Role of Aspirational Learning on Evolution of Cooperation in Static and Dynamic Networks

MS Project Report submitted in partial fulfillment of the requirements for the degree of

Master of Science

by

Anuran Pal

Roll No : **16MS152**

to the

Department of Physical Sciences

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH, KOLKATA

Supervised by

Prof. Supratim Sengupta

Professor, Department of Physical Sciences, IISER Kolkata

July 2021

Declaration of Authorship

July 22, 2021

I, Mr. Anuran Pal, Registration No. 16MS152, a student of Department of Physical Sciences of the 5 year integrated BS-MS Dual Degree Programme of IISER Kolkata, hereby declare that the M.S. Project Report titled “Role of Aspirational Learning on Evolution of Cooperation in Static and Dynamic Networks” is my own work and it has not been submitted for any degree/diploma or any other academic award anywhere before. All due references have been made.

Anuran Pal.

Anuran Pal
Department of Physical Sciences
Indian Institute of Science Education and Research, Kolkata
Mohanpur 741246, West Bengal, India.

Certificate from the Supervisor

July 22, 2021

This is to certify that the M.S. Project Report titled "Role of Aspirational Learning on Evolution of Cooperation in Static and Dynamic Networks" submitted by Anuran Pal, Roll No. 16MS152, a student of Department of Physical Sciences of the 5-year integrated BS-MS Dual Degree Programme of IISER Kolkata, is his work under my supervision. This is also to certify that neither the M.S. Project Report nor any part of it has been submitted for any degree/diploma or any other academic award anywhere before, and all due references have been made. In my opinion, this fulfills the requirements for the award of the degree of Master of Science.

Prof. Supratim Sengupta
Professor
Department of Physical Sciences
Indian Institute of Science Education and Research, Kolkata
Mohanpur 741246, West Bengal, India.

Abstract

Anuran Pal

Role of Aspirational Learning on Evolution of Cooperation in Static and Dynamic Networks

Altruistic acts come at a cost for the individuals who undertake them, but, cooperative societies perform much better than that of selfish ones. This dilemma has been approached using different models exploiting the availability of information to all the players. In this project, I have used the public goods games framework to look closely at a particular low information decision making process. Individuals embedded in a network, not privy to any information from the neighbours, aspire for a payoff based on previous experiences only, parametrized by a population level variable, learning ability. Categorizing players based on their satisfaction and the resulting action reveal periodic cycling between the different categories. It is observed that static network consisting of slow learners display sustained cooperative behaviour for short intervals. Incorporating cooperation or satisfaction supportive rewiring mechanism eliminates the defective phase altogether and the population display cooperative behaviour irrespective of the learning ability. Network with slow learners capable of modifying ties with ease result in the polarization of the population's expectations.

Acknowledgements

I express my sincere gratitude to Prof. Supratim Sengupta for his contribution in igniting my interests in Evolutionary Dynamics. His patience in dealing with my ever fluctuating interests is very much appreciated. I am really grateful to have worked under him for my project.
I would like to acknowledge the Department of Science and Technology, Government of India, for the generous financial support, which was awarded to me via the INSPIRE Fellowship.
I am grateful to the wonderful IISER community. This campus truly became my home for the past 5 years and I would cherish these moments for the rest of my life. Special thanks to Abhigyan, Arani, Balu, Bobb, Nilanjana, Naru, Raichu, Sid and Taniya. These few have received the worst from me but never shied away from sharing the best with me.
Most importantly, I would like to thank my mother and my late father for believing in me when times became tough, and for providing unconditional support so that I could blossom without constraints.

Contents

Chapter 1

Introduction

Cooperation is an ubiquitous feature in our world. Individuals engaging in cooperation have to pay a cost, but defection on the other hand reap higher benefits by exploiting the good nature of cooperators. So, it is natural for logical and myopic individuals to defect and receive higher benefits from the surrounding cooperators. This apparent illogical decision-making, although not optimum for the individual players, is perfect for the society. The sum of the benefits obtained by the individuals in the entire population is maximized if cooperation persists in the network. Otherwise, the population is filled with defective players who do not donate neither do they receive any benefits, famously called "tragedy of the commons". How do individuals in systems ranging from animals in societies to cells in tissues, display cooperation when individual interests do not align with the group's wellbeing?

Simple reactive strategies like tit-for-tat, generous-tit-for-tat, win-stay-lose-shift among others are able to show the persistence of cooperation in two player games [6]. Multiplayer games are modelled using the framework of Public Goods Games (PGG), where cooperators donate to the public goods and receive returns, but defectors exploit the public goods without contributing. This formalism is able to capture the social multiplayer dilemma that individuals in a population face. Several Evolutionary Game Theory (EGT) models have been developed to tackle this dilemma and have been successfull in showing that cooperation can persist if the payoff-maximizing agents have an underlying startegy update mechanism. Many such mechansims utilize the information that the players gather from their surrounding environment, like action [4], empathy [10], wealth [3] etc. More direct approaches involve cooperation enforcing mechanisms like punishment [1]. But all these methods rely on the basic assumption that the players have knowledge about their surroundings. There might be restricted situations when individuals are barred from accessing confidential information of other players, or individuals simply might not want to access because of the associated cost of procuring such knowledge. Thus, it is imperative to search for strategy update mechanisms which do not rely on information of the neighbours and decide based on some internal metric.

1.1 Relevant Literature

Karandikar et al.[2] proposed a model where individuals in 2 player repeated prisoner's dilemma set expectations for payoffs, which if met, result in players having a higher probability of repeating the actions. The expectation for the payoffs, termed aspiration updated itself with the received payoff depicting the procedure of learning from the past. Posch et al.[9] analyzed the dynamics of the players with the floating aspiration levels with a deterministic win-stay-lose-shift strategy in the entire 2 player game parameter space. The authors mentioned that a major drawback of their model was the absence of stochasticity which is crucial for such learning algorithms. In a model called the Bush-Mosteller reinforcement learning model (1951), Macy and Flache[5] demonstrated that aspirations fixed in a particular interval can result in cooperative behaviour. They also analyzed the response of players with the ability to change their aspirations dictated by the parameter habituation. Slowly changing aspiration due to small values of habituation with a low starting aspiration could sustain cooperative behaviour. A minor modification of the Bush-Mosteller algorithm was proposed by Masuda and Nakamura[7], which allowed a bigger range of habituation values to cooperative dynamics in a 2 player prisoner's dilemma game. An extension of the above study ([12]) in a well mixed population also revealed the importance of the initial aspiration levels for very low habituation values. A more realistic strategy update mechanism heurestic using the PGG framework was propsed by Roca and Helbing[11] and was shown to display cooperative behaviour in a population on a planar lattice with the players having the ability to change location. Here, instead of the aspirations, the players were allowed to hold a decaying memory of their maximum and minimum payoffs obtained in the past. An extra factor, greediness chose the aspiration to be somewhere in between the two extreme payoffs. The study showed that moderately greedy players performed better than the rest and were able to sustain cooperation. Pichler and Shapiro[8] extended this model into a network. In this project I have studied the model put forward by Masuda and Nakamura[7] in detail. The microscopic mechanisms that allow the players to cooperate has been explored. Furthermore, I have shown that allowing players to restructure their environments allow cooperators to dominate even for high values of habituation.

Chapter 2

Model

Agents occupy the nodes in a network, they update their strategies and play public goods games in every round. These agents have access only to their respective payoff histories and are not privy to any information about the network or their neighbours. The agents decide on their future actions based on their past experiences.

2.1 Modified Public Goods Games

Each node represents an agent who plays $(k + 1)$ modified PGG with all her k neighbors, once as a focal player and the remaining k times as a neighbor of a focal player, in each round of a PGG. If an agent decides to cooperate as a focal player, she allocates a fixed amount of money (m) for each one of her connected neighbors, which after getting multiplied by a synergy factor r, is distributed to her neighbours, regardless of her neighbours' actions. A selfish agent does not donate any money to her neighbors. Thus, the payoff of an individual player is:

$$\begin{aligned} \pi_{iC} &= r \times n_{iC} - k_i && i\text{'th player is a cooperator} \\ \pi_{jD} &= r \times n_{jC} && j\text{'th player is a defector} \end{aligned} \tag{2.1}$$

where, n_{iC} denotes the total number of cooperator neighbors and k_i denotes the degree of the i'th player.

Standard PGG and other social dilemmas like Staghunt and Snowdrift have also been considered and are mentioned in A

2.2 Internal Metrics

An individual will implement an action based on the results that she obtained due to her action in the previous interactions. Her expectations might (or might not) be met and accordingly she will opt to repeat (or change) her action in the subsequent interaction. Expectations of each and every player is governed by the payoffs that the agent had received in the past, appropriately balanced by a factor called the habituation(h), also considered to be the learning ability.

$$\begin{aligned} A_i(t) &= (1-h)\ A_i(t-1) \ +\ h\ \Pi_i(t-1) \\ A_i(t) &= (1-h)^{t-1}\ A_i(1) \ +\ h\ \Sigma_{z=1}^{t-2}\ (1-h)^{t-z}\ \Pi_i(z) \ +\ h\ \Pi_i(t-1) \end{aligned} \tag{2.2}$$

$A_i(t)$, the aspired payoff of the i'th individual at time t dictates an agent's payoff expectation at time t. High values of h will result in the player expecting her payoff to be very close to the one that she obtained in the last round, whereas, low values of h imply that a player is slower to react and consider payoffs beyond their previous payoff before expecting for the future benefits.
The players obtain their respective payoffs after a round of PGG in the network, which when compared to their respective aspired payoffs, determine if they want to continue with their response in the last round. The satisfaction of the individuals is captured by the following form,

$$s_i(t) \ =\ \tanh(\ \beta\ (\Pi_i(t)\ -\ A_i(t))) \tag{2.3}$$

β is the measure of the ith player's sensitivity to the difference in the aspired and the received payoff. The tanh function restricts the satisfaction, $s_i(t) \in (0,1)$. If the payoffs obtained are larger than what was aspired, then the player has a positive value of satisfaction, otherwise the player is dissatisfied. The satisfaction of each player determines if repeating the last action will be fruitful or changing it will be beneficial i.e. the player exhibits a probabilistic win-stay-lose-shift strategy.

$$p_i(t+1) = \begin{cases} p_i(t) + (1-p_i(t))\ s_i(t) & act_i(\mathrm{t}) = C \ \text{ and } s_i(t) \geq 0, \\ p_i(t) + p_i(t)\ s_i(t) & act_i(\mathrm{t}) = C \ \text{ and } s_i(t) < 0, \\ p_i(t) - p_i(t)\ s_i(t) & act_i(\mathrm{t}) = D \ \text{ and } s_i(t) \geq 0, \\ p_i(t) - (1-p_i(t))\ s_i(t) & act_i(\mathrm{t}) = D \ \text{ and } s_i(t) < 0 \end{cases} \tag{2.4}$$

where $p_i(t+1)$ gives the probability of cooperation of agent i in round $t+1$ and $act_i(t)$ is the strategy employed by the i'th agent at time t. Every player is susceptible to make an error with a probability called the implementation error ($\epsilon = 0.02$). This accounts for the trembling hand effect which is an additional stochasticity to the system . A seperate mechanism has also been considered to account for the additional stochasticity and has been discussed in the chapter 3.3. It is evident from 2.4 that very sensitive players swap their actions whenever their payoffs are lower than their aspirations. Whereas, players with low β are incapable of affecting their probability

of cooperation substantially and fail to react fast enough.
The simulations consider both habituation (h) and sensitivity (β) to be population wide features i.e. everyone learns and responds with a similar intensity.

2.3 Rewiring

The decision heurestic is also observed for a network where individuals are capable of changing their connections i.e. a dynamic network. The dynamic nature is brought about by the agents deciding to create and break ties with other agents in the system based on certain rules. Population with sensitive players can have players switching between defection and cooperation multiple times in a short interval because of the form of aspiration (2.2) and probability of cooperation (2.4). Thus, to have a considerable effect on the behaviour of the system due to rewiring, the agents are also required to make and break ties with high frequencies, which is accomplished by letting re fraction of all possible pairs of players ($\binom{n}{2}$) in the network to change their status. Two seperate algorithms are used to carry out rewiring in the population.

High Information Rewiring based on the Action of the Players

In this rewiring formalism, an agent is assumed to have gained access to the information about her neighbours' action in the last game after she fixes her probability of cooperation for the next round. This allows a preferential rewiring scheme (adapted from [3]) to be implemented favouring the creation and maintenance of connection with cooperators. Links are retained based on the action of a randomly chosen participant of the link. The link is kept with a probability $p_r (= 0.87)$ or $p_b (= 0.3)$ if cooperation or defection was the response of the selected player in the last round respectively. If the selected pair of nodes are not connected, then the probability of establishing a connection depends on both the players. A link is established with a high probability $p_m (= 0.93)$ if both had cooperated in the previous round. If either or both of the nodes had defected in the last round, a new link can still be established with probability $p_e (= 0.3)$ and $p_s (= 0.2)$ respectively.

