Two Coalitional Models for Network Formation and Matching Games

by

Simina Brânzei

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2011

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis comprises of two separate game theoretic models that fall under the general umbrella of network formation games. The first is a coalitional model of interaction in social networks that is based on the idea of social distance, in which players seek interactions with similar others. Our model captures some of the phenomena observed on such networks, such as homophily driven interactions and the formation of small worlds for groups of players. Using social distance games, we analyze the interactions between players on the network, study the properties of efficient and stable networks, relate them to the underlying graphical structure of the game, and give an approximation algorithm for finding optimal social welfare. We then show that efficient networks are not necessarily stable, and stable networks do not necessarily maximise welfare. We use the stability gap to investigate the welfare of stable coalition structures, and propose two new solution concepts with improved welfare guarantees.

The second model is a compact formulation of matchings with externalities. Our formulation achieves tractability of the representation at the expense of full expressivity. We formulate a template of solution concept that applies to games where externalities are involved, and instantiate it in the context of optimistic, neutral, and pessimistic reasoning. Then we investigate the complexity of the representation in the context of many-to-many and one-to-one matchings, and provide both computational hardness results and polynomial time algorithms where applicable.

Acknowledgements

First, I would like to thank my advisor, Kate Larson, for her enthusiasm and unwavering support, for the numerous useful discussions, feedback, guidance, and for giving me the freedom to pursue my own interests, while providing with everything I needed to complete this work.

I thank the readers of this thesis, Shai Ben-David and Alejandro Lopez-Ortiz for reviewing this work and for teaching excellent courses.

I thank all my co-authors, from UW and outside, for making research collaboration a lot of fun. In alphabetical order: Margareta Ackerman, Shai Ben-David, Tariq Elahi, Clara Forero, Ian Goldberg, Nicholas Jennings, Kate Larson, David Loker, Tomasz Michalak, and Talal Rahwan.

I thank Nicholas Jennings for the invited visit to the University of Southampton and Hanna Makaruk for the invited talk to the AMS session at the University of New Mexico. I also thank Hanna for the very insightful correspondence throughout the years.

I thank Margareta Ackerman for her friendship, advice, support, and encouragement, without which I would not be here today. I thank all my other friends and colleagues in graduate school, including Ahmed Ataullah, John Champaign, Lachlan Dufton, Clara Forero, Joshua Gorner, Greg Hines, Richard Jang, Hyunggu Jung, Georgia Kastidou, Ana Krulec, Martin Macko, Lisa Nagy, Lubos Steskal, and all my friends outside the AI group for many happy moments - their friendship has meant a lot to me.

Last but not least, I thank my parents, Luminita and Gheorghe Brânzei, and grandparents, Ioana and Iordache Ciobanu, for their unconditional love and support. This thesis is dedicated to them.

Table of Contents

List of Algorithms

List of Figures

List of Tables

Chapter 1

Introduction

Game theory provides a rich mathematical framework for analysing interactions among self interested parties. Coalitional games study the dynamics of players that interact to accomplish more together than they would individually. The central questions in coalitional game theory are how the players should cooperate and how the payoffs should be divided among the members of a coalition.

In recent years there has been growing interest in the field and many new classes of games have been formalised. One of the main driving forces behind the development of new games was the emergence of the internet, with applications such as internet architecture, routing, peer-to-peer systems, and viral marketing. As Scott Shenker famously said, *"The Internet is an equilibrium, we just need to identify the game"* [40]. Another important line of research that contributed to the development of coalitional games was inspired from social and economic networks ([17], [29]). Social networks influence all aspects of everyday life, such as where people live, work, what music they listen to, and with whom they interact. Early work on social networks was done by Milgram in the 1960's and his experiments indicated that any two people in the world are connected by a path of average length six. This idea is also known as the "six degrees of separation" hypothesis. Since then, researchers observed that many natural networks, such as the web, biological networks, networks of scientific collaboration, exhibit the same properties as the web of human acquaintances. The emergence of online communities such as Facebook, MySpace, and LinkedIn has enabled a much more detailed analysis of real networks.

Several areas of interest can be identified in network formation research. One line of work includes the study of random graph models that exhibit the same characteristics as those observed in real social networks, such as heavy-tailed degree distribution and the small world phenomenon. Among these, we mention the small world model of Kleinberg [31] and preferential attachment models [29]. A separate line of work focuses on strategic network formation, in which the nodes represent self interested entities with different incentives for maintaining or cutting links in the network. Research in network formation literature focuses on understanding issues such as how the structure of the network influences the behaviour of the players, what type of equilibria arise, which players are influential, and which networks are efficient. The study of network formation games has in turn led to important theoretical concepts and new analysis tools.

This thesis comprises of two separate game theoretic models that fall under the general umbrella of network formation games. The first is a coalitional model of interaction in social networks that is based on the idea of social distance, in which players seek interactions with similar others. We rely on a centrality-based measure of utility and analyze the interactions between the players from the point of view of coalitional game theory. The second is a compact model of matchings with externalities. Our formulation achieves tractability of the representation at the expense of full expressivity. We formulate a template of solution concept that applies to games where externalities are involved, and instantiate it in the context of optimistic, neutral, and pessimistic reasoning. Then we investigate the complexity of the representation in the context of many-to-many and one-to-one matchings, and provide both computational hardness results and polynomial time algorithms where applicable.

1.1 Contributions

In this thesis we propose and analyze two separate game theoretic models, that fall under the general umbrella of social and economic network formation.

The first main contribution is formulating and analyzing the class of social distance games, a coalitional model of interaction in social networks. We define the game using a network centrality utility measure, investigate the social welfare and stability of social distance games using the core solution concept, provide a polynomial-time algorithm to approximate optimal social welfare that also satisfies a notion of fairness, and formulate two alternative stability solution concepts with improved welfare guarantees. Besides the technical contribution, social distance games are among the first models to use coalitional game theory for analyzing interactions in network formation.

The second main contribution is formulating the first compact model of externalities for matchings, and formalizing a general stability notion, called Γ stability, which applies to any coalitional game with externalities. We instantiate Γ stability in the context of many-to-many and one-to-one matchings. In many-to-many matchings, we consider optimism, neutrality, pessimism, and contractual stability. In one-to-one matchings, we investigate optimism, neutrality, and pessimism. We provide polynomial-time algorithms in the case of pairwise stability under pessimism and neutrality, and hardness results for optimistic pairwise stability and the core solution concept under all the stability notions considered. We also note that our model of matchings with compact externalities is one of the very first models of compact externalities in the computer science literature.

1.2 Outline

The rest of the thesis is organized as follows. In Chapter 2 we present background material on graph theory and measures of centrality from social networks. Then we formulate the coalitional model of interaction and study its properties using social welfare, the core solution concept, and two refined coalitional solution concepts with improved welfare guarantees.

In Chapter 3 we present some background material on matchings and introduce the compact model of matchings with externalities. We study the model in many-to-many and one-to-one matchings, provide several different properties, complexity results, and polynomial algorithms where applicable.

In Chapter 4, we review our conclusions and discuss ideas for future work.

Chapter 2

A Coalitional Model of Interaction in Social Networks

2.1 Introduction

In this chapter we formulate a model of interaction on social networks using coalitional game theory and the notion of social distance. Our game captures the idea that social networks exhibit homophily, and so players prefer to maintain ties with other players who are close to them. Using social distance games, we analyze the interactions between players on the network, study the properties of efficient and stable networks, relate them to the underlying graphical structure of the game, and give an approximation algorithm for finding optimal social welfare. We then show that efficient networks are not necessarily stable, and stable networks do not necessarily maximise welfare. We use the stability gap to investigate the welfare of stable coalition structures, and propose two new solution concepts with improved welfare guarantees.

2.2 Background

2.2.1 Coalitional Games

Coalitional game theory focuses on how groups of players can accomplish more together than they would individually [45]. One of the main representations in coalitional game theory is that of characteristic function form.

Definition 1 (Characteristic Function Form). *A game in characteristic function form is a pair* (N, v)*, where*

- N *is a set of players*
- v *is a real-valued function on the family of subsets of* N*, such that* $v(S)$ *represents the value that the members of coalition* $S \subseteq N$ *can achieve together.*

Coalitional games can be divided in two broad categories, namely transferable utility (TU) and nontransferable utility (NTU) games. Transferable utility games assume that the players can freely redistribute among themselves the payoffs of the coalition. In that case, $v(S)$ can be specified with a unique real value. In nontransferable utility games, the players are not allowed to make transfers, and so the characteristic function v returns a set of payoff vectors, rather than a single value. Nontransferable utility games are the most general class of games in characteristic function form and contain the class of transferable utility games.

Definition 2 (Nontransferable Utility Game). *A coalitional game with non-transferable utility is a pair* (N, v) *such that*

- N *is a set of players*
- $v : 2^N \rightarrow 2^{\mathbb{R}^{|S|}}$ *associates each coalition* $S \subset N$ *with a set of value vectors* $v(S) \subseteq \mathbb{R}^{|S|}$*, which is the set of different payoffs that coalition* S *can achieve for each of its members.*

The set of players, N, is also known as the *grand coalition*.

Definition 3 (Coalition Structure). *A coalition structure is a partition* $P = (C_1, \dots, C_m)$ *of the grand coalition into coalitions* C_i *such that* $\cup_{i=1}^{m} C_i = N$ *and* $C_i \cap C_j = \emptyset, \forall i \neq j$.

Hedonic games are a general subclass of nontransferable utility games that have been studied in economics [8] and more recently received attention in computer science [20, 10, 4]. In hedonic games, the utility of a player in a coalition is completely determined by the identity of the other members of that coalition.

Definition 4 (Hedonic Game). *A hedonic game is a pair* $(N; (\succeq_i)_{i \in N})$*, where for every player* $i \in N$*,* $\succeq_i$ *is a reflexive, complete, and transitive binary relation on the subsets of* N *that contain* i.

Alternatively, hedonic games can be defined as nontransferable utility games in which for every coalition S, $v(S)$ contains exactly one payoff vector.

Definition 5. *A hedonic game is a nontransferable utility coalitional game (N, v) in which $v(S)$ is a unique vector such that for every player $i \in S$, $v_i(S)$ is the payoff of player i while a member of coalition S.*

Hedonic games require an exponential representation, as every player has to keep a list of length $2^{|N|} - 1$ of preferences over all the possible coalitions in N. Thus any questions related to the complexity of determining outcomes in such games are trivially exponential in the number of players. Several compact representations of hedonic games have been studied, among which additively separable hedonic games.

Definition 6 (Additively Separable Hedonic Game)**.** *An additively separable hedonic game is a hedonic game (N, v), where the payoff of a player i as a member of coalition S can be computed additively as $v_i(S) = \sum_{j \in S} v_i(j)$, where $v_i(j)$ is the amount that player i receives from player j.*

Additively separable hedonic games can also be represented as a weighted directed graph [10], such that every player is a node in the graph, and the weight of the edge (i, j) is the amount that player i receives from j when they are in the same coalition.

The central questions studied in hedonic games and more generally, in coalitional game theory, are related to social welfare and stability. We define social welfare and the required stability concepts for games with coalition structures.

Definition 7 (Social Welfare)**.** *The social welfare of a coalition structure $P = (C_1, \ldots, C_m)$ is*

$$SW(P) = \sum_{i=1}^{m} \sum_{k \in C_i} v_k(C_i)$$

where $v_k(C_i)$ is the utility of player k while a member of coalition C_i.

We will sometimes refer to the utility of player k while a member of coalition C_i in partition P as $v_k(P)$.

One the most important stability solution concepts in coalitional game theory is the core.

