Stochastic Dynamics of Cooperation under Different Learning Rules

António Medeiros Fernandes
antonio.m.fernandes@tecnico.lisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

Abstract

Understanding the conundrum of cooperative action in nature has been declared one of the century's grand challenges. Natural selection is the main engine of the evolution of species and promotes the survival of the fittest, yet cooperative interactions prevail in living systems. Interactions between individuals, either as dyadic or collective processes, are commonly modelled through the lens of game theory, based on relative strategies payoffs. However, these interactions may not always be driven by a selfish agenda such as fitness maximization or rational decision making. In this thesis, we analyze different strategy revision rules that may take place in social systems resorting to the analysis of large-scale Markov processes. We describe the cooperation dynamics emerging from the interplay of different strategy revision paradigms: social learning (our baseline), conformity and counterfactual thinking. We show that conformity creates bistable dynamics irrespectively of the various dilemmas we analyzed. This outcome is only altered whenever one increases the heterogeneity levels of the population. Counterfactual thinking, in turn, while fostering cooperation in the entire population, fails to promote cooperation among those that also learn through counterfactual reasoning. Moreover, we show that, whereas a slight prevalence of individuals resorting to counterfactual thinking is enough to modify the dynamical nature of each dilemma, a fair amount of conformity-driven agents will be required to create an analogous impact.

Keywords: Emergence of cooperation, Evolutionary Game Theory, Conformity, Counterfactual Thinking

1. Introduction

In this world of selfish individuals, when should one expect the emergence of cooperative action? This problematic of has been widely studied and yet it remains one of the biggest mysteries about animal behaviour [1]. In fact, the understanding of cooperative behavior in societies has been declared as one of the grand scientific challenges of the 21st century and has been intriguing the scientific community in various areas such as sociology, biology, mathematics, physics and others [2, 3].

In the animal world, natural selection has been introduced and popularized as one of the fundamental key mechanisms of evolution, along with genetic mutation. From a Darwinist perspective, the survival of individuals is highly dependent on their adaptability to the demands of the environment and it may be understood as a competitive process, within a certain community, where the most successful individuals and the best suited to their surroundings tend to dominate. Thereby, the fact that natural selection favors the fittest individuals implies an innate selfishness that greatly challenge the concept of cooperation. In other words, from a Darwinist point of view, cooperation is costly, and, as weird as it may sound to us, humans, this means that cooperation is unnatural and rather odd than rational [4, 5]. So we now come across with this paradox: natural selection, which promotes the survival of the fittest, is the main engine of the evolution of species, yet we verify that cooperative interactions prevail in most living systems [6]. How can we then explain the emergence of cooperation?

Classical Game Theory was developed extensively in the 1950s, originally for application in economics and social science, but later in the 70s a group of biologists started to recognize how similar the games that have been studied were to the interactions between agents within a certain community. The Evolutionary Game Theory (**EGT**) was then born when John Maynard Smith and George Price laid its foundations, providing a powerful theoretical framework to model Darwinian competition [7], known as social learning (**SL**). Traditionally, interactions have been modeled in terms of one-shot, symmetric two-person dilemmas of Cooperation [8], however many real-life interactions

Games	NSG		NSH		CRD	
	C	D	C	D	C	D
$1 \le k < M$	$-\frac{c}{M}$	0	$-c$	0	$b(1-r) - cb$	$b(1-r)$
$M \le k$	$b - \frac{c}{k}$	b	$\frac{kFc}{N} - c$	$\frac{kFc}{N}$	$b - cb$	b

Table 1: Payoff values for the N-person snowdrift game (**NSG**), the N-person Stag-Hunt game (**NSH**) and the Collective-risk dilemma (**CRD**).

involve individual's collective action, in groups composed of more than two agents, which brings a much more rich and complex dynamics to the system. This type of interactions are best described in the framework of N-person games [9, 10, 11], such as the N-person Snowdrift game (**NSG**), the N-person Stag-Hunt (**NSH**), the Public Good Games (**PGG**) and the Collective Risk Dilemma (**CRD**).

Regardless of the model, under the simplest co-operation dynamics, interacting agents may adopt one of two pure strategies: cooperation (C) or defection (D). While cooperative individuals contribute to the collective welfare at a personal cost, c, in order to provide a certain benefit, b, to another individual (where we imply $b > c$), defectors choose not to contribute. In the case of collective action, it is common that no benefit is produced unless its costs are shared by a minimum amount of cooperators [9], which we will define as the threshold $1 \le M \le N$, with N the groups size. The final outcomes, or payoffs, for each strategy are now dependent on the social interaction rules, that differ on the metaphor used to describe each game. Specifically for N-person games above mentioned, the payoff values Π_S, with $S \in \{C, D\}$, may be organized, for both C and D, according to Table 1.

In **EGT**, the payoff individuals obtain from all interactions is translated onto individual fitness as an average payoff, that is characteristic for each strategy. In a biological context, one may see fitness as a quantification of the reproduction success: the more successful individuals are, the more they can spread their strategy, which will then increase its frequency in the population [8]. In some cases, researchers assume populations to be infinite [12], which, although simpler, neglects the stochastic effects characteristic of all evolutionary and social processes. Instead, here we focus on finite and well-mixed populations, that is, finite populations in which every individual has the same probability of interaction with one another in the population. Moreover, we will consider the presence of a small mutation rate μ, to avoid the absorbing property of the monomorphic states, that is, states where all individuals follow the same strategy.

Whereas in the 2-person framework fitness may

be achieved directly from the payoff values, in collective action one must take a more complex formulation to define it [13]. Specifically, in a well-mixed and finite population with size Z, engaged in groups sized N, the fitness of strategies C and D, respectively, f_C and f_D, may be given by the hypergeometric distributions, for a certain number k of cooperators, as follows:

$$f_C(k) = \binom{Z-1}{N-1}^{-1} \sum_{j=0}^{N-1} \binom{k-1}{j}\binom{Z-k}{N-j-1} \Pi_C(j+1) \quad (1)$$

$$f_D(k) = \binom{Z-1}{N-1}^{-1} \sum_{j=0}^{N-1} \binom{k}{j}\binom{Z-k-1}{N-j-1} \Pi_D(j) \quad (2)$$

The evolutionary dynamics of games are commonly studied through replicator equations [12], which provides a convenient view for the evolutionary process, although assuming infinite populations and a deterministic dynamic process [14]. The growth of the fraction of cooperators, x, in time may be given by a first order differential equation, the so-called replicator equation, as [12]:

$$\dot{x} = x(1-x)(f_C - f_D) \equiv g(x) \quad (3)$$

where the designation $g(x)$ refers to the gradient of selection, as it indicates the most likely outcome of evolution under **SL** [11, 13, 15]. When moving to finite populations, however, the description of the evolutionary dynamics of the propagation of strategies must now take into account the stochastic effects.

