THE UNIVERSITY OF CHICAGO

ON THE PROBABILTY OF EXISTENCE OF STABLE NETWORKS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

COMMITTEE ON ECONOMICS

BY
MARTIN SANTAMARIA

CHICAGO, ILLINOIS
DECEMBER 2013

To my wife and two children

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I thank my comitte Professor Myerson, Professor Szentes and specially Professor Hugo Sonnenschein for their generous continuos support on this project. I also would like to thank all the participants in the economic theory working group for valuable comments and discussion. Thanks go also to Robert Herbst and Julie Less for their help through out the years. I thank Yakym Pirozhenko for developing the codes for the simulations. To all my friends and specially my wife Bike and my two children Ozgur and Inti for encouragement and pacience. Finally, to my dad, my mom and my sister, for their sacrifice of not seeing me for so long.

ABSTRACT

Network formation models are a class of models for which the structure of the network is endogenously determined. Strategic network formation treats each node as a player that has preferences over network structures. The solution concept of pairwise stability is widely used in the strategic network formation literature.Intutively a network is pairwise stable if no two players have an incentive to add a link between them and no player has an incentive to sever an existing link that connects him to another player. Despite intuitevily appealing and widely used in application pairwise stable networks may fail to exist. In this thesis, we address this problem using three different methodologies. First, we introduce two assumptions on preferences that guarantee existence of pairwise stable networks. The single crossing condition guarantees existence of pairwise stable networks by ruling out improving cycles of networks. Under the link monotonicity condition, existence can be shown using a fixed point argument. Second,we estimate how likely it is for a randomly selected network formation problem to have no pairwise stable networks. We use a probabilistic method where each player independently draws a utility for each possible network from a uniform$(0,1)$ probability distribution. We ask, what is the probability that at least one pairwise stable network exists? We use Monte Carlo simulations to estimate this probability. For $n \leq 10$ the results show that this probability is very close to 1, i.e. the probability of writing a model where no solution exists is very close to zero. This suggest that as long as these results are valid for $n > 10$, the examples where pairwise stable networks fail to exist should not be of practical concern in applications. We then consider network formation problems (NFP) consisting of two elements: a *neighborhood structure (*or *deviation structure)* and a *defeated relation*. A neighborhood of a network is the set of all other networks that differ from the initial network by a deviation rule. A neighborhood structure is the collection of neighborhoods. A defeated relation is a binary relation over the set of all pairs

of networks that are neighbors. A network is said to be *stable* if it defeats every network in its neighborhood. Finally we ask, given a neighborhood structure, if we pick a defeated relation at random, what is the probability that at least one stable network will exist? First, we show that if a neighborhood is defined by differing in only one link, then the probability that at least one stable network exists converges to $1-e^{-1}$ as the number of players in the network grows. Next, we consider the class of *regular* neighborhood structures where all networks have the same number of neighbors. We show that if the size of the neigborhood doesn't grow "too fast", then the probability that there is at least one stable network approaches unity as the number of players grows. We show how apply this ideas for defeated relations on finite sets with the same regularity property.

CHAPTER 1

INTRODUCTION

Networks, i.e. collections of bilateral relationships between objects, permeate our economic, biological and social life. Economically relevant examples include connections by which job opportunities are shared [1], co authorship of scholarly papers [1], and informal insurance agreements [2]. Other examples outside of economics include the internet, neural networks, electrical grids, and movie collaborations by actors(See [**?**,3,4]). Certain common structures have been observed in available network data and to understand why common properties emerge in different contexts we must understand how these networks are formed.

Networks formation models consider a fundamental question of understanding networks. Namely, how are networks formed? Consider a group of economic agents making decisions on how to form bilateral relationships. Suppose that we are given some commonly accepted rules describing how these relationships shall be made. The payoffs that each agent receives depends not only on his set of relationships but potentially on the entire network of bilateral relationships. Suppose also that agents are acting strategically in order to maximize their own payoffs, then agents should weigh the cost and benefits of adding and substracting links when called upon to form such a relationship.

As an example of the type of situations analyzed by strategic network formation, consider a person Ann who receives a friend request from Bob on some social media like facebook or twitter. To decide whether to accept or decline the friend request Ann must weigh the benefits and costs of such decision given the entire structure of the network. She might look at the friends that Bob has on facebook, maybe even look at Bob's friends friends and so on. Let's suppose that Ann accepts Bob's request. The payoffs for everyone can be changed by this decision. For example, Sue might not find it worthwhile to remain a friend of Bob given his new friendhip with Ann.

As an example of the kind of question that motivates network formation theory, consider the so-called "small-world networks." These networks satisfy a set of properties.[1]Well-documented examples include social networks, the internet and actor-movie collaborations. What is the reason that two networks, generated in different complex environments, such as the decision process for hiring actors for movies and the way in which web pages are linked to one another on the internet both exhibit small-world properties? Why is it that, even though the nodes are very different in each of these applications, the ways in which they are related with each other have common properties. Why is the internet, as a network, shaped the way it is? To attempt to answer these questions we must understand the process by which these connections are made and severed.

These example suggest that, at least in some interesting situations the network structure is not fixed and agents in the network make decisions about the network that links them. As for example when two scholar decide to write a paper together or two people decide to become facebook friends.This dynamic aspect of network formation makes it natural to ask what kind of network structures will be stable and which are likely to change. Will there be cycles of networks?

The first attempt to model network structure endogenously were random networks models(See Hofstad [6] and Bollobas [7]). These models assume some probability distribution over the set of all possible networks. As a example consider the erdos-renyi random graph model(See Erdos and Renyi [8,9]) where each link has independently of all other links in the network the same given probability p of being formed. While these models pro-

1. This properties have to do mainly with short average shortest path and relatively high clustering coefficient For further discussion and references on this topic see Watts [5], [4] and Buchanan [3] and references there.

vide useful insights into the way networks are created, they are not useful for economic applications where agents form and sever links strategically.

In a seminal paper, Jackson and Wolinsky [1] introduced a strategic approach to modeling endogenous network formation. Now, links are not added at random but are the strategic decision of the nodes. Typically, additional assumptions about the network formation process are made. For example, we could assume a network formation process where each node declares the set of nodes she wants to be connected with, and two nodes are connected if they both name each other. However, the structure of the network that emerges is sensitive to the details of this process.

To avoid making assumptions about the specific process, Jackson and Wolinsky [1] define a stability notion they call pairwise stability. Intuitively, a network is pairwise stable if no two players have an incentive to add a link between them and no player has an incentive to sever an existing link that connects him to another player. This notion of stability has the advantage of not specifying the details of the network formation game and just focusing on nodes not having an incentive to add or sever one link at a time.

Although widely used in applications,, it is well known that pairwise stable networks may fail to exist (See Jackson and Watts [10]).To gain intuition on why pairwise stable networks might fail to exist we consider the following example.

There are 3 players A, B and C who can form friendship relationships. Players have preferences depend on the whole network structure of friendship. For example A may like to be friends with B only if B is not friend of C and so on. Suppose that players preferences on friendship relations are as follows.

A's preferences are such that A hates C and likes B . Specifically, A prefers any network where he is not a friend of C to a network where he is a friend with C. Moreover, assume A likes B but not as much as he hates C. Thus, A only agrees to form a friendship with B if B is not a friend of C. C likes A more than he dislikes B, so C only agrees to become B's friend if B is a friend of A. B likes both A and C so he is happy to form friendships with both of them no matter what.

If preferences are as above, then no stable networks exist. To see this point, begin with a situation where there are no friendship links. A and B will both agree to become friends. Now suppose that after A and B have become friends, B (who likes both A and C) proposes to form a friendship with C. C agrees since B is already a friend of A.

However, after the relationship between B and C is formed, A (who hates C) unilaterally finishes his relationship with B. Then, since B is no longer a friend of A, C would sever his relationship with B, bringing us back to the situation with no friendships. The existence of this cycle is enough to conclude that there are no stable networks since all other possible networks among A, B, and C contain a link between A and C and hence cannot be stable because A hates C.

The example hence demontrates not only the lack of existence of pairwise stable networks but also the existence of cycles. The relationship between existence of stable networks and cycles has been studied in Jackson and Watts([10]). They show that existence of cycles is a necessary but not sufficient condition for pairwise stable networks to fail to exist.

CHAPTER 2

SUFFICIENT CONDITIONS FOR EXISTENCE OF PAIRWISE STABLE NETWORKS

In this chapter, we introduce the general network formation model that we will consider which is due to Jackson and Woilnsky. We define two conditions on preference profiles: single crossing and link monotonicity. Single crossing is sufficient condition to rule out cycles, which implies it is also a sufficient condition for existence of pairwise stable networks. Link monotonicity is shown to be sufficient for existence using a fixed point argument.

2.1 Network Formation Model

To model network formation in a strategic way, we need to define the costs and benefits that each player faces as a function of the network. There is a set of players N which will be hold fixed throught the paper. The set of all networks with N as the vertex set is denoted G. Given a network $g \in G$ be denote by $k(g)$ the number of links in g. If i and j are nodes the link between i and j is denoted by ij. We denote by $g + ij$ the network that is created by adding to g the link ij and by g_{-i} the network obtained by deleting player i and all his links from the network g.

Each player has preferences over networks represented by the utility function $u_i : G \to \mathbb{R}$. A utility profile is the collection of utility functions for each player $u = (u_1, \cdots, u_n)$.

Let us say that two networks are adjacent if they differ in only one link. That is, g and g' are adjacent if either$g' = g + ij$ for some $ij \notin g$ or $g' = g - ij$ for some $ij \in g$.

A network $g \varepsilon G$ defeats an adjacent network g' if either

- $g' = g + ij$, and $u_i(g + ij) \geq u_i(g)$and $u_j(g + ij) \geq u_j(g)$with at least one strict inequality, or

- $g' = g - ij$, and $u_i(g') > u_i(g)$

A network is pairwise stable if it is not defeated by an adjacent network.

2.1.1 Improving path and improving cycles of networks

An improving path is a sequence of distinct networks $\{g_1, g_2, \ldots, g_K\}$such that each network with $k < K$ is adjacent to and defeated by the subsequent network g_{k+1}. If we also have that $g_1$defeats $g_K$then we have an improving cycle. It is well know that if no improving cylces exists then a pairwise stable network must exist, see [10].

2.2 Single Crossing property

We define a condition on the utility profile called monotonicity. We show that this condition rules out improving cycles and hence is a sufficient condition for the existence of pairwise stable networks. We also provide some economic examples where the condition is satisfied.

Definition 1. A profile of utility functions $u = (u_1, \cdots, u_n)$ satisfies (weak) single crossing if for all $g \in G$, $ij, pk \notin g$ $ij \neq pk$ we have

$$u_t(g + ij) \geqq u_t(g) \Longrightarrow u_t(g + pk + ij) \geqq u_t(g + pk)$$

for $t = i, j$.

We also consider the following stronger version of single crossing which is useful in proving the results below.

Definition 2. A profile of utility functions $u = (u_1, \cdots, u_n)$is said to satisfy strong single crossing if for all $g \in G$, $ij \notin g$, for all g'such that $g \subset g'$we have

$$u_t(g+ij) \geqq u_t(g) \Longrightarrow u_t(g'+ij) \geqq u_t(g')$$

for $t = i, j$.

The difference is that the weak version only requires that if, given a network g ,both i and j would rather add the link between them, then adding a link between other two nodes to g should not reverse i and j preferences.

On the other hand, the strong version requires that if, given a network g ,both i and j would rather add the link between them, then adding any number of links between other nodes to g should not reverse i and j willingness to add the link.

We begin by showing that the two assumptions are indeed equivalent.

Lemma 3. *If a profile of utility function $u = (u_1, \cdots, u_n)$ satisfies weak single crossing then it satisfies strong single crossing, hence both definitions are equivalent.*

Proof. Let u be a utility profile satisfying weak single crossing and assume that $u_t(g+ij) \geqq u_t(g)$for $t = i, j$ then we must show that for all g'such that $g \subset g'$,i.e. for all $g' = g + pk + rs + ... + qt$ we have

$$u_t(g'+ij) \geqq u_t(g')$$

Let $G_1(g) = \{g' \in G(N) : g' = g + e, \text{ for some link } e \notin g\}$By weak single crossing the condition is true for all $g_1 \in G_1(g)$. Define

$$G_2(g) = \{g' \in G : g' = g_1 + e, \text{ for some link } e \notin g_1 \text{and some } g_1 \in G_1(g)\}$$

, since the inequality is true for all $g_1 \in G_1(g)$then by weak SC it is also true for all $g_2 \in G_2(g)$ and so on for g_3 etc... □

Proposition 4. *If a profile of utility functions $u = (u_1, \cdots, u_n)$ induces an improving cycle of graphs thenu violates strong single-crossing.*

Proof. Let $C = g_1, \cdots, g_k$ be an improving cycle. Assume that u satisfies strong single crossing and, without loss of generality, that $g_1$is the network with the least number of links in the cycle C. Notice that g_1 must be unique. Then we can write $g_2 = g_1 + ij$ for some ij. Since C is a cycle there is at least one other place along the cycle where the link ij is substracted. If u satisfies strong single crossing it must be that the network g_t where ij is severed has less than or equal number of links than $g_1$which is a contradiction to $g_1$having the least number of links in C. □

Proposition implies that a pairwise stable network will always exist under the single crossing condition.

2.3 Examples

Next we consider some examples of utility profiles that satisfy the single crossing property and discuss the condition. Within this framework, introduced by Goyal and Roshi [11] we also discuss an economic application that satisfies this set of assumptions.

2.3.1 Playing the field

Consider network formation games in which the marginal returns from links for every player can be expressed in terms of the number of links of the player and the aggregate number of links of the rest of the players(Goyal and Roshi). Recall that g_{-i} is the network obtained by deleting player i and all his links from the network g and $L(g_{-i}) = \sum_{j \neq i} n_j(g)$. Thus the marginal gross returns of player i satisfy the playing the field property if for any network $g \in G(N)$ and any additional link $ij \notin g$

$$u_i(g + ij) - u_i(g) = \Phi(n_i(g), L_i(g)) - c$$

2.3.2 *Local Spillovers*

Consider a class of games in which the marginal gross returns of a player from a link are a function only of the number of links of the player and that of the potential partner. Formally, the payoffs of player i satisfy local spillovers if for any network $g \in G(N)$ and any additional link $ij \notin g$ we have

$$u_i(g+ij) - u_i(g) = \Psi\left(n_i(g), n_j(g)\right) - c$$

Next we show that both playing the field and local spillover games satisfy single crossing provided that marginal payoffs of adding a link funtions are increasing in both arguments.

Lemma 5. *local spillovers satisfies single crossing if and only if* $\Psi\left(n_i(g), n_j(g)\right)$ the marginal payoff for i of adding a link with j is increasing in both $n_i(g)$and $n_j(g)$

Proof. Assume that $u_i(g+ij) - u_i(g) > 0$. Which basically means that $\Psi\left(n_i(g), n_j(g)\right) > c$. We must show that for all $pk \notin g$ it holds that

$$u_i(g+pk+ij) - u_i(g+pk) > 0$$

Since obviously we have that $n_i(g+pk) \geq n_i((g))$ and $n_j(g+pk) \geq n_j((g))$ and Ψ is increasing in both arguments we have

$$u_i(g+pk+ij) - u_i(g+pk) \geq u_i(g+ij) - u_i(g) > 0$$

□

Next we consider an economic application suggested by [11].

2.3.3 *Provision of a pure Public good(Goyal and Joshi).*

There are n players, each of whom is deciding on a level of output, x_i, to produce of a pure public good. Given each player's output, the utility of player i is: $u(x) = x_i + \sum_{j \neq i} x_j$. A collaboration link between two players is an agreement to share knowledge about the production of the public good. Let $c > 0$ be the fixed investment required from each player in such a link. In any network g the cost of producing output x_i is given by, $f(x_i, g) = \frac{1}{2}\left(\frac{x_i}{n_i(g)}\right)$, where we set $n_i(g) = \eta_i(g) + 1$and $\eta_i(g)$ denotes the number of links that i has under network g. Given a network g, player i will choose output to maximize utility net of production costs. This yields an optimal output of $x_i(g) = n_i(g)^2$. Therefore, the reduced form gross payoff of player i is:

$$u_i(g) = \frac{1}{2} n_i(g)^2 + \sum_{j \neq i} n_j(g)^2 - c n_i(g)$$

Hence it follows that the public goods example satisfies local spillovers, moreover both own links and potential partner number of links increase the gross benefit of forming the link.

2.4 Link Monotonicity condition

The condition can be expalined intuitevily as assuming that for each possible link e there is one natural number k_e such that to add e is profitable for both players involved in e if and only if the network g has more than k_e links. Hence we call the property link monotonicity since it implies that once a pair of players finds it profitable to add a link when the network has k links then they will continue to do so for all networks with more than k links.

Remember that $k(g)$denotes the number of links in g.

Definition 6. A profile of utilities$u$ is said to be link monotonic if it satisfies that for each possible link ij there exist an integer $0 \leq k_{ij} \leq M$ where $M = N(N-1)/2$ such that for all $g \in G$ we have:

1. for all $ij \in g$: $k(g) > k_{ij} \Longleftrightarrow g - ij \prec g$

2. for all $ij \notin g$: $k(g) > k_{ij} \Longleftrightarrow g \prec g + ij$

We assume throughout this section that preferences satisfy link monotonicity. Under this assumption we show that pairwise stable networks exist by using a fixed point theorem argument.

We define the function for $0 \leq k \leq M$ given by $f(k) = \left|\left\{ij : k_{ij} \leq k\right\}\right|$, i.e. $f(k)$ is the number of links such that $k_{ij} \leq k$ and hence for any network with k links would be prefereed to be added by both players involved.

Next we show that under link monotonicity pairwise stable networks exist is anf only if f has a fixed point.

Lemma 7. *If a pairwise stable network exists then there exist a* $0 \leq k^* \leq M$ *such that* $f(k^*) = k^*$.

Proof. Let g^* be a pairwise stable network and suppose g^*has k^* links $0 < k^* < M2^{M-1}$. Then, since $g^*$is pairwise stable we have for each $ij \in g^*$ $g^* - ij \prec g^*$. It follows by link monotonicity that $k_{ij} \leq k^*$. Hence if a network g^* is pws with k^* links then the set $\left\{ij : k_{ij} \leq k^*\right\}$has k^* elements and so $f(k^*) = k^*$ □

Lemma 8. *If there exist a* $0 \leq k^* \leq M$ *such that* $f(k^*) = k^*$*then there is at least one pairwise stable network*

Proof. Let k^*be a fixed point of f and consider the the network $g^*$where the links active in g^* are those in the set$\left\{ij : k_{ij} \leq k^*\right\}$. By definition of k^*such a network would have

k^*links. Therefore we have that for all $ij \in g^*$ $k_{ij} \leq k^*$and by link monotonicity this implies $g^* - ij \prec g^*$ and for all $ij \notin g^*$ we have $k_{ij} > k^*$ which again by link monotonicity implies $g^* + ij \prec g^*$. Therefore g^* is pairwise stable. □

The consequence of the two previous lemmas is that under link monotonicity proving that at least one pairwise stable network exists is equivalent to proving that $f(k)$ has a fixed point.

We use the following fixed point theorem by Knaster-Tarski (See Knaster [12] and Tarski [13]).

Theorem 9. *Let L be a complete lattice and let $g : L \to L$ be an order-preserving function. Then the set of fixed points of g in L is also a complete lattice.*

Using the Knaster-Tarski theorem we can show the the fixed point exist by showing the following remark is true.

Remark 10. Let the function $f : K \to K$ be defined as above, then f satisfies the hypothesis of the theorem.

Proof. The domain K of f is given by the set of integers k such that $0 \leq k \leq M$ which is a complete lattice. f is order preserving since by definition we have that for $k_1 < k_2$ it must be that $\{ij : k_{ij} \leq k_1\} \subseteq \{ij : k_{ij} \leq k_2\}$ and hence the second set has weakly more elements. □

The results are summarized in the following theorem.

Theorem 11. *Let u be a profile of utility functions satisfing link-monotonicity, then a pairwise stable equilibrium exists.*

2.5 Conclusion

This chapter is a contribution to the literature that looks at the problem of lack of existence of pairwise stable networks. We define two sufficient conditions for existence of pairwise stable networks. The weak (strong) single crossing condition requires that once a pair of players i and j finds it profitable to add a link between each other given a network g, then if we add one (many) link(s) to g between any two other players, this will not reverse the fact that i and j would rather add the link between each other.

Link monotonicity states that preferences are such that for each possible pair of nodes i, j there exist a natural number k_{ij} such that given a network g with $k > k_{ij}$ then if $ij \in g$ both i and j prefer the network g to $g - ij$ and if $ij \notin g$ then both i and j would rather add the link between each other. Using the observation that any pairwise stable network whith more than k_{ij} links must contain a link between i and j, we construct a function such that (under link monotonicity) a pairwise stable network exists if and only if the function has a fixed point.

We also provide some economic applications discussed in the literature that satisfy the single crossing condition.

One direction for future work includes finding more general conditions on preferences that are sufficient for the existence of at least one pairwisse stable network.

CHAPTER 3

MONTE CARLO ESTIMATION OF PROBABILITY OF EXISTENCE OF PAIRWISE STABLE NETWORKS

In this chapter we define a simple model to generate random utilities and use montecarlo simulations to measure the probability that at least one pairwise stable network exist. Our results show that for $3 \leq N \leq 10$ if we assign utilities to networks at random based on iid draws of the same continuous distribution then the probability that for a given realization stable networks fail to exist is statiscally non distinguishable from 0. This supports the use of pairwise stability in applications.

3.1 Random Utilities Model

We are interested in providing a measure of how likely it is that we find a utility profile such that no pairwise stable network exists. To do so we define a random utility model. Once the different utilities that different players obtain from different networks become random variables, we have a well-defined probability that at least one pairwise stable network exists.

The random utility model we consider is very simple and is the same as the one used in the literature on finding the probability that a game with random payoffs contains a pure strategy Nash equilibrium (See Dresher [14], Papavassilopoulos [15] and William and Stanford [16]).

For each player $i \in N$ and each network $g \in G(N)$ the utility $u_i(g)$ is a random variable with probability distribution f. We assume that f is the same for all utilities in the model and hence utilities are identically distributed. We also assume that the random variable $u_i(g)$ is independent from all other random variables in the model. For the purpose of our simulations we will assume that f is the density function of a $Uniform(0,1)$ random variable.

For each pair of adjacent networks g and $g' = g + ij$, the probability that g' defeats g is given by

$$Pr\left(\left\{u_i\left(g'\right) \geq u_i\left(g\right)\right\} \cap \left\{u_j\left(g'\right) \geq u_j\left(g\right)\right\}\right)$$

Using the identically distributed assumption, we have that

$$Pr\left(\left\{u_i\left(g'\right) \geq u_i\left(g\right)\right\}\right) = Pr\left\{u_j\left(g'\right) \geq u_j\left(g\right)\right\}$$

and since the distribution is continuous we have $Pr\left(\left\{u_i\left(g'\right) = u_i\left(g\right)\right\}\right) = 0$ which implies

$$Pr\left(\left\{u_i\left(g'\right) \geq u_i\left(g\right)\right\}\right) = \frac{1}{2}$$

and, using the independence assumption, we can write the probability that g' defeats g as

$$Pr\left(\left\{u_j\left(g'\right) \geq u_j\left(g\right)\right\}\right)^2 = \frac{1}{4}$$

From which it follows that the probability for any network with k links on n nodes to be stable is given by

$$\left(\frac{3}{4}\right)^{m-k}\left(\frac{1}{4}\right)^{k} = \left(\frac{3}{4}\right)^{m}\left(\frac{1}{3}\right)^{k}$$

where $m = n\left(n-1\right)/2$ is the number of possible links on n nodes.

Let us remark that the probability that a given network is stable under our model is a decreasing function on the number of links on the network, the most likely network to be stable is the empty network, and the least likely is the complete network. This asymmetry follows from the requirement of pairwise stability that mutual consent to add a link is required but severing a link can be done unilaterally. This partially accounts for the difficulty in obtaining the exact probability that at least one stable network exist for arbitrary number of nodes N.

To gain intuition on why this is the case let us define the random variable x defined as the number of stable networks and consider its distribution. For the remainder of the paper, all dependence of the random variables and sets on N the number of nodes on the networks considered will not be explicitly noted.

Let Γ_k denote the set of all networks with k links and let x_k denote the random variable number of stable networks with k links. We can write $x = \sum_{k=0}^{m} x_k$ and hence the distribution of x is given by the convolution of the x_k's. Moreover, given our independence assumption, we can easily establish that each of the x_k has a binomial distribution with $N_k = \binom{m}{2k}$ trials and probability of success given by $p_k = \left(\frac{1}{4}\right)^m (3)^k$.

It follows that the distribution of x is the distribution of the sum of binomial distributions with different number of trials and probabilities of success. The distribution of the sum of binomial distributions with different parameters was studied in Butler(ref [17]). The author gives algorithms for drawing from such a distribution and discusses different methods for approximation. However, there is no well-known closed-form solution for such a random variables distribution.

The difficulty of analytically solving the random utility model can be clarified by considering the event that a given network g is stable $\{g$ is stable$\}$. Such an event is the intersection, for all neighboring networks g' of g, of the events $\{g$ dominates $g'\}$. Consider two of such events that have a player in common, i.e., consider events $\{g$ dominates $g+ij\}$ and $\{g$ dominates $g+ik\}$. Two such events will be correlated because if i likes g more than $g+ij$, then it is more likely that i likes g also more than $g+ik$. Formally we have that $P\left(u_i^g > u_i^{g+ij} \mid u_i^g > u_i^{g+ik}\right) \neq P\left(u_i^g > u_i^{g+ij}\right)$ and hence the events are correlated. The fact that any of such two events are correlated makes closed-form solutions difficult.

3.2 Simulations

To estimate the probability that at least one stable network exists, we use Monte Carlo simulations. The reason for using simulations is, as sketched in the previous section, that there are no well-known closed-form solutions for the distribution of the number of stable networks. The results presented assume uniform$U(0,1)$. The basic mechanics of the simulations involve two steps. In the first step, we generate a realization of utilities drawing from the uniform iid model. In the second step, given the utilities for each player for each network, we check whether a pairwise stable network exists or not.

We repeat this process many times and compute the frequency of realizations (or iterations) in which we obtained at least one pairwise stable network and use that as an estimate of the probability that at least one pairwise stable network exists. We also compute (whenever possible) the standard error of the estimates.

The simulations were implemented using Go, and the code is presented in the Appendix.

3.2.1 Results

The results are presented in the following table. The first column has the number of nodes in the networks, the second has the number of iterations, and the third has how many of those iterations were such that the code found a stable network. The estimated probability that at least one stable network exists on the given number of nodes can be computed by dividing the third by the second column.Standard errors are included in parenthesis whenever possible.

Number of Players/Nodes	# of simulations	# of simulations where stable found	Standard Error
N=3	1Million	932507	0.00002
N=4	100000	94120	0.00007
N=5	100000	99804	0.00003
N=6	10000	10000	*
N=7	10000	10000	*
N=8	10000	10000	*
N=9	1000	1000	*
N=10	100	100	*

3.3 Conclusion

This paper contributes to the literature that looks at the problem of lack of existence of pairwise stable networks. We consider a random utility model where the utility for each player for each possible network is indepndently drawn from the Uniform(0,1) distribution. In the context of this models we ask what is the probability that at least one pairwise stable network exists. We use simulations to obtain estimates of this probability for networks with 3 throught 10 players.

The results of the simulations show that for the range of parameters considered finding a utility profile such that no pairwise stable networks exist is hard . For networks with more than 6 players the code could not find any utility profile such that no pairwise stable network exists. For networks with 3, 4 and 5 players the estimates of the probability are very close to one, although statistically significantly different from one.

These results provide support for the use of pairwise stability in strategic network formation applications. Future work would try to generalize the results in this paper to networks of bigger size.

CHAPTER 4

PROBABILITY OF EXISTENCE OF STABLE NETWORKS

As we have seen, examples where stable networks fail to exist are not hard to come by. In this chapter, we consider more general stability solutions and leave pairwise stability behind. Any definition of stability must include a notion of neighborhood among networks, i.e., it has to specify a list other networks g' a given network g shall be compared with, when trying to determine if g is stable. In addition to the neighborhood any stability notion must sepcify a criterion by which networks defeat each other. We call this criterion a defeated relation defined for pairs of neighboring networks. A network is stable if it defeats all of it's neighboring networks.

Given a neighborhood structure, we consider a simple model to generate a defeated relation at random. For each pair of neighboring networks, we toss a fair coin to determine which network defeats the other. Moreover, we assume that the toin coss are independent from each other.

We ask, given a neighborhood structure among networks, what is the probability that if we select one defeated relation at random according to this process, then there exists at least one network that defeats all of it's neighbors? In other words, what is the probability that at least one stable network exists? We also study how this probability depends on the rule we select to determine which networks are comparable.

In this chapter, first we analyze the one link deviation structure and show that the problem of lack of existence doesn't go away as the network size rows. Then we show that this result is not robust if we allow for more general deviation structures. We consider regular deviation structures where each network is compared with same number of neighboring networks and show that if the number of networks doesn't grow too fast, under our random model, the probability that at least one stable network exists converges to one as the size of the network approaches infinity.

Then we consider applying the same tools developed to consider a situation where the elements we consider as outcomes are not networks and characterize the probability of existence of stable outcomes and show that this can be used as an alternative proof to results in Papavassilopoulos [15] to find the probability of existence of pure strategy nash equilibria in strategic form games where utilities are defined randomly.

4.1 Reduced Form Approach to Network Formation

We begin by introducing the model we will consider in this chapter. We call it *reduced form approach to network formation problems* and define the notion of stability that we will consider. This reduced form approach is defined by two basic ingredients:

First, a criterion of what networks are feasible deviations from each other, i.e. for each given network g, what are the other networks that according to the criterion are to be compared with g. We call this set of comparable networks the "neighborhood" of g and write $N(g)$. We assume that the neighborhood criterion is symmetric, i.e. if g' is a neighbor of g then g is a neighbor of g'.

Formally, we summarize all the neighborhoods into one "deviation graph" Q. The graph Q is defined as a graph whose vertex set is the set of all networks on n nodes and there is an edge between two networks g and g' if $g'$is a neighbor of g.

For example, suppose that we consider a network formation problem where the deviation rule only allows players to add or subtract one link at a time. Then the neighborhood of any network g on n nodes is composed of $m = n(n-1)/2$ other networks of the form $g \pm \text{link} ij$ for some nodes$i,\ j \in N$, and it can be shown that Q , the deviation graph, is isomorphic to the m-dimensional hypercube. On the other hand, if the deviation rule allows any number of links to be added or subtracted at the same time, then Q is the complete graph on 2^m nodes.

The second ingredient of the reduced form approach to network formation is a criterion by which to compare a network g to its neighbors in $N(g)$. We call this criterion a defeated relation (see [18]) and assume that is satisfies that for all networks g' in the neighborhood of g either "g' defeats g" or "g defeats g'". Then we define a network to be *stable* if it defeats all of its neighbors.

Formally, let G_n denote the set of all networks over n players. Define a *network formation problem* as a couple $(Q, \prec)$ where:

- Q is a "deviation graph" with vertex set G_n. Two networks g and g' are linked or neighbors in Q written as gQg', if they are feasible deviations from one another. Let $N(g) = \{g' \in G_n : gQg'\}$ denote the neighborhood of g.
- $\prec$ is a defeated relation over the set G_n such that for all $g, g' \in G_n$ such that gQg' either $g \prec g'$ or $g' \prec g$.

A network is said to be stable with respect to problem $(Q, \prec)$ if it defeats all of its neighbors. Formally, a network g is *stable* relative to NFP $(Q, \prec)$, if we have $g' \prec g \ \ , \forall g' \in N(g)$.

1

Next we consider some examples of how this defeated relation is derived. Many applications use a utility-based approach. This approach assumes each player has preferences over networks represented by the utility function $u_i : G_n \to \mathbb{R}$ and derives the defeated relation using some form of preference aggregation function that maps the individual utilities into a defeated relation among networks.

---

1. One can think of 1) and 2) together as an orientation $Q^{\prec}$ obtained by defining that the link between g and g' points from g into g' if and only if g' defeats g under $\prec$. This interpretation leads to an alternative definition of stability: we can say that a network is stable if it is a sink of $Q^{\prec}$.

An example of a utility-based stability theory is the concept of pairwise stability (see [1,18]). Let us say that two networks are neighbors if they differ by only one link. A network g' defeats a neighbor g if either

- $g' = g - ij$ and $u_i\left(g'\right) > u_i\left(g\right)$,or
- $g' = g + ij$ and $u_i\left(g'\right) \geq u_i\left(g\right)$ and $u_j\left(g'\right) \geq u_j\left(g\right)$, with at least one inequality holding strictly.

A network is pairwise stable if and only if it is not defeated by a neighboring network.

Another example of how one could aggregate preferences to define a defeated relation is by using the same notion of neighborhood and saying a network g' defeats a neighbor g if

$$u_i\left(g'\right) + u_j\left(g'\right) \geqq u_i\left(g\right) + u_j\left(g\right)$$

Another way to construct the defeated relation is to define g' defeats a neighbor g if

$$\sum_{i \in N} u_i\left(g'\right) \geqq \sum_{i \in N} u_i\left(g\right)$$

this notion was defined in Jackson and Wollinsky [1] as 'strong efficiency'.

4.2 Random Defeated Relation over Networks

To provide a measure of how likely it is for stable networks to exist we follow the probabilitic approach. We define some random model by which instances of the problem are drawn and then compute the probability that at least one stable network will exist. Even if preferences of economic application are not generated at random this probability can be interpreted as the fraction of total instances such that stable networks exist and hence provide some measure of how likely it is to run into this problem in practice.

In this section we introduce the random model that we will use in the rest of the paper. As we discussed in the previous section, network formation theories are sometimes constructed by defining a utility function over the set of all networks for each agent in the network and then defining some aggregation rule that maps utility profile to a defeated relation over the set of all networks.To generate random instances we could then use a random model where the utility functions are chosen at random in a similar way as it is done in other literatures [14–16,19,20]. This approach has the advantage of being used by other literatures and of being intuitive.

Notice that in such a model, given a realtization of utilities we can construct through the aggregation rule a defeated relation on the set of networks. Hence each random utility model will induce a random defeated relation model where we draw at random from the set of all defeated relations. Therefore, we could work directly with random models definned at the level of the defeated relation.

In this case we don't need to assume anything about the aggregation rule that maps utility profiles into defeated relations. Also this model can be thought as including theories that are not generated by utilities. Moreover, as shown in the following sections, under some (mainly independence) assumptions.the defeated relation is tractable and assymptotic results are possible.

. We define the uniform random being defeated relation as follows: hold the possible deviations graph fixed and chose from all possible theories at random uniformly. Formally, define the set of all possible defeated relations over Q and assign to each of the $m2^{m-1}$ possible defeated relations a probability of being drawn given by $\frac{1}{m2^{m-1}}$. This model is equivalent to assuming that for each possible pair of neighboring networks $g, g' \in G_N$ we independently toss a fair coin and if the coin lands heads then g defeats g' and if the coin lands tails then g' defeats g.

Then, we have that independently of all other comparisons of networks

$$P\left(g \prec g'\right) = P\left(g' \prec g\right) = 1/2$$

This shows that probabilities to the events $\left\{g \text{ defeats } g'\right\}$ and hence the probability that a stable network exists are well defined.

Formally, let $E_g$denote the event that g is stable, i.e. the event where g dominates all of it's neighbors .

$$E_g = \bigcap_{g' \in N(g)} \left\{g' \prec g\right\}$$

We are interested in finding the probability that there exists at least one stable graphs which we denote P_s. Using our notation, we can write

$$P_s = Pr\left(\bigcup_{g \varepsilon G_N} \bigcap_{g' \in N(g)} \left\{g' \prec g\right\}\right)$$

4.3 Probability of Existence of Stable Network

In this section we give some answers to the following question. Given a deviation structure Q if we pick a being defeated relation at random according to the model introduced in the previous section what is the probability that a stable network exists? As argued before, one might think of this problem as the problem of counting the orientations of Q with no sink, a problem studied in Bubley and Dyer [21].

First, we provide a general characterization of the probability of existence in terms of a graph theoretic properties of the deviations graph Q . We show that the probability that at least one stable network exists is completely determined by the independence sequence of Q. Next, we provide asymptotic characterizations of the probability for two different definitions of feasible deviations.

4.3.1 Probability of Existence and Independent Sets of Q

This section follows closely the arguments presented in Bubley and Dyer [21]. Consider the model of random defeated relations defined in the previous section. Let E_g denote the event that g is stable, i.e. it dominates all of it's neighbors. We are interested in finding the probability that there exists at least one stable network which we denote P_s, i.e.

$$P_s = Pr\left(\cup_{g \varepsilon G_N} E_g\right)$$

Similarly, for a set of networks G let E_G denote the event that all networks in G are stable simultaneously.

Using the principle of inclusion-exclusion (See Riordan [22]) we can write the probability of the union as follows

$$P_s = \sum_{k=1}^{2^m} (-1)^{k-1} \sum_{\{G \subset G_N : |G| = k\}} Pr(E_G)$$

where $m = n(n-1)/2$ is the number of possible links on a network with n agents.

Let us say that a set of networks forms an independent set if all networks in the set are not neighbors with each other. The next Lemma provides a characterization of the set of networks that can be simultaneously stable.

Lemma 12. *Given a deviation graph Q and set of networks G, there exist a being defeated relation $\prec$ such that the set of stable networks for instance $(Q, \prec)$is G if and only if G is an independent set of Q.*

Proof. This follows directly from the assumption that either g defeats g' or g' defeats g but not both. $\square$

Let I_k denote the cardinality of the set of all independent sets of networks with k members in Q, the greatest integer α such that $I_\alpha > 0$ is called the stability number of Q. The sequence of integers $(I_k)_{k=1}^{\alpha}$ is the independence sequence of Q.

The independence of the events $g \prec g'$ and $g \prec g''$ and the observation that two neighbors can't be stable at the same time gives us that $Pr(E_G) = 0$ if G is not an independent set and $Pr(E_G) = \left(\frac{1}{2}\right)^{mk}$ if $G \in I_k$so we can use Lemma 1 and write:

$$P_s = \sum_{k=1}^{2^m} (-1)^{k-1} I_k \left(\frac{1}{2}\right)^{mk}$$

This expression allows us to compute the probability of existence of a stable network given that we know the independence sequence of the deviations graph Q.

To fix the ideas, we will show how to apply the formula to an example where $n = 3$ and only networks that differ by one link are neighbors. In that case $m = 3$ and there are $2^3 = 8$ different networks with 3 nodes, so $I_1 = 8$. It is not hard to see that $I_2 = 16$, $I_3 = 8$ and $I_4 = 2$and $I_k = 0$ for $k = 5,6,7,8$. Therefore the probability that a stable networks exists is given by

$$P_{pws} = 1 - 16\frac{1}{2^6} + 8\frac{1}{2^9} - 2\frac{1}{2^{12}}$$

In general, it is a hard problem to find all the values of the $I_k$for $n > 3$ so we will find upper and lower bounds for I_k and then let the size of the network grow to give a asymptotic results.

Another application of the formula is when we allow all networks to be neighbors with each other in which case Q is the complete graph. In this case the only independent sets are singletons and the probability $P_s$would be

$$P_s = 2^m \left(\frac{1}{2}\right)^{2^m - 1}$$

By taking logs we get that

$$log_2(P_s) = m + 1 - 2^m$$

from which it follows that as n goes to infinity $P_s$will converge to 0.

4.3.2 Probability of Existence for One Link Deviations

Consider a deviation structure where only one link at a time deviations are allowed, i.e. define two networks g and g' to be neighbors if there exists a pair $i, j \in N$ such that $g' = g + ij$. The deviations graph associated with this definition is denoted Q_m. It is easy to see that Q_m is isomorphic to the m-hypercube (hence the notation) and hence the independence sequence is not known for arbitrary n (See Galvin [23]).

Suppose we pick a defeated relation among neighboring networks at random using the model described before, and let $P_n$denote the probability that at least one stable network among n players exists . As argued in the previous section using an inclusion-exclusion argument and letting $m_n = \frac{n(n-1)}{2}$we can write

$$P_n = \sum_{k=1}^{2^{m_n}} (-1)^{k-1} I_k^n \left(\frac{1}{2}\right)^{m_n k}$$

We show that the probability converges converges to $1 - e^{-1}$ as the number of players goes to infinity.

Theorem 13. *When only one link at a time deviations are allowed the probability that there is at least one stable network P_n converges to $1 - \frac{1}{e}$ as the number of players n goes to infinity.*

The proof is contained in the appendix. Whether this is large or small is left to the reader to decide but it shows that at least in the context of the random defeated relation model the problem of existence does not go away as the size of the network grows.

4.3.3 Probability of Existence for Regular Deviations

We have just showed that in the one link deviations model $0 < \lim_{n\to\infty} P_n < 1$ and in this section we introduce a more general model to explore how robust this property is to changes in the deviation structure. We find that this property is in some sense unique to the one link deviation structure and that most other deviation structures the probability of at least one stable networks converges either to 1 or to 0.

The model we consider allows for a more general type of deviations where we only impose that every networks is comparable to the same number of networks E_n. The behavior of the integer $E_n$as n grows large will be the most important determinant of the asymptotic properties of P_n.

Notice this assumption is indeed more general than the the one in the previous section since the one link deviation satisfies the assumption with $E_n = m_n$ where(as before) $m_n = n(n-1)/2$ is the number of possible links among n players.

We get the following result

Theorem 14. *Assume that $lim_{n\to\infty}\frac{E_n}{2^{m_n}} = 0$ and that E_n is well behaved so that $lim_{n\to\infty}m_n - E_n$ exists, then one of three cases hold:*

1. There exist an N, such that $n > N$ implies $E_n = m_n + a$ in which case $lim_{n\to\infty}P_n = 1 - e^{-2^{-a}}$
2. $lim_{n\to\infty}m_n - E_n = \infty$ in which case $lim_{n\to\infty}P_n = 1$.
3. $lim_{n\to\infty}m_n - E_n = -\infty$ in which case $lim_{n\to\infty}P_n = 0$.

For a proof see the appendix.

Intuitively, the theorem shows that there is a threshold function m_n such that if the number of comparable networks grows faster than this function then the probability of existence of a stable network will converge to zero. If the number of comparable networks grows slower than m_n then the probability will converge to one and if it converges at assymptotically the same rate, then we will have the probability converging to a number between zero and one.

The theorem can be interpreted as saying that as long as you don't allow for "too many" feasible deviations then the lack of existence of stable networks is a problem only for small networks, i.e. for networks with a small number of nodes. As the number of nodes in the network grows large, then the chance of ending up with a defeated relation with no stable networks becomes negligible.

On the other hand, if your network formation theory allows for "too many" feasible deviations then as the number of nodes in the network grows you will most certainly run into situations where stable networks fail to exist.

4.4 Stability for finite sets

In this section we introduce a similar notion of stability for finite sets. The elements of this set will be referred to as outcomes. Let V_n denote the set of all outcomes indexed by n. Define a Outcome Choice Problem as a couple $(Q, \prec)$ where:

- Q is a "deviation graph"with vertex set V_n. Two outcomes a and a' are linked or neighbors in Q written as *aQa*', if they are comparable in the theory and let $N(a) = \{a' \in V_n : aQa'\}$ denote the neighborhood of outcome$a$.

- $\prec$ is a defeated relation over the set V_n such that for all $a, a' \in V_n$ such that aQa' we have either $a \prec a'$ or $a' \prec a$.

An outcome is said to be stable with respect to problem $(Q, \prec)$ if it defeats all of its neighbors. Formally, an outcome a is *stable* relative to NFP $(Q, \prec)$, if we have $a' \prec a$,$\forall a' \in N(a)$. Consider as before a finite set V_n of possible outcomes and a adjacency relation on V_n defining which pair of alternatives are comparable. Let $Q_n$denote the "deviations graph" indexed by n.

Assume that the size of the neighborhood of every outcome is E_n, i.e. $E_n = |N(a)| = |N(a')|$ for all $a, a' \varepsilon V_n$. Under this assumption, all outcomes have the same number of alternative outcomes to be compared with and there are no outcomes that are compared with more alternatives than others. In the particular case where outcomes are networks studied above, this symmetry assumption is satisfied by the one link deviation adjacency but it is not satisfied if we allow a player to delete as many links as he wants unilaterally[explain this better] as in Nash stability(for a definition see [18]).

Consider the random orientation model where every possible orientation of $Q_n$is given equal probability of being chosen. Denote by $P_n$the probability that the chosen orientation has at least one sink. We have shown by inclusion exclusion argument that $P_n = \sum_{k=1}^{V_n} (-1)^{k-1} \frac{I_k^n}{2^{kE_n}}$, where I_k^n is the number of independent sets of vertices of $Q_n$with k elements. The main result of this section is the following theorem.

Theorem 15. *Assumeme that* $lim_{n\to\infty} \frac{V_n}{2^{E_n}} = c$ *and that* $lim_{n\to\infty} \frac{E_n}{V_n} = 0$ *then* $lim_{n\to\infty} P_n = 1 - e^{-c}$

The proof of the theorem is included in the appendix.

4.4.1 Example: Prob of Existence of pure strategy Nash equilibria.

In this section we show how to apply the theorem to compute the probability that at random game as defined in for example [15] contains at least one pure strategy Nash equilibrium. As mentioned above this problem has already been analyzed elsewhere and we intend this as an example and alternative proof of the existing literature on the topic.

Suppose we have a game with N agents and each agent has p actions available. Consider the random utility model where each $u_i(a)$is independent(from all other RV in the model) identically distributed random variable with a continuous probability distribution given by f. We want to find the probability that randomly selected game has a pure strategy Nash equilibrium.

We proceed by defining a graph $G_N$with vertex set given by the strategy profiles of the game and say that two strategy profiles are adjacent if they differ in only one coordinate, i.e. strategy profiles a and a' are adjacent if there exists some player i such that we can write $a' = \left(a'_i, a_{-i}\right)$. Then we define an orientation by saying that for two adjacent strategy profiles a and a', a dominates a' if $u_i\left(a_i, a_{-i}\right) \geq u_i\left(a'_i, a_{-i}\right)$ it follows from the definitions that a strategy profile is a pure strategy Nash equilibrium if and only if it is a sink of this orientation. Moreover, it follows from the independence assumption that for any pair of adjacent strategy profiles a and a' we have that $Pr\left(a \to a'\right) = Pr\left(a' \to a\right) = 1/2$ independently of all other pair of adjacent strategies and hence the random orientation model induced by the random utility model is equivalent to the random defeated relation considered above.

Therefore, since there are p^Ndifferent strategy profiles and each profile is adjacent to $N\left(p-1\right)$ other profiles we can apply the theorem with $V_n = p^N$ and $E_N = \left(p-1\right)^N$and so the probability that at least on pure strategy Nash equilibrium exists will be determined by $lim_{n\to\infty}\frac{p^N}{2^{N(p-1)}}$ which equals 1 if $p = 2$ and equals zero for $p \geq 3$. Hence we obtain the following theorem:

Theorem 16. *In the random game model with N players and each player has p actions let P_N denote the probability that at least one pure strategy Nash equilibrium exists. Then we have that $lim_{n\to\infty}P_n = 1 - e^{-0} = 0$ for $p \geq 3$ and $lim_{n\to\infty}P_n = 1 - e^{-1}$* for $p = 2$.

4.5 Conclusion

We define a broad class of Networks Formation Problems and consider the solution concept of stability. In general stable networks could fail to exists and we provide a measure of how likely it is that there is at least one stable network. We consider the class of network formation theories defined by two basic components: the notion of which networks are feasible deviations from each other and a criterion by which to rank comparable networks.

We hold the definition of comparable networks fixed and consider a simple random model to create a random defeated relation among networks. This allows us to compute the probability that at least one stable network exists. We show that if we consider theories were only adding or subtracting one link at a time is valid then if we pick a theory at random that probability converges to $1-e^{-1}$as the number of players grows to infinity. Whether this number is large or small is left up to the reader.

Next, we focus on how sensitive our result is to changes in the definition of comparable networks. We assume that all networks have the same number of neighbors and show that unless the number of comparable networks is asymptotically equal (plus a constant) to the number of comparable networks in the one link deviation the probability will converge either to zero or one. In this sense the result for one link deviations is rather sensitive to changes in the definition of comparable networks.

Finally, we consider abstract finite sets were the elements are not necessarily networks and derive asymptotic formulas for the probability that at least one element of the set defeats all of its neighbors. As an application we show that this result can be applied to the problem of finding the probability that at least one pure strategy equilibrium exists considered in [15].

Further research would be to answer similar question for the notion of pairwise stability defining the random process at the level of the utility functions and to apply the results in the paper to matching problems.

4.6 Appendix 1: Proof of results one link deviations

In this appendix we prove the following theorem.

Theorem 17. *When only one link at a time deviations are allowed, the probability that there is at least one stable network P_n ,converges to $1-\frac{1}{e}$ as the number of players in the network n goes to infinity.*

We begin with a technical lemma about the independent sequence of the m-dimensional hypercube.

Lemma 18. *Let I_k^m denote the number of independent sets of size k in a m-dimensional hypercube then we get that:*

$0 \leqq \frac{1}{k!} - \frac{I_k^m}{2^{km}} \leq \frac{m+1}{2^{m+1}(k-2)!}$ *for* $2 \leq k \leq \frac{2^m}{m+1} + 1$

Proof. First, to construct a lower bound for I_k^m we define a procedure to find an independent set of size k. Pick any network g, this network will have m neighbors, so to find the second network of the set we are free to chose from any of the $2^m - m - 1$ remaining networks. Among this we will assume we pick one that has no neighbors in common with g. Since the two networks have no neighbors in common the size of their neighborhoods must be $2m$ which gives us $2^m - 2(m-1)$available choices to pick the third network. Notice that if the neighborhoods overlap we actually have more choices for the third network and hence the estimate of I_k^mwe get out of this procedure is indeed a lower bound. Counting all possible choices gives

$$\prod_{j=1}^{k} 2^m - j(m+1)$$

To obtain our lower bound we divide by the $k!$ different ways in which we end up with the same set, i.e. disregarding the order in which we construct the set and we get that for

$2 \leq k \leq \frac{2^m}{m+1} + 1$

$$I_k^m \geq \frac{1}{k!} \prod_{j=1}^{k} 2^m - j\,(m+1)$$

Note that if k becomes any larger, then we start multiplying negative numbers on the left hand side, and our initial argument for this inequality breaks down. From the inequality, we get:

$$\frac{I_k^m}{2^{km}} \leq \frac{1}{k!} \prod_{j=1}^{k-1} 1 - j\frac{(m+1)}{2^m}$$

Next we bound the right hand side

$$\frac{1}{k!} \prod_{j=1}^{k-1} 1 - j\frac{(m+1)}{2^m} \leq \frac{1}{k!}\left(1 - \sum_{j=1}^{k-1} j\frac{(m+1)}{2^m}\right)$$

The inequality above is proved using induction.

For example, $(1-x)(1-y) > 1-x-y = 1-(x+y)$ for $x, y > 0$ and so $(1-x)(1-y)(1-z) > (1-(x+y))(1-z) > (1-(x+y+z))$.

The right hand side above can be further simplified to

$$\frac{1}{k!}\left(1 - \frac{k\,(k-1)}{2}\frac{(m+1)}{2^m}\right) = \frac{1}{k!} - \frac{m+1}{2^{m+1}}\frac{1}{(k-2)!}$$

which rearranged proves the right inequality in the Lemma. The left inequality is proved similarly using as the upper bound on $I_k^m$the number of all subsets of size k chosen from 2^m elements $\binom{2^m}{k}$. □

To prove that $\underset{n\to\infty}{lim} P_n = 1 - \frac{1}{e}$ we must show that for every $\varepsilon > 0$ there exists a N such that $n > N$ implies that $\left|\left(1 - \frac{1}{e}\right) - P_n\right| < \varepsilon$. Chose an arbitrary $\varepsilon > 0$ and let $m_n = n\,(n-1)/2$. Observe the following:

1. $\underset{n\to\infty}{lim} \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k!} = 1 - \frac{1}{e}$

2. The sequence $\sum_{k=1}^{n} \frac{1}{k!}$converges (to e).

3. $\lim\limits_{n\to\infty} e\frac{n+1}{2^{n+1}} = 0$

Thus we can choose $N_1$where $n > N_1$implies $\left|\left(1-\frac{1}{e}\right) - \sum_{k=1}^{2^{m_n}} \frac{(-1)^{k-1}}{k!}\right| < \frac{\varepsilon}{3}$. We can choose N_2 where $n > N_2$ implies $\frac{1}{M!} + \frac{1}{(M+1)!} + \cdots + \frac{1}{2^{m_n}!} < \frac{\varepsilon}{3}$ where $M_n = \frac{2^{m_n}}{m_n+1} + 2$ (Cauchy Theorem). We can chose $N_3$where $n > N_3$implies $e\frac{n+1}{2^{n+1}} < \frac{\varepsilon}{3}$. Let $N = max(N_1, N_2, N_3)$so that $n > N$ implies that:

$$
\begin{aligned}
&\left|\left(1-\frac{1}{e}\right) - P_n\right| \\
&\leq \left|\left(1-\frac{1}{e}\right) - \sum_{k=1}^{2^{m_n}} \frac{(-1)^{k-1}}{k!}\right| + \left|\sum_{k=1}^{2^{m_n}} \frac{(-1)^{k-1}}{k!} - P_n\right| \\
&< \frac{\varepsilon}{3} + \left|\sum_{k=1}^{2^{m_n}} \frac{(-1)^{k-1}}{k!} - P_n\right|
\end{aligned}
$$

By the triangle inequality and point 1 above.

Next, we focus on the second term, using the formula for P_n in terms of the independence sequence we get:

$$
\begin{aligned}
\left|\sum_{k=1}^{2^{m_n}} \frac{(-1)^{k-1}}{k!} - P_n\right| &< \left|\sum_{k=1}^{2^{m_n}} (-1)^{k-1} \left(\frac{1}{k!} - \frac{I_k^{m_n}}{2^{km_n}}\right)\right| \\
&\leq \sum_{k=1}^{2^{m_n}} \left|\frac{1}{k!} - \frac{I_k^{m_n}}{2^{km_n}}\right| \\
&\leq \frac{m_n+1}{2^{m_n}} \left(\frac{1}{0!} + \cdots + \frac{1}{(M-1)!}\right) + \left(\frac{1}{M!} + \cdots + \frac{1}{2^{m_n}!}\right) \\
&< e\frac{m_n+1}{2^{m_n}} + \frac{\varepsilon}{3} \leq \frac{2}{3}\varepsilon
\end{aligned}
$$

Hence $\left|\left(1-\frac{1}{e}\right) - P_n\right| < \varepsilon$ and so $P_n \to 1 - \frac{1}{e}$.

4.7 Appendix 2: Proof for finite sets

We want to proof the following theorem for finite sets of outcomes, where V_n is the number of elements in the set and E_n is the number of alternatives that each otucome is compared to.

Theorem. *Assume that* $lim_{n\to\infty}\frac{V_n}{2^{E_n}} = c$ *and that* $lim_{n\to\infty}\frac{V_n}{E_n} = 0$ *then* $lim_{n\to\infty}P_n = 1-e^{-c}$

All the other theorems in the paper are special cases of this. For example, theorem 17 proved in the previous appendix where the outcomes are networks and networks are compared accroding to the one link deviations criterion, is a special case of theorem where we have$V_n = 2^m$ and $E_n = m$ for $m = n(n-1)/2$.

We begin by noticing that

$$\left|\left(1-e^{-c}\right)-P_n\right| \leq \left|\left(1-e^{-c}\right)-\left(c-\frac{c^2}{2!}+\frac{c^3}{3!}-\cdots+\frac{c^{V_n}}{V_n!}\right)\right|+ \\ +\left|\left(c-\frac{c^2}{2!}+\frac{c^3}{3!}-\cdots+\frac{c^{V_n}}{V_n!}\right)-\sum_{k=1}^{V_n}(-1)^{k-1}\frac{I_k^n}{2^{kE_n}}\right|$$

and the first term can be made arbitrarily small by choice of n (Taylor theorem).

To prove the theorem we must show we can make the second term arbitrarily small for n big enough.. Begin by noticing that:

$$\left|\left(c-\frac{c^2}{2!}+\frac{c^3}{3!}-\cdots+\frac{c^{V_n}}{V_n!}\right)-P_n\right|=$$
$$=\left|\left(c-\frac{V_n}{2^{E_n}}\right)+\left(\frac{I_2^n}{2^{2E_n}}-\frac{c^2}{2!}\right)+\cdots+\left(\frac{I_{V_n}^n}{2^{V_nE_n}}-\frac{c^{V_n}}{V_n!}\right)\right|\leq$$
$$\leq\sum_{k=1}^{V_n}\left|\frac{c^k}{k!}-\frac{I_k^n}{2^{kE_n}}\right|$$

Next, we get

$$\begin{gathered}\sum_{k=1}^{V_n}\left|\frac{c^k}{k!}-\frac{I_k^n}{2^{kE_n}}\right|\leq\\ \sum_{k=1}^{V_n}\left|\frac{1}{k!}\left[c^k-\left(\frac{V_n}{2^{E_n}}\right)^k\right]\right|+\\ +\sum_{k=1}^{V_n}\left|\frac{1}{k!}\left(\frac{V_n}{2^{E_n}}\right)^k-\frac{I_k^n}{2^{kE_n}}\right|\end{gathered}\tag{4.7.1}$$

We will work with the two terms above separately and show that by choosing n large enough both terms can be made arbitrarily small. First, we focus on the first term in 4.7.1 . We want to show that:

Proposition. *For every $\varepsilon>0$ there exists an N,such that for all $n>N$ we have* $\sum_{k=1}^{V_n}\left|\frac{1}{k!}\left[c^k-\left(\frac{V_n}{2^{E_n}}\right)^k\right]\right|<\varepsilon$

Proof. First we write

$$\left|c^k-\left(\frac{V_n}{2^{E_n}}\right)^k\right|=\left|c-\left(\frac{V_n}{2^{E_n}}\right)\right|\left|c^{k-1}+c^{k-2}\left(\frac{V_n}{2^{E_n}}\right)+\cdots+c\left(\frac{V_n}{2^{E_n}}\right)^{k-2}+\left(\frac{V_n}{2^{E_n}}\right)^{k-1}\right|$$

Because $lim_{n\to\infty}\frac{V_n}{2^{E_n}} = c$ we know that for every $\varepsilon > 0$ there exists a $N_0 > 0$ such that $\forall n > N_0 : \left|\frac{V_n}{2^{E_n}} - c\right| < \varepsilon$If we let $\varepsilon = c$ then we know there exist some $N_c > 0$ such that $\forall n > N_c : \frac{V_n}{2^{E_n}} < 2c$ and hence for all $n > N_0$ it holds that

$$\left|c^{k-1} + c^{k-2}\left(\frac{V_n}{2^{E_n}}\right) + \cdots + c\left(\frac{V_n}{2^{E_n}}\right)^{k-2} + \left(\frac{V_n}{2^{E_n}}\right)^{k-1}\right| < 2kc^{k-1}$$

So that given $\varepsilon > 0$ we let $N_1 = max\{N_0, N_c\}$ and we get that for all $n > N_1$:

$$\sum_{k=1}^{V_n}\left|\frac{1}{k!}\left[c^k - \left(\frac{V_n}{2^{E_n}}\right)^k\right]\right| < \sum_{k=1}^{\infty}\varepsilon\frac{2kc^{k-1}}{k!} = \varepsilon 2e$$

□

Next we show a similar proposition regaring the second term in 4.7.1.To prove the proposition we break the sum in the second term in 4.7.1 into three terms. The first term, the second up to $M_n \equiv \frac{V_n}{E_n+1} + 1$ and the remaining from $M_n$up:

$$\sum_{k=1}^{V_n}\left|\left[\frac{1}{k!}\left(\frac{V_n}{2^{E_n}}\right)^k - \frac{I_k^n}{2^{kE_n}}\right]\right| = \left|\left(\frac{V_n}{2^{E_n}}\right) - \frac{I_1^n}{2^{E_n}}\right| + \tag{4.7.2}$$
$$+\sum_{k=2}^{M_n}\left|\left[\frac{1}{k!}\left(\frac{V_n}{2^{E_n}}\right)^k - \frac{I_k^n}{2^{kE_n}}\right]\right| + \sum_{k=M_n+1}^{V_n}\left|\left[\frac{1}{k!}\left(\frac{V_n}{2^{E_n}}\right)^k - \frac{I_k^n}{2^{kE_n}}\right]\right|$$

The first term is identically equal to zero since $I_1^n = V_n$ for all n.

To show that the second term in 4.7.2can be made arbitrarily small be use the following lemma:

Lemma. *For* $2 \le k \le \frac{V_n}{E_n+1} + 1$*we have:*

$$\frac{1}{k!}\left(\frac{V_n}{2^{E_n}}\right)^k - \frac{I_k^n}{2^{kE_n}} \le \frac{1}{2}\frac{(E_n+1)}{V_n}\frac{1}{(k-2)!}\left(\frac{V_n}{2^{E_n}}\right)^k$$

Proof. Proof: Use the lower bound for I_k^n given by $\frac{1}{k!}V_n\left[V_n-(E_n+1)\right]\cdots\left[V_n-(k-1)(E_n+1)\right]$and divide by 2^{kE_n} to obtain

$$\frac{1}{k!}\left(\frac{V_n}{2^{E_n}}\right)^k\left[1-\frac{(E_n+1)}{V_n}\right]\cdots\left[1-(k-1)\frac{(E_n+1)}{V_n}\right]\leq\frac{I_k^n}{2^{kE_n}}$$

Next we use the inequality $(1-x_1)\cdots(1-x_k)>1-(x_1+\cdots+x_k)$and obtain that

$$\frac{1}{k!}\left(\frac{V_n}{2^{E_n}}\right)^k\left[1-\frac{(E_n+1)}{V_n}(1+\cdots+(k-1))\right]\leq\frac{I_k^n}{2^{kE_n}}$$

which implies that

$$\frac{1}{k!}\left(\frac{V_n}{2^{E_n}}\right)^k-\frac{I_k^n}{2^{kE_n}}\leq\frac{1}{2}\frac{(E_n+1)}{V_n}\frac{1}{(k-2)!}$$

□

Using the previous lemma we can now show the following.

Proposition. *Given $\varepsilon>0$ there exists a $N>0$ such that for all $n>N$ we have* $\sum_{k=2}^{M_n}\left|\left[\frac{1}{k!}\left(\frac{V_n}{2^{E_n}}\right)^k-\frac{I_k^n}{2^{kE_n}}\right]\right|<\varepsilon$.

Proof. We use the bound proved in the previous lemma to get that
$\sum_{k=2}^{M_n}\left|\left[\frac{1}{k!}\left(\frac{V_n}{2^{E_n}}\right)^k-\frac{I_k^n}{2^{kE_n}}\right]\right|\leq\frac{1}{2}\frac{(E_n+1)}{V_n}\sum_{k=2}^{M_n}\frac{1}{(k-2)!}\left(\frac{V_n}{2^{E_n}}\right)^k$

By assumption we can pick n big enough so that $\frac{V_n}{2^{E_n}}<2c$ and hence we get$\sum_{k=2}^{M_n}\left|\left[\frac{1}{k!}\left(\frac{V_n}{2^{E_n}}\right)^k-\frac{I_k^n}{2^{kE_n}}\right]\right|<\frac{1}{2}\frac{(E_n+1)}{V_n}4c^2\sum_{k=2}^{M_n}\frac{1}{(k-2)!}(2c)^{k-2}$and the right hand side is strictly smaller than $\frac{1}{2}\frac{(E_n+1)}{V_n}4c^2e^{4c^2}$which converges to zero by the assumption that $lim_{n\to\infty}\frac{V_n}{E_n}=0$. The last term in the sum converges to zero since it's strictly smaller than $\sum_{k=M_n+1}^{V_n}\left|\frac{1}{k!}\left(\frac{V_n}{2^{E_n}}\right)^k\right|\leq\sum_{k=M_n+1}^{\infty}\frac{1}{k!}(2c)^k$and the right hand side converges to zero since it is the tail of a series that converges(to e^{2c}). □

Insert conclusion here.

REFERENCES

[1] M. O. Jackson and A. Wolinsky, "A strategic model of social and economic networks," *Journal of Economic Theory*, vol. 71, pp. 44–74, October 1996.

[2] Y. Bramoulle and R. Kranton, "Risk-sharing networks," *Journal of Economic Behavior and Organization*, vol. 64, no. 3-4, pp. 275–294, 2007.

[3] M. Buchanan, *Nexus: Small Worlds and the Groundbreaking Theory of Networks*. W. W. Norton, 2003.

[4] D. Watts, *Small Worlds: The Dynamics of Networks Between Order and Randomness*. Princeton studies in complexity, Princeton University Press, 1999.

[5] D. Watts, *Six Degrees: The Science of a Connected Age*. W. W. Norton, 2004.

[6] R. V. D. Hofstad, "Random graphs and complex networks," in *In preparation*, 2008.

[7] B. Bollobas, *Random Graphs*. Cambridge University Press, 2001.

[8] P. Erdos and A. Renyi, "On random graphs i," *Publ. Math. Debrecen*, vol. 6, p. 290, 1959.

[9] P. Erdos and A. Renyi, "On the evolution of random graphs," *Publ. Math. Inst. Hungary. Acad. Sci.*, vol. 5, pp. 17–61, 1960.

[10] M. O. Jackson and A. Watts, "The existence of pairwise stable networks," *Seoul Journal of Economics*, vol. 14, p. 2001.

[11] S. Goyal and S. Joshi, "Unequal connections," *International Journal of Game Theory*, vol. 34, pp. 319–349, October 2006.

[12] B. Knaster, "Un théorème sur les fonctions d'ensembles.(french),"

[13] A. Tarski, "A lattice-theoretical fixpoint theorem and its applications," *Pacific journal of Mathematics*, vol. 5, no. 2, pp. 285–309, 1955.

[14] M. Dresher, "Probability of a pure equilibrium point in n-person games," *Journal of Combinatorial Theory*, vol. 8, no. 1, pp. 134 – 145, 1970.

[15] G. Papavassilopoulos, "On the probability of existence of pure equilibria in matrix games," *Journal of Optimization Theory and Applications*, vol. 87, pp. 419–439, 1995. 10.1007/BF02192572.

[16] William and Stanford, "A note on the probability of k pure nash equilibria in matrix games," *Games and Economic Behavior*, vol. 9, no. 2, pp. 238 – 246, 1995.

[17] K. Butler and M. Stephens, "The distribution of a sum of binomial random variables," tech. rep., DTIC Document, 1993.

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

[19] P. Favardin, D. Lepelley, and J. Serais, "original papers: Borda rule, copeland method and strategic manipulation," *Review of Economic Design*, vol. 7, no. 2, pp. 213–228, 2002.

[20] H. C. Huang and V. C. H. Chua, "Analytical representation of probabilities under the iac condition," *Social Choice and Welfare*, vol. 17, pp. 143–155, 2000. 10.1007/s003550050011.

[21] R. Bubley and M. Dyer, "Graph orientations with no sink and an approximation for a hard case of #sat," in *Proceedings of the eighth annual ACM-SIAM symposium on Discrete algorithms*, SODA '97, (Philadelphia, PA, USA), pp. 248–257, Society for Industrial and Applied Mathematics, 1997.

[22] J. Riordan, *An Introduction to Combinatorial Analysis,*. John Wiley, 1958.

[23] D. Galvin, "The independent set sequence of regular bipartite graphs," *Discrete Mathematics*, vol. 312, no. 19, pp. 2881–2892, 2012.