Definition 8 (Core)**.** *A coalition structure $P = (C_1, \ldots, C_m)$ is in the* core *if there is no coalition $B \subseteq N$, $B \neq C_i, \forall i = \overline{1, m}$ such that $\forall i \in B$, $v_i(B) \geq v_i(P)$ and for some $j \in B$ the inequality is strict: $v_j(B) > v_j(P)$. If all the inequalities are strict then P is in the* weak core*.*

If coalition structure P is in the core, then P is resistant against group deviations. No coalition can deviate and improve the utility of at least one member, while not degrading the other ones. If B exists, then it is called a *blocking coalition*.

The core is a strong notion of stability that can be empty in many games. Thus we define several weaker notions of stability which are often studied in the context of hedonic games.

Definition 9 (Nash Stability). *A coalition structure $P = (C_1, \ldots, C_m)$ satisfies* Nash stability *if no player i can improve their utility by leaving their current coalition $C(i)$ and either* (1) *moving to a different coalition $C^{'}(i) \in P$, or* (2) *forming the singleton coalition $\{i\}$.*

Individual stability is a stronger variant of Nash stability, which stipulates that a player cannot enter a new coalition unless they receive approval from the members of that coalition.

Definition 10 (Individual Stability). *A coalition structure $P = (C_1, \ldots, C_m)$ satisfies* individual stability *if no player i can improve their utility by leaving their current coalition $C(i)$ and either* (1) *forming the singleton $\{i\}$, or* (2) *moving to a different coalition $C^{'}(i) \in P$, assuming that none of the players in $C^{'}(i)$ degrade upon i's arrival.*

An even stronger version of Nash stability is contractual individual stability, which stipulates that a player can deviate from their current coalition to a new one only if they have the approval of both their current coalition members and the members of the coalition they are entering.

Definition 11 (Contractual Individual Stability). *A coalition structure $P = (C_1, \ldots, C_m)$ satisfies* individual stability *if no player i can improve their utility by leaving their current coalition $C(i)$ and either* (1) *forming the singleton $\{i\}$, or* (2) *moving to a different coalition $C^{'}(i) \in P$, assuming that none of the players in $C^{'}(i)$ degrade upon i's arrival and none of the players in $C(i)$ are hurt by i leaving the coalition.*

A variant of core stability is inner stability, which requires that no group deviations occur within the same coalition.

Definition 12 (Inner Stability). *A coalition structure $P = (C_1, \ldots, C_m)$ satisfies* inner stability *if there exists no blocking coalition $B \subset C_i$, for some $C_i \in P$.*

If a coalition structure exhibits inner stability then any blocking coalition must "cross coalition boundaries" by drawing members from at least two different coalitions. In certain situations, the cost of crossing coalition boundaries can be prohibitive.

2.2.2 Social Networks

We first introduce some standard notions from graph theory that are used in the remainder of this chapter.

Definition 13 (Diameter)**.** *The* diameter *of a graph G is the longest shortest path between any two vertices of G.*

Definition 14 (Induced Subgraph)**.** *An* induced subgraph *of a graph $G = (V, E)$ is a graph $H = (V', E')$, such that $V' \subseteq V$, and for all vertices $s, v \in V'$, the edge $(s, v) \in E'$ if and only if $(s, v) \in E$.*

Centrality measures are some of the most important and frequently used measures of social network structure. Centrality captures the relative importance of a node in the network. Next we present several widely used measures of centrality, namely degree centrality, betweenness centrality, closeness centrality, and eigenvector centrality. As usual in graph theory, we denote a graph by $G = (V, E)$, where V is the set of nodes and E is the set of edges. We denote by $|V| = n$ the number of nodes in the graph.

Definition 15 (Degree Centrality)**.** *The* degree centrality *of a vertex v in graph $G = (V, E)$ is*

$$C_D(v) = \frac{deg(v)}{n-1}$$

Betweenness centrality is a measure that assigns a high score to vertices that are situated on many shortest paths. Intuitively, such nodes are bridges between many other nodes. Such bridges are of special value to individuals, and their significance in social networks has been documented by many studies, such as [24].

Definition 16 (Betweenness Centrality)**.** *The* betweenness centrality *of a vertex v in the graph $G = (V, E)$ is*

$$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

Closeness centrality is a measure that assigns a high score to vertices which tend to be close to every other vertex. Closeness centrality is positively correlated with other measures, such as the degree of a vertex.

Definition 17 (Closeness Centrality)**.** *The* closeness centrality *of a vertex v in the graph $G = (V, E)$ is*

$$C_C(v) = \frac{1}{\sum_{t \in V \setminus \{v\}} d(v, t)}$$

There exist many alternate measures of closeness centrality. Among them, we mention a popular variant introduced by Latora and Machiori [33], which is well-defined on disconnected graphs.

Definition 18 (Latora-Marchiori Closeness Centrality)**.** *The* Latora-Marchiori closeness centrality *of a vertex v in the graph $G = (V, E)$ is*

$$C_C(v) = \sum_{t \in V \setminus \{v\}} \frac{1}{d(v,t)}$$

2.3 The Model

In this section we introduce social distance games and key concepts from the literature. We assume the reader is familiar with basic graph theoretic notions including shortest path, complete graph, and induced subgraph.

Our utility formulation is based on the concept of social distance, which is the number of hops required to reach one node to another, and has become famous since Milgram's study on six degrees of separation. The utility function reflects the principle of homophily, that similarity breeds connection and people tend to form communities with similar others [36]. Homophily has been repeatedly observed in many real world networks, such as marriage, friendship, work, and voluntary organizations. In addition, it can be observed that large groups of people interacting at the same time, such as at a party, easily break into smaller components that ensure high centrality for each of the participants. Our model of interaction predicts some of these phenomena.

Definition 19. *A* social distance game *is represented as a simple unweighted graph $G = (N, E)$ where*

- *$N = \{x_1, \ldots, x_n\}$ is the set of players*
- *The utility of a player x_i in coalition $C \subseteq N$ is*

$$u(x_i, C) = \frac{1}{|C|} \sum_{x_j \in C \setminus \{x_i\}} \frac{1}{d_C(x_i, x_j)}$$

where $d_C(x_i, x_j)$ is the shortest path distance between x_i and x_j in the subgraph induced by coalition C on the graph G. If x_i and x_j are disconnected in C, then $d_C(x_i, x_j) = \infty$.

If we view the inverse social distance as similarity, the utility of a player in a coalition is defined as the *average similarity* of the player to that coalition. A singleton player always receives zero. To be consistent, we define the similarity of a player to himself as zero, and so when computing the utility of a player in a coalition C, we divide by the size of C.

Our utility formulation is related to all the centrality measures introduced (**??**). In particular, it is a variant of closeness centrality, which is well defined on disconnected sets and normalized in the interval $[0, 1]$. Moreover, it is related to several other classical measures from network analysis, such as degree, closeness, betweenness, and eigenvector centrality [23], all of which are used to determine how a node is embedded in the network. Our main goal is to understand the dynamics generated by homophily driven communities, and so this utility function has a number of desirable properties that reflect the sociable nature of the players.

Property 1. *A player prefers direct connections over indirect ones. In general, the player prefers a connection by a factor inversely proportional with the distance to that connection.*

Property 2. *Adding a close connection positively affects a player's utility.*

Moreover, our improvement function reflects diminishing returns. An additional friend benefits everyone, but the added benefit depends on how many friends the player already has.

Property 3. *Adding a distant connection negatively affects a player's utility.*

Property 3 states that players who want to be central in their coalition experience loss in social status because of distant connections.

Property 4. *All things being equal, players favour larger coalitions.*

In this work, we assume coalitions are connected, since a disconnected coalition can improve everyone's utility by splitting into its connected components, and that the input graph is connected, since disconnected graphs can be analyzed componentwise. We emphasize a coalition is defined on the subgraph it induces on the original graph.

2.4 Social Welfare

We are interested in understanding the properties of social welfare maximising structures in social distance games. These structures can be viewed as the best outcomes for the society overall. The next properties follow immediately from the definition of the model.

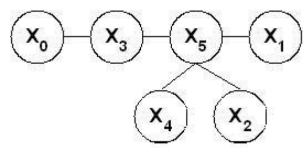

Figure 2.1: In the grand coalition, $u(x_0) = \frac{1}{5}(1+1/2+3\cdot 1/3) = \frac{1}{2}$, $u(x_1) = \frac{1}{5}(1/2+4\cdot 1) = \frac{9}{10}$. In partition $(\{0,3\},\{1,2,4,5\})$, $u(x_0) = u(x_3) = \frac{1}{2}$, $u(x_1) = \frac{1}{4}(1+2\cdot 1/2) = \frac{1}{2}$, $u(x_2) = u(x_4) = \frac{1}{4}(1+2\cdot 1/2) = \frac{1}{2}$, $u(x_5) = \frac{3}{4}$.

Property 5. *On complete graphs, the grand coalition is the only welfare maximising coalition structure.*

Property 6. *The welfare of any coalition structure is bounded by $n-1$. The upper bound is only attained by the grand coalition on complete graphs.*

From Property 5, the grand coalition is welfare-maximising on complete graphs. However, the grand coalition is still welfare maximising on complete bipartite graphs (such as a star), which can be significantly sparser. In addition, complete bipartite graphs have diameter two, and thus the grand coalition is both optimal and gives utility at least 1/2 to each player.

Theorem 1. *The grand coalition maximises social welfare on complete bipartite graphs.*

Proof. A bipartite graph is a triplet $G = (\top, \bot, E)$, where $\top$ and $\bot$ are disjoint sets of vertices, and $E = \{(x,y) | x \in \top, y \in \bot\}$ are the edges. Let $|\top| = m$ and $|\bot| = n$. Let $x \in \top$ and $y \in \bot$. In the grand coalition $K_{m,n}$

$$u(x) = \frac{n + (m-1)/2}{m+n} \quad \text{and} \quad u(y) = \frac{m + (n-1)/2}{m+n}$$

The social welfare of the grand coalition is:

$$SW(K_{m,n}) = \sum_{x\in\top} u(x) + \sum_{y\in\bot} u(y) = \frac{m+n-1}{2} + \frac{mn}{m+n} \tag{2.1}$$

Let $P = \{C_1, \ldots, C_k\}$ be another partition, where $|C_i| = m_i + n_i$, and m_i and n_i are the number of nodes from $\top$ and $\bot$, respectively. Each coalition C_i is a complete bipartite graph, and so it satisfies Equation 2.1. Moreover $m_i \geq 1$ and $n_i \geq 1$, since otherwise C_i

would have welfare zero.

$$SW(P) = \sum_{i=1}^{k} SW(C_i) = \sum_{i=1}^{k} \left(\frac{m_i + n_i - 1}{2} + \frac{m_i n_i}{m_i + n_i} \right)$$

From $\sum_{i=1}^{k} m_i = m$ and $\sum_{i=1}^{k} n_i = n$,

$$SW(P) = \frac{m + n - k}{2} + \sum_{i=1}^{k} \left(\frac{m_i n_i}{m_i + n_i} \right)$$

We claim that $K_{m,n}$ has a better welfare than P:

$$\frac{m + n - k}{2} + \sum_{i=1}^{k} \left(\frac{m_i n_i}{m_i + n_i} \right) < \frac{m + n - 1}{2} + \frac{mn}{m + n}$$

When $k = 2$, the inequality is:

$$\frac{m_1 n_1}{m_1 + n_1} + \frac{m_2 n_2}{m_2 + n_2} \leq \frac{(m_1 + m_2)(n_1 + n_2)}{m_1 + m_2 + n_1 + n_2}$$

Equivalently, $(m_1 m_2 - m_2 m_1)^2 \geq 0$, which holds trivially. Using this fact, the welfare of P can be improved by joining coalitions C_{k-1} and C_k:

$$SW(C_1, \ldots C_{k-2}, \{C_{k-1} \cup C_k\}) > SW(C_1, \ldots, C_k)$$

We can transform P into $K_{m,n}$ by iteratively joining pairs of coalitions. Since social welfare increases with every join, it follows that the grand coalition maximises welfare. ∎

2.4.1 An Approximation of Optimal Welfare

Finding the optimal welfare partition can be shown to be NP-hard on social distance games via a reduction from Partition into Triangles. In this section we give an algorithm to approximate optimal welfare within a factor of two. The algorithm decomposes the graph into connected components, such that each component has diameter less than or equal to two, and no component is a singleton. We call this type of partition a diameter two decomposition of the graph.