To study this deterministic propagation of strategies, methods of statistical physics may be applied, where a comprehensive understanding of statistical physics of human cooperation is given [5, 15, 16, 17]. In fact, to analyze an evolutionary process, one can adopt the Fermi function distribution from statistical mechanics [15], that may be described according to the imitation process: a certain individual A interacting with another player B, respectively adopting strategies S_A and S_B, with $S_B \ne S_A$, will revise and replace its strategy by its opponent strategy with a certain probability given by:

$$p_{S_A \to S_B}(k) \equiv \frac{1}{1 + e^{-\beta(f_{S_B}(k) - f_{S_A}(k))}} \quad (4)$$

2

where β, which in physics corresponds to an inverse temperature, controls the intensity of selection [16, 18, 19], and k is the number of cooperators. Notice that, by taking $\beta_{SL} \ll 1$, the process described is reduced to the frequency-dependent Moran process under weak selection [14].

As we enter in a discrete domain, the description of the system's evolutionary dynamics is now given according to the formalism of the Markov Processes [16]. In specific, we take a stochastic birth–death process in which, for a population sized Z, we will have $Z + 1$ states, as each state represents the number of cooperators k. Making use of the switching probability in equation 4, one may compute the transition probabilities, fully characterizing the transition matrix $\mathcal{T}$ that describes such a Markov chain, which is tri-diagonal in this specific case. The transition probabilities, i.e., the probabilities to increase (T^+) or decrease (T^-) the number of cooperators, in the population of size Z, is given by:

$$T^{\pm}(k) = \frac{k}{Z}\frac{Z-k}{Z-1}\left[1 + e^{\mp\beta(f_C(k)-f_D(k))}\right]^{-1} \quad (5)$$

From this, we may finally compute the most likely direction of the evolutionary process, in terms of the gradient of selection G, as:

$$G(x) \equiv T_+(x) - T_-(x) \quad (6)$$

which, in the limit of large Z, is equivalent to the replicator equation 3 for infinite populations.

From the transition matrix $\mathcal{T}$, we may also compute the stationary distribution, $\bar{s}$, which basically defines the fraction of time the system spends on each state, or in other words, defines the probability of finding the system at a certain state k at any time. It may be obtained simply by computing the left eigenvector of the transition matrix with eigenvalue 1, that is, it must satisfy $\bar{s} = \mathcal{T}\bar{s}$.

These two quantities will be of foremost importance in the study of population's dynamics, as they allow us to predict the most likely outcomes of evolution, in this case under selection. To evaluate and analyze these predictions, we may either calculate the expected fraction of cooperators, simply as $EFC = \sum_{k=0}^{Z} k \cdot \bar{s}(k)$, or the overall probability of achieving success, in the form of the group achievement, defined as $\eta_G = \sum_{k=0}^{Z} \bar{s}(k) \cdot a_G(k)$, where we take the multivariate hypergeometric sampling to compute a_G, the (average) fraction of groups that overcome the threshold of M contributors. Notice, however, that the EFC analysis may be escorted by misleading conclusions, namely for coordination dynamics, as it does not take into account the overall shape of the stationary distribution.

2. Mathematical Framework

Until now, most of population dynamics in Evolutionary Game Theory had its mathematical framework based on social learning. However, besides the pursuit of highest payoffs in evolutionary social dilemmas is risky, individual's interactions are not always driven by a selfish agenda such as fitness maximization [20]. In this theme of updating rules, some different learning rules have already been analyzed, such as conformity (**C**) [20, 21] and counterfactual thinking (**CT**) [22]. For further usage, we will define the set of available learning rules as $\mathcal{L}$ (in this case we will have $\mathcal{L} = \{SL, C, CT\}$).

As some individuals might not opt to imitate those with higher fitness, they may ignore the relative payoffs of the game and focus on choosing the most common strategy among their partners. Intuitively, conformists players tend to minimize their risk while ensuring they still receive some kind of payoff, not much lower than the average. To take into account well-mixed populations, one can very elegantly define a framework for conformity by simply making a pairwise comparison with the previously defined **SL** dynamics. In fact, the imitation process may be obtained as before, but instead of fitness-dependent, it is now frequency-dependent, as follows:

$$p_{S_A \to S_B}^{[C]}(k) \equiv \frac{1}{1 + e^{-\beta^{[C]}(x_{S_B}^{(Z-1)}(k) - x_{S_A}^{(Z-1)}(k))}} \quad (7)$$

where we explicitly define $\beta^{[C]} \neq \beta(\equiv \beta^{[SL]})$ as the intensity of conformity. Here, $x_S^{(Z-1)}$ represents the fraction of players following strategy S, in a population with $Z - 1$ individuals.

Making use of this new switching probabilities, the conformity transition probabilities, again by analogy to equations 5, may be written as:

$$T_{\pm}^{[C]}(k) = \frac{k}{Z}\frac{Z-k}{Z-1}\left[1 + e^{\mp\beta^{[C]}(x_C - x_D)}\right]^{-1} \quad (8)$$

which will define a novel gradient of selection, now specific for conformity:

$$G^{[C]}(k) \equiv T_+^{[C]}(k) - T_-^{[C]}(k) \quad (9)$$

Notice that here we changed the notation from $T^{\pm}$ to $T_{\pm}^{[L]}$, as these probabilities are now dependent on different learning rules $L \in \mathcal{L}$.

Counterfactual thinking is a human cognitive ability where individuals capture the process of reasoning about hypothetical past events, specifically what would have happened if those events occurred, hence taking into account the next player moves. Therefore, the counterfactual thinker may be already considered a Theory of Mind (**ToM**) first

degree agent, differently from a social learner or a conformist, that stand at the lower level of **ToM**. In its simplest form, **CT** may be modelled as an initial form of myopic best response rule at the population level, by considering the fitness of the agent in a system configuration that did not, but could have occurred [22]. From a pairwise comparison to the **SL** framework, the switching probability may be now defined as:

$$p_{S_A \to S_B}^{[CT]}(k) \equiv \frac{1}{1 + e^{-\beta^{[CT]}(f'_{S_B}(k) - f'_{S_A}(k))}} \quad (10)$$

where $\beta^{[CT]}$ stands for the intensity of counterfactual thinking. The probability increases non-linearly with the different between the fitness the individual would have had if it had chosen strategy S_B, designated as $f'_{S_B}(k)$, and the fitness it actually got by playing S_A, designated as $f'_{S_A}(k)$.

