Selfish Creation of Realistic Networks

Anna Melnichenko

Universitätsdissertation
zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

in der Wissenschaftsdisziplin
Theoretische Informatik

eingereicht an der
Digital-Engineering-Fakultät
der Universität Potsdam

Datum der Disputation: 28. März 2022

Betreuer

Prof. Dr. Tobias Friedrich
Hasso Plattner Institute, University of Potsdam

Gutachter

Prof. Dr. Stefano Leucci
University of L'Aquila

Prof. Dr. Matúš Mihalák
Maastricht University

Online veröffentlicht auf dem
Publikationsserver der Universität Potsdam:
https://doi.org/10.25932/publishup-54814
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-548141

Abstract

Complex networks like the Internet or social networks are fundamental parts of our everyday lives. It is essential to understand their structural properties and how these networks are formed. A game-theoretic approach to network design problems has become of high interest in the last decades. The reason is that many real-world networks are the outcomes of decentralized strategic behavior of independent agents without central coordination. Fabrikant, Luthra, Maneva, Papadimitriou, and Schenker [Fab+03] proposed a game-theoretic model aiming to explain the formation of the Internet-like networks. In this model, called the Network Creation Game, agents are associated with nodes of a network. Each agent seeks to maximize her centrality by establishing costly connections to other agents. The model is relatively simple but shows a high potential in modeling complex real-world networks. In this thesis, we contribute to the line of research on variants of the Network Creation Games. Inspired by real-world networks, we propose and analyze several novel network creation models. We aim to understand the impact of certain realistic modeling assumptions on the structure of the created networks and the involved agents' behavior.

The first natural additional objective that we consider is the network's robustness. We consider a game where the agents seek to maximize their centrality and, at the same time, the stability of the created network against random edge failure.

Our second point of interest is a model that incorporates an underlying geometry. We consider a network creation model where the agents correspond to points in some underlying space and where edge lengths are equal to the distances between the endpoints in that space. The geometric setting captures many physical real-world networks like transport networks and fiber-optic communication networks.

We focus on the formation of social networks and consider two models that incorporate particular realistic behavior observed in real-world networks. In the first model, we embed the anti-preferential attachment link formation. Namely, we assume that the cost of the connection is proportional to the popularity

of the targeted agent. Our second model is based on the observation that the probability of two persons to connect is inversely proportional to the length of their shortest chain of mutual acquaintances.

For each of the four models above, we provide a complete game-theoretical analysis. In particular, we focus on distinctive structural properties of the equilibria, the hardness of computing a best response, the quality of equilibria in comparison to the centrally designed socially optimal networks. We also analyze the game dynamics, i.e., the process of sequential strategic improvements by the agents, and analyze the convergence to an equilibrium state and its properties.

Zusammenfassung

Komplexe Netzwerke, wie das Internet oder soziale Netzwerke, sind fundamentale Bestandteile unseres Alltags. Deshalb ist es wichtig, ihre strukturellen Eigenschaften zu verstehen und zu wissen, wie sie gebildet werden. Um dies zu erreichen, wurden in den letzten Jahrzehnten spieltheoretische Ansätze für Netzwerkdesignprobleme populär. Der Grund dafür ist, dass viele reale Netzwerke das Ergebnis von dezentralem strategischem Verhalten unabhängiger Agenten ohne zentrale Koordination sind. Fabrikant, Luthra, Maneva, Papadimitriou und Schenker [Fab+03] haben ein solches spieltheoretisches Modell vorgeschlagen, um die Entstehung von internetähnlichen Netzwerken zu erklären.

In diesem Modell, dem sogenannten Network Creation Game, repräsentieren die Agenten die Knoten eines Netzwerks. Jeder Agent versucht, durch den Kauf von Verbindungen zu anderen Agenten seine Zentralität im erzeugten Netzwerk zu maximieren. Dieses Modell ist relativ einfach, aber es hat ein großes Potenzial, reale Netzwerke modellieren zu können. In der vorliegenden Arbeit tragen wir zur aktuellen Forschungsrichtung, die sich der Untersuchung von Varianten der Network Creation Games widmet, bei. Inspiriert von realen Netzwerken, schlagen wir verschiedene neuartige Netzwerkbildungsmodelle vor und analysieren diese. Wir wollen hierbei die Auswirkungen bestimmter realistischer Modellierungsannahmen auf die Struktur der erstellten Netzwerke und das Verhalten der beteiligten Agenten verstehen.

Die erste natürliche zusätzliche Modellierungsannahme, die wir betrachten, ist ein Fokus auf die Robustheit des erzeugten Netzwerks. In diesem Modell haben die Agenten das Ziel, ihre Zentralität zu maximieren und gleichzeitig das erstellte Netzwerk robust gegenüber zufällige Verbindungsausfälle zu machen.

Das zweite neue Modell, das wir hier betrachten, bezieht eine zu Grunde liegende Geometrie mit ein. Hierbei entspricht jeder Agent einem Punkt in einem gegebenen Raum und die Länge einer Netzwerkverbindung entspricht der Distanz zwischen den jeweiligen Endpunkten in diesem Raum. Diese geometrische Variante erlaubt die Modellierung vieler realer physischer Netzwerke, wie z.B. Transportnetzwerke und Glasfaserkommunikationsnetzwerke.

Des Weiteren fokussieren wir uns auf die Bildung von sozialen Netzwerken und betrachten zwei Modelle, die ein bestimmtes realistisches Verhalten einbeziehen, das in realen sozialen Netzwerken beobachtet werden kann. Das erste Modell basiert auf einer anti-präferentiellen Kantenerzeugung. Dabei nehmen wir an, dass die Kosten einer Verbindung proportional zur Popularität des Agenten am anderen Endpunkt sind. Das zweite betrachtete Modell basiert auf der Beobachtung, dass die Wahrscheinlichkeit, dass zwei Personen verbunden sind, proportional zur Länge ihrer kürzesten Kette von gegenseitigen Bekanntschaften ist.

Für jedes der vier oben genannten Modelle liefern wir eine komplette spieltheoretische Analyse. Insbesondere fokussieren wir uns auf charakteristische strukturelle Eigenschaften der spieltheoretischen Gleichgewichte, die Komplexität der Berechnung einer optimalen Strategie und die Qualität der Gleichgewichte im Vergleich zu den zentral entworfenen sozial optimalen Netzwerken. Außerdem analysieren wir auch die Spieldynamik, d.h. den Prozess von sequentiellen verbessernden Strategieänderungen der Agenten. Dabei untersuchen wir die Konvergenz zu einem Gleichgewichtszustand und die Eigenschaften solcher Konvergenzprozesse.

Acknowledgments

My path from the the first day at Hasso Plattner Institute until the publication of the thesis was associated with many changes: another country, language, culture, research topic, and new people. However, this was also the most exciting period of my life. And I am very grateful and endlessly love all the changes that have happened to me and around me.

First of all, I would like to thank two people without whom there would be neither beginning nor end. I want to express my sincere gratitude to my supervisor Tobias Friedrich for the continuous support of my study, for his trust, and for believing in a second chance. Thank you for guiding me all the time, even when the future seemed unclear.

My sincere thanks to Pascal Lenzner for being the best adviser ever. His optimism, inspiring ideas, continuous motivation, and immense knowledge have lightened my way from the first day of my study. And many thanks to Pascal's family. Marie, August, and Linda, thanks for being so great!

I thank all members of the Algorithm Engineering group for all the fun we have had. I cannot list all 30+ names here, but each of you is in my heart. My many thanks to Louise for being the game theory soul mate, for all our conference trips together, and always being up for new adventures. I thank Gregor for his unique attitude and always kind words. Many thanks to Karen and her family, Philipp and Nathan, for being my best friends. And my thanks to Vanja, Philipp, Katerina, Janosh, Martin, and Ziena for being so great!

Also, I want to acknowledge Davide Bilò for each joint research project, from an idea until the last second of the submission deadlines. Many thanks for all the interesting scientific discussions and knowledge I learned from you.

Of course, my journey to mathematics and computer science would not be possible without the support of my parents. Mom and dad, thank you for your love and faith in me. Last but not least, I would like to thank my husband. Evgeny, you not only stayed supportive and patient when it was almost impossible but helped me to make the most unbelievable dreams come true.

Contents

1 Introduction

1.1 Motivation

Networks accompany all fields of our life. Social networks represent our relationships with other people; transport networks are essential vessels delivering people and goods; the World Wide Web represents a global information space; mobile networks and the Internet help us keep in touch with other people. Clearly, it is important to keep-up the functionality and efficiency of these networks to provide economic stability and the well-being of our world. Two core questions are facing modern science: understanding the process of real-world network formation and a relationship between the network behavior and its structure.

Many exciting results and valuable progress in resolving these problems have been performed in Operations Research, Economics, and Theoretical Computer Science [GK11; Jac10]. High interest in network design problems has also kindled the interdisciplinary field of Network Science [Bar16], which is devoted to analyzing and understanding real-world networks. Most of the results and studied models assume that the considered networks are controlled by a central authority that can perform any kind of structural modifications to the network and that has all data about the network's current state. However, many real-world networks such as social networks and the Internet are maintained and developed by independent agents where each agent aims to satisfy her personal interests.

Therefore, the focus of this thesis is a game-theoretic approach to network formation. Here, independent agents corresponding to nodes establish connections between each other to minimize their payoff. The union of all agents' strategies defines the outcome of the game, i.e., the network structure. Equilibrium networks, i.e., states of the game where no agent can perform an improvement of her strategy, is of particular interest since it can be seen as a result of a decentralized process of network formation. Moreover, an analysis of the quality of equilibria in the game can help to understand how the lack of a central authority can influence a system's infrastructure and functionality.

The pioneering work by Myerson [Mye77] joined two different fields, Graph Theory and Game Theory, and introduced a new form of a game where agents are nodes of a network who can establish connections in pairwise cooperation to minimize their cost. The idea of strategic network formation became very popular and has been considered and analyzed in many scientific works. One of the seminal papers has been introduced by Fabrikant, Luthra, Maneva, Papadimitriou, and Shenker [Fab+03]. Their model, called the Network Creation Game (NCG), aims at modeling the formation of the Internet-like networks. In this game, agents are nodes who can establish incident costly connections to maximize their centrality at the minimum price. For its simplicity but at the same time intriguing properties of equilibria, the game attracted much attention in the Algorithmic Game Theory field. Later, Chun et al. [Chu+04] performed a series of experiments with various modifications of the game. They showed the potential of the NCG in modeling real-world networks by varying the cost function. In this thesis, we extend the research on the NCGs and show what basic principles are necessary ingredients for the NCG to model a behavior of real-world networks. Moreover, we perform a rigorous theoretical analysis of the games' outcomes.

1.2 Thesis Outline

In Chapter 2 we introduce the necessary game-theoretic notation and give a detailed definition of the original Network Creation Game. Moreover, this chapter provides an overview of the related recent results on the network formation models.

In Chapter 3, we investigate a network creation model where agents aim to form connections that guarantee robustness against random edge failure. The model aims to understand the structure and quality of the selfishly created networks and to compare the results with the outcomes of the classical Network Creation Game outcomes. In particular, we will show that similar to the NCG, the equilibria have bounded diameter, but the class of optimum and equilibrium networks is much more diverse compared to the classic NCG.

In Chapter 4 we present a generalized Network Creation Game with arbitrary edge weights. We mainly focus on the geometric setting where the weights equal the metric distances between the nodes induced by an underlying geometry. The model naturally replicates real-world infrastructure networks like road networks and electricity networks. In contrast to the classical NCG with uniform edge weights, even the problem of of settling the existence of an equilibrium state is a big challenge in the model with weighted edges. We provide a rigorous analysis of the game for different metric spaces and answer questions regarding the convergence to an equilibrium state, the hardness of computing the best strategy, and the quality of the equilibrium networks. The most surprising result of this chapter is that the quality of the worst equilibrium does not depend on the geometry.

Chapters 5 and 6 present a novel class of NCGs with dynamic edge costs. Both models are inspired by real-world social networks. In Chapter 5 we consider a model based on a natural principle that the cost for the formation of an edge is proportional to the popularity of the targeted node. We consider different versions of the game by incorporating locality, i.e., we assume that that agents can only form edges in their local neighborhood. In addition to the standard game-theoretical analysis focusing on the properties and the quality of equilibria, we study the game dynamics. In particular, we analyze the convergence speed to a stable state if agents sequentially perform improving edge additions starting from some particular initial configurations.

The definition of the model presented in Chapter 6 incorporates several natural properties of social networks. We assume that edge formation is bilateral and that the edge cost is proportional to the distance between the nodes in the network without this edge. This assumption is based on the observation that a connection between two friends of a common friend is more likely than between two more remote persons in real-world social networks. We analyze the structure and quality of the pairwise stable networks in the game and show theoretically and experimentally that the stable networks mimic many real-world network properties like low diameter, high clustering, and a power-law degree distribution.

We conclude with Chapter 7 where we emphasize the most important findings and observations of this thesis and give an overview on the most promising future research directions and open questions.

2 Preliminaries

Algorithmic Game Theory is a relatively young research area that attracted much attention because it advances a dialogue between Theoretical Computer Science and the real world. The pioneering work by von Neumann and Morgenstern [MV44] introduced Utility Theory which allowed modeling the behavior of selfish agents. Each agent chooses a strategy to maximize her utility function, which is a mapping of the observed outcome of the game to a real number. In the last decades, focus on different concepts, questions, and application domains gave birth to different research directions within Algorithmic Game Theory. We focus on only one if this directions in this work, namely *Network Creation Games (NCG)*. For an overview of other areas, we refer to the books by Myerson [Mye91], Nisan et al. [Nis+07], and by Shoham and Leyton-Brown [SL08].

In the following, we will define the original Network Creation Game [Fab+03] and all terms and notions necessary to understand the thesis.

We assume that the reader is familiar with basics concepts of Graph Theory and refer to the book by West [Wes01] for omitted definitions. However, we recap certain key notation. We use the standard graph-theoretic notation and denote a network on a set of nodes V and an edge set E as $G = (V, E)$. We use uv to denote an undirected edge between the nodes u and v in a network. The degree of a node u, i.e., its number of neighbors, in a network G is denoted as $deg_G(u)$. A node of degree one is called a leaf. For a network $G = (V, E)$, the network $G + uv$ (respectively, $G - uv$) denotes the network G in which the edge uv has been added (respectively, deleted). We denote as $N_k(u)$ the neighborhood of a node u (nodes at distance at most k), and we denote as $B_k(u)$ the set of nodes at distance exactly k from u. We highlight some special networks that we will need later on: a *clique* $\mathbf{K}_n$ on n nodes is a network where every two distinct nodes are adjacent; a *star* S_n with n nodes is a network that contains one central node connected to $n - 1$ leaves.

2.1 The Network Creation Game

In 2003, Fabrikant, Luthra, Maneva, Papadimitriou, and Shenker [Fab+03] proposed a model of decentralized network formation. The model is called Network Creation Game, and we will refer to it as the *classical Network Creation Game (NCG)*. It has been proposed as a game-theoretic model of real-world network formation with no central coordination. The model can be seen as a simplified variant of the Connection Game by Jackson & Wolinsky [JW96]. Formally, the NCG is defined as follows. Given a set of n agents $V = \{v_1, \ldots, v_n\}$ corresponding to nodes of a network. Each agent can individually decide to what subset of nodes she buys the edges in order to maximize her centrality. More precisely, let $\mathbf{s} := (S_{v_1}, \ldots, S_{v_n})$ denote the *strategy profile*, where $S_u \subseteq V \setminus \{u\}$ corresponds to the *pure strategy* of an agent u. The strategy S_u specifies the edges *owned* by agent u. Every edge has price $\alpha \in \mathbf{R}_{>0}$, where α is some fixed parameter of the game. The strategy profile uniquely specifies the network $G(\mathbf{s}) = (V, E)$ with $E = \{uv | u, v \in V, u \in S_v \vee v \in S_u\}$. If $v \in S_u$, then we call agent u the *owner* of the undirected edge uv, and u has to pay the full edge price. Note that if $v \in S_u$ and $u \in S_v$, then both agents have to pay the full edge price. However, in this case, one of the agents could improve her current situation in the network by not buying the edge uv, which implies that every edge has exactly one owner in any equilibrium or social optimum network.

We will omit the reference to $\mathbf{s}$ when it is clear from a context. Note that all edges are undirected, but we use directed edges on the figures to illustrate the edge ownership (an arrow directs away from its owner).

Let $d_G(u, v)$ be the hop-distance between two nodes u and v in G, which is equal to the number of edges on the shortest path between u and v in G. If there is no u-v path, then $d_G(u, v) = +\infty$. We use $d_G(u, U)$ to denote the sum of distances from the node u to all nodes in $U \subseteq V$ in G, i.e., $d_G(u, U) := \sum_{v \in U} d_G(u, v)$. We call $d_G(u, V)$ the *distance cost* and $\alpha \cdot |S_u|$ the *edge cost* of agent u. Each agent aims at minimizing her cost, that is the agent's distance cost plus her edge cost:

$$cost(u, \mathbf{s}) := \alpha \cdot |S_u| + d_{G(\mathbf{s})}(u, V).$$

When a strategy profile is clear from the context, we can establish a bijection between the network $G(\mathbf{s})$ and the pair $(V, \mathbf{s})$. In this case, we can use $\mathbf{s}$ and G interchangeably. Hence, the cost of an agent u is denoted as $cost(u, G)$.

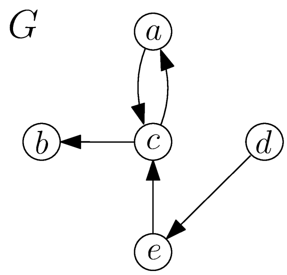


Figure 2.1: An instance of the classical Network Creation Game. Strategies of the agents are $S_a = \{c\}$, $S_b = \emptyset$, $S_c = \{a, b\}$, $S_d = \{e\}$, $S_e = \{c\}$. Hence, agents have the following payoffs: $cost(a, G) = \alpha + 8$, $cost(b, G) = 8$, $cost(c, G) = 2\alpha + 5$, $cost(d, G) = \alpha + 9$, and $cost(e, G) = \alpha + 6$.

For an illustration of the model definition, consider Figure 2.1.

Let $\mathbf{s}$ be any strategy profile and $G := G(\mathbf{s})$ is the corresponding network. Consider agent u with her strategy S_u in G. Let $\mathbf{s}'$ be a strategy profile obtained from StP after agent u changed her strategy from S_u to some strategy S'_u. We say S'_u is an *improving response* for the agent u if the strategy change from S_u to S'_u strictly decreases her cost, i.e., $cost(u, \mathbf{s}') < cost(u, \mathbf{s})$. We call this strategy change as an *improving move*. We say that agent i is playing a *best response* if she has no improving move.

For a given input of the game, consider a set of all possible strategy profiles S. A subset of S specified by some formal rule is called a *solution concept*. The standard solution concept for the Network Creation Game is the *pure Nash equilibrium (NE)* [Nas51]. A strategy profile $\mathbf{s}$ is in NE if no agent can unilaterally update her strategy to strictly decrease her cost, i.e., if each agent is playing her best response. Formally, a strategy profile $\mathbf{s}$ is in NE if for any agent $u \in V$, given strategy profile $\mathbf{s}$ and every $\mathbf{s}'$ where only the strategy of agent u is modified, $cost(u, \mathbf{s}) \leq cost(u, \mathbf{s}')$.

The concept of Nash equilibrium assumes that each agent can replace her strategy by an arbitrary other strategy. However, this assumption makes computing a best response computationally hard[1]. To simplify the computation of an improving step and to make the model more realistic, several other weaker solution concepts have been introduced in the line of recent research:

1 A computational hardness of computing a best response has been shown in the original NCG paper [Fab+03]

- *Greedy equilibrium (GE).* A network is in greedy equilibrium if no agent can unilaterally improve her strategy by adding, deleting, or swapping one own incident edge. This concept has been introduced by Lenzner in [Len12].

- *β-approximate equilibrium.* A network is in β-approximate equilibrium if no agent can unilaterally change her strategy to improve her cost by more than a factor of β. The concept has been first defined by Chien and Sinclair in [CS11] for the class of congestion games. Note that this concept can be used in combination to the mentioned equilibrium variant.

We introduce a new solution concept, called *add-only equilibrium (AE).* We say that a network is in add-only equilibrium if no agent can unilaterally improve her strategy by adding one incident edge. This concept is even simpler than Greedy equilibrium since every agent can compute her best response in $O(n)$ steps by trying all possibilities. Observe, that the add-only agent's behavior naturally models social networks where a connection between two nodes means that two persons know each other.

Apart from the above mentioned solution concepts, the asymmetric swap equilibrium [MS12], where improving moves are restricted to a swap of one edge owned by an agent, the swap equilibrium [Alo+13] where improving moves are restricted to a swap of one incident edge, and the strong equilibrium [AFM09], where no coalition of agents can improve their strategies via a joint strategy change, have been studied.

Aside from the above defined equilibrium concepts, there is the concept of pairwise stability [JW96]. Formally, a network $G = (V, E)$ is *pairwise stable* if and only if the following conditions hold:

1. for any edge $uv \in E$, $cost(u, G - uv) \geq cost(u, G)$ and $cost(v, G - uv) \geq cost(v, G)$;

2. for every edge $uv \notin E$, $cost(u, G + uv) \geq cost(u, G)$ or $cost(v, G + uv) \geq cost(v, G)$.

In the simplest terms, a network is pairwise stable if every edge in G is beneficial for both endpoints, and for every edge not in G, at least one of its endpoints blocks the edge because it does not decrease her cost. This concept is prevalent in economics research. Moreover, it is a natural concept for modeling real-world social networks where an edge between two nodes means a connection between two persons.

If the concept is clear from the context, we call a network *stable* if it is in the named equilibrium.

For an illustration of the concepts that we will need later on, consider the networks in Figure 2.2.

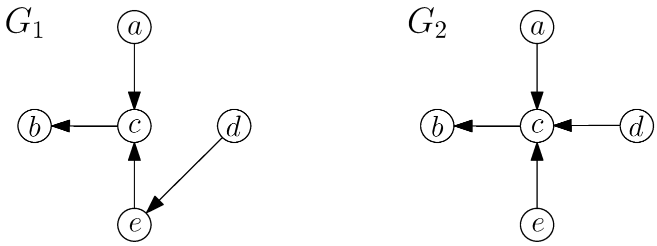


Figure 2.2: Instances of different equilibrium concepts. The first network G_1 is in add-only equilibrium for $\alpha = 1$. However, it is not in greedy equilibrium because agent d can improve her strategy by the swap de to dc. The network G_2 is in Nash equilibrium for $\alpha = 1$, and thus, it is also in add-only and greedy equilibrium. Both networks (if we ignore the edge ownership) are pairwise stable for $\alpha = 2$ since no edge can be deleted unilaterally and no edge can be added because it is not beneficial for one of the endpoints.

2.2 Game-Theoretic Analysis

We aim to evaluate the quality of a considered model and understand whether the corresponding equilibrium networks meet the expectations about the model. The first problem that we need to study is the existence of equilibria for any input of the game. In the classical NCG, the objective to maximize the centrality in a network leads to a trivial observation that a star is an equilibrium network for any $\alpha \geq 1$, whereas a clique is an equilibrium for any $\alpha < 1$. However, in the general case, the question is not trivial, and there are variants of the NCG which do not guarantee the existence of pure NE for any value of α [MSW06].

As soon as the existence of NE is confirmed, the next natural question occurs: Is equilibrium unique in the game? If not, how bad or good can the game outcomes be? It leads us to the question of the quality of equilibrium networks in the game. The standard way to measure the efficiency of networks in the

NCG is an objective function called *social cost*, which is the sum of the costs of all agents:

$$\mathrm{SC}(G) = \sum_{u \in V} cost(u, G).$$

To measure the loss of efficiency due to selfishness of agents, two measures have been established, the *Price of Anarchy (PoA)* [KP99] and the *Price of Stability (PoS)* [Ans+08a]. Let $worst_n$ (respectively, $best_n$) be the highest (respectively, lowest) possible social cost of an equilibrium network obtained by n agents. Let, OPT_n be a network of n nodes with the minimum social cost. Then we define:

- the Price of Anarchy is $\frac{worst_n}{\mathrm{SC}(OPT_n)}$,
- the Price of Stability is $\frac{best_n}{\mathrm{SC}(OPT_n)}$.

Simply put, the Price of Anarchy evaluates how bad selfishly created networks can be in contrast to a socially optimal network designed by a central authority. Bounding the PoA is an important but challenging problem. A long line of research [Alb+14; ÀM17; ÀM18; ÀM19; BL20; Dem+12; Fab+03; MMM15; MS13] has established that the PoA of the NCG is constant for almost all α, and it is conjectured that this holds for all α [Alo+13; Fab+03; MMM15; MS12].

The Price of Stability considers the ratio between the social cost of an equilibrium and a social optimum in a more optimistic way. For instance, a statement that a game has the $PoS = 1$ implies that there always exists an equilibrium state with minimum possible social cost. It holds for the classical NCG for $\alpha \geq 2$, as a center-sponsored star is in equilibrium and a social optimum [Fab+03].

Another important question to study is the hardness of computing a best response. Computing the best possible strategy of an agent in the NCG was shown to be NP-hard [Fab+03], and this also holds for many NCG variants [CL15; MS12]. However, restricted variants with efficient best response computation also exist [Alo+13; BG00; Fri+17; Len12].

To measure the quality of stable networks we analyze their graph theoretic properties, like the network diameter, structural properties, and the average clustering coefficient.

2.3 Game Dynamics

So far, we focused on the analysis of equilibrium states. In this case, the game is considered a one-shot game. It means that for a given initial state, agents can simultaneously perform only one improving move. However, the NCGs allow us to model a process of network formation if we take a more constructive sequential view of the game. More precisely, we want to study the process of how the networks evolve over time if we start from a non-stable configuration and let the agents sequentially play improving moves. If the process converges to a state where no agent can perform an improving move, an equilibrium is found. Hence, these game dynamics are an algorithm for computing stable networks. However, it is a challenging problem to prove (or disprove) convergence to a stable state. To specify the game dynamics, the initial configuration, the agents' activation order, the type of the move (improving or best move) should be defined.

Strategic games can be classified according to their dynamics. We focus on one well-known class: games with the finite improving property. Formally, a game has the *finite improving property (FIP)* if starting from any strategy vector, any sequence of improving moves is finite. It is equivalent to the game being an *ordinal potential game* [MS96]. According to the definition, a game is an ordinal potential game if there exists a generalized ordinal potential function that maps a strategy profile to a real number and has the property that if the active agent's cost decreases, then the potential function decreases as well. If a game with a finite set of feasible strategies comes to the initial state after several steps of an improving response dynamic, then the process is cyclic. The cycle is called an *improving response cycle (IRC)* (respectively, *best response cycle (BRC)*) in case of the improving (respectively, best) response dynamics. The existence of an IRC or BRC disproves that the respective game has an ordinal potential function.

The FIP is a potent property that not only guarantees convergence of game dynamics but also proves the existence of an equilibrium state. However, the speed of convergence of such distributed local search can be exponential [FPT04; SV08].

For the classical NCG and some variants of the game, it has been proven [KL13; Len11] that there is no ordinal potential function, i.e., natural convergence protocols like iterated improving response dynamics have no convergence guarantee.

2.4 Related Work

This section will list the latest game-theoretic models for the formation of networks with realistic properties. Note that there can be two ways of modeling real-world networks: either by considering models with realistic properties or by developing models that guarantee that the stable networks in this model mimic real-world networks.

Extensive research, e.g. [AJB99; BA99; Bar16; Bro+00; Kle00; LKF05; NBW11], on real-world networks from many different domains like communication networks, social networks, and metabolic networks. has revealed the astonishing fact that most of these real-world networks share the following basic properties:

- *Small-world property:* The diameter and average distances are at most logarithmic in number of nodes.
- *High clustering:* Two nodes with a common neighbor have a high probability of being neighbors, i.e., there is an abundance of triangles and small cliques. More formally, let $\Delta(v)$ denote the number of triangles in G that contain the node v. The *local clustering coefficient* $CC(v)$ of node v in G is the probability that two randomly selected neighbors of v are neighbors, i.e., $CC(v) := \frac{2\Delta(v)}{deg(v)(deg(v)-1)}$ if $deg(v) \geq 2$, and 0 otherwise. The *average local clustering coefficient* CC of a network G with n nodes is the average of the local clustering coefficients over all nodes v, i.e., $CC(G) = \frac{1}{n}\sum_{v \in V} CC(v)$. It was observed that the average local clustering coefficient in real-world networks depends on the node's degree k and roughly is k^{-1}, and that the average local clustering coefficient is relatively high [Bar16].
- *Power-law degree distribution:* if the probability that a node has degree k is proportional to $k^{-\beta}$, the degree distribution follows a power-law. For $2 \leq \beta \leq 3$, networks with a power-law degree distribution are called *scale-free networks.*

The phenomenon that real-world networks from different domains are very similar calls for a scientific explanation, i.e., there is a high interest in formal models that generate networks with the above properties from simple rules.

Many models that generate networks with the above properties following a simple protocol have been proposed. Most prominent models are the preferential attachment model [BA99], Chung-Lu random graphs [CL02], hyperbolic random graphs [FK15; Kri+10], and geometric inhomogeneous random graphs [BKL19]. However, all these models describe a purely random process while many real-world networks evolved by an interaction of rational agents. For example, in (online) social networks [Jac10] the selfish agents correspond to people or firms that choose carefully with whom to maintain a connection. Thus, a model with higher explanatory power should consider rational selfish agents who use and modify the network to their advantage [Pap01]. It explains the need for a game-theoretic model to explain the process of real-world network formation.

So far, the NCG can explain the small-world property. That is, it has been proven that the diameter of all equilibrium networks is small [Dem+12]. Other game-theoretic models usually embed the desired properties in the agents' cost function to guarantee the real-world properties for any game outcome. For example, the island connection model [JR05] assumes that groups of agents are based on islands and that the edge cost within an island is much lower than across islands. This yields equilibria with low diameter and high clustering but no realistic degree distribution.

Another promising model considers the World Wide Web creation [Kou+15]. In this model, nodes strategically choose outgoing links and click probabilities to maximize their utility, which is proportional to the traffic on the node and the node's quality. The resulting Nash equilibria have many features of real-world content networks.

In the next model, called the network navigation game [Gul+15], agents are randomly sampled points in the hyperbolic plane. Each agent selfishly decides which edges to add to ensure the performance of greedy routing in the formed network. The equilibrium networks indeed have a power-law degree distribution and high clustering. However, the main reason for this is not the agents' strategic behavior but the fact that the agents correspond to uniformly sampled points in the hyperbolic plane.

Since the classical NCG has been introduced, many researchers tried to modify the model to incorporate more realistic constraints. One natural direction is to model a formation of physical networks by embedding weighted networks in the model [Alb+14; Dem+09]. In the particular case, when the edge weights satisfy the triangle inequality, we have a geometric network. One fundamental property of many efficient geometric networks is the low stretch factor. The stretch is the maximum of the ratio of the shortest path length in the network and the geometric distance over all pairs of nodes. Strategic network formation where each agent seeks to minimize her total stretch has been studied in [AQ19; MSW06; MSW11]. A more detailed literature overview related to the topic can be found in Chapter 4.

One of the main critiques of the classical NCG is that it neglects the locality of the agents' knowledge about the network. Usually, agents cannot observe the entire network to calculate their best strategy optimally. Several versions of the definition of the NCGs with locality have been proposed. In [Bil+14b; Bil+16; YY20] the authors consider agents with local knowledge, i.e., when agents can observe only a local part of the network and have to perform a strategy calculation based on their worst case assumption about the unseen part of the network. A more optimistic definition of locality has been considered in [CL15] where agents have a global knowledge about the network but can perform moves only within their local neighborhood.

The model can be augmented by restricting the allowed strategy changes, i.e., by choosing a specific solution concept. For example, pairwise stability is a natural way to model real-world social networks, while add-only equilibria can mimic network formation in co-authorship networks. Other solution concepts, as well as the related work, have been introduced earlier in this chapter.

Apart from the mentioned models, there is one model which incorporates one of the real-world network properties in the utility function, called the clustering coefficient network formation model [BK11]. This approach can guarantee a high local clustering coefficient of the agents in the equilibrium networks. However, it has been shown that the NEs can have a large diameter, i.e., the small-world property does not hold.

3 On Selfish Creation of Robust Networks

Robustness is one of the critical properties of nowadays networks. However, robustness cannot be simply enforced by design or regulation since many important networks, most prominently the Internet, are not created and controlled by a central authority. Instead, Internet-like networks emerge from strategic decisions of many selfish agents. Despite this fact, the Internet seems robust against node or edge failures, which hints that a socially beneficial property like network robustness may emerge from selfish behavior.

To investigate this phenomenon we present in this chapter a simple model for selfish network creation, which explicitly incorporates agents striving for a central position in the network while at the same time protecting themselves against random edge-failure. We show that networks in our model are more diverse than in the original NCG model. We also show the versatility of our model by adapting various properties and techniques from the non-robust versions. In combination with the interesting structural properties of stable networks claiming that the amount of edge-overbuilding due to the adversary is limited, we prove an upper bound on the Price of Anarchy of $O(1 + \alpha/\sqrt{n})$. Another positive statement for our model provides a constant bound for the Price of Stability that implies that the selfish agents' behavior can lead to efficient and stable networks. Moreover, we analyze the computational hardness of finding best possible strategies and investigate the game dynamics of our model.

3.1 Model and Notation

We consider the Network Creation Game (NCG) by Fabrikant et al. [Fab+03] augmented with the uniform edge-deletion adversary from Kliemann [Kli11] and we call our model *Adversary NCG* (Adv-NCG). More specifically, in an Adv-NCG there are n selfish agents which correspond to the nodes of an undirected multi-graph $G = (V, E)$. A pure strategy S_u of agent $u \in V(G)$ is any multi-set over elements from $V \setminus \{u\}$. If v is contained k times in S_u then this encodes that agent u wants to create k undirected edges to node v. Moreover we say that u is the owner of all edges to the nodes in S_u. Given an n-dimensional vector of pure strategies for all agents, then the union of all the edges encoded in all agents' pure strategies defines the edge set E of the multi-graph G.

The agents prepare for an adversarial attack on the network after creation. This attack deletes one edge uniformly at random. Hence, agents try to minimize the attack's impact on themselves by minimizing their *expected cost*. Let $G - e$ denote the network G where edge e is removed. We define the *expected distance cost* for an agent u as

$$\delta_G(u, V) = \frac{1}{|E|} \sum_{e \in E} d_{G-e}(u, V(G-e)) = \frac{1}{|E|} \sum_{e \in E} \sum_{v \in V} d_{G-e}(u, v).$$

The expected cost of agent u in a network $G(\mathbf{s}) = (V, E)$ with the edge price α is defined as

$$cost(u, G) = \alpha \cdot |S_u| + \delta_G(u, V),$$

where $\alpha \cdot |S_u|$ is the *edge cost* of the agent u. Thus, compared to the original NCG [Fab+03], the distance cost term is replaced by the expected distance cost with respect to uniform edge deletion.

The social cost of a network $G(\mathbf{s})$ is defined in a standard way as

$$SC(G(\mathbf{s})) = \sum_{u \in V} cost(u, G(\mathbf{s})).$$

3.2 Related Work

For more than two decades network security is a hot topic in the game theory field. Many models and approaches have been introduced (see [EB19; LX12; Mor+20; Roy+10] for an overview) but most of them focus on the interaction between two agents, attacker and defender, on a given network. The defender chooses its strategy to keep the network stable against the strategic attacks of the opponent. However, this approach assumes that the defender can observe and control the entire network, while our main focus is on the network formation process where agents can modify only incident connections. To the best of our knowledge, only a few network formation models have been studied that incorporate edge-failure.

The paper by Meirom et al. [MMO15] is the only centrality network connection model[2] which incorporates edge-failures. The authors consider two types of agents, major-league and minor-league agents, which maintain that the network remains 2-connected while trying to minimize distances, which are a linear combination of the length of a shortest path and the length of a best possible vertex disjoint backup path. Under some specific assumptions, e.g. that there are significantly more minor-league than major-league agents, they prove various structural properties of equilibrium networks and investigate the corresponding game-dynamics. In contrast to this, we will investigate a much simpler model with homogeneous agents which is more suitable for analyzing networks created by equal peers. Our results can be understood as zooming in on the sub-network formed by the major-league agents (i.e., top tier ISPs).

In reachability models the service quality of an agent is simply defined as the number of reachable agents and distances are ignored completely. For reachability models the works of Kliemann [Kli11; Kli17; KSS17] and the paper by Goyal et al. [Goy+16] explicitly incorporate a notion of network robustness in the utility function of every agent. All models consider an external adversary who strikes after the network is built. In [Kli11; Kli17] the adversary randomly removes a single edge and the agents try to maximize the expected number of reachable nodes post attack. Two versions of the adversary are analyzed: edge removal uniformly at random or removal of the edge which hurts the society of agents most. For the former adversary a constant Price of Anarchy is shown, whereas for the latter adversary this positive result is only true if edges can be created unilaterally. In [KSS17] the authors extended the previous Kleimann's adversary model to the node-failure case but focus on the swap equilibrium concept only. In [Goy+16] nodes are attacked (and killed) and this attack spreads virus-like to neighboring nodes unless they are protected by a firewall. Interestingly also this model has a low Price of Anarchy and the authors prove a tight linear bound on the amount of edge-overbuilding due to the adversary. A similar model is considered in [Che+19]. Here, once a network is formed, the attack kills one node chosen uniformly at random and then spreads through the network according to the independent cascade model. The authors mostly focus on the structural properties of equilibria, in particular on the edge density. Another interesting result from the paper is that there are equilibrium

2 In centrality models for network formation the agents' service quality in the created network depends on their centrality measure, i.e., the distances to other nodes.

networks with a social welfare that differs by a constant factor from the social welfare of equilibrium without attack.

Despite many similarities between the Adv-NCG and the listed models, the Adv-NCG is a different model. We can illustrate this by a direct comparison with Kliemann's model. The main difference between Kliemann's model and the Adv-NCG is the 2-edge-connectedness. It means, that in Kliemann's model any agent has an individual cost $C_v(S) = |S_v| \cdot \alpha$, if the network is 2-edge-connected. It follows, that a double-clique where each node is connected with each other node by a double edge is a NE in our model for very small α, but it is not a NE in Kliemann's model. Conversely, for any small α a cycle where every agent owns exactly one edge is a NE in Kliemann's model, but it is not a NE for small α in the Adv-NCG.

Also, the Adv-NCG is completely different from the NCG model, even for the NCG restricted on the class of 2-edge-connected networks. This can be illustrated by the following example. Consider the network G depicted in Fig 3.1 (a). It is

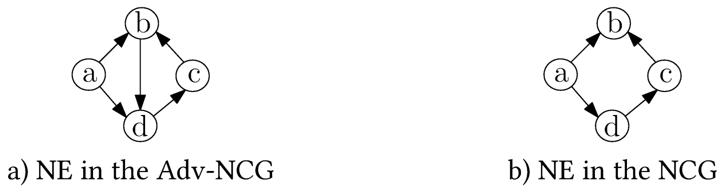


Figure 3.1: Comparison of NE for the NCG and the Adv-NCG models

a NE for the Adv-NCG if $\frac{11}{15} \leq \alpha \leq \frac{7}{5}$. This network is not a NE in the NCG, because agent b could delete the edge bc and thereby decrease her cost from $\alpha + 3$ to 4 if $1 < \alpha \leq \frac{7}{5}$.

On the other hand, it is easy to see, that network G', depicted in Figure 3.1 (b), is a NE in the NCG, but it is not a NE in the Adv-NCG if $1 \leq \alpha \leq \frac{7}{5}$.

The very recent paper by Echzell et al. [Ech+20] considers a notion of robustness in a network creation from a wholly different perspective. The authors present a model where agents seek to maximize the amount of flow to all other agents in a created network. Two versions of the model are considered, where agents maximize the maximum flow sent to other agents, and where agents maximize the average amount of flow in the obtained network. In both versions, agents have a fixed budget to establish connections. Even though the Adv-NCG is comparable with the above model assuming the agents' budget being at least two, these two models are entirely different since the quality of a strategy is measured by different means.

3.3 Optimal Networks

Clearly, every optimal network must be 2-edge connected. Thus, every optimal network must have at least n edges, e.g., imagine a cycle. We first prove the intuitive fact that if edges get more expensive, then the number of edge in the optimum networks cannot increase.

▶ **Theorem 3.1.** Consider two positive values α and α' such that $\alpha' > \alpha$. Let $G = (V, E)$ and $G' = (V, E')$ be optimal networks on n nodes in the Adv-NCG for α and α', respectively. Then $|E| \geq |E'|$. ◀

Proof. We prove the statement by contradiction. Assume that $\alpha' > \alpha$ holds and that network $G' = (V, E')$ has strictly more edges than network $G = (V, E)$. Let $\Delta = |E'| - |E|$ denote this difference.

On the one hand, since G' is an optimal network for edge price α', then the social cost of (G', α') must be at most the social cost of (G, α'). For an edge cost parameter α and a network G, we denote the total edge cost of G in as $edge(G, \alpha)$, the total distance cost as $dist(G)$, and the social cost of G as $\mathrm{SC}(G, \alpha)$. Thus we have

$$\begin{aligned} \mathrm{SC}(G', \alpha') &\leq \mathrm{SC}(G, \alpha') \\ edge(G', \alpha') + dist(G') &\leq edge(G, \alpha') + dist(G) \\ edge(G', \alpha') - edge(G, \alpha') &\leq dist(G) - dist(G') \\ \Delta \cdot \alpha' &\leq dist(G) - dist(G'). \end{aligned}$$

On the other hand, since G is an optimal network for α, we have

$$\begin{aligned} \mathrm{SC}(G, \alpha) &\leq \mathrm{SC}(G', \alpha) \\ edge(G, \alpha) + dist(G) &\leq edge(G', \alpha) + dist(G') \\ dist(G) - dist(G') &\leq edge(G', \alpha) - edge(G, \alpha) \\ dist(G) - dist(G') &\leq \Delta \cdot \alpha. \end{aligned}$$

Hence, we have $\Delta\alpha' \leq dist(G) - dist(G') \leq \Delta\alpha$, which implies $\alpha' \leq \alpha$. This contradicts our assumption that $\alpha' > \alpha$. ■

▶ **Remark 3.2.** Note that the above proof works for any function $d : V \times V \longrightarrow \mathbb{R}_+$ and any edge cost function of the form $\alpha \cdot f(x)$ where $f : V \times V \longrightarrow \mathbb{R}_+$, that is, in particular also for the NCG. ◀

In the following, we show that the landscape of optimum networks is much richer in the Adv-NCG, compared to the NCG where the optimum is either a clique or a star, depending on α. In particular, we prove that there are $\Omega(n^2)$ different optimal topologies. We consider the following types of networks: Here

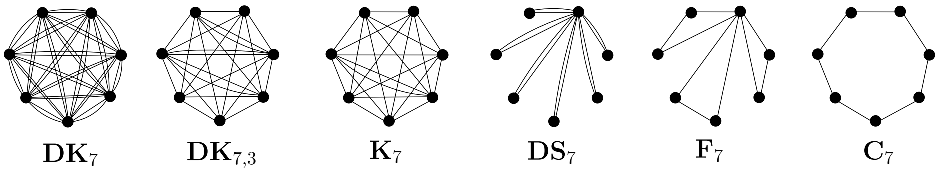


Figure 3.2: Different candidates for optimum networks.

$\mathbf{DK}_n$ is a clique of n nodes where we have a double edge between all pairs of nodes. Let $\mathbf{DK}_{n,k}$ be a n node clique with exactly k pairs of nodes which are connected with double edges. Thus, $\mathbf{DK}_{n,0} = \mathbf{K}_n$ and $\mathbf{DK}_{n,\binom{n}{2}} = \mathbf{DK}_n$. Moreover, let $\mathbf{F}_n$ denote a fan graph on n nodes which is a collection of triangles that all share a single node, and let $\mathbf{DS}_n$ denote a star on n nodes where all connections between the center and the leaves are double edges. Finally, let $\mathbf{C}_n$ be a cycle of length n.

Clearly, if $\alpha = 0$, then the optimum network on n nodes must be a double clique $\mathbf{DK}_n$ since in this network no edge deletion by the adversary has any effect. Moreover, since edges are for free, the distances between the nodes are minimized.

Now consider what happens if one pair of agents, say u and v, are just connected via a single edge instead of a double edge. The probability that the adversary removes this edge is $\frac{1}{n(n-1)-1}$. The removal would cause an increase in distance cost of 1 between u and v and between v and u. Thus, if $\alpha < \frac{2}{n(n-1)-1}$, then, for agent u and v, it would be better to buy another edge between each other. Thus, we have the following observation.

▶ **Observation 3.3.** If $0 \leq \alpha \leq \frac{2}{n(n-1)-1}$, then $\mathrm{OPT_n} = \mathbf{DK}_n$. ◀

▶ **Lemma 3.4.** If $\frac{2n(n-1)}{(\binom{n}{2}+k)(\binom{n}{2}+k+1)} \leq \alpha \leq \frac{2n(n-1)}{(\binom{n}{2}+k)(\binom{n}{2}+k-1)}$, for $1 \leq k \leq \binom{n}{2} - 1$, then the network $\mathbf{DK}_{n,k}$ is a social optimum. For $\frac{4}{\binom{n}{2}+1} \leq \alpha < 2 - \frac{2}{\binom{n}{2}}$, a clique $\mathbf{K}_n$ is a social optimum. ◀

Proof. Consider network $\mathbf{DK}_{n,k} = (V, E)$ with k double edges. Let u_i denote a node with exactly i incident single edges. We have

$$cost(u_i, \mathbf{DK}_{n,k}) = \alpha \cdot |S_{u_i}| + \frac{(|E| - i)(n-1) + i \cdot n}{|E|} = \alpha \cdot S_{u_i} + (n-1) + \frac{i}{|E|},$$

because the distance between u_i and any other node increases only if the adversary deletes any of u_i's incident single edges. The social cost of $\mathbf{DK}_{n,k}$ is $\mathrm{SC}(\mathbf{DK}_{n,k}) = \sum_{i=0}^{\binom{n}{2}-k} a_i \cdot cost(u_i, \mathbf{DK}_{n,k})$, where a_i is the number of nodes having exactly i incident single edges. Note, that $\sum_{i=0}^{\binom{n}{2}-k} a_i = n$. Thus, $\mathrm{SC}(\mathbf{DK}_{n,k})$ is

$$\alpha|E| + \sum_{i=0}^{\binom{n}{2}-k} a_i \left(n - 1 + \frac{i}{|E|}\right) = \alpha|E| + n(n-1) + \frac{1}{|E|} \sum_{i=0}^{\binom{n}{2}-k} a_i \cdot i.$$

Now we simplify the above cost function. Consider the induced sub-graph $G = (V, E')$ of network $\mathbf{DK}_{n,k}$ which contains only the single edges of $\mathbf{DK}_{n,k}$. By the Handshake Lemma, we obtain $\sum_{i=0}^{\binom{n}{2}-k} a_i \cdot i = \sum_{u \in V} deg_G(u) = 2|E'| = 2(\binom{n}{2} - k)$, where $deg_G(u)$ is u's degree in G. Hence, we have

$$\mathrm{SC}(\mathbf{DK}_{n,k}) = \alpha|E| + n(n-1) + \frac{2(\binom{n}{2} - k)}{\binom{n}{2} + k}.$$

Now, if look at the cost difference between $\mathbf{DK}_{n,k-1}$ and $\mathbf{DK}_{n,k}$ we get

$$\mathrm{SC}(\mathbf{DK}_{n,k}) - \mathrm{SC}(\mathbf{DK}_{n,k-1}) = \alpha - \frac{2(\binom{n}{2} - k + 1)}{\binom{n}{2} + k - 1} + \frac{2(\binom{n}{2} - k)}{\binom{n}{2} + k}.$$

The cost difference between $\mathbf{DK}_{n,k}$ and $\mathbf{DK}_{n,k+1}$ is

$$\mathrm{SC}(\mathbf{DK}_{n,k}) - \mathrm{SC}(\mathbf{DK}_{n,k+1}) = -\alpha - \frac{2(\binom{n}{2} - k)}{\binom{n}{2} + k} + \frac{2(\binom{n}{2} - k - 1)}{\binom{n}{2} + k + 1}.$$

Thus, if $\alpha = \alpha_k$ where

$$\frac{2(\binom{n}{2} - k)}{\binom{n}{2} + k} - \frac{2(\binom{n}{2} - k - 1)}{\binom{n}{2} + k + 1} \leq \alpha_k \leq \frac{2(\binom{n}{2} - k + 1)}{\binom{n}{2} + k - 1} - \frac{2(\binom{n}{2} - k)}{\binom{n}{2} + k},$$

then upgrading a single edge to a double edge or downgrading a double edge to a single edge in $\mathbf{DK}_{n,k}$ does not decrease the social cost.

We observe that $\mathrm{OPT_n}$ has diameter 1. Indeed, consider a network $G'' = (V, E'')$ that has diameter 2. Then there are two agents u and v with the expected distance at least 2 between each other. The new edge uv makes their expected distance equal to $1 + \frac{1}{|E''|+1}$. This yields a decrease in social distance cost of at least $2\left(2 - \left(1 + \frac{1}{|E''|+1}\right)\right) = 2 - \frac{2}{|E''|+1} \geq 2 - \frac{2}{n}$. Since $\alpha_k < \frac{8}{n^2} < 2 - \frac{2}{n}$ for $n \geq 3$, it follows that $\mathrm{OPT_n}$ for $\alpha = \alpha_k$ has diameter 1.

Since for α_k $\mathrm{OPT_n}$ has diameter 1, we know that we can obtain $\mathrm{OPT_n}$ from $\mathbf{DK}_{n,k}$ by either downgrading some double edges to single edges or by upgrading some single edges to double edges. We have chosen α_k such that downgrading one double edge to a single edge or upgrading one single edge to a double edge does not decrease the social cost. Since downgrading or upgrading edges only affects the distance costs of the incident agents it follows that if downgrading or upgrading one edge does not decrease the social cost, then downgrading or upgrading more than one edge cannot decrease the social cost as well. Thus, $\mathbf{DK}_{n,k}$ is the optimal network for α_k.

Finally, if $\frac{4}{\binom{n}{2}+1} \leq \alpha < 2 - \frac{2}{\binom{n}{2}}$, a social optimum has diameter 1 and has no double edges, i.e., it is a clique. ■

We also remark that we conjecture that Figure 3.2 resembles a snapshot of optimum networks for increasing α from left to right. In fact, extensive simulations indicate that the optimum changes from G_n to DS_n and then, for slightly larger α to $\mathbf{F}_n$. After this the cycles in the fan-graph increase and get fewer in number until, finally, for $\alpha \in \Omega(n^3)$ the cycle appears as optimum.

3.4 Computing Best Responses and Game Dynamics

In this section we investigate computational aspects of the Adv-NCG. First we analyze the hardness of computing a best response and the hardness of computing a best possible multi-swap. Then we analyze a natural process for finding an equilibrium network by sequentially performing improving moves.

