game dynamics takes place on top of structured populations and networks, in which nodes account for players and links represent the existence of game interactions. Moreover, further including real structural patterns of large systems [9] (scale-free distribution for the number of contacts a player has [10–13], the small-world properties [14,15], nonzero density of triads [16,17], etc.) provides also high cooperative outputs.

On the other hand, in most cases a real population—be it a biological or a social system—is not isolated and interactions take place at and between different levels (or layers) following different rules [18,19]. Think, for instance, of an economical system, where different levels account for different competitive markets and their interdependencies (developers, manufactures, providers). The rules governing the interactions at one layer are not necessarily the same as those driving the dynamics at another layer; admittedly, within each layer competition should exist while this is not necessarily the case for interlayer interactions. Thus, a natural question arises as to whether the observed degree of interdependency in real systems is a relevant factor for the emergence and survival of cooperative behavior.

The previous interdependency, which is also referred to as multiplexity, can be easily incorporated into the framework of any dynamical process by coupling two or more networked populations in which links between individuals of the same population involve a different dynamical relationship to those established between members of different populations [20–23]. In this paper, we focus on the case in which an evolutionary PD game drives the interactions between agents

of the same population. In turn, the existence of links between agents of different populations allow the two networks to interact. We will assume that the latter interactions are ruled by the Snowdrift (SD) game. In this way, defection is punished when facing other defectors outside the original population, thus balancing the evolutionary advantage that defectors find by exploiting cooperators in their respective populations.

We henceforth analyze what new emergent behavior results from the multilevel nature. As a first step in understanding multilevel structures, we investigate a system made up by two populations that interact through a number of links connecting nodes located at each subsystem. Let us remark that, in this simple framework, the two populations do not overlap (i.e., each individual belongs to only one of the two populations considered). Exact analytical calculations can be carried out for the case in which the population of each layer is well mixed, through the nonlinear analysis of the two-coupled-variable replicator equation for the strategic densities in both layers. Our results show the emergence of a *polarized* state in which all the individuals in one of the populations cooperate while all in the other population defect. In addition we find *quasipolarized* states, so that all the agents in one population are defectors, while most of the other one cooperate. Moreover, we also numerically show that the previous results hold for the case of networked populations. As we will discuss later on, our findings provide mechanisms for the rise and survival of cooperation and for its control.

Let us first describe the evolutionary dynamics of two interacting populations of size $N_1$ and $N_2$ (see Fig. 1). Two agents belonging to the same population $\alpha$ ($=1,2$) play a PD game so that a cooperator facing a cooperator (defector) in population $\alpha$ obtains a payoff $R = 1$ ($S = 0$). On the other hand, a defector facing a cooperator (defector) obtains a benefit of $T = b > 1$ ($P = r \geqslant 0$). The games played between agents of different populations follow the same parametrization except for the situation in which two defectors meet. In this case, the associated punishment is negative, $P = \epsilon < 0$, thus interpopulation games follow the SD formulation. Importantly, the strategists' competition for replication only occurs among own-population players. That is to say that there is no "interbreeding" (as it happens for different species, in biological contexts) or "strategic diffusion" (as for functionally
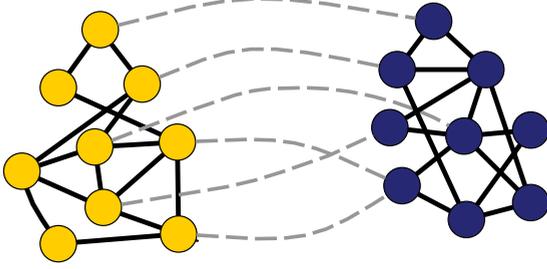
FIG. 1. (Color online) Schematic representation of the setup considered. Two networked populations of (in this case equal) sizes $N_1 = N_2 = 9$ interact through six dashed interpopulation links whereas individuals belonging to the same population interact via solid edges. The evolutionary game dynamics governing the interactions of dashed and solid links are different as described in the text.

heterogeneous layers in social or economical contexts) among the individuals of different populations. In terms of imperfect (and/or irrelevant) knowledge, the strategists from a population are unaware of the replicating success of strategies in the other population (and/or this information is irrelevant for its replication).

## II. WELL-MIXED POPULATION LIMIT

To start with, consider the case in which agents of the same population (layer) are well mixed. Let us also assume that both $N_1$ and $N_2$ are large enough (i.e., $N_1, N_2 \gg 1$). Under these simple assumptions, an exact analytical description via the analysis of the phase portrait of the two-dimensional replicator equation for two-by-two matrix games is possible. In our well-mixed population approximation an individual in population $\alpha$ has $N_\alpha - 1$ neighbors inside this population. Moreover, for interactions between the two layers, we suppose that any pair of nodes (each one of a different population) is present with probability $p$. Thus, the number of interpopulation links is equal to $pN_1N_2$.

Let us call $x_\alpha$ the fraction of cooperators in the population $\alpha$. The replicator equations for the evolutionary game dynamics

are

$$
\begin{aligned}
\dot{x}_1 &= x_1(1 - x_1)\{(N_1 - 1)[x_1(1 - b + r) - r] \\
&\quad + N_2 p[x_2(1 - b + \epsilon) - \epsilon]\} \\
\dot{x}_2 &= x_2(1 - x_2)\{(N_2 - 1)[x_2(1 - b + r) - r] \\
&\quad + N_1 p[x_1(1 - b + \epsilon) - \epsilon]\}.
\end{aligned}
\tag{1}
$$

The results of the theoretical analysis (see appendix) of these coupled deterministic equations are illustrated in Fig. 2 for the symmetric (thus nongeneric) case $N_1 = N_2$, and the simple weak ($r = 0$) PD game for those intrapopulation encounters. Below we will comment on the main qualitative changes for the generic case [i.e., whenever both the size proportion $\beta = N_1/N_2 \neq 1$ and general PD ($r > 0$) game for intrapopulation interactions apply].

