



City Research Online

City St George's, University of London

Citation: Aizouk, R. A. (2022). Modelling conflicting individual preference: Target sequences and graph realization. (Unpublished Doctoral thesis, City, University of London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/29193/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Modelling conflicting individual preference:
target sequences and graph realization

Raneem Ayman Aizouk

Doctor of Philosophy



City, University of London
Department of Mathematics

November 3, 2022

Contents

List of Figures	4
List of Tables	6
0.1 COVID-19 Impact Statement	8
1 Introduction	1
1.1 Thesis Outline	1
1.2 Summary of contributions and publications	3
1.3 Modelling evolution in a population	3
1.4 Game theory	4
1.5 Evolutionary game theory	6
1.6 Evolutionarily Stable Strategies (ESSs)	7
1.7 The Replicator dynamics	8
1.8 Evolutionarily Game Theory in a Finite Population	10
1.9 Graph Theory	11
1.10 Evolution on graphs	16
1.11 The dynamic network model applications	17
2 Graphical Sequences	21
2.1 What is a graphical sequence?	21
2.2 The transition graph	23

2.3	Sufficient and Necessary conditions for a graphical sequence	24
2.4	Applications of the theorem	26
2.5	Vertex classifications	28
3	A dynamic network population model	30
3.1	Modelling populations	30
3.2	The Markov Chain Model	32
3.3	The minimal set	36
3.3.1	Properties of the minimal set	36
3.4	The deficit	38
4	The size of the minimal set	41
4.1	The largest minimal set	41
4.2	The minimal set for the arithmetic sequence	47
4.3	The all or nothing sequence	63
5	Game theoretical Models and the Minimal Set	68
5.1	The Game	69
5.2	Example: showing that leaving the minimal set can be optimal . . .	70
5.3	The game theoretical model with target sequence score equal to 1 .	75
5.3.1	Score 1 games	75
5.3.2	The sequence 1, 1, 1	76
5.3.3	States and transitions	77
5.3.4	Finding Nash equilibria NE	81
6	New dynamic network population models	85
6.1	General sequence results	89
6.1.1	Vertex classifications	89
6.2	Properties of the terminal set	96

6.3	The Reverse Havel- Hakimi Theorem	100
6.4	The terminal set over the arithmetic sequence	104
6.5	An illustrative game example: the arithmetic target 4, 3, 2, 1, 0	109
6.6	The maximum possible deviation for a graph from the terminal set of the arithmetic sequence	112
7	Game theoretical examples: reaching the terminal set	120
7.1	The game	120
7.2	A game theoretical example: $T=4, 3, 2, 1, 0$	122
7.3	Suggested Strategies	127
7.4	Examples showing that our alternative strategies are not always optimal	129
7.4.1	Strategy LT is not always optimal	130
7.4.2	Neither LR nor LRA are always optimal.	130
7.5	Game simulations	132
7.5.1	Game simulation of the sequence $t = 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0$	133
7.5.2	Game simulations for the sequence $t= 2, 2, 2, 2, 2.$	139
8	Discussion	141
9	Appendix	146
9.1	The biggest minimal set	146
9.2	MATLAB program:	148
9.3	The game theoretical model, sequence 1, 1, 1	161
9.4	Python	164
	Bibliography	174

List of Figures

1.1	Complete graphs	14
1.2	Empty graph	14
1.3	The complement graph	14
1.4	Regular graph	15
1.5	Cycle graph	15
1.6	Path graph	15
1.7	Sub graph	16
1.8	Two Bipartite graphs	16
2.1	Graphs a, b, and c that realise sequence $t=2, 2, 2, 2$	23
2.2	Graph G' realises sequence s' before adding vertex w in the proof of H-H Theorem	25
2.3	Adding vertex w to G' in the proof of H-H Theorem	26
4.1	Transitions between the elements of $K(min)$ of the sequence 1, 1, 1.	44
4.2	The links formed between two set of nodes following the linking process mentioned in theorem 4.2.5.	55
4.3	The analysis of the odd and even values of the all or nothing and arithmetic sequences	67
5.1	Transition probabilities of the sequence 1, 1, 1	77

5.2	Transition graph of the sequence 1, 1, 1 showing the possible elements of the minimal set	79
5.3	Transition graph of the first NE solution of sequence 1, 1, 1	83
6.1	In the proof of lemma 6.2.5, this figure shows the mechanism for exchanging links between two considered nodes	99
6.2	In the proof of lemma 6.2.5, the new sets of edges were employed after exchanging their links	100
6.3	Example showing that, for the arithmetic sequence $n = 12$, the cases of node $i \in V_{S_1}$ when it follows lemma 6.4.2.	108
6.4	Example showing that, in the arithmetic sequence $n = 13$, the cases of node $i \in V_{S_1}$ when it follows lemma 6.4.2.	109
6.5	All the possible graphs/states an individual can achieve moving one step from the empty graph	110
7.1	LT is not always an optimal strategy in the example sequence $t = 7, 6, 5, 4, 3, 2, 1, 0$	131
7.2	LR and LRA are not always optimal strategies in the sequence 9, 8, 8, 8, 6, 3, 1, 1, 1, 1.	132

List of Tables

3.1	The eight possible graphs involving individuals v_1, v_2, v_3 with a target $(2, 1, 0)$ respectively.	34
3.2	The minimal set involving individuals v_1, v_2, v_3 with a target $(2, 1, 0)$ respectively	37
4.1	The sequences with the biggest minimal set size where n is the number of individuals.	45
4.2	In this table we show the size of the minimal set, its logarithm and the corresponding term from the given approximation.	63
4.3	The sequences and minimal set sizes for the maximal score sequences for $n = 2, \dots, 9$	63
4.4	The values of the minimal set for the all or nothing sequences with m_1 “all” vertices and m_2 “nothing” vertices.	65
4.5	In this table, we show the logarithm of the size of the minimal sets for the all or nothing and arithmetic sequences.	65
4.6	The values of the minimal set for the all or nothing sequences with m_1 “all” vertices and m_2 “nothing” vertices.	66
6.1	Example sequence of the R.H.H and the H.H	103

6.2	This table presents all 26 possible non-terminal graph states α_i and the 4 terminal graphs T_i for the target $t = 4, 3, 2, 1, 0$	111
7.1	All possible transitions between pairs of states of a game starting from the empty graph to graphs within the terminal set.	125
7.2	The sequence 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0 is examined here, with a focus on the deviations when nodes $i = 4, \dots, 10$ are selected to alternate.	135
7.3	This table displays the group's overall score considering the sequence 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0.	136
7.4	This table shows the deviations of the nodes in the sequence 2, 2, 2, 2, 2 following different strategies.	139
9.1	Sequences with the biggest minimal set size where n is the number of individuals, M.S size: refers to the minimal set size	146

0.1 COVID-19 Impact Statement

This statement is provided for the aid and benefit of future readers to summarise the impact of the COVID-19 pandemic on the scope, methodology, and research activity associated with this thesis. The academic standards for a research degree awarded by City, University of London and for which this thesis is submitted remain the same regardless of this context. Title of the research project: Modelling conflicting individual preference: target sequences and graph realization
.....

- Summary of how the research project, scope or methodology has been revised because of COVID-19 restrictions We held our meetings online (skype / zoom) due to covid 19 where we used tablets to further discuss the work established.
- Summary of how research activity and/or data collection was impacted because of COVID-19 restrictions, and how any initially planned activity would have fitted within the thesis narrative. My research work basically was not affected by Covid 19 restrictions, and I was able to do my research and collect the necessary data from home without difficulty. The difficulties I faced were feeling separated from my colleagues friends as I was stuck at home and it was very difficult just to conduct research. And my family, who financially supports me, endured major income reductions in the UAE, which had a temporary impact on my ability to pay tuition expenses (which has been paid later on).
- Summary of actions or decisions taken to mitigate for the impact of data collection or research activity that was prevented by COVID-19 Professor Mark Broom was able to provide me with the help required and follow up on

any issues that arose because our work was more theoretical and proof-based.

- Summary of how any planned work might have changed the thesis narrative, including new research questions that have arisen from adjusting the scope of the research project We generally stuck to the research plan we laid forth at the start of the program.

Abstract

In this thesis, we build on a model introduced by Mark Broom and Chris Cannings in 2013. A group of individuals form pairwise contacts between themselves, and each has a potentially different target number of desired connections. How close can this target come to being realised? It is known that under a reasonable dynamic process, the graph of links evolves to one that minimises the total deviation from the target. The sets of all such graphs and associated sequences which are the sequences of node degrees are termed the minimal sets. We revisited the minimal set of sequences and obtained a formula that generates the size of that set for a given arithmetic sequence. We investigated the all or nothing sequences which are sequences of n individuals with target of $n - 1$ or 0 , and discovered a recurrence relation for such a formula considering the size of the minimal set of sequences. We investigated a game-theoretical version of the model, where individuals strategically chose the specific link. Showing that optimal play could lead to the minimal set being left, thus answering an open question from earlier work. We revisited the game over the sequence $1, 1, 1$ and found a general expression for the payoff functions for possible strategy combinations. A new set of six more solutions showed the complexity of this simple case, with the possibility of more solutions open. We introduced two models allowing individuals to reject, accept or break links suggested by others. Following the original model dynamical process and considering each new model rule led to a non-transient set of graphs, termed the terminal set. We defined the terminal set and demonstrated a general sequence's vertex classifications and related properties. We introduced the Reverse Havel Hakimi theorem/algorithm and considered examples of game simulations

over the terminal set that reached it, and investigated the individual's best strategy to maximise their payoff at every state using backwards induction. Together, these results have provided us with some predictions about which networks are likely to form following different proposed social models during a particular period of time in an evolutionary process.

Chapter 1

Introduction

This chapter is considered as the thesis introduction; we shall discuss the fundamental principles on which this thesis is based, including game theory and evolutionary graph theory. Additionally, it will provide a summary of the thesis chapters and instructions on how to read the thesis.

1.1 Thesis Outline

In this section, we summarise the thesis chapters in which we considered a dynamical population network which is a snapshot of a dynamically evolving process. Since the dynamical process will be performed over a short snapshot life time t , it will be recognised as a portion of an evolutionary process. The outcomes will be incorporated in the larger evolutionary process and the payoffs of the players will play a role in a more sophisticated evolutionary process that occurs across an infinite number of populations, with individuals classified into a range of target classes. As a result, the games presented will be performed over a short snapshot life time t , with each target class member allocated to the same target of social interactions. We examined and investigated models with a fixed number of vertices that represent individuals interacting with one another, and the set of edges that

represent the links between these individuals.

Chapter 1 provides a general introduction to game theory, evolutionary game theory and graph theory, as well as some applications of each, as well as a general introduction to the dynamic network population model. Chapter 2 defined graphical sequences and discussed the sufficient and necessary conditions for a graphical sequence, which are summarised in the Havel Hakimi theorem, as well as some applications of the theorem.

Chapter 3 describes the model, including the target sequence as well as the general concept of the minimal set and its properties.

Chapter 4 considers the minimal set size, including two key classes of target sequence.

Chapter 5: discusses the game theoretical model introduced in [10] and further developed in [12] to resolve an open question from [10].

Chapter 6 introduced the terminal set over two different models B and C that were derived from model A, as well as their concluded properties. It also introduced the Reverse Havel Hakimi Theorem RHH, the RHH algorithm and the score of a sequence that could be found by this algorithm, and an algorithm that classifies the type of the sequences type. We studied the models more closely for two key classes of target sequences, and we used the sequence 4,3,2,1,0 as an example. Furthermore, it demonstrated the maximum possible deviation for a graph from the terminal set of the arithmetic sequence that describes the group's worst behaviour. Chapter 7 presents many games versions considering the terminal set over model B and C.

Chapter 8 is a discussion of our results and future work.

Chapter 9 is the appendix, and it contains some more detailed calculations and programs.

1.2 Summary of contributions and publications

This thesis has been divided into 9 chapters, chapters 4, 5, 6 and 7 cover the actual original new research that we conducted as part of my PhD thesis. The original research was a collaboration between me and my supervisor, Mark Broom in which we were meeting weekly to discuss the progress of the work. I've carried out this research under his guidance, and supervision. In terms of publications, I participated as the first author for two papers, which have been published in peer-reviewed journals. I contributed as the first author in the paper entitled: Modelling conflicting individuals preference: target sequence and graph realisation, see [2]. This paper was published in "Discrete and Continuous Dynamical systems series B" (2022). The paper entitled "The game-theoretical modelling of a dynamically evolving network: revisiting the target sequence 1, 1, 1 was submitted and reviewed for publications, but it has yet to be published. A third paper is still being prepared for submission; we have finalised the mathematical framework and concepts and need to make the final adjustments for submission. We intend to work on a fourth paper which will examine the games over the terminal set.

1.3 Modelling evolution in a population

In the 18th century, many scientists didn't accept the idea that species could evolve. Lamarck was the first scientist to propose that there were living species that had evolved from their forefathers. An example of this adaptation is the giraffe's long neck, which helps it acquire food [28]. Others disagreed with Lamarks' theory, claiming that the long neck was due to food competition [29]. Darwin developed a more realistic vision of species development in the mid-nineteenth century. His theory that all living things descended from common ancestors is

now widely acknowledged and regarded as a fundamental scientific principle [19]. Parents who are physically fit and have procreative qualities have a better chance of having more children than those who are not. Those with a procreative disadvantage are more likely to vanish. These traits are passed from one generation to the following generations. That was presented by Darwin's book "The Origin of Species" in 1885, which is considered to be the foundation of evolution, see [19]. A mathematical theory of evolution to a significant extent requires consideration of the interactions of individuals, where fitness depends upon the choices of others so we must consider models involving game theory.

1.4 Game theory

Game theory is commonly employed in different fields, including the study of both human and animal behaviour. It helps economists to better understand the economic behaviours, which include those of consumers, markets and companies. The social sciences also frequently employ game theory. People have always played games, see [90, 20, 89] for more details, they tried to choose the best strategies to maximise the chance of victory. Game theory studies the mathematical models of individuals' strategies used while interacting to optimise the best outcome possible [80, 58, 44, 68, 72, 24]. Individuals might use mixed or pure strategies (explained below), but the chosen strategy will not just affect the individual itself, but it will also affect the whole population in the long term. John von Neumann invented modern game theory in 1944 and published a book with Morgenstern that explains "The games of economics behaviour" see [82], where they said that any economic game could be defined as the outcome of a game between two or more players, and defined the elements of the game which have been discussed in detail in [11] as well as follows:

1. Players: they meet in pairwise competitions with others and make strategic decisions. The most common number of players is two, however it can be more than two.
2. Strategy: this is the complete specification of the chosen action in every possible position; it's classified as follows:
 - Pure strategy: defined as the move a player will make in each iteration, and there can be a finite or infinite number of pure strategies.
 - Mixed strategy: for the set S_1, S_2, \dots, S_n of finitely pure strategies, a mixed strategy is defined as a probability vector $\mathbf{P} = (p_1, p_2, \dots, p_n)$ where p_i is the probability of playing strategy S_i in the current game.
3. Payoffs: in other words, rewards and /or punishments that occur for different strategy combinations. The fitness of the individual is considered to be the payoff of the game, plus some additional background fitness.

One limitation of John von Neumann's version is that it focused on finding optimal strategies for one type of game called a zero-sum game. However many important games, for example the Prisoner's Dilemma, are not zero-sum games. John Nash found a way to analyse any finite game, and this became the main tool to define the solution of a non-cooperative game with two or more players. A Nash equilibrium [58] occurs in a game of two or more players when each individual uses a strategy that is the best response to the other. As a result, no individual can increase their payoff by changing to a different strategy. For two players, suppose that player 1 plays strategy S against player 2 playing T ; $E_i(S, T)$ then represents the payoff to player i . The pair of strategies (S, S) is a Nash equilibrium if neither player can improve their payoff by changing strategy, i.e. $E_1(S, S) \geq E_1(T, S)$ and $E_2(S, S) \geq E_2(S, T)$.

1.5 Evolutionary game theory

In this section, we will go over the evolutionary game theory (EGT), for more details see [51, 52, 18, 46, 94, 38, 39, 66, 11] is a mathematical way of modelling the Evolution of biological populations that was initially developed for well-mixed (non-spatial) populations. It originated in 1973 with John Maynard Smith and George R.Price, where they defined and analysed the strategies used and demonstrated the mathematical criteria that can be used to predict the result of competing strategies [52].

Classical and evolutionary game theory attempt to explain different phenomena, evolutionary game theory is focusing more on the dynamics of strategy change [59], which is influenced by the frequency of the competing strategies in the population [22]. Hamilton was a pioneer in evolutionary game theory as he proposed one of the first works (see[32]). The application of game theory to population evolution was later considered in more detail by John Maynard Smith and George Price [52]. Recently, evolutionary game theory has become an effective and influential methodology in modelling biology, and it has proved to be a popular approach.

Evolutionary game theory analyses Darwinian mechanisms. Nowak and May [64, 65] studied the Prisoners' Dilemma on lattices to further illustrate the problems of the evolution of cooperative behaviour. In the model they introduced, individuals will be selected to interact with another neighbour, and the successful individual will occupy the site. This work encouraged other scientists to investigate evolutionary game dynamics on lattices further [63, 62, 21, 45, 57, 87, 54, 36, 69, 73]. More generally, spatial structure can be represented as a graph [49, 84], and we shall revisit this idea in later sections. Individuals are represented as vertices on a graph, and their relations are represented as a set of edges. Individuals were selected randomly to form or break a link with another neighbour to receive its

payoff.

1.6 Evolutionarily Stable Strategies (ESSs)

A primary mathematical concept in biological games is the Evolutionarily Stable strategy (ESSs) introduced by John Maynard Smith and George R. Price in 1972/3 [50, 52]. If employed by a population, it is a strategy that can withstand attempted invasion from all other strategies. All sufficiently rare alternative strategies will be out-competed. This concept was anticipated in 1967 by Hamilton, when he proposed the idea of an unbeatable strategy in his paper on sex ratios in games (see[32]). The ESS is a stronger version of the Nash equilibrium (note that here we have a symmetric game, so a Nash equilibrium is one that is a best reply against itself), so every ESS corresponds to a Nash equilibrium, but some Nash equilibrium are not *ESS*. Broom and Rychtar considered the idea of finding *ESS* for any given matrix in a systematic way and mentioned some important properties of *ESS* in their book "Game-Theoretical Models in Biology" in [11], following work developed earlier by John Haigh, see [30]. For an infinite population with two competitive players, every individual is equally likely to meet each other, considering, F_I as the fitness of individuals using strategy I, and $E(I, J)$ is the payoff to an individual using strategy I against an opponent using strategy J , i.e. for the payoffs from the previous section $E_1(I, J) = E(I, J)$ and $E_2(I, J) = E(J, I)$, we will have the following: Strategy I is an ESS if when almost all members of the population adopt I, if the fitness of these types is greater than those of mutant type that appears in a positive proportion of the population. Here we consider a population consisting mainly of strategy I with a small proportion ϵ of a mutant type J :

$$F_I = (1 - \epsilon)E(I, I) + \epsilon E(I, J) \tag{1.1}$$

$$F_J = (1 - \epsilon)E(J, I) + \epsilon E(J, J) \quad (1.2)$$

To prevent invasion from the individuals using strategy $J \neq I$ we require that: $(1 - \epsilon)E(I, I) + \epsilon E(I, J) > (1 - \epsilon)E(J, I) + \epsilon E(J, J)$. Since ϵ is very small, we can ignore the terms with ϵ , which means strategy I is evolutionarily stable against J whenever $E(I, I) > E(J, I)$.

However if $E(I, I) = E(J, I)$ then we do need to consider the small fraction of games against J , when in addition we require $E(I, J) > E(J, J)$.

A strategy is an ESS if it is evolutionarily stable against all $J \neq I$.

The first condition holding against all $J \neq I$ is equivalent to a strict Nash equilibrium [34].

1.7 The Replicator dynamics

The replicator dynamics is another primary approach that analyzes non-spatial evolutionary games and was developed mainly for symmetric games with finitely many strategies. The replicator dynamics is a system of ordinary differential equations describing strategy frequencies over time, assuming reproductive rates proportional to game payoffs [85, 38, 39, 26, 66, 25, 76].

Consider an infinitely large population consisting of n types. To define the replication process, we find a deterministic equation which leads on to the evolutionary equilibria of the population. The replicator equation is defined as:

$$\dot{x}_i = x_i[F_i - \phi], \quad (1.3)$$

where x_i is the frequency of type i , the population's state is given by the vector $x = (x_1, x_2, \dots, x_n) \in S_n$, F_i is the fitness of individuals of type i , and ϕ is the population's average fitness. The replicator equation says that the frequency of individuals using a strategy will decrease if their fitness is lower than the population's

average fitness and will increase if it is higher. To find evolutionary equilibria we can solve the equations where the derivatives are set to equal zero. Consider the case of two pure strategies and assume that the frequency of individuals playing strategy A is x and for strategy B is $1 - x$, with the payoff matrix:

$$\begin{array}{c|cc}
 & A & B \\
 \hline
 A & a & b \\
 B & c & d
 \end{array} \tag{1.4}$$

The entries a, b, c, d represent the payoffs; for instance, d is the payoff of an individual using strategy B with the opponent playing strategy B as well.

$$F_A = xa + (1 - x)b, \tag{1.5}$$

$$F_B = xc + (1 - x)d. \tag{1.6}$$

The average fitness of the population is given by $\phi = xF_A + (1 - x)F_B$.

Thus, the change in frequency of type A individuals is:

$$x_A = x[F_A - \phi], \tag{1.7a}$$

$$= x[F_A - xF_A - (1 - x)F_B], \tag{1.7b}$$

$$= x(1 - x)[F_A - F_B], \tag{1.7c}$$

$$= x(1 - x)[x(a - b - c + d) + b - d]. \tag{1.7d}$$

By solving the equation $x_A = 0$, we will get the equilibrium points which are:

$x^* = 1$, $x^* = 0$, $x^* = \frac{d-b}{a-b-c+d}$. Hisashi Ohtsuki and Martin A. Nowak derived

a differential equation representing the replicator equation on a spatial structure, see [67]. Cox, Durrett, and Perkins [17], provided another strict spatial analogue

of replicator dynamics in their study of evolutionary dynamics on square lattices of dimension 3. They showed that, when they scale time and space, the particle systems of the voter model, a discrete dynamics related to the Moran process described in Section 1.8, converged to a solution to a particular partial differential equation. This PDE describes how strategies can transmit through space, and can be explained as a replicator equation.

1.8 Evolutionarily Game Theory in a Finite Population

For the sake of simplicity, we are analysing a snapshot of time for a limited set of members from an evolutionary process. Finite populations can be modelled using the Moran process, a biological stochastic process of evolutionary game theory in finite populations with frequency-dependent fitness/evolution in a group [55], [56].

The Moran process follows a group of N individuals with i of them following strategy A and $N - i$ following strategy B. The fitnesses of each type can depend upon the value of i ; often this depends upon pairwise games against players selected at random, following the payoff matrix given in Table 1.4. Individuals that use strategy A have the fitness $f_i = a(i - 1) + b(N - i)$, and the fitness of individuals using strategy B is $g_i = ci + d(N - i - 1)$. When an individual has a fitness advantage at a particular time, it is more likely to be chosen for reproduction, the offspring generated will then replace another randomly selected individual. As a result, the number of individuals of either type can be changed by no more than one at each step of the Moran process while the population size remains constant. State i of the population represents when we have i type A's and $N - i$ type B's; a transition could happen from state i to states $i - 1, i, i + 1$. Assuming there are no further

mutations in the population implies that one type will fixate (totally dominate) the population eventually, there being two absorbing states $i = 0, i = N$.

To calculate a type's fixation probability, we can define the transition probability of transitioning from a state where there are i type A individuals to a state where there are j type A individuals,

$$p_{i,j} = \begin{cases} ib_A(i) * \frac{N-i}{N-1} & j = i + 1 \\ (N-i)b_B(i) * \frac{i}{N-1} & j = i - 1 \\ 1 - p_{i,i+1} - p_{i,i-1} & j = i \end{cases}$$

where $p_{i,j}$ denotes the probability of transition from state i to state j and $b_A(i) = \frac{f_i}{if_i+(N-i)g_i}$, $b_B(i) = \frac{g_i}{if_i+(N-i)g_i}$. If we assumed neutral selection with the same fitness for both types we will have $p_{i,i+1} = p_{i,i-1} = \frac{i}{N} * \frac{N-i}{N-1}$.

1.9 Graph Theory

Definition 1.9.1. A vertex (or node) is the basic component of graph formations according to graph theory. The structure of the vertices varies depending on the use for which the graph is created; for example, a population network is a graph in which the vertices represent specific people. The concept “node” and “vertex” will be used interchangeably.

Definition 1.9.2. An edge (or link) of a network (or graph) is the connection between two vertices of the network. Edges can be directed or undirected.

Definition 1.9.3. A graph $G = (V, E)$ is defined as a set of vertices and a set of unordered pairs $E \subseteq V * V$ with $(i, i) \notin E$ in a simple graph, see [5].

Vertices are considered indistinguishable if they have no special type over and

above properties resulting from the graph itself, e.g. degree (see [10] and [33] for more information). When considering some evolutionary process for graphs where vertices represent individuals, see [49], and the edges represent the links between them, there will be several levels of complexity. For directed multigraphs, the link can join the vertex to itself, and we call this a loop.

Definition 1.9.4. Multiple edges in graph theory are two or more edges that are incident to the same two vertices in an undirected graph or two or more edges that share the same head vertex and tail vertex in a directed graph.

Definition 1.9.5. A simple graph is an undirected graph which does not have loops or multiple edges. In a simple graph of order n , the maximum degree of any vertex is $n - 1$.

Definition 1.9.6. A tree is an undirected graph in which any two vertices are connected by precisely one path of edges.

Graph theory in mathematics means the study of graphs (see [92, 5, 7, 15, 86, 6]). Graphs form one of the prime objects of study in discrete mathematics and underlie many computer programs that make modern communication and technological processes possible. Here are a few concrete examples of a graphs:

- Countries are nodes and air routes are edges.
- Humans are nodes and social media between them represent edges.
- Cells are nodes and blood vessels are the edges.

We will present some types of graphs; for more details, see [33].

- complete graphs: The complete graph on n vertices. Denote it by K_n , it is a graph of order n where each pair of graph vertices is connected by an edge. $uv \in E$ for all u and $v \in V$, see Figure 1.1.

- Empty graphs or edgeless graphs: The empty graphs on n isolated vertex, denoted by E_n , is the graph of order n where E is the empty set, see Figure 1.2.
- The complement or inverse of a graph G is a graph \bar{G} where both have the same vertex set such that two distinct vertices of \bar{G} are adjacent if and only if they are not adjacent in G meaning the edges of \bar{G} do not present in G , see Figure 1.3.
- Regular graphs: a regular graph is a graph where each vertex has the same number of neighbours; i.e. every vertex has the same degree. Complete graphs of order n are regular of degree $n-1$, and empty graphs are regular of degree zero, see Figure 1.4.
- Cycles: a cycle graph or circular graph is a graph that consists of a single cycle, or in other words, some number of vertices (at least 3) connected in a closed chain, see Figure 1.5.
- path graphs: The path graph or line graph is a tree with two nodes of vertex degree 1, and the other nodes of vertex degree 2. A path graph is therefore a graph that can be drawn so that all of its vertices and edges lie on a single straight line, see Figure 1.6.
- Sub graphs: A sub graph H of a graph G is a graph whose vertex set is a subset of the vertex set of G , so $V(H) \subseteq V(G)$ and whose edges is a subset of the edge set of G , so $E(H) \subseteq E(G)$. In this case we conclude that $H \subseteq G$ and we say G contains H , see Figure 1.7.
- Bipartite graphs: A graph G is Bipartite if its vertex set can be divided into two disjoint sets X and Y , called the parts of the graph, and every edge

connects a vertex from set X with a vertex from set Y , see Figure 1.8. You can see more details about the figures below in [33].

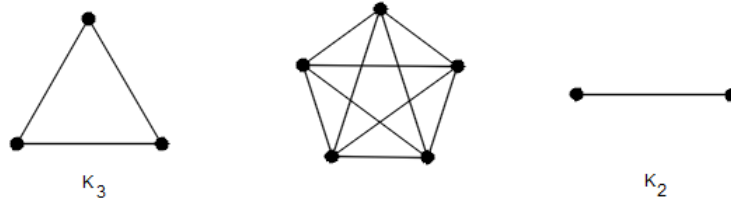


Figure 1.1: Complete graphs

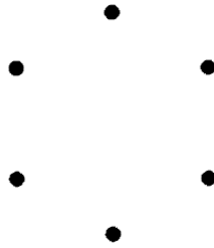


Figure 1.2: Empty graph

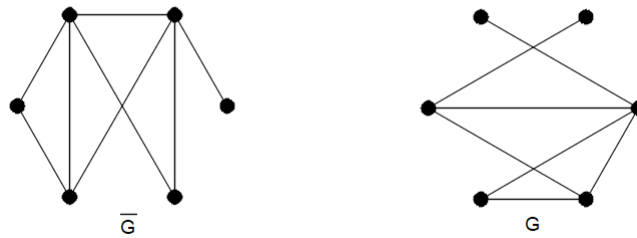


Figure 1.3: The complement graph

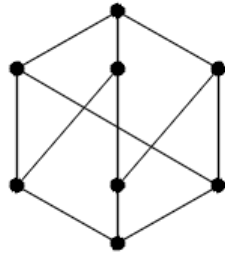


Figure 1.4: Regular graph

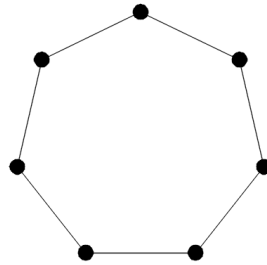


Figure 1.5: Cycle graph



Figure 1.6: Path graph

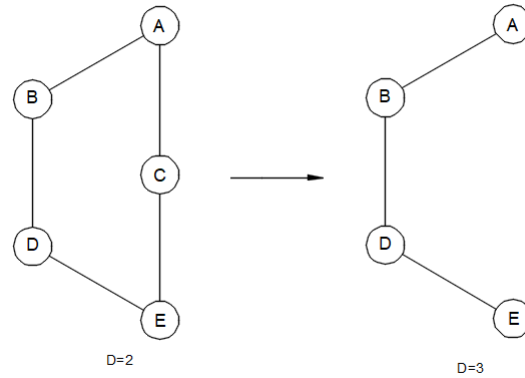


Figure 1.7: Sub graph

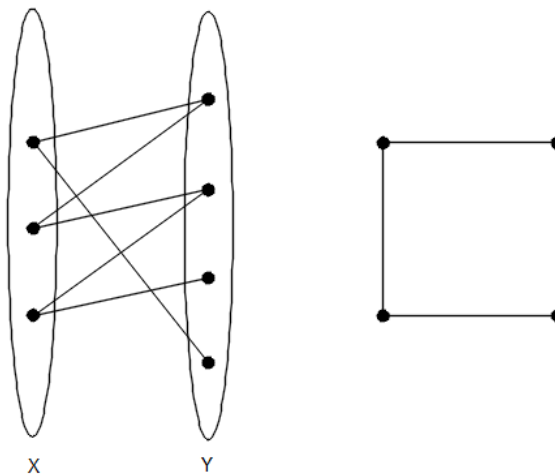


Figure 1.8: Two Bipartite graphs

1.10 Evolution on graphs

Evolutionary game theory (EGT) was introduced in 1973 by John Maynard Smith and George Price, see [52]. In 2005, Lieberman and colleagues developed evolutionary graph theory, see [49], which was a development of the original Moran population model in [55] to investigate the ability of mutants to take over residents in a finite population. Since then, many studies have been conducted in

this area [23, 95, 49, 1]. Individuals are represented as vertices in (EGT), and their relations represent the edges; their interactions will form a direct, weighted graph $G(V, E)$ at time τ . By taking into account births and deaths throughout time, a change in the population's overall composition will occur at time $\tau + 1$, which will be represented by $G(V_{\tau+1}, E_{\tau+1})$. The fixation probability has been shown to depend greatly on the structure of the population (the graph) and on the specific evolutionary dynamics, the latter much more than for well-mixed populations. Evolution on graphs has been an interesting topic for many scientists. Martin A. Nowak reviewed the basic approach for evolutionary games on graphs, see [4].

1.11 The dynamic network model applications

In the model introduced by Mark Broom and Chris Cannings [?], they considered a network of interactions between the distinguishable vertices set $v = (v_1, v_2, v_3, \dots, v_n)$ which can represent animals or humans, and the social or geographical relations between them are represented as a set of edges each of which links a pair of individuals (vertices); $X = (x_{ij})$ where $i, j = 1, 2, \dots, n$ (see [10]). The vector $e_\tau = (e_{1\tau}, e_{2\tau}, \dots, e_{n\tau})$ refers to the sequence e_τ and here at any time τ an individual v_i has a specific number of links $e_{i\tau}$, where each individual has a unique target of links t_i to form by the end of this dynamical process. This procedure is considered for a snapshot of time from an original evolutionary process. The set of vertices or individuals in the models represents a group of number of animals (or humans) that will establish social and geographical relations in between them (edges). The desire for a certain level of social contact, or sociability, is relevant to a wide range of animal species living in groups including humans, and typically varies between individuals. Studying sociability, it is of interest to

try to measure the advantages of social relationships and estimate how long such relationships last, which may be influenced by such factors as an individual's age, gender, dominance. Jackson and Wolinsky [43] discussed the influence of human social networks on behaviour and economic effects. Social networks are frequently used to communicate information and allocate goods and services that are not traded in markets. It is possible to list invitations to events and other forms of friendship exchange among these goods, as well as details on available employment opportunities, business ventures, and similar things. Their research included theoretical analysis of the role of social networks in markets and exchange. They studied a case example closely related to ours, specifically to model B and C, individuals are represented as a set of vertices and the relationships in between of them represented as a set of edges. A self-interested individuals had the ability to create or break links, forming a connection requires agreement from both parties, breaking a link does not.

However, they investigated the issue of graph endogeneity in particular scenarios, including cooperative games using the Shapley value, where their interest is in determining which networks are likely to emerge in different contexts. While we concentrated on the network structure and the properties of the final set of graphs obtained by following model B and C that has been termed by the terminal set. Additionally, we concentrated on two particular types of sequences in our games: the arithmetic sequences and the all-or-nothing sequences which have been the key sequences of the initial model A, which we will be building this thesis on. They classified the individuals into target classes from which they may select one individual from each class, taking into account that the dynamical process is a component of the evolutionary process. We explored suggested strategies to investigate the behaviour of the individuals in the terminal set.

Related to this, the theory of biological markets created by Noe and colleagues [60, 61]. They established the properties of human markets by relating them to social systems like the peacock's mating system, in which members of one group (often males, as in the case of peacocks) gain by being picked by members of another (females, peahens, for the peacocks).

Among dolphins [16] two degrees of social alliances were noticed when observing dolphins in Shark Bay, Australia, with 14 out of 400 dolphins forming highly unstable alliances, while others formed more solid alliances. According to the research, the complexity of bottlenose dolphins' social relationships may be related to their large brain size. Variation in social alliances might be due to several factors, such as in male baboons (*Papio Cynocephalus*), where they develop alliances as a conditional strategy, which is mostly utilised by mid-ranking males against high-ranking males. Giraffes [13], tend to develop solid social relations in societies with short path lengths throughout the network.

In sheep [78], a social network was observed and investigated to see the level of their proximity to others when grazing in a group, and whether their tendency to move away from the group was caused by a desire to graze preferred vegetation (long grass). It was discovered that the trade-off between keeping close to one another and grazing further apart on longer grass was relatively small, and so individuals relocating further away mainly reflected other factors, including sociability.

An interesting related problem is the stable marriage problem, see [27]; which is connected to our work but different. They examined a society of two equal-sized sets for men and women, with only heterosexual partnerships taken into account. The marriage will be unstable if one of the spouses could find a better fit than the one they currently have. As a result, we see that the connections between the

pairings fluctuate depending on the individuals' preferences, which is comparable to models A, B, and C in which individuals would break or form links with others based on their preferences until they are on target. Our work would be closer to the marriage problem by only having links allowed on a bipartite graph. Then a difference is that we have a different target number of social relation while their target is often 1.

Chapter 2

Graphical Sequences

In this chapter, we will define the graphical sequence and outline the sufficient and necessary conditions of a sequence to be graphical. This concept will be presented as a theorem (the Havel-Hakimi theorem), and we will discuss several applications of it.

2.1 What is a graphical sequence?

Definition 2.1.1. The distance between the target sequence t and sequence e is given by: $Z(t, e) = \sum_{i=1}^n |t_i - e_i|$.

Definition 2.1.2. The deviation of the set of a graph over vertices $v = v_1, v_2, \dots, v_n$ from the target sequence $t = t_1, t_2, \dots, t_n$ is defined as $Z(e(G), t) = \sum_{i=1}^n |t_i - e_i|$

Definition 2.1.3. The sequence e_1, e_2, \dots, e_n is graphical if you can find a graph G which achieves that sequence exactly.

Definition 2.1.4. $deg(v_i)$ the degree of a vertex (v_i) is the number of links going out of the vertex, which connects the vertex with other vertices.

Definition 2.1.5. A finite sequence $e_1, e_2, e_3, \dots, e_n$ of non-negative integers is called a degree sequence of graph G if the vertices of G , v_1, v_2, \dots, v_n , are such that $\deg(v_i) = e_i$ for all $i = 1, 2, 3, \dots, n$ (see [92, 75]).

Thus, if a sequence $(S) : e_1, e_2, e_3, \dots, e_n$ is a degree sequence of a graph, then the sequence S is called a graphical sequence. Obviously, not every sequence of non-negative integers can be graphical. We can summarize the obvious conditions for a graphical sequence:

- $\sum_{i=1}^n e_i$ must be even.
- $n - 1 \geq e_i \geq 0$.
- The target sequence is finite so the corresponding graph (if it exists) must be finite.