3.4.1 Hardness of Best Response Computation

We first introduce useful properties for ruling out multi-buy or multi-delete moves. The proof is similar to the proof of Lemma 1 in [Len12].

▶ **Proposition 3.5.** If an agent cannot decrease her expected cost by buying (deleting) one edge in the Adv-NCG, then buying (deleting) $k > 1$ edges cannot decrease the agent's expected cost. ◀

▶ **Lemma 3.6.** Consider an agent u in a network $G = (V, E)$. If $1 - \frac{1}{|E|+1} < \alpha < 1 + \frac{1}{|E|(|E|-1)}$ and if the agent u is not an endpoint of any double-edge in G, then buying the minimum number of edges such that u's expected distance to all nodes in $V \setminus N_1(u)$ is at most 2 and to nodes in $N_1(u)$ is $1 + \frac{1}{|E|}$ is u's best response. ◀

Proof. Consider a network $G = (V, E)$ where u's expected distance to all nodes in $V \setminus N_1(u)$ is 2 and to all nodes in $N_1(u)$ is $1 + \frac{1}{|E|}$.

Buying an additional edge to some $v \in N_1(u)$ in G creates a double edge, i.e., it decreases u's expected distance to v by $1 + \frac{1}{|E|} - \frac{|E|+1}{|E|+1} = \frac{1}{|E|}$. Buying an edge towards a node $w \notin N_1(u)$ decreases u's expected distance to w by $1 - \frac{1}{|E|+1} > \frac{1}{|E|}$. Thus if $\alpha > 1 - \frac{1}{|E|+1}$, then buying a single edge does not decrease u's expected cost. Thus by Proposition 3.5, agent u cannot improve her expected cost in G by buying more than one edge.

Swapping an edge to some $v \in N_1(u)$ decreases u's expected distance to v by $\frac{1}{|E|}$ but increases u's expected distance to some $w \in N_1(u)$ by $1 - \frac{1}{|E|}$. Swapping an edge towards a node $w \notin N_1(u)$ decreases u's expected distance to w by $1 - \frac{1}{|E|}$ but increases the expected distance to $w \in N_1(u)$ by at least $1 - \frac{1}{|E|}$.

Assume u has bought the minimum number of edges such that u's expected distance to all nodes in $V \setminus N_1(u)$ is 2 and to all nodes in $N_1(u)$ is $1 + \frac{1}{|E|}$. Then deleting an edge uv for some $v \in N_1(u)$ increases u's expected distance to v by at least $1 + \frac{1}{|E|(|E|-1)}$ since after deleting the edge the expected distance between u and v is $2 + \frac{1}{|E|-1}$. Thus, if $\alpha < 1 + \frac{1}{|E|(|E|-1)}$, then deleting a single edge does not decrease u's expected cost. Thus, by Proposition 3.5, agent u cannot decrease her expected cost by deleting more than one edge. ■

Now we show that computing the best possible strategy-change is intractable.

▶ **Theorem 3.7.** The problem of computing a best response is NP-hard in the Adv-NCG. ◀

Proof. We prove both statements by reduction from MINIMUM-m-CONNECTED k-DOMINATING SET (Min-(m, k)-CDS) [Sha+07] which is defined as follows: Given a network $G = (V, E)$ and two natural numbers m and k, find a subset $S \subseteq V$ of minimum size such that every vertex in V/S is adjacent to at least k nodes in S and the induced sub-graph of S is m-connected.

More precisely, we provide the reduction from Min-$(1, 2)$-CDS. Consider an instance $G' = (V \cup \{u\}, E)$. We prove that the best response for agent u corresponds to the minimum 1-connected 2-dominating set S. For any $1 - \frac{1}{|E|+1} < \alpha < 1 + \frac{1}{|E|(|E|-1)}$, and since no other agent has an edge to u, then, by Lemma 3.6, the best response for u is to buy edges to all the nodes in the S. Indeed, in that case the expected eccentricity of u will be at most 2 since every node $w \notin S$ is adjacent to at least two nodes in S. Moreover, since S is connected, u has expected distance at most $1 + \frac{1}{|E|}$ to all nodes in S. ■

3.4.2 Game Dynamics

For the game dynamics of the Adv-NCG we prove the strongest possible negative result, which essentially shows that there is no hope for convergence if agents stick to performing improving moves only. In particular, we prove that the order of the agents' moves or any tie-breaking between different improving moves does not help for achieving convergence. This result is even stronger than the best known non-convergence results for the NCG [KL13].

▶ **Theorem 3.8.** The Adv-NCG is not weakly acyclic. ◀

Proof sketch. We provide a best response cycle $G_1, \ldots, G_7$, where $G_1 = G_7$ and G_{i+1} is obtained from G_i by an improving move of one agent in G_i. Our best response cycle has the special features that in every step of the cycle there is exactly one agent who can perform an improving move and that this improving move is unique. Thus, starting with G_1, any sequence of improving moves must be infinite. The best response cycle on 10 agents with $\alpha = 10.3$ is depicted in Figure 3.3. We omit the quite lengthy proof that the shown best response cycle behaves as indicated. ∎

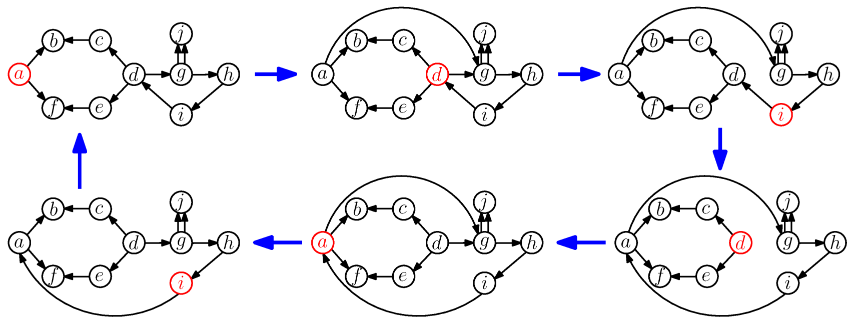


Figure 3.3: A best response cycle for $\alpha = 10.3$, where in every step only the red agent has an improving move and this improving move is unique.

3.5 Analysis of Networks in Nash Equilibrium

In this section we establish the existence of networks in Nash equilibrium for almost the whole parameter space. Moreover, we investigate structural properties which allow us to provide bounds on the Price of Stability and the Price of Anarchy.

We start with the existence result.

▶ **Theorem 3.9.** The networks $\mathbf{DK}_n$ and DS_n are in pure Nash Equilibrium if $\alpha \leq \frac{1}{n(n-1)-1}$ and $\alpha \geq 1 - \frac{1}{2n-1}$, respectively. ◀

Proof. We start with proving that the double clique network $\mathbf{DK}_n$ where every agents owns an edge to all other agents is in Nash Equilibrium if $\alpha \leq \frac{1}{n(n-1)-1}$.

Clearly, in $\mathbf{DK}_n$ no agent can improve her expected distance cost by buying one or more edges since each agent already has the minimum possible expected distance cost. The same holds true for performing edge-swaps: since the edge cost stays the same and the expected distance cost increases by swapping edges, no agent can improve by swapping one or more edges. It remains to analyze the edge deletions. In $\mathbf{DK}_n$ every agent has expected distance cost $n-1$. Deleting $1 \leq k \leq n-1$ edges yields expected distance cost of

$$\frac{(n(n-1)-2k)(n-1)+kn}{n(n-1)-k} = n-1+\frac{k}{n(n-1)-k} \geq n-1+\frac{1}{n(n-1)-1}.$$

Thus, if $\alpha \leq \frac{1}{n(n-1)-1}$, then deleting one or more edges is not an improving move.

Next, we show that the double star DS_n with an arbitrary edge-ownership is in Nash equilibrium if $\alpha \geq 1-\frac{1}{2n-1}$.

Clearly, no agent can delete edges since this destroys the 2-edge-connectedness of the network and hence induces infinite cost. Moreover, no agent can swap edges since this does not change the edge cost but increases the expected distance cost. Thus, we are left to analyze edge purchases. Clearly the center of the double star cannot buy edges to decrease its cost. Hence, we analyze edge additions by non-center nodes of DS_n. Every such agent has expected distance cost of $1+2(n-2)=2n-3$.

Let u be a non-center agent and let S_u be u's current strategy in $G := DS_n$. Assume that agent u can change her strategy from S_u to S_u' and thereby strictly decrease her cost. Let DS_n' be the network G after u's strategy-change from S_u to S_u'. We claim that if $\alpha > \frac{1}{n-1}$ and if G' contains at least three edges between the center node and u or if there are at least two edges between u and some other non-center node v, then agent u has a strategy S_u'', which strictly outperforms strategy S_u' and where the corresponding network G'' has exactly two edges between the center node and u and at most one edge between u and any other non-center node. Thus, we can assume that if agent u has an improving strategy-change, then there exists an improving strategy-change towards a strategy which buys only additional single edges towards other non-center nodes. After proving the above claim, we will prove that no such improving strategy-change exists if $\alpha \geq 1-\frac{1}{2n-1} > \frac{1}{n-1}$, which then implies that DS_n is in Nash Equilibrium for all $\alpha \geq 1-\frac{1}{2n-1}$.

Now we prove the claim: We first show that strategy S_u' can be improved if G' contains at least three edges between the center node and u. In this case this implies that u owns at least one edge to the center node and that agent u could remove one edge from G' to ensure that at least two edges between the agent and the center node remain. Let G'' be the network G' after the edge-removal, and let S_u'' be the new strategy of u. This removal would save α in the edge cost. If u has no single edges towards any non-center vertex, then her expected distance cost in G'' would not increase compared to her expected distance cost in G' by the edge-removal since all edges on all her shortest paths are backed up by another parallel edge. Since $\alpha > 0$, strategy S''_u strictly outperforms strategy S'_u. If u has $1 \leq k \leq n-2$ single edges towards k different non-center nodes in G', then the edge-removal of one edge between u and the center node increases the probability that one of the k edges is destroyed by the adversary. The probability increases by

$$\frac{k}{m(m-1)} \leq \frac{n-2}{2n(2n-1)} < \frac{1}{n-1},$$

where $m \geq 2n$ is the number of edges in G'. Thus, agent u's expected distance cost in G'' increases by at most $\frac{1}{4n}$ compared to her expected distance cost in G'. Since $\alpha > \frac{1}{n-1}$, it follows that S''_u strictly outperforms S'_u. If G'' contains more than three edges between u and the center vertex, then we can apply the above argument iteratively to obtain a strategy S''_u which strictly outperforms S'_u and a corresponding network G'' which has exactly two edges between u and the center vertex.

Now we show that strategy S'_u can be improved if G' contains at least two edges between u and some other non-center node v. Note that in this case all edges between u and v are bought by agent u.

It is possible that in the network G' there is no edge or only one edge between u and the center node. If there is no edge between u and the center node, then agent u could swap two edges from v to the center node and thereby strictly decrease her cost. This is true since this swap would decrease u's expected distance to every node $w \neq v$ by at least 1, and it only increases her expected distance to v by 1. If there is exactly one edge between u and the center node, then agent u could swap one edge from v to the center node and thereby decrease her cost. This swap may create a single edge towards v, but if this edge is attacked by the adversary, then this only increases u's distance to v by 1, whereas in G' an attack on the single edge between u and the center node increases u's distances to $n-2$ nodes by at least 1. Hence, if there is no edge or only one edge between u and the center node, then in both cases there is a strategy S''_u which strictly outperforms strategy S'_u and where the corresponding network G'' has exactly two edges between u and the center node. Thus, we will assume in the following that there are exactly two edges between u and the center vertex and at least two edges between u and some non-center node v.

Let G'' be the network obtained from network G' by removing one of the edges between u and v and let S''_u be u's strategy obtained by removing v from S'_u. If there are at least two edges between u and v in G'', then an analogous argument as above yields that S''_u strictly outperforms S'_u if $\alpha > \frac{1}{n-1} > \frac{n-3}{2n(2n+1)}$. Note that in this case G' has at least $2n+1$ many edges. If there is a single edge between u and v in G'', then the number of non-center nodes to which u has a single edge increases by 1 from k to $k+1$ for some $0 \leq k \leq n-3$. Thus, her expected distance cost compared to network G' increases by

$$\frac{k+1}{2n-1} - \frac{k}{2n} = \frac{2n+k}{2n(2n-1)} \leq \frac{1}{2n-1} + \frac{n-3}{2n(2n-1)} < \frac{1}{n-1},$$

which implies that S''_u strictly outperforms strategy S'_u if $\alpha > \frac{1}{n-1}$.

Having settled the claim, we now analyze the case where a non-center agent u buys $1 \leq k \leq n-2$ single edges to k other non-center nodes. In this case u's expected distance cost is

$$\frac{2(n-1)(k+1+2(n-2-k)) + k(k+2(n-2-k+1))}{2(n-1)+k}$$
$$=2n-3-k+\frac{k}{2(n-1)+k}.$$

Since $-k + \frac{k}{2(n-1)+k} \geq -1 + \frac{1}{2(n-1)+1} = -1 + \frac{1}{2n-1}$ for $1 \leq k \leq n-2$, it follows that the expected distance cost after buying $1 \leq k \leq n-2$ single edges is at least $2n-3-1+\frac{1}{2(n-1)+1}$. Thus if $\alpha \geq 1 - \frac{1}{2n-1}$, then buying one or more single edges is not an improving move for any non-center agent.

Since $\alpha \geq 1 - \frac{1}{2n-1} > \frac{1}{4n}$, it follows that no non-center vertex can buy one or more edges in network DS_n to strictly decrease her cost. ■

3.5.1 Relation between the Diameter and the Social Cost

We prove a property which relates the diameter of a network with its social cost. With this, we prove that one of the most useful tools for analyzing NE in the NCG [Fab+03] can be carried over to the Adv-NCG.

Before we start, we analyze the diameter increase induced by removing a single edge in a 2-edge-connected network.

▶ **Lemma 3.10.** Let $G = (V, E)$ be any 2-edge-connected network having diameter D and let $G - e$ be the network G where some edge $e \in E$ is removed. Then the diameter of $G - e$ is at most $2D$. ◀

Proof. Let $e = uv$ be the edge which is removed from a diameter D network G. Consider any shortest path P in G which uses edge e somewhere along the path. Let x and y be the endpoints of path P and we assume that u and v are the endpoints of e which are closer to x and y, respectively.

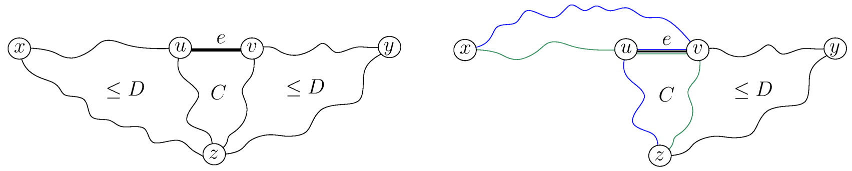


Figure 3.4: Left: paths P_{xz} and P_{zy} do not contain e. Right: path P_{xz} goes over P_{xv}, e, and P_{uz} (blue line) or it goes over P_{xu}, e, and P_{vz} (green line).

Since G is 2-edge-connected, we can find a smallest induced cycle C in G which includes edge e. Let z be a node in the cycle C which has maximum distance to edge e, that is, maximum distance to both u and v simultaneously. There exists a shortest path P_{xz} in G which connects x and z, and there is a shortest path P_{zy} in G which connects z and y. Both paths have length at most D. See Figure 3.4(left) for an illustration. Observe, that it is impossible that both paths P_{xz} and P_{zy} contain edge e, since otherwise there must be a shorter path between x and z or z and y. If both of the paths do not contain edge e, then $P_{xz} \cup P_{zy}$ is a path between nodes x and y in the network $G - e$ and it has a length at most $2D$.

Finally, consider a situation where exactly one of the paths contains edge e in G and let this be path P_{xz}. Thus, we have that $P_{xz} = P_{xv} \cup \{v, u\} \cup P_{uz}$ or $P_{xz} = P_{xu} \cup \{u, v\} \cup P_{vz}$ in graph G and both of the paths P_{uz} and P_{vz} are parts of the cycle C. See Figure 3.4(right) for an illustration of the network. By choice of z, it follows that $P_{xv} \cup P_{vz}$ or $P_{xu} \cup P_{uz}$ is a path between x and z in graph $G - e$ which has length at most D. Since the path P_{zy} does not contain edge e, it follows that it can be used in network $G - e$. Since P_{zy} has length at most D, we have that the distance between nodes x and y in $G - e$ is at most $2D$. ■

Next, we will focus on edges which are part of cuts of the network of size 2. Remember that a bridge is an edge whose removal from a network increases the number of connected components of that network. Let $G = (V, E)$ be any

2-edge-connected network. We say that an edge $e \in E$ is a 2-*cut-edge* if there exists a cut of G of size 2 which contains edge e. Equivalently, e is a 2-cut-edge of G if its removal from G creates at least one bridge in $G - e$. We now bound the number of 2-cut-edges in any 2-edge-connected network G. This is an important structural result, since this proves that the amount of edge-overbuilding due to the adversary is sharply limited.

▶ **Lemma 3.11.** Any 2-edge-connected network G with n nodes can have at most $2(n-1)$ edges which are 2-cut-edges. ◀

Proof. Let e be any 2-cut-edge in network G. By definition, the removal of e creates one or more bridges in $G - e$. Let $b_1, \ldots, b_l$ denote those bridges. Note, that $b_1, \ldots, b_l$ also must be 2-cut-edges in G. Moreover, it follows that there must be a shortest cycle C in G which contains all the edges $e, b_1, \ldots, b_l$. If there is more than one such cycles, then fix one of them. We call the fixed cycle C a cut-cycle.

Notice that any 2-cut-edge corresponds to exactly one cut-cycle in the network and that every cut-cycle contains at least two 2-cut-edges. We show in the following that if any cut-cycle in the network contains at least three 2-cut-edges, then we can modify the network to obtain strictly more 2-cut-edges and strictly more cut-cycles. This implies that the number of 2-cut-edges is maximized if the number of cut-cycles is maximized, and every cut-cycle contains exactly two 2-cut-edges.

Now we describe the procedure which converts any network with at least one cut-cycle containing at least three 2-cut-edges into a modified network with a strictly increased number of cut-cycles and 2-cut-edges (see Figure 3.5).

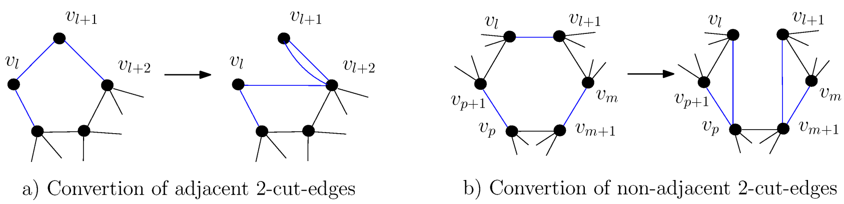


Figure 3.5: Increasing the number of 2-cut-edges by splitting up a cut-cycle.

Let G be any network with at least one cut-cycle $C = v_1, \ldots, v_k, v_1$ containing at least three 2-cut-edges. If there are two adjacent 2-cut-edges $v_l v_{l+1}, v_{l+1} v_{l+2}$ in cycle C, then delete the 2-cut-edge $v_l v_{l+1}$ and insert two new edges $v_l v_{l+2}$ and

$v_{l+1}v_{l+2}$. First of all, note that these new edges have not been present in network G before the insertion since otherwise $\{v_l v_{l+1}, v_{l+1}v_{l+2}\}$ cannot be a cut of G. We claim that both new edges are 2-cut-edges and that the cycle C is divided into two new cut-cycles $v_1, \ldots, v_l, v_{l+2}, \ldots, v_k, v_1$ and $v_{l+1}, v_{l+2}, v_{l+1}$. Indeed, there are at least two bridges $v_{l+1}v_{l+2}$ and $v_k v_{k+1}$ in the cut-cycle C after deleting $v_l v_{l+1}$, and both of them end up in different new cut-cycles. Hence, deleting any of the newly inserted edges $v_l v_{l+2}$ or $v_{l+1}v_{l+2}$ implies that $v_k v_{k+1}$ or $v_{l+1}v_{l+2}$ becomes a bridge. Thus, both new edges are 2-cut-edges and both of the new cycles are cut-cycles.

If there are three pairwise non-adjacent 2-cut-edges $v_l v_{l+1}, v_m v_{m+1}, v_p v_{p+1}$ in cycle C, then delete one 2-cut-edge $v_l v_{l+1}$ and insert two new edges $v_l v_{p+2}$ and $v_{l+1}v_{m+1}$. Analogous to above, both new edges cannot be already present in G and both are 2-cut-edges because deleting any of them renders edge $v_m v_{m+1}$ or $v_p v_{p+1}$ a bridge. Moreover, cut-cycle C is divided into two new cut-cycles.

Finally, we claim that the maximum number of cut-cycles in any n-node network G is at most $n-1$. Since we know that every such cut-cycle contains exactly two 2-cut-edges this then implies that there can be at most $2(n-1)$ 2-cut-edges in any network G.

Now we prove the above claim. Note that applying our transformation does not disconnect the network. Thus, we know that network G after all transformations is connected. Now we iteratively choose any cut-cycle C in G and we delete the two 2-cut-edges contained in C. This deletion increases the number of connected components of the current network by exactly 1. We repeat this process until we have destroyed all cut-cycles in G. Note that deleting edges from G may create new cut-cycles, but we never destroy more than one of them at a time. Thus, since each iteration increases the number of connected components of the network by 1, it follows that there can be at most $n-1$ iterations since network G with n nodes cannot have more than n connected components. ■

▶ **Remark 3.12.** Lemma 3.11 is tight, since a path of length $n-1$, where all neighboring nodes are connected via double edges, has exactly $2(n-1)$ 2-cut-edges. ◀

Now we relate the diameter with the social cost.

▶ **Theorem 3.13.** Let $G = (V, E)$ be any NE network on n nodes having diameter D and let $\mathrm{OPT}_{\mathrm{n}n}$ be the corresponding optimum network for some $\alpha > 0$. Then $\frac{\mathrm{SC}(G)}{\mathrm{SC}(\mathrm{OPT}_{\mathrm{n}n})} \in \mathcal{O}(D)$. ◀

Proof. Since $\mathrm{OPT_n}$ is 2-edge-connected, it has at least n edges. Moreover, the minimum expected distance between each pair of nodes in $\mathrm{OPT_n}$ is at least 1. Thus, we have that $\mathrm{SC}(OPT_n) \in \Omega(\alpha \cdot n + n^2)$.

Now we analyze the social cost of the NE network $G = (V, E)$. We will evaluate the edge cost and the distance cost of G separately.

Since G has diameter D, and since G is 2-edge-connected, Lemma 3.10 implies that the expected distance between each pair of nodes in G is at most $2D$. Thus, we have that $O(n^2 \cdot D)$.

Now we analyze the edge cost. By Lemma 3.11 we have at most $2n$ many 2-cut-edges in G. Buying all those edges yields cost of at most $2n \cdot \alpha$.

We proceed with bounding the number of non-2-cut-edges in G. We consider an agent v and analyze how many non-2-cut-edges agent v can have bought. We claim that this number is in $O(\frac{nD}{\alpha})$, which yields the total edge cost of $O(nD)$ for agent v. Summing up over all n agents, this yields the total edge cost of $O(n^2D)$ for all non-2-cut-edges of G. This implies an upper bound of $O(\alpha \cdot n + n^2D)$ on the social cost of G which finishes the proof.

Now we prove our claim. Fix any non-2-cut-edge $e = vw$ of G which is owned by agent v. Let $V_e \subset V$ be the set of nodes of G to which all shortest paths from v traverse the edge e.

We first show that removing the edge e increases agent v's expected distance to any node in V_e to at most $4D$. By Lemma 3.10, removing edge e increases the diameter of G from D to at most $2D$. Since e is a non-2-cut-edge, we have that $G - e$ is still 2-edge-connected. Thus, again by Lemma 3.10, it follows that agent v's expected distance to any other node in $G - e$ is at most $4D$.

However, removing edge e not only increases v's expected distance towards all nodes in V_e, instead, since $G - e$ has a fewer edges than G, agent v's expected distance to *all* other nodes in $V \setminus (V_e \cup \{v\})$ increases as well. We now proceed to bound this increase in the expected distance cost.

We compare agent v's expected distance cost in network G and in network $G - e$. Let m denote the number of edges in G. Thus, $G - e$ has $m - 1$ many edges. For network G agent v's expected distance cost is

$$\delta_G(v, V) = \frac{1}{m} \sum_{f \in E} d_{G-f}(v, V) = \frac{1}{m} \sum_{f \in E \setminus \{e\}} d_{G-f}(v, V) + \frac{d_{G-e}(v, V)}{m}.$$

In network $G - e$, we have $\delta_{G-e}(v, V) = \frac{1}{m-1} \sum_{f \in E \setminus \{e\}} d_{G-e-f}(v, V)$. Now we upper bound the increase in expected distance cost for agent v due to removal of edge e from G. The expected distance cost difference $\delta_{G-e}(v,V) - \delta_G(v,V)$ is

$$\frac{1}{m-1}\sum_{f\in E\setminus\{e\}} d_{G-e-f}(v,V) - \left(\frac{1}{m}\sum_{f\in E\setminus\{e\}} d_{G-f}(v,V) + \frac{d_{G-e}(v,V)}{m}\right)$$
$$= \sum_{f\in E\setminus\{e\}}\left(\frac{d_{G-e-f}(v,V)}{m-1} - \frac{d_{G-f}(v,V)}{m}\right) - \frac{d_{G-e}(v,V)}{m}.$$

We have that $d_{G-e-f}(v,V) \le d_{G-f}(v,V) + |V_e|\cdot 4D$, since in $G-e-f$ only the distances to nodes in V_e increase, compared to the network $G-f$ and since e is a non-2-cut-edge in G. Moreover, by Lemma 3.10, the distances to nodes in V_e in $G-e-f$ increase to at most $4D$ for each node in V_e. Thus, we have that

$$\begin{aligned}
&\delta_{G-e}(v,V) - \delta_G(v,V)\\
&= \sum_{f\in E\setminus\{e\}}\left(\frac{d_{G-e-f}(v,V)}{m-1} - \frac{d_{G-f}(v,V)}{m}\right) - \frac{d_{G-e}(v,V)}{m}\\
&\le \sum_{f\in E\setminus\{e\}}\left(\frac{d_{G-f}(v,V) + |V_e|4D}{m-1} - \frac{d_{G-f}(v,V)}{m}\right)\\
&= |V_e|4D + \sum_{f\in E\setminus\{e\}}\frac{d_{G-f}(v,V)}{m(m-1)} \le |V_e|4D + \sum_{f\in E\setminus\{e\}}\frac{2D\cdot n}{n(m-1)}\\
&\le |V_e|4D + 4D = (|V_e|+1)4D.
\end{aligned}$$

Since G is in Nash Equilibrium, we know that removing edge e is not an improving move for agent v. Thus, we have that

$$\alpha \le (|V_e|+1)4D \iff |V_e| \ge \frac{\alpha}{4D} - 1.$$

Hence, for all non-2-cut-edges e which are bought by agent v, we have that $|V_e| \in \Omega(\frac{\alpha}{D})$. Since all these sets V_e are disjoint, it follows that v can have bought at most $\frac{n}{\Omega(\frac{\alpha}{D})} \in O(\frac{nD}{\alpha})$ many non-2-cut-edges. ■

3.5.2 Price of Stability and Price of Anarchy

▶ **Theorem 3.14.** For $\alpha \leq \frac{1}{n(n-1)-1}$, PoS = 1. For $\frac{1}{n(n-1)-1} < \alpha < \frac{2}{n(n-1)-1}$, the PoS is strictly larger than 1, and for $\alpha > \frac{1}{n(n-1)-1}$, the PoS is at most 2. ◀

Proof. By Theorem 3.9 and Observation 3.3, network $\mathbf{DK}_n$ is optimal, and it is a Nash equilibrium when $\alpha \leq \frac{1}{n(n-1)-1}$. Thus, the Price of Stability is 1 for this range of α.

Consider the case when $\frac{1}{n(n-1)-1} < \alpha < \frac{2}{n(n-1)-1}$. Network $\mathbf{DK}_n$ is the unique optimum for this range of alpha but any agent can delete an edge and thereby increase her expected distance cost by $\frac{1}{n(n-1)-1}$. Thus, if $\frac{1}{n(n-1)-1} < \alpha \leq \frac{2}{n(n-1)-1}$, the edge-deletion is an improving move which shows that $\mathbf{DK}_n$ is not a NE, and the PoS is strictly larger than 1.

We observe that for $\alpha < \frac{2}{n(n-1)-1}$, any NE has diameter 1. Indeed, consider a network $G = (V, E)$ such that there are two agents u and v at distance 2. The expected distance between the agents in G is at least 2. Hence, the addition of the edge uv decreases the expected distance by $2 - \frac{|E|+2}{|E|+1} = 1 - \frac{1}{|E|+1}$. Therefore, since the cost of the new edge $\alpha < \frac{2}{n(n-1)-1} \leq 1 - \frac{1}{n} \leq 1 - \frac{1}{|E|+1}$, the addition of uv is profitable. It contradicts the assumption that G is in NE. It implies that the PoS can be upper bounded by the ratio

$$\frac{\mathrm{SC}(\mathbf{K}_n)}{\mathrm{SC}(\mathbf{DK}_n)} = \frac{\frac{1}{2}\alpha \cdot n(n-1) + n^2}{\alpha \cdot n(n-1) + n(n-1)} \leq 2.$$

The third part of the statement follows from Theorem 3.9 and the simple lower bound on the expected social cost of the optimum from the proof of Theorem 3.13. Thus, for $\alpha > 1 - \frac{1}{2n-1}$ the PoS is at most $\frac{cost(DS_n)}{n\alpha+n^2} = \frac{2(n-1)\alpha+2(n-1)^2}{n\alpha+n^2} \leq 2$. ■

We now show how to adapt two techniques from the NCG for bounding the diameter of equilibrium networks to our adversarial version. This can be understood as a proof of concept showing that the Adv-NCG can be analyzed as rigorously as the NCG. However, carrying over the currently strongest general diameter bound of $2^{O(\sqrt{\log n})}$ due to Demaine et al. [Dem+12], which is based on interleaved region-growing arguments seems challenging due to the fact that we can only work with expected distances.

We start with a simple diameter upper bound based on [Fab+03].

▶ **Theorem 3.15.** The diameter of any NE network is in $O(\sqrt{\alpha})$. ◀

Proof. We prove the statement by contradiction. Assume that there are agents u and v in network G with $d_G(u,v) \geq 4\ell$, for some ℓ. Since expected distances cannot be shorter than distances in G, it follows that u's expected distance to v is at least 4ℓ. If u buys an edge to v for the price of α then u's decrease in expected distance cost is at least $\frac{|E|}{|E|+1}(4\ell - 1 + 4\ell - 3 + \cdots + 1) = \frac{|E|}{|E|+1} 2\ell^2$.

Thus, if $d_G(u,v) > 4\sqrt{\alpha}$, then $u's$ decrease in expected distance cost by buying the edge uv is at least $\frac{|E|}{|E|+1} 2\alpha > \alpha$. Thus, if the diameter of G is at least $4\sqrt{\alpha}$, then there is some agent who has an improving move. ∎

Together with Theorem 3.13 this yields the following statement:

▶ **Corollary 3.16.** The Price of Anarchy in the Adv-NCG is in $O(\sqrt{\alpha})$. ◀

Next, we show how to adapt a technique by Albers et al. [Alb+14] to get a stronger statement, which implies constant PoA for $\alpha \in O(\sqrt{n})$.

▶ **Theorem 3.17.** The Price of Anarchy in the Adv-NCG is in $O\left(1 + \frac{\alpha}{\sqrt{n}}\right)$. ◀

Proof. We use Theorem 3.13 and give an improved bound on the expected diameter of any NE network. Let d be the expected diameter of the network. Consider two nodes u and v which have expected distance d.

Let B be the set of nodes in the network which are at expected distance of $d' = \left\lfloor \frac{d-1}{8} \right\rfloor$ from node u. First, we analyze the change in expected distance cost of agent v if she buys an edge towards u. Consider any node $w \in B$. By Lemma 3.10 we have that without edge vu agent v has expected distance of at least $\frac{d}{2} - d'$ towards w. After buying the edge vu, agent v's expected distance to w is at most $\frac{(1+d')|E|+d}{|E|+1}$. Thus, agent v's expected distance to w decreases by at least

$$\frac{d}{2} - d' - \left(\frac{|E|d' + |E| + d}{|E| + 1}\right) \geq \frac{d}{2} - 2d' - 2 > \frac{d-8}{4}.$$

It implies that after buying the edge vu agent v's expected distance cost decreases by at least $\frac{d-8}{4}|B|$. Since G is in NE, it follows that $\alpha \geq (\frac{d-8}{4})|B|$.

Now consider node u which has expected distance of at most d' to any node B. Thus, by Lemma 3.10, and since $d_G(u,v) \geq \frac{d}{2}$, we know that there must be nodes $w \in B$ with $d_G(u,w) = \frac{d'}{2}$. Let B' be a set of all nodes w in B with $d_G(u,w) \leq \frac{d'}{2}$. For any node $w \in B'$ denote

$$S_w := \{x \mid w \text{ is the last node in } B' \text{ on a shortest path from } u \text{ to } x\}.$$

If S_w is non-empty, then $d_G(u,w)$ is $\frac{d'}{2}$. Since there are $n - |B'|$ nodes outside of B' it follows that there must be some node w with $|S_w| \geq \frac{n-|B'|}{|B'|}$. If u buys the edge $\{u,w\}$, then u's expected distance cost decreases by at least

$$\left(\frac{d'}{2} - \frac{|E| + \frac{d'}{2}}{|E|+1}\right)|S_w| \geq \left(\frac{d'}{4} - \frac{1}{2}\right)|S_w|.$$

Since G is an NE, it follows that $\alpha \geq \left(\frac{d'}{4} - \frac{1}{2}\right)|S_w| \geq \left(\frac{d'}{4} - \frac{1}{2}\right)\frac{n-|B'|}{|B'|}$. By rearranging we get

$$|B'|2\alpha \geq |B'|\left(\alpha + \left(\frac{d'}{4} - \frac{1}{2}\right)\right) \geq \left(\frac{d'}{4} - \frac{1}{2}\right)n,$$

where the first inequality holds since $\alpha \geq d > \frac{d'-2}{4}$ because G is in NE. Thus, we have $|B| \geq |B'| \geq (d'-2)\frac{n}{8\alpha}$.

From $\alpha \geq (\frac{d-8}{4})|B|$, we get $\alpha \geq \left(\frac{d-8}{4}\right)\left(\frac{d'-2}{8\alpha}\right)n \iff 8\alpha^2 > (\frac{d}{2} - 2)(d'-2)n$. Since $\frac{d}{2} > d'$ we have

$$8\alpha^2 \geq (d'-2)^2 n \iff \sqrt{\frac{8}{n}}\alpha \geq d' - 2 \geq \frac{d-1}{8} - 3.$$

Hence, we have $25 + \frac{8\sqrt{8}\alpha}{\sqrt{n}} \geq d$. ■

▶ **Theorem 3.18.** The Price of Anarchy of the Adv-NCG is at least 2 and for very large α this bound is tight. ◀

Proof. Consider an arbitrary large α, e.g., $\alpha = 2^n$. In that case the optimum network must be a cycle whereas, by Theorem 3.9, the double-star network DS_n is in Nash Equilibrium for this α. Since DS_n has $2(n-1)$ edges and since in this range of alpha the edge cost term dominates the social cost, the lower bound follows. The tight upper bound for large α follows from Lemma 3.11, since 2-cut-edges cannot be deleted without creating a bridge. ■

3.6 Conclusion

We presented a simple and accessible model for selfish network creation incorporating both centrality and robustness aspects. In essence we proved that many properties and techniques can be carried over from the non-adversarial NCG, and we indicated that the landscape of optimum and equilibrium networks in the Adv-NCG is much more diverse than without adversary. As for the NCG, proving strong upper or lower bounds on the PoA is very challenging. Especially surprising is the hardness of constructing higher lower bounds than in the NCG since by introducing suitable gadgets it is always possible to enforce that no agent wants to swap or delete edges. A non-constant lower bound on the PoA seems possible if α is linear in n.

It would also be interesting to consider different adversaries. An obvious candidate for this is node-removal model. Another promising choice is a local adversary, where every agent considers that some of her incident edges may fail. This local perspective combined with a centrality aspect could explain why many selfishly built networks have a high clustering coefficient.

Another direction is to consider the swap version [Ehs+15; MS10] of the Adv-NCG, especially in the case where all agents own at least 2 edges. We note in passing, that the swap-version of the Adv-NCG is not a potential game. The following improving move cycle shows that even if agents are only allowed to perform multi-swaps, then infinite sequences of improving moves are possible.

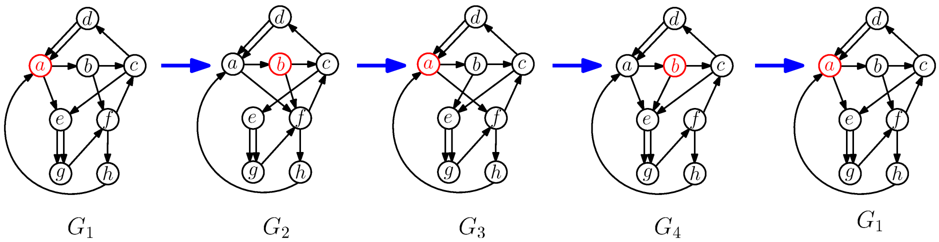


Figure 3.6: Improving response cycle for the swap version.

Moreover creating equilibrium networks having diameter 4 is already very challenging.

4 Geometric Network Creation Games

Network formation is studied from two different perspectives in the game theory filed. One branch of the research is about Network Creation Games, and the second is about Network Design Games (NDGs), e.g., [Ans+08a; Ans+08b]. In contrast to NCGs, in NDGs a given network with weighted edges serves as the host network and every agent has a pair of terminal nodes in the host network she wants to connect. For this, agents select a connecting path in the host network and pay a cost proportional to the length of the path for its usage. If edges are used by several agents, then the cost of the edge is split among these agents.

Thus in NCGs the distances between all pairs of nodes are important, whereas in NDGs the focus is on simply connecting the terminal pairs. Moreover, the former assume a complete unweighted host network, whereas the latter assume a weighted not necessarily complete host network. Hence, NCGs are suitable to model the formation of social networks or the AS-level network of the Internet, where using the hop-distance is more natural and where agents want to be central, i.e., close to all other agents. But, since NCGs crucially rely on an unweighted host network, these models cannot be used to investigate the creation of physical communication networks, where edges, e.g., fiber-optic cables, have lengths. NDGs are well-equipped to model the creation of physical communication networks between given terminal pairs, e.g., a network connecting many clients to a server or access point, where only connectivity matters. However, NDGs are not suited for studying settings where the agents are interested in communicating with all other agents and where agents are restricted to buying only incident edges.

To overcome these shortcomings of NCGs and NDGs, we propose and investigate a model which is a generalization of NCGs but which also shares some aspects with NDGs and therefore allows to model the creation of physical communication networks where the goal is to achieve an efficient communication between all pairs of nodes at low cost. That is, we are interested in the decentralized creation of *edge-weighted* networks which minimize the pairwise distances between agents and the total cost of all built edges. This can be seen as the game-theoretic analogue of the well-known Network Design Problem [GJ02; JLK78; MW84], where a weighted network and budgets for buying edges and the total routing cost between all pairs are given and the goal is to select a sub-network which respects both budgets. For this, we consider a variant of the NCG, where the given host network is an arbitrary weighted network and the prices for buying and using an edge are proportional to its weight. For example, with this we can model the realistic geometric setting where agents have a position in some metric space and the given weighted host network uses the distance between the positions of the involved agents as edge weights. To the best of our knowledge, this is the first variant of a NCG with weighted edges.

In stark contrast to the state-of-the-art for the unit-weight version, where the Price of Anarchy is conjectured to be constant and where resolving this is a major open problem, we prove a tight non-constant bound on the Price of Anarchy for the metric version and an asymptotically tight bound for the non-metric case. Moreover, we analyze the existence of equilibria, the computational hardness, and the game dynamics for several natural metrics. As we discussed earlier, the model we propose can be seen as the game-theoretic analogue of the classical Network Design Problem. Thus, low-cost equilibria of our game correspond to decentralized and stable approximations of the optimal network design.

4.1 Model and Notation

We consider a generalization of the Network Creation Game by Fabrikant et al. [Fab+03]. In our game, called the *Generalized Network Creation Game (GNCG)*, we consider a given host network $H = (V, E(H))$, which is a complete undirected weighted network on n nodes $v_1, \ldots, v_n$ with arbitrary non-negative edge weights $w : E(H) \to \mathbb{R}^+$.

Every node of H corresponds to a selfish agent who wants to participate in the network formation. As in the original NCG, agents strategically decide which subset of incident edges to buy, i.e., a strategy S_u of an agent u is any node subset of $V \setminus \{u\}$ towards which agent u wants to create edges. We assume that the edge price of any edge uv is proportional to its weight $w(u,v)$. In particular, we assume that the edge price for any edge uv is $\alpha \cdot w(u,v)$, where $\alpha > 0$ is a fixed parameter of the game which allows to model different trade-offs between the cost for buying and for using edges.

The strategy profile $\mathbf{s} = (S_{v_1}, \ldots, S_{v_n})$ uniquely determines a sub-network $G(\mathbf{s}) = (V, E(\mathbf{s}))$ of the host network $H = (V, E(H))$, where $E(\mathbf{s}) = \{uv \mid u \in V, v \in S_u\}$.

Let $d_G(u,v)$ be the distance between two nodes u and v in the network $G = (V, E)$, which is equal to the total weight of the shortest path between u and v, or $+\infty$ if such a path does not exist. To simplify the notation we will use $w(u, U) := \sum_{uv \in E : v \in U} w(u,v)$ as the sum of the weights of the edges between u and $U \subseteq V$ in G. Then $d_G(u, V)$ is the *distance cost* and $\alpha \cdot w(u, S_u)$ is the *edge cost* of the agent u.

Given any strategy profile $\mathbf{s}$ and its corresponding network $G(\mathbf{s})$, then the *cost* of agent u in $G(\mathbf{s})$ is defined as

$$cost(u, G(\mathbf{s})) = \alpha \cdot w(u, S_u) + d_{G(\mathbf{s})}(u, V).$$

The *social cost* of network $G(\mathbf{s})$, denoted $\mathrm{SC}(G(\mathbf{s}))$, is defined as the sum of the cost of all agents, i.e., $\mathrm{SC}(G(\mathbf{s})) = \sum_{u \in V} cost(u, G(\mathbf{s}))$.

For any host network H, we say that the *social optimum sub-network* $\mathrm{OPT_n} = \mathrm{OPT_n}(H)$ of H is the network $G(\mathbf{s}^*) = (V, E(\mathbf{s}^*))$ which minimizes $cost(G(\mathbf{s}^*))$ among all possible strategy profiles. Thus, *OPT* minimizes $\alpha \cdot \sum_{(u,v) \in E(\mathbf{s}^*)} w(u,v) + \sum_{u \in V} d_{G(\mathbf{s}^*)}(u, V)$.

As a solution concept we consider the *pure Nash Equilibrium (NE)* and the *Greedy Equilibrium (GE)*. Additionally, we say that $G(\mathbf{s})$ is in *β-approximate NE (β-NE)* if no agent u can change her strategy to decrease her cost to less than $1/\beta \cdot cost(u, G(\mathbf{s}))$. A *$\beta$-approximate GE ($\beta$-GE)* is defined analogously.

Model Variants

Besides the GNCG, where the game is played on a complete host network H with arbitrary non-negative edge weights, we also consider several interesting special cases (See Figure 4.1 for an overview). In the *metric GNCG* (M–GNCG) the edge weights of H satisfy the triangle inequality. Besides the general metric version, we consider three versions where the edge weights of H are defined by specific metrics. In the simplest case, the 1-2–GNCG, the edge weights of H are restricted to the set $\{1, 2\}$. We also consider the variant where the metric edge weights of H are derived from the shortest path distances in a given weighted tree, the T–GNCG. Finally, we consider the variant $\mathbf{R}^d$–GNCG, where the agents are points in $\mathbf{R}^d$ and the edge weights of H correspond to their p-norm distances. The original

Model	host graph is a clique with
GNCG	arbitrary non-neg. edge weights
M–GNCG	metric edge weights
$\mathbf{R}^d$–**GNCG**	edge weights defined by distances with p-norm in $\mathbf{R}^d$
T–GNCG	edge weights defined by distances in a given weighted tree
1-2–**GNCG**	edge weights in $\{1, 2\}$
1-∞–**GNCG**	edge weights in $\{1, \infty\}$
NCG	unweighted edges

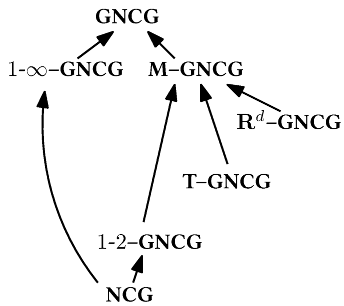


Figure 4.1: Left: Model overview. Right: Model relations. Arrows point from special case to more general model.

NCG [Fab+03] where H is an unweighted clique, is the most restricted special case of the M–GNCG. In the literature a non-metric special case of the GNCG, where the edge weights are restricted to the set $\{1, \infty\}$ was proposed [Dem+09]. We call this variant the 1-∞–GNCG.

4.2 Related Work

There is a huge body of literature both on variants of the Network Creation Game and on Network Design Games and it is impossible to give a full account. Instead, we focus on proposed models which share core features of our model and discuss how they are related to our approach. See also Table 4.1.

A weighted version of the NCG has been proposed by Albers et al. [Alb+14]. In contrast to our model, they consider a version with a specific amount of traffic between each pair of agents but distances are still measured by counting hops. Much closer to the GNCG model is the model investigated by Demaine et al. [Dem+09], where a NCG on a general unweighted host network is introduced. This corresponds to the special case of our model where only the edge weights 1 and ∞ are allowed. Edge weight ∞ encodes that a particular edge cannot be bought. The authors prove a general upper bound on the PoA of $O(\sqrt{\alpha})$ and for $\alpha \geq n$ they show that the PoA is in $\Omega\left(\min\{\sqrt{\alpha/n}, n^2/\alpha\}\right)$ and at most $\min\left\{O(\sqrt{n}), n^2/\alpha\right\}$ which yields a tight non-constant PoA bound if $\alpha \in \omega(n^{1.5})$ and $\alpha \in o(n^2)$ holds. The highest tight PoA bound as a function of α therefore is $\Theta(\sqrt[5]{\alpha})$ and is achieved for $\alpha = n^{\frac{5}{3}}$. Unfortunately, the proof techniques in [Dem+09] crucially rely on edge weights in $\{1, \infty\}$ and can therefore not be carried over to our model. However, their lower bound construction yields a lower bound of $\Omega(\sqrt[5]{\alpha})$ for the general non-metric case of our model. Also related is the work of Bilò et al. [Bil+12] who investigated the max-version of the NCG [Dem+12], where agents try to minimize their maximum distance, on a general unweighted host network.

One of the distinctive features of our model is the non-uniform edge price. A few other models with this feature have been proposed, e.g., [MMO14; MMO15], but they all use unit-weight edges. In the model by Cord-Landwehr et al. [CMH14] agents can choose different quality levels of an edge for different prices, i.e., the paid price influences the edge length. With this, the model is incomparable to our approach.

Also related are network formation games where not centrality but some other property is the goal of each agent. There are games where agents simply want to be connected to all other agents, e.g. [BG00; Goy+16; Kli11]. Among them, the work by Eidenbenz et al. [EKZ06] is closely related to our work. In their wireline strong connectivity game agents are points in the Euclidean plane who strategically buy incident edges to create a connected network. The edge price equals the length of the edge. This is similar to our model in the Euclidean plane with $\alpha = 1$ but the focus on connectivity changes the game completely. Another related geometric game was proposed by Moscibroda et al. [MSW06; MSW11]. Also there the agents are points in some metric space but agents pay a fixed price for each edge and try to minimize the total stretch towards all other agents. Gulyás et al. [Gul+15] considered a network formation game in the hyperbolic plane where agents strive for maximum navigability. This is also a geometric model but drastically different from our approach.

Network Design Games have been proposed in [Ans+08a; Ans+08b]. Their most important feature is that they are potential games [MS96], which already shows the contrast to Network Creation Games. Interestingly, Hoefer & Krysta [HK05] proposed and analyzed a geometric version.

There are many classical optimization problems related to network design, e.g., see the survey by Magnanti & Wong [MW84]. Many of them are NP-complete, e.g., all the problems labeled “ND” in [GJ02]. Our model is closely related to the Network Design Problem [Sco67] and the Optimum Communication Spanning Tree Problem (ND7 in [GJ02]). In particular, finding the social optimum network corresponds to a variant of the Network Design Problem, where, instead of having a budget for buying edges to minimize the routing cost, the total sum of edge costs and routing costs is to be minimized.

4.3 Our Contribution

In this paper we investigate the Network Creation Game on edge-weighted host networks. This variant allows modeling the decentralized creation of networks, like fiber-optic communication networks or many variants of overlay networks, by selfish agents, e.g., ISPs. In such settings, the nodes in a network have a physical location and the edge weights and also the cost for creating and maintaining them depend on these locations. In particular, we focus on specific natural metrics, e.g., network and tree metrics as well as the geometric setting where the agents correspond to points in $\mathbf{R}^d$.

We show that computing a best response strategy is NP-hard for all variants of our model and we prove for the 1-2–GNCG that deciding if a given strategy profile is in NE is NP-hard as well. The latter is the first result of this type in the realm of NCGs. Moreover, we prove that all our models do not have the finite improvement property. On the positive side, we give an efficient algorithm for computing a social optimum network for the 1-2–GNCG and we show how to trivially obtain the social optimum in the T–GNCG.

Our main focus is a rigorous study of the quality of the induced equilibrium networks of our models. For this we show that NE exist in the 1-2–GNCG and the T–GNCG and that the more general M–GNCG always admits a $(\alpha + 1)$-approximate NE. The main contribution of this chapter is a collection of bounds on the Price of Anarchy, i.e., we bound the loss in social welfare due to selfishness and to the lack of central coordination. We prove a tight PoA bound of $(\alpha+2)/2$ for the M–GNCG and the T–GNCG. This bound is remarkable, since it is non-constant and much higher than the previously known upper bounds for the NCG or the inherently non-metric 1-∞–GNCG. This shows that allowing weighted edges completely changes the picture. Moreover, in contrast, settling the PoA for the original NCG, which is a special case of all our models, is a major open problem in the field. For the model variant which is closest to the NCG, the 1-2–GNCG, we prove a tight constant bound on the PoA for $\alpha \leq 1$ and show that the PoA is in $\mathcal{O}(\sqrt{\alpha})$ for $\alpha > 1$. Hence, this model behaves very similar to the NCG. For the variant with points in $\mathbf{R}^d$, the $\mathbf{R}^d$–GNCG, we show a 1-dimensional

Table 4.1: Overview of our results and comparison with related work.