The analysis of Fig. 2 shows a rather natural nonlinear resolution of the conflict introduced by fitness-punishment ($\epsilon$) to interpopulations defective encounters. Briefly said, even-symmetric ($x_1 = x_2$) states D (both populations are fully defective) and C (fully cooperative populations) are both, for any $b > 1^+$, unstable against perturbations in all directions, and stability resides instead on odd-symmetric polarized states [A (all D in population 1 and all C in population 2) and its symmetric transformed B (all C in population 1 and all D in population 2)] for strictly positive temptation $b$ less than a bound $b^{\text{up}}(\epsilon; p) = 1 - p\epsilon$ [see Fig. 2(a)]. At this critical (bifurcation) value of $b$ the interior nullclines $\dot{x}_1 = 0$ and $\dot{x}_2 = 0$ [see Fig. 2(b)] touch states A and B respectively. Increasing the value of the temptation $b$ above $b^{\text{up}}$ the polarized states lose their stability in favor of the quasipolarized states [A′ (all D in 1 and mostly C in 2) and its symmetric B′ ], which detach from A and B and become attractors. At $b = b^c = 1 - \frac{p\epsilon}{1-p}$ the interior nullclines coincide [see Fig. 2(c)] becoming a line (A′B′) of marginally stable equilibria. Finally, for $b > b^c$ [see Fig. 2(d)] the global attractor is the interior even-symmetric state E, the intersection of the interior nullclines, which keeps approaching, as $b$ increases, the neighborhood of the high $b$ limit attractor, say the state D of fully defective populations.

This scenario remains qualitatively unchanged for strictly positive values of the parameter $r$, provided $0 < r < -p\epsilon$, the only change being that the bifurcation value $b^c$ where the
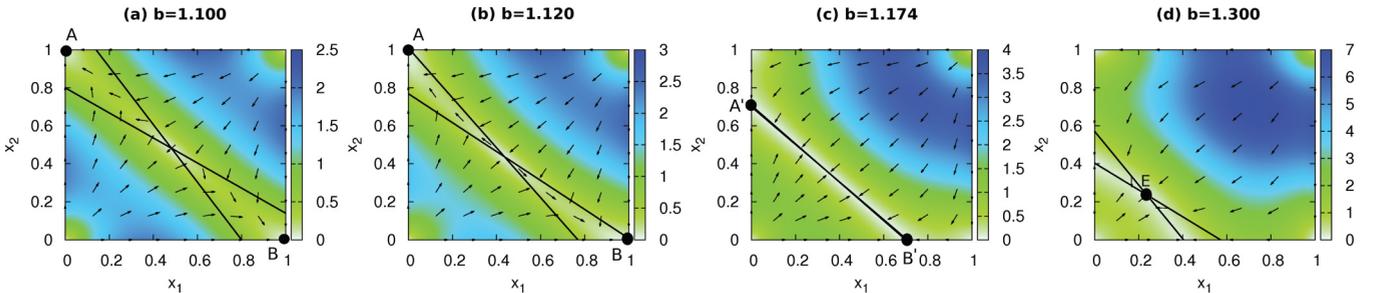


FIG. 2. (Color online) Phase portrait of replicator equation (1) for the symmetric case ($\beta = 1$) and weak PD ($r = 0$) intrapopulation game, for different values of the temptation $b$, with $p = 0.3$, and $\epsilon = -0.4$. The direction of velocity field is indicated by the arrows, and its modulus by the colors. We also plot the interior nullclines. For low values of $b$ (a), the polarized states A and B are attractors. They lose stability at $b = b^{\text{up}}$ (b), in favor of the quasipolarized states A′ and B′. These in turn destabilize at $b = b^c$ (c) when the nullclines coincide in a line of marginally stable equilibria. From there on, the interior equilibrium E becomes the global attractor (d).

quasipolarized states loose stability becomes

$$b^c = 1 + \frac{r - p\epsilon}{1 - p}. \tag{2}$$

In other words, the weak PD limit ($r = 0$) for the intrapopulation game is structurally stable with respect to (small enough) positive parametric variations of the game parameter $r$. For $r > -p\epsilon$, the scenario changes drastically: D is now a stable equilibrium, but still, for $b < b^{\mathrm{up}}$ (which doesn't depend on $r$), the polarized states are also stable equilibria. Only for larger $b > b^{\mathrm{up}}$ values of the temptation, D becomes the unique global attractor. Summarizing the results for the symmetric case, the attractor states for increasing values of $b$ from $b = 1^+$ follow the sequence

$$\mathrm{A,B} \xrightarrow{b^{\mathrm{up}}} \mathrm{A',B'} \xrightarrow{b^c} \mathrm{E}. \tag{3}$$

when $0 \leqslant r < -p\epsilon$ while, when $r > -p\epsilon$, the sequence is

$$\mathrm{D,A,B} \xrightarrow{b^{\mathrm{up}}} \mathrm{D}. \tag{4}$$

For the general case $N_1 \neq N_2$, the lack of the population interchange symmetry modifies some of the features seen in the symmetric case. Without loss of generality, we assume that $\beta = N_1/N_2 > 1$. On one hand, the lower bound of $r$ for the stability of the fully defective state D becomes now $r = -\beta p\epsilon$. On the other hand, the bifurcation values at which the polarized states lose their stability are now different,

$$b_B^{\mathrm{up}} = 1 - \frac{p\epsilon}{\beta} < b_A^{\mathrm{up}} = 1 - \beta p\epsilon, \tag{5}$$

as well as the bifurcation values (provided they exist) at which quasipolarized states destabilize, $b_B^c < b_A^c$, where

$$b_B^c = 1 + \frac{r^2 - (p\epsilon)^2}{(r + \beta p\epsilon) - p(\beta r + p\epsilon)} \tag{6}$$

$$b_A^c = 1 + \frac{\beta(r^2 - (p\epsilon)^2)}{(\beta r + p\epsilon) - p(r + \beta p\epsilon)}. \tag{7}$$