Although these conditions are necessary, they are not sufficient. For example, the sequence 3, 3, 3, 1 has the following:

- $n = 4$
- $\sum e_i = 10$
- $e_i \leq 3$

The sequence satisfies the conditions mentioned above but is not graphical.

Note 2.1.6. We may have more than one graph realise the same sequence for a graphic sequence.

Example 2.1.7. The graphic sequence $t = 2, 2, 2, 2$ has three graphs which achieve it, labelled (a, b, c) in Figure 2.1 below:

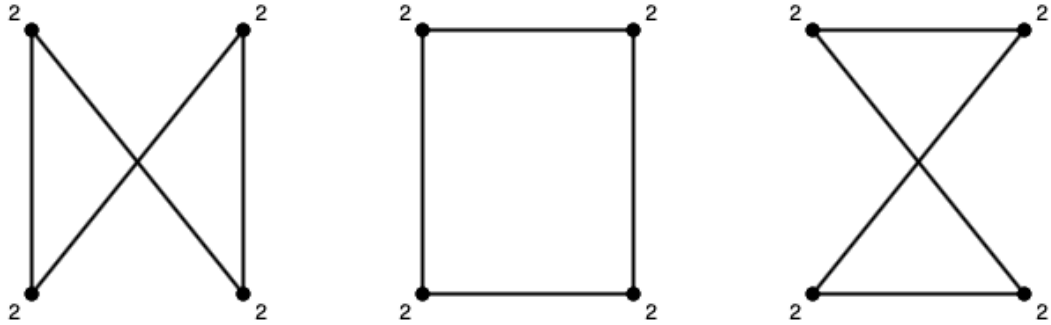


Figure 2.1: Graphs a, b, and c that realise sequence $t=2, 2, 2, 2$.

2.2 The transition graph

Starting the linking formation process described in general in Section 1.11 and further more in Section 3.1 between the individuals will transfer them from one state to another, trying to get the closest possible to the target sequence t until they reach a state of the minimal set X_t , see Section 3.3.1. Broom and Cannings in [9], showed that there is always a path of transitions from a graph state X to another with the same deviation or lower. We will add or break a link between two individuals at each step. For this to happen, one of the two individuals must be moving in the right direction to obtain his target. We can classify the transitions into two types:

1. The degree of both vertices evolve to be closer to their ideal value, so the deviation drops by 2.
2. One vertex moves towards the target, and the other moves away, then the deviation is not affected, and it will stay as it is.

Definition 2.2.1. H_n is the set of all graphical sequences which are non-increasing. H_n^* is the expansion of H_n with all the permutations of these, see [9].

Definition 2.2.2. [9] Assume that the sequences $u, v \in H_n$ and $\exists i, j; u_i = v_i + 1, u_j = v_j + 1$ or $u_k = v_k$. Considering the transitions described earlier, a transition

could happen from u to v if $u_i < t_i$ or $u_j < t_j$, and from v to u if $v_i > t_i$ or $v_j > t_j$. We define the transition graph $T_n = \{H_n^\dagger(t), F_n(t)\}$ where $F_n(t)$ is the set of possible transitions defined above; these edges are directed.

2.3 Sufficient and Necessary conditions for a graphical sequence

Havel-Hakimi Theorem A non-negative sequence $t : t_1, t_2, t_3, \dots, t_n$ of size n , and non increasing integers; $n \geq 2$, $t_1 \geq 1$, is graphic if and only if the sequence $t' = t_2 - 1, t_3 - 1, \dots, t_{t_1+1} - 1, \dots, t_n$ where $t_2 - 1 \geq t_3 - 1 \geq t_{t_1+1} - 1 \dots \geq t_n$ is graphic (see [9], [31], [37]). Here t' is obtained by deleting the largest element t_1 in t and subtracting 1 from the next t_1 elements.

Proof:

If $n = 1$, the statement is trivial (as there is only one vertex, it is not possible to have any edges, so the degree sequence must be zero).

- First we prove sufficiency. Assume $t = (t_2 - 1, t_3 - 1, \dots, t_{k+1} - 1, t_{k+2}, \dots, t_{n-1}, t_n)$ is graphic. So a graph G' realises t' . Thus there are k vertices in G' having degrees $(t_2 - 1, t_3 - 1, \dots, t_k - 1, t_{k+1} - 1)$, we denote this set of vertices v_2, \dots, v_{k+1} by T . Add a new vertex v_1 adjacent to the vertices v_2, \dots, v_{k+1} . By adding v_1 we are adding new edges. A new graph G will be created. Then $\deg(v_1) = k = t_1$, and the degree of each of v_2, \dots, v_{k+1} has increased by 1, so that the degree sequence of G is (t_1, t_2, \dots, t_n) .
- Second we prove necessity. We will produce G' that realises t' , by taking v_1 then by removing it we will get our graph. Suppose $t = (t_1, t_2, \dots, t_n)$ is graphical, in decreasing order and non-negative. Then there will be a graph G with the vertices v_1, v_2, \dots, v_n for which t is the degree sequence, so

$\deg(v_i) = t_i$, for $i = 1, \dots, n$. We will denote the neighbours of v_1 by $N(v_1)$.

1. If $N(v_1) = T$ then the graph G' is obtained by deleting v_1 (and its incident edges) which realise the sequence t' .
2. Suppose that v_1 is not adjacent to all of the vertices $T = (v_2, \dots, v_{k+1})$. We show how to modify the graph G by changing the neighbours of v_1 and change G to get a new graph with the same degree sequence t , where v_1 has one more neighbour in T . We repeat this process until we obtain a graph with degree sequence t . Then deleting v_1 from this graph gives a new graph with degree sequence t' .

If $N(v_1) \neq T$, there is at least one vertex $x \in T$, and $z \notin T$ where v_1 is not adjacent to x and v_1 is adjacent to z . Since $\deg(x) \geq \deg(z)$ there will be some extra vertex $y \in V(G)$ such that y is adjacent to x and y is not adjacent to z . Now we delete the edges (v_1z, xy) and we add the edges (v_1x, yz) . Repeat until this process converts G to G^* where $N(v_1) = T$. Then we delete v_1 from G^* to obtain the desired G' realising t' .

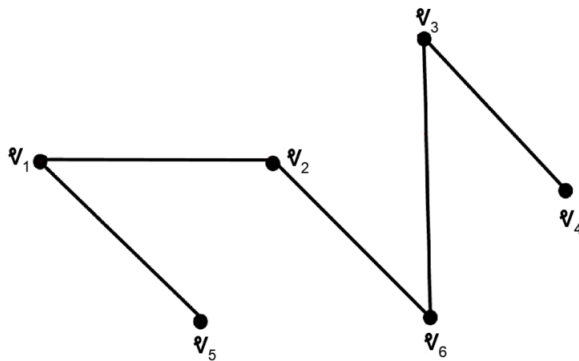


Figure 2.2: Graph G' realises sequence s' before adding vertex w .

Now add the vertex w to the Figure 2.2 which will be adjacent to the vertices v_1, v_2, v_3, v_4 .

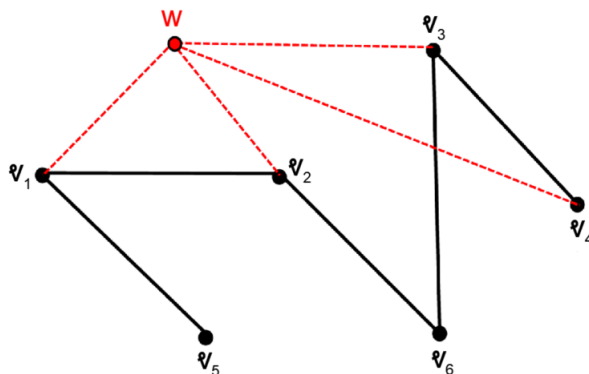


Figure 2.3: Adding vertex w to G'

2.4 Applications of the theorem

Example 2.4.1. Sufficiency Example: Consider the sequence $S = 4, 3, 3, 3, 2, 2, 1$. To apply the H-H theorem we will delete the greatest element in the sequence which is $t_1 = 4$ and we subtract 1 from the next four elements as follows: $t_2 - 1 = 2, t_3 - 1 = 2, t_4 - 1 = 2, t_5 - 1 = 1 \implies S' = 2, 2, 2, 1, 2, 1$.

Reorder the sequence $S' = 2, 2, 2, 1, 2, 1$ in decreasing order to get $2, 2, 2, 2, 1, 1$. We notice that this sequence S' is a graphic sequence that has a graph G' realises it. Then the original sequence S is graphic.

Note 2.4.2. A sequence of zeroes is always graphical as all the vertices achieve their target for the empty graph.

Definition 2.4.3. The Havel-Hakimi algorithm (see [9]) shows how we can test if a sequence is graphical as follows:

1. Sort the sequence in decreasing order.
2. Remove the first (greatest) element. Suppose that the first element was k . Subtract 1 from the k following terms.

3. Repeat this process until we end up with a sequence of zeros (graphic) or a number that cannot be removed without creating a negative entry (not graphic).
4. If a negative number is obtained, stop. Otherwise, repeat from step 1.

Example 2.4.4. Is the sequence 3, 3, 3, 2, 2, 1, 0 graphic or not?

To find the answer we perform following steps:

- Remove the first 3 and then subtract 1 from the next three terms to get 2, 2, 1, 2, 1, 0.
- Write the sequence in decreasing order so we will get 2, 2, 2, 1, 1, 0.
- Remove 2 and subtract 1 from the next two terms to get 1, 1, 1, 1, 0.
- Remove 1 and subtract 1 from the next term to get 0, 1, 1, 0.
- Reorder the sequence 0, 1, 1, 0 in decreasing order to get 1, 1, 0, 0.
- Remove 1 and subtract one from the next term to get 0, 0, 0.

This is a zero sequence which tells us the original sequence is graphical.

Example 2.4.5. We will check if the sequence 3, 3, 3, 1 is graphic by applying the Havel- Hakimi algorithm.

The sequence 3, 3, 3, 1 is graphical \iff 2, 2, 0 is graphical, and the sequence 2, 2, 0 is graphical \iff 1, -1 is graphical, which is clearly false. In general, for non-graphical sequences, we eventually achieve at least one -1 value; thus, it can't be in the degree sequence for any graphical sequence, so the sequence 3, 3, 3, 1 is not graphical.

Example 2.4.6. The sequence 3, 3, 2, 1, 1, 0 is a graphical sequence for the following reason:

$3, 3, 2, 1, 1, 0 \mapsto 2, 1, 0, 1, 0 \mapsto 2, 1, 1, 0, 0 \mapsto 0, 0, 0, 0$.

The sequence $3, 3, 2, 1$ is not graphic for the following reason:

$3, 3, 2, 1 \mapsto 2, 1, 0 \mapsto 0, -1$.

Note that in some versions of the model, the individuals can choose which link to form or break, which makes a big difference in the model. Our models have a specific type and structure, but they relate to other authors' models. They can be thought of as a particular class of the economic models of Jackson and colleagues see [40, 41, 42, 43]

2.5 Vertex classifications

In this section, we will talk about the vertex classifications (See [10]), which represent the individual status at each stage, each of them trying to reach its target by breaking or forming a new connection with other individuals following the random process described in Section 3.3.1.

Graphs that can realise the target sequences are graphic sequences.

Definition 2.5.1. Individual $v_i \in V$ is in deficit when $e_i \neq t_i$.

Definition 2.5.2. The score of target sequence t is $s(t) = \min_G Z(e(G), t)$; $G \in \mathbf{G}$ where \mathbf{G} is the set of all graphs. *i.e.* the score is the minimal total deviation of all graphical sequences from the target sequence t .

We will now define a general concept, the minimal set $K(\min)$.

Definition 2.5.3. $K(\min)$ is the set of graphs which achieve the score.

Definition 2.5.4. $J(\min)$ is the set of sequences of graphs that achieve the score. For each element of $J(\min)$, there will be a corresponding set of at least one element of $K(\min)$.

Here we are interested in non graphic sequences and we will be considering the graphs in $K(\min)$.

If individual i had a number of edges e_i at time τ , considering a target sequence t in which individual i has a target t_i , then we will discuss the following cases:

- vertex i is Neutral if $e_i = t_i$ (it would neither connect to nor break an edge)
- vertex i is a Joiner if $e_i < t_i$ (would wish to form a new edge).
- vertex i is a Breaker if $e_i > t_i$ (it would wish to break one of its edges).

In [10] they classified the vertices into the following sets:

1. Vertices can be either Joiner or Breaker in some elements but necessarily Neutral for some elements of $K(\min)$, those are $\in S_A$.
2. Vertices which are either Joiners or Neutral for all elements of $K(\min)$, those are $\in S_J$.
3. Vertices which are Breakers or Neutral for all elements of $K(\min)$, those are $\in S_B$.
4. Vertices which are Neutrals only, those are $\in S_N$.

Note 2.5.5. by Lemma 4.1 in [10], we cannot have vertices of type (1) and vertices of type (4) in the same sequence.

If the sequence is graphic, then there is a unique element of $K(\min)$, and all of its vertices will be Neutrals. In this case, we cannot have Joiners or Breakers, and all vertices are of type (4).

Chapter 3

A dynamic network population model

In this chapter, we will present the dynamical model introduced in previous work [8], in which players try to accomplish their goals by randomly breaking and forming links with other players. Furthermore, we will determine the minimal set and its properties.

3.1 Modelling populations

Consider a network of interactions between the set of distinguishable vertices $V = (v_1, v_2, v_3, \dots, v_n)$, each of which represents a distinct individual, and the set of distinct edges $X = (x_{ij})$, where $i, j = 1, 2, \dots, n$ represent the social relations between them and each edge connects a pair of distinct individuals (see[10]), with $x_{ij} = 1(0)$ representing the presence (absence) of an edge between individuals i and j . The considered population models are studied over a short snapshot of the entire evolutionary process, hence death and birth are not taken into account. For the considered period of time, each individual has unique number of social re-

quirements, which are represented by the target; this target vary from individual to another depending on many factors, including age. For example, people emphasise emotionally strong relationship and report smaller social networks as they age, showing increasing of social selectivity [14, 48]. Similar findings have been noted in primates [88, 74], whales [91], rodents [47, 93] and deer [3], while elderly animals interact with fewer social partners and spend less time in social bonding. Sine the considered time is relatively short compared to the evolutionary process, an individual's target will remain constant over the examined period of time and will not vary. However, considering the entire evolutionary process, the target number of social interactions may vary. A vital fitness element in the evolutionary process is the total number of offspring a single individual potentially produce in its lifetime. In insects, female reproductive capacity is largely determined by the number of ovarioles. However, that will not have any impact on our models given the short time frame we are considering. The strategic interaction take place in discrete time thus, we have a time-dependent matrix $X(\tau)$ describing the links, and this in turn generates the vector $e_\tau = (e_{1\tau}, e_{2\tau}, \dots, e_{n\tau})$, the *sequence* e_τ of $X(\tau)$, where at time τ an individual i has $e_{i\tau}$ links. In this thesis, each vertex will have a target number of desired links it would like to have by the end of the linking formation process t_i . Following a dynamic process, described precisely in Section 3.2 for the unique target case, see [8], individuals are selected at random to change their number of links, where an individual below (above) their target would form (break) a link/ links at random. Individuals on target would make no changes. Individuals (vertices) are fixed, while the links in between them change randomly. An individual i will try to get as close as possible to the target t_i by breaking or forming random links with other group members, this follows the Markov chain process. The presence or absence of a link may benefit both, neither or one of the

two individuals involved, depending on their desired total number of links. If the *target sequence*, can be achieved by a graph, then it is possible that all individuals can be on target simultaneously, and no further changes can be made, the target sequence is called a *graphical sequence* see Section 2. A lot of work has been done on graphical sequences, for example see: [35, 31, 37, 53, 75]. In our studies as well as general sequences, we concentrated on two types of sequences in order to build on the previous work [?] that was focused on them and for simplicity. Those sequence are:

Definition 3.1.1. The n -element *arithmetic sequence* has the form $(n - 1, n - 2, n - 3, \dots, 1, 0)$ for general n .

Definition 3.1.2. An *all or nothing sequence* is a sequence of n elements, m_1 of which have target $n - 1$ and $m_2 = n - m_1$ of which have target 0.

Future study can further investigate other sequence categories, such as sequences with all targets equal to one and indeed generic sequences.

3.2 The Markov Chain Model

In Section 3.1 we denoted the set of links by the matrix \mathbf{X} , where $x_{ij} = x_{ji} = 1$ if v_i and v_j are linked, and $x_{ij} = x_{ji} = 0$ otherwise. A sequence of moves occurs, where a random individual is selected to potentially update their set of links. Each move starts with an individual being selected to do the update at random (i.e. with probability $1/n$). It has n distinct (pure) choices, i.e. it can change any given edge (remove if present or add if absent) with any of the other $n - 1$ individuals, or to make no change. Thus, a number of moves might lead the individual from multiple beginning states to the same final state. The probabilities of future events depend on the current state of the system. The events are connected one to the other; they

form a Markov chain see [83, 70]. The system must have a set of distinct states, with detectable transitions between them. As in [10, 12], we denote the probability that individual i chooses to change edge x_{ij} , conditional on it being selected to be the updating individual, by u_{ij} , with u_{ii} representing the probability of no change. Given that there are n individuals that could be chosen, the strategies for any situation can be written as an $n \times n$ matrix \mathbf{U} , where all row sums equal 1. Our strategy matrix \mathbf{U} depends upon the current state \mathbf{X} , and so the full set of strategies is denoted by $\mathbf{U}_{\mathbf{X}}$. For \mathbf{x}^* which differs from \mathbf{x} at a single position ($x_{ij} = 0, x_{ij}^* = 1$ or $x_{ij} = 1, x_{ij}^* = 0$) for a specific i, j ,

$$P(\mathbf{X}_{\tau+1} = \mathbf{x}^* | \mathbf{X}_{\tau} = \mathbf{x}) = \frac{u_{ij}(\mathbf{x}) + u_{ji}(\mathbf{x})}{n}.$$

v_i will have a number of edges $e_{i\tau}$ to other individuals, so that the above process generates an evolving sequence $e_{\tau} = (e_{1\tau}, e_{2\tau}, e_{3\tau}, \dots, e_{n\tau})$. In particular in [?] a model where there were no strategy choices except to always try to improve the deviation for any given individual was introduced. Thus any individual below (above) target added (removed) a link, selecting out of the available choices with equal probability. For the case $x_{ij} = 0, x_{ij}^* = 1$ this led to

$$P(X_{\tau+1} = x^* | X_{\tau} = x) = \begin{cases} \frac{1}{n} \frac{1}{n-1-e_i} + \frac{1}{n} \frac{1}{n-1-e_j} & e_i < t_i, e_j < t_j \\ \frac{1}{n} \frac{1}{n-1-e_i} & e_i < t_i, e_j \geq t_j \\ \frac{1}{n} \frac{1}{n-1-e_j} & e_i \geq t_i, e_j < t_j \\ 0 & e_i \geq t_i, e_j \geq t_j \end{cases}$$

and for the case $x_{ij} = 1, x_{ij}^* = 0$

$$P(X_{\tau+1} = x^* | X_{\tau} = x) = \begin{cases} \frac{1}{n} \frac{1}{e_i} + \frac{1}{n} \frac{1}{e_j} & e_i > t_i, e_j > t_j \\ \frac{1}{n} \frac{1}{e_i} & e_i > t_i, e_j \leq t_j \\ \frac{1}{n} \frac{1}{e_j} & e_i \leq t_i, e_j > t_j \\ 0 & e_i \leq t_i, e_j \leq t_j \end{cases}$$

Example 3.2.1. Consider the target sequence $t = 2, 1, 0$. Here $n = 3$, and we will have $2^3 = 8$ distinct graphs in total which are shown in Table 3.1 below:

a=2,1,0							
number	v_1	v_2	v_3	0-1	0-2	1-2	sequence
1	N	B	B	Y	Y	Y	2,2,2
2	J	B	B	Y	N	Y	1,2,1
3	N	N	B	N	Y	Y	2,1,1
4	J	N	B	Y	Y	N	1,1,2
5	J	N	B	Y	N	N	0,1,1
6	J	J	B	N	Y	N	1,0,1
7	J	N	N	N	N	Y	1,1,0
8	J	J	N	N	N	N	0,0,0

Table 3.1: The eight possible graphs involving individuals v_1, v_2, v_3 with a target $(2, 1, 0)$ respectively. Columns (2-4) show the vertex type at that specific graph (Neutral, Breaker, Joiner), the next three columns (5 – 7) tells us if there is a link between the vertices or not (Yes, No) and the final column shows the sequence of the graph.

The transition probabilities between the eight graphs are given by the following matrix:

$$P = \begin{pmatrix} 1/3 & 1/6 & 1/3 & 1/6 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 0 & 1/6 & 0 & 1/2 & 0 \\ 0 & 0 & 2/3 & 0 & 0 & 0 & 1/3 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/6 & 1/6 & 0 & 0 \\ 0 & 1/6 & 0 & 1/6 & 1/3 & 0 & 0 & 1/3 \\ 0 & 0 & 1/2 & 1/6 & 0 & 0 & 0 & 1/3 \\ 0 & 0 & 1/3 & 0 & 0 & 0 & 2/3 & 0 \\ 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/3 & 1/3 \end{pmatrix}$$

We will find the deviation of the eight graphs and the associated sequences shown in the table as follows:

$$d_1 = |(2 - 2)| + |(2 - 1)| + |(2 - 0)| = 3 \text{ for the sequence } 2, 2, 2$$

$$d_2 = |(1 - 2)| + |(2 - 1)| + |(1 - 0)| = 3 \text{ for the graphic sequence } 1, 2, 1$$

$$d_3 = |(2 - 2)| + |(1 - 1)| + |(1 - 0)| = 1 \text{ of the sequence } 2, 1, 1$$

$$d_4 = |(1 - 2)| + |(1 - 1)| + |(2 - 0)| = 3 \text{ for the sequence } 1, 1, 2$$

$$d_5 = |(0 - 2)| + |(1 - 1)| + |(1 - 0)| = 3 \text{ of the sequence } 0, 1, 1$$

$$d_6 = |(1 - 2)| + |(0 - 1)| + |(1 - 0)| = 3 \text{ for the sequence } 1, 0, 1$$

$$d_7 = |(1 - 2)| + |(1 - 1)| + |(0 - 0)| = 1 \text{ for the sequence } 1, 1, 0$$

$$d_8 = |(0 - 2)| + |(0 - 1)| + |(0 - 0)| = 3 \text{ for the sequence } 0, 0, 0$$

By observing the deviations of the sequences we note that we have sequences with two different deviations (3 and 1). The score of the sequence 2, 1, 0 is the minimal value of the deviation which equals to 1.

Thus $K(\min)$ here contains two elements (those with minimal score, d_3 and d_7).

Note 3.2.2. To find the minimal deviation we have considered all the possible graphs of n individuals, this is very computationally expensive and thus, we will introduce a simpler method called the adjusted Havel-Hakimi algorithm.

3.3 The minimal set

As the transitions always move from one state to another with a lower or the same score, it may take many steps for the system to reach one of the graphical sequences u where $d(u, t)$ is minimal, i.e. it belongs to the set of all sequences which realises the minimal deviation $J(\min) = \{u \in H_n, Z(u, t) = s(t)\}$, where $s(t)$ is the score of sequence u of target t and $J(\min)$ is the minimal deviation of a sequence u with a target t , (see [10],[9] for more details).

3.3.1 Properties of the minimal set

In this model, individuals will perform many state transitions to reduce their deviation until the minimal set $J(\min)/K(\min)$ is reached when no further movement can occur. Then we have the following:

1. There will be no links between two Breakers.
2. Any two Joiners will have a link between them.
3. If a Neutral node is joined to a Breaker node, it means that the Neutral node is linked to every Joiner else. Similarly, if a Neutral is not joined to some Joiners that means it is not joined to any Breakers.
4. If we have a node $v \in S_N$ that is linked to nodes $\in S_B$ and another node $s \in S_N$ is joined to nodes $\in S_B$ then v and s are joined.
5. If we have a node $v \in S_N$ is not joined to some Joiners and another node $s \in S_N$ is not joined to some other Joiners then v and s are not joined.

Note 3.3.1. $J(\min)$ contains at least one sequence with no Joiners and at least one sequence with no Breakers.

$J(\min)$: is the Minimal Set of t (see [9] for further details)

Note 3.3.2. For any minimal graph the Joiners come before the Breakers in terms of the ordered target as any Joiner will have a degree at least as large as any Breaker, and because a Joiner's target is greater than its degree and a Breaker's is less, the Joiner's target is larger than the Breaker's. See [9, 10].

Example 3.3.3. Considering the sequence 2, 1, 0 from example 3.2.1 and Table 3.1. The following table will list the sequences of the minimal set with vertex classifications.

t=2, 1, 0							
number	v_1	v_2	v_3	0-1	0-2	1-2	sequence
3	N	N	B	Y	Y	N	2,1,1
7	J	N	N	Y	N	N	1,1,0

Table 3.2: The minimal set involving individuals v_1, v_2, v_3 with a target (2, 1, 0) respectively. The columns (2 – 4) show the vertex type at that specific graph (Neutral, Breaker, Joiner). The next three columns (5 – 7) tell us if a link exist between the vertices or not and the final column shows the sequence of the graph.

Definition 3.3.4. In [9] the adjusted Havel-Hakimi algorithm says: we can find the score of a sequence by following the steps below:

- Sort the sequence in decreasing order.
- If the first (greatest) element is k remove it. Subtract 1 from the k following terms.
- If a negative number appears (only -1 is possible) turn it to zero and add one to the score.
- Repeat this process until we end up with a sequence of zeros.

Definition 3.3.5. The Havel Hakimi score

In [9] we denote the $H - H$ score by $SHH(t)$ of a sequence (t) . The $SHH(t)$

defined as the total number of added 1's during the process of the adjusted $H - H$ algorithm.

Example 3.3.6. Find the $SHH(t)$ of the following sequence 3, 2, 1, 0.

1. The sequence 3, 2, 1, 0 is already sorted in decreasing order
2. Delete the greatest term (3) then subtract 1 from the next 3 terms to get 1, 0, 0(+1).
3. Delete the biggest term (1) and then delete (1) from the next term. That will lead to a negative number (-1) in the resulting sequence, so we will add it up to zero and the score will increase by (+1). Thus, the score in this case is 2.

Here we added (1) to the (-1) to get zero in terms for us to apply the adjusted H-H algorithm to find the score and to avoid negative numbers.

Example 3.3.7. The target 4, 3, 2, 1, 0 has a score 2, with eight elements in $K(\min)$. The elements of $K(\min)$ represent graphs, we write them below in the form of their corresponding sequences, distinguishing between the two graphs that have the same sequence. We label them G_1 to G_8 in the following order $\{4, 3, 2, 2, 1\}$, $\{3, 3, 2, 2, 0\}$, $\{3, 3, 2, 1, 1(1)\}$, $\{2, 3, 2, 1, 0\}$, $\{4, 3, 2, 1, 2\}$, $\{3, 3, 2, 1, 1(2)\}$, $\{4, 2, 2, 1, 1\}$ and $\{3, 2, 2, 1, 0\}$, the state $\{3, 3, 2, 1, 1(1)\}$ is the graph where the vertex with target 4 is linked to the vertex with target 0, (see this example in [10]).

3.4 The deficit

For a target sequence $t = (t_i); i = 1, \dots, n$, the conjugate vector $v = (v_i)$ of t was defined by $v_i = \#j : t_j \geq i$ (where $\#$ means “the number of”). There, and in the working below, we define f_k as $f_k = \sum_{i=1}^{\mu} (t_i + 1 - v_i)$, with $f_0 = 0$.

Definition 3.4.1. For a target sequence $t = (t_i); i = 1, \dots, n$, the number of targets for which $t_i \geq i$, denoted by λ , is called the Durfee number.

A sequence is graphic if and only if the sequence sum is even and $f_k \leq 0$ for $k = 1, \dots, \lambda$, [9].

Definition 3.4.2. The deficit of the sequence t was defined as the summation $\sum_i^\lambda d_i$ where $\lambda \neq i : t_i \geq i$ is the Durfee number see [9], [10].

Definition 3.4.3. The deficit of the sequence t is $Max_{0 \leq k \leq \lambda} f_k$.

Note 3.4.4. Its worth mentioning that the deficit defined in this section is distinct to the (in deficit) concept defined earlier in Definition 2.5.1.

Example 3.4.5. Find the deficit of the sequence 6, 6, 3, 3, 3, 0, 0.

$t_1 = 6, t_2 = 6, t_3 = 3, t_4 = 3, t_5 = 3, t_6 = t_7 = 0$, so that

$v_1 = 5, v_2 = 5, v_3 = 5, v_4 = 2, v_5 = 2, v_6 = 2, v_7 = 0$.

Now we will work out $f_k = (t_i + 1 - v_i)$

$$1) t_1 + 1 - v_1 = 2$$

$$2) t_2 + 1 - v_2 = 2$$

$$3) t_3 + 1 - v_3 = -1$$

We stop here as the Durfee number = 3.

Now to find the deficit we have $d = Max_{0 \leq k \leq \lambda} f_k$

$$f_0 = 0,$$

$$f_1 = f_0 + f_1 = 0 + 2 = 2,$$

$$f_2 = f_2 + f_1 + f_0 = 2 + 2 + 0 = 4,$$

$$f_3 = f_3 + f_2 + f_1 + f_0 = -1 + 2 + 2 = 3.$$

Thus, the deficit is $d = 4$.

Lemma 3.4.6. The score is equal to the deficit (the deficit plus 1) if and only if

$\sum_{i=\mu+1}^n t_i + \mu^2 - \sum_{i=1}^\mu \min(v_i, t_{i+1})$ is even (odd), where $\mu \leq \lambda$ is the peak (see [10]).

Example 3.4.7. The target sequence 7, 6, 3, 2, 2, 2, 0, 0 has deficit 4 and score 4.

The target sequence 7, 6, 3, 3, 3, 0, 0, 0 has deficit 5 and score 6.

Chapter 4

The size of the minimal set

This chapter will consider the minimal set(s) in more details. In particular, we are interested in the size of the minimal set for two special types of sequences and the size of the largest potential minimal set in general for a given number of vertices n .

4.1 The largest minimal set

We consider two types of the minimal set, namely $J(\text{min})$ and $K(\text{min})$. We will mainly consider $J(\text{min})$ for the following reasons:

1. Dealing with $J(\text{min})$ is simpler than $K(\text{min})$, partly due to the difference in the size of the minimal sets. Every sequence (in $J(\text{min})$ and in general) has at least one graph (in $K(\text{min})$ and in general), which makes the size of $K(\text{min})$ is greater or equal to that of $J(\text{min})$. The number of graphs for n vertices is $2^{\binom{n}{2}}$. So asymptotically, the logarithm of this number is of order n^2 , whilst the number of sequences is bounded above by n^n , the logarithm of which is of order $n \ln(n)$ (many sequences are not graphical and ties in the target further reduce this).

2. The minimal set is defined in terms of sequences, not graphs. We saw this from the definition of the score and the fact that it can be found using the modified Havel-Hakimi algorithm [9]. Whether any given graph/sequence achieves the score is thus established by considering the sequence. When is the minimal set likely to be large? It might be thought that $J(\min)$ should be at its largest when the score is the largest. What sequences yield this maximal score? We show this in the theorem below. Note that, although the main focus of this paper is $J(\min)$, as discussed above, we shall consider $K(\min)$ below; as for the particular sequences involved, this is actually an easier set to find. The maximal score $ms(n)$ for n vertices occurs as follows.

Theorem 4.1.1. *1. If $n = 2m$, the maximal score occurs for target $\{(n - 1)^m, 0^m\}$, where x^y denotes a list of y x 's, and the score is $ms(n) = m^2$.*

2. If $n = 2m + 1$ the maximal score occurs for target $\{(n - 1)^{m+1}, 0^m\}$ (and for its dual $\{(n - 1)^m, 0^{m+1}\}$), and the score is $ms(n) = m(m + 1)$. In each case the size of the minimal set $K(\min)$ is $2^{ms(n)}$.

Proof: For given target \mathbf{t} (elements listed in descending order) there is at least one graph that achieves the score. Choose such a graph, and count the edges of each vertex, listing these in vector \mathbf{s} . Clearly the elements of \mathbf{s} are non-increasing (except perhaps in tied positions from the target, and so there is an equivalent sequence where they are non-increasing), otherwise swapping the links between some pair of vertices reduces the deviation, which is not possible since the graph achieves the score. It is shown in Theorem 4.11 from [10] that for our graph, any vertices that are short of target (Joiners) must precede those achieving their target (Neutrals) which in turn must precede those over target (Breakers). Now consider the sequence where we replace the target for all Joiners by $n - 1$, the target for all Breakers by 0, and the target for Neutrals by some number of $n - 1$ s preceding

some number of 0s. This has a larger score than \mathbf{t} (unless it is already of this type). Thus the maximum score sequence(s) must be of form $\{(n-1)^x, 0^{n-x}\}$. This sequence has score $x(n-1) - x(x-1) = x(n-x)$. It is easy to see this by connecting all pairs of $n-1$ vertices and splitting all pairs of 0 vertices; the existence (or otherwise) of a link between any 0 and $n-1$ pair has no effect on the deviation. Thus the largest possible scores are given by the targets stated in the theorem. Now consider the size of the minimal set $K(\min)$. For $n = 2m$ consider $\{(m-1)^m, 0^m\}$ which is graphic, achieved by a unique graph with a subgraph K_m , a subgraph with m vertices and no edges, and no edges between the two subgraphs. The score of our target being m^2 , this graphic sequence is a member of the minimal set for our target sequence. We can add any set of r edges between the two subgraphs and obtain a graph in this minimal set. Adding any edges between the 0 vertices or removing any edges from between the other vertices increases the deviation, meaning that any resulting graph is not in the minimal set. Thus we have 2^{m^2} states in the minimal set. For $n = 2m + 1$ we consider the graph $\{m^{m+1}, 0^m\}$, with subgraph K_{m+1} , which is graphic and a member of the minimal set of target $\{(2m)^{m+1}, 0^m\}$. Using the same reasoning as above, we obtain a minimal set of size $2^{m(m+1)}$. \square .

We note, however, that the largest $K(\min)$ does not necessarily occur for the sequence with the largest score. Consider the $n = 3$ case. From Figure 4.1 we see that there are six states within the minimal set for the sequence 1, 1, 1; however, the sequences with the maximal score are 220 and 200 with four states in the minimal set.

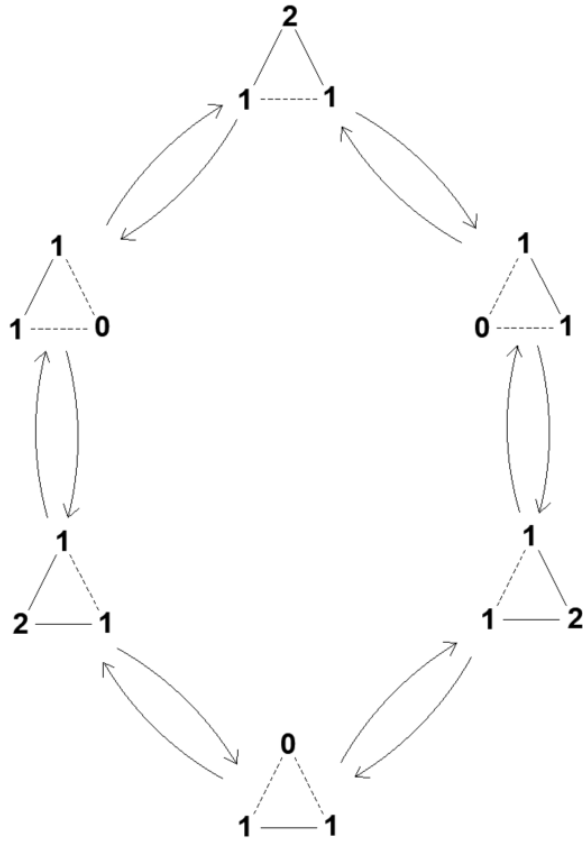


Figure 4.1: Transitions between the elements of the minimal set of graphs $K(min)$ of the sequence 1, 1, 1 showing every possible move between each of the six elements.

— represents a link between the corresponding two nodes,
 - - - represents the absence of such a link.

Sequences of this type have been considered before in [10]. Markov process defined in Section 3.2 was considered, and the stationary distribution over the population states (the set of recurrent states, which is $K(min)$ and which formed a single class) in this case was found. We noticed that the corresponding process of the strategic game would be much more difficult to analyse, because of the large number of states and possible transition distributions between these states (which occurs depending upon strategic choices); all $ms(n)$ edges represent a change that would improve the deviation of one of the two participating individuals. We note that although the

above "all or nothing" targets have a large minimal set of graphs, their minimal set of sequences is not necessarily large, due to the large number of ties within the target (there are only two distinct values, $n - 1$ and 0). Thus reordering within the sequences as discussed in Section 2.5 means that many apparently distinct sequences are actually identical. Thus it will not be the case generally that sequences with the largest score will have the largest minimal set $J(\min)$. But which sequences will have the largest $J(\min)$? This is a question that we cannot answer in general, though we can investigate it for the smallest number of individuals. In particular we have found the minimal sets for all sequence lengths for $2 \leq n \leq 7$ using a Matlab program see Appendix 9.2.

n	score	$ J(\min) $	sequence	Ar $ J(\min) $	Ar score	AoN $ J(\min) $	AoN score
2	1	2	(1,0)	2	1	2	1
3	2	3	(2,2,0)*	2	1	3	2
3	2	3	(2,0,0)*				
4	3	7	(3,2,0,0)	7	2	7	4
4	3	7	(3,3,1,0)				
4	4	7	(3,3,0,0)*				
4	2	7	(3,2,1,0) ⁺				
5	4	20	(4,3,1,0,0)	7	2	13	6
5	4	20	(4,4,3,1,0)				
6	7	84	(5,5,4,1,0,0)	30	3	34	9
7	9	262	(6,6,5,1,1,0,0)	30	3	76	12
7	9	262	(6,6,5,5,1,0,0)				

Table 4.1: The sequences with the biggest minimal set size where n is the number of individuals, $|J(\min)|$: refers to the minimal set size. We also consider two special classes of sequence, the arithmetic (Ar, the sequence ⁺ above is an example) and all-or-nothing (AoN, sequences * above are examples) sequences discussed later, for comparison.