Model	**PoA**	**Existence of Equilibrium**
NCG*	$\mathcal{O}(\sqrt{\alpha})$ [Fab+03], $o(n^{\epsilon})$ [Dem+12]	NE exists [Fab+03]
1-∞–GNCG*	$\Theta(\sqrt[5]{\alpha})$ [Dem+09]	?
1-2–GNCG		
$\alpha < \frac{1}{2}$	$= 1$ (Thm. 4.17)	NE exists (Thm. 4.17)
$\frac{1}{2} \leq \alpha < 1$	$= \frac{3}{\alpha+2}$ (Thm. 4.16+4.15)	NE exists (Thm. 4.12)
$\alpha = 1$	$= \frac{3}{2}$ (Thm. 4.16+4.4)	
$1 < \alpha < 3$	$\mathcal{O}(\sqrt{\alpha})$ (Thm. 4.20)	AE is $(\alpha+1)$-GE (Thm. 4.5), GE is 3-NE (Thm. 4.7)
$\alpha \geq 3$		NE exists (Thm. 4.18)
T–GNCG	$= \frac{\alpha+2}{2}$ (Thm. 4.25+4.4)	NE exists (Cor. 4.22)
$\mathbb{R}^d$–GNCG		
$0 < \alpha < 2$	$\Theta(1)$ (Thm. 4.30+4.4),	AE is $(\alpha+1)$-GE (Thm. 4.5)
$\alpha \geq 2$	$\Theta(\alpha)$ (Thm. 4.29+4.4)	GE is 3-NE (Thm. 4.7)
1-norm	$\geq 1 + \frac{\alpha}{2+\alpha/(2d-1)}$ (Thm. 4.31), $\leq \frac{\alpha+2}{2}$ (Thm. 4.4)	
M–GNCG	$= \frac{\alpha+2}{2}$ (Thm. 4.25+4.4)	AE is $(\alpha+1)$-GE (Thm. 4.5), GE is 3-NE (Thm. 4.7)
GNCG	$\geq \frac{\alpha+2}{2}$ (Thm. 4.25), $\leq \alpha+1$ (Thm. 4.33)	$(\alpha+1)$-NE exists (Prop.4.32)

Table 4.2: Overview of our results and comparison with related work.

Model	Complexity	FIP
NCG*	BR NP-hard [Fab+03]	no [KL13]
1-∞–GNCG*	BR NP-hard (Cor. 4.1)	no (Cor. 4.1)
1-2–GNCG	BR NP-hard (Cor. 4.1), Dec. NE NP-hard (Thm. 4.10)	no (Cor. 4.1)
T–GNCG	BR NP-hard (Thm. 4.23)	no (Thm. 4.24)
$\mathbb{R}^d$–GNCG p-norm, $p \geq 2$ 1-norm	BR NP-hard (Thm. 4.26)	 ? no (Thm. 4.27)
M–GNCG	BR NP-hard (Cor. 4.1) Dec. NE NP-hard (Thm. 4.10)	no (Cor. 4.1)
GNCG	BR NP-hard (Cor. 4.1) Dec. NE NP-hard (Thm. 4.10)	no (Cor. 4.1)

Notation used in the tables: **BR** – a problem of computing a best response; **Dec. NE** – a problem of deciding if a given strategy profile is in NE; **FIP** – finite improving property;

construction which delivers a lower bound for the PoA that asymptotically meets the upper bound of $O(\alpha)$. Moreover, we show how to embed the lower bound construction from the T–GNCG. This yields a tight PoA bound if d tends to infinity. Finally, for the most general case, the GNCG, we show that the PoA is between $(\alpha + 2)/2$ and $\alpha + 1$, i.e., the bound is asymptotically tight.

See Table 4.1 and Table 4.2 for an overview over the majority of our results and the most relevant results for the earlier models which are marked with the star symbol. All results on the Price of Anarchy with an equality sign are tight bounds.

4.4 Preliminaries

We start by clarifying the relation of the models we investigate. Figure 4.1 shows which models are special cases of other models. These relationships and the facts that computing a best response strategy is NP-hard for the NCG [Fab+03] and that the NCG does not have the FIP [KL13] directly yields the following corollary.

▶ **Corollary 4.1.** Computing a best response strategy is NP-hard for the 1-2–GNCG, 1-∞–GNCG, the M–GNCG and the GNCG. Additionally, these models do not have the FIP. ◀

Let $t \geq 1$. We say that a sub-network G of H is a *t-spanner* if $d_G(u,v) \leq td_H(u,v)$ for every pair of vertices $u,v \in V$. Next, we show a useful property, which holds for any host network.

▶ **Lemma 4.2.** For any host network H any add-only equilibrium is a $(\alpha + 1)$-spanner. ◀

Proof. First, we consider edges uv with $w(u,v) = d_H(u,v)$, that is, a shortest path between u and v in the host network H uses the direct edge. We claim for such pairs u and v that in any NE network G we have $d_G(u,v) \leq (\alpha + 1)d_H(u,v) = (\alpha + 1)w(u,v)$. To see this, assume towards a contradiction that $d_G(u,v) > (\alpha+1)d_H(u,v)$, which implies that $(u,v) \notin E(G)$. Now consider what happens if agent u buys the edge uv: Agent u additionally has to pay $\alpha \cdot w(u,v)$ for creating the edge and then her distance to v is guaranteed to be $w(u,v)$. Thus her total cost for buying the edge uv and reaching node v is $(\alpha+1)w(u,v)$. Since $d_G(u,v) > (\alpha+1)w(u,v)$, buying the edge uv is an improving move for agent u.

Now we consider two arbitrary agents u and v in G and let $P_{uv} = x_1, x_2, \ldots, x_k$ with $u = x_1$ and $x_k = v$ be a shortest path between u and v in the host network H. It follows that $d_H(u,v) = w(x_1,x_2)+w(x_2,x_3)+\cdots+w(x_{k-1},x_k)$. Since P_{uv} is a shortest path in H and since any subpath of a shortest path must be a shortest path itself, it follows that for all pairs x_i and x_{i+1}, with $1 \leq i \leq k-1$, the equality $w(x_i,x_{i+1}) = d_H(x_i,x_{i+1})$ holds. Thus, in any NE G on the host network H we have that $d_G(x_i, x_{i+1}) \leq (\alpha + 1)w(x_i, x_{i+1})$ holds for all $1 \leq i \leq k-1$. Thus, the distance between u and v in any NE G is

$$\begin{aligned} d_G(u,v) &\leq (\alpha+1)w(x_1,x_2) + (\alpha+1)w(x_2,x_3) + \cdots + (\alpha+1)w(x_{k-1}x_k) \\ &= (\alpha+1)d_H(u,v). \end{aligned}$$

Since we considered only additions of edges, the statement holds for a more general add-only equilibrium concept. ■

With a similar technique we get an analogous statement for the social optimum network OPT.

▶ **Lemma 4.3.** The social optimum network $\mathrm{OPT_n}(H)$ is a $\left(\frac{\alpha}{2}+1\right)$-spanner for any connected host network H. ◀

Proof. The proof is analogous to the proof of Lemma 4.2. Let $OPT(H)$ be the sub-network of H which minimizes the social cost. We start by considering edges uv in $OPT(H)$ where $w(u,v) = d_H(u,v)$, that is, a shortest path between u and v in the host network H uses the direct edge. We claim for such pairs u and v that in $OPT(H)$ we have

$$d_{OPT(H)}(u,v) \leq \left(\frac{\alpha}{2}+1\right)d_H(u,v) = \left(\frac{\alpha}{2}+1\right)w(u,v).$$

To see this, assume towards a contradiction that $d_{OPT(H)}(u,v) > \left(\frac{\alpha}{2}+1\right) \cdot w(u,v)$, which implies that $(u,v) \notin E(OPT(H))$. Now consider what happens if the edge (u,v) is added to $OPT(H)$: The social cost increases by $\alpha \cdot w(u,v)$ for creating the additional edge. Moreover, the creation of the edge uv ensures that the distance between u and v is $w(u,v)$. Thus, the distance from u to v is decreased by more than $\left(\frac{\alpha}{2}+1\right)w(u,v) - w(u,v) = \left(\frac{\alpha}{2}\right)w(u,v)$. The same holds true for the distance from v to u. Thus, the total distance decrease induced by the addition of the edge uv to $OPT(H)$ is more than $2\left(\frac{\alpha}{2}\right)w(u,v) = \alpha \cdot w(u,v)$. Since the total distance decrease is strictly larger than the edge cost of the edge uv, this implies that the network $OPT(H)$ augmented by the edge uv has strictly less social cost than $OPT(H)$. This contradicts the assumption that $OPT(H)$ minimizes the social cost.

Now we consider two arbitrary agents u and v in $OPT(H)$. Let $P_{uv} = x_1, \ldots, x_k$ with $u = x_1$ and $x_k = v$ be a shortest path between u and v in the host network H. It follows that $d_H(u,v) = w(x_1,x_2) + w(x_2,x_3) + \cdots + w(x_{k-1},x_k)$. Since

P_{uv} is a shortest path in H and since any subpath of a shortest path must be a shortest path itself, it follows that for all pairs x_i and x_{i+1}, with $1 \le i \le k-1$, the equality $w(x_i, x_{i+1}) = d_H(x_i, x_{i+1})$ holds. Thus, in $OPT(H)$ we have that $d_{OPT(H)}(x_i, x_{i+1}) \le (\frac{\alpha}{2}+1)w(x_i, x_{i+1})$ holds for all $1 \le i \le k-1$. Thus, the distance between u and v in $OPT(H)$ is

$$d_{OPT(H)}(u,v) \le \left(\frac{\alpha}{2}+1\right)w(x_1,x_2) + \left(\frac{\alpha}{2}+1\right)w(x_2,x_3) + \dots + \left(\frac{\alpha}{2}+1\right)w(x_{k-1}x_k) = \left(\frac{\alpha}{2}+1\right)d_H(u,v).$$ ■

4.5 Host Networks with Metric Weights

In this section we investigate the NCG on complete host networks with edge weights which satisfy the triangle inequality. After giving some general results, we focus on specific natural metrics.

4.5.1 General Results for the M–GNCG

▶ **Theorem 4.4.** The PoA in the M–GNCG is at most $\frac{\alpha+2}{2}$ for any host network for which a NE exists. ◀

Proof. Let G be a NE and let u and v be two distinct nodes. Let x and x^* be two Boolean variables such that $x = 1$ if and only if uv is an edge of G and $x^* = 1$ if and only if uv is an edge of the social optimum OPT. We prove the claim by showing that

$$\sigma := \frac{\alpha \cdot w(u,v) \cdot x + 2d_G(u,v)}{\alpha \cdot w(u,v) \cdot x^* + 2d_{OPT}(u,v)} \le \frac{\alpha+2}{2}.$$

Essentially σ is the ratio of the social cost contribution of every pair of nodes in the NE and in OPT. If the ratio for every pair of nodes is bounded by $(\alpha+2)/2$ then this also holds for their sum.

Now we prove the claim. If $x = 1$ then $d_G(u,v) = w(u,v)$ and hence

$$\sigma \le (\alpha+2) \cdot \frac{w(u,v)}{2d_{OPT}(u,v)} \le (\alpha+2) \cdot \frac{w(u,v)}{2w(u,v)} = \frac{\alpha+2}{2}.$$

If $x = 0$ and $x^* = 1$ then $\sigma \le 2(\alpha+1)/(\alpha+2) \le (\alpha+2)/2$ since, by Lemma 4.2, $d_G(u,v) \le (\alpha+1)w(u,v)$.

It remains to prove $\sigma \leq (\alpha+2)/2$ when $x = 0$ and $x^* = 0$. This means that there is a vertex z with $z \neq u$ and $z \neq v$ along a fixed shortest path in OPT between u and v. Clearly, both edges As G is a NE, neither u nor v has an incentive to buy the edge towards z. If u bought the edge uz at the price of $\alpha \cdot w(u,z)$, her distances towards z would be at most $w(u,z)$ and, by the triangle inequality, her distance towards v would be at most $w(u,z) + d_G(z,v)$. Since this is not an improvement, we have

$$\begin{aligned} d_G(u,z) + d_G(u,v) &\leq \alpha \cdot w(u,z) + d_{G+(u,z)}(u,z) + d_{G+(u,z)}(u,v) \\ &\leq (\alpha+2)d_{OPT}(u,z) + d_G(z,v) = (\alpha+2)w(u,z) + d_G(z,v) \end{aligned}$$

and hence

$$d_G(u,z) + d_G(u,v) \leq (\alpha+2)d_{OPT}(u,z) + d_G(z,v). \tag{4.1}$$

Analogously for agent v, we get

$$\begin{aligned} d_G(v,z) + d_G(v,u) &\leq \alpha \cdot w(v,z) + d_{G+(v,z)}(v,z) + d_{G+(v,z)}(v,u) \\ &\leq (\alpha+2)d_{OPT}(v,z) + d_G(z,u) \leq (\alpha+2)w(v,z) + d_G(z,u), \end{aligned}$$

which yields

$$d_G(v,z) + d_G(v,u) \leq (\alpha+2)d_{OPT}(v,z) + d_G(z,u). \tag{4.2}$$

By summing up the inequalities (4.1) and (4.2), we obtain

$$2d_G(u,v) \leq (\alpha+2)d_{OPT}(u,z) + (\alpha+2)d_{OPT}(v,z) = (\alpha+2)d_{OPT}(u,v).$$

Note that the inequalities (4.1) and (4.2) hold as well if any of two edges, uz or vz, exist in G. Clearly, in case both edges are in G, $d_G(u,v) \leq d_{OPT}(u,v)$. Therefore, also the last case yields $\sigma \leq (\alpha+2)/2$. ■

Existence of Nash Equilibria

It is an interesting open question if NE always exist for the M–GNCG. Here we prove a weaker result which essentially states that there always is an outcome of the add-only game where no agent can greedily improve (i.e., by a single addition, deletion, or swap) by a high multiplicative factor.

▶ **Theorem 4.5.** Any add-only equilibrium network in the M–GNCG is in $(\alpha+1)$-GE. ◀

Proof. Consider a network $G = (V, E)$ which is in AE. Clearly, G always exists since any complete host network H is in AE. By the definition of a $(\alpha+1)$-GE we need to evaluate the maximal improvement of the cost function which can be made by a deletion or swap of any edge in G.

First, we consider a deletion. Compare the cost function value of some agent $u \in V$ before and after an improving deletion of one of her edge $e = uv \in E(G)$:

$$\frac{cost(u,G)}{cost(u,G')} = \frac{\alpha \cdot w(u,S_u) + d_G(u,V)}{\alpha \cdot w(u,S'_u) + d_{G'}(u,V)},$$

where G' is the network obtained from G by applying new strategy $S'_u = S_u \setminus \{v\}$, i.e. $E(G') = E(G) \setminus \{uv\}$. In the worst case, the deletion of the edge uv does not change the distance between the nodes u and v, i.e., $w(u,v) = d_{G'}(u,v)$. This yields

$$\begin{aligned}\frac{cost(u,G)}{cost(u,G')} &= \frac{\alpha \cdot w(u,S_u) + d_G(u,V)}{\alpha \cdot w(u,S'_u) + d_G(u,V)}\\ &= \frac{\alpha \cdot w(u,v) + \alpha \cdot w(u,S_u \setminus \{v\}) + d_G(u,V)}{\alpha \cdot w(u,S_u \setminus \{v\}) + d_G(u,V)}\\ &= 1 + \frac{\alpha \cdot w(u,v)}{\alpha \cdot w(u,S_u \setminus \{v\}) + d_G(u,V)}\\ &\leq 1 + \frac{\alpha \cdot w(u,v)}{d_G(u,V)} \leq 1 + \frac{\alpha \cdot w(u,v)}{w(u,v)} = 1 + \alpha. \end{aligned} \tag{4.3}$$

Now we consider an improvement which can be made by one swap. Let agent $u \in V(G)$ be an agent that can improve her cost by swap an edge uv to uw, and let G_{swap} be the new network. Compare the cost function after the swap with the cost value after the sequential addition of the edge uw and the deletion of the edge uv. Let G_{add} and G_{del} be the corresponding networks. Thus,

$$E(G_{swap}) = E(G_{del}) = E(G_{add}) \setminus \{uv\} = (E(G) \cup \{uw\}) \setminus \{uv\}.$$

Then, by the inequality (4.3) and because G is in AE, we have:

$$cost(u, G_{swap}) = cost(u, G_{del}) \geq \frac{1}{\alpha+1} cost(u, G_{add}) \geq \frac{1}{\alpha+1} cost(u, G). \quad (4.4)$$

Finally, by (4.3) and (4.4), we get that G is in $(\alpha + 1)$-GE. ■

In the above proof, we observed that a complete host network H is in $(\alpha+1)$-GE. It is easy to show an even stronger result that the host network is in $(\alpha + 1)$-NE.

▶ **Proposition 4.6.** Every host network H is in $(\alpha+1)$-NE in the M−GNCG. ◀

Proof. Since H is a clique, the only improving move an agent can perform is an edge deletion. Consider a subnetwork H' obtained from H after an agent u improved her strategy by deleting some edges. Since the deletion of edges increases the distance cost, we have that $cost(u, H') \geq d_{H'}(u, V) \geq d_H(u, V)$. Hence, we get

$$\frac{cost(u, H)}{cost(u, H')} \leq \frac{\alpha \cdot w(u, S_u) + d_H(u, V)}{d_H(u, V)} \leq \frac{(\alpha+1) d_H(u, V)}{d_H(u, V)} = \alpha + 1. \quad ■$$

Now, we adapt the technique from [Len12] to relate GE and β-NE.

▶ **Theorem 4.7.** In the M−GNCG every network in GE is in 3-NE. ◀

Proof. We prove the claim by a "locality gap preserving" reduction to the *Uncapacitated Metric Facility Location* problem (UMFL). Roughly speaking, in UMFL we are given a set of facilities, each of which has a non-negative opening cost, a set of clients, and a distance between each client and each facility (the distances satisfy the triangle inequality). The task in UMFL it to open a set of facilities and assign each client to the closest opened facility in such a way that the overall cost – i.e., the overall cost of the opened facilities plus the overall sum of client-to-assigned-facility distances – is minimized. Since it was shown in [Ary+04] that the locality gap of UMFL is 3, that means that any UMFL solution that cannot be improved by a single move, i.e., by opening, closing or swapping one facility, is a 3-approximation of the optimal solution.

Consider a network $G = (V, E)$. Let $u \in V$ be an agent in (G, α) and let $Z \subset V$ be the set of nodes which own an edge to u. Consider the sub-network $G' = (V, E')$ of G which does not contain edges owned by the agent u. Denote $S(u)$ be the set of u's pure strategies in (G', α). We construct an instance $I(G')$

for UMFL from the network G' as follows: let $F = C = V \setminus \{u\}$, where F is the set of facilities, C is the set of clients; we define for all facilities $f \in Z \cap F$ the opening cost $c(f)$ to be 0, and $c(f) = \alpha \cdot w(f,u)$ for all other facilities. We define distances for all $i \in F, j \in C$ to be $d_{ij} = d_{G'}(i,j) + w(i,u)$. If there is no path connecting i and j in G', then $d_{ij} = +\infty$. Since the distances and the edge weights in G' are metric, the distances in $I(G')$ satisfy the triangle inequality.

Now we construct a map $\pi : S(u) \longrightarrow S_{\text{UMFL}}$, where S_{UMFL} is the set of solutions of the UMFL for the instance $I(G')$, as follows: for any $S \in S(u)$, define $\pi(S) = S \cup Z$ and for any $F_S \in S_{\text{UMFL}}$, $\pi^{-1}(F_S) = F_S \setminus Z$. Since the opening cost for any $f \in Z \cap F$ is 0, we can assume that S_{UMFL} contains only solutions F_S such that $Z \subseteq F_S$. Then the strategy $S' = \pi^{-1}(F_S)$ exists, and for any two strategies $S_1 \neq S_2$, $\pi(S_1) \neq \pi(S_2)$. Therefore, the map π is a bijection. To prove the statement of the theorem we need to show that if agent u cannot improve her strategy by adding, deleting or swapping one edge, then the corresponding solution $F_S = \pi(S)$ for UMFL cannot be improved by opening, closing or swapping one facility.

First, we show that the cost of agent u is equal to the cost of the corresponding UMFL solution F_S. Indeed,

$$\begin{aligned}
cost(u, G(S)) &= \alpha \cdot w(u,S) + \sum_{v \in V \setminus \{u\}} \left(\min_{x \in S \cup Z} (d_{G'}(x,v) + w(u,x)) \right) \\
&= \alpha \cdot w(u, S \setminus Z) + 0 \cdot w(u,Z) + \sum_{v \in V \setminus \{u\}} \left(\min_{x \in S \cup Z} d_{xv} \right) \\
&= \sum_{f \in F_S \setminus Z} c(f) + \sum_{f \in Z} c(f) + \sum_{v \in C} \left(\min_{x \in F_S} d_{xv} \right) \\
&= cost(F_S).
\end{aligned}$$

Next we show that $F_s = \pi(S)$ cannot be decreased by opening, closing or swapping one facility. For the sake of contradiction, assume that the solution F_S can be improved by a single step. Denote F'_S be an improved solution. Note that no facility $z \in Z$ is included in an opening, closing or swapping step. Indeed, by construction, $Z \subseteq F_S$ and the opening cost of each facility in Z is zero, hence we only need to consider closing of one of the facilities from Z in F'_S. If there is a facility $z \in F_S \setminus F'_S$, then there is at least one client $c \in C$ such that $d_{cz} \leq \min_{f \in F_S} d_{cf}$, thus, closing the facility z does not decrease $cost(F_S)$ and, therefore, $z \in F'_S$. Thus, and because π is a bijection, we have that there is a strategy $S' = \pi^{-1}(F'_S)$ such that $S' \neq S$. Therefore, we have $cost(u, G(S')) = cost(F'_S) < cost(F_S) = cost(u, G(S))$. Hence, there is the better strategy S' for the agent u, which contradicts with the assumption that there is no one step improvement of the strategy S.

Finally, applying the result by Arya et al.[Ary+04], we get

$$cost(u, G(S)) \leq 3\, cost(u, G(S^*))$$

where S^* is an optimal strategy in (G', α). ■

4.5.2 1-2-networks

Here we consider the M–GNCG for the special case where for every pair of nodes u and v we have either $w(u, v) = 1$ or $w(u, v) = 2$. We call an edge of weight 1 or 2 a *1-edge* or *2-edge*, respectively. We call such networks *1-2-networks*.

Studying 1-2-networks is especially interesting since this class of host networks is the simplest generalization of the unweighted host networks from the NCG and the edge weights are guaranteed to satisfy the triangle inequality. 1-2-networks are commonly used as the simplest non-trivial metric special case, e.g., when studying the TSP [AMP18; BK06; Kar72], and hence they are a natural starting point.

We start with a simple statement about 1-edges. We show that for $\alpha < 1$ any NE must contain all the 1-edges from the host network. If $\alpha = 1$, then there always exists a NE which contains all 1-edges.

▶ **Lemma 4.8.** For $\alpha = 1$ in any NE network in the 1-2–GNCG buying any additional 1-edge is cost neutral for the buyer. For $\alpha < 1$ any NE network contains all 1-edges from the host network. ◀

Proof. Consider a network G which is in NE in the 1-2–GNCG. Assume there is an edge uv of weight 1 which is not in G. Thus, $d_G(u, v) \geq 2$. Then buying the edge by one of its endpoint costs α while the distance cost decreases by at least $2 - 1$. Hence, if $\alpha < 1$, the decrease of the distance cost exceeds the increase in the edge cost, which means that this is an improving move for the buying agent. If $\alpha = 1$, the cost for the buying agent does not change. ■

Hardness

Here we discuss the hardness of deciding if a given strategy profile is in NE for the 1-2−GNCG. Note that the NP-hardness of computing a best response strategy for some agent, which is guaranteed by Corollary 4.1, does not directly imply the NP-hardness of the NE decision problem.

First, we take a detour via the Vertex Cover problem. A *vertex cover* of an undirected network G is a subset C of nodes of G such that, for every edge uv of G, $u \in C$ or $v \in C$. It is well-known that computing a minimum vertex cover of a subcubic network is NP-hard. For the sake of completeness, we start with the result that claims that there is no polynomial time algorithm that decides if there is a vertex cover of less size than a given vertex cover.

▶ **Lemma 4.9.** Unless P=NP, there is no polynomial time algorithm that, given a network G and a vertex cover of G of size k, decides whether G admits a vertex cover of size at most $k-1$. ◀

Proof. We prove the claim by showing that the existence of such an algorithm would imply the existence of a polynomial time algorithm for computing a vertex cover of G of size at most $k-1$, assuming it exists. Therefore, by reiterating the algorithm at most k times, we might be able to compute a minimum vertex cover of G in polynomial time, thus proving that P=NP. Let C be a vertex cover of G of size k.

The algorithm works as follows. First of all, we query the algorithm to understand whether G admits a vertex cover of size strictly better than k. In case of a "no" answer, we know that C is an optimal vertex cover and therefore G does not admit a vertex cover of size $k-1$. So, we assume that the algorithm answers "yes". This implies that there is a vertex cover of size $k-1$. In the following we show how to compute a vertex cover of size (at most) $k-1$ in polynomial time.

Let $G-v$ be the network obtained from G without the vertex v (and all the edges incident to v). For every vertex v of C, we query the algorithm using the network $G-v$ and the cover $C-v$ (so we want to know whether $G-v$ has a vertex cover of size $k-2$). If all the k answers returned by the algorithm are "no", then $V(G) \setminus C$ is a vertex cover of size strictly smaller than k. Indeed, the answer "no" for v means that there is no vertex cover of size $k-1$ that contains v. However, since a vertex cover of size $k-1$ exists, such a vertex cover has to contain the entire neighborhood of v (otherwise some edges incident in v would remain uncovered).

To complete the proof, we assume that the algorithm has answered "yes" for at least one vertex of C, say v. This means that there is a vertex cover of size (at most) $k-1$ that contains v. We build such a vertex cover by adding v to the vertex cover of size (at most) $k-2$ that is computed recursively on $G-v$ and $C-v$. Clearly, the running time of this algorithm is polynomial in the number of vertices of the graph. ■

▶ **Theorem 4.10.** Unless P=NP, there is no polynomial time algorithm that decides whether a strategy profile is in NE for the 1-2–GNCG. ◀

Proof. The reduction is from the Vertex Cover problem and $\alpha = 1$. More precisely, we define both a 1-2-network and a strategy profile such that every agent but one is playing her best response and computing a best response of the remaining agent is equivalent to computing a minimum vertex cover.

We define the network $G = (V, E)$ such that there is one vertex node $a_i \in V$ for each vertex v_i of the Vertex Cover instance, and two edge nodes p_j and p'_j in V for each edge e_j of the Vertex Cover instance. Finally, there is a new node u, that is neither a vertex node nor an edge node. There is an edge of weight 1 between vertex node a_i and each edge node p_j, p'_j if and only if v_i is an endvertex of e_j. Furthermore, there is an edge of weight 1 between every pair of vertex nodes. All the other edges have weight 2. See Figure 4.2 for the construction.

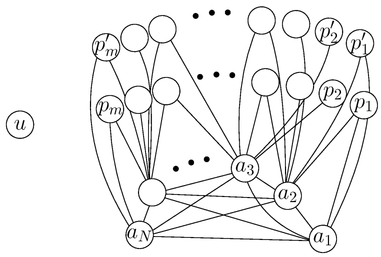


Figure 4.2: Illustration of the construction used in the reduction. All depicted edges have weight equal to 1; the missing edges are all of weight equal to 2.

Consider the strategy profile in which each edge of weight 1 is bought by any of the two agents that are incident to the edge, while u is buying all the edges towards vertex nodes that correspond to a vertex cover of size k w.r.t. the Vertex Cover instance computed using any polynomial time algorithm. Note that by Lemma 4.8 and since $\alpha = 1$, buying a 1-edge is neutral for the incident agents.

First of all, we observe that the eccentricity of each node, i.e., the maximum distance between a node to all other nodes, is at most 3. Therefore, every agent other than u is actually playing a best response. We claim that for any improving move of u, there exists another improving move in which agent u buys only the edges towards the vertex nodes that correspond to a vertex cover of size at most $k-1$ w.r.t. the Vertex Cover instance. The claim then would follow from Lemma 4.9.

Consider any improving move S_u for u. We prove the claim by first showing the existence of an alternative improving move consisting only of edges towards vertex nodes. Indeed, if u bought an edge towards an edge node in S_u, w.l.o.g. say p_j, then u would not buy the edge towards any vertex node a_i such that v_i is an endpoint of e_j in the Vertex Cover instance. This is simply because the edge ua_i would improve distances to more nodes than the edge up_j. Moreover, either p'_j would be at distance 4 from u or u would have also bought the edge towards p'_j. In either case, u would want to delete the edge towards p_j – as well as the edge towards p'_j, if she has bought it – and in buying the edge towards a vertex node a_i, with v_i being an endpoint of e_j, thus, decreasing her overall cost by at least 1.

Now we show that for any improving move S_u in which u buys only edges towards vertex nodes, there is another improving move in which u buys only edges towards vertex nodes that correspond to a vertex cover of the Vertex Cover instance. Indeed, if this is not the case, then there exist two nodes, say p_j and p'_j, which are at distance 4 from u. Let a_i be a vertex node such that v_i is an endpoint of e_j. Clearly, the distance from u to a_i is 3. Therefore, by buying the edge towards a_i the cost of u would decrease by at least 1.

As a consequence, we can restrict the strategy space for agent u only to improving moves that correspond to vertex covers of the Vertex Cover instance. Let k' be the number of edges bought by u in any strategy of the restricted strategy space for u, and let N and m be the number of nodes and edges of the Vertex Cover instance, respectively, The cost of u is equal to $2k' + 2k' + 3(N - k') + 6m = 3N + 6m + k'$. Since N and m are fixed, we observe that the cost of u is minimized when k' is minimized. Hence, any improving move for u would define a vertex cover of size of at most $k-1$. ■

1-2-networks for $\alpha \leq 1$

Here we study the 1-2–GNCG with $\alpha \leq 1$. We prove that in this case a NE network always exists. In contrast to the corresponding result for the original NCG [Fab+03] we do not prove this via a generic construction. Moreover, we provide a simple algorithm which computes a social optimum network in polynomial time and we provide tight bounds on the PoA.

Existence of Nash Equilibria

In the following we prove an interesting connection between existence of a NE for the 1-2–GNCG with $\alpha \leq 1$ and k-spanners. The weight of a k-spanner is the total sum of its edge weights. The following results are inspired by Lemma 4.3.

▶ **Lemma 4.11.** Let $\frac{1}{2} \leq \alpha \leq 1$, and let G be a $\frac{3}{2}$-spanner of minimum weight. Then G contains all the edges of the host network H of weight 1 and has a diameter of at most 3. ◀

Proof. Let uv be any edge of H. Since $d_G(u,v) \leq 3/2 \cdot d_H(u,v) \leq 3/2 \cdot w(u,v)$ and all edge weights are in the set $\{1,2\}$, we have that $d_G(u,v) \leq 3$; furthermore, if $w(u,v) = 1$, then $d_G(u,v) = 1$, i.e., uv is contained in G. Therefore, G contains all the edges of H of weight 1 and has a diameter of at most 3. ■

▶ **Theorem 4.12.** Let $\frac{1}{2} \leq \alpha \leq 1$ and let G be a $\frac{3}{2}$-spanner of minimal weight. There is an edge ownership assignment in G such that G is in NE. ◀

Proof. The claim is proved by contradiction. Consider any edge ownership assignment in G which induces strategy profile $\mathbf{s}$ and assume there is an agent $u \in V$ who can improve on her strategy S_u in $\mathbf{s}$. We will show that if there is a better strategy S'_u for agent u, then $|S'_u| \leq |S_u| - 1$ and that S'_u contains strictly less 2-edges than S_u. Then we prove that for any edge uv, which would be removed by agent u in the strategy change from S_u to S'_u, we can exchange the ownership of its endpoint such that the new owner v cannot improve on her strategy, or we can apply a combination of the two strategies S'_u and S'_v to G which yields a new network which is a 3/2-spanner with less total weight, which contradicts that G is a 3/2-spanner of minimum weight. Therefore, the edge ownership can be chosen such that network G is in NE.

First, we prove that $|S'_u| \leq |S_u| - 1$ and that S'_u contains less 2-edges than S_u. Towards this we claim that the change from S_u to S'_u can only consist of a change of the 2-edges which are bought by u and, if $\alpha = 1$, possibly the removal of some 1-edges. This is true since by Lemma 4.11 we have that all 1-edges are contained in G and removing any 1-edge is not an improving move. In particular, removing a 1-edge is a cost neutral move if $\alpha = 1$. Using the latter, we can define a new strategy S''_u which is identical to S'_u but still has all the 1-edges which are contained in S_u. Thus, $S''_u \setminus S'_u$ only consists of 1-edges which are cost neutral for agent u under strategy S'_u. Hence, $cost(u, S''_u) = cost(u, S'_u)$ and we have $|S''_u| \geq |S'_u|$. Let $S_{u+} = \{v \in V : v \in S''_u \setminus S_u\}$ be the set of nodes to which new edges have been added, $S_{u-} = \{v \in V : v \in S_u \setminus S''_u\}$ be the set of nodes to which the edges have been deleted and let G'' be the network obtained from G by exchanging agent u's strategy S_u with S''_u. Since the diameter of G is 3, then, after changing the strategy from S_u to S''_u, only distances between u and nodes at hop-distance 2 from u might increase. Indeed, if there is a node at hop-distance 3, then a shortest path to this node contains only 1-edges. Since only 2-edges will be deleted after the strategy change, the path of 1-edges does not change. Thus, if there is a node $v \in S_u \setminus S''_u$ such that $d_{G''}(u,v) \geq 4$ or $d_{G''}(u,x) \geq 4$, where x is at distance 1 from v, then the deletion of v from strategy S_u is not an improvement for u. This means that for any node $v \in V$ we have $d_{G''}(u,v) \leq 3$. Therefore, the new strategy S''_u decreases agent u's edge cost by $2\alpha \cdot (|S_u| - |S''_u|)$, increases her distance to all nodes in S_{u-} by 1 and decreases her distance by 1 for $|S_{u+}|$ many nodes. Since we assume that $cost(u, S_u) > cost(u, S''_u)$, then

$$0 > -2\alpha \cdot (|S_u| - |S''_u|) + |S_{u+}| - |S_{u-}| = (2\alpha + 1)(|S_{u+}| - |S_{u-}|),$$

thus, $|S_{u+}| \leq |S_{u-}| - 1$. Hence, $|S''_u| \leq |S_u| - 1$, i.e., S''_u contains strictly less 2-edges than S_u. Since $|S''_u| \geq |S'_u|$, we have $|S'_u| \leq |S_u| - 1$.

Let G' be the network obtained from G by exchanging agent u's strategy S_u with strategy S'_u. Since the number of edges in G' is strictly less than the number of edges in G and since G' has strictly less 2-edges than G, it follows that if the diameter of G' is 3, then G' is a 3/2-spanner of total weight less than the total weight of G and we get a contradiction. But it might happen that there are at least two nodes $x, y \in V$ at distance 4 in G'. Note that if the distance between x and y increased because of removing the edge uv ,i.e., $d(x,y) = d(x,u)+w(u,v)+d(v,y)$, then $w(u,v) = 2$. Indeed, if uv was a 1-edge, then the distance between u and y as well as the distance between u and v would increase by 1. Therefore, the 1-edge uv would not be a neutral edge and its removing is not an improving move, i.e.,

$v \in S'_u$. Hence, any edge whose deletion influences the distance between not only its endpoints must be a 2-edge.

Note that for any $v \in V$ we have $d_{G''}(u,v) \leq 3$. Also note that G'' and G' only differ in 1-edges bought by agent u, and the removal of any such 1-edge by u increases the distance only to the edge's endpoint. Hence, for any $v \in V$ we have $d_{G'}(u,v) \leq 3$. Since the diameter of G is 3 and since for any $v \in V$ we have $d_{G'}(u,v) \leq 3$, then x must be a neighbor of u which is connected by a 1-edge and $y \in S_u \setminus S'_u$. For each such edge uy we can invert the ownership.

Now we prove that after the inversion of the edge ownership for each edge uy, for all $y \in S_u \setminus S'_u$, any improving strategy for y contains u. Assume towards a contradiction that U is the non-empty set of nodes y, which have an improving strategy S'_y which does not contain u. We apply all improving strategies S'_y, for all $y \in U$, and S'_u to G and obtain a new network G^*. Note that if there are two nodes x, y such that there is a 2-edge $xy \in E(G)$, the edge can be removed by one of the endpoints, say x. This move does not influence the strategy of the agent y, since otherwise there must be a node $v \in V$, which is at distance 1 from x and $d_G(y,v) = w(y,x) + w(x,v) = 2 + 1$, and then we could assign the ownership of xy to agent y and then the edge xy would not be removed from G. Therefore, all the strategies can intersect only in pairs of nodes that want to add the same edge.

Note that for any $y \in U$ we have $S_{y+} \cap S_{u+} = \emptyset$ and $S_{y-} \cap S_{u-} = \{uy\}$, and for all $v \in V$, we have $d_{G^*}(u,v) \leq 3$ and $d_{G^*}(y,v) \leq 3$. The number of edges in G is

$$|E(G)| = \Big|\tilde{E} \cup \bigcup_{y \in U} S_{y-} \cup S_{u-}\Big| = |\tilde{E}| + \sum_{y \in U} |S_{y-}| + |S_{u-}| - |U|,$$

where $\tilde{E} \subset E(G)$ is a set of edges which are both in G and in G^*. On the other hand,

$$\begin{aligned}
|E(G^*)| &= \Big|\tilde{E} \cup \bigcup_{y \in U} S_{y+} \cup S_{u+}\Big| \\
&\leq |\tilde{E}| + \sum_{y \in U} |S_{y+}| + |S_{u+}| \leq |\tilde{E}| + \sum_{y \in U} \big(|S_{y-}| - 1\big) + |S_{u+}| - 1 \\
&= |\tilde{E}| + \sum_{y \in U} |S_{y-}| + |S_{u+}| - |U| - 1 < |E(G)|.
\end{aligned}$$

Hence, since only 2-edges were modified, the new network G^* is a $3/2$-spanner with less weight than the weight of the spanner G, which contradicts that G is a $3/2$-spanner with minimum weight. Therefore, the edge ownership can be chosen such that the network G is in NE. ■

Optimal networks

Now we consider how to compute a social optimum network.

Algorithm 1: social optimum for the 1-2–GNCG

```
input A complete network G = K_n;
while there is 1-1-2 triangle in G do
    Remove the edge of weight 2 from the triangle;
```

▶ **Theorem 4.13.** For any $\alpha \leq 1$, Algorithm 1 produces an optimal network in polynomial time. ◀

Proof. Let G^* be an optimal network. We first prove that there is an optimal network of diameter 2. We assume that G^* has diameter strictly greater than 2. Let u and v be the nodes at distance greater than or equal to 3 in G^*. We show that $G^* + (u, v)$ is also an optimal network. Indeed, the cost of adding the edge to the network is at most $2\alpha \leq 2$, while the sum of the all-to-all distances decreases by at least 2 as the distance between u and v decreases by at least 1.

Next, we show that the social optimum contains all 1-edges. Indeed, if one 1-edge, say uv, were missing in G^*, then $G^* + (u, v)$ would be a network which is cheaper than G, because its edge cost is at most 1 plus the edge cost of G, while its distance cost is at most the distance cost of G minus 2.

Now, observe that the network G produced by Algorithm 1 has diameter equal to 2 and contains all 1-edges. The claim follows by observing that every network of diameter 2 that contains all the 1-edges has to contain all the edges of G.

Finally, it is easy to see that the algorithm runs in polynomial time. Indeed, a network contains $O(n^3)$ triangles, and the algorithm checks each triangle at most once. ■

Price of Anarchy

We start with the following technical lemma observing a relation between stable networks and the corresponding optimum.

▶ **Lemma 4.14.** Consider $0 < \alpha \leq 1$. Let G^* be the social optimum obtained by Algorithm 1 and let G be a stable network. Then $E(G) \subseteq E(G^*)$. Moreover, $d_G(u,v) = 2$ for every 1-edge $uv \notin E(G)$ and $d_G(u,v) \leq 3$ for every 2-edge $uv \notin E(G^*)$. ◀

Proof. We observe that G^* contains all the 1-edges and has diameter 2. So, every 1-edge contained in G is also contained in G^*. Let uv be a 1-edge that is not contained in G. We have $d_G(u,v) = 2$, as otherwise u could buy the edge towards v to improve her cost by at least 1.

Let uv be a 2-edge of G. We show by contradiction that uv is also contained in G^*. Assume that uv is not contained in G^*. Since G^* has diameter 2, there exists a node x such that ux and vx are two 1-edges. First of all, we observe that G cannot contain both the edges ux and vx as otherwise the agent that is buying the 2-edge uv would remove such an edge without increasing any point-to-point distance in the network and thus saving 2α of her edge cost. We split the proof into two cases, according to whether exactly one of the two edges between ux and vx is contained in G, or not, and we show how to obtain a contradiction in either case.

We consider the case in which either ux or vx is an edge of G. W.l.o.g., we assume that ux is an edge of G. Since $d_G(v,x) = 2$, there is a node, say y, such that xy and yv are two 1-edges of G. If the edge uv is bought by v, then v can improve her cost by swapping the edge uv with the edge xv. By this the edge cost decreases by α and no distances from v towards all the other nodes increases. Therefore, the edge uv is bought by agent u. Because G is stable, there is a node z such that the unique shortest path from u to z passes through v, as otherwise u would never have bought the edge towards v. Therefore, we have

$$1 + d_G(x,z) = d_G(u,z) + d_G(x,z) \geq d_G(u,v) + d_G(v,z) + 1 = d_G(v,z) + 3$$

which implies that $d_G(x,z) \geq d_G(v,z) + 2$. But in this case, x can improve on her cost by buying the 1-edge towards v. By this, her edge cost increases by at most 1 while both the distances towards v and z decrease by at least 1. Hence, G could not be stable.

We consider the case in which neither ux nor vx is an edge of G. Since $d_G(u,x) = d_G(v,x) = 2$, there are two nodes, say y and z, such that uy, yx, xz, and zv are four 1-edges in G. We claim that $d_G(u,z) = d_G(v,y) = 2$. We prove the claim for $d_G(u,z)$ as the proof for $d_G(v,y)$ uses similar arguments. The claim is proved by contradiction. If $d_G(u,z) = 1$, then the agent buying the edge uv may remove such an edge, without increasing any point-to-point distance, and thus saving a cost of 2α. If $d_G(u,z) \geq 3$, then u can improve on her cost by buying the 1-edge towards x. By this, the edge cost increases by at most 1 while both the distances towards x and z decrease by 1. As a consequence, there is a vertex w such that the unique shortest path from u to w in G passes through v, as otherwise u would never bought the edge towards v. Therefore, we have

$$2 + d_G(x,w) = d_G(u,x) + d_G(x,w) \geq d_G(u,v) + d_G(v,w) + 1 = d_G(v,z) + 3$$

which implies that $d_G(x,w) \geq d_G(v,z)+1$. But in this case x can improve on her cost by buying the 1-edge towards v. Indeed, by this the edge cost increases by at most 1 while both the distances towards v and w decrease by at least 1. Hence, G could not be stable.

To complete the proof, it remains to show that $d_G(u,v) \leq 3$ for every 2-edge uv that is not in G^*. Let uv be a 2-edge that is not in G^*. Since $E(G) \subseteq E(G^*)$, uv is not contained in G. We prove by contradiction that $d_G(u,v) \leq 3$. For the sake of contradiction, assume that $d_G(u,v) \geq 4$. Since G^* has diameter 2, there is a vertex x such that ux and vx are two 1-edges. Since $d_G(u,x), d_G(v,x) \leq 2$, both edges ux and xv are missing from G. Furthermore, there are two nodes, say y and z, such that uy, yx, xz, and zv are 1-edges in G. Because $d_G(u,v) \geq 4$, we have that $d_G(u,z) \geq 3$. In this case, u can improve her cost by buying the edge towards x. By this, the edge cost increases by $\alpha \leq 1$ while all the distances towards x, z, and v decrease by 1. ■

► **Theorem 4.15.** For $1/2 \leq \alpha < 1$, the PoA is at most $3/(\alpha+2)$. ◄

Proof. First of all, we observe that both the social optimum and any NE contain all the 1-edges.

Let G be a NE and let u and v be two distinct nodes. Let x and x^* be two Boolean variables such that $x = 1$ iff uv is an edge of G, and $x^* = 1$ iff uv is an edge of the social optimum $\mathrm{OPT_n}$. We prove the claim by showing that

$$\sigma := \frac{\alpha \cdot w(u,v)x + 2d_G(u,v)}{\alpha \cdot w(u,v)x^* + 2d_{\mathrm{OPT_n}}(u,v)} \leq \frac{3}{\alpha+2}.$$

First of all, we observe that if $w(u,v) = 1$, then $x = x^* = 1$. Furthermore, if

$x = x^* = 1$, then $\sigma = 1$. Therefore, we only need to prove the claim for the case in which $w(u,v) = 2$ and x and x^* that cannot be both equal to 1.

Let G' be the network induced by all the 1-edges. We observe that if $d_{G'}(u,v) = 2$, then neither $\mathrm{OPT_n}$ nor G contains the edge uv since G' is a sub-network of both $\mathrm{OPT_n}$ and G. Therefore, we assume that $d_{G'}(u,v) \geq 3$. In this case, we have that $x^* = 1$: indeed, the addition of the edge uv to $\mathrm{OPT_n}$ would increase the edge cost by 2α, but would decrease the overall sum of all-to-all distances by at least 2. Similarly, if $d_{G'}(u,v) \geq 4$, then $x = 1$. Since we are considering the case in which x and x^* cannot be both equal to 1, but $x^* = 1$, it follows that $x = 0$. Therefore, $d_{G'}(u,v) = 3$ and thus, $d_G(u,v) = 3$. Hence, $\sigma \leq 6/(2\alpha + 4) = 3/(\alpha + 2)$. The claim follows. ■

We proceed with a lower bound on the PoA which matches the upper bounds given in Theorem 4.4 and Theorem 4.15.

▶ **Theorem 4.16.** For every constant $\epsilon > 0$,

$$PoA \geq \begin{cases} 3/2 - \varepsilon & \text{if } \alpha = 1; \\ 3/(\alpha + 2) - \varepsilon & \text{if } 1/2 \leq \alpha < 1. \end{cases}$$

◀

Proof. We prove the lower bound for $\alpha = 1$ first. Consider the host network H contains a clique K of N nodes formed by 1-edges only. Each node v of the clique is the center of star X_v made of 1-edges only and whose leaves are N new nodes. Finally, there is a new node, that we call u, that is connected to every other node by a 1-edge. Thus, the overall number of nodes of the host network is $n = N^2 + N + 1$. All other edges are 2-edges.

We observe that the social optimum corresponds to exactly the subnetwork induced by the 1-edges. Therefore, the edge cost of the social optimum is $O(N^2)$, while the distance cost is at most $2N^4 + 2N^2$. Therefore, the social cost of the social optimum is at most $2N^4 + O(N^2)$.

We claim that the sub-network induced by all the 1-edges, except for those among u and the leaves of each of the X_v, is a NE. Indeed, since the resulting network has diameter 3, no agent has an incentive to buy a 2-edge. Furthermore, no agent has an incentive in removing a 1-edge. Finally, neither u nor any leaf of any star X_v has an incentive in buying the leftover 1-edge connecting them. The edge cost of the stable network is $O(N^2)$, while the distance cost is at least

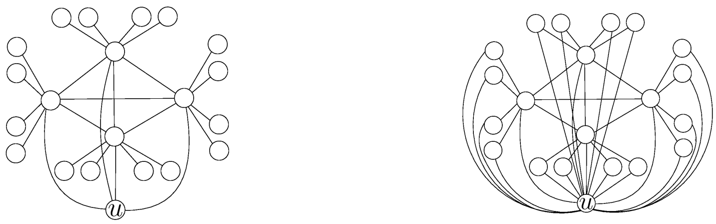


Figure 4.3: NE and optimal networks for $N = 4$. All depicted edges are 1-edges. On the right hand side, the optimal network for $\alpha = 1$ is depicted. The network on the left hand side is a NE for every $1/2 \leq \alpha \leq 1$.

$3N^2(N(N-1)) = 3N^4 - 3N^3$, since any leaf of any star X_v has all the $N(N-1)$ leaves of the other stars at distance 3. Therefore, the social cost of the stable network is $3N^4 - \Theta(N^3)$. The claim for $\alpha = 1$ follows by choosing a sufficiently large value of N.

Now, we prove the claim for $1/2 \leq \alpha < 1$. Let the network on the left hand side of Figure 4.3 be the host network (only 1-edges are depicted). A trivial upper bound on the social optimum is the cost of the entire host network. Therefore, the edge cost of the social optimum is upper bounded by $2\alpha\frac{(N^2+N+1)(N^2+N)}{2} = \alpha N^4 + \Theta(N^3)$, while the distance cost is upper bounded by $2(N^2+N+1)(N^2+N) = 2N^4 + \Theta(N^3)$. Therefore, the cost of the social optimum is $(\alpha+2)N^4 + \Theta(N^3)$. Once again, the sub-network induced by all the 1-edges is a NE as, for $\alpha < 1$, any NE contains all the 1-edges. Furthermore, since the resulting network has diameter 3, no agent has an incentive to buy a 2-edge as $\alpha \geq 1/2$. As already proved for the case $\alpha = 1$, the social cost of the NE network is $3N^4 - \Theta(N^3)$. The claim now follows by choosing a sufficiently large value of N. ■

Now we show that selfishness does not lead to a loss in social welfare if α is small enough.

▶ **Theorem 4.17.** For any $0 \leq \alpha < \frac{1}{2}$ the PoA is equal to 1. ◀

Proof. We claim that, if $0 \leq \alpha < \frac{1}{2}$, any NE network is equal to the optimal network produced by Algorithm 1. To show this we need to prove that a NE network does not contain 1-1-2 triangles but contains all the other edges.

Consider a network which is in NE. It is easy to see that all 1-edges are contained in G because otherwise, the addition of any such edge improves distance cost by at least 1 and increases the edge cost by $\alpha < 1$. Moreover, if there is a 2-edge uv, which is not in G and does not form a 1-1-2 triage in $G + uv$, then the addition of uv improves the distance cost by $d_G(u, v) - w(u, v) \geq 3 - 2 = 1$ and increases the edge cost by $2\alpha < 1$, i.e., it implies an improvement for an owner of the edge. Finally, it is clear that G does not contain 1-1-2 triangles because removing a 2-edge from such triangle does not change the distance cost. Therefore, G is equal to the social optimum obtained by Algorithm 1. ■

1-2-networks for $\alpha > 1$

In this section we show that the 1-2–GNCG for $\alpha > 1$ behaves very similar to the original NCG.