Let us note that the polarized state A, where the defective population is of larger size, turns out to have a wider range of stability, as well as a larger basin of attraction, than the state B. The results of the complete analysis of the replicator equation (1) are summarized in Table I, where we show the sequences of attractors coexisting in phase space. The seven scenarios (a)–(g) correspond to different ranges of values of the parameters $r$, $\beta$, $p$, and $\epsilon$ (see Appendix for further details).

### III. EVOLUTION ON RANDOM NETWORKS

From the previous analysis of well-mixed populations, one sees that polarized and quasipolarized states appear as generic attractors of the evolutionary dynamics for wide ranges of model parameters, which in turn has the effect of enhancing in a remarkable way the asymptotic levels of cooperation in the two-population system. On the other hand, for structured populations, where individuals interact with their neighbors as dictated by a given network of contacts, it is known that under some assumptions cooperation is enhanced, a phenomenon called network reciprocity [10,24].

While for well-mixed populations, the stability of polarized states extends down to $b = 1^+$, one should expect that at small $b > 1$ values, the enhancement of cooperative fluctuations

TABLE I. Sequence of attractors in phase space for Eq. (1), as $b$ increases from $b = 1^+$. The arrow indicates a bifurcation at the $b$ value that appears over the arrow. The scenarios (a)–(g) correspond to different ranges of values of the parameters $r$, $\beta$, $p$, and $\epsilon$, which are made explicit in the appendix. Note that except for the scenario (a), that corresponds to $r > -\beta p\epsilon$, polarized and quasipolarized states dominate the asymptotic behavior.

| Scenario | Sequence |
|---|---|
| (a) | $\mathrm{D,A,B} \xrightarrow{b_B^{\mathrm{up}}} \mathrm{D,A} \xrightarrow{b_A^{\mathrm{up}}} \mathrm{D}$ |
| (b) | $\mathrm{A,B} \xrightarrow{b_B^{\mathrm{up}}} \mathrm{A} \xrightarrow{b_A^{\mathrm{up}}} \mathrm{A'} \xrightarrow{b_A^c} \mathrm{E}$ |
| (c) | $\mathrm{A,B} \xrightarrow{b_B^{\mathrm{up}}} \mathrm{A} \xrightarrow{b_A^{\mathrm{up}}} \mathrm{A'}$ |
| (d) | $\mathrm{A,B} \xrightarrow{b_B^{\mathrm{up}}} \mathrm{A,B'} \xrightarrow{b_B^c} \mathrm{A} \xrightarrow{b_A^{\mathrm{up}}} \mathrm{A'}$ |
| (e) | $\mathrm{A,B} \xrightarrow{b_B^{\mathrm{up}}} \mathrm{A,B'} \xrightarrow{b_B^c} \mathrm{A} \xrightarrow{b_A^{\mathrm{up}}} \mathrm{A'} \xrightarrow{b_A^c} \mathrm{E}$ |
| (f) | $\mathrm{A,B} \xrightarrow{b_B^{\mathrm{up}}} \mathrm{A,B'} \xrightarrow{b_A^{\mathrm{up}}} \mathrm{A',B'} \xrightarrow{b_B^c} \mathrm{A'}$ |
| (g) | $\mathrm{A,B} \xrightarrow{b_B^{\mathrm{up}}} \mathrm{A,B'} \xrightarrow{b_A^{\mathrm{up}}} \mathrm{A',B'} \xrightarrow{b_B^c} \mathrm{A'} \xrightarrow{b_A^c} \mathrm{E}$ |

due to network reciprocity in the defective population 1 destabilizes the polarized states below some critical value $b^{\mathrm{low}}$. Moreover, one should also expect $b^{\mathrm{low}}$ to decrease with the parameter $p$, because higher values of $p$ increase the payoff that a (defector) individual in population 1 obtains from encounters with (cooperator) individuals of population 2, thus decreasing the resilience of cooperative fluctuations (network reciprocity) in population 1. In other words, for low values of $b$, the interaction between populations acts against network reciprocity. These expectations are fully confirmed by the results from simulations of the evolutionary dynamics in populations with a random network structure of intrapopulation contacts, using the discrete version of replicator dynamics.

In Fig. 3 we show the average cooperation $\langle c \rangle$ level (over a sample of 200 different realizations) on the two-
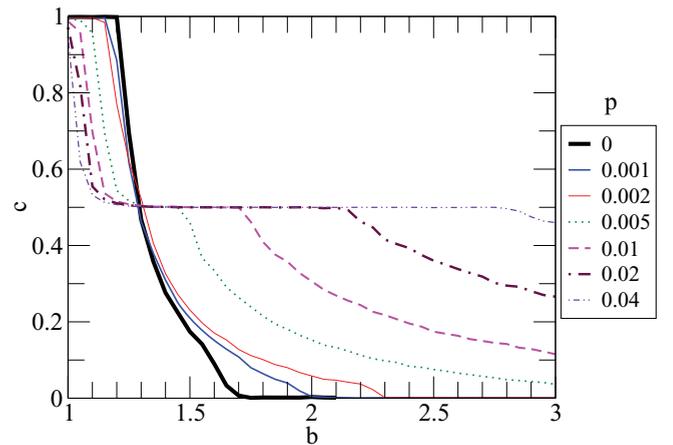