In Table 4.1 we show the size of the largest minimal set, the corresponding target sequence and its score. We also show the size of the minimal set and the score for two special classes of sequences, the all or nothing and the arithmetic sequences.

Example 4.1.2. Consider the following sequences

1. 5, 5, 5, 0, 0, 0
2. 5, 5, 4, 1, 0, 0
3. 5, 4, 3, 2, 1, 0

The sequence 5, 5, 5, 0, 0, 0 has a score= 9, which is the largest score, the sequence 5, 4, 3, 2, 1, 0 has a score= 3 which is the smallest score, and the sequence 5, 5, 4, 1, 0, 0 has a score= 7, an intermediate value of the score. Thus we might think that the sequence 5, 5, 5, 0, 0, 0 should have the largest minimal set out of the three, and sequence 5, 4, 3, 2, 1, 0 must have the smallest one. The first sequence has the top three numbers, and the last three numbers are all the same; the second sequence has the top two numbers, and the last two numbers are the same, but the middle two numbers take different values; the third sequence has all the numbers different from each other. Considering the ties in each target sequence, then when we find the minimal set, we must reorder the first three numbers and the last three numbers of the sequences belonging to the minimal set of the target sequence 5, 5, 5, 0, 0, 0. We must reorder the first two numbers and the last two numbers of the sequences belonging to the minimal set of the target sequence 5, 5, 4, 1, 0, 0. The target sequence 5, 4, 3, 2, 1, 0 does not have any tie. Thus the number of the sequences that are equivalent to the target sequence (1) will be more the ones that are equivalent to the target sequence (2), and hence the target sequence (1) will have more elements to be omitted from the eventual sequences of the minimal set. Thus $J(\min)$ of the target sequence 2 is larger than the one of the target sequence 1; however, its score is smaller. In fact, it has the biggest minimal set for the $n = 6$ case. In general, something similar occurs for all values of n , as for all or nothing sequences, the number of ties is increasing. We have also seen

in the above example and from Table 4.1, that the arithmetic sequence, which is the only type with no ties, also does not achieve the maximum minimal set size. It's score is relatively low, and we hypothesise that in general a combination of high score and low number of ties is what gives a large minimal set. Note that we see in Section 4.3 that the minimal set $J(\min)$ of the all or nothing sequences are generally bigger than the arithmetic one (although we have not been able to prove this). In the following two sections we consider the two special classes of sequences, which are the arithmetic and all or nothing sequences.

4.2 The minimal set for the arithmetic sequence

Definition 4.2.1. The n -element *arithmetic sequence* has the form $(n - 1, n - 2, n - 3, \dots, 1, 0)$ for general n .

In this section we consider the n -element arithmetic sequence $(n - 1, n - 2, n - 3, \dots, 1, 0)$ for general n as mentioned in Definition 4.2.1. In particular, we demonstrate an exact formula for the size of the minimal set $J(\min)$ for this sequence, in the process identifying its precise membership. We will label the set as J_n and thus the size of the set as $|J_n|$. We know that the set $n = 2m + 1$ has the same size as that for $n = 2m$, and that the elements of the set for $n = 2m + 1$ can easily be derived from those for $n = 2m$ (see [10] Theorem 4.2). Thus we only need to solve the problem for $n = 2m$ for general m . We can find the score of $n = 2m$ by the following lemma:

Lemma 4.2.2. *The score for $n = 2m$ is m .*

Proof: $2m - 1, 2m - 2, \dots, m, m - 1, \dots, 0$ is an arithmetic sequence. Applying the adjusted H-H algorithm on this sequence as follows, we obtain:

$$2m - 1, 2m - 2, \dots, m - 1, \dots, 0 (+1)$$

$2m - 3, 2m - 4, \dots, 0, 0 (+1)$

$2m - 5, 2m - 6, \dots, 0, 0, 0 (+1)$

and so on. After the r th step, we obtain the sequence: $2m - 2r - 1, 2m - 2r - 2, \dots, 1, 0, \dots, 0$ where there are r zeros with a cumulative total of r 1's added. The leading term reduces by 2 for each further step with an addition of 1 to the H-H algorithm's value. We end with a sequence of m zero's with a value of m . Thus the score is m . \square

Lemma 4.2.3. *The members of the minimal set for $n = 2m$ have the first m elements in S_J and the last m elements in S_B , in the terminology of [10].*

Proof: We shall divide the sequence $2m - 1, 2m - 2, \dots, m, m - 1, \dots, 0$ into two sets:

$S_1 = 2m - 1, 2m - 2, \dots, m$ and $S_2 = m - 1, m - 2, \dots, 0$.

We further denote:

β_i : The links within S_1 (or S_2) for $i = 0, 1, 2, \dots, 2m - 1$,

α_i : The links between the sets S_1 and S_2 for $i = 0, 1, 2, \dots, 2m - 1$.

1. The deviation of the sequence, written as D , can be found as follows.

$$\begin{aligned}
 D &= \sum_{i=m}^{2m-1} |i - \alpha_i - \beta_i| + \sum_{i=0}^{m-1} |i - \alpha_i - \beta_i| \geq \\
 &\sum_{i=m}^{2m-1} (|(i - \alpha_i)| - |\beta_i|) + \sum_{i=0}^{m-1} (|\beta_i + \alpha_i| - |i|) \geq \\
 &\sum_{i=m}^{2m-1} ((i - \alpha_i) - \beta_i) + \sum_{j=0}^{m-1} (\beta_j + \alpha_j - j) = \\
 &\sum_{i=m}^{2m-1} i - \sum_{i=m}^{2m-1} \alpha_i - \sum_{i=m}^{2m-1} \beta_i + \sum_{i=0}^{m-1} \beta_i + \sum_{i=0}^{m-1} \alpha_i - \sum_{i=0}^{m-1} i.
 \end{aligned}$$

We know that

$$\sum_{i=0}^{m-1} \alpha_i = \sum_{i=m}^{2m-1} \alpha_i,$$

since the α_i s represent the links between S_1 and S_2 and so the summation is common between the two sets. Thus

$$\begin{aligned} D &\geq \sum_{i=0}^{2m-1} i - 2 \sum_{i=0}^{m-1} i - \sum_{i=m}^{2m-1} \beta_i + \sum_{i=0}^{m-1} \beta_i = \\ &(2m(2m-1)/2) - (2m(m-1)/2) - \sum_{i=m}^{2m-1} \beta_i + \sum_{i=0}^{m-1} \beta_i = \\ &m^2 - \sum_{i=m}^{2m-1} \beta_i + \sum_{i=0}^{m-1} \beta_i. \end{aligned}$$

The only way to achieve the score m is to have that all $\beta_i = m - 1$ for $i \geq m$ and all $\beta_i = 0$ for $i < m$.

Thus we have proved that to be in the minimal set we must have $\beta_i = m - 1$ in S_1 and $\beta_i = 0$ in S_2 .

2. This then leaves the target number of links between the elements of S_1 and S_2 , as $i = 0, 1, \dots, m - 1$ for S_1 and $j = 1, 2, \dots, m$ for S_2 . We will substitute for the values of β_i for all nodes as follows:

$$D = \sum_{i=m}^{2m-1} |(i - \alpha_i) - (m - 1)| + \sum_{i=0}^{m-1} |(i - \alpha_i) - 0|.$$

Denoting $j = i - m + 1$, we have

$$D = \sum_{j=1}^m |\alpha_{j+m-1} - j| + \sum_{i=0}^{m-1} |\alpha_i - i|.$$

Now suppose that for some k such that $0 \leq k \leq m$ we have $k < \alpha_{k+m-1}$,

then

$$\begin{aligned}
D &= \sum_{j=1}^m |\alpha_{j+m-1} - j| + \sum_{i=0}^{m-1} |\alpha_i - i| = \\
&(\alpha_{k+m-1} - k) + \sum_{j \neq k}^m |j - \alpha_{j+m-1}| + \sum_{i=0}^{m-1} |i - \alpha_i| \geq \\
2(\alpha_{k+m-1} - k) - \sum_{j=1}^m \alpha_{j+m-1} + \sum_{j=1}^m |j| + \sum_{i=0}^{m-1} \alpha_i - \sum_{i=0}^{m-1} |i| = \\
2(\alpha_{k+m-1} - k) + (m(m+1)/2) - (m(m-1)/2) &= 2(\alpha_{k+m-1} - k) + m > m.
\end{aligned}$$

That means that the sequence is not in the minimal set, and so $k \geq \alpha_{k+m-1}$ i.e. no element in S_1 can be a Breaker, i.e. all elements of S_1 are in S_J .

3. For $1 \leq k \leq m$ we will now assume that $\alpha_k < k$.

$$\begin{aligned}
D &= \sum_{j=1}^m |\alpha_{j+m-1} - j| + \sum_{i=0}^{m-1} |\alpha_i - i| = \\
&\sum_{j=1}^m |\alpha_{j+m-1} - j| + (k - \alpha_k) + \sum_{i \neq k}^{m-1} |\alpha_i - i| \geq \\
\sum_{j=1}^m |j| - \sum_{j=1}^m |\alpha_{j+m-1}| + 2(k - \alpha_k) + \sum_{i=1}^{m-1} |\alpha_i| - \sum_{i=1}^{m-1} |i| = \\
2(k - \alpha_k) + m &> m,
\end{aligned}$$

so again that means the sequence is not in the minimal set. Thus $\alpha_k \geq k$ i.e. no element in S_2 can be a Joiner, i.e. all elements of S_2 are in S_B \square .

Lemma 4.2.4. *If a member of the minimal set has elements where for $i < j$ in S_J we have $t_i \geq t_j$ and $e_i \leq e_j$, then the sequence with these two numbers swapped, so vertex i (j) has element e_j (e_i), is also in the minimal set. Similarly for any two members of S_B .*

Proof: Consider the sequences $S : e_1, e_2, \dots, e_i, \dots, e_j, \dots, e_n$

and $S' : e_1, e_2, \dots, e_j, \dots, e_i, \dots, e_n$, the same sequence with e_i and e_j swapped.

We will define

$$D_s = \sum_{i=0}^n |t_i - e_i|, \quad D'_s = \sum_{k \neq i \text{ or } j}^n |t_k - e_k| + |t_i - e_j| + |t_j - e_i|.$$

We thus have that

$$D'_s - D_s = |t_i - e_j| + |t_j - e_i| - |t_i - e_i| - |t_j - e_j|.$$

We wish to prove that $D'_s - D_s \leq 0$. To do that we will discuss the following cases:

1. If $e_j \geq e_i \geq t_i \geq t_j$

$$D'_s - D_s = (e_j - t_i) + (e_i - t_j) - (e_i - t_i) - (e_j - t_j) = 0, \text{ so } D'_s = D_s.$$

2. If $e_j \geq t_i \geq e_i \geq t_j$

$$D'_s - D_s = (e_j - t_i) + (e_i - t_j) - (t_i - e_i) - (e_j - t_j) = 2e_i - 2t_i \leq 0, \text{ so}$$

$$D'_s \leq D_s.$$

3. If $e_j \geq t_i \geq t_j \geq e_i$

$$D'_s - D_s = (e_j - t_i) + (t_j - e_i) - (t_i - e_i) - (e_j - t_j) = 2t_j - 2t_i \leq 0, \text{ thus}$$

$$D'_s \leq D_s.$$

4. If $t_i \geq e_j \geq e_i \geq t_j$

$$D'_s - D_s = (t_i - e_j) + (e_i - t_j) - (t_i - e_i) - (e_j - t_j) = -2e_j + 2e_i \leq 0, \text{ thus}$$

$$D'_s \leq D_s.$$

5. If $t_i \geq e_j \geq t_j \geq e_i$

$$D'_s - D_s = (t_i - e_j) + (t_j - e_i) - (t_i - e_i) - (e_j - t_j) = -2e_j + 2t_j \leq 0, \text{ thus}$$

$$D'_s \leq D_s.$$

6. If $t_i \geq t_j \geq e_i \geq e_j$

$D'_s - D_s = (t_i - e_j) + (t_j - e_i) - (t_i - e_i) - (t_j - e_j) = 0$, thus $D'_s = D_s$ in this case. Thus $D'_s - D_s \leq 0$ in all cases, so that the result is true. \square .

Lemma 4.2.5. *From Lemmas 4.2.2, 4.2.3 and 4.2.4 we have that the vertex deviations, denoted ϵ_i in the terminology of [10], must satisfy $\epsilon_i \leq m + 1 - i$ for $i \leq m$ and $\epsilon_i \leq i - m$ for $i > m$ for all i .*

Proof:

1. From the proof of Lemma 4.2.4 we have the following: $1 \leq i \leq m$, $\beta_i = m - 1$, so $e_i \geq m - 1$.

From Lemma 4.2.4 we have $e_i \leq 2m - i$.

Thus the deviation $\epsilon_i = |t_i - e_i| = 2m - i - e_i \leq 2m - i - (m - 1) = m + 1 - i$.

2. From the proof of Lemma 4.2.4 we have: $m + 1 \leq i \leq 2m$, $\beta_i = 0$, $e_i \leq m$.

From Lemma 4.2.4 again we have: $e_i \geq 2m - i$, so the Deviation $\epsilon_i = e_i - (2m - i) \leq m - (2m - i) = i - m$. \square .

Lemma 4.2.6. *The following inequality holds.*

$$\sum_{j=1}^l \epsilon_j + \sum_{j=2m-l+1}^{2m} \epsilon_j \geq l \quad l = 1, \dots, m. \quad (4.1)$$

Proof: From Lemma 4.2.4 we have:

$$\sum_{j=1}^l \epsilon_j = \sum_{j=1}^l (2m - j - e_j), \quad \sum_{j=2m-l+1}^{2m} \epsilon_j = \sum_{j=2m-l+1}^{2m} (e_j - (2m - j)).$$

Denoting $s = 2m + 1 - j$,

$$\sum_{j=2m-l+1}^{2m} (2m - j) = \sum_{s=1}^l (s - 1).$$

Adding the two terms from the LHS of Inequality (4.1) we get:

$$\begin{aligned}
& \sum_{j=1}^l (2m-j) - \sum_{s=1}^l (s-1) - \sum_{j=1}^l e_j + \sum_{j=2m-l+1}^{2m} e_j = \sum_{j=1}^l (2m-2j+1) - \sum_{j=1}^l e_j + \sum_{j=2m-l+1}^{2m} e_j = \\
& l(2m+1) - 2 \sum_{j=1}^l j - \sum_{j=1}^l e_j + \sum_{j=2m-l+1}^{2m} e_j = \\
& l(2m+1) - (2l(l+1)/2) - \sum_{j=1}^l e_j + \sum_{j=2m-l+1}^{2m} e_j = \\
& l(2m+1-l-1) - \sum_{j=1}^l e_j + \sum_{j=2m-l+1}^{2m} e_j = l(2m-l) - \sum_{j=1}^l e_j + \sum_{j=2m-l+1}^{2m} e_j.
\end{aligned}$$

We will divide $j = 1, 2, \dots, 2m$ into three sets: set A which contains the first l elements, set B which contains the middle $2m - 2l$ elements and set C which contains the last l elements. We can make $\sum_{j=1}^l e_j$ as small as possible by forcing all the elements from set A to be connected to elements in set B and force the elements in set A to be connected to each other.

We can similarly make $\sum_{j=2m-l+1}^{2m} e_j$ as small as possible by forcing all the elements in set C to be broken from the elements in set and B, and the elements in set C to be broken from each other.

We denote the following:

L_{AA} : are the links between the elements in set A (each edge is counted twice, representing a link for two A individuals).

$L_{AB} = L_{BA}$: are the links between the elements in set A and set B.

$L_{AC} = L_{CA}$: are links between the elements in set A and set C.

$L_{BC} = L_{CB}$: are the links between the elements in set B and set C.

L_{CC} : are the links between the elements in set C (each edge is counted twice, representing a link for two C individuals).

We have the following:

$$\sum_{j=1}^l e_j + \sum_{j=2m-l+1}^{2m} e_j = l(2m-l) - (|L_{AA}| + |L_{AB}| + |L_{AC}|) + (|L_{CA}| + |L_{CB}| + |L_{CC}|) \geq$$

$$l(2m-l) - |L_{AA}| - |L_{AB}| + |L_{CB}| + |L_{CC}| \geq l(2m-l) - l(l-1) - l(2m-2l) = l.$$

□.

Lemma 4.2.7. *For any sequence that satisfies Lemmas 4.2.5 and 4.2.6 and has $\sum_{i=1}^m \epsilon_i = m$, then there is a graph which has this sequence.*

Proof: We have the following sequence which represents the index of the nodes of our target sequence: $1, 2, \dots, m, m+1, \dots, 2m-1, 2m$ so that the target of node i is $2m-i$. We will divide this sequence into two sets: $S_1 : 1, 2, \dots, m$ and $S_2 : 2m, 2m-1, \dots, m+1$, where we have switched the order of S_2 for convenience as we see below. The required number of links for each node respectively will be as follows:

$$S_1 : 2m-1 - \epsilon_1, 2m-2 - \epsilon_2, \dots, m+1 - \epsilon_m.$$

$$S_2 : 0 + \epsilon_{2m}, 1 + \epsilon_{2m-1}, \dots, m-2 + \epsilon_{m+2}, m-1 + \epsilon_{m+1}.$$

Firstly: from the statement of the lemma, we have the following:

$$\sum_{j=1}^l \epsilon_j + \sum_{j=2m-l+1}^{2m} \epsilon_j \geq l,$$

and

$$\sum_{j=1}^m \epsilon_j + \sum_{j=m+1}^{2m} \epsilon_j = m.$$

From these we can conclude that

$$\sum_{j=l+1}^m \epsilon_j + \sum_{j=m+1}^{2m-l} \epsilon_j \leq m-l.$$

Secondly: as per Lemma 4.2.4 we have $m - 1$ links for each individual in S_1 to the other S_1 elements and zero links between the nodes in S_2 . In addition we will connect each node i from S_1 with nodes $m + 1$ up to $2m - i$ in S_2 . This is illustrated in Figure 4.2.

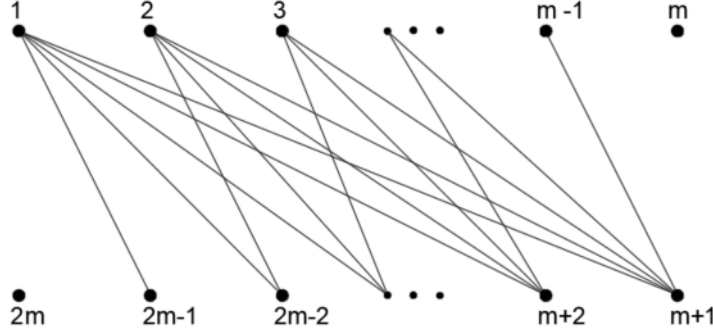


Figure 4.2: This figure shows the links between node $i \in S_1$ and node $j \in S_2$; $S_1 : 1, 2, \dots, m - 1, m$. and $S_2 : 2m, 2m - 1, \dots, m + 2, m + 1$ following the linking processes mentioned in the second step (secondly) in Lemma 4.2.7 (other links between the two sets are absent). All elements of the set S_2 are split from each other. The elements of S_1 are linked. Thus, we have illustrated the situation before the third step(thirdly).

By this procedure we will get the following required extra links for each node in S_1 and S_2 respectively:

$$S_1 : 1 - \epsilon_1, 1 - \epsilon_2, \dots, 1 - \epsilon_{m-1}, 1 - \epsilon_m.$$

$$S_2 : \epsilon_{2m}, \epsilon_{2m-1}, \dots, \epsilon_{m+2}, \epsilon_{m+1}.$$

Thirdly: now we need to consider the value of $\epsilon_i + \epsilon_{2m-i}$?

Here we will discuss the various possibilities:

1. Suppose that $\epsilon_i + \epsilon_{2m-i} = 1$, then there are two possibilities:
 - If $\epsilon_i = 0$ and $\epsilon_{2m-i} = 1$, this means each node has achieved its desired target.
 - If $\epsilon_i = 1$ and $\epsilon_{2m-i} = 0$, both nodes in each opposite pair will need one

more link, in which case we will link these two nodes, so they would then have achieved their target.

2. If $\epsilon_i + \epsilon_{2m-i} > 1$ we will denote $\epsilon_i + \epsilon_{2m-i} = k$ and $\epsilon_{2m-i} = x$. Given that

$$\sum_{j=l+1}^m \epsilon_j + \sum_{j=m+1}^{2m-l} \epsilon_j \leq m - l,$$

there are at least $k - 1$ pairs where $\epsilon_j + \epsilon_{2m-j} = 0$ with $j > i$. In this case we will link the node $2m - i$ in S_2 with x nodes from S_1 corresponding to the highest j values for which $\epsilon_i + \epsilon_{2m-i} = 0$. By this connection all these x nodes in S_1 will reach their desired target and the node $2m - i$ in S_2 will be satisfied as well. That will leave $k - x - 1$ nodes from S_1 where $\epsilon_j + \epsilon_{2m-j} = 0$ requiring one more link and node i with $k - x - 1$ too many links. To have them achieve their target we will connect these $k - x - 1$ nodes in S_1 to the opposite node vertically in S_2 . This will remove the extra one's in the deviation for those nodes, but leave their partners with one link too many. Now we consider node i in S_1 which has $(k - x - 1)$ extra links. To remove these links we will break $(k - x - 1)$ links between node i in S_1 to these $(k - x - 1)$ nodes in S_2 . Then i and $2m - i$ have achieved their target, and for all $j \geq i$ either j and $2m - i$ are on target or $\epsilon_j + \epsilon_{2m-j} = 0$. Finally: repeat step 2 for subsequent i . For the final step $\epsilon_j + \epsilon_{2m-j} = k$, there must be exactly $k - 1$ of $\epsilon_j + \epsilon_{2m-j} = 0$ with $j \leq i$, following the previous procedure, then we are left with all nodes on target, and finally that we have a sequence in the minimal set. \square . Now let us define Q_k^i as the number of ways of picking a sequence of $2i$ numbers q_1, q_2, \dots, q_{2i} which sum to k where the following conditions are satisfied:

$$\begin{cases} q_j \leq i + 1 - j & j \leq i \\ q_j \leq j - i & j > i \end{cases} \quad (4.2)$$

and

$$\sum_{j=1}^l q_j + \sum_{j=2i-l+1}^{2i} q_j \geq l + k - i \quad (4.3)$$

In particular Lemma 4.2.7 implies that

$$|J_{2m}| = Q_m^m.$$

Lemma 4.2.8. $Q_k^i = \sum_{j=\max(k-i+1,0)}^k (j+1)Q_{k-j}^{i-1}$, for $k \leq i$, $Q_0^i = 1$.

Proof:

1. Let W_k^i be the set of all sequences of length $2i$ which sum to k and satisfy Inequalities 4.2 and 4.3. Thus by the definition of Q_k^i we have $|W_k^i| = Q_k^i$.

2. Let $X = q_1, q_2, \dots, q_{2i} \in W_k^i$ where $q_1 + q_{2i} = j_0$ and $k - i + 1 \leq j_0$. We will show that this implies that $X' = (q_2, \dots, q_{2i-1}) \in W_{k-j_0}^{i-1}$.

X is of length $2i \Rightarrow X'$ is of length $2i - 2 = 2(i - 1)$.

$\sum_{j=1}^{2i} q_j = k \Rightarrow \sum_{j=2}^{2i-1} q_j = k - j_0$; $q_j \leq i + 1 - j$ for $j \leq i \Rightarrow q_j \leq (i - 1) + 1 - (j - 1)$ for $j - 1 \leq i - 1$. Similarly $q_j \leq (j - 1) + 1 - (i - 1)$ for $j - i > i - 1$ so $q_j = q'_{j-1}$ for $2 \leq j \leq 2i - 1$. Thus the entries of X' satisfy Inequality (4.2).

Now we must show that the sequence X' also satisfies Inequality (4.3).

We have

$$\sum_{j=1}^l q_j + \sum_{j=2i-l+1}^{2i} q_j \geq l + k - i \Rightarrow \sum_{j=2}^l q_j + \sum_{j=2i-l+1}^{2i-1} q_j + q_1 + q_{2i} \geq l + k - i.$$

Using $q_1 + q_{2i} = j_0$, we have

$$\sum_{j=2}^l q_j + \sum_{j=2i-l+1}^{2i-1} q_j \geq l + (k - j_0) - i.$$

Letting $j' = j - 1$ then we have

$$\sum_{j'=1}^{l-1} q'_j + \sum_{j'=2i-l}^{2i-2} q'_j \geq l + (k - j_0) - i.$$

Since $q'_j \geq 0$ for any value of j' , we have

$$\sum_{j'=1}^l q'_j \geq \sum_{j'=1}^{l-1} q'_j$$

and

$$\sum_{j'=2i-l+1}^{2i-2} q'_j \geq \sum_{j'=2i-l}^{2i-2} q'_j \Rightarrow \sum_{j'=1}^l q'_j + \sum_{j'=2i-l+1}^{2i-1} q'_j \geq l + (k - j_0) - i$$

which means that Condition (4.3) holds. Thus we have that $X' \in W_{k-j_0}^{i-1}$.

3. We now show the reverse implication to the above, namely that if $X = q_2, \dots, q_{2i-1} \in W_{k-j_0}^{i-1}$, $q_1 + q_{2i} = j_0$ and $k - i + 1 \leq j_0$ then $X' = q_1, q_2, \dots, q_{2i-1}, q_{2i} \in W_k^i$.

It is clear that X' is of length $2i$ and $\sum_{j=1}^{2i} q_j = k$. For $2 \leq j \leq i$, $q_j \in W_{k-j_0}^{i-1}$ satisfies $q_j \leq (i - 1) + 1 - (j - 1) = j + 1 - j$ and $i < j \leq 2i - 1$ satisfies $q_j \leq (j - 1) - (i - 1) = j - i$. Since we also have $q_1, q_{2i} \leq i$ then all of the entries of X' satisfy Inequality (4.2).

Now we must show that the sequence X' also satisfies Inequality (4.3).

For $l = 1$, we have $q_1 + q_{2i} = j_0 \geq k - i + 1$.

For $2 \leq l \leq i$ we have:

$$\sum_{j=1}^l q_j + \sum_{j=2i-l+1}^{2i} q_j = \sum_{j=2}^l q_j + \sum_{j=2i-l+1}^{2i-1} q_j + (q_1 + q_{2i}) \geq ((l-1) + (k-j_0) - (i-1)) + j_0 = l + k - i$$

which means that Condition (4.3) holds.

4. Let $X = q_2, q_3, \dots, q_{2i-1} \in W_{k-j_0}^{i-1}$. and $X' = q_1, q_2, \dots, q_{2i} \in W_k^i$. Since $q_1 + q_{2i} = j_0$, $q_1 \geq 0$, $q_{2i} \geq 0$ and $j_0 \leq k \leq i$ then we have $j_0 + 1$ ways to transform X to X' . Thus all sequences of $W_{k-j_0}^{i-1}$ can be transformed into a sequence of W_k^i in $j_0 + 1$ unique ways.

5. Finally we note that no two sequences in W_{k-j}^{i-1} or two sequences one each from $W_{k-j_1}^{i-1}$ and $W_{k-j_2}^{i-1}$ can generate the same sequence in W_k^i .

If x_1 and x_2 are two such distinct elements then they differ in at least one position, so then any super sequence of them must also differ similarly.

We shall solve the recurrence relation in Lemma 4.2.8 to get a solution for $|J_{2m}|$. We note that from the above we have a way of identifying all of its elements (any sequence with deviations satisfying Lemma 4.2.7).

Theorem 4.2.9. *We have the following formulae:*

$$Q_k^i = \binom{2i+k+1}{k} - 2 \binom{2i+k+1}{k-1}. \quad (4.4)$$

From Equation (4.4) and $|J_{2m}| = Q_m^m$ we will have:

$$|J_{2m}| = \binom{3m+1}{m} - 2 \binom{3m+1}{m-1}. \quad (4.5)$$

Proof

$$Q_k^i = \sum_{j=0}^k (j+1) Q_{k-j}^{i-1} \Rightarrow$$

$$Q_{k-1}^i = \sum_{j=0}^k (j+1) Q_{(k-1)-j}^{i-1} \Rightarrow$$

$$Q_k^i - Q_{k-1}^i = \sum_{j=0}^k (j+1) Q_{k-j}^{i-1} - \sum_{j=0}^{k-1} (j+1) Q_{(k-1)-j}^{i-1} \quad (4.6)$$

$$= Q_k^{i-1} + \sum_{j=1}^k (j+1)Q_{k-j}^{i-1} - \sum_{j=0}^{k-1} (j+1)Q_{(k-1)-j}^{i-1} \quad (4.7)$$

$$= Q_k^{i-1} + \sum_{j=0}^{k-1} (j+2)Q_{k-1-j}^{i-1} - \sum_{j=0}^{k-1} (j+1)Q_{(k-1)-j}^{i-1} \quad (4.8)$$

$$= Q_k^{i-1} + \sum_{j=0}^{k-1} [(j+2)Q_{(k-1)-j}^{i-1} - (j+1)Q_{(k-1)-j}^{i-1}] \quad (4.9)$$

$$= Q_k^{i-1} + \sum_{j=0}^{k-1} Q_{k-1-j}^{i-1} = \sum_{j=0}^k Q_{k-j}^{i-1} = \sum_{j=0}^k Q_j^{i-1} \quad (4.10)$$

$$\Rightarrow (Q_k^i - Q_{k-1}^i) - (Q_{k-1}^i - Q_{k-2}^i) = Q_k^{i-1} \quad (4.11)$$

$$\Rightarrow Q_k^i - 2Q_{k-1}^i + Q_{k-2}^i = Q_k^{i-1}. \quad (4.12)$$

This is a standard second order recurrence relation (for $i = 1$) which has the solution: $a_n = a_n^h + a_n^p$, where $a_n = a_n^h$ is the solution for the homogeneous case and a_n^p is a particular solution.

Firstly: we will find the solution for the homogeneous case a_n^h which is given using $m^2 - 2m + 1 = 0$. Solving this equation we get a double root where $m = 1$, thus the general solution has the form $A_ik + B_i$. Secondly: we will consider a particular solution (for simplicity of final form, we shall actually consider the summation of two parts) a_n^p . We will first check

$$Q_k^i = \binom{2i+k+1}{k}, \quad (4.13)$$

by substituting into Equation (4.12). Thus

$$\begin{aligned}
& \binom{2i+k+1}{k} - 2\binom{2i+(k-1)+1}{k-1} + \binom{2i+(k-2)+1}{k-2} \\
&= \frac{(2i+k+1)!}{k!(2i+1)!} - 2\frac{(2i+k)!}{(k-1)!(2i+1)!} + \frac{(2i+k-1)!}{(k-2)!(2i+1)!} \\
&= \frac{(2i+k-1)!}{k!(2i+1)!} [(2i+k+1)(2i+k) - 2k(2i+k) + k(k-1)] \\
&= \frac{(2i+k-1)!}{k!(2i+1)!} [2i(2i+1)] = \frac{(2i+k-1)!}{k!(2i-1)!} = \binom{2i+k-1}{k} = Q_k^{i-1},
\end{aligned}$$

as required. Similarly, checking a second term

$$Q_k^i = \binom{2i+k+1}{k-1} \quad (4.14)$$

it is easy to see that Equation (4.12) is also satisfied as required. We will consider the expression in (4.13) minus twice that in (4.14) as our particular solution. Thus our solution will have the form

$$Q_k^i = A_i k + B_i + \binom{2i+k+1}{k} - 2\binom{2i+k+1}{k-1}.$$

Now we will find the values of A_i and B_i :

$Q_1^i = 2i \Rightarrow A_i = 0$ and we have $Q_0^i = 1 \Rightarrow B_i + 1 + 0 = 1 \Rightarrow B_i = 0$. We thus have the expression from Equation (4.4) as required, and the theorem follows directly from this. \square .

Example 4.2.10. We have the size of the minimal set given by:

$$|J_{2m}| = \binom{3m+1}{m} - 2\binom{3m+1}{m-1}.$$

For $m = 1$ we will have $|J_2| = \binom{4}{1} - 2\binom{4}{0} = 4 - 2 = 2$.

For $m = 2$ we will have $|J_4| = \binom{7}{2} - 2\binom{7}{1} = 21 - 14 = 7 \dots$ etc, see Table 4.1 where

we give further results for the size of the arithmetic sequences.

It is worth recalling that, in the arithmetic sequence we have $|J_{2m+1}| = |J_{2m}|$ as we can obtain any sequence of J_{2m+1} from the corresponding sequence of J_{2m} by adding a node with target m in the middle of the $2m$ element sequence then adding one to all higher target nodes (as shown in [10]), see for example Table 4.1 where $|J_3| = |J_2|$ and $|J_5| = |J_4|$ etc.

Using Stirling's approximation formula

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n,$$

we can investigate the asymptotic behaviour of the above sequence.

$$\binom{3m+1}{m} - 2\binom{3m+1}{m-1} = \frac{(3m+1)!2}{(2m+2)!m!} \approx$$

$$\frac{2e}{\sqrt{2\pi}} \sqrt{\frac{3m+1}{m(2m+2)}} \exp\{(3m+1)\ln(3m+1) - m\ln m - (2m+2)\ln(2m+2)\}.$$

The exponent term is clearly the dominant one, and for large m this term is approximated by $2m \ln(3\sqrt{3}/2)$, and so $\ln |J_{2m}| \approx 2m \ln(3\sqrt{3}/2)$.

Thus for even n we have an expression for the size of the minimal set for a particular class of target sequence, and so we know that the largest minimal set must increase at least at rate $\exp(n \ln(3\sqrt{3}/2))$ or 2.598^n . We can see a comparison between the natural logarithm of the precise values of the minimal set formulae from Equation (4.5) and the above approximation in Table 4.2.

$ J_{2m} = \binom{3m+1}{m} - 2\binom{3m+1}{m-1}$	$\ln J_{2m} $	$\ln J_{2m} \approx 2m \ln(3\sqrt{3}/2)$	% error
$ J_2 = 2$	$\ln J_2 = 0.70$	$\ln J_2 \approx 1.90$	171.4
$ J_4 = 7$	$\ln J_4 = 1.95$	$\ln J_4 \approx 3.82$	95.9
$ J_8 = 143$	$\ln J_8 = 4.96$	$\ln J_8 \approx 7.64$	54.0
$ J_{14} = 21318$	$\ln J_{14} = 9.97$	$\ln J_{14} \approx 13.37$	34.1
$ J_{18} = 690690$	$\ln J_{18} = 13.45$	$\ln J_{18} \approx 17.19$	27.8
$ J_{100} = 5.90065579 * 10^{38}$	$\ln J_{100} = 89.27$	$\ln J_{100} \approx 95.48$	7.0

Table 4.2: In this table we show the size of the minimal set, its logarithm and the corresponding term from the approximation from above. We see here that whilst this approximation is poor as expected for small m , the relative error (third column minus second column, divided by second column) decreases with m , i.e. the larger m , the more accurate value we will get from the approximation formula we gave.

4.3 The all or nothing sequence

Here we will consider the *all or nothing* sequence $(n-1)^{m_1}, 0^{m_2}$ that we defined in Definition 3.1.2 and we considered in Section 4.1. For example the sequences in Table 4.3 are those of this type, where the number of $n-1$ and 0 target vertices are in equal number for even n , and there is one more $n-1$ target for odd n . These are the sequences which yield the maximum score from Theorem 4.1.1. Using the Matlab program described in Section 9.2 we obtain the following sequence of sizes of the minimal set: 2, 3, 7, 13, 34, 36, 221, 557.....(1) We know from the proof of

Sequence	minimal set size
{ 1, 0 }	2
{ 2, 2, 0 }	3
{ 3, 3, 0, 0 }	7
{ 4, 4, 4, 0, 0 }	13
{ 5, 5, 5, 0, 0, 0 }	34
{ 6, 6, 6, 6, 0, 0, 0 }	76
{ 7, 7, 7, 7, 0, 0, 0, 0 }	221
{ 8, 8, 8, 8, 8, 0, 0, 0, 0 }	557

Table 4.3: The sequences and minimal set sizes for the maximal score sequences for $n = 2, \dots, 9$.