▶ **Theorem 4.18.** For $\alpha \geq 3$ any star network is in NE. ◀

Proof. Consider a star network S_n that has $n - 1$ edges. Assume that the central node u is an owner of all edges in the star. Then u cannot improve her strategy. Let v, z be two leaf nodes. The only possible strategy improvement for a leaf node is an edge addition. In the worst case $w(v, u) = w(z, u) = 2$ and $w(z, v) = 1$, thus, adding an edge zv improves the distance only between the edge endpoints by 3 and costs $\alpha \geq 3$. Therefore, there is no strategy improvement for any agent. This implies that S_n is in NE. ■

Price of Anarchy

We use the proof technique from [Fab+03] to show that the PoA may be bounded by the same value as in the original proof. We start with the bounding the social cost of the NE.

▶ **Lemma 4.19.** Consider any NE network G in the 1-2–GNCG. If G has diameter D, then its social cost is at most $O(D)$ times the social cost of the optimal network OPT. ◀

Proof. First, we evaluate the social cost of the optimal network. Since the network should be connected and each edge has length at least 1, the total cost is in $\Omega(\alpha \cdot n + n^2)$.

Now we analyze the social cost of $G = (V, E)$ which is in NE. The distance cost is trivially in $O(Dn^2)$ since each pair of nodes is in distance D in G. To evaluate the edge cost we consider cut edges, whose removal disconnects G. There are at most $n - 1$ cut edges in the network, thus, the edge cost is at most $O(\alpha(n - 1))$

plus the edge cost of the non-cut edges. Now consider a node v in G which has at least one non-cut edge. We claim that the number of the non-cut edges of v is at most $n(2D+1)/\alpha$, thus, the total edge cost of all non-cut edges in G is at most $2n^2(2D+1)$, which implies that $cost(G) \in O(\alpha(n-1)+2n^2(2D+1)+Dn^2) = O(\alpha n + Dn^2)$. And therefore we conclude that the ratio between $cost(G)$ and the cost of the optimal solution is in $O(D)$.

Consider an edge $e = uv \in E(G)$, owned by the agent u. Let V_e be the set of nodes w, such that the edge e is in the shortest path from u to w. Let G' be the network G without the edge e. Since G is stable, we have

$$0 \le cost_{G'}(u) - cost_G(u) \le -\alpha \cdot w(u,v) + (d_{G'}(u,v) - w(u,v)) \cdot |V_e|.$$

We claim that $d_{G'}(u,v) \le 2D$. Indeed, consider a cycle which consists of the shortest path from u to v in G' and the edge uv. Let $v' \in V_e$ be the node which is furthest away from node v in the cycle and let $u'v'$ be its incident edge in the cycle such that $u' \notin V_e$. The node u' exists because clearly, not all nodes of the cycle are in V_e. Since $u' \notin V_{uv}$, we have $d_{G'}(u,u') = d_G(u,u') \le D$, and since $v' \in V_e$ we get $d_{G'}(v,v') = d_G(v,v') \le D$. Thus, $d'_G(u,v) \le d_{G'}(u,u') + w(u',v') + d_{G'}(v',v) \le 2D+2$. Therefore, we get

$$0 \le -\alpha \cdot w(u,v) + (d_{G'}(u,v) - w(u,v)) \cdot |V_e| \le -\alpha + (2D+2-1)|V_e|.$$

It follows that $|V_e| \ge \alpha/(2D+1)$. Thus, the total number of non-cut edges of v in G is at most $n(2D+1)/\alpha$. This completes the proof. ■

With the above lemma we can easily get the following.

▶ **Theorem 4.20.** The 1-2–GNCG with $\alpha > 1$ has a $PoA \in O(\sqrt{\alpha})$ for any host network for which a NE exists. ◀

Proof. Using Lemma 4.19, we only need to prove that the diameter of the NE is at most $\sqrt{\alpha}$. We consider a pair of nodes u, v in the network G, which is in NE and has diameter D. Assume that $d_G(u,v) = D$. Since G is in NE, the addition of the edge uv does not yield an improvement for agent u. Thus, $0 \le cost_{G+(u,v)}(u) - cost_G(u)$. Let $P := v = v_1, v_2, \ldots, v_{m-1}, v_m = u$ be the shortest u–v path in G and let $k = D/5$. We observe that the distances from u to $v_1, \ldots, v_k$ will all change after the addition of the edge uv. Thus, taking into account that each edge has length at most 2, we have:

$$
\begin{aligned}
0 &\leq cost_{G+uv}(u) - cost_G(v) \\
&\leq \alpha \cdot w(u,v) + \sum_{i=1}^{k} \big(w(u,v) + d_{G+uv}(v,v_i) - d_G(u,v_i)\big) \\
&\leq 2\alpha + \sum_{i=1}^{k} \big(2i - (D-2k)\big) \leq 2\alpha + \sum_{i=1}^{k} (4k - D) \leq 2\alpha - \frac{D^2}{25}
\end{aligned}
$$

It follows that $D \in O(\sqrt{\alpha})$. ■

We are convinced that also other proof techniques from the NCG can be carried over to the 1-2–GNCG. Thus, the PoA should be constant for almost all α and in $o(n^{\varepsilon})$ for the remaining range.

4.5.3 Tree Metrics

This section is devoted to the study of tree metrics. We assume that the host network $H = (V, E)$ is defined as the *metric closure* of an edge-weighted tree T. More precisely, $w(u,v) = d_T(u,v)$ for every two nodes u and v.

Existence of Nash Equilibria

The first result is about the structure of any NE. Differently for general metrics and 1-2-networks, we prove that any NE in T–GNCG is as sparse as possible.

▶ **Theorem 4.21.** In the T–GNCG any NE is a tree. ◀

Proof. Consider a network $G = (V, E)$ which is in NE. For the sake of contradiction, we assume that G contains a cycle. Clearly, this cycle has at least one edge, say uv, which is not contained in the tree T. Without loss of generality, assume u be the owner of the edge uv. Consider a vertex $x \in V$ such that the edge xv is in the unique shortest u-v–path in T. Note that $xv \notin E(G)$, otherwise swapping the edge uv to ux is an improving move that contradicts with G being in NE. Consider two possible situations: $d_G(u,x) > w(u,x)$ and $d_G(u,x) = w(u,x)$.

In case $d_G(u,x) > w(u,x)$, consider a network $G' = (V, E')$ obtained from G by swapping the edge uv to ux by agent u. Denote by $Z = \{z \in V : d_G(u,z) < d_{G'}(u,z)\}$ the set of nodes to which the distance from u has increased. Note that

Z is not an empty set because $v \in Z$. Since G is in NE, $\alpha \cdot w(u,x) + w(u,x) + d_{G'}(u,Z) \geq \alpha \cdot w(u,v) + d_G(u,x) + d_G(u,Z)$, whereas the left part of the inequality is at most $\alpha \cdot w(u,x) + w(u,x) + |Z| \cdot w(u,x) + d_G(x,Z)$, and the right part is equal to $\alpha \cdot w(u,v) + d_G(u,x) + |Z| \cdot w(u,v) + d_G(v,Z)$. Since $d_G(u,x) > w(u,x)$, we get

$$\alpha \cdot w(u,x) + |Z| \cdot w(u,x) + d_G(x,Z) > \alpha \cdot w(u,v) + |Z| \cdot w(u,v) + d_G(v,Z).$$

In case $d_G(u,x) = w(u,x)$, consider deletion of the edge uv. Since uv is in the cycle, the deletion does not increase any distance in G' to infinity. We use the same notation: $G' = (V, E')$ is a network after modification, Z is a set of nodes to which the distance from u has increased. As before, $Z \neq \emptyset$ because otherwise deletion of the edge uv is an improving move for agent u. Since G is stable,

$$\begin{aligned}|Z| \cdot w(u,x) + d_G(x,Z) \geq d_{G'}(u,Z) &\geq \alpha \cdot w(u,v) + d_G(u,Z)\\ &= \alpha \cdot w(u,v) + |Z| \cdot w(u,v) + d_G(v,Z).\end{aligned}$$

Adding positive term $\alpha \cdot w(u,x)$ to the left part of the inequality, we get the same inequality as in the previous case:

$$\alpha \cdot w(u,x) + |Z| \cdot w(u,x) + d_G(x,Z) > \alpha \cdot w(u,v) + |Z| \cdot w(u,v) + d_G(v,Z).$$

Note that the agent x does not buy the edge xv. Therefore, for the new network $G'' = (V, E'')$ obtained after the modification, $\alpha \cdot w(x,v) + |Z| \cdot w(x,v) + d_G(v,Z) \geq cost(x, G'') \geq cost(x, G) \geq d_G(x,Z)$. We sum up this inequality with the inequality we obtained by analyzing swapping and deletion, and we get

$$\begin{aligned}\alpha(w(x,v) + w(u,x)) + |Z|(w(x,v) + w(u,x)) + d_G(x,Z) + d_G(v,Z)\\ > \alpha \cdot w(u,v) + |Z| \cdot w(u,v) + d_G(v,Z) + d_G(x,Z).\end{aligned}$$

Simplifying and taking into account that $w(u,x) + w(x,v) = w(u,v)$, we get that $(\alpha + |Z|) \cdot w(u,v) > (\alpha + |Z|) \cdot w(u,v)$, that is a contradiction. Therefore, G has no cycles. Obviously, G is connected since otherwise any agent can improve her strategy by adding edges to other components. This implies that G is a tree. ■

The next result follows by observing that the tree T that defines the metric is the network with cheapest total edge cost that preserves all the distances of the host network at the same time.

▶ **Proposition 4.22.** In the T–GNCG the tree T which defines the metric is both the social optimum and in NE. ◀

Proposition 4.22 claims that the cheapest NE is also a social optimum. This is equivalent to say that the Price of Stability – the ratio between the cost of the cheapest NE and the cost of a social optimum – is 1.

Hardness

We prove that the problem of computing the best response of an agent is NP-hard for the (T–GNCG).

▶ **Theorem 4.23.** It is NP-hard to compute a best response of an agent in the T–GNCG. ◀

Proof. We perform the proof by a reduction from the Minimum Set Cover problem, which is well-known to be NP-hard. The problem is defined as follows: for a given universe $U = \{1, 2, \ldots, k\}$ and a collection of non-empty subsets $\mathcal{X} = \{X_1, \ldots, X_m\}$ such that for any $1 \leq i \leq m$ we have $X_i \subseteq U$ and $\bigcup_{i=1}^{m} X_i = U$. It is required to find minimum number of subsets covering U.

We define the corresponding instance of the best response problem in the T–GNCG with $\alpha = 1$ as follows: Consider a tree $T = (V, E_T)$ which defines metric such that

$$V = \{u, c\} \cup \{a_1, \ldots, a_m\} \cup \{b_1, \ldots, b_m\} \cup \{p_1, \ldots, p_k\}$$

and

$$E = \{cu\} \cup \bigcup_{i=1}^{m} (\{b_i u, ca_i\} \cup \{a_i p_j |\, p_j \in X_i\}),$$

where each p_j represents one element of the universe U and each a_i corresponds to one subset X_i. All nodes $c, b_1, \ldots, b_m$ are connected with u and each edge ub_i has length $\frac{1}{2}(L - \beta)$, whereas the edge uc is of length $L - \varepsilon$. Each of set nodes $a_1, \ldots, a_m$ is connected with c by an edge of length ε. Furthermore, all edges between the element nodes $p_1, \ldots, p_m$ and the set nodes are of length L and each set element node is connected with only one set node. We assume throughout the proof that $L >> \varepsilon$ holds. Moreover, we assume that each edge $b_i u$ is owned by the respective node b_i. Finally, note that agent u does not own any edges in G. See the right side Figure 4.4 for the illustration of the constructed network.

Consider a network $G = (V, E)$ such that each node $c, b_1, \dots, b_m$ is an owner of the edge connecting it with the node u. For all $i = 1, \dots, m$ there is an edge $b_i a_i$ of length $\frac{1}{2}(L - \beta) + (L - \varepsilon) + \varepsilon = \frac{1}{2}(L - \beta) + L$. Also each element node p_j is connected with some set node a_i iff the element is in the set corresponding to the set node. Note that $L \le w(a_i, p_j) \le L + 2\varepsilon$. See Figure 4.4(left) for the illustration of the constructed network G and the metric tree T.

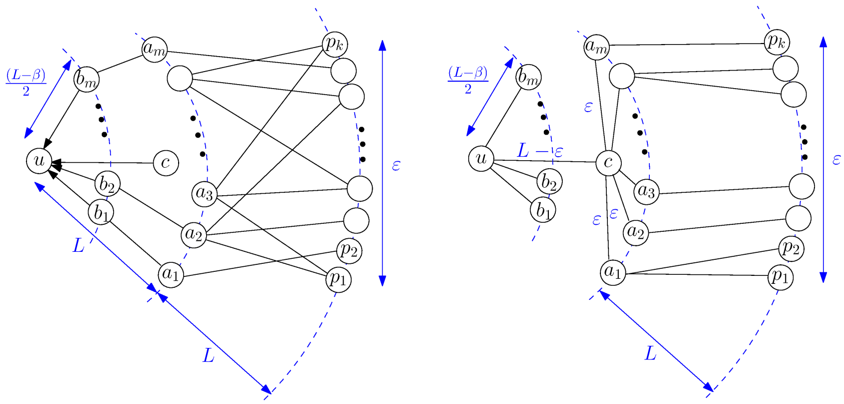


Figure 4.4: Illustration of the construction used in the reduction. The left figure shows the constructed network G. On the right figure the given metric tree T is shown .

We claim that the best response S_u^* of the agent u contains only the set nodes corresponding to a solution of the Minimum Set Cover problem.

First, we prove that S_u^* includes only set nodes. Assume that there is a node $p_j \in S_u^*$. Consider a node a_i corresponding to the set X_i containing element p_j. If $a_i \in S_u^*$, then the agent u can delete the edge up_j because this move decreases her edge cost by $w(u, p_j) = 2L$ and increases her distance cost by at most 2ε, thus, her total cost decreases by at least $L - 2\varepsilon$. If $a_i \notin S_u^*$, then swap of the edge up_j to ua_i improves agent's cost by at least $L - 2\epsilon - \beta$ because it decreases her edge cost by L, decreases distance cost to at least one node, which is a_i, by $L - \beta$ and increases distance only to the node p_j by 2ε.

Next we show that all the nodes in S_u^* correspond to a set cover. Indeed, if there is a node p_j such that none of its incident set nodes is in S_u^*, then buying an edge to one of the corresponding set node a_i decreases distance between the node u and both nodes a_i, p_j by at least $2(2L - \beta - L) = 2(L - \beta)$. Hence, the social cost of the agent u is improves by at least $L - 2\beta$.

Finally, we show that S_u^* corresponds to a minimum set cover. Consider two strategies S_u^1, S_u^2 corresponding to two different set covers of all elements. Assume $\Delta := |S_u^2| - |S_u^1| > 0$. Thus, the difference in the agent's u cost with strategy S_u^1 compared with the strategy S_u^2 is

$$-\Delta \cdot L + (L - \beta)\Delta + 2k\varepsilon = -\Delta\beta + 2k\varepsilon < 0. \qquad \blacksquare$$

Dynamic Properties

The following theorem shows that the network dynamics consisting of best responses only may never converge to a NE for the T–GNCG and thus also for the M–GNCG.

▶ **Theorem 4.24.** The T–GNCG is not a potential game. ◀

Proof. Consider the weighted tree depicted in Figure 4.5 (left). With this tree defining the metric distances, we can construct an improving response cycle[3] of length 4 for $\alpha = 1$. See Figure 4.5 (right)).

In the first step ($G_1 \longrightarrow G_2$), the agent a_0 swaps the edge a_0a_3 to a_0a_4. It is an improving move because the edge a_0a_3 saves the distance from a_0 to only two nodes, a_3 and a_4. Then this move changes the cost of a_0 by

$$\begin{aligned} cost(a_0, G_2) - cost(a_0, G_1) \leq& -\alpha \cdot w(a_0, a_3) + \alpha \cdot w(a_0, a_4) + d_{G_2}(a_0, a_3) \\ &- d_{G_1}(a_0, a_3) + d_{G_2}(a_0, a_4) - d_{G_1}(a_0, a_4) \\ =& - 12 + 18 + 18 - 12 + 18 - 32 < 0. \end{aligned}$$

In the next step ($G_2 \longrightarrow G_3$), the edge a_2a_3 is deleted by a_2. Since the edge a_2a_3 contains in a shortest path from a_2 to a_3 only. Hence, it is an improving move because

$$\begin{aligned} cost(a_2, G_3) - cost(a_2, G_2) &\leq -w(a_2, a_3) + d_{G_3}(a_2, a_3) - d_{G_2}(a_2, a_3) \\ &= -2 \cdot 10 + 18 < 0. \end{aligned}$$

In the step $G_3 \longrightarrow G_4$, the agent a_0 swaps the edge a_0a_4 back to a_0a_3. This move

3 It is enough to show that an improving response cycle exists. However, it is possible to show that the cycle is a best-response cycle.

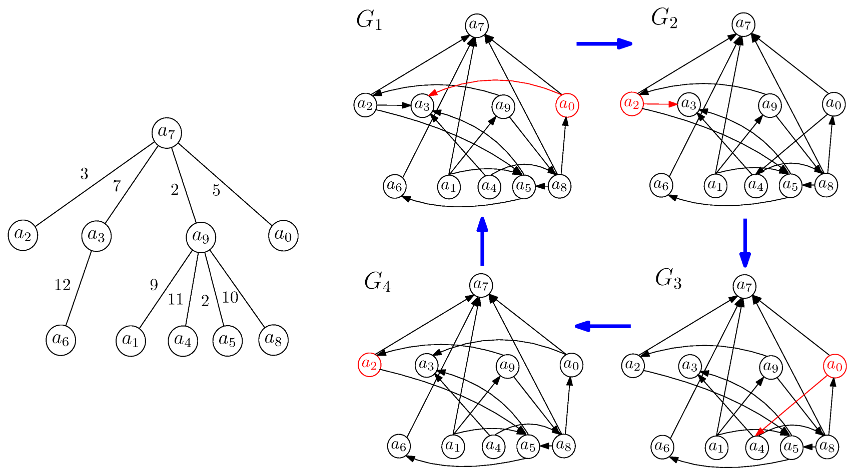


Figure 4.5: Left: Tree which defines the metric. Right: Best response cycle. The active agent is colored red. The direction of the edges denotes the owner.

improves distance to the node a_3 and increase the distance to a_4 only. Therefore,

$$\begin{aligned} cost(a_0, G_4) - cost(a_0, G_3) \leq & - w(a_0, a_4) + w(a_0, a_3) + d_{G_4}(a_0, a_4) \\ & - d_{G_3}(a_0, a_4) + d_{G_4}(a_0, a_3) - d_{G_3}(a_0, a_3) \\ = & - 18 + 12 + 32 - 18 + 12 - 28 < 0. \end{aligned}$$

The last step ($G_4 \rightarrow G_1$) closes the cycle. Here the agent a_2 creates the edge to a_3. It costs 10 and improves the distance to the nodes a_3 and a_4 by 20, i.e., it is an improving move. ■

Price of Anarchy

We prove that the upper bound given in Theorem 4.4 for more general metric instances is tight for tree metrics and thus also for network metrics.

▶ **Theorem 4.25.** The PoA in the T–GNCG is at least $\frac{\alpha+2}{2} - \varepsilon$, for any $\varepsilon > 0$. ◀

Proof. Let S_n^* be the weighted tree which defines the metric distances. The tree S_n^* is a star and contains $n-2$ edges of weight $2/\alpha$ and one edge uv of weight 1, where u is the center of the star. See Figure 4.6 (left).

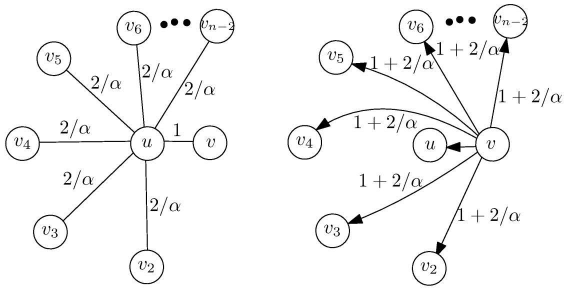


Figure 4.6: Left hand side figure shows a tree T, on the right figure is shown the stable network S_n.

The star S_n^* with the central node u minimizes the social cost, which is

$$\begin{aligned}&SC(S_n^*)\\&= \alpha \cdot w(u, V) + w(v, V) + (n-2)(w(u, V) + (n-2)\cdot 2/\alpha) + (w(u, V) + n - 2)\\&= (2n+\alpha-2)\cdot w(u, V) = (2n+\alpha-2)\cdot((n-2)2/\alpha+1).\end{aligned}$$

Let S_n be a spanning star of the host network such that a center of S_n is the node v. The star S_n contains one edge of weight 1 and $(n-2)$ edges of weight $(1+2/\alpha)$. Moreover, we assume that the central vertex v is an owner of all edges in S_n. See Figure 4.6 (right).

We claim that S_n is in NE. Indeed, the central agent v cannot improve her strategy because all other nodes are leaves. No leaf owns an edge and, therefore, a leaf agent can possibly improve her strategy only by adding edges to other leaves. For any leaf agent $x \neq u$ buying the edge xu costs $\alpha \cdot 2/\alpha = 2$ and improves her distances only towards u by $2 + 2/\alpha - 2/\alpha = 2$. Hence, buying xu is not an improvement. Buying any other edge costs $\alpha \cdot 4/\alpha = 4$ for agent x and improves her distance only to the endpoint of the new edge by $2+4/\alpha-4/\alpha = 2$. At the same time, the agent u cannot improve her strategy by buying edges to the leafs because it improves distance to each v_i by 2 and increases edge cost by the same value for each new edge. Hence, no agent has an improving strategy change and it follows that S_n is in NE. The social cost of S_n is

$$SC(S_n) = (2n+\alpha-2)\cdot w(v, V) = (2n+\alpha-2)\cdot((n-2)(1+2/\alpha)+1).$$

Then, for sufficiently large n, the ratio between he social costs of the NE network S_n and the optimum S_n^* is $\frac{\alpha+2}{2} - \varepsilon$. ■

4.5.4 Points in $\mathbb{R}^d$

In this section we consider the M–GNCG with the assumption that all nodes are points in $\mathbf{R}^d$ and that distances are measured via the p-norm, i.e., for any two points $u = (u_1, \ldots, u_d), v = (v_1, \ldots, v_d)$ the weight of the corresponding edge between them is defined as

$$w_p(u,v) := \left(\sum_{i=1}^{d} |u_i - v_i|^p \right)^{1/p}.$$

Further, we omit the subscript p if its value does not play any role.

Hardness

We start with investigating the hardness of computing the best response of an agent in the $\mathbf{R}^d$–GNCG.

► **Theorem 4.26.** It is NP-hard to compute a best response of an agent in the $\mathbf{R}^d$–GNCG under any p-norm. ◄

Proof. We perform the proof by a reduction from the Minimum Set Cover problem analogously to the proof of the Theorem 4.23. We define the corresponding instance of the best response problem in the $\mathbf{R}^d$–GNCG with $\alpha = 1$ as follows: Consider a network $G = (V, E)$ such that

$$V = \{u\} \cup \{a_1, \ldots, a_m\} \cup \{b_1, \ldots, b_m\} \cup \{p_1, \ldots, p_k\}$$

and

$$E = \bigcup_{i=1}^{m} \{b_i u, b_i a_i\} \cup \bigcup_{i=1}^{m} \bigcup_{p_j \in X_i} \{a_i p_j\},$$

where each p_j represents one element of the universe U and each a_i corresponds to one subset X_i. We locate nodes on the plane such that all points $a_1, \ldots, a_m$ are at the same distance L from u and equally spaced on the circle segment of length equal to some arbitrary small value $\varepsilon > 0$. All points $p_1, \ldots, p_k$ are equispaced on the circle segment of the same length ε and are at distance $2L$ from u. We assume throughout the proof that $L >> \varepsilon$ holds. In addition, we have a set of nodes $\{b_1, \ldots, b_m\}$ at distance $\frac{1}{2}(L-\beta)$, where $\frac{1}{3}L > \beta > k\varepsilon$, such that each b_i, for $1 \leq i \leq m$ lies on the line through the nodes u and a_i and is connected to nodes u and a_i. By construction we get that $d_G(b_i, a_i) = \frac{1}{2}(L-\beta) + L$, $d_G(u, a_i) = 2L - \beta$ and $d_G(u, p_j) = 3L - \beta$. Moreover, we assume that each edge $b_i u$ is owned by the respective node b_i. Finally, note that agent u does not own any edges in G. See Figure 4.7 for the illustration of the constructed network.

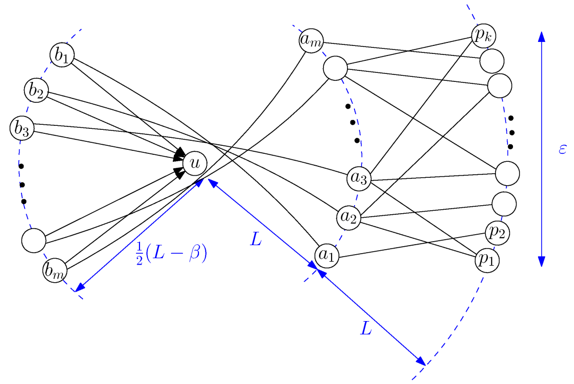


Figure 4.7: Illustration of the construction used in the reduction.

We claim that the best response of the agent u corresponds to the solution of the Minimum Set Cover problem.

Consider a best response strategy S_u^* of agent u and the corresponding network G^* which is the network G augmented by the edges which agent u buys according to strategy S_u^*. Thus, $G^* = (V, E^*)$, where $E^* = E(G) \cup \bigcup_{v \in S_u^*} \{uv\}$.

First of all, we show that $S_u^* \neq \emptyset$ must hold, since adding at least one edge ua_i to network G is already an improvement for agent u. This edge costs L and decreases agent u's distance to a_i by $L - \beta$. Moreover, since X_i is non-empty there must be a neighboring p-node of node a_i. Let p_j be this node. Buying edge ua_i changes agent u's distance to p_j from at least $3L - \beta$ to at most $2L + \varepsilon$, which is an improvement of at least $L - \beta - \varepsilon$. Hence, the total improvement for agent u is at least $L - 2\beta - \varepsilon > 0$.

Now we prove that agent u always prefers buying edges to a_i nodes over edges to p_j nodes. Assume that $p_j \in S_u^*$ and let a_r be a set node which is adjacent to p_j. Since $\bigcup X_i = U$, such a node a_r must exist by construction. If there is a set node $a_i \in S_u^*$ which is adjacent to p_j in G^*, then agent u can simply delete the edge up_j because it decreases her edge cost by at least $2L$ and increases her distance cost by at most ε, since only the distance to the nodes p_j can increase by at most ε, since for any $1 \leq r \leq m$ holds $w(a_i, a_r) \leq \varepsilon$.

Since $L >> \varepsilon$, deleting up_j would be an improving move for agent u. If there is no node $a_i \in S_u^*$ such that $a_i p_j \in E(G^*)$, then the swap of the edge up_j to ua_r improves agent u's cost by at least $2L - \varepsilon - \beta$ since this move improves u's distance to at least node a_r by $L - \beta$, increases u's distance only to the nodes p_j by ε, respectively, and improves u's total edge cost by L. Hence, S_u^* cannot contain p_j nodes.

Next, we show that every p_i node is adjacent to some node $a_i \in S_u^*$, i.e., that the corresponding set of subsets $\{X_i \mid a_i \in S_u^*\}$ is a set cover of U. For the sake of contradiction, assume that there is a node p_j for which there is no node $a_i \in S_u^*$ such that $a_i p_j \in E(G^*)$. Clearly, there must be a path from u to p_j in G^* since otherwise agent u would have infinite cost. Let $a_r \notin S_u^*$ be any set node, for which $a_r p_j \in E(G^*)$. Such a node a_r must exist, since $\bigcup X_i = U$. Thus, we have that $d_{G^*}(u, p_j) \geq 3L - \beta$, since there is a path from u to p_j via b_r and a_r. We claim that agent u could buy the edge ua_r and thereby strictly decrease her cost. The edge ua_r costs L and decreases agent u's distances to a_r by $L - \beta$ and to each of the nodes p_j by at least $L - \beta$. Thus, this yields a cost decrease for agent u. Note, that this implies that agent u can improve on any strategy S_u, where the corresponding set of subsets $\{X_i \mid a_i \in S_u\}$ does not cover all elements of U. Thus, the set of subsets $\{X_i \mid a_i \in S_u^*\}$ must be a set cover of U.

We finish the proof by showing that the best response strategy of agent u corresponds to a minimum set cover of the given set cover instance. For this, consider two arbitrary strategies S_u^1 and S_u^2 of agent u, such that the corresponding sets $\{X_i \mid a_i \in S_u^1\}$ and $\{X_i \mid a_i \in S_u^2\}$ both cover all elements of U. Now we show that if $|S_u^1| < |S_u^2|$ then u's cost with strategy S_u^1 is strictly less than u's cost with strategy S_u^2. This implies that agent u's best response strategy S_u^* corresponds to a minimum set cover.

Let $\Delta = |S_u^2| - |S_u^1|$. Hence, the difference between agent u's edge cost with strategy S_u^1 and u's edge cost with strategy S_u^2 is exactly $-\Delta \cdot L$. Since both strategies correspond to set covers and since $w(a_1, a_m) = \varepsilon$, the distances of u to any p_j node under the strategies S_u^1 and S_u^2 can differ by at most ε. Moreover, with strategy S_u^1 agent u has distance L to exactly $|S_u^1|$ many a_i nodes and distance $2L - \beta$ to all the other a_i nodes. Analogously, with strategy S_u^2 agent u has distance L to $|S_u^2|$ many a_i nodes and distance $2L - \beta$ to the other a_i nodes. Thus, the total difference in agent u's cost with strategy S_u^1 compared with strategy S_u^2 for agent u is $-\Delta \cdot L + k\varepsilon + \Delta(L - \beta) = -\Delta\beta + k\varepsilon < 0$, where the inequality holds since $\Delta \geq 1$ and by construction we have $\beta > k\varepsilon$. ■

Dynamic Properties

Also for the $\mathbf{R}^d$–GNCG we investigate whether best response dynamics are guaranteed to converge, i.e., if the game has the finite improvement property.

▶ **Theorem 4.27.** The $\mathbf{R}^d$–GNCG with the 1-norm does not have the finite improvement property. ◀

Proof. We prove the statement by providing an improving response cycle[4], shown in Figure 4.8. We use the following node positions: $a_0 = (3, 0), a_1 =$

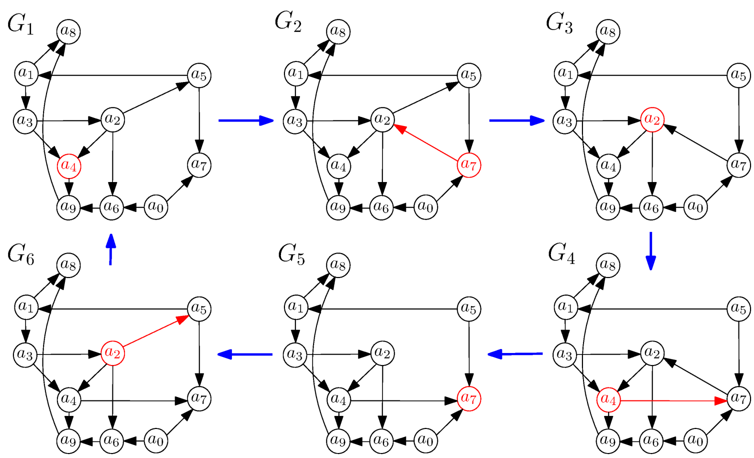


Figure 4.8: Improving response cycle for the $\mathbf{R}^d$–GNCG with 1-norm.

$(0, 3), a_2 = (2, 2), a_3 = (0, 2), a_4 = (1, 1), a_5 = (4, 3), a_6 = (2, 0), a_7 = (4, 1), a_8 = (1, 4), a_9 = (1, 0)$. Distances are measured by the 1-norm.

On the first step from G_1 to G_2, the agent a_7 buys the edge to a_2 because it costs 3 and improves the distance to a_2 and a_3 by 4. Next, a_2 removes the edge a_2a_5. It increases the distance to a_5 by 2 only, while the edge cost decreases by 3. On the next step, the edge a_4a_7 is created since it costs 3 and improves the

4 It even holds that each improving move is a best improving move.

distance to a_7 and a_5 by 4 in total. $G_4 \longrightarrow G_5$: agent a_7 deletes a_7a_2. It decreases the edge cost by 3 and increases the distance to a_2 by 2 only. Next, a_2 creates the edge a_2a_5 because it improves the distance to a_5 by 4 and costs 3. Finally, a_4 deletes a_4a_7 since the edge costs 3, and the move increases the distance to only a_7 by 2. ■

We are convinced that the above best response cycle can be adapted to arbitrary p-norms.

▶ **Conjecture 4.28.** The $\mathbf{R}^d$–GNCG with any p-norm does not have the finite improvement property. ◀

Price of Anarchy

From Theorem 4.4 it follows that in the $\mathbf{R}^d$–GNCG the PoA is at most $\frac{\alpha+2}{2}$. It turns out that settling the PoA for the $\mathbf{R}^d$–GNCG is a challenging problem. We prove some first steps in this direction and show that the PoA approaches the upper bound for the 1-norm if the number of dimensions grows.

We start with a lower bound which is strictly larger than 1 for the PoA in case of an arbitrary p-norm and independent of number of nodes n and dimension d.

▶ **Theorem 4.29.** For $\alpha \geq 2$, the PoA in the $\mathbf{R}^d$–GNCG is in $\Omega(\alpha)$. ◀

Proof. To prove the claim we show a NE network and the corresponding optimal network in the 1-dimensional space, i.e., path metric. It is clear that the obtained result holds for arbitrary $d \geq 1$ and is independent on the p-norm.

We assume $n = \lfloor \alpha \rfloor$ and $\alpha \geq 2$. Consider a path network P_{n+1} on $n+1$ nodes $\{v_0, v_1, \ldots, v_n\}$. We arrange lengths of the edges as follows: $w(v_0, v_1) = 1$, and $\forall i \in \{2, \ldots, n\}$, $w(v_{i-1}, v_i) = \frac{2}{\alpha} \cdot \left(1 + \frac{2}{\alpha}\right)^{i-2}$. Then the host network H is a complete network containing P_{n+1} such that for any pair of nodes u, v $w(u,v) = d_{P_{n+1}}(u,v)$. Clearly, P_{n+1} corresponds to a social optimum.

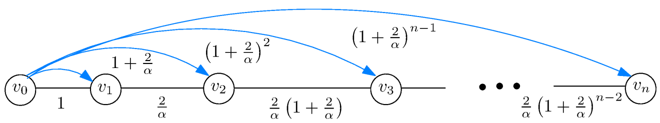


Figure 4.9: Lower bound network. The black network corresponds to the social optimum P_{n+1}, the blue network S_{n+1} is in NE.

Consider a star network $S := S_{n+1}$ on the same set of nodes with the following edge weights

$$w(v_0, v_i) = d_P(v_0, v_i) = \sum_{j=1}^{i} w(v_{j-1}, v_j) = 1 + \frac{2}{\alpha} \cdot \sum_{j=2}^{i} \left(1 + \frac{2}{\alpha}\right)^{j-2}$$
$$= 1 + \frac{2}{\alpha} \cdot \left(\left(1 + \frac{2}{\alpha}\right)^{i-1} - 1\right) \cdot \frac{\alpha}{2} = \left(1 + \frac{2}{\alpha}\right)^{i-1}.$$

See Figure 4.9 for the construction.

First, we show that the star network S_n is in NE. Since no deletions or swaps are possible, we need to prove that no addition of any new edge is profitable for its owner. Indeed, consider a agent v_i, which can buy an edge (v_i, v_j) such that $j \leq i-1$. This move decreases distance cost of the agent by $w(v_i, v_0) + w(v_0, v_j) - w(v_i, v_j) = 2w(v_0, v_j) = 2\left(1 + \frac{2}{\alpha}\right)^{j-1}$. At the same time, the edge cost increases by

$$\alpha \cdot w(v_i, v_j) = \alpha \cdot \sum_{k=j+1}^{i} w(v_{k-1}, v_k) = \alpha \cdot \frac{2}{\alpha}\left(1 + \frac{2}{\alpha}\right)^{j-1} \cdot \left(\left(\frac{2}{\alpha} + 1\right)^{i-j} - 1\right) \cdot \frac{\alpha}{2}$$
$$= \alpha \cdot \left(1 + \frac{2}{\alpha}\right)^{j-1}\left(\left(\frac{2}{\alpha} + 1\right)^{i-j} - 1\right) \geq \alpha \cdot \left(1 + \frac{2}{\alpha}\right)^{j-1} \cdot \left(\frac{2}{\alpha} + 1 - 1\right)$$
$$\geq \alpha \cdot \left(1 + \frac{2}{\alpha}\right)^{j-1} \cdot \frac{2}{\alpha} = 2 \cdot \left(1 + \frac{2}{\alpha}\right)^{j-1}.$$

Thus, this move is not an an improvement for the agent v_i. If $j > i$, distance cost is $2\left(1 + \frac{2}{\alpha}\right)^{i-1}$, whereas the edge cost is

$$\alpha \cdot w(v_i, v_j) = \alpha \cdot \sum_{k=i+1}^{j} w(v_{k-1}, v_k) = \alpha \cdot \left(1 + \frac{2}{\alpha}\right)^{i-1}\left(\left(\frac{2}{\alpha} + 1\right)^{j-i} - 1\right)$$
$$\geq \alpha \cdot \left(1 + \frac{2}{\alpha}\right)^{i-1}\left(\frac{2}{\alpha} + 1 - 1\right) \geq \alpha \cdot \left(1 + \frac{2}{\alpha}\right)^{i-1} \cdot \frac{2}{\alpha}$$
$$= 2 \cdot \left(1 + \frac{2}{\alpha}\right)^{i-1}.$$

Thus, the star network is in NE. The social cost of S_{n+1} is

$$\begin{aligned}
(2n+\alpha)\cdot\sum_{uv\in E(S_n)} w(u,v) &= (2n+\alpha)\sum_{i=1}^{n}\left(1+\frac{2}{\alpha}\right)^{i-1}\\
&= (2n+\alpha)\cdot\frac{\alpha}{2}\left(\left(1+\frac{2}{\alpha}\right)^{n}-1\right)\\
&\text{since } n\geq\alpha-1, \text{ we have}\\
&\geq (2n+\alpha)\cdot\frac{\alpha}{2}\left(1+\frac{2}{\alpha}(\alpha-1)-1\right)\\
&= (2n+\alpha)\cdot\frac{\alpha}{2}\left(2-\frac{2}{\alpha}\right)\\
&\geq (3\alpha-2)\cdot\alpha,\ \text{for } \alpha\geq n\geq 2.
\end{aligned}$$

The social optimum is a path network P_{n+1}. An edge cost is $\alpha\cdot w(v_0,v_n) = \alpha\left(1+\frac{2}{\alpha}\right)^{n-1}$. To calculate the distance cost we count for each edge how many shortest paths it participates, i.e., its betweenness centrality.

$$\begin{aligned}
\sum_{v\in V} d_{P_n}(v,V) &= 2\sum_{i=1}^{n} w(v_{i-1},v_i)i(n-i+1) = 2\sum_{k=0}^{\lfloor n/2\rfloor}(n-2k)\sum_{i=k+1}^{n-k} w(v_{i-1},v_i)\\
&= 2n\sum_{i=1}^{n} w(v_{i-1},v_i) + 2\sum_{k=1}^{\lfloor n/2\rfloor}(n-2k)\sum_{i=k+1}^{n-k} w(v_{i-1},v_i)\\
&= 2n\cdot\left(1+\frac{2}{\alpha}\right)^{n-1} + \frac{4}{\alpha}\sum_{k=1}^{\lfloor n/2\rfloor}(n-2k)\sum_{i=k+1}^{n-k}\left(1+\frac{2}{\alpha}\right)^{i-2}\\
&= 2n\cdot\left(1+\frac{2}{\alpha}\right)^{n-1} + 2\sum_{k=1}^{\lfloor n/2\rfloor}(n-2k)\left(1+\frac{2}{\alpha}\right)^{k-1}\left(\left(1+\frac{2}{\alpha}\right)^{n-2k}-1\right)\\
&\leq 2n\cdot\left(1+\frac{2}{\alpha}\right)^{n-1} + 2n\sum_{k=1}^{\lfloor n/2\rfloor}\left(1+\frac{2}{\alpha}\right)^{n-k-1}\\
&= 2n\cdot\left(1+\frac{2}{\alpha}\right)^{n-1} + (\alpha+2)\cdot n\cdot\left(1+\frac{2}{\alpha}\right)^{n-2}\left(1-\left(1+\frac{2}{\alpha}\right)^{-\lfloor n/2\rfloor}\right)
\end{aligned}$$

$$
\begin{aligned}
&< 2n \cdot \left(1+\frac{2}{\alpha}\right)^{n-1} + (\alpha+2)\cdot n \cdot \left(1+\frac{2}{\alpha}\right)^{n-2} \cdot \left(1-\left(1+\frac{2}{\alpha}\right)^{-1}\right) \\
&\leq 2n \cdot \left(1+\frac{2}{\alpha}\right)^{n-1} + 2n \cdot \left(1+\frac{2}{\alpha}\right)^{n-2} \\
&= 2n \cdot \left(1+\frac{2}{\alpha}\right)^{n-2} \cdot \left(2+\frac{2}{\alpha}\right).
\end{aligned}
$$

Thus, an upper bound of the social cost of the social optimum is

$$
cost(P_n) \leq \left(1+\frac{2}{\alpha}\right)^{n-2} \cdot \left(\alpha\left(1+\frac{2}{\alpha}\right) + 2n\left(2+\frac{2}{\alpha}\right)\right).
$$

Note that the first factor in the above inequality is upper bounded by a constant value. Indeed,

$$
\left(1+\frac{2}{\alpha}\right)^{n-2} = \left(1+\frac{2\cdot\frac{n-2}{\alpha}}{n-2}\right)^{n-2} \leq \exp\left(\frac{n-2}{\alpha}\right) \leq \exp(2).
$$

Next, we simplify the second factor with the assumption $n = \lfloor \alpha \rfloor$:

$$
\alpha\left(1+\frac{2}{\alpha}\right) + 2n\left(2+\frac{2}{\alpha}\right) \leq 5\alpha + 6.
$$

Therefore, for any $\alpha \geq 3$, there is a set of $n = \lfloor \alpha \rfloor$ points such that the PoA ratio is at least:

$$
\frac{cost(S_n)}{cost(P_n)} \geq \frac{(3\alpha-2)\cdot\alpha}{(5\alpha+6)\exp(2)} \in \Omega(\alpha).
$$

Hence, $PoA \in \Omega(\alpha)$. ■

The above lower bound construction also works for small values of α. In the next statement we show that the PoA is strictly larger than 1 for any $\alpha > 0$.

▶ **Theorem 4.30.** In the $\mathbf{R}^d$–GNCG under any p-norm with $p \geq 1$ the PoA is strictly larger than 1. More precisely,

$$
\text{PoA} \geq \frac{3\alpha^3 + 24\alpha^2 + 40\alpha + 24}{\alpha^3 + 10\alpha^2 + 32\alpha + 24} > 1\,.
$$

◀

Proof. We prove the claim for the 1-dimensional case, then the result immediately follows for higher dimensions.

Consider the same construction as in the proof of Lemma 4.29 restricted to 4 nodes v_0, v_1, v_2, v_3. The ratio between the cost of the star network, which is in NE, and the social optimum is

$$\frac{cost(S_4)}{cost(P_4)} = \frac{(6+\alpha)\cdot\frac{\alpha}{2}\cdot\left(\left(1+\frac{2}{\alpha}\right)^3-1\right)}{(6+\alpha)\left(1+\frac{2}{\alpha}\right)^2+\frac{4}{\alpha}} = \frac{3\alpha^3+24\alpha^2+40\alpha+24}{\alpha^3+10\alpha^2+32\alpha+24}.$$

■

In contrast to other p-norms, where $p \geq 2$, the 1-norm allows us to embed a reduced version of our lower bound construction from the T–GNCG. With increasing number of dimensions we can embed more and more of our construction. The following statement shows that for arbitrary large d the lower bound of the PoA approaches the upper bound of $\frac{\alpha+2}{2}$.

▶ **Theorem 4.31.** In a 1-norm d-dimensional space the PoA is at least

$$1 + \frac{\alpha}{2+\alpha/(2d-1)}.$$

◀

Proof. Consider a set of $n = 2d+1$ points $v_0 = (0,\dots,0), v_1 = (1,0,\dots,0)$, $v_2 = (-\frac{2}{\alpha}, 0\dots,0)$, $v_{i+1} = (\underbrace{0,\dots,0}_{i-1}, \frac{2}{\alpha}, 0,\dots,0), v_{i+d} = (\underbrace{0,\dots,0}_{i-1}, -\frac{2}{\alpha}, 0,\dots,0)$ for $i \in \{2,\dots,d\}$.

On the one hand it is easy to see that the star network S_n^* with a center in v_0 is an optimal network. On the other hand, the star network S_n with its center in v_1 such that v_1 is the owner of all edges in the star, is in NE. Indeed, since the distances are measured via the 1-norm, the construction is exactly the same as in the proof of Theorem 4.25. See Figure 4.10 for the construction in $\mathbf{R}^3$. Thus,

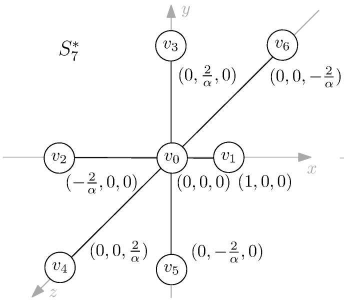


Figure 4.10: Construction of the optimal network S_7^* and the stable network S_7 in 3-dimensional 1-norm space.

the PoA is bounded by:

$$\frac{cost(S_n)}{cost(S_n^*)} = \frac{(2 \cdot (2d+1) + \alpha - 2)\left(\frac{\alpha+2}{\alpha} \cdot (2d-1) + 1\right)}{(2 \cdot (2d+1) + \alpha - 2)\left(\frac{2}{\alpha} \cdot (2d-1) + 1\right)} = 1 + \frac{\alpha}{2 + \dfrac{\alpha}{2d-1}}.$$

■

4.6 General Weighted Host networks

In this section we consider a more general model where a complete host network $H = (V, E)$ with arbitrary edge weights $w\colon V \times V \to \mathbb{R}_+$ is given.

4.6.1 Existence

It is a challenging question to show an existence NE in the general non-metric setting. Also the approximation techniques for the M–GNCG cannot be easily adopted. However, Proposition 4.6 can be directly generalized if we consider a spanning sub-network H' of the host network H such that each edge in H' participates in at least one shortest path in H.

▶ **Proposition 4.32.** Let $H = (V, E)$ be a host network in the GNCG. A spanning sub-network $H' = (V, E')$ of H, where $E' = \{uv \mid w(u,v) = d_H(u,v)\}$, is in $(\alpha+1)$-NE. ◀

4.6.2 Price of Anarchy

▶ **Theorem 4.33.** For any host network for which a NE exists, the PoA in the GNCG is at most $\alpha + 1$. ◀

Proof. Consider a host network $H = (V, E(H))$, a stable network $G = (V, E)$ and an optimum network $G^* = (V, E^*)$. To simplify a notation in the proof, for any two nodes $x, y \in V$, we denote the shortest path in the network G as $\pi_G(x, y)$.

We partition the set of edges E in two sets. Let $E_1 := \bigcup_{uv \in E^*} \{\pi_G(u, v)\}$ be a set of edges in G appearing in the shortest u-v path $\pi_G(x, y)$ in G for each $uv \in E^*$, i.e., for every edge uv in the social optimum, E_1 contains all edges from the shortest path between u and v in G. We denote the rest of the edges in G as $E_2 := E \setminus E_1$.

First we analyze the cost of all edges in E_1. Since G is a $(\alpha + 1)$-spanner, we have:

$$\alpha \cdot w(E_1) = \alpha \cdot w\left(\bigcup_{uv \in E^*} \pi_G(u, v)\right) \leq \alpha \sum_{uv \in E^*} d_G(u, v)$$
$$\leq \alpha(\alpha + 1) \sum_{uv \in E^*} d_H(u, v) \leq \alpha(\alpha + 1) w(E^*). \quad (4.5)$$

Consider all agents who own at least one edge in E_2. We denote this set as $V_2 := \{v : \exists uv \in E_2 \text{and } uv \text{ is owned by } u\} \subseteq V$. Next, we compute the costs for all agents in V_2. Consider an agent $u \in V_2$. Since G is a NE, deleting all edges to the nodes in $V_2 \cap S_u$ is not an improving move for u. Hence,

$$\alpha w(u, V_2 \cap S_u) + d_G(u, V) \leq d_{G - E_2}(u, V),$$

where $G - E_2$ is the network obtained after the deletion. Note that if we sum up all inequalities for all $u \in V$ that own at least one edge in E_2, we get an upper bound on the total cost of all such agents:

$$\alpha w(E_2) + \sum_{u \in V_2} d_G(u, V) \leq \sum_{u \in V_2} d_{G - E_2}(u, V).$$

Next we will evaluate the left part of the inequality, i.e., the agents' distance cost after deleting all edges from E_2 set. For this we will show that the obtained network $G - E_2$ is an $(\alpha + 1)$ spanner.

Consider a node $v \neq u$ from V. Note that d_{G-E_2} is equal to the distance between u and v in the network G restricted on the edge set E_1. Consider a shortest path $\pi_{G^*}(u,v)$ in the optimum network G^*. By definition, for each edge $xy \in \pi_{G^*}(u,v)$, set E_1 contains a shortest path $\pi_G(x,y)$ of the length

$$d_{G-E_2}(x,y) \leq (\alpha+1)d_H(x,y) \leq (\alpha+1)d_{G^*}(x,y).$$

Thus,

$$\begin{aligned} d_{G-E_2}(u,v) &\leq \sum_{xy\in\pi_{G^*}(u,v)} d_{G-E_2}(x,y) \\ &\leq (\alpha+1)\cdot \sum_{xy\in\pi_{G^*}(u,v)} d_{G^*}(x,y) = (\alpha+1)d_{G^*}(u,v). \end{aligned}$$

The total cost of the agents in the set V_2 is

$$\alpha w(E_2) + \sum_{u\in V_2} d_G(u,V) \leq \sum_{u\in V_2} d_{G-E_2}(u,V) \leq (\alpha+1)\sum_{u\in V_2} d_{G^*}(u,V). \tag{4.6}$$

Note that the total distance cost in the equilibrium network G is $\sum_{u\in V} d_G(u,V) = \sum_{u\in V\setminus V_2} d_G(u,V) + \sum_{u\in V_2} d_G(u,V)$. Since G is a $(\alpha+1)$-spanner (by Lemma 4.2), we evaluate the second term as follows:

$$\sum_{u\in V\setminus V_2} d_G(u,V) \leq (\alpha+1)\sum_{u\in V\setminus V_2} d_H(u,V) \leq (\alpha+1)\sum_{u\in V\setminus V_2} d_{G^*}(u,V) \tag{4.7}$$

A combination of inequalities 4.5, 4.6, and 4.7 let us evaluate the PoA as follows:

$$\begin{aligned} \frac{SC(G)}{SC(G^*)} &= \frac{\alpha w(E_1) + \sum_{u\in V\setminus V_2} d_G(u,V) + \alpha w(E_2) + \sum_{u\in V_2} d_G(u,V)}{\alpha w(E^*) + \sum_{u\in V} d_{G^*}(u,V)} \\ &\leq \frac{\alpha(\alpha+1)w(E^*) + (\alpha+1)\sum_{u\in V\setminus V_2} d_{G^*}(u,V) + (\alpha+1)\sum_{u\in V_2} d_{G^*}(u,V)}{\alpha w(E^*) + \sum_{u\in V} d_{G^*}(u,V)} \\ &= \frac{(\alpha+1)(\alpha w(E^*) + \sum_{u\in V} d_{G^*}(u,V))}{\alpha w(E^*) + \sum_{u\in V} d_{G^*}(u,V)} = \alpha+1. \end{aligned}$$

■

Even though the above upper bound for the PoA asymptotically meets the lower bound (Theorem 4.25), it would be highly interesting to close the gap. Several attempts to find an instance of equilibrium such that its cost is strictly higher than $\frac{\alpha+2}{2}$ have failed. Therefore, we conjecture that the PoA for GNCG should be the same as the PoA for M–GNCG.

► **Conjecture 4.34.** For any host network for which a NE exists, the PoA for the GNCG is $\frac{\alpha+2}{2}$. ◄

4.7 Conclusion

In this section we have analyzed the Network Creation Game on weighted complete host networks. We think this is a significant step towards a more realistic game-theoretic model for the decentralized creation of networks, like fiber-optic or overlay networks. We showed that the weighted version of these games behaves similarly to the unit-weight NCG in terms of the hardness of computing a best response and in its dynamic properties. However, the Price of Anarchy is radically different. Whereas in the original NCG the PoA is conjectured to be constant and actually proven to be constant for almost all α, we have shown that the PoA, even for the restricted metric case of the T–GNCG, is linear in α. Since α is a parameter for adjusting the trade-off between edge cost and distance cost, this implies that for settings where the edge cost dominates, i.e., if α is high, coordination is needed to guarantee socially efficient outcomes.

For understanding the impact of coordination, the next step should be to analyze the Price of Stability, i.e., the social cost ratio of the best equilibrium network and the social optimum. Another challenging task is to prove or refute that pure Nash equilibria always exist and to find a way to guide the agents to stable states with preferably low social cost. Besides this, naturally Conjecture 4.34, calls for further investigation.

5 Selfish Network Creation with Degree Dependent Edge Cost

The original Network Creation Game as well as many of its follow up versions have the drawback that edges are treated uniformly, i.e., every edge has the same cost. This common parameter heavily influences the outcomes and the analysis of these games. We take a radical departure from this assumption by proposing and analyzing a variant of the Network Creation Game in which the agents follow an anti-preferential attachment rule. In particular, the cost of an edge between agent u and v which is bought by agent u is proportional to v's degree in the network, i.e., edge costs are proportional to the degree of the other endpoint involved in the edge. Thus, we introduce individual prices for edges and at the same time we obtain a simple model which is parameter-free.

Our model is inspired by social networks in which the nodes usually have very different levels of popularity which is proportional to their degree. In such networks connecting to a celebrity usually is expensive. Hence, we assume that establishing a link to a popular high degree node has higher cost than connecting to an unimportant low degree node. Moreover, in social networks links are formed mostly locally, e.g., between agents with a common neighbor, and it rarely happens that links are removed, on the contrary, such networks tend to get denser over time [LKF05]. This motivates two other extensions of our model which consider locality and edge additions only.

5.1 Model Definition

We consider unweighted undirected networks $G = (V, E)$, where V is the set of nodes and E is the set of edges of G. Since edges are unweighted, the distance $d_G(u, v)$ between two nodes u, v in G is the number of edges on a shortest path between u and v. For a given node u in a network G let $N_k(u)$ be the set of nodes which are at distance at most k from node u in G and let $B_k(u)$ be the set of nodes which are at exactly distance k from node u (the distance-k ball around u).

We investigate a natural variant of the Network Creation Game (NCG) which we call the *Degree Price Network Creation Game (degNCG)*. The definition of the game is similar to the NCG. The only crucial difference is the topology-dependent edge cost. More precisely, we assume that the cost of an edge is proportional to the degree of the endpoint, which is not the edge owner.

That is, if agent u buys the edge uv then u's cost for this edge is proportional to node v's degree. For simplicity we will mostly consider the case where the price of edge uv for agent u is exactly v's degree without counting edge uv, i.e., it corresponds to the cost of the connection before the edge is added. Thus the cost of agent u in network $G(\mathbf{s})$ is

$$cost(u, G(\mathbf{s})) = \sum_{v \in S_u} (\deg_{G(\mathbf{s})}(v) - 1) + d_{G(\mathbf{s})}(u) \quad .$$

Note that in contrast to the NCG our variant of the model does not depend on any parameter and the rather unrealistic assumption that all edges have the same price is replaced with the assumption that buying edges to well-connected nodes is more expensive than connecting to low degree nodes. Another significant difference is that the cost of edges is dynamic, i.e., any network change can influence the cost of many edges. See Figure 5.1 for an example.

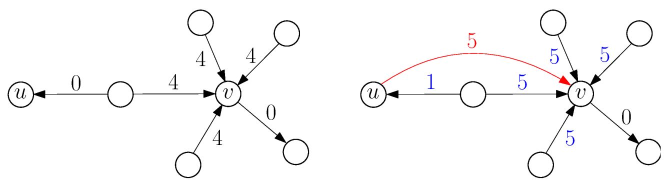


Figure 5.1: Edge cost before and after the addition of the edge uv.

As a solution concept we consider the pure Nash Equilibrium (NE). Observe that in the degNCG we assume that agents can buy edges to every node in the network. Especially in modeling large social networks, this assumption seems unrealistic. To address this, we also consider a restricted version of the model which includes locality, i.e., where only edges to nodes in distance at most k, for some fixed $k \geq 2$, may be bought. We call this version the *k-local degNCG* (denoted as deg$k$NCG) and its pure Nash equilibria are called k-local NE (kNE). We measure the quality of a network $G(\mathbf{s})$ by its *social cost*, which is simply the sum over all agents' costs, i.e., $cost(G(\mathbf{s})) = \sum_{u \in V} cost_u(G(\mathbf{s}))$.

The use case of modeling social networks indicates another interesting version of the degNCG, which we call the *degree price add-only game (degAOG)* and the k-local version deg$k$AOG. In these games, agents can only add edges to the network, whereas removing edges is impossible. This mirrors social networks where an edge means that both agents know each other. Hence, losing a connection would mean forgetting the prior known person, which is not realistic.

To simplify the notation, we say that a result holds for the deg(k)NCG (resp. deg(k)AOG) if it holds for the global and the k-local version of the *degree price game* (resp. *degree price add-only game*).

5.2 Related Work

A distinctive feature of the model discussed in this chapter is its topology dependent edge cost, i.e., the model is parameter-free. Removing the parameter α by restricting the agents to edge swaps was studied in [Alo+13; MS12]. The obtained results are similar to the results for the original Network Creation Game: the best known upper bound on the PoA is $2^{O(\sqrt{\log n})}$, there cannot exist a potential function [Len11], and computing a best response is NP-hard. However, allowing only swaps leads to the unnatural effects that the number of edges cannot change and that the sequential version heavily depends on the initial network.

To the best of our knowledge, there are very few related papers that analyze the NCG with non-uniform edge price. In [CMH14] agents can buy edges of different quality, which corresponds to their length, and the edge price depends on the edge quality. Distances are measured in the induced weighted network. Closer to our model is [MMO14] where heterogeneous agents, important and unimportant ones, are considered, and both classes of agents have different edge costs. Here, links are formed with bilateral agreement [CP05; JW96] and important nodes have a higher weight, which increases their attractiveness.

In our model we assume that the edge cost is proportional to the degree of its non-owning endpoint, like in social networks, connections to celebrities are less likely than links to less famous persons. The idea of having nodes with different popularity is discussed in the so-called celebrity games [Àlv+16; ÀM16]. There, nodes have a given popularity, and agents buy fixed-price edges to get highly popular nodes within some given distance bound. Hence, this model differs heavily from our model.

The idea to use node popularity in the link-formation mechanism to describe a network formation process is actively studied outside of the game theory community. The Preferential Attachment model [BA99] is known for claims that the resulting networks generated by the mechanism exhibit real-world network properties. Here, each new arriving node creates an edge to an existing node with probability proportional to the node's degree. There is a vast literature on variations of the preferential attachment model. We refer to the book by Van der Hofstand [Van16] for an overview. Note that we use the reverse rule in our model, i.e., we assume that agents avoid linking with high-degree nodes. This approach has two arguments of support. First, as we discussed earlier, this assumption is natural for social networks. Second, a recent paper [DPS20] investigated that a network growing with a mix of preferential and anti-preferential attachment where the anti-preferential mechanism dominates results in a network with real-world properties.

Among other things, in this chapter, we actively study the influence of locality on the structure of equilibria by comparing the global and the local versions of the game. There are several versions for augmenting the NCG with locality that have been proposed and analyzed recently. It was shown that the PoA may deteriorate heavily if agents only know their local neighborhood or only a shortest path tree of the network [Bil+14a; Bil+14b], and even if strategy changes are restricted to edge swaps [YY20]. In contrast, a global view with a restriction to only local edge-purchases yields only a moderate increase of the PoA [CL15]. Therefore, we focus on the last, more ”optimistic", case where agents can observe the entire network while performing local moves.

5.3 Results Overview

We introduce and analyze the first parameter-free variants of Network Creation Games [Fab+03] which incorporate non-uniform edge cost. In almost all known versions the outcomes of the games and their analysis heavily depend on the edge cost parameter α. We depart from this by assuming that the cost of an edge solely depends on structural properties of the network, in particular, on the degree of the endpoint to which the edge is bought. Essentially, our models incorporate that the cost of an edge is proportional to the popularity of the node to which it connects. This appears to be a realistic feature, e.g., for modeling social networks.

On the first glance, introducing non-uniform edge cost seems to be detrimental

Model	Complexity	PoA	Diameter	Dynamics
degNCG	BR NP-hard (Thm.5.1)	$\mathcal{O}(1)$ (Thm.5.14)	≤ 3 (Cor.5.6)	no FIP (Thm.5.18)
deg$k$NCG	BR NP-hard (Thm.5.3)	$k = 2$: $\mathcal{O}(\sqrt{n})$ (Cor.5.16)	$k = 2$: $\mathcal{O}(\sqrt{n})$ (Thm.5.8)	no FIP (Thm.5.18)
		$k \geq 3$: $\mathcal{O}(1)$ (Thm.5.14)	$k = 3$: ≤ 5 (Thm.5.9)	
			$k \geq 4$: ≤ 3 (Cor.5.7)	
degAOG	BR NP-hard (Thm.5.1)	$\Theta(n)$ (Thm.5.17)	≤ 3 (Cor.5.6)	FIP (Sec.5.6.2)
deg$k$AOG	BR NP-hard (Thm.5.3)	$\Theta(n)$ (Thm.5.17)	$k = 2$: $\mathcal{O}(\sqrt{n})$ (Thm.5.8)	FIP (Sec.5.6.2)
			$k = 3$: ≤ 5 (Thm.5.9)	
			$k \geq 4$: ≤ 3 (Cor.5.7)	

Table 5.1: Overview of the properties of the deg(k)NCG and the deg(k)AOG.

to the analysis of the model. However, in contrast to this, we give a simple proof that the PoA of the degNCG is actually constant. Besides this strongest possible bound on the PoA, which we also generalize to arbitrary linear functions of a node's degree and to the 3-local version, we prove a PoA upper bound of $O(\sqrt{n})$ for the deg2NCG, where agents are restricted to act within their 2-neighborhood, and we show for that computing a best response strategy is NP-hard for deg$k$NCG. Moreover, we investigate the dynamic properties of the deg(k)NCG and prove that improving response dynamics may not converge to an equilibrium, that is, there cannot exist a generalized ordinal potential function.

We contrast these negative convergence results by analyzing a version where agents can only add edges, i.e., the deg(2)AOG, where convergence of the sequential version is trivially guaranteed, and by analyzing the speed of convergence for different agent activation schemes. The restriction to only edge additions has severe impact on the PoA, yielding a $\Theta(n)$ bound. However, we show that the diameter of equilibrium networks is constant for any $k \geq 3$. Hence, unlike in the degNCG model, the non-constant PoA is explained not by the upper bound for the network diameter but by the high edge cost of the worst-case equilibrium. We also show that the impact of the add-only restriction on the social cost is low, if round-robin dynamics starting from a path are considered, where agents buy their best possible single edge in each step.

See Table 5.1 for a comparison of the results for the different versions of the model.

5.4 Hardness

In this section we investigate the computational hardness of computing a cost minimizing strategy, i.e., a best response, in the deg(k)NCG and deg(k)AOG.

First, we provide a hardness reduction for the general version of the model. Note that the construction used in the reduction has diameter 4. It implies that the reduction works for the local version of the model where agents can add edges to nodes within a distance four or more.

▶ **Theorem 5.1.** The problem of computing a best response is NP-hard in the deg(k)NCG and the deg(k)AOG for $k \geq 4$. ◀

Proof. We provide a polynomial time reduction from the Exact Cover by 3-Sets problem which is known to be NP-complete [Pap03]. The problem is defined as follows: given a collection $\mathcal{A} = \{A_1, \ldots, A_\ell\}$ of subsets of a universal set $U = \{1, \ldots, n\}$, such that $n = 3m$ for some $m \in \mathbb{N}$ and $|A_i| = 3$ for all i. The problem is to determine whether there are m sets in A that are pairwise disjoint and give U in their union. We consider the corresponding instance of the best response problem in the degNCG. Consider a graph $G = (V, E)$ such that

$$V = \{u\} \cup \bigcup_{i=1}^{9}\{x_i\} \cup \bigcup_{i=1}^{n}\{v_i^1, v_i^2, v_i^3\} \cup \bigcup_{i=1}^{n}\bigcup_{r=1}^{9}\{p_{ir}^1, p_{ir}^2, p_{ir}^3\} \cup \bigcup_{i=1}^{\ell}\{a_i\}$$

where each a_i corresponds to a set A_i and each set $\{v_i^1, v_i^2, v_i^3\}$ corresponds to one element $i \in U$.

$$\begin{aligned} E =& \{ux_i \mid 1 \leq i \leq 9\} \cup \left\{x_k p_{ir}^j \mid i \in U, j = 1,2,3 \text{ and } 1 \leq k, r \leq 9\right\} \\ & \cup \left\{v_i^j a_j \mid i \in A_j, j = 1,2,3\right\} \cup \left\{v_i^j p_{ir}^j \mid j = 1,2,3,\ i \in U \text{ and } 1 \leq r \leq 9\right\}. \end{aligned}$$

We assume that each edge ux_i, $1 \leq i \leq 9$, is owned by the corresponding agent x_i. See Figure 5.2 for illustration of the construction. Note that in this instance all set nodes a_i have a degree 9, all element nodes v_j^r have a degree at least 10 and the extra nodes p_{jk}^r are of a degree 10.

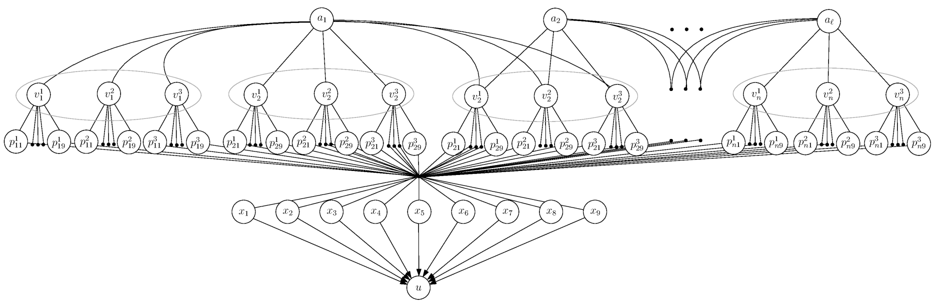


Figure 5.2: Illustration of the construction used in the reduction.

Now we prove that the best response S_u^* of the agent u contains only the nodes from $\{a_1, \ldots, a_\ell\}$ corresponding to the solution of the Exact cover by 3-sets. Assume to the contrary, that there is a node $v_i^r \in S_u^*$. In case there is a set-node a_j such that $i \in A_j$, the agent u can improve her strategy S_u^* by removing v_i^r

since it improves her edge cost by $9 + k$ (where k is the number of set nodes connected with v_i^r; note that $k \geq 1$). Also, this move increases the distance by at most k. In case there is no set node in S_u^* such that the corresponding set contains the element i, then the agent u can improve the strategy by swapping v_i^r with some set node a_j such that $i \in A_j$. Indeed, this move changes the edge cost by $9 - (9 + k)$ and the distance cost increases by 1 to the node v_i^r and $k - 1$ set nodes adjacent to it, i.e., the cost improves by $-k + k - 1 = -1$. Thus, there are no element nodes in the best response of agent u.

Now, assume that there is a node $p_{ij}^r \in S_u^*$. In case none of the set nodes which "covers" the element j is in S_u^*, there is a better strategy $S_u' = S_u^* \setminus \{p_{ij}^r\} \cup \{a_k\}$ where a_k corresponds to the set A_k containing the element $i \in U$. Indeed, in the worst case, such a strategy move changes the distance to p_{ij}^r and a_k only. Thus, the distance cost decreases by at least 1, whereas the edge cost decreases by 1 since $deg(p_{ij}^r) = 10 = deg(a_k) + 1$ in G. In case there is a set node $a_k \in S_u^*$ which "covers" the element i, node p_{ij}^r can be removed from the u's strategy S_u^*. Indeed, it improves the edge cost by 9 and increases the distance only to node p_{ij}^r, i.e., it improves agent u's cost by 8. Therefore, the best response S_u^* contains only set nodes. Moreover, if there are two set nodes $a_i, a_j \in S_u^*$ which cover the same element $k \in U$, then one of two set nodes, say a_j, can be removed from S_u^*. Note that this move can increase distances to sets of element nodes corresponding to the elements from $A_j \setminus A_i$ and the set nodes adjacent to them. In case such a set node a_k exists, i.e., there is a set node $a_k, k \neq j$ to which the distance from u increases after removing the edge ua_j, then there is no other set node $a_{k'} \in S_u^*$ such that $A_k \cap A_{k'} \neq \emptyset$. Thus, u can improve on the strategy S_u^* by replacing a_j with a_k. It does not change the edge cost of u and improves the distance cost by at least 3, i.e., it is an improving strategy change. In case a_j is the only set node to which the distance increases after removing the edge ua_j, there is the better strategy $S_u' = S_u^* \setminus \{a_j\}$ which gives an improvement of the edge cost by 9 and increases the distance cost to only two sets of element nodes and a_j by at most 8 in total.

This implies that S_u^* corresponds to the solution of the Exact Cover by 3-Sets problem. Let m' be the number of nodes in S_u^*. Then the cost of u is $9 \cdot m' + m' + 3(\ell - m') + 2 \cdot 3m' + 3 \cdot 3(n - m') + 2 \cdot 27n + 9 = 4m' + 3\ell + 63n + 9$. Since ℓ and n are fixed parameters, the cost is minimized if m' is minimized. Thus, if there is any better strategy for agent u, this implies an existence of a better solution for the Exact Cover by 3-Sets problem. ■

To complete this section, we need to prove the NP-hardness of computing a best response in the deg$k$AOG and the deg$k$NCG for k equals two and three. For this case we consider a problem similar to the Exact Cover by 3-Sets which is called Exact-q-Set Cover and prove that it is NP-hard via a reduction from Dominating Set in q-regular graphs. The Exact-q-Set Cover problem is defined as follows: Given a universal set $U = \{1, 2, \ldots, n\}$ and a collection of sets $\mathcal{A} = \{A_1, \cdots A_\ell\}$ where $\forall A_i \in \mathcal{A}, A_i \subseteq U$ and $|A_i| = q$ with $q \geq 4$; the task is to find the minimum subset of $\mathcal{A}$ which covers U.

▶ **Lemma 5.2.** Exact-q-Set Cover is NP-hard. ◀

Proof. We give a polynomial time reduction from Dominating Set in q-regular graphs, which is known to be NP-hard [AK00], to Exact-$(q+1)$-Set Cover. Given a q-regular graph $G = (V, E)$ with $V = \{1, \cdots, n\}$, construct an instance of Exact-$(q+1)$-Set Cover as follows: the universe is $U = V$ and the collection of sets is $\mathcal{A} = \{A_1, \cdots A_n\}$ such that A_i consists of node i and the nodes adjacent to node i. Since G is q-regular, we have that $|A_i| = q+1$. Now S is an optimal solution for Exact-$(q+1)$-Set Cover for the instance $(U, \mathcal{A})$ if and only if $R = \{i \mid A_i \in S\}$ is a minimum dominating set for the graph G. ■

▶ **Theorem 5.3.** Computing a best response in the deg$k$NCG and deg$k$AOG is NP-hard for $k \in \{2, 3\}$. ◀

Proof. We prove the theorem by giving a polynomial time reduction from the NP-hard Exact-q-Set Cover problem to the problem of computing a best response of an agent u in an instance of the deg$k$NCG. Consider an instance I of the Exact-q-Set Cover problem with universal set $U = \{1, 2, \ldots, n\}$ and a collection of sets $\mathcal{A} = \{A_1, \cdots A_\ell\}$, where $A_i \subseteq U$ for all $A_i \in \mathcal{A}$ and $|A_i| = q$. We create a corresponding instance $G = (V, E)$ of the best response problem in the deg$k$NCG, where the network $G = (V, E)$ is defined as follows:

$$V = \bigcup_{i=1}^{n} \{v_i\} \cup \bigcup_{i=1}^{n} \left\{p_i^1, \ldots, p_i^{q+1}\right\} \cup \bigcup_{i=1}^{\ell} \{a_i\} \cup \bigcup_{i=1}^{k-1} \{x_i\} \cup \{u\},$$

where each v_i corresponds to $i \in U$ and each a_i corresponds to $A_i \in \mathcal{A}$. The edge set E is defined as follows:

$$E = \left\{v_i a_j \mid i \in A_j\right\} \cup \left\{v_i p_i^r \mid i \in U \text{ and } 1 \leq r \leq q+1\right\}$$

$$\cup \left\{a_j x_{k-1} \mid 1 \le j \le \ell\right\} \cup \{x_i x_{i+1} \mid 1 \le i \le k-2\} \cup \{u x_1\} \; .$$

Moreover, we assume that the edge ux_1 is owned by agent x_1. See Figure 5.3.

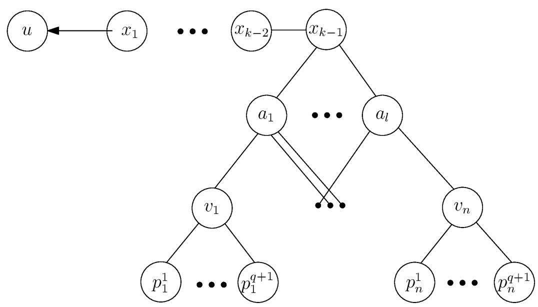


Figure 5.3: Illustration of the construction used in the reduction.

In case $k = 2$, agent u can build edges only to the set-nodes. In case $k = 3$, agent u can create an edge to x_2 or to the set-nodes. Trivially, the best move for the agent is to buy edges to the set-nodes rather than to the node x_2. Indeed, one edge ux_2 costs $\ell+1$ and improves the distance to $1+\ell+n+n(q+1)$ nodes by one, while a set of edges to a set cover of size $m \le \ell$ costs $(q+1)m$ and improves the distance to $m+n+n(q+1)$ nodes by 2. Note that we assume $m \le \ell \le n$, otherwise, if $\ell > n$, we can slightly modify the network by creating $\lceil \frac{\ell}{n} \rceil$ copies of each v_i node and all its incident p_i^r nodes. Hence, the improvement obtained by the agent u after adding edges to the set cover nodes is $2n(q+2)+m-qm \ge 2n(q+2)+n(1-q) = n(q+2)+3n$ that is strictly larger than the total improvement $n(q+1)$ after adding ux_2 edge.

Note that the degree of every node in $B_k(u)$ is exactly $q+1$ and in $B_{k+1}(u)$ at least $q+2$. Now we observe that if agent u is playing her best response this means agent u has to create a minimum number of edges to nodes in $B_k(u)$ such that every node v_i is at distance 2 from u. This is true since otherwise there exists a node $a_j \in B_k(u)$ to which buying an edge costs $q+1$ and the resulting improvement of the distance cost is at least $q+3$ which contradicts the fact that u is playing her best response. Moreover, since u does not own any edges, no deletions or edge swaps are possible.

Let $S_u^* \subseteq N_k(u)$ denote agent u's best response. Since S_u^* has the minimum size such that all v_i nodes are at distance 2 to agent u, the union of the sets $A_j \in \mathcal{A}$ corresponding to the chosen nodes $a_j \in S_u^*$ forms the minimum size solution for the Exact-q-Set Cover instance I.

Since in the above reduction the best response of agent u consists of only buying edges, this implies NP-hardness for the deg$k$AOG. ■

5.5 Analysis of Equilibria

We start with the most fundamental statement about equilibria which is their existence. We use the center sponsored spanning star S_n for the proof.

▶ **Theorem 5.4.** The center sponsored star S_n is a (k)NE for the deg(k)NCG and the deg(k)AOG for any k. ◀

Proof. In the center sponsored spanning star S_n the center agent cannot delete or swap any edge since this would disconnect the network. Since the center already has bought the maximum number of edges, no edge purchases are possible. Moreover, no leaf agent can profit from buying any number of edges because only edges to other leaves can be bought, which is a 2-local move. Such edges have cost of 1 which equals the maximum possible distance improvement. Thus, no agent has an improving move for any k which implies that S_n is in (k)NE. ■

5.5.1 Bounding the Diameter of Equilibrium Networks

In this section we investigate the diameter of (2)NE networks. As in most NCGs, bounding the diameter plays an important role in bounding the PoA.

▶ **Theorem 5.5.** Consider a generalization of the degAOG and the degNCG where the price of an edge uv bought by agent u is a linear function of v's degree in G, i.e., the edge price is $\beta \cdot\ deg_{G(s)}(v) + \gamma$, where $\beta, \gamma \in \mathbb{R}$. Then the diameter of any NE network is constant. ◀

Proof. We consider a NE network $G = (V, E)$ and assume that the diameter D of G is at least 4. Then there exist two nodes $a, b \in V$, such that $d_G(a, b) = D$. Therefore, the distance cost of an agent a in G is at least $D+|B_1(b)|(D-1)+|N_2(a)|$. Thus, if agent a buys the edge ab, then it improves a's distance cost by at least $D-1+|B_1(b)|(D-3)$. Since the network G is in NE, the distance cost improvement must be at most the agent u's cost for buying the edge ab:

$$D - 1 + |B_1(b)|(D - 3) \leq \beta \cdot deg_G(b) + \gamma$$
$$\iff D - 1 + (D - 3) \cdot deg_G(b) \leq \beta \cdot deg_G(b) + \gamma \ .$$

Solving for D under the assumption $deg_G(b) \geq 1$ yields

$$D \leq \frac{(\beta+3)deg_G(b)+\gamma+1}{deg_G(b)+1} < \beta+3+\frac{\gamma+1}{deg_G(b)+1} \in O(1). \tag{5.1}$$

■

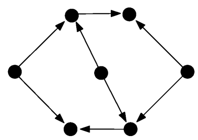


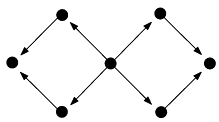
Figure 5.4: A NE in the degNCG with $D = 3$.

Note that in our version of the degNCG and the degAOG we use the edge cost function with parameters $\beta = 1$ and $\gamma = -1$. Hence, the inequality 5.1 in the proof of Theorem 5.5 and the NE example in Figure 5.4 yields the following:

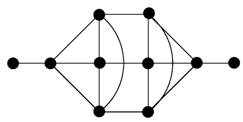
▶ **Corollary 5.6.** The diameter of any NE network in the degAOG and degNCG is at most 3 and this upper bound is tight. ◀

Since in the proof of Theorem 5.5 in the case of $\beta = 1$ and $\gamma = -1$ buying an edge to a node in distance 4 suffices, we get the following statement.

▶ **Corollary 5.7.** For $k \geq 4$, any kNE network has diameter at most 3. ◀



A 2NE in the deg2NCG with $D = 4$.



A 2NE in the deg2AOG with $D = 5$.

Figure 5.5: Examples of 2NE networks.

The examples in Figure 5.5 show that the diameter in the 2-local version, i.e., in the deg2NCG and the deg2AOG, can exceed 3. We prove a higher upper bound on the diameter for the 2-local versions.

▶ **Theorem 5.8.** The diameter of any 2NE network is in $O(\sqrt{n})$. ◀

Proof. Consider a 2NE network $G = (V, E)$ with $|V| = n$ and let $D \geq 2$ denote its diameter. Consider two nodes $a, b \in V$ such that $d_G(a, b) = D$ and a shortest-path tree $T_a = (V, E_a)$ which is rooted at node a (see Figure 5.6).

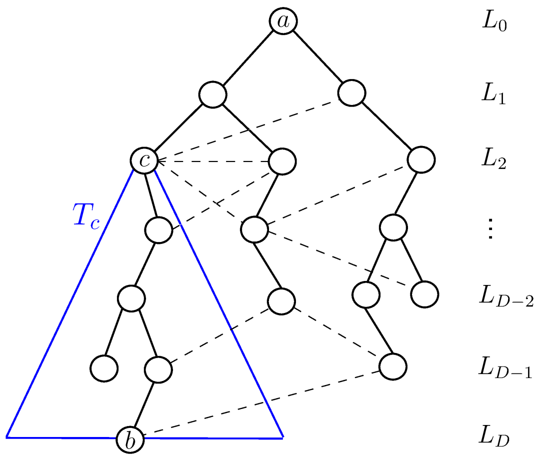


Figure 5.6: The shortest-path tree T_a. Dashed lines denote edges of G which are not in the tree, i.e., the non-tree edges.

The height of T_a is D and there must be a subtree T_c which contains node b and which has node c as root, where c is chosen such that $d_G(a, c) = 2$ and c belongs to the path from a to b in T_a. Since the height of T_c is $D - 2$, it follows that the number of nodes in T_c must be at least $D - 1$. Let $|T_x|$ denote the number of nodes in the subtree of T_a rooted at node x. Hence, we have $|T_c| \geq D - 1$.

Note, that if agent a buys any edge ax in network G then this improves a's distance cost by at least $|T_x|$. Since G is in 2NE, we know that buying the edge ac is not an improving move for agent a which implies that $|T_c|$ is at most the cost of the edge ac which is equal to $deg_G(c)$. Since $|T_c| \geq D - 1$ it follows that $deg_G(c) \geq D - 1$.

Let L_i denote the set of nodes which are in distance i from the root a in the tree T_a. For example $L_0 = \{a\}, c \in L_2$ and $b \in L_D$. Thus, we have $D - 1 \leq deg_G(c) \leq |L_1| + (|L_2| - 1) + |L_3|$.

Analogously, since G is in 2NE, we have that no agent v_i in layer L_i on the $c - b$ path in T_a can decrease her cost by buying an edge to a node in layer L_{i+2} which is a neighbor of a neighbor in T_a. With analogue reasoning as above we get $D - (i - 1) \leq deg_G(v_i) \leq |L_{i-1}| + (|L_i| - 1) + |L_{i+1}|$.

Note that not only agents from lower layers cannot improve by buying edges towards nodes in upper layers but also agents from upper layers cannot improve by buying edges towards nodes in lower layers. Thus we have

$$D - (i-1) \leq deg_G(v_i) \leq |L_{i-1}| + (|L_i| - 1) + |L_{i+1}|$$

and

$$D - (i-1) \leq deg_G(v_{D-i}) \leq |L_{D-i-1}| + (|L_{D-i}| - 1) + |L_{D-i+1}|$$

for any $2 \leq i \leq \lfloor \frac{D}{2} \rfloor - 1$. Summing up all inequalities yields:

$$2 \sum_{i=2}^{\lfloor \frac{D}{2} \rfloor - 1} \big(D - (i-1)\big) \leq 3\left(\sum_{i=1}^{D} |L_i| - (D-1)\right).$$

For the left side we have

$$\frac{3D^2}{4} - 4D - 3 < \left(\left\lfloor \frac{D}{2} \right\rfloor - 2\right)\left(2D + 1 - \left\lfloor \frac{D}{2} \right\rfloor\right) = 2 \sum_{i=2}^{\lfloor \frac{D}{2} \rfloor - 1} \big(D - (i-1)\big)$$

and the right side gives $3\left(\sum_{i=1}^{D} |L_i| - (D-1)\right) \leq 3n - 3D + 3$, which yields

$$\frac{3D^2}{4} - 4D - 3 < 3n - 3D + 3 \Rightarrow D < \frac{2}{3}\left(1 + \sqrt{9n+19}\right) \in O(\sqrt{n}).$$

■

Clearly, the statement above holds for any kNE network, but we can prove a stronger statement that the diameter of 3NE networks is constant.

▶ **Theorem 5.9.** The diameter of any 3NE network is at most 5. ◀

Proof. Consider a 3NE network $G = (V, E)$. Assume to the contrary, that G has diameter at least 6. Then there are two nodes a and b at distance 6. Consider a node v at distance 3 from a on the a-b shortest path. For $i = 2, 3, 4$, let V_i be a set of nodes incident to v and at distance i from a. Since $d_G(a, b) = 6$, the distance between b and any node $x \in V_i$ is at least $6 - i$. Now we consider two possible moves, that is the addition of av edge by agent a and the addition of bv edge by b. The difference of the cost obtained by a after the move is at most $deg_G(v) - |V_3| - 2|V_4| - 4$ that is at least 0 since G is in 3NE. From the other side, the cost difference for the agent b after the addition of the edge bv is at most

$deg_G(v) - |V_3| - 2|V_2| - 4$ that is again at least 0. Summing up two inequalities gives us:

$$8 + 2|V_2| + 2|V_3| + 2|V_4| \leq 2deg_G(v). \tag{5.2}$$

Since $|V_2| + |V_3| + |V_4| = deg_G(v)$, we get a contradiction. Therefore, the diameter of the 3NE networks is at most 5. ■

Our extensive experiments and attempts on constructing a high diameter 2NE indicate that the diameter of all 2-local equilibria is at most 5. It leads us to the conjecture that all Nash equilibria have a constant diameter. This argument is also supported by the fact that a high diameter yields high degrees of all nodes on the diameter shortest path. Therefore, there is always an agent who can improve her distance to many nodes "incident" to the shortest path. However, a high degree means a high price of connection that makes the proof of the conjecture quite challenging.

▶ **Conjecture 5.10.** The diameter of any 2NE network is constant. ◀

5.5.2 Price of Stability

For analyzing the Price of Stability, we have to investigate the network which has the minimum possible social cost.

▶ **Lemma 5.11.** The center sponsored spanning star S_n is a social optimum in the deg(k)NCG and the deg(k)AOG for any $k \geq 2$. ◀

Proof. Consider an optimal network $G = (V, E)$ with m edges and n nodes. As G has to be connected, we have $m \geq n - 1$. Now, all the pairs which are not connected by an edge are at distance of at least 2, and there are $n(n-1) - 2m$ many such pairs. Adding the remaining $2m$ pairs with distance 1 yields the distance cost of $2(n(n-1) - 2m) + 2m = 2n(n-1) - 2m$ which is also the lower bound on the distance cost of any graph. Since S_n has diameter 2 and all edges cost zero because every leaf node has degree 1, the social cost of the center sponsored spanning star S_n meets the above lower bound. ■

We have shown in the proof of Theorem 5.4 that the center sponsored spanning star S_n is in (k)NE for any k. With Lemma 5.11 this yields the following PoS result.

▶ **Corollary 5.12.** The Price of Stability in the deg(k)NCG and the deg(k)AOG is 1. ◀

5.5.3 Price of Anarchy

For investigating the quality of the equilibria of our games, we first adapt an important lemma by Fabrikant et al. [Fab+03] to our setting.

▶ **Lemma 5.13.** If a (k)NE network G in the deg(k)NCG has diameter D, then its social cost is at most $O(D)$ times the minimum possible social cost. ◀

Proof. The minimum possible social cost is at least $n^2-1 \in \Omega(n^2)$, as the network is connected and every pair of nodes is at least in distance 1. To bound the social cost of the (k)NE in G, we bound the social distance cost and social edge cost separately. A trivial upper bound for the social distance cost is n^2D, since G has diameter D.

For bounding the social edge cost we first consider bridges of G, which are edges whose removal will disconnect G. There are at most $n-1$ bridges, so the total edge cost of all bridges is at most $\Delta(n-1) \in O(n^2)$ with $\Delta = \max_{v \in V} deg_G(v)$. Now we will argue that the cost of all non-bridges bought by any agent u is in $O(nD)$, which implies that the total edge cost of all edges is in $O(n^2D)$. This yields an upper bound on the social cost of $O(n^2D + n^2 + n^2D) = O(n^2D)$, completing the proof.

Consider an agent u and fix agent u's shortest path tree T_u, that is, we fix a shortest path from u to all other nodes in G. Let uv be any non-bridge edge bought by agent u. Let R_v be the set of nodes w, where the shortest path from u to w in T_u contains the node v.

We first argue that the distance between u and v is at most $2D$ if agent u would remove the edge uv. Note that removing uv cannot disconnect the network, since uv is not a bridge. Let xw be the edge on some shortest path from u to v in $G = (V, E \setminus uv)$ where $x \notin R_v$ and $w \in R_v$. As the diameter of G is D and since $x \notin R_v$ there must be a path of length at most D between u and x in $G = (V, E \setminus uv)$. Moreover, there exists a path of length at most $D-1$ between v and w in G. This is true since $d_G(u, w) \leq D$. Since x is a neighbor of w, it follows that the distance between every node $z \in R_v$ and x is at most D. Thus, removing the edge uv increases the diameter to at most $2D$ and agent u's total distance cost by at most $2D|R_v|$.

We know that G is in (k)NE in the deg(k)NCG. Hence, buying the edge uv must be profitable for agent u, that is, $\deg_G(v) - 1 \leq 2D|R_v|$. Let $S(u)$ be the set of nodes to which agent u bought a non-bridge edge. Summing up the inequalities for all nodes in $S(u)$ yields

$$\sum_{v\in S(u)} \left(\deg_G(v) - 1\right) \le 2D \sum_{v\in S(u)} |R_v| < 2nD,$$

where the last inequality holds since all sets R_v are disjoint. This implies that the total edge cost in any (k)NE network G is at most $2n^2D + \Delta(n-1)$, which concludes the proof. ■

From Corollary 5.6 and 5.7, and Theorem 5.9 we know that the diameter of any (k)NE is constant for all $k \ge 3$ in the deg(k)NCG. Also, from Lemma 5.13 we know that the social cost of any NE network G is at most $\mathcal{O}(D(G))$ times the minimum possible social cost. This implies the following statement.

▶ **Theorem 5.14.** The PoA is in $\mathcal{O}(1)$ in the deg(k)NCG for $k \ge 3$. ◀

A straightforward adaptation of Lemma 5.13 together with Theorem 5.5 yields:

▶ **Corollary 5.15.** The Price of Anarchy in the variants of the deg(k)NCG where the price of any edge uv bought by agent u is linear in v's degree in G, is constant. ◀

Theorem 5.8 and Lemma 5.13 yields the following statement.

▶ **Corollary 5.16.** The Price of Anarchy is in $\mathcal{O}(\sqrt{n})$ in the deg2NCG . ◀

We conclude this section with analyzing the PoA in the deg(k)AOG. The trivial upper bound for the PoA is in $\mathcal{O}(n)$, since the edge cost and the distance cost of an equilibrium are in $\mathcal{O}(n^2)$. By Lemma 5.11, the center sponsored star is a social optimum and has social cost in $\Omega(n)$. Therefore, the implication holds. Note that a clique provides a matching lower bound since it is in (k)NE in the deg(k)AOG for any k. We obtained the following result.

▶ **Theorem 5.17.** The PoA is in $\Theta(n)$ in the deg(k)AOG for any $k \ge 2$. ◀

5.6 Dynamics

In this section we consider the dynamic properties of the sequential version of the deg(k)NCG and the deg(k)AOG. Namely, we consider improving response dynamics (IRD), which starts with some initial strategy vector $\mathbf{s}$ and its corresponding initial network $G(\mathbf{s})$ and then agents are activated one at a time according to some activation scheme, e.g., a random, adversarially chosen move order, or round-robin activation. At each activation, an active agent is allowed to myopically update her current strategy. The respective agent will do so only if the new strategy yields strictly less cost than her current strategy. For the deg(2)AOG we will also consider the *best single edge dynamics*, which is a special case of the improving response dynamics, in which active agents buy the best possible single edge if this strictly decreases their current cost.

5.6.1 Dynamics in the deg(k)NCG

We investigate the convergence properties of the deg(k)NCG and prove that the deg(k)NCG may not converge under improving move dynamics.

▶ **Theorem 5.18.** For any $k \geq 2$, the deg(k)NCG does not admit the FIP, which implies that these games cannot have a generalized ordinal potential function. ◀

Proof. We prove the statement by providing an improving response cycle, which is a cyclic sequence of networks where neighboring networks differ only by the strategy of one agent and this strategy change is an improving response for the respective agent. See Figure 5.7.

The improving move cycle consists of six steps $G_1, \ldots, G_6$ and the transition from step G_i to $G_{i+1 \bmod 6}$ is an improving local move by some agent. Since all these improving moves are 2-local, this proves the statement for both the deg$k$NCG and the degNCG for any $k \geq 2$.

In network G_1 agent e has edge cost of 1 and distance cost of 23 which yields a total cost of 24. By buying the edge to node i and removing the edge to node h (e "swaps" her edge from h to i), agent e can decrease her cost by 1 since her edge cost in G_2 is 3 and her distance cost is 20.

In network G_2 agent b has edge cost of 6 and distance cost of 14 which yields a total cost of 20. By buying the edge to node h and removing the edge to node i, agents b's edges cost decreases to 4 and her distance cost increases to 15, which yields a total decrease of 1.

In network G_3 agent j has edge cost of 4 and distance cost of 19 giving a total cost 23. By performing the swap from i to h, agent j's edge cost does not change but her distance cost decreases by 1.

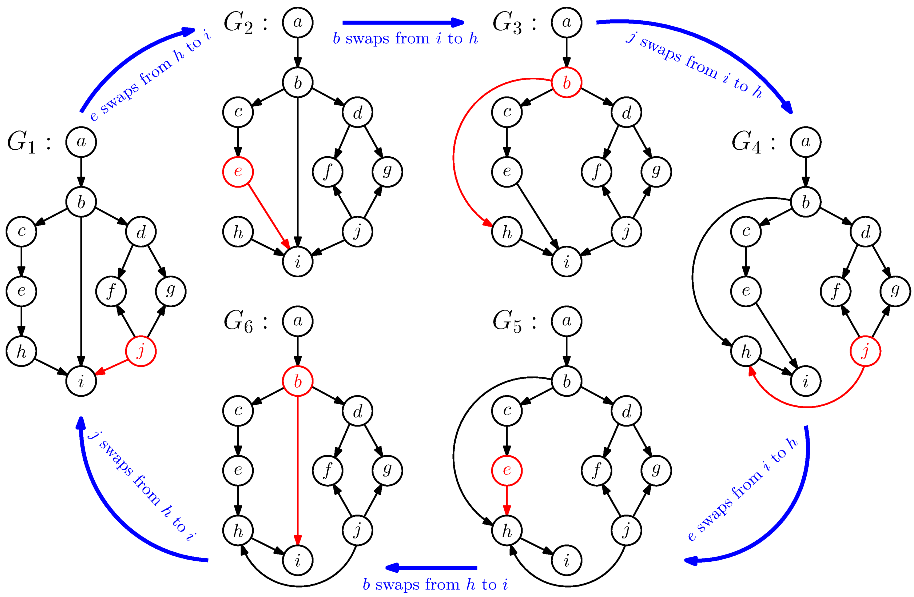


Figure 5.7: Example of an improving response cycle for the deg(k)NCG.

To complete the improving move cycle, note that, except for the edge ownership of edge $\{h, i\}$, network G_4 is isomorphic to G_1, network G_5 is isomorphic to G_2 and G_6 is isomorphic to G_3. Since neither agent h nor agent i are part of the improving move cycle, this implies that the described improving moves exist in these networks as well. That is, in network G_4 agent e swaps her edge from i back to h. In network G_5 agent b swaps her edge from h back to i and in network G_6 agent j swaps her edge from h back to i. ■

► **Remark 5.19.** The presented improving response cycle in Figure 5.7 is not a best response cycle for the deg(k)NCG since in network G_3 agent j has a strictly better local move: Buying the edge to agent h and swapping her edge from i to e. ◄

5.6.2 Dynamics in the Add-Only Model

We consider dynamics in the deg(k)AOG. First of all, since agents can only add edges, the deg(k)AOG trivially has the FIP, i.e., it is an ordinal potential game with the number of bought edges serving as a generalized ordinal potential function.

Since convergence is guaranteed, we focus on investigating the speed of convergence and the quality of the obtained networks. For the latter, Observation 5.17 yields a devastating result. However, we contrast this for the deg2AOG by proving that if round-robin best single edge dynamics starting on a path as initial network are used, then the social cost is actually close to the best possible achievable social cost.

▶ **Theorem 5.20.** Let $P_n = \{v_1 \cdots v_n\}$, a path of length n with v_1 and v_n as leaf nodes, be the initial network for the deg(k)AOG:

1. If in any step the active agent is chosen uniformly at random then IRD in the deg(k)AOG converge in $O(n^3)$ steps in expectation.
2. If in any step the active agent and her improving response is chosen adversarially (see Algorithm 2) then IRD in the deg(k)AOG converge in $\Theta(n^2)$ steps.
3. If round-robin best single edge dynamics are used in the deg2AOG, the process converges in at most $O(n \log n)$ steps to a network with diameter $O(1)$.

◀

Proof. 1. Consider the following procedure: At any time $t \geq 1$ let $G_t = (V, E_t)$ be the graph in the process where $G_0 = P_n$, and let $v \in V$ be an active node chosen uniformly at random. If v in G_t has an improving move then v adds a profitable edge. Otherwise it does nothing and another node is chosen. The process repeats until no agent has an improving move.

Consider the stochastic process $\{X_t\}_{t \geq 0}$, where $X_t = |E_t|$. Now if the graph G_t is not in equilibrium, then the probability $Pr(X_t < X_{t+1}) \geq \frac{1}{n}$ and $Pr(X_t = X_{t+1}) = 1 - Pr(X_t < X_{t+1})$. This implies that the expected number of steps until an improving response is played in the process $\{X_t\}_{t \geq 0}$ is dominated by the geometric random variable $G(\frac{1}{n})$. In the process $\{X_t\}_{t \geq 0}$ the absorbing state corresponds to a (k)NE network. Since $n - 1 \leq X_t \leq \frac{n(n-1)}{2}$, for any t, it follows that the expected number of steps needed to reach a (k)NE is at most $O(n^3)$.

2. Consider the adversarial scheme in Algorithm 2 which is illustrated in Figure 5.8. First we prove that each addition of an edge induced by this scheme is an improving response.

   Consider line 2 to 5 of the algorithm. At each step of the loop the degree of any node is at most $\lceil \frac{n}{2} \rceil$ and the distance improves by at least $\lceil \frac{n}{2} \rceil + 1$. Thus every edge addition in the loop is an improving response for the activated node. Now in the next loop, line 6–7, only node $v_{\lceil \frac{n}{2} \rceil}$ is active and each of its additions is an improving response since all bought edges have a cost of 2 and the distance cost improvement is $n - i + 1 \geq 3$. The edge added in line 8 is an improving response as it costs 3 and the distance cost improvement is at least $n - 2$.

   Now after this step we note that the network is in 2NE. Indeed, any two nodes from $\{v_1, \ldots, v_{n-2}\}$ are in distance at most 2 to each other and adding an edge to v_{n-1} or v_n improves the distance cost by only 1. Note that adding the edges $v_{n-1}v_{n-3}$,$v_{n-1}v_{\lceil \frac{n}{2} \rceil - 1}$, $v_{n-1}v_{n-3}$, or $v_{n-1}v_{\lceil \frac{n}{2} \rceil - 1}$ still may be an improvement. But buying an edge to $v_{\lceil \frac{n}{2} \rceil - 1}$ costs $n - 3$ and the distance cost improvement is at most $n - 4$. At the same time adding $v_{n-1}v_{n-3}$ shortcuts the distance to the nodes v_{n-3} and v_{n-4}, thus this move improves the distance cost by 2 and has a cost of at least 3. Hence the graph is in 2NE after line 8.

   Next, we prove that the remaining edge additions are improving responses in the degAOG. In line 11 the added edge yields an improvement for the active agent since it costs 2 and improves the distance to node v_{n-1} by 2 and to node v_n by 1. In the next loop, line 13 to 15, node v_{n-1} is active and it buys an edge to $v_{\lceil \frac{n}{2} \rceil + 3}$ and to every third node starting from there. These additions are improving responses since the degree of each node v_i is 3 and adding an edge to them improves v_{n-1}'s distance to v_i, v_{i-1} and v_{i-2} by 2, 1 and 1, respectively. After this step the diameter of the network is 3 and the degree of every node, except v_n, is at least 3, hence no agent can improve further, which implies that the network is in (k)NE, where $k \geq 3$.

   Note that already in the first loop, line 2 to 5, $\Theta(n^2)$ many improving responses are played. Therefore, the algorithm can be simplified, and steps after the first loop can be chosen arbitrarily.

3. Let $G = P_n$. For $i, j = 1, \ldots, n$, if $i < j$, we call an edge v_iv_j a *forward edge*,

Algorithm 2: Adversarial Scheme for the deg(k)AOG

1 **input** undirected path $P_n = (V, E)$ from v_1 to v_n;
2 **for** $i := 1$ *to* $\lceil \frac{n}{2} \rceil - 3$ **do**
3   activate v_i and add the edge $v_i v_{i+2}$ to E ;
4   **for** $j = i - 1$ *to 1* **do**
5   activate v_j and add the edge $v_j v_{i+2}$ to E;
6 **for** $i = \lceil \frac{n}{2} \rceil + 1$ *to* $n - 2$ **do**
7   activate $v_{\lceil \frac{n}{2} \rceil - 1}$ and add the edge $v_{\lceil \frac{n}{2} \rceil - 1} v_i$;
8 activate v_n and add the edge $v_n v_{n-2}$ to E;
9 **if** *the model under consideration is the deg2AOG* **then**
10   go to line 16;
11 activate $v_{\lceil \frac{n}{2} \rceil}$ and add the edge $v_{\lceil \frac{n}{2} \rceil} v_{n-1}$;
12 $i := \lceil \frac{n}{2} \rceil + 3$;
13 **while** $i \leq n - 5$ **do**
14   activate v_{n-1} and add the edge $v_{n-1} v_i$ to E;
15     $i := i + 3$;
16 **return** Graph $G = (V, E)$;

otherwise, if $i > j$, a *backward edge*. We consider round-robin activation in the order $v_1, v_2, \ldots, v_n$.

Starting with $v_1, v_2, \ldots, v_{n-4}$ every agent v_i can buy a forward edge to node v_{i+2} since $deg_G(v_{i+2}) = 2$ and the improvement in distance cost is $n-(i+1)$. Another possibility for v_i is to buy a backward edge to v_{i-3} then it improves her distance to every second node in $v_1, \ldots, v_{i-4}$ and to node v_{i-3} by 1, i.e., the distance cost improvement is at most $\frac{i-2}{2}$. At the same time buying a backward edge to v_{i-4} costs 4 and improves the distance cost by $i - 4$. Hence, buying a forward edge is the best single edge addition for v_i if $n - (i + 1) - 2 \geq (i - 4) - 4$. Thus the best single edge addition of agent $v_{\lceil \frac{n+5}{2} \rceil}$ is an edge to $v_{\lceil \frac{n-3}{2} \rceil}$.

Now agent $v_{\lceil \frac{n+5}{2} \rceil + 1}$ can buy an edge to $v_{\lceil \frac{n+5}{2} \rceil + 2}, v_{\lceil \frac{n+5}{2} \rceil - 3}, v_{\lceil \frac{n+5}{2} \rceil - 4}$ or to $v_{\lceil \frac{n+5}{2} \rceil - 5}$. Observe that buying a backward edge to the node with the lowest possible index maximizes the distance cost improvement. Adding the edge $v_{\lceil \frac{n+5}{2} \rceil} v_{\lceil \frac{n+5}{2} \rceil - 5}$ costs 5 and it improves the distance towards all

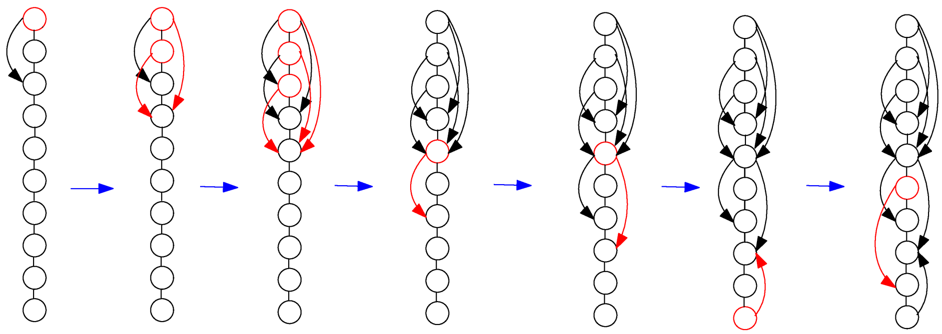


Figure 5.8: Example of an improving move sequence induced by Algorithm 2. (a) is an 2NE in the deg2AOG and (b) is (k)NE in deg(k)AOG, where $k \geq 3$.

nodes $v_1, \ldots, v_{\lceil \frac{n+5}{2} \rceil - 5}$ by 1. This yields a higher improvement than buying a forward edge since $\lceil \frac{n+5}{2} \rceil - 4 + 1 - 5 \geq n - (\lceil \frac{n+5}{2} \rceil + 1) - 2 \iff 2 \cdot \lceil \frac{n+5}{2} \rceil \geq n+5$. Thus, adding $(v_{\lceil \frac{n+7}{2} \rceil}, v_{\lceil \frac{n-3}{2} \rceil})$ is a best single edge purchase.

For nodes $v_{\lceil \frac{n+7}{2} \rceil}$ and $v_{\lceil \frac{n+9}{2} \rceil}$ adding an edge to the same node $v_{\lceil \frac{n-3}{2} \rceil}$ saves $\lceil \frac{n-3}{2} \rceil + 1$ and $\lceil \frac{n-3}{2} \rceil + 2$ in distance cost, respectively, and the edge cost is 5 and 6, respectively. For all the other nodes with higher index buying an edge to node $v_{\lceil \frac{n-3}{2} \rceil}$ saves $\lceil \frac{n-3}{2} \rceil + 3 + (x-2)$ whereas edge costs are $x + 4$, where x is the number of nodes which already bought a backward edge to $v_{\lceil \frac{n-3}{2} \rceil}$. Since buying to a node with larger index or buying a forward edge is inferior, it follows that buying an edge to $v_{\lceil \frac{n-3}{2} \rceil}$ is a best single edge purchase for all agents with index larger than $\lceil \frac{n+5}{2} \rceil$.

We call nodes having a degree linear in n a *high degree node*. So node $v_{\lceil \frac{n-3}{2} \rceil}$ becomes a high degree node with $deg_G(v_{\lceil \frac{n-3}{2} \rceil}) = n - \lceil \frac{n+5}{2} \rceil + 1 + 4 = \lfloor \frac{n+5}{2} \rfloor$. Consider the diameter of the network $D(G)$ after the first round. The length of the shortest path from v_1 to $v_{\lceil \frac{n-3}{2} \rceil}$ is roughly halved by the forward edges and since every node between $v_{\lceil \frac{n+5}{2} \rceil}$ and v_n buys an edge to $v_{\lceil \frac{n-3}{2} \rceil}$ the longest shortest path between the nodes $\{v_{\lceil \frac{n-3}{2} \rceil}, \ldots, v_n\}$ is 3 so $D(G)$ decreases from $n-1$ to $\lceil \frac{n+9}{4} \rceil$.

In the next round the best single improving edge of the first nodes v_i, where $i \leq \lceil \frac{n-11}{2} \rceil$ is an edge to v_{i+4}, since every edge costs 4 and distance cost improvement is $n - (i + 3)$. Buying a backward edge to v_{i-8} has cost 6 and improves the distance cost by $i - 8 + 1$.

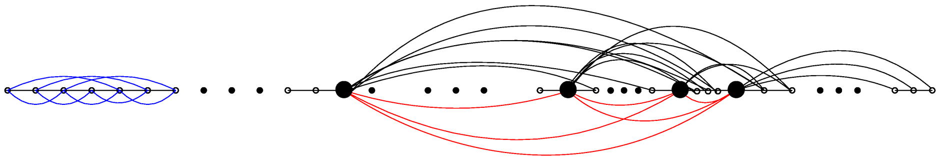


Figure 5.9: A sketch of the graph G in round k. The forward edge region, where the best single edge addition is to buy a forward edge, is shown in blue. The large nodes show the high degree nodes and the red edges belong to the clique formed by the high degree nodes.

Let the *forward edge region* of any round be the set of nodes which have a best single edge purchase which buys a forward edge, cf. Figure 5.9. In round k no node up to node $v_{\lceil \frac{n+2^k+2}{2} \rceil}$ will buy a backward edge, except when the forward edge would connect to a high degree node, since a forward edge from v_i to v_{i+2^k} costs $2k$ and improves the distance cost by $n - (i + 2^k - 1)$ while a backward edge from v_i to $v_{i-2^{k+1}}$ costs $2k + 2$ and the distance cost improvement is $i - 2^{k+1} + 1$. It follows that the diameter of the forward edge region halves every round.

Now we consider the connection between high degree nodes during each round. In the second round the first backward edge is bought by agent $v_{\lceil \frac{n-11}{2} \rceil}$, since buying a forward edge to the high degree node $v_{\lceil \frac{n-3}{2} \rceil}$ is clearly too expensive. When we are in round k a new high degree node v_i, with $i \geq \lceil \frac{n+21}{2} \rceil - 3 \cdot 2^{k+1}$ to which other nodes buy backward edges will occur.

Since there are at most 2^k nodes between the first node who starts to buy a backward edge and the high degree node from the last round and to the next high degree nodes there are at most $3 \cdot 2^{k-1}$, $3 \cdot 2^{k-2}$ etc. nodes which lie in-between, there can be at most $O(\log n)$ high degree nodes. Consider the high degree node with lowest index in round $k \geq 2$ and consider the step where the high degree node with the next higher index is active. At this point, all nodes between the two high degree nodes have bought a backward edge to the high degree node with lowest index, which implies that the high degree node with the second lowest index can buy a backward edge as well. Since not all nodes with lower index than the high degree node with lowest index have bought a forward edge to the high degree node with lowest index, it follows that the best single edge of the high degree node with the second lowest index connects to the high degree node with lowest index. Analogously, all high degree nodes with higher index will also buy an edge to the high degree node with lowest index. Thus, all high degree nodes after any round will form a clique. It follows that the distance of any node with higher index than the current high degree node with lowest index to the latter node is at most 3.

Since the diameter of the forward edge region halves in every round there can be at most $O(\log n)$ rounds until the diameter of the network is in $O(1)$. At this point all nodes with lower index than the lowest index of the clique nodes can possibly buy an improving forward edge to all clique nodes but no improving backward edge can be bought by any agent. Thus, there can be $O(\log n)$ additional rounds and the total number of additional edges is in $O(n \log n)$.

In total the dynamics needed $O(\log n)$ rounds and $O(n \log n)$ best single edge improvements are made.

■

We contrast the upper bounds by showing that convergence in $O(n)$ many improving responses is possible.

► **Theorem 5.21.** Let $P_n = \{v_1 \cdots v_n\}$ be the path of length n, with v_1 and v_n as leaf nodes, then there exists a sequence of improving responses which takes

1. $n - 2 + \frac{n-7}{3}$ steps to obtain a NE network in the degAOG;
2. $n - 1$ steps to obtain a 2NE network in the deg2AOG.

◄

Proof. 1. Consider a path $P_n = v_1, \cdots, v_n$. We activate v_1 and sequentially buy the edges $v_1 v_i$ for each $3 \leq i \leq n-2$. Afterwards we activate v_{n-1} and sequentially buy the edges $v_{n-1} v_{i+3}$ with $0 \leq i \leq n-8$. Each edge addition is an improving move since v_1 pays 2 to improve the distance cost by at least 3 and v_n pays 3 and improves her distance cost by 4 to the nodes v_{i+2}, v_{i+3} and v_{i+4}. After that no edge addition is possible as each edge costs at least 3 whereas the possible distance cost improvement is at most 3. Thus the network is in NE after $n - 2 + \frac{n-7}{3} = \Theta(n)$ many steps.

2. Consider the following activation scheme in which node v_1 buys an edge to each node v_i where $3 \leq i \leq n-2$. Every single edge purchase is an improving response because the edge cost is 2 and the distance cost improvement is at least 3. Next, we activate v_n and add the edge $v_n v_{n-2}$. This is improving for v_n, since the edge costs 3 and it decreases the distance cost by $n-2$.

   All nodes $\{v_1, \ldots, v_{n-2}\}$ are in distance at most 2 towards each other, thus there are no possible further improving edge additions between these agents. Buying an edge to v_{n-1} and v_n costs 2 and 1, respectively, whereas the distance cost improvement is also 2 and 1, respectively. The nodes v_{n-1} and v_n do not want to buy an edge to v_1 as it costs $n-3$ and the distance cost improvement is $n-4$. There is no other local improving response for both nodes as each node in their 2-neighborhood costs 3 and the distance cost improvement is just 1. Therefore the network is in 2NE after $n-1$ many steps.

■

Finally, we investigate the quality of the (2)NE networks which can be obtained by improving move dynamics starting from the path P_n. For this we introduce a measure which is similar to the Price of Anarchy. Let G_0 be any initial connected network and let $Z(G_0)$ be the set of networks which can be obtained via improving response dynamics in the deg(2)AOG. Let $Best(G_0) \in Z(G_0)$ be the reachable network with the minimum social cost among all networks in $Z(G_0)$. We can now measure the quality of any network $G \in Z(G_0)$ by investigating the ratio $\rho(G, G_0) = \frac{cost(G)}{cost(Best(G_0))}$.

▶ **Theorem 5.22.**

1. Let G be any network in $Z(G_0)$ then $\rho(G, G_0) \in \mathcal{O}(n)$.
2. There is a network $G \in Z(P_n)$ with $\rho(G, P_n) \in \Theta(n)$ in the deg(2)AOG.
3. Let G^* be the network obtained by the round-robin best single edge dynamics in the deg2AOG, then we have $\rho(G^*, P_n) \in \mathcal{O}(\log n)$.

◀

Proof. 1. To provide an upper bound on $\rho(G, G_0)$ we evaluate an upper bound on the social cost of G. Let D be the diameter of G. Then the total distance cost of all agents is at most $Dn^2 \leq n^3$. The total edge cost of all agents is in $\mathcal{O}(n^3)$. The social cost of $Best(G_0)$ is at least $n(n-1)$, since the diameter of $Best(G_0)$ is at least 1. Hence, $\rho(G, G_0) \in \mathcal{O}(n)$.

2. The upper bound follows from the first part of the proof. The matching lower bound follows from the example network G given in the second part of the proof of Theorem 5.20. G has $\Theta(n^2)$ many edges which all have cost in $\Theta(n)$. It yields that the social cost of G is in $\Theta(n^3)$. Now consider the activation scheme in the proof of Theorem 5.21. In the constructed equilibrium network G', the total edge cost is in $\Theta(n)$. Thus, the social cost of G' is in $\Theta(n^2)$, which is an upper bound on the social cost of $Best(P_n)$. Hence, $\rho(G, P_n) \in \Omega(n)$.

3. We prove the upper bound on $\rho(G^*, P_n)$ by providing an upper bound on the social cost of G^* and the lower bound on $Best(P_n)$.

   From the third part of Theorem 5.20 we know that the total number of edges and the diameter of G^* is in $\mathcal{O}(n \log n)$ and $\mathcal{O}(1)$, respectively. Since the diameter of G^* is in $\mathcal{O}(1)$, the total distance cost can be upper bounded by $\mathcal{O}(n^2)$. The total edge cost of G^* is upper bounded by $\mathcal{O}(n^2 \log n)$, since there are $\mathcal{O}(\log n)$ many high degree nodes with degree $\Theta(n)$, and almost all edges are bought towards these high degree nodes and thus each have cost in $\Theta(n)$. Hence, G^* has a social cost in $\mathcal{O}(n^2 \log n)$.

   On the other hand, a trivial lower bound on the social cost of $Best(P_n)$ is $n(n-1)$, which then yields $\rho(G^*, P_n) \in \mathcal{O}(\log n)$.

■

5.7 Conclusion

We have introduced natural variants of the classic NCG, which have the distinctive features that they are parameter-free and at the same time incorporate non-uniform edge costs. Besides proving that computing a best response is NP-hard and that improving response dynamics may never converge to an equilibrium, we have also established that the degNCG has a constant Price of Anarchy. This strong statement holds whenever the edge price is any linear function of the degree of the non-owner endpoint of the edge or if agents are allowed to buy edges to nodes in their 3-neighborhood. For the version which includes stronger locality, i.e., the deg2NCG, we have shown that the PoA is in $O(\sqrt{n})$ and, as a contrast, for the add-only version the PoA is in $\Theta(n)$. We also demonstrate how to circumvent the latter negative result by using suitable activation schemes on a sparse initial network.

Studying the bilateral version of our model, where both endpoints of the edge have to agree and pay proportionally to the degree of the other endpoint for establishing an edge, is an obvious future research direction. For this version, we have already established that most of our proofs can be easily adapted, which implies that our results, with minor modifications, still hold. Another interesting extension would be to consider an edge price function which depends on the degree of *both* involved nodes. This could be set up such that edges between nodes of similar degree are cheap and edges become expensive when the degree of both nodes differs greatly.

6 Selfish Creation of Social Networks

In this chapter, we propose and analyze a simple and very general game-theoretic model which is inspired by real-world social networks. It was observed that in real-world social networks connections are often established by recommendations from common acquaintances or by a chain of such recommendations. Thus establishing and maintaining a contact with a friend of a friend is easier than connecting to complete strangers. This explains one of the core properties of real-world networks, the high clustering, i.e., the abundance of triangles. To model this situation in terms of the NCGs, we consider a network creation game where selfish agents bilaterally form costly links to increase their centrality. The cost of each link is proportional to the distance of the endpoints before establishing the connection. We provide results for generic cost functions, which essentially only must be convex functions in the distance of the endpoints without the respective edge. For this broad class of cost functions, we provide many structural properties of equilibrium networks and prove (almost) tight bounds on the diameter, the Price of Anarchy and the Price of Stability. Moreover, as a proof-of-concept we show via experiments that the created equilibrium networks of our model indeed closely mimic real-world social networks. We observe degree distributions that seem to follow a power-law, high clustering, and low diameters. Hence, our model promises to be the first game-theoretic network formation model which predicts networks that exhibit all core properties of real-world networks.

6.1 Model and Notation

We consider a new model, called the *Social Network Creation Game (SNCG)*, which is related to the bilateral network creation game [CP05]. The set of n selfish agents V corresponds to the nodes of a network and the agents' strategies determine the edge-set of the formed network G. Each agent u tries to optimize a cost function $cost(u, G(\mathbf{s}))$, which depends on the structure of the network G.

In real-world social networks new connections are formed by a bilateral agreement of both endpoints while an existing connection can be unilaterally removed by any one of the involved endpoints. Following this idea, we consider only single edge additions with consent of both endpoints or single edge deletions as possible (joint) strategy changes of the agents. As equilibrium concept we adopt the well-known solution concept called *pairwise stability* [JW96]. Intuitively, a network G is pairwise stable if every edge of G is beneficial for both endpoints of the edge and for every non-edge of G, at least one endpoint of that edge would not decrease her cost by creating the edge. More formally, $G = (V, E)$ is pairwise stable if and only if the following conditions hold:

1. for every edge $uv \in E$, we have $cost(u, G - uv) \geq cost(u, G)$ and $cost(v, G - uv) \geq cost(v, G)$;
2. for every non-edge $uv \notin E$, we have $cost(u, G + uv) \geq cost(u, G)$ or $cost(v, G + uv) \geq cost(v, G)$.

Created edges are bidirectional and can be used by all agents, but the cost of each edge is equally shared by its two endpoints.

The main novel feature of our model is the definition of the cost of any edge $uv \in E$, which is proportional to the distance of both endpoints without the respective edge, i.e., proportional to $d_{G-uv}(u, v)$. This is motivated by the fact that, in social networks, the probability of establishing a new connection between two parties is inversely proportional to their degree of separation. More precisely, let $\sigma : \mathbb{N} \longrightarrow \mathbb{R}_+$ be a monotonically increasing convex function such that $\sigma(0) = 0$. The cost of the edge uv in network G is equal to

$$c_G(uv) = \begin{cases} \sigma(d_{G-uv}(u, v)) & \text{if } d_{G-uv}(u, v) \neq +\infty, \\ \sigma(n) & \text{otherwise,} \end{cases}$$

where $d_G(u, v)$ is the hop-distance between u and v in $G = (V, E)$, i.e., the number of edges in a shortest path between u and v in G. We assume that $d_G(u, v) = +\infty$ if no path between u and v exists in G.

We call an edge uv a *k-edge* if $d_{G-uv}(u, v) = k$, and a *bridge* (or *n-edge*) if $d_{G-uv}(u, v) = +\infty$. If the network is clear from the context, we will sometimes omit the reference to G and we still simply write $c(uv)$ to denote the cost of edge uv. Note that by definition, any bridge in G, i.e., any edge whose removal would increase the number of connected components of G, has cost $\sigma(n) > \sigma(n - 1)$ and thus any bridge has higher cost than any other non-bridge edge. The latter

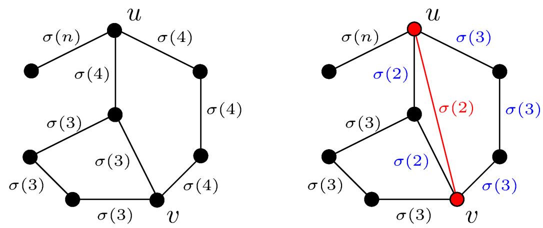


Figure 6.1: Edge costs before and after the edge uv is added.

property can be understood as an incentive towards more robust networks. Note, that the addition or removal of an edge in network G can also influence the cost of other edges in G. See Figure 6.1 for an example.
The resulting cost for an agent u in the network G is the sum of the cost of all edges incident to u and the sum of distances to all other agents:

$$cost(G,u) := \frac{1}{2} \sum_{v \in N_G(u)} c_G(uv) + \sum_{v \in V} d_G(u,v),$$

where $N_G(u)$ is the set of all neighbors of u in G. The quality of the created network is measured by its *social cost*, which is denoted by $cost(G) := \sum_{u \in V} cost(u,G)$ and measures the total cost of all agents.

As in the original bilateral network creation game [CP05], we restrict our study to connected networks, as any pairwise stable non-connected network has an unbounded social cost[5].

6.2 Related Work

The SNCG is a variant of the bilateral network creation game (BNCG) [CP05]. In contrast to the original unilateral NCG, for the BNCG it was shown that the PoA of the BNCG is in $\Theta(\min\{\sqrt{\alpha}, n/\sqrt{\alpha}\})$ and that equilibrium networks having a diameter in $\Theta(\sqrt{\alpha})$ exist [CP05; Dem+12]. The original NCG was dedicated to model real-world networks like peer-to-peer networks and social networks. However, the main downside of these classical models is that they do not predict a realistic degree distribution or high clustering. One NCG variant was proposed where agents try to maximize their local clustering instead of their

5 E.g., any network with no edges and $n \geq 3$ is pairwise stable.

centrality [BK11]. This model yields various sparse equilibrium networks with high clustering but these can have a large diameter and a homogeneous degree distribution.

The SNCG incorporates a robustness aspect since bridge-edges are expensive. This fits to a recent research trend for studying robust network formation [Che+19; Ech+20; Goy+16; MMO15].

Despite the variety of studied network formation models, to the best of our knowledge, no simple game-theoretic model exists, which predicts a low diameter, a power-law degree distribution and high clustering in its equilibrium networks. We are also not aware of any simulation results in this direction. However, there are two promising but very specialized candidates in that direction. The first candidate, which is particularly tailored to the web graph [Kou+15], yields directed equilibrium networks that share many features of real-world content networks. The second candidate uses a game-theoretic framework and hyperbolic geometry to generate networks with real-world features. In the network navigation game [Gul+15], agents correspond to randomly sampled points in the hyperbolic plane and they strategically create edges to ensure greedy routing in the created network. It is shown that the equilibrium networks indeed have a power-law degree distribution and high clustering. However, the main reason for this is not the strategic behavior of the agents but the fact that the agents correspond to uniformly sampled points in the hyperbolic plane. It is known that the closely related hyperbolic random graphs [Kri+10] indeed have all core properties of real-world networks.

6.3 Properties of Equilibrium Networks

In this section we prove structural properties satisfied by all connected pairwise stable networks that will be useful in proving our main results. We first provide a nice property satisfied by the function σ following from its convexity.

▶ **Proposition 6.1.** Fix a positive real value x. Let $x_1, \ldots, x_k$, with $0 \leq x_i \leq x$, be $k \geq 2$ positive real values and let $\lambda_1, \ldots, \lambda_k$, with $\lambda \in [0, 1]$, such that $x = \sum_{i=1}^{k}(\lambda_i x_i)$. Then $\sigma(x) \geq \sum_{i=1}^{k}\left(\lambda_i \sigma(x_i)\right)$. ◀

Proof. We show that $\sigma(x_i) \leq \frac{x_i \sigma(x)}{x}$. This is enough to prove the claim since

$$\sum_{i=1}^{k} \big(\lambda_i \sigma(x_i)\big) \leq \frac{\sigma(x)}{x} \sum_{i=1}^{k} (\lambda_i x_i) \leq \sigma(x).$$

Let $x_i = \bar{\lambda}_i x$, i.e., $\bar{\lambda}_i = \frac{x_i}{x}$, and observe that $\bar{\lambda}_i \in [0,1]$. By convexity of σ we have that

$$\sigma(x_i) = \sigma\big((1-\bar{\lambda}_i)0 + \bar{\lambda}_i x\big) \leq (1-\bar{\lambda}_i)0 + \bar{\lambda}_i \sigma(x) = \frac{x_i}{x}\sigma(x).$$ ■

In the next statement we claim that nodes can be incident to at most one expensive edge. Hence, the number of such edges is limited. This property implies a high number of triangles (induced cycles of size 3) in stable networks.

▶ **Proposition 6.2.** In any pairwise stable network, any node has at most one incident edge of cost at least $\sigma(4)$. If $2\sigma(2) \leq \sigma(3)$ holds, any node in a pairwise stable network has at most one incident edge of cost at least $\sigma(3)$. ◀

Proof. Let uv and vw be two distinct edges of G that are incident to v. We prove the claim by showing that at most one of these edges can have a cost of at least $2\sigma(2)$. This implies the claim since, by Proposition 6.1, $\sigma(4) \geq 2\sigma(2)$.

Note that the edge $uw \notin E(G)$ because otherwise, it implies that costs of all edges equal $\sigma(2)$. If both edges uv and vw have a cost of at least $2\sigma(2)$ each, then G is not pairwise stable as, by adding the edge uw, the total edge cost of both agent u and agent w does not increase, while the total distance cost of each of the two agents decreases by at least 1. In fact, the edge cost of each edge uv, uw, and vw in $G + uw$ is equal to $\sigma(2)$. ■

Next, we establish that all pairwise stable networks contain almost no bridges. This property can be observed in most real-world social networks.

▶ **Theorem 6.3.** Any pairwise stable network contains at most three bridges. ◀

Proof. We prove that any network G with four or more bridges cannot be pairwise stable. First, we show that in this case there are two bridges at a distance of at most $n/2 - 2$ in G, i.e., two nodes u_1 and u_2 that are incident to 2 distinct bridges $e_1 = u_1v_1$ and $e_2 = u_2v_2$ such that $d_G(u_1, u_2) \leq \frac{n}{2} - 2$. We observe that this is enough to prove the claim. Indeed, w.l.o.g., let $d_G(v_1, v_2) = 2 + d_G(u_1, u_2)$.

We have that $d_G(v_1, v_2) \leq \frac{n}{2}$. By adding the edge v_1v_2 to G the total distance cost of both agents v_1 and v_2 decreases by at least 1, while their total edge cost differs by $\sigma(n/2) - \frac{1}{2}\sigma(n) \leq 0$ as, by Proposition 6.1, $\sigma(n) \geq 2\sigma(n/2)$. Hence, G cannot be pairwise stable.

We now complete the proof by showing that there are two bridges at a distance of at most $\frac{n}{2} - 2$.

Let $\mathcal{T}$ be the *block-cut tree* decomposition of G, i.e., a decomposition of G into maximal *2-connected components* and *cut nodes*.[6] Notice that each bridge uv is represented in $\mathcal{T}$ as a 2-connected component that is connected with the two cut nodes u and v. Let $\mathcal{T}'$ be a minimal connected subtree of $\mathcal{T}$ that contains exactly four 2-connected components that are bridges, and let G' be the sub-network of G whose block-cut tree decomposition is represented by $\mathcal{T}'$. Let $e_i = u_iv_i$, with $i = 1, 2, 3, 4$, be the four bridges of G'. We denote by $d(e_i, e_j) = \min\left\{d_{G'}(u_i, u_j), d_{G'}(u_i, v_j), d_{G'}(v_i, u_j), d_{G'}(v_i, v_j)\right\}$ the distance in G' between the two bridges e_i and e_j. We shall prove that $\min_{1 \leq i < j \leq 4} d(e_i, e_j) \leq \frac{n}{2} - 2$.

Since each edge $e_i = u_iv_i$ is a bridge, we can assume that it is represented by an edge u_iv_i in $\mathcal{T}'$. Let $P_{i,j}$ denote the (unique) simple path in $\mathcal{T}'$ between the representatives of the edges e_i and e_j. The proof divides into two complementary cases, depending on the structures of the paths $P_{i,j}$, with $1 \leq i < j \leq 4$.

The first case is when at least two paths in $\{P_{i,j} \mid 1 \leq i < j \leq 4\}$ are node disjoint. W.l.o.g., we assume that $P_{1,2}$ is node disjoint w.r.t. $P_{3,4}$, and that the overall number of nodes of the 2-connected components corresponding to the nodes in $P_{1,2}$ is at most $n/2$. It is well-known that a 2-edge-connected network with n nodes has diameter of at most $\frac{2}{3}n < n$ [CS92]. Hence we have that

$$\min_{1 \leq i < j \leq 4} d(e_i, e_j) \leq d(e_1, e_2) \leq \frac{2}{3}\left(\frac{n}{2} - 2\right) < \frac{n}{2} - 2.$$

The second case is when there are no two node disjoint paths in $\{P_{i,j} \mid 1 \leq i < j \leq 4\}$. This can happen only if there is exactly one 2-connected component,

6 A node x of a connected network G is a *cut node* if its removal from G results in a network that is not connected. A *2-connected network* is a connected network with no cut node. A *2-connected component* of G is a maximal (w.r.t. node insertion) 2-connected sub-network of G. A *block-cut tree* $\mathcal{T}$ of G is a tree where each tree node represents either a cut node or a 2-connected component of G. More precisely, there is an edge between the representative of a cut node x of G and the representative of a 2-connected component G' of G if and only if x is a node of G'.

say C that is traversed by all the four paths $P_{i,j}$. Let $n_C := |C|$ denote the number of nodes of C in G'. For $i = 1, 2, 3, 4$, let n_i be the overall number of nodes of the 2-connected components corresponding to the nodes in the (unique) simple path in $\mathcal{T}'$ from the node that represents e_i to the node that represents the 2-connected components right before C. Clearly, $n_1, n_2, n_3, n_4 \geq 2$ and $n_C \leq n-4$. W.l.o.g., we assume that $n_1 \leq n_2 \leq n_3 \leq n_4$. We prove that $d(e_1, e_2) \leq \frac{n}{2} - 2$. It is well known that the diameter of any 2-connected network with n nodes is at most $\lceil \frac{n-1}{2} \rceil$ [CS92]. We divide the proof into three sub-cases. The first sub-case is when $n_1 = n_2 = 2$. This implies that

$$d(e_1, e_2) \leq \left\lceil \frac{n_C - 1}{2} \right\rceil \leq \left\lceil \frac{n - 4 - 1}{2} \right\rceil \leq \frac{n}{2} - 2.$$

The second sub-case is when $n_1 = 2$ and $n_2 > 2$. In this case, we have that $n_2, n_3, n_4 \geq 4$. Moreover, $n_C + n_1 + n_2 + n_3 + n_4 \leq n + 4$ from which we derive $n_C \leq n + 2 - 3n_2$. The diameter of C is at most $\lceil (n_C - 1)/2 \rceil$ as C is 2-connected. Moreover, all the other 2-connected components traversed by $P_{1,2}$, except for e_1 and e_2 and C, form a 2-edge-connected networks of diameter at most $\frac{2}{3}(n_2 - 1)$. Therefore, we have that

$$\begin{aligned} d(e_1, e_2) &\leq \left\lceil \frac{n_C - 1}{2} \right\rceil + \frac{2}{3}(n_2 - 1) \leq \frac{n_C}{2} + \frac{2}{3}n_2 - \frac{2}{3} \\ &\leq \frac{n}{2} + 1 - \frac{3}{2}n_2 + \frac{2}{3}n_2 - 1 \leq \frac{n}{2} - \frac{5}{6}n_2 \\ &< \frac{n}{2} - 2. \end{aligned}$$

The third and last sub-case is when $n_1, n_2 > 2$. In this case, we have that $n_1, n_2, n_3, n_4 \geq 4$. Moreover, $n_C + n_1 + n_2 + n_3 + n_4 \leq n + 4$ from which we derive

$$n_C \leq n + 4 - 2(n_1 + n_2) \leq n + 4 - \frac{4}{3}(n_1 + n_2) - \frac{2 \cdot 8}{3} = n - \frac{4}{3}(n_1 + n_2) - \frac{4}{3}.$$

The diameter of C is at most $\lceil (n_C - 1)/2 \rceil$ as C is 2-connected. Moreover, all the other 2-connected components traversed by $P_{1,2}$, except for e_1 and e_2 and C, form two 2-edge-connected networks of diameter at most $\frac{2}{3}(n_1 - 1)$ and $\frac{2}{3}(n_2 - 1)$, respectively. Therefore, we have that

$$\begin{aligned} d(e_1, e_2) &\leq \left\lceil \frac{n_C - 1}{2} \right\rceil + \frac{2}{3}(n_1 - 1) + \frac{2}{3}(n_2 - 1) \\ &\leq \frac{n_C}{2} + \frac{2}{3}(n_1 + n_2) - \frac{4}{3} \\ &\leq \frac{n}{2} - \frac{2}{3}(n_1 + n_2) - \frac{2}{3} + \frac{2}{3}(n_1 + n_2) - \frac{4}{3} \\ &= \frac{n}{2} - 2. \end{aligned}$$

This completes the proof. ■

The following theorem shows an upper bound on the diameter of any pairwise stable network that only depends on the cost of edges which close a triangle.

▶ **Theorem 6.4.** The diameter of any pairwise stable network in the SNCG is at most $\sigma(2) + 1$. ◀

Proof. Consider a pairwise stable network G of diameter D. Let $v_0, v_1, \dots, v_D$ be a diametral path of G. Consider the addition of the edge between $v_{\lfloor D/2 \rfloor - 1}$ and $v_{\lfloor D/2 \rfloor + 1}$ to network G. Each node $v_0, \dots, v_{\lfloor D/2 \rfloor - 1}$ becomes 1 unit closer to $v_{\lfloor D/2 \rfloor + 1}$; similarly, each node $v_{\lfloor D/2 \rfloor + 1}, \dots, v_D$ becomes 1 unit closer to $v_{\lfloor D/2 \rfloor - 1}$. In both cases, the distance cost of the considered agent decreases by at least $\lfloor D/2 \rfloor$. Since the network is pairwise stable, both agents $v_{\lfloor D/2 \rfloor - 1}$ and $v_{\lfloor D/2 \rfloor + 1}$ have no incentive in buying the considered edge. Therefore, $\sigma(2)/2 - \lfloor D/2 \rfloor \geq 0$. Since $\lfloor D/2 \rfloor \geq \frac{D-1}{2}$, we get $D \leq \sigma(2) + 1$. ■

Finally, we prove an upper bound on the cost of non-bridge edges. This implies that all pairwise stable networks contain only small minimal cycles, i.e., cycles where the shortest path between two nodes in the cycle is along the cycle.

▶ **Proposition 6.5.** In a pairwise stable network, for all $k \notin \{2, 3, n\}$, the cost of any k-edge is $\sigma(k) < n\sigma(2)$. If $\sigma(2) \leq \frac{1}{2}\sigma(3)$ holds, for all $k \notin \{2, n\}$, the cost of any k-edge is $\sigma(k) \leq n\sigma(2)$. ◀

Proof. Consider a pairwise stable network G. Assume to the contrary that there is a non-bridge k-edge uv in G of cost $\sigma(k)$. Consider the deletion of the edge uv by one of its endpoints, say u. Let V_v (resp., V_u) be a subset of nodes such that all shortest paths between u (resp., v) and any node in V_v (resp., V_u) goes through the edge uv. By Proposition 6.2, all other edges incident to either u or v are 2-edges and 3-edges. As a consequence, the deletion of the edge uv does not increase the cost of the edges incident to u and v. Therefore, the edge cost of u (resp., v) decreases by $\frac{1}{2}c_G(uv)$, while the distance cost of u increases by at most $|V_v|(d_{G-uv}(u,v)-1) = |V_v|(k-1)$ (resp., $|V_u|(k-1)$). Since G is pairwise stable, the u's (resp., v) cost difference is greater than zero, i.e., $-\sigma(k)/2+|V_v|(k-1) \geq 0$ (resp., $-\sigma(k)/2 + |V_u| \cdot (k-1) \geq 0$). We sum up the two inequalities and get $\sigma(k) \leq (|V_v| + |V_u|) \cdot (k-1)$. Note that $V_v \cap V_u = \emptyset$. Indeed, if there is $x \in V_v \cap V_u$, then $d_G(u,x) = 1 + d_G(v,x) = 1+1+d_G(u,x)$, i.e., $0 = 2$. Therefore, $\sigma(k) \leq n(k-1)$.

If we assume $\sigma(2) \leq \frac{1}{2}\sigma(3)$, each node has at most one incident 3-edge according to Proposition 6.2. Similarly to the above proof, we obtain $\sigma(3) \leq 2n$.

Now we consider the addition of a 2-edge $u'v$ in the minimal cycle of length $(k+1)$ that contains the edge uv by the node v and a neighbor u' of u in the cycle. First, we consider $k \geq 4$. For both endpoints, this move improves the distance to at least $\frac{k+1-3}{2}$ nodes in the cycle. Moreover, by Proposition 6.2, all other of v's incident edges are 2- or 3-edges; therefore v has at least one neighbor v' that is not in the cycle and $d_G(u',v') \geq k-1 \geq 3$ (otherwise, it would not be a k-edge). Analogously, u' has a neighbor outside of the cycle at distance at least 3 from v. This implies that both endpoints of the edge $u'v$ will improve their distance to at least $\frac{k-2}{2}+1$ nodes by 1 after adding the edge. Since G is pairwise stable, this move is not profitable, i.e., $\frac{\sigma(2)}{2} - \frac{k-2}{2} - 1 \geq 0$. Hence, $\sigma(2) \geq k$. Combining this inequality with the inequality $\sigma(k) \leq n(k-1)$ from the first part of the proof, we get $\sigma(k) < n\sigma(2)$, if $k \geq 4$. If $k = 3$ and $\sigma(2) \leq \frac{1}{2}\sigma(3)$, the addition of a 2-edge can improve the distance to only one node. Since we assume that G is pairwise stable, $\frac{1}{2}\sigma(2) - 1 \geq 0$, i.e., $\sigma(2) \geq 2$. Combining this inequality with the above inequality for 3-edges, we get $\sigma(3) \leq 2n \leq n\sigma(2)$. The statement follows. ∎

6.4 Equilibrium Existence and Social Optima

Clique and fan networks play an important role since, as we will prove, the former are social optima when $\sigma(2) \leq 2$, while the latter are social optima when $\sigma(2) \geq 2$. Furthermore, we also show that complete networks are pairwise stable whenever $\sigma(2) \leq 2$, while (almost) fan networks are pairwise stable whenever $\sigma(2) \geq 2$.

As a reminder of the definition: The *clique graph* of n nodes is denoted by $\mathbf{K}_n$. A *fan graph* $\mathbf{F}_n$ with n nodes consists of a star with $n-1$ leaves $v_0, \ldots, v_{n-2}$ augmented with all the edges of the form $v_{2i}v_{2i+1}$, for $i = 0, \ldots, \lfloor \frac{n-2}{2} \rfloor$, where all indices are computed modulo $n-2$ (see Figure 6.2 for examples). In other words, $\mathbf{F}_n$, with n odd, is a star augmented with a perfect matching w.r.t. the star leaves, while $\mathbf{F}_n$, with n even, consists of $\mathbf{F}_{n-1}$ augmented with an additional node that is connected to the star center and any star leaf.

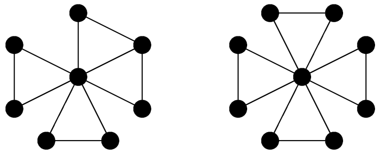


Figure 6.2: Two examples of fan networks.

▶ **Theorem 6.6.** If $\sigma(2) < 2$, then $\mathbf{K}_n$ is the unique social optimum. If $\sigma(2) > 2$, then $\mathbf{F}_n$ is the unique social optimum. Finally, if $\sigma(2) = 2$ any network of diameter 2 and containing only 2-edges is a social optimum.[7] ◀

Proof. Let G be a social optimum and assume that G contains m edges. We provide a lower bound $LB(m)$ on the social cost of G. The edge cost of G is lower bounded by $m\sigma(2)$; moreover, such a lower bound is matched only by networks whose edges are all 2-edges. The distance cost of G is lower bounded by $2(n(n-1)-2m)+2m = 2n(n-1)-2m$ as there are $2m$ ordered pairs of nodes that are at distance 1 (two pairs for each of the m edges) and $n(n-1)-2m$ ordered pairs of nodes that are at distance greater than or equal to 2. We observe that the lower bound on the distance cost is matched only by networks of diameter 2. Therefore, the social cost is lower bounded by

$$LB(m) := 2n(n-1) + (\sigma(2) - 2)m$$

and the value $LB(m)$ can be matched only by networks of diameter 2 whose edges are all 2-edges.

Clearly, when $\sigma(2) < 2$, we have that $LB(m)$ is minimized when m is maximized; furthermore, $\mathbf{K}_n$ is the only network whose social cost matches the lower

7 Hence, for $\sigma(2) = 2$, $\mathbf{K}_n$ and $\mathbf{F}_n$ are also social optima.

bound. Hence, for $\sigma(2) < 2$, $\mathbf{K}_n$ is the unique social optimum. For the case $\sigma(2) = 2$, we have that the value $LB(m) = 2n(n-1)$ is matched by any network of diameter 2 that contains only 2-edges. Hence, all and only such networks are social optima.

Now we prove the theorem statement for the remaining case in which $\sigma(2) > 2$. We have to show that $\mathbf{F}_n$ is the unique social optimum. First of all we observe that $LB(m)$ is minimized when m is minimized. Since the social cost of $\mathbf{F}_n$ matches the lower bound $LB(m')$ with $m' = n-1+\lceil\frac{n-1}{2}\rceil$, we have that $m \leq n-1+\lceil\frac{n-1}{2}\rceil$. Moreover, the social cost of the network is bounded only if the network is connected. This implies that G is connected, which in turn implies that $m \geq n-1$. Therefore, to prove that $\mathbf{F}_n$ is the unique social optimum, it is enough to prove that any network with m edges, with $n-1 \leq m \leq n-1+\lceil\frac{n-1}{2}\rceil$, has a social cost that is strictly larger than $2n(n-1)+(\sigma(2)-2)\big((n-1)+\lceil\frac{n-1}{2}\rceil\big)$, unless it is isomorphic to $\mathbf{F}_n$.

For the rest of the proof we can also assume that $n \geq 4$. In fact, for $n = 2$ it is clear that $\mathbf{F}_2 = \mathbf{K}_2$ is the social optimum since this is the only connected network of 2 nodes. Moreover, for $n = 3$ we have that $\mathbf{F}_3 = \mathbf{K}_3$ is again the only social optimum. Indeed, $\mathbf{F}_3$ has a social cost of $3\sigma(2)+6$, while the unique other connected network – i.e., the path of length 2 – has a social cost of $2\sigma(3)+8$. Since by Proposition 6.1, $\sigma(3) \geq \frac{3}{2}\sigma(2)$ (indeed $3 \geq 1\cdot 2 + 0.5\cdot 2$), it follows that $2\sigma(3)+8 \geq 3\sigma(2)+8 > 3\sigma(2)+6$.

Moreover, we can also assume that G is 2-edge-connected. Indeed, let G be a network that contains a bridge, say uv. W.l.o.g., let $u' \neq v$ be a neighbor of u, whose existence is guaranteed since $n \geq 4$. The distance cost of $G+u'v$ is strictly smaller than the distance cost of G. Moreover, the edge cost of $G+u'v$ is at most the edge cost of G. In fact, the cost of the bridge uv in G is at least $\sigma(4)$, while the cost of the two edges uv and $u'v$ in $G+u'v$ is at most $2\sigma(2)$ and, by Proposition 6.1, we have that $\sigma(4) \geq 2\sigma(2)$. As a consequence, the social cost of $G+u'v$ is strictly smaller than the social cost of G.

We divide the proof into two cases, depending on whether $m < n-1+\lceil\frac{n-1}{2}\rceil$ or not.

We consider the case in which $m < n-1+\lceil\frac{n-1}{2}\rceil$. Consider the subgraph H of G that is induced by 2-edges only. Such a subgraph contains $k \geq 1$ connected components, h of which are singleton nodes. Let $C_1,\ldots,C_{k-h}$ be the non-trivial connected components of H. Each C_i contains $n_i \geq 3$ nodes and m_i edges, where $m_i \geq n_i - 1 + \lceil\frac{n_i-1}{2}\rceil$. Indeed, each C_i can be generated starting from a triangle, i.e., $\mathbf{K}_3$, and by iteratively adding either one or two nodes so as the resulting induced subgraph contains at least one more triangle than before.

Clearly, at each step, we add either one node and at least two edges or two nodes and at least three edges. Obviously, the number of edges is minimized when we add two new nodes and exactly three edges at each step. Therefore, when n_i is odd, i.e., n_i-3 is even, we add at least three edges for every two nodes; when n_i is even, i.e., n_i-3 is odd, we add at least three edges for every two nodes except one node and at least two edges for the remaining node. As a consequence, when n_i is odd, we have $m_i \geq 3+3\frac{n_i-3}{2} = n_i - 1 + \lceil \frac{n_i-1}{2} \rceil$; when n_i is even, we have $m_i \geq 3+3\frac{n_i-4}{2}+2 = n_i - 1 + \lceil \frac{n_i-1}{2} \rceil$. In either case, $m_i \geq n_i - 1 + \lceil \frac{n_i-1}{2} \rceil$.

First of all, we observe that $n = h + \sum_{i=1}^{k-h} n_i$. Furthermore, since we are assuming $m < n-1+\lceil \frac{n-1}{2} \rceil$ it must be the case that $k \geq 2$. Indeed, for $k=1$, h would be equal to 0 and therefore $m_1 \geq n-1+\lceil \frac{n-1}{2} \rceil$. Finally, since we are assuming that G is 2-edge-connected, there are at least k edges of G each of which connects a node of one connected component with a node of another connected component. Clearly, the cost of each of these k edges is at least $\sigma(3)$ each. Therefore, since by Proposition 6.1, we have that $\sigma(3) \geq \frac{3}{2}\sigma(2)$, the overall edge cost of the network G is lower bounded by

$$\begin{aligned} k\sigma(3) + \sigma(2)\sum_{i=1}^{k-h} m_i &\geq \frac{3}{2}k\sigma(2) + \sigma(2)\sum_{i=1}^{k-h}\left(\frac{3}{2}(n_i-1)\right) \\ &= \frac{3}{2}k\sigma(2) + \frac{3}{2}(n-k)\sigma(2) \\ &> \left(n-1+\left\lceil \frac{n-1}{2} \right\rceil\right)\sigma(2). \end{aligned}$$

As the distance cost of G is lower bounded by $2n(n-1)-2m$, the overall social cost of G is strictly larger than $2n(n-1)+(\sigma(2)-2)\big((n-1)+\lceil \frac{n-1}{2} \rceil\big)$, i.e., the social cost of $\mathbf{F}_n$. Therefore, no network with $m < n-1+\lceil \frac{n-1}{2} \rceil$ can be a social optimum.

We now consider the remaining case in which $m = n-1+\lceil \frac{n-1}{2} \rceil$ and show that G is isomorphic to $\mathbf{F}_n$. First of all we observe that the social cost of G cannot be smaller than the social cost of $\mathbf{F}_n$ as the social cost of $\mathbf{F}_n$ matches the value $LB(m)$. This implies that $\mathbf{F}_n$ is a social optimum. For the sake of contradiction, assume that G is not isomorphic to $\mathbf{F}_n$. We show that G must satisfy some structural properties, based on three important observations. The first observation is that

G consists of only 2-edges and has diameter 2, as otherwise the social cost of G would be strictly larger than the value $LB(m)$. The second observation is that G has minimum degree equal to 2. Indeed, any network of minimum degree greater than or equal to 3 would have at least $m \geq \frac{3}{2}n > n-1+\lceil\frac{n-1}{2}\rceil$ edges. The third and last observation is that G cannot have a node of degree $n-1$, otherwise G would contain a spanning star, and hence, would be isomorphic to $\mathbf{F}_n$. As a consequence of these three observations, G has the following structure:

- it contains a node v that is connected to only two nodes, say u and u';
- it contains the edge uu' (otherwise the edges uv and $u'v$ would not be 2-edges);
- u and u' form a dominating set (otherwise the diameter of G would be at least 3);
- there is at least one node that is connected to u but not to u' and one node that is connected to u' but not to u (otherwise u or u' would have degree $n-1$) .

Let A be the set of nodes different from u' that are connected to u but not to u'; similarly, let B be the set of nodes different from u that are connected to u' but not to u. Finally, let C be the set of nodes different from v that are connected to both u and u' (see Figure 6.3). We observe that $|A|+|B|+|C|=n-3$. Each node

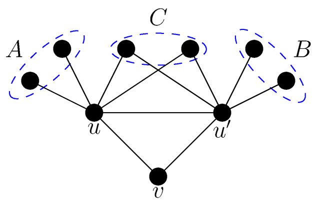


Figure 6.3: The structure of a 2-edge-connected network of diameter 2 with minimum degree 2 that is not isomorphic to $\mathbf{F}_n$. The node v has degree 2, the sets A and B are non-empty, while C can be empty. We have $|C| \geq 1$ if no edge ab with $a \in A$ and $b \in B$ exists. All the edges of the network are 2-edges only if each node $a \in A$ (resp., $b \in B$) is connected either with a node of C or another node of A (resp., B).

a of A must be connected to another node of $A \cup C$ as otherwise the edge ua would not be a 2-edge; similarly, each node b of B must be connected to another node of $B \cup C$ as otherwise the edge $u'b$ would not be a 2-edge. Furthermore, since G has diameter 2, it must be the case that for any two nodes $a \in A$ and $b \in B$ either ab is an edge of G or there is common neighbor $z \in C$ such that az and zb are both edges of G.

If we assume the existence of an edge ab for $a \in A$ and $b \in B$, then we can lower bound the number of edges of G with $m \geq 1 + n + |C| + \left\lceil \frac{|A|}{2} \right\rceil + \left\lceil \frac{|B|}{2} \right\rceil \geq n + 1 + \left\lceil \frac{n-3}{2} \right\rceil > n - 1 + \left\lceil \frac{n-1}{2} \right\rceil = m$.

If we assume that no edge ab with $a \in A$ and $b \in B$ exists, then $|C| \geq 1$ and we can lower bound the number of edges of G with $m \geq n + |C| + |A| + |B| = n + (n-3) > n - 1 + \left\lceil \frac{n-1}{2} \right\rceil = m$, as $n \geq 3 + |A| + |B| + |C| > 6$.

In either case, we have obtained a contradiction. Hence, $\mathbf{F}_n$ is the unique social optimum when $\sigma(2) > 2$. ■

Now we prove the existence of pairwise stable networks. For this we consider a modified fan graph $\mathbf{F}'_n$ that is equal to $\mathbf{F}_n$ if n is odd. If n is even, $\mathbf{F}'_n$ consists of $\mathbf{F}_{n-1}$ and one additional node connected to the center.