FIG. 3. (Color online) Average level of cooperation in the two-population system as a function of $b$, for different values of the fraction $p$ of interpopulation contacts. Other parameters are $r = 0$, $\epsilon = -0.4$, $N_1 = N_2 = 10^3$. The two populations have a random (Erdös-Rényi) network of contacts with average degree $\langle k \rangle = 6$. See the text for further details.
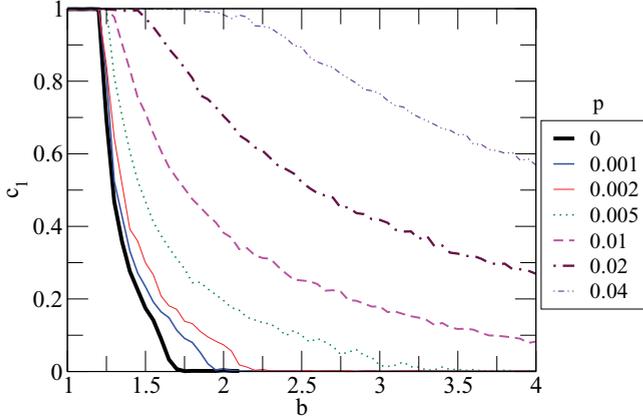
FIG. 4. (Color online) Average level of cooperation in the population 1 as a function of $b$, for different values of the fraction $p$ of interpopulation contacts. The population 1 (of size $N_1 = 10^3$) has been coupled to a smaller population 2 ($N_2 = 10^2$). While initial strategies in population 1 are equiprobables (random initial conditions), the population 2 starts from the absorbent state of full defection. Other parameters are $r = 0$, $\epsilon = -0.4$. Both populations have a random (Erdös-Rényi) network of contacts with average degree $\langle k \rangle = 6$.

population system as a function of $b$ for different values of $p$ and parameters as indicated. The two populations have a random (Erdös-Rényi [9]) network of contacts with average degree $\langle k \rangle = 6$. In the initial conditions, the individuals of both populations were chosen cooperators with probability 1/2. The plateau at $\langle c \rangle = 1/2$ points out the asymptotic polarized state. Moreover, the states with $\langle c \rangle < 1/2$ correspond to quasipolarized regimes where all the individuals in one population are defectors, while those with $\langle c \rangle > 1/2$, at values of $b < b^{\text{low}}$, result from states where all the individuals in one population are cooperators. This represents a type of quasipolarized states that were not found to be attractors of the dynamics for well-mixed populations. The comparison with the average cooperation level for noninteracting populations ($p = 0$ in Fig. 3) confirms that for low values of $b$ the interpopulation interaction acts against network reciprocity. From a complementary perspective, the networked populations show new attractors, impossible to be so for coupled well-mixed populations because they are the effect of network reciprocity. On the other hand, for larger values of $b$, the populations' coupling favors the achievement of substantial levels of cooperation, well beyond the typical values of $b$ for which network reciprocity ceases to be effective, this being an effect already present in the well-mixed case. This clarifies further the confluent effects of these two different mechanisms of cooperation enhancement.

Finally, the robustness of polarized and quasipolarized states suggests the use of the coupling to a defective population as an engineered (control) procedure to induce high levels of cooperation in a target population. To check for this possibility, we have coupled a large population 1 with random (equiprobable in strategies) initial conditions to a smaller defective population 2. In Fig. 4 we show the asymptotic average level of cooperation in a target population of size $N_1 = 10^3$ for different values of the average number, $N_2 p$, of interpopulation contacts per individual of the target population.

## IV. CONCLUSION

Summarizing, two PD populations SD coupled in conditions of strict inbreeding (no interpopulation strategic diffusion) evolve easily to polarized and quasipolarized strategic probability densities in the well-mixed thermodynamical limit of the evolutionary replicator dynamics. This happens also when population structure is a complex network of contacts, where other mechanisms (known as network reciprocity) of enhanced cooperation also operate. The confluence of both mechanisms has been analyzed in depth showing that polarization opposes network reciprocity at small values of the temptation parameter, while both act (synergy) together enhancing cooperation in one of the layers for higher temptation values. This phenomenon, that could be rationalized as the effect of incorporating a punishment to defective interpopulation encounters, illustrates the remarkable effects that structural multiplexity introduces in evolutionary dynamics.

## APPENDIX: PHASE PORTRAIT ANALYSIS OF THE TWO-VARIABLE REPLICATOR EQUATION

The replicator equation that describes the continuum time evolution of the cooperator fractions $x_1(t)$, $x_2(t)$ in subpopulations 1 and 2 can be written as

$$\dot{x}_1 = \mathcal{F}_1(x_1, x_2), \quad \dot{x}_2 = \mathcal{F}_2(x_1, x_2), \tag{A1}$$

where the velocities $\mathcal{F}_{1,2}$, after time rescaling, are explicitly given as

$$\begin{aligned}
\mathcal{F}_1(x_1, x_2) &= x_1(1 - x_1)\{\beta[x_1(1 - b + r) - r] \\
&\quad + p[x_2(1 - b + \epsilon) - \epsilon]\}, \\
\mathcal{F}_2(x_1, x_2) &= x_2(1 - x_2)\{[x_2(1 - b + r) - r] \\
&\quad + \beta p[x_1(1 - b + \epsilon) - \epsilon]\}.
\end{aligned} \tag{A2}$$