Theorem 2. *Diameter two decompositions guarantee to each player utility at least* $1/2$.

Proof. Let G be a graph, P a diameter two decomposition of G, C a coalition in P, and $x_i \in C$ a player. Let a and b the number of players in distance one and two from x_i, respectively. The diameter of C is at most two, hence $|C| = a + b + 1$. The utility of x_i in C is:

$$u(x_i, C) = \frac{a + b/2}{|C|} = \frac{2a + b}{2(a + b + 1)} \tag{2.2}$$

C is not a singleton, thus x_i has at least one direct neighbour in C, and so $a \geq 1$. Using Equation 2.2, we get $u(x_i) \geq 1/2$. ■

The diameter two decomposition is an approximation of optimal welfare that satisfies at the same time a notion of *fairness*: *every* player is guaranteed to receive more than half of their best possible value. In general, welfare maximising and core stable partitions do not necessarily ensure that every player receives at least $1/2$.

Algorithm 1 Fair Approximation of Optimal Welfare

$T \leftarrow$ Minimum-Spanning-Tree(G);
$k \leftarrow 1$;
while $|T| \geq 2$ **do**
    $x_k \leftarrow$ Deepest-Leaf(T);
    $C_k \leftarrow \{\text{Parent}(x_k)\}$;
    for all $y \in \text{Children}(\text{Parent}(x_k))$ **do**
        $C_k \leftarrow C_k \cup \{y\}$;
    end for
    // Remove vertices C_k and their edges from T
    $T \leftarrow T - C_k$;
    $k \leftarrow k + 1$;
end while
// If the root is left, add it to the current coalition
if $|T| = 1$ **then**
    $C_k \leftarrow C_k \cup \{\text{Root}(T)\}$;
end if
return $(C_1, \ldots, C_k)$;

Algorithm 2.4.1 finds a diameter two decomposition of the graph. Let T be a minimum spanning tree of G. Iteratively remove stars from T, starting from the bottom leaves. In each iteration i, let x_i be a leaf of maximum depth. Place x_i and $Parent(x_i)$, the parent of x_i, in the same coalition C_i, together with all the direct children of $Parent(x_i)$. The tree remains connected after each such removal, since otherwise there must have been a leaf of greater depth than x_i. In the last iteration, k, there may be only one node left, namely $Root(T)$, the root of the tree. In that case, add $Root(T)$ to coalition C_k. Note that $Root(T)$ is still distance at most two from all the nodes in C_k.

Computing the minimum spanning tree requires time $O(n)$, where n is the size of the graph. A stack of nodes in decreasing order by depth in the tree can be precomputed in $O(n)$. Each iteration requires popping one element from the stack, visiting its parent together with any potential siblings, and removing the visited nodes from the tree. Since every node is visited exactly once, and the number of operations per node is constant, the algorithm runs in $O(n)$ steps.

Corollary 1. *The optimal welfare partition attains at least $n/2$, where n is the number of vertices.*

Proof. Immediate given that Algorithm 2.4.1 always returns a diameter two decomposition, with welfare at least $n/2$. ∎

A variation of this algorithm can be used to obtain a partition which in addition to giving utility at least 1/2 to each player, also satisfies internal stability. Also, note that the constant diameter holds more generally for partitions satisfying internal stability.

2.5 The Core

Group stability is an important concept in coalitional games. No matter how many desirable properties a coalition structure satisfies, if there exist groups of players that can deviate and improve their utility by doing so, then that configuration can be easily undermined. We investigate properties of the core in social distance games.

Property 7. *On complete graphs, the grand coalition is the only core stable coalition structure.*

Theorem 3. *The grand coalition is core-stable on complete bipartite graphs.*

Proof. We show that the grand coalition $K_{m,n}$ is not blocked. Assume by contradiction that there exists a blocking coalition B. Denote $p = |\top \cap B|$ and $q = |\bot \cap B|$, where $p + q = |B|$, and let $x \in B_\top$ and $y \in B_\bot$. For B to be blocking, it must be the case that

$$u(x, K_{m,n}) \leq u(x, B) \text{ and } u(y, K_{m,n}) \leq u(y, B)$$

This is equivalent to:

$$\frac{2n + m - 1}{m + n} \leq \frac{2q + p - 1}{p + q} \text{ and } \frac{2m + n - 1}{m + n} \leq \frac{2p + q - 1}{p + q}$$

By summing these inequalities it follows that $m+n \leq p+q$, which is a contradiction, since B is a strict subset of $K_{m,n}$. ■

There exist social distance games for which the core is empty, such as the game in Figure 2.1. The grand coalition is blocked by $\{1, 2, 4, 5\}$, partition $(\{0, 3\}, \{1, 2, 4, 5\})$ is blocked by $\{1, 2, 3, 4, 5\}$. Similar examples exist for the weak core. However, if the graph is a tree, the weak core always exists and can be found in polynomial time.

Theorem 4. *Algorithm 2.4.1 returns a weak core partition when the graph is a tree.*

Proof. (sketch) Let $P = (C_1, \dots, C_k)$ be the diameter two decomposition returned by the algorithm. If G is a tree, then every coalition C_i is a star. Note that the leaves of G are not involved in any blocking coalition, because they already receive $1/2$, the maximum amount possible for them. The parent node of any leaf is in coalition with all its direct neighbours, except for its own parent. Since the leaves refuse to block, the parents of the leaves cannot hope to strictly improve utility, and so they will not block either. Iteratively, it can be verified that no vertex can be part of a blocking coalition, and so P is core-stable in the weak sense. ■

The partition returned by Algorithm 2.4.1 also satisfy some weaker notions of stability.

Lemma 1. *The partition returned by Algorithm 2.4.1 satisfies Nash stability and contractual individual stability.*

Next we present a variant of Algorithm 2.4.1 which, in addition to ensuring fairness, satisfies inner stability.

Theorem 5. *Algorithm 2.5 returns a partition a fair approximation of optimal welfare which satisfies inner stability.*

Algorithm 2 Innerly Stable Approximation of Optimal Welfare

$k \leftarrow -1$;
while $N \geq \emptyset$ **do**
    $k \leftarrow k + 1$;
    if there is $z \in N, deg(z) = 1$ **then**
        $x \leftarrow N(z)$; *// x is z's only neighbour*
        $S \leftarrow \{y \in N(x) | deg(y) = 1\}$;
        // S are x's neighbours of degree one, including z
        $C_k \leftarrow \{x\} \cup S$;
    else
        Pick random neighbours $x, y \in N$;
        $S \leftarrow \{z \in N(x) \cap N(y) | deg(z) = 2\}$;
        // S are the nodes connected only to x and y
        $C_k \leftarrow \{x\} \cup \{y\} \cup S$;
    end if
    $N \leftarrow N \setminus C_k$;
end while
return $(C_1, \ldots, C_k)$;

Proof. Let $P = (C_1, \ldots C_k)$ be the coalition structure returned by Algorithm 2.5. By construction, every coalition C_i has one of the following forms:

- *Star*: $C_i = \{x, z_1, \ldots, z_p\}$. C_i clearly gives utility at least 1/2 to all its members. Agent x cannot be involved in a blocking coalition, since they are connected to all the players in C_i, and so any blocking coalition $B \subset C_i$ would decrease x's utility. None of the players z_j can block in the absence of x, since the z_j's are not connected to each other. Thus in this case, C_i satisfies inner stability.

- *Union of two stars with identical leaves but different centers*: $C_i = \{x, y, z_1, \ldots, z_p\}$, where $E(x, y) = 1, E(x, z_j) = 1$, and $E(y, z_j) = 1$, for all $j \in \{1, \ldots, p\}$. Agents x and y are connected to all the members of C_i, and so any blocking coalition would decrease their utility. The players z_j are not connected to each other, and so none of the z_j s can block without x and y. Again, C_i satisfies inner stability. It is immediate to verify that C_i has diameter at most two, and so it gives utility at least 1/2 to all of its players.

At the end of every iteration, Algorithm 2.5 updates the set N to remove from consideration coalition C_i. By induction, the algorithm decomposes the graph into a partition with the required properties. ■

We note that unlike Algorithm 2.4.1, the partition returned by Algorithm 2.5 does not necessarily belong to the core when the graph is a tree. In addition, the coalitions returned by Algorithm 2.5 tend to be smaller than those returned by Algorithm 2.4.1, and so the appropriate algorithm may depend on the situation at hand. Algorithm 2.5 can be implemented to run in $O(n^2)$ using adjacency lists.

2.5.1 Core Partitions are Small Worlds

A small world network is a graph in which most nodes can be reached from any other node using a small number of steps through intermediate nodes. The expected diameter of small world networks is $O(\ln(n))$. Most real networks display the small world property, and examples range from genetic and neural networks to the world wide web [5]. In this model, core stable partitions divide the players into small world coalitions, regardless of how wide the original graph was. We obtained an upper bound of 14 on the diameter of any coalition in the core. The tight bound is likely even lower.

Theorem 6. *The diameter of any coalition belonging to a core partition is bounded by a constant.*

Proof. Let C be a coalition belonging to a core partition. Denote $|C| = c$, D the diameter of C, and x_0, y two players with $d_C(x_0, y) = D$. Divide C into sets $(C_0 = \{x_0\}, C_1, \ldots, C_D)$, and define

- $A = C_0 \cup C_1 \cup \ldots \cup C_{\lfloor \frac{D}{4} \rfloor}$
- $B = C_{\lfloor \frac{D}{4} \rfloor + 1} \cup \ldots \cup C_{\lfloor \frac{D}{2} \rfloor}$
- $\Gamma = C_{\lfloor \frac{D}{2} \rfloor + 1} \cup \ldots \cup C_{\lfloor \frac{3D}{4} \rfloor}$
- $\Delta = C_{\lfloor \frac{3D}{4} \rfloor + 1} \cup \ldots \cup C_D$

where $|A| = \alpha$, $|B| = \beta$, $|\Gamma| = \gamma$, $|\Delta| = \delta$, and $\alpha + \beta + \gamma + \delta = c$.