Theorem 4.1.1 that in the minimal set all the nodes with target 0 are split from each other and all nodes with target $(n - 1)$ are connected to each other, and then any combination of links between nodes with target $n - 1$ and nodes with target 0 is in the minimal set. This can be represented as an $m_1 \times m_2$ matrix with a 1 if there is a link and zero otherwise. The row-sums and column-sums of this matrix then have a 1-1 correspondence with the sequences of the minimal set. For any collection of row-sums and column-sums, adding $m_1 - 1$ to the row sums and listing them in decreasing order, followed by listing the column sums in decreasing order, gives the equivalent sequence in the minimal set. The problem of finding the size of the minimal set for this class of target corresponds precisely to finding the number of distinct sets of row sums and column sums for matrices where all entries are 0 and 1. This problem was considered in [8], and is addressed in the on-line encyclopedia of integer sequences [81] as sequence number A029894 (strictly speaking the above sequence 2, 3, 7, 34, ... is sequence A327913, and A029894 is given in tabular form as it is two-dimensional). [81] gives the following formulae, which we use to generate Table 4.4 of minimal set sizes (an extended version of the table shown in A327913 $T(m_1, m_2)$, for sequence $(n - 1)^{m_1}, 0^{m_2}$ using a Matlab program. $T(m_1, m_2) = F(m_1, m_1, 0, m_2)$ where the function F satisfies the following iterative formula:

$$F(b, c, t, w) = \sum_{i=0}^b \sum_{j=\lceil (t+i)/w \rceil}^{\min(t+i, c)} F(i, j, t + i, w - 1) \quad w > 0,$$

where $F(b, c, 0, 0) = 1$, $F(b, c, t, 0) = 0$, $t > 0$. For example:

$$T(2, 2) = F(2, 2, 0, 2) = F(1, 1, 0, 1) + F(2, 1, 1, 1) + F(2, 2, 0, 1) = F(0, 0, 0, 0) + F(0, 0, 0, 0) + F(1, 1, 0, 0) + F(0, 1, 0, 0) + F(0, 0, 0, 0) + F(1, 1, 0, 0) + F(2, 2, 0, 0) = 7.$$

Note that $T(1, 1)=2$, $T(2, 1)=3$, $T(2, 2)=7$, $T(3, 2)=13$, $T(3, 3)=34$, $T(4, 3)=76$, $T(4, 4)= 221$, $T(5, 4)=557$ are the first entries in the leading diagonal, and its

m_1/m_2	0	1	2	3	4	5	6	7	8
0	1	1	1	1	1	1	1	1	1
1	1	2	3	4	5	6	7		9
2	1	3	7	13	22	34	50	70	95
3	1	4	13	34	76	152	280	482	787
4	1	5	22	76	221	557	1264	2630	5108
5	1	6	34	152	557	1736	4766	11812	26930
6	1	7	50	280	1264	4766	15584	45356	119999
7	1	8	70	482	2630	11812	45356	153228	465673
8	1	9	95	787	5108	26930	119999	465673	1611189
9	1	10	125	1230	9362	57270	293089	1294838	5060227

Table 4.4: The values of the minimal set for the all or nothing sequences with m_1 “all” vertices and m_2 “nothing” vertices.

neighbour from Table 4.4, which are of course the same values of the minimal sets size of the sequences as given in Table 4.3.

$\ln J_{2m} $ of the AoN	$\ln J_{2m} $ of the Ar	$\ln J_{2m} \approx 2m \ln(3\sqrt{3}/2)$
$\ln J_8 = 5.4$	$\ln J_8 \approx 4.9$	$\ln J_8 \approx 7.64$
$\ln J_{10} = 7.5$	$\ln J_{10} \approx 6.6$	$\ln J_{10} \approx 9.55$
$\ln J_{12} = 9.7$	$\ln J_{12} \approx 8.3$	$\ln J_{12} \approx 11.46$
$\ln J_{14} = 11.9$	$\ln J_{14} \approx 10$	$\ln J_{14} \approx 13.67$
$\ln J_{16} = 14.3$	$\ln J_{16} \approx 11.7$	$\ln J_{16} \approx 15.28$

Table 4.5: In this table, we show the logarithm of the size of the minimal sets for the all or nothing and arithmetic sequences.

$\ln J_{2m} $ of the AoN sequence	$\ln J_{2m} $ of the Arithmetic sequence
$\ln J_2 \approx 0.7$	$\ln J_2 \approx 0.7$
$\ln J_4 \approx 1.9$	$\ln J_4 \approx 1.9$
$\ln J_6 \approx 3.5$	$\ln J_6 \approx 3.4$
$\ln J_8 \approx 5.4$	$\ln J_8 \approx 4.9$
$\ln J_{10} \approx 7.5$	$\ln J_{10} \approx 6.6$
$\ln J_{12} \approx 9.7$	$\ln J_{12} \approx 8.3$
$\ln J_{14} \approx 11.9$	$\ln J_{14} \approx 10$
$\ln J_{16} \approx 14.3$	$\ln J_{16} \approx 11.7$

Table 4.6: The values of the minimal set for the all or nothing sequences with m_1 “all” vertices and m_2 “nothing” vertices.

From Table 4.6 we can see that for small numbers of individuals the (largest) minimal set for the all or nothing sequence is larger than for the arithmetic sequence, and that the gap is increasing. We know from the approximation at the end of Section 4.2 that the logarithm of the arithmetic sequence has a leading term that is of order n , though a more careful examination of the approximation shows that it is not a purely linear function, as it involves a term in the logarithm of n too. In Figure 4.3 we can see that the logarithm of the arithmetic sequence gradually increases in slope, approaching its limiting term, and we can see that it appears that the logarithm of the (dominant) all or nothing sequence does the same. Thus we can conjecture that it also has a leading term that is of order n . We have fitted a line to each of these data sets for illustration for the lowest values of n (distinct lines for odd and even n in each case) for illustration. In general the all or nothing sequence appears to increase approximately linearly at a faster rate than the arithmetic sequence, indeed the slope in Figure 4.3 is already steeper than the limiting case for the arithmetic sequence, and so it appears that the all or nothing sequence has a larger minimal set than the arithmetic one in general.

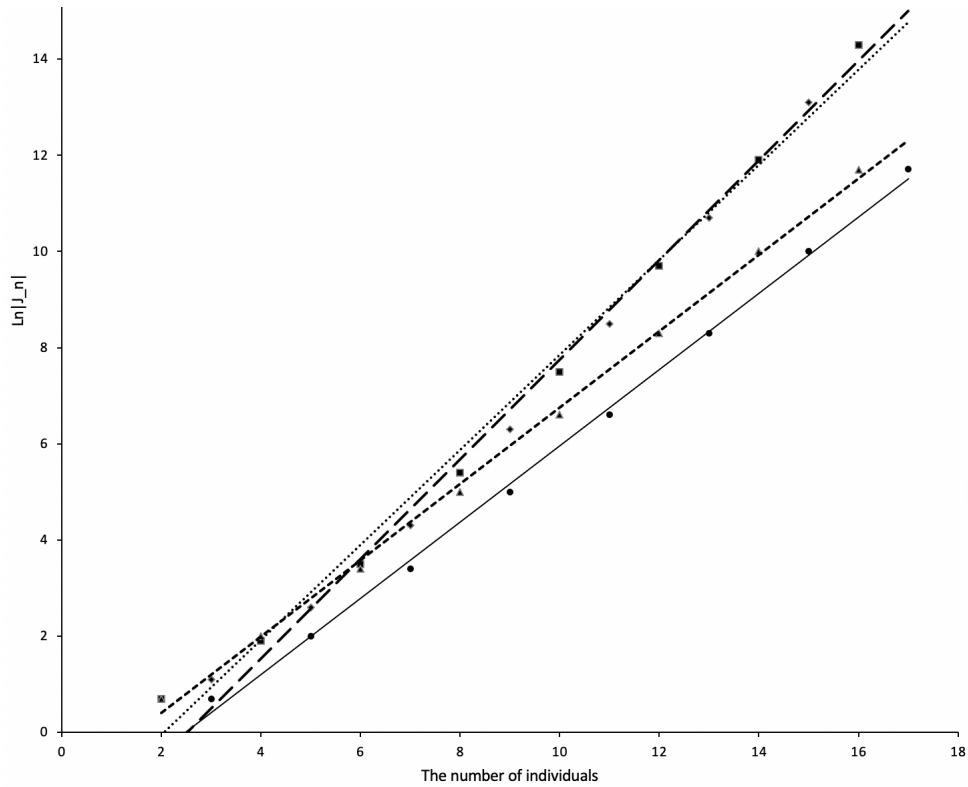


Figure 4.3: In this figure we consider two types of sequences; the all or nothing and the arithmetic sequences, and fitted two lines for each; one for the even values and the other for the odd values of the sequence.

Line 1: represented by \blacksquare refers to the even terms of the all or nothing sequences, $Y = 0.9875x - 2.025$.

Line 2: represented by \diamond — — refers to the odd terms of the all or nothing sequences, $Y = 1.0357n - 2.6071$.

Line 3: represented by \blacktriangle - - - refers to the even terms of the arithmetic sequences, $Y = 0.7935n - 1.1786$.

Line 4: represented by \bullet ——— refers to the odd terms of the arithmetic sequences, $Y = 0.7935n - 0.3851$.

Chapter 5

Game theoretical Models and the Minimal Set

In this chapter we consider strategic games on a dynamically interacting model introduced By Broom and Cannings in [8] and mentioned in Section 3.3.1. The games considered in this thesis are “memoryless” due to the characteristic of the Markov chain, so that if one knows the process’s current state (the graph), then knowledge of prior states is not required to provide the best prediction of its future, and we are able to considerably reduce the number of states that must be considered, and follow a similar methodology to, for example [71]. A strategy is hence a choice of transitions that would be made from each state, and we compare these against alternative strategies of the same type, typically a change of transition from one graph only. It would be possible to consider an alternative, based upon the assumption that a state is a combination of both the current graph and the time (or even also some further history of the process) so that alternative strategies would involve a change of transition only at that time point. This would lead to a potentially different, but in practice much more complex, problem. In some of these games, the number of graphs exceeds 40, so making the equilibrium analy-

sis for each local transition a costly task. A Markov steady state is a stochastic matrix's eigenvector. That is, if I take a probability vector and multiply it by the probability transition matrix and obtain the same exact probability vector, it was a steady state. Since the Nash equilibrium is defined based on a single interaction, we will define Nash equilibria here as the steady states of Markov chain.

5.1 The Game

In the games that we will discuss, only a single individual has the right to make a change at each time step, then transition from a current state to another will happen. $\mathbf{U}_{\mathbf{X}}$ denotes the set of strategies for all situations. If following these strategies leads to a unique stationary distribution over the states \mathbf{x} (this will happen if there is always some non-zero probability of any transition in the minimal set, but does not happen in general, as shown in [10]; in particular it happens for the game we are interested in below) which we can denote by $\pi(\mathbf{X})$, then the payoff to individual i can be written as follows:

$$R_i(\mathbf{U}_{\mathbf{X}}) = - \sum_{\mathbf{X}} \epsilon_i(\mathbf{X}) \pi(\mathbf{X}), \quad (5.1)$$

where $\epsilon_i(\mathbf{X})$ is the deviation of v_i in state \mathbf{X} . Individuals can try to improve their payoffs by changing their strategy. We only allow *local changes*, where individual i changes the i th row of $\mathbf{U}_{\mathbf{X}}$ for a single \mathbf{X} only (as opposed to *global changes*, where v_i is able to change the i th row of $\mathbf{U}_{\mathbf{X}}$ for any number of states simultaneously). Considering all possible changes by any single individual i , a strategy set is a Nash equilibrium if under all allowable changes by $i : \mathbf{U}_{\mathbf{X}} \rightarrow \mathbf{U}_{\mathbf{X}}^i$

$$R_i(\mathbf{U}) \geq R_i(\mathbf{U}^i) \quad i = 1, \dots, n.$$

5.2 Example: showing that leaving the minimal set can be optimal

Suppose that we have n individuals, one of them v_1 with target $n-2$ and the others with target zero, i.e we have the target sequence: $n-2, 0, 0, \dots, 0$. We assume that all individuals with target zero will always break links when they have the opportunity (this is clearly beneficial as then they will be on target almost all of the time and have a maximum of a single link otherwise). Similarly individual v_1 with target $n-2$ would always form links when it is below its target of $n-2$ (this individual will almost always be below target as whilst it must be selected to add any links, all of its potential partners will break such links when they are selected) and may or may not add the final link if it has the opportunity to do so whilst connected to $n-2$ individuals. We will show that in this game it is better for v_1 to form a link when already on target causing the graph to leave the minimal set. Consider the states S_0, S_1, \dots, S_{n-1} , all of them involving no links between individuals with target zero, where in state S_k v_1 has links to precisely k other individuals. Given all have target 0 and will break their one possible link (with v_1) when they have the chance, it does not matter which of the k individuals v_1 is connected to, and so we represent this by a single state. Assume that if v_1 is selected in S_{n-2} , then it will choose to form the final link with probability p , and that if v_1 is selected in S_{n-1} then it will choose to break a link with probability q (and again it does not matter which is selected in this particular game). The transition matrix P for this process is given in Matrix (5.2) and can be used to

find the stationary distribution, $\pi_0, \pi_1, \pi_2, \dots, \pi_{n-1}$ where $0 \leq p \leq 1, 0 \leq q \leq 1$:

$$P = \begin{pmatrix} (n-1)/n & 1/n & 0 & \cdot & \cdot & \cdot & 0 \\ 1/n & (n-2)/n & 1/n & \cdot & \cdot & \cdot & 0 \\ 0 & 2/n & (n-3)/n & 1/n & \cdot & \cdot & 0 \\ 0 & 0 & 3/n & (n-4)/n & 1/n & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & (n-3)/n & 2/n & 1/n & 0 \\ 0 & \cdot & \cdot & \cdot & (n-2)/n & (2-p)/n & p/n \\ 0 & \cdot & \cdot & \cdot & 0 & (n-1+q)/n & (1-q)/n \end{pmatrix}. \quad (5.2)$$

The payoff $R(p, q)$ to v_1 for this game with n individuals is given by the formula:

$$R(p, q) = - \sum_{i=0}^{n-1} \pi_i |i - (n-2)| = - \sum_{i=0}^{n-2} \pi_i (n-2-i).$$

1. To find the stationary distribution π we have $\pi = \pi * P$, which yields the following equations:

$$\pi_0 = \frac{n-1}{n} \pi_0 + \frac{1}{n} \pi_1 \Rightarrow \pi_0 = \pi_1, \pi_1 = \frac{1}{n} \pi_0 + \frac{n-2}{n} \pi_1 + \frac{2}{n} \pi_2 \Rightarrow \pi_2 = \frac{1}{2!} \pi_0.$$

For $m \leq n-2$ we have:

$$\pi_{m-1} = \frac{1}{n} \pi_{m-2} + \frac{n-m}{n} \pi_{m-1} + \frac{n}{m} \pi_m \Rightarrow m \pi_{m-1} = \pi_{m-2} + m \pi_m. \quad (5.3)$$

Let us assume that

$$\pi_k = \frac{1}{k!} \pi_0. \quad (5.4)$$

Substituting Equation (5.4) into Equation (5.3) we obtain

$$\frac{m}{(m-1)!}\pi_0 = \frac{1}{(m-2)!}\pi_0 + m\pi_m \implies \pi_m = \frac{1}{m!}\pi_0.$$

Thus, since $\pi_1 = \pi_0$, we have that $\pi_m = \pi_0/m!$ for $m < n-1$. We will now find π_{n-1} which depends upon p, q and π_0 as follows:

$$\begin{aligned} \pi_{n-2} &= \frac{1}{n}\pi_{n-3} + \frac{2-p}{n}\pi_{n-2} + \frac{n-1+q}{n}\pi_{n-1} \implies \\ \frac{n-2+p}{n}\pi_{n-2} &= \frac{1}{n(n-3)!}\pi_0 + \frac{n-1+q}{n}\pi_{n-1} \implies \\ \pi_{n-1} &= \frac{p}{n-1+q} \frac{\pi_0}{(n-2)!}. \end{aligned}$$

To find the value of π_0 we have $\sum_{m=0}^{n-1} \pi_m = 1 \implies$

$$\pi_0 = \frac{1}{\sum_{m=0}^{n-2} \frac{1}{m!} + \frac{p}{n-1+q}(n-2)!}.$$

2. We now proceed to find the payoff of individual v_1 by applying the following formula for the payoff following Equation (5.1):

$$R(p, q) = - \sum_{m=0}^{n-1} \pi_m |m - (n-2)| = - \sum_{m=0}^{n-3} \pi_m (n-2-m).$$

Thus

$$R(p, q) = - \left[\sum_{m=0}^{n-2} (n-m-2) \frac{1}{m!} + \frac{p}{(n-1+q)(n-2)!} \right] \pi_0 \implies$$

$$R(p, q) = - \frac{(n-2) \sum_{m=0}^{n-3} 1/m! + \sum_{m=1}^{n-3} 1/(m-1)! + p/(n-1+q)(n-2)!}{\sum_{m=0}^{n-2} 1/m! + p/((n-1+q)(n-2)!)}.$$

(5.5)

Denoting $S_x = \sum_{m=0}^x 1/m!$ and $r = p/(n - 1 + q)$ in (5.5) we obtain:

$$R(p, q) = -\frac{(n-2)S_{n-3} + S_{n-4} + r}{S_{n-2} + r}. \quad (5.6)$$

Finding the derivative of Equation (5.6) with respect to r we obtain

$$\frac{dR(p, q)}{dr} = \frac{-S_{n-2} + (n-2)S_{n-3} + S_{n-4}}{(S_{n-2} + r)^2} \quad (5.7)$$

Now we will discuss the sign of the numerator of Equation (5.7), which we shall denote by O_n . We have $O_n = -S_{n-2} + (n-2)S_{n-3} - S_{n-4}$. Recalling that $S_x = \sum_{m=0}^x 1/m!$, we have $S_0 = 1, S_1 = 2, S_2 = 2.5, S_3 = 2.6, \dots$, which leads to the following terms for O_n ; $O_2 = -1, O_3 = -1, O_4 = 0.5, O_5 = 2.9, \dots$. In general O_n is positive for all larger values of n . We can see this since S_n is bounded above by $e < 3$, we have that (for $n \geq 2$), $2 \leq S_n < 3$ and so $O_n = -S_{n-2} + (n-2)S_{n-3} - S_{n-4} > -3 + 2(n-2) - 3 = 2n - 10 > 0$ whenever $n \geq 6$ (equivalent to $n-4 \geq 2$). Thus for sufficiently large n , O_n is positive. We have that $dR/dr < 0$ when $n < 4$ and $dR/dr > 0$ when $n \geq 4$. From the definition of r it is clear that r is maximised ($r = 1/(n-1)$) when $p = 1, q = 0$ and minimised ($r = 0$) when $p = 0$ and q can take any value. Thus $r = 1/(n-1)$, and so $p = 1$, is the optimal strategy whenever $n \geq 4$ and we thus have that v_1 should take the extra link when it has the chance, which leads to the minimal set being left, whenever $n \geq 4$.

Example 5.2.1. Consider the following $n = 5$ individuals: S_0, S_1, S_2, S_3, S_4 one of them with the target $n - 2 = 3$ and the others have the target 0.

Each individual will try to reach his desired target.

- Considering the transition matrix P find the stationary distribution $p_0, p_1, p_2, p_3, p_4; p_i = p[S_i]$.
- Show that $R(1, 0) > R(0, q)$ by applying the formula:

$$R(p, q) = -[(n - 2)S_{n-3} + S_{n-4} + r] * [1/S_{n-2} + r] \text{ for } n = 5, q = 1.$$

The solution:

1. To find the stationary distribution we have $\pi = P * \pi$: where P is given as follows:

$$P = \begin{pmatrix} 4/5 & 1/5 & 0 & 0 & 0 \\ 1/5 & 3/5 & 1/5 & 0 & 0 \\ 0 & 2/5 & 2/5 & 1/5 & 0 \\ 0 & 0 & 3/5 & 1/5 & 1/5 \\ 0 & 0 & 0 & 4/5 & 1/5 \end{pmatrix}$$

By applying $\pi = P * \pi$: we find:

$$4/5 * \pi_0 + 1/5 * \pi_1 = \pi_0,$$

$$1/5\pi_0 + 3/5\pi_1 + 2/5\pi_2 = \pi_1,$$

$$1/5\pi_1 + 2/5\pi_2 + 3/5\pi_3 = \pi_2,$$

$$1/5\pi_2 + 1/5\pi_3 + 4/5\pi_4 = \pi_3,$$

$$1/5\pi_3 + 1/5\pi_4 = \pi_4,$$

$$\implies \pi_0 = \pi_1,$$

$$\pi_2 = 1/2\pi_0,$$

$$\pi_3 = 1/6\pi_0,$$

$$\pi_4 = 1/24\pi_0.$$

We have $\sum_{m=0}^{n-1} \pi_m = 1$, thus we have $\pi_0 = 24/65, \pi_1 = 24/65, \pi_2 = 12/65, \pi_3 = 4/65, \pi_4 = 1/65$.

2. To find the payoff we will apply the formula: $R(p, q) = -[(n - 2)S_{n-3} + S_{n-4} + r] * [1/S_{n-2} + r]$.

The optimal strategy occur when r is maximised or minimised i.e at $p = 0, q = 1$ or $p = 1, q = 0$.

$$\implies R(1, 0) = -(24/65 + 24/65 * 2 + 12/65 * 1 + 4/65 * 0 + 1/65 * 1) =$$

$$-133/65 = -2.04615.$$

$$R(0, 1) = -(3 * 3/8 + 2 * 3/8 + 3/16 * 1 + 3/48 * 0 + 0) = -2.0625.$$

which means $R(1, 0) > R(0, 1)$ so that $p = 1$, is the optimal strategy.

5.3 The game theoretical model with target sequence score equal to 1

5.3.1 Score 1 games

As we have seen, not all degree sequences can be realised as undirected graphs and the Havel-Hakimi algorithm [31, 37] characterises those that can. When the sequence is graphic, all individuals will eventually be on target (i.e. the score will be zero) and this situation is of minimal interest, as then the process will evolve no further. We are more interested in non-graphical sequences where not all the individuals can be on target simultaneously, and more precisely, we want to find the sequences, and associated graphs, with the lowest possible deviation from the original required target, which we classified into a set called the minimal set, denoted by $J(\min)$ for the sequences and $K(\min)$ for the graphs.

To reach the minimal set, the competitors should use rational decisions that prioritise them in getting the desired number of links to achieve their targets. We will consider a particular case of the type of target sequence considered in [12], those as close as possible to the regular graphical sequences, with a score equal to 1. In this model, within the minimal set only one individual is off-target and so only this one individual has an interest in changing its number of links. In particular in any state within $K(\min)$ only one individual will attempt to improve its payoff by changing its strategy from any state, and so we know which individual will make the next change irrespective of the order of selected individuals, as all

others will "pass". We will revisit an example game that was considered in [12].

5.3.2 The sequence 1, 1, 1

Let us consider the target sequence 1, 1, 1, which has $n = 3$ individuals. It is easy to identify the transitions between the states in this case, and we can see that the transition graph will have 6 vertices which are shown in Figure 4.1. When the individual in deficit is chosen, if it has too many links (a Breaker) it will break one of its links, and if it has too few (a Joiner) it will form a link. In particular the individual to be broken with/connected to will not be chosen at random, as in the transition probabilities displayed in Section 3.2, but will be selected strategically. This particular sequence is of interest as (it was previously considered that it was) the simplest "non-trivial" sequence for the game-theoretical model. When considering all possible sequences with up to four individuals, we would have two types of situation. Firstly there are those with straightforward strategies within the minimal set (however, the game from [2] where optimal play can lead to the minimal set being left) in which only one individual ever has a strategic choice to make; for example for 2, 0, 0 the 0 individuals can have at most a single link, and their optimal strategy is always to break it. Secondly when this does not occur, we will have a game that has a considerable complexity in analysing the optimal behaviour, for example the sequence 3, 2, 1, 0 with $J(\min)=7$ and $K(\min)=8$ which was studied in [10]. It was shown there that there are multiple pure Nash equilibria, and the analysis of mixed equilibria would have been very complex. Similarly the sequence 2, 2, 0, 0 which has $J(\min)=4$ but $K(\min)=16$, which would lead to complex computation of the resulting stationary distribution over the 16 states. The sequence 1, 1, 1 is the most straightforward nontrivial sequence found (and likely possible) with $J(\min)=2$ and $K(\min)=6$. Note that the sequence 1, 1, 1, 1, 1

is vastly more complex, as analysis from [12] has shown. In this work we show that even 1, 1, 1 is more complex than previously thought.

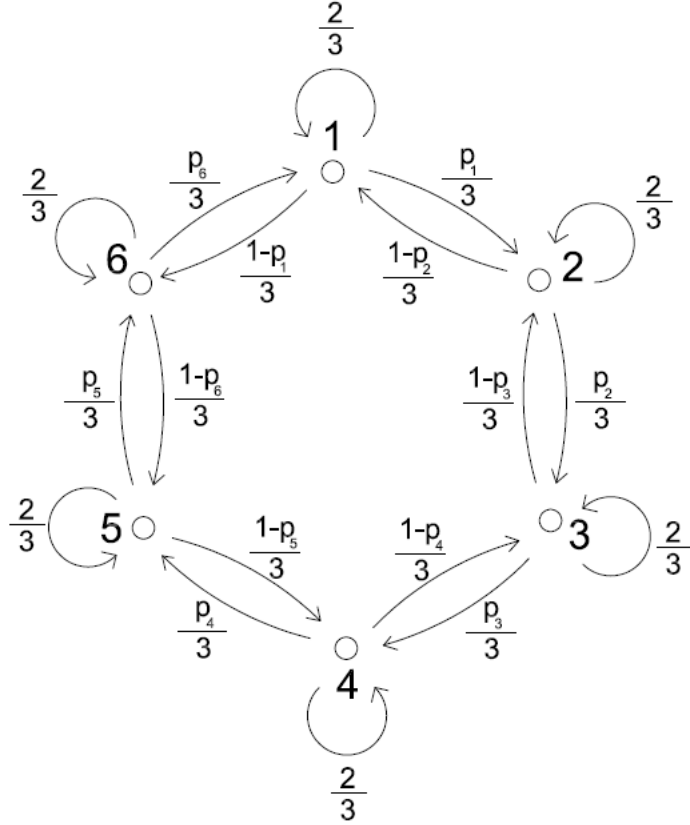


Figure 5.1: Transition graph of the sequence 1, 1, 1 showing the minimal set in every possible move at each of the six vertices. Vertex 1 represents the sequence 2, 1, 1, i.e. individual v_1 is the one in deficit and moves clockwise to sequence 101 with probability $\frac{p_1}{3}$ and anticlockwise to 110 with probability $\frac{1-p_1}{3}$, it will stay in the same state with probability $\frac{2}{3}$. Similarly, vertex 2 represents sequence 1, 0, 1, vertex 3 represents sequence 1, 1, 2, vertex 4 represents sequence 0, 1, 1, vertex 5 represents sequence 1, 2, 1 and vertex 6 represents sequence 1, 1, 0.

5.3.3 States and transitions

Sequence 1, 1, 1 has a transition graph with 6 states. If a vertex not in deficit is selected to update its links, it will make no change. If the vertex in deficit is selected, it will either have two links, one of which it will break, or no links, in which

case it will form one of those available. Thus there are two possible transitions. A collection of probabilities p_1, \dots, p_6 corresponds to a set of strategies to be played by all individuals. So all combinations of probabilities correspond to all possible strategy sets for this game. We define $\pi[i]; i = 1, \dots, 6$ as the stationary probability distribution for our Markov chain, which will (under the usual assumptions) be the long term probability distribution of the occupancy of the states. The deficit vertex individual in-state i will move clockwise with probability p_i or anticlockwise with probability $(1 - p_i)$, the value of p_i being a strategic choice. The numbered vertices represent the sequence/graph described in the caption (for example, vertex 1 represents v_1 connected to both other individuals, with the link between the others broken), and the corresponding possible transitions are shown as an arrow moving clockwise and anticlockwise. Transitions occur as follows:

The individual that is potentially to move is chosen at random. This leads to one of the two following cases when in any state i :

1. When the selected individual is neutral, it will decide not to make any transition and to stay at the same position; this occurs with a probability of $\frac{1}{3}$ for each such individual, giving a total probability of $\frac{2}{3}$ of the state remaining unchanged see Figure 5.1.
2. When the selected individual is in deficit, it will have the choice to move clockwise and (link/break) with the individual who is in deficit at the corresponding vertex with probability p_i or move anticlockwise and (link/break) with the other individual with probability $(1 - p_i)$. We thus have:

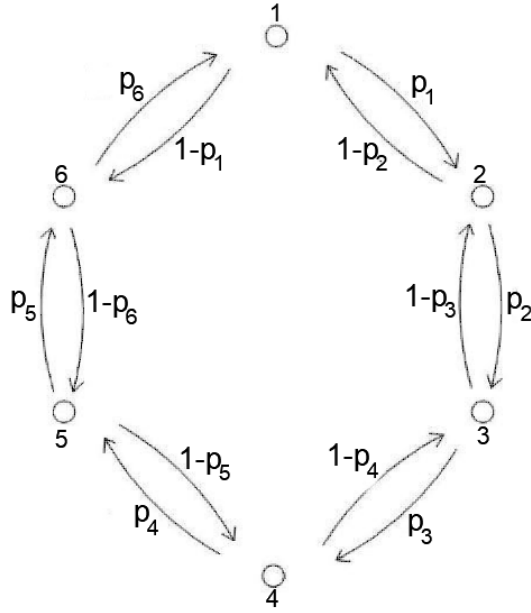


Figure 5.2: Transition graph of the sequence 1, 1, 1 showing the minimal set in every possible move at each of the six vertices. Vertex i will move clockwise to vertex $i + 1$ with probability p_i . Vertex i will move anticlockwise to vertex $i - 1$ with probability $1 - p_i$.

$$\pi[1] = \pi[6] * p_6/3 + \pi[1] * 2/3 + \pi[2] * (1 - p_2)/3; \quad (5.8)$$

$$\pi[2] = \pi[1] * p_1/3 + \pi[2] * 2/3 + \pi[3] * (1 - p_3)/3; \quad (5.9)$$

$$\pi[3] = \pi[2] * p_2/3 + \pi[3] * 2/3 + \pi[4] * (1 - p_4)/3; \quad (5.10)$$

$$\pi[4] = \pi[3] * p_3/3 + \pi[4] * 2/3 + \pi[5] * (1 - p_5)/3; \quad (5.11)$$

$$\pi[5] = \pi[4] * p_4/3 + \pi[5] * 2/3 + \pi[6] * (1 - p_6)/3; \quad (5.12)$$

$$\pi[6] = \pi[5] * p_5/3 + \pi[6] * 2/3 + \pi[1] * (1 - p_1)/3. \quad (5.13)$$

This leads to the transition matrix P :

$$P = \begin{bmatrix} \frac{2}{3} & \frac{p_1}{3} & 0 & 0 & 0 & \frac{1-p_1}{3} \\ \frac{1-p_2}{3} & \frac{2}{3} & \frac{p_2}{3} & 0 & 0 & 0 \\ 0 & \frac{1-p_3}{3} & \frac{2}{3} & \frac{p_3}{3} & 0 & 0 \\ 0 & 0 & \frac{1-p_4}{3} & \frac{2}{3} & \frac{p_4}{3} & 0 \\ 0 & 0 & 0 & \frac{1-p_5}{3} & \frac{2}{3} & \frac{p_5}{3} \\ \frac{p_6}{3} & 0 & 0 & 0 & \frac{1-p_6}{3} & \frac{2}{3} \end{bmatrix}. \quad (5.14)$$

The stationary distribution π can be found by solving $\pi * P = \pi$, in conjunction with the knowledge that

$$\pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 + \pi_6 = 1. \quad (5.15)$$

We found the stationary distribution values in terms of $p_1, p_2, p_3, p_4, p_5, p_6$ along with the payoff of the individuals, see Appendix 9.3. However, only one individual is in deficit, and only that individual would choose to move. Thus the state will be unchanged until eventually that individual is selected. Since staying probabilities are the same in any state, we could effectively assume that the individual is chosen immediately see Figure 5.3; here, the figure is written using the above equivalence for simplicity.

The payoff for individual v_1 is minus the probability that it is in deficit, i.e. is the probability that the Markov chain is in states 1 or 4. Thus the payoff for v_1 is: $-(\pi[1] + \pi[4])$, and so similarly the payoffs to individuals v_2 and v_3 are $-(\pi[2] + \pi[5])$ and $-(\pi[3] + \pi[6])$ respectively. In [12] three Nash equilibria were found. Since the system will move clockwise if $p_i = 1$ for all i and anticlockwise if $p_i = 0 \forall i$, in each of these situations each state occurs with a frequency of $1/6$, so the payoff to each individual is $-1/3$. These two sets of strategies are pure Nash equilibria.

They also found a mixed Nash equilibrium when $p_i = 1/2 \forall i$, again with payoffs of $-1/3$ to all individuals. In [12] they showed that when $p_i = 1/2$, the solution is not stable since minor deviations from $1/2$ would subsequently favour strategies that deviate further by running simulations.

5.3.4 Finding Nash equilibria NE

We used the payoffs as described above to investigate the Nash equilibria of the game. A Nash equilibrium occurs if and only if no individual has an incentive to change its strategy. For the choice at vertex 1:

$$p_1 \text{ is a NE iff } \begin{cases} \frac{d(\pi_1 + \pi_4)}{dp_1} \geq 0 & p_1 = 0 \\ \frac{d(\pi_1 + \pi_4)}{dp_1} = 0 & 0 < p_1 < 1 \\ \frac{d(\pi_1 + \pi_4)}{dp_1} \leq 0 & p_1 = 1 \end{cases} \quad (5.16)$$

Recalling that the expressions above in the derivative are minus the payoff to v_1 , these conditions ensure that no change in p_1 enables v_1 to gain a higher payoff. Equivalent conditions hold for the other five cases. We found the following Nash equilibria:

1. $p_i = 1$ for every i , every movement (transition) of an individual will be clockwise. For $p_i = 1$ we have

$$\frac{d(\pi_1 + \pi_4)}{dp_1} = \frac{d(\pi_1 + \pi_4)}{dp_4} = \frac{d(\pi_2 + \pi_5)}{dp_5} = \frac{d(\pi_2 + \pi_5)}{dp_2} = \frac{d(\pi_3 + \pi_6)}{dp_3} = \frac{d(\pi_3 + \pi_6)}{dp_6} = -2,$$

which means that $p_i = 1$ is a Nash equilibrium as in conditions 5.16.

2. $p_i = 0$ for every i , every movement (transition) of an individual will be

anticlockwise. For $p_i = 0$ we have

$$\frac{d(\pi_1 + \pi_4)}{dp_1} = \frac{d(\pi_1 + \pi_4)}{dp_4} = \frac{d(\pi_2 + \pi_5)}{dp_5} = \frac{d(\pi_2 + \pi_5)}{dp_2} = \frac{d(\pi_3 + \pi_6)}{dp_3} = \frac{d(\pi_3 + \pi_6)}{dp_6} = 2,$$

which means that $p_i = 0$ is a Nash equilibrium from conditions 5.16.

In cases (1) and (2) each state occurs with frequency $1/6$ so the cost to each individual is $1/3$. If any individual at any state tried to act differently, then the system will return that individual back to the previous state, and the system will oscillate giving a cost of $1/2$ to the individual who switches play. Thus these two sets of choices are strict Nash equilibria.

3. $p_i = 1/2$ for all i , for $p_i = 1/2$ we have

$$\frac{d(\pi_1 + \pi_4)}{dp_1} = \frac{d(\pi_1 + \pi_4)}{dp_4} = \frac{d(\pi_2 + \pi_5)}{dp_5} = \frac{d(\pi_2 + \pi_5)}{dp_2} = \frac{d(\pi_3 + \pi_6)}{dp_3} = \frac{d(\pi_3 + \pi_6)}{dp_6} = 0,$$

and we again have a Nash equilibrium from conditions 5.16.

These results reproduce the Nash equilibria from the simpler analyses from [12]. For the solutions (4) -(9) below, we have a set of cyclic solutions which satisfies the following:

$$\frac{d(\pi_1 + \pi_4)}{dp_1} = \frac{d(\pi_1 + \pi_4)}{dp_4} = \frac{d(\pi_2 + \pi_5)}{dp_5} = \frac{d(\pi_2 + \pi_5)}{dp_2} = \frac{d(\pi_3 + \pi_6)}{dp_3} = \frac{d(\pi_3 + \pi_6)}{dp_6} = 0$$

for every solution respectively, and we thus have Nash equilibria as well by conditions 5.16.

4. $p_1 = 1, p_2 = 1, p_3 = 0, p_4 = 0, 0 < p_5 < 1, 0 < p_6 < 1$, see Figure 5.3.

5. $0 < p_1 < 1, p_2 = 1, p_3 = 1, p_4 = 0, p_5 = 0, 0 < p_6 < 1$.

6. $0 < p_1 < 1, 0 < p_2 < 1, p_3 = 1, p_4 = 1, p_5 = 0, p_6 = 0$.

7. $p_1 = 0, 0 < p_2 < 1, 0 < p_3 < 1, p_4 = 1, p_5 = 1, p_6 = 0$.

8. $p_1 = 0, p_2 = 0, 0 < p_3 < 1, 0 < p_4 < 1, p_5 = 1, p_6 = 1.$

9. $p_1 = 1, p_2 = 0, p_3 = 0, 0 < p_4 < 1, 0 < p_5 < 1, p_6 = 1.$

Figure 5.3 illustrates solution (4) showing the movement of each individual. Solutions (5) - (9) are simple rotations of Figure 5.3.

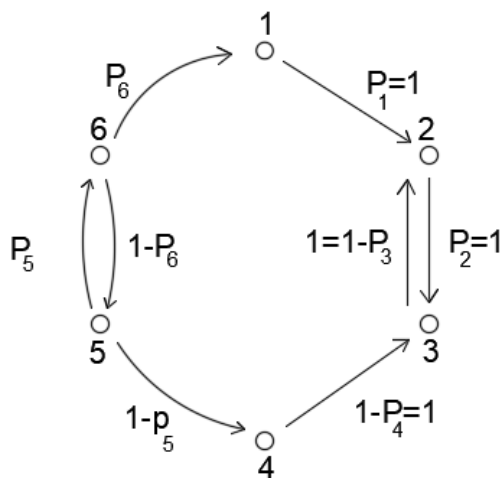


Figure 5.3: Transition graph of the sequence 1, 1, 1 showing the first solution of the cyclic set, solution (4): Straight arrows \searrow : when $p_i = 1$, individual (i) will move clockwise. Straight arrows \swarrow : when $p_i = 0$, individual (i) will move anti-clockwise. Curved arrows: for $0 < p_i < 1$ the movement of individual (i) to the next state could happen in either direction, clockwise with probability p_i or anti-clockwise with probability $1 - p_i$.