The unit square $0 \leqslant x_1, x_2 \leqslant 1$ is the invariant set of interest here. To follow the phase portrait variation of a two-degrees of freedom nonlinear system such as Eq. (A1) is pretty straightforward for one-parameter variations. We are dealing with a model where $b$, $r$, $\epsilon$, $\beta$, and $p$ are free model parameters, each one inside their natural range (i.e., $b > 1^+$, $0 \leqslant r \leqslant 1$, $\epsilon < 0^-$, $\beta \geqslant 1$, and $0 \leqslant p \leqslant 1$). In our systematics below, we will consider continuum variation of $b$, from $b = 1^+$ up to infinity, at fixed values of the other parameters and so we will obtain the critical (bifurcation) points $b^*(\epsilon, r; \beta, p)$, where the phase portrait of the evolution experiences *qualitative* changes: the direction of increasing temptation $b$ is most often considered in recent literature on PD games. But we will pay due attention also to variations of the parameter $r$ and find two important critical values that do not depend on the value of the temptation $b$ so that different scenarios of phase transitions (inside the well-mixed population approximation

to the thermodynamical limit $N_1, N_2 \to \infty$) as $b$ varies do appear. Finally, we choose also $\beta$ as an interesting (e.g., for control applications) parameter to vary, and find also two critical values that are temptation independent, which in turn increases the number of those scenarios.

The best visualization of the velocity field is a phase portrait where fixed (equilibrium) points and nullclines are also plotted, as in Fig. 1 in the main text. A nullcline is the locus of points defined by $\mathcal{F}_i(x_1, x_2) = 0$ for some $i$. The nullclines that correspond to $\mathcal{F}_1(x_1, x_2) = 0$ are the straight lines

$$x_1 = 0, \tag{A3}$$

$$x_1 = 1, \tag{A4}$$

$$x_2 = \frac{-x_1 \beta(b - 1 - r) - (\beta r + p\epsilon)}{p(b - 1 - \epsilon)}, \tag{A5}$$

while those that correspond to $\mathcal{F}_2(x_1, x_2) = 0$ are

$$x_2 = 0, \tag{A6}$$

$$x_2 = 1, \tag{A7}$$

$$x_2 = \frac{-x_1 \beta p(b - 1 - \epsilon) - (r + \beta p\epsilon)}{(b - 1 - r)}. \tag{A8}$$

The possible equilibria are the crossing points of any line from the first group with any other line from the second one, so there are nine candidates. Moreover, only solutions in the unit square, $0 \leqslant x_1, x_2 \leqslant 1$, interest us and this excludes two of the crossing points (see below) leaving the following seven possibilities, namely the four corners of the unit square

   (i)   $A = (0,1)$,
   (ii)   $B = (1,0)$,
   (iii)   $C = (1,1)$,
   (iv)   $D = (0,0)$,

and those whose location depends on parameter values,

   (i) We call $A'$ the crossing point of nullclines (A3) and (A8), whose coordinates are $x_1(A') = 0$ and

$$x_2(A') = \frac{-(r + \beta p\epsilon)}{(b - 1 - r)}. \tag{A9}$$

   (ii) We call $B'$ the crossing point of nullclines (A5) and (A6), so that $x_2(B') = 0$ and

$$x_1(B') = \frac{-(\beta r + p\epsilon)}{\beta(b - 1 - r)}. \tag{A10}$$

   (iii) Finally, we call $E$ the crossing of (A5) and (A8). Its coordinates are obtained as:

$$x_1(E) = \frac{(b - 1 - r)(\beta r + p\epsilon) - p(b - 1 - \epsilon)(\beta r + p\epsilon)}{\beta[(p(b - 1 - \epsilon)^2 - (b - 1 - r)^2]}, \tag{A11}$$

$$x_2(E) = \frac{(b - 1 - r)(r + \beta p\epsilon) - p(b - 1 - \epsilon)(r + \beta p\epsilon)}{(p(b - 1 - \epsilon)^2 - (b - 1 - r)^2}. \tag{A12}$$

The (missing in the list) crossings of (A4)–(A8), and of (A5)–(A7), are easily seen to be always outside the unit square for the range of parameters considered. Also inside this range, the nongeneric event of *nullclines' coincidence* only could happen provided $\beta = 1$ and $r > -p\epsilon$, at a value $b^c(\epsilon, r; \beta = 1, p) = 1 + \frac{r - p\epsilon}{1 - p}$. Only then, the exotic (forced by symmetry) situation in which there is a segment of marginally stable equilibria occurs.

To determine the bifurcation points, one uses the spectral analysis of tangent space perturbations around equilibria. The linearized evolution of small perturbations around the fixed point $\mathbf{x}^*$ is given by the matrix

$$\begin{pmatrix} \frac{\partial \mathcal{F}_1}{\partial x_1} & \frac{\partial \mathcal{F}_1}{\partial x_2} \\ \frac{\partial \mathcal{F}_2}{\partial x_1} & \frac{\partial \mathcal{F}_2}{\partial x_2} \end{pmatrix}_{\mathbf{x} = \mathbf{x}^*} \tag{A13}$$

In what follows, the presentation of the results from the phase portrait analysis of the nonlinear coupled ODE (A2) tries to rationalize them in terms of evolutionary game theoretic concepts, within a thermodynamical limit (statistical physics) perspective.

### 1. Symmetric case: $N_1 = N_2(= N)$.

For simplicity, as well as to illustrate neatly the systematics that we follow, we analyze first the case of equal population sizes. For this case, where populations are identical (though distinguishable), the population interchange symmetry imposes that phase portrait is invariant under permutation of coordinates $(x_1 \leftrightarrow x_2)$, a nongeneric property that limits severely the possible scenarios. The stability analysis of the equilibria shows that there are two generic scenarios for the sequence of bifurcations that appear when $b$ increases from $1^+$ up to infinity.