Low Information Rewiring based on the Satisfaction of the Players

A satisfied player will react differently than an unsatisfied player. The constraint of restricted access to information of the neighbours do not stop a player from knowing if another player is linked to her or not. A player can decide the fate of her connections based on these two factors. Every pair of players ($\binom{N}{2}$) have a probability (re) to be rewired for the next round. The two participating players have individual probabilities of wanting to retain/create a link based on their satisfaction after the last game. This respects the low information constraint while allowing pairs to modify ties sensibly. The willingness of the two participating players wanting to make ties are indpendent of each other and is denoted as $p_{linked?,satisfied?}$. For example, $p_{1,0}$ is the probability that an unsatisfied player agrees to preserve the connection that was present in the previous round, whereas, $p_{0,1}$ is the probability that a satisfied player is inclined towards creating the connection. The probability of a link being created is dependent on the willingness of the two participating players. Thus, the probability that a connection will be present for the next round is a product of the willingness of the two participating players.

2.4 Update Mechanism

Each round starts with the agents setting their aspirations for the upcoming games using 2.2. The players then engage in the public goods games with their neighbours and receive their payoffs according to 2.1. The players compare their received payoffs with the aspired ones (2.3) and modify their probability of cooperation for the next round based on their satisfaction using 2.4. This marks the end of a round and the players start the next round by evaluating their aspiration. In case of dynamic networks, players can be allowed to rewire anytime after the games are played, because the strategy update mechanism does not require information about neighbours, and the high information rewiring demands the action of a player and her neighbours in the last game.

Initialization

The underlying structure of the population consists of 500 individuals as the nodes in an Erdős Rényi network with a 0.3 probability of link formation. Each individual is assigned a probability of half to cooperate in the first round, so that their initial response can be considered as random. The aspiration for the first round of public goods games is set to $\frac{k_i \times m \times (r-1)}{2}$ where, k_i is the degree of the player and m is the unit of transaction which is set to 1 (refer to B).

Categorization

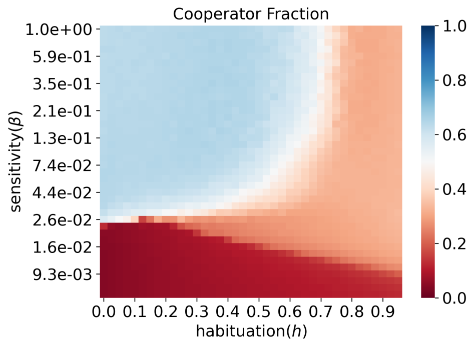
An obvious distinction between the players in the network at the end of every round is marked by their action in the round. A further categorization based on the satisfaction coupled with the action is helpful in revealing the detailed behaviour of the model. A player choosing to cooperate (defect) in the next round based on her satisfaction in the current round can be classified into either satisfied cooperator, SC (satisfied defector, SD) or an unsatisfied cooperator, UC (unsatisfied defector, UD).

Chapter 3

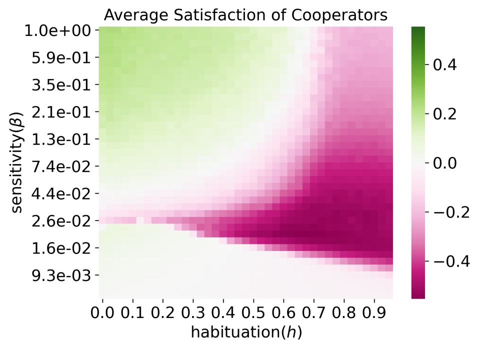
Behaviour in Static Network

The model consists of two parameters which are considered to be globally and temporally static over the entire simulation - habituation (h) and sensitivity (β).

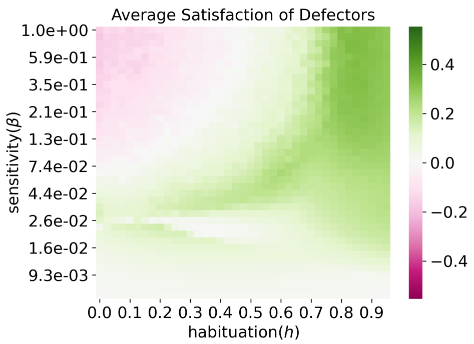
3.1 Dependence on Habituation



(a)



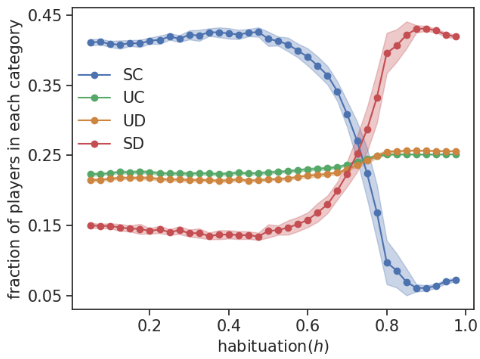
(b)



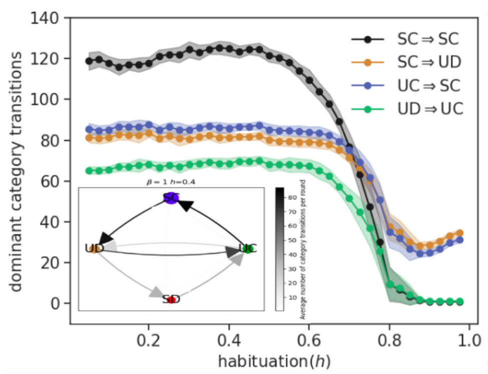
(c)

FIGURE 3.1: Heatmap showing the (a) cooperator fraction; average satisfaction of (b) cooperators and (c) defectors as a function of habituation (h) and sensitivity (β) for $N = 500$. Each pixel value is obtained by first averaging from round 250 to 750 for each trial and then averaged over 5 trials.

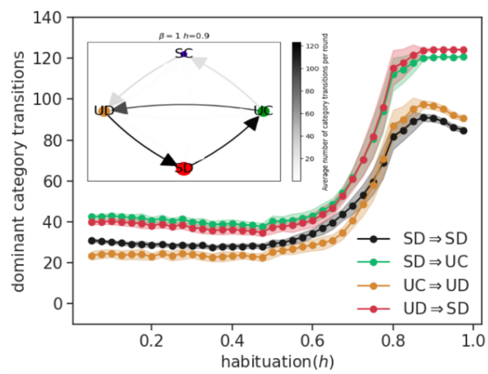
The heatmap plots show the existence of 3 different types of dynamics that occur in the network when the two parameters are varied. The players in the network display substantial cooperative behaviour (see 3.1(a)) only when the learning rate (h) is below a certain critical threshold (h_c) and the players are sensitive enough to the difference in their received and aspired payoff. Habituation values higher than the h_c are not able to sustain cooperation due to the lack of memory of the past payoffs while determining the aspirations. The cooperator satisfaction heatmap in 3.1(b) shows that satisfied slow learners opt to cooperate which in turn leads to a positive satisfaction. Whereas, fast learning cooperators get acquainted to the payoffs very fast and shift their strategy.



(a)



(b)



(c)

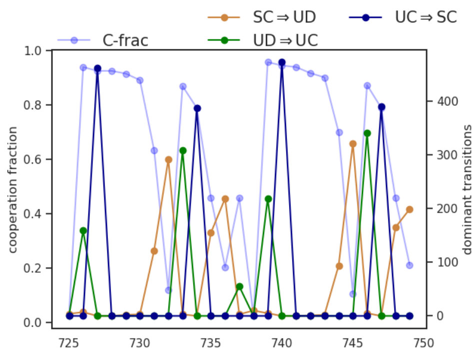
FIGURE 3.2: (a) Change in fraction of players in each category as a funciton of habituation (h). Major transitions for population with (b) low habituation and (c) high habituation. All the simulations are time averaged from round 500 to 1000 over 25 trials. The shaded region denotes one sigma deviation from the mean. The inset shows the dominant cycles through the categories for (b) $h = 0.4$ and (c) $h = 0.9$. The inset diagrams can be a real asset for visualising the dynamics that is described in the next section.

Very low values of β make the players insensitive to the offset of aspiration and received payoff (see 3.1(b) 3.1(c)). Thus, the probability of cooperation does not get modified sufficiently and the system saturate towards a defective dynamics, explained in B. Thus, $\beta = 1$ is considered in the rest of the text unless mentioned otherwise. The dynamics of the system can be appreciated better if they are looked through the lens of the categorization that is discussed in 2.4. Satisfied cooperators (SC) dominate when the population consists of slow learners, whereas satisfied defectors (SD) dominate in populations with players having high values of habituation as can be seen in 3.2(a). The transition of players through the different categories reveal how the category fractions are established. Low values of habituation display high $SC \rightarrow SC$ transitions (see 3.2(b)). $SC \rightarrow UD$ and $UC \rightarrow SC$ are the next dominant transitions followed closely by $UD \rightarrow UC$. This reveals the inter-category cycle, $SC \rightarrow UD \rightarrow UC \rightarrow SC$, as the second most dominant cyclic transition. At high habituation values the $SD \rightarrow UC$ and $UD \rightarrow SD$ transitions are much higher compared to the SD self loops (see 3.2(c)). The inter-category loop $SD \rightarrow UC \rightarrow UD \rightarrow SD$ is the most dominant cycle followed by the $SD \rightarrow SD$ self transitions.

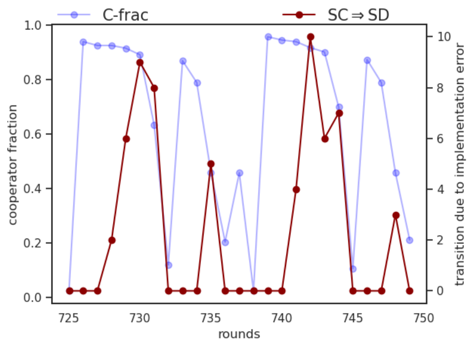
3.1.1 Cooperative Phase

The averaged out transition rates between the categories in 3.2(b) show that $SC \rightarrow SC$ self transitions are the most dominant transitions in the cooperative regime, reinforcing the result of positive average satisfaction of cooperators in 3.1(b). The inter-category transition is a by-product of cooperator collapse at regular intervals. There are two factors responsible for the collapse of cooperation at regular intervals. The obvious reason is the aspiration nearing the payoff for all the SC individuals. After a few rounds of cooperation the aspiration of the players are very close to the received payoffs (see 3.3(c). This is where the second factor, implementation error, comes into play. Approximately $N\epsilon$ number of players misimplement their actions in every round. If a cooperator in a cooperative neighbourhood defects for once due to the trembling hand (like the players in 3.3(b)), then she will not revert back to cooperation as long as her neighbours do not change strategy, because defection yields a higher payoff than cooperators in the same neighbourhood. These two effects push the system towards a tipping point characterized by high aspirations but reducing payoffs resulting in the collapse of cooperation (3.3(d)). As soon as a cooperative player with aspiration close to her payoff experiences one of her cooperative neighbour shift strategy, she also follows suit after getting unsatisfied. This produces a chain reaction of $SC \rightarrow UD$ transitions in the network. This further reduces the payoff in the network resulting in further dissatisfaction and strategy shift to UC. A synchronized cooperation by the majority of the players yield a high payoff for everyone, which is substantially higher than the new lowered aspiration level of all the inividual players. This restarts the

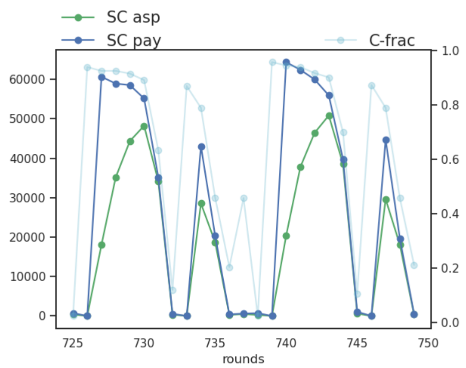
SC self transitions in the system completing the cycle $SC \rightarrow UD \rightarrow UC \rightarrow SC$ as can be seen in 3.3(a). The inter-category cycle assist the players in reducing their aspirations, such that satisfied cooperation can restart in the network. The cyclic transition is seldom perfect because all the players in the network are rarely in perfect synchrony. The desynchronization between the players are brought about by some of the participants being satisfied after defecting and turning into SDs, such as the players who misimplemented during the cooperative interval. Repeated transitions between UC and UD with the help of SD lower the aspiration substantially which helps regain the lost synchrony, and finally the majority of the players cooperate and turn to SC.