Let us consider solution (4), where $p_1 = 1, p_2 = 1, p_3 = 0, p_4 = 0, 0 < p_5 < 1, 0 < p_6 < 1$. Here we alternate between states 2 and 3, with a payoff of 0 to v_1 and $-1/2$ to both v_2 and v_3 . We note that this case is a NE due to the following reasoning: If we are in states 1 and 4 then v_1 would not change strategy, as currently they already receive the maximum payoff of 0 (though actually any change would leave v_1 with payoff 0 too).

If we are in state 5 (state 6) then it makes no difference what strategy the individual in deficit v_2 (v_3) will use as they will end up in states 2 and 3 after enough moves. In state 2, if v_2 picked $p_2 = 0$ instead of $p_2 = 1$ the system will instead alternate

between states 1 and 2, leaving the payoff to v_2 unchanged. If some value $0 < p_2 < 1$ was chosen, there would be alternation between state 2 and one of states 1 and 3, with the same result. Similar reasoning holds for the choice of v_3 in state 3. We note from the above that a change in strategy at any vertex does not improve the payoff of the corresponding individual (otherwise it would not be in Nash equilibrium), but it does not make it worse either, and this explains why all the derivatives from condition 5.16 are equal to 0, as opposed to those in solution (1), for example we note that reflecting this figure through its axes, e.g. through the axis from vertex 1 to vertex 4, does not provide any new solutions, as this always leads to one of the other six (for the reflection above, this leads from solution (4) to solution (7)).

We have tried to find other sets of Nash equilibria, or to show that they could not exist, but an exhaustive search has proved too complicated. We have been able to demonstrate the absence of Nash equilibria in certain subsets of the parameters, but there are many combinations outstanding.

Chapter 6

New dynamic network population models

In this chapter, we will introduce two new dynamic models building on the model considered earlier in Section 1.11 in the thesis. We discussed the strategies with the original model, below we will consider three models A, B, and C as follows:

1. Model A: This is the original model from Mark Broom and Chris Cannings in [8] which we have been considering. In this model, an individual is chosen at random, and it will try to improve its deviation from the required target by forming or breaking links with a randomly selected neighbour. In this case, the other recipient individual must accept that decision as it does not have the right to reject it. The individuals will be selected with equal probability. This process follows a Markov chain. For the case $x_{ij} = 0$, $x_{ij}^* = 1$ this led

to

$$P(X_{t+1} = x^* | X_t = x) = \begin{cases} \frac{1}{n} \frac{1}{n-1-u_i} + \frac{1}{n} \frac{1}{n-1-u_j} & u_i < t_i, u_j < t_j \\ \frac{1}{n} \frac{1}{n-1-u_i} & u_i < t_i, u_j \geq t_j \\ \frac{1}{n} \frac{1}{n-1-u_j} & u_i \geq t_i, u_j < t_j \\ 0 & u_i \geq t_i, u_j \geq t_j \end{cases}$$

and for the case $x_{ij} = 1, x_{ij}^* = 0$

$$P(X_{t+1} = x^* | X_t = x) = \begin{cases} \frac{1}{n} \frac{1}{u_i} + \frac{1}{n} \frac{1}{u_j} & u_i > t_i, u_j > t_j \\ \frac{1}{n} \frac{1}{u_i} & u_i > t_i, u_j \leq t_j \\ \frac{1}{n} \frac{1}{u_j} & u_i \leq t_i, u_j > t_j \\ 0 & u_i \leq t_i, u_j \leq t_j \end{cases}$$

With notation as described in Section 3.1.

2. Model B: here if individuals would like to form a link, it would need the consent of both parties, but if the chosen individual wants to break a link, it will not require the agreement of the other individual.

We will discuss the node choices/situations based on the type as follows:

- Breaker nodes will keep improving their target by imposing breaking links with others until they are on target. Once an individual is on target, it cannot be a Breaker anymore; having this procedure, we will have all the Breakers turn into either Joiner or Neutral nodes; this process will continue as long as we have Breaker nodes.
- We note that Joiner and Neutral nodes might lose more links if a Breaker

is chosen and decides to break their link. Thus in model B it is possible for nodes to finish further below target than they start.

If initiated from any graph, this model will then eventually reach a state with no Breaker nodes. The transitions for this model are as follows:

For the case $x_{ij} = 0, x_{ij}^* = 1$

$$P(X_{t+1} = x^* | X_t = x) = \begin{cases} \frac{1}{n} \frac{1}{n-1-u_i} + \frac{1}{n} \frac{1}{n-1-u_j} & u_i < t_i, u_j < t_j \\ 0 & u_i < t_i, u_j \geq t_j \\ 0 & u_i \geq t_i, u_j < t_j \\ 0 & u_i \geq t_i, u_j \geq t_j \end{cases}$$

and for the case $x_{ij} = 1, x_{ij}^* = 0$

$$P(X_{t+1} = x^* | X_t = x) = \begin{cases} \frac{1}{n} \frac{1}{u_i} + \frac{1}{n} \frac{1}{u_j} & u_i > t_i, u_j > t_j \\ \frac{1}{n} \frac{1}{u_i} & u_i > t_i, u_j \leq t_j \\ 0 & u_i \leq t_i, u_j > t_j \\ 0 & u_i \leq t_i, u_j \leq t_j \end{cases}$$

3. Model C: In this model, any change/addition or removal of a link, between two vertices must have the consent of both of them; thus, only pairs of Joiners and pairs of Breakers can change by forming/ breaking links among each other; as any other possible change will be declined by the other party. For example: if a Neutral node is selected, it will choose not to take any further actions.

(a) This model could be initiated from a general graph containing Breaker,

Joiner and Neutral nodes. We guarantee to split a connection between two Breaker nodes by having both of them agree to do so, but we need the consent of pairs which involve a Joiner or Neutral (which will never be given) to break a link between them. This procedure will be repeated until all Breakers are satisfied and on target; however, we might still have Breakers in a terminal graph, for example: if we have one remaining Breaker node in the graph sequence, no further movement is allowed for this individual.

- (b) This model could be initiated from an empty graph with zero links; in this case, there will be no Breakers during the game or in a terminal graph. The game will begin with Joiner and Neutral nodes; the only possible movements here are for the Joiners to link to each other while the Neutral nodes will reject any links.

The transitions for the model are as follows:

For the case $x_{ij} = 0, x_{ij}^* = 1$

$$P(X_{t+1} = x^* | X_t = x) = \begin{cases} \frac{1}{n} \frac{1}{n-1-u_i} + \frac{1}{n} \frac{1}{n-1-u_j} & u_i < t_i, u_j < t_j \\ 0 & u_i < t_i, u_j \geq t_j \\ 0 & u_i \geq t_i, u_j < t_j \\ 0 & u_i \geq t_i, u_j \geq t_j \end{cases}$$

and for the case $x_{ij} = 1, x_{ij}^* = 0$

$$P(X_{t+1} = x^* | X_t = x) = \begin{cases} \frac{1}{n} \frac{1}{u_i} + \frac{1}{n} \frac{1}{u_j} & u_i > t_i, u_j > t_j \\ 0 & u_i > t_i, u_j \leq t_j \\ 0 & u_i \leq t_i, u_j > t_j \\ 0 & u_i \leq t_i, u_j \leq t_j \end{cases}$$

In both models, we will reach a stage where no further movements are allowed; here, we know that we have reached an absorbing state, which we shall call a terminal graph. In the terminal graph, we might/ might not have Breaker nodes, depending on the chosen model and the initial situation. The terminal sets of a target sequence t is the collection of the terminal graphs $K(\text{term})$, along with associated sequences $J(\text{term})$ of the target sequence t . A sequence is in the terminal set if and only if it is a terminal graph sequence.

6.1 General sequence results

6.1.1 Vertex classifications

The fundamental property of a terminal graph is that there are no further allowable moves; in this section, we will mention some other properties of the terminal sets:

1. In models B and C, if a sequence is graphical, any graph that achieves the sequence will be an element of the terminal set.

Proof: we assume that a sequence $d = (d_1, \dots, d_n)$ is a graphical sequence with $0 \leq d_1 \leq \dots \leq d_n \leq n - 1$. By following the H-H algorithm we will add a particular set of edges, this sequence of links has a non zero probability of being picked, thus at least one graph (G) will achieve the degree sequence

where $d(G) = d$; in this case, the vertices are all on target (N). Thus no further movements are allowed, which classifies the graph as an element of the terminal set.

2. The all or nothing sequences have precisely one element in the terminal set of sequences $J(\text{term})$.

Proof: assume we have an all or nothing sequences S consisting of the sets: S_1 the set of all zero nodes in S , S_2 the set of all the $n - 1$ nodes in S , and we further define the set S_3 as the set of links connecting nodes from the set S_1 to the set S_2 .

- (a) Starting from an empty graph for both models, the terminal set will contain one element as follows: the set S_3 has zero links in the initial graph, which will not change throughout the game, and all the nodes in set S_1 are on target; thus, they will reject any suggested links from Joiners. However, all pairs of Joiner nodes in set S_2 will connect.
- (b) Starting from a general graph in model B, the terminal set will have one element as follows: nodes of set S_1 will break all the links and reject any proposed links. Thus, the set S_1 will have no links between its members and S_3 will have zero links in the terminal graph. All the nodes in set S_2 will end up connected.
- (c) In model C, starting from a general graph, there will be again one element. The nodes in S_1 will agree to split from each other but cannot break any connection with the nodes of the set S_2 as they will always reject any split. Therefore, we will end up with the original elements in the set S_3 , which will not change throughout the interaction and zero links in the set S_1 , while the nodes in the set S_2 will again all connect.

3. In the terminal set of model B, a sequence t that realises a terminal graph G , will have its vertices classified as one of the mentioned sets below. However, these classifications will vary depending on the initial graph that has been chosen:

V_J : The set of all Joiners in the sequence t .

V_{S_1} : the set of all nodes, which are sometimes Joiners and sometimes Neutrals.

V_N : the set of all nodes which are always Neutrals.

Now we will discuss the different cases of the initial graphs:

- Case 1: If the game started with zero connections between the individuals (empty graph), the set of nodes at this stage will be classified as N or J . However, when the individuals start interacting, the Joiners will try to form links with others to minimise their deviation; Neutrals will always reject to connect as they are already on target. In contrast, the Joiners will accept a link, which will lead to all the Joiners being linked in a terminal graph.

Here it is possible to have vertices which remain Joiners irrespective of the linking sequence, elements of V_J . Others will change between Neutral and Joiner nodes based on the sequence linking chosen V_{S_1} . The remaining nodes will be always Neutral nodes under any connections V_N .

In case 1 nodes of a terminal graph will be elements of V_{S_1} , V_N or V_J .

- Case 2: suppose that we start with a general graph that might have different types of vertices; we will discuss the following:
 - (a) Some individuals who start the game with a number of connections equivalent to its target (Neutral nodes); they will always stay

- on target and so are members of V_N after any linking formation. Others will turn into Joiner nodes if a Breaker can split with it in a different linking formation, with no further opportunity to form other links; those who act as Neutrals sometimes and as Joiners at other times depending on the sequence linking are members of V_{S_1} .
- (b) Individuals who require more links to be on target after any sequence linking are always Joiners nodes; those are elements of V_J . Others who start the game as Joiners and can achieve their target following specific nodes connecting will be classified as some times Joiner and some times Neutral nodes; those will be elements of V_{S_1} . The individuals who start the game as Joiners and always achieve the required target under any link formation; and so will be in V_N .
- (c) Individuals who start the game with extra links to the required target t will take any chance to break connections with others until they become Neutral nodes without having any restrictions in this model. Once they become Neutrals, some nodes will always act as Neutral nodes by the end of the game regardless of any linking formation, so those are elements of V_N . Other nodes will have different statuses, either remaining Neutrals under certain sequence linking or becoming Joiner nodes if a Breaker decides to break with them under a different sequence linking, thus can be elements of V_{S_1} . Some nodes will become Joiners by the end of this process under any linking formation, those can be elements of V_J .

In case 2 nodes of a terminal graph will be in V_{S_1} , V_N , or V_J .

- Case 3: starting from a full graph, where all the nodes are connected to each other:

- (a) Nodes with target $n - 1$ will be Neutral nodes before the interactions. Those who stay Neutral nodes following any possible linking will be in V_N , and the others whose type change and become Joiner nodes following other linking can be in V_{S_1} .
- (b) Nodes with $t_i < n - 1$ will keep breaking extra links until they become Neutral nodes. The nodes who always stay Neutral, whatever the change in the link formations, will be in V_N . Nodes that fluctuate between Neutral and Joiner's nodes depending on the sequence linking are in V_{S_1} . And nodes that become always Joiners will be in V_J .

Generally, in model B, we don't have Breaker nodes in a terminal graph since any Breaker will keep forcing a split until reaching the required target (or becoming a Joiner).

4. Consider the terminal set of Model C, a sequence $t = t_1, t_2, \dots, t_n$ that achieves a terminal graph G . Its vertices will be classified as one of the below-mentioned sets, vertex classifications will vary depending on the starting graph that has been chosen:

V_J : The set of all Joiners in the sequence S.

V_{S_1} : the set of all nodes, which are sometimes Neutrals and sometimes Joiners.

V_N : the set of all nodes which are always Neutrals.

V_B : the set of all nodes which are always Breakers.

V_{S_2} : the set of all nodes, which are sometimes Neutrals and sometimes Breakers.

Now we will discuss the cases of different initial graphs:

- Case 1: If we started with zero connections between the individuals (empty graph), this would be a special case of case 2; the difference is that in this case, we will not have nodes $\in V_{S_2}$ and V_B in the terminal graphs as we will show. When the individuals start interacting, the Joiner nodes will try to form links with others; Neutral ones will reject to create or break links. In contrast, the Joiners will accept it, leading to all the Joiners being linked. This turns some nodes into always Neutral nodes ($\in V_N$), others will stay as always Joiner nodes ($\in V_J$), while other individuals will have affected types by depending upon the linking sequence, and those will be V_{S_1} . We note that in case 1, the terminal graphs will not have nodes in V_B or V_{S_2} , all the nodes will be in V_N, V_J and V_{S_1} .
- Case 2: We can potentially start from a general graph; for a given starting graph, we will have the following possible cases:
 - Individuals on target (N) initially will reject to form/break any link at any interaction. Those will be classified in V_N .
 - Suppose the individuals are short of target initially. Some of them will have type which will vary depending on the sequence linking; sometimes acting as Joiner nodes and sometimes as Neutral nodes, i.e. in V_{S_1} . Other nodes will always stay Joiners and hence in V_J , and some of them will always turn to Neutral nodes in V_N in the terminal graph.
 - Individuals who start the game with extra links to the required target (B) will take any chance to break relations with other Breakers. This process will continue until they become Neutral nodes if there are enough Breakers to split with; if they can achieve their target

for any linking sequence, they will be classified as always Neutral nodes and so in V_N . If the number of Breaker nodes is less than the additional links of the individuals, they will stay short of the target no matter what linking formation they choose and will be classified as always Breaker nodes (V_B). Other Breaker nodes will sometimes become Neutrals and sometimes not, so they classified in being in V_{S_2} . In this case, nodes of the terminal set are in V_N, V_B, V_J, V_{S_2} or V_{S_1} . We note that no vertex can be sometimes a Joiner, and sometimes as a Breaker (and sometimes Neutral) *given its initial number or linking* since wherever the target is achieved, all changes will stop.

- Case 3: starting from a full graph, where all the nodes are connected; with target $0 \leq t_i \leq n - 1$ for $i = 1, 2, \dots, n$, we will have two types of vertices in the initial graph which are Neutral or Breaker:
 - (a) Nodes with target $n - 1$ will always stay Neutral nodes and will resist any change, so they classified in V_N .
 - (b) Breaker nodes with $t_i < n - 1$. These can be those who can always break links with other Breakers until they become Neutral nodes, i.e. V_N .
 - (c) Breaker nodes with $t_i < n - 1$ who do not have enough Breakers to split will always be in V_B .
 - (d) There can be Breaker nodes with $t_i < n - 1$ who are on target some times and short of target in other times depending on the linking sequence, i.e. classified as in V_{S_2} . We note that in case 3, in the terminal graphs, we do not have nodes in V_J and V_{S_1} .

Lemma 6.1.1. *In model B, and Model C starting from an empty graph, a sequence/graph is a member of the terminal set if and only if it contains Neutral and Joiner vertices only, and any pair of Joiners is joined.*

Proof:

1. Assume we are in the terminal set, and graphs have Breaker, Joiner, and Neutral nodes. Based on the model B definition, Breakers will keep imposing breaking links until they are Neutral nodes considering the sequence is of length n with $t_i \leq n - 1$. Thus, having a Breaker node in the graph means it is not terminal as that will allow some individuals to have further movements, which contradicts our assumption. Thus the terminal set of model B cannot include Breakers. However, it consists of Neutral and Joiner nodes.

Assume we are in the terminal set, and we have the Joiner nodes u and v that are not connected. In this case, we can always create an additional link between them which contradicts our assumption of being in the terminal set. That implies all the Joiners must be connected in the terminal set.

2. Assume all nodes are either Joiners or Neutrals in a set, and all the Joiners are connected; in this case, we cannot have more movements as this will have the Joiners trying to link to the Neutral nodes. Since adding a link will require the consent of both of them, the Neutral nodes will decline any proposed connection. This means that we are in the terminal set, and no further movement could exist.

6.2 Properties of the terminal set

In this section we will mention some properties of the terminal set that help us to further classify the vertices into V_N, V_J, V_{S_1} .

Lemma 6.2.1. *For the set of terminal graphs associated with any target sequence $t = t_1, t_2, \dots, t_n$ of length n in model B , and terminal graphs initiated with an empty graph, we can divide the individuals (vertices) into the following classes:
 A_n : the number of nodes which are always Joiners.*

B_n : the number of nodes with target $n - 1$.

Any vertex i with target t_i such that $t_i \leq A_n + B_n$ is an always Neutral node.

Proof Let us assume the contrary, that the node s_i which has a target $t \leq A_n + B_n$ is not Neutral $\implies \exists s_j$ which is always a Joiner, or it has a target $n - 1$ and not linked to the node $s_i \implies s_j$ is a Joiner, and s_i is a Joiner. They are not connected, and that's a contradiction; by the definition of the terminal set, all the Joiner nodes must be joined together $\implies s_i$ is always a Neutral node.

Example 6.2.2. for the sequence $t = 8, 8, 7, 5, 4, 4, 3, 0, 0$

we have $n = 9, A_n = 2, B_n = 2$ for any node i with target $t_i \leq A_n + B_n$ thus any node with target $t_i \leq 3$ is always a Neutral node. Thus, the node with target 3 is a Neutral node.

We will refer to n_0 as the number of the nodes with target $t \geq 1$

Lemma 6.2.3. *In model B , with terminal graphs initiated with an empty graph, vertices with target $n_0 - 1$ are $\in V_{S_1}$ except when all vertices of the sequence have target $\geq n_0 - 1$.*

Proof since we have n_0 nodes have a target greater than or equal to 1, the maximum number of links any vertex can have is $n_0 - 1$, so nodes with target $n_0 - 1$ will be Neutral if we connect it first to the remaining $n_0 - 1$ nodes in the sequence.

A $n_0 - 1$ node will be a Joiner if at least one node v_i is not connected to it.

Note 6.2.4. Any vertex that has a target $t_i > n_0 - 1$ will be $\in V_j$.

That is obvious since the greatest number of links any node can have is $n_0 - 1$.

Lemma 6.2.5. *In model B, and model C initiated with an empty graph, considering a sequence $t = t_1, t_2, \dots, t_n$ that reach a terminal graph G , if a node $t_i \in V_N$ then $\forall j \geq i; t_j \in V_N$.*

Proof In the sequence $t = t_1, t_2, \dots, t_n$, assume that $t_i \in V_N$ and $t_j \in V_J; t_i > t_j$. That implies that node t_i is linked to n other nodes we divided them into three sets B, C, and D, and the Joiner node t_j linked to $t_l < t_j$ other nodes we divided them into the sets A and B, which described as follows:

A: the set of vertices connected to j and flexible to swap their links with i has the size $|A|$.

B: the set of vertices connected to j and i simultaneously; has the size $|B|$.

C: the set of vertices connected to i and flexible to swap their links with j ; and has the size $|C|$.

D: the set of vertices connected to i and not flexible to swap; has the size $|D|$. see Figure 6.1. Thus we will have:

$$|B| + |C| + |D| = t_i \tag{6.1}$$

$$|A| + |B| < t_j \tag{6.2}$$

$$|B| + |C| = t_j \tag{6.3}$$

From equation 6.1 and 6.3 we have:

$$|D| = t_i - |C| - |B| = t_i - t_j \tag{6.4}$$

Since we assumed that j is a Joiner, it must be connected to all other Joiners in

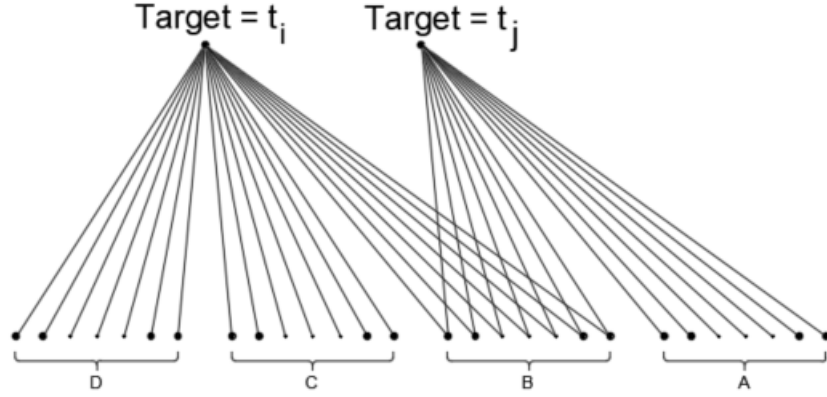


Figure 6.1: A: the set of vertices connected to j and flexible to swap, has the size $|A|$.
 B: the set of vertices connected to j and i simultaneously and has the size $|B|$.
 C: the set of vertices connected to i , can exchange links and has the size $|C|$.
 D: the set of vertices connected to i , cant swap links and has the size $|D|$.

a terminal graph. We have $j \geq i$ and $t_i \geq t_j$, swapping the links of the sets C and A between t_i and t_j will lead to t_i losing $|C| = t_j - B$ links by breaking it and linking it to node t_j , and gaining $|A| < t_j - B$ links.

In the new terminal graph, G' see Figure 6.2, which has been formed from swapping the mentioned links, we will have node j as an always Neutral node since it has been linked to t_j links coming from i . However, node i will be a Joiner as it will be missing at least one link, as follows: $t_i - t_j + (t_j - 1) = t_i - 1$ which is a contradiction since we assumed from the beginning that i is always a Neutral node. Thus $j \in V_N$ and $t_i = t_j$.

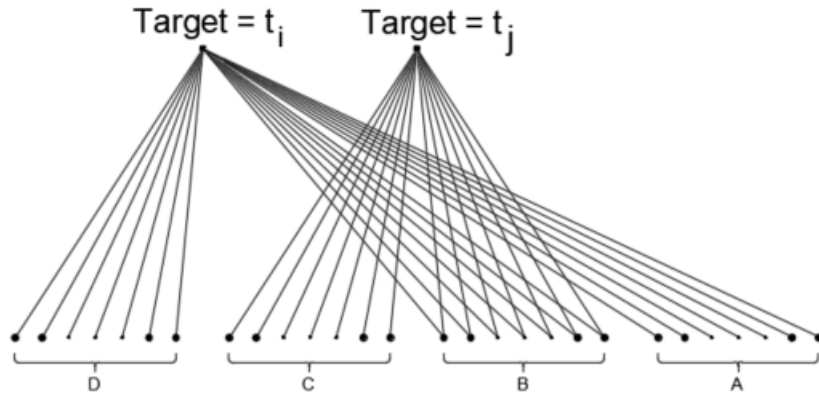


Figure 6.2: A, B, D: The sets of links currently connected to the node t_i .
 C, B: The sets of links that are currently connected to the node t_j .
 A swap of links happened between sets C and A.

6.3 The Reverse Havel- Hakimi Theorem

This section presents a novel theorem termed as the Reverse Havel- Hakimi Theorem. This theorem is derived from the Havel- Hakimi Theorem, which is used to determine whether a sequence is graphical or not, thereby determining whether a simple graph exists which achieves the target sequence.

Theorem 6.3.1. *The Reverse Havel Hakimi Theorem (RHH): let $t = (t_1, t_2, \dots, t_n)$ be a finite list of non negative and non increasing integers. list S is graphic if and only if the sequence $t' = (t_1 - 1 \geq t_2 - 1 \geq \dots \geq t_n - 1 \geq t_{n+1} \geq \dots \geq t_{n-1})$ is graphic.*

Proof

- First we prove sufficiency: If RHH leaves this sequence all Neutral then it is clear that the sequence is graphic.
- Second we prove necessity, using the induction method.

1. A graphic sequence of length (1) or (2) (the sequences 0 or 1, 1) clearly leaves the nodes as Neutrals.
2. Assume RHH leaves the nodes of a graphic sequence of length n Neutral.
3. Let the sequence $t = (t_1, t_2, \dots, t_{n+1})$ of length $n + 1$ be a graphic of length $n+1$.
 - If the node t_{n+1} is linked to all t_i nodes such that $1 \leq i \leq n + 1$ then we construct $G - t_{n+1}$ which would have degree $(t_1 - 1, t_2 - 1, \dots, t_n - 1)$ which by our assumption we know RHH leaves it all Neutral, thus by adding t_{n+1} also will leave t Neutral.
 - Assume that t_{n+1} is not linked to all t_i such that $1 \leq i \leq n + 1$ in G , let t_j be the largest vertex such that $j > t_{n+1}$ and linked to t_{n+1} and let t_x be the vertex that t_{n+1} is not linked to such that $1 \leq x \leq t_{n+1}$ and since t is non increasing sequence we have $t_x \geq t_j$. If $t_x = t_j$, we swap the two vertices and we are done. if $t_x > t_j$, then there is a vertex w ; w linked to t_x but not connected to t_j then we can removed the edges $t_x w$ and $t_j w$. This produce a new graph G' with a degree sequence and has the node t_{n+1} linked to t_x . We repeat until t_{n+1} is linked to all t_i nodes; $1 \leq i \leq t_{n+1}$

Definition 6.3.2. The Reverse Havel-Hakimi algorithm

We will introduce this algorithm which tests if a sequence is graphical or not by applying the following steps:

1. Sort the sequence in decreasing order.
2. Remove the last (smallest) node. Suppose that the last node has a target k . Subtract one from the first k node.
3. Repeat this process until we reach a state where we cannot create further links between the individuals. By then, either we end up with a sequence of

zeros (graphic) or a number that represent the deviation between the current sequence and the target sequence (not graphic).

Theorem 6.3.3. *Let $t = t_1, t_2, t_3, \dots, t_i, \dots, t_n$ be a non-increasing and non-negative sequence which achieves a terminal graph G with score k ; to determine whether the node t_i is in V_N, V_J , or V_S we will define the sequence t^* as follows:*

1. *Delete the top $i - 1$ nodes from the sequence t .*
2. *Apply the R-H-H algorithm for the remaining nodes.*

If the sequence t^ is graphical, there will be no more Joiners and the node t_i will always be a Neutral node; however, if the sequence t^* is not graphical, there will be a new set of Joiners t_A^* . If the target of the top remaining node t_0 in t_A^* is such that $t_0 \geq t_i$, then t_i is a Neutral node; otherwise, t_i can be a Joiner.*

Proof Let t_i be the node in the sequence S that we want to investigate.

T_1 : the set of the top $i - 1$ nodes excluding node t_i .

T_2 : the remaining nodes excluding node t_i .

t_A^* : The new targets of the Joiners of T_2 after applying the R-H-H.

Case 1: $\exists t_j \in T_1$ such that $t_j \leq t_{i-1} \implies \forall t_k \in T_2$ we will have $t_k \leq t_{i-1} \implies |t_A^*| \leq t_{i-1}$ and $\forall t_k \in T_2, t_k \not\geq t_i \implies t_i$ can be a Joiner.

Case 2: $\exists t_j \in T_1$ such that $t_j \geq t_i$ then let $\max t_A^* = t_0 \implies t_j \geq t_0 + |t_A^*| - 1, \forall t_j \in T_1$. As otherwise t_j can not be an element of the top nodes set T_1 . Assume that $t_0 \geq t_i$ then $t_j \geq t_i + |t_A^*| - 1$ as there are only $t_i - 1 + |t_A^*|$ Joiners left after R-H-H $\implies t_j$ can have only $t_i - 2 + |t_A^*|$ links not including t_i .

$t_j - (t_i - 2) - |t_A^*| \geq t_i + |t_A^*| - 1 - (t_i - 2) - |t_A^*| = 1$ thus t_j can make one more link to t_i . $\implies t_i$ gets $t_i - 1$ links from T_1 . As $t_0 \geq t_i \implies t_0$ can not be Neutral after connecting all to $T_2 \implies t_0$ must be connected to t_i . Thus, t_i must be Neutral.

Corollary 6.3.4. *If $t_i = 1$ and $t_0 \geq 1$ then t_i must be Neutral.*

Proof assume $t_i = 1$ and $t_0 \geq 1$ but t_i is not Neutral node $\implies t_i$ is a Joiner $\implies G - t_i$ nodes are all Neutrals \implies applying RHH will leave $t_0 = 0$. However, we know that $t_0 \geq 1$ so we have a contradiction which implies that t_i must be a Neutral node.

Example 6.3.5. In this example we will apply the Reverse Havel Hakimi (R.H.H) and the Havel Hakimi algorithm (H.H) for the sequence 9, 7, 7, 7, 5, 4, 4, 4, 4 as follows:

R.H.H	H.H
977754444	977754444
86665444 0	66643333, 1
755554400	5533222,10
654444000	422211,100
544330000	1111,10000
433300000	11,1000000
322000000	100000000
210000000	100000000
100000000	100000000

Table 6.1: R.H.H: the reverse Havel Hakimi algorithm.

H.H: the Havel Hakimi algorithm.

The numbers after the comma represents the missing links of the individual after forming every possible link with others.

We note that in this example both algorithm give the same result.

Definition 6.3.6. The S_{RHH} is the summation of the excess targets in t .

Theorem 6.3.7. *The RHH-score equals the score of the sequence; i.e. the RHH algorithm produces a minimal graph.*

Proof. Let S_n be the set of non negative integer sequences of length n . For $u, v \in S_n$ we have $z(u, v) = \sum_i |u_i - v_i|$. Consider the RHH algorithm above. For $u \in S_n$ this will produce a unique value, defining the function $sRHH(u)$. For some v , $z(u, v) = 1$ then the vectors u and v are identical in every entry except

one, where their elements differ by 1. Suppose that $v_k = u_k + 1$. Applying the algorithm either

- $k = n$ then u, v are identical except at index k where they differ by 1.
- $k < n, v_k < v_{k-1}$ and subsequently the vectors are identical except at index k where they differ by 1.
- $k < n, v_k = v_{k-1}$ and subsequently the vectors are identical except at precisely one index between k and $n - 1$.

Thus the difference of 1 is preserved. This process is repeated until we can't create any further links between the nodes and the top term/ terms has a 1 leftover as an extra link in one sequence and not the other (sometimes there will be an identical number of zeros with one in each sequence) after which the sequences will be identical, but with one more 1 leftover in one than the other. Thus $|s_{RHH(u)} - s_{RHH(v)}| = 1$. For all $u, v \in S_n$, $z(u, v) \geq |s_{RHH(u)} - s_{RHH(v)}|$ since if we have $u, v \in S_n$ with $z(u, v) = n - k$ there is a path $v = z_n, z_{n-1}, \dots, z_k = U$ where $z(z_i, z_{i+1}) = 1$. $z(u, v) = \sum_{n-k}^n z(z_i, z_{i+1}) = \sum_{n-k}^n |s_{RHH(z_i)} - s_{RHH(z_{i+1})}| \geq |\sum_{n-k}^n (s_{RHH(z_i)} - s_{RHH(z_{i+1})})|$, with equality iff all $(s_{RHH(z_i)} - s_{RHH(z_{i+1})}) = 1$ or all $(s_{RHH(z_i)} - s_{RHH(z_{i+1})}) = -1$. Any graphic sequence w has $s_{RHH(w)} = 0$, so $z(w, u) \geq s_{RHH(u)}$ and since $d(RHH(u), u) = s_{RHH(u)}$ we have $RHH(u)$ as a minimal graphic sequence for u .

6.4 The terminal set over the arithmetic sequence

In Sections 6.1.1 and 6.2 we investigated the properties of the terminal set for a general sequence, and we classified the different types of vertices. Now we will investigate the properties of the terminal set of the arithmetic sequences and the vertex classifications for model B and model C started from an empty graph.

Lemma 6.4.1. *In the terminal set of an arithmetic sequence of model B and Model C, started with an empty graph, the nodes with targets: $t_i = 0, 1$ and 2 are always Neutrals.*

Proof For a terminal graph G of order n , assume that a node with target $t_i = 1$ is not always Neutral $\implies t_i = 1$ is not linked to node $t_i = n - 1$. We know that node $t_i = n - 1$ is an always Joiner \implies Graph G is not terminal as a new link between $t_i = 1$ and $t_i = n - 1$ can be made; this is a contradiction. Thus $t_i = 1$ must be a Neutral node.

Now assume $t_i = 2$ is not always Neutral, here we will discuss two cases:

- If the node with target $t_i = 2$ has no links, similar to the above, this will lead to a contradiction that graph G is not terminal as a new link between $t_i = 2$ and $t_i = n - 1$ can be made $\implies t_i = 2$ must have at least one link.
- if $t_i = 2$ has one link and is not linked to $n - 1$ then we have the similar argument. Assume that $t_i = 2$ and $t_i = n - 1$ are linked \implies node $t_i = 2$ is not linked to $t_i = n - 2$ but node $t_i = n - 2$ cannot be Neutral if it is not linked to every node else except $t_i = 0$ and so would still form a link with $t_i = 2 \implies G$ is not terminal, thus $t_i = 2$ must be an always Neutral node.

Lemma 6.4.2. *In the terminal graphs of an arithmetic sequence of model B and Model C starting from an empty graph, a node $i \in V_{S_1}$ in the following cases:*

1. *In an arithmetic sequence of order, $n = 2m + 1$, a node i such that $3 \leq i \leq 2m - 1$ is $\in V_{S_1}$.*
2. *In any arithmetic sequence of order $n = 2m$, a node i such that $3 \leq i \leq 2m - 2$ is $\in V_{S_1}$.*

Proof

1. by applying the Havel Hakimi algorithm on an arithmetic sequence of order $n = 2m + 1$ where the n node has target $2m$, we will find that the top nodes of targets $2m, 2m - 1, \dots, m + 1$ are $\in V_J$ and the lower nodes of targets $m, m - 1, \dots, 2, 1, 0$ are $\in V_N \implies$ the sequence has reached a graph of the terminal set where all the upper nodes are Joiners and all the lower nodes are Neutrals.

Second: let i be a vertex in the lower part such that $3 \leq i \leq m$. We want to show that there is a terminal graph G where i is a Joiner.

Algorithm: let $f(v) = 2m - v - 1$ for $m + 1 \leq v \leq 2m - 2$ and let

$$g(v) = \begin{cases} f(v) & f(v) < i \\ f(v) + 1 & f(v) \geq i \end{cases}$$

For $2m - i - 1 \leq v \leq 2m - 2$ we link all vertices $\{u \geq g(v); u \neq i, u \neq v\}$, we will have v is Neutral here are exactly v nodes connected to it. And $g(v) \in V_N$ since $\exists v^*; 2m - 2 \geq v^* > v$ are all linked to it. For $m + 1 \leq v \leq 2m - i - 1$, we link all vertices $\{u \geq g(v); u \neq v\}$, and we will have $v \in V_N$ again for the same reason as above. In this case vertex $g(v)$ would need an extra link so we will link it with the node of target $2m$ to make it Neutral, thus $g(v)$ is linked with all the nodes $v^*; v \leq v^* \leq 2m$.

We add a link between $2m, 2m - 1, i$ the graph then will be terminal as every node can be Neutral except for the nodes $\{2m, 2m - 1, i\} \implies$ we have a graph in the terminal set where $m \leq i \leq 3$ is a Joiner.

Third: the node with target $2m-1$ will be Neutral if we connect it to every node except itself and 0 , and will be a Joiner in any other case.

From the above we conclude that for the arithmetic sequence of order $n =$

$2m + 1$ a node $3 \leq i \leq 2m - 1$ can be sometime Neutral and sometime a Joiner.

2. For the arithmetic sequence of order $2m$ we will apply the same algorithm in 1) with $f(v) = (2m - 1) - v - 1$ and

$$g(v) = \begin{cases} f(v) & f(v) < i \\ f(v) + 1 & f(v) \geq i \end{cases}$$

Example 6.4.3. We will consider the case when the sequence has $n = 2m$ nodes, for instance, the sequence with $n = 12$, for any node $3 \leq i \leq 6$. Following Lemma 6.4.2, we can show that node i is a member of V_{S_1} . Let us assume $i = 3$, and we will show that $i \in V_{S_1}$ as follows:

If $v = 9$, $f(v) = (2m - 1) - v - 1 \implies f(9) = 11 - 9 - 1 = 1$ which means we have to link node 9 to all nodes starting from node 1 except for node $i = 3$.

if $v = 7$, $f(v) = (2m - 1) - v - 1 \implies f(7) = 11 - 7 - 1 = 3$, in this case $f(7) = i$ which is not ideal as we want to avoid linking any node $m + 1 \leq j \leq 2m - 2$ to node i , so following Lemma 6.4.2 in this case, we defined $g(v) = f(v) + 1 \implies g(7) = 4$ so that we have to link node 7 to all nodes starting from node 4 as in Figure 6.3, that leads to having nodes 9 and 7 as Neutral nodes in the current sequence formation and node 3 as a Joiner node. Repeating this process for nodes 6 and 8, we will have nodes 0, 1, 2, 4, 5, 6, 7, 8, 9 as Neutral nodes leaving nodes 3, 10, 11 as Joiner nodes for this particular case.

It is clear that the node $i = 3$ can act as a Neutral node following another sequence formation, e.g. following the H-H linking formation; thus, $3 \in V_{S_1}$

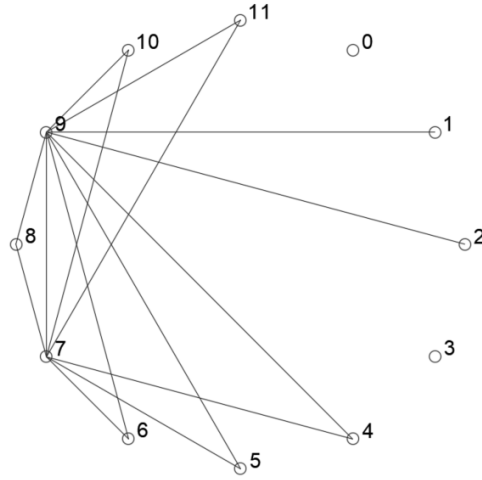


Figure 6.3: This figure shows that the arithmetic sequence S with $n = 12$ has node $i = 3$ acting as a Joiner node and nodes 7 and 9 as Neutral nodes without linking to node $i = 3$ in the current linking formation.

Node 7 is linked to the following nodes: 4, 5, 6, 8, 9, 10, 11.

Node 9 is linked to the following nodes: 1, 2, 4, 5, 6, 7, 8, 10, 11.

Example 6.4.4. We will consider the case when the sequence has $n = 2m + 1$ nodes, for instance, $n = 13$, for any node $3 \leq i \leq 7$, following Lemma 6.4.2. We can show that node i is in V_{S_1} . Assume $i = 3$, and we will show that it is $\in V_{S_1}$ as follows: If $v = 10$, $f(v) = 2m - v - 1 \implies f(10) = 12 - 10 - 1 = 1$ which means we have to link node 10 to all nodes starting from node 1 except node $i = 3$. If $v = 8$, $f(v) = 2m - v - 1 \implies f(8) = 12 - 8 - 1 = 3$, in this case $f(8) = i$ which is again not ideal as we want to avoid linking any node $m + 2 \leq j \leq 2m - 1$ to node i , so following Lemma 6.4.2 in this case we defined $g(v) = f(v) + 1 \implies g(8) = 4$ so that we have to link node 8 to all nodes starting from node 4 as in Figure 6.4, that leads to having nodes 8 and 10 as Neutral nodes in the current sequence formation and node 3 as a Joiner node. Repeating this process for nodes 7 and 9, we will have the nodes 0, 1, 2, 4, 5, 6, 7, 8, 9, 10 are Neutral nodes leaving nodes 3, 11, 12 as Joiner nodes for this particular case. Repeating this process for nodes 6 and 8, we

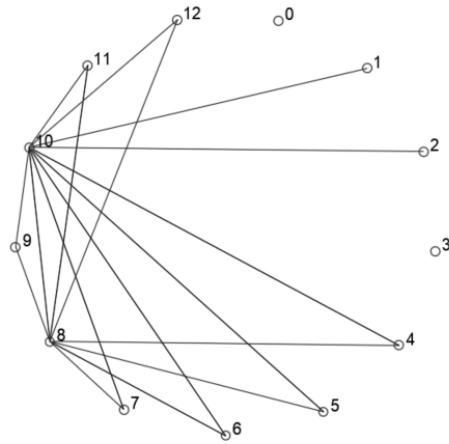


Figure 6.4: This figure shows that the arithmetic sequence S with $n = 13$ has node $i = 3$ acting as a Joiner node and nodes 8 and 10 as Neutral nodes without linking to node $i = 3$ in the current linking formation.

Node 8 is linked to the following nodes: 4, 5, 6, 7, 9, 10, 11, 12.

Node 10 is linked to the following nodes: 1, 2, 4, 5, 6, 7, 8, 9, 11, 12.

will have the nodes 0, 1, 2, 4, 5, 6, 7, 8, 9 as Neutral nodes leaving nodes 3, 10 and, 11 as Joiner nodes for this particular case. It is again clear that node $i = 3$ can act as a Neutral node following another sequence formation, such as following the H-H linking formation; thus, $3 \in V_{S_1}$.

6.5 An illustrative game example: the arithmetic target 4, 3, 2, 1, 0

Consider the arithmetic sequence $S = 4, 3, 2, 1, 0$ as a target sequence starting at an empty graph on $n = 5$ nodes. Figure 6.5 shows the interaction process of randomly nodes forming a link towards achieving the target. Here we consider all the possibilities for an individual to select one of the other Joiners to link to.

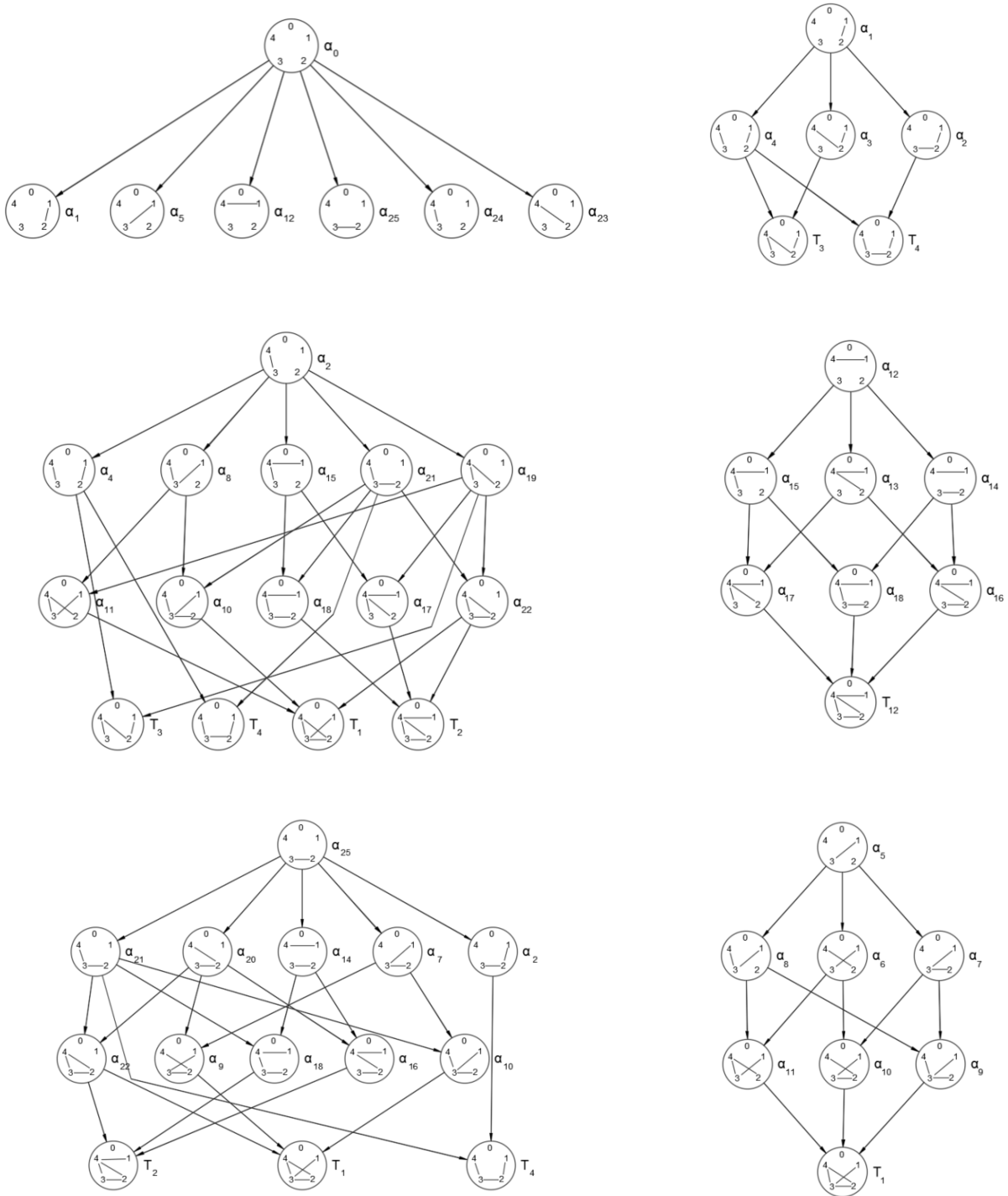


Figure 6.5: States α_1 , α_5 , α_{12} , α_{23} , α_{24} , and α_{25} represent all the possible graphs/states an individual can achieve moving one step from the empty graph when chosen at random to form a link. Other figures/ states show the later interaction process between the individuals after visiting the mentioned states to reach the terminal graphs T_1, T_2, T_3 , and T_4

α_i	x_{1-2}	x_{1-3}	x_{1-4}	x_{2-3}	x_{2-4}	x_{3-4}
α_0	0	0	0	0	0	0
α_1	1	0	0	0	0	0
α_2	1	0	0	1	0	0
α_3	1	0	0	0	1	0
α_4	1	0	0	0	0	1
α_5	0	1	0	0	0	0
α_6	0	1	0	0	1	0
α_7	0	1	0	1	0	0
α_8	0	1	0	0	0	1
α_9	0	1	0	1	1	0
α_{10}	0	1	0	1	0	1
α_{11}	0	1	0	0	1	1
α_{12}	0	0	1	0	0	0
α_{13}	0	0	1	0	1	0
α_{14}	0	0	1	1	0	0
α_{15}	0	0	1	0	0	1
α_{16}	0	0	1	1	1	0
α_{17}	0	0	1	0	1	1
α_{18}	0	0	1	1	0	1
α_{19}	0	0	0	0	1	1
α_{20}	0	0	0	1	1	0
α_{21}	0	0	0	1	0	1
α_{22}	0	0	0	1	1	1
α_{23}	0	0	0	0	1	0
α_{24}	0	0	0	0	0	1
α_{25}	0	0	0	1	0	0
T_1	0	1	0	1	1	1
T_2	0	0	1	1	1	1
T_3	1	0	0	0	1	1
T_4	1	0	0	1	0	1

Table 6.2: This table presents all 26 possible non-terminal graph states α_i and the 4 terminal graphs T_i for the target $t = 4, 3, 2, 1, 0$. Here X_{i-j} represents the edge between node i and node j ; if $X_{i-j} = 1$ then there is a link between i, j , if $X_{i-j} = 0$ there is no link.

That will give six possible states $X_{1-2}, X_{1-3}, X_{1-4}, X_{2-3}, X_{2-4}, X_{3-4}$ with one link formed in each of them following the first successful linking (and all individuals know which will reject a link so will not attempt a connection that will be rejected). Individuals are picked with equal probabilities to attempt to connect with others.

Individuals will repeat this procedure until a terminal graph is reached; by then, no further movement is allowed. In this example, we have 4 terminal graphs T_1, T_2, T_3 , and T_4 , see Figure 6.5.

We represent the 25 possible states to reach the terminal set, which contains four elements of the sequence $S = 4, 3, 2, 1, 0$ in Table 6.2, where zero means no edge between the opposite nodes and one implies a link between the correspondent nodes as in Table 6.2.

This example has been considered for Model A, see [10], where they analysed the optimal behaviour for the target sequence $4, 3, 2, 1, 0$ over the minimal set under both models, to determine multiple pure Nash equilibria (PNEs), sixteen in all. Thus, our work in this paper is consistent with the previous work, but we focused on the process under models B and C over the terminal set.

We note that the terminal set of the sequence $J = 3, 2, 1, 0$ is not the same as the terminal set of the sequence $J^* = 4, 3, 2, 1, 0$, which wasn't the case for the minimal set under model A, see [10] which showed that both sequences have the same minimal set. In Section 7.2 we consider a game-theoretical version of this case.

6.6 The maximum possible deviation for a graph from the terminal set of the arithmetic sequence

This section will show the maximum possible deviation that a graphical sequence in the terminal set could have. That would happen by forming the least possible links between the individuals in each case, leaving the vertices with no further possible links to create with others. To obtain this:

1. First, we present an algorithm to make the deviation worse or at least the

same as that of a graph with deviation that is potential maximally. The algorithm shows us that the maximal graphs (a graph whose score is optimised by this algorithm) must have the top part forming a clique and the bottom part being optimally linked (have the same score as H-H). The algorithm does not discuss the maximum possible score (which could be obtained by changing the top set's ratio to the bottom set). In the proof, we showed that if the deviation of the lower set is not optimal (not minimal), then we can still increase it.

The algorithm: To maximise the deviation of graph G, we will divide the nodes into two sets, shown below. Let G be a terminal graph reaching a deviation $s=t+d$; t is the missing links of the upper set, and d is the missing links of the lower set. The two sets are divided as follows:

- (a) lower set sequence $B = 0, 1 \dots n - a - 1$, should follow either the H-H algorithm or other linking rules, which lead to an equivalent minimal total deviation of the members of set B.

Proof: if the nodes of the lower set B are connected using the H-H algorithm, it will have a deviation of $\lfloor (n - a)/2 \rfloor$ (or $\lfloor (n - a - 1)/2 \rfloor$), while if the nodes of the lower set B are connected using another way of connecting, it will have a deviation of $l \geq k$. For sufficiently large a we can always find links to the top part for any nodes in the bottom part short of their target. Forming a link between the two sets will decrease the deviation of the total graph by 2; we can create k links when we use the H-H algorithm; thus, if s is the number of missing links from the upper and the lower set, forming k links it would be $s - 2k$ while using another way of linking, we will have total missing links is s' , thus by forming l links it would be $s' - 2l$. $s - 2k \leq s' - 2l$, thus, if we do

not use the H-H algorithm or an algorithm that provide a deviation of k to connect the lower nodes, more links will be formed with the upper set, decreasing the total deviation than in our indicated case.

- (b) For the upper set $T = n - 1, n - 2, \dots, n - a$, the link connections will form a clique eventually, provided that $a < \frac{n+4}{3}$ and so every node in T must be connected to all others. We will have a condition on the elements of the top set to form a clique. $\lfloor (n-a)/2 \rfloor$ (or $\lfloor (n-a-1)/2 \rfloor$).

For the top set $T = n - 1, \dots, n - a$ we will link all the nodes together to form a clique, that will leave the set with d_1 missing links which represents the deviation of the set T. This linking decreased the deviation of the smallest node $n - a$ by $a - 1$ links, thus it will become $n - a - (a - 1) = n - 2a + 1$. In the worst-case scenario, the node $(n - a)$ must be able to link to all other missing links from the bottom set, $\lfloor \frac{(n-a)}{2} \rfloor$. That means we must have $n - 2a + 1 - \lfloor (n - a)/2 \rfloor = \lceil \frac{n-3a}{2} + 1 \rceil \geq 0 \implies \frac{n-3a}{2} + 1 > -1 \implies a < \frac{n+4}{3}$. If $\lceil \frac{n-3a}{2} + 1 \rceil = 0$, then node $n - a$ will be a Neutral node, forming a clique in the top set and connecting to all other missing links in the lower set. In this case, $a = (n + 2)/3$ represents the minimum value we can consider for a in the top set in our algorithm. Considering this condition, we will have the top set forming a clique always.

2. To optimise the division between the sets A and B which will give the maximum deviation. Suppose we have n vertices divided into the top a vertices and the bottom $n - a$ vertices as previously discussed. The bottom $n - a$ vertices, are connected to achieve the "minimum deviation" with each vertex being either Neutral or Joiner. We will proceed by taking all targets from

$n - a - 1$, then apply Havel Hakimi algorithm to this sequence, then we switch links for breaks. We thus have $(n - a)/2$ missing links (or $(n - a - 1)/2$ if $n - a$ is odd).

The top a nodes then have target $(n - 1) + (n - 2) + \dots + (n - a) = an - a(a + 1)/2$. We then connect all of the top a to each other, which will reduce the total target by $a(a - 1)$. We reduce it further by connecting the bottom nodes which are missing $(n - a)/2$ links, so the bottom nodes are all on target and the number of missing links is

$$an - a(a + 1)/2 - a(a - 1) - (n - a)/2 = an - 3a^2/2 - n/2 + a. \quad (6.5)$$

When $n - a$ is odd the formula will become:

$$an - 3a^2/2 - n/2 + a + 1/2. \quad (6.6)$$

To find the maximum of this we differentiate with respect to a , thus we will get $n - 3a + 1$ i.e. the value of a which gives the maximum possible number of links in a graph of a terminal set satisfies

$$a = \lfloor (n + 1)/3 \rfloor \text{ or } \lceil (n + 1)/3 \rceil. \quad (6.7)$$

That will lead us to discuss three different cases based on the number of nodes n to define the formula of the maximum possible deviation of a graph from the terminal set in each case as follows:

1. A sequence of $n = 3k$ nodes.

Case(1): In this case we have $n = 3k$ vertices divided into top part a vertices and bottom part $n - a = 3k - a$ vertices. Thus, from equations 6.6, 6.5 and

6.7 we have the maximum deviation at either $a = k$ or $a = k + 1$ Now we will discuss the two cases:

Case(1-1): In this case we have $n = 3k$ vertices divided into top part $a = k$ vertices and bottom part $n - a = 2k$ vertices. We will find the formed links in each part in this case.

The number of the formed links in the lower/ upper part is given by following formulas respectively:

$$E_A(G) = \frac{\frac{2k(2k-1)}{2} - k}{2} = k^2 - k \quad (6.8)$$

$$E_B(G) = \frac{k(k-1)}{2} + k = \frac{k^2 + k}{2} \quad (6.9)$$

The number of formed links in both parts $E_1(G) =$ the sum of equations (6.8) and (6.9) i.e:

$$E_1(G) = \frac{3k^2 - k}{2} \quad (6.10)$$

Now the number of possible formed links of a graph (G) of a sequence $n = 3k$ is:

$$E_2(G) = \binom{3k}{2} = \frac{9k^2 - 3k}{2} \quad (6.11)$$

$$\implies MD(G) = E_2(G) - 2E_1(G) = \frac{3k^2 - k}{2}$$

Case(1-2): In this case $n = 3k$ vertices divided into top part $a = k + 1$ vertices and bottom part $n - a = 2k - 1$ vertices. We will find the formed links in each part as follows:

The number of the formed links in the lower/ upper part is given by the following formulae respectively:

$$E_A(G) = k^2 - 2k + 1. \quad (6.12)$$

$$E_B(G) = \frac{k^2 + 3k - 2}{2} \quad (6.13)$$

The number of formed links in both parts $E_1(G)$ = the sum of equations (6.12) and (6.13)

$$\implies E_1(G) = \frac{3k^2 - k}{2} \quad (6.14)$$

Given the equation 6.11 $\implies MD(G) = E_2(G) - 2E_1(G) = \frac{3k^2 - k}{2}$

We notice that the cases (1-1), (1-2) are giving the same form of the maximum possible deviation $MD(G) = \frac{3k^2 - k}{2}$ which is the worst possible score in the case of $n = 3k$.

2. A sequence of $n = 3k + 1$ nodes.

Case(2): In this case we have $n = 3k + 1$ vertices divided into top part a vertices and bottom part $n - a = 3k - a$ vertices. Thus, from equations 6.6, 6.5 and 6.7 we have the maximum deviation when $a = k + 1$.

Case (2-1): In this case we have $n = 3k + 1$ vertices divided into top part $a = k + 1$ vertices and bottom part $n - a = 2k$ vertices. We will find the set of formed links in each part at this case.

The number of the formed links in the lower/ upper part is given by following formulas respectively:

$$\implies E_A(G) = \frac{\frac{2k(2k-1)}{2} - k}{2} = k^2 - k \quad (6.15)$$

$$\implies E_B(G) = \frac{k(k+1)}{2} + k = \frac{k^2 + 3k}{2} \quad (6.16)$$

The number of formed links in both parts $E_1(G)$ = the sum of 6.15 and

$$6.16 = (k^2 - k) + \left(\frac{k^2+3k}{2}\right)$$

$$\implies E_1(G) = \frac{3k^2 + k}{2} \quad (6.17)$$

Now the number of possible formed links of a graph (G) of a sequence $n = 3k + 1$ is:

$$E_2(G) = \binom{3k+1}{2} = \frac{9k^2 + 3k}{2} \quad (6.18)$$

$$\implies MD(G) = E_2(G) - 2E_1(G) = \frac{3k^2 + k}{2} \quad (6.19)$$

The cases (2-1): gives the worst possible score when $n = 3k + 1$.

3. A sequence of $n = 3k + 2$.

Case(3): In this case we have $n = 3k + 2$ vertices. Thus from the equations 6.5, 6.6 and 6.7 we find that $a = k + 1$.

Case (3-1): In this case we have $n = 3k + 2$ vertices divided into top part $a = k + 1$ vertices and bottom part $n - a = 2k + 1$ vertices. We will find the set of formed links in each part at this case.

$$E_A(G) = \frac{\frac{2k(2k+1)}{2} - k}{2} = k^2 \quad (6.20)$$

$$E_B(G) = \frac{k(k+1)}{2} + k = \frac{k^2 + 3k}{2} \quad (6.21)$$

The number of formed links in both parts $E_1(G)$ = the sum of 6.20 and 6.21 = $(k^2) + \left(\frac{k^2+3k}{2}\right)$

$$\implies E_1(G) = \frac{3k^2 + 3k}{2} \quad (6.22)$$

Now the number of possible formed links of a graph (G) of a sequence $n =$

$3k + 2$ is:

$$E_2(G) = \binom{3k+2}{2} = \frac{9k^2 + 9k + 2}{2} \quad (6.23)$$

$$\implies MD(G) = E_2(G) - 2E_1(G) = \frac{3k^2 + 3k + 2}{2} \quad (6.24)$$

Example 6.6.1. The arithmetic target sequence $t = 8, 7, 6, 5, 4, 3, 2, 1, 0$ has the form $n = 3k = 9$; $k = 3$. By forming all the potential links following model B or C, we will get a terminal graph (G) in which no further movement between the individuals is possible. The maximum possible deviation from the target sequence t is given by $MD(G) = \frac{3k^2 - k}{2} = \frac{24}{3} = 12$ missing links.

Example 6.6.2. The arithmetic target sequence $t = 9, 8, 7, 6, 5, 4, 3, 2, 1, 0$ has the form $n = 3k + 1 = 10$; $k = 3$. By forming all possible links following model B or C, a terminal graph (G) will be reached. The maximum possible deviation from the target sequence t is given by: $MD(G) = \frac{3k^2 + k}{2} = \frac{24}{3} = 15$ missing links.

Example 6.6.3. The arithmetic sequence $t = 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0$ has the form $n = 3k + 2 = 11$; $k = 3$. By forming all possible links following model B or C, a terminal graph (G) will be reached. The maximum possible deviation from the target sequence t is given by: $MD(G) = \frac{3k^2 + 3k + 2}{2} = \frac{38}{2} = 19$ missing links.

Chapter 7

Game theoretical examples: reaching the terminal set

This chapter considers many game-theoretical versions of model B and model C over the terminal set. Individuals strategically choose the link to form or break following the procedure described in Chapter 6.

7.1 The game

The game over the minimal set considering model A has been discussed in Section 5.1. In this section, we will start by defining the game over the terminal set considering model B and model C. Suppose that individual i is selected randomly to potentially change one of their edges by breaking/forming a link with another individual in each state until reaching the desired target; a state is denoted by the edge set X as usual.

For n individuals, we have n distinct (pure) choices in general, an individual i can choose a possible change from the other $n - 1$ edges when they need to improve their target, or make no change if they are on target. Starting with model C: if

individual i was under target such that $e_i < t_i$ and attempts to form a link with individual j , given that $x_{ij} = 0, x^*_{ij} = 1$, the probability of individual i forming that link with j , v_{ij} , is given by:

$$v_{ij} = \begin{cases} 1 & e_j < t_j \\ 0 & e_j \geq t_j \end{cases}$$

That is because in model C if the individual i is under target and got chosen to improve its deviation from the target, it can form a link with another individual j only when j is under target too. In the other case where individual i is above the target such that $e_i > t_i$ and attempts to break a link with a chosen individual j , given that $x_{ij} = 1, x^*_{ij} = 0$, then the probability of individual i breaking that link with j , v_{ij} , is given by:

$$v_{ij} = \begin{cases} 1 & e_j > t_j \\ 0 & e_j \leq t_j \end{cases}$$

That is because in model C, if individual i is above target and wanted to break a link with individual j , the only way this is possible is when individual j is above target too. If the individual was on target and decided not to make any movement then in this case we denote $v_{ii} = 1$.

In model B, if an individual i is chosen and is over target $e_i > t_i$ and attempts to break a link with a chosen individual j , given that $x_{ij} = 1, x^*_{ij} = 0$, then they can break that link unconditionally, so that $v_{ij} = 1$ always. If individual i is under target such that $e_i < t_i$ and chooses to form a link with individual j , given that

$x_{ij} = 0, x^*_{ij} = 1$, then the probability of that link being broken j , v_{ij} is given by:

$$v_{ij} = \begin{cases} 1 & e_j < t_j \\ 0 & e_j \geq t_j \end{cases}$$

If the individual was on target and decided not to make any movement then again in this case we have $v_{ii} = 1$.

We note that for model A, as all links and breaks must be accepted, we have $v_{ij} = 1$ always.

Recall that the probability of individual i choosing individual j is $u_{ij}(X)$, $j \in (0, \dots, n - 1)$. For any two states x, x^* which differs from each other in one iteration, where $x_{ij} = 0, x^*_{ij} = 1$ or $x_{ij} = 1, x^*_{ij} = 0$.

$$P(X_{t+1} = x^* | X_t = x) = \frac{u_{j(x)} * v_{ij(x)} + u_{ji(x)} * v_{ji(x)}}{n}.$$

Following this Markov chain we must eventually reach the terminal set, and we need to find the final distribution over the states X , which we denote by $\pi^*(X)$. This is a vector of probabilities of length equal to the number of states in the terminal set. Then the payoff is denoted as the negative of their expected long term deviation $R_i(U_X) = - \sum \epsilon_i(X) \pi^*(X)$ where $\epsilon_i(X)$ is the deviation of i in state X .

7.2 A game theoretical example: $T=4, 3, 2, 1, 0$

Game theory is considered as an appropriate framework for studying mathematical models and investigating the strategic interactions of rational agents. However, the process as described may be a component of an evolutionary process, but it is not an entirely evolutionary process as explained in Section 7.5. In this section, we consider the target sequence 4, 3, 2, 1, 0 in which $n = 5$ and each individual is assigned to a target number of social links to achieve over model B and model C

starting from an empty graph. The strategic interactions will be to form/ break links as follows: Links are only added when both partners consent; thus, each Joiner individual will take the chance to create a connection with other neighbours whenever possible. Once it is on target, it can simply decline any proposed links from others (if it succeeds in achieving this). We consider the target sequence $t = 4, 3, 2, 1, 0$, which has also been considered extensively in the original dynamical process, see [10]. As no individual will form a link or allow another to link to it, if that will lead to it go over target, then there will be no attempts to break a link, both models are effectively the same. If an individual is short of the target, it will attempt to link to another group member that is not linked to it. As the composition of the population changes, the preferences of the individuals may also change. For this game, we noticed that it is better for the nodes 3, 4 that are $\in V_A$ and V_S to link to the individuals with the smallest original target. However, the always Neutral nodes $1, 2 \in V_N$ are more flexible in choosing, as they will reach their target by the end of this game no matter what strategy has been chosen (note that individual 0 will never form or accept a link). As mentioned above, we will start our game from an empty graph (zero links between the individuals). In the first iteration, a random individual will be chosen randomly to form a connection; after that, we will have six possible states with a single link formed in each $X_{1-2}, X_{1-3}, X_{1-4}, X_{2-3}, X_{2-4}, X_{3-4}$ each one having only one link formed. Following this process, we will have 25 possible intermediate graphs (states) during the linking process ending up with four terminal graphs. Table 6.2 shows all different possible graphs in this game.

α_i	v_1	v_2	v_3	v_4	P_4	P_3	P_2	P_1
α_0	$\alpha_1, \alpha_{11}, \alpha_{12}$	$\alpha_1, \alpha_{23}, \alpha_{25}$	$\alpha_{11}, \alpha_{24}, \alpha_{25}$	$\alpha_{12}, \alpha_{23}, \alpha_{24}$	-983/576	-101/144	0	0
X_{1-2}								
α_1	ϕ	α_2, α_3	$\alpha_2, \{\alpha_4\}$	$\alpha_3, \{\alpha_4\}$	-5/2	-3/2	0	0
α_2	ϕ	ϕ	T_4	T_4	-3	-1	0	0
α_3	ϕ	ϕ	T_3	T_3	-2	-2	0	0
α_4	ϕ	T_3	T_3	ϕ	-2	-2	0	0
X_{1-3}								
α_5	ϕ	α_6, α_7	α_6, α_8	α_7, α_8	-2	0	0	0
α_6	ϕ	α_9	α_{10}	α_9, α_{10}	-2	0	0	0
α_7	ϕ	α_9	α_9, α_{11}	α_{11}	-2	0	0	0
α_8	ϕ	α_{10}, α_{11}	α_{10}	α_{11}	-2	0	0	0
α_9	ϕ	ϕ	T_1	T_1	-2	0	0	0
α_{10}	ϕ	T_1	ϕ	T_1	-2	0	0	0
α_{11}	ϕ	T_1	T_1	ϕ	-2	0	0	0
X_{1-4}								
α_{12}	ϕ	α_{13}, α_{14}	α_{14}, α_{15}	α_{13}, α_{15}	-1	-1	0	0
α_{13}	ϕ	α_{16}	α_{16}, α_{17}	α_{17}	-1	-1	0	0
α_{14}	ϕ	α_{16}	α_{18}	α_{16}, α_{18}	-1	-1	0	0
α_{15}	ϕ	α_{17}, α_{18}	α_{18}	α_{17}	-1	-1	0	0
α_{16}	ϕ	ϕ	T_2	T_2	-1	-1	0	0
α_{17}	ϕ	T_2	T_2	ϕ	-1	-1	0	0
α_{18}	ϕ	T_2	ϕ	T_2	-1	-1	0	0
X_{2-4}								
α_{23}	$\alpha_3, \alpha_6, \alpha_{13}$	α_3, α_{20}	$\alpha_6, \alpha_{20}, \alpha_{19}$	α_{13}, α_{19}	-77/48	-13/16	0	0
α_3	ϕ	ϕ	T_3	T_3	-2	-2	0	0
α_6	ϕ	α_9	α_9, α_{11}	α_{11}	-2	0	0	0
α_{13}	ϕ	α_{16}	α_{16}, α_{17}	α_{17}	-1	-1	0	0
α_{20}	α_{16}, α_9	ϕ	α_{22}, α_9	α_{16}, α_{22}	-3/2	-1/2	0	0
α_{19}	$\alpha_{11}, \alpha_{17}, T_3$	α_{22}, T_3	α_{11}, α_{22}	α_{17}	-77/48	-13/16	0	0
α_9	ϕ	ϕ	T_1	T_1	-2	0	0	0
α_{11}	ϕ	T_1	T_1	ϕ	-2	0	0	0
α_{16}	ϕ	ϕ	T_2	T_2	-1	-1	0	0
α_{17}	ϕ	T_2	T_2	ϕ	-1	-1	0	0
α_{22}	T_2, T_1	ϕ	T_1	T_2	-3/2	-1/2	0	0
X_{3-4}								
α_{24}	$\alpha_4, \alpha_8, \alpha_{15}$	$\alpha_{21}, \alpha_4, \alpha_{19}$	$\alpha_8, \{\alpha_{21}\}$	$\alpha_{15}, \{\alpha_{19}\}$	-245/144	-101/144	0	0
α_{21}	$\alpha_{10}, \alpha_{18}, T_4$	α_{22}, T_4	α_{10}	$\alpha_{18}, \{\alpha_{22}\}$	-29/16	-29/48	0	0
α_8	ϕ	α_{10}, α_{11}	α_{10}	α_{11}	-2	0	0	0

α_4	ϕ	T_4	T_4	T_3	-5/2	-3/2	0	0
α_{15}	ϕ	α_{17}, α_{18}	α_{18}	α_{17}	-1	-1	0	0
α_{19}	$\alpha_{11}, \alpha_{17}, T_3$	α_{22}, T_3	$\alpha_{11}, \{\alpha_{22}\}$	α_{17}	-77/48	-13/16	0	0
α_{18}	ϕ	T_2	ϕ	T_2	-1	-1	0	0
α_{10}	ϕ	T_1	ϕ	T_1	-2	0	0	0
α_{22}	T_1, T_2	ϕ	T_1	T_2	-3/2	-1/2	0	0
α_{11}	ϕ	T_1	T_1	ϕ	-2	0	0	0
α_{17}	ϕ	T_2	T_2	ϕ	-1	-1	0	0
X_{2-3}								
α_{25}	$\alpha_2, \alpha_7, \alpha_{14}$	α_2, α_{20}	$\alpha_7, \{\alpha_{21}\}$	$\alpha_{14}, \{\alpha_{20}, \alpha_{21}\}$	-15/8	-29/48	0	0
α_2	ϕ	ϕ	T_4	T_4	-3	-1	0	0
α_7	ϕ	α_9	α_{10}	α_9, α_{10}	-2	0	0	0
α_{14}	ϕ	α_{16}	α_{18}	α_{16}, α_{18}	-1	-1	0	0
α_{20}	α_9, α_{16}	ϕ	α_9	$\alpha_9, \{\alpha_{22}\}$	-2	-1/2	0	0
α_{21}	$\alpha_{10}, \alpha_{18}, T_4$	T_4, α_{22}	α_{10}	$\alpha_{18}, \{\alpha_{22}\}$				
α_9	ϕ	ϕ	T_1	T_1	-2	0	0	0
α_{10}	ϕ	T_1	ϕ	T_1	-2	0	0	0
α_{16}	ϕ	ϕ	T_2	T_2	-1	-1	0	0
α_{18}	ϕ	T_2	ϕ	T_2	-1	-1	0	0
α_{22}	T_1, T_2	ϕ	T_1	T_2	-3/2	-1/2	0	0

Table 7.1: This table shows all the possible transitions that occur between pairs of states during the game on the path from the empty graph to graphs within the terminal set, together with the strategies that individuals chose in each position.

α_i : The non-terminal graph state i (see Table 6.2 for the specific links of the graph).

T_i : The terminal graph i (again see Table 6.2).

v_i : The entry in this column represents the possible states to move to if the i th individual is to choose.

ϕ : This represents the choice not to move, which occurs when individual i is on target.

P_i : The expected payoff of individual v_i at the corresponding state α_i (evaluated using backwards induction from the terminal set, and assuming if an individual is ambiguous about its transition, it makes each choice with equal probability).

$\{\}$: This represents when individual v_i will choose the strategy in the curly bracket with probability =0

Moreover, when a random state is reached, and an individual is selected to form a link, we might have the possibility to move to more than one state or terminal

graph in the following round. For example, in the state α_{23} when individual 1 is in deficit to form a link, it will have the possibility to move to three different states, which are α_3, α_6 and α_{13} . The selected individual can choose the link to form, and here, the individual might prefer to make one move rather than another, or it might have no preference. This will depend upon the expected payoffs that it will receive for each move. If an individual has no preference, we will assume that there is some non-zero probability of selecting all possible moves. If it has a preference, it will choose the preferred choice with probability 1 (any other option will be selected with probability 0). After following the random process, the game ends when no more possible links can be formed. The total payoff is taken as an average of all payoffs in the game. Table 7.1 explains the set of possible movements and payoffs of the individuals in each state. By the end of this game, we will obtain the terminal set $T = \{T_1, T_2, T_3, T_4\}$ and define the optimal strategy for each individual to maximise their payoff at every state using backwards induction. Following the process of having the top nodes linking to the nodes with the lowest target will increase the top nodes' possibility of having more links with other individuals, which will reduce its deviation and will lead to a better payoff. We notice from Table 7.1 that when individual v_4 is linked to v_1 (lowest target), occurring in states $\{\alpha_{12}, \alpha_{13}, \alpha_{14}, \alpha_{15}, \alpha_{16}, \alpha_{17}, \alpha_{18}\}$, we have $P_4 = -1$ which is the best possible payoff for v_4 in this game. If v_4 is linked to individual v_2 , as in states $\{\alpha_{23}, \alpha_3, \alpha_6, \alpha_{11}, \alpha_{20}, \alpha_{19}, \alpha_9, \alpha_{22}\}$, or to I_3 as in $\{\alpha_{24}, \alpha_{21}, \alpha_8, \alpha_4, \alpha_{19}, \alpha_{10}, \alpha_{22}, \alpha_{11}\}$, then the payoff of v_4 $\{P_4\}$ will have worse values than -1 . Let us consider another state in this game from Table 7.1 where the individuals have a preference for one strategy over others. We notice that individuals will have the same type of preference; for instance in state $\{\alpha_{25}\}$, v_3 prefers to have a link with v_1 and move to state α_7 rather than having a link with v_4 and move to state α_{21} , in the same

state α_{25} individual v_4 will choose to link to v_1 and move to state α_{14} and not to link with individuals v_2 or v_3 , which would lead to states α_{20} and α_{21} respectively. When the selected individual is in V_N , here for nodes $\{1, 2\}$, then no matter what individual they will link to, they will eventually achieve their target. We see this in Table 7.1, where the payoffs of nodes 1 and 2, P_1 and P_2 , are always equal to zero. Thus, those individuals will not prefer linking to any one node over another.

7.3 Suggested Strategies

The previous example suggested some strategies to possibly follow. This section introduces new strategies for the individuals to follow more generally. We will consider game examples where all individuals but one follow one strategy at each step while the other individual uses a different strategy. In particular we will consider whether, out of strategy set, the population strategy is optimal for the individual or if it can be "invaded" by an alternative strategy. In our study, we will consider an ESS analysis following matrix games, so allowing for many potentially different mutants. An alternative idea would be to consider the memory of the individuals at each time step, so the individual can make a better judgement of the required action and recall the history of movements they went through before the current state. However, we did not consider this situation here as our models were based on a previously established model and we built on that methodology. We saw in the 4, 3, 2, 1, 0 example from Section 7.2, that connecting to the individuals with the lowest target was a good strategy. In particular we used three variants of the lowest target strategy. Here, we will introduce the strategies chosen as follows:

1. The least target strategy LT: in this strategy, any node in deficit to link with others will always link to the node with the lowest possible target.

2. The least remaining strategy LR: in this strategy, any node in deficit to link with other neighbour nodes will always choose to link to the node with the lowest remaining links required to achieve its target at that stage.
3. The least ratio strategy LRA: For a general sequence t , we will consider the dynamical process described in Chapter 6 applied to the nodes of this sequence, and in particular two nodes with remaining targets n and m , which we refer to simply as nodes n and m below, with original targets N and M , respectively, that are still Joiners at this stage of the game.
 s_m : is the set representing the number of the nodes that have not linked to the node m yet.
 s_n : is the number of the nodes that have not linked to the node n yet.
 v : A node in deficit selected to form a link, which can link to either n or m .
For node v to follow the LRA strategy, it considers $\frac{m}{s_m}$ and $\frac{n}{s_n}$ and selects the one with the smaller value.
4. The greatest target strategy GT: In this strategy, any node in deficit will always link to the node with the greatest possible target.
5. The greatest remaining strategy GR: In this strategy any node in deficit will link to the node with the greatest number of missing links to achieve its target at that stage of the game.

We thought about these strategies for the following reasons:

- GR and GT, we assumed that these two strategies give the best overall deviation for the group. We noted some aspects of the similarity between those two strategies and the H-H algorithm.
- LT and LR: we thought about those strategies as they reduce the individuals' deviation as seen in the example sequence 4, 3, 2, 1, 0. Since the probability

of losing the chance to connect to the nodes with lower targets will be higher, they will try to link to the nodes with lower targets first.

- LRA: Consider, as above, a node v which can link to nodes with remaining targets n and m , and s_n and s_m nodes, respectively, which can link to them. If node v decided to link with node n , then the probability of losing the link with node m (assuming all the nodes which could link to m would if they got the chance) is: $P(A) := (s_m - 1)/s_m * (s_m - 2)/(s_m - 1) * \dots * (s_m - n)/(s_m - m + 1) = 1 - m/s_m$ If the node v decided to link to node m instead, then the probability of losing the link with node n (assuming all the nodes which could link to n would if they got the chance) is: $P(B) := (s_n - 1)/s_n * (s_n - 2)/(s_n - 1) * \dots * (s_n - n)/(s_n - n + 1) = 1 - n/s_n$ Thus, to decide the best link, we have to compare between m/s_m and n/s_n . If $m/s_m < (>)n/s_n \Rightarrow p(A) > (<)P(B)$ in other words the chance of losing node m is greater (less) than losing node n ; in the this case its better for v to link to the node with less ration .

7.4 Examples showing that our alternative strategies are not always optimal

Strategies can be determined based on many perspectives such as history. In this case, we take into account the history of interactions with previous players. Another factor is reputation; a player with a better reputation is more likely to receive help. Many other criteria, such as spatial or other clustering, can be examined, but in our study, we focused on the individuals' target. However, none of the strategies we examined are optimal in all cases, and we believe that categorising such a strategy would be difficult. We will present many cases demonstrating that

the strategies we investigated were not optimal.

7.4.1 Strategy LT is not always optimal

Assume that we have the sequence $T = 6, 5, 4, 3, 2, 1, 0$, LT is not always optimal. Suppose the first four random selections were for the node 5, following the LT strategy, 5 will link to 1, 2, 3 and 4 respectively. Then suppose 7 is then picked three times to play; it will link to 2, 3, and 4. Finally suppose that 6 is selected once; it will link to 3. We have the following links at this stage: $5 - 1, 5 - 2, 5 - 3, 5 - 4, 7 - 2, 7 - 3, 7 - 4, 6 - 3$, in which individuals 1, 2 and 3 are Neutrals. Individual 4 is connected to 5 and 7, and it is still short of two links. If Node 4 is chosen, the only choice is to link to node 6, and it will stay short of one link even after making that connection. Node 6 has one link with node 3, and it will be short of target no matter what other links it forms. Thus for any linking formation, 6 will eventually link to 4. Similarly, 6 will link to 7. 5 needs one link to reach its target, and that can be to 6 or 7. Let us assume everyone is picking the lowest target.

If 4 is picked, it will connect to 6.

If 5 is picked, it will connect to 6.

If 7 is picked, it will connect to 5, which will leave 5 Neutral.

Thus if 6 is picked, if it links to 5 then eventually it will connect to 4 and 7, but if it picks 4 (or 7) it may or may not get to link to 5; it will not get it if and only if 7 is picked to choose before 5 or 6. Thus, 6 should pick 5 and not 4, i.e. not follow LT in this state.

7.4.2 Neither LR nor LRA are always optimal.

Consider the sequence $t = 9, 8, 8, 8, 8, 6, 3, 1, 1, 1, 1$ where individuals will play a game to reach the terminal set using either the LR and LRA strategies (in our

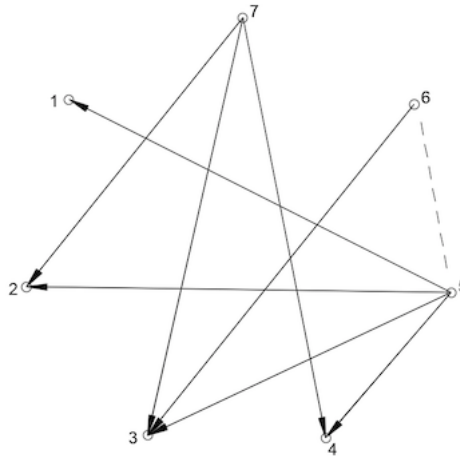


Figure 7.1: This figure illustrates that choosing the LT strategy for node $i = 6$ in the arithmetic sequence 7, 6, 5, 4, 3, 2, 1, 0 is not always optimal. \rightarrow : Refers to a link between the pair of individuals.
 - - - - Refers to the optimal link node 6 can have.

example both will give the same moves to play). Suppose the first four random selections were using each of the four 8s – these then connect to the 1s in turn (under either rule), which make all the 1s on target. Suppose the 6 is picked the next five times; it will connect to the 3 and then to the four 8s (again under either rule). Suppose the 9 is picked; it can link to the 8s or 6; we will discuss the best choice as follows. The 6 has one missing link to achieve the target and 5 nodes that are not linked to it; 9 is the only possible choice for 6 as all other nodes are on target. The 3 has two missing links and nine nodes that are not linked to it. However, any of the 8s would connect to the 3 if given the chance. Thus LR and LRA would say to pick the 6, but it is best to pick the 3, assuming that the rules are followed subsequently.

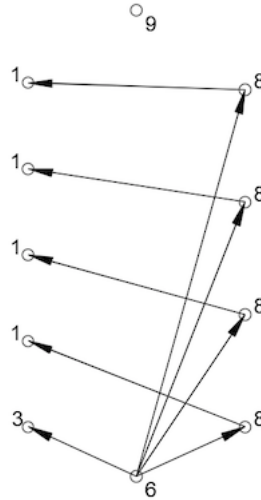


Figure 7.2: This diagram shows that for the sequence 9, 8, 8, 8, 6, 3, 1, 1, 1, 1, it is not always optimal to use the LR or LRA strategy. Following the linking sequence shown, after node 9 is selected LR and LRA would link to 3, but the optimal choice is 6, assuming the rules are subsequently followed. \rightarrow Refers to a link between the pair of individuals.

7.5 Game simulations

In this section, we consider two game simulations. In general, we consider an infinite population that is divided into a number of target classes, where members of each target class member are assigned the same target number of social relations. Groups are formed at random consisting of one member of each target class (this could be thought of as a social group with for example a dominant male and female, and subordinate individuals fulfilling different roles, for example).

The individuals in each target class can play different strategies. In the games we considered, we start by assuming a specific set of strategies to be played by all players. Individuals in all classes but one are then assumed to play this given strategy, whilst we allow the remaining class to consider an alternative strategy (as is standard in the analysis of multi-role games, see [11]). Solutions of this type are ESSs of the multi-role game defined. We are examining individuals' different

strategies to check whether they have a better alternative than the one they picked and see if we can find an optimal strategy in general. We did not consider birth, death, and replacement events at the single life time t in the game. Each of these simulations could be interpreted as a social process occurring within a short lifetime considered as a snapshot from an evolutionary process, the outcome of which determines fitness will be integrated into the evolutionary process. The fitness values tell us only what might happen from generation 1 to generation 2. In the second generation, following replacement events, there will be a different distribution of strategies and potentially different payoffs. As the populations in the evolutionary process are infinite, then small differences between mean payoff values matter. The stochastic variation is caused by the number of simulations performed.

7.5.1 Game simulation of the sequence $t = 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0$

This section consider a game where each player tries to improve their own target over the arithmetic target sequence $t = 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0$. At each time step, one player will be chosen to play several strategies, while the others will stick to a single strategy. We utilised python to run simulations of 100000 trials to find the potential deviations in each scenario. However, running these simulations for large populations or networks (e.g. > 100) is highly computationally expensive thus we did not do this. The results of these simulations are displayed in Tables 7.3 and 7.2. To discover how each iteration's approach performed, see Section 9.8. Allowing the selected individual to play either the population (resident) or an alternate (mutant) way enabled us to analyse their decisions and better grasp the game's overall outcomes. However, we focused on the impact of a certain target's strategy selection while playing the game, precisely when all the individuals play the same strategy. Then we compared the values in the main diagonals when all

individuals follow the same strategy to the values when one individual follows a different strategy; if the number in the main diagonal is larger, the population will be stable against that invasion. We can consider the GT and GR strategies in the last two columns as cooperative behaviours, as they reduce group deviation, while the LT, LR, and LRA strategies in the first three columns can be thought of as defecting behaviours, as they tend to optimise individual behaviour at the expense of the group. The numbers on the main diagonal represent group fitness, whereas the numbers off the main diagonal represent invasion plans. When the target individual we're monitoring was playing strategy B then chose to play strategy 'A' and profits from it, it's said that strategy 'A' had invaded another strategy 'B'. Stable strategies that can resist invasion from a range of individuals will be of interest to us. A dominant strategy invades all other strategies and resists all of them. It is worth mentioning that we are using only the the five strategies previously described so our results do not indicate a wider stability against all possible alternatives.

Player 4	LT	LR	LRA	GT	GR
LT	0	0	0	0	0.00004
LR	0	0	0	0	0.00001
LRA	0	0	0	0	0
GT	0	0	0	0	0.00239
GR	0	0	0	0	0.001419
Player5	LT	LR	LRA	GT	GR
LT	0.0011	0.00779	0.0085	0	0.00251
LR	0.0008	0.0076	0.008	0	0.0022
LRA	0.0009	0.00713	0.0075	0	0.00215
GT	0.01274	0.03092	0.04689	0	0.08704
GR	0.12489	0.000001	0.04758	0	0.05693
Player6	LT	LR	LRA	GT	GR
LT	0.3164	0.34049	0.343149	0.04539	0.025838
LR	0.31676	0.33401	0.34175	0.04908	0.026036
LRA	0.314219	0.33335	0.33574	0.0463	0.025459
GT	0.9896	0.96601	1.16216	0.69098	0.43139

GR	0.99539	0.96369	1.16057	0.52707	0.37018
Player7	LT	LR	LRA	GT	GR
LT	1.22347	1.250847	1.265849	0.119551	0.11192
LR	1.22047	1.235515	1.24278	0.130204	0.11333
LRA	1.21281	1.235367	1.242447	0.124273	0.110549
GT	2.42922	2.381531	2.743464	1.037359	0.817492
GR	2.42692	2.384385	2.743464	0.893895	0.73773
Player8	LT	LR	LRA	GT	GR
LT	2.333726	2.358718	2.369837	0.244738	0.331415
LR	4.537371	2.335007	2.334183	0.255179	0.335913
LRA	4.54027	2.334050	2.334351	0.254017	0.32752
GT	4.418941	0.255335	3.972314	1.130119	1.261858
GR	3.737865	0.335666	3.985187	1.069198	1.175966
Player9	LT	LR	LRA	GT	GR
LT	3.33655	3.446351	3.473516	0.532385	0.822996
LR	3.438482	3.436052	3.432149	0.532743	0.828338
LRA	3.430697	3.432144	3.430779	0.532358	0.819563
GT	4.962515	4.983138	5.07025	1.096696	1.748921
GR	4.963939	4.988118	5.081973	1.082666	1.683261
Player10	LT	LR	LRA	GT	GR
LT	4.53506	4.53948	4.54114	1.009230	1.539738
LR	4.520815	4.504275	4.49983	1.00920	1.543411
LRA	4.508736	4.499296	4.49669	1.009165	1.538757
GT	6.12286	6.106747	6.114093	1.043490	2.259248
GR	6.125179	6.114746	6.127229	1.04263	2.216735

Table 7.2: The sequence 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0 is examined here, with a focus on the deviations when nodes $i = 4, \dots, 10$ are selected to alternate. The columns: LT, LR, LRA, GT, GR stand for population strategies, whereas the rows indicate which strategy player i will use at each time step.

Player5	LT	LR	LRA	GT	GR
LT	11.85776	11.83604	11.8417	9.32448	9.6857
LR	11.87094	11.8619	11.84164	9.26168	9.67338
LRA	11.85025	11.8421	11.8403	9.34474	9.69274
GT	10.62004	10.6112	10.63794	5	6.098239
GR	10.6037	10.6001	10.6257	5.315	6.23986
Player 6	LT	LR	LRA	GT	GR
LT	11.857504	11.8468	11.8509	7.54882	8.31486
LR	11.8553	11.84898	11.8472	7.436	8.2858
LRA	11.8491	11.84506	11.8452	7.5336	8.32434
GT	11.759282	11.670498	11.794616	4.9999	6.22084
GR	11.755824	11.668856	11.79182	5.069151	6.24066
Player 7	LT	LR	LRA	GT	GR
LT	11.85966	11.86251	11.857122	6.401502	7.14503
LR	11.85384	11.850126	11.84512	6.329344	7.12741
LRA	11.8496	11.84963	11.851508	6.3722	7.1483
GT	12.5889	12.525106	12.567722	5	6.2868
GR	12.59616	12.527978	12.567722	5.003643	6.2426
Player 8	LT	LR	LRA	GT	GR
LT	11.855640	11.865448	11.863684	5.505616	6.258158
LR	11.854138	11.852752	11.848988	5.487629	6.25718
LRA	11.858068	11.849188	11.8513	5.49259	6.260088
GT	12.753054	5.48764	12.735254	4.999	6.323382
GR	12.749314	6.257008	12.73693	5.000078	6.242818
Player9	LT	LR	LRA	GT	GR
LT	11.852258	11.864442	11.860504	4.9	7.609515
LR	11.855068	11.856038	11.847249	5	5.899348
LRA	11.8555	11.850874	11.847304	5	5.892338
GT	12.763886	12.758942	12.755354	4.999	6.320618
GR	12.761216	12.754538	12.752876	4.9	6.24323

Table 7.3: This table displays the group's overall score when individual i is chosen to alternate their strategy considering the sequence 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0. In particular, individuals $i = 5, \dots, 9$ are the ones who got chosen to change their strategies. The columns: LT, LR, LRA, GT, GR stand for population strategies, whereas the rows indicate which strategy player i will use at each time step.

Table 7.2 displays the participants' deviations from their targets while using the described strategies in a group of 11 players. We follow a single player considering the range of strategies while all others use a given strategy. We briefly discuss the outcomes from Tables 7.2 and 7.3 and will go into greater detail when player 5 is in deficit and can switch strategies in order to achieve the smallest possible deviation from the target at that stage of the game, as well as explore certain Prisoner Dilemma scenarios as follows.

Look at the first column, where everyone else in the group is using the LT strategy, and player 5 is alternating between the strategies. We can see that it is difficult to determine which of the least strategies is performing better for individual 5, and since they all have a very close deviation value and the numbers are almost equal, no conclusion can be drawn from them. Hence, we ran a larger number of trials (10 million) times to better understand the behaviour, the deviations of player 5 following each of the least strategies had the values respectively: 0.0008679, 0.0008429, 0.0008476. That's why these numbers don't match the numbers in the Table 7.2. Looking at these simulations we see that LR and LRA both invade LT. In column 2, we notice that the LRA strategy invades the LR strategy. In column 3, considering the least strategies only, when player 5 plays the LRA strategy the same as other neighbours, their deviation will be minimal, and there will be no need to change the strategy. In column 4, we can see that while the population is using the GT strategy, player five will be able to use any method to achieve the same deviation. In column 5, the GR strategy has been invaded by any of the least strategies: LT, LR and LRA, and the best for him will be the LRA strategy. In general, we notice that the LRA is a dominant strategy, as it invades all others and resists invasion. In Table 7.2, looking at the main diagonal, where everyone is playing the same least strategy, we can see that the LRA is the optimum strategy

among all the least strategies for player 5. However, looking at the main diagonal in the total score Table 7.3, where all individuals are playing the same least strategy, we notice that LRA is the optimum strategy among all the least strategies for the group collectively. Hence we can conclude that this does not represent a Prisoner's Dilemma type game.

Moving on to the table of player 6, in column 1, we can see that the LRA strategy invades all other strategies. It would be better for player 6 to stick to the LRA strategy to obtain the minimum possible deviation. However, that is not the case for the entire group, since Table 7.3 shows that the GT strategy is the best for them. At the same time, GT is the most vulnerable to invasion. If individual 6 chose to play the group strategy, the deviation would be worse than if it switched to one of the least strategies. This is a classical Prisoner's Dilemma.

Consider when player 7 is the alternating player, see Table 7.2. Looking at the main diagonal, we can see that all the least strategies have very close values, making it difficult to determine whether or not there is a Prisoner's Dilemma among them. In general, it is best for player 7 to play one of the least strategies. However, in the following example, we can define a prisoner Dilemma: Look at the 2 x 2 game matrix, where the players select from the LT and GT strategies as follows:

		The group strategies	
		<i>LT</i>	<i>GT</i>
Player 7	<i>LT</i>	1.22347	0.119551
	<i>GT</i>	2.42922	1.037359

If individual 7 decides to play LT while the population plays GT, the best possible deviation is 0.119551. When player 7 decides to play the population strategy Gt it will get the second-best deviation which is 1.037359. In the other case, when node 7 decides to play the population strategy LT it will get the third-best deviation which is 1.22347. The worse possible deviation is 2.42922 when individual7 play

GT and the population play LT. This is a Prisoner's Dilemma.

7.5.2 Game simulations for the sequence $t= 2, 2, 2, 2, 2$.

1st 2	LT	LR	LRA	GT	GR
LT	0	0	0	0.68184	0
LR	0	0	0	0.59328	0
LRA	0	0	0	0.6712	0
GT	0	0	0	0.56283	0
GR	0	0	0	0.69395	0
2nd 2	LT	LR	LRA	GT	GR
LT	0	0	0	0	0
LR	0	0.1983	0.20129	0.19874	0.20015
LRA	0	0.1997	0	0	0
GT	0	0.20049	0	0.42068	0.42199
GR	0	0.20032	0	0.42261	0.00571
3rd 2	LT	LR	LRA	GT	GR
LT	0.15916	0.40243	0.16011	0.20431	0.03125
LR	0.15996	0.40159	0.15818	0.17267	0.02987
LRA	0.16244	0.4008	0.16011	0.20063	0.03082
GT	0.2057	0.42517	0.20217	0.1594	0.08383
GR	0.20962	0.43548	0.21103	0.2135	0.07503
4rth 2	LT	LR	LRA	GT	GR
LT	0.41794	0.41923	0.41718	0.41242	0.41968
LR	0.42063	0.53272	0.53217	0.53393	0.53363
LRA	0.42213	0.5335	0.41797	0.15989	0.42052
GT	0.41569	0.53055	0.42885	0	0
GR	0.42287	0.53103	0.4229	0	0.16613
5th 2	LT	LR	LRA	GT	GR
LT	0.55828	0.53306	0.56156	0	0.0967
LR	0.55709	0.53659	0.56259	0	0.09646
LRA	0.56097	0.53267	0.56101	0	0.09719
GT	0.68551	0.59386	0.68412	0	0.20407
GR	0.66076	0.59423	0.68313	0	0.18053

Table 7.4: This table shows the deviations of the nodes in the sequence 2, 2, 2, 2, 2 following different strategies, where each node is assigned a numbered position and all nodes are chosen to alternate. The columns LT, LR, LRA, GT, GR represent population strategies, while the rows indicate which strategy player i will use at each time step.

In this example game of the sequence 2, 2, 2, 2, 2, we ran 100000 simulations using

python to examine the players' behaviour in each case. Since we have ties in this example, each 2 has been assigned to a numbered place in the program, thus the first 2 will be assigned to the first node position and summoned as node one, and so on. The simulations results are summarised in Table 7.4, where we can see that the LT and GT strategies behave almost the same. When the 1st 2 and 5th 2 are the ones who change their strategies while the rest of the group plays the same strategy at the same time. In the table, when node 3 plays LT, and all other nodes play GT it will get a deviation equal to 0.2057, which is very close to the deviation when node 3 plays GT and all other nodes play LT as it will get a deviation equals to 0.20431. In general, looking at the three tables, we notice that GT is a stable strategy as it has the lowest deviation among all the deviations. Hence, no strategy invades it in the column of GT. Looking at the main diagonal, we notice that the least strategies are not dominant in the game, and the greatest strategies are taking over them. By running some simulations we got the following:

Least Remaining: [0, 1, 2, 3, 4]

The target list is: [2, 2, 2, 2, 2]

Average Score of 1000000 graphs is given as follows:

[0.0, 0.0, 0.3336779999999999, 0.666836, 0.6657740000000001]

considering 1000000 trials we found that the average sum of 1000000 graphs is: 1.666288.

It shows that LR tends towards leaving one of the nodes out. There is a higher probability of it leaving a node out.

Chapter 8

Discussion

In this thesis, we further developed the model introduced initially by Broom and Cannings in [8]. We considered a group of individuals represented as vertices. These individuals are not identical, for example they may have different reproductive capacities, social demands. depending on other factors including age, location or gender. Thus, each individual was assigned to their own desired target. The time scale considered was relatively short, and no individuals were born or died; however, the relations between these individuals, represented as a set of edges, were constantly changing, governed by the individuals preferences. This interactions occur over simple undirected graphs, and the process led to be *minimal set* where the collective deviation of the group from the target is at its smallest value. The set of all sequences that achieve the minimal deviation from the target sequence and the associated graphs are the minimal set. We found a general formula to find the size of the minimal set and its precise composition for the arithmetic sequence. We also discovered a recurrence relation for the size of the minimal set for the all or nothing sequence by showing its equivalence to an existing problem. In another part of the thesis, we considered the game-theoretical model developed in [10] where the question of whether it is never optimal to leave the minimal set (and

so considering games only on the minimal set was in some way sufficient) was left open. Following the chosen strategies and starting at a (perhaps randomly) chosen initial state, transitions representing the individual's choices will continue indefinitely, leading to the distribution over the states following the unique stationary distribution (or if there is no such distribution, a weighted average over stationary distributions on irreducible subsets of the states). Thus we can find the payoffs for any set of strategies and potentially Nash equilibria for the game. In Section 5 we give a counter-example showing that leaving the minimal set is an optimal strategy for one of the players, thus considering the minimal set is not always sufficient to solve the game. Jackson and Wolinsky [43] studied a case example closely related to ours as mentioned earlier, where individuals' payoffs depend upon the network, taking into consideration the incentive for individuals to form networks. In this case, individuals can form or break links with others following the same process as in our model B and C. They classified stable and efficient networks into two models. They used two ways of deciding the payoffs to the players: the equal split rule and another rule that emerged naturally from the players' bargaining. They demonstrated that evolutionary dynamics on graphs could favour cooperation over defection when the cost ratio exceeds the average connectivity favourably.

In [12] a particular case of the model was discussed where the sequence has a score equal to 1, this case being the closest to the graphical sequence where the score is equal to zero, and all the individuals are on target. Many example sequences were considered. Furthermore, they demonstrate a methodology for analysing score one games. We revisited an exceptional example sequence 1, 1, 1 considered in that paper in this work. Broom and Cunnings defined three Nash equilibria for this sequence, and at first sight, it appeared that these were the only ones yielding a rare simple but non-trivial case for this game. For this sequence, we considered the

individuals' strategies in more depth (the strategy of each individual comprises the set of all of their choices at the states for which they are in deficit). We identified strategies that led to another six cyclic Nash equilibria for the game and the original three solutions, but found no other solutions. That does not mean that no such solutions exist, and a complete game analysis would be very complicated. Thus although the game over this sequence initially seemed relatively simple, there has turned out to be significant complexity. All target sequences thus seem to yield one of the following.

Trivial games over the minimal set, where the number of players which can be in deficit is 0 (i.e. graphical sequences with no game at all).

Two potentially off target players in deficit (leading to simple pairwise swaps between the individuals).

More than two players potentially off target but involving a collection of such simple swaps (such as the sequence 2,0, 0).

Complex games involving three or more interacting potentially off target individuals, with our example as potentially the simplest of these.

In Sections 6 and 7, we introduced two new models derived from the original model with some differences between the models that have been clarified. We showed that the process evolved to a state-terminal set of graphs, and were able to classify the models' vertices considering different initial graphs to reach the terminal set along with defining other properties of the terminal set that further simplified the vertex categorisations of the sequences associated with terminal graphs. We investigated the reverse Havel Hakimi Theorem, and we showed that it tests a sequence to show if it is graphical or not (similar to the Havel Hakimi Theorem). We found the reverse Havel Hakimi algorithm, which determines the classifications of the nodes of the terminal set. Moving to the arithmetic sequence of model B and

C starting from an empty graph, we classified the nodes of such sequences. We considered some game-theoretical versions of models B and C, starting from an empty graph over the terminal set with suggested strategies to follow. We studied game simulations of two sequences using python, and we considered the payoffs at each state to suggest possible strategies.

There are a number of potential ways to develop this work in the future. Thinking of the original model, an individual may have a preference to connect with some individuals over others, so that some links will be preferable to others, so-called "transitivity preferences". Another interesting possibility is to only allow certain links to be formed or indeed force certain links to be formed (there may be people you cannot interact with, or have to interact with even if you would prefer not to). We may have a spatially distributed population, and individuals may only form links with their close neighbours. An interesting question is whether this complexity is inherent in such games or relies on specific assumptions of our model. We could consider variants of this game that might be more amenable to analysis in future work. We have already considered alternative models, Models B and C, in which individuals must agree to form a link in both models before it can be established, but in Model B, either individual can break a link without having any restrictions, whereas breaking a link requires the consent of both parties in Model C. Further work can be done for these models, like further investigating the properties of the models for general and arithmetic sequences, for example looking for a formula of the size of the terminal set for different sequences, and finding the elements of the terminal set of a given sequence. We investigated an example that showed that we can get an optimal solution by leaving the minimal set in section 5; possible future work is to investigate if we can find better strategies by leaving the terminal set. A further potential future research direction is to embed

the models A, B, and C into an evolutionary process such as the Moran process by considering two linking strategies 1 and 2 for individuals in a population. The Moran process determines the evolution of the population, but at each stage, we play the linking game to evaluate the corresponding payoffs over individuals for the associated population mix. We can determine the transition probabilities between the states and thus the associated fixation probabilities in the population. This will allow us to consider what the best linking strategies might be. Will the same strategies be effective in the evolutionary context, where a variety of population compositions must be considered, as in the original game?

Chapter 9

Appendix

9.1 The biggest minimal set

We have worked out the minimal set for $n = 7, 6, 5, 4, 3, 2, 1$ using Matlab, see Section 9.2. Considering the data we obtained from the program, we created Table 9.1 summarising the outcome of the program with the biggest size minimal set for each n , the score and the deficit.

n	score	<i>deficit</i>	<i>M.Ssize</i>	sequence
2	1	1	2	1,0
3	2	2	3	2,2,0
3	2	2	3	2,0,0
4	3	3	7	3,2,0,0
4	3	3	7	3,3,1,0
4	4	4	7	3,3,0,0
4	2	2	7	3,2,1,0
5	4	4	20	4,3,1,0,0
5	4	4	20	4,4,3,1,0
6	7	7	84	5,5,4,1,0,0
7	9	9	262	6,6,5,1,1,0,0
7	9	9	262	6,6,5,5,1,0,0

Table 9.1: Sequences with the biggest minimal set size where n is the number of individuals, M.S size: refers to the minimal set size

We noted that the score and the deficit are equal for all largest minimal sets

where $n \leq 7$ in the steps. It would be interesting to see if the score and deficit are always the same for the largest minimal set, although we have not been able to find a method to handle this problem. The results highlight the importance of this.

Example 9.1.1. Suppose we have the following sequences, see [4.1.2](#).

- 1) 5, 5, 5, 0, 0, 0
- 2) 5, 5, 4, 1, 0, 0

The first sequence has the top three numbers all the same and the second sequence with the top two numbers the same. When we are looking for the minimal set we can reorder the first three top numbers for sequences (1) and the first two top numbers for sequences (2) in any position as any order will give the same result, so the reordering in the first sequence will give more sequences than the second sequence which means more elements to be omitted from the eventual count of sequence. That is why in fact $J(\min)$ is larger for the second sequence even though the score is smaller. From the minimal set $K(\min)$ of the first sequence contains more elements than that for the second one even though, $J(\min)$ is still larger in the second sequence. However, a sequence with a larger score will tend to have larger minimal sets ($J(\min)$ and $K(\min)$).

- Out of the all or nothing sequences, the biggest minimal set contains $n-1$ s and zeros only, and the biggest scores are those given by the all or nothing system as shown in [\[2\]](#). Perhaps $K(\min)$ might be bigger for the all or nothing sequences.
- What might the patterns for for the target with the largest minimal set look like? It may further decrease as follows: $n - 1, n - 1, n - 2, n - 2, n -$

$3, n - 3, \dots, 2, 2, 1, 1, 0, 0$ (where there will be a single large drop from the large to small numbers at the appropriate point) if we considered a long enough sequence, as this will have a large score and less numbers of ties. As multiplicity increases e.g. three values of the same number, the number of ties increases a lot.

9.2 MATLAB program:

Matlab has been used widely in many fields to find the results quickly and efficiently, see [79] and [77]. The Matlab program we present below is to find the minimal set, the score and the deficit for a target sequence, though it is impractical for large target sequences as the time it takes to perform the required results increases very quickly with n . The program solved the problems of finding the minimal set for n with values ranging from $n = 2$ to $n = 7$. The elapsed time for $n = 6$ was 27 minutes and it took approximately 4 days for the program to perform the calculations for $n = 7$, so we stopped there. If we attempted to find the result for $n = 8$ it would take an extremely long time in order for the program to perform the necessary calculations. We started our work by creating all the required functions, see Figures 9.1, 9.2, 9.3, 9.4, 9.5, 9.6, 9.7 below. The next step was to call these functions to the main program, see Figure 9.7 where the full task of finding the desired minimal set will be completed here. The following is a brief description of each part of the Matlab program:

1. New target score, see Figure 9.1

We used the Havel-Hakimi algorithm to calculate the score of a target sequence as follows:

- Order the given sequence in decreasing order.

- Delete the first (greatest) element k and subtract (1) from the next k elements.
- Replace any negative value (-1 is the only possible negative value) and turn it to zero.
- Finally, cumulative the number of negative values (all -1) and assign them to the score as the final result.

2. Is graphic a, see Figure 9.2:

In this program we have used the Havel-Hakimi algorithm to check each sequence if it is graphic or not. First, we found the score using the program function in 9.1 by calling it to this program. Next, we apply the H-H algorithm to get all of the graphical sequences of the target sequence t and these are the ones which achieve the score.

3. myperm, see Figure 9.3:

Here we are taking into consideration if there (is/are) (a tie/ties) in the target sequence.

The program gets the graphical sequence S and the target t as input. So, it will define what is the position of the (tie/ties) and then go to the graphical sequences and apply the permutation with respect to the position of the ties, then choose the one with decreasing order.

4. Remove ties, see Figure 9.4

We aim here to find the graphical sequences after removing the ties. Suppose that the target sequence has a tie or more, in this case we can not take all the permutations of the graphical sequence.

Example: The sequence 2, 2, 1 is a non graphical sequence. If we think about its permutations we will write $\{1, 2, 2\}, \{2, 1, 2\}, \{2, 2, 1\}$ which is correct

for general sequences without consideration of the ties. However, with a tie sequence we have to take all the permutation of graphical sequence (G) in the minimal set taking into consideration the ties order. Here we have a tie 2, 2 which has the first and the second position in the sequence, thus for the elements of the minimal set which are $\{1, 2, 1\}$, $\{2, 1, 1\}$, $\{2, 2, 2\}$, the only allowed permutations will be between the elements in the first and second position of a sequence of the minimal set. Now we will remove the tie by taking the permutation in the first and second position of each sequence (we must keep all the resulting sequences in decreasing order with respect to the tie position only) and after removing the tie we will get the following final sequences $\{2, 1, 1\}$, $\{2, 2, 2\}$.

5. Compute deficit, see Figure 9.5: To find the deficit of a target sequence :

First we find the vector v with v_i defined as: the number of elements where $t_j \geq i$.

The second step is to find the Durfee number $\lambda = \max_i; t_i \geq i$.

Now, after finding the vector v and the Durfee number we can proceed to find the deficit $f_k = \sum_{i=1}^{\mu} (t_i + 1 - v)$ with $f_0 = 0$ where the summation here runs from 1 to the Durfee number. The deficit will be the f_k with the greatest value.

6. Create Decreasing Sequences n, see Figure 9.6

This program computes the deficit, score, and graphical sequences for a given target sequence $n = 1, 2, 3, \dots, 7$. Similar results cannot be calculated for $n = 8$ in this program because [there will be a large number of permutations that the computer must consider and arrange in a matrix before finding the minimal set, and here we did not have enough memory for those kinds of numbers], as we mentioned earlier.

7. Full program, see Figure 9.7

After finding all of the required functions in terms of program completion, we can create a full program that computes the deficit, the score of a target sequence, and finds the full set of graphical sequences taking into account the following cases:

- (a) When the sequence includes a tie/ ties.
- (b) After removing the ties from the sequence it will show another set of graphical sequences, leaving only those in decreasing order with respect to the tie.

We also calculated the elapsed time to find the result.

Listing 9.1: Finding the target score of a target sequence

```
1 function T= newtargetScore(target)
2
3 n = numel(target);
4 T = 0;
5
6 for t=1:n
7     q = target(1);
8     target(1) = [];
9
10    if (numel(target)>0)
11        target(1:q) = target(1:q) - 1;
12
13        T = T+ sum( target(1:q) < 0);
14        ptarget(1:q)=abs(target(1:q));
15        target(1:q)=1/2*(ptarget(1:q)+target(1:q));%every ...
           positive number stay positive and every negative ...
```

```

        number become zero
16
        target(1:q) = abs(target(1:q));
18
        target = sort(target, 'desc');
20     end
21 end

```

Listing 9.2: Finding the minimal set of a target sequence (t)

```

1 function R = isgraphica(a,target)
2     T = sum(abs(target-a)); % sum of absolute differences
3     n = numel(a); %compute the number of element n
4     tScore=newtargetScore(target);
5     if (T==tScore) % perform HH algorithm
6         S = a;
7         while ( ~all(S==0) && (numel(S)>0) )
8             S = sort(S, 'desc');
9             if ( any(S < 0) && any(S > n) )
10                R=0; %fault
11                return;
12            end
13            k = S(1); %take first element
14            S(1) = []; %remove first element
15
16            S(1:k) = S(1:k) - 1; %decrease the first k element
17
18        end
19        S = sort(S, 'desc');
20
21        if (all(S==0) && ~isempty(S))

```

```

22         R = 1; %graphic
23         return;
24     else
25         R = 0; %non graphic
26         return ;
27     end
28
29 else
30     R = 0;%non graphic
31 end

```

Listing 9.3: Define what are the ties position in the target sequence

```

1 function P = myPerm3(S,t)
2 un = unique(t);
3 P = [];
4 for i = 1:numel(un)
5     mask = un(i) == t;
6     idx = find(mask);
7
8     if numel(idx)>1
9         p = perms(idx);
10
11         for j = 1:size(p,1)
12             s = S;
13             s(idx) = s(p(j,:));
14
15             non_descr = true;
16
17             k = 2;
18             while (non_descr==true) && (k<numel(s))

```

```

19         if ( s(k-1) < s(k))
20             non_descr=false;
21         end
22         k=k+1;
23     end
24
25     if (non_descr==true)
26         P = [P;s];
27     end
28 end
29 end
30 end

```

Listing 9.4: Finding the graphical sequences after removing the ties

```

1 function A = removeTies(G,t)
2
3 A = [];
4
5 n = size(G,1); %total numbers of graphic sequences
6 num = unique(t); %the numbers (without rep) in the target
7
8 for i=1:n
9     incr = false;
10    for k=1:numel(num) %number of unique digits
11        mask = t == num(k); %finds the positions of num(k) in t
12
13        if (sum(mask)>1) % checking if the number is repeating ...
14            at least twice
15            seq = G(i,mask); %I get the numbers in the graphic ...
16            sequence in correspondance of the ties

```

```

15         for q = 1:numel(seq)-1 %checking if the seq is in ...
           non-increasing order
16             if (~incr)
17                 if (seq(q)<seq(q+1))
18                     incr=true;
19                 end
20             end
21         end
22     end
23 end
24
25     if (~incr)
26         A = vertcat(A,G(i,:));
27     end
28 end
29
30
31 end

```

Listing 9.5: Finding the deficit of a target score(t)

```

1 function [d,s] = computeDeficit(t)
2 n = numel(t);
3 % compute v vector
4 v=zeros(n,1);
5 lambda=0;
6 for i=1:n
7     v(i)=sum(t>=i);
8 end
9 j=1;
10 while (lambda==0) && (j<=n)

```

```

11     if (t(j)<j)
12         lambda=j-1;
13     end
14     j=j+1;
15 end
16 % compute f vector
17 f=zeros(lambda+1,1);
18 for k=1:lambda
19     f(k)=t(k)+1-v(k);
20     if(k>1)
21         f(k)=f(k)+f(k-1);
22     end
23 end
24 d = max(f);
25 num1=mod(sum(t),2);
26 num2=mod(d,2);
27 if (num1==num2)
28     s = d;
29 else
30     s = d+1;
31 end

```

Listing 9.6: Calculating the score, deficit, and minimal set for a target sequence has length N

```
1 n = 7;
2 [T,score,deficit] = createDescSequences(n);
3
4 h = fopen(['results_n=' int2str(n) '.txt'],'w'); % file handle
5
6 for i=1:numel(score)
7     t = T(i,:);
8     fprintf(h,'Current target sequence: ');
9
10    fprintf(h,'%d ',t); %each element of t will
11    %be written in the file as '# '
12
13    fprintf(h,'\n');
14
15    fprintf(h,'Sequence score: %d\n',score(i));
16    fprintf(h,'Sequence deficit: %d\n',deficit(i));
17
18    minimalSet = graphTheory(t);
19
20    fprintf(h,'Minimal set size: %d\n\n',size(minimalSet,1));
21
22    for j = 1:size(minimalSet,1) %write in the file each row
23        % of the minimal set independently
24
25        fprintf(h,'%d ',minimalSet(j,:));
26        fprintf(h,'\n');
27    end
28    fprintf(h,'\n-----\n\n');
29 end
```

```
30
31 fclose(h);
```

Listing 9.7: A full program finding all previous results of target sequence considering a tie sequence

```
1  t= [ 4 3 2 1 0];
2  tic
3  n=numel(t);
4  [defecit,score]=computeDefecit(t);
5  fprintf('defecit: %d\n',defecit);
6  fprintf('score .: %d\n',score);
7  %A=nmultichoosek(0:n-1,n);
8  %A=A(:,end:-1:1);
9  m = 0;%min(t);
10 M = n-1;%max(t);
11 num = 0:n-1;
12 A = zeros( numel(num)^n ,n);
13 A(1,:) = m;
14 scores = zeros(size(A,1),1);
15 scores(1) = newtargetScore(A(1,:));
16 for i = 2:size(A,1)
17     reminder=true;
18     for j=n:-1:1
19         if (reminder)
20             curr = A(i-1,j);
21             mask = num==curr;
22             next = circshift(mask,1);
23             A(i:end,j) = num(next);
24             if (next(1)==1)
25                 reminder=true;
```

```

26         else
27             reminder=false;
28         end
29     end
30 end
31 scores (i) = computeScore(A(i,:));
32 end
33
34 %%
35
36 %%
37
38 R=zeros(size(A,1),1);
39 for j=1:size(A,1)
40     seq=A(j,:);
41     R(j)=isgraphica(seq,t);
42 end
43     idx=find(R);
44     graphicsequences=A(idx,:);
45
46 %% remove wrt the ties
47
48 disp('Full set')
49 graphicsequences
50 disp('After removing ties')
51
52 graphicsequences = removeTies(graphicsequences,t)
53
54 %% all possible permutations
55 for i = 1:numel(idx)
56     %fprintf('%d out of %d\n',i,numel(idx));
57     permA = myPerm3(graphicsequences(i,:),t);

```

```
58
59     for j = 1:size(permA,1)
60         seq = permA(j,:);
61         r = isgraphica(seq,t);
62
63         if (r==1)
64             graphicsequences = [graphicsequences;seq];
65         end
66     end
67 end
68 graphicsequences=unique(graphicsequences,'rows','stable')
69 toc
```

9.3 The game theoretical model, sequence 1, 1, 1

Following from Section 5.3.2, we use in the calculations of this section the following alternative representation to minimise the number of subscripts used: $x = p_1, y = p_2, z = p_3, f = p_4, g = p_5, h = p_6$. Substituting these values in the row vector $\pi * p = \pi$ where π is the stationary distribution, and replacing the last (redundant) row by the condition that the row vector entries (probabilities) are non-negative and sum to 1, yields

$$\begin{bmatrix} -1 & -y+1 & 0 & 0 & 0 & h \\ x & -1 & -z+1 & 0 & 0 & 0 \\ 0 & y & -1 & -f+1 & 0 & 0 \\ 0 & 0 & z & -1 & -g+1 & 0 \\ 0 & 0 & 0 & f & -1 & -h+1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} * \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Solving this set of equations by the inverse matrix method we obtained the values of the stationary distribution $\pi_1, \pi_2, \pi_3, \pi_4, \pi_5$ and π_6 in term of x, y, z, f, g and h given below: Denoting the common denominator term for all components of the stationary distribution by

$$\begin{aligned} A = & (4f - 4gf + 4g - 2hf + 2ghf - 4gh + 4h - 2fx + 2fgx - 2gx + 2hfx - 2ghxf \\ & + 2ghx - 4hx + 4x - 2yf + 2gyf - 2gy + 2ghy - 2hy + 2fxy - 2fgxy + 2gxy - 2ghxy \\ & + 2hxy - 4xy + 4y - 4fz + 2fgz - 2gz + 2fhz - 2fghz + 2ghz - 2hz + 2fzx - 2fhxz \\ & + 2hxz - 2xz + 2fyz - 2fgyz + 2gyz - 2ghyz + 2hyz - 2fxyz - 2hxyz + 2xyz - 4yz + 4z - 6), \end{aligned}$$

the stationary distribution terms are:

$$\begin{aligned}
\pi_1 &= \frac{f-fg+g-gh-fy+fgy-gy+ghy+y-fz+fgz-gz-fghz+ghz+fyz-fgyz+gyz-ghyz-yz+z-1}{A}, \\
\pi_2 &= \frac{f-fg+g-fh+fgh-gh+h+fhx-fghx-hx-fz+fgz-gz+fhz-fghz+ghz-hz-fhxz+hxz+z-1}{A}, \\
\pi_3 &= \frac{f-fg+g-fh+fgh-gh+h-fx+fgx-gx+fhx-fghx+ghx-hx+x+fx-y-fgy+gxy-ghxy-xy-1}{A}, \\
\pi_4 &= \frac{g-gh+h-gx+ghx-hx+x-gy+ghy-hy+gxy-ghxy+hxy-xy+y+gyz-ghyz+hyz-hxyz-yz-1}{A}, \\
\pi_5 &= \frac{h-hx+x-hy+hxy-xy+y-fz+fhz-hz+fxz-fhxz+hxz-xz+hyz-fxyz-hxyz+xyz-yz+z-1}{A}, \\
\pi_6 &= \frac{f-fg-fx+fgx+x-fy+fgy+fyx-fgyx-xy+y-fz+fxz-xz+fyz-fgyz-fxyz+xyz-yz+z-1}{A}.
\end{aligned}$$

For the general game the payoff of the individuals is defined in [10] as the negative of their expected long term deviation, thus the payoff to individual i is:

$$R_i(U_X) = - \sum_x \epsilon_i(x) * \pi(x)$$

where $\epsilon_i(x)$ is the deviation of i in state X , and $\pi(x)$ is the stationary distribution over X .

We then have the following payoffs of the individuals in the sequence 1, 1, 1: the payoff of v_1 is $-(\pi_1 + \pi_4)$, the payoff of v_2 is $-(\pi_2 + \pi_5)$ and the payoff of v_3 is $-(\pi_3 + \pi_6)$.

We then found the derivatives of mean deviation, the negative of the payoffs, of each of the three individuals, each the sum of two π_i terms, at the two states where they are in deficit with respect to the variable that govern the transition at the corresponding state. Starting by differentiating $(\pi_1 + \pi_4)$ with respect to x we get:

$$\frac{d(\pi_1+\pi_4)(x,y,z,f,g,h)}{dx} = -2f^2g^2h^2z - 4f^2g^2hyz - 2f^2g^2y^2z - 2f^2gh^2z^2 - 4f^2ghyz^2 -$$

$2f^2gy^2z^2 - 4fg^2h^2yz - 4fg^2hy^2z - 4fgh^2yz^2 - 4fghy^2z^2 - 2g^2h^2y^2z - 2gh^2y^2z^2 +$
 $2f^2g^2hy + 4f^2g^2hz + 2f^2g^2y^2 + 4f^2g^2yz + 2f^2gh^2z + 8f^2ghyz + 4f^2ghz^2 + 6f^2gy^2z +$
 $4f^2gyz^2 + 2f^2hyz^2 + 2f^2y^2z^2 + 2fg^2h^2y + 2fg^2h^2z + 4fg^2hy^2 + 10fg^2hyz + 4fg^2y^2z +$
 $6fgh^2yz + 4fgh^2z^2 + 10fghy^2z + 14fghyz^2 + 6fgy^2z^2 + 2fhy^2z^2 + 2g^2h^2y^2 + 2g^2h^2yz +$
 $4g^2hy^2z + 4gh^2y^2z + 4gh^2yz^2 + 6ghy^2z^2 - 2f^2g^2h - 4f^2g^2h - 4f^2g^2y - 2f^2g^2z -$
 $4f^2ghy - 8f^2ghz - 4f^2gy^2 - 12f^2gyz - 2f^2gz^2 - 4f^2hyz - 2f^2hz^2 - 4f^2y^2z - 4f^2yz^2 -$
 $2fg^2h^2 - 8fg^2hy - 4fg^2hz - 4fg^2y^2 - 6fg^2yz - 2fgh^2y - 6fgh^2z - 6fghy^2 - 28fghyz -$
 $8fghz^2 - 16fgy^2z - 10fgyz^2 - 4fhy^2z - 6fhyz^2 - 6fy^2z^2 - 2g^2h^2y - 4g^2hy^2 - 4g^2hyz -$
 $2g^2y^2z - 2gh^2yz - 2gh^2z^2 - 14ghy^2z - 10ghyz^2 - 4gy^2z^2 - 2hy^2z^2 + 2f^2g^2 + 4f^2gh +$
 $8f^2gy + 6f^2gz + 2f^2hy + 4f^2hz + 2f^2y^2 + 8f^2yz + 2f^2z^2 + 4fg^2h + 6fg^2y + 2fg^2z +$
 $2fgh^2 + 14fghy + 16fghz + 10fgy^2 + 26fgyz + 4fgz^2 + 2fhy^2 + 12fhyz + 4fhz^2 +$
 $12fy^2z + 10fyz^2 + 4g^2hy + 2g^2y^2 + 2g^2yz + 4gh^2y + 4gh^2z + 8ghy^2 + 22ghyz + 4ghz^2 +$
 $10gy^2z + 6gyz^2 + 4hy^2z + 4hyz^2 + 4y^2z^2 - 4f^2g - 2f^2h - 4f^2y - 4f^2z - 2fg^2 - 8fgh -$
 $16fgy - 10fgz - 6fhy - 8fhz - 6fy^2 - 20fyz - 4fz^2 - 2g^2y - 2gh^2 - 12ghy - 8ghz -$
 $6gy^2 - 14gyz - 2gz^2 - 2hy^2 - 8hyz - 2hz^2 - 8y^2z - 6yz^2 + 2f^2 + 6fg + 4fh + 10fy +$
 $8fz + 4gh + 8gy + 4gz + 4hy + 4hz + 4y^2 + 12yz + 2z^2 - 4f - 2g - 2h - 6y - 4z + 2$. In the
 same manner we found $\frac{d(\pi_1 + \pi_4)(x,y,z,f,g,h)}{df}$. We notice that finding the derivatives is
 a complicated procedure which produces a long formula. Therefore we used the
 symmetry of this dynamical system which made the calculations easier. Due to
 the symmetry in the dynamical system we will have:

$$\begin{aligned}
 \frac{d(\pi_1 + \pi_4)(x, y, z, f, g, h)}{dx} &= \frac{d(\pi_2 + \pi_5)(h, x, y, z, f, g)}{dy} = \frac{d(\pi_3 + \pi_6)(y, z, f, g, h, x)}{dh} \\
 &= \frac{d(\pi_1 + \pi_4)(x, y, z, f, g, h)}{df} = \frac{d(\pi_2 + \pi_5)(h, x, y, z, f, g)}{dg} = \frac{d(\pi_3 + \pi_6)(y, z, f, g, h, x)}{dz}.
 \end{aligned}$$

We thus have all of the terms required to find Nash equilibria following Equation

5.16.

9.4 Python

See Figure 9.8; the goal of this program is to find a selected individual deviation and a group overall deviation while playing a game in which individuals are represented as nodes in a target sequence, as discussed in Section 6, Table 7.2, Table 7.3 and Table 7.4. We used five different strategies defined earlier in Section 7.3, which are LT, LR, LRA, GR, GR; the entire group of individuals will play one strategy, with the exception of one selected individual (though a larger number could be chosen) who will use each of the other strategies. The user will select the length of the sequence to be investigated, as well as the node that will alternate between the strategies and the group strategy. In our game, we ran simulations of 100000 trails. Each node is assigned a numbered position, so even if the sequence contains ties, the program will treat them as distinct nodes based on their position.

Listing 9.8: Game simulations of a general sequence

```
1 import random
2 import copy
3
4 showEveryLink = False
5 count = 0
6 strategyBool = [[False, [], "Least Target"], [False, [], "Least ...
    Remaining"], [False, [], "Least Ratio"], [False, [], "Greatest ...
    Target"], [False, [], "Greatest Remaining"]]
7
8 def targetFunction(targetList):
9     flag = True
10    possibleConnection = []
11    connectionList = []
12
```

```

13     for i in range(len(targetList)):
14         connectionList.append([])
15
16     for i in range(len(targetList)):
17         possibleConnection.append(targetList[:])
18     for i in range(len(possibleConnection)):
19         possibleConnection[i].remove(i+1)
20
21
22     possibleChoice = targetList[:]
23
24     while (flag):
25
26         # Choose a target
27         randomInt = random.randint(0, len(possibleChoice) - 1)
28         targetChoice = possibleChoice[randomInt]
29
30
31         #Chooses a second target to form a link with
32
33         if (strategyBool[0][0] == True ):
34             if (strategyBool[0][1].count(targetChoice) == 1):
35                 secondChoice = leastTarget(targetChoice, ...
36                                         possibleConnection)
37
38         if (strategyBool[1][0] == True ):
39             if (strategyBool[1][1].count(targetChoice) == 1):
40                 secondChoice = ...
41                                     leastRemainingTarget(targetChoice, ...
42                                                         possibleConnection, connectionList)
43
44         if (strategyBool[2][0] == True ):

```

```

42     if (strategyBool[2][1].count(targetChoice) == 1):
43         secondChoice = ...
44         leastRatioTarget(targetChoice,possibleConnection,
45                             connectionList)
46
47     if (strategyBool[3][0] == True ):
48         if (strategyBool[3][1].count(targetChoice) == 1):
49             secondChoice = ...
50             greatTarget(targetChoice,possibleConnection)
51
52     if (strategyBool[4][0] == True ):
53         if (strategyBool[4][1].count(targetChoice) == 1):
54             secondChoice = greatRemain(targetChoice, ...
55             possibleConnection,
56             connectionList)
57
58     #Form the link
59     connectionList[targetChoice-1].append(secondChoice)
60     connectionList[secondChoice-1].append(targetChoice)
61     possibleConnection[targetChoice - 1].remove(secondChoice)
62     possibleConnection[secondChoice - 1].remove(targetChoice)
63
64     if (len(connectionList[secondChoice-1]) == secondChoice):
65         for i in possibleConnection[secondChoice-1]:
66             possibleConnection[i-1].remove(secondChoice)
67         possibleConnection[secondChoice-1] = []
68
69     if (len(connectionList[targetChoice-1]) == targetChoice):
70         for i in possibleConnection[targetChoice-1]:
71             possibleConnection[i-1].remove(targetChoice)
72         possibleConnection[targetChoice-1] = []

```

```

71
72     newPossibleChoice = copy.deepcopy(possibleChoice)
73     for i in newPossibleChoice:
74         if possibleConnection[i-1] == []:
75             possibleChoice.remove(i)
76             # _____
77             # Code to show each individual connection
78             if (showEveryLink == True):
79                 print("-----")
80                 print("Current connections: ")
81                 print(connectionList)
82                 print("Node picked: ")
83                 print(targetChoice)
84                 print("Node to be linked with: ")
85                 print(secondChoice)
86                 print("-----")
87
88             if (len(possibleChoice)==0):
89                 flag = False
90
91     degreeList =[]
92     scoreList = []
93     for i in range(len(connectionList)):
94         degreeList.append(len(connectionList[i]))
95         scoreList.append(i+1-degreeList[i])
96
97     return connectionList,degreeList, scoreList
98
99 def leastTarget(target, possibleConnection):
100     return possibleConnection[target-1][0]
101
102 def greatTarget(target, possibleConnection):

```

```

103     return possibleConnection[target-1][-1]
104
105 def greatRemain(targetChoice, possibleConnection, connectionList):
106     # Chooses a second target to form a link with
107
108     maximum = possibleConnection[targetChoice - 1][0] -
109     len(connectionList[possibleConnection[targetChoice - 1][0] ...
110         - 1])
111     maxIndex = 0
112
113     for i in range(len(possibleConnection[targetChoice - 1])):
114         if (possibleConnection[targetChoice - 1][i] -
115             len(connectionList[possibleConnection[targetChoice ...
116                 - 1][i] - 1]) > maximum):
117             maxIndex = i
118             maximum = possibleConnection[targetChoice - 1][i] ...
119                 - len(
120                 connectionList[possibleConnection[targetChoice ...
121                     - 1][i] - 1])
122
123     return possibleConnection[targetChoice - 1][maxIndex]
124
125 def leastRemainingTarget(targetChoice, possibleConnection, ...
126     connectionList):
127     # Chooses a second target to form a link with
128
129     minimum = possibleConnection[targetChoice - 1][0] -
130     len(connectionList[possibleConnection[targetChoice - 1][0] ...
131         - 1])
132     minIndex = 0
133
134     for i in range(len(possibleConnection[targetChoice - 1])):

```

```

129     if (possibleConnection[targetChoice - 1][i] -
130         len(connectionList[possibleConnection[targetChoice ...
            - 1][i] - 1]) < minimum):
131         minIndex = i
132         minimum = possibleConnection[targetChoice - 1][i] ...
            - len(
133             connectionList[possibleConnection[targetChoice ...
                - 1][i] - 1])
134
135     return possibleConnection[targetChoice - 1][minIndex]
136
137 def
138 leastRatioTarget(targetChoice, possibleConnection, connectionList):
139     # Chooses a second target to form a link with
140
141     minTarget = possibleConnection[targetChoice - 1][0]
142     minimum = (minTarget - len(connectionList[minTarget - 1])) ...
        / len(possibleConnection[minTarget - 1])
143     minIndex = 0
144
145     for i in range(len(possibleConnection[targetChoice - 1])):
146         minTarget = possibleConnection[targetChoice - 1][i]
147         if ((minTarget - len(connectionList[minTarget - ...
            1]))/len(possibleConnection[minTarget - 1]) < minimum):
148             minIndex = i
149             minimum = (minTarget - ...
                len(connectionList[minTarget - ...
                    1]))/len(possibleConnection[minTarget - 1])
150
151     return possibleConnection[targetChoice - 1][minIndex]
152
153 print("Enter list size:")

```

```

154 arraySize = int(input())
155
156 print("Enter number of trials:")
157 numberOfTrials = int(input())
158
159 print("Use Least Target Method? 0 for no, 1 for yes")
160 dummy = int(input())
161 if (dummy!=0):
162     count += 1
163     strategyBool[0][0]=True
164
165 print("Use Least Remaining Method? 0 for no, 1 for yes")
166 dummy = int(input())
167 if (dummy!=0):
168     count += 1
169     strategyBool[1][0]=True
170
171 print("Use Least Ratio Method? 0 for no, 1 for yes")
172 dummy = int(input())
173 if (dummy!=0):
174     count += 1
175     strategyBool[2][0]=True
176
177 print("Use Great Target Method? 0 for no, 1 for yes")
178 dummy = int(input())
179 if (dummy != 0):
180     count += 1
181     strategyBool[3][0] = True
182
183 print("Use Great Remain Method? 0 for no, 1 for yes")
184 dummy = int(input())
185 if (dummy != 0):

```

```

186     count += 1
187     strategyBool[4][0] = True
188
189 if count==0:
190     exit()
191
192 print("Show graphs? 0 for no, 1 for yes")
193 graphTrue = int(input())
194
195 print("Show every link? 0 for no, 1 for yes")
196 dummy = int(input())
197 if (dummy!=0):
198     showEveryLink=True
199
200 #create target list
201 targetList = []
202 for i in range(arraySize):
203     targetList.append((i+1))
204
205 strat = None
206 if count == 1:
207     for i in range(len(strategyBool)):
208         if strategyBool[i][0] == True:
209             strat =strategyBool[i]
210             strat[1] = targetList
211             break
212 else:
213     targetListCopy = targetList[:]
214     while (count-1 >0):
215         for i in range(len(strategyBool)):
216             if strategyBool[i][0] == True:
217                 strat = strategyBool[i]

```

```

218         print("Write which nodes follows " + ...
                strategyBool[i][2] +
219         " strategy, separate each node with a space, ...
                write 'skip' to skip
                writing the nodes for this strategy:")
220
221     dummy = input()
222     if dummy != "skip":
223         strategyBool[i][1] = [int(item) for item ...
                                in dummy.split()]
224         for entry in strategyBool[i][1]:
225             targetListCopy.remove(entry)
226             count -= 1
227     for i in range(len(strategyBool)):
228         if strategyBool[i][0] == True and ...
                len(strategyBool[i][1]) == 0:
229             strat = strategyBool[i]
230             strat[1] = targetListCopy
231             break
232
233     print("\nNOW SIMULATING:\n")
234
235     leastT, averageDegree, scoreLeastT = targetFunction(targetList)
236     if(graphTrue == 1):
237         print(leastT)
238         print("-----")
239     for i in range(numberOfTrials-1):
240         leastT, degreeLeastT, scoreLeastT = targetFunction(targetList)
241         averageDegree = [x + y for x, y in zip(averageDegree, ...
                                                degreeLeastT)]
242     if (graphTrue == 1):
243         print(leastT)
244         print("-----")

```

```

245
246 averageDegree = [x/numberOfTrials for x in averageDegree]
247 averageScore = [y - x for x, y in zip(averageDegree, targetList)]
248
249 print("Average Score of "+ str(numberOfTrials) + " graphs are:")
250 print(averageScore)
251 print("Average SUM of "+ str(numberOfTrials) + " graphs are:")
252 print(sum(averageScore))
253
254 file_object = open('results.txt', 'a')
255 for x in strategyBool:
256     if x[0] == True:
257         file_object.write(x[2] + ": ")
258         file_object.write(str(x[1]))
259         file_object.write('\n')
260 file_object.write("Average Score of "+ str(numberOfTrials) + " ...
graphs are:")
261 file_object.write(str(averageScore))
262 file_object.write('\n')
263 file_object.write("Average SUM of "+ str(numberOfTrials) + " ...
graphs are:")
264 file_object.write(str(sum(averageScore)))
265 file_object.write('\n')
266 file_object.write("Number of trials:" + str(numberOfTrials))
267 file_object.write('\n')
268 file_object.write("-----\n")
269
270 file_object.close()
271
272 input()

```

Bibliography

- [1] William Aiello, Fan Chung, and Linyuan Lu. Random evolution in massive graphs. In *Handbook of massive data sets*, pages 97–122. Springer, 2002.
- [2] Raneem Aizouk and Mark Broom. Modelling conflicting individual preference: target sequences and graph realization. *Discrete & Continuous Dynamical Systems-B*, 2022.
- [3] Gregory F Albery, Tim H Clutton-Brock, Alison Morris, Sean Morris, Josephine M Pemberton, Daniel H Nussey, and Josh A Firth. Ageing red deer alter their spatial behaviour and become less social. *Nature Ecology & Evolution*, 6(8):1231–1238, 2022.
- [4] Benjamin Allen and Martin A Nowak. Games on graphs. *EMS surveys in mathematical sciences*, 1(1):113–151, 2014.
- [5] Norman Biggs, E Keith Lloyd, and Robin J Wilson. *Graph Theory, 1736-1936*. Oxford University Press, 1986.
- [6] Béla Bollobás and Bela Bollobas. *Modern graph theory*, volume 184. Springer Science & Business Media, 1998.
- [7] John Adrian Bondy and Uppaluri Siva Ramachandra Murty. *Graph theory with applications*, volume 290. Macmillan London, 1976.

- [8] Mark Broom and Chris Cannings. A dynamic network population model with strategic link formation governed by individual preferences. *Journal of theoretical biology*, 335:160–168, 2013.
- [9] Mark Broom and Chris Cannings. Graphic deviation. *Discrete Mathematics*, 338(5):701–711, 2015.
- [10] Mark Broom and Chris Cannings. Game theoretical modelling of a dynamically evolving network i: general target sequences. *Journal of Dynamics and Games*, 4(4):285–318, 2017.
- [11] Mark Broom and Jan Rychtár. *Game-theoretical models in biology*. CRC Press, 2013.
- [12] Chris Cannings and Mark Broom. Game theoretical modelling of a dynamically evolving network : target sequences of score 1. *Journal of Dynamics & Games*, 2019.
- [13] Kerryn D Carter, Rachel Brand, John K Carter, Bryan Shorrocks, and Anne W Goldizen. Social networks, long-term associations and age-related sociability of wild giraffes. *Animal Behaviour*, 86(5):901–910, 2013.
- [14] Susan T Charles and Laura L Carstensen. Social and emotional aging. *Annual review of psychology*, 61:383–409, 2010.
- [15] Wai-Kai Chen. *Applied graph theory*, volume 13. Elsevier, 2012.
- [16] Richard C Connor, Michael R Heithaus, and Lynne M Barre. Superalliance of bottlenose dolphins. *Nature*, 397(6720):571–572, 1999.

- [17] J. Theodore Cox, Richard Durrett, and Edwin A. Perkins. *Voter model perturbations and reaction diffusion equations*. Number 349 in Astérisque. Société mathématique de France, 2013.
- [18] Ross Cressman. *The stability concept of evolutionary game theory: a dynamic approach*, volume 94. Springer Science & Business Media, 2013.
- [19] Charles Darwin. *On the origin of species, 1859*. Routledge, 2004.
- [20] Mary-Ann Dimand and Robert W Dimand. *The history of game theory, volume 1: from the beginnings to 1945*. Routledge, 1996.
- [21] Richard Durrett and Simon Levin. The importance of being discrete (and spatial). *Theoretical population biology*, 46(3):363–394, 1994.
- [22] David Easley and Jon Kleinberg. Networks, crowds, and markets: reasoning about a highly connected world. *Significance*, 9(1):43–44, 2012.
- [23] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.
- [24] Robert Fagen. Animal play games of angels, biology, and Brian. *The future of play theory*, pages 23–44, 1995.
- [25] Dean P Foster and H Peyton Young. A correction to the paper stochastic evolutionary game dynamics. *Theoretical Population Biology*, 51(1):77, 1997.
- [26] Drew Fudenberg and David K Levine. *The theory of learning in games*, volume 2. MIT press, 1998.
- [27] David Gale and Lloyd S Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.

- [28] Michael T Ghiselin. The imaginary lamarck: a look at bogus ‘history’ in schoolbooks. *The Textbook Letter*, 5(4), 1994.
- [29] Steven Jay Gould. Another look at lamarck. *New Scientist*, 84(1175):38–40, 1979.
- [30] John Haigh. Game theory and evolution. *Advances in applied probability*, 7(1):8–11, 1975.
- [31] Seifollah Louis Hakimi. On the realizability of a set of integers as degrees of the vertices of a graph. *SIAM Journal Applied Mathematics*, 1962.
- [32] William D Hamilton. Extraordinary sex ratios. *Science*, 156(3774):477–488, 1967.
- [33] John Michael Harris, Jeffrey L Hirst, and Michael J Mossinghoff. *Combinatorics and graph theory*, volume 2. Springer, 2008.
- [34] John C Harsanyi. Oddness of the number of equilibrium points: a new proof. *International Journal of Game Theory*, 2(1):235–250, 1973.
- [35] Werner Hässelbarth. Die verzweigtheit von graphen. *Match*, 16:3–17, 1984.
- [36] Christoph Hauert and Michael Doebeli. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature*, 428(6983):643–646, 2004.
- [37] Václav Havel. A remark on the existence of finite graphs. *Casopis Pest. Mat.*, 80:477–480, 1955.
- [38] Josef Hofbauer and Karl Sigmund. *Evolutionary games and population dynamics*. Cambridge University press, 1998.

- [39] Josef Hofbauer and Karl Sigmund. Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, 40(4):479–519, 2003.
- [40] Matthew O Jackson. The stability and efficiency of economic and social networks. In *Advances in economic design*, pages 319–361. Springer, 2003.
- [41] Matthew O Jackson. *Social and economic networks*. Princeton University press, 2010.
- [42] Matthew O Jackson. An overview of social networks and economic applications. *Handbook of social economics*, 1:511–585, 2011.
- [43] Matthew O Jackson and Asher Wolinsky. A strategic model of social and economic networks. *Journal of economic theory*, 71(1):44–74, 1996.
- [44] Mohammed Ali Khan and Yeneng Sun. Non-cooperative games with many players. *Handbook of game theory with economic applications*, 3:1761–1808, 2002.
- [45] Timothy Killingback and Michael Doebeli. Spatial evolutionary game theory: Hawks and doves revisited. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 263(1374):1135–1144, 1996.
- [46] Tim Kohler. Evolution of the social contract. *Complexity*, 2(6):6–6, 1997.
- [47] Svenja B Kroeger, Daniel T Blumstein, and Julien GA Martin. How social behaviour and life-history traits change with age and in the year prior to death in female yellow-bellied marmots. *Philosophical Transactions of the Royal Society B*, 376(1823):20190745, 2021.

- [48] Frieder R Lang and Laura L Carstensen. Close emotional relationships in late life: further support for proactive aging in the social domain. *Psychology and aging*, 9(2):315, 1994.
- [49] Erez Lieberman, Christoph Hauert, and Martin A Nowak. Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316, 2005.
- [50] John Maynard Smith. Game theory and the evolution of fighting. *On evolution*, pages 8–28, 1972.
- [51] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge University Press, 1982.
- [52] John Maynard Smith and George R Price. The logic of animal conflict. *Nature*, 246(5427):15–18, 1973.
- [53] Russell Merris and Tom Roby. The lattice of threshold graphs. *J. Inequal. Pure Appl. Math*, 6(1):1–21, 2005.
- [54] Joshua Mitteldorf and David Sloan Wilson. Population viscosity and the evolution of altruism. *Journal of theoretical biology*, 204(4):481–496, 2000.
- [55] Patrick Alfred Pierce Moran. Random processes in genetics. In *Mathematical proceedings of the cambridge philosophical society*, volume 54, pages 60–71. Cambridge University Press, 1958.
- [56] Patrick Alfred Pierce Moran. The statistical processes of evolutionary theory. *The statistical processes of evolutionary theory.*, 1962.
- [57] Mayuko Nakamaru, Hiroyuki Matsuda, and Yoh Iwasa. The evolution of cooperation in a lattice-structured population. *Journal of Theoretical Biology*, 184(1):65–81, 1997.

- [58] John Nash. Non-cooperative games. *Annals of mathematics*, pages 286–295, 1951.
- [59] Jonathan Newton. Evolutionary game theory: A renaissance. *Games*, 9(2):31, 2018.
- [60] Ronald Noë. Biological markets: partner choice as the driving force behind the evolution of mutualisms. *Economics in nature: Social dilemmas, mate choice and biological markets*, pages 93–118, 2001.
- [61] Ronald Noë and Peter Hammerstein. Biological markets: supply and demand determine the effect of partner choice in cooperation, mutualism and mating. *Behavioral ecology and sociobiology*, 35(1):1–11, 1994.
- [62] Martin A Nowak, Sebastian Bonhoeffer, and Robert M May. More spatial games. *International Journal of Bifurcation and Chaos*, 4(01):33–56, 1994.
- [63] Martin A Nowak, Sebastian Bonhoeffer, and Robert M May. Spatial games and the maintenance of cooperation. *Proceedings of the National Academy of Sciences*, 91(11):4877–4881, 1994.
- [64] Martin A Nowak and Robert M May. Evolutionary games and spatial chaos. *Nature*, 359(6398):826–829, 1992.
- [65] Martin A Nowak and Robert M May. The spatial dilemmas of evolution. *International Journal of bifurcation and chaos*, 3(01):35–78, 1993.
- [66] Martin A Nowak and Karl Sigmund. Evolutionary dynamics of biological games. *science*, 303(5659):793–799, 2004.

- [67] Hisashi Ohtsuki, Christoph Hauert, Erez Lieberman, and Martin A Nowak. A simple rule for the evolution of cooperation on graphs and social networks. *Nature*, 441(7092):502–505, 2006.
- [68] Bezalel Peleg and Peter Sudhölter. *Introduction to the theory of cooperative games*, volume 34. Springer Science & Business Media, 2007.
- [69] Carlos Pérez Roca, José A Cuesta, and Angel Sánchez. Effect of spatial structure on the evolution of cooperation. 2009.
- [70] Bernard Philippe, Youcef Saad, and William J Stewart. Numerical methods in Markov chain modeling. *Operations research*, 40(6):1156–1179, 1992.
- [71] William H Press and Freeman J Dyson. Iterated prisoner’s dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences*, 109(26):10409–10413, 2012.
- [72] John M Roberts, Malcolm J Arth, and Robert R Bush. Games in culture. *American anthropologist*, 61(4):597–605, 1959.
- [73] Carlos P Roca, José A Cuesta, and Angel Sánchez. Evolutionary game theory: temporal and spatial effects beyond replicator dynamics. *Physics of life reviews*, 6(4):208–249, 2009.
- [74] Alexandra G Rosati, Lindsey Hagberg, Drew K Enigk, Emily Otali, Melissa Emery Thompson, Martin N Muller, Richard W Wrangham, and Zarin P Machanda. Social selectivity in aging wild chimpanzees. *Science*, 370(6515):473–476, 2020.
- [75] Ernst Ruch and Ivan Gutman. The branching extent of graphs. *J. Combin. Inform. System Sci*, 4(4):285–295, 1979.

- [76] Sebastian J Schreiber. Urn models, replicator processes, and random genetic drift. *SIAM Journal on Applied Mathematics*, 61(6):2148–2167, 2001.
- [77] Ronald W Shonkwiler and James Herod. *Mathematical biology: an introduction with Maple and Matlab*. Springer Science & Business Media, 2009.
- [78] Angela M Sibbald and Russell J Hooper. Sociability and the willingness of individual sheep to move away from their companions in order to graze. *Applied Animal Behaviour Science*, 86(1-2):51–62, 2004.
- [79] Ole Sigmund. A 99 line topology optimization code written in matlab. *Structural and multidisciplinary optimization*, 21(2):120–127, 2001.
- [80] Branislav L Slantchev. Game theory: Elements of basic models. *University of California, San*, 2009.
- [81] Neil J. A. Sloane. The on-line encyclopedia of integer sequences, 2003.
- [82] Cedric Austen Bardell Smith. Theory of games and economic behaviour. by John Von Neumann and Oskar Morgenstern. pp. xviii, 625. 66s. 6d. 1944.(princeton up; humphrey milford). *The Mathematical Gazette*, 29(285):131–133, 1945.
- [83] Daniel W Stroock. *An introduction to Markov processes*, volume 230. Springer Science & Business Media, 2013.
- [84] György Szabó and Gabor Fath. Evolutionary games on graphs. *Physics reports*, 446(4-6):97–216, 2007.
- [85] Peter D Taylor and Leo B Jonker. Evolutionary stable strategies and game dynamics. *Mathematical biosciences*, 40(1-2):145–156, 1978.
- [86] Richard J Trudeau. *Introduction to graph theory*. Courier Corporation, 2013.

- [87] Minus Van Baalen and David A Rand. The unit of selection in viscous populations and the evolution of altruism. *Journal of theoretical biology*, 193(4):631–648, 1998.
- [88] Hans C Veenema, Berry M Spruijt, Willem Hendrik Gispen, and JARAM Van Hooff. Aging, dominance history, and social behavior in java-monkeys (macaca fascicularis). *Neurobiology of aging*, 18(5):509–515, 1997.
- [89] Paul Walker. An outline of the history of game theory. Technical report, 1995.
- [90] Eliot Roy Weintraub. *Toward a history of game theory*, volume 24. Duke University Press, 1992.
- [91] Michael N Weiss, Daniel W Franks, Deborah A Giles, Sadie Youngstrom, Samuel K Wasser, Kenneth C Balcomb, David K Ellifrit, Paolo Domenici, Michael A Cant, Samuel Ellis, et al. Age and sex influence social interactions, but not associations, within a killer whale pod. *Proceedings of the Royal Society B*, 288(1953):20210617, 2021.
- [92] Douglas Brent West. *Introduction to graph theory*, volume 2. Prentice Hall Upper Saddle River, NJ, 1996.
- [93] Tina W Wey and Daniel T Blumstein. Social cohesion in yellow-bellied marmots is established through age and kin structuring. *Animal Behaviour*, 79(6):1343–1352, 2010.
- [94] Bin Wu, Da Zhou, Feng Fu, Qingjun Luo, Long Wang, and Arne Traulsen. Evolution of cooperation on stochastic dynamical networks. *Plos one*, 5(6):e11187, 2010.

- [95] Christopher C Yang, Xiaodong Shi, and Chih-Ping Wei. Discovering event evolution graphs from news corpora. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 39(4):850–863, 2009.