($s_1$) If $r > r_c = -p\epsilon$ there is only one bifurcation at $b^{\mathrm{up}}(r, \epsilon, \beta = 1, p) = 1 - p\epsilon$. For $b < b^{\mathrm{up}}$, the phase portrait has three stable equilibria with their own basins of attraction: D, A, and B. The equilibria C, $A'$, and $B'$ are unstable, and E is outside the unit square. At $b = b^{\mathrm{up}}$, A and B destabilize (through collision with $A'$ and $B'$ that exit the unit square) becoming saddle equilibria, and D becomes the unique global attractor for $b > b^{\mathrm{up}}$. This translates into the following sequence of attractors when temptation increases from $1^+$:

$$\mathrm{D,A,B} \;\xrightarrow{b^{\mathrm{up}}}\; \mathrm{D}. \tag{A14}$$

($s_2$) If $r < r_c(p, \epsilon)$, however, D is always unstable, and there are two bifurcations at $b^{\mathrm{up}}$ and $b^c$ (and note that $b^{\mathrm{up}} < b^c$). For $b < b^{\mathrm{up}}$ the equilibria C, D are sources, E is a saddle, and A and B are attractors, becoming saddle equilibria at $b^{\mathrm{up}}$ where $A'$ and $B'$ enter into the unit square. For $b^{\mathrm{up}} < b < b^c$ $A'$ and $B'$ are the only attractors. At $b^c$ the segment $A'B'$ of marginally stable equilibria is the limit set for all trajectories (nullcline's coincidence). For $b > b^c$ E becomes the unique (and even-symmetric) global attractor. This last bifurcation restores the symmetry of the asymptotic evolution that was spontaneously broken at lower $b$ values. The sequence of stationary limiting (point) densities is

$$\mathrm{A,B} \;\xrightarrow{b^{\mathrm{up}}}\; \mathrm{A',B'} \;\xrightarrow{b^c}\; \mathrm{E}. \tag{A15}$$

Note that the condition $r = r_c(p, \epsilon)$ that separates the regimes where the equilibrium D is unstable ($r < -p\epsilon$) or attractor ($r > -p\epsilon$), corresponds to the exact compensation of the surplus $rN$ of defective intrapopulation interactions of a defector and the punishment $p\epsilon N$ it receives from interpopulation interactions. Below this critical value, full defection is unstable to cooperative fluctuations. But, as we have just seen, even in case the punishment from coupling is weaker than surplus, polarized states have their own basins of

attraction, away from whole defection, at low values of $b > 1^+$. This can be rationalized from the role that punishment plays in our no interbreeding, punishing defective coupling setting. Populations' strategic polarization emerge as stable generic asymptotic states of evolution, even when defectors can afford external punishment (D being then fully stable): The duplex (two coupled populations) has always the option to become polarized or quasipolarized provided the initial conditions belong to its basin of attraction.

### 2. General case: $N_1 \neq N_2$.

The parameter $p$ determines the fraction of interpopulation-to-intrapopulation interactions any agent plays per unit time in the symmetric ($N_1 = N_2$) case. This fraction changes to $\beta p$ and $p/\beta$ ($\beta > 1$) for small and large populations respectively, when symmetry of population interchange is absent. This combination of parameters regulates how important to the replicating power (fitness) of an individual the interpopulation coupling is, and we then see that for the largest population the effective coupling $p/\beta$ is smaller. This makes the polarized state A (where population 1 is defective) more robust than the polarized state B, and provided both are attractors, the basin of attraction of A is correspondingly larger. This is a major qualitative change in the phase portrait of the velocity field of evolution in the absence of symmetry. The concomitant change is the shift, and in more extreme cases the disappearance, of the bifurcations associated to the quasipolarized equilibria A' and B' (i.e., $b^{\mathrm{up}}_{A,B}$ and $b^c_{A,B}$)

$$b^{\mathrm{up}}_B(r,\epsilon;\beta,p) = 1 - (p/\beta)\epsilon, \qquad (A16)$$

$$b^{\mathrm{up}}_A(r,\epsilon;\beta,p) = 1 - \beta p\epsilon, \qquad (A17)$$

$$b^c_B(r,\epsilon;\beta,p) = 1 + \frac{r^2 - (p\epsilon)^2}{(r + \beta p\epsilon) - p(\beta r + p\epsilon)}, \quad (A18)$$

$$b^c_A(r,\epsilon;\beta,p) = 1 + \frac{\beta(r^2 - (p\epsilon)^2)}{(\beta r + p\epsilon) - p(r + \beta p\epsilon)}. \quad (A19)$$

Note that the minimum of this set of values is $b^{\mathrm{up}}_B$, its maximum is $b^c_A$, and that the relative order of the other two values is parameter dependent. Several generic scenarios of phase portrait variations naturally follow from these major effects, when the population interchange symmetry is absent. Still, let us remark that the evolutionary attractiveness of the odd-symmetric polarized (A and B) and quasipolarized (A' and B') asymptotic densities still dominates ample regions of parameter space.

