Evolution of cooperation in networked heterogeneous fluctuating environments

Viktor Stojkoski[1,2], Marko Karbevski[2,3,4,5], Zoran Utkovski[6], Lasko Basnarkov[7], Ljupco Kocarev[2,7]

[1] *SS. Cyril and Methodius University, Faculty of Economics,, blvd. Goce Delcev 9V, 1000 Skopje, North Macedonia*

[2] *Macedonian Academy of Sciences and Arts, P.O. Box 428, 1000 Skopje, North Macedonia*

[3] *Sorsix International, Dame Gruev 18, Skopje 1000, Republic of Macedonia*

[4] *Institute of Mathematics, Faculty of Natural Sciences and Mathematics, Ss. Cyril and Methodius University, Arhimedova 3, 1000 Skopje, Republic of Macedonia*

[5] *Sorbonne Université, 4 Place Jussieu, 75005 Paris, France*

[6] *Fraunhofer Heinrich Hertz Institute, Einsteinufer 37, 10587, Berlin, Germany*

[7] *SS. Cyril and Methodius University, Faculty of Computer Science and Engineering, P.O. Box 393, 1000 Skopje, Macedonia*

Abstract

Fluctuating environments are situations where the spatio-temporal stochasticity plays a significant role in the evolutionary dynamics. The study of the evolution of cooperation in these environments typically assumes a homogeneous, well mixed population, whose constituents are endowed with identical capabilities. In this paper, we generalize these results by developing a systematic study for the cooperation dynamics in fluctuating environments under the consideration of structured, heterogeneous populations with individual entities subjected to general behavioral rules. Considering complex network topologies, and a behavioral rule based on generalized reciprocity, we perform a detailed analysis of the effect of the underlying interaction structure on the evolutionary stability of cooperation. We find that, in the presence of environmental fluctuations, the cooperation dynamics can lead to the creation of multiple network components, each with distinct evolutionary properties. This is paralleled to the freezing state in the Random Energy Model. We utilize this result to examine the applicability of our generalized reciprocity behavioral rule in a variety of settings. We thereby show that the introduced rule leads to steady state cooperative behavior that is always greater than or equal to the one predicted by the evolutionary stability analysis of unconditional cooperation. As a consequence, the implementation of our results may go beyond explaining the evolution of cooperation. In particular, they can be directly applied in domains that deal with the development of artificial systems able to adequately mimic reality, such as reinforcement learning.

Keywords: Cooperation, Fluctuating environments, Generalized reciprocity

1. Introduction

Cooperation is ubiquitous in nature and is essential to the functioning of a large number of biological systems. This phenomenon has played a fundamental role in many of the major transitions in biological evolution, and cooperative interactions are required for many levels of networked organizations ranging from single cells to groups of animals and, ultimately, humans. In evolutionary biology, theoretical models provide evidence that cooperative behavior can evolve and persist if a certain mechanism is at work, for example, kin selection, group (multilevel) selection, and different forms of reciprocity [1, 2]. In addition, it has been recognized that the network structure plays an important role in the emergence of cooperation, as underlying network properties such as network heterogeneity, scale-freeness, etc., crucially determine the outcome of multiple dynamical phenomena.

A standard approach for examining the effect of different mechanisms on the cooperation dynamics in complex networks is through evolutionary game theory where the individual entities interacting in a network are given a set of strategies that they can choose from, and a set of payoffs (changes in the individual resource endowment) that result from interactions with other entities based on their chosen strategies. Traditional game theory has predominantly focused on additive payoffs, which yields the world of two-player iterative social dilemmas, such as the iterated prisoner's dilemma, or multi-player (i.e. public goods) games [3].

In contrast to interaction models with additive payoffs, the temporal evolution of the individual resource endowments in natural and social systems is often characterized by a multiplicative process [4]. In evolutionary biology and ecology, these systems are studied within the general framework of *fluctuating environments* , which refers to spatio-temporal stochasticity in the environmental conditions [5]. Environmental fluctuations can be integrated in models for evolutionary games in various ways, including stochastic network structures [6, 7, 8, 9], fluctuating population size [10, 11, 12, 13, 14, 15] and random payoffs [16, 17, 18, 19]. In our concept, we follow the last strand of works, and model environmental fluctuations as randomly varying payoff values.

In environments subjected to fluctuations in the payoff structure, the non-ergodicity of the fluctuation-generating process may have non-trivial effects on the evolutionary dynamics [5]. In particular, when the growth in the resource endowments is governed by a multiplicative process, the resulting fluctuations exhibit a net-negative effect on the time-averaged growth of the resources of the individual entities, while having no effect on the ensemble growth rate. Since, on the long run the ensemble properties are never observed, the time-averaged growth is the only relevant quantity for evolutionary performance [20, 21]. Recently, it has been discovered that this yields an evolutionary behavior which essentially differs from the one observed in standard models, i.e., leads to evolutionary dynamics where cooperators can coexist with defectors without any additional auxiliary mechanism at place [16]. For example, on this basis, it has been hypothesized that repeated pooling and sharing of resources which previously exhibit a fluctuating growth may constitute a fundamental mechanism for the evolution of cooperation in a well-mixed population. The rationale is that, by reducing the amplitude of fluctuations, pooling and sharing increases the steady state growth rate at which the cooperating entities self-reproduce [18, 19, 17].

The research on the evolution of cooperation in fluctuating environments typically focuses

on the following simplifying assumptions: i) the population is homogeneous, i.e. all entities are characterized by the same individual traits, ii) the interaction takes place in a well-mixed population, and iii) the individual entities can decide only between two behavioral strategies. The first assumption constraints each entity to have the same physical characteristics and to be subjected to the same fluctuation process. In reality, however, entities are *heterogeneous*: they are endowed with different capabilities and experience non-identical randomness. These traits can represent the level of skills or income profiles in economic societies [22, 23] or sex differences in development, physiology, morphology and/or life-history in biological systems [24]. The second assumption implies that entities spend, on average, an equal time interacting with each other member of the population. In practice, individual entities are most often interacting through a social *network* of contacts (or within a neighborhood in a biological network), thus favoring interactions with a certain group. It has been recognized that such network structures crucially determine the outcome of a multitude of dynamical phenomena, including cooperation [25]. The last assumption reduces the actions of the individual entities to a set of two possible outcomes. In other words, this assumption does not allow for learning to take place in the population. While all these assumptions allowed for powerful analytical findings, they can significantly reduce practical applications.

In this paper, we revisit the above assumptions and develop a systematic approach for investigating the cooperation dynamics in fluctuating environments by considering structured, heterogeneous populations with individual entities subjected to more general behavioral rules. We consider an interaction model on a complex network based on pooling and sharing of resources undergoing multiplicative growth that models environmental fluctuations. We know from [18, 19], that in a well-mixed homogeneous population, cooperation is always favored under this model. When moving from a well-mixed population to a more general network structure, in Ref. [26] it was shown that cooperation remains the sole evolutionary stable strategy (given that the population is homogeneous), but the performance of the population is uniquely determined by the underlying interaction topology. Complementing these works, here we investigate the joint effects of heterogeneous population and network topology. A crucial observation from the analysis is that, when population heterogeneity is introduced to the model, complex behavior emerges and the stability of cooperation is dependent on both the network topology and the behavioral update rules employed by the individual entities.

By analyzing the model from an evolutionary perspective, we derive the criteria required for unconditional cooperation to be evolutionary stable within any interaction structure. More importantly, by applying these criteria to simple and tractable analytical examples, as well as to more complex numerical examples, we are able to show that the interaction structure may yield a non-trivial effect on the behavior in multiplicative dynamics. In particular, the interaction structure can induce creation of multiple connected components, each consisting of unconditional cooperators. Even though the multiplicative nature of the process makes the distribution of resources among the cooperating entities to have a an infinite average in the time limit, the connected components instigate dynamics in which some unconditional cooperators are much "richer" than others. This effect is related to the freezing transition in the Random Energy model of Derrida [27], and in our model occurs as a result of the heterogeneous individual capabilities, in combination with the network structure [28].

3

The evolutionary stability analysis assumes that each individual entity implements strategies that allow only for unconditional cooperator or unconditional defector behavior. As such it does not allow us to examine the effect of the behavioural learning update rules employed by the individual entities. Due to the non-ergodicity, which arises as a consequence of the fluctuating environment, standard rules cannot be easily translated to our model [20]. Instead, one should use a suitable modification of the rule which is studied.

Here we develop a simple behavioral update and examine its implications. The introduced rule is related to the concept of *generalized reciprocity*, which itself is rooted in the principle of "help anyone if helped by someone" [29]. In particular, under our rule the individual entities search for their optimal strategies solely by comparing their observed growth rate with the one that is expected under unconditional defection. From a game-theoretic perspective the rule may be framed in the context of continuous games, where entities may be able to chose from a continuum of strategies [2], or as a stochastic decision making process, with random payoffs determined by the entity's experience [3]. Direct parallels can be made to state-of-the-art reinforcement learning techniques based on novelty search [30]. We thereby show that our rule leads to steady state cooperative behavior that is always greater than or equal to the one predicted by evolutionary stability analysis of unconditional cooperation. In fact, for a certain regime of parameter values we are able to analytically solve the model and show that then the strict inequality holds.

The rest of the paper is organized as follows. In Section 2 we introduce the system model, together with a brief summary of the properties of pooling and sharing in a heterogeneous environment. In Section 3 we study the model from the perspective of evolutionary game theory, and derive the criteria for evolutionary stability of unconditional cooperation on complex networks. Here, we also solve several analytically tractable examples that elucidate the role of the network structure. We also examine numerically the extent to which different complex network topologies/structures promote cooperative behavior. In Section 4 we present a behavioral strategy based on generalized reciprocity and study the properties of the resulting cooperation dynamics. The last section briefly summarizes our findings.

2. Model

2.1. Interaction structure

In our model we consider a population of N individual entities interacting through M pools. The pools may represent various compositions that facilitate the sharing of resources. For instance, in biological systems the pools may constitute structures that ease the transport of nutrients [31]. Similarly, in economics the pools may be seen as equivalent to baskets that are used to share commodities in community-supported agriculture [32]. In the model, the interaction structure is described by a connected bipartite random graph, whose adjacency matrix $\mathbf{B}$, with binary edge variables $B_{im} \in \{0, 1\}$, captures the participation of individual i in pool m ($B_{im} = 1$, indicating participation of i in pool m). The bipartite graph representation, although not essential for our concept, offers a principled way of capturing wider information regarding the group composition and network interactions [33, 34, 35] than standard unipartite graphs, where each node behaves as both an individual entity and a pool through which resources are shared [33]. As a result, bipartite graphs are a convenient tool for modeling pooling and sharing on networks. In particular, each unipartite graph can
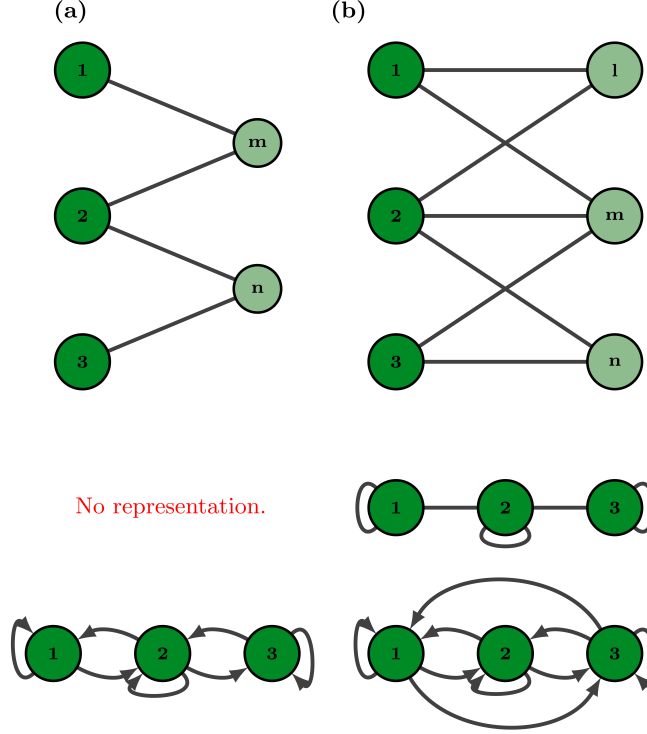
Figure 1: **Bipartite graph representation.** **(a)** First row: a bipartite graph consisting of tree entities and two pools; Second row: its unipartite representation; Third row: the weighted replacement graph. **(b)** Same as **(a)**, only now there are three pools and the connections between the entities and pools are slightly different.

be mapped into a bipartite graph by considering a replacement graph procedure, while the opposite is not true in general, as illustrated in Fig. 1. In this representation each node behaves as both an individual entity and a pool through which resources are shared [33]. We refer to [33, 35, 26, 36] for more details on the procedure for constructing a replacement graph.

