DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS SERIES B
Volume **xx**, Number **0x**, July **xxxx**

doi:xxxx

pp. **x–xx**

MODELLING CONFLICTING INDIVIDUAL PREFERENCE: TARGET SEQUENCES AND GRAPH REALIZATION

Raneem Aizouk and Mark Broom

Department of Mathematics,
City, University of London,
Northampton Square,
London EC1V 0HB, UK.

Abstract. This paper will consider a group of individuals who each have a target number of contacts they would like to have with other group members. We are interested in how close this can be to being realized. Considering the group's long-term outcome under reasonable dynamics on the number of contacts. We formulate this as a graph realization problem for undirected graphs, with the individuals as vertices and the number of desired contacts as the vertex degree. It is well known that not all degree sequences can be realized as undirected graphs, and the Havel-Hakimi algorithm characterizes those that can. When we ask how close the degree sequences can be to realization, we ask for graphs that minimize the total deviation between what is desired and possible. The set of all such graphs and all such associated sequences are termed the minimal sets. Broom and Cannings have previously considered this problem in many papers, and it is hard to tackle for general target sequences. This paper revisited the minimal set in general, investigating two particular classes of sequence in particular. We considered the n-element arithmetic sequence (n-1, n-2, ... 1,0) for general n, including obtaining a formula that generates the size of the minimal set for a given arithmetic sequence, and the all or nothing sequences, where targets are either 0 or n-1, where a recurrence relation for such a formula was found. Further, we consider the question of the size of the minimal set of sequences in general. We consider a strategic version of the model where the individuals are involved in a multiplayer game, each trying to achieve their target, and show that optimal play can lead to the minimal set being left, thus answering an open question from earlier work.

1. **Introduction.**

1.1. **General evolving models.**

The original game theoretical models of populations considered infinite well-mixed populations [13, 14, 32, 25, 17, 18] where all individuals can interact. There are important differences between finite and infinite populations; however, see, for example, [37]. Real populations are also not homogeneous but have a population structure, which has been incorporated in various ways. A population may evolve on a simple graph $G = (V, E)$, see [24], where vertices represent individuals and edges connections between them. Here the population is assumed to be comprised of (at least) two types of individuals. The games are played between neighbours, with the composition of the population changing according to the outcomes of these games; perhaps the essential property of such populations is the fixation probability, the probability that a randomly placed mutant will eventually replace the resident population [2, 8]. Other models consider situations where the graph grows as new individuals are introduced to the population following some reproductive process. For example, Southwell and Cannings in [33, 34, 35] consider a model where at each time point t every vertex produces an "offspring". The current edges are retained, and new edges involving the new vertices (amongst themselves and linked to the old vertices) are formed following some rules.

In real populations, both graph structure and population composition change. A work showing how the interactions between the individuals affect the structure, as well as the types occupying the vertices, has been done, e.g. see [36]. In [27, 28] the rates at which links are formed or broken depend upon the types of individuals involved. For a review see [29] (see also[1]). A lot of this work considers the problem of whether cooperative behaviour can evolve. The details vary, but Allen and Nowak [1] note that a common theme is that it is easier to achieve cooperation when the cooperators can group in a way that can exclude defectors, at least to some extent. In the current paper, we consider strategic network formation. Such models date back to 1975 [4]. More recently, Jackson and Wolinsky [23] considered a scenario where an individual's payoffs depend upon the network. And they considered incentives of individuals to form networks (see also [19, 20], and [21] for a review of such models, as well as a demonstration of some simple examples, which illustrate some complexities and interesting features of these models). A variety of scenarios have been considered, depending upon the allowable changes, whether individuals can condition their play on potential changes of others and whether links can be formed unilaterally; this latter situation [3, 11] is also how our model works.

1.2. **A dynamic network population model.**

Consider a network of interactions between the distinguishable vertices set $V = (1, 2, 3, ...)$, and the edges which link pairs of individuals (vertices) such that $X = (x_{ij})$ where $i, j = 1, 2, ..., n$ denotes the edges for any given graph (see [7]), with $x_{ij} = 1(0)$ representing the presence (absence) of an edge between individuals i and j. Edges can be formed or broken, and we follow a process in discrete time, which we describe below. We thus have a time-dependent matrix $X(t)$ describing the links, and this in turn generates the vector $e_t = (e_{1t}, e_{2t}, ..., e_{nt})$, the *sequence* e_t of $X(t)$, where at time t an individual i has e_{it} links. In [9] each vertex had a preferred range of links $[m_i, M_i]$ to other individuals, where $0 \leqslant m_i \leqslant M_i \leqslant n-1$. In this paper we will only consider $m_i = M_i = t_i$ where t_i is the unique target of individual i. Following a dynamic process, described precisely in Section **??**. For the unique target case, in [9] individuals were selected at random to change their number of links. The individual below (above) their target range would form (break) one of its links at random; an individual within its target range would make no change. Thus, we have a fixed set of individuals (vertices) while the links change randomly. The process follows a Markov chain, with vertices breaking or forming random connections to optimize their total number; for the unique target case that we will consider, individual i will aim to get as close as possible to the target t_i. Depending on their desired total number of links, the presence or absence of a specific link may benefit both, neither, or precisely one of the two involved.

Suppose the collection of targets, the *target sequence*, can be achieved by a graph. In that case, it is possible that all individuals can be on target simultaneously when no further changes happen (indeed, it is proved in [6] that the situation must eventually be reached). In this case, the target sequence is called a *graphical sequence*. A lot of work has been done on these sequences, for example in ([15], [12], [16], [26], and [30]). We shall discuss graphical sequences in Section 2.

1.3. **Paper outline.** In this paper, we consider a model where "evolution" takes place on the class of simple graphs with a fixed number of vertices that represent individuals interacting with each other, and the set of edges represents the links between these individuals. Section **??** has been a general introduction to our dynamic network population model. Section 2 describes the model, including the target sequence and the general concept of the minimal set and its properties. In Section 3 we consider the size of the minimal set, including for two key classes of target sequence. In Section 4 we discuss the game theoretical model introduced in [7] and further developed in [10] to resolve an open question from [7]. Finally Section 5 is a discussion of our results and future work.

2. The Model.

2.1. **Graphic Sequences.** In this section, we will present many key definitions for this paper, in particular, that of a graphical sequence, and we give sufficient and necessary conditions for a sequence to be graphic, which we will introduce as a theorem (The Havel-Hakimi Theorem), as well as mentioning some applications.

2.2. **What is a graphical sequence?**

Definition 2.1. The degree of a vertex (V_i) is the number of links going out of the vertex which connects the vertex with other individuals $deg(v_i)$.

Definition 2.2. A finite sequence $(t_1, t_2, t_3, \ldots, t_n)$ of non-negative integers is called a degree sequence of a graph G if the vertices of G can be labelled $v_1, v_2, \ldots, v_n$ such that $deg(v_i) = t_i$ for all $i = 1, 2, 3, \ldots, n$ [38].

Definition 2.3. The sequence $(e_1, e_2, \ldots, e_n)$ is graphical if there exists a graph which achieves that sequence exactly. Thus, if a sequence $S : (e_1, e_2, e_3, \ldots, e_n)$ is a degree sequence of some graph the sequence S is called a graphical sequence. Not every sequence of non-negative integers can be graphical. Below we define two important classes of sequence which we shall consider later in the paper.

Definition 2.4. The n-element *arithmetic sequence* has the form $(n-1, n-2, n-3, \ldots, 1, 0)$ for general n.

Definition 2.5. 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.

2.2.1. *Sufficient and Necessary conditions for a graphical sequence.*

Theorem 2.6 (the Havel-Hakimi Theorem)**.** *A sequence $S : (t_1, t_2, t_3, \ldots, t_n)$ of non-negative and non-increasing integers with $n \geq 2$, $t_1 \geq 1$, is graphic if and only if the sequence S' : $(t_2 - 1, t_3 - 1, t_{t_{i+1}} - 1, \ldots, t_n)$ is graphic.*

See [6, 12, 16] for more on graphic sequences. Here S' is obtained by deleting the largest element in S and subtracting one from the next largest elements. We note that it may be that $t_2 - 1 \geq t_3 - 1 \geq t_{t_{i+1}} - 1 \ldots \geq t_n$ does not hold, in which case we would normally re-order the terms to again be a non-increasing sequence. This is necessary for the repeated implementation of the Havel-Hakimi result to a sequence often carried out.

2.2.2. *Vertex classification.* In this section we will discuss *vertex classifications* (see [7]). We shall see that for specific target sequences, following the random process described in Section 4.1, we will eventually reach a situation (the minimal set, as discussed later) where some individuals might never be under or over target, and we shall classify them by their long-term possibilities. We shall consider a target sequence t and the sequence of a graph v, $e(v)$.

Definition 2.7. The distance between two sequences (a) and (b) is giving by: $Z(a, b) = \sum_{i=1}^{n} |a_i - b_i|$.

Definition 2.8. The deviation of a graph (and its associated sequence) is the distance from that sequence to the target sequence t.

Definition 2.9. The score of t is $s(t) = \min_{v \in G} Z(e(v), t)$.

The score is the minimal value of the deviation for all graphical sequences of the target t.
The *adjusted Havel-Hakimi algorithm*finds the score of a sequence by following the steps:

- Sort the sequence in decreasing order.
- Remove the first (greatest) element. If this greatest element was k, subtract 1 from the k following terms.
- If a negative number appears (only -1 is possible) change it to zero and add one to the running total (which starts at 0).
- Repeat this process until we end up with a sequence of zeros.

The score is then the final total of summed 1's. We will now define two central concepts for our paper, the minimal sets of graphs and sequences.

Definition 2.10. The minimal set of graphs, K(min), is the set of graphs which achieve the score.

Definition 2.11. The minimal set, J(min), is the set of sequences of graphs which achieve the score.

For sequences, we must be careful about ties. We say a sequence of integers contains a *tie* if an integer belonging to that sequence is repeated in a row, in other words if two consecutive individuals have the same target. For example in the target sequence $\{2, 2, 1\}$ there is a tie in the first and second positions. For members of the minimal set permutations within the tied positions are equivalent, for example $\{2, 1, 1\}$ and $\{1, 2, 1\}$ for the above sequence. To find J(min) we thus disallow any sequence that is not non-increasing in the tie positions. For each element of J(min) there will be a corresponding set of at least one element of K(min). For graphic sequences the score is 0, as the target can be achieved by a graph, and consequently J(min) has a single element (the target sequence itself). Here we are interested in non-graphic sequences, where not all vertices can achieve their target simultaneously. For any given graph, a vertex can be in one of three states:

- vertex i is Neutral if $e_i = t_i$ (it would neither wish to form nor break an edge);
- vertex i is a Joiner if $e_i < t_i$ (it 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).

It was shown in [7] that all vertices are Neutral for at least some elements of J(min). This then implies that every vertex is of one of the following four types:

1. vertices which can be either Joiner or Breaker in some elements, and necessarily must be Neutral for some other elements, of J(min). We denote the set of these as S_A.
2. vertices which are either Joiners or Neutral for all elements of J(min). We denote the set of these as S_J.
3. vertices which are Breakers or Neutral for all elements of J(min). We denote the set of these as S_B.
4. vertices which are Neutrals only. We denote the set of these as S_N.

By Lemma 4.1 in [7], We know that we cannot have vertices from S_A and vertices from S_N for the same sequence. Clearly, if the sequence is graphic then the only sequence within J(min) corresponds to members of K(min) where the target is achieved for all vertices, i.e. all of its vertices will be Neutral. In this case we cannot have Joiners or Breakers and all vertices are of type 4.

2.3. **Properties of the minimal set.** In order to reach the minimal set, there will be a number of steps and transitions which reduce the deviation of the graph until the minimal set J(min)/K(min) is reached. By noting that no further improvement can occur we can conclude 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 is joined to a Breaker it means that Neutral is joined to every Joiner. Similarly, if a Neutral is not joined to some Joiners that means it is not joined to any Breakers.
4. If we have a Neutral N which is joined to some Breakers and another Neutral $N*$ is joined to some other Breakers then N and $N*$ are joined.

5. If some Neutral N is not joined to some Joiners and another Neutral $N*$ is not joined to some other Joiners then N and $N*$ are not joined.

Note 2.12. $J(min)$ contains at least one member with no Joiners and at least one with no Breakers (see [6] for further details).

Note 2.13. It was shown in [7] that (for a non-increasing target sequence) that members of S_J must have lower index than members of S_N or S_A, which must have lower index than members of S_B. Some sets can be missing, indeed as stated above both S_A and S_N cannot occur, but any present must always satisfy the above conditions.

3. **The size of the minimal set.** In this section we shall consider the minimal set(s) more generally, and in particular we are interested in the size of the minimal set. We consider two special cases, and the size of the largest minimal set in general for a given number of vertices n.

3.1. **The largest minimal set.** We consider two types of the minimal set, namely J(min) and K(min). We shall mainly consider J(min) for the following reasons:

1. Dealing with J(min) is simpler than K(min), partly due to the difference in the size of the minimal sets. Every sequence (in J(min), and in general) has at least one graph (in K(min), and in general) which makes K(min) larger than J(min). The number of graphs for n vertices is $2^{\binom{n}{2}}$, and 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 $nln(n)$ (many sequences are not graphical and ties in the target further reduces 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 [6]. Whether any given graph/sequence achieves the score is thus established through consideration of 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 following theorem. 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 easy set to find. The maximal score $ms(n)$ for n vertices occurs as follows.

Theorem 3.1. *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 [7] 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 all. 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 1 we see that there are six states within the minimal set for the sequence 111; however, the sequences with the maximal score are 220 and 200 with four states in the minimal set.

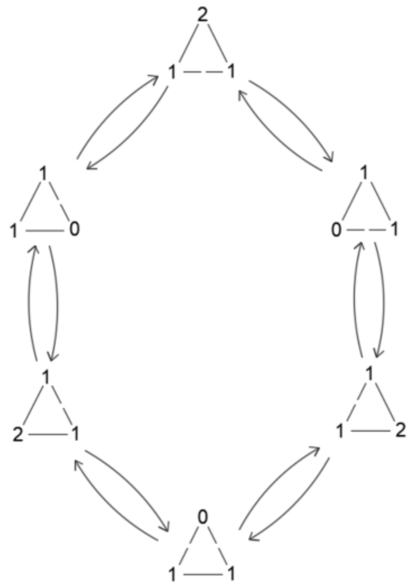

FIGURE 1. Transitions between the elements of the minimal set of graphs K(min) of the sequence 111 showing every possible move between each of the six elements.
——— represents a link between the corresponding two nodes.
- - - - represents the absence of such a link.

We have considered sequences of this type before in [7]. We considered a Markov process with individuals not achieving their targets always choosing out of the potential improving moves uniformly. And found the stationary distribution over the population states (the set of recurrent states, which is K(min) and which formed a single class). We note that for the strategic game considered in Section 4, the corresponding process would be much more difficult to analyse. There are many 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 a 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 2.2.2 means that many different sequences are identical. Therefore 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)? We cannot answer this question 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.

In Table 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 sequence

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 1. The sequences with the biggest minimal set size where n is the number of individuals, —J(min)— size: 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.

described later. From Table 1 we see that the largest minimal set is not in general for the all-or-nothing sequences, except for when there are very few nodes. Why is this? Let us consider the case with 6 nodes as an example.

Example 3.2. Consider the following sequences

- 1) 555000
- 2) 554100
- 3) 543210

The first sequence has the largest score (9), the third sequence has a much smaller score (3), and the second sequence has an intermediate value of the score (7). Thus we might think that (1) should have the largest minimal set of the three and (3) the smallest.

The first sequence has the largest three numbers and the smallest three numbers all the same; the second sequence has the largest and smallest two numbers the same, but the middle two take different values; the third sequence has all numbers different. When we find the minimal set, we must reorder the first and last three numbers for sequences in the minimal set of (1) and the first and last two numbers for sequences in the minimal set of (2) to account for ties as discussed in Section 2.2.2 (for sequence (3) there are no ties to consider). Thus there will be more equivalent sequences for (1) than (2), and therefore (1) will have more elements to be omitted from the minimal set. Thus J(min) is larger for the second sequence even though the score is smaller. 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 ever increasing. We have also seen in the above example, and from Table 1, that the arithmetic sequence, which is the only type with no ties, also does not achieve the maximum minimal set size. It has a relatively low score, and we hypothesise that, in general, a combination of a high score and a low number of ties gives a large minimal set. Note that we see in Section 3.3 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 sequence, arithmetic and all or nothing, that we have discussed above.

3.2. **The minimal set for the arithmetic sequence.** In this section we consider the $n-$ *element* arithmetic sequence $(n-1, n-2, n-3, \ldots, 1, 0)$ for general n as mentioned in definition 2.4. In particular, we demonstrate an exact formula for the size of the minimal set J(min) for this sequence in the process of identifying its precise membership. We will label the set as J_n and thus the size of the set as $\mid J_n \mid$.

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 [7] 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 3.3. *The score for $n = 2m$ is m.*

Proof: $2m - 1, 2m - 2, \ldots, m, m - 1, \ldots, 0$ is an arithmetic sequence. Applying the adjusted H-H algorithm on this sequence as follows, we obtain:
$2m - 1, 2m - 2, \ldots, m - 1, \ldots, 0$ $(+1)$
$2m - 3, 2m - 4, \ldots, 0, 0$ $(+1)$
$2m - 5, 2m - 6, \ldots, 0, 0, 0$ $(+1)$
and so on. After the rth step, we obtain the sequence: $2m - 2r - 1, 2m - 2r - 2, \ldots, 1, 0, \ldots, 0$ where there are r zeros with a cumulative total of r 1's added. The leading term reduces by 2 for each further step and 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 3.4. *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 [7].*

Proof: We shall divide the sequence $(2m - 1, 2m - 2, \ldots, m, m - 1, \ldots, 0)$ into two sets: $S_1 = \{2m - 1, 2m - 2, \ldots, m\}$ and $S_2 = \{m - 1, m - 2. \ldots, 0\}$.
We further denote:
β_i: The links within S_1 (or S_2) for $i = 0, 1, 2, \ldots, 2m - 1$,
α_i: The links between the sets S_1 and S_2 for $i = 0, 1, 2, \ldots, 2m - 1$.

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

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

We know that
$$\sum_{i=0}^{m-1} \alpha_i = \sum_{i=m}^{2m-1} \alpha_i,$$
since the α_is represent the links between S_1 and S_2 and so the summation is common between the two sets.

Thus
$$D \geq \sum_{i=0}^{2m-1} i - 2\sum_{i=0}^{m-1} i - \sum_{i=m}^{2m-1} \beta_i + \sum_{i=0}^{m-1} \beta_i =$$
$$(2m(2m - 1)/2) - (2m(m - 1)/2) - \sum_{i=m}^{2m-1} \beta_i + \sum_{i=0}^{m-1} \beta_i =$$
$$m^2 - \sum_{i=m}^{2m-1} \beta_i + \sum_{i=0}^{m-1} \beta_i.$$
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 $i = 0, 1, ..., m-1$ for S_1 and $j = 1, 2, ..., m$ 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 $0 \le k \le m$ we have $k < \alpha_{k+m-1}$, then

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

$$(\alpha_{k+m-1} - k) + \sum_{j \neq k}^{m} |j - \alpha_{j+m-1}| + \sum_{i=0}^{m-1} |i - \alpha_i| \ge$$

$$2(\alpha_{k+m-1} - k) - \sum_{j=1}^{m} \alpha_{j+m-1} + \sum_{j=1}^{m} |j| + \sum_{i=0}^{m-1} \alpha_i - \sum_{i=0}^{m-1} |i| =$$

$$2(\alpha_{k+m-1} - k) + (m(m+1)/2) - (m(m-1)/2) = 2(\alpha_{k+m-1} - k) + m > m.$$

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

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

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

$$\sum_{j=1}^{m} |\alpha_{j+m-1} - j)| + (k - \alpha_k) + \sum_{i \neq k}^{m-1} |\alpha_i - i| \ge$$

$$\sum_{j=1}^{m} |j| - \sum_{j=1}^{m} |\alpha_{j+m-1}| + 2(k - \alpha_k) + \sum_{i=1}^{m-1} |\alpha_i| - \sum_{i=1}^{m-1} |i| =$$

$$2(k - \alpha_k) + m > m,$$

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

Lemma 3.5. *If a member of the minimal set has elements where for $i < j$ in S_J we have $t_i \ge t_j$ and $e_i \le e_j$, 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 $S : e_1, e_2, \ldots, e_i, \ldots, e_j, ..., e_n$
and $S' : e_1, e_2, \ldots, e_j, \ldots, e_i, \ldots, e_n$, the same sequence with e_i and e_j swapped. We will define

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

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 $D'_s - D_s \le 0$. To do that we will discuss the following cases:

1. If $e_j \ge e_i \ge t_i \ge t_j$
   $D'_s - D_s = (e_j - t_i) + (e_i - t_j) - (e_i - t_i) - (e_j - t_j) = 0$, so $D'_s = D_s$.
2. If $e_j \ge t_i \ge e_i \ge t_j$
   $D'_s - D_s = (e_j - t_i) + (e_i - t_j) - (t_i - e_i) - (e_j - t_j) = 2e_i - 2t_i \le 0$, so $D'_s \le D_s$.
3. If $e_j \ge t_i \ge t_j \ge e_i$
   $D'_s - D_s = (e_j - t_i) + (t_j - e_i) - (t_i - e_i) - (e_j - t_j) = 2t_j - 2t_i \le 0$, thus $D'_s \le D_s$.

4. If $t_i \geq e_j \geq e_i \geq t_j$
   $D'_s - D_s = (t_i - e_j) + (e_i - t_j) - (t_i - e_i) - (e_j - t_j) = -2e_j + 2e_i \leq 0$, thus $D'_s \leq D_s$.
5. If $t_i \geq e_j \geq t_j \geq e_i$
   $D'_s - D_s = (t_i - e_j) + (t_j - e_i) - (t_i - e_i) - (e_j - t_j) = -2e_j + 2t_j \leq 0$, 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. □.

Lemma 3.6. *From Lemmas 3.3, 3.4 and 3.5 we have that the vertex deviations, denoted ϵ_i in the terminology of [7], 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 3.5 we have the following: $1 \leq i \leq m$, $\beta_i = m-1$, so $e_i \geq m-1$.
   From Lemma 3.5 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 3.5 we have: $m + 1 \leq i \leq 2m, \beta_i = 0, e_i \leq m$.
   From Lemma 3.5 again we have: $e_i \geq 2m - i$, so the Deviation $\epsilon_i = e_i - (2m - i) \leq m - (2m - i) = i - m$. □.

Lemma 3.7. *The following inequality holds.*

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

Proof: From Lemma 3.5 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 (1) we get:

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

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

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

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

We will divide $j = 1, 2, \ldots, 2m$ 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 connect 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 the 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 3.8. *For any sequence that satisfies Lemmas 3.6 and 3.7 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: $(1, 2, \ldots m, m+1, \ldots, 2m-1, 2m)$ so that the target of node i is $2m-i$. We will divide this sequence into two sets: $(S_1 : 1, 2, \ldots m)$ and $S_2 : (2m, 2m-1, \ldots m+1)$, 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:
$S_1 : 2m-1-\epsilon_1, 2m-2-\epsilon_2, \ldots, m+1-\epsilon_m$.
$S_2 : 0+\epsilon_{2m}, 1+\epsilon_{2m-1}, \ldots, m-2+\epsilon_{m+2}, m-1+\epsilon_{m+1}$.
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 3.5 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 2. By this procedure we will get the following required extra links for each node in S_1 and S_2 respectively:
$S_1 : 1-\epsilon_1, 1-\epsilon_2, \ldots, 1-\epsilon_{m-1}, 1-\epsilon_m$.
$S_2 : \epsilon_{2m}, \epsilon_{2m-1}, \ldots, \epsilon_{m+2}, \epsilon_{m+1}$.
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

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

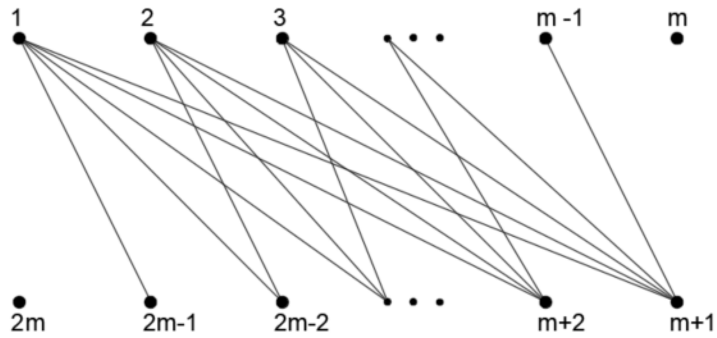

FIGURE 2. This figure shows the links between node i from S_1 and node j in S_2, where $S_1 : 1, 2, \ldots, m-1, m.$ and $S_2 : 2m, 2m-1, \ldots, m+2, m+1$ at the second step of the process in Lemma 3.8 (other links between the two sets are absent). These are in addition to the links between all elements of S_1. All elements of S_2 are split from each other. We have thus illustrated the situation prior to the third step.

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 $j \leqslant i$, following the previous procedure, then we are left with all nodes on target, and finally that we have a sequence in the minimal set. □.

Now let us define Q^i_k as the number of ways of picking a sequence of $2i$ numbers $q_1, q_2, \ldots. q_{2i}$ which sum to k where the following conditions are satisfied:

$$\begin{cases} q_j \leq i+1-j & j \leq i \\ q_j \leq j-i & j > i \end{cases} \tag{2}$$

and

$$\sum_{j=1}^{l} q_j + \sum_{j=2i-l+1}^{2i} q_j \geq l+k-i \tag{3}$$

In particular Lemma 3.8 implies that

$$\mid J_{2m} \mid = Q^m_m.$$

Lemma 3.9. $Q^i_k = \sum_{j=\max(k-i+1,0)}^{k} (j+1) Q^{i-1}_{k-j}$, *for* $k \leqslant i$, $Q^i_0 = 1$.

Proof:

1. Let W^i_k be the set of all sequences of length $2i$ which sum to k and satisfy Inequalities 2 and 3. Thus by the definition of Q^i_k we have $|W^i_k| = Q^i_k$.
2. Let $X = (q_1, q_2, \ldots, q_{2i}) \in W^i_k$ where $q_1 + q_{2i} = j_0$ and $k-i+1 \leqslant j_0$. We will show that this implies that $X' = (q_2, \ldots, q_{2i-1}) \in W^{i-1}_{k-j_0}$.

   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 (2).
Now we must show that the sequence X' also satisfies Inequality (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 (3) holds.
Thus we have that $X' \in W^{i-1}_{k-j_0}$.

3. We now show the reverse implication to the above, namely that if $X = (q_2, \ldots, q_{2i-1}) \in W^{i-1}_{k-j_0}$, $q_1 + q_{2i} = j_0$ and $k-i+1 \leqslant j_0$ then $X' = (q_1, q_2, \ldots, q_{2i-1}, q_{2i}) \in W^i_k$.

   It is clear that X' is of length $2i$ and $\sum_{j=1}^{2i} q_j = k$. For $2 \leqslant j \leqslant i$, $q_j \in W^{i-1}_{k-j_0}$ satisfies $q_j \leq (i-1)+1-(j-1) = j+1-j$ and $i < j \leqslant 2i-1$ satisfies $q_j \leq (j-1)-(i-1) = j-i$. Since we also have $q_1, q_{2i} \leqslant i$ then all of the entries of X' satisfy Inequality (2).
   Now we must show that the sequence X' also satisfies Inequality (3).
   For $l = 1$, we have $q_1 + q_{2i} = j_0 \geq k-i+1$.
   For $2 \leqslant l \leqslant i$ 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 (3) holds.
4. Let $X = (q_2, q_3, \ldots, q_{2i-1}) \in W^{i-1}_{k-j_0}$. and $X' = (q_1, q_2 \ldots, q_{2i}) \in W^i_k$.
   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 $W^{i-1}_{k-j_0}$ can be transformed into a sequence of W^i_k in j_0+1 unique ways.
5. Finally we note that no two sequences in W^{i-1}_{k-j} or two sequences one each from $W^{i-1}_{k-j_1}$ and $W^{i-1}_{k-j_2}$ can generate the same sequence in W^i_k.
   If x_1 and x_2 are two such distinct elements then they differ in at least one position, so then any supersequence of them must also differ similarly.

We shall solve the recurrence relation in Lemma 3.9 to get a solution for $\mid J_{2m} \mid$. We note that from the above we have a way of identifying all of its elements (any sequence with deviations satisfying Lemma 3.8.

Theorem 3.10. *We have the following formulae:*

$$Q^i_k = \binom{2i+k+1}{k} - 2\binom{2i+k+1}{k-1}. \tag{4}$$

From Equation (4) and $\mid J_{2m} \mid = Q_m^m$ *we will have:*

$$|J_{2m}| = \binom{3m+1}{m} - 2\binom{3m+1}{m-1}. \tag{5}$$

Proof:

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

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

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

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

$$Q_k^i - 2Q_{k-1}^i + Q_{k-2}^i = Q_k^{i-1}. \tag{6}$$

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 $A_i k + B_i$.

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

$$Q_k^i = \binom{2i+k+1}{k}, \tag{7}$$

by substituting into Equation (6). Thus

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

as required. Similarly, checking a second term

$$Q_k^i = \binom{2i+k+1}{k-1} \tag{8}$$

it is easy to see that Equation (6) is also satisfied as required. We will consider the expression in (7) minus twice that in (8) 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) as required, and the theorem follows directly from this. $\square$.

Example 3.11. 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 $|J_4| = \binom{7}{2} - 2\binom{7}{1} = 21 - 14 = 7$...etc, see Table 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 [7]), see for example Table 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 (5) and the above approximation in Table 2.

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

TABLE 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.

3.3. **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 2.5 and we considered in Section 3.1. For example the sequences in Table 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 3.1. Using the Matlab program described in Section 3.1 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 Theorem 3.1 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

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

problem was considered in [9], and is addressed in the on-line encyclopedia of integer sequences [31] as sequence number A029894 (strictly speaking the above sequence $2, 3, 7, 34, \dots$ is sequence A327913, and A327913 is given in tabular form as it is two-dimensional). [31] gives the following formulae, which we use to generate Table 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. where the function F satisfies the following iterative formula:

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

where $F(b,c,0,0) = 1$, $F(b,c,t,0) = 0$, $t > 0$. For example:
$T(2,2) = F(2,2,0,2) = F1,1,0,1) + F(2,1,1,1) + F(2,2,0,1) = F(0,0,0,0) + F(0,0,0,0) + F(1,1,0,0) + F(0,1,0,0) + F(0,0,0,0) + F(1,1,0,0) + F(2,2,0,0) = 7$. Note that T(1,1)=2,

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. The values of the minimal set for the all or nothing sequences with m_1 "all" vertices and m_2 "nothing" vertices.

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 neighbour, from Table 4, which are of course the same values of the minimal sets size of the sequences as given in Table 3.

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

TABLE 5. Here we show the logarithm of the size of the minimal sets for the all or nothing and arithmetic sequences.

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

TABLE 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 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 3.2 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 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 one. 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 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.

4. **The Game-Theoretical Model and the Minimal set.** In this section we consider the game-theoretical model first introduced in [7] and examine whether consideration of the minimal set is enough to investigate the optimal behaviour, a previously open question. We show below with a simple example that it is not, but first we need to define the game. In order to do this, we need to define the underlying Markov chain which governs how links are formed and broken.

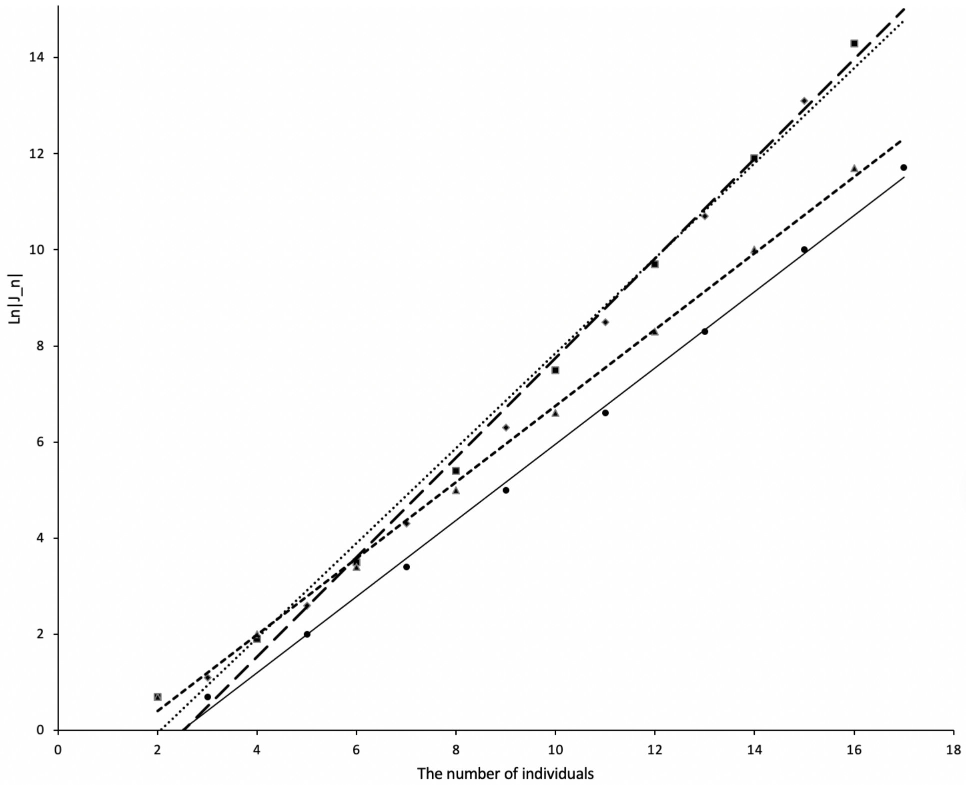

FIGURE 3. In this figure we considered two types of sequences; the all or nothing and the arithmetic sequences, and fitted two lines for each; one for the even values and the other for the odd terms of the sequence.
Line 1:represented by ■ ...... refers to the even terms of the all or nothing sequences, $Y = 0.9875x - 2.025$.
Line 2: represented by ◇___ ___ refers to the odd terms of the all or nothing sequences, $Y = 1.0357x - 2.6071$.
Line 3: represented by ▲_ _ _ _ refers to the even terms of the arithmetic sequences, $Y = 0.7935x - 1.1786$.
Line 4: represented by •_______ refers to the odd terms of the arithmetic sequences, $Y = 0.7935x - 0.3851$.

4.1. **The Markov Chain Model:** In Section 1.2 we denoted the set of links by the matrix $\mathbf{X}$, where $x_{ij} = x_{ji} = 1$ if I_i and I_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.

As in [7, 10], 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 $\mathbf{U_X}$. For an $\mathbf{x}^*$ 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,

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

I_i will have a number of edges e_{it} to other individuals, so that the above process generates an evolving sequence $e_t = (e_{1t}, e_{2t}, e_{3t}, ...e_{nt})$. In particular in [5] 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

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

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

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

4.1.1. *The Game.* $\mathbf{U_X}$ 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 [7]; 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:

$$R_i(\mathbf{U_X}) = -\sum_{\mathbf{X}} \epsilon_i(\mathbf{X})\pi(\mathbf{X}), \tag{9}$$

where $\epsilon_i(\mathbf{X})$ is the deviation of I_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 ith row of $\mathbf{U_X}$ for a single $\mathbf{X}$ only (as opposed to *global changes*, where I_i is able to change the ith row of $\mathbf{U_X}$ for any number of states simultaneously). Considering all possible changes by any single individual, a strategy set is a Nash equilibrium under if, under all allowable changes by $i : \mathbf{U_X} \to \mathbf{U^i_X}$

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

4.2. **Example showing that leaving the minimal set can be optimal.** Suppose that we have n individuals, one of them I_1 with target $n-2$ and the others with target zero, i.e we have the target sequence:$\{n-2, 0, 0, \ldots.., 0\}$. 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 I_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/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 sometimes optimal to leave the minimal set, in particular that I_1 should form a link when already on target.

Consider the states $S_0, S_1, \ldots, S_{n-1}$, all of them involving no links between individuals with target zero, where in state S_k I_1 has links to precisely k other individuals. Given all have target 0 and will break their one possible link (with I_1) when they have the chance, it does not matter which of the k individuals I_1 is connected to, and so we represented this by a single state.

Assume that if I_1 is selected in S_{n-2}, then it will choose to form the final link with probability p, and that if I_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 Markov process is given in Matrix (10) and can be used to find the stationary distribution,

$\pi_0, \pi_1, \pi_2, ..., \pi_{n-1}$ where and $0 \leq p \leq 1$, $0 \leq q \leq 1$:

$$P = \begin{pmatrix} (n-1)/n & 1/n & 0 & . & . & . & 0 \\ 1/n & (n-2)/n & 1/n & . & . & . & 0 \\ 0 & 2/n & (n-3)/n & 1/n & . & . & 0 \\ 0 & 0 & 3/n & n-4/n & 1/n & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & . & . & n-3/n & 2/n & 1/n & 0 \\ 0 & . & . & . & (n-2)/n & (2-p)/n & p/n \\ 0 & . & . & . & 0 & (n-1+q)/n & (1-q)/n \end{pmatrix}. \tag{10}$$

The payoff $R(p,q)$ to I_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:

$$\pi_{m-1} = \frac{1}{n}\pi_{m-2} + \frac{n-m}{n}\pi_{m-1} + \frac{n}{m}\pi_m \Rightarrow m\pi_{m-1} = \pi_{m-2} + m\pi_m. \tag{11}$$

Let us assume that

$$\pi_k = \frac{1}{k!}\pi_0. \tag{12}$$

Substituting Equation (12) into Equation (11) 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:

$$\pi_{n-2} = \frac{1}{n}\pi_{n-3} + \frac{2-p}{n}\pi_{n-2} + \frac{n-1+q}{n}\pi_{n-1} \Rightarrow$$
$$\frac{n-2+p}{n}\pi_{n-2} = \frac{1}{n(n-3)!}\pi_0 + \frac{n-1+q}{n}\pi_{n-1} \Rightarrow$$
$$\pi_{n-1} = \frac{p}{n-1+q}\frac{\pi_0}{(n-2)!}.$$

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

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

2. We now proceed to find the payoff of individual I_1 by applying the following formula for the payoff following Equation (9):

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

Thus

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

$$R(p,q) = -\frac{(n-2)\sum_{m=0}^{n-3} 1/m! + \sum_{m=1}^{n-3} 1/(m-1)! + p/(n-1+q)(n-2)!}{\sum_{m=0}^{n-2} 1/m! + p/((n-1+q)(n-2)!}. \tag{13}$$

Denoting $S_x = \sum_{m=0}^{x} 1/m!$ and $r = p/(n-1+q)$ in (13) we obtain:

$$R(p,q) = -\frac{(n-2)S_{n-3} + S_{n-4} + r}{S_{n-2} + r} \tag{14}$$

Finding the derivative of Equation (14) with respect to r we obtain

$$\frac{dR(p,q)}{dr} = \frac{-S_{n-2} + (n-2)S_{n-3} + S_{n-4}}{(S_{n-2} + r)^2} \tag{15}$$

Now we will discuss the sign of the numerator of Equation (15), 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 $S_0 = 1, S_1 = 2, S_2 = 2.5, S_3 = 2.6, \ldots$, which leads to the following terms for O_n; $O_2 = -1, O_3 = -1, O_4 = 2.5, O_5 = 6.9, \ldots$. In general O_n is positive for all larger values of n. 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 ($r = 1$) when $p = 1, q = 0$ and minimised ($r = 0$) when $p = 0$ and q can take any value. Thus $r = 1$, and so $p = 1$, is the optimal strategy whenever $n \geq 4$ and we thus have that I_1 should take the extra link when it has the chance, which leads to the minimal set being left, whenever $n \geq 4$.

5. **Discussion.** In this paper we have further developed the model of conflicting individual preference introduced in [5]. In this model each individual is assigned to a vertex in a graph, with relations between these individuals represented by the set of edges between the vertices. Every individual has a specific target number of links they would like to form with their neighbours. The composition of the population does not change, but the relations between the individuals vary over time following a dynamic process, the transition probabilities of which are governed by strategic choices of the individuals, as introduced in [7].

An important concept of these models is that of the minimal set (of graphs or their associated sequences), where collectively the total population deviation, a measure of how far they are from satisfying all individuals' objectives, is satisfied. Up to now only some general properties of these were known, and the core focus of the paper was to consider this in more precise detail. In particular 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 general these are not the largest minimal sets, and we have also investigated all minimal sets for a small number of vertices. There are also interesting questions to be examined, even for these sequences, e.g. in the long run will the minimal set of the arithmetic or all or nothing sequence be the larger (there is support for this being the latter)?

We considered the game theoretical model developed in [7] 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. In Section 4 we have shown a counter-example where 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. The study of social networks and their role in shaping human activities has been a subject of increasing interest in recent years. A good review of work on how networks influence social and economic activity as well as how they can be modelled and analysed (up to the time of its publication) is provided in [22]. There is hope that using such methods will lead to a better understanding of the patterns of human interactions, and will open the door for future applications. Jackson and Wolinsky [23] studied a case closely related to ours, where individuals' payoffs depend upon the network, taking into consideration the incentive for individuals to form networks.

In their model self-interested individuals could break links unilaterally, although the formation of links required the consent of both individuals. Thus the detailed analysis was rather different to ours, and they concentrated on proving general results under these assumptions. The key aspects that we consider in this paper do align well with their ideas, however. Firstly for our game, minus the graph deviation is the value as they define it, and similarly the minimal set is precisely the set of graphs which they call strongly efficient. We note that we have no stable states under their definition, unless the sequence is graphic.

There are a number of potential ways to develop this work in the future. Perhaps an individual prefers 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 be able to form links with their close neighbours.

Alternatively, we might consider different link formation/ breaking rules. In our current work the selected individual can form or break a link without the receiving individual having any say in this. What if links could only be added or broken with the permission of the receiver, or if any individual can break a link but both participants must agree to form one,as in [23]? Work on the former case is in its early stages.

REFERENCES

[1] Benjamin Allen and Martin A Nowak. Games on graphs. *EMS surveys in mathematical sciences*, 1(1):113–151, 2014.

[2] Tibor Antal, Sidney Redner, and Vishal Sood. Evolutionary dynamics on degree-heterogeneous graphs. *Physical review letters*, 96(18):188104, 2006.

[3] Venkatesh Bala and Sanjeev Goyal. A noncooperative model of network formation. *Econometrica*, 68(5):1181–1229, 2000.

[4] Scott A Boorman. A combinatiorial optimization model for transmission of job information through contact networks. *The bell journal of economics*, pages 216–249, 1975.

[5] 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.

[6] Mark Broom and Chris Cannings. Graphic deviation. *Discrete Mathematics*, 338(5):701–711, 2015.

[7] 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.

[8] Mark Broom and Jan Rychtář. An analysis of the fixation probability of a mutant on special classes of non-directed graphs. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 464(2098):2609–2627, 2008.

[9] Mark Broom and Jan Rychtár. *Game-theoretical models in biology*. Chapman and Hall/CRC, 2013.

[10] Chris Cannings and Mark Broom. Game theoretical modelling of a dynamically evolving network : Target sequences of score 1. *Journal of Dynamics & Games*, 7(1):37, 2020.

[11] Bhaskar Dutta and Matthew O Jackson. The stability and efficiency of directed communication networks. *Review of Economic Design*, 5(3):251–272, 2000.

[12] Seifollah Louis Hakimi. On the realizability of a set of integers as degrees of the vertices of a graph. *SIAM Journal Applied Mathematics*, 1962.

[13] William D Hamilton. The genetical evolution of social behaviour. ii. *Journal of theoretical biology*, 7(1):17–52, 1964.

[14] William D Hamilton. Extraordinary sex ratios. *Science*, 156(3774):477–488, 1967.

[15] Werner Hässelbarth. Die verzweigtheit von graphen. *Match*, 16:3–17, 1984.

[16] Václav Havel. A remark on the existence of finite graphs. *Casopis Pest. Mat.*, 80:477–480, 1955.

[17] Josef Hofbauer and Karl Sigmund. Dynamical systems and the theory of evolution, 1988.

[18] Josef Hofbauer, Karl Sigmund, et al. *Evolutionary games and population dynamics*. Cambridge university press, 1998.

[19] Matthew O Jackson. The stability and efficiency of economic and social networks. In *Advances in economic design*, pages 319–361. Springer, 2003.

[20] Matthew O Jackson. *Social and economic networks*. Princeton university press, 2010.

[21] Matthew O Jackson. An overview of social networks and economic applications. In *Handbook of social economics*, volume 1, pages 511–585. Elsevier, 2011.

[22] Matthew O Jackson. An overview of social networks and economic applications. In *Handbook of social economics*, volume 1, pages 511–585. Elsevier, 2011.

[23] Matthew O Jackson and Asher Wolinsky. A strategic model of social and economic networks. *Journal of economic theory*, 71(1):44–74, 1996.

[24] Erez Lieberman, Christoph Hauert, and Martin A Nowak. Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316, 2005.

[25] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge university press, 1982.

[26] Russell Merris and Tom Roby. The lattice of threshold graphs. *J. Inequal. Pure Appl. Math*, 6(1):1–21, 2005.

[27] Jorge M Pacheco, Arne Traulsen, and Martin A Nowak. Active linking in evolutionary games. *Journal of theoretical biology*, 243(3):437–443, 2006.

[28] Jorge M Pacheco, Arne Traulsen, and Martin A Nowak. Coevolution of strategy and structure in complex networks with dynamical linking. *Physical review letters*, 97(25):258103, 2006.

[29] Matjaž Perc and Attila Szolnoki. Coevolutionary games—a mini review. *BioSystems*, 99(2):109–125, 2010.

[30] Ernst Ruch and Ivan Gutman. The branching extent of graphs. *J. Combin. Inform. System Sci*, 4(4):285–295, 1979.

[31] Neil J. A. Sloane. sloane: lineencyclopedia of integer sequence.

[32] Maynard Smith and George Robert Price. The logic of animal conflict. *Nature*, 246(5427):15–18, 1973.

[33] Richard Southwell, Chris Cannings, et al. Some models of reproducing graphs: I pure reproduction. *Applied Mathematics*, 1(03):137, 2010.

[34] Richard Southwell, Chris Cannings, et al. Some models of reproducing graphs: Ii age capped vertices. *Applied Mathematics*, 1(04):251, 2010.

[35] Richard Southwell, Chris Cannings, et al. Some models of reproducing graphs: Iii game based reproduction. *Applied Mathematics*, 1(05):335, 2010.

[36] György Szabó and Gabor Fath. Evolutionary games on graphs. *Physics reports*, 446(4-6):97–216, 2007.

[37] Christine Taylor, Drew Fudenberg, Akira Sasaki, and Martin A Nowak. Evolutionary game dynamics in finite populations. *Bulletin of mathematical biology*, 66(6):1621–1644, 2004.

[38] Douglas Brent West et al. *Introduction to graph theory*, volume 2. Prentice hall Upper Saddle River, NJ, 1996.