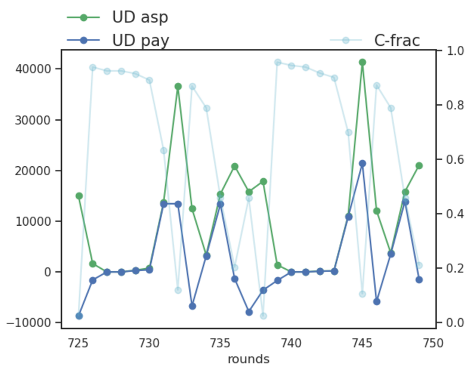
(a)



(b)



(c)

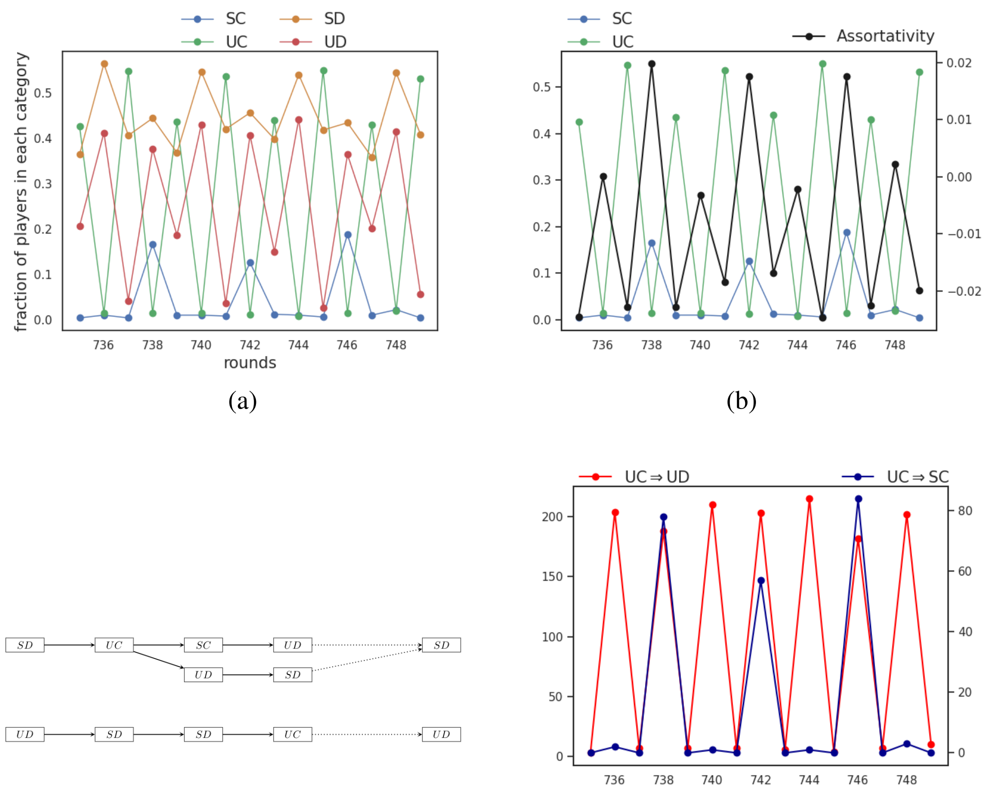


(d)

FIGURE 3.3: (a) Time evolution of dominant transitions and the cooperator fraction. (b) Key transitions from $SC \rightarrow SD$. The payoff and the aspiration which give rise to (c) SC and (d) UD. All the plots are obtained from a single realization with $h = 0.4$ and $\beta = 1$.

3.1.2 Defective Phase

Large habituation values imply that the aspiration of the players change rapidly by adjusting their values close to the last received payoff. This when coupled with the implementation error of the players result in periodic cycling between the categories



(a) (b)

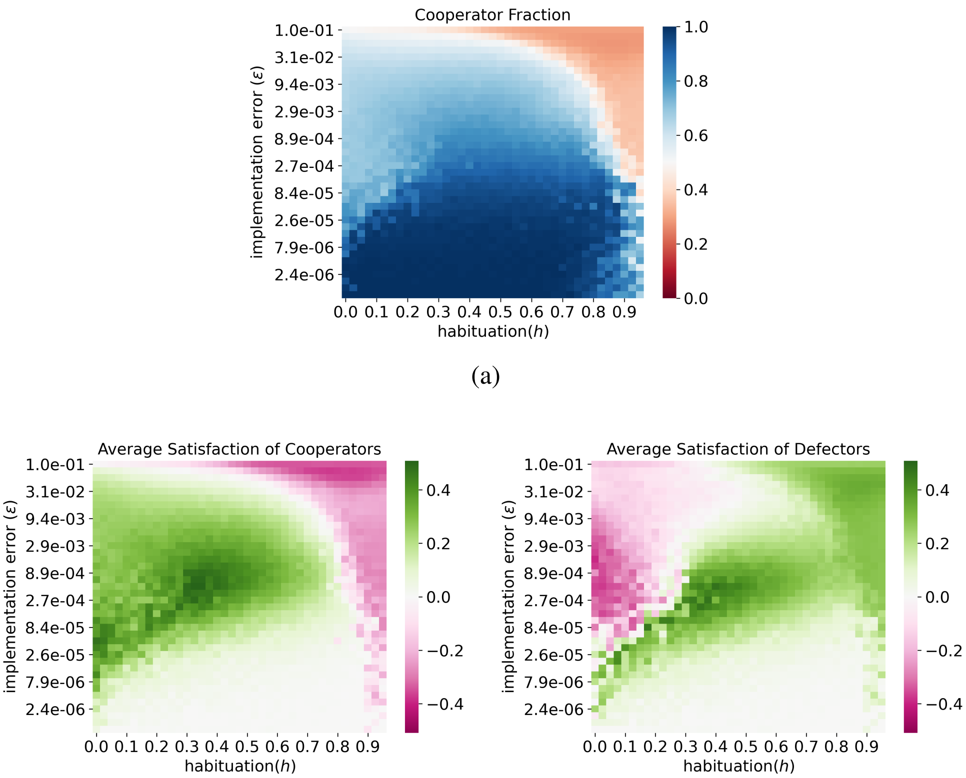
(c) (d)

FIGURE 3.4: (a) Time evolution of fraction of players in each category. (b) Flow diagram of the category transitions. (c) Time evolution of assortativity in the network measured by labeling the nodes as cooperators and defectors. (d) The bifurcation in the transitions giving rise to a 3-cycle transtion and a 4-cycle transition. All plots are obtained from a single realization with $h = 0.9$ and $\beta = 1$

as seen in 3.4(a). The dominant cycle is $SD \rightarrow UC \rightarrow UD \rightarrow SD$. Another cycle with SC as a component ($SD \rightarrow UC \rightarrow SC \rightarrow UD \rightarrow SD$) activates when the large transition from $SD \rightarrow UC$ produce regions in the network with an abundance of cooperators. This results in some of the cooperators to have enough cooperator neighbours such that they become satisfied. In response, these players opt to cooperate in the next round i.e. $UC \rightarrow SC$, while the other UCs select to defect in the next round, and follow the 3 member cycle (3.4(d)). Comparitively higher homophily (measured with respect to C and D,) in the network can be seen when players occupy the SC category, whereas, negative assortativity is seen when players are UCs (3.4(b)). This hints at cooperative domains in the network produced in the initial $SD \rightarrow UC$ transition. The UCs which are exploited more are the ones responsible for lowering the homophily, shift to UD. While, the UCs who had higher number of cooperator neighbours became satsified and repeat the action. The transition of $UC \rightarrow UD$ lower the payoff of some SDs making them unsatisfied, in response to which they cooperate in the next round. In the same round, players in the secondary cycle transition from $SC \rightarrow UD$, after realising that the supporting cooperators betrayed in the last round. The new UCs still being unsatisfied choose to defect, following the 3 member cycle for the second time. This dissatisfies neighbouring SDs and the whole 4 member cycle restarts from the next round. Thus, even though the averaged out dynamics show that the dominant transition is a 3 member cycle, the periodicity of the dynamics in this phase is demarcated by the weaker 4 member cycle. The strength of the 3 member cycle is much stronger because it is activated after every 2 rounds. A schematic diagram of the two cycles is displayed in 3.4(c).

3.2 Necessity and Effect of Mistakes

An obvious question to ask is - what is the reason for having a trembling hand effect (ϵ), if players use a "probability" of cooperation to arrive at their decisions?



(b) (c)

FIGURE 3.5: Heatmap of (a) cooperator fraction (b) average satisfaction of cooperators and (c) average satisfaction of defectors as a function of implementation error (ϵ) and habituation (h)

Consider an instant of time when all the faultless players in a network are satisfied with their actions i.e. the payoffs are higher than their aspiration. In the absence of errors, the aspiration assymptotically reaches the payoff, $A_i \longrightarrow \Pi_i^- \; \forall i$. The payoff received by individual players do not change in time because all the players are satisfied and do not make errors. This arrests the system in a local minima where, $s_i \longrightarrow 0^+ \; \forall i$. Such a stagnation is evaded if the players make mistakes ($\epsilon > 0$) or if the players have noisy aspirations (discussed in 3.3). The heatmaps in 3.5 show that the critical habituation, h_c, increases as the error rate reduces. Very low error rates correspond to vanishingly small satisfaction of both cooperators and defectors. As discussed above, satisfaction close to 0 for prolonged periods of time correspond to stagnant configurations. Thus, at very low error rates the system keeps halting at stable configurations, and modify only when a player makes an error, after which it again stabilizes in some other configuration.

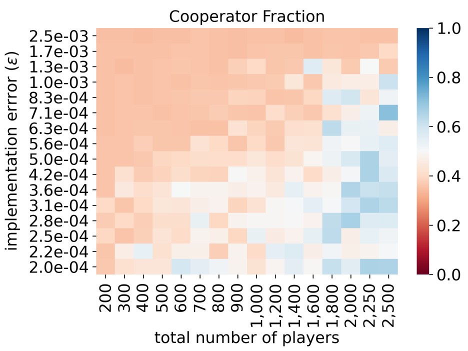


FIGURE 3.6: Heatmap of the cooperator fraction as a function of the implementation error (ϵ) and total number of players in the network with $h = 0.9$

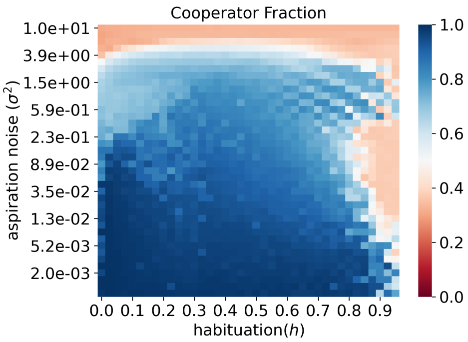
It is also interesting to note that below the error rate threshold the defective dynamics cease to exist, hinting at the importance of mistakes driving the cyclic dynamics at high habituation values. It is straightforward to guess that a network with different number of players in the network will require a different error rate to arrive at a similar dynamics. Considering the presence of defection at high habituation as a marker for a comparable dynamics in networks of different sizes, the heatmap 3.6 show that bigger networks require higher error rates to unlock phase 2 cyclic dynamics. This tells us that the defective dynamics require misimplementation of action by several players well distributed across the network.

3.3 Replacing Mistakes by Noisy Aspiration

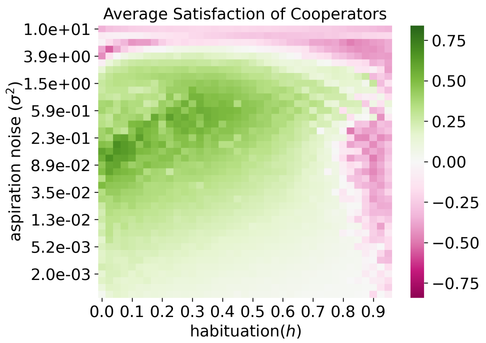
While calculating the aspiration the players are assumed to make miscalculations. This miscalculation is modelled using random numbers generated from a normal distribution. The standard deviation of the distribution act as the measure of the magnitude of error the players make. The modified aspiration is of the form:

$$A_i(t) = (1 - h)A_i(t-1) + h\Pi_i(t-1) + \mathcal{N}(0, \sigma^2)$$

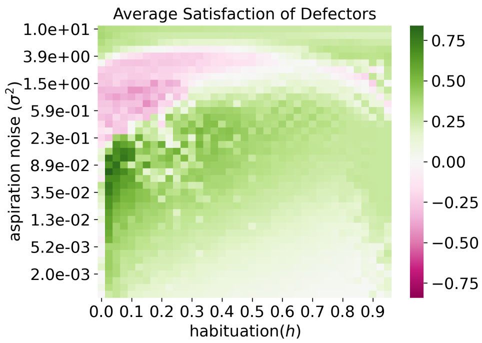
Thus, aspiration with noise takes the role of additional stochasticity in the system instead of implementation error. Significant number of players start becoming unsatisfied whenever the difference between the aspiration and the payoff reach an order of the standard deviation. Heatmaps of the cooperator fraction and the satisfaction of the players are shown to demonstrate the similarity between the dynamics generated by the two different sources of stochasticity (compare 3.7 with 3.5) .



(a)



(b)



(c)

FIGURE 3.7: Heatmap of (a) cooperator fraction (b) average satisfaction of cooperators and (c) average satisfaction of defectors as a function of aspiration noise (σ^2) and habituation (h)

However, there are subtle differences in the dynamics because of the two different methods of incorporating stochasticity. Implementation error has an uniform distribution of errors over time, whereas, aspiration noise affect the actions of players more when players have their aspiration close enough to the payoff. These subtle differences can be seen to amplify in the heatmaps (3.7) at high noise/error rates and extreme habituation values.

Cooperative Phase

The time average abundance of categories compared for different values of habituation for a sensible aspiration noise ($\sigma^2 = 0.1$, such that both the phases are captured) reveal

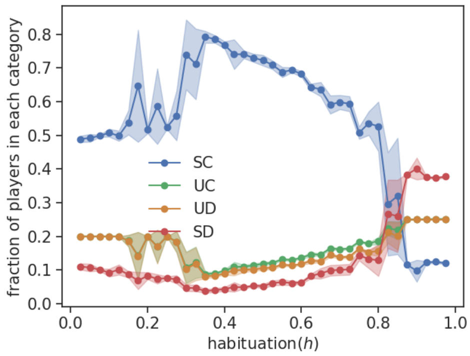


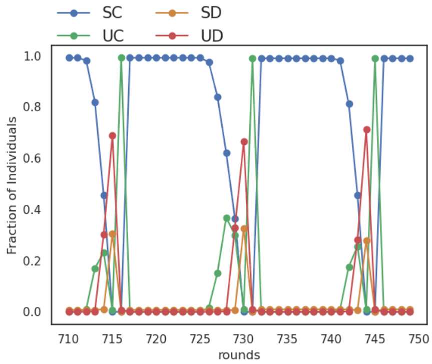
FIGURE 3.8: Change in fraction of players in each category with habituation (h). All the simulations are time averaged from round 1000 to 1500 over 25 trials. The shaded region denotes one sigma deviation from the mean. The aspiration noise is taken to be $\sigma^2 = 0.1$

a somewhat similar trend (3.8). The cooperator fraction for low and moderate values of habiutation are higher. This is because, unlike implementation error, players do not change their strategies when their aspirations are far below their payoffs. So the only factor responsible for the periodic collapse of cooperation is the slowly rising noisy aspiration. This lets the players cooperate for longer time periods 3.9(a)

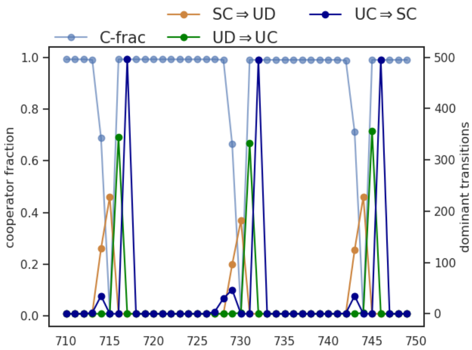
Another key difference is the mechanism for periodic collapse of cooperation at moderate habituation. In case of implementation error, the occurence of $SC \rightarrow SD$ transition was crucial for the collapse (see 3.3(b)). But for noisy aspirations, SC players start becoming unsatisfied with low absolute values, leading them to UC (3.9(c)). These small values of dissatisfaction slowly reduce the probability of cooperation of UCs until one of them decides to defect. At this juncture the aspiration of all the players are almost equal to their payoffs, so as soon as one of the UC/SC player decides to flip her action to defection (UD), cooperation collapses. The usual cycle that reset the players' aspirations takes effect in this scenario as well but spread out in time (3.9(b)). The reduction in the cooperator fraction for low values of the habituation in 3.8 is explained in C.

Defective Phase

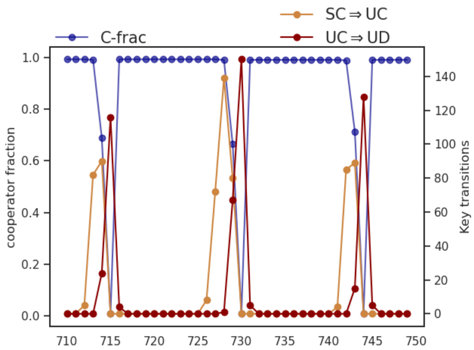
The network displays a similar dynamic equilibrium for habituation values above the critical habituation. The dynamic is more uniform (see 3.10(a)) compared to its cousin, with implementation error as the stochasticity, showed in 3.4(a). A major difference is the dominance of $UC \rightarrow SC$ over $UC \rightarrow UD$ during the bifurcation. But, the 3 member cycle catches up, thanks to its re-activation after 2 rounds because of the peripheral UCs shift to defection (3.10(b)).



(a)

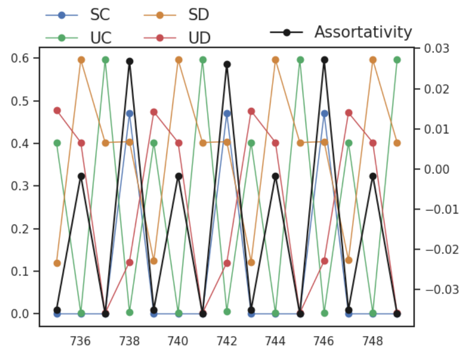


(b)

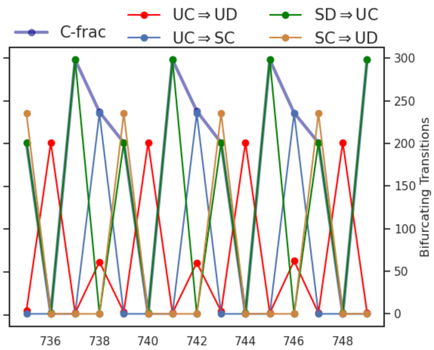


(c)

FIGURE 3.9: (a) Time evolution of dominant transitions and the cooperator fraction. (b) The dominant transitions that restore cooperation by lowering the aspirations. (c) Key transition $SC \rightarrow UC$ that introduces defection into the marginal cooperative network. All the plots are obtained from a single realization with $h = 0.5$



(a)



(b)

FIGURE 3.10: (a) Time evolution of fraction of players in each category and the evolution of assortativity in the network measured by labeling the nodes as cooperators and defectors. (b) The dominant transitions driving the 4 period cycle at high habituation values. All plots areobtained from a single realization with $h = 0.9$

Chapter 4

Behaviour in Dynamic Network

Allowing players to modify links increases the complexity of dynamics, but under certain favourable conditions can also bolster the establishment of cooperation.

4.1 High information rewiring

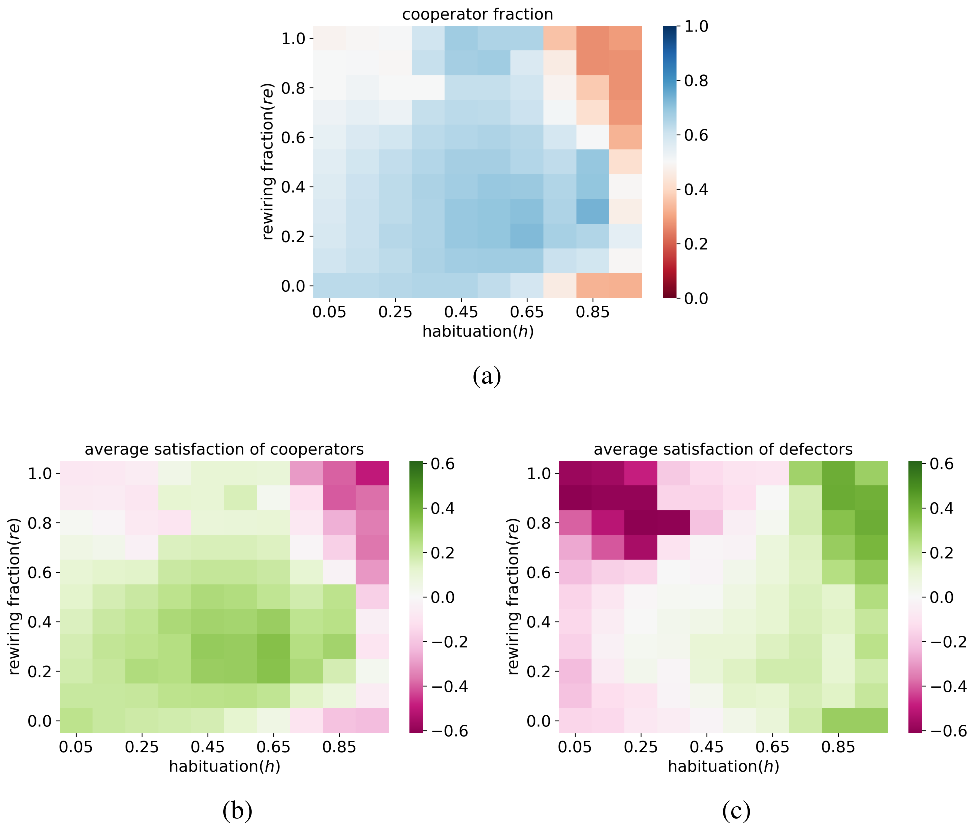


FIGURE 4.1: Heatmaps showing the (a) cooperator fraction; average satisfaction of (b) cooperators and (c) defectors as a function of habituation (h) and rewiring probability (re) for $N = 500$. Each pixel value is obtained by first averaging from round 250 to 750 for each trial and then averaging over 5 trials.

This rewiring procedure allow cooperators to make/retain bonds with a higher probability compared to defectors as described in 2.3. Cooperators dominate the population even for high values of habituation if the players are allowed to rewire at a reasonable rate($re < 0.5$) as can be seen in 4.1(a).

Low Levels of Rewiring

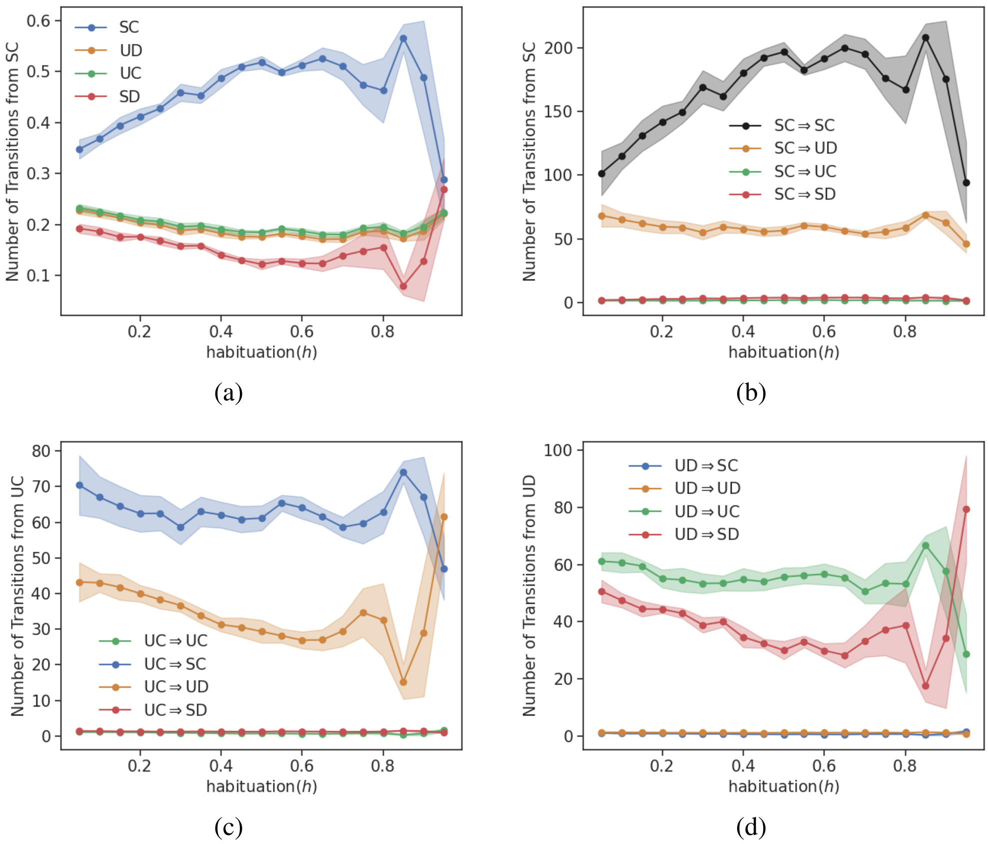


FIGURE 4.2: (a) Variation of time averaged abundance of categories with habituation. Category transitions from (b) SC, (c) UD and (d) UC with $re = 0.3$.

Rewiring help cooperators make more ties with other cooperators increasing the payoffs and subsequently making them more satisfied which can be seen in 4.1(b). The low levels of rewiring ensure that defectors do not exploit cooperative players by forming connections at the same rate as cooperators themselves. High satisfaction values leading to cooperation hints at $SC \rightarrow SC$ self transitions which can be seen in 4.2(b). The aspiration replenishing cycle, $SC \rightarrow UD \rightarrow UC \rightarrow SC$ remains the second most dominant cycle, although, $UC \rightarrow UD$ transitions and SD indirectly help in synchronizing the players lowering the aspirations enough to restart the SC

self transitions. The SC self transitions are remarkably high at moderate habituation compared to the static network because of the additional factor of $C-C$ link formation.