We provide an example for this property in Fig. 1 where we draw two bipartite graphs, each composed of three entities. In addition, the first graph is composed of two pools (Fig. 1a), whereas in the second graph there is an additional third pool (Fig. 1b). In the second row of the figure we give the simple graph representation. For the first bipartite graph there is no such representation, but the second graph can be represented in a standard representation that allows for self-loops. Both bipartite graphs can be described as a unipartite only if we include more complex replacement graph representation that allows for edge directions and weights (bottom row of the figure).

We consider discrete-time dynamics and track the evolution of the individual resource endowments $y_i(t)$, for all $i \in \mathcal{N}$. The resource growth is governed by a multiplicative process, where in each period the resource growth of the i-th entity is modeled by a drift μ_i and a noise amplitude σ_i. Note that we consider heterogeneous population where μ_i and σ_i differ across the different members of the population. After the growth phase, the individual resources are a subject to a pooling phase where each entity i distributes a fraction $p_i(t)/d_i$ of its resources to each of the $d_i = \sum_m \mathrm{B}_{im}$ pools where it participates. Finally, every pool shares

a fraction $1/u_m$, with $u_m = \sum_i \mathrm{B}_{im}$, of the total pooled resources to each of its members. Formally, the evolution of the individual resources can be expressed as

$$y_i(t + \Delta t) = \underbrace{\sum_j A_{ij} p_j(t) y_j(t) \left[\Delta t + \mu_j \Delta t + \sigma_j \varepsilon_j(t) \sqrt{\Delta t} \right] - p_i(t) y_i(t) \Delta t}_{\text{resources due to pooling and sharing}}$$

$$+ \underbrace{y_i(t) + (1 - p_i(t)) y_i(t) \left[\mu_i \Delta t + \sigma_i \varepsilon_i(t) \sqrt{\Delta t} \right]}_{\text{individual resources}}, \qquad (1)$$

where $\varepsilon_j(t)$ is a standardized Gaussian random variable and $\mathbf{A}$ is a transition matrix of the network with entries $\mathrm{A}_{ij} = \sum_m^M \frac{B_{im}}{u_m} \frac{B_{jm}}{d_j}$ which determine the total allocated resources from entity j to entity i. The variable $p_i(t) \in [0, 1]$ models the level of cooperation displayed by an individual entity in a particular period. The case of $p_i(t) = 0$ for all t represents the situation when the entity always defects, i.e., the unconditional defector case. On the other hand, the case of $p_i(t) = 1$ for all t is the unconditional cooperator setting. In a more general setting, the entities propensity too cooperate may vary over time based on their environment and thus $p_i(t)$ may be interpreted as an internal state that captures the propensity of an entity to engage in pooling and sharing of its resources.

Following [17], we write Eq. (1) in a continuous form by setting $\Delta t \to 0$

$$\mathrm{d}y_i = \left[\sum_j \mathrm{A}_{ij} p_j y_j - p_i y_i \right] \mathrm{d}t + \sum_j \mathrm{A}_{ij} p_j y_j \left(\mu_j \mathrm{d}t + \sigma_j \mathrm{d}W_j \right) + (1 - p_i) y_i \left(\mu_i \mathrm{d}t + \sigma_i \mathrm{d}W_i \right), \qquad (2)$$

where $(\mathrm{d}W_i)_{i \in \mathcal{N}}$ are independent Wiener increments, i.e. $W_i(t) = \int_0^t \mathrm{d}W_i$. This representation allows us to utilize techniques such as Itô calculus for the theoretical analysis of the model. As in [17], when working with numerical estimations, we will resort the discrete formulation, with the note that in that case the temporal differences will be rescaled such that $\Delta t = 1$. This, as we will see, allows for an easy Markov representation of the discrete model.

The resulting interaction structure is directly related to games of public goods on networks. The main difference is that in this model the fluctuations arise inherently from the individual traits owned by the entities, whereas in a public goods game the randomness is a feature of each pool. In other words, in public goods games the stochasticity comes as a result of the process of pooling, here there is randomness even without pooling [33, 37, 3]. Another similar model is the Bouchaud–Mezard model of wealth reallocation and its extensions [38, 39, 28]. The essential difference is that in our description, reallocation is done after the growth phase.

2.2. Growth rate

In a situation with additive resource dynamics, the relevant quantity for measuring the payoff of an individual entity is the steady state time-averaged absolute change of wealth [21]. This is not the case with our model, since the fluctuations introduced by the multiplicative process governing the resource dynamics make the system non-stationary, and hence non-ergodic. Under these circumstances, the relevant quantity for measuring the payoff is the steady state time-averaged growth rate of wealth, $g_i(y_i(t), t)$. In evolutionary biology, this

observable is more known as the *geometric mean fitness* for the accumulated payoff (i.e. resources) of a particular phenotype [5]. Formally, it is defined as

$$g_i(y_i(t), t) = \frac{1}{t} \log\left(\frac{y_i(t)}{y_i(0)}\right), \tag{3}$$

where $y_i(0)$ is the initial amount of resources. In what follows, we are going to assume $y_i(0) = 1$.

When the internal state for cooperation of each entity i is $p_i^* = 0$, we have the unconditional defector situation and the solution to the model represents N independent geometric Brownian motion (GBM) trajectories. Then, the steady state time-average growth rate can be easily found using Itô calculus as

$$g_i^{\mathcal{D}} = \lim_{t \to \infty} g_i^{\mathcal{D}}(y_i(t), t) = \mu_i - \frac{\sigma_i^2}{2}, \tag{4}$$

where the superscript $\mathcal{D}$ denotes the case in which no individual entity pools its resources.

On the other hand, when every entity cooperates unconditionally, i.e. $p_i(t) = 1$ for all i and t, the growth rate of each entity converges to the same value,

$$g^{\mathcal{C}} = \lim_{t \to \infty} g_i(y_i(t), t) = \langle \mu v \rangle_{\mathcal{N}} - \frac{1}{2N} \langle v^2 \sigma^2 \rangle_{\mathcal{N}}, \tag{5}$$

where the superscript $\mathcal{C}$ denotes that every entity pools its resources, $\langle \cdot \rangle_{\mathcal{N}}$ is the population average and v is an index given by the right-eigenvector associated with the largest eigenvalue of $\mathbf{A}$ normalized in a way such that $\sum_i v_i = N$.

For cooperation to be evolutionary favored by the population, the quantity in Eq. (5) should be larger than the one of Eq. (4). This is always the case if the population is homogeneous, i.e. the individual entities are subject the same drift and noise amplitude [26]. In a heterogeneous population where the individual entities exhibit different traits (i.e. they are subject to different drifts and noise amplitudes), the evolution of cooperation is dependent on both the network structure and the individual traits. We elaborate on this in more detail in the sequel.

3. Evolutionary stability

3.1. Preliminaries

We begin the analysis by considering the circumstance in which every individual entity can behave either as an unconditional cooperator ($p_i(t) = p_i^* = 1$ for all t), or as an unconditional defector ($p_i(t) = p_i^* = 0$ for all t), and examine the Evolutionary Stable State (ESS). As discussed in Nowak [1], an ESS of unconditional cooperation is a situation in which a large population of cooperators cannot be invaded by defectors under deterministic selection dynamics. ESS has been widely used for determining the level of cooperative behavior in various interaction structures [40].

In a simple interaction structure consisting of two entities i and j, the possible outcomes are described by a payoff matrix

$$
\begin{array}{c c}
 & j \\
 & \begin{array}{c c} C & D \end{array} \\
i \quad \begin{array}{c} C \\ D \end{array} & \begin{array}{|c|c|} \hline \pi_{CC} & \pi_{CD} \\ \hline \pi_{DC} & \pi_{DD} \\ \hline \end{array}
\end{array}
\tag{6}
$$

where π_{ab} denotes the payoff of entity i under strategy a when entity j has chosen strategy b. As discussed, in our case the strategies $p_i^* = 1$ and $p_i^* = 0$ translate to being an unconditional cooperator (C), respectively defector (D), and the growth rates of the entities translate to the corresponding payoffs in the payoff matrix. This implies that unconditional cooperation is ESS if $\pi_{CC} > \pi_{DC}$. In interaction structures involving multiple entities, unconditional cooperation is ESS if the overall benefit of cooperation for an individual entity is greater than the benefit of defection, assuming that all other entities cooperate unconditionally. In other words, in the model with networked pooling and sharing, full cooperation will be ESS if and only if

$$
g^{\mathcal{C}} > \max_i g_i^{\mathcal{D}}.
\tag{7}
$$

3.2. Growth rate

In a more general setting, unconditional cooperation may only be favored by a certain subset $\mathcal{C}_l$ of the population, whereas for the other part it is optimal to behave as unconditional defectors. This intermediate case can lead to the emergence of multiple connected components where each component has its own growth rate depending on the network parameters. Recall, a connected component is a subset of the network in which any two nodes (entities) are connected to each other by paths, and which is connected to no additional nodes.

To provide a better intuition on this behavior, in Fig. 2 we illustrate three situations which lead to different growth rate outcomes by considering a simple network composed of five entities and two pools. In particular, in Fig. 2a we assume that unconditional cooperation is evolutionary stable for all entities, Therefore, in this case the growth rate of each individual entity converges to (9). On the other hand, in, Fig. 2b we set the growth rate of entity 3, which acts as bridge by being the only one to pool its resources in both pools, to be large enough so as this entity behaves as an unconditional defector. This creates two separate components $\mathcal{C}_1 = \{1, 2\}$ and $\mathcal{C}_2 = \{4, 5\}$, which pool and share resources only between themselves. Notice that the entities also pool part of their resources to entity 3 but they are not shared back. This implies that the resource dynamics, and hence growth rates, between the components are independent. If the growth rates of each component are different, this leads to great discrepancies in the observed resources between the cooperators belonging to separate components, with the resources of the entities in one component being negligible in comparison to the resources of the entities in the other component. This is a distinguishing feature of our model and occurs solely as a result of the heterogeneous individual capabilities, in combination with the network structure. The observed effect is formally equivalent to the freezing state in the Random Energy model of Derrida which has been extensively studied
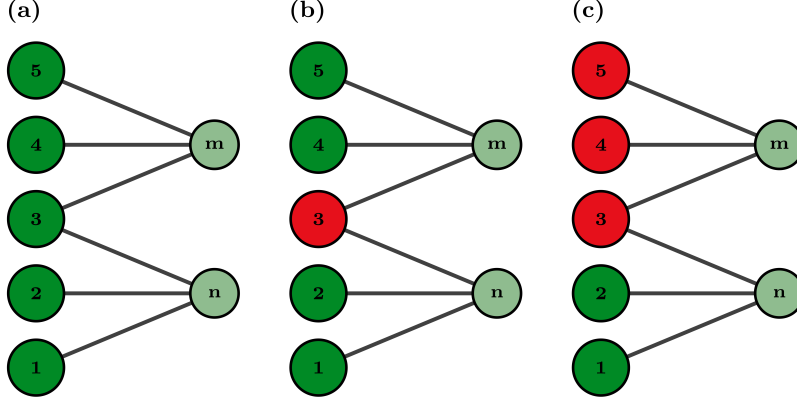
8

Figure 2: **Creation of cooperative components. (a-c)** Network composed of five entities and two pools. Green nodes are the entities which prefer unconditional cooperation, whereas the red nodes depict entities which favor defection.

in the statistical mechanics community [27]. Finally, in the last example, besides entity 3 we also set the entities in component $\mathcal{C}_2$ to be defectors. Then, there is only one connected component and each unconditional cooperator will have the same steady state growth rate. Hence, in this circumstance there is no freezing state among the entities belonging to the unconditional cooperator set.