▶ **Theorem 6.7.** For $\sigma(2) \geq 2$, a modified fan graph $\mathbf{F}'_n$ is a pairwise stable network, otherwise a clique $\mathbf{K}_n$ is the unique pairwise stable network. ◀

Proof. First we show that if $\sigma(2) < 2$, any pairwise stable network is a clique. Assume to the contrary that there is a network G that is pairwise stable but at least one edge is missing. Adding any missing 2-edge costs $\sigma(2)/2$ for both its endpoints and improves the distance cost by at least 1. Thus, if $\sigma(2) < 2$, this move is an improvement.

Next, we prove that the modified fan graph is pairwise stable for $\sigma(2) \geq 2$. Assume $n \geq 3$, otherwise $\mathbf{F}'_n$ is trivially stable. If n is odd, deletion of any edge from $\mathbf{F}'_n$ increases the edge cost by $(\sigma(n) - 2\sigma(2))/2 > 0$ and increases distance between its endpoints by 1. Thus, any edge removal is not an improvement. On the other hand, buying any edge which is not present in $\mathbf{F}'_n$ costs $\sigma(2)/2$ and improves the distance only between its endpoints, i.e., it improves the distance cost by 1. Hence, since $\sigma(2) \geq 2$, an addition of any extra edge to the modified fan graph is not an improvement.

If n is even, then the pairwise stability of $\mathbf{F}'_n$ follows from an analogous proof as for odd n and from the observation that creating an edge with a leaf is not profitable. ■

6.5 PoA and PoS

Here we prove upper and lower bounds on the PoA and PoS.

▶ **Theorem 6.8.** The PoA of the SNCG is in $O\left(\min\{\sigma(2), n\} + \frac{\sigma(n)}{n \max\{\sigma(2), n\}}\right)$. For the class of 2-edge-connected networks the PoA is in $O\left(\min\{\sigma(2), n\}\right)$. ◀

Proof. By Theorem 6.6 and Theorem 6.7, we only need to focus on the case $\sigma(2) \geq 2$. Indeed, when $\sigma(2) < 2$, $\mathbf{K}_n$ is the unique pairwise stable network as well as the unique social optimum and, therefore, the PoA is equal to 1.

For the rest of the proof we assume that $\sigma(2) \geq 2$. By Theorem 6.6, $\mathbf{F}_n$ is a social optimum of cost $\Omega\left(n^2 + \sigma(2)n\right) = \Omega\left(n \max\{\sigma(2), n\}\right)$. Consider a pairwise stable network G of maximum social cost for a given number of nodes n. Let D be the diameter of G. A trivial upper bound for the distance cost of the network is $n(n-1) \cdot D$. By Theorem 6.4, the network diameter is at most $\sigma(2) + 1$, hence the distance cost of G is at most $(\sigma(2) + 1) \cdot n(n-1)$.

Now we will show an upper bound for the edge cost. Let k_i denote the number of i-edges in G. By Theorem 6.3, G has at most 3 bridges. Hence, for any pairwise stable network we have that $k_n \leq 3$; if the network is additionally 2-edge-connected, then $k_n = 0$. We consider two cases, depending on whether $2\sigma(2) \leq \sigma(3)$ or not.

We consider the case $2\sigma(2) \leq \sigma(3)$. By Proposition 6.2, each node has at most one incident i-edge where $3 \leq i < n$. Moreover, by Proposition 6.5, we have $\sigma(i) \leq n\sigma(2)$ for any $i \geq 3$. Then the overall edge-cost of the network is at most

$$\begin{aligned}
&k_2 \cdot \sigma(2) + \sum_{i=3}^{n-1} (\sigma(i) \cdot k_i) + k_n \sigma(n) \\
&\leq \left(\frac{n(n-1)}{2} - \sum_{i=3}^{n-1} k_i \right) \cdot \sigma(2) + n\sigma(2) \sum_{i=3}^{n-1} k_i + k_n \sigma(n) \\
&\leq \sigma(2) \cdot \frac{n(n-1)}{2} + (n-1)\sigma(2) \cdot \frac{n}{2} + k_n \sigma(n) \\
&\leq \sigma(2) n^2 + k_n \sigma(n).
\end{aligned}$$

As a consequence, the PoA is at most in $\frac{\sigma(2)n^2 + k_n\sigma(n)}{\Omega\left(n \max\{\sigma(2), n\}\right)}$. Thus, the PoA is in

$O\left(\min\{\sigma(2), n\} + \frac{\sigma(n)}{n \max\{\sigma(2),n\}}\right)$, while for the class of 2-edge-connected networks, the PoA is in $O\big(\min\{\sigma(2), n\}\big)$.

If $2\sigma(2) > \sigma(3)$, the upper bound for the edge cost from Proposition 6.5 holds for all i-edges such that $4 \le i \le n-1$. To estimate the number of 3-edges, we use result from extremal graph theory. By Mantel's Theorem [Man07] we have $k_3 \le \frac{n^2}{4}$. Then the edge cost of G is at most

$$
\begin{aligned}
& k_2 \cdot \sigma(2) + k_3 \cdot \sigma(3) + \sum_{i=4}^{n-1} (\sigma(i) \cdot k_i) + k_n \sigma(n) \\
\le & \left(\frac{n^2}{2} - \sum_{i=3}^{n-1} k_i \right) \cdot \sigma(2) + \frac{n^2}{4} \sigma(3) + n\sigma(2) \frac{n}{2} + k_n \sigma(n) \\
< & \ \sigma(2) \cdot n^2 + \frac{n^2}{4} \cdot \sigma(2) + k_n \sigma(n) \\
= & \ \frac{5}{4} \sigma(2) \cdot n^2 + k_n \sigma(n).
\end{aligned}
$$

As in the previous case, we get that the PoA is in $O\left(\min\{\sigma(2), n\} + \frac{\sigma(n)}{n \max\{\sigma(2),n\}}\right)$, while for the class of 2-edge-connected networks, the PoA is $O\big(\min\{\sigma(2), n\}\big)$. ∎

It is worth noticing that the high inefficiency of worst case pairwise stable networks in Theorem 6.8 follows from the existence of bridges in a network. The PoA is much better in bridge-free pairwise stable networks. Such networks can for example evolve via edge additions starting from a 2-edge-connected network. A real-world example for this would be co-authorship networks of authors with at least two papers.

We now prove lower bounds on the PoA. We start with the construction of a pairwise stable 2-edge-connected network with high social cost and a diameter in $\Omega(\sigma(2))$.

▶ **Lemma 6.9.** There are 2-edge-connected pairwise stable networks with $n = \Omega(\sigma(2))$ nodes, social cost in $\Omega\big(\sigma(2)n^2\big)$, and diameter of at least $\frac{\sigma(2)}{4}$. ◀

Proof. Let $k \ge 2$ be an integer and $n = 2\left\lceil \frac{\sigma(2)}{8} \right\rceil k + 1$. The pairwise stable network G of n nodes is obtained from a spider graph $\mathcal{S}$ with center x, with $\left\lceil \frac{\sigma(2)}{8} \right\rceil k + 1$

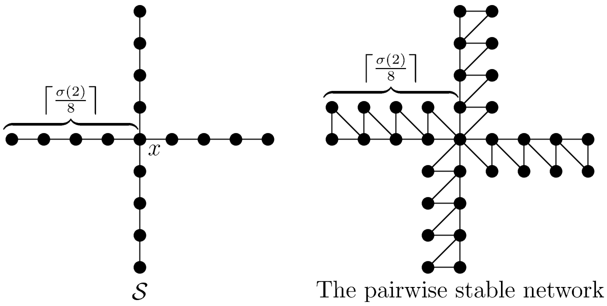


Figure 6.4: A high diameter pairwise stable network for $k = 4$ and $25 \leq \sigma(2) \leq 32$. Left: the spider $\mathcal{S}$ with k legs of length $\left\lceil \frac{\sigma(2)}{8} \right\rceil = 4$ each. Right: the pairwise stable network obtained by augmenting $\mathcal{S}$ with $\left\lceil \frac{\sigma(2)}{8} \right\rceil k$ new nodes.

nodes, and having k legs of length $\left\lceil \frac{\sigma(2)}{8} \right\rceil$ each, where we add a new node for each edge of $\mathcal{S}$ that is connected to both endpoints of the respective edge (see Figure 6.4 for an example).

The distance cost of G is at least n times the distance cost of the center x, i.e.,

$$n \sum_{i=1}^{\left\lceil \frac{\sigma(2)}{8} \right\rceil} (2ki) \geq kn \left\lceil \frac{\sigma(2)}{8} \right\rceil^2 = \Omega\big(\sigma(2)n^2\big).$$

Hence, the social cost of G is $\Omega\big(\sigma(2)n^2\big)$. Finally, the diameter of G is equal to $2\left\lceil \frac{\sigma(2)}{8} \right\rceil \geq \frac{\sigma(2)}{4}$.

We now prove that G is pairwise stable. First of all, no agent wants to delete an edge of G as, by doing so, the agent would save $\sigma(2)$ by the removal of the edge, but the cost of another edge bought by the same agent would increase from $\sigma(2)$ to $\sigma(n)$. Since $n \geq 5$, we have that $\sigma(n) > 2\sigma(2)$, by Proposition 6.1.

Now we show that no two agents u and v want to add the edge uv. Let i and j be the distances from u and v to the center x of the spider, respectively. W.l.o.g., we assume that $j \geq i$. We divide the proof into two cases according to whether u and v are in the same leg of the spider or not.

If u and v are in the same leg of the spider (we assume that this is the case if $i = 0$), then, by adding uv the edge cost of u would increase by $\frac{1}{2}\sigma(j - i)$, while the distance cost of u would decrease by at most

$$\begin{aligned}
&2\left(\left\lceil\frac{\sigma(2)}{8}\right\rceil - j\right)(j-i-1) + 2\sum_{\ell=1}^{\left\lfloor\frac{j-i-1}{2}\right\rfloor}\ell \\
&\leq 2\left(\frac{\sigma(2)}{8} + 1 - j\right)(j-i) + \frac{1}{4}(j-i)^2 \\
&\leq \frac{\sigma(2)}{4}(j-i) \leq \frac{1}{2}\sigma(j-i),
\end{aligned}$$

where the last inequality holds by Proposition 6.1, while the last but one inequality holds because $j \geq 2$. Thus, u would not agree to adding the edge uv.

If u and v are in different legs of the spider, then, by adding uv the edge cost of u would increase by $\frac{1}{2}\sigma(i+j)$, while the distance cost of u would decrease by at most

$$\mu := 2\left(\left\lceil\frac{\sigma(2)}{8}\right\rceil - j\right)(j+i-1) + 2\sum_{\ell=1}^{\left\lfloor\frac{j+i-1}{2}\right\rfloor}\ell.$$

When $i = j = 1$, the value μ is upper bounded by $\frac{\sigma(2)}{4}$, and therefore it is not convenient for u to add the edge uv. When $j \geq 2$, i.e., $i + j \geq 3$, the value μ is upper bounded by

$$\begin{aligned}
\mu &\leq 2\left(\frac{\sigma(2)}{8} + 1 - j\right)(j+i-1) + \frac{1}{2}(j+i-1)^2 \\
&\leq \frac{\sigma(2)}{4}(j+i) \leq \frac{1}{2}\sigma(i+j),
\end{aligned}$$

where the last inequality holds by Proposition 6.1. Hence, u would not agree on adding uv. The claim follows. ■

The pairwise stable networks of Lemma 6.9, depicted in Figure 6.4, asymptotically reach the upper bound for the diameter of pairwise stable networks. Moreover, they allow us to prove asymptotically matching lower bounds to the PoA for the class of 2-edge-connected networks.

▶ **Theorem 6.10.** The PoA of SNCG is in $\Omega\left(\frac{\sigma(n)}{n\max\{\sigma(2),n\}}\right)$. For the class of 2-edge-connected networks the PoA is in $\Omega\big(\min\{\sigma(2),n\}\big)$. ◀

Proof. First we prove the lower bound for the class of 2-edge-connected networks. By Lemma 6.9, there is a pairwise stable network with $n = \Omega(\sigma(2))$ nodes and social cost in $\Omega\big(\sigma(2)n^2\big)$. Now, if we assume that $\sigma(2) > 2$, then, by Theorem 6.6, we have that $\mathbf{F}_n$ is a social optimum. The social cost of $\mathbf{F}_n$ is at most

$$2n(n-1) + (\sigma(2) - 2)\frac{3}{2}(n-1) = O\big(n^2 + \sigma(2)n\big).$$

Therefore, the PoA is in $\Omega\big(\min\{\sigma(2), n\}\big)$.

Concerning the lower bound for the general case, consider the modified fan graph $\mathbf{F}'_n$ from Theorem 6.7 for even n. This network is pairwise stable for $\sigma(2) \geq 2$ and has a social cost of $O\big(\sigma(2)n + \sigma(n)\big)$. Thus, the PoA is in $\Omega\Big(\frac{\sigma(n)}{n\max\{\sigma(2),n\}}\Big)$. ■

We conclude this section by showing bounds for the PoS.

▶ **Theorem 6.11.** The PoS of the SNCG when $\sigma(2) \leq 2$ or n is odd is 1. The PoS of the SNCG when $\sigma(2) > 2$ and n is even:

- at most $\frac{11}{8}$ if $\sigma(3) \geq 6$ and $\sigma(2) \leq \frac{n}{2} - 4$;
- at most $\frac{17}{12}$ if $\sigma(2) \geq \frac{2n}{3}$;
- $O\left(\frac{\sigma(n)}{n\max\left\{\sigma(2),n\right\}}\right)$, otherwise.

◀

Proof. For the cases in which $\sigma(2) \leq 2$ or n is odd we have that the PoS is 1, since, from Theorem 6.6 and Theorem 6.7, there always exists a pairwise stable network which is also a social optimum.

It remains to prove the theorem statement for the case where n is even and $\sigma(2) > 2$. We observe that the fan graph $\mathbf{F}_n$ is not pairwise stable in this case as the node of degree 3 can remove the edge towards a node of degree 2. An upper bound for the PoA is delivered by the modified fan graph $\mathbf{F'}_n$ from the proof of Theorem 6.7. Then the PoS ratio is at most

$$\frac{\mathrm{SC}(\mathbf{F'}_n)}{\mathrm{SC}(\mathbf{F}_n)} = \frac{(n-2)\frac{3}{2}\sigma(2) + \sigma(n) + 2n^2 - 5n + 4}{(n-2)\frac{3}{2}\sigma(2) + 2\sigma(2) + 2n^2 - 5n + 2}$$

$$= O\left(\frac{\sigma(n)}{n \max\left\{\sigma(2), n\right\}}\right).$$

Unfortunately, the above upper bound is infinitely large for a sufficiently large cost of a bridge $\sigma(n)$. However, under some additional assumptions it is possible to show that the PoS is constant.

We will show that for $\sigma(3) \geq 6$ and $\sigma(2) \leq n/2$, there are pairwise stable networks with no bridges, small diameter and containing 2-edges only.

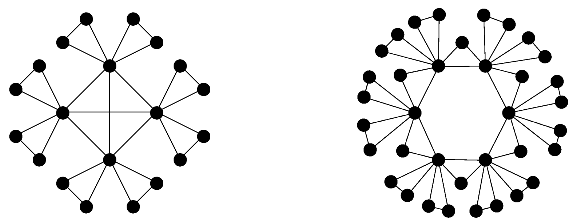


Figure 6.5: Two instances of the pairwise stable networks. The left figure shows an example of pairwise stable with 20 nodes for the case $\sigma(2) \leq \frac{n}{2}$ and $\sigma(3) \geq 6$. On the right side is an example with 36 nodes for the case $\sigma(2) \geq \frac{2n}{3}$.

Consider a clique $\mathbf{K}_4$. Connect each node of the clique with $\frac{n-4}{8}$ edge-disjoint triangles (if $n \bmod 8 \neq 0$, we uniformly distribute triangles among nodes of the clique such that the number of incident triangles differs by at most 1 among any two nodes of the clique). See Figure 6.5 (left) for an illustration of the construction.

We will now prove that this network is pairwise stable. Consider an agent in the clique. She cannot deviate from her current strategy because deleting any of her incident edges from the clique increases her distance to at least $1 + \frac{n-4}{4} - 2$ nodes by 1 and decreases her edge cost by $\frac{\sigma(2)}{2} \leq \frac{n}{4} - 2$. The addition of any edge to a node of a triangle would cost $\frac{\sigma(2)}{2}$ and would improve her distance by only 1. No agent located at a node of a triangle can delete any of her incident edges because it creates a high-cost bridge. The addition of an edge with a clique is also not profitable for the agent from the clique as we observed earlier. Also, the addition of an edge to a node from another triangle either costs $\sigma(2)/2$ and improves the distance cost by only 1, or it costs $\sigma(3)/2$ and improves the distance cost by 3. Hence, we conclude that the construction is pairwise stable.

The social cost of the construction asymptotically approaches the value $\frac{3}{2} \cdot$

$(n-4)\sigma(2)+6\sigma(2)+\frac{11}{4}n^2$, while the social cost of a fan graph is $(n-2)\frac{3}{2}\sigma(2)+2\sigma(2)+2n^2-5n+2$, i.e., the PoS goes to $\frac{11}{8}$ when n goes to infinity.

Now we consider the case when $\sigma(2) > \frac{2n}{3}$. We construct the following pairwise stable network G. Consider a cycle of length 6 consisting of edge-disjoint triangles. Each second node of the inner cycle is connected with $\frac{n-12}{12}$ edge-disjoint triangles (if $n \bmod 12 \neq 0$, we uniformly distribute the triangles among nodes of the cycle such that the number of triangles differs by at most 1). See Figure 6.5 (right) for an illustration of the construction.

We will show that G is pairwise stable for $\sigma(2) \geq 2n/3$. We start with the edge deletions. Clearly no agent can delete any edge from a triangle outside of the central 6-cycle because it creates a bridge of cost $\sigma(n)$. Also, an agent located in the central 6-cycle cannot delete any of her edges since it either creates a bridge or increases her cost of another incident edge from $\sigma(2)$ to $\sigma(5)$. In the second case, the deletion of the edge decreases the agent's edge cost by $\sigma(2)$ and at the same time the edge cost is increased by $\sigma(5)/2 \geq \sigma(4)/2 \geq \sigma(2)$ (by Proposition 6.1 and monotonicity of σ-function) since the edge cost of another edge is increased by this deletion.

Now we will show that no pair of non-neighboring agents u, v wants to add a new edge uv. Note, there are two types of nodes in the central 6-cycle: nodes of degree 2 (outer cycle), and nodes of degree $4+\frac{n}{6}-2$ (inner cycle). Let u be a node in the cycle with degree $\left(2+\frac{n}{6}\right)$. If v is a similar node of the same degree, then one of two cases holds:

- node v is at distance 2 from u. Then u's cost changes by $\frac{1}{2}\sigma(2)-2\cdot\left(\frac{n}{6}-2\right)-3$, since it improves agent u's distance by 1 to two nodes from the cycle and to two high-degree nodes. This changes agent u's cost by $\frac{1}{2}\sigma(2)-\frac{n}{3}+1 > 0$, since $\sigma(2) \geq 2n/3$. Thus, the edge addition is not profitable for agent u.

- node v is at distance 3 from u. Then u's cost changes by $\frac{1}{2}\sigma(3)-2\cdot\left(\frac{n}{6}-2\right)-4$, since u decreases her distance to only three nodes in the cycle and only one high-degree node by 2. The change of u' cost is $\frac{1}{2}\sigma(3)-\frac{n}{3} \geq 0$, because $\sigma(3) \geq \sigma(2) \geq 2n/3$. Hence the move is not profitable for u.

If v is any other node in the network, i.e., any 2-degree node, then creating the edge uv would cost at least $\sigma(2)/2$ and decrease the distance to at most $\frac{n}{6}$ nodes. Hence it is enough to have $\sigma(2) \geq \frac{n}{3}$ to prevent u from creating any edge with a 2-degree node. Therefore, there is no improving edge addition between a $\left(2+\frac{n}{6}\right)$-degree node and any other node in G.

Now we consider 2-degree nodes. Let u be a 2-degree node from the central cycle. As we know from the previous case, v cannot be a $\left(2+\frac{n}{6}\right)$-degree node. If v is another 2-degree node in the central cycle, the addition of the edge (u,v) significantly decreases the distance to other nodes only if v is at distance 3 or 4 from u. In the first case, u pays $\frac{1}{2}\sigma(3)$ and improves her distances by at most $4+\frac{n}{6}-2$. Since $\frac{n}{6}+2<\frac{n}{3}<\frac{1}{2}\sigma(2)\leq\frac{1}{2}\sigma(3)$, this move is not profitable for u. In the second case, edge (u,v) costs $\frac{1}{2}\sigma(4)$ and improves her distance cost by at most $2\cdot\left(\frac{n}{6}-2\right)+5$. Since $\frac{1}{2}\sigma(4)\geq\sigma(2)\geq\frac{2n}{3}>\frac{n}{3}+1$, this move is not profitable.

Analogously, if v is a 2-degree node outside of the central cycle, then the maximum profit u achieves if v is at distance 4 from u. This move costs $\frac{1}{2}\sigma(4)$ for agent u and decreases her distance to $\frac{n}{6}-3$ nodes by 1 and to three other nodes by 5 in total. Since $\frac{1}{2}\sigma(4)\geq\sigma(2)>\frac{n}{6}-2$, the edge will not be added.

Finally, let u be a 2-degree node that is not in the central cycle, i.e., is a 2-degree node in a vane. Clearly, creating an edge with any other 2-degree node connected to the same center is not profitable for u. Consider a 2-degree node v in a vane at distance at least 3 from u. If the distance between u and v is 3, the addition of uv improves the distance to at most two nodes but costs $\frac{1}{2}\sigma(3)$, and therefore is not profitable. If v is at distance 4 from u, the addition of the edge uv costs $\frac{1}{2}\sigma(4)\geq\sigma(2)\geq\frac{2n}{3}$ and improves the distance cost by at most $\frac{n}{3}+2<\frac{2n}{3}$. If v is at distance 5 from u, the addition of uv costs $\frac{1}{2}\sigma(5)>\sigma(2)\geq\frac{2n}{3}$ and improves the distance cost by at most $2\cdot\left(\frac{n}{6}-4\right)+11<\frac{2n}{3}$. Since we checked all possible agents' improving moves, G is pairwise stable.

Network G provides the following upper bound to the PoS

$$\frac{\mathrm{SC}(G)}{\mathrm{SC}(\mathbf{F}_n)}=\frac{\frac{3\cdot 6}{2}\left(\frac{n}{6}-2\right)\sigma(2)+6\sigma(2)+\frac{17}{6}n^2-13n-258}{(n-2)\frac{3}{2}\sigma(2)+2\sigma(2)+2n^2-5n+2}\leq\frac{17}{12}.$$ ■

6.6 Dynamics of the SNCG

So far we have considered the SNCG as a one-shot game, i.e., we only have specified the strategy space of the agents and then focused on analyzing the equilibria of the game. In this section we focus on a more constructive sequential view of the game. As our goal is to mimic real-world social networks, we want to study the process of how such networks evolve over time. For this, we consider some initial network and then we activate the agents sequentially. An active agent will try to decrease her current cost by adding (jointly with another agent) or deleting an edge in the current network. If this process converges to a state where no agent wants to add or delete edges, then a pairwise stable network is found. Hence, such so-called improving move dynamics are a way for actually finding equilibrium states of a game. Such dynamics are guaranteed to converge if and only if the strategic game has the *finite improvement property (FIP)*, i.e., if from any strategy vector any sequence of improving strategy changes must be finite. This is equivalent to the game being an ordinal potential game [MS96]. We start with the negative result that the convergence of improving move dynamics is not guaranteed for the SNCG.

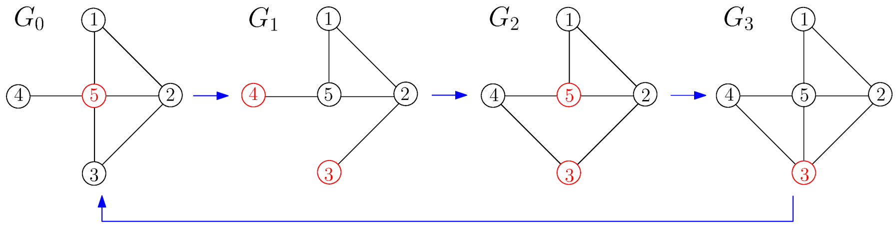


Figure 6.6: A cyclic sequence of improving moves. The active agent or pairs of agents in each step are highlighted.

▶ **Theorem 6.12.** The SNCG does not have the FIP. ◀

Proof. Let σ be any function such that $2 < \sigma(2) < \frac{1}{2}\sigma(3) + 1$ and $\sigma(3) < \frac{1}{2}\sigma(5) + 2$.[8] Consider the cyclic sequence of improving moves depicted on Figure 6.6 in which G_i and $G_{(i+1) \mod 4}$ differ by exactly one edge. We will show that each step of the cycle is an improving response.[9]

$G_0 \to G_1$: agent 5 deletes the edge towards agent 3 because her edge cost decreases by $\frac{1}{2}\sigma(2) > 1$ while her distance cost increases by 1.

$G_1 \to G_2$: the edge between agents 3 and 4 is created. For each of the two agents 3 and 4, the distance cost decreases by 2 while the edge cost increases by $\sigma(3) - \frac{1}{2}\sigma(5) < 2$. In fact, in G_1 both agents pay their share for a bridge of cost $\sigma(5)$ while in G_2 both agents pay their share for two edges of cost $\sigma(3)$ each.

8 Observe that there are several functions σ that satisfy these additional constraints (for example $\sigma(x) = \frac{3}{2}x$).

9 It actually holds that each improving move is also a best response.

$G_2 \longrightarrow G_3$: the edge between agents 3 and 5 is created. In fact, for agent 3 the edge cost varies by $\frac{3}{2}\sigma(2) - \sigma(3) \leq 0$ while the distance cost decreases by 1. For agent 5 the edge cost varies by $2\sigma(2) - \sigma(2) - \frac{1}{2}\sigma(3) = \sigma(2) - \frac{1}{2}\sigma(3) < 1$ while the distance cost decreases by 1.

$G_3 \longrightarrow G_0$: agent 3 deletes the edge towards agent 4 because her edge cost decreases by $\frac{1}{2}\sigma(2) > 1$ while her distance cost increases by 1. ■

The above negative result for the sequential version of the SNCG should not be overrated. In fact, when simulating the sequential process it almost always converges to a pairwise stable network. We will now discuss such simulations.

6.6.1 Experimental Results

We will illustrate that starting from a sparse initial network, the sequential version of the SNCG converges to a pairwise stable network with real-world properties, like low diameter, high clustering and a power-law degree distribution. We will measure the clustering with the *average local clustering coefficient (CC)*, that is a commonly used measure in Network Science [Bar16]. The clustering coefficient is the probability that two randomly chosen neighbors of a randomly chosen node in the network are neighbors themselves. More formally, let $\deg(v)$ denote the degree of v in G and let $\Delta(v)$ denote the number of triangles in G that contain v as a node. The *local clustering coefficient* $CC(v)$ of node v in G is the probability that two randomly selected neighbors of v are neighbors, i.e., $CC(v) := \frac{2\Delta(v)}{deg(v)(deg(v)-1)}$ if $deg(v) \geq 2$, and 0 otherwise. Clearly, $0 \leq CC(v) \leq 1$. The CC of a network G with n nodes is the average of the local clustering coefficients over all nodes v, i.e., $CC(G) = \frac{1}{n}\sum_{v \in V} CC(v)$.

We will also illustrate power-law degree distributions via log-log plots and a comparison with a perfect power-law distribution.

For all experiments[10] we choose $\sigma(x) = 2\lfloor \log_2(n) \rfloor \cdot x^\alpha$, where $n \in \mathbb{N}$ (the number of agents) and $\alpha \in \mathbb{R}_{\geq 1}$ (the exponent) are input parameters. Clearly, this function satisfies all constraints we have in the definition of the game, i.e., it is convex, monotone, and $\sigma(0) = 0$.

Note that by Theorem 6.4 the upper bound for the diameter of pairwise stable networks is $\sigma(2) + 2$ and thus we have to define $\sigma(2)$ to be growing with n to avoid a constant diameter. Using $\sigma(x) = 2\log_2(n) \cdot x^\alpha$ as a proof-of-concept

10 The source code we used can be found at https://github.com/melnan/distNCG.git.

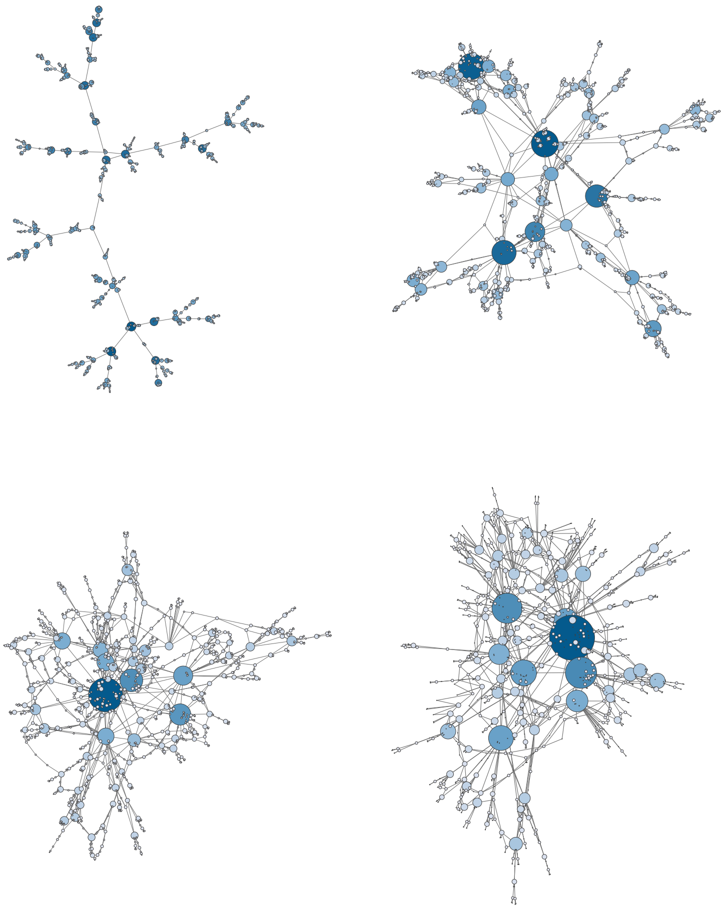


Figure 6.7: Snapshots of networks obtained by the iterative best move dynamic starting from a random spanning tree with $n = 1000$ and $\alpha = 3$. Each plot from the top left to the bottom right shows the current network after 1000 steps each. The top left plot shows the initial tree; the bottom right plot shows is the final pairwise stable network. The size of the nodes is proportional to the node degrees.

ensures a diameter upper bound of $O(\log n)$ that is in line with the observed diameter bounds in many real-world networks [Bar16]. We emphasize that also other edge cost functions with similar properties yield similar results.

In each step of our simulations one agent is activated uniformly at random and this agent then performs the best possible edge addition (jointly with the other endpoint if the respective agent agrees) or edge deletion. If no such move exists then the agent is marked, otherwise the network is updated, and all marked agents become unmarked and we repeat. The process stops when all agents are marked.

In our experiments, we always start from a sparse initial network, i.e., a cycle or a random spanning tree, to simulate an evolving social network, i.e., agents are initially connected with only very few other agents, and the number of new connections grows over time. See Figure 6.7 for showcase snapshots from this process.

Additional experiments starting with sparse Erdös-Renyi random networks support our intuition that the network initialization does not matter as long as the networks are sparse and the average distances are large, i.e., the resulting stable structures have the same structural properties as starting from random trees or cycles. However, for example, starting from a star network yields drastically different results. Moreover, if the initial structure is a fan graph, the algorithm stops immediately since a fan is a stable network as stated in Theorem 6.7. This shows that for the initial networks both sparseness and large average distances are crucial.

Figure 6.8 shows the box-and-whiskers plot for the average clustering coefficient of the pairwise stable networks obtained by the algorithm for $n = 1000$ with respect to the value of the power coefficient α. The upper and lower whiskers show the maximal and the minimal average clustering coefficient over 20 runs. The bottom and top of the boxes are the first and the third quartiles; the middle lines are the median values. The plot explicitly shows that pairwise networks generated by the best move dynamic for a polynomial edge-cost function have a high clustering coefficient. The results indicate that the clustering coefficient correlates with the power coefficient α.

Figure 6.9 shows the degree distribution for the resulting pairwise stable networks for $n = 3000$. We supplemented each plot with a plot of a perfect power-law distribution $P(k) \sim k^{-\gamma}$. All our experiments show that the power-law exponent γ is between 2 and 3, which indicates that our generated pairwise stable networks are indeed scale-free.

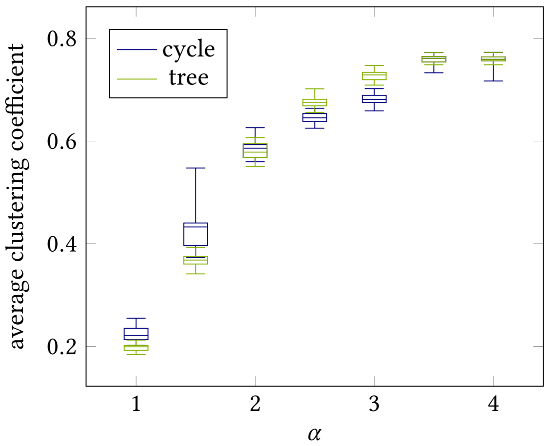


Figure 6.8: Average clustering coefficient of pairwise stable networks obtained by the best move dynamic for $n = 1000$ over 20 runs with $\sigma(x) = 18x^{\alpha}$. Blue: results of the process starting from a cycle; green: starting from a random tree.

Finally, Figure 6.10 illustrates the correlation between the node degree and the local clustering coefficient of nodes with the respective degree. All plots show that the local clustering coefficient is an inverse function of the node degree. In Network Science, a local clustering following the law $\sim k^{-1}$ is considered as an indication of the network's hierarchy that is a fundamental property of many real-world networks[RB03].

Table 6.1 shows a comparison of an experimentally generated network with 3000 nodes for $\alpha = 2$ and $\alpha = 3$, and real-world social networks.

In summary, we conclude from our proof-of-concept experiments that the best move dynamic of the SNCG generates pairwise stable networks that have very similar properties as real-world social networks.

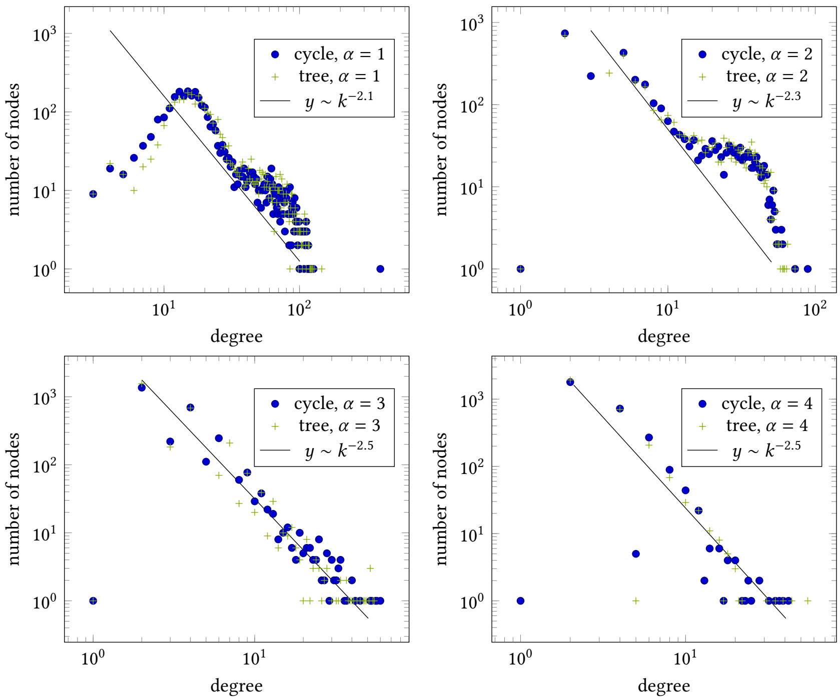


Figure 6.9: Log-log plot of the degree distribution of pairwise stable networks obtained by the best move dynamic for $n = 3000$ with $\sigma(x) = 18x^{\alpha}$. Blue: results for the process starting from a cycle; green: starting from a random tree. Black: a fitted perfect power law distribution.

6.7 Unilateral SNCG

It is natural to assume that pairwise stability is the right solution concept to model social networks since each new connection is formed in both sides' agreement. However, it is an interesting question how this assumption influences the behavior of the model. Moreover, the unilateral link formation can naturally model the weaker social connections, e.g., one link represents that two agents know each other but have no relationship. In this case, one agent can contribute to the connection price while the other agent can later use the link, e.g., to spread her information to other agents.

In this section, we will show that a change of the solution concept can lead to

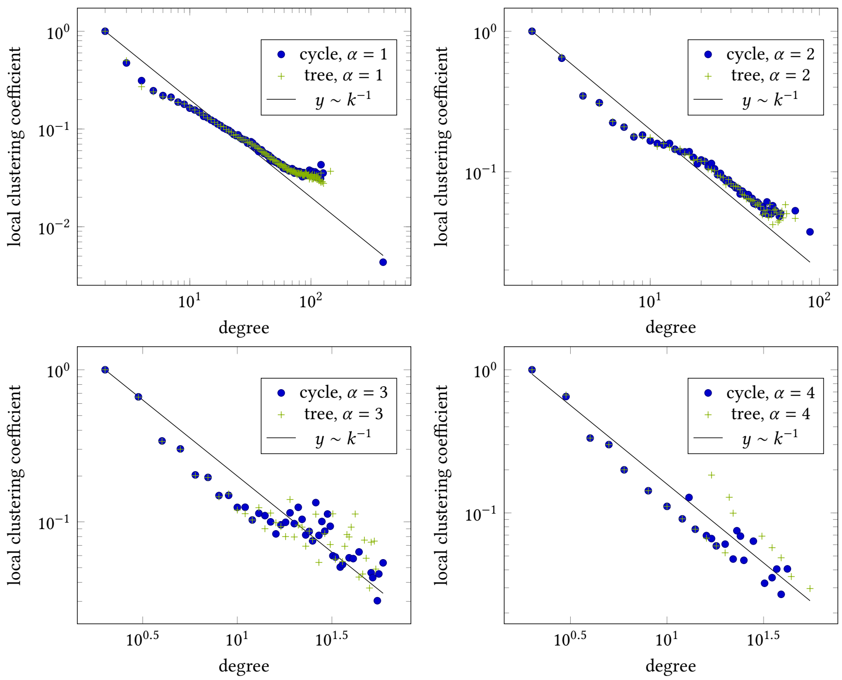


Figure 6.10: Log-log plot of the local clustering coefficient of nodes of a given degree in pairwise stable networks obtained by the best move dynamic for $n = 3000$ where $\sigma(x) = 18x^{\alpha}$. Blue: results starting from a cycle; green: starting from a random tree. Black line: the function $2/k$.

drastically different results for the model. Formally, we consider the unilateral SNCG where each edge is fully paid by its owner only, i.e., by one of its endpoints. Therefore, the cost for an agent u in the network G is the sum of the cost of all edges owned by u and the sum of distances to all other agents:

$$cost(u, G) := \sum_{v \in S_u} c_G(uv) + \sum_{v \in V} d_G(u, v),$$

where S_u is the strategy set of u in G.

As equilibrium concepts we consider the Nash equilibrium, the Greedy equilibrium, and the add-only equilibrium. All other definitions, like the PoA and the PoS, change accordingly.

	SNCG, $\alpha = 2$	SNCG, $\alpha = 3$	e-F [LM12]	ADV [RA15]	HAMST [RA15]
$\lvert V\rvert$	3000	3000	4039	2280	1348
$\lvert E\rvert$	18059	6019	88234	5251	6642
Diameter	8	11	8	11	6
avg distance	3.69	5.17	3.69	3.85	3.2
max degree	72	55	1045	148	273
avg degree	12	4.013	43.7	4.61	9.85
avg CC	0.415	0.67	0.617	0.2868	0.54

Table 6.1: Comparison of basic structural properties of pairwise stable networks of the SNCG and real-world social networks. The networks **e-F (ego-Facebook)**, **ADV (ADVOGATO)**, and **HAMST (HAMSTERSTER)** are (snippets of) online social networks.

6.7.1 Computational Hardness

The first result of this section shows the main difference between the unilateral and bilateral versions of the model. Under the pairwise stability concept, the strategy changes are restricted to single edge additions and deletions. Therefore, an agent's best response can be computed in polynomial time, e.g., by a greedy algorithm. In contrast to this, the unilateral version allows an arbitrary change of the agent's strategy, which leads to NP-hardness of the problem.

▶ **Theorem 6.13.** Computing a best response is NP-hard in the unilateral SNCG if $\sigma(2) \geq 1$. For $\sigma(2) < 1$, a best response can be computed in polynomial time. ◀

Proof. First, note that for $\sigma(2) < 1$ it is always profitable for any agent to add any new incident edge at distance two since it costs $\sigma(2)$ and decreases the distance by at least 1. Hence, computing a best response is trivial since any agent would want to connect with all other agents.

For the case $\sigma(2) \geq 1$ we provide a polynomial time reduction from the Minimum Set Cover problem which is NP-hard. The problem is defined as

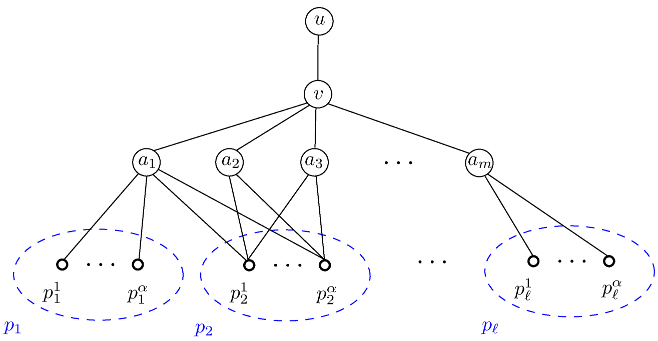


Figure 6.11: Illustration of the construction used in the reduction.

follows. Given a universe $U \subseteq \mathbb{N}$, $|U| = \ell$ and a collection of non-empty subsets $A = \{A_1, \ldots, A_m\}$ such that $\bigcup_{i=1}^m A_i = U$. The problem is to find a minimum number of subsets covering U. To simplify the notation we denote $\alpha := \lceil \sigma(2) \rceil$.

We consider the following instance of the best response problem. Given a graph $G = (V, E)$ such that

$$V = \{u, v\} \cup \{a_1, \ldots, a_m\} \cup \bigcup_{j=1}^{\ell} \bigcup_{k=1}^{\alpha} \{p_j^k\}, \text{ and}$$

$$E = \{uv\} \cup \bigcup_{i=1}^{m} \{va_i\} \cup \bigcup_{i=1}^{m} \bigcup_{p_j \in A_i} \bigcup_{k=1}^{\alpha} \{a_i p_j^k\}.$$

Here each node a_i corresponds to a set A_i, and each set $\{p_j^k\}_{k=1}^{\alpha}$ corresponds to one element p_j of U. Node u is connected only with the node v and an edge uv is owned by v. Each set-node a_i is connected with all α nodes $\{p_j^k\}_{k=1}^{\alpha}$ if and only if the corresponding element of U is in A_i. See Figure 6.11 for an illustration of the construction.

We will show that the best response S_u^* of agent u includes only nodes from $\{a_i\}_{i=1}^m$ corresponding to a solution of the Minimum Set Cover problem.

Note that S_u^* contains only set-nodes. Indeed, a connection with any element-node p_j^k costs at least $\sigma(2)$ and improves distance only to the node p_j^k. Thus, for $\sigma(2) \geq 1$ agent u can either delete the edge up_j^k (if there is a node $a_i \in S_u^*$ such that the corresponding set A_i contains the element p_j), or swap to a set-node a_i which is connected with p_j^k.

Next we show that all nodes in S_u^* corresponds to a set cover, i.e., each node p_j^k has at least one common neighbor with u. If there is a node p_j^k which has no neighboring nodes in S_u^*, then there are at least α nodes at distance 3 from u which are connected to the same set-node a_i. Thus, u can add an edge ua_i and improve her cost by at least $|\sigma(2) - \lceil\sigma(2)\rceil - 1| \geq 1$.

Now we prove that the set of nodes S_u^* corresponds to a Minimum Set Cover problem solution. The cost for the agent u under some strategy S^* is

$$\begin{aligned} cost(u, G(S_u^*, S_{-u})) &= |S_u^*| \cdot \sigma(2) + \sum_{a_i \in S_u^*} d(u, a_i) \\ &\quad + \sum_{j,k} d(u, p_j^k) + \sum_{a_i \notin S_u^*} d(u, a_i) \\ &= |S_u^*| \cdot \sigma(2) + 1 + |S_u^*| + 2l \cdot \lceil\sigma(2)\rceil + 2(m - |S_u^*|). \end{aligned}$$

Consider two strategies S_u^1, S_u^2 such that the corresponding set covers cover all elements of U and $|S_u^1| < |S_u^2|$. Since both strategies correspond to a set cover, the difference of the agent's cost is in the edge cost and distance to the set nodes, i.e., $cost(u, G(S_u^2, S_{-u})) - cost(u, G(S_u^1, S_{-u})) = (|S_u^2| - |S_u^1|) \cdot \sigma(2) - (|S_u^2| - |S_u^1|) \geq 0$ for $\sigma(2) \geq 1$. Therefore, the strategy set with minimum size is the best response, and thus, corresponds to the solution of the Minimum Set Cover problem. ■

6.7.2 Properties of Equilibrium Networks

In this section, we show basic structural properties of equilibria in the unilateral SNCG. Most of the statements are similar to the results we observed for pairwise stable networks.

▶ **Proposition 6.14.** All non-bridge edges of an add-only equilibrium have cost at most $\sigma(3)$ in the unilateral SNCG. ◀

Proof. Let uv be a non-bridge edge in an equilibrium network G of cost at least $\sigma(4)$. W.l.o.g., uv is bought by agent u. Let vw, with $u \neq w$, be another edge incident to v. Such a node always exists because uv is a non-bridge edge. The edge uw is not in G as otherwise the cost of uv would be $\sigma(2)$. If agent u adds also the edge uw, then the cost for having bought two edges uv and uw would be $2\sigma(2)$ that is less or equal $\sigma(4)$ by Proposition 6.1; furthermore, the distance to w would decrease by at least 1. As a consequence, non-stable network has a non-bridge edge whose cost is at least $\sigma(4)$. ■

In the next statement we will show that if the edge cost is low or very high, agents have no incentive to buy costly edges, i.e., all edges are 2-edges. The result hols for all greedy equilibria.

▶ **Proposition 6.15.** In the unilateral SNCG, if $2\sigma(2) \leq \sigma(3)$ or $\sigma(3) > 2(n-3)$, then any non-bridge edge costs $\sigma(2)$ in a GE network. ◀

Proof. Let uv be a non-bridge edge in a GE network G and, w.l.o.g., we assume that u has bought the edge uv. By Proposition 6.14, the cost of uv is at most $\sigma(3)$. If the edge cost $\sigma(3)$, it implies that there is at least one neighbor x of v at distance 2 from u. If u buys an edge ux it improves her distance cost by at least 1, and changes the edge cost by at least $2\sigma(2) - \sigma(3)$. If $2\sigma(2) \leq \sigma(3)$ holds, buying the edge is an improving move.

Now we show that there are induced cycles of size 4 in G if $\sigma(3) > 2(n-3)$. We will prove the claim by showing that every agent in NE owns at most one 3-edge, and thus, for sufficiently large $\sigma(3)$ it is profitable to delete the expensive edge.

Let u be the owner of the edge uv of cost $\sigma(3)$. Let C be a set of nodes $x \in N(u)$ such that the edge cost of the edge ux increases after the deletion of uv, i.e., $C = \{x \in S_u | c_{G-uv}(ux) > c_G(ux)\}$. We will show that $|C| = 0$. Indeed, C includes no endpoints of 2-edges. Moreover, if there is a node $x \in C$ such that $c_G(ux) = \sigma(3)$, then there is a cycle $u - x - w - v$ of size 3 which contains both edges uv and ux. Then u can improve her cost by adding an edge uw because it improves her edge cost by $3\sigma(2) - 2\sigma(3) \leq 0$ and the distance cost by at least 1. Therefore, u owns no edges whose cost depends on the edge uv, that implies $|C| = 0$. It yields that the deletion of the edge uv changes u's cost by at most $-\sigma(3) + 2(n-3)$. Thus, if $\sigma(3) > 2(n-3)$ any 3-edge will be deleted by its owner. ■

In the next proposition we claim that any network which is in NE has a structure of a 2-edge-connected component with spikes[11].

▶ **Proposition 6.16.** In the unilateral SNCG, any add-only equilibrium contains only bridges that form spikes. ◀

Proof. For the sake of contradiction, assume an add-only equilibrium network G having a bridge uv such that both endpoints u and v are not leaves. Clearly, $n \geq 4$,

11 A spike is a bridge edge such that one of its endpoints is a leaf.

otherwise any bridge is a spike. W.l.o.g. let u be an owner of the edge uv. Since v is not a leaf, there is a node $w \in N(v)$ such that u and w are in different connected components in $G - uv$. Then, agent u can buy the edge uw to improve her payoff. Indeed, this move decreases the edge cost by $\sigma(n) - 2\sigma(2)$ and improves the distance cost by at least 1. Therefore, it is an improving move. ■

▶ **Proposition 6.17.** In the unilateral SNCG, the diameter of any add-only equilibrium network is at most $\frac{2}{3}\sigma(2) + 1$. ◀

Proof. Consider an add-only equilibrium network G with diameter $D(G)$. Let u, v be two nodes at distance $D(G)$. Consider a $u - v$ shortest path, and let x be a node on this path such that $d_G(x, u) = 2$. If u buys the edge ux, this increases the edge cost by at most $\sigma(2)$ and decreases the distance cost by at least $D(G) + (D(G) - 3)/2$ because by Propositions 6.14 and 6.16 each edge, except for the spike edges, closes at least one 3- or 4-cycle. Since G is in add-only equilibrium, $D(G) \leq \frac{2}{3}\sigma(2) + 1$. ■

Note that for any agent in the unilateral SNCG, the addition of two edges to two endpoints of an existing edge in the network brings a similar improvement as adding one of the edges in the add-only NCG. Indeed, if two edges close a triangle, then this move always costs $2\sigma(2)$ for the agent. Therefore, some techniques for the original NCG can be applied to our model.