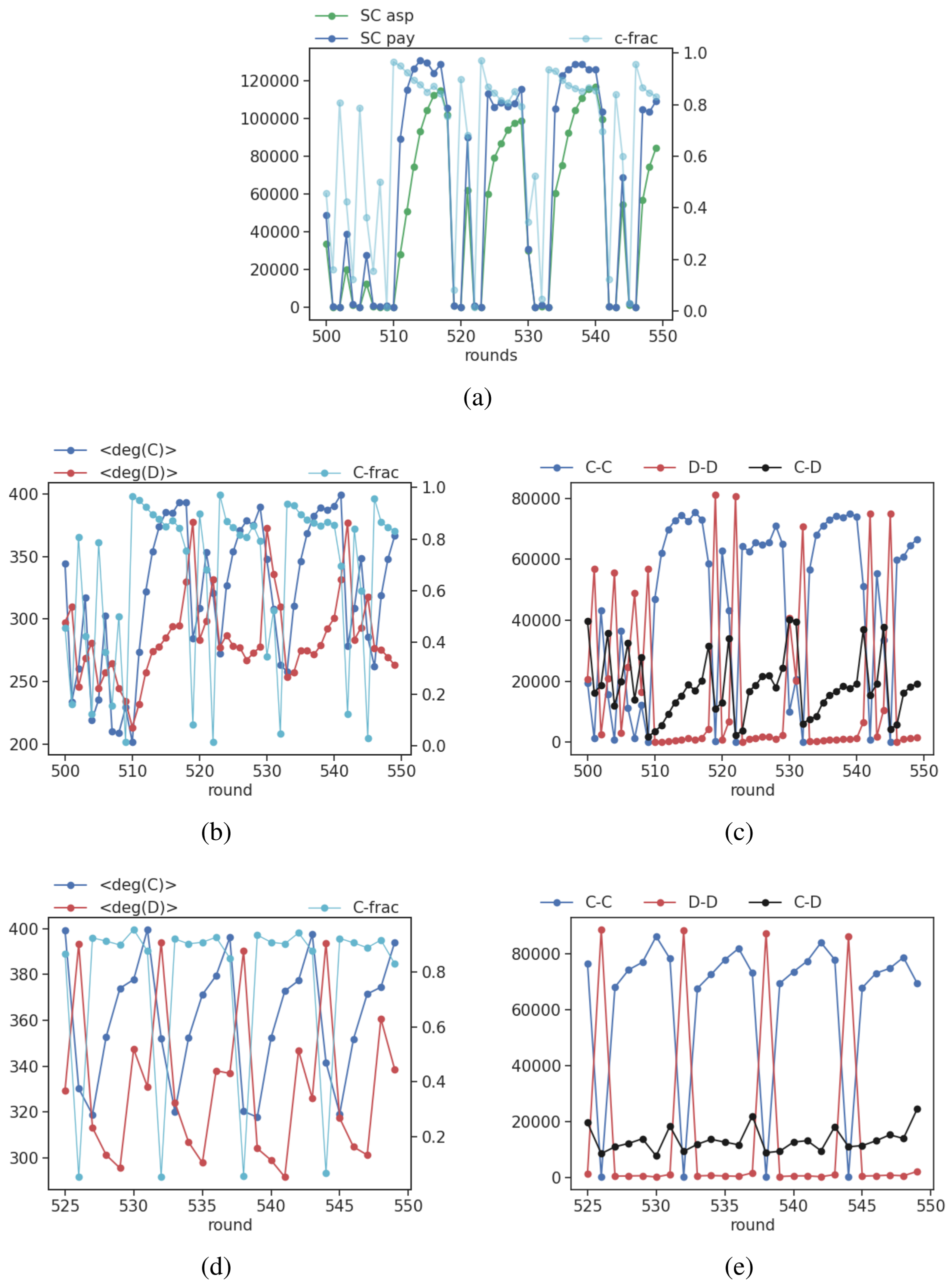


FIGURE 4.3: (a) Slow rise of aspiration of the SCs at $h = 0.4$(b) The average degree of cooperators increase rapidly followed by a slow increase of the defector average degree, which rapidly increase the (c) $C-C$ links and slowly increases $C-D$ links in the network, finally resulting in the collapse of cooperation at $h = 0.4$. The (d) increase in the cooperator degree and (e) the $C-C$ links are the only mechanisms that enable cooperation at $h = 0.9$. (a), (b) and (c) are obtained from a single realization with $re = 0.3$. (d) and (e) are also obtained from a single realization with $re = 0.3$.

Surprisingly, for low values of habituation, the population exhibits low levels of

SC self transitions. On inspecting the temporal behaviour of the players through the categories in 4.4, it reveals that the players display sustained cooperation less often. A different dynamics is observed to be at play, which is driven mainly due to segregation of the population into two groups. One of the group has players with high aspiration, while the other group has players with achievable aspirations. This dynamic equilibrium is more prominent at high values of rewiring and is discussed in the next subsection (4.7(d)).

An explanation for the cooperative behaviour at high habituation values can be found in the temporal evolution of $C-C$, $C-D$ links and the average degree of cooperators and defectors (see 4.3(d) and 4.3(e)). The rise in cooperator degree helps in maintaining cooperation for short intervals, although the aspiration rises rapidly. Irrespective of the value of habituation, high cooperation is correlated with high values of $C-C$ links and the high average degree of cooperators. In case of low habituation, due to the slow aspiration increase (4.3(a)), the cooperative phase is maintained for a longer time period while sustaining an increase in $C-D$ links (4.3(c)).

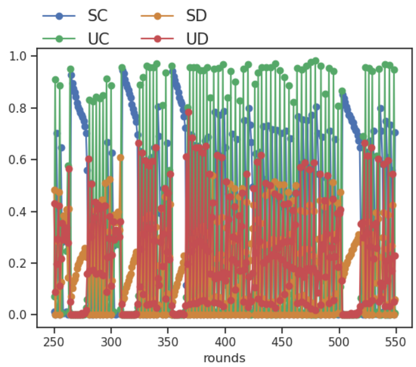


FIGURE 4.4: Temporal dynamics of the occupancy of different categories by the players at $re = 0.3$ and $h = 0.1$

High Levels of Rewiring

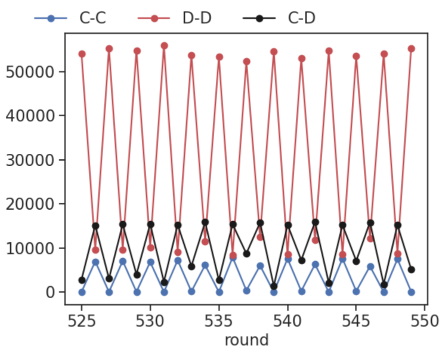
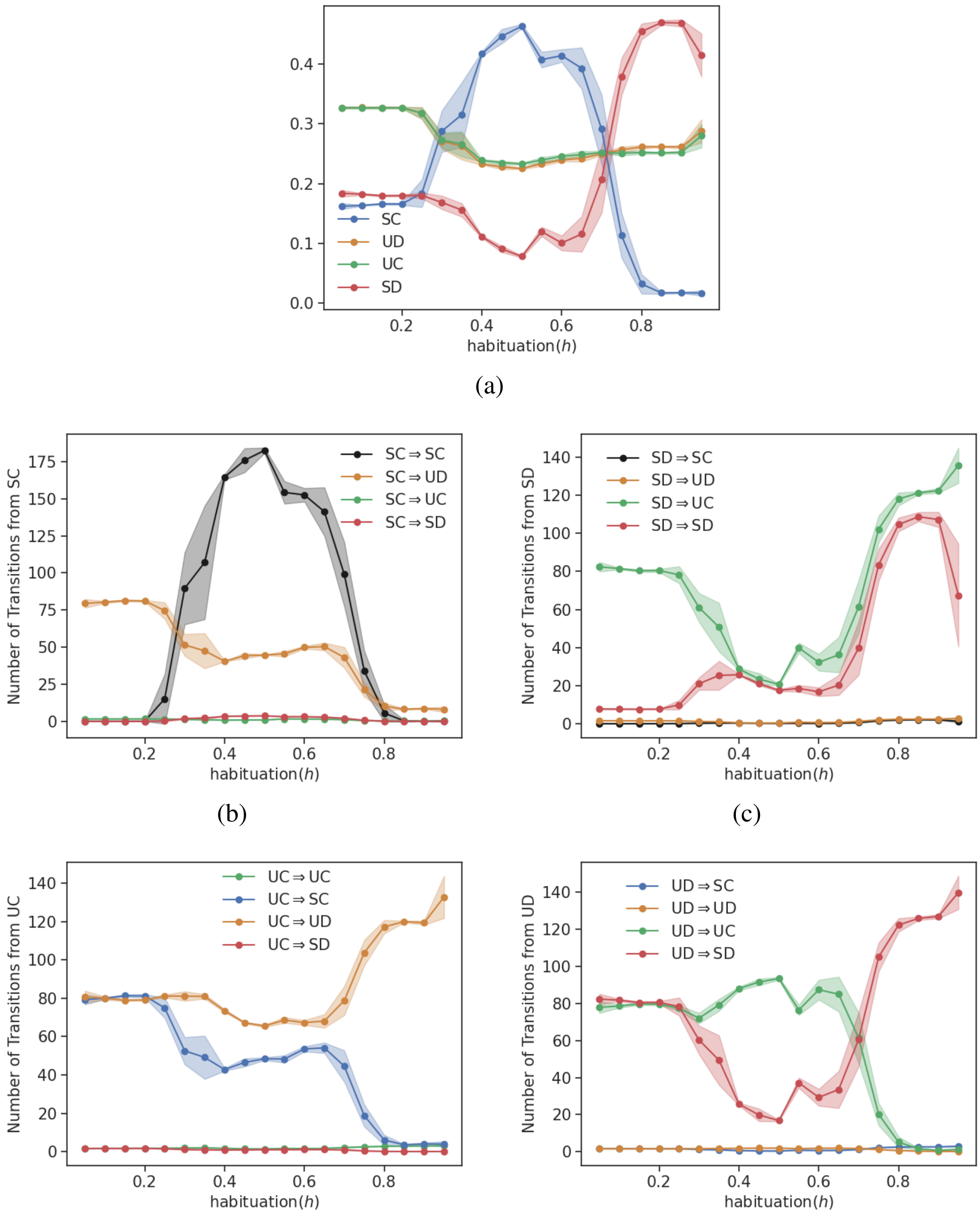


FIGURE 4.5: Temporal dynamics of the links for $h = 0.9$ and $re = 1.0$

At high rewiring rates almost every pair of player is given the opportunity to abstain / make / destroy / retain the link, making the network very unstable. Still, moderate values of habituation are capable of maintaining cooperation (4.6(a)). High habituation value display defection as the dominant action amongst the players because of the presence of $SD \rightarrow UC \rightarrow UD \rightarrow SD$ cycle, followed closely by SD self transitions (4.6). The extensive rewiring of the network allow defectors to take advantage of the cooperators by forming links. Increase in the $C-C$ links is perfectly correlated with an increase in the $C - D$ links (4.5) which reduces the overall payoff of cooperators. Combined with the fast evolving aspiration, the players become unsatisfied and choose to defect in the next round. This dynamics is very similar to the one that was observed for high habituation in static network and implementation error again has a crucial role in sustaining the cyclic dynamics.



(a) (b) (c) (d) (e)

FIGURE 4.6: (a) Fraction of players in each category. Transitions from (b) SC, (c) SD, (d) UC and (e) UD.

4.1.1 Polarized Population

Dynamics at low values of habituation with very high rewiring rates consist of both the dominant inter-category cyclic transitions i.e. $SC \rightarrow UD \rightarrow UC \rightarrow SC$ and $SD \rightarrow UC \rightarrow UD \rightarrow SD$ (4.7(a) and 4.7(b)). The the two cycles share the category of the unsatisfied players (UC and UD). The cycles are synchronized such that all the players occupy UC at the same time (4.7(c)).

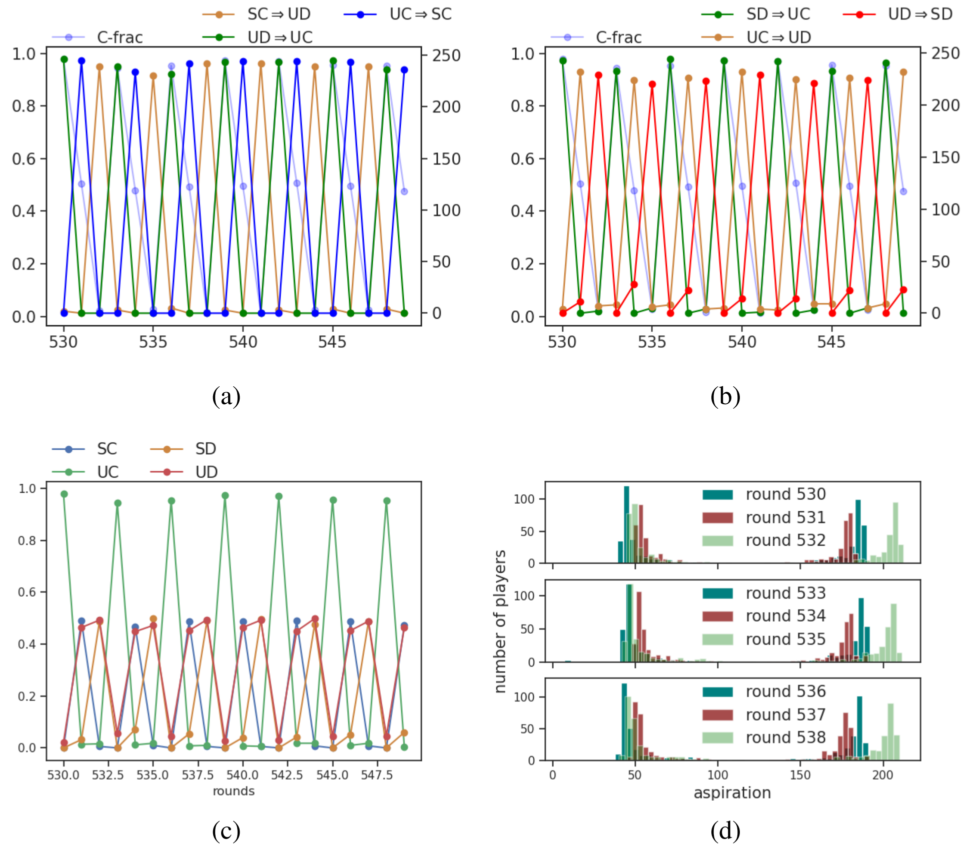


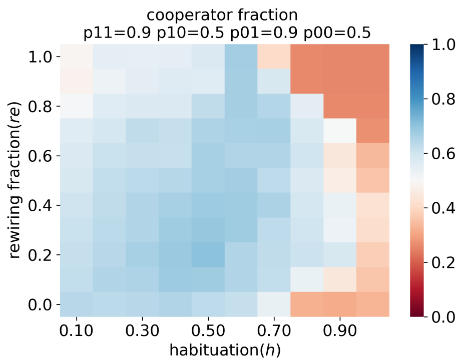
FIGURE 4.7: (a)$SC \rightarrow UD \rightarrow UC \rightarrow SC$ cycle and (b) $UD \rightarrow SD \rightarrow UC \rightarrow UD$ cycle. (c) Temporal evolution of fraction of players in each category. (d) Histogram of the aspiration of all the players from round 530 to round 538. All the plots are obtained from a single realization with parameters $re = 1.0$ and $h = 0.1$

In the next instant of time, half of the population occupy SC when the other half occupy UD. Followed by SCs transitioning to UD and the UDs in the previous round transtioning to SD. Finally, they meet again at UC. This equilibrium dynamics hints at two distinct group of players (4.7(d)). One group of players have realistic aspirations and becomes satisfied once everyone cooperates at UC making them transition to SC in the next round. The apirations for the players of the other group can be achieved only if they are surrounded by cooperators while being a defector. The presence of SC allowed the UD players to become satisfied while sticking to defection. The low value of habituation is incapable of drastically changing the aspirations, leaving the system in the above dynamics.

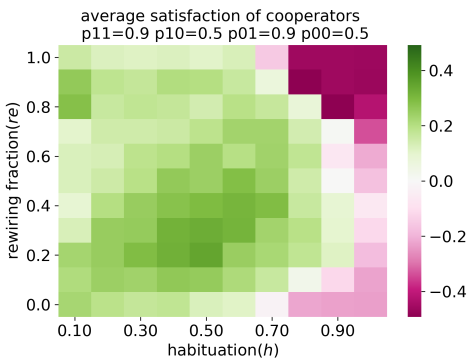
The evolution of the system is diffiult to explain in equal detail over the entire paramater space, since it combines the effects of rewiring as well as strategy shifts in the network. The 3 types of dynamics that the system exhibit coexist at the overlapping parameter values. Still, it is exciting to see cooperators being favoured, on an average, over the entire habituation range for moderate rewiring rates.

4.2 Low Information Rewiring

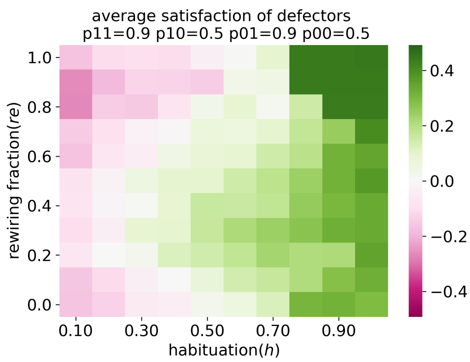
$p_{11} = 0.9, p_{10} = 0.5, p_{01} = 0.9, p_{00} = 0.5$



(a)



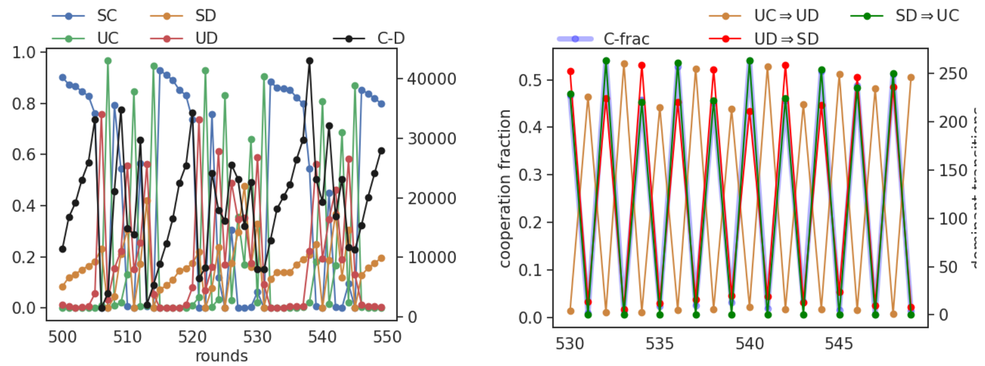
(b)



(c)

FIGURE 4.8: Heat map showing the (a) cooperator fraction; average satisfaction of (b) cooperators and (c) defectors as a function of habituation (h) and rewiring probability (re) for $N = 500$. Each pixel value is obtained by first averaging from round 250 to 750 for each trial and then averaging over over 5 trials. The willingness probabilities are $p_{11} = 0.9, p_{10} = 0.5, p_{01} = 0.9, p_{00} = 0.5$.

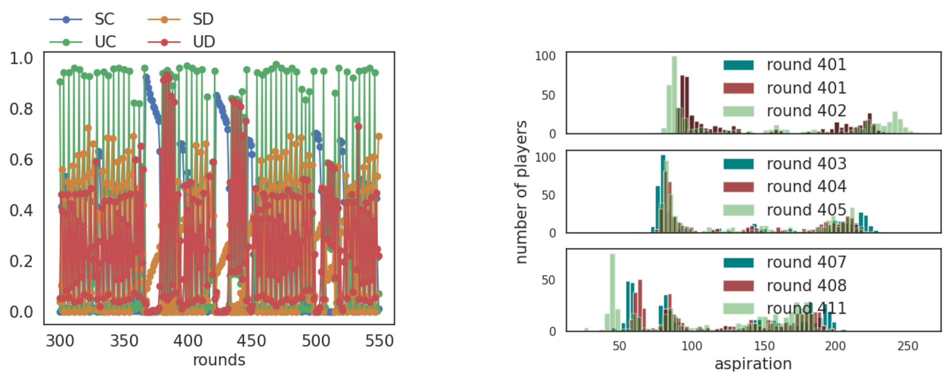
The heatmap 4.8(a) shows that rewiring enable cooperation for populations with higher values of habituation, provided satisfied players have a higher propensity to



(a) (b)

FIGURE 4.9: Temporal dynamics of the (a) occupancy of different categories by the players and C-D links at $re = 0.3$ & $h = 0.4$ and (b) cooperator fraction and major transitions at $re = 1.0$ & $h = 0.9$.

make and retain ties than the unsatisfied ones. Similar features of satisfied cooperators opting for cooperation can also be seen in 4.8(b) and 4.8(c). A minor difference is the involvement of SC in the aspiration rejuvenating cycle. If considerable number of defectors (especially satisfied) are present when majority of the players transition from UC to SC, then $C - D$ links form, making the cooperators susceptible to exploitation 4.9(a). This results in an immediate collapse of cooperators to UD. Beside this new cycle, the original $UC \rightarrow UD \rightarrow UC$ repeatitions reset the aspirations to a lower value for satisfied cooperation to restart.

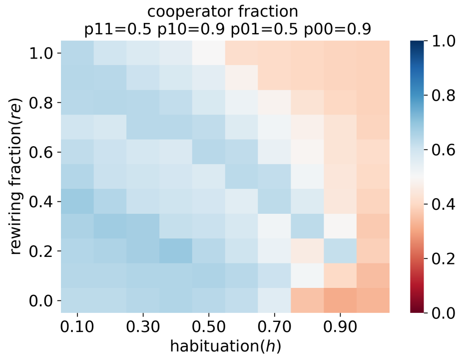


(a) (b)

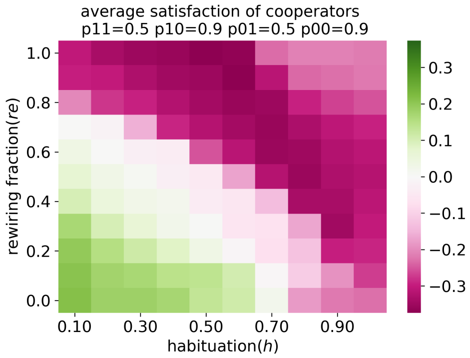
FIGURE 4.10: (a) Temporal dynamics of fraction of players in each category. (b) Histogram of the aspiration of all the players. The plots are obtained in a single realization with $re = 1.0$ and $h = 0.1$

At high habituation values and high rewiring rate the population experiences the similar $SD \rightarrow UC \rightarrow UD \rightarrow SD$ transitions (4.9(b)). Interestingly, populations with low habituation display cooperative behaviour at very long intervals (4.10(a)), similar to what we observed in 4.4. The system has a population with polarized aspiration for prolonged periods of time. But, the group with higher aspiration slowly reduce their aspirations with time4.10(b). As soon as the two aspirations are small and closeby, the players in the network cooperate. The cooperation after the resetted aspiration again polarizes the aspiration of the population.

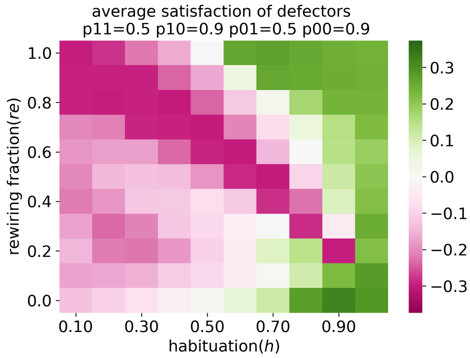
$$p_{11} = 0.5, p_{10} = 0.9, p_{01} = 0.5, p_{00} = 0.9$$



(a)



(b)



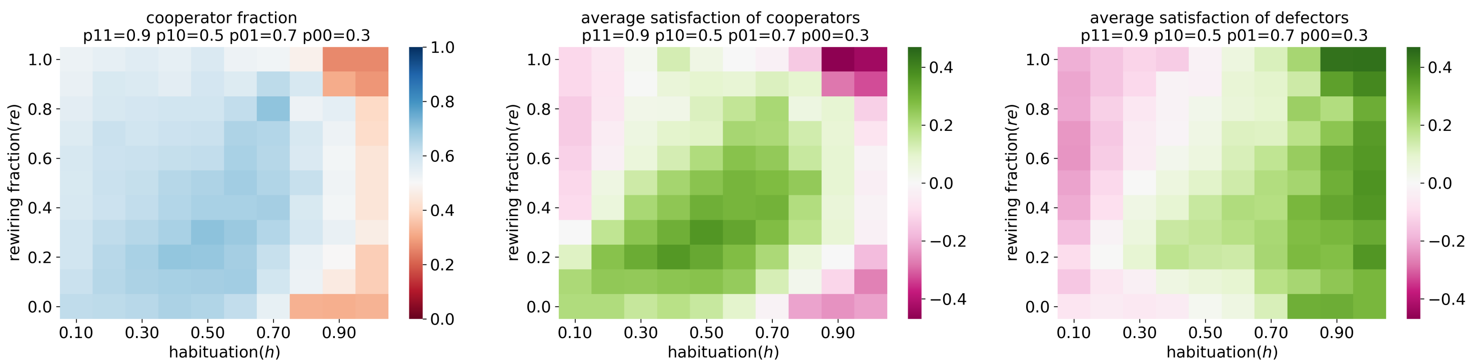
(c)

FIGURE 4.11: Heat map showing the (a) cooperator fraction; average satisfaction of (b) cooperators and (c) defectors as a function of habituation (h) and rewiring probability (re) for $N = 500$. Each pixel value is obtained by first averaging from round 250 to 750 for each trial and then averaging over over 5 trials. The willingness probabilities are $p_{11} = 0.5, p_{10} = 0.9, p_{01} = 0.5, p_{00} = 0.9$.

Reversing the situation i.e. making unsatisfied players more willing to make ties compared to the satisfied players, stop cooperation from proliferating even at moderate rewiring. High rewiring make the chances slimmer for cooperators, with cooperators staying unsatisfied for the entire range of habituation.

Highly connected networks support cooperation for larger ranges of habituation (see A.3). But selectively allowing cooperators or satisfied players to make more ties help the system to attain the same fate.

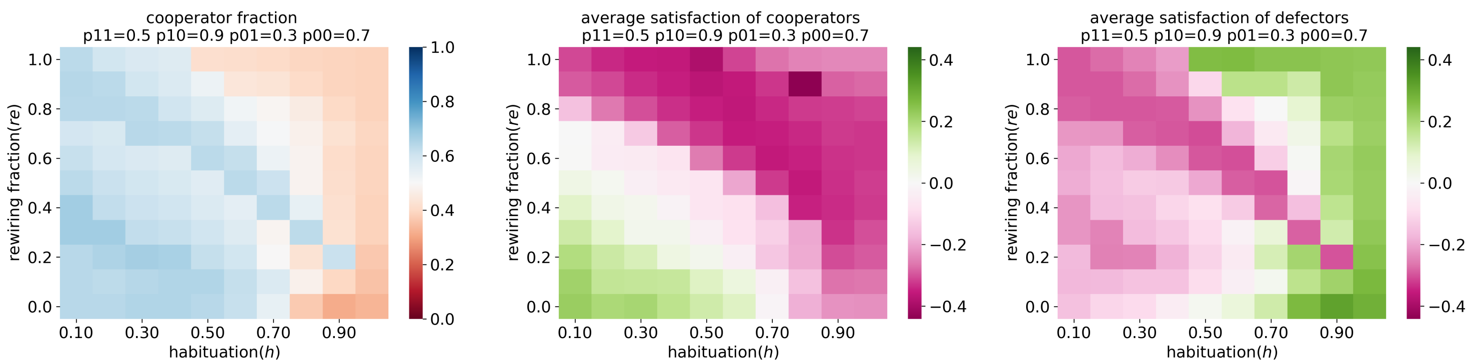
$$p_{11} = 0.9, p_{10} = 0.5, p_{01} = 0.7, p_{00} = 0.3$$



(a) (b) (c)

FIGURE 4.12: Heat map showing the (a) cooperator fraction; average satisfaction of (b) cooperators and (c) defectors as a function of habituation (h) and rewiring probability (re) for $N = 500$. Each pixel value is obtained by first averaging from round 250 to 750 for each trial and then averaging over over 5 trials. The willingness probabilities are $p_{11} = 0.9, p_{10} = 0.5, p_{01} = 0.7, p_{00} = 0.3$.

$$p_{11} = 0.5, p_{10} = 0.9, p_{01} = 0.3, p_{00} = 0.7$$



(a) (b) (c)

FIGURE 4.13: Heat map showing the (a) cooperator fraction; average satisfaction of (b) cooperators and (c) defectors as a function of habituation (h) and rewiring probability (re) for $N = 500$. Each pixel value is obtained by first averaging from round 250 to 750 for each trial and then averaging over over 5 trials. The willingness probabilities are $p_{11} = 0.5, p_{10} = 0.9, p_{01} = 0.3, p_{00} = 0.7$.

The information of link presence in the previous round does not have a substantial effect. The main factor can thus be attributed to the satisfaction of players.

Chapter 5

Conclusion

Evolutionary game theory models have tried to explain the evolution of cooperation using different types of models, where individuals in a population gather information from the surrounding and decide on their actions. In reality, procuring such information is hard and comes with an associated cognitive load. In some cases it is impossible since the setting of the interactions do not allow the participants to gain insight about their environments. Such restricted scenarios require the players to make decisions based on their past experiences. This study has shown that it is possible for players to maintain cooperation in a network if the players have an expectation which does not fluctuate very rapidly. The players in the network cooperate in brief intervals, because the aspirations of the players get adjusted to the recent payoffs. If individuals in a network get accustomed fast to their last payoff, then, the system reaches a dynamic equilibrium dominated by defection. Errors, while implementing action or during the setting of the aspiration, preserve sanity of the model by incorporating additional stochasticity in the system. Both the mode of additional stochasticity incorporation give identical results in a sensible range of the parameter values.

Allowing the players in the network to rewire ties with other players help in cooperators getting the upper hand, even if aspirations fluctuate rapidly. The modification of links based on the action of the players or based on the satisfaction both prefer the proliferation of cooperators, provided cooperators are preferentially selected to make more ties in the former and satisfied players have a higher propensity for connections in case of the latter.

Appendix A

Appendix

A.1 Behaviour of a static network in a Public Goods Games

In the usual public goods games, a game is centered around each player who acts as a focal player. All of her neighbours take part in the game with cooperators contributing a fixed amount but defectors freeload. The contributions are then multiplied by the synergy factor and redistributed equally among all the players. Thus, each game yields the following payoff for the players:

$$\begin{aligned} \Pi_{Ci,fp}(t) &= r\ m\ \frac{n_{C,fp}}{k_{fp}} - m \\ \Pi_{Di}(t) &= r\ m\ n_{C,fp} \end{aligned} \tag{A.1}$$

where, $n_{C,fp}$ is the number of cooperator neighbours the focal player has. Thus, every round consists of N games centering each player once.

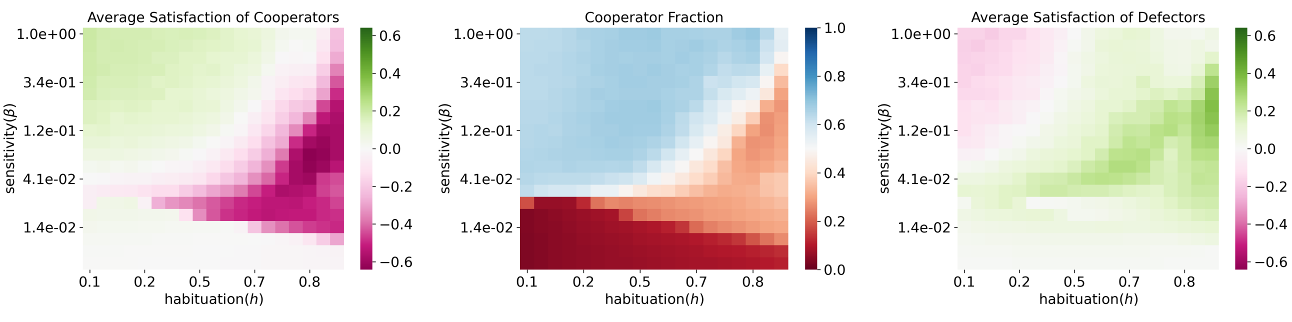
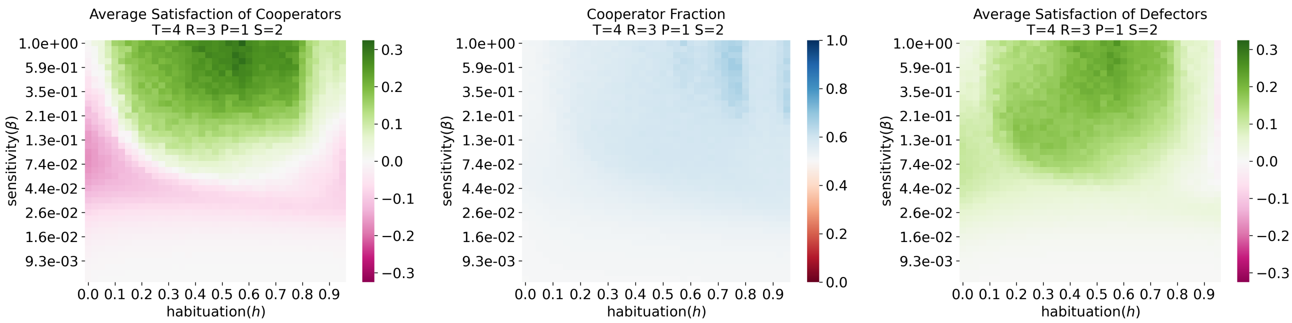


FIGURE A.1: Heat map showing the (a) cooperator fraction; average satisfaction of (b) cooperators and (c) defectors as a function of habituation (h) and sensitivity (β) for $N = 500$. Each pixel value is obtained by first averaging from round 250 to 750 for each trial and then averaging over over 5 trials.

A.2 Behaviour of the static model in other games

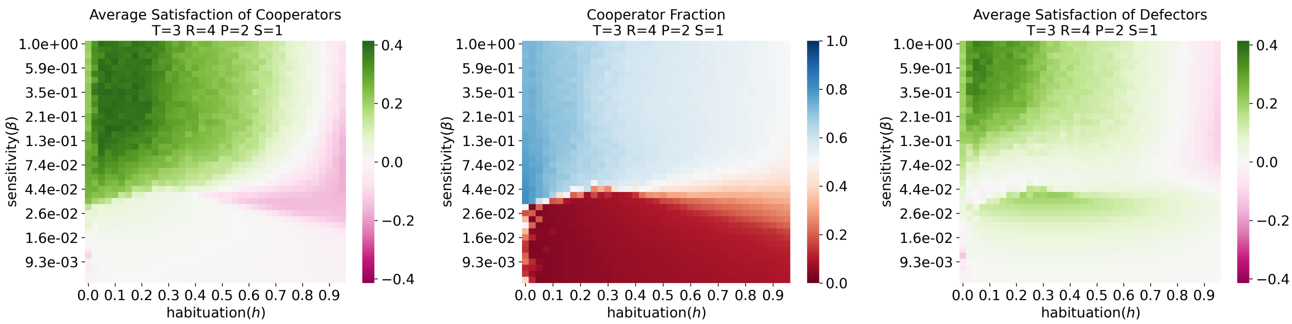
Snowdrift Game



(a) (b) (c)

FIGURE A.2: Heat map showing the (a) cooperator fraction; average satisfaction of (b) cooperators and (c) defectors as a function of habituation (h) and sensitivity (β) for $N = 500$ when players play snowdrift games with their neighbours. Each pixel value is obtained by first averaging from round 250 to 750 for each trial and then averaging over 5 trials.

Staghunt Game



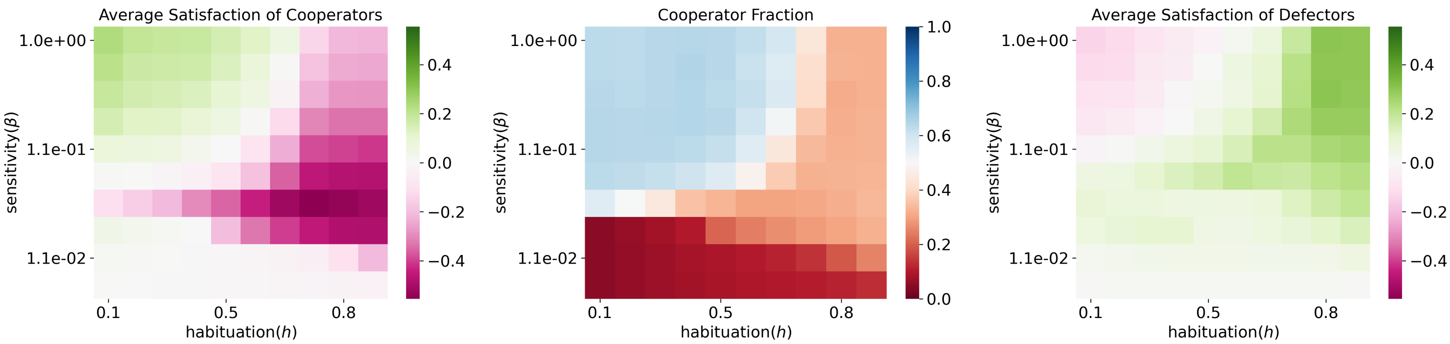
(a) (b) (c)

FIGURE A.3: Heat map showing the (a) cooperator fraction; average satisfaction of (b) cooperators and (c) defectors as a function of habituation (h) and sensitivity (β) for $N = 500$ when players play staghunt games with their neighbours. Each pixel value is obtained by first averaging from round 250 to 750 for each trial and then averaging over over 5 trials.

The two other social dilemmas show that moderate values of habituation can sustain cooperative behaviour no matter the dilemma.

A.3 Behvaiour of the static model on other networks

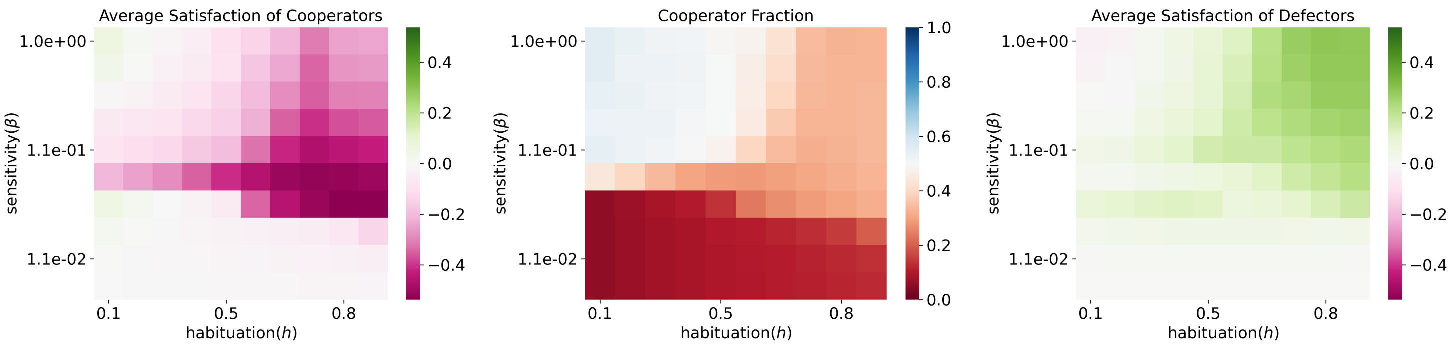
Barabasi Albert, $m = 75$



(a) (b) (c)

FIGURE A.4: Heat map showing the (a) cooperator fraction; average satisfaction of (b) cooperators and (c) defectors as a function of habituation (h) and sensitivity (β) for $N = 500$. Each pixel value is obtained by first averaging from round 250 to 750 for each trial and then averaging over over 5 trials.

Watts Strogatz, $k = 75$, $p = 0.5$



(a) (b) (c)

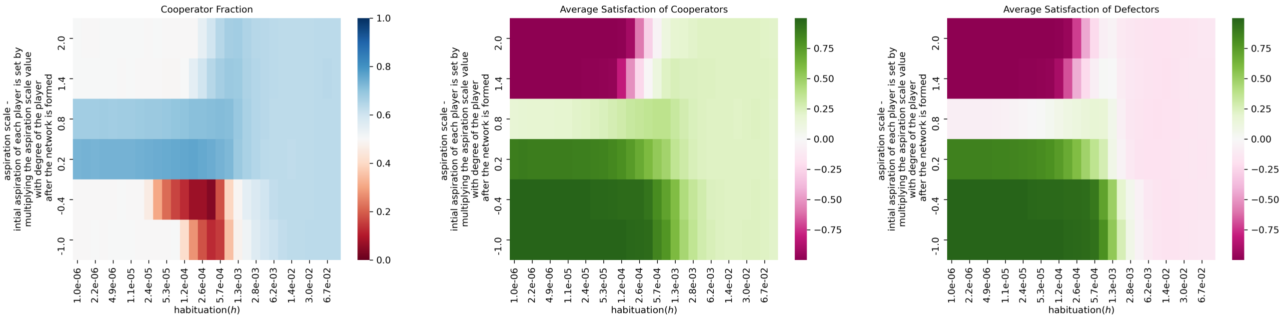
FIGURE A.5: Heat map showing the (a) cooperator fraction; average satisfaction of (b) cooperators and (c) defectors as a function of habituation (h) and sensitivity (β) for $N = 500$. Each pixel value is obtained by first averaging from round 250 to 750 for each trial and then averaging over over 5 trials.

Effect of connectivity in the network

FIGURE A.6: Heat map showing the (a) cooperator fraction; average satisfaction of (b) cooperators and (c) defectors as a function of habituation h and the connectivity in between players. Each pixel value is obtained by first averaging from round 250 to 750 for each trial and then averaging over over 5 trials.

The connectivity is meaured by the probability of edge formation in an Erdős Rényi network. The plot shows that a highly connected network increases the value of critical habituation h_c.

A.4 Dependence on initial aspiration



(a) (b) (c)

FIGURE A.7: Heat map showing the (a) cooperator fraction; average satisfaction of (b) cooperators and (c) defectors as a function of habituation *h* and the initial aspiration. Each pixel value is obtained by first averaging from round 250 to 750 for each trial and then averaging over over 5 trials.

Very low values of habituation restrict the players from exploring aspirations away from their initial aspiration. This can in turn stop the players from responding sensibly to the environment.

Appendix B

Appendix

B.1 Unit of transaction and sensitivity

Let the modified variables and constants for a game with unit transaction m(≠1) be denoted by tilde, $\tilde{X}$. Modifying the transaction unit essentially changes the payoff and the initial aspiration level:

$$\begin{aligned} \tilde{\Pi}_i(t) &= m\Pi_i(t) \\ \tilde{A}_i(1) &= mA_1(1) \end{aligned} \tag{B.1}$$

The following is the evaluation of aspiration for a few rounds starting from the initial round.

$$\begin{aligned} \tilde{A}_i(1) &= mA_i(1) \\ &\text{game 1 played} \\ \tilde{S}_i(1) &= \tanh(\tilde{\beta}(\tilde{\Pi}_i(1) - \tilde{A}_i(1))) \\ &= \tanh(\tilde{\beta}m(\Pi_i(1) - A_i(1))) \\ &\text{Probability of cooperation updated} \\ \tilde{A}_i(2) &= (1-h)\tilde{A}_i(1) + h\tilde{\Pi}_i(1) \\ &= m((1-h)A_i(1) + h\Pi_i(1)) \\ &= mA_i(2) \\ &\text{game 2 played} \\ \tilde{S}_i(2) &= \tanh(\tilde{\beta}(\tilde{\Pi}_i(2) - \tilde{A}_i(2))) \\ &= \tanh(\tilde{\beta}m(\Pi_i(2) - A_i(2))) \\ &\text{Probability of cooperation updated} \\ &\text{and so on.} \end{aligned} \tag{B.2}$$

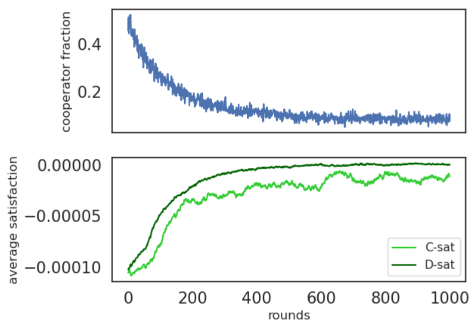
Thus, the general form of the equation can be written as:

$$\begin{aligned}
\tilde{A}_i(t+1) &= (1-h)\tilde{A}_i(t) + h\tilde{\Pi}_i(t) \\
&= m((1-h)A_i(t) + h\Pi_i(t)) \\
&= mA_i(t+1) \\
&\text{game is played} \\
\tilde{S}_i(t) &= \tanh(\tilde{\beta}(\tilde{\Pi}_i(t) - \tilde{A}_i(t))) \\
&= \tanh(\tilde{\beta}m(\Pi_i(t) - A_i(t))) \\
&\text{Probability of cooperation updated}
\end{aligned} \tag{B.3}$$

Thus, it is evident that keeping the habituation fixed and modifiying the learning rate β by a factor of m will keep the dynamics unchanged.

$$\tilde{\beta} = \beta/m \tag{B.4}$$

B.2 Insensitive players



(a)

FIGURE B.1: The simulation is averaged over 100 trials and the plots are generated by averaging the data generated in a moving window of 10 rounds.

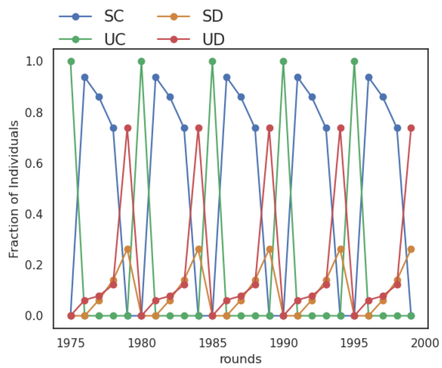
Players who are insensitive ($\beta \to 0$) to the difference between their payoffs and aspirations produce satisfaction values whose absolute values are close to 0 (2.3). Such low values of satisfaction restricts the players from sufficiently modifying their cooperation probabilities in response to the received payoffs. Nevertheless, the dissimilarity of the payoffs between cooperators and defectors gradually build up in the population. A cooperator surrounded by N defectors contribute Nm to her neighbours but does not get anything in return. A defector on the other hand does not gain or lose anything when surrounded by defectors. Similarly, in a cooperative environmnent, a free riding defector earns more than a cooperator. This imbalance produces minute difference in the satisfaction corresponding to cooperation and defection which gradually lower the probability of cooperation as can be seen in B.1(a).

Appendix C

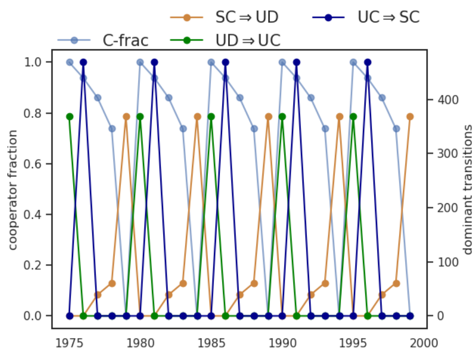
Appendix

C.1 Regular cycles at low habituation with noisy aspirations

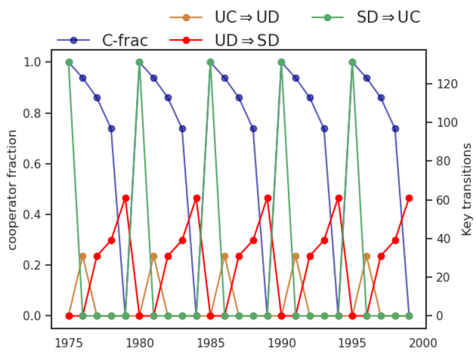
There exists two different types of cooperative dynamics for habituation lower than the critical value (refer to 3.8). The cooperative phase observed for low habituation values display extremely regular dynamics. The system gets stuck in a very regular cooperative dynamics demonstrated in C.1(a).



(a)



(b)



(c)

FIGURE C.1: (a) Temporal evolution of the players surfing through each category. (b) Dominant transitions and (c) Key transitions responsible for the collapse and rejuvination of cooperation.

The players who stay as SD for the longest, attain unachievable aspiration values. When the entire system cooperates after being unsatisied i.e. UC, these few players still stay unsatisfied and move to defection, UD (C.1(c)), while the rest merrily move on to SC (C.1(b)). The UD individuals become satisfied and continue defection but drag a few SC to UD making them into SD the following round. The $SC \rightarrow UD \rightarrow SD$ transition happens for two consecutive rounds, after which the rest of the cooperators follow the dominant cycle $SC \rightarrow UD \rightarrow UC \rightarrow SC$ (C.1(b)) because of critically lowered payoffs. Thus, the initial budge provided by the high aspiring $UC \rightarrow UD$ indivduals, periodically pull the entire system out of cooperation. Disturbing the system periodically reveal that there are several such configurations where the system can stagnate (not shown).

Bibliography

[1] Dirk Helbing et al. "Punish, but not too hard: how costly punishment spreads in the spatial public goods game". In: *New Journal of Physics* (2010).

[2] Karandikar Rajeeva et al. "Evolving aspirations and cooperation". In: *Journal of Economic Theory* 72 (1998), pp. 292–331.

[3] Pathak Spandan et al. "How strategy environment and wealth shape altruistic behaviour: Cooperation rules affecting wealth distribution in dynamic networks". In: *Proc. R. Soc. Lond. B* (2020).

[4] Traulsen Arne et al. "Stochastic dynamics of invasion and fixation". In: *Physical Review E* (2006).

[5] Flache Andreas Macy W. Michael. "Learning dynamics in social dilemmas". In: *PNAS* (2002), pp. 7229–7236.

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

[7] Nakamura Mitsuhiro Masuda Naoki. "Numerical analysis of a reinforcement learning model with the dynamic aspiration level in the iterated Prisoner's dilemma". In: *Journal of Theoretical Biology* (2011), pp. 55–62.

[8] Shapiro M. Avi Pichler Elgar. "Public goods games on adaptive coevolutionary networks". In: *Chaos* (2017).

[9] Sigmund Karl Posch Martin Pichler Alexander. "The efficiency of adapting aspiration levels". In: *Proc. R. Soc. Lond. B* (1999), pp. 1427–1435.

[10] Plotkin B. Joshua Radzvilavicius L. Arunas Stewart J. Alexander. "Evolution of empathetic moral evaluation". In: *eLife* (2019).

[11] Helbing Dirk Roca P. Carlos. "Emergence of social cohesion in a model society of greedy, mobile individuals". In: *PNAS* (2011), pp. 11370–11374.

[12] Masuda Naoki Tanabe Shoma. "Evolution of cooperation facilitated by reinforcement learning with adaptive aspiration levels". In: *Journal of Theoretical Biology* (2012), pp. 151–160.