Differently from **SL** or **C**, the revision process does not occur through imitation, rather not requiring an interaction between agents. In this reasoning, and making use of the new switching probability, we may define the transition probabilities for **CT** may be obtain as follows:

$$T_-^{[CT]}(k) = \frac{k}{Z} \left[1 + e^{-\beta^{[CT]}(f_D(k-1) - f_C(k))} \right]^{-1} \quad (11)$$

$$T_+^{[CT]}(k) = \frac{Z-k}{Z} \left[1 + e^{-\beta^{[CT]}(f_C(k+1) - f_D(k))} \right]^{-1} \quad (12)$$

which will define the deterministic gradient of selection for counterfactual thinking:

$$G^{[CT]}(k) \equiv T_+^{[CT]}(k) - T_-^{[CT]}(k) \quad (13)$$

Notice that, differently from other learning rules, these transition probabilities do not have a second fraction of individuals term, hence a relative higher order of magnitude is expected for **CT**.

Having established the foundations for each the different rules in particular, we now ought to find a formulation that set all these learning rules together. In real-world populations, all individuals, regardless of their way to reason, may belong to the same community, either mixed with different-heuristic individuals, that is, in an homogeneous population, or displayed in separated communities, that is, in an heterogeneous population. Each of these organization types will unlock a whole new populations dynamic, from which very interesting insights may be taken.

3. Homogeneous Populations

Let us start by assuming that individuals are equivalent as they may resort to any of the available heuristics to their strategies. They may resort to different learning rules, yet all share the same probabilities of using each, and hence the equivalent

characteristic between them. In this case, populations may be addressed as homogeneous. Formally speaking, this is the same as redefine the system's gradient of selection to take into account all the different learning rules as:

$$G(k) = \sum_{L \in \mathcal{L}} \delta^{[L]} G^{[L]}(k) \quad (14)$$

where $\delta^{[L]}$ stands for the probability that each individual, at each time-step, adopts $L \in \mathcal{L}$.

We notice that the well-mixed property of the population is not violated, as expected when homogeneously mixing different learning rules in a population. From this point on, we may now evaluate the dynamics of different populations configuration under different scenarios.

We consider a well-mixed population of conformity-driven agents under a social learning environment, that is, having $\mathcal{L} = \{SL, C\}$. As both these learning rules are associated with dynamic revision processes, and assuming a similar magnitude for both switching probabilities, it wouldn't be unreasonable to expect similar orders of magnitude for the values of both $G^{[SL]}$ and $G^{[C]}$. Very roughly, this means the conformity impact on the overall population dynamics is dimensionally equivalent to the social learning's impact, hence one expects to verify progressively stronger effects of conformity as we increase its probability on the population.

To analyze the conformity effects, we start by building a pure **SL** population ($\delta^{[C]} = 0$) and we progressively increase the value of the probability of conformity until reaching a pure conformity population ($\delta^{[C]} = 1.0$). For each step, we compute the gradient of selection, as an explicit function of the fraction of cooperators x, and analyze the system's dynamics. Recall that, in this setup, we assume a constant fixed value for $\delta^{[C]}$, characteristic of the population, and therefore the G variations with the probability of conformity should not be considered as part of the population's evolution. An example of a study of this kind can be found in Figure 1.

Analyzing the conformity overall impact, results match the expectations: the variations on the system's dynamics progressively increase with increasing probability of conformity, and hence the effects of conformity apparently become more visible for larger $\delta^{[C]}$. This result suggests that, differently to **CT**, it is required a considerable prevalence of conformity-driven agents to substantially change the social game dynamics, as the conformity gradient is dimensionally equivalent to the social learning gradient.

Regarding the modifications itself, conformity may completely change the gradient of selection,
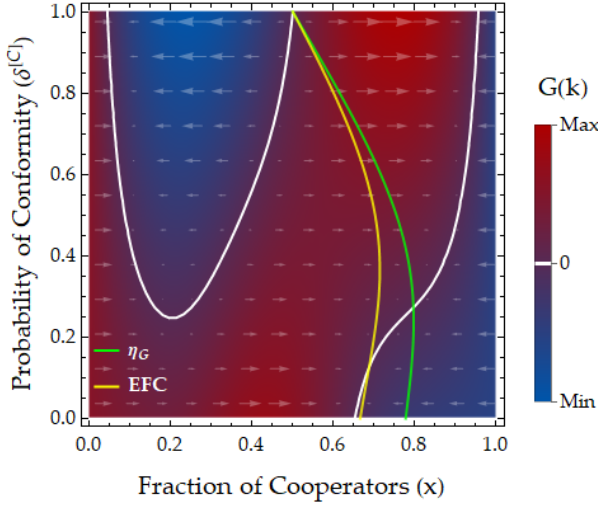
Figure 1: Gradient of selection of a **NSG** with $(N, M, b, c) = (10, 5, 1, 0.2)$, having $\mathcal{L} = \{SL, C\}$, for a homogeneously well-mixed population sized $Z = 50$, with $\mu = \frac{1}{Z}$ and $\beta^{[SL]} = \beta^{[C]} = 1.0$. It is presented in a relative terms.

depending on the original **SL** evolutionary dynamics. In the example above shown, increasing the number of conformity-driven individuals will progressively twist the original **NSG** co-existence dynamic onto a strong coordination dynamic, characterized by two stable fix points, on the monomorphic states of the system, and a single unstable fix point at exactly $x = 0.5$, which, in turn, is characteristic of the pure conformity dynamic. Attending to the conformity formulation, as it is not explicitly dependent on the game's configuration, we may admit this insight to be generalized, meaning that, irrespective of the **SL** dynamics, conformity will always promote a coordination dynamics, an effect that is more intense the more conformists we have in the population. Even though this dynamic strongly depends on the intensity of **C**, in its formal definition, conformity promotes a coordination dynamic, hence configurations with a value of $\beta^{[C]}$ that promotes co-existence will lose their physical meaning. As that, we are more prone to assume that conformity always promotes coordination, notwithstanding the social game's dynamics.

Interestingly, either the group achievement and the *EFC* progressively move towards the internal fix point, a result that is, again, independent on the social game's configuration. However, if in the pure **SL** dynamic we have $\eta_G < 0.5$, we could expect conformity to be beneficial in the overall cooperation dynamics, whereas, if we have $\eta_G > 0.5$, conformity would generally become prejudicial to the cooperation dynamics. This insight may suggest that, although the pure-conformity dynamic is very unique and independent on social game, the effects on cooperation optimization may have

either a positive or a negative impact regarding cooperation maximization, depending whether the **SL** dynamics provides, or not, a worst scenario in terms of cooperation, when compared to the pure conformity coordination dynamics.