▶ **Theorem 6.18.** For $0.5 \leq \sigma(2) < 0.5n^{1-\varepsilon}, \varepsilon \geq 1/\log n$, the diameter of any add-only equilibrium network in the unilateral SNCG is constant. Namely, the diameter is at most $4.667 \cdot 3^{\lceil 1/\varepsilon \rceil} + 7$. ◀

Proof. Consider a network G which is stable against the new edge additions. Towards a contradiction, assume that the diameter $D(G)$ is non-constant. Define the NCG on the same set of agents and the same initial strategy profile with the edge cost parameter $\alpha = 2\sigma(2)$. By Theorem 10 in [Dem+12] any network G which is stable against single edge additions has constant diameter for $1 \leq \alpha < n^{1-\varepsilon}, \varepsilon \geq 1/\log x$. More precisely, the diameter is at most $4.667 \cdot 3^{\lceil 1/\varepsilon \rceil} + 7$. Thus, if $D(G)$ is not constant, there is an improving edge addition by some agent u in the NCG. Then the agent wins at least the same distance improvement by adding the same edge and an edge to the adjacent endpoint in the SNCG for the same edge price. Hence, we have a contradiction with the assumption that G is in add-only equilibrium. ■

6.7.3 Existence of Equilibria

In this section we prove the existence of equilibrium networks for any feasible σ-function.

▶ **Theorem 6.19.** For $\sigma(2) < 1$ a clique $\mathbf{K}_n$ is the unique NE. ◀

Proof. An agent u that is at distance 2 from v always profits from buying an edge to v as, by doing so, the distance towards v would decrease by 1 while the cost of the edge would be at most $\sigma(2)$. Hence, if $\sigma(2) < 1$ all nodes are in distance 1 from each other in a NE network, i.e., any NE is a clique. ■

▶ **Theorem 6.20.** For $\sigma(2) \geq 1$, a center sponsored star $\mathbf{S_n}$ is a NE. ◀

Proof. It is enough to observe that no leaf u wants to buy an edge towards another leaf v. Indeed, this move costs $\sigma(2)$ and decreases the distance cost by 1 only. ■

Next we prove that the unilateral SNCG for sufficiently high edge cost has a unique stable configuration.

▶ **Theorem 6.21.** If $\sigma(2) > n - 2$ and $\sigma(3) \geq 2\sigma(2)$, the center sponsored star S_n is the unique GE and NE. ◀

Proof. Consider an equilibrium network G. By Proposition 6.15, G contains no 3-edges since $\sigma(3) \geq 2\sigma(2)$. We will prove that there are also no triangles in G. For this we will show that if a triangle exists, there is at least one edge whose deletion does not change the cost of other edges for its owner. That would mean that such an edge can be deleted if its cost $\sigma(2)$ is strictly larger than the increase of the distance cost. Towards a contradiction, we assume that none of the edges can be deleted without an increase in the edge cost. Denote $S(u_1, u_2)$ the set of endpoints of the edges owned by u_1 such that the deletion of the edge u_1u_2 increases its costs, i.e., $S(u_1, u_2) = \{x \in S_{u_1} | c_{G-u_1u_2}(u_1x)) > c_G(u_1x))\}$. Thus, our assumption above means that for any edge $u_1u_2 \in E(G)$, we have $S(u_1, u_2) \neq \emptyset$.

Note that in any triangle u-v-z in the network G there is at least one node, say z, which owns only one edge, say zv, of the triangle. Since $\sigma(3) > 2\sigma(2)$, it follows that $S(z, v)$ contains no endpoints of 3-edges. If a deletion of the edge zv changes the cost of other edges owned by z, then there is at least one edge zx owned by z such that nodes z, v, x form a triangle z-v-x in G. Since by our assumption $S(z,x) \neq \emptyset$, then z also owns an edge $zy \neq zv$ such that z, x, y form a triangle in G. Hence, zv can be deleted without the cost increase for other edges owned by z, i.e., $S(z,v) = \emptyset$, and we have a contradiction. Note that the deletion is profitable for v because it increases the distances to at most $n-2$ nodes by 1. Hence, if $\sigma(3) \geq 2\sigma(2)$ and $\sigma(2) > n-2$, network G contains no cycles.

Finally, we need to show that G has diameter 2. If the diameter is more than 2, there is an edge uv owned by u such that v is not a leaf in G. Then the agent u can buy an edge of cost $\sigma(2)$ and thereby decrease her cost by at least $\sigma(n) - 2\sigma(2) + 1$, and thus, make an improvement. It implies that the center sponsored star is the only stable network. ■

6.7.4 PoA and PoS

Since the definitions of the social cost in the unilateral and bilateral SNCG are identical, Theorem 6.6 holds for the unilateral version as well.

As we proved in Theorem 6.20, the central sponsored star $\mathbf{S_n}$ is a NE, while a clique or a fan network is an optimum by Theorem 6.6. Moreover, by Theorem 6.19, any NE as well an optimum is a clique for low values of $\sigma(2)$. This immediately implies a lower bound for the PoA. We will show that the lower bounds asymptotically meets the upper bound.

▶ **Theorem 6.22.** The PoA in the unilateral SNCG is in $\Theta\left(\frac{\sigma(n)}{\max\{\sigma(2),n\}}\right)$ for $\sigma(2) \geq 2$, and it is in $\Theta\left(\frac{\sigma(n)}{n\sigma(2)}\right)$ for $1 \leq \sigma(2) < 2$. For $\sigma(2) < 1$, the PoA is 1. ◀

Proof. First, we provide a general upper bound for the social cost of equilibrium networks for $\sigma(2) \geq 1$, and then we will compare the result with the social cost of the corresponding optimum with respect to the range of $\sigma(2)$.

Consider an equilibrium network G of maximum social cost for a given number of nodes n. The distance cost of G can be upper bounded by $n(n-1)\left(\frac{2}{3}\sigma(2)+1\right)$ by Proposition 6.17. Since $\sigma(2) \geq 1$, the expression is at most $\frac{5}{3}n(n-1)\sigma(2)$.

Next, we will show an upper bound for the edge cost of G. Let k_i be the number of i-edges in G. By Proposition 6.14, G contains only 2-, 3-, and n-edges. Moreover, if $2\sigma(2) \leq \sigma(3)$, all edges are either 2-edges, or spikes.

Consider the case $2\sigma(2) \leq \sigma(3)$. As we noticed, for any $i \geq 3$, it holds that for $i \neq n$, $k_i = 0$. Clearly, $k_2 \leq \frac{n(n-1)}{2}$, and there can be at most $k_n \leq n-1$ spikes.

Then the edge cost of the network is at most

$$k_2 \cdot \sigma(2) + k_n \cdot \sigma(n) < n^2\sigma(2) + n\sigma(n).$$

In case $2\sigma(2) > \sigma(3)$, the equilibrium network can contain 3-edges, as well as 2- and n-edges. By Mantel's Theorem [Man07], $k_3 \leq \frac{n^2}{4}$. Then G's edge cost is at most

$$\begin{aligned} k_2 \cdot \sigma(2) + k_3 \cdot \sigma(3) + k_n \cdot \sigma(n) &< \frac{n^2}{2}\sigma(2) + \frac{n^2}{4}\sigma(3) + n\sigma(n) \\ &< \frac{n^2}{2}\sigma(2) + \frac{n^2}{2}\sigma(2) + n\sigma(n) = n^2\sigma(2) + n\sigma(n). \end{aligned}$$

Therefore, in any case, the edge cost of G is at most $n^2\sigma(2) + n\sigma(n)$. By the convexity of the sigma function, $n\sigma(2) \leq 2\sigma(n)$ (Proposition 6.1). Then in combination with the upper bound for the distance cost, we get:

$$\begin{aligned} \mathrm{SC}(G) &< n^2\sigma(2) + n\sigma(n) + \frac{5}{3}n(n-1)\sigma(2) < \frac{8}{3}n^2\sigma(2) + n\sigma(n) \\ &\leq \frac{16}{3}n\sigma(n) + n\sigma(n) = \frac{19}{3}n\sigma(n). \end{aligned}$$

To prove the upper bound on the PoA, we need to consider two cases: if $1 \leq \sigma(2) < 2$, and if $\sigma(2) \geq 2$. In the first case, a social optimum is a clique $\mathbf{K}_n$ of cost $n(n-1)(\sigma(2)+1)$ (by Theorem 6.6), then the PoA is at most

$$\frac{\mathrm{SC}(G)}{\mathrm{SC}(\mathbf{K}_n)} \in \mathcal{O}\left(\frac{n\sigma(n)}{n^2\sigma(2)}\right) = \mathcal{O}\left(\frac{\sigma(n)}{n\sigma(2)}\right).$$

In case $\sigma(2) \geq 2$, by Theorem 6.6, $\mathbf{F}_n$ is a social optimum of cost $\Omega(n^2+n\sigma(2)) = \Omega(n \cdot \max\{n, \sigma(2)\})$, then the PoA is at most

$$\frac{\mathrm{SC}(G)}{\mathrm{SC}(\mathbf{F}_n)} \in \mathcal{O}\left(\frac{\sigma(n)}{\max\{n, \sigma(2)\}}\right).$$

Note that both bounds are asymptotically tight because the upper bound for equilibrium networks is achieved on a center sponsored star $\mathbf{S}_n$. Finally, by Theorem 6.19, any equilibrium, as well any optimum, is a clique if $\sigma(2) < 1$, hence, the PoA is 1. ■

▶ **Theorem 6.23.** The PoS is equal to 1 for $\sigma(2) < 1$ and $2 \leq \sigma(2) \leq n-3$. For $1 \leq \sigma(2) < 2$, the PoS equals 2. For $n-2 < \sigma(2) \leq \frac{1}{2}\sigma(3)$, the PoS equals the PoA$\in \Theta\left(\frac{\sigma(n)}{\max\{\sigma(2),n\}}\right)$. ◀

Proof. First, we will show that for $1 \leq \sigma(2) \leq n-3$, a fan network $\mathbf{F}_n$ is a NE. Consider $\mathbf{F}_n$ with a central node u. We assume that u's strategy profile is empty, i.e., u owns no edges in $\mathbf{F_n}$. If n is even, there is a non-central node with three incident edges. We assume that this node owns only one edge towards u.

Consider a triangle $x - u - v$ in $\mathbf{F}_n$. W.l.o.g. x owns two edges xu and xv, i.e., $S_x = \{u, v\}$. Agent x cannot improve her strategy by deleting any of her edges since it creates a bridge of cost $\sigma(n)$ or disconnects the network. Also, any new strategy S'_x with more than two edges costs at least $\sigma(2)(|S'_x| - |S_x|)$ and improves distances only to the new neighbors by 1 each. Since $\sigma(2) \geq 2$, it is not profitable for x to increase the number of owned edges. Also a swap of any edge xu or xv can either increase the distance cost, or can be a neutral move.

Agent v owns only one edge. This edge cannot be deleted because it increases the distance to at least $n-3$ nodes by 1 (since v can have three neighbors). Hence, this move is not an improvement if $\sigma(2) \leq n-3$. Clearly, any better strategy for v includes the edge to the central node u. Since for $\sigma(2) \geq 1$ the addition of any new edge is not profitable, agent v cannot improve on her strategy.

Finally, the central node u owns no edges but is connected with all nodes. Then u cannot change her strategy. Therefore, $\mathbf{F_n}$ is a NE for $1 \leq \sigma(2) \leq n-3$. Hence, $\mathbf{F}_n$ is a social optimum and is in NE for $2 \leq \sigma(2) \leq n-3$. This implies that the PoS is 1 [12]. Moreover, as stated in Theorem 6.19 and Theorem 6.6, a clique $\mathbf{K}_n$ is an equilibrium and optimum network. The first part of the statement follows.

For $1 \leq \sigma(2) < 2$, by Theorem 6.6, the social optimum is a clique, then

$$\text{PoS} = \sup_n \frac{\text{SC}(\mathbf{F}_n)}{\text{SC}(\mathbf{K}_n)} = \sup_n \frac{(n-1)\left(\frac{3}{2}\sigma(2) + 2n - 3\right)}{n(n-1)(\sigma(2)+1)} = 2.$$

Finally, for $n-2 < \sigma(2) \leq \frac{1}{2}\sigma(3)$, by Theorem 6.21, a center sponsored star $\mathbf{S}_n$ is the unique NE. Hence, we have that PoS = PoA $\in \Theta\left(\frac{\sigma(n)}{\max\{\sigma(2),n\}}\right)$. ■

12 Note that if n, the number of agents, is odd, then there is symmetric fan network such that all non-central nodes have degree two. Hence, the deletion of any edge by a non-central node is not profitable when $\sigma(2) \leq n-2$. This implies that the PoS is 1 for $\sigma(2) \leq n-2$.

6.7.5 Game Dynamics

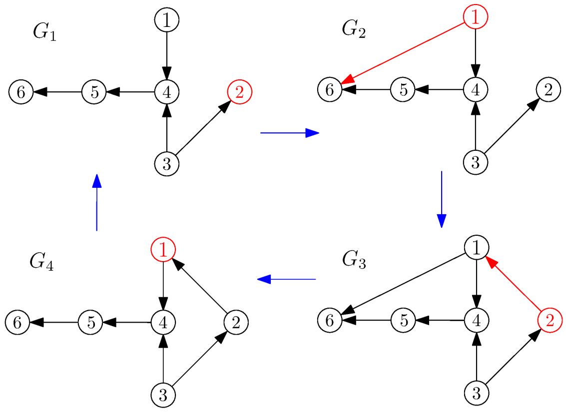


Figure 6.12: Improving response cycle for the distNCG

▶ **Theorem 6.24.** The unilateral SNCG is not a potential game. ◀

Proof. The correctness of the statement can be proven by showing the existence of an improving response cycle, which is a cyclic sequence of networks where in each step, only one agent changes her strategy by applying one of the permitted moves of the game while improving his overall cost. The cycle consists of four consecutive strategy changes G_1 to G_4, shown in Figure 6.12, where two agents take turns adding or deleting edges. We assume $\sigma(x) = x$.

In G_1, we have $cost(1, G_1) = 23$. By adding the edge between 1 and 6, agent 1 can improve her cost by 8 as the cost of the edge between 1 and 4 decreases by 9 and the path to agent 6 shortens. In G_2, the cost of agent 2 is 13. By adding the edge connecting 1 and 2, the cost of agent 2 decreases by 2. In G_3, it holds that $cost(1, G_3) = 13$. Agent 1 can now delete the edge to node 6. Deleting the edge decreases the cost of agent 1 by 1. In G_4, agent 2 has cost 14 and can decrease her cost by 1 by deleting the edge to node 2. The resulting network is G_1. ■

6.8 Conclusion and Discussion

We introduced the SNCG, a promising game-theoretic model of strategic network formation. We emphasize that our model is based on only four simple principles:

(1) agents are selfish, (2) each agent aims at increasing her centrality, (3) new connections are most likely to appear between friends of friends rather than between more remote nodes, and (4) connections are costly. All principles are motivated by modeling real-world social networks.

This chapter's main focus was on the bilateral version of the game where any new edge is formed if it is profitable for both endpoints only, and any edge can be deleted unilaterally. This variant perfectly mimics the formation of connections in social networks. However, there are real-world network formation scenarios where the links between the nodes are unilateral. Therefore, it is also essential to study the unilateral variant of the game where each edge is controlled by its owner only. Moreover, a study of both game variants helps to understand the impact of the chosen solution concept. In the following, we will compare properties of equilibrium networks in the unilateral and bilateral SNCG. An overview of the results for both versions is provided in Table 6.2.

An important common property of both model variants is the dependence of the results on the cost of 2-edges and bridges only. More precisely, the structure and efficiency of selfishly created networks depend on the cost of forming a link between two friends of a common friend, i.e., the cost of forming a triangle. A good illustration of this is the upper bound for the diameter of a stable network G. In both models $D(G) \in \mathcal{O}(\sigma(2))$. However, it is worth noticing that greedy equilibria in the unilateral SNCG have constant diameter for almost any value of $\sigma(2)$ (Theorem 6.21 and Theorem 6.18). In contrast, the bilateral version allows stable networks with high diameter (Lemma 6.9). This phenomenon can be explained by the fact that in the bilateral version, deletions are mostly costly because each node is the owner of all its incident edges, while in the unilateral version, any network of high diameter most probably has an edge whose deletion does not increase the cost of other edges and therefore makes the deletion profitable for its owner.

Another common property of the two versions is a high number of triangles in equilibrium networks. This property looks promising for proving a high clustering property. Unfortunately, both games admit stable networks with the average clustering coefficient going to zero when the number of agents goes to infinity. See Figure 6.13 for an illustration of such networks.

Due to the centrality measure in the agent's cost function, the best stable state (as well as a social optimum) in both versions is a star-like construction, i.e., one central node incident to all other nodes. Certainly, this construction does not mimic real-world networks, but, unfortunately, in the unilateral case, agents cannot achieve any other stable state if the cost of one link is expensive (Theorem 6.21). In contrast, the bilateral version offers a variety of equilibria for any family of sigma function, e.g., the construction from Lemma 6.9 is pairwise stable for high values of $\sigma(2)$ as well as the fan graph.

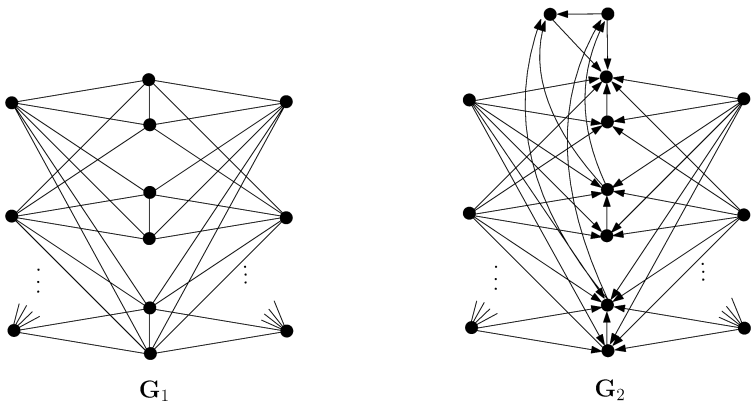


Figure 6.13: Two instances of the game with the low clustering coefficient. The left side network $\mathbf{G}_1$ is pairwise stable for $\sigma(x) := x$. It contains $\frac{n}{2}$ pairwise connected central nodes and $\frac{n}{2}$ side nodes, each of which connected with both endpoints of the central node pairs. The right side network $\mathbf{G}_2$ is identical to $\mathbf{G}_1$ but contains two more nodes. All side nodes are the owner of the edges towards the central nodes. Each central node is either an owner of an edge to its pair or owns two edges to the top nodes. $\mathbf{G}_2$ is in greedy equilibrium for $\sigma(x) := x$. The average clustering coefficient for both networks asymptotically approaches $\frac{1}{n} \to 0$ when $n \to \infty$.

One of the weaknesses of the unilateral game is the potential in forming bridges. By the definition of the game, bridges lead to high social cost, and thus, to the extreme inefficiency of some equilibrium network. Therefore, the star network, which is in equilibrium in the unilateral SNCG, delivers a high value for the PoA in the game. In contrast, the bilateral version allows at most three bridges in a stable network, leading to the much better quality of the selfishly created networks (see the results on the PoA and the PoS in Table 6.2).

As in most of the Network Creation Games, the bilateral and the unilateral SNCG are not potential games. However, the games show radically different behavior in their dynamic. The pairwise stable networks achieved by the best response dynamic show much better properties than the greedy equilibria reached

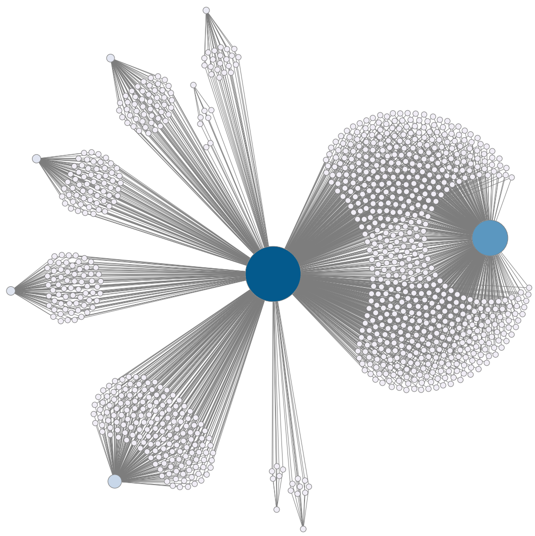


Figure 6.14: An instance of a stable network generated by the best response dynamics in the unilateral SNCG starting from a random tree.

in the same setting. In particular, our empirical results illustrate that the best move dynamic of the bilateral SNCG converges to equilibrium networks that share fundamental properties with real-world networks, like a power-law degree distribution, a high clustering, and a low diameter. Unfortunately, the unilateral version does not show the same results. Figure 6.14 shows a typical greedy equilibrium achieved by the best response dynamics starting from a random tree. It has a star-like construction with one central node connected with all other nodes. Indeed, high degree nodes often have high centrality in a network. Therefore, in the unilateral case, any other node prefers to form one edge to the most central node and one edge to keep the edge cost equal to $\sigma(2)$. In contrast, in the bilateral version, the central node can reject any new incoming edge if it does not improve its centrality.

To summarize, we see the bilateral version as the most natural and promising model of social network formation that yields networks with all core properties of real-world networks. Future work could systematically study the influence of our model parameters on the obtained network features and prove that the sequential network creation process indeed converges to real-world-like networks with high probability. As for the unilateral SNCG, it would be interesting to find a mechanism that prevents the formation of constant diameter equilibria.

Table 6.2: Comparison of the unilateral and bilateral SNCG.

	SNCG	**unilateral SNCG**
equilibrium existence	always exists (Thm.6.7)	always exists (Thm.6.19, 6.20)
complexity	BR can be computed in polynomial time	BR is NP-hard (Thm.6.13)
structural properties of equilibria	each node incident to at most one induced cycle of length ≥ 4 (Prop.6.2); ≤ 3 bridges (Thm.6.3); $D(G) \leq \sigma(2) + 1$ (Thm.6.4);	no induced cycles of length ≥ 4 (Prop.6.14) star is a NE (Thm.6.20, 6.21) $D(G) \leq \frac{2}{3}\sigma(2) + 1$ (Thm.6.17)
PoA	$O\left(\min\{\sigma(2), n\} + \frac{\sigma(n)}{n \max\{\sigma(2),n\}}\right)$, $\Omega\left(\frac{\sigma(n)}{n \max\{\sigma(2),n\}}\right)$ 2-edge-connected networks: $\Theta\left(\min\{\sigma(2), n\}\right)$. for $\sigma(2) < 2$, PoA= 1 (Thm. 6.8, 6.10)	for $\sigma(2) < 1$, PoA= 1 for $\sigma(2) \geq 2$: $\Theta\left(\frac{\sigma(n)}{\max\{\sigma(2),n\}}\right)$; for $1 \leq \sigma(2) < 2$: $\Theta\left(\frac{\sigma(n)}{n\sigma(2)}\right)$; (Thm.6.22)
PoS	for $\sigma(2) \leq 2$ or n is odd: PoS= 1; for $\sigma(3) \geq 6$, $\sigma(2) \leq \frac{n}{2} - 4$: PoS$\leq \frac{11}{8}$; for $\sigma(2) \geq \frac{2n}{3}$: PoS$\leq \frac{17}{12}$; otherwise: $O\left(\frac{\sigma(n)}{n \max\left\{\sigma(2),n\right\}}\right)$ (Thm.6.11)	for $\sigma(2) < 1$, $2 \leq \sigma(2) \leq n-3$: PoS= 1; for $1 \leq \sigma(2) < 2$: PoS= 2; for $n - 2 < \sigma(2) \leq \frac{1}{2}\sigma(3)$: PoS$\in \Theta\left(\frac{\sigma(n)}{\max\{\sigma(2),n\}}\right)$ (Thm.6.23)
Game Dynamics	no FIP (Thm.6.12)	no FIP (Thm.6.24)

7 Conclusions & Outlook

In this thesis, we proposed and analyzed several variants of the Network Creation Game with a particular focus on a realistic scenarios. Our main aim was to understand the behavior and the structure of the equilibria in the game under realistic assumptions. Each chapter already contains a conclusion and overview of future directions for the respectively studied model. So in this chapter, we will continue the discussion from a more general perspective.

We have observed that the Network Creation Game with the incorporated robustness aspect behaves similarly to the non-adversarial version. However, to achieve stability against a random edge failure, agents have to create sufficiently more connections. This leads to a more diverse landscape of optimum and equilibrium states compared to the model without adversary.

Moreover, we have learned that incorporating edge weights and geometry into the Network Creation Game leads to many surprising results. Namely, the inefficiency of the worst-case equilibrium in comparison to a social optimum. Interestingly, this negative result is independent of the geometry and holds for a metric and non-metric case. This implies that centralized coordination is necessary for the outcome of the game to be socially efficient. However, in some cases, the problem is complicated by the fact that computing an optimal centrally designed network is hard. In particular, in Euclidean space, computing the social optimum is related to a minimum weight geometric t-spanner problem which is known to be NP-hard [CC13].

We presented a model with degree-dependent edge cost. Here, we observed that the anti-preferential attachment rule of the edge formation guarantees a constant diameter and efficient equilibria in the game. Hence, to successfully imitate the real-world networks, the model requires additional restrictions to avoid the constant diameter equilibria in the game. For this, we propose to consider the bilateral version or to incorporate a more general edge-cost function, e.g., a linear or polynomial function.

The last introduced model seems the most promising for modeling of real-world social networks. Based on several natural assumptions, i.e., bilateral edge formation and distance-dependent edge cost, the obtained equilibria have many real-world properties. Namely, we proved that the small-worlds property holds for any pairwise stable network for suitable edge-cost functions. Our empirical results showed that the game outcomes have a power-law degree distribution and high clustering.

So far, we have focused on different objectives separately. However, many assumptions can be combined such that the obtained model takes the best properties from the parent models. For example, the geometric version can be incorporated into any other model. In particular, the geometric version in combination with the distance-dependent edge cost. This model combination incorporates the principle that agents are more likely to establish new connections to close acquaintances and prefer connections to the agents with common interests, i.e., to the agents within a close metric distance in the host network.

In Chapter 5 and 6 we aimed to model real-world social networks from two different perspectives: assuming that cost of the connections depends on the popularity of the targeting endpoint or the distance to the target before the connection was established. Both objectives are natural and should be combined in one model.

So far, we studied the local version for the Degree Price Network Creation Game only. However, locality is a crucial realistic assumption that can be considered in combination with other model settings. Indeed, in large networks, agents cannot observe the entire network but only have a partial knowledge of the network structure, usually within their neighborhood. Several definitions of locality have been proposed. The pessimistic scenario where agents have only partial information about the world outside of their neighborhood [Bil+14a; Bil+14b] sounds realistic but leads to dramatically inefficient networks. The more optimistic approach (which has been incorporated in our Social Network Creation Game) is one where the agents have all information about the network but can perform only local steps [CL15]. This scenario guarantees a better quality of the stable networks. However, a complete knowledge about the network structure is unrealistic if we talk about complex networks like the Internet. The core question in this setting is what minimum information about the network is necessary to enable selfish agents to find efficient equilibria.

For the distance-dependent edge cost model, we provided the analysis for a general case when the dependence between the edge price and its cost for an agent is expressed by a convex function. To the best of our knowledge, this generalization has never been studied before for the original Network Creation Game, however, this question is interesting to study. For example, some structural properties can be easily extended to the class of linear functions.

In this thesis, we paid much attention to the edge cost function while the distance cost, representing a centrality measure, keeps unchanged. It would be interesting to study other measures to understand the influence of other forces that motivate the agents to form costly connections.

Given the above list of interesting open problems for future research, we are optimistic that the study of variants of the NCG will go on, and that it will yield many novel insights into the formation of real-world networks.

Bibliography

[AFM09] Nir Andelman, Michal Feldman, and Yishay Mansour. **Strong price of anarchy**. *Games and Economic Behavior* 65:2 (2009), 289–317 (see page 8).

[AJB99] Réka Albert, Hawoong Jeong, and Albert-László Barabási. **Internet: Diameter of the World-wide Web**. *Nature* 401:6749 (1999), 130 (see page 12).

[AK00] Paola Alimonti and Viggo Kann. **Some APX-completeness Results for Cubic Graphs**. *Theoretical Computer Science* 237 (2000), 123–134 (see page 97).

[Alb+14] Susanne Albers, Stefan Eilts, Eyal Even-Dar, Yishay Mansour, and Liam Roditty. **On Nash Equilibria for a Network Creation Game**. *ACM Transactions on Economics and Computation* 2:1 (2014), 1–27 (see pages 10, 13, 35, 42).

[Alo+13] Noga Alon, Erik D. Demaine, Mohammad T. Hajiaghayi, and Tom Leighton. **Basic Network Creation Games**. *SIAM Journal on Discrete Mathematics* 27:2 (2013), 656–668 (see pages 8, 10, 91).

[Àlv+16] Carme Àlvarez, Maria J. Blesa, Amalia Duch, Arnau Messegué, and Maria Serna. **Celebrity Games**. *Theoretical Computer Science* 648 (2016), 56–71 (see page 91).

[ÀM16] Carme Àlvarez and Arnau Messeguè. **Max Celebrity Games**. In: *In Proceedings of the Workshop on Algorithms and Models for the Web Graph (WAW'16)*. Springer. 2016, 88–99 (see page 91).

[ÀM17] Carme Àlvarez and Arnau Messegué. **Network creation games: Structure vs anarchy**. *Computing Research Repository (arXiv)* (2017). URL: https://arxiv.org/abs/1706.09132 (see page 10).

[ÀM18] Carme Àlvarez and Arnau Messegué. **On the Constant Price of Anarchy Conjecture**. *Computing Research Repository (arXiv)* (2018). URL: https://arxiv.org/abs/1809.08027 (see page 10).

[ÀM19] Carme Àlvarez and Arnau Messegué. **On the Price of Anarchy for High-Price Links**. In: *Proceedings of the Conference on Web and Internet Economics (WINE'19)*. Springer, 2019, 316–329 (see page 10).

[AMP18] Anna Adamaszek, Matthias Mnich, and Katarzyna Paluch. **New Approximation Algorithms for (1, 2)-TSP**. In: *Proceedings of the 45 International Colloquium on Automata, Languages and Programming (ICALP'18)*. 2018, 9:1–9:14 (see page 54).

[Ans+08a] Elliot Anshelevich, Anirban Dasgupta, Jon Kleinberg, Eva Tardos, Tom Wexler, and Tim Roughgarden. **The Price of Stability for Network Design with Fair Cost Allocation**. *SIAM Journal on Computing* 38:4 (2008), 1602–1623 (see pages 10, 39, 43).

[Ans+08b] Elliot Anshelevich, Anirban Dasgupta, Éva Tardos, and Tom Wexler. **Near-Optimal Network Design with Selfish Agents**. *Theory of Computing* 4:1 (2008), 77–109 (see pages 39, 43).

[AQ19] Mohammad Ali Abam and Mahnaz Sadat Qafari. **Geometric Spanner Games**. *Theoretical Computer Science* 795 (2019), 398–407 (see page 14).

[Ary+04] Vijay Arya, Naveen Garg, Rohit Khandekar, Adam Meyerson, Kamesh Munagala, and Vinayaka Pandit. **Local Search Heuristics for k-median and Facility Location Problems**. *SIAM Journal on Computing* 33:3 (2004), 544–562 (see pages 52, 54).

[BA99] Albert-László Barabási and Réka Albert. **Emergence of Scaling in Random Networks**. *Science* 286:5439 (1999), 509–512 (see pages 12, 13, 92).

[Bar16] Albert-László Barabási. **Network Science**. Cambridge University Press, 2016 (see pages 1, 12, 140, 142).

[BG00] Venkatesh Bala and Sanjeev Goyal. **A Noncooperative Model of Network Formation**. *Econometrica* 68:5 (2000), 1181–1229 (see pages 10, 43).

[Bil+12] Davide Bilò, Luciano Gualà, Stefano Leucci, and Guido Proietti. **The Max-Distance Network Creation Game on General Host Graphs**. In: *Proceedings of the Conference on Web and Internet Economics (WINE'12)*. LNCS. Springer, 2012, 392–405 (see page 43).

[Bil+14a] Davide Bilò, Luciano Gualà, Stefano Leucci, and Guido Proietti. **Locality-based Network Creation Games**. In: *Proceedings of the ACM Symposium on Parallelism in Algorithms and Architectures (SPAA'14)*. ACM, 2014, 277–286 (see pages 92, 162).

[Bil+14b] Davide Bilò, Luciano Gualà, Stefano Leucci, and Guido Proietti. **Network Creation Games With Traceroute-based Strategies**. In: *Proceedings of the International Colloquium on Structural Information and Communication Complexity (SIROCCO'14)*. Springer. 2014, 210–223 (see pages 14, 92, 162).

[Bil+16] Davide Bilò, Luciano Gualà, Stefano Leucci, and Guido Proietti. **Locality-based Network Creation Games**. *ACM Transactions on Parallel Computing (TOPC)* 3:1 (2016), 1–26 (see page 14).

[BK06] Piotr Berman and Marek Karpinski. **8/7 Approximationn Algorithm for (1,2)-TSP**. In: *Proceedings of the 17 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'06)*. 2006, 641–648 (see page 54).

[BK11] Michael Brautbar and Michael Kearns. **A clustering coefficient network formation game**. In: *Proceedings of the International Symposium on Algorithmic Game Theory (SAGT'11)*. Springer. 2011, 224–235 (see pages 14, 120).

[BKL19] Karl Bringmann, Ralph Keusch, and Johannes Lengler. **Geometric Inhomogeneous Random Graphs**. *Theoretical Computer Science* 760 (2019), 35–54 (see page 13).

[BL20] Davide Bilò and Pascal Lenzner. **On the Tree Conjecture for the Network Creation Game**. *Theory of Computing Systems* 64:3 (2020), 422–443 (see page 10).

[Bro+00] Andrei Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, and Janet Wiener. **Graph Structure in the Web**. *Computer Networks* 33:1 (2000), 309–320 (see page 12).

[CC13] Paz Carmi and Lilach Chaitman-Yerushalmi. **Minimum weight Euclidean t-spanner is NP-hard**. *Journal of Discrete Algorithms* 22 (2013), 30–42 (see page 161).

[Che+19] Yu Chen, Shahin Jabbari, Michael J. Kearns, Sanjeev Khanna, and Jamie Morgenstern. **Network Formation under Random Attack and Probabilistic Spread**. In: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'19)*. 2019, 180–186 (see pages 17, 120).

[Chu+04] Byung-Gon Chun, Rodrigo Fonseca, Ion Stoica, and John Kubiatowicz. **Characterizing selfishly constructed overlay routing networks**. In: *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM'04)*. Vol. 2. IEEE. 2004, 1329–1339 (see page 2).

[CL02] Fan Chung and Linyuan Lu. **The Average Distances in Random Graphs with Given Expected Degrees**. *Proceedings of the National Academy of Sciences of the United States of America* 99:25 (2002), 15879–15882 (see page 13).

[CL15] Andreas Cord-Landwehr and Pascal Lenzner. **Network Creation Games: Think Global - Act Local**. In: *Proceedings of the International Symposium on Mathematical Foundations of Computer Science (MFCS'15)*. 2015, 248–260 (see pages 10, 14, 92, 162).

[CMH14] Andreas Cord-Landwehr, Alexander Mäcker, and Friedhelm Meyer auf der Heide. **Quality of Service in Network Creation Games**. In: *Proceedings of the Conference on Web and Internet Economics (WINE'14)*. Springer. 2014, 423–428 (see pages 43, 91).

[CP05] Jacomo Corbo and David Parkes. **The Price of Selfish Behavior in Bilateral Network Formation**. In: *Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC'05)*. ACM, 2005, 99–107 (see pages 91, 117, 119).

[CS11] Steve Chien and Alistair Sinclair. **Convergence to approximate Nash equilibria in congestion games**. *Games and Economic Behavior* 71:2 (2011), 315–327 (see page 8).

[CS92] Lou Caccetta and W. F. Smyth. **Graphs of Maximum Diameter**. *Discrete Mathematics* 102:2 (1992), 121–141 (see pages 122, 123).

[Dem+09] Erik D. Demaine, Mohammad Taghi Hajiaghayi, Hamid Mahini, and Morteza Zadimoghaddam. **The Price of Anarchy in Cooperative Network Creation Games**. *SIGecom Exchanges* 8:2 (2009), 1–20 (see pages 13, 42, 43, 45).

[Dem+12] Erik D. Demaine, Mohammad Taghi Hajiaghayi, Hamid Mahini, and Morteza Zadimoghaddam. **The Price of Anarchy in Network Creation Games**. *ACM Transactions on Algorithms* 8:2 (2012), 13 (see pages 10, 13, 34, 43, 45, 119, 150).

[DPS20] Umberto De Ambroggio, Federico Polito, and Laura Sacerdote. **On Dynamic Random Graphs with Degree Homogenization via Anti-preferential Attachment Probabilities**. *Physica D: Nonlinear Phenomena* 414 (2020), 132689 (see page 92).

[EB19] S Rasoul Etesami and Tamer Başar. **Dynamic Games in Cyber-Physical Security: An Overview**. *Dynamic Games and Applications* 9:4 (2019), 884–913 (see page 16).

[Ech+20] Hagen Echzell, Tobias Friedrich, Pascal Lenzner, and Anna Melnichenko. **Flow-Based Network Creation Games**. In: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'20)*. ijcai.org, 2020, 139–145 (see pages 18, 120).

[Ehs+15] Shayan Ehsani, Saber Shokat Fadaee, MohammadAmin Fazli, Abbas Mehrabian, Sina Sadeghian Sadeghabad, Mohammad Ali Safari, and Morteza Saghafian. **A Bounded Budget Network Creation Game**. *ACM Transactions on Algorithms* 11:4 (2015), 34 (see page 37).

[EKZ06] Stephan Eidenbenz, Sritesh Kumar, and Sibylle Zust. **Equilibria in Topology Control Games for Ad hoc Networks**. *Mobile Networks and Applications* 11:2 (2006), 143–159 (see page 43).

[Fab+03] Alex Fabrikant, Ankur Luthra, Elitza Maneva, Christos H. Papadimitriou, and Scott Shenker. **On a Network Creation Game**. In: Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC'03). ACM, 2003, 347–351 (see pages iii, v, 2, 5–7, 10, 15, 16, 28, 34, 40, 42, 45–47, 58, 66, 92, 104).

[FK15] Tobias Friedrich and Anton Krohmer. **On the Diameter of Hyperbolic Random Graphs**. In: *Proceedings of the 15 International Colloquium on Automata, Languages and Programming (ICALP')*. Springer. 2015, 614–625 (see page 13).

[FPT04] Alex Fabrikant, Christos Papadimitriou, and Kunal Talwar. **The complexity of pure Nash equilibria**. In: *Proceedings of the 36 Annual ACM Symposium on Theory of Computing (STOC'04)*. 2004, 604–612 (see page 11).

[Fri+17] Tobias Friedrich, Sven Ihde, Christoph Keßler, Pascal Lenzner, Stefan Neubert, and David Schumann. **Efficient Best Response Computation for Strategic Network Formation Under Attack**. In: *Proceedings of the International Symposium on Algorithmic Game Theory (SAGT'17)*. 2017, 199–211 (see page 10).

[GJ02] Michael R. Garey and David S. Johnson. **Computers and Intractability**. Vol. 29. WH Freeman New York, 2002 (see pages 40, 43).

[GK11] Anupam Gupta and Jochen Könemann. **Approximation algorithms for network design: A survey**. *Surveys in Operations Research and Management Science* 16:1 (2011), 3–20 (see page 1).

[Goy+16] Sanjeev Goyal, Shahin Jabbari, Michael J. Kearns, Sanjeev Khanna, and Jamie Morgenstern. **Strategic Network Formation with Attack and Immunization**. In: *Proceedings of the Conference on Web and Internet Economics (WINE'16)*. 2016, 429–443 (see pages 17, 43, 120).

[Gul+15] András Gulyás, József J Bíró, Attila Kőrösi, Gábor Rétvári, and Dmitri Krioukov. **Navigable Networks as Nash Equilibria of Navigation Games**. *Nature Communications* 6 (2015), 7651 (see pages 13, 43, 120).

[HK05] Martin Hoefer and Piotr Krysta. **Geometric Network Design with Selfish Agents**. In: *Proceedings of the Annual International Conference on Computing and Combinatorics (COCOON'05)*. 2005, 167–178 (see page 43).

[Jac10] Matthew O Jackson. **Social and Economic Networks**. Princeton university press, 2010 (see pages 1, 13).

[JLK78] David S Johnson, Jan Karel Lenstra, and AHG Rinnooy Kan. **The Complexity of the Network Design Problem**. *Networks* 8:4 (1978), 279–285 (see page 40).

[JR05] Matthew O Jackson and Brian W Rogers. **The economics of small worlds**. *Journal of the European Economic Association* 3:2-3 (2005), 617–627 (see page 13).

[JW96] Matthew O. Jackson and Asher Wolinsky. **A Strategic Model of Social and Economic Networks**. *Journal of Economic Theory* 71:1 (1996), 44–74 (see pages 6, 8, 91, 118).

[Kar72] Richard M Karp. "Reducibility Among Combinatorial Problems." In: *Complexity of Computer Computations*. Springer, 1972, 85–103 (see page 54).

[KL13] Bernd Kawald and Pascal Lenzner. **On Dynamics in Selfish Network Creation**. In: *Proceedings of the ACM Symposium on Parallelism in Algorithms and Architectures (SPAA'13)*. ACM, 2013, 83–92 (see pages 11, 24, 46, 47).

[Kle00] Jon Kleinberg. **The Small-world Phenomenon: An Algorithmic Perspective**. In: *Proceedings of the Annual ACM Symposium on Theory of Computing (STOC'00)*. ACM, 2000, 163–170 (see page 12).

[Kli11] Lasse Kliemann. **The Price of Anarchy for Network Formation in an Adversary Model**. *Games* 2:3 (2011), 302–332 (see pages 15, 17, 43).

[Kli17] Lasse Kliemann. **The Price of Anarchy in Bilateral Network Formation in an Adversary Model**. *Algorithmica* 77:3 (2017), 921–941 (see page 17).

[Kou+15] Georgios Kouroupas, Evangelos Markakis, Christos Papadimitriou, Vasileios Rigas, and Martha Sideri. **The Web Graph as an Equilibrium**. In: *Proceedings of the International Symposium on Algorithmic Game Theory (SAGT'15)*. Springer. 2015, 203–215 (see pages 13, 120).

[KP99] Elias Koutsoupias and Christos Papadimitriou. **Worst-case Equilibria**. In: *Proceedings of the Symposium on Theoretical Aspects of Computer Science (STACS'99)*. Springer-Verlag, 1999, 404–413 (see page 10).

[Kri+10] Dmitri Krioukov, Fragkiskos Papadopoulos, Maksim Kitsak, Amin Vahdat, and Marián Boguñá. **Hyperbolic geometry of complex networks**. *Physical Review E* 82 (3 2010), 036106 (see pages 13, 120).

[KSS17] Lasse Kliemann, Elmira Shirazi Sheykhdarabadi, and Anand Srivastav. **Swap equilibria under Link and Vertex Destruction**. *Games* 8:1 (2017), 14 (see page 17).

[Len11] Pascal Lenzner. **On Dynamics in Basic Network Creation Games**. In: *Proceedings of the International Symposium on Algorithmic Game Theory (SAGT'11)*. Vol. 6982. Springer, 2011, 254–265 (see pages 11, 91).

[Len12] Pascal Lenzner. **Greedy Selfish Network Creation**. In: *Proceedings of the Conference on Web and Internet Economics (WINE'12)*. Lecture Notes in Computer Science. Springer, 2012, 142–155 (see pages 8, 10, 23, 52).

[LKF05] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. **Graphs over Time: Densification Laws, Shrinking Diameters and Possible Explanations**. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery in Data Mining (SIGKDD'05)*. ACM. 2005, 177–187 (see pages 12, 89).

[LM12] Jure Leskovec and Julian J Mcauley. **Learning to Discover Social Circles in Ego Networks**. In: *Advances in Neural Information Processing Systems*. 2012, 539–547 (see page 146).

[LX12] Xiannuan Liang and Yang Xiao. **Game Theory for Network Security**. *IEEE Communications Surveys & Tutorials* 15:1 (2012), 472–486 (see page 16).

[Man07] Willem Mantel. **Problem 28**. *Wiskundige Opgaven* 10:60-61 (1907), 320 (see pages 132, 153).

[MMM15] Akaki Mamageishvili, Matúš Mihalák, and Dominik Müller. **Tree Nash Equilibria in the Network Creation Game**. *Internet Mathematics* 11:4-5 (2015), 472–486 (see page 10).

[MMO14] Eli A. Meirom, Shie Mannor, and Ariel Orda. **Network Formation Games with Heterogeneous Players and the Internet Structure**. In: *Proceedings of the ACM Conference on Economics and Computation (EC'14)*. ACM. 2014, 735–752 (see pages 43, 91).

[MMO15] Eli A. Meirom, Shie Mannor, and Ariel Orda. **Formation Games of Reliable Networks**. In: Proceedings of the IEEE International Conference on Computer Communications (INFOCOM'15). IEEE. 2015, 1760–1768 (see pages 16, 43, 120).

[Mor+20] Nils Morawietz, Niels Grüttemeier, Christian Komusiewicz, and Frank Sommer. **Colored Cut Games**. In: *Proceedings of the 40th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS'20)*. Ed. by Nitin Saxena and Sunil Simon. Vol. 182. Leibniz International Proceedings in Informatics (LIPIcs). Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2020, 30:1–30:17 (see page 16).

[MS10] Matúš Mihalák and Jan Christoph Schlegel. **The Price of Anarchy in Network Creation Games is (Mostly) Constant**. In: *Proceedings of the International Symposium on Algorithmic Game Theory (SAGT'10)*. Springer, 2010, 276–287 (see page 37).

[MS12] Matúš Mihalák and Jan Christoph Schlegel. **Asymmetric Swap-Equilibrium: A Unifying Equilibrium Concept for Network Creation Games**. In: *Proceedings of the International Symposium on Mathematical Foundations of Computer Science (MFCS'12)*. 2012, 693–704 (see pages 8, 10, 91).

[MS13] Matúš Mihalák and Jan Christoph Schlegel. **The Price of Anarchy in Network Creation Games Is (Mostly) Constant**. *Theory of Computing Systems* 53:1 (2013), 53–72 (see page 10).

[MS96] Dov Monderer and Lloyd S. Shapley. **Potential Games**. *Games and Economic Behavior* 14:1 (1996), 124–143 (see pages 11, 43, 139).

[MSW06] Thomas Moscibroda, Stefan Schmid, and Roger Wattenhofer. **On the Topologies Formed by Selfish Peers**. In: *Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC'06)*. ACM, 2006, 133–142 (see pages 9, 14, 43).

[MSW11] Thomas Moscibroda, Stefan Schmid, and Roger Wattenhofer. **Topological Implications of Selfish Neighbor Selection in Unstructured Peer-to-peer Networks**. *Algorithmica* 61:2 (2011), 419–446 (see pages 14, 43).

[MV44] Oskar Morgenstern and John Von Neumann. **Theory of games and economic behavior**. Princeton university press, 1944 (see page 5).

[MW84] Thomas L. Magnanti and Richard T. Wong. **Network Design and Transportation Planning: Models and Algorithms**. *Transportation Science* 18:1 (1984), 1–55 (see pages 40, 43).

[Mye77] Roger B Myerson. **Graphs and cooperation in games**. *Mathematics of Operations Research* 2:3 (1977), 225–229 (see page 2).

[Mye91] Roger B. Myerson. **Game Theory: Analysis of Conflict**. Harvard University Press, 1991 (see page 5).

[Nas51] John Nash. **Non-cooperative games**. *Annals of Mathematics* (1951), 286–295 (see page 7).

[NBW11] Mark Newman, Albert-László Barabási, and Duncan J Watts. **The Structure and Dynamics of Networks**. Princeton University Press, 2011 (see page 12).

[Nis+07] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani. **Algorithmic Game Theory**. Cambridge University Press, 2007 (see page 5).

[Pap01] Christos H. Papadimitriou. **Algorithms, Games, and the Internet**. In: *Proceedings of the 01 Annual ACM Symposium on Theory of Computing (STOC')*. 2001, 749–753 (see page 13).

[Pap03] Christos H Papadimitriou. **Computational Complexity**. John Wiley and Sons Ltd., 2003 (see page 94).

[RA15] Ryan A. Rossi and Nesreen K. Ahmed. **The Network Data Repository with Interactive Graph Analytics and Visualization**. In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI'15)*. 2015. URL: http://networkrepository.com (see page 146).

[RB03] Erzsébet Ravasz and Albert-László Barabási. **Hierarchical organization in complex networks**. *Physical Review E* 67:2 (2003), 026112 (see page 143).

[Roy+10] Sankardas Roy, Charles Ellis, Sajjan Shiva, Dipankar Dasgupta, Vivek Shandilya, and Qishi Wu. **A Survey of Game Theory as Applied to Network Security**. In: *Proceedings of the 43rd Hawaii International Conference on System Sciences (HICSS'10)*. IEEE. 2010, 1–10 (see page 16).

[Sco67] AJ Scott. **A Programming Model of an Integrated Transportation Network**. In: *Papers of the Regional Science Association*. Vol. 19. 1. Springer. 1967, 215–222 (see page 43).

[Sha+07] Weiping Shang, Pengjun Wan, Frances Yao, and Xiaodong Hu. **Algorithms for Minimum-connected-tuple Dominating Set Problem**. *Theoretical Computer Science* 381:1–3 (2007), 241–247 (see page 24).

[SL08] Yoav Shoham and Kevin Leyton-Brown. **Multiagent systems: Algorithmic, Game-theoretic, and Logical Foundations**. Cambridge University Press, 2008 (see page 5).

[SV08] Alexander Skopalik and Berthold Vöcking. **Inapproximability of pure Nash equilibria**. In: *Proceedings of the 40 Annual ACM Symposium on Theory of Computing (STOC'08)*. 2008, 355–364 (see page 11).

[Van16] Remco Van Der Hofstad. **Random Graphs and Complex Networks**. Vol. 1. Cambridge university press, 2016 (see page 92).

[Wes01] Douglas Brent West. **Introduction to graph theory**. Vol. 2. Prentice hall Upper Saddle River, 2001 (see page 5).

[YY20] Shotaro Yoshimura and Yukiko Yamauchi. **Network Creation Games with Local Information and Edge Swaps** (2020), 349–365 (see pages 14, 92).

List of Publications

Materials from the following publications have been used in this thesis:

References

[1] **Geometric Network Creation Games**. In: *Proceedings of the ACM Symposium on Parallelism in Algorithms and Architectures (SPAA'19)*. 2019, 323–332. Joint work with Davide Bilò, Tobias Friedrich, and Pascal Lenzner.

[2] **Selfish Creation of Social Networks**. In: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI'21)*. 2021, 5185–5193. Joint work with Davide Bilò, Tobias Friedrich, Pascal Lenzner, and Stefanie Lowski.

[3] **On Selfish Creation of Robust Networks**. In: *Proceedings of the International Symposium on Algorithmic Game Theory (SAGT'16)*. Springer Berlin/Heidelberg, 2016, 141–152. Joint work with Ankit Chauhan, Pascal Lenzner, and Martin Münn.

[4] **Selfish Network Creation with Non-Uniform Edge Cost**. In: *Proceedings of the International Symposium on Algorithmic Game Theory (SAGT'17)*. 2017, 160–172. Joint work with Ankit Chauhan, Pascal Lenzner, and Louise Molitor.