A first scenario, similar to the first one seen above for the symmetric case, is found when $r > r^A_c(\epsilon; \beta, p) = -\beta p\epsilon$. In this scenario, the fully defective state D is stable for all $b > 1$ values. For very low values of $b$, A and B are also stable. Due to asymmetry, the instabilities of A and B occur at different bifurcation values, $b^{\mathrm{up}}_B < b^{\mathrm{up}}_A$, so that state B destabilizes first when $b$ increases from $b = 1^+$, as expected, i.e.,

(i) If $r^A_c < r$ there are only two bifurcations at $b^{\mathrm{up}}_B < b^{\mathrm{up}}_A$. For all $b > 1^+$, C is unstable and E is outside the unit square. For $b < b^{\mathrm{up}}_B$, the states D, A, and B are attractors. At $b^{\mathrm{up}}_B$, B collides with the unstable B' that exits the unit square, then becoming a saddle with unstable direction corresponding to defective fluctuations in cooperative population 1. The same happens *mutatis mutandi* ($1 \leftrightarrow 2$ interchange) to A at $b^{\mathrm{up}}_A$, leaving finally D (for $b > b^{\mathrm{up}}_A$) as the global attractor.

$$\mathrm{D,A,B} \overset{b^{\mathrm{up}}_B}{\to} \mathrm{D,A} \overset{b^{\mathrm{up}}_A}{\to} \mathrm{D}. \qquad (A20)$$

At $r = r^A_c$, for a defective individual in population 2, and state D, the internal surplus coupling punishment balance exactly compensates. This means that changing to cooperator makes no difference to its replicating power, and thus a zero eigenvalue appears in the spectrum of the Jacobian (linear stability) matrix of the fully defective state D. Inside the range $r < r^A_c$, D is always unstable faced with cooperative fluctuations in the smaller population. Further down in surplus ($r$) values, at $r = r^B_c = -(p/\beta)\epsilon$, D becomes also unstable faced with cooperative fluctuations in the large population. In other words, when decreasing $r$ from large (compared to $r^A_c$) positive values of intrapopulation surplus, to $0^+$ (weak PD limit), there are two critical values, where qualitative changes of the phase portrait occur, which coincide with the change of stability of D from stable ($r > r^A_c$) to saddle ($r^B_c < r < r^A_c$), to source ($r < r^B_c$).

Provided $r < r^A_c$, if one considers the high $b$ ($\to \infty$) limit, one easily finds that it can be either mixed type (state E, interior to the unit square) or quasipolarized (state A', on the vertical $x_1 = 0$) regarding its convergence to virtually full defection. The transition between these two qualitatively different high temptation limit behaviors, for given values of $\epsilon, p$, and $r$, is controlled by the value of the population ratio $\beta$ and it occurs at the critical value

$$\beta^A_c(\epsilon, r;\ p) = \frac{p(r - \epsilon)}{r - p^2\epsilon}. \qquad (A21)$$

At this value of the population ratio, the bifurcation value $b^c_A$ (where A' collides with state E, this one entering into the unit square) formally diverges, so that the collision occurs (or doesn't), depending on the value of the population ratio $\beta$, for fixed value of $p$, $r$, and $\epsilon$.

On the other side, the bifurcation value at $b^c_B$ only occurs provided $r < r^B_c$, but its relative order with respect to $b^{\mathrm{up}}_A$ depends also on the value of $\beta$ with a critical value at

$$\beta^B_c(\epsilon, r;\ p) = \frac{-p\epsilon(p^2\epsilon - r) - \sqrt{p^2\epsilon^2(p^2\epsilon - r)^2 - 4p^2\epsilon(r - \epsilon)(p^2\epsilon^2 - r^2)}}{2p^2\epsilon(r - \epsilon)}. \qquad (A22)$$

The different possible combinations of all the previous possibilities give the following scenarios.

(ii) If $r^B_c < r < r^A_c$, then the stable linear manifold of the saddle point D ($x_2 = 0$) does not allow B' to be a stable

equilibrium, while its unstable direction ($x_1 = 0$) pushes evolution to polarized A or quasipolarized A′ states; C is a always a source for all $b > 1$. Two different scenarios are realized depending on the interpopulation ratio value, $\beta$.

(ii$_1$) If $\beta > \beta_c^A$ [see Eq. (A21)], bifurcations only occur at $b_B^{\mathrm{up}} < b_A^{\mathrm{up}}$. At $b_B^{\mathrm{up}}$, the collision of B and the unstable exiting B′ occurs, while at $b_A^{\mathrm{up}}$, it takes place the collision of A with the entering state A′. The corresponding sequence of attracting equilibria is given by

$$\mathrm{A,B} \ \overset{b_B^{\mathrm{up}}}{\to} \ \mathrm{A} \ \overset{b_A^{\mathrm{up}}}{\to} \ \mathrm{A'}. \qquad (A23)$$

(ii$_2$) If $\beta_c^A > \beta$, besides the bifurcations described in (ii$_1$), there is an additional bifurcation at $b_A^c$, where A′ collides with state E that enters into the unit square. The corresponding sequence of attracting equilibria is given by

$$\mathrm{A,B} \ \overset{b_B^{\mathrm{up}}}{\to} \ \mathrm{A} \ \overset{b_A^{\mathrm{up}}}{\to} \ \mathrm{A'} \ \overset{b_A^c}{\to} \ \mathrm{E}. \qquad (A24)$$

The presence or absence of the bifurcation $b_A^c$ determines whether the approach to the high temptation limit is via mixed interior type E state, or edge quasipolarized type A′ state, so that for values of $\beta$ below critical ($\beta_c^A$), virtually full defection $(1^-, 1^-)$ is approached with nonzero cooperation levels in both populations as $b$ diverges.

(iii) If $r < r_c^B$, both quasipolarized states A′ and B′ enter into the unit square at $b_B^{\mathrm{up}}$ and $b_A^{\mathrm{up}}$, respectively. B′ always destabilizes at $b_B^c$ ($> b_B^{\mathrm{up}}$ always) to become a saddle through collision with the exiting unstable interior equilibrium E. This may happen before [as in (iii$_1$) and (iii$_2$) below] or after [as in (iii$_3$) and (iii$_4$)] the entrance of A′ at $b_A^{\mathrm{up}}$ depending on $\beta$

value (relative to $\beta_c^B$). And finally note that the bifurcation at $b_A^c$ only occurs for $\beta < \beta_c^A$, as analyzed above, to arrive at the following possible four scenarios.