To analyze the steady state growth rate g_i^* of the entities in component $\mathcal{C}_l$ we utilize a mean-field approach together with Itô calculus, as is done in [26]. The mean-field approach assures that each connected component of cooperators $\mathcal{C}_l$ will have a convergent growth rate given by the growth rate of the average $\langle y \rangle_{\mathcal{C}_l}$ constructed by the resource trajectories of the entities belonging to this component.

To prove this claim, we define the rescaled resources of each cooperating entity i as $\hat{y}_i(t) = \frac{y_i(t)}{\langle y \rangle_{\mathcal{C}_l}}$. Due to the normalization, in steady state, this observable also converges to a constant value $\hat{y}_i^*$ that is greater than zero and less then the number of unconditional cooperators. Hence, the growth rate of each entity i in the set of unconditional cooperators can be written as

$$
\begin{aligned}
\lim_{t\to\infty} g_i(y_i(t), t) &= \lim_{t\to\infty} \frac{1}{t} \log\left(\frac{y_i(t)}{y_i(0)} \right) \\
&= \lim_{t\to\infty} \frac{1}{t} \log\left(\langle y(t) \rangle_{\mathcal{C}_l} \cdot \hat{y}_i(t) \right) \\
&= \lim_{t\to\infty} \frac{1}{t} \log\left(\langle y(t) \rangle_{\mathcal{C}_l} \right) + \lim_{t\to\infty} \frac{1}{t} \log\left(\hat{y}_i(t) \right) \\
&= \lim_{t\to\infty} \frac{1}{t} \log\left(\langle y(t) \rangle_{\mathcal{C}_l} \right) \\
&\doteq \lim_{t\to\infty} g(\langle y(t) \rangle_{\mathcal{C}_l}, t).
\end{aligned}
$$

Consequently, one can use Itô's lemma to directly calculate the cooperative time-average growth rate. Formally, the lemma states that the differential of an arbitrary one-dimensional

function $f(\mathbf{y}, t)$ governed by an Itô drift-diffusion process is given by

$$df(\mathbf{y}, t) = \frac{\partial f}{\partial t} dt + \sum_i \frac{\partial f}{\partial y_i} dy_i + \frac{1}{2} \sum_i \sum_j \frac{\partial^2 f}{\partial y_i \partial y_j} dy_i dy_j. \tag{8}$$

In the case of $g(\langle y \rangle_{\mathcal{C}_l}, t)$, we have that $f(t, \mathbf{y}) = \log(\langle y \rangle_{\mathcal{C}_l})$. Then, the first and second derivative of f with respect to y_i and y_j are easily calculated as $\frac{\partial f}{\partial y_i} = \frac{1}{N_{\mathcal{C}_l}} \frac{1}{\langle y \rangle_{\mathcal{C}_l}}$ and $\frac{\partial^2 f}{\partial y_i \partial y_j} = -\frac{1}{N_{\mathcal{C}_l}^2} \frac{1}{\langle y \rangle_{\mathcal{C}_l}^2}$, where $N_{\mathcal{C}_l}$ is the number of cooperators in steady state. Moreover, this transformation makes the differential $df(\mathbf{y}, t)$ ergodic, and since we are looking at steady state averages, dy_i and $dy_i dy_j$ can be substituted with their expected values $\langle dy_i \rangle$ and $\langle dy_i dy_j \rangle$. To estimate these expectations we utilize the independent Wiener increment property $\langle dW_i^2 \rangle = dt$, and define $z_j^{[\mathcal{C}_l]} = \sum_{k \in \mathcal{C}_l} A_{kj}$. Further, we omit terms of order dt^2 as they are negligible. As a result, we obtain that

$$\langle dy_i \rangle = \left[\sum_{k \in \mathcal{C}_l} A_{ik}(1 + \mu_k) y_k - y_i \right] dt,$$

and,

$$\langle dy_i dy_j \rangle = \sum_{k \in \mathcal{C}_l} A_{ik} A_{jk} \sigma_k^2 y_k^2 dt.$$

By inserting the estimates in Eq. (8) we can approximate the time-average growth rate as

$$g(\langle y \rangle_{\mathcal{C}_l}, t) = \frac{1}{N_{\mathcal{C}_l}} \sum_{i \in N_{\mathcal{C}_l}} \hat{y}_i(t) \left[(1 + \mu_i) z_i^{[\mathcal{C}_l]} - 1 \right] - \frac{1}{2} \frac{1}{N_{\mathcal{C}_l}^2} \sum_{i \in \mathcal{C}_l} (z_i^{[\mathcal{C}_l]} \hat{y}_i(t))^2 \sigma_i^2. \tag{9}$$

Eq. (9) describes the growth of $\langle y \rangle_{\mathcal{C}_l}$ as a function of $\hat{y}_i(t)$. To derive its steady state behavior, notice that we can remove the unconditional defectors from the cooperative dynamics, set $\Delta t = 1$, and rewrite Eq. (1) as

$$y_i(t+1) \approx \sum_{j \in \mathcal{C}_l} A_{ij} y_j(t) \left[1 + \mu_j + \sigma_j \varepsilon_j(t) \right]. \tag{10}$$

When Eq. (10) is divided by the population average resources $\langle y(t+1) \rangle_{\mathcal{C}_l}$ and is written in matrix form, the steady state dynamics for the rescaled resources can be approximated as

$$\hat{\mathbf{y}}^* = \lim_{t \to \infty} \hat{\mathbf{y}}(t) \approx \lim_{t \to \infty} \frac{\mathbf{A}_{\mathcal{C}_l} \mathbf{y}(t)}{\langle \mathbf{A}_{\mathcal{C}_l} \mathbf{y}(t) \rangle_{\mathcal{C}_l}}, \tag{11}$$

where $\mathbf{A}_{\mathcal{C}_l}$ is a reduced version of the transition matrix $\mathbf{A}$ in which includes only the rows and columns associated to the entities in the set $\mathcal{C}_l$. By the power method, this leads to

$$\hat{y}_i^* = v_i^{[\mathcal{C}_l]}, \tag{12}$$

where $v_i^{[\mathcal{C}_l]}$ is the i-th element of the right-eigenvector of $\mathbf{A}_{\mathcal{C}_l}$ associated with the largest eigenvalue $\lambda_{\mathcal{C}_l}$ normalized in a way such that $\sum_i v_i^{[\mathcal{C}_l]} = N_{\mathcal{C}_l}$.

By inserting the estimates of Eq. (12) in Eq. (9) we get that the steady state cooperative growth rate is

$$g_{\mathcal{C}_l} = (\lambda_{\mathcal{C}_l} - 1) + \langle x^{[\mathcal{C}_l]} \mu \rangle_{\mathcal{C}_l} - \frac{1}{2} \frac{1}{N_{\mathcal{C}_l}} \langle (x^{[\mathcal{C}_l]})^2 \sigma^2 \rangle_{\mathcal{C}_l}, \tag{13}$$

where $x_i^{[\mathcal{C}_l]} = v_i^{[\mathcal{C}_l]} z_i^{[\mathcal{C}_l]}$, is a network centrality index whose relation with the drifts and amplitudes ultimately determines the cooperative growth rate.

10

3.3. Determining ESS

To determine the ESS in situations when there is a possibility that unconditional defectors and cooperators may coexist, we consider an alternate-projection method [41]. In particular, notice that the ESS conditions may be reformulated as

$$g_i^* = g_{\mathcal{C}_l}, \; i \in \mathcal{C}_l,$$
$$0 = \left(g_{\mathcal{C}_l} - g_i^{\mathcal{D}}\right)\left(1 - p_i^*\right),$$

resulting in a nonlinear system of $2N$ equations with $2N$ variables (g_i^* and p_i^*) in total. Hence, a simplified, iterative approach based on the alternate projection method can be used for finding the steady state solution. We follow [2] and summarize the method as follows:

1. Set $p_i^* = 0$ and $g_i^* = g_i^{\mathcal{D}}$ for all i not satisfying the condition (5). Set $p_i^* = 1$ for the remaining entities. To estimate their growth rate solve Eq. (13) for each component of unconditional cooperators $\mathcal{C}_l$ and set for each $i \in \mathcal{C}_l$, $g_i = g_{\mathcal{C}_l}$.
2. For all i satisfying $g_i^* < g_i^{\mathcal{D}}$ in the obtained solution, set $p_i^* = 0$ and $g_i^* = g_i^{\mathcal{D}}$. For each remaining component of cooperators, solve again the corresponding growth rate.
3. Repeat steps 1. and 2. until there are no $g_i^* < g_i^{\mathcal{D}}$.

3.4. Simple examples

To provide an illustrative representation for the evolutionary properties of the model in what follows we study three simple examples. The purpose of the first example is to show the evolutionary dynamics in the simplest population structure consisting of only two entities. The second example extends the population size to an arbitrary size of N entities. The last example shows how the position of the entities in an ordinary network structure can impact the individual payoff.

Example 1:. The first situation that we consider is a replication of the example studied in Ref. [17]. Concretely, we assume an interaction structure of two entities i and j who have the option to share their resources through one pool m, as illustrated in the right panel of Fig. 3.