Let $x_1 \in C_1$. Agent x_0 is connected to all of C_1, and so it is connected to x_1. C is in the core, thus coalition $\{x_0, x_1\}$ is not blocking, and so at least one of x_0, x_1 obtains utility

1/2. Observe that players x_0 and x_1 prefer B, Γ, and Δ to be small, because the players in those sets are distant and contribute to decreasing x_0 and x_1's utility. Bounding the utilities of x_0 and x_1, we obtain:

$$\begin{aligned}\frac{1}{2} \leq \max\{u(x_0), u(x_1)\} &\leq \frac{1}{c}\left(\alpha + \frac{\beta}{\lfloor \frac{D}{4} \rfloor} + \frac{\gamma}{\lfloor \frac{D}{2} \rfloor} + \frac{\delta}{\lfloor \frac{3D}{4} \rfloor}\right) \\ &\leq \frac{1}{c}\left(\alpha + \frac{\beta}{(D-4)/4} + \frac{\gamma}{(D-4)/2} + \frac{\delta}{3(D-4)/4}\right)\end{aligned}$$

Denote $D' = D - 4$ and assume $D \geq 8$:

$$\frac{1}{2} \leq \frac{1}{c}\left(\alpha + \frac{4\beta}{D'} + \frac{2\gamma}{D'} + \frac{4\delta}{3D'}\right)$$

or equivalently,

$$\alpha + \beta + \gamma + \delta \leq 2\alpha + \frac{8\beta}{D'} + \frac{4\gamma}{D'} + \frac{8\delta}{3D'} \tag{2.3}$$

Similarly, let $x_D \in C_D$. Agent x_D can be directly connected only to players in C_D and C_{D-1}. There must exist a path from x_0 to x_D that passes through C_{D-1}, thus there is $x_{D-1} \in C_{D-1}$ neighbour of x_D. Then

$$\begin{aligned}\frac{1}{2} \leq \max\{u(x_D), u(x_{D-1})\} \leq \frac{1}{c}\Bigg(\delta + \frac{\gamma}{D - 1 - \lfloor \frac{3D}{4} \rfloor} + \\ + \frac{\beta}{D - 1 - \lfloor \frac{D}{2} \rfloor} + \frac{\alpha}{D - 1 - \lfloor \frac{D}{4} \rfloor}\Bigg)\end{aligned}$$

Use D' to get simplified (coarser) bounds:

$$\frac{1}{2} \leq \max\{u(x_D), u(x_{D-1})\} \leq \frac{1}{c}\left(\delta + \frac{4\gamma}{D'} + \frac{2\beta}{D'} + \frac{4\alpha}{3D'}\right)$$

or equivalently,

$$\alpha + \beta + \gamma + \delta \leq 2\delta + \frac{8\gamma}{D'} + \frac{4\beta}{D'} + \frac{8\alpha}{3D'} \tag{2.4}$$

Summing Inequalities 2.3 and 2.4:

$$\left(2 - \frac{12}{D'}\right)(\beta + \gamma) \leq \frac{8}{3D'}(\alpha + \delta)$$

that is:

$$(D' - 6)(\beta + \gamma) \leq \frac{4}{3}(\alpha + \delta) \tag{2.5}$$

Finally, take the "perspective" of the middle players. Let $x_{\lfloor \frac{D}{2} \rfloor + 1} \in C_{\lfloor \frac{D}{2} \rfloor + 1}$. There is a path from $x_{\lfloor \frac{D}{2} \rfloor + 1}$ to x_0 that passes through $C_{\lfloor \frac{D}{2} \rfloor}$, thus there exists $x_{\lfloor \frac{D}{2} \rfloor} \in C_{\lfloor \frac{D}{2} \rfloor}$ connected to $x_{\lfloor \frac{D}{2} \rfloor + 1}$.

$$\frac{1}{2} \leq \max\{u(x_{\lfloor \frac{D}{2} \rfloor}), u(x_{\lfloor \frac{D}{2} \rfloor + 1})\} \leq \frac{1}{c}\left(\beta + \gamma + \frac{4\alpha}{D'} + \frac{4\delta}{D'}\right)$$

that is,

$$\alpha + \beta + \gamma + \delta \leq 2\beta + 2\gamma + \frac{8\alpha}{D'} + \frac{8\delta}{D'}$$

if and only if

$$(D' - 8)(\alpha + \delta) \leq D'(\beta + \gamma) \tag{2.6}$$

Multiplying inequalities 2.5 and 2.6, we get:

$$(D' - 8)(D' - 6) \leq \frac{4}{3}D'$$

Solving for integer D' and taking into account that $D \geq 8$, it follows that $D' \leq 10$, and so $D \leq 14$. ■

2.6 Stability Gap

We observed from the empty core game in Figure 2.1 that maximum welfare partitions are not always stable, and from Figure 2.2 that stable partitions do not necessarily maximise welfare. We analyse the loss of welfare that comes from being in the core using the notion of stability gap [10], which is the ratio between the best possible welfare and the welfare of a core stable partition (if it exists). The worst case and best case performance of a core member are measured by the minimum and maximum stability gap, respectively. The stability gap parallels the prices of anarchy and stability [40].

Let P^* be a welfare maximising coalition structure and P^C a member of the core. In

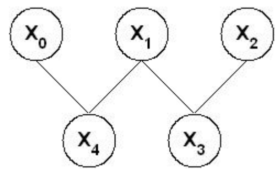

Figure 2.2: The core and welfare maximising partitions do not coincide. The core is $(\{x_0, x_1, x_2, x_3, x_4\})$. Welfare is maximised by $(\{x_0, x_1, x_4\}, \{x_2, x_3\})$ and $(\{x_0, x_4\}, \{x_1, x_2, x_3\})$

the best case, P^C is also a welfare maximiser. Since it is possible that no core member attains the welfare of P^*, the quantity

$$\text{Gap}_{\min}(G) = \frac{SW(P^*)}{\min_{P \in \text{Core}(G)} SW(P)}$$

measures the worst case ratio. We show the worst case bound for the stability gap of social distance games is $\Theta(\sqrt{n})$. In order to get the bound, we use a few lemmas.

Lemma 2. *If a player has utility greater than* $1/2$ *in a coalition, then all its direct neighbours from the same coalition have utility at least* $1/4$.

Proof. Let C be a coalition and $x_i, x_j \in C$ two neighbouring players such that $u(x_i, C) > 1/2$, and $x_j \in C$ a direct neighbour of x_i. The utilities of x_i and x_j are

$$u(x_i, C) = \frac{1}{|C|}\left(1 + \sum_{z \in C \setminus \{x_i, x_j\}} \frac{1}{d(x_i, z)}\right) > \frac{1}{2}$$

and

$$u(x_j, C) = \frac{1}{|C|}\left(1 + \sum_{z \in C \setminus \{x_i, x_j\}} \frac{1}{d(x_j, z)}\right)$$

For every $z \neq x_i, x_j$, the shortest path distance between x_j and z passes, in worst case, through x_i. Thus the inequality $d(x_j, z) \leq d(x_i, z) + 1$ holds, and we can bound the utility of x_j as follows:

$$u(x_j, C) \geq \frac{1}{|C|}\left(1 + \sum_{z\in C\setminus\{x_i,x_j\}} \frac{1}{d(x_i,z)+1}\right) \geq$$
$$\frac{1}{|C|}\left(1 + \sum_{z\in C\setminus\{x_i,x_j\}} \frac{1}{2d(x_i,z)}\right) = \frac{1}{2}u(x_i, C) > \frac{1}{4}$$

■

Lemma 3. *If P is a partition in the core and x, y two players adjacent in G, then at least one of x, y attains utility greater than or equal to $1/2$ in P.*

Proof. Assume by contradiction that none of the conditions hold. Then there exists adjacent pair x, y with $u(x, P) \leq 1/2$ and $u(y, P) < 1/2$, or viceversa. But then coalition $\{x, y\}$ is blocking, since it gives utility $1/2$ to each player, contradition with P being core stable. ■

Corollary 2. *The welfare of any non-singleton coalition in a core stable partition is at least $1/4$ of the maximum possible for that coalition.*

Proof. Let P be the core-stable partition, and T_i, D_i, and F_i denote the sets of players in C_i with utilities strictly greater than $1/2$, equal to $1/2$, and strictly less than $1/2$, respectively. Denote $t_i = |T_i|$, $d_i = |D_i|$, $f_i = |F_i|$, and $c_i = |C_i|$. The sets T_i, D_i, and F_i are disjoint, T_i is non-empty, and $T_i \cup D_i \cup F_i = C_i$.

From Lemmas 2 and 3, each player $y \in F_i$ is surrounded only by players $x \in T_i$. Since $u(x) > 1/2$, we have that $u(y) > 1/4$. The welfare of C_i satisfies the inequality:

$$SW(C_i) > \frac{t_i}{2} + \frac{d_i}{2} + \frac{f_i}{4} > \frac{t_i + d_i + f_i}{4} = \frac{c_i}{4} \tag{2.7}$$

■

Theorem 7. *Let $G = (N, E)$ be a game with nonempty core. Then $Gap_{\min}(G)$ is in worst case $\Theta(\sqrt{n})$.*

Proof. Let partition P^* be welfare maximising and $P = (C_1, \ldots, C_k, S_1, \ldots, S_s)$ in the core, where $c_i \geq 2$, $i = \overline{1, k}$, and $|S_j| = 1$, $j = \overline{1, s}$. Let T_i, D_i, and F_i denote the sets of players in C_i with utilities strictly greater than 1/2, equal to 1/2, and strictly less than 1/2, respectively. Denote $t_i = |T_i|$, $d_i = |D_i|$, $f_i = |F_i|$, and $c_i = |C_i|$. The sets T_i, D_i, and F_i are disjoint, T_i is non-empty, and $T_i \cup D_i \cup F_i = C_i$.

From Lemmas 2 and 3, each singleton S_i can only be connected to players in $T_1, \ldots, T_k$. In addition, $u(x_{ij}) \leq \frac{c_i-1}{c_i}$, $\forall x_{ij} \in T_i \subset C_i$. Thus each such x_{ij} can be connected (in G) to at most $c_i - 1$ singletons, because otherwise x_{ij} and the singletons would form a blocking coalition. The number of singletons, s, satisfies the inequality:

$$s \leq \sum_{i=1}^{k} t_i(c_i - 1)$$

With Corollary 2 this gives:

$$\begin{aligned} \mathrm{Gap}_{\min}(G) \leq \frac{SW(P^*)}{SW(P)} \\ = \frac{SW(P^*)}{\sum_{i=1}^{k} SW(C_i, P)} < \frac{n}{\left(\frac{\sum_{i=1}^{k} c_i}{4}\right)} = \frac{4n}{n-s} \end{aligned}$$

Consider n fixed and find an upper bound for s, which in turn gives an upper bound for $\frac{4n}{n-s}$. Observe that each player in coalition C_i has at most $c_i - 2$ singletons connected to it, thus C_i has at most $c_i(c_i - 2)$ singletons:

$$s \leq \sum_{i=1}^{k} c_i(c_i - 2) \leq \sum_{i=1}^{k} c_i^2 \leq (\sum_{i=1}^{k} c_i)^2 = (n-s)^2$$

From $s < n$ and $s \leq (n-s)^2$, it follows that $s \leq n - \sqrt{n} + 1/2$, and so:

$$\mathrm{Gap}_{\min}(G) \leq \frac{4n}{n-s} \leq \frac{4n}{\sqrt{n} - 1/2} \in O(\sqrt{n})$$

To complete the proof, we give a $\Theta(\sqrt{n})$ example. Let G be such that $N = C \cup S$, $C = \{x_1, \ldots, x_c\}$ is a clique, $S = \{y_1, \ldots, y_s\}$ an independent set, with $|C| = c$ and $|S| = s$. In addition, each player $y \in S$ is connected to exactly one $x \in C$, and each $x \in C$ has exactly $c - 2$ direct neighbous in S, which we denote as $S(x)$. Thus $s = c(c-2)$, and so $n = c^2 - c$. Note that while n cannot be any integer, there are arbitrarily large n of this form. Solving for c gives

$$c = \sqrt{n + \frac{1}{4}} + \frac{1}{2} \tag{2.8}$$

The optimal welfare partition, P^*, is such that each player $x \in C$ forms a coalition with its set $S(x)$. Agent x has utility $\frac{c-2}{c-1}$, while each player in $S(x)$ gets $1/2$. Thus,

$$SW(P^*) = c\left(\frac{c-2}{c-1}\right) + c(c-2)\frac{1}{2} = \frac{(c-2)c(c+1)}{2(c-1)}$$

Partition P^* is not stable because C is blocking, but $P = (C, \{y_1\}, \ldots, \{y_s\})$ is in the core. Each player in C gets $\frac{c-1}{c}$, while everyone in S obtains zero:

$$SW(P) = \sum_{x \in C} \frac{c-1}{c} = c - 1$$

Using Equation 2.8,

$$\mathrm{Gap}_{\min}(G) = \frac{(c-2)c(c+1)}{2(c-1)^2} \approx \sqrt{n} - 1 \in \Theta(\sqrt{n})$$

■

Better bounds for the gap can be obtained depending on the underlying graph model. Here we consider dense graphs, which are common in social networks, and show their stability gap is small.

