Evolution of Cooperation in Arbitrary Complex Networks

Bijan Ranjbar-Sahraei[1], Haitham Bou Ammar[2], Daan Bloembergen[1], Karl Tuyls[3] and Gerhard Weiss[1]

[1]Maastricht University, The Netherlands
{b.ranjbarsahraei, daan.bloembergen, gerhard.weiss}@maastrichtuniversity.nl
[2]University of Pennsylvania, PA, United States
haithamb@seas.upenn.edu
[3]University of Liverpool, United Kingdom
k.tuyls@liverpool.ac.uk

ABSTRACT

This paper proposes a new model, based on the theory of nonlinear dynamical systems, to study the evolution of cooperation in arbitrary complex networks. We consider a large population of agents placed on some arbitrary network, interacting with their neighbors while trying to optimize their fitness over time. Each agent's strategy is continuous in nature, ranging from purely cooperative to purely defective behavior, where cooperation is costly but leads to shared benefits among the agent's neighbors. This induces a dilemma between social welfare and individual rationality. We show in simulation that our model clarifies why cooperation prevails in various regular and scale-free networks. Moreover we observe a relation between the network size and connectivity on the one hand, and the resulting level of cooperation in equilibrium on the other hand. These empirical findings are accompanied by an analytical study of stability of arbitrary networks. Furthermore, in the special case of regular networks we prove convergence to a specific equilibrium where all agents adopt the same strategy. Studying under which scenarios cooperation can prevail in structured societies of self-interested individuals has been a topic of interest in the past two decades. However, related work has been mainly restricted to either analytically studying a specific network structure, or empirically comparing different network structures. To the best of our knowledge we are the first to propose a dynamical model that can be used to analytically study arbitrary complex networks.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Multiagent systems

General Terms

Theory, Experimentation

Keywords

Evolution of Cooperation, Repeated Games on Graphs

1. INTRODUCTION

Many real-world systems can be described as networks in which nodes represent individual decision makers and edges represent the interactions that occur between them. Examples are wide-spread, ranging from neural networks in the human brain to communication networks of hubs and routers that make up the World Wide Web. Next to these tangible networks, the past decade has seen a strong increase in interest in the study of social networks, mainly from a sociological or economic perspective examining how ideas (or likewise new technologies) can spread through social communities [6, 14]. This work focusses on the latter, analyzing the evolution of cooperation in societies of self-interested decision-makers.

A central question in this line of research is how cooperation can be sustained in a population despite the fact that cooperative behavior is costly. Consequently, the paradoxical emergence of mutually beneficial interactions among selfish individuals has attracted a lot of interest from the Game Theory community in past decades (e.g., [4, 13, 17, 24]). Most notably, research has focused on the widely used example of the Prisoner's Dilemma in order to find out under what conditions rational decision makers can be incentivized to cooperate so as to achieve a better joint outcome [1].

The standard Prisoner's Dilemma model is limited by its binary nature: only two discrete actions are available - pure cooperation and pure defection - whereas in many real settings cooperation can be better thought of as being a continuous trait [17]. Moreover, this simplification may hide many interesting dynamics of cooperative behavior, preventing a full analysis of such settings. Therefore, this work focusses on a continuous version of the Prisoner's Dilemma in which the strategy space depicts the individual's level of cooperation, ranging from fully cooperative to fully defective.

We consider a large population of individuals (or agents) placed on an arbitrary complex network, interacting with their neighbors according to the continuous Prisoner's Dilemma while trying to optimise their fitness over time. The agents update their strategy by adopting, with some probability, the strategy of one of their neighbors depending on their difference in fitness [11]. We propose a continuous dynamical model that predicts the level of cooperation of each of the nodes in this network over time, allowing to study convergence and stability in arbitrary networks. The deterministic characteristics of this model make it computationally less complex than traditional binary choice models, allowing to study a broader range of scenarios numerically. We employ our model to analyze the evolution of cooperation in a variety of regular and scale-free networks. The analysis shows that cooperation is promoted for a range of network settings, and indicates a relation between network size and connectivity on the one hand, and the resulting level of cooperation in equilibrium on the other. Moreover, we analytically study the stability of the model for arbitrary networks. For the special case of regular networks, we additionally prove convergence to

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*

a specific equilibrium where all agents adopt the same strategy.

The remainder of the paper is organized as follows. First, Section 2 positions this work within the large body of related research. Preliminaries on networks, game theory and system modeling and control are provided in Section 3, laying the foundation for the proposed model, which is introduced in Section 4. Section 5 shows numerical experiments highlighting the transient and long-term behavior of the model. Section 6 presents stability analysis, and Section 7 concludes.

2. RELATED WORK

Many studies have analyzed how behavior evolves in a society of individuals. On a macro level researchers have studied population dynamics governed by evolutionary rules such as survival of the fittest; a well known example of such a dynamical system are the replicator dynamics, a system of differential equations describing how different strategic types evolve in a large population under evolutionary pressure [12, 19]. On a micro level studies have looked at the behavioral change of single individuals as a result of their interactions with the environment or with others; notably this line of research includes learning approaches such as reinforcement learning [27] and recently also multi-agent reinforcement learning [5, 29]. These two branches share an interesting common ground: the replicator dynamics have been related to infinitesimal time versions of several reinforcement learning algorithms such as learning automata and Q-learning [3, 28].

However, the study of population dynamics usually assumes well-mixed populations where all pairs of individuals have an equal chance of interacting, omitting the possibility of spacial structure. Recently, researchers have started to investigate situations in which such spacial structure is present, where the population is represented by a network indicating with whom each individual can interact. Their aim is usually to establish structural network criteria under which a 'beneficial' outcome is reached for the population as a whole, e.g. which criteria lead to a cooperative outcome in the Prisoner's Dilemma when played on a network. Nowak and May [20] were the first to study the Prisoner's Dilemma in a population of myopic individuals placed on a grid, and interacting only with their eight neighbors. They found that under an imitate-best-neighbor rule, cooperators and defectors can survive simultaneously in the network. In a similar fashion, Santos and Pacheco [24] investigate imitation dynamics on scale-free networks and show that cooperation becomes the dominant strategy in such networks. Ohtsuki *et al.* [21] look at various network topologies and find a link between the cost-benefit ratio of cooperation and the average node degree for certain imitation-based update rules. Hofmann *et al.* [13] simulate various update rules in different network topologies and find that the evolution of cooperation is highly dependent on the combination of update mechanism and network topology. Cooperation can also be promoted using some incentivising structure in which defection is punishable [4, 26], or in which players can choose beforehand to commit to cooperation for some given cost [10]. Both incentives increase the willingness to cooperate in scenarios where defection would be individually rational otherwise. Allowing individuals to choose with whom to interact may similarly sustain cooperation, e.g. by giving individuals the possibility to break ties with 'bad' neighbors and replacing them with a random new connection. Zimmermann and Eguíliz [32] show how such a mechanism may promote cooperation, albeit sensitive to perturbations.

Finally, attempts have been made to bridge these two views, uniting the well-mixed population model of the replicator dynamics with networked interaction structures. Kearns and Suri [16] extend evolutionary game theory to networks and show that evolutionarily stable strategies (ESS) are preserved assuming a random network and adversarial mutant set or vice versa. Ohtsuki *et al.* [22] show that moving from a well-mixed population (or complete network) to a regular network keeps the structure of the replicator dynamics intact, only transforming the payoff function to account for local competition. They observe the coexistence of cooperators and defectors under such settings. This last line of research is most closely related to the work presented in this paper, however our model is more general and not limited to specific network structures. Moreover, the proposed model captures the continuous nature of cooperation in real world settings, in contrast to the binary choice between pure cooperation and defection assumed in most related work presented here. This allows us to study the evolution of cooperation in a broader setting.

3. PRELIMINARIES

This section introduces elementary background on Networks, Game Theory, and System Modeling and Control that forms the foundation of the work presented in this paper. For an in-depth discussion of these fields the interested reader is referred to [14], [9], and [18], respectively.

3.1 Networks

Networks, in the most general sense, can be seen as patterns of interconnections between sets of entities [6]. These entities make up the nodes in the network, whereas the edges represent how those entities interact or how they are related. Formally, a network can be represented by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{W})$ consisting of a non-empty set of nodes (or vertices) $\mathcal{V} = \{v_1, \ldots, v_N\}$ and an $N \times N$ adjacency matrix $\mathcal{W} = [w_{ij}]$ where non-zero entries w_{ij} indicate the (possibly weighted) connection from v_i to v_j. If $\mathcal{W}$ is symmetrical the graph is said to be undirected, indicating that the connection from v_i to v_j is equal to the connection from v_j to v_i. In social networks, for example, one might argue that friendship is usually mutual and hence undirected; this is the approach followed in this paper as the interaction is a game, naturally involving both parties. In general however the relationship between nodes can be asymmetrical resulting in a directed graph. The neighborhood $\mathcal{N}$ of a node v_i is defined as the set of vertices it is connected to, i.e. $\mathcal{N}(v_i) = \cup_j v_j : w_{ij} > 0$. The node's degree deg$(v_i)$ is given by the cardinality of its neighborhood.

Several types of networks can be distinguished based on their structural properties. In a regular network all nodes have exactly the same degree, e.g. a ring is a regular network of degree 2. In the special case of a fully connected network the degree equals $n - 1$, meaning that all nodes are connected to all other nodes. In contrast, many large social, technological or biological networks exhibit a heavy-tailed degree distribution following a power law [2]. In these so-called scale-free networks the majority of nodes will have a small degree, while simultaneously there will be relatively many nodes with very large degree. Another model used to describe real-world networks is the small-world model, that exhibits short average path lengths between nodes, and high clustering [30]. Finally, random networks are defined by a probability p with which two arbitrary nodes are connected [7].

3.2 Game Theory

Game theory models strategic interactions in the form of games. Each player has a set of actions, and a preference over the joint action space that is captured in the received payoffs. For two-player games, the payoffs can be represented by a bi-matrix (A, B), that gives the payoff for the row player in A, and the column player in

$$\begin{pmatrix} a_{11}, b_{11} & a_{12}, b_{12} \\ a_{21}, b_{21} & a_{22}, b_{22} \end{pmatrix} \qquad \begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} R, R & S, T \\ T, S & P, P \end{pmatrix} \end{array}$$

Figure 1: General payoff bi-matrix (A, B) for two-player two-action games (left), and the Prisoner's Dilemma (right).

B, see Figure 1. In this example, the row player chooses one of the two rows, the column player simultaneously chooses one of the columns, and the outcome of this joint action determines the payoff to both. The goal for each player is to come up with a strategy (a probability distribution over its actions) that maximizes his expected payoff in the game. A strategy that maximizes the payoff given fixed strategies for all opponents is called a best response to those strategies.

The players are thought of as rational, in the sense that each player purely tries to maximize his own payoff, and assumes the others are doing likewise. Under this assumption, the Nash equilibrium concept can be used to study what players will reasonably choose to do. A set of strategies forms a Nash equilibrium if no single player can do better by unilaterally switching to a different strategy [9]. In other words, each strategy in a Nash equilibrium is a best response against all other strategies in that equilibrium.

The canonical payoff matrix of the Prisoner's Dilemma is given in Figure 1, where $T > R > P > S$. In this game, jointly both players would be best off cooperating and receiving reward R, however individually both are tempted by the higher payoff T, leaving the other with the sucker payoff S. As both reason like this, they end up in the less favorable state of mutual defection, receiving as punishment $P < R$, hence the dilemma.

In this work the players are nodes in the network, repeatedly playing a game with their neighbors. The players have no knowledge of the underlying game, however this repeated interaction allows for adaptation, i.e. to learn a better strategy over time based on the payoff received. The game used in this paper is more general than the games presented in this section, in that the players can have a continuous strategy defining their level of cooperation, and payoffs are calculated accordingly (see Section 4). The dilemma remains, however, which is the main focus of the analysis in Sections 5 and 6.

3.3 System Modeling and Control

First, models used for representing dynamical systems are introduced. Both control and stability, being the basis of the analysis performed in this paper, are then detailed.

3.3.1 Modeling Dynamical Systems

A model can be regarded as an accurate mathematical representation of the (nonlinear) dynamics of a system. Essentially, the goal is the discovery of (nonlinear) differential equations describing the transient behavior of some state variables in a system. Typically, state representations are collected in a state vector $\mathbf{x} = [x_1, x_2, \ldots, x_N]^{\mathrm{T}}$ and control variables (i.e., actions applied to affect the state vector) are collected in a vector $\mathbf{u} = [u_1, u_2, \ldots, u_q]^{\mathrm{T}}$ where x_i and u_i denote the i^{th} state and input respectively. A linear and time invariant system (LTI) can thus be represented by

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$$

where $\mathbf{A}$ and $\mathbf{B}$ correspond to the dynamic and control matrices, respectively.

When the system dynamics are nonlinear and/or time varying, as is the case in this paper, the state space model has to be extended to a more general form

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_N \end{bmatrix} = \begin{bmatrix} f_1(t; x_1, \ldots, x_N, \mathbf{u}) \\ f_2(t; x_1, \ldots, x_N, \mathbf{u}) \\ \vdots \\ f_N(t; x_1, \ldots, x_N, \mathbf{u}) \end{bmatrix}$$

where the change in the state variables is a nonlinear mapping of the state variables and the control action. Moreover, each state variable is governed by its own dynamics. Compactly this can be written in matrix form as $\dot{\mathbf{x}} = \mathbf{f}(t; \mathbf{x}, \mathbf{u})$.

3.3.2 Stability & Control of Dynamical Systems

One of the main goals in control theory is the manipulation of the system's inputs to follow a reference over time. In other words, this manipulation feeds back the difference between the state variable $\mathbf{x}$ and the reference point $\mathbf{x}_{\mathrm{ref}}$ at any instance in time. Such a rule, where $\mathbf{u} = l(\mathbf{x}, \mathbf{x}_{\mathrm{ref}})$, is called a feedback controller. Controller design is a wide-spread field and its discussion is beyond the scope of this paper. Interested readers are referred to [18]. Here, the main interest is in stability and convergence analysis of dynamical systems as this work builds on such results to study certain aspects of the proposed model (see Section 6). Stability is studied in the vicinity of equilibrium points (i.e., points where $\dot{\mathbf{x}} = 0$). In the following, $\mathcal{B}(\bar{\mathbf{x}}, \epsilon)$ denotes an open ball centered at $\bar{\mathbf{x}}$ with a radius ϵ, that is the set $\{\mathbf{x} \in \mathbb{R}^d : ||\mathbf{x} - \bar{\mathbf{x}}|| < \epsilon\}$, where $|| \cdot ||$ represents the L_2-norm. The following definition of stability can be stated:

DEFINITION 1 (STABILITY). *An equilibrium point $\boldsymbol{x}_e$ of a nonlinear system is said to be stable, if for all $\epsilon > 0$ there exists a $\delta > 0$ such that:*

$$\bar{\boldsymbol{x}} \in \mathcal{B}(\boldsymbol{x}_e, \delta) \implies \boldsymbol{f}(t; \bar{\boldsymbol{x}}, \boldsymbol{0}) \in \mathcal{B}(\boldsymbol{x}_e, \epsilon) \text{ for all } t \geq 0.$$

Stability can be studied under different contexts, including Lyapunov [23], Gershgorin [8, 15], and Jacobian analysis [18]. Defining and analyzing Lyapunov functions for general networks is complex. Therefore, this paper adopts methods from Gershgorin and Jacobian analysis for the stability study. Next, each of these are briefly detailed.

Gershgorin Theorem.

One of the most important ingredients in stability analysis is the sign of the dynamic matrix eigenvalues. In general, these can not be computed in a closed form and numerical analysis is required. Gershgorin disks, however, provide a bound estimate for the eigenvalues of strictly diagonally dominant matrices. Gershgorin's theorem, used later in Section 6.1, is presented next without proof.

THEOREM 1 (GERSHGORIN DISKS [8]). *Let $\boldsymbol{A} = [a_{ij}]$ be a dominant $n \times n$ matrix. Then the eigenvalues of $\boldsymbol{A}$ are located in the union of n disks:*

$$\Big\{\lambda : \Big|\lambda - a_{ii}\Big| \leq \sum_{j \neq i} |a_{ij}|\Big\}, \;\; i, j \in \{1, 2, \ldots, n\}$$

Jacobian Analysis.

Typically, nonlinear dynamical systems exhibit multiple equilibria (i.e. points where $\dot{\mathbf{x}} = 0$). It is important in such cases to study stability in the neighborhood of an equilibrium point. Typically, the Jacobian matrix $\mathbf{J}$ of the system is first computed in its general form according to

$$\mathbf{J} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_N}{\partial x_1} & \frac{\partial f_N}{\partial x_2} & \cdots & \frac{\partial f_N}{\partial x_N} \end{bmatrix}$$

$\mathbf{J}$ is then evaluated at the equilibria and its eigenvalues are assessed for stability. Namely, if the real part of the eigenvalues of $\mathbf{J}$ were negative then the system is stable, else it is said to be unstable.

4. DYNAMICAL MODELING

This section details the continuous dynamical model proposed in this work. Firstly, the model of a 2-player Continuous Action Iterated Prisoner's Dilemma (CAIPD) is derived. This model is then generalized to the N-player case.

4.1 2-Player CAIPD

In the 2-player continuous action iterated prisoner's dilemma (CAIPD) each player can choose a level of cooperation from a continuous set of strategies.

Let $x_i \in [0,1]$ denote the strategy of the i^{th} player with $i \in \{1,2\}$ representing each player. Here, $x_i = 0$ corresponds to full defection, while $x_i = 1$ represents full cooperation. A player pays a cost cx_i while the opponent receives a benefit bx_i, with $b > c$. It is clear that a defector (i.e., $x_i = 0$) pays no cost and distributes no benefits. The fitness of player i, $F(x_i)$, can be thus defined as:

$$F(x_i) \quad = \quad -cx_i + bx_j \tag{1}$$

Using (1), the difference between the fitnesses of two players can be derived as

$$\begin{aligned} \Delta F_{ji} &= F(x_j) - F(x_i) \\ &= -c(x_j - x_i) - b(x_j - x_i) \\ &= (-c - b)(x_j - x_i) \end{aligned}$$

Following the imitation dynamics [11], where each player switches to a neighboring strategy with a certain probability, the following strategy evolution law is introduced:

$$x_i(k+1) \quad = \quad (1 - p_{ij})x_i(k) + p_{ij}x_j(k), \tag{2}$$

where k represents the iteration number and $p_{ij} = \text{sig}(\beta\Delta F_{ji})$, with $\text{sig}(\beta\Delta F_{ji}) = 1/(1 + \exp(-\beta\Delta F_{ji}))$ and $\beta > 0$. In words, Equation 2 states that in iteration k a player switches to a neighboring strategy with a probability p_{ij}.

The change $\Delta x_i(k) = x_i(k+1) - x_i(k)$ in strategies between two iterations $k+1$ and k can be rewritten as:

$$\begin{aligned} \Delta x_i(k) &= x_i(k+1) - x_i(k) \\ &= p_{ij}\Big(x_j(k) - x_i(k)\Big) \end{aligned}$$

Assuming infinitesimal changes and using Taylor expansion, the strategy adaptation law of player i can be written as

$$\dot{x}_i(t) \quad = \quad p_{ij}(x_j(t) - x_i(t)), \tag{3}$$

In essence, the adaptation law of Equation (3) shows that for high values of p_{ij}, the i^{th} player switches its strategy to the opponent's strategy, while for low values of p_{ij} it keeps its own strategy.

4.2 N-Player CAIPD

Having introduced the 2-player CAIPD, this section details the more general N-player case. The N-player CAIPD is defined for a group of N players on a weighted graph (Section 3.1) $\mathbb{G} = (\mathcal{V}, \mathcal{W})$ where $\mathcal{V} = \{v_1, v_2, \ldots, v_N\}$ represents the set of nodes (i.e., each player is represented by a node), and $\mathcal{W} = [w_{ij}]$ denotes the symmetric weighted adjacency matrix, where $w_{ij} \in \{0,1\}$ is a binary variable describing the connection between players i and j, $\forall(i,j) \in \{1,2,\ldots,N\} \times \{1,2,\ldots,N\}$. Further, w_{ii} is assumed to be zero for all $i \in \{1,2,\ldots,N\}$.

Let $x_i \in [0,1]$ denote the cooperation level of node v_i. A network with value $\mathbf{x}$ and topology $\mathbb{G}$ is defined as $\mathbb{G}_{\mathbf{x}} = (\mathbb{G}, \mathbf{x})$ with $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$. Supposing that each node of the network $\mathbb{G}_{\mathbf{x}}$ is a dynamic player with

$$\dot{x}_i \quad = \quad h_i(\mathbf{x}), \tag{4}$$

the network $\mathbb{G}_{\mathbf{x}}$ can be regarded as a dynamical system in which the value $\mathbf{x}$ evolves according to the network dynamics $\dot{\mathbf{x}} = \mathbf{H}(\mathbf{x})$.

Having a general form for the network dynamics, the next step is to determine $h_i(\cdot)$ in Equation 4 for all $i \in \{1,2,\ldots,N\}$. To determine $h_i(\cdot)$ in its full form, firstly, the i^{th} node/player fitness needs to be derived. Generalizing the 2-player CAIPD, the following can be computed:

$$F(x_i) \quad = \quad -\text{deg}(v_i)cx_i + b\sum_{j=1}^{N} w_{ij}x_j \tag{5}$$

where $\text{deg}(v_i)$ is the degree of node v_i. In Equation 5, the i^{th} player pays a cost of cx_i for each of its neighbors j (i.e., $-\text{deg}(v_i)cx_i$) and receives a benefit of bx_j for all its neighbors j (i.e., $b\sum_{j=1}^{N} w_{ij}x_j$, with $w_{ij} \in 0,1$ indicating whether i and j are connected). Therefore, the difference between the i^{th} and j^{th} player fitnesses can be written as:

$$\begin{aligned} \Delta F_{ji} &= F(x_j) - F(x_i) \\ &= c\left(\text{deg}(v_i)x_i - \text{deg}(v_j)x_j\right) + b\left(\sum_{k=1}^{N}(w_{jk} - w_{ik})x_k\right) \end{aligned}$$

Given that the probability of strategy adaptation $p_{ij} = \text{sig}(\beta\Delta F_{ji})$, similar to (2) the evolution law for i^{th} player in the network can be derived as

$$x_i(k+1) = \frac{1}{\text{deg}(v_i)}\left[\sum_{j=1}^{N}(1 - p_{ij}w_{ij})x_i(k) + p_{ij}w_{ij}x_j(k)\right] \tag{6}$$

The difference equations in (6) can again be converted to differential equations by assuming infinitesimal k and applying Taylor expansion, which leads to

$$\dot{x}_i(t) \quad = \quad \frac{1}{\text{deg}(v_i)}\left[\sum_{j=1}^{N} p_{ij}w_{ij}\Big(x_j(t) - x_i(t)\Big)\right]$$

Therefore, $h_i(\cdot)$, introduced in (4), is:

$$h_i(\mathbf{x}) \quad = \quad \frac{1}{\text{deg}(v_i)}\left[\sum_{j=1}^{N} p_{ij}w_{ij}\Big(x_j(t) - x_i(t)\Big)\right] \tag{7}$$

Applying the same techniques to the overall network, the following *network* dynamics can be derived:

$$\dot{\mathbf{x}} = \mathbf{H}(\mathbf{x}) = \mathbf{DA}(\mathbf{x})\mathbf{x} \tag{8}$$

where

$$
\mathbf{A}(\mathbf{x}) = \begin{bmatrix} -\sum_{j=1}^{N} p_{1j}w_{1j} & w_{12}p_{12} & \cdots & w_{1N}p_{1N} \\ w_{21}p_{21} & -\sum_{j=1}^{N} p_{2j}w_{2j} & \cdots & w_{2N}p_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1}p_{N1} & w_{N2}p_{N2} & \cdots & -\sum_{j=1}^{N} w_{Nj}p_{N,j} \end{bmatrix}
$$

with $p_{ij} = 1 - p_{ji}$, and $\mathbf{D} = \text{diag}(1/\deg(v_i))$ for $i = \{1, 2, \ldots, N\}$.

Having derived the dynamical model capable of representing arbitrary complex networks, next a detailed experimental analysis on various networks is performed.

5. EXPERIMENTS

The dynamical behavior of the proposed model is studied on two types of networks. Scale-free and regular networks constitute the benchmarks on which experiments are carried out. The scale-free networks are generated using the Barabási-Albert model [2], where node addition follows the preferential attachment scheme. Namely, starting from a connected subset of nodes, every additional node is connected to m existing nodes with a probability proportional to the degree of those nodes. In other words, nodes having higher number of neighbors have a higher probability for attaining even more neighbors. The regular networks follow the simple rule that every node in the network has exactly the same degree, i.e. $\deg(v_i) = \deg(v_j), \forall i, j \in \{1, 2, \ldots, N\}$.

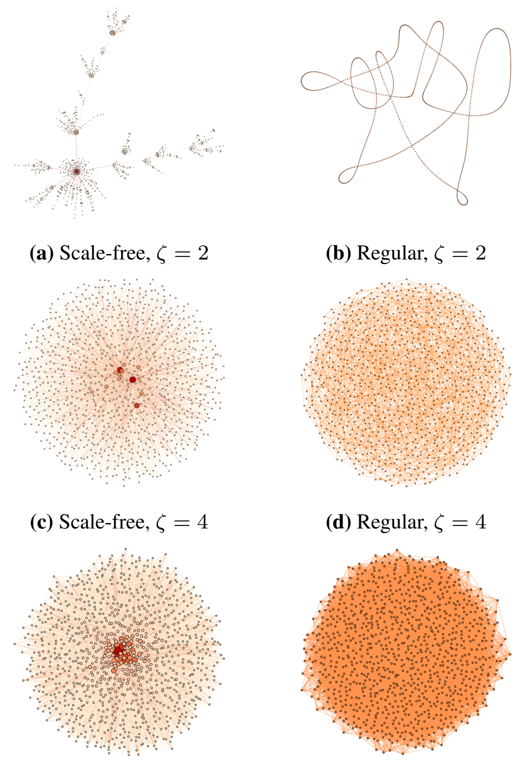
A pictorial illustration of the change in the network's topology as a function of its average node degree ζ is shown in Figure 2. Figures 2(a), 2(c) and 2(e) illustrate this change for scale-free networks following a power law degree distribution. It is clear that as ζ increases so does the network's link density. In regular graphs, however, all nodes have equal degrees leading to an evenly distributed link density with an increase in ζ, as shown in Figures 2(b), 2(d) and 2(f).

Two sets of experiments are conducted. Firstly, Section 5.1 studies the effect of the network's connectivity/topology on cooperation promotion and convergence. The transient behavior of scale-free networks is analyzed. Results show that: (1) the proposed model promotes cooperation, and (2) convergence is achieved under different parametric settings. Secondly, in Section 5.2 the interest is in the long-term behavior on a broader spectrum of networks. Both scale-free and regular networks are considered. Results again confirm convergence and cooperation promotion in both types of networks.

In all simulations, each individual is initiated to either pure cooperation (i.e., $x_i = 1$) or pure defection (i.e., $x_i = 0$). The cost of cooperation is set to $c = 1$ and the benefit, b, is set to 4. Finally, $\beta = 1$ is used for the sigmoidal function defining the probability of strategy adaptation.

5.1 Transient Behavior

The transient behavior of scale-free networks is investigated on different network topologies, initial conditions, and sizes. It is essential to note that the proposed model incurs factors of magnitude of lower computational complexity compared to the traditional techniques available elsewhere, see Section 2. This allows to simulate the model exactly. In other words, the behavioral results, shown in Figure 3, are attainable through only one run of the model.



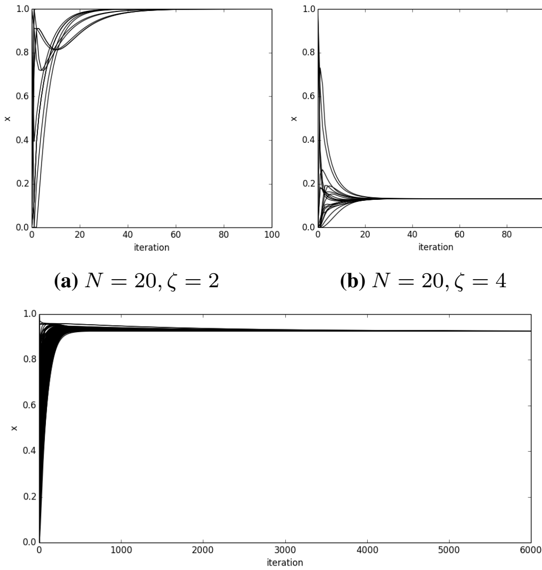
(a) Scale-free, $\zeta = 2$ **(b)** Regular, $\zeta = 2$ **(c)** Scale-free, $\zeta = 4$ **(d)** Regular, $\zeta = 4$ **(e)** Scale-free, $\zeta = 20$ **(f)** Regular, $\zeta = 20$

Figure 2: Scale-free and Regular graphs with various average degrees ζ.

State variable trajectories (i.e, x_i, for $i \in \{1, \ldots, n\}$) are plotted over time for different sizes and average degrees of the network. Firstly, trajectory results on relatively small networks with varying average degree ζ are shown in Figures 3(a) and 3(b). It is clear that trajectories converge to pure cooperation for small average degrees (i.e., $\zeta = 2$). The percentage of cooperators decreases with an increase of ζ, see Figure 3(b). It is interesting to note, however, that contrary to the binary case (Section 2), some percentage of cooperation can still be sustained due to the continuous nature of the model.

The transient behavior is also investigated for larger (i.e., $N = 1000$) scale-free networks, as shown in Figure 3(c), where similar conclusions can be drawn. Finally, the expected cooperation level among all network nodes is investigated for large (i.e., $N = 1000$) scale-free networks with different structural characteristics. For different initial fractions (i.e., 20%, 50% and 80%) of cooperators and defectors that are randomly placed at different positions in the network, the cooperation level is calculated using $E(\mathbf{x}) = \frac{1}{N}\sum_{i=1}^{N} x_i$. Averaged cooperation levels are then collected and plotted in Figures 4(a)-(d) for $\zeta = 2$, 6, 10 and 20, respectively. It is again clear that lower average degrees ζ promote cooperation, whereas higher values lead to defection.

Another interesting aspect of the model is interpretability. Generally speaking, internal node dynamics are hardly interpretable from game theoretic models. Using our proposed model, studying the transient behavior of the network can uncover internal dynamics and insights about the system. For instance, in Figure 3(a) damped oscillations can be seen in the cooperation levels of several individuals. Such analysis might uncover correlations between



(a) $N = 20, \zeta = 2$ **(b)** $N = 20, \zeta = 4$

(c) $N = 1000, \zeta = 6$

Figure 3: Transient behavior of sample Barabási-Albert scale-free graphs with various average degrees.

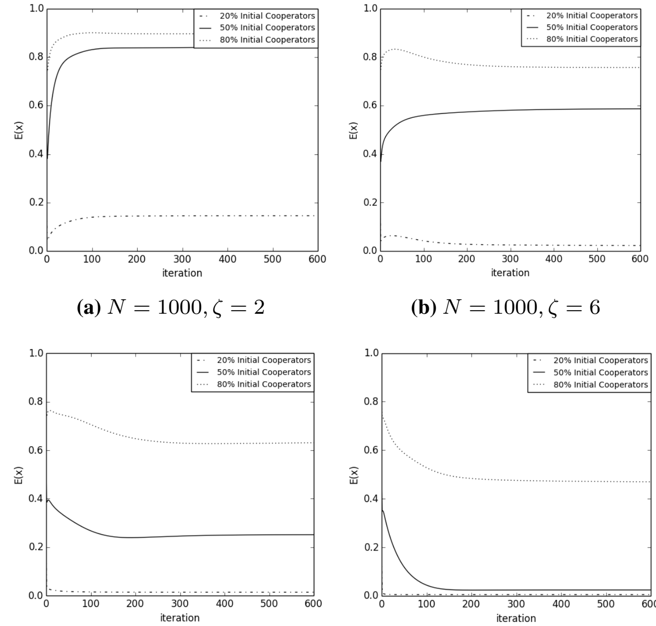
the nodes' dynamical behaviors and their positions on the graph. Furthermore, a link to the discrete case can be made. Intuitively, the continuous model is related to the study of the game theoretic models in expectation. In discrete terms, if the value of cooperation attained from the proposed continuous model is, say, 0.9 (i.e., for all nodes $x_i = 0.9$), e.g., Figure 4(a), then in 90% of the discrete algorithm runs cooperators will dominate while in the other 10% defectors win.

5.2 Long Term Behavior

Clearly from the previous experiments, the proposed model converges. Here, we are interested in analyzing the long-term dynamical behavior. Specifically, the values to which state variables eventually converge are studied in both regular as well as scale-free networks. Networks of different sizes $N \in \{20, 100, 200, \ldots, 800\}$ and average node degrees $\zeta \in \{2, 4, \ldots, 20\}$ are considered.

For each specific setting, 100 random initial populations with equal number of cooperators and defectors are simulated and the average final level of cooperation is computed. Results are reported in Figures 5(a) and 5(b) for scale-free and regular networks, respectively. These results not only show the convergence values, they also indicate that for both network types, highest cooperation promotion is attained for low average degree and large populations. Note that the slightly lower final cooperation level for regular graphs of degree $\zeta = 2$ can be explained by the fact that a connected regular network of degree 2 can only be a cycle. Therefore it misses many of the structural properties of higher degree networks such as clustering, which hinders the evolution of cooperation. Moreover, by comparing Figures 5(a) and 5(b), it can be seen that the highest cooperation level achieved in scale-free networks (i.e., $E(\mathbf{x}) = 0.75$) is larger in value compared to that achieved in regular networks (i.e., $E(\mathbf{x}) = 0.65$). This result is in accordance with the notion that scale-free networks promote cooperation due to their power law degree distribution [24].

Finally, we compare our model with the binary iterated prisoner's dilemma model introduced in [11], with game parameters



(a) $N = 1000, \zeta = 2$ **(b)** $N = 1000, \zeta = 6$

(c) $N = 1000, \zeta = 10$ **(d)** $N = 1000, \zeta = 20$

Figure 4: Average cooperation level in scale-free networks with $N = 1000$ and various average degrees.

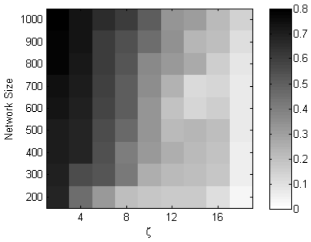
$c = 1, b = 4$ and $\beta = 1$. This model is also similar to the models in [24, 25] except for minor differences in the design of the probability function. Figure 5(c) shows that using this binary model, cooperators have significantly less chance to survive in scale-free networks. In regular networks, cooperation goes extinct in all scenarios. Therefore, it is clear that our proposed dynamical model is better able to explain why cooperation can prevail in complex social networks.

6. STABILITY ANALYSIS

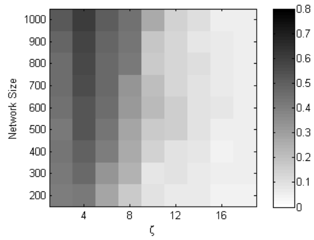
Having performed the above experimental analysis, various properties of the model/network are apparent. For instance, it is clear that the model converges when all players adopt the same strategy. In this section, we amend these experimental results by developing mathematical tools explaining such behavior. Gershgorin disks are used as the first step to analyze the long term stability of any arbitrary network. Due to the time varying nature of the state matrix $\mathbf{DA}(.)$, insights on the required conditions for system stability of the system will be provided. Finally, Taylor linearization and Jacobian analysis are used to prove equilibrium point stability for regular networks.

6.1 Long Term Stability

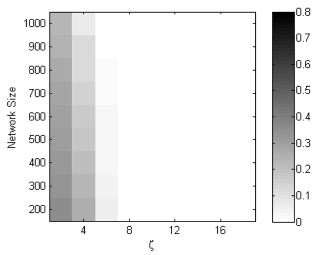
To analyze long term stability of the proposed model we start by studying the eigenvalues of the system. This study builds on the Gershgorin theorem, introduced in Section 3.3, for estimating eigenvalues of a matrix. From (7) and (8) it is clear that the rows of the state matrix sum to zero and therefore, the matrix is diagonally dominant. Theorem 1 can be used to find the ranges of the



(a) Scale-Free Network



(b) Regular Network



(c) Scale-Free network with binary model

Figure 5: Long term expected cooperation level in Regular networks.

eigenvalues as

$$\Big|\lambda - a_{ii}\Big| \leq \sum_{j\neq i} |a_{ij}| \;\; i \in \{1, \dots, N\}$$

$$-\sum_{j\neq i} |a_{ij}| + a_{ii} \leq \lambda \leq \sum_{j\neq i} |a_{ij}| + a_{ii}$$

Therefore, it can be seen that for any row of matrix $\mathbf{DA}(.)$ we have

$$-2\sum_{i\neq j} w_{ij}p_{ij} \leq \lambda \leq 0$$

Knowing that $w_{ij} \in \{0, 1\}$ and $p_{ij} \in [0, 1]$, it can be concluded that the eigenvalues are nonpositve. According to [31] it is plausible that when the $\mathbf{DA}(.)$ matrix changes sufficiently slow and has small variations, the eigenvalues can be used to analyze its stability (i.e., system $\mathbf{x}$ converges to a nonzero but bounded vector). Further study of the general time varying $\mathbf{DA}(.)$ matrix is beyond the main scope of this paper. However, next we provide a proof of equilibrium point stability for the special case of regular networks.

6.2 Regular Networks

We start with the following proposition:

PROPOSITION 1 (REGULAR NETWORKS). *For a regular network $\mathbb{G}_{\boldsymbol{x}}^{\mathfrak{Reg}}$, $\boldsymbol{x}^\star = [\alpha, \dots, \alpha]^{\mathrm{T}}$ for $\alpha \in [0, 1]$ is a* stable *equilibrium point.*

PROOF. It is clear that the point $\mathbf{x}^\star = [\alpha, \dots, \alpha]^{\mathrm{T}}$ is an equilibrium point of Equation 8 since $\dot{\mathbf{x}} = \mathbf{DA}(\mathbf{x})|_{\mathbf{x}^\star} = 0$. Next, a proof that $\mathbf{x}^\star$ is stable has to be achieved to conclude the proposition. For that we start by computing the Jacobian matrix $\mathbf{J}$. Please note that in what comes next a regular network is adopted and therefore all nodes have the same degree d. Starting with the diagonal entries $[\mathbf{DA}]_{ii}$, the following is derived:

$$\nabla_{x_i}\Big[-\frac{1}{d}\sum_{j=1}^{N} w_{ij}p_{ij}\Big] = -\frac{1}{d}\nabla_{x_i}\sum_{j=1}^{N}\mathrm{sig}(\beta\Delta F_{ji})w_{ij}$$

$$= -\frac{1}{d}\sum_{j=1}^{N}\nabla_{x_i}\mathrm{sig}(\beta\Delta F_{ji})w_{ij}$$

$$= -\frac{\beta}{d}\sum_{j=1}^{N} w_{ij}\mathrm{sig}(\beta\Delta F_{ji})(1-\mathrm{sig}(\beta\Delta F_{ji}))$$

$$\nabla_{x_i}\Big[F(x_j) - F(x_i)\Big]$$

$$= -\frac{\beta}{d}\sum_{j=1}^{N} w_{ij}\mathrm{sig}(\beta\Delta F_{j,i})(1-\mathrm{sig}(\beta\Delta F_{j,i}))\Big[bw_{ij} + cd\Big]$$

Evaluating the above around the equilibrium point yields:

$$\nabla_{x_i}\Big[-\frac{1}{d}\sum_{j=1}^{N} w_{ij}p_{ji}\Big]\Big|_{\mathbf{x}^\star} = -\frac{1}{4d}\sum_{j=1}^{N} w_{ij}\Big[bw_{ij} + cd\Big]$$

Next, the derivations of the off-diagonal entries $[\mathbf{DA}]_{ij}, i \neq j$ are detailed:

$$\nabla_{x_j}\frac{1}{d}w_{ij}p_{ij} = w_{ij}\nabla_{x_j}\mathrm{sig}(\beta\Delta F_{ji})$$

$$= \frac{\beta}{d}w_{ij}\mathrm{sig}(\beta\Delta F_{ji})(1-\mathrm{sig}(\beta\Delta F_{ji}))\nabla_{x_j}\Delta F_{ji}$$

$$= \frac{\beta}{d}w_{ij}\mathrm{sig}(\beta\Delta F_{ji})(1-\mathrm{sig}(\beta\Delta F_{ji}))\nabla_{x_j}\Big[F(x_j) - F(x_i)\Big]$$

$$= \frac{\beta}{d}w_{ij}\mathrm{sig}(\beta\Delta F_{ji})(1-\mathrm{sig}(\beta\Delta F_{ji}))\Big[-cd - bw_{ij}\Big]$$

Evaluating the above around $\mathbf{x}^\star$ yields:

$$\nabla_{x_j}\Big[\frac{1}{d}w_{ij}p_{i,j}\Big]\Big|_{x^\star} = \frac{\beta}{4d}w_{ij}\Big(-cd - bw_{ij}\Big)$$

Therefore, the Jacobian of $\mathbf{DA}$ can be written as:

$$\mathbf{J} = \frac{\beta}{4d}\begin{bmatrix} -\sum\limits_{j=1}^{N} w_{1j}(cd + bw_{1j}) & \dots & -w_{1N}(cd + bw_{1N}) \\ \vdots & \ddots & \vdots \\ -w_{N1}(cd + bw_{N1}) & \dots & -\sum\limits_{j=1}^{N} w_{Nj}(cd + bw_{Nj}) \end{bmatrix}$$

An equilibrium point $\mathbf{x}^\star$ is stable if and only if all eigenvalues of the Jacobian matrix have negative real parts or, equivalently, if the Jacobian matrix is negative definite. Considering that $\mathbf{J}$ is a symmetric matrix, $\nu^{\mathrm{T}}\mathbf{J}\nu$, where $\nu = [\nu_1, \nu_2, \dots, \nu_N]^{\mathrm{T}}$ can be any non-zero column vector of N real numbers, can be calculated as:

$$\nu^{\mathrm{T}}\mathbf{J}\nu = \frac{\beta}{4d}\begin{bmatrix} -\sum_{j=1}^{N} w_{1j}(cd + bw_{1j})(\nu_1 + \nu_j) \\ -\sum_{j=1}^{N} w_{2j}(cd + bw_{2j})(\nu_2 + \nu_j) \\ \vdots \\ -\sum_{j=1}^{N} w_{Nj}(cd + bw_{Nj})(\nu_N + \nu_j) \end{bmatrix}^T \begin{bmatrix} \nu_1 \\ \nu_2 \\ \vdots \\ \nu_N \end{bmatrix}$$

$$= -\frac{\beta}{4d}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}(cd + bw_{ij})(v_i + v_j)^2 < 0$$

Therefore $\nu^{\mathrm{T}}\mathbf{J}\nu$ is always negative for any non-zero column vector ν. Therefore, $\mathbf{J}$ is a negative definite matrix. This concludes that all eigenvalues of $\mathbf{J}$ are negative, which means the equilibrium $\mathbf{x}^\star$ is a stable equilibrium. $\square$

This leads to the following theorem for regular networks:

THEOREM 2 (STABILITY IN REGULAR NETWORKS). *For a regular network* $\mathbb{G}_{\boldsymbol{x}}^{\mathfrak{Reg}}$, $\boldsymbol{x}^\star = [\alpha, \ldots, \alpha]^{\mathrm{T}}$ *for* $\alpha \in [0,1]$ *is a* stable *equilibrium point.*

7. CONCLUSIONS

This paper has introduced a novel model to study the evolution of cooperation in arbitrary complex networks. This model offers three advantages over existing approaches. (1) To the best of our knowledge, it is the first dynamical model that can be used to analytically study arbitrary complex networks. (2) It allows for a continuous level of cooperation, in line with many real settings that do not merely offer a binary choice between cooperation and defection. (3) The deterministic nature of the proposed model makes it computationally less complex than binary choice models, allowing to study a broader range of scenarios numerically.

Numerical simulations have been performed on a variety of regular and scale-free networks. These show that the proposed model promotes cooperation for a broad range of settings, clarifying why cooperation can prevail in natural societies. The results also indicate a positive relation between the network size and the level of cooperation in equilibrium on the one hand, and an inverse relation between connectivity and cooperation on the other. Additionally, a study of stability has been presented for arbitrary networks; moreover, regular networks have been shown to converge to a specific equilibrium where all agents adopt the same strategy.

8. REFERENCES

[1] R. Axelrod and W. D. Hamilton. The evolution of cooperation. *Science*, 211:1390–6, 1981.

[2] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–12, Oct. 1999.

[3] T. Börgers and R. Sarin. Learning through reinforcement and replicator dynamics. *Journal of Economic Theory*, 77(1), November 1997.

[4] R. Boyd, H. Gintis, and S. Bowles. Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science*, 328(5978):617–620, 2010.

[5] L. Buşoniu, R. Babuška, and B. De Schutter. A comprehensive survey of multi-agent reinforcement learning. *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews*, 38(2):156–172, 2008.

[6] D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, 2010.

[7] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci*, 5:17–61, 1960.

[8] D. G. Feingold and R. S. Varga. Block diagonally dominant matrices and generalizations of the Gershgorin circle theorem. *Pacific J. Math.*, 12:1241–1250, 1962.

[9] R. Gibbons. *A Primer in Game Theory*. Pearson Education, 1992.

[10] T. A. Han, L. M. Pereira, F. C. Santos, and T. Lenaerts. Good agreements make good friends. *Scientific reports*, 3, 2013.

[11] C. Hauert and G. Szabó. Game theory and physics. *American Journal of Physics*, 73:405, 2005.

[12] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 2002.

[13] L.-M. Hofmann, N. Chakraborty, and K. Sycara. The Evolution of Cooperation in Self-Interested Agent Societies: A Critical Study. *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, pages 685–692, 2011.

[14] M. O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008.

[15] R. L. Johnston. Gershgorin theorems for partitioned matrices. *Linear Algebra and its Applications*, 4(3):205–220, July 1971.

[16] M. Kearns and S. Suri. Networks preserving evolutionary equilibria and the power of randomization. *Proceedings of the 7th ACM conference on Electronic commerce - EC '06*, pages 200–207, 2006.

[17] T. Killingback and M. Doebeli. The continuous prisoner's dilemma and the evolution of cooperation through reciprocal altruism with variable investment. *The American Naturalist*, 160(4):421–438, 2002.

[18] W. Levine, editor. *The Control Handbook*. CRC Press, Boca Raton, FL, 1996.

[19] J. Maynard Smith and G. R. Price. The logic of animal conflict. *Nature*, 246(2):15–18, 1973.

[20] M. A. Nowak and R. M. May. Evolutionary games and spatial chaos. *Nature*, 359(6398):826–829, 1992.

[21] H. Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441(7092):502–505, 2006.

[22] H. Ohtsuki and M. A. Nowak. The replicator equation on graphs. *Journal of theoretical biology*, 243(1):86–97, 2006.

[23] S. Park, J. Deyst, and J. P. How. Performance and lyapunov stability of a nonlinear path-following guidance method. *AIAA Journal on Guidance, Control, and Dynamics*, 30(6):1718–1728, November-December 2007.

[24] F. Santos and J. Pacheco. Scale-Free Networks Provide a Unifying Framework for the Emergence of Cooperation. *Physical Review Letters*, 95(9):1–4, Aug. 2005.

[25] F. C. Santos, M. D. Santos, and J. M. Pacheco. Social diversity promotes the emergence of cooperation in public goods games. *Nature*, 454(7201):213–216, 2008.

[26] K. Sigmund, C. Hauert, and M. A. Nowak. Reward and punishment. *Proceedings of the National Academy of Sciences*, 98(19):10757–10762, 2001.

[27] R. Sutton and A. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge MA, 1998.

[28] K. Tuyls, P. 't Hoen, and B. Vanschoenwinkel. An evolutionary dynamical analysis of multi-agent learning in iterated games. *Journal of Autonomous Agents and Multi-Agent Systems*, 12(1):115–153, 2006.

[29] K. Tuyls and G. Weiss. Multiagent learning: Basics, challenges, prospects. *AI Magazine*, 33(3):41–53, 2012.

[30] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *nature*, 393(6684):440–442, 1998.

[31] J. J. Zhu. A note on extension of the eigenvalue concept. *Control Systems, IEEE*, 13(6):68–70, 1993.

[32] M. G. Zimmermann and V. M. Eguíluz. Cooperation, social networks, and the emergence of leadership in a prisoner's dilemma with adaptive local interactions. *Physical Review E*, 72(5), Nov. 2005.