(iii$_1$) If $\max(\beta_c^A, \beta_c^B) < \beta$, then $b_B^c < b_A^{\mathrm{up}}$, and $b_A^c$ is absent

$$\mathrm{A,B} \ \overset{b_B^{\mathrm{up}}}{\to} \ \mathrm{A,B'} \ \overset{b_B^c}{\to} \ \mathrm{A} \ \overset{b_A^{\mathrm{up}}}{\to} \ \mathrm{A'}. \qquad (A25)$$

(iii$_2$) If $\beta_c^B < \beta < \beta_c^A$, then $b_B^c < b_A^{\mathrm{up}}$, and $b_A^c$ occurs

$$\mathrm{A,B} \ \overset{b_B^{\mathrm{up}}}{\to} \ \mathrm{A,B'} \ \overset{b_B^c}{\to} \ \mathrm{A} \ \overset{b_A^{\mathrm{up}}}{\to} \ \mathrm{A'} \ \overset{b_A^c}{\to} \ \mathrm{E}. \qquad (A26)$$

(iii$_3$) If $\beta_c^A < \beta < \beta_c^B$, then $b_A^{\mathrm{up}} < b_B^c$, and $b_A^c$ is absent

$$\mathrm{A,B} \ \overset{b_B^{\mathrm{up}}}{\to} \ \mathrm{A,B'} \ \overset{b_A^{\mathrm{up}}}{\to} \ \mathrm{A',B'} \ \overset{b_B^c}{\to} \ \mathrm{A'}. \qquad (A27)$$

(iii$_4$) If $\beta < \min(\beta_c^A, \beta_c^B)$, then $b_A^{\mathrm{up}} < b_B^c$, and $b_A^c$ occurs

$$\mathrm{A,B} \ \overset{b_B^{\mathrm{up}}}{\to} \ \mathrm{A,B'} \ \overset{b_A^{\mathrm{up}}}{\to} \ \mathrm{A',B'} \ \overset{b_B^c}{\to} \ \mathrm{A'} \ \overset{b_A^c}{\to} \ \mathrm{E}. \qquad (A28)$$

This analysis provides the three-dimensional phase diagram $(r, \beta, b)$ for fixed, though arbitrary, $\epsilon$ and $p$. It exhibits a wealth of different macroscopic phases separated by critical lines and surfaces. It shows that polarized and quasipolarized phases dominate wide regions in parameter space. This illustrates the effects of interpopulation trade of fitness (even under the simplest possible structure of interpopulation and intrapopulation contacts) on the evolution of PD replicators. Note that if one uses the adjusted replicator equation introduced in Ref. [1], the stability analysis of the fixed points of the dynamics could show some differences, as already remarked in Ref. [25] concerning the evolutionary dynamics of the *Battle of the Sexes* game.

[1] J. Maynard-Smith, *Evolution and the Theory of Games* (Cambridge University Press, Cambridge, 1982).

[2] P. Kollock, Annu. Rev. Sociol. **24**, 183 (1998).

[3] E. Pennisi, Science **309**, 93 (2005).

[4] E. Pennisi, Science **325**, 1196 (2009).

[5] G. Szabó and G. Fáth, Phys. Rep. **446**, 97 (2007).

[6] C. P. Roca, J. Cuesta, and A. Sánchez, Phys. Life Rev. **6**, 208 (2009).

[7] M. Perc and A. Szolnoki, BioSystems **99**, 109 (2010).

[8] C. Gracia-Lázaro, A. Ferrer, G. Ruíz , A. Tarancón , J. A. Cuesta, A. Sánchez, and Y. Moreno, Proc. Natl. Acad. Sci. (USA) **109**, 12922 (2012).

[9] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).

[10] F. C. Santos and J. M. Pacheco, Phys. Rev. Lett. **95**, 098104 (2005).

[11] J. Gómez-Gardeñes, M. Campillo, L. M. Floría, and Y. Moreno, Phys. Rev. Lett. **98**, 108103 (2007).

[12] J. Gómez-Gardeñes, J. Poncela, L. M. Floría, and Y. Moreno, J. Theor. Biol. **253**, 296 (2008).

[13] M. Tomassini, E. Pestelacci, and L. Luthi, Int. J. Mod. Phys. C **18**, 1173 (2007).

[14] G. Abramson and M. Kuperman, Phys. Rev. E **63**, 030901 (2001).

[15] F. Fu, L.-H. Liu, and L. Wang, Eur. Phys. J. B **56**, 367 (2007).

[16] A. Pusch, S. Weber, and M. Porto, Phys. Rev. E **77**, 036120 (2008).

[17] S. Assenza, J. Gómez-Gardeñes, and V. Latora, Phys. Rev. E **78**, 017101 (2008).

[18] M. Szell, R. Lambiotte, and S. Thurner, Proc. Natl. Acad. Sci. USA **107**, 13636 (2010).

[19] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela, Science **328**, 876 (2010).

[20] S. V. Buldyrev *et al.*, Nature (London) **464**, 1025 (2010).

[21] R. Parshani, S. Buldyrev, and S. Havlin, Proc. Natl. Acad. Sci. USA **108**, 1007 (2011).

[22] Z. Wang, A. Szolnoki, and M. Perc, Europhys. Lett. **97**, 48001 (2012).

[23] J. Gómez-Gardeñes, I. Reinares, A. Arenas, and L. M. Floría, Sci. Rep. **2**, 620 (2012).

[24] L. M. Floría, C. Gracia-Lázaro, J. Gómez-Gardeñes, and Y. Moreno, Phys. Rev. E **79**, 026106 (2009).

[25] A. Traulsen, J. C. Claussen, and C. Hauert, Phys. Rev. Lett. **95**, 238701 (2005).