Theorem 8. *The stability gap of every graph with m edges, where*

$$m \geq \left(\frac{1-\varepsilon^2}{2}\right)n^2 - \left(\frac{1-\varepsilon}{2}\right)n \tag{2.9}$$

is at most $\frac{4}{1-\varepsilon}$, where $\varepsilon \in [0,1]$.

Proof. For $\varepsilon = 0$, the inequality simply states that the number of edges is non-negative. For $\varepsilon = 1$, it requires that the graph is complete: $m \geq \frac{n(n-1)}{2}$.

We observe that the singletons in the core form an independent set, since otherwise they could organize themselves into coalitions and improve their welfare by doing so. Let α be the the independence number of the graph. From the fact that a graph with independence number α has the number of edges, m, bounded as follows:

$$m \leq \binom{n}{2} - \binom{\alpha}{2},$$

we obtain that α satisfies the inequality

$$\alpha \leq \frac{1}{2} + \sqrt{\frac{1}{4} + n(n-1) - 2m}$$

From Inequality 2.9, we get that $\alpha \leq \varepsilon n$, and so any core configuration has at most εn singletons. From Corollary 2, the welfare of the remaining $(1-\varepsilon)n$ players is at least $\frac{1-\varepsilon}{4}n$, and so the gap is bounded as follows:

$$\text{Gap}_{\text{min}}(G) < \frac{n}{(1-\varepsilon)n/4} = \frac{4}{1-\varepsilon}$$

For example, take $\varepsilon = 1/2$. Then every graph with at least $m \geq \frac{3}{8}n^2 - \frac{n}{4}$ edges has a gap bounded by eight. ■

The Erdős-Rényi random graph model is perhaps the best known and widely studied method for generating random graphs [15].

Theorem 9. *The expected stability gap of graphs generated under the Erdős-Rényi $G(n,p)$ graph model is bounded by $\frac{4}{1-2\log(n)/n}$ whenever $p \geq 1/2$.*

Proof. It is known that the expected independence number of $G(n, 1/2)$ graphs is $\alpha \leq 2\log(n)$, and in general, $G(n,p)$ graphs have $\alpha \leq 2\log(n)$ whenever $p \geq 1/2$. The gap can be bounded as follows:

$$\text{Gap}_{\text{min}}(G) < \frac{n}{(n - 2\log(n))/4} = \frac{4}{1 - 2\log(n)/n}$$

For large graphs, the bound is four. ■

2.7 Alternative Solution Concepts

From Theorem 7, the stability gap can be as high as $\Theta(\sqrt{n})$. In this section we consider several variations of the core with improved social support.

2.7.1 Stability Threshold

Recall that a player achieves his best possible utility in a coalition with his direct neighbours and no-one else. Moreover, the improvement function satisfies diminishing returns, and so the higher a player's utility, the harder it is to improve it. The stability threshold is descriptive of situations where players naturally stop seeking improvements once they achieved a minimum value. This is a well-known assumption observed experimentally as a form of bounded rationality: choosing outcomes which might not be optimal, but will make the players sufficiently happy.

We analyse stability for a threshold of $k/(k+1)$, which is equivalent to a player forming a coalition with k of his direct neighbours. In this case, there can be at most $k-1$ singletons neighbouring any player with utility at least $1/2$ in the core, since otherwise the singletons can block with that player.

Theorem 10. *Let $G = (N, E)$ be an induced subgraph game with nonempty core of threshold $k/(k+1)$. Then $Gap_{\min}(G) \leq 4$ if $k = 1$, and $Gap_{\min}(G) \leq 2k$ if $k \geq 2$.*

Proof. We use the notation in Corollary 2. Let $P = (C_1, \ldots, C_m, S_1, \ldots, S_s)$ be a core-stable partition with threshdold k, where coalitions S_i are singletons. For each player in $T_i \subset C_i$ there can be at most $k-1$ singletons, and the welfare of C_i satisfies the inequality

$$SW(C_i) > \frac{t_i}{2} + \frac{d_i}{2} + \frac{f_i}{4} \tag{2.10}$$

Let $T = \cup_{i=1}^m T_i$, $D = \cup_{i=1}^m D_i$, $F = \cup_{i=1}^m F_i$, $S = \cup_{i=1}^m S_i$ and denote $|T| = t$, $|D| = d$, $|F| = f$, $|S| = s$. Given that $SW(S_i) = 0$, we can use Inequality 2.10 to obtain

$$SW(P) = \sum_{i=1}^{m} SW(C_i) > \sum_{i=1}^{m} \frac{t_i}{2} + \frac{d_i}{2} + \frac{f_i}{4} = \frac{t}{2} + \frac{d}{2} + \frac{f}{4} \tag{2.11}$$

We also have that

$$n = t + d + f + s \leq t + d + f + (k-1)t = kt + d + f$$

Since $SW(P^*) < n$, we get

$$\mathrm{Gap}_{\min}(G) = \frac{SW(P^*)}{SW(P)} < \frac{kt + d + f}{\left(\frac{t}{2} + \frac{d}{2} + \frac{f}{4}\right)} = \frac{4kt + 4d + 4f}{2t + 2d + f}$$

If $k \geq 2$, the inequality guarantees that $\text{Gap}_{\min}(G) < 2k$. If $k = 1$, the best bound we get is $\text{Gap}_{\min}(G) < 4$. Note, however, that for $k = 1$, $\text{Gap}_{\max}(G) = 2$, since the diameter two decomposition of G is core-stable with threshold $k = 1$. ■

2.7.2 The "No Man Left Behind" Policy

From Corollary 2, the core guarantees average utility greater than $1/4$ to every non-singleton coalition. Thus the reason for which the core welfare can be low is because there exist networks in which many players are left alone in equilibrium.

Here we view the formation of core stable structures as a process that starts from the grand coalition and stabilizes through rounds of coalitions splitting and merging. While the search for equilibrium can begin from any partition, we observe that initializing with the grand coalition is natural in many situations. For example, at the beginning of any joint project, a group of people gather to work on it. However, as the project progresses, they may form subgroups based on the compatibilities and strength of social ties between them. We formulate a simple social rule that players have to follow when merging or splitting coalitions. That is, whenever a new group forms, it cannot leave behind any player working alone. We call this rule the "No Man Left Behind" policy. The "No Man Left Behind" code of conduct is well known in the army and refers to the fact that no soldier can be left alone in a mission or abandoned in case of injury.

Lemma 4. *If $P = (C_1, \ldots, C_m)$ is stable under the "No Man Left Behind" policy, then for any two direct neighbours $x, y \in C_i \in P$ it cannot be the case that both x and y receive less than $1/2$.*

Proof. Assume, similarly to the proof for Lemma 3, that there exist adjacent players $x, y \in C_i$ such that $u(x) < 1/2$ and $u(y) < 1/2$. Agents x and y can block to improve their utility. However, by the "No Man Left Behind" Policy, they cannot deviate as a pair if they leave singletons behind. Let $Z(x)$ and $Z(y)$ be the players that become singletons when x and y deviate, where x is connected to the players in $Z(x)$, and y to $Z(y)$. Thus x and y can deviate with coalition:

$$B' = \{x, y\} \cup Z(x) \cup Z(y)$$

Both x and y are in distance at most two from all the players in B', and so $u(x, B') \geq 1/2$ and $u(y, B') \geq 1/2$. ■

Theorem 11. *Let G be a game which is stable under the "No Man Left Behind" policy. Then $\text{Gap}_{\min}(G) < 4$.*

Proof. Let $P = (C_1, \ldots, C_m)$ be a stable partition. From the requirements of the policy, each coalition C_i has size at least two. From Lemma 4 and Corollary 2, we get that

$$SW(P) = \sum_{i=1}^{k} SW(C_i, P) > \sum_{i=1}^{k} \frac{c_i}{4} = \frac{n}{4}$$

and so

$$\text{Gap}_{\min}(G) \leq \frac{SW(P^*)}{SW(P)} < \frac{n}{(n/4)} = 4$$

■

2.8 Related Work

Social distance games are a compact model that can be placed in the general context of hedonic games [7]. Alon *et al.* [2] propose a graph-based model and uncover the relationship between the existence of Nash equilibrium and the graph's diameter. Bloch and Jackson [6] analyse network formation games among players whose payoffs depend on the structure of the network, using the stability notions of Nash equilibrium and pairwise stability. In their formulation, players derive utility from forming links to other players in the network, but have to pay explicitely for maintaining those links. Aadithya *et al.* [1] propose efficient algorithms for computing a Shapley value-based network centrality. Suri and Narahari [48] use the Shapley value to select the top k nodes in a social network. Finally, there exists a rich body of literature investigating the small world phenomenon and the properties of the networks in which it occurs. We mention the seminal work of Kleinberg [31] in this area.

Chapter 3

Matchings with Compact Externalities

3.1 Introduction

Matchings are an important theoretical abstraction which have been extensively studied in several fields, including economics, combinatorial optimization, and computer science. Matchings are often used to model markets, and examples include the classical marriage problem, firms and workers, schools and students, hospitals and medical interns [13]. Previous matching literature focused primarily on one-to-one and one-to-many models [42]. More recently, however, attention has been paid to more complex models of many-to-many matchings due to their relevance to many real-world markets. For example, most labour markets involve at least a few many-to-many contracts [18]. One of the central questions in matching games is stability. There exists a rich literature on the stability of matchings [42] which dates back to the classical stable marriage problem [21].

More realistic matching models should take into account the fact that in many settings the utility of a player is influenced not only by their own choices, but also by the choices that other players make. Such an influence is called an externality. For instance, companies care not only about the employers they hire themselves, but also about the employers hired by other companies. This aspect is crucial to how competitive a company is on the market, and so externalities must be considered to completely understand such situations. Modeling matchings with externalities is computationally challenging, as fully expressive representations require exponential space.

Our contributions can be summarized as follows. We formulate the first compact model of externalities in the computer science literature, in which the influence of a match on each player is computed additively. Additive functions are a natural tradeoff between expressivity and computational attractiveness, and have been used in many other domains, such as additive hedonic games [8]. This enables computationally bounded players to reason about the dynamics of the game. Second, we formulate a general stability concept, in which a potentially deviating group evaluates the reaction of the outside players before deciding whether to deviate or not. Then we consider several solution concepts that fall under this general notion of stability, namely the neutral, optimistic, pessimistic, and contractual setwise-stable sets. We study the properties of these stable sets and show how they are related to each other, analyze their complexity and provide polynomial-time algorithms where applicable. Finally we discuss social welfare and some natural examples of matchings that can be expressed in our model.

3.2 Background

Perhaps the best known instance of a matching problem is the classical stable marriage problem, which is commonly stated as follows [42]:

[The Stable Marriage Problem]:
Let M and W be two disjoint and equally sized sets of men and women, such that each person has a strict ordering over the members of the opposite sex. The problem is to marry the men and the women such that no two people of the opposite sex prefer each other to their current partners. If no such pair exists, the marriages are considered to be stable.

The preference lists may be such that one would prefer to remain single, rather than being married with someone they don't care for. The following instance is an example of the stable marriage problem.