For simplicity, we are going to assume that i has a larger individual steady state growth rate, (i.e. $g_i^{\mathcal{D}} > g_j^{\mathcal{D}}$ and examine the situation from the point of view of entity i, In this case, the payoff matrix for i reads

$$
\begin{array}{cc|c|c|}
& & \multicolumn{2}{c}{j} \\
& & C & D \\
\cline{3-4}
\multirow{2}{*}{i} & C & \langle\mu\rangle_{\mathcal{N}} - \frac{\langle\sigma^2\rangle_{\mathcal{N}}}{4} & 0 \\
\cline{3-4}
& D & \mu_i - \frac{\sigma_i^2}{2} & \mu_i - \frac{\sigma_i^2}{2} \\
\cline{3-4}
\end{array}
\tag{14}
$$

which, after some reordering, implies that full unconditional cooperation is ESS if

$$\mu_j - \frac{\sigma_j^2}{4} > \mu_i - \frac{3\sigma_i^2}{4}.$$

In the special case when $\mu_j = \mu_i$, the condition reduces to $3\sigma_i > \sigma_j$. This result is displayed in Fig. 3 where we plot the average steady state propensities for cooperation within the
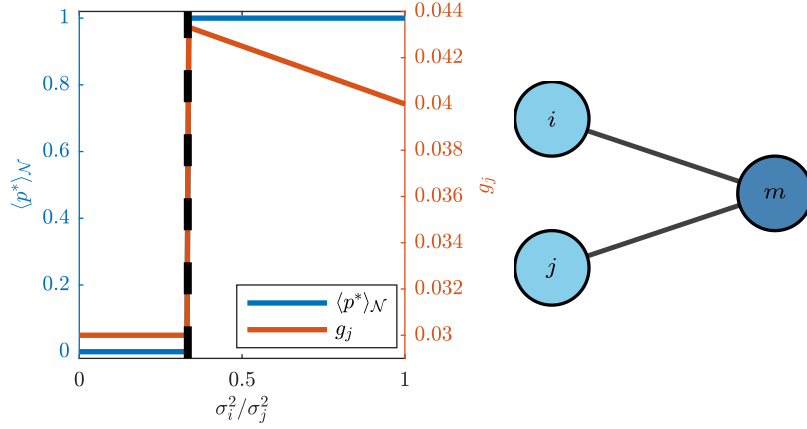
11

Figure 3: **Example 1.** Left panel: average fraction of unconditional cooperators $\langle p^* \rangle_{\mathcal{N}}$ and steady state growth rate g_j of entity j as a function of the noise amplitude ratio σ_i^2/σ_j^2. The dashed line is the threshold after which cooperation is evolutionary stable. The right panel depicts the interaction structure.
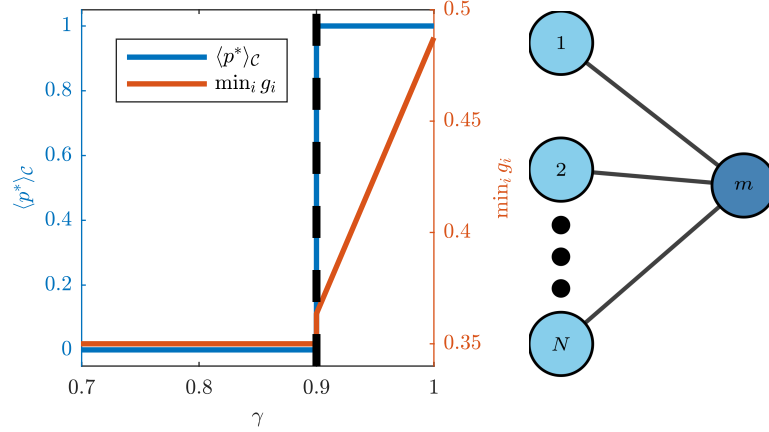

Figure 4: **Example 2.** Left panel: average fraction of unconditional cooperators $\langle p^* \rangle_{\mathcal{C}}$ and minimum steady state growth rate for the entities in the unconditional cooperator set $\mathcal{C}$ as a function of the fraction of population belonging to that set γ. The dashed line is the threshold for cooperation to be ESS among the entities in $\mathcal{C}$, whereas the right panel describes the interactions. In this case $N = 12$, $\mu = 0.5$ and $\sigma^2 = 0.3$.

population and the steady state growth g_j of entity j as a function of the square of the amplitude ratio σ_i/σ_j. We observe that at the critical point at $\sigma_i^2/\sigma_j^2 = 1/3$ the growth rate of entity j is largest and afterwards it is decreasing linearly due to the increase in the noise amplitude σ_i of entity i.

Example 2:. The second example extends the interaction to an arbitrary number of N entities, as depicted in the right panel of Fig. 4. This is the well-mixed situation which has been extensively utilized for determining the performance of a particular mechanism in the evolution of cooperation. As such this allows us to link our model to previous findings.

To ease the analysis, we assume that a fraction of the population has an individual growth rate larger than $g^{[\mathcal{N}]}$ and therefore the entities belonging to this group behave as unconditional defectors. The other fraction, γ, has an equal drift μ and noise amplitude σ.

As a consequence, in this example there is only one component of possible unconditional cooperators $\mathcal{C}$ and for each entity i in it we have that $v_i^{\mathcal{C}} = 1$ and $z_i^{[\mathcal{C}]} = \lambda_{\mathcal{C}} = \gamma$. Hence, the cooperative growth rate of this component is

$$g^{\mathcal{C}} = \gamma - 1 + \gamma\mu - \gamma\frac{\sigma^2}{2N}. \tag{15}$$

Clearly, under evolutionary dynamics, unconditional cooperation by the entities in the potential cooperator set will be favored only if

$$g^{\mathcal{C}} > \mu - \frac{\sigma^2}{2},$$

which, when rearranged in terms of γ, yields

$$\gamma > \frac{1 + \mu - \frac{\sigma^2}{2}}{1 + \mu - \frac{\sigma^2}{2N}}. \tag{16}$$

Fig. 4 visualizes the dependence of the minimum steady state growth rate of the potential unconditional cooperators as a function of the fraction of potential cooperators γ. It is easily noticed that after condition (16) is satisfied, the growth rate of the entities in the set $\mathcal{C}$ increases, which means that cooperation is an ESS for them.

Eq. (16) is a direct pooling and sharing counterpart to the evolutionary rules for cooperation that have been both numerically and analytically studied [1]. In particular, notice that when there is no noise ($\sigma = 0$), then unconditional cooperation is never favored. As such, the derived inequality is directly related to the result in [16] where the evolutionary dynamics in fluctuating environments were studied in a broader setting. The main difference is that our model allows us to easily infer the role of network structures in fluctuating environments as well as to study the presence of different behavioral update rules.

Example 3:. In the last example we once more examine a structure consisting of N entities which now do not construct a well-mixed population and instead interact through 2 pools, m and n. A fixed number of entities $N - 2$ interact solely through pool m and one entity, j, interacts only through pool n. In addition, there is one entity, i, which connects the network by pooling its resources in both m and n. This interaction structure is described in the right panel of Fig. 5.

We assume that a fraction $1 - \gamma$ of the individuals interacting solely through pool m have a large enough growth rate to behave as an unconditional defectors, and thus study the circumstances under which the remaining $N_{\mathcal{C}} = \gamma(N-2) + 2$ entities behave as unconditional cooperators $\mathcal{C}$. In this case the interaction matrix reads

$$\mathbf{A}_{\mathcal{C}} = \begin{bmatrix} \frac{1}{N-1} & \frac{1}{N-1} & \cdots & \frac{1}{N-1} & \frac{1}{2(N-1)} & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ \frac{1}{N-1} & \frac{1}{N-1} & \cdots & \frac{1}{N-1} & \frac{1}{2(N-1)} & 0 \\ \frac{1}{N-1} & \frac{1}{N-1} & \cdots & \frac{1}{N-1} & \frac{1}{2(N-1)} + \frac{1}{4} & \frac{1}{2} \\ 0 & 0 & \cdots & 0 & \frac{1}{4} & \frac{1}{2} \end{bmatrix}.$$
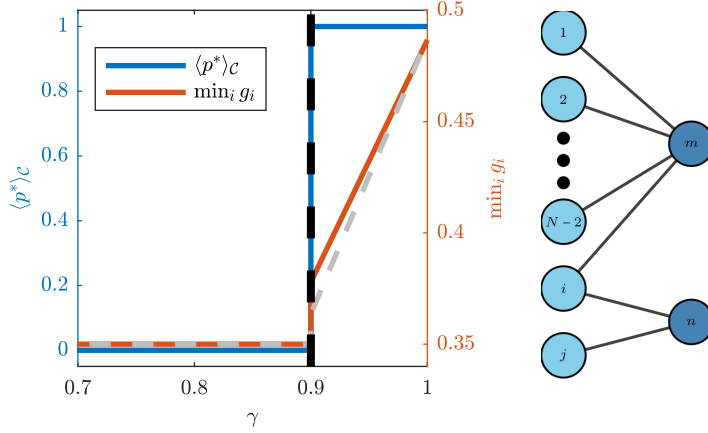
Figure 5: **Example 3.** Left panel: average fraction of unconditional cooperators $\langle p^* \rangle_{\mathcal{C}}$ and minimum steady state growth rate for the entities in the cooperator set $\mathcal{C}$ as a function of the fraction of potential unconditional cooperators γ. The dashed black line is the threshold for cooperation to be ESS among the entities in $\mathcal{C}$, whereas the grey line in the background gives the growth rate for the corresponding well-mixed system. The right panel describes the interactions. In this example $N = 12$, $\mu = 0.5$ and $\sigma^2 = 0.3$.

Moreover, the largest eigenvalue of $\mathbf{A}_{\mathcal{C}}$ is

$$\lambda_{\mathcal{C}} \approx \frac{|4\gamma - 3| + 4\gamma + 3}{8}, \tag{17}$$

with corresponding right-eigenvector entries $v_k^{[\mathcal{C}]} \approx \frac{(\gamma(N-2)+2)w_k^{[\mathcal{C}]}}{\sum_l w_l^{[\mathcal{C}]}}$ approximated for large N and a fixed $\gamma > 0$ by

$$w_k^{[\mathcal{C}]} = \begin{cases} \frac{|4\gamma-3|+4\gamma-1}{2}, & \text{if } k = i, \\ 1 & \text{if } k = j, \\ \frac{|4\gamma-3|(N-2)+4\gamma(N-4)+(11-3N)}{2(N-2)} & \text{otherwise,} \end{cases} \tag{18}$$

and $z^{[\mathcal{C}]}$ index entries

$$z_k^{[\mathcal{C}]} = \begin{cases} \frac{\gamma(N-2)+1}{2(N-1)} + \frac{1}{2} & \text{if } k = i, \\ 1 & \text{if } k = j, \\ \frac{\gamma(N-2)+1}{N-1}, & \text{otherwise.} \end{cases} \tag{19}$$

To illustrate the differences between this example and Example 2., we set each entity belonging to the potential set of unconditional cooperators an equal growth rate, $\mu_k = \mu$ and noise amplitude $\sigma_k = \sigma$.

Combining Eqs. (17), (18) and (19), one can derive an analytical expression for the cooperative growth rate (9). However, the resulting mathematical expression is rather complicated. Instead of writing it down, we visualize it in Fig. 5. The figure shows the cooperative growth rate as a function of γ. In the same figure, with the light gray line we plot the cooperative growth rate of the corresponding well-mixed system (Example 2). We observe that once unconditional cooperation is achieved by the entities in $\mathcal{C}$, the presence of a network structure effectively increases the growth rate of each individual, as compared to the well-mixed

14

situation. Moreover, even if the population size increases indefinitely, i.e., if we take the limit as $N \to \infty$ of Eqs. (17), (18) and (19), the cooperative growth rate will remain dependent on the network structure and the fraction of unconditional cooperators γ. This can be easily deduced from Eq. (17), as the equation is not dependent on the population size, but rather on the fraction of potential unconditional cooperators which interact solely through pool m and the non-trivial interaction between entities i and j. This serves as an intuitive example that the interaction structure may have a disproportionate effect on the observed individual behavior.

3.5. Random graphs

In a more general, random graph structure, besides the size of the unconditional defector set, the positioning of the entities in the network also determines the ESS for each other entity. While Eq. (9) and the alternate projection method give the solution for the level of cooperation in any arbitrary situation, the extent to which different random graphs are able to support cooperation in the presence of defectors can not be deduced explicitly.

To provide an insight on this phenomena, we consider four different types of random graph models i) Random d-regular (RR) graph [42], ii) Erdos-Renyi (ER) random graph [43], iii) Watts-Strogatz (WS) random graph [44], and iv) Barabasi-Albert (BA) random scale-free network [45] and study their robustness in the presence of arbitrary positioned and/or number of defectors. In an RR graph each entity is characterized with the same degree d, whereas in an ER graph two entities share an edge with probability d/N. Both types of random graphs yield homogeneous degree distributions and low clustering coefficients. The WS random graph, on the other hand is an extension of the two aforementioned graphs which is able to capture higher levels of clustering. In fact, this random graph type lies in-between the ER and RR graph as its creation starts with an initial structure of an RR graph and then each edge is re-wired with a fixed probability. Finally, the BA graph is constructed through a preferential attachment mechanism for generating random graphs. As such it yields a power law degree distribution coupled with high clustering.

To assess the robustness of the types of random graphs we conduct the following experiment. We begin by generating a random graph through the typical algorithms that are used for this procedure. Afterwards, we choose a random fraction $1 - \gamma$ of the population. We set the growth rate of the entities in it to be large enough so as they behave as unconditional defectors. We characterize the entities in the other set, the potential unconditional cooperators, with an equal drift μ and noise amplitude σ. Under these circumstances, we estimate the steady state growth rates of each entity i belonging to the cooperative cluster $\mathcal{C}_l$ as

$$g^{\mathcal{C}_l} = (\lambda_{\mathcal{C}_l} - 1) + \mu \langle x^{[\mathcal{C}_l]} \rangle_{\mathcal{C}_l} - \frac{\sigma^2}{2N_{\mathcal{C}_l}} \langle (x^{[\mathcal{C}_l]})^2 \rangle_{\mathcal{C}_l}$$
$$= (\mu + 1) \lambda_{\mathcal{C}_l} - \frac{\sigma^2}{2N_{\mathcal{C}_l}} \langle (x^{[\mathcal{C}_l]})^2 \rangle_{\mathcal{C}_l} - 1, \tag{20}$$

using the the procedure described in Section 3.3. We gather the average number of unconditional cooperators among the entities in the potential unconditional cooperator set, $\langle p^* \rangle_{\mathcal{C}}$. To get the typical behavior of a particular random graph we average across random graph instances and across defector selection samples.

15

Numerical results are summarized in Fig. 6a where we plot the average p_i^* among the entities in the cooperator set as a function of γ. We observe that for every γ the RR graph presents itself as the most supportive for cooperation as it requires the lowest amount of entities in the unconditional cooperator set for cooperation to exist. It is followed by the WS graph, while under this criterion the ER and BA graphs have similar performance and are the least robust random graphs.

The reason for this behavior becomes apparent if we look at the evolution of $\lambda_{\mathcal{C}}$ and $\langle (x^{[\mathcal{C}_l]})^2 \rangle_{\mathcal{C}_l}$ as a function of the fraction of entities belonging in the potential unconditional cooperator set. Since $\mu > \sigma^2$, $\lambda_{\mathcal{C}}$ is decisive in the determination for the level of cooperation, the networks with a larger eigenvalue will be better promoters of cooperation. In this aspect, we observe in Fig. 6b that the RR graph always has the largest eigenvalue, followed by the WS graph. As γ increases, the differences in $\langle (x^{[\mathcal{C}_l]})^2 \rangle_{\mathcal{C}_l}$ become apparent, whereas $\lambda_{\mathcal{C}}$ slowly converges to the same value for each network since in each random graph case as it becomes a stochastic matrix. Because $\langle (x^{[\mathcal{C}_l]})^2 \rangle_{\mathcal{C}_l}$ is related to the amplitude of the noise, the networks with lower value will be better promoters of cooperation. As depicted in Fig. 6c, the RR graph again is the graph with the largest magnitude of the observable, but its eigenvalue is already converged to the maximum value. Thus it is less affected by the properties of $\langle (x^{[\mathcal{C}_l]})^2 \rangle_{\mathcal{C}_l}$. We arrive at similar conclusions when inspecting the properties of the other graphs. We hereby note that in a structure where the noise amplitude has a larger magnitude, $\langle (x^{[\mathcal{C}_l]})^2 \rangle_{\mathcal{C}_l}$ will play a bigger role in determining the steady state cooperative behavior.

To provide robustness of the network properties that we discussed in terms of the relationship between $\lambda_{\mathcal{C}}$ and $\langle (x^{[\mathcal{C}_l]})^2 \rangle_{\mathcal{C}_l}$ and γ, in Fig 7 we provide the corresponding standard deviations, $S_{\lambda_{\mathcal{C}}}$ and $S_{\langle (x^{[\mathcal{C}_l]})^2 \rangle_{\mathcal{C}_l}}$ of the numerical results as a function of the fraction of entities belonging in the potential unconditional cooperator set. They have a significantly smaller magnitude than the averages shown in Fig. 6, thus indicating that the averaged results represent an adequate representation for the typical behavior of the graphs.

4. Generalized reciprocity

So far, we addressed evolutionary behavior in which pooling and sharing is the sole cooperative mechanism. We showed that, as a consequence of the non-ergodicity, if certain conditions are satisfied, unconditional cooperation can be evolutionary stable even without the presence of any additional decision making mechanisms. In fact, the presence of additional auxiliary mechanisms should yield dynamics that at least complement the evolution of cooperation. To provide an initial insight on the role of these mechanisms, here we examine the cooperative behavior in the presence of a state-based generalized reciprocity update rule.

Generalized reciprocity suggests that cooperation can emerge as a consequence of previous positive experience with not necessarily the same group of opponents, i.e. it is based on the rule of "help anyone if helped by someone". This is significantly different from the two main forms of reciprocity, direct and indirect, which explain the emergence of cooperation either as a result of repeated encounters between the same group of entities or as an attempt to build positive reputation for future interactions [46, 47].

The main presumption which favors generalized reciprocity over other reciprocal mechanisms is that individual entities following this rule can be said to exhibit a simple state-based behavior. Due to this straightforward behavioral update rule, generalized reciprocity has
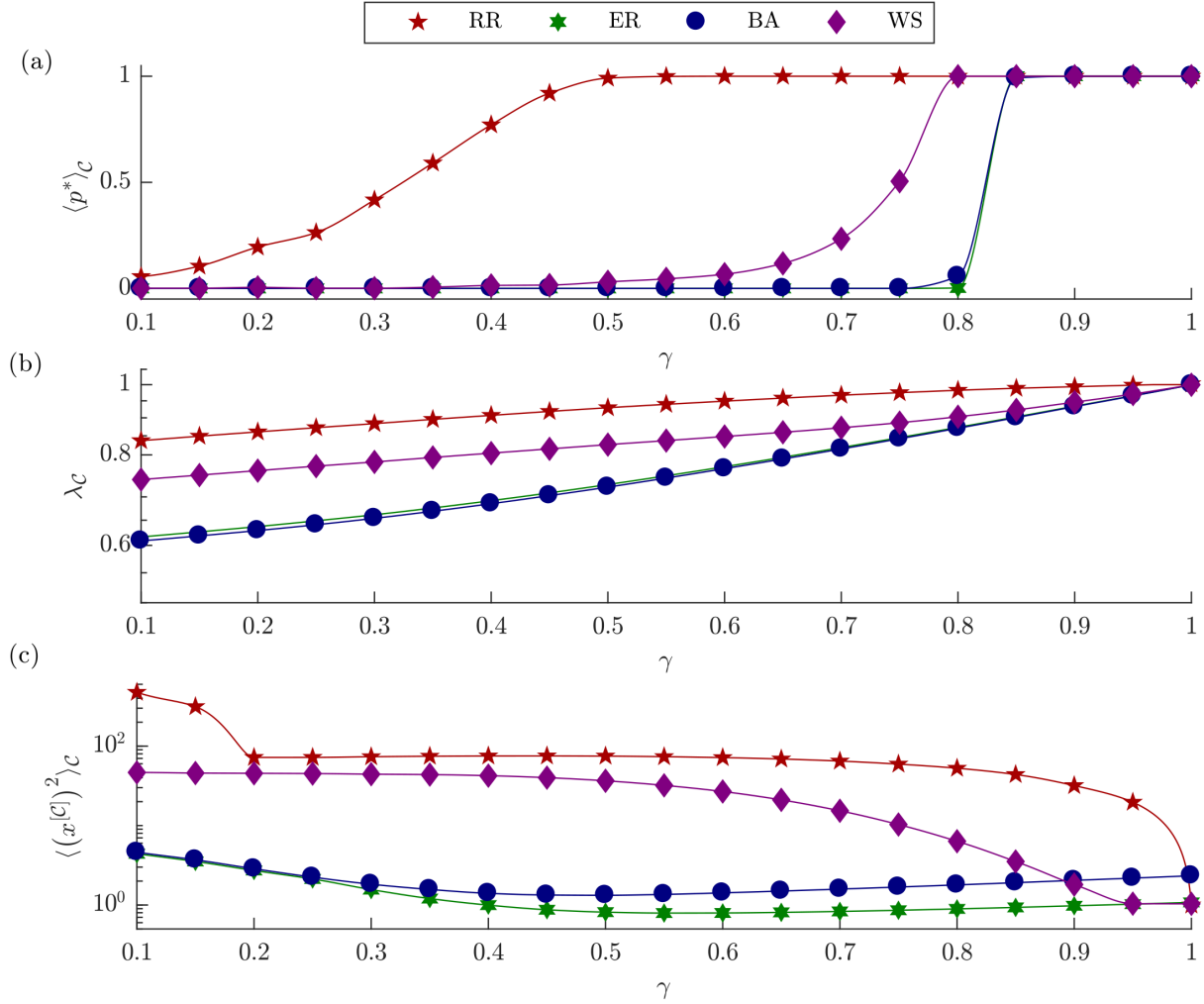
Figure 6: **The role of complex networks in heterogeneous populations.** **(a)** Average steady state propensity for cooperation $\langle p^* \rangle_{\mathcal{C}}$ as a function of the fraction of potential unconditional cooperators γ for RR, ER BA and WS graphs. **(b)** Average largest eigenvalue $\lambda_{\mathcal{C}}$ of the matrix $A_{\mathcal{C}}$ as a function of γ for the same networks. **(c)** Average $\langle (x^{[\mathcal{C}_l]})^2 \rangle_{\mathcal{C}_l}$ as a function of γ. In the simulation $\mu = 0.25$ and $\sigma^2 = 0.3$. The results are averaged across 100 different unconditional defector choices and 100 graph realizations with each graph having 2000 nodes and an average degree of 4.

been observed in a wide range of animal and human societies each manifesting different level of cognitive prowess and interacting in various environments [48, 49, 50, 51, 52].

While the extent to which generalized reciprocity is able to evolve as a sole cooperation mechanism has been a subject to an active debate, recent studies have shown that once this mechanism is present in a system, it induces dynamics which assist the stability of cooperation [53, 54, 55, 56, 57]. Concretely, in [2, 58] it was argued that state-based generalized reciprocity effectively prevents the individual entities from being exploited by unconditional defectors, whereas in [3] it was suggested that this prevention is accompanied with maximization of the level of cooperation displayed by each entity. Nevertheless, the theoretical work done so far has assumed that the entities interact in an additive environment where
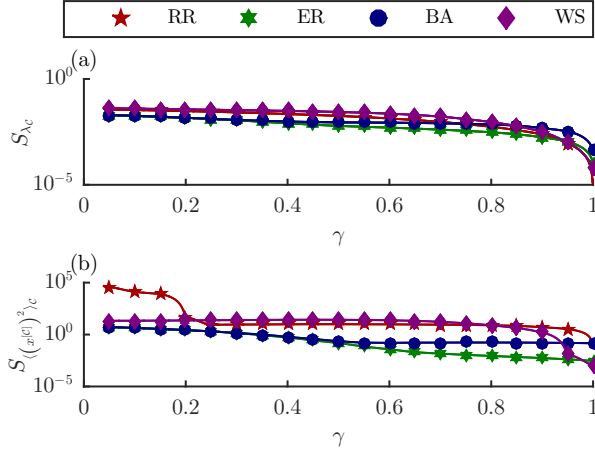
17

Figure 7: **Standard deviation of the network properties.** (a) Standard deviation $S_{\lambda_\mathcal{C}}$ of the largest eigenvalue $\lambda_\mathcal{C}$ of the matrix $A_\mathcal{C}$ as a function of γ for the same networks. (c) Standard deviation $S_{\langle (x^{[\mathcal{C}_l]})^2 \rangle_{\mathcal{C}_l}}$ of $\langle (x^{[\mathcal{C}_l]})^2 \rangle_{\mathcal{C}_l}$ as a function of γ. In the simulation $\mu = 0.25$ and $\sigma^2 = 0.3$. The results are averaged across 100 different unconditional defector choices and 100 graph realizations with each graph having 2000 nodes and an average degree of 4.

ensemble averages are a good approximation for the stochastic behavior. In this section we extend the rule to account for possible resource dynamics driven by a multiplicative noise.

4.1. Behavioral update

To study the individual behavior under a simple generalized reciprocity rule we consider an update based on the entity's estimate for its growth rate

$$p_i \left(t + \Delta t \right) = f_{i,t} \left[g_i(y_i(t), t) \right], \tag{21}$$

where $f_{i,t} : \mathbb{R} \to [0, 1]$ is monotonic (nondecreasing). A plausible choice would be the sigmoid (logistic) function

$$f_{i,t}(\omega) = \left[1 + \exp\{ (-\kappa_i(t))(\omega - \omega_i) \} \right]^{-1}, \tag{22}$$

where the midpoint ω_i is given by the steady state of $g_i(y_i(t), t)$ without pooling, i.e., $g_i^{\mathcal{D}}$.

We remark that we purposefully allow for the steepness of the sigmoid function to be an unbounded monotonically increasing function of t, so as to account for the time-dependence in the variance of $g_i(y_i(t), t)$. For simplicity, we focus on the special case when $\kappa(t) = t^\alpha$ where α is a positive parameter that captures the learning rate of the entities. In particular, $\alpha < 1$ corresponds to convergence towards equilibrium propensities to cooperate at a rate lower than the elapsed time t, whereas $\alpha > 1$ provides the opposite dynamics. The variable $\kappa(t)$ in our model is directly related to the intensity of choice parameter used in standard Choice theory (it is the analogue of the inverse temperature in statistical mechanics) [59].

We point out that the introduced rule provides a simple description for the cooperative behavior in a wide range of interaction structures. The advantage of the behavioral update lies in its simplicity since (21) implies that an entity only has to know its current amount

18

of resources in order to determine the next action. This is significantly different from other behavioral update rules where entities are required to optimize over domains depending on both opponent behavior and possible future interactions [58].

4.2. Model properties

In the analysis performed in Section 3 we always ended up with dynamics that have a steady state. However, the introduction of the state-based generalized reciprocity rule may in fact disturb this property of the model and we may end up with complex dynamics whose study is out of the scope of the paper.

To greatly ease the analysis of generalized reciprocity here we consider two situations. In the first situation, we examine the individual cooperative behavior under the assumption that there is a steady state growth rate g_i^* for every entity i. In the second situation, we study the properties of the model numerically and derive an analytical solution for the model in the circumstance when $\alpha < 1$ since, as it will be shown, there is always a steady state.

Cooperative behavior:. Let us assume that there is a steady state cooperative behavior p_i^* and growth rate g_i^* for each entity i conforming to the rule (21). In addition, we assume that the growth rate of each entity is set such that the system is non-degenerate. By non-degenerate we mean that $g_i^{\mathcal{D}} \neq g^{\mathcal{C}}$ for all i. Then, we can derive the following properties of the model

i. *Prevention of exploitation:* – It can be easily shown that $g_i^* < g_i^{\mathcal{D}}$ is an impossible situation for any entity following the behavioral update rule. To see this, assume that entity i follows (21) and has $g_i^* < g_i^{\mathcal{D}}$. Then, the behavioral update rule indicates that the steady state propensity for cooperation of this entity is $p_i^* = 0$. Subsequently, this implies that

$$y_i(t + \Delta t) \geq y_i(t)\left[1 + \mu_i \Delta t + \sigma_i \varepsilon_i(t)\sqrt{\Delta t}\right]. \tag{23}$$

Taking the limit as $t \to \infty$ we get that $g_i^* \geq g_i^{\mathcal{D}}$, thus contradicting our initial assumption. This is a favorable property of state-based generalized reciprocity which has been also observed in additive dynamics [2]. It significantly differs from other forms of generalized reciprocity since it has been shown that they are prone to exploitation [56].

ii. *Sufficient condition for existence of unconditional cooperators:* – The behavioral update rule coupled with the monotonicity and unboundedness of $\kappa_i(t)$ imply that if $g_i^* > g_i^{\mathcal{D}}$ then $p_i^* = 1$. Moreover, it follows that a necessary condition for $p_i^* < 1$ is $g_i^* = g_i^{\mathcal{D}}$.

Altogether, in steady state the entities may thus be attributed to two (disjoint) sets, $\mathcal{D}_{gen} = \left\{d \in \mathcal{N} : p_d^* < 1\right\}$ and $\mathcal{C}_{gen} = \{c \in \mathcal{N} : p_c^* = 1\}$, depending on the steady state propensity for cooperation p_i^*. The entities in $\mathcal{D}_{gen}$ are further characterized by $g_d^* = g_d^{\mathcal{D}}$, while the entities y_c in $\mathcal{C}_{gen}$ have an unknown growth rate $g_c \geq g_c^{\mathcal{D}}$, which is dependent on the network parameters. We will refer to the entities in the sets $\mathcal{D}_{gen}$ and $\mathcal{C}_{gen}$ as "generalized defectors", respectively "generalized cooperators", with an intention to emphasize their role in the pooling and sharing mechanism.

Numerical experiment:. The behavior of the individual growth rates ultimately depends on the magnitude of α. This parameter indicates the speed at which every entity reacts to the environment: if an entity has a smaller-than-expected growth, it will converge towards unconditional defection faster (proportionally with α), and otherwise towards unconditional cooperators. Even if the fraction of resources that is pooled and shared is small, the amount may be non-negligible for some of the entities which behave as unconditional defectors in evolutionary sense and, in fact, it may significantly increase their growth rate, thus making them "generalized cooperators". In other words, while the rule asserts that the entities which experience lower growth than their own eventually end up as generalized defectors, the number of generalized defectors may in general be smaller than the number of defectors inferred through evolutionary stability analysis.

To illustrate this effect we make use of the network given in Fig. 2. Concretely, we initialize the drift and amplitudes of the entities so as under evolutionary analysis $p_3^* = 0$ and cooperation is unstable for every other entity. Then, we simulate the coupled dynamics of Eq. (1) and Eq. (21) for 2×10^4 time steps and record the propensities for cooperation among the entities in the potential generalized cooperator set at the last point in time.

Fig. 8a provides a boxplot for the propensities for cooperation for each entity. In general, we observe two different regimes depending on α. The first appears when $\alpha < 1$. In this case, the propensity to cooperate p_i for each entity i appears to be at a value less than 1 but larger than 0. As α increases, the propensities to cooperate for each $i \neq 3$ also increase, eventually converging to $p_i = 1$. In contrast, the propensity of entity 3 decreases and converges to $p_3 = 0$. We assert that for the lower values of α the system is not in steady state and is eventually converging to the steady state where $p_i^* = 1$ for all $i \neq 3$, $p_3^* = 0$ and $g_i^* = g_3^{\mathcal{D}}$. This is evidenced in Fig. 8b where we plot the numerically observed growth rate for each entity as a function of α. Concretely, we notice that the observed growth rate of each $i \neq 3$ is larger than their defector growth rate, thus implying that $p_i^* = 1$. In contrast, the growth rate of 3 is smaller than its defector growth rate. As a consequence, the propensity to cooperate of this entity must be converging to 0. This slow convergence is illustrated in Fig. 8c where we visualize the dynamics over time of each $p_i(t)$ for various α.

The second regime occurs when $\alpha > 1$. In this situation we observe that the propensities for cooperation for each entity i vary from one simulation to another. One explanation for this phenomena is the fact that when $\alpha \geq 1$ the steepness of the logistic curve diverges with a rate that is at least as fast as the divergence of the resources of each entity. This reduces the domain of all $p_i(t)$ to $\{0, 1\}$ faster. Since growth rate of each entity is highly stochastic we may observe non-equilibrium dynamics even though we numerically find that on average the growth rate of each potential generalized cooperator is larger than expected (Fig. 8b).

Analytical solution:. The general case of $\alpha < 1$ in any interaction structure can be easily analytically solved and thus the observations explained. To see this, let i represent the strongest entity in the network in the sense that $g_i^{\mathcal{D}} > g^{\mathcal{C}}$ and $g_i^{\mathcal{D}} > g_j^{\mathcal{D}}$ for all $j \neq i$. We claim that the limit of the growth of each entity is at least as large as $g_i^{\mathcal{D}}$. For every finite time t we approximate the growth rate $g_j(t)$ of each entity $j \neq i$ and notice that this quantity is bounded from below by the the growth induced by the shared resources from the strongest entity. Since the defection is slow enough, these shared resources are enough to bring the growth of each $j \neq i$ to $g_i^{\mathcal{D}}$. The formal proof is given in the Appendix.
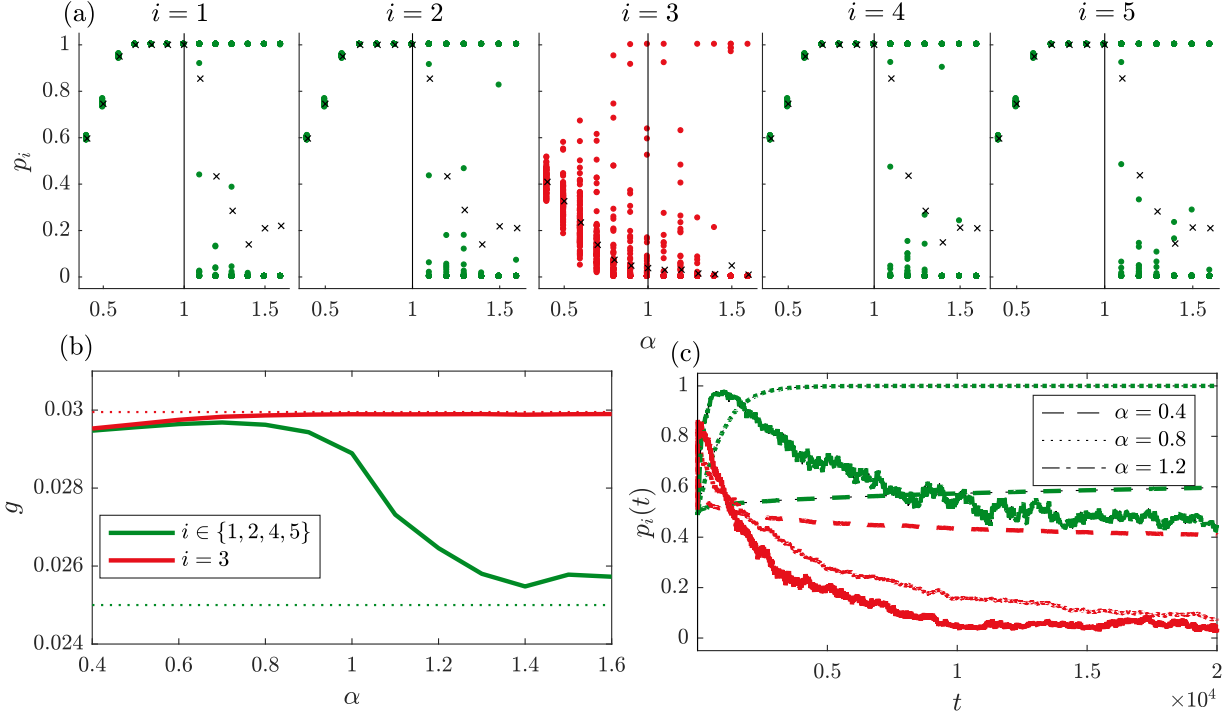
Figure 8: **State-based generalized reciprocity in fluctuating environments.** (a) Observed propensities for cooperation $p_i(t)$ for each entity at $t = 2 \times 10^4$ over 100 realizations. The circles are the results from each realization, whereas the crosses are the estimated averages across realizations for a given α. (b) Numerically observed growth rate averaged across realizations. (c) Dynamics over time for the average $p_i(t)$ across realizations. The green lines are the averages for entities $i \in \{1, 2, 4, 5\}$ whereas the red lines are the averages for $i = 3$. (a-c) The parameters are set to $\mu_i = 0.03$ for all i, $\sigma = \sqrt{0.01}$ for $i \neq 3$ and $\sigma = 0.01$ for $i = 3$.

The previous discussion leads to the situation where each entity exhibits a growth rate equal to the growth rate of the strongest entity $g_i^{\mathcal{D}}$. In turn, from the cooperative behavior properties *i.* and *ii.* we infer that $p_j^* = 1$ for all $i \neq j$ and $p_i^* < 1$.

5. Conclusion

We investigated the evolution of cooperation in networked heterogeneous fluctuating environments. We found out that the fluctuations induce evolutionary behavior which may lead to emergence of components within the population structure. The properties of each component are characterized solely by the interaction topology and the traits possessed by the entities belonging in it. Thus, we may observe great disparities in the owned resources between cooperating entities belonging to different components.

Moreover, by introducing a simple behavioral update rule we showed that state-based generalized reciprocity enhances the promotion of cooperation in fluctuating environments. The rule induces dynamics under which each entity is prevented from exploitation, i.e., each i observes a growth rate that is at least the same size as its own growth when behaving as an unconditional defector. More importantly, by construction, the entities whose observed growth rates at time t that are greater than in the unconditional defector situation, also

21

display greater propensities to cooperate. For the regime in which the learning rate is slower than the temporal dynamics, we analytically derived the exact behavior of each entity. When the learning rate is faster than the temporal dynamics we numerically observed complex dynamics. We argue that the observed dynamics is a consequence of the noise generating process having a larger impact on the evolutionary result and may produce situations with no stable state. In these situations, there is a never ending cycle in which the behavior of each entity alternates between unconditional cooperation and unconditional defection. While a detailed study of the dynamics is required, we believe that its implications are non-trivial and may contribute in explaining events such as the co-evolution of life cycles and multicellularity [60, 61, 62].

The implementation of our results goes beyond explaining the evolution of cooperation. In particular, the introduced rule is directly related to the concept of novelty search where individual entities decide their next actions on the basis of previous experience [30]. Novelty search is omnipresent in reinforcement learning and has been utilized in developing machines that efficiently mimic human behavior. In this aspect, we argue that the results discovered here behave as a building block in constructing machines which interact in a fluctuating environment.

6. Appendix

Here we provide the mathematical details of our claim in Section 4.2 c. Analytical solution.

Theorem 1. *Let $y_k(t)$ be a solution to Eq. (2) with $\alpha < 1$ and assume that g_k^* exists for all k. In addition, assume that i represents the strongest entity in the sense that $g_i^{\mathcal{D}} > g^{\mathcal{C}}$ and $g_i^{\mathcal{D}} > g_j^{\mathcal{D}}$ for all $j \neq i$. Then $g_k^* = g_i^{\mathcal{D}}$ for all k. Moreover, $p_j^* = 1$ for $j \neq i$ and $p_i^* < 1$.*

Proof. We argue by contradiction. Namely, let $g_j^* < g_i^{\mathcal{D}}$, and let entities i and j take part in a common pool, i.e. $A_{ji} > 0$ and $A_{ij} > 0$. For any $T > T_0 > 0$ the solution to (2) can be written as

$$y_j(T) = \int_{T_0}^{T} A_{ji} p_i(t) y_i(t) \mathrm{d}t - p_j(t) y_j(t) \mathrm{d}t + \left(\int_{T_0}^{T} G_{T,T_0}(t) \right),$$

where $G_T(t)$ is given as

$$G_{T,T_0}(t) = \frac{y_j(T_0)}{T - T_0} + \sum_{k \neq i} A_{jk} p_k(t) y_k(t)(\mathrm{d}t + \mu_k \mathrm{d}t + \sigma_k \mathrm{d}W_k) +$$
$$+ A_{ji} p_i(t) y_i(t)(\mathrm{d}t + \hat{\mu}_i \mathrm{d}t + \sigma_i \mathrm{d}W_i) + (1 - p_j(t)) y_j(t)(\mu_j \mathrm{d}t + \mathrm{d}W_j).$$

The integral of this expression gives the total amount of resources of entity j gathered from sources other than i, as well as resources from i with a changed drift, $\hat{\mu}_i := \mu_i - 1$ between times T_0 and T. Therefore it is non-negative. We can use this fact to arrive at a lower bound of $g_j(T) = \frac{1}{T} \log(y_j(T))$ for $T - 1 > T_0 \gg 0$,

$$g_j(T) \geq \frac{1}{T} \log \left(\int_{T_0}^{T} (A_{ji} p_i(t) y_i(t) - p_j y_j(t)) \, \mathrm{d}t \right)$$
$$\geq \frac{1}{T} \log \left(\int_{T_0}^{T} (A_{ji} p_i(t) y_i(t) - y_j(t)) \, \mathrm{d}t \right),$$

with the second inequality following from the fact that $p_j \le 1$. The fact that the logarithms of the two integrals are well defined stems from the fact that because $g_j < g_i^{\mathcal{D}}$ we have $A_{ji}p_i(t)y_i(t) - y_j(t) > 0$ for all $t > T_0$ with T_0 set to be large enough. The details are analogous to the ones in the justification of the passage to the limit that can be found below in inequality (25).

We can further relax the inequality by applying the Gronwall-Bellman lemma [63, 64] (or, alternatively, the assumption of the existence of a finite g_i^*) to the properties of the update rule (21) of $p_i(t)$. Formally, the lemma implies that the norm $\|z(t)\|$ of a solution to a differential equation $dz = A(z(t), t)z\,dt + B(z(t), t)z\,dW$ with continuous bounded coefficients $|A| + |B| < M$ is bounded from above by a function of the form $t \mapsto \theta_1 \exp(\theta_2 t)$. In other words, $\|z(t)\| \le \theta_1 \exp(\theta_2 t)$ for some coefficients $\theta_1, \theta_2 \ge 0$ and all $t > 0$. Since the differential equation defining $y(t)$ is with bounded coefficients, we can apply the lemma to $y_i(t)$ to deduce that $g_i(t)$ is also bounded. In our case we have

$$p_i(t) = \frac{1}{1 + \exp[-t^\alpha(g_i(t) - g_i^{\mathcal{D}})]} \ge \frac{1}{\xi + \exp(\eta t^\alpha)},$$

with constants $\xi, \eta \ge 0$ dependent only on the realization. This reduces the lower bound to

$$g_j(T) \ge \frac{1}{T} \log\left(\int_{T_0}^T \left(\frac{A_{ji}}{\xi + \exp(\eta t^\alpha)} y_i(t) - y_j(t) \right) \mathrm{d}t \right)$$

$$\ge \frac{1}{T} \log\left(\int_{T-1}^T \left(\frac{A_{ji}}{\xi + \exp(\eta t^\alpha)} y_i(t) - y_j(t) \right) \mathrm{d}t \right).$$

Next, by applying Jensen's inequality to $x \mapsto \log(x)$ we further get that

$$g_j(T) \ge \frac{1}{T} \int_{T-1}^T \log\left(\frac{A_{ji}}{\xi + \exp(\eta t^\alpha)} y_i(t) - y_j(t) \right) \mathrm{d}t. \tag{24}$$

By taking the limit of this expression we get that

$$g_j^* \ge g_i^*,$$
$$\ge g_i^{\mathcal{D}},$$

with the last inequality following from the prevention of exploitation property described previously. Altogether, this is the desired contradiction. Since the graph is connected, the property $g_j^* \ge g_i^{\mathcal{D}}$ is propagated to each entity. Now if $g_i^* > g_i^{\mathcal{D}}$, we would have $p_i^* = 1$, and since all $p_j^* = 1$ by $g_j^* > g_j^{\mathcal{D}}$, the growth g_i^* would be equal to $g^{\mathcal{C}}$, which is impossible since $g_i^* \ge g_i^{\mathcal{D}} > g^{\mathcal{C}}$ where the first inequality is the prevention of exploitation property. $\square$

Let us finish by justifying the passage to the limit of the right hand side in the inequality (24). First notice that for $k \in \{i, j\}$ we have $y_k(t) = \exp(tg_k^* + \varphi_k(t))$ for some functions

$\varphi_k(t)$ such that $\lim_{t\to\infty} \frac{\varphi_k(t)}{t} = 0$. Since $g_i^* \geq g_i^{\mathcal{D}}$, we have,

$$
\begin{aligned}
y_i(t) &= \exp(tg_i^* + \varphi_i(t)) \\
&\geq \exp\left(tg_i^{\mathcal{D}} + \varphi_i(t)\right), \\
&\geq \exp\left(tg_j^* + t\varepsilon + \varphi_i(t)\right), \\
&\geq_{+\infty} \exp\left(tg_j^* + t\frac{\varepsilon}{2} + \varphi_j(t)\right), \\
&= \exp\left(t\frac{\varepsilon}{2}\right)y_j(t),
\end{aligned}
\tag{25}
$$

for all large t, and for any small $\varepsilon > 0$ so that $g_i^{\mathcal{D}} > g_j^* + \varepsilon$. Indeed, there exists $t_\varepsilon > 0$ such that $t\frac{\varepsilon}{2} \geq |\varphi_j(t)| + |\varphi_i(t)|$ for all $t > t_\varepsilon$. In turn, for large T,

$$
\frac{y_j(T)(\xi + \exp(\eta T^\alpha))}{y_i(T)A_{ji}} \leq \frac{\xi + \exp(\eta T^\alpha)}{A_{ji}\exp\left(T\frac{\varepsilon}{2}\right)} \xrightarrow[T\to\infty]{} 0,
$$

implying that

$$
\frac{1}{T}\log\left(\frac{A_{ji}}{\xi + \exp(\eta T^\alpha)}y_i(T) - y_j(T)\right) =
$$

$$
= \underbrace{\frac{1}{T}\log\left(\frac{A_{ji}}{\xi + \exp(\eta T^\alpha)}y_i(T)\right)}_{\xrightarrow[T\to\infty]{} g_i^*} + \underbrace{\frac{1}{T}\log\left(1 - \frac{y_j(T)(\xi + \exp(\eta T^\alpha))}{y_i(T)A_{ji}}\right)}_{\xrightarrow[T\to\infty]{} 0}.
$$

Finally, applying the change of variables $t' := t - T + 1$ we obtain the expression

$$
\frac{1}{T}\int_{T-1}^{T}\log\left(\frac{A_{ji}}{\xi + \exp(\eta t^\alpha)}y_i(t) - y_j(t)\right)dt
$$

$$
= \int_0^1 \frac{1}{T}\log\left(\frac{A_{ji}}{\xi + \exp(\eta(t' + T - 1)^\alpha)}y_i(t' + T - 1) - y_j(t' + T - 1)\right)dt',
$$

to which we can apply the dominated convergence theorem [65] as $T \to \infty$ thereby completing the justification of the passage to the limit.

References

[1] M. A. Nowak, Five rules for the evolution of cooperation, Science 314 (2006) 1560–1563.

[2] Z. Utkovski, V. Stojkoski, L. Basnarkov, L. Kocarev, Promoting cooperation by preventing exploitation: The role of network structure, Physical Review E 96 (2017) 022315.

[3] V. Stojkoski, Z. Utkovski, L. Basnarkov, L. Kocarev, Cooperation dynamics of generalized reciprocity in state-based social dilemmas, Physical Review E 97 (2018) 052305.

[4] D. Meder, F. Rabe, T. Morville, K. H. Madsen, M. T. Koudahl, R. J. Dolan, H. R. Siebner, O. J. Hulme, Ergodicity-breaking reveals time optimal economic behavior in humans, arXiv preprint arXiv:1906.04652 (2019).

[5] B.-E. Sæther, S. Engen, The concept of fitness in fluctuating environments, Trends in ecology & evolution 30 (2015) 273–281.

[6] A. Antonioni, M. Tomassini, Network fluctuations hinder cooperation in evolutionary games, PloS one 6 (2011).

[7] D. Melamed, B. Simpson, Strong ties promote the evolution of cooperation in dynamic networks, Social networks 45 (2016) 32–44.

[8] M. G. Zimmermann, V. M. Eguíluz, M. San Miguel, Coevolution of dynamical states and interactions in dynamic networks, Physical Review E 69 (2004) 065102.

[9] J. J. Jordan, D. G. Rand, S. Arbesman, J. H. Fowler, N. A. Christakis, Contagion of cooperation in static and fluid social networks, PloS one 8 (2013).

[10] H. Behar, N. Brenner, Y. Louzoun, Coexistence of productive and non-productive populations by fluctuation-driven spatio-temporal patterns, Theoretical population biology 96 (2014) 20–29.

[11] B. Houchmandzadeh, Fluctuation driven fixation of cooperative behavior, Biosystems 127 (2015) 60–66.

[12] B. Houchmandzadeh, M. Vallade, Selection for altruism through random drift in variable size populations, BMC evolutionary biology 12 (2012) 61.

[13] W. Huang, C. Hauert, A. Traulsen, Stochastic game dynamics under demographic fluctuations, Proceedings of the National Academy of Sciences 112 (2015) 9064–9069.

[14] C. S. Gokhale, C. Hauert, Eco-evolutionary dynamics of social dilemmas, Theoretical population biology 111 (2016) 28–42.

[15] G. W. Constable, T. Rogers, A. J. McKane, C. E. Tarnita, Demographic noise can reverse the direction of deterministic selection, Proceedings of the National Academy of Sciences 113 (2016) E4745–E4754.

[16] F. Stollmeier, J. Nagler, Unfair and anomalous evolutionary dynamics from fluctuating payoffs, Physical Review Letters 120 (2018) 058101.

[17] O. Peters, A. Adamou, The evolutionary advantage of cooperation, arXiv preprint arXiv:1506.03414 (2015).

[18] G. Yaari, S. Solomon, Cooperation evolution in random multiplicative environments, The European Physical Journal B 73 (2010) 625–632.

[19] T. Liebmann, S. Kassberger, M. Hellmich, Sharing and growth in general random multiplicative environments, European Journal of Operational Research 258 (2017) 193–206.

[20] O. Peters, M. Gell-Mann, Evaluating gambles using dynamics, Chaos: An Interdisciplinary Journal of Nonlinear Science 26 (2016) 023103.

[21] O. Peters, The ergodicity problem in economics, Nature Physics 15 (2019) 1216–1221.

[22] F. Guvenen, Learning your earning: Are labor income shocks really very persistent?, American Economic Review 97 (2007) 687–712.

[23] X. Gabaix, J.-M. Lasry, P.-L. Lions, B. Moll, The dynamics of inequality, Econometrica 84 (2016) 2071–2111.

[24] A. Forsman, On the role of sex differences for evolution in heterogeneous and changing fitness landscapes: insights from pygmy grasshoppers, Philosophical Transactions of the Royal Society B: Biological Sciences 373 (2018) 20170429.

[25] S. N. Dorogovtsev, A. V. Goltsev, J. F. Mendes, Critical phenomena in complex networks, Rev. Mod. Phys. 80 (2008) 1275.

[26] V. Stojkoski, Z. Utkovski, L. Basnarkov, L. Kocarev, Cooperation dynamics in networked geometric brownian motion, Phys. Rev. E 99 (2019) 062312.

[27] B. Derrida, Random-energy model: Limit of a family of disordered models, Physical Review Letters 45 (1980) 79.

[28] T. Gueudré, A. Dobrinevski, J.-P. Bouchaud, Explore or exploit? a generic model and an exactly solvable case, Physical review letters 112 (2014) 050602.

[29] M. Taborsky, J. G. Frommen, C. Riehl, Correlated pay-offs are key to cooperation, Philosophical Transactions of the Royal Society B: Biological Sciences 371 (2016) 20150084.

[30] J. Lehman, K. O. Stanley, Exploiting open-endedness to solve problems through the search for novelty, in: Proceedings of the Eleventh International Conference on Artificial Life (Alife XI), MIT Press, 2008.

[31] M. Roper, M. J. Dayel, R. E. Pepper, M. Koehl, Cooperatively generated stresslet flows supply fresh fluid to multicellular choanoflagellate colonies, Physical Review Letters 110 (2013) 228104.

[32] A. Sapir, P. Aghion, G. Bertola, M. Hellwig, J. Pisani-Ferry, D. Rosati, J. Viñals, H. Wallace, M. Buti, M. Nava, et al., An agenda for a growing Europe: The Sapir report, OUP Oxford, 2004.

[33] M. Perc, J. Gómez-Gardeñes, A. Szolnoki, L. M. Floría, Y. Moreno, Evolutionary dynamics of group interactions on structured populations: a review, Journal of the royal society interface 10 (2013) 20120997.

[34] J. Gómez-Gardeñes, D. Vilone, A. Sánchez, Disentangling social and group hetero-geneities: Public goods games on complex networks, EPL (Europhysics Letters) 95 (2011) 68003.

[35] J. Peña, Y. Rochat, Bipartite graphs as models of population structures in evolutionary multiplayer games, PloS one 7 (2012) e44514.

[36] M. Tomassini, A. Antonioni, Public goods games on coevolving social network models, FrP 8 (2020) 58.

[37] F. C. Santos, M. D. Santos, J. M. Pacheco, Social diversity promotes the emergence of cooperation in public goods games, Nature 454 (2008) 213.

[38] J.-P. Bouchaud, M. Mézard, Wealth condensation in a simple model of economy, Physica A: Statistical Mechanics and its Applications 282 (2000) 536–545.

[39] J.-P. Bouchaud, On growth-optimal tax rates and the issue of wealth inequalities, Journal of Statistical Mechanics: Theory and Experiment 2015 (2015) P11011.

[40] M. Perc, J. J. Jordan, D. G. Rand, Z. Wang, S. Boccaletti, A. Szolnoki, Statistical physics of human cooperation, Physics Reports 687 (2017) 1–51.

[41] S. Boyd, J. Dattorro, et al., Alternating projections, EE392o, Stanford University (2003).

[42] B. Bollobás, Modern graph theory, volume 184, Springer Science & Business Media, 2013.

[43] P. Erdos, A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hung. Acad. Sci 5 (1960) 17–60.

[44] D. J. Watts, S. H. Strogatz, Collective dynamics of 'small-world'networks, Nature 393 (1998) 440.

[45] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.

[46] R. Axelrod, W. D. Hamilton, The evolution of cooperation, Science 211 (1981) 1390–1396.

[47] R. Alexander, The biology of moral systems, Routledge, 2017.

[48] C. Rutte, M. Taborsky, Generalized reciprocity in rats, PLoS Biol 5 (2007) e196.

[49] K. L. Leimgruber, A. F. Ward, J. Widness, M. I. Norton, K. R. Olson, K. Gray, L. R. Santos, Give what you get: capuchin monkeys (cebus apella) and 4-year-old children pay forward positive and negative outcomes to conspecifics, PloS one 9 (2014) e87035.

[50] N. Gfrerer, M. Taborsky, Working dogs cooperate among one another by generalised reciprocity, Scientific reports 7 (2017) 1–6.

[51] M. Y. Bartlett, D. DeSteno, Gratitude and prosocial behavior: Helping when it costs you, Psychol. Sci. 17 (2006) 319–325.

[52] L. Stanca, Measuring indirect reciprocity: Whose back do we scratch?, J. Econ. Psychol. 30 (2009) 190–202.

[53] T. Pfeiffer, C. Rutte, T. Killingback, M. Taborsky, S. Bonhoeffer, Evolution of cooperation by generalized reciprocity, Proceedings of the Royal Society of London B: Biological Sciences 272 (2005) 1115–1120.

[54] D. J. Rankin, M. Taborsky, Assortment and the evolution of generalized reciprocity, Evolution 63 (2009) 1913–1922.

[55] G. S. van Doorn, M. Taborsky, The evolution of generalized reciprocity on social interaction networks, Evolution 66 (2012) 651–664.

[56] M. A. Nowak, S. Roch, Upstream reciprocity and the evolution of gratitude, Proceedings of the Royal Society B: Biological Sciences 274 (2006) 605–610.

[57] T. Ito, R. Suzuki, T. Arita, Evolution of four forms of reciprocity in the prisoner's dilemma game, Artificial Life and Robotics 24 (2019) 140–146.

[58] V. Stojkoski, Z. Utkovski, E. André, L. Kocarev, The role of multiplex network structure in cooperation through generalized reciprocity, Physica A: Statistical Mechanics and its Applications 531 (2019) 121805.

[59] J. Moran, A. Fosset, D. Luzzati, J.-P. Bouchaud, M. Benzaquen, By force of habit: Self-trapping in a dynamical utility landscape, Chaos: An Interdisciplinary Journal of Nonlinear Science 30 (2020) 053123.

[60] K. Hammerschmidt, C. J. Rose, B. Kerr, P. B. Rainey, Life cycles, fitness decoupling and the evolution of multicellularity, Nature 515 (2014) 75.

[61] P. B. Rainey, B. Kerr, Cheats as first propagules: a new hypothesis for the evolution of individuality during the transition from single cells to multicellularity, Bioessays 32 (2010) 872–880.

[62] Y. Pichugin, C. S. Gokhale, J. Garcia, A. Traulsen, P. B. Rainey, Modes of migration and multilevel selection in evolutionary multiplayer games, Journal of theoretical biology 387 (2015) 144–153.

[63] R. Khasminskii, Stochastic stability of differential equations, volume 66, Stochastic Modelling and Applied Probability Book Series, Springer Science & Business Media, 2011.

[64] N. Touzi, Optimal stochastic control, stochastic target problems, and backward SDE, volume 29, Springer Science & Business Media, 2012.

[65] R. Durrett, Probability: theory and examples, volume 49, Cambridge university press, 2019.