Contrary to conformity, counterfactual reasoning is characterized by a static revision process rather than a dynamic one. As a consequence, **CT** comes with a much larger impact on the overall gradient of selection for homogeneous populations, when compared to the **SL** case. This means that a small prevalence of **CT** individuals, that is, a small value of $\delta^{[CT]}$, is enough to drastically change the population's dynamics, specifically towards highly cooperative standards, a result that was previously seen in [22]. To control and balance these effects, when analyzing the population's dynamics, we will assume $\beta^{[SL]} > \beta^{[CT]}$, significantly increasing the overall strength of **SL** in order to balance the gradient of selection for both learning rules.

Proceeding as before, to analyze the **CT** effects, we start by building a pure **SL** population ($\delta^{[CT]} = 0$) and we progressively increase the value of the probability of **CT** until reaching a pure **CT** population ($\delta^{[CT]} = 1.0$), and for each step we compute the gradient of selection. An example of this type of study may be found below in Figure 2.
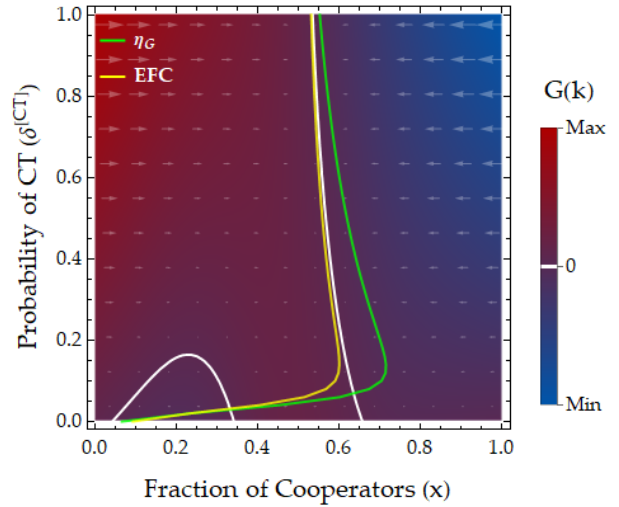

Figure 2: Gradient of selection of a **NSH** with $(N, M, F, c) = (10, 5, 9.5, 0.2)$, having $\mathcal{L} = \{SL, CT\}$, for a homogeneously well-mixed population sized $Z = 50$, with $\mu = \frac{1}{Z}$, $\beta^{[SL]} = 5.0$ and $\beta^{[CT]} = 1.0$. It is presented in a relative terms.

In opposition to conformity, and meeting the expectations, a small prevalence of **CT**-driven agents is enough to aggressively modify the population's dynamics, specifically being responsible for pushing the entire population to highly cooperative standards, as one may infer from both η_G and *EFC*. However, although a small incidence **CT** individuals is beneficial in terms of cooperation, after reach-

ing the maximum value for η_G, increasing $\delta^{[CT]}$ will progressively prejudice the cooperative standards. Therefore we may complement the conclusions previously taken in [22] by stating that a small prevalence of **CT**-driven individuals does in fact promote highly cooperative standards, but only until a certain point, from which adding more **CT** agents becomes a disadvantage.

Counterfactual thinking, in its pure form, and specifically for the **NSH**, seems to foster a co-existence dynamic, which, while still dependent on the game's configuration, is generalizable for other games. However, according to [22], in collective action problems, and specifically for **NSH**, one should expect counterfactual reasoning to promote coordination dynamics. Although it appears to be a contradiction, one must be aware of the dependency on the intensity of **CT**: in the same way as for conformity, while for relatively lower values of $\beta^{[CT]}$ individuals resorting to **CT** promote a co-existence dynamic, for higher intensities a coordination dynamic is favoured. This result is generic for any dynamic and should be taken into consideration when performing studies of this kind.

4. Heterogeneous Populations

When a population is organized in multiple smaller communities of same-heuristic individuals, a different approach shall be taken. Following a framework similar to the one presented in [23], we will divide our population according to each individual's learning rule, that is fixed and unchangeable, treating each sub-population independently, however always considering all interactions between different sub-populations are possible. The impact between individuals from different sub-populations will be weighted by the homophily, $0 \leq h \leq 1$, that may be defined as follows: when $h = 0$ anyone in the population may influence and be influenced by anyone else, while when having $h = 1$ individuals are restricted to influence (and be influenced) by those of the same sub-population. One may notice that the first case is equivalent to assume the population is well-mixed, while the latter case highlights the definition of heterogeneous populations. In this sense, homophily may be seen as a measure of the level of heterogeneity, from an well-mixed arrangement ($h = 0$), to a pure heterogeneous one ($h = 1$).

This new framework requires its own proper def-inition, as the previously presented formulation (in section 2) is very limited. In this study, we will focus on analyzing the dynamics, first having $\mathcal{L} = \{SL, C\}$, and secondly with $\mathcal{L} = \{SL, CT\}$.

Starting with the first case, we consider a population of $Z = Z^{[SL]} + Z^{[C]}$ individuals, with $Z^{[SL]}$ social learners and $Z^{[C]}$ conformists. Individuals resorting to each learning rule are given an initial endowment, $b^{[SL]}$ for **SL** and $b^{[C]}$ for **C**, that will remain fixed, and play the N-person game in study, thus engaging in groups sized N. Each type of co-operators will contribute for the public good with $c^{[SL]} = c \cdot b^{[SL]}$ and $c^{[C]} = c \cdot b^{[C]}$, respectively for **SL** and **C**. Having this new more complex formulation, the payoffs for each game must be updated, however always respecting each game's description. For the **NSG** and the **CRD**, the new payoffs may be defined as follows:

- **NSG**, with $(N, M, \bar{b}, b^{[SL]}, b^{[C]}, c)$:

$$\Pi_C^{[SL]}(i) = \Pi_D(i) - \frac{c^{[SL]}}{i_C^{[SL]} + i_C^{[C]}}\Theta - \frac{c^{[SL]}}{M}(1 - \Theta) \quad (15)$$

$$\Pi_D^{[SL]}(i) = b^{[SL]}\Theta \quad (16)$$

- **CRD**, with $(N, M, \bar{b}, b^{[SL]}, b^{[C]}, c, r)$:

$$\Pi_C^{[SL]}(i) = \Pi_D(i) - c^{[SL]} \quad (17)$$

$$\Pi_D^{[SL]}(i) = b^{[SL]}(\Theta + (1 - r)(1 - \Theta)) \quad (18)$$

where $i = \{i_C^{[SL]}, i_C^{[C]}\}$, with $i_C^{[L]}$ the number of co-operators with $L \in \mathcal{L}$. Moreover, here we have $\Theta \equiv \Theta(c^{[SL]}i_C^{[SL]} + c^{[C]}i_C^{[C]} - c\bar{b}M)$ as the Heaviside function, where $\Theta(k) = 1$ whenever $k \geq 0$ and $\Theta(k) = 0$ otherwise, and where $Z\bar{b} = Z^{[SL]}b^{[SL]} + Z^{[C]}b^{[C]}$ is the average endowment.

In opposition to **SL**, conformity does not rely on fitness in the individual's strategic decision, rather depending in the number of individuals following each strategy. In this sense, one may generally define $\Pi_C^{[C]} = \frac{i_C^{[C]} + i_C^{[SL]}}{Z - 1}$ and $\Pi_D^{[C]} = \frac{i_D^{[C]} + i_D^{[SL]}}{Z - 1}$, where $i_D^{[L]} = Z^{[L]} - i_C^{[L]}$, for $L \in \mathcal{L}$.

As conformity populations are not engaged in groups of size N, the conformity fitnesses are reduced to their respective payoffs. As that, we just have to compute **SL** fitnesses, which come naturally by using the multivariate hypergeometric sampling without replacement, which according to [23], may be formulated as shown in the equations below:

$$f_C^{[SL]}(i) = \binom{Z-1}{N-1}^{-1} \sum_{j^{[SL]}=0}^{N-1} \sum_{j^{[C]}=0}^{N-1-j^{[SL]}} \binom{i^{[SL]}-1}{j^{[SL]}}\binom{i^{[C]}}{j^{[C]}}\binom{Z-i^{[SL]}-i^{[C]}}{N-1-j^{[SL]}-j^{[C]}}\Pi_C^{[SL]}(j^{[SL]}+1, j^{[C]}) \quad (19)$$

$$f_D^{[SL]}(i) = \binom{Z-1}{N-1}^{-1} \sum_{j^{[SL]}=0}^{N-1} \sum_{j^{[C]}=0}^{N-1-j^{[SL]}} \binom{i^{[SL]}}{j^{[SL]}}\binom{i^{[C]}}{j^{[C]}}\binom{Z-1-i^{[SL]}-i^{[C]}}{N-1-j^{[SL]}-j^{[C]}}\Pi_D^{[SL]}(j^{[SL]}, j^{[C]}) \quad (20)$$

$$T^{[SL]}_{C\to D\setminus D\to C}(i) \equiv T^{[SL]}_{\mp}(i) = \frac{i^{[SL]}_{C\setminus D}}{Z}\left\{ \frac{i^{[SL]}_{D\setminus C} + (1-h)i^{[C]}_{D\setminus C}}{Z^{[SL]} - 1 + (1-h)Z^{[C]}}\left(1 + e^{\beta^{[SL]}(f^{[SL]}_{C\setminus D}(i) - f^{[SL]}_{D\setminus C}(i))^{-1}}\right)\right\} \quad (21)$$

$$T^{[C]}_{C\to D\setminus D\to C}(i) \equiv T^{[C]}_{\mp}(i) = \frac{i^{[C]}_{C\setminus D}}{Z}\left\{ \frac{i^{[C]}_{D\setminus C} + (1-h)i^{[C]}_{D\setminus C}}{Z^{[C]} - 1 + (1-h)Z^{[SL]}}\left(1 + e^{\beta^{[C]}(f^{[C]}_{C\setminus D}(i) - f^{[C]}_{D\setminus C}(i))^{-1}}\right)\right\} \quad (22)$$

Assuming an asynchronous updating process in the evolutionary dynamics, from these fitnesses we may compute all the required transition probabilities, as shown in the equations 21 and 23.

From these transition equations, we may compute the transition matrix, now adapted to a multidimensional space. To do so, we index all the possible states with an integer number, for which conversion we define a bijective function V such that $p = V(i)$ and $q = V(i')$, and, consequently, $i = V-1(p)$ and $i' = V-1(q)$. Afterwards, we may write the transition matrix as $\mathcal{T}_{qp} = T_{i\to i'}$, where $T_{i\to i'}$ is the corresponding transition probability from configuration i to i'. Having computed $\mathcal{T}$, one may finally obtain the stationary distribution following the usual procedure, and to the obtained vector it is applied the inverse of V so one may finally obtain the stationary distribution directly as a function of i, as $\bar{s} \equiv \bar{s}(i)$ [24].

Lastly, also from the transition probabilities, we may generally define the 2-dimensional heterogeneous gradient of selection for social learning and conformity as follows:

$$\nabla(i) = \{G^{[SL]}(i), G^{[C]}(i)\} \quad (23)$$

In this new setup, the group achievement must also be updated as follows:

$$\eta_G(i) = \sum_{j^{[SL]}=0}^{Z^{[SL]}} \sum_{j^{[C]}=0}^{Z^{[C]}} a_G(j) \cdot \bar{s}(j) \quad (24)$$

with $j = \{j^{[SL]}, j^{[C]}\}$.

Having all this quantities, have now conditions to proceed with the evaluation of the gradient of selection. We start by investigating the effects of homophily in evolutionary dynamics for populations with an equal amount of social learners and conformists, that is, having $Z^{[SL]} = Z^{[C]}$. A study of this kind is presented in Figure 3, specifically performed for the **NSG**.

Firstly, we notice that, having null homophily, one can recover, in some way, the typical homogeneous conformity dynamics, characterized by the existence of three fix points, two of them being stable on both the monomorphic states, while the other is unstable at the center. This insight comes more clear if we take into consideration the symmetry of the dynamics, which in this case may be considered practically symmetrical.

Attending to the stationary distribution, the stable equilibrium relative to the cooperation monomorphic state is clearly much stronger than the other one, as the system spends much more time in the first case. Moreover, the population dynamics is not much sensitive to most lower homophily values, when, even having $h = 0.6$, the dynamic is practically the same as for $h = 0$. However, as homophily increases, we will eventually reach a threshold from which the system dynamics strongly changes, especially when h is maximum. In this case, the cooperation monomorphic fixed point has slightly moved to lower $i^{[SL]}_C$, while the other two fix points have now disappeared. This result may suggest that an increase of the homophily may lead to a new dynamics where the only most probable outcome, after a sufficient amount of time, is the configuration which, in this particular case, promotes highly cooperative standards, specifically representing a C-dominance dynamic. Even though the group achievement is lower than in the other cases with lower h, which could indicate a worst performance in terms of cooperation optimization, we no longer have the tendency to move towards the defection monomorphic state, having only one strong attraction point.

Regarding the group achievement, we notice the relations between group achievement and the homophily are not linear. For instance, from the available results for the **NSG**, we infer a very light increase of η_G from $h = 0$ to $h = 0.6$, whereas for the homophilic setup the group achievement reaches its lowest value. We are, then, led to suggest that the group achievement goes, not only as a function of the game's configuration, as we saw previously on homogeneous populations, but also as a function of the homophily, that is, $\eta_G \equiv \eta_G(v, h)$ with v the set of variables that identify each game.

Contrary to conformity, even in homogeneous populations, counterfactual thinking has shown to be a very interesting and complex learning rule, being able to either foster coordination or coexistence, depending on $\beta^{[CT]}$, hence becoming a very flexible heuristic.

Moving to the second case, like before, we start by building a population sub-divided onto two sub-populations, one of social learners and the other constituted by counterfactual thinkers, thus having
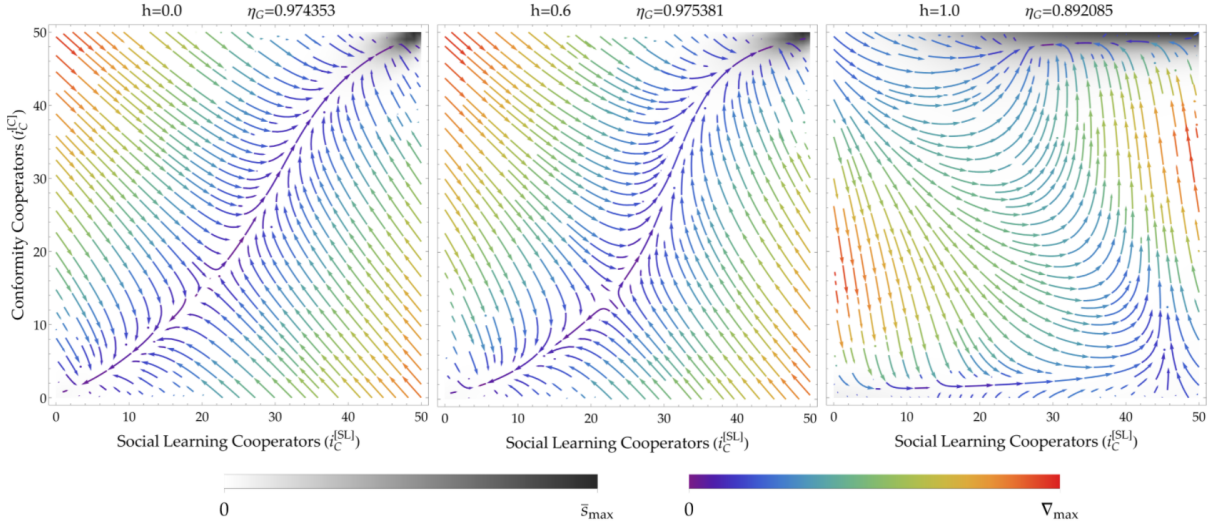
Figure 3: Gradient of selection of a **NSG**, with $(N, M, b^{[SL]}, b^{[C]}, c) = (10, 5, 1, 1, 0.1)$, having $\mathcal{L} = \{SL, C\}$, for a heterogeneous population sized $Z = 100$, with $Z^{[SL]} = Z^{[C]} = 50$, $\mu = \frac{1}{Z}$, $\beta^{[SL]} = \beta^{[C]} = 1.0$. We show that an increasing the heterogeneity levels progressively corrupts the conformity bistability onto a C-dominance dynamic.

$\mathcal{L} = \{SL, CT\}$. Following the same reasoning as for conformity, we consider our population has a size of $Z = Z^{[SL]} + Z^{[CT]}$ individuals, from where $Z^{[SL]}$ are social learners and $Z^{[CT]}$ follow **CT**, that will play a given N-person game. Individuals resorting to either **SL** or **CT** are given an initial endowment $b^{[SL]}$ and $b^{[CT]}$, respectively, and, being cooperators, they are compelled to contribute, proportionally to their respective endowments, with $c^{[SL]} = c \cdot b^{[SL]}$ and $c^{[CT]} = c \cdot b^{[CT]}$. Having counterfactual thinking based on fitness, the payoffs relative to **SL** for each the N-person games (equations 15-18) remain unchanged, only substituting $[C]$ by $[CT]$ in the learning rules.

From these payoffs, we again may compute the fitnesses for **SL**, but again substituting $[C]$ by $[CT]$ on equations 19 and 20. In turn, for **CT**, the fitness may be obtained in the same was as previously but swapping $[SL]$ with $[CT]$, however now taking into consideration the **CT** formulation:

$$f_C^{[CT]}(\boldsymbol{i}) = \begin{cases} f_C^{[SL]}(i_C^{[SL]}, i_C^{[CT]}) & C \to D \\ f_C^{[SL]}(i_C^{[SL]}, i_C^{[CT]} + 1) & D \to C \end{cases} \quad (25)$$

$$f_D^{[CT]}(\boldsymbol{i}) = \begin{cases} f_D^{[SL]}(i_C^{[SL]}, i_C^{[CT]} - 1) & C \to D \\ f_D^{[SL]}(i_C^{[SL]}, i_C^{[CT]}) & D \to C \end{cases} \quad (26)$$

Having this, and again considering an asynchronous updating process in the evolutionary dynamics, we may compute the required transition probabilities. For **SL**, we may retake equation 21, but substitute $[C]$ by $[CT]$, and for **CT**, attending to its static updating process, we may have:

$$T_{\mp}^{[CT]}(\boldsymbol{i}) = \frac{i_{C\backslash D}^{[CT]}}{Z} \left\{ \left(1 + e^{\beta^{[CT]}(f_{C\backslash D}^{[CT]}(\boldsymbol{i}) - f_{D\backslash C}^{[CT]}(\boldsymbol{i}))^{-1}} \right) \right\} \quad (27)$$

From the definition above we may already observe the **CT** dynamic is not explicitly dependent on the homophily. Let us keep this in mind for the remaining analysis.

Similarly as before, from these transition probabilities, we may generally define the 2-dimensional heterogeneous gradient of selection for social learning and counterfactual thinking again by replacing $[C]$ by $[CT]$ on equation 23.

In addition to the previously defined group achievement, it will also be important to compute the EFC, however, having different sub-populations, one may compute it specifically for each sub-population, which can be done for any learning rule $L \in \mathcal{L}$ as follows:

$$EFC_L = \frac{i_C^{[L]}}{Z^{[L]}} \cdot \bar{s} \quad (28)$$

Attending to previous studies, as in [22], when having homogeneous populations, one would expect that a small prevalence of individuals resorting to **CT** strongly favours cooperation. In this sense, and before proceeding with our investigation on the effects of homophily in the evolutionary dynamics, it would be interesting to evaluate the overall cooperative performance of the dynamics of populations when varying the different sub-populations sizes. To do so, for a population with fixed size $Z = Z^{[SL]} + Z^{[CT]}$, we vary the sub-populations size proportion $(Z^{[SL]}/Z^{[CT]})$ and for each value we may compute all the previously described cooperation evaluation tools. An example of this study, for the **CRD**, is shown on Figure 4.

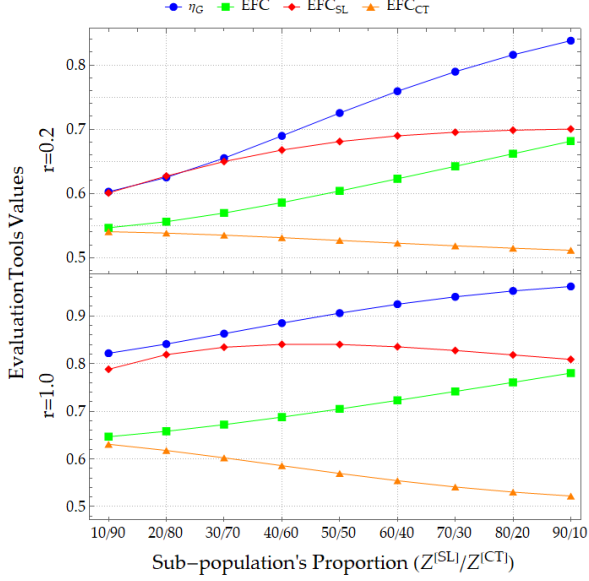For the lower risk, notice that all quantities, with exception to the EFC_{CT}, progressively increase to-

Figure 4: η_G, EFC, EFC_{SL} and EFC_{CT} versus the proportion between the sizes of the different sub-populations. For all the cases we have $Z = 100$, $\mu = \frac{1}{Z}$, $h = 0.9$ and $\beta^{[SL]} = \beta^{[CT]} = 5$, having $\mathcal{L} = \{SL, CT\}$ for the **CRD** with $(N, M, b^{[SL]}, b^{[CT]}, c, r) = (10, 5, 1, 1, 0.01, r)$, where $r = \{0.2, 1.0\}$. We show that, while responsible for fostering the population dynamics towards high cooperative standards, CT does not concerned about improving its own cooperation performance.

wards their maximum value as we increase the proportion until reaching $Z^{[SL]}/Z^{[CT]} = 90/10$. Surprisingly, it is exactly at this point where we find the minimum value for the EFC_{CT}, which reveals a very exciting result: counterfactual reasoning, while responsible for fostering the population dynamics towards much more cooperative standards, does not concerned about improving the cooperation standards of the ones of its kind. Furthermore, these results go along with the previously obtained insights for homogeneous populations, where we saw that after increasing the prevalence of **CT** agents, after reaching a critical point, when η_G is maximum, progressively leads to a decrease of the group achievement.

This previous insight may also be inferred from the analysis on on the effects of homophily in the population dynamics. Contrary to the homogeneous case, we will not focus on balancing the different learning rules through their respective intensities, as we intend to see how they react in their true form. An example of this investigation, specifically for the **CRD**, is presented in Figure 5.

The results suggest that increasing the homophily does not significantly affects the overall population dynamics. Attending to the definitions for the transition probabilities for **CT**, this result go as expected, not only because counterfactual thinking has a much stronger dynamics, when compared

to **SL**, but also it is independent on the homophily. As a consequence, the only changes on $\nabla(i)$ with increasing homophily will come from the differences on the x-component, relative to $G^{[SL]}$.

Even though the overall dynamics in terms of the gradient of selection does not significantly change, the system internal equilibria does change with the homophily, and, with that, the stationary distribution. Specifically, increasing the levels of heterogeneity progressively pushes the stationary distribution, along with the stable fix point, towards the **SL** edge of $i_C^{[SL]} \approx Z^{[SL]}$. This means that, even though we do not have any improve on the cooperative dynamics of **CT**, the presence of **CT**'s, in a full homophilic scenario, nudges specifically the entire sub-population of **SL** towards extremely high cooperative standards, matching the previous results from Figure 4.

5. Conclusions

The conformity's dynamics, alike we saw for homogeneous populations, is very specific and easily identifiable, generally not depending on the social game involved. In its pure form, it is characterized by the coordination dynamic, with two stable equilibria, for each the monomorphic states, and one unstable internal fix point. When moving to heterogeneous populations, however, homophily shows itself as able to corrupt the overall expected conformity dynamics, being able to either promote a C-dominance or a D-dominance dynamics, depending on the social game involved. Moreover, it is to be noted that, to reach such dynamics, it will be necessary a fair amount of conformity-driven agents to significantly modify the social learning dynamics, as the conformity gradient is dimensionally equivalent to the social learning gradient.

Counterfactual thinking, in turn, shows to be very versatile, as it may either promote a coexistence or a coordination dynamics, depending on the values of its intensity. Moreover, **CT** is highly dependent on the game's configuration, although generally acting in favor of maximizing cooperation. In fact, and contrary to conformity, a small prevalence of **CT**-driven individuals is enough to foster highly cooperative standards, although only until a certain critical point, where the group achievement is maximum, and from which adding more **CT** agents becomes progressively more disadvantageous for cooperative action.

When introducing heterogeneity in the population, and to complement the previous insights, we may specify that **CT** individuals, while responsible for fostering the population dynamics towards much more cooperative standards, does not concerned about improving the co-operation standards of the ones of its kind. Moreover, in con-
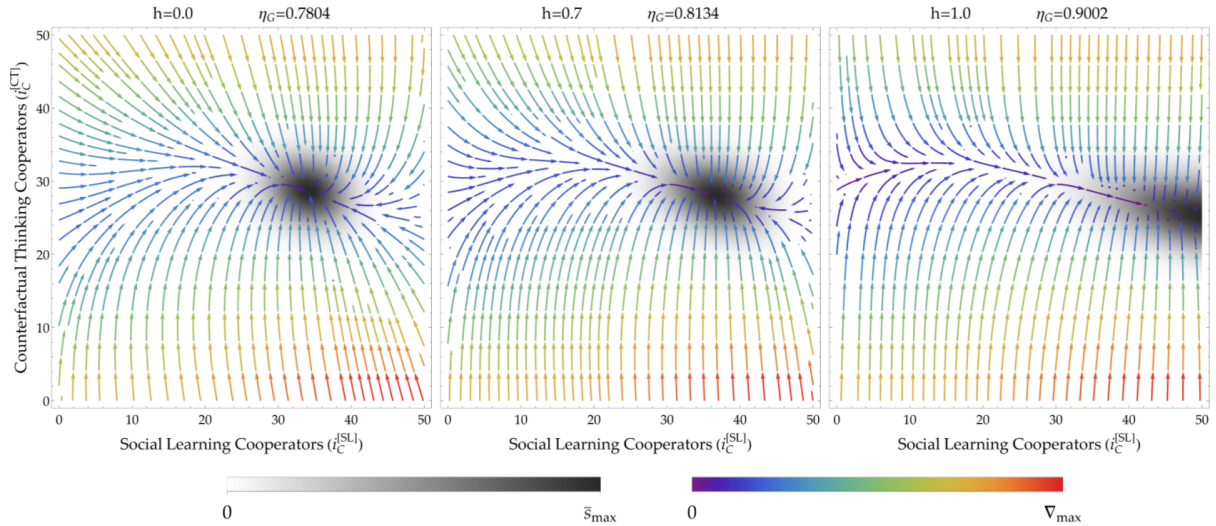
9

Figure 5: Gradient of selection of a **NSG**, with $(N, M, b^{[SL]}, b^{[CT]}, c) = (10, 5, 1, 1, 0.1)$, having $\mathcal{B} = \{SL, CT\}$, for a heterogeneous population sized $Z = 100$, with $Z^{[SL]} = Z^{[CT]} = 50$, $\mu = \frac{1}{Z}$, $\beta^{[SL]} = \beta^{[CT]} = 1.0$. We show that an increase on the heterogeneity levels leads to an cooperative enhancement among social learners, while among counterfactual thinkers this is not necessarily true.

trary to conformity's framework, higher heterogeneous levels generally promote higher cooperative outcomes, or, in the other hand, well-mixed populations lead to worst performance in terms of cooperation maximization.

Acknowledgements

References

[1] R. Axelrod. *The Evolution Of Cooperation*. Basic Books, Inc., Publishers, 1984.

[2] D. D. Kennedy and C. Norman. What don't we know? *Science*, 309(5731):75–75, 2005.

[3] E. Pennisi. How did cooperative behavior evolve? *Science*, 309(5731):93–93, 2005.

[4] Martin Nowak. Five rules for the evolution of cooperation. *Science (New York, N.Y.)*, 314:1560–3, 2007.

[5] M. Perc, J. Jordan, D. Rand, Z. Wang, S. Boccaletti, and A. Szolnoki. Statistical physics of human cooperation. *Physics Reports*, 687:1–51, 2017.

[6] Martin A. Nowak. *Evolutionary Dynamics: Exploring the Equations of Life*. Harvard University Press, 2006.

[7] J. Maynard-Smith and G. R. Price. The logic of animal conflict. *Nature*, 246 (5427):15–18, 1973.

[8] Marta Daniela de Almeida Santos. *Evolutionary dynamics of collective action in structured populations*. PhD thesis, Universidade de Lisboa, 2012.

[9] Max O. Souza, Jorge M. Pacheco, and Francisco C. Santos. Evolution of cooperation under n-person snowdrift games. *J Theor Biol*, 260(4):581–588, 2009.

[10] C. S. Gokhale and A. Traulsen. Evolutionary games in the multiverse. *Proc Natl Acad Sci USA*, 107(12):5500–5504, 2010.

[11] F. Santos and J. Pacheco. Risk of collective failure provides an escape from the tragedy of the commons. *Proc Natl Acad Sci USA*, 108:10421 – 10425, 2011.

[12] Josef Hofbauer, Karl Sigmund, et al. *Evolutionary games and population dynamics*. Cambridge university press, 1998.

[13] Jorge M. Pacheco, Francisco C. Santos, Max O. Souza, and Brian Skyrms. Evolutionary dynamics of collective action in n-person stag hunt dilemmas. *Proc. Royal Soc. B: Biol. Sci.*, 276(1655):315–321, 2009.

[14] Martin Nowak, Akira Sasaki, Christine Taylor, and Drew Fudenberg. Emergence of cooperation and evolutionary stability in finite populations. *Nature*, 428:646–50, 05 2004.

[15] A. Traulsen, M. Nowak, and J. Pacheco. Stochastic dynamics of invasion and fixation. *Phys Rev E*, 74:011909, 2006.

[16] Vítor V Vasconcelos, Fernando P Santos, Francisco C Santos, and Jorge M Pacheco. Stochastic dynamics through hierarchically embedded markov chains. *Phys Rev Lett*, 118(5):058301, 2017.

[17] Flávio L Pinheiro, Francisco C Santos, and Jorge M Pacheco. Linking individual and collective behavior in adaptive social networks. *Phys Rev Lett*, 116(12):128702, 2016.

[18] Sven Van Segbroeck, Francisco C Santos, Tom Lenaerts, and Jorge M Pacheco. Selection pressure transforms the nature of social dilemmas in adaptive networks. *New J Phys*, 13(1):013007, 2011.

[19] Arne Traulsen, Jorge M Pacheco, and Martin A Nowak. Pairwise comparison and selection temperature in evolutionary game dynamics. *J Theor Biol*, 246(3):522–529, 2007.

[20] Attila Szolnoki and Matjaž Perc. Conformity enhances network reciprocity in evolutionary social dilemmas. *Journal of The Royal Society Interface*, 12(103):20141299z, 2015.

[21] Attila Szolnoki and Matjaž Perc. Leaders should not be conformists in evolutionary social dilemmas. *Scientific Reports*, 6:23633, 03 2016.

[22] Luís Moniz Pereira and Francisco C. Santos. Counterfactual thinking in cooperation dynamics. *Studies in Applied Philosophy, Epistemology and Rational Ethics*, page 69–82, 2019.

[23] V. Vasconcelos, Francisco Santos, and Jorge Pacheco. A bottom-up institutional approach to cooperative governance of risky commons. *Nature Clim. Change*, 3:797–801, 09 2013.

[24] V. Vasconcelos, Francisco Santos, Jorge Pacheco, and Simon Levin. Climate policies under wealth inequality. *Proc Natl Acad Sci USA*, 111, 01 2014.