Example 1. *Let $M = \{m_1, m_2, m_3\}$ and $W = \{w_1, w_2, w_3\}$, such that the men's preference lists are:*

- $P(m_1) = (w_2, w_1, w_3)$
- $P(m_2) = (w_1, w_3, w_2)$
- $P(m_3) = (w_1, w_2, w_3)$

The women's preference lists are:

- $P(w_1) = (m_1, m_3, m_2)$
- $P(w_2) = (m_3, m_1, m_2)$
- $P(w_3) = (m_1, m_3, m_2)$

The matching given by $\mu = \{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$ *is unstable, because pair* (m_1, w_2) *is blocking. However, there exists a stable matching, namely* $\mu^{'} = \{(m_1, w_1), (m_3, w_2), (m_2, w_3)\}$.

The problem always admits a stable set of marriages [21]. We briefly describe the Gale-Shapley algorithm, which returns such a stable solution in polynomial time.

[The Gale-Shapley Algorithm]:
Let every man propose to his favourite woman. Each woman who receives more than one proposal rejects all, except for her favourite, among those men who proposed to her. However, the woman does not accept the man yet, but rather keeps him on a "maybe" list, to account for the possiblity that someone better may come later. In the second stage, the men who were rejected proceed to propose to their second choices. Each woman receiving proposals chooses her favourite among the new proposers and the maybe man from the previous iteration. She rejects everyone, except for her favourite, among those who proposed to her, which is marked as the new "maybe". The men who were rejected in the second stage proceed to the next stages, up to at most n *stages, where* n *is the number of men and women, respectively. In the end, every woman is married to her favourite man, among those who proposed to her, and the set of marriages formed can be shown to be stable.*

Matching problems have been studied in depth since the introduction of the stable marriage problem. The marriage problem is also known as a one-to-one matching. However, there exist many realistic scenarios in which one-to-one matchings are not an appropriate modeling tool. For example, assigning students to colleges, medical interns to hospitals, and workers to companies, are examples of many-to-one matchings. The many-to-one matching problem is also known as the stable admission model, or the hospital/residents problem. In this case, every student has preferences over universities, and every university has preferences over the subsets of students. A many-to-one matching of students to universities is stable if there is no pair (u, a) consisting of a university u and a student a, such that student a would prefer university u to their current one, and university u would prefer to enroll student a, possibly after letting go of some of its current students. The many-to-one matching problem can be solved efficiently using a generalization of the Gale-Shapley algorithm.

Finally, consider a set of consultants and companies, such that every company wants to hire several consultants, and a consultant can work for more than one company. This is an example of a many-to-many matching problem. Most job markets are known to contain at least a few many-to-many contracts [18]. Many-to-many matchings are the least understood class of matching problems and will be studied in more detail in this thesis.

Here we consider a generalization of matching problems, in which players are affected not only by the choices that they make themselves, but also by the choices that other players make. Such influences, called externalities, are well studied in economics. In the most general representation of matchings with externalities, players have preferences over the set of all the possible matchings. While such a representation is complete, it requires the players to store an exponential number of preference profiles. Next we formulate a representation which is polynomial in the number of players and study its properties.

3.3 The Model

We start with the following definition:

Definition 20 (Matching Game)**.** *A* matching game *is represented as a tuple* $G = (M, W, \Pi)$*, where*

- (M, W) *is a bipartite graph, where* $N = M \cup W$ *is the set of players,* $M = \{m_1, \ldots, m_{|M|}\}$ *and* $W = \{w_1, \ldots, w_{|W|}\}$ *are disjoint and on opposite sides of the bipartite graph.*
- Π *is a real valued function such that* $\Pi(m, w|m)$ *is the* value *that player* m *receives from forming an edge with player* w*,* $\Pi(m, w|w)$ *is the value that* w *receives from forming an edge with player* m*, and* $\Pi(m, w|z)$*, where* $z \notin \{m, w\}$ *is the* externality *that player* z *receives when* m *and* w *form the edge* (m, w)*.*

We denote by a *match* (m, w) the formation of the edge between players m and w, and by a *matching* a set of matches. We assume the formation of a match requires the consent of both parties, while severing a match can be done unilaterally by any of its endpoints. The “empty matching” contains no matches, while the “complete matching” contains all the possible matches.

Definition 21 (Utility)**.** *Given a matching* A *over* N*, the* utility *of a player* z *in* A *is:*

$$u(z, A) = \sum_{(m,w) \in A} \Pi(m, w|z)$$

That is, a player's utility is the sum of values it receives from matches it participates in, along with the sum of all externalities that arise due to the matchings of other players. When referring to some matching A, we write (m, w) to denote the existence of a match between players m and w, and by (z) to denote that player z is not involved in any matches.

Example 2. *Let $M = \{m\}, W = \{w_1, w_2\}$, and Π given as follows:*

- $\Pi(m, w_1|w_1) = \Pi(m, w_1|m) = 2$
- $\Pi(m, w_2|w_2) = 1$
- $\Pi(m, w_2|m) = 5$
- $\Pi(m, w_1|w_2) = -1$
- $\Pi(m, w_2|w_1) = 3$

Matching $A = \{(m, w_1)\}, \{w_2\})$ matches m with w_1, giving the players utilities:

- $u(m, A) = \Pi(m, w_1|m) = 2$
- $u(w_1, A) = \Pi(m, w_1|w_1) = 2$
- $u(w_2, A) = \Pi(m, w_1|w_2) = -1$

Similarly, matching $A^{'} = \{(m, w_1), (m, w_2)\}$ gives the players utilities:

- $u(m, A^{'}) = \Pi(m, w_1|m) + \Pi(m, w_2|m) = 7$
- $u(w_1, A^{'}) = \Pi(m, w_1|w_1) + \Pi(m, w_2|w_1) = 5$
- $u(w_2, A^{'}) = \Pi(m, w_1, w_2) + \Pi(m, w_2|w_2) = 0$

3.4 Gamma-Stability

In this section we define a very general notion of stability, namely Γ-stability, in which a deviating group evaluates the reaction of the players outside the group when deciding whether to engage in the deviation or not. We first introduce some preliminary notions.

Definition 22 (Deviation)**.** *Given a matching A and a set of players $B \subseteq N$, a* deviation *of B from A is a new matching, $A^{'}$, which is identical to A, except that each player in B performs at least one of the following: severs a match with another player in N or forms a new match with a player in B .*

In other words, for every $(m, w) \in A$, if $m \in B$ or $w \in B$, then (m, w) may or may not exist in $A^{'}$, while if $m, w \in N \setminus B$, then (m, w) is always in $A^{'}$. Furthermore, for every $(m, w) \notin A$, (m, w) may be in $A^{'}$ only when $m, w \in B$.

If group B can deviate and improve the utility of at least one of its members, without degrading any other member, then B is called a *blocking coalition*. It is important to note that for a coalition to be deviating, we require that all of its members perform *some* action, either in the form of severing edges or in forming new matches. We contrast this with the following possibility:

- There exists a set of players, B, that can perform a deviation as described above, such that each player in B forms a new match or severs an old match.
- There exists player $z \notin B$ that performs no action, but improves upon B's deviation.

In this analysis, we say that coalition B is deviating. However, coalition $B \cup \{z\}$ is not deviating, since z performs no action. A consequence of this requirement is that for B to be a blocking coalition, all of its members must perform some action.

Definition 23 (Γ-response)**.** *Let Γ be a function such that for every set of players $B \subseteq N$ and matching A, the* Γ-response $\Gamma(B, A, A^{'})$ *is the matching that $N \setminus B$ forms upon B's deviation, $A^{'}$, from A.*

Thus the Γ-response represents the reaction of the rest of the society upon the deviation of a group.

Definition 24 (Γ-Stability)**.** *Matching A is* Γ-stable *if there exists no set of players $B \subseteq N$ that can deviate to a different matching $A^{'}$, and improve the utility of at least one member of B while not degrading the other members of B in the Γ-response $\Gamma(B, A, A^{'})$.*

If such a coalition B exists, then we say that matching A is *blocked* by coalition B through matching $A^{'}$.

In addition, we differentiate between *weak stability*, in which all the members of a blocking coalition must improve strictly with the deviation, and *strong stability*, in which at least one member must improve while the others do not degrade.

The solution concepts discussed in this paper refer to strong stability, with the only exception of the pairwise stable set, for which we require strict improvements for all the deviators, as is done in the classical model of one-to-one matchings.

The Γ function may be hard to compute, and since the players are limited by bounded rationality [43], they may be unable to engage in very complex computations to determine the Γ-response. Thus in this paper, we will consider several natural heuristics that a deviating group can use to evaluate the consequences of a possible deviation. We note that the recursive core of Koczy [32] for games in partition function form can be seen as an approximation of the Γ-response. There, the response $\Gamma(B, A, A')$ to a deviation A' by coalition B contains the set of all equilibria in the residual game restricted to the players in $N \setminus B$, if any. If the residual game has no stable structure, then $\Gamma(B, A, A')$ contains all the possible partitions in the residual. Next we define some weaker notions of stability.

Definition 25 (Γ-Individual Rationality). *Matching A is Γ-individually rational if it has no blocking coalition of size one.*

Thus under Γ-individual rationality, no single player can improve by severing matches, knowing that the reaction of the rest of the players upon their deviation is given by the function Γ.

Lemma 5. *The empty matching satisfies Γ-individual rationality for all functions Γ.*

Definition 26 (Γ-Pairwise Stability). *Matching A is Γ-pairwise stable if it has no blocking coalitions of size one or two.*

That is, A satisfies Γ-individual rationality and has no blocking pairs under the Γ notion of stability. It is immediate that if a matching A satisfies Γ-stability, it also satisfies Γ-pairwise rationality, and if A satisfies Γ-pairwise rationality, it also satisfies Γ-individual rationality.

We also introduce a notion of fairness, that stipulates a minimal acceptable level of utility. First observe that a utility of zero is achievable by all the players in the empty matching. Thus a matching giving non-negative utility to all the players conveys the idea that nobody is hurt in that matching. We emphasize that fairness can be achieved in any game.

Definition 27 (Fairness). *A matching is* fair *if it gives non-negative utility to every player.*

In matching games without externalities, fairness is equivalent to Γ-individual rationality. A player would never accept negative utility, since they can always get zero by severing their matches.

Fairness and individual rationality do not coincide in the presence of externalities. There may exist individually rational matchings with negative utilities, since a player can be hurt by matches they have no control over. Moreover, there can exist fair matchings which are not individually rational, where a player may increase utility by severing some negative matches.

In the next two sections we analyze the stability of many-to-many and one-to-one matchings using several notions of stability that fall under the Γ-stability template, where the players compute approximations of the Γ-response upon their deviation using neutral, pessimistic, and optimistic reasoning.

3.5 Many-to-Many Matchings

Many-to-many matchings are the most general, yet the least understood class of matching problems. In many-to-many matchings, each player can maintain contracts with arbitrarily many players from the other side of the market.

3.5.1 Neutral Stability

The setwise-stable set is a solution concept which models interactions between groups of players and has been introduced by Sotomayor [46].

Definition 28 (Setwise-Stable Set)**.** *Matching A is setwise-stable if no coalition B can block assuming that $N \setminus B$ will not react to their deviation. The setwise-stable set is the set of all such matchings.*

The setwise-stable set falls under the Γ notion of stability, where Γ is the identity function. That is, $\Gamma(B, A, A') = A'$, for all $B \subset N$ and all matchings A, A'. Thus from the point of view of a deviating group, the deviation has no consequences. The rest of the society will be organized exactly the same way, and the deviators are allowed to keep any matches with the outside players whenever they wish to do so. Because this solution concept assumes the outside players do not respond to the deviation, we also refer to it as the *neutral setwise-stable set*, to differentiate it from other variants we studied. The matchings that belong to the setwise stable set can be characterized as follows:

Theorem 12. *Any matching in the setwise-stable set of a game has the form:*

$$A = \{(m, w) | \Pi(m, w|m) \geq 0 \textit{ and } \Pi(m, w|w) \geq 0\}$$

Proof. Let A be a stable matching and (m, w) a match in G. If $\Pi(m, w|m) < 0$, then (m, w) cannot belong to A, since m will sever the match and improve their utility by doing so. The same reasoning applies for the case where $\Pi(m, w|w) < 0$. Thus any match $(m, w) \in A$ satisfies $\Pi(m, w|m) \geq 0$ and $\Pi(m, w|w) \geq 0$. ∎

Note that the setwise-stable set can be empty as we will show in the following example.

Example 3 (Empty Setwise-Stable Set)**.** *Let $M = \{m\}$, $W = \{w_1, w_2\}$, with Π given as follows:*

- $\Pi(m, w_i|m) = 0$
- $\Pi(m, w_i|w_i) = \varepsilon$
- $\Pi(m, w_1|w_2) = \Pi(m, w_2|w_1) = -\Delta$, *where* $\Delta > \varepsilon > 0$

The empty matching is blocked by (m, w_1), matching $\{(m, w_1), (w_2)\}$ is blocked by (m, w_2), matching $\{(m, w_2), (w_1)\}$ is blocked by (m, w_1), and the complete matching $\{(m, w_1), (m, w_2)\}$ is blocked by the empty matching.

Theorem 13. *Any matching in the neutral setwise-stable set is* fair *under positive externalities.*

Proof. Let A be a matching in the neutral setwise-stable set and $z \in N$ any player. Matching A satisfies individual rationality, thus the sum of values that z receives in A must be positive, since otherwise z could improve utility by cutting all its matches. The externality values are positive, i.e. $\Pi(m, w|z) > 0$, $\forall(m, w)$, and so:

$$u(z, A) = \sum_{(x,z)\in A} \Pi(x, z|z) + \sum_{(m,w)\in A; m,w\neq z} \Pi(m, w|z) \geq 0$$

∎

Proposition 1. *The complete matching is setwise-stable under positive values and positive externalities.*

Proposition 2. *The empty matching is setwise-stable under negative values and negative externalities.*

Theorem 14. *The unique candidate for the setwise-stable set under positive values and negative externalities is the complete matching.*

Proof. Let A be an incomplete matching and m, w two players that do not form a match in A. Since $\Pi(m, w|m) > 0$ and $\Pi(m, w|w) > 0$, the pair $\{m, w\}$ can deviate by forming the match (m, w). Both players are allowed to keep all their other matches in A, and so their utility strictly increases after the deviation. Thus no incomplete matching can be setwise-stable. ■

For example, if the externalities are drawn from $(-\varepsilon, 0)$, where ε is small compared to the values, the complete matching is setwise-stable. If the externality values are $-\infty$, the empty matching is blocking for the complete matching, case in which the setwise-stable set is empty.

Theorem 15. *The unique candidate for the setwise stable set with negative values and positive externalities is the empty matching.*

Proof. Any non-empty matching will contain some match (m, w) which is negative for m and w. These players will have an incentive to cut (m, w) since they can improve their utility by doing so. Thus the empty matching is the only candidate for the setwise stable set. ■

The following game contrasts the case where the empty matching belongs to the setwise-stable set with the case where it does not.

Example 4. *Let $G = (M, W, \Pi)$ be such that $\Pi(m, w|m) = \Pi(m, w|w) = -\varepsilon < 0$, for all m, w and $\Pi(m, w|z) = \delta > 0$ for all m, w and $z \neq m, w$. If $\varepsilon \gg \delta$, the only matching satisfying setwise stability is the empty matching, since the formation of any match is very expensive for the participating players and not compensated by the utility obtained from the externalities. Otherwise, if $\delta \gg \varepsilon$, the complete matching is blocking for the empty matching, and so the setwise-stable set is empty.*

To prove the next hardness result regarding the complexity of computing setwise-stable matchings, we provide a reduction from the *KNAPSACK* problem [22].

KNAPSACK

***Input*:** Tuple $I = \langle U, s, v, B, K\rangle$, where $U = \{u_1, \ldots, u_n\}$ is a finite set, $s(u) \in \mathbb{Z}^+$ the size of element $u \in U$, $v(u) \in \mathbb{Z}^+$ the value of element $u \in U$, $B \in \mathbb{Z}^+$ a size constraint, and $K \in \mathbb{Z}^+$ a value goal.

***Question*:** Does there exist a subset $U^{'} \subseteq U$ such that

$$\sum_{u \in U'} s(u) \leq B \text{ and } \sum_{u \in U'} v(u) \geq K? \tag{3.1}$$

Theorem 16. *Checking nonemptiness of the setwise-stable set is NP-hard.*

Proof. We provide a reduction from *KNAPSACK*. Given instance $I = \langle U, s, v, B, K \rangle$, we construct a matching game G such that G has a nonempty setwise-stable set if and only if I has a solution. Let $M = \{x_1, \ldots, x_n, m_1, m_2\}$, $W = \{y_1, \ldots, y_n, w\}$, and Π have the following non-zero entries:

- $\Pi(x_i, y_i | m_1) = -s(u_i)$ and $\Pi(x_i, y_i | m_2) = v(u_i)$, $\forall i \in \{1, \ldots, n\}$
- $\Pi(m_1, w | m_1) = -B$
- $\Pi(m_2, w | m_2) = K - \sum_{u_i \in U} v(u_i)$
- $\Pi(x_j, w | x_j) = -1$ and $\Pi(m_i, y_j | y_j) = -1$, $\forall i \in \{1, 2\}$ and $\forall j \in \{1, \ldots, n\}$

Since the instances with $\sum_{u_i \in U} v(u_i) < K$ are trivially solvable, we are interested in those cases where $\sum_{u_i \in U} v(u_i) \geq K$, and so $\Pi(m_2, w | m_2) \leq 0$. If the *KNAPSACK* instance I has a solution $U^{'}$, we claim the following matching belongs to the setwise-stable set:

$$A = \left(\bigcup_{u_i \in U'} \{(x_i, y_i)\} \right) \cup \{(m_1), (m_2), (w)\} \tag{3.2}$$

First note that $U^{'}$ satisfies the conditions in (3.1), and so the utilities of the players in A are:

- $u(m_1, A) = \sum_{u_i \in U'} \Pi(x_i, y_i | m_1) = \sum_{u_i \in U'} -s(u_i) \geq -B$
- $u(m_2, A) = \sum_{u_i \in U'} \Pi(x_i, y_i | m_2) = \sum_{u_i \in U'} v(u_i) \geq K$
- $u(x_i, A) = u(y_i, A) = u(w, A) = 0$, $\forall i \in \{1, \ldots, n\}$

All the players, except possibly for m_1 and m_2, obtain their best possible utility in A. Thus any blocking coalition, B, would have to contain at least one of m_1 and m_2. Since blocking requires each member of B to perform an action, it follows that m_1, m_2 can only be involved in blocking A by forming a new match. Recall that $\Pi(m_i, y_j|y_j) = -1$, and so players y_j will never accept a match with either m_1 or m_2. Thus the only matches that m_1 and m_2 could form, if deviating, are (m_1, w) and (m_2, w), respectively.

The best utility that m_1 can get when matched with w is attained in $A_1^* = \{(m_1, w), (x_1), \ldots, (x_n), (y_1), \ldots, (y_n), (m_2)\}$, where the candidate for the blocking coalition is $B_1 = \{m_1, w\} \cup \left(\bigcup_{u_i \in U'} \{x_i, y_i\}\right)$:

$$u(m_1, A_1^*) \quad = \Pi(m_1, w|m_1) = -B \leq \textstyle\sum_{u_i \in U'} -s(u_i) = u(m_1, A)$$

The best utility that m_2 can get when matched with w is attained in $A_2^* = \{(m_2, w), (x_1, y_1), \ldots, (x_n, y_n), (m_1)\}$, where the candidate for the blocking coalition is $B_2 = \{m_2, w\} \cup \left(\bigcup_{u_i \notin U'} \{x_i, y_i\}\right)$:

$$\begin{aligned} u(m_2, A_2^*) &= \Pi(m_2, w|m_2) + \sum_{i=1}^{n} \Pi(x_i, y_i|m_2) = \left(K - \sum_{u_i \in U} v(u_i)\right) + \sum_{u_i \in U} v(u_i) \\ &= K \leq \sum_{u_i \in U'} v(u_i) = u(m_2, A) \end{aligned}$$

Thus m_1 and m_2 cannot improve by deviating from A. Since the other players have no incentive to deviate, it follows that A belongs to the setwise-stable set.

Conversely, if the setwise-stable set of G is non-empty, let A be a stable matching. First note that A must satisfy neutral individual rationality, and so it cannot contain any match with negative value for one of the endpoints. Thus the only non-zero matches that can be included in A are a subset of $\left(\bigcup_{u_i \in U} \{(x_i, y_i)\}\right)$. In addition, any matches of the form (x_i, y_j), with $i \neq j$, can be removed from A without losing stability. Thus without loss of generality, A can be written as a union of the matching in Equation 3.2 for some subset $U' \subseteq U$. From A stable, coalitions $\{(m_1, w)\}$ and $\{(m_2, w)\}$ are not blocking, thus Inequalities 3.3 and 3.3 hold. Equivalently, the conditions in (3.1) are satisfied, and so U' is a solution for the *KNAPSACK* instance. ■

Theorem 17. *Checking setwise-stable set membership is coNP-complete.*

Proof. We show that the complementary problem, of deciding whether a matching does not belong to the setwise-stable set of a game, is NP-complete. Given matching A, one can nondeterministically guess pair $\langle B, A_B \rangle$ such that matching A is blocked by coalition B through matching A_B. Verifying that $\langle B, A_B \rangle$ is blocking for A can be done in $O(n^3)$, by computing for each member z of B, the utility in A and in A_B (assuming that $N \setminus B$ does not react to the deviation). To prove NP-hardness, we provide a reduction from *KNAPSACK*. Given instance $I = \langle U, s, v, B, K \rangle$, let $G = (M, W, \Pi)$ be a matching game with $M = \{x_1, \dots, x_n, m_1, m_2\}$, $W = \{y_1, \dots, y_n, w_1, w_2\}$, and Π with non-zero entries:

- $\Pi(x_i, y_i | m_1) = -s(u_i)$ and $\Pi(x_i, y_i | w_1) = v(u_i)$, $\forall i \in \{1, \dots, n\}$
- $\Pi(m_2, w_2 | m_1) = -B - \varepsilon$ and $\Pi(m_2, w_2 | w_1) = K - \varepsilon$, for some $0 < \varepsilon < 1$
- $\Pi(m_1, w | w) = \Pi(m, w_1 | m) = -1$, for all $w \in W \setminus \{w_1\}$ and $m \in M \setminus \{m_1\}$

Let matching $A = \{(m_2, w_2), (m_1), (w_1), (x_1), \dots, (x_n), (y_1), \dots, (y_n)\}$. The utilities of the players are:

- $u(x_i, A) = \Pi(m_2, w_2 | x_i) = 0$ and $u(y_i, A) = \Pi(m_2, w_2 | y_i) = 0$, $\forall i \in \{1, \dots, n\}$
- $u(m_1, A) = \Pi(m_2, w_2 | m_1) = -B - \varepsilon$
- $u(w_1, A) = \Pi(m_2, w_2 | w_1) = K - \varepsilon$
- $u(m_2, A) = \Pi(m_2, w_2 | m_2) = 0$ and $u(w_2, A) = \Pi(m_2, w_2 | w_2) = 0$

All the players, except for m_1 and w_1 obtain their maximum utility in A. In addition, m_1 or w_1 can only block by forming the match (m_1, w_1), since all other matches with one of these players as an endpoint are unfeasible. We claim that I has a solution if and only if A has a blocking coalition. If I has a solution $U^{'} \subseteq U$, consider the grand coalition N and matching:

$$A_B = \left(\bigcup_{u_i \in U'} \{(x_i, y_i)\} \right) \cup \left(\bigcup_{u_i \notin U'} \{(x_i, y_{i+1}), (x_{i+1}, y_i)\} \right) \cup \{(m_1, w_1), (m_2), (w_2)\},$$

where $x_{n+1} = x_1$ and $y_{n+1} = y_1$. The utilities of the players in A_B are:

- $u(x_i, A_B) = 0$ and $u(y_i, A_B) = 0$, $\forall i \in \{0, \dots, n\}$
- $u(m_1, A_B) = -\sum_{u_i \in U'} s(u_i) \geq -B > -B - \varepsilon = u(m_1, A)$
- $u(w_1, A_B) = \sum_{u_i \in U'} v(u_i) \geq K > K - \varepsilon = u(w_1, A)$
- $u(m_2, A_B) = u(w_2, A_B) = 0$

Thus the grand coalition is blocking under neutrality, since it can form matching A_B and (weakly) improve the utilities of all its members by doing so.

Conversely, assume that A is blocked by a coalition B through some matching A_B. Note that all the players, except for m_1 and w_1, obtain their maximum possible utility in A. Player m_1 cannot block without player w_1, since all other matches (m_1, w), for $w \neq w_1$ are unfeasible; similarly for w_1. Thus any blocking coalition must include players m_1, w_1 and match (m_1, w_1). In addition, for at least one of m_1, w_1 to strictly improve after the deviation, the edge (m_2, w_2) must be removed from A_B. The conditions for improvement in A_B are:

$$u(m_1, A_B) = \sum_{(x_i,y_i)\in A_B} \Pi(x_i, y_i|m_1) = \sum_{(x_i,y_i)\in A_B} -s(u_i) \geq u(m_1, A) = -B - \varepsilon \qquad (3.3)$$

and

$$u(w_1, A_B) = \sum_{(x_i,y_i)\in A_B} \Pi(x_i, y_i|w_1) = \sum_{(x_i,y_i)\in A_B} v(u_i) \geq u(w_1 A) = K - \varepsilon \qquad (3.4)$$

Since $v(u_i)$ and $s(u_i)$ are integers, Inequalities (3.3) and (3.4) are equivalent to

$$\sum_{(x_i,y_i)\in A_B} s(u_i) \leq B$$

and

$$\sum_{(x_i,y_i)\in A_B} v(u_i) \geq K$$

Thus $U' = \{u_i \in U \,|\, (x_i, y_i) \in A_B\}$ is a solution for the I instance. ■

3.5.2 Optimistic Stability

Optimistic stability is another instantiation of Γ-stability, in which a blocking coalition hopes that the rest of the players will do the best for the deviators. Optimism has been considered before in game theoretic models [32].

Definition 29 (Optimistic Setwise-Stable Set)**.** *A matching A belongs to the optimistic setwise-stable set if there is no coalition B that can block under the assumption that the players in $N \setminus B$ react in the best possibly way for B.*

In our model, optimism translates to the following. When coalition B considers deviation A' from A, every player $z \in B$ evaluates the deviation assuming the players in $N \setminus B$

will organize themselves in the best possible way for z. That is, z hopes that $N \setminus B$ will cut any matches with a negative influence on z and form all the matches with a positive influence on z, including initiating those with z as an endpoint (which z would accept). According to z's evaluation, this leads to matching $A_z^{''}$. If z improves (or does not degrade) in $A_z^{''}$ compared to A, then z agrees to participate in blocking.

Observe that since the Π function is asymmetric, it may be the case that the members of a blocking coalition B have inconsistent wishes. That is, a match $(m, w) \in N \setminus B$ that is positive for a player $z_1 \in B$ may be negative for another player $z_2 \in B$. Agent z_1 will hope that $N \setminus B$ keeps (m, w), while player z_2 hopes the same match will be cut.

Optimism can be attributed to bounded rationality and incomplete knowledge. In our model, a player is aware of the influence of every match on their utility. However, the same player might not keep track of the influence of all the matches over all the players, and thus would not be able to compute every utility upon the deviation. Moreover, even if the players are aware of the entire Π function, they do not engage in sophisticated reasoning about the feasibility of possible reactions. Instead, every player computes their own version of Γ, and uses that to decide whether to participate in the deviation.

Under optimism, any player not receiving their maximum utility can randomly sever a match to destabilize the matching, hoping they will get the best possible as a result of this destabilization.

Theorem 18. *Any matching in the optimistic setwise-stable set is a union of two disjoint matchings* $(M^{'}, W^{'}) \cup (M^{''}, W^{''})$*, where* $M^{'} \cup M^{''} = M$*,* $W^{'} \cup W^{''} = W$*, and every player in* $(M^{'}, W^{'})$ *obtains their highest possible utility, while* $(M^{''}, W^{''})$ *is the empty matching.*

Proof. Let A be a matching in the optimistic setwise-stable set and z a player. Since A is stable, z is not blocking. Thus it must be that z is either unmatched, case in which he cannot deviate, or z is matched but already obtains the highest possible utility, and so has no incentive to deviate. Thus N can be partitioned in two subsets, $N^{'}$, the players that obtain in A their maximal utility, and $N^{''}$, the unmatched players. ■

While it may appear that optimistic players will never accept the formation of negative edges in equilibrium, the following example shows that this is not necessarily the case.

Example 5. *Let* (M, W, Π) *be a game such that* $M = \{m_1, m_2\}$*,* $W = \{w_1, w_2\}$*, and* Π *as follows:*

- *Player* m_1*:* $\Pi(m_1, w_1|m_1) = 1$, $\Pi(m_1, w_2|m_1) = 1$

- *Player* m_2*:* $\Pi(m_2, w_2|m_2) = 1$
- *Player* w_1*:* $\Pi(m_2, w_2|w_1) = -W$*,* $\Pi(m_1, w_2|w_1) = 1$*, and* $\Pi(m_1, w_1|w_1) = \Pi(m_2, w_1|w_1) = -2W$*, where* $W \gg 1$
- *Player* w_2*:* $\Pi(m_2, w_2|w_2) = 1$

The unique matching that belongs to the optimistic setwise-stable set is $A = \{(m_1, w_2), (m_2, w_2), (w_1)\}$*, which gives player* w_1 *utility* $u(w_1, A) = \Pi(m_1, w_2|w_1) + \Pi(m_2, w_2|w_1) = 1-W$*. Player* w_1 *attains their best possible utility in matching* $A' = \{(m_1, w_2), (m_2), (m_2)\}$*, with* $u(w_1, A') = \Pi(m_1, w_2|w_1) = 1 \gg 1 - 2W$*. Any matching other than* A *can be blocked by coalition* $\{m_1, m_2, w_2\}$*, since all these players obtain their maximum utility in* A*. In addition,* w_1 *cannot sever any matches in* A *due to being a singleton, and also cannot block by forming new matches, because those edges would decrease* w_1*'s utility even further.*

Theorem 19. *Checking nonemptiness of the optimistic setwise-stable set is NP-complete.*

Proof. First note that checking nonemptiness of the optimistic setwise-stable set is in NP. Given a matching A, the utility of each player $z \in N$ in A can be computed in $O(n^2)$. In addition, for every pair $(m, w) \in (M, W)$, we can again compute in $O(n^2)$ the best case utilities of m and w when matched with each other. Verifying if A is stable can be done by iterating over all pairs (m, w) and checking in $O(1)$ if both m and w can (weakly) improve by deviating under optimism, compared to their current utility in A. The reduction is similar to that in Theorem 16. ■

Next we consider the problem of computing the optimistic setwise-stable set from the perspective of parametrized complexity. A problem is *fixed parameter tractable (FPT)* with respect to some parameter if there exists an algorithm which given an instance of size n and parameter value k, computes a solution in $O(f(k)n^{O(1)})$, where f is a computable function.

Given a matching game $G = (M, W, \Pi)$, classify the set of possible matches, $M \times W$, into the following categories:

- *Required:* $R = \{(m, w) : \Pi(m, w|m) \geq 0, \Pi(m, w|w) \geq 0, \text{ and } \Pi(m, w|m) + \Pi(m, w|w) > 0\}$
- *Feasible Positive:* $FP = \{(m, w) : \Pi(m, w|m) = \Pi(m, w|w) = 0 \text{ and } \Pi(m, w|z) \geq 0, \forall z \neq m, w\}$

- *Feasible Negative:* $FN = \{(m, w) : \Pi(m, w|m) = \Pi(m, w|w) = 0$ and $\Pi(m, w|z) \leq 0, \forall z \neq m, w$ and the inequality is strict for at least one $z\}$
- *Feasible Controversial:* $FC = \{(m, w) : \Pi(m, w|m) = \Pi(m, w|w) = 0$ and there exist $z_1 \neq z_2$ such that $\Pi(m, w|z_1) > 0$ and $\Pi(m, w|z_2) < 0\}$
- *Forbidden:* $F = \{(m, w) : \Pi(m, w|m) < 0$ or $\Pi(m, w|w) < 0\}$

Any matching, A, belonging to the setwise-stable set must include all the required matches, since the endpoints of such a match will block otherwise, and cannot contain any forbidden matches, which would be severed by at least one of the endpoints. In addition, matching $A_1 = A \cup FP$ is stable, since it does not degrade any player and possibly improves the utility of some players. Similarly, matching $A_2 = A \cup FP \setminus FN$ is also stable. Thus, if the optimistic setwise-stable set is non-empty, it must contain a matching that has all the matches in FP, but none of FN. Thus the only difficulty lies in choosing a subset of the feasible controversial edges. By parameterizing the number of such edges, we can obtain an FPT algorithm for determining non-emptiness of the optimistic setwise-stable set (Algorithm 3.5.2).

Algorithm 3 FPT: Optimistic Setwise-Stable Matching

$A \leftarrow \emptyset$;
for all $(m, w) \in (M, W)$ **do**
  // Include (m, w) *if required or feasible positive*;
  if $(m, w) \in R$ **or** $(m, w) \in FP$ **then**
    $A \leftarrow A \cup \{(m, w)\}$;
  end if
end for
// Check all possible combinations of feasible controversial matches (FC);
for all $FC' \subseteq FC$ **do**
  $A^* \leftarrow A \cup FC'$;
  if *stable*(A^*) **then**
    return A^*;
  end if
end for
// The optimistic setwise-stable set is empty;
return $NULL$;

Theorem 20. *Checking non-emptiness of the optimistic setwise-stable set is FPT with respect to the number of feasible controversial matches.*

Proof. Let k be the number of feasible controversial matches. From previous observations, if the optimistic setwise-stable set is non-empty, it must contain a matching A, such that $R \subseteq A$, $FP \subseteq A$, $A \cap FN = \emptyset$, and $A \cap F = \emptyset$. Matching A may contain some feasible controversial matches, and Algorithm 3.5.2 checks all the possible subsets of FC. Thus if a stable matching exists, the algorithm will find it. The algorithm must be preceded by a pre-processing step that computes, for every pair (m, w), the best case utilities of players m, w if they are matched with each other. The pre-processing step can be completed in $O(n^4)$, checking the stability of a matching in $O(n^3)$, and so the runtime of Algorithm 3.5.2 is $O(n^4 + 2^k n^3)$. ■

The following two theorems consider the case where externalities are positive, and negative, respectively.

Theorem 21. *The unique candidate for the optimistic setwise-stable set under negative externalities is the empty matching.*

Proof. Let A be a non-empty matching satisfying optimistