Evolution of Networked Populations in Spatial Iterated Prisoner's Dilemma with Strategy Switching

João Calmeiro

Abstract

We focused our study in understanding how populations of individual strategies co-evolve when restricted to an underlying connection topology of a scale-free network generated by the Barabási-Albert algorithm with an average degree $\langle k \rangle \approx 3.985$ and a mean cooperation rate of 0.618857 , according to three different switching mechanisms. A preliminary analysis shows that, for an *Axel Rank*, the dominant strategies prevail through a change in tournament style; and that strategies that Win consistently, achieve high *Elo Ranks*. We find a positive correlation between the *Axel rank* and the Δ to *Elo Rank*. The evolution of strategies is accompanied by an expected rise in mean cooperation rate for *Axel* and unexpected rise for *Elo*. In the end we verify the *Emergence of Cooperation*.

I. Introduction

"Cooperation in organisms, whether bacteria or primates, has been a difficulty for evolutionary theory since Darwin".
— Axelrod, The Evolution of Cooperation(1984).

In the late 1970's, Robert Axelrod, a man deeply concerned about international politics – and especially the risk of nuclear war – conducted a computer tournament structured as 200 turns round-robin (and repeated 5 times to reduce score variability) with the objective of identifying the conditions under which cooperation could emerge in a Prisoner's Dilemma (PD). He invited game theorists to submit strategies for playing an Iterated Prisoner's Dilemma (IPD) game with the following payoff matrix [6]:

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array}, \tag{1}$$

where D represents Defection, C represents Cooperation, and the terms R, S, P, and T, are called the Reward, Sucker, Punishment, and Temptation.

In a PD, the matrix elements satisfy the following rank ordering: $S < P < R < T$. For the repeated version (ie the IPD), usually the constraint $T + S < 2R$ is also assumed to prevent the strategy profile of alternating Cooperation and Defection giving a greater reward than mutual Cooperation.

If the players know the game is played exactly N times, then the only possible Nash equilibrium is to always defect[1]. Similar reasoning applies if the game length N is unknown but has a known upper limit. For cooperation to emerge between game theoretic rational players, the total number of rounds N must be unknown to the players. In this case the *Always Defect* Nash equilibrium may no longer be a strictly dominant strategy.

Humans frequently face Prisoner's Dilemmas in the every day life when they have to pick between selfishness and altruism, to work hard or laze. The applications go beyond human or animal societies.

II. Motivation

A simple application of the PD, although informative, is not very interesting. Even with the inclusion of randomness into the strategies or tournament structure, by playing out repeated PDs, the stochastic effects should taper off. A more interesting application is the study of how a repeated PD evolves should the agents (eg the players) be allowed to switch strategies. In this context, this work explores how the dynamics of a PD change under different switching mechanisms.

The full study of the dynamics is too broad to be covered here, so we focus only on a few aspects (that vaguely mirror social interactions).

First, there is a myriad of variables we could consider, but we focus mainly on two metrics the: *payoff* and the *cooperation rate*.

Second, we consider the effects of *how* the agents are paired in a match. Again, this topic is too broad so

[1] We can infer this conclusion from a simple inductive reasoning that one might as well defect on the last turn, since the opponent will not have a chance to later retaliate. Therefore, both will defect on the last turn. Thus, the player might as well defect on the second-to-last turn, since the opponent will defect on the last turn no matter what is done, and so on.

we focus on the effects of a pairing based on a network – where the nodes are the agents, and the edges are the pairings allowed.

And third, how the strategies evolve when the agents are allowed to adopt the strategy of one of its neighbours.

III. Theoretical Background

In this section we give a brief overview of the concepts approached by this project. The topics are expanded where needed, but the purpose is more informative rather than an exhaustive description.

3.1. Scale-free Networks

A scale-free network, is one such that the fraction $P(k)$ of nodes having k connections to other nodes is

$$P(k) \sim k^{-\gamma}\,,$$

where typically $\gamma \in (2,3)$.

The study of such networks as approximations for real-world networks is disputed [8]; nonetheless, here it serves its purely academic purpose.

Several mechanisms for generating scale-free networks exist. However, in the implementation of this project, we made use of the Barabási–Albert (BA) model. It generates random scale-free networks using the method of preferential attachment: the more connections a certain node has, the likelier it is to form additional connections, commonly called the *rich get richer* method.

3.2. Selection Methods

Selection methods are greatly context dependent. What is advantageous (or not) depends on the specific purpose of the analysis.

In the context of this project we want to study which Prisoner's Dilemma strategies "survive" when players are allowed to change strategy between tournaments. Consequently, the method of selection is based on the scores of the players at the end of a tournament.

Specifically, we define that at the end of a tournament a player has a certain probability of adopting the strategy of its best scoring first-neighbour: the higher the score difference, the higher the probability. (The details of this probability are discussed in § 3.3.1.)

The restriction to the first-neighbours is in keeping with the application of a scale-free network model. If a player had unrestricted access to the information about the strategy of the best scoring node, this would quickly become the default. While farther kth-neighbours can be consider while maintaining the underlying purpose of a network, such analysis is beyond the scope of this project.

3.3. Elo Rating System

The choice of scoring system to base the rates mentioned in (2) can have drastic effects on the final results (as will be seen in § 5.1.2). One obvious (and usual) choice is to consider the scores given by the PD (as in (1)); another possible choice is to consider an external scoring system, with different properties, and see how the results fare in comparison.

In the scope of this project, we chose to also consider the Elo Rating System, having an important appearance of the logistic function (§ 3.3.1) in its expected score function, and being one of the most popular systems of rating in use.

3.3.1. Logistic Function

The logistic function is given by

$$f(x) = \frac{L}{1+e^{-\beta(x-x_0)}}\,,$$

where β is the logistic growth rate, x_0 is the midpoint of the Sigmoid curve, and L is the maximum value. It can be found in many applications, for example, in modelling population growth, the spread of disease, and Fermi-Dirac statistics in physics. Notably, when $L=1$, it is a smooth function valued in $(0,1)$. We apply this function as a rule for the transition of strategies: given a pair of players, A and B, the probability that player A with strategy s_A will transition into using strategy s_B is given by:

$$P(s_A \to s_B) = \frac{1}{1+10^{\frac{R_A-R_B}{N}}}\,, \tag{2}$$

where R_A, and R_B are the rates of the players (under some scoring system), and N is a coefficient appropriate for the rating scale. (Where N includes both β and the switch from base e to base 10.)

Similar probabilities, based on payoffs, expected payoffs, and other metrics, can be found throughout the literature. [1, 10]

Regarding the coefficient N, for $R_A - R_B > 0$, if $N \to \infty$, then $P(s_A \to s_B) \to \frac{1}{2}$, ie the updating rule is stochastic, resembling a coin flip; if $N \to 0$, then $P(s_A \to s_B) \to 0$, or conversely $P(s_B \to s_A) \to 1$, ie the updating rule is deterministic in always choosing the strategy of the higher rated player.

3.3.2. Elo Rating System

The Elo rating system, originally invented as an improved chess ranking system by Arpad Elo (1903–1992), is nowadays applied in numerous areas such as sports and video games. The Elo rating is a comparative and self correcting rating system that calculates for every player – based on performances – a numerical rating R that changes over time according to the outcome of matches.

After every turn, a player takes or gives a certain number of points to the opponent, in proportion to the difference in rating between the two players (ie $R_A - R_B$).

The key point about the Elo rating is that it is related to the log-odds of players winning games, so that

$$400 \log_{10}(\text{odds(A beats B)}) = R_A - R_B\,.$$

That is, there is an expected score for player A:

$$E_A = \frac{1}{1 + 10^{(R_B - R_A)/400}}\,,$$

and for player B:

$$E_B = \frac{1}{1 + 10^{(R_A - R_B)/400}}\,,$$

where the new rating, R'_A, is then given by

$$R'_A = R_A + K(S_A - E_A)\,,$$

and likewise for player B, where $S_A = 0,\, 0.5,\, 1$ for a loss, a draw, or a win, respectively, and K is some factor (discussed bellow).

The expected score is sometimes written as

$$E_A = \frac{Q_A}{Q_A + Q_B}\,,$$

and

$$E_B = 1 - E_A = \frac{Q_B}{Q_A + Q_B}\,,$$

where $Q_A = 10^{R_A/N}$, and $Q_B = 10^{R_B/N}$.

***K*-factor** Two players automatically wager some of their Elo points before their match. The size of this wager is known as K-factor and it regulates how quickly the ratings change in response to new information. A high K-factor makes the Elo rate very sensitive to recent results, so ratings jump around a lot based on each match outcome. A low K-factor make Elo rates to be slow in changing, since every match carries comparatively little weight.[2]

Finding the appropriate the K-factor is a challenge, since it logically varies by type of game, for example, in games where players (or teams) are expected to play many games (for example baseball) the K-factor should be smaller.[3]

3.3.3. Elo Payoff Matrix

In the context of a game's payoff, the Elo rating has a symmetric, zero-sum payoff matrix.

Consider that the Row player is higher Elo rated than the Column player, then, from the Row player's perspective, we have the following cases:

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} D & L \\ W & D \end{pmatrix} \end{array}, \tag{3}$$

where, W is a Win, L is a Loss, and D is a Draw.

In case of either L or D, the Row player loses rating, while the Column player gains rating; therefore, in order to maximise the chances of gaining rating, both players are incentivised to play Defect. In our case, that implies that the probability (2) using this method of rating will favour strategies that Defect more than Cooperate (see § 6.1.3).

3.4. Moran Process

A Moran Process is a simple stochastic process commonly used to describe finite populations, where individuals are selected based on some characteristic.

In each iteration, a random individual is chosen for reproduction (proportionally to some fitness function), and a random individual is chosen for death[4]. Under this process the population size is constant, but always results in the annihilation of the less favourable characteristics.

3.5. Axelrod's Tournament

Here we provide a small glossary to clarify some of the terms used throughout [16].

- **Action.** An action is either to Cooperate or Defect.
- **Strategy.** A strategy is a set of instructions that dictate how to choose an action.
- **Player.** A player is a single agent employing some strategy.
- **Play.** A play is a single player choosing an action.
- **Turn.** A turn is a one-shot interaction between two players (ie 2 plays).
- **Match.** A match is a consecutive number of turns.

[2] In current systems used by some chess federations the K-factor is a function of the number of games played in addition to the player's current rating [11].

[3] Whereas, for predictable tournaments, the K-factor should be large. An unfortunate example of this would be, according to the author, the Portuguese First League.

[4] The same individual can be chosen for reproduction and for death in the same iteration.

- **Win.** A win is attributed to the player with the highest total score in a match.
- **Round-robin.** A round-robin is the set of matches for all combinations between a given set of players.
- **Repetition.** A repetition is the application of another round-robin.
- **Tournament.** A tournament is a given number of repetitions. (So as to smooth out stochastic effects.)
- **Ranking.** A ranking is the ordering of a set of players by their median score.

IV. Procedure

In this project we study how the set of used strategies evolves when players are allowed to change strategy between different tournaments.

We keep record of two types of scoring. A *primary* scoring, based on the payoffs of the Prisoner's Dilemma, that determines who wins a certain tournament; this score is reset between each tournament, in similar fashion to how "season" scores are reset in sports. A *secondary* scoring, based on the Elo rating system; this score is kept between tournaments.

4.1. Implementation

We made use of the `axelrod` library for *Python*. It includes implementations for a vast number of strategies, and functions to simulate the tournament with different parameters and functionalities. However, since we also aim to study tournaments with an Elo rating system, and with custom criteria for switching strategies, some extra functionalities had to be implemented to accommodate our goals.
The challenge is to choose the set of parameters and strategies for this project. We proceed so.

4.2. Prisoner's Dilemma

4.2.1. PRST Parameters

Two IPD constraints were already introduced in the previous section. To reinforce the dilemma for the IPD, an interesting choice of parameters $\{P, R, S, T\}$ should respect the defector's extra income (in comparison with mutual cooperation) $T-R$, being less than the relative loss of the cooperator $R-S$. So, in historical fashion we select the same parameters Axelrod picked for his tournaments: $\{P, R, S, T\}=\{1, 3, 0, 5\}$.

Expanding the dilemma for varying payoff parameters has been studied in the academic literature [1] and is beyond the scope of this project.

4.2.2. Tournament Parameters

A tournament has 2 essential parameters, the number of turns, and the number of repetitions.

Number of Turns Strategies can have quite varied behaviours: while some are deterministic in their plays, others can adapt their subsequent plays in response to previous turns. That is, some strategies need to "learn" what plays are effective against a particular opponent; consequently, these types of strategies need a minimum number of turns to reach some steady-state or equilibrium (if such state is possible).

To get a feeling of the adequate number of turns to use throughout the analysis (ie when/if a steady-state is achieved), we find the ranks per turn for all the strategies competing in a tournament.

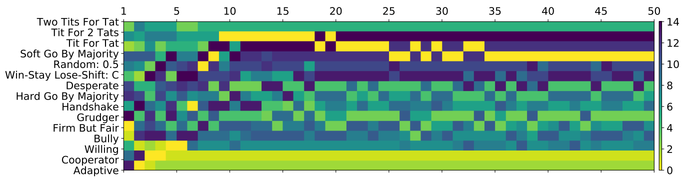


Figure 1. Ranks per turn, for all the competing strategies in the tournament. (Lower is better.)

In fig. 1 we see a lot of ranking variability at a lower number of turns (eg fewer than 5); after about the 20 turns mark, the rank variability seem to taper off for some strategies. The strategies either stabilise – mostly in lower or higher rankings (namely Bully, Pavlov, Tit For Tat variants) – or they alternate between two or three ranks. We also observe that, in general, as the turn number is increased, the strategies tend to keep their rankings for longer periods (in number of turns).

In consideration of the previous analysis, we opted for a total turn number of 50.

Number of Repetitions Since our selection of strategies (see § 4.4) does not include highly stochastic strategies, a very high number of repetitions is unnecessary; thus we see no reason to change the default implementation, and opt for a number of repetition of 10.

4.3. Elo Rating

The K-factor should be chosen as too balance the desire that incorrect scores should adjust as quickly as possible against a desire not to have too much volatility in scores.

Since per tournament we perform a large number of turns per match, and repetitions thereof, the K-factor should be on the smaller end.

From experiments with varying values for these parameter, we saw little variance in final ranking by Elo rates; therefore we used the standard values of $K = 10$ and $N = 400$.

4.4. Strategies

From an extensive list [7, 16] of strategies studied in the PD literature, a selection of **15** strategies were chosen, considering historical importance, simplicity, and variation in type (ie non retaliation, adaptive, heuristic, group strategies, and Tit-for-Tat variants).

- **Adaptive.** Starts with CCCCCCDDDDD, play the strategy that has worked best, recalculated each turn.
- **Cooperator.** The blind optimist - Cooperates on every move.
- **Bully.** Starts by defecting and then does the opposite of opponent's previous move (ie the opposite of Tit for Tat).
- **Firm but Fair.** Cooperates on the first move, and cooperates except after receiving a sucker payoff.
- **Grudger.**[5] Cooperates, until the opponent defects, and thereafter always defects.
- **Handshake.** Starts with CD. If the opponent plays the same CD, cooperate forever, else defect forever.
- **Hard Majority.** Defects on the first move, and defects if the number of defections of the opponent is greater than or equal to the number of times it has cooperated, else cooperates.
- **Desperate.** First move is random. After that, only cooperates after mutual defection.
- **Willing.** First move is random. After that, only defects after mutual defection.
- **Pavlov.** Cooperates on the first move. If payoff $= \{R, T\}$ in the last round then repeats last choice, otherwise shifts choice.
- **Random.** Makes a random move.
- **Soft Majority.** Cooperates on the first move, and cooperates as long as the number of times the opponent has cooperated is greater than or equal to the number of times it has defected, else it defects.
- **Tit for Tat.** *An eye for an eye.* Cooperates on the first move, then copies the opponent's last move.[6]
- **Tit for Two Tats.** Cooperates on the first move, and defects once only after two defects from opponents.
- **Two Tits for Tat.** Starts by cooperating and replies to each defect by two defections.

V. Preliminary Analysis

In this section we make a preliminary analysis of the strategies employed.

The presence of a network restricts the match-ups available; therefore, by first studying the unrestricted match-ups, we can then analyse the underlying effect of the network in the overall results.

For this purpose we perform a twofold analysis: first, we pit the strategies directly against each other in round-robin tournament; second, we simulate a natural selection process by means of a Moran Process.

Unless stated otherwise, the tournaments are run with the parameters described in § IV.

5.1. Round-robin Tournament

This subsection comprehends the results for the round-robin tournament.

5.1.1. Payoffs

From fig. 4 and fig. 2, we can see that the payoff is highly dependent on the match-ups between strategies. For example, in the case of *Grudger*, it has a higher payoff due to beneficial confrontations versus some strategies (eg versus *Bully*) that makeup for bad match-ups versus others (eg versus *Handshake*), while, for example, *Tit for Tat* and its variants have a more stable payoff versus all strategies. This is not unexpected, since, historically, *Tit for Tat* variant strategies have performed well in PD tournaments.

At a glance, fig. 3 reveals that the Elo rating system lowers the global variance of the payoffs – the heatmap is globally less bright. This behaviour is easily confirmed verifying that the strategies that, at some point in fig. 2, have the highest payoffs (eg *Handshake* or *Bully*) are not as yellow anymore. This lowering of payoff variance is expected since the Elo rate adjustment values should tend to zero, specially for confrontationally stable strategies.

[5] This strategy, submitted by Morton Davis, came 8th in Axelrod's original first tournament.

[6] Developed by Anatol Rapoport, showing just three lines of code, was the winning strategy in not only the first but also the second Axelrod tournament in which other strategies were already aware of the Tit for Tat reputation and strategy predictability.

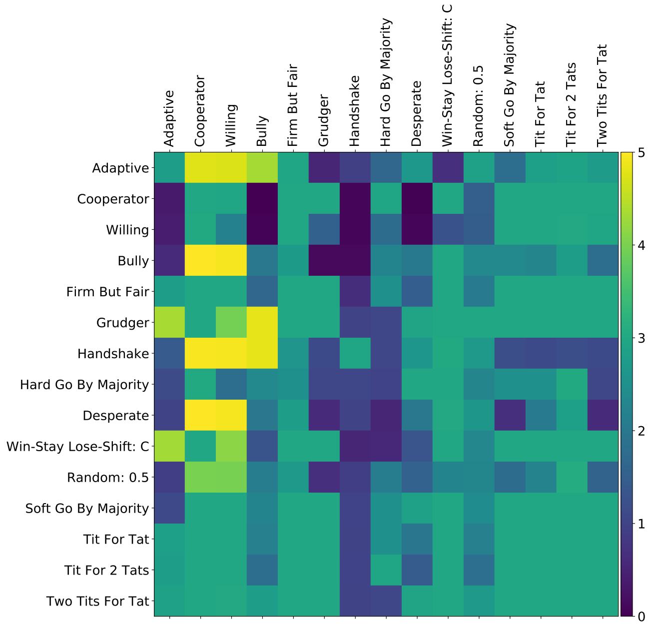


FIGURE 2. Heatmap for the preliminary tournament with the usual PD payoff matrix. Shown are the mean payoffs of each row strategy vs each column strategy. (Higher is better.)

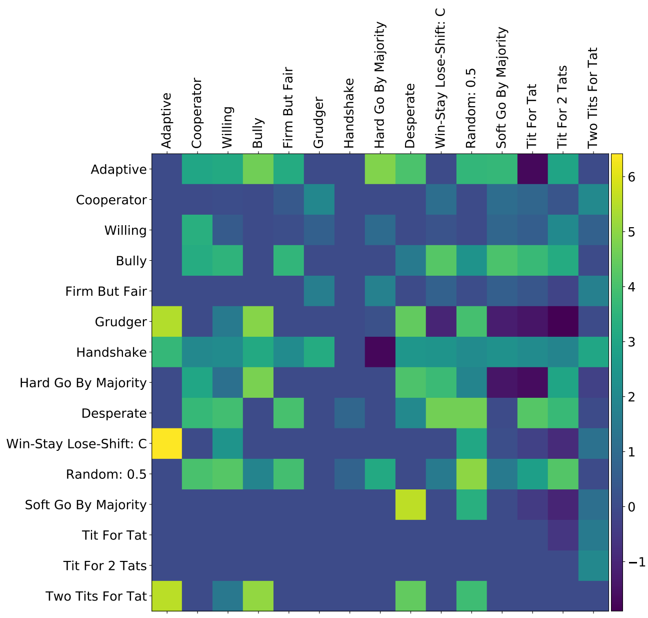


FIGURE 3. Heatmap for the preliminary tournament using the Elo score payoff matrix. Shown are the mean payoffs of each row strategy vs each column strategy. (Higher is better.)

5.1.2. Rankings

In table 1 and fig. 5 we exhibit a comparison of the ranks given by the mean scores of the PD versus the ranks given by the Elo ratings. We observe that all ranks change (some with drastic changes, like *Tit for Tat*) with the median change of 3 ranks.

The payoffs of the PD are static (cf § 4.2.1); so strategies keep accumulating points at the same rate, whereas the Elo rating is self correcting (cf § 3.3.2); so the payoffs taper off, that is, initially, all strategies start with a rating of 1200, then, for each repetition, the previous rate is taken into account, eventually settling on the "true" ratings. These differences account for the disparity in ranking between the two methods.

TABLE 1. Ranking by mean score (payoff) of the PD: Axel Rank; and ranking by the Elo rates: Elo Rank. Rank Δ is the rank change (up or down) from Axel to Elo Rank.

Strategy	Ranking		Δ
	Axel	Elo	
Grudger	1	4	-3.0
Two Tits For Tat	2	5	-3.0
Tit For Tat	3	11	-8.0
Adaptive	4	2	2.0
Soft Go By Majority	5	8	-3.0
Win-Stay Lose-Shift: C	6	9	-3.0
Tit For 2 Tats	7	15	-8.0
Firm But Fair	8	12	-4.0
Handshake	9	1	8.0
Bully	10	6	4.0
Random: 0.5	11	10	1.0
Hard Go By Majority	12	3	9.0
Desperate	13	7	6.0
Cooperator	14	14	-0.0
Willing	15	13	2.0
Mean Rank Δ			4.27
Median Rank Δ			3

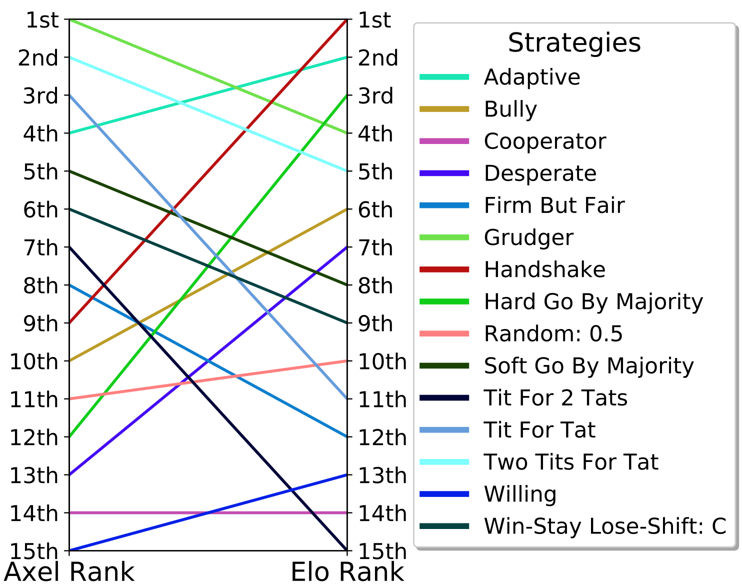


FIGURE 5. Parallel plot comparing the tournament rankings by mean score (payoff) of the PD: Axel Rank; and ranking by the Elo rates: Elo Rank.

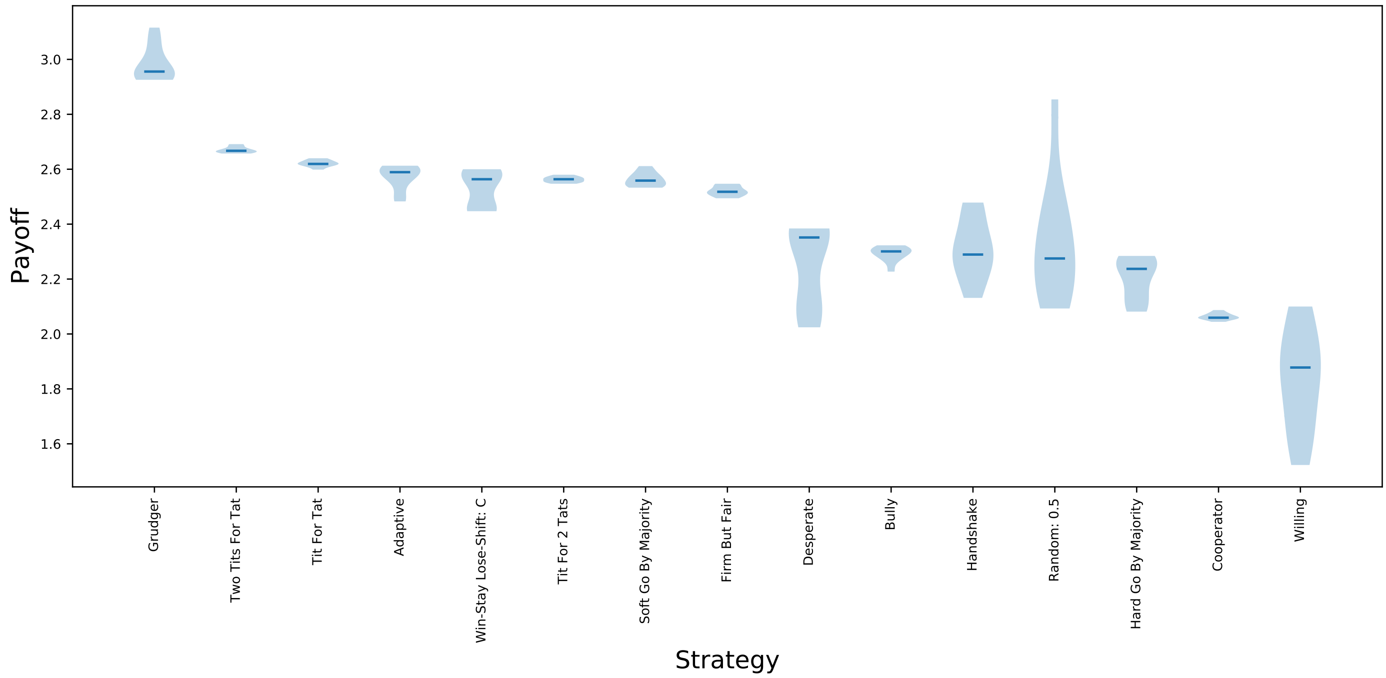


FIGURE 4. Preliminary tournament PD payoffs. Shown is the mean payoff and its variance, for each strategy.

5.1.3. Cooperation

In table 2 and fig. 6 we have the cooperation rate for all the strategies. We confirm some expected results: *Cooperator* has a 100% rate, and *Random* has a 50% rate. Overall, we see that the set of chosen strategies tend to cooperate more than defect; this fact should be kept in mind when interpreting cooperation rates further in the report.

TABLE 2. Normalised (per turn and per repetition) cooperation rates for each strategy.

Strategy	Cooperation Rate
Cooperator	1.000000
Willing	0.946714
Tit For 2 Tats	0.844714
Firm But Fair	0.832714
Tit For Tat	0.768857
Win-Stay Lose-Shift: C	0.746429
Soft Go By Majority	0.693143
Two Tits For Tat	0.691429
Grudger	0.558286
Random: 0.5	0.502143
Bully	0.494429
Adaptive	0.454286
Hard Go By Majority	0.364429
Desperate	0.351571
Handshake	0.033714
Mean cooperation	0.618857
Median Cooperation	0.691429

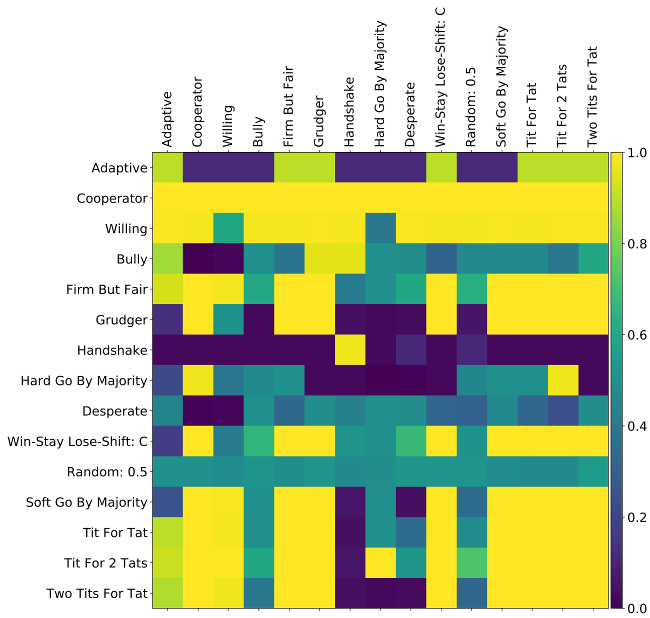


FIGURE 6. Heatmap for the preliminary tournament cooperation rate. Shown are the cooperation rates between each row strategy vs each column strategy.

5.1.4. Win–Loss Flow

Since some strategies win[7] out consistently (eg *Bully* wins vs *Cooperator*), we can consider the strategies from a Predator–Prey point of view, where a kind of "food chain" can be established. However, since we do not have a strict predatory relationship between all the strategies, as a comparative note we plot a chord diagram (fig. 7) of the payoff matrix (fig. 2) that enables a comparison of the winnings for each strategy (ie to

[7]A strategy is said to win when, at the end of a N-turn match, it has higher total payoff than its opponent.

visualy assess the proportion of confrontational payoffs contributing to each strategy global payoff).

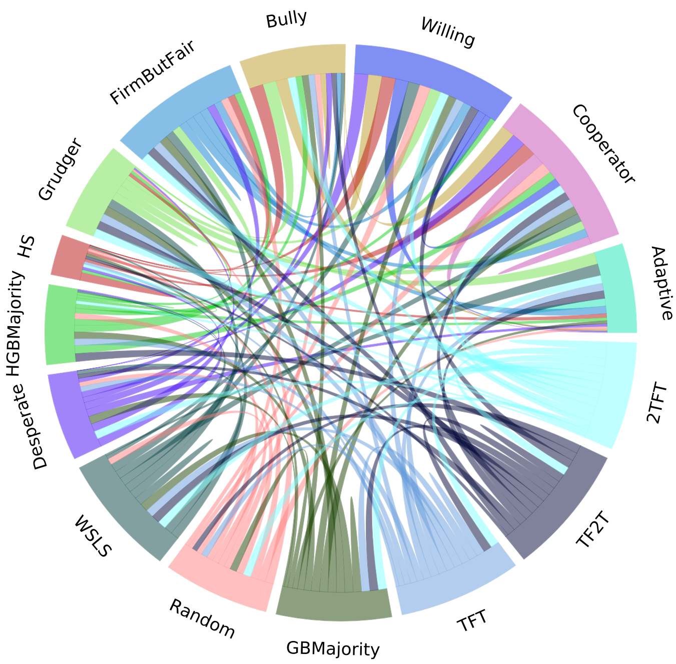


FIGURE 7. Chord diagram of the payoff matrix for the PD scores. The circular layout of the chord diagram is divided in sections, each representing a different strategy. Each strategy is then composed by 15 arcs (connecting any two strategies, including itself). The thickness of each arc is proportional to the payoff of the corresponding competing coloured strategy.

The interconnectedness of the results in fig. 7 is expected since, from fig. 2 we can see that the majority of payoffs are in the (approximate) range of 2 to 3.

Strategies with a lot of colour variety (eg *HGBMajority* and *Willing*) lose (ie have lower matching payoff), on average, with many other strategies. Strategies that win a lot (ie have higher corresponding matching payoff) fill much of their arc with their own colour (eg *TFT* and *TF2T*). The arc-length of each strategy's section evaluates the payoff loss to the others (eg *Handshake* being short giving off little payoff).

The purpose of implementing a chord diagram is to help the reader pick up on the relative contribution (arc thickness) of each connection (match-up). However, having 15 strategies difficults this task as thicknesses are small and arcs tangle. Moreover, it does not provide precise information about the winning frequency.

After some thought about §§ 3.3.3 and 5.1.2 is given, the importance of evaluating the number of wins per strategy is recognisable: the Elo rating – given its essence – stabilises the rates, and it favours Defection over Cooperation (ie the Reward payoff is less beneficial than the Temptation payoff); thus, winning more often becomes an important factor in increasing the Elo rate.

Unlike fig. 7, fig. 8 enables us to point out the global losers (eg *TFT*, *TF2T*, and *Cooperator*) and the global winners (eg *Handshake* and *Adaptative*). Uncoincidentally, they line up accordingly with the Elo Ranking results in table 1 and fig. 5.

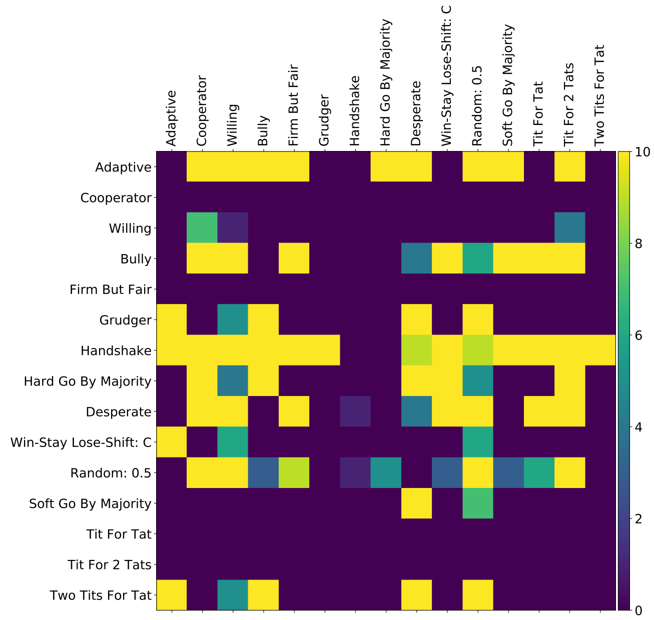


FIGURE 8. Heatmap for the preliminary tournament win count. Shown are the number of wins over a maximum of 10 (the number of repetitions in a tournament).

We hypothesise that strategies with a globally high number of wins (eg *Handshake* and *HGBMajority*), when subject to the Elo rating system, and a restriction of the match-up pairings, such as a network, should fare even better.

5.2. Moran Process

Table 3. Number of wins (ie last surviving strategy) for each of the strategies in 200 Moran Processes.

Strategy	Wins
Grudger	34
Two Tits For Tat	21
Tit For 2 Tats	21
Win-Stay Lose-Shift: C	19
Handshake	19
Tit For Tat	16
Soft Go By Majority	15
Firm But Fair	14
Bully	11
Adaptive	11
Desperate	7
Random: 0.5	5
Willing	4
Cooperator	2
Hard Go By Majority	1

We ran several Moran Processes for the set of strategies. The results are presented in table 3. We observe that the frequency of winners is in concordance with the results in fig. 4 and fig. 5.

5.3. Preliminary Conclusions

Through testing in the round-robin tournament, and Moran Process, we observe that the resulting winning strategies are consistent in both situations when considering the *Axel Rank* (fig. 5 and table 3), suggesting that in all-vs-all situations the dominant strategies – by PD scores – prevail through a change in the method of determining the winner, that is, from a round-robin tournament style, to an evolutionary/evolving style.

Comparing the win count in fig. 8 with the *Elo Ranking* in table 1, we can categorically match not only the global winners (*Handshake*, *Adaptative*, and *HGBMajority*) with the higher Elo ranked but also the global losers (*Cooperator* and *Firm But Fair*) with the lower Elo ranked.

Finally, for comparing the Axel and Elo rankings we use table 1 to construct fig. 9 which provides the needed perspective for detecting a pattern that was unnoticeable till now (although its explanation is not new) – A positive correlation between Axel rank and its Δ from the Elo rank. This result agrees with the theoretical principle that Elo rating stabilises the ratings.

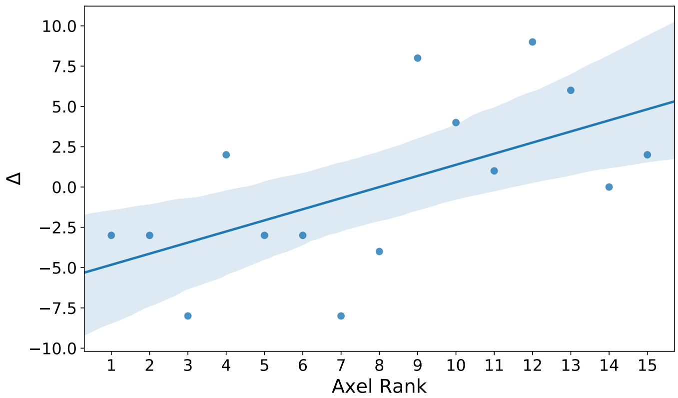


Figure 9. Scatter Plot of the ranked *Axel Rank* strategies vs Δ, from table 1, and their fitted line.

It's only natural to ask how these preliminary results would be affected from the influence of an underlying interaction structure of a network. This study should not only be more interesting but also more realistic as multi-player systems players do not usually interact with all other players. We now possess a foundation to analyse the overall results.

VI. Evolving Strategies

In this section we show the results across 10 scale-free networks, each with 525 nodes – 35 players for each one of the 15 strategies.

In each simulation we consider – beyond just the probability in (2) – the limiting cases of $N \to \infty$, and $N \to 0$, that is, we consider the three cases:

$$P(s_A \to s_B) = \frac{1}{1 + 10^{\frac{R_A - R_B}{400}}}\,, \tag{4a}$$

$$P(s_A \to s_B) = \frac{1}{2}\,, \tag{4b}$$

$$P(s_A \to s_B) = \begin{cases} 0\,, & R_A > R_B\,, \\ 1\,, & R_A < R_B\,. \end{cases} \tag{4c}$$

We aim to answer the following questions:

Q1 Survivors. What are the surviving strategies? Do they match the winners found in the Preliminary Analysis (§ V)?

Q2 Cooperation. Is Cooperation or Defection favoured by surviving strategies? How does the Mean Cooperation Rate vary?

Q3 PD Score vs Elo Rating. How does the Elo rating and the PD payoffs fare in comparison on determining the final (stationary) fraction of the population for each strategy?

Q4 Dominance of Global Winners. Do the global winners (such as defined in § 5.1.4) dominate the network when subject to the Elo rating?

6.1. Results

6.1.1. Axelrod

In fig. 10 we present the results for the cooperation and for the population when using the PD Scores (ie payoffs) as the determining factor in the probability of switching strategies, that is, the PD payoffs are the R_A and R_B in (4).

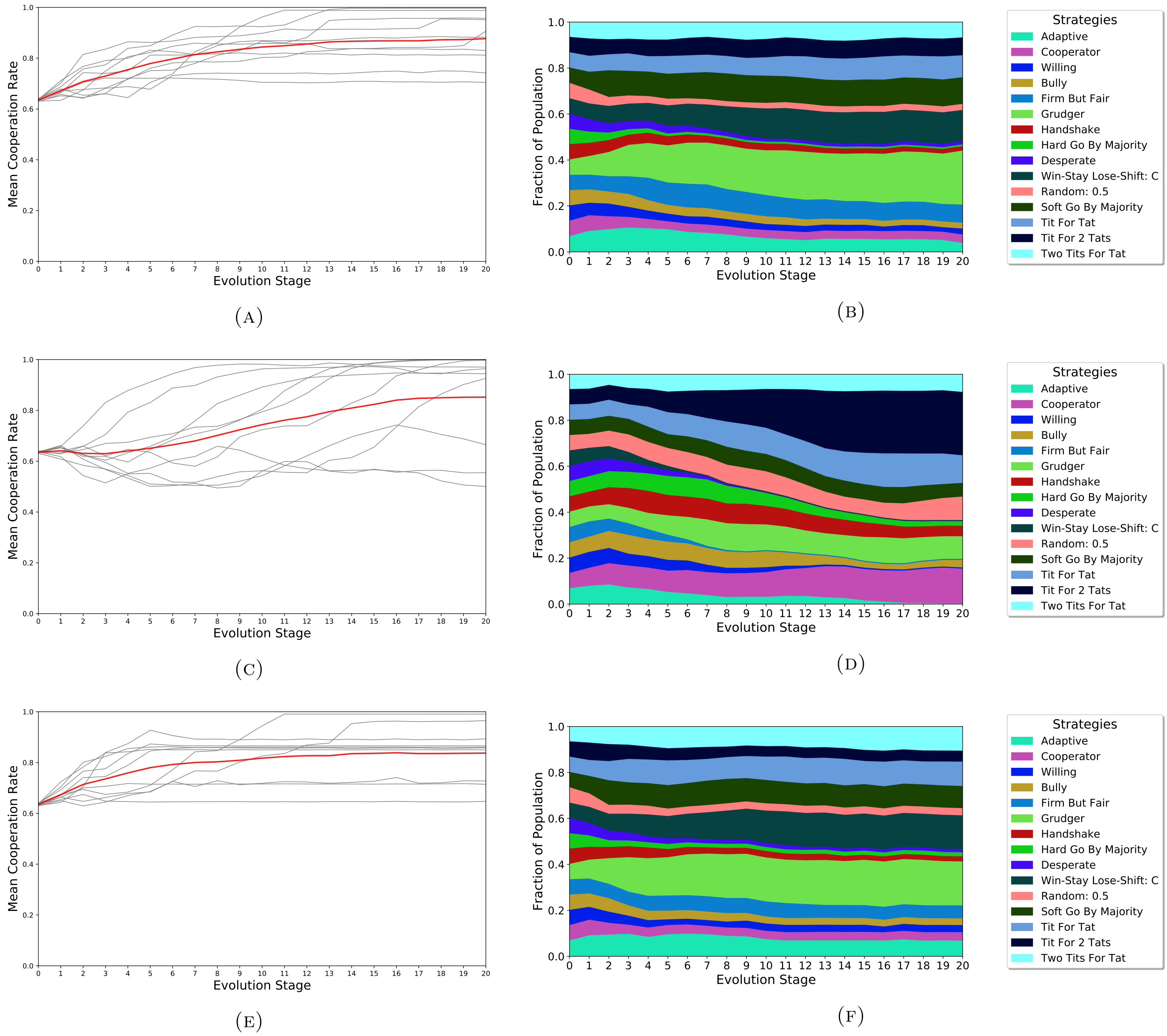


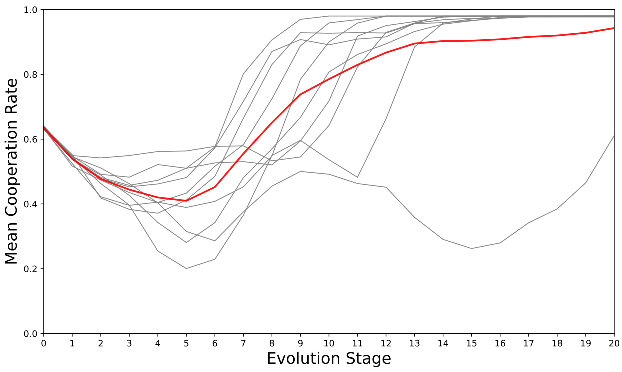
Figure 10. Results for the simulations, using 10 networks (see table A.1), and the probabilities in (4): figs. 10a and 10b for probability (4a), figs. 10c and 10d for probability (4b), and figs. 10e and 10f for probability (4c). In figs. 10a, 10c and 10e are the Mean Cooperation Rates per evolution stage, for each of the networks (grey). (In red is the mean.) In figs. 10b, 10d and 10f are the Fraction of Population per evolution stage, across the networks, for all strategies.

In figs. 10a, 10c and 10e we see an increase in the Mean Cooperation Rate, rising to a similar value. In essence, the rise of the Mean Cooperation Rate is aligned with the *Emergence of Cooperation* found in the literature [2].

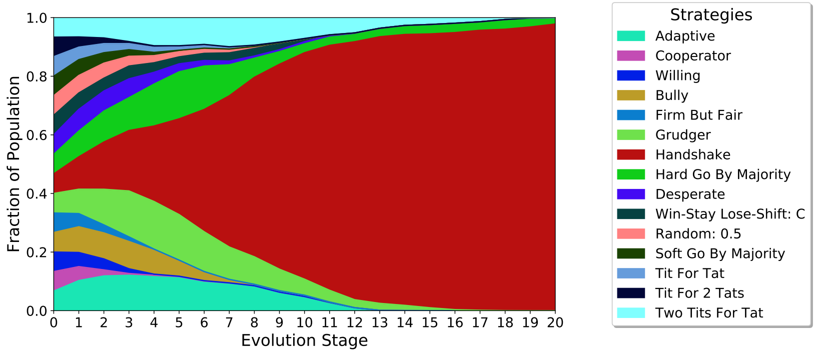
Despite figs. 10b and 10f displaying some slight differences in final populations, all strategies reach stationarity; no strategy strongly dominates the network, and the "winners" all possess high cooperation rates.

6.1.2. Elo Rating

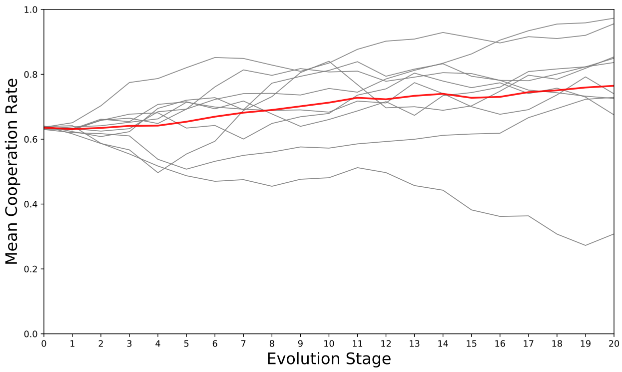
In fig. 11 we show the results for the cooperation and for the population when using the Elo Ratings in the determining factor in the probability of switching strategies, that is, the Elo ratings are the R_A and R_B in (4).



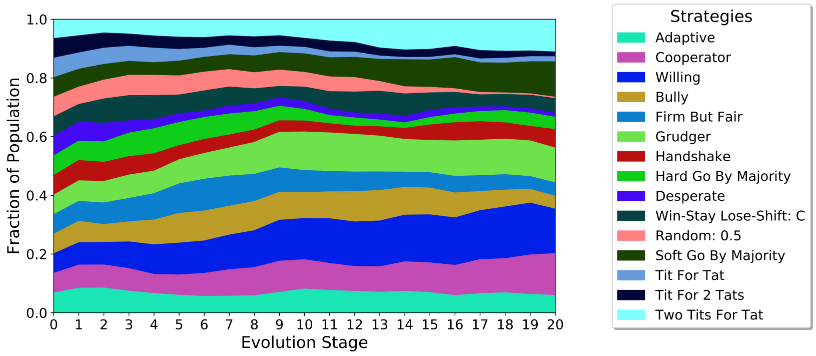
(A)



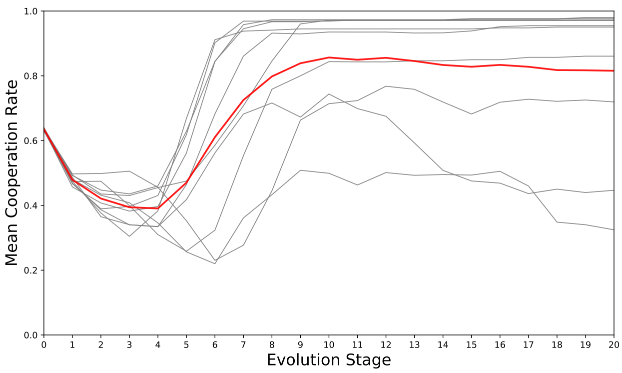
(B)



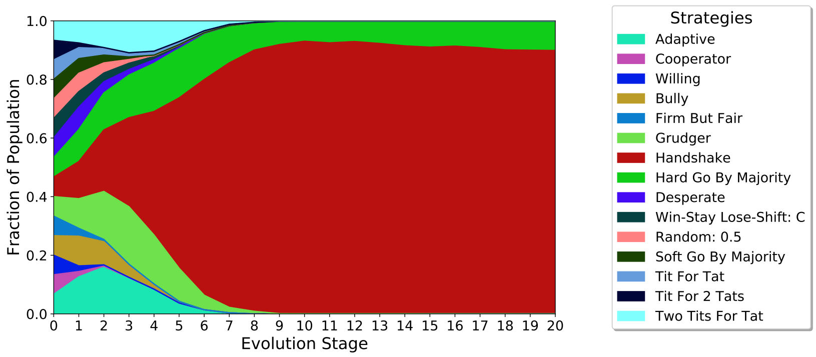
(C)



(D)



(E)



(F)

Figure 11. Results for the simulations, using 10 networks (see table A.1), and the probabilities in (4): figs. 11a and 11b for probability (4a), figs. 11c and 11d for probability (4b), and figs. 11e and 11f for probability (4c). In figs. 11a, 11c and 11e are the Mean Cooperation Rates per evolution stage, for each of the networks (grey). (In red is the mean.) In figs. 11b, 11d and 11f are the Fraction of Population per evolution stage, across the networks.

An odd and interesting shape of Mean Cooperation Rate is found in figs. 11a and 11e. The initial decrease is in accordance with defectors being favoured by Elo Rating. In the early evolution stages the global winning strategies (fig. 8) govern most of the population.

At juvenile stages the Mean Cooperation Rate increases, revealing an interesting occurrence from the *Handshake* strategy as it becomes the dominant strategy across the networks (figs. 11b and 11f).

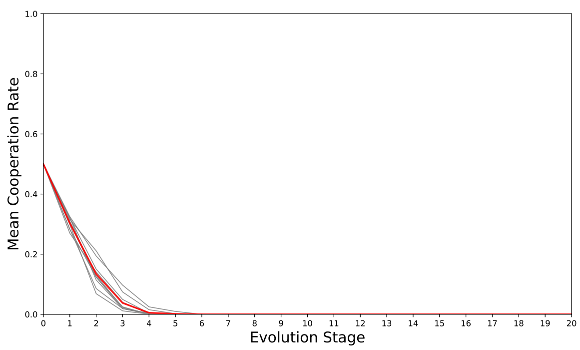
This is due to the effects of unintentional self-boosting[8]: as the *Handshake* strategy is likely to defect, it gains an initial boost in rating, and as more nodes adopt the *Handshake* strategy, they start to Cooperate with each other, boosting the strategy further.

In the final evolution stages, only *Handshake* and *HGBMajority* are competing. Given their nature, they are bound to play exactly the same way in every set of their confrontations, so each strategy gets the same confrontation payoff at every repetition (50 turns). The payoff difference lies on the mutual cooperation between *Handshake* and mutual defection between *HGBMajority*.

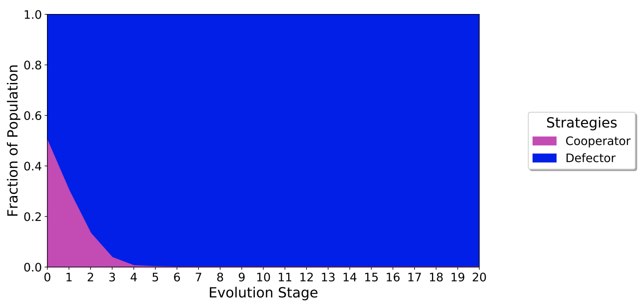
Since Elo rates higher a "weaker" player winning against a "stronger" one, rather than the reverse, we should expect in later stages a stabilisation of the two populations of strategies in coexistence.

For fig. 11c – as expected – the result is similar to fig. 10c, ie a slight rise in the Mean Cooperation Rate.

6.1.3. Elo Rating – Cooperator vs Defector



(A)



(B)

Figure 12. Mean Cooperation Rate and Fraction of Population, using probability (4a), for a population of only *Cooperator* and *Defector* strategies. In fig. 12a we see that the Mean Cooperation Rate decreases until no cooperation exists; that is, *Defector* is the only surviving strategy (fig. 12b).

In this section we perform a brief test to empirically show that Elo benefits Defection over Cooperation, using a simple population of *Cooperator* and *Defector* strategies.

Consistent with § 6.1.2, we see in fig. 12 that Defection wins over Cooperation. Around the sixth evolution stage Defection has already won over the population in the network; this is similar to figs. 11a, 11b, 11e and 11f, whereafter Cooperation is only restored due to the effects already discussed in § 6.1.2.

VII. Discussion

In fig. 13 we see that the surviving strategies (fraction of the population) vaguely resemble the ranking by *Axel Rank* (decreasing top to bottom) – both when using probability A and probability C. For probability B the values are expectedly different from either A or B (there being little effective difference between A and B).

The situation for § 6.1.2 similarly follows: the preliminary results for initial stages where the population follows the *Elo Rank*, and the hypothesis at the end of § 5.1.4 of global winners being favoured. However, it is then taken over by the effects of self-boosting.

Regarding the Mean Cooperation Rate, for § 6.1.1, all three switching probability schemes increase up to a steady-state with a high value – *Emergence of Cooperation*. For § 6.1.2, an initial downfall is exhibited as Defection is favoured. As some strategies die, this tendency breaks and we observe the self-cooperation boost from some strategies.

Curiously, the top survivors by *Axel Rank* do not possess the highest values of Cooperation Rate (eg *Grudger*: 0.56, *SGBMajority*: 0.69, and *Tit for Tat*: 0.77). Whereas, *Willing* with 0.95, barely survives each tournament. This result can (cautiously) be extrapolated to a controlled multi-player network realistic scenario with the conclusion we should not be blind Cooperators with our peers. The results also indicate, in the specific conditions tested, that some Cooperation is necessary to survive.

VIII. Future Work

Prisoner's Dilemma is still a current research area with nearly 18000 papers during the past five years [12].

The first steps in furthering the research are computational in nature, eg the inclusion of a greater number of nodes, evolution stages, and networks.

[8] Other tournaments have used actual intentional self-boosting. In 2004, Graham Kendall won the 20th-anniversary IPD competition on a tournament that allowed submission of multiple strategies. The University of Nottingham lecturer submitted 60 programs that relied on their collusion in order to achieve the highest number of points for a single program. They were designed to recognise each other through a series of five to ten moves at the start [3]. This strategy ended up taking the top three positions in the competition, as well as a number of positions towards the bottom [13].

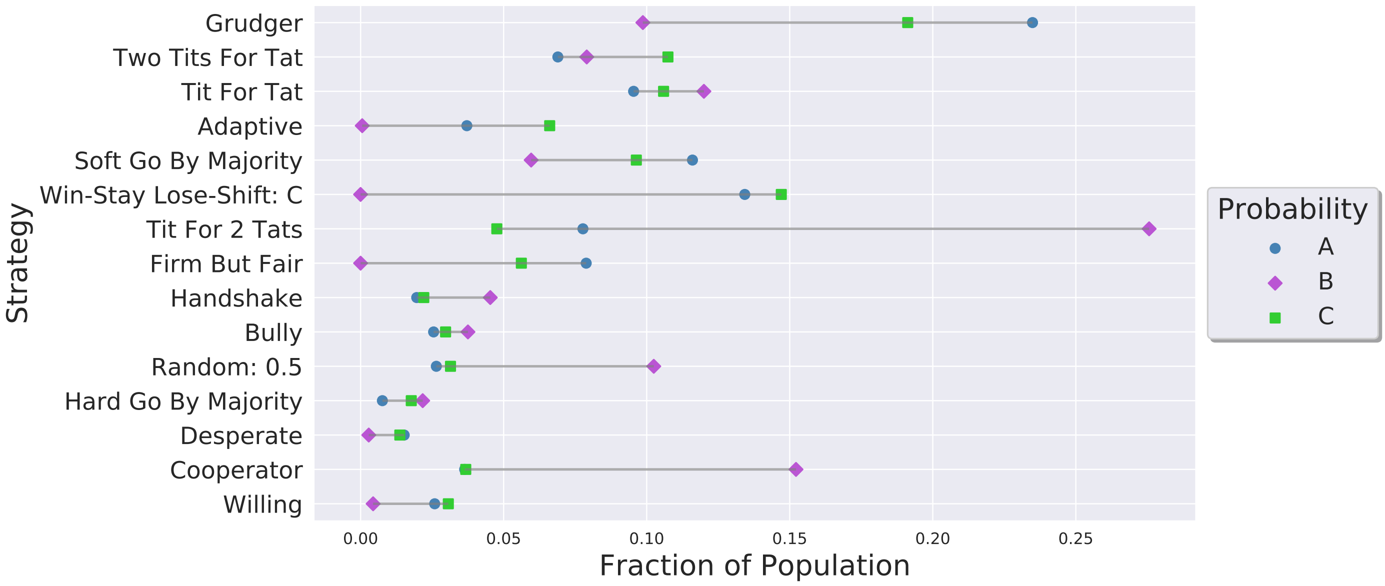


FIGURE 13. Final (evolution stage) fraction of population for § 6.1.1. Probability A, B, and C refer to (4a), (4b), and (4c), respectively. The strategies (left axis) are ordered according to the *Axel Rank* in table 1.

In regards to the networks, we only considered scale-free networks generated by the Barabási-Albert algorithm; further research includes different scale-free generating algorithms, and different network types altogether.

Different sets of strategies could also be tested. Namely: sets of more stochastic or more deterministic strategies; sets of different overall initial Cooperation rates; and sets of "interesting" strategies, eg the *Handshake* strategy led to a self-boosting effect in § 6.1.2.

We could test how the results change for a much lower cooperation rate: would we still observe an *Emergence of Cooperation*? It would be interesting[9] to repeat the analysis with the restriction of no self-interaction of strategies.

Parameter wise, we could further study the effects under different payoff values (PRST), and also apply a more quantitative method of picking the Elo parameters (K-factor and N). As the K-factor regulates ratings variation, it would be compelling to manipulate it in to find the value ranges it would achieve a better fit for fig. 13.

In this work, a qualitative analysis was chosen in order to better serve its exploratory purpose. However, it could be easily expanded to place a greater emphasis on a quantitative analysis (eg studying trends and variances of the Cooperation Rate, calculating the correlation coefficient for the Elo rank differences, Replicator Equation dynamics).

[9] And quite readily as the author's code is properly constructed to be prepared for that and a multitude of other variations. As interesting as this project may be, time is limited.

A. Implementation

The algorithm used to generate the scale-free network was the Barabási-Albert algorithm (§ 3.1). We made use of the implementation of this algorithm in the `networkx` module. The PD tournaments were simulated using the `axelrod` module (§ 4.1).

In table A.1, the "networkx" and "axelrod" columns refer to the *seed* parameter in the respective functions.

In all networks $\langle k \rangle \approx 3.985$ (average connectivity).

Table A.1. Seed values for the networks and tournaments used in §§ 6.1.1 and 6.1.2.

`networkx`	311241	582907	392722	34510	994611	479412	159459	537154	161788	742156
`axelrod`	99816	52652	401970	218664	928531	261145	313441	155399	696413	59651

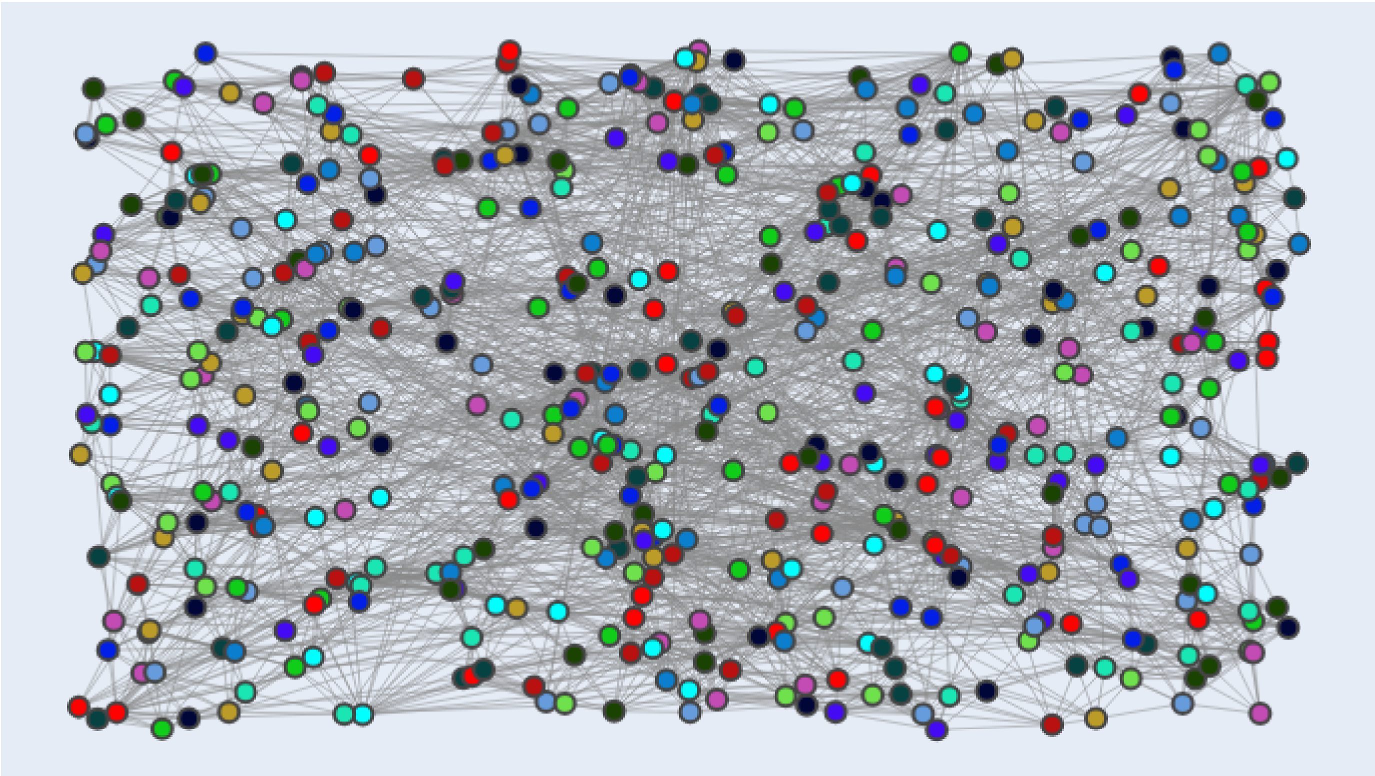


Figure 14. Representative Network chart generated for this project. The different colours represent the different strategies.

References

[1] F. Santos et all. "Evolutionary dynamics of social dilemmas in structured heterogeneous populations". In: *Proc. Natl. Acad. Sci. USA* 103.3490-3494 (2006). url: https://www.jstor.org/stable/173638.

[2] F. Santos et all. "Scale-Free Networks Provide a Unifying Framework for the Emergence of Cooperation". In: *Physical Review Letters* (2005). url: https://pdfs.semanticscholar.org/824f/6e30654db760358e59719779fb7285d18332.pdf.

[3] Graham et all. "The Iterated Prisoner's Dilemma: 20 Years On". In: May 2007. isbn: 978-981-270-697-3. doi: 10.1142/9789812770684_0001.

[4] R Axelrod. "Effective choice in the prisoner's dilemma". In: *The Journal of Conflict Resolution* 24.3 (1980), pp. 379–403. url: https://www.jstor.org/stable/173638.

[5] R Axelrod. *Effective choice in the prisoner's dilemma.* 7th ed. Elsevier, 2013.

[6] R Axelrod. "The Evolution of Cooperation". In: *The RAND Journal of Economics* 15.2 (1984), pp. 305–309. URL: https://www.jstor.org/stable/173638.

[7] B Benson. *A prisoner's dilemma cheat sheet.* URL: https://medium.com/thinking-is-hard/a-prisoners-dilemma-cheat-sheet-4d85fe289d87.

[8] A. Broido and A. Clauset. "Scale-free networks are rare". In: (2018). URL: https://arxiv.org/abs/1801.03400.

[9] A Elo. *The Rating of Chess Players, Past and Present.* 1st ed. Ishi Press, 2008.

[10] Szabo G. and G Fath. "Evolutionary games on graphs". In: *Physics Reports* 446.(4-6) (2007), pp. 97–216. DOI: http://dx.doi.org/10.1002/andp.19053221004.

[11] M. E. Glickman and A.C. Jones. "Rating the Chess Rating System". In: *Chance* 12.21-28 (1999). URL: http://math.bu.edu/people/mg/research/chance.pdf.

[12] *Google Scholar Article Search.* Accessed: 2019-12-10. URL: https://scholar.google.com/scholar?q=prisoner%5C%27s+dilemma&hl=en&as_sdt=0%5C%2C5&as_ylo=2014&as_yhi=2019.

[13] J. Humble. *Prisoner's Dilemma Tournament.* URL: http://www.prisoners-dilemma.com/results/cec04/ipd_cec04_full_run.html.

[14] J. von Neumann and O. Morgenstern. "Theory of Games and Economic Behavior". In: *Princeton University Press* (1944).

[15] K Sigmund. "Games of Life". In: *Oxford Univ. Press* (1993).

[16] The Axelrod project developers. *Axelrod: Axelrod-Python/Axelrod: v4.7.0.* Apr. 2016. URL: http://dx.doi.org/10.5281/zenodo.35171.

[17] P. A. Whigham and G.Dick. "Evolutionary Dynamics on Graphs: The Moran Process". In: *Lecture Notes in Computer Science book series* 4247 (2006). URL: https://link.springer.com/chapter/10.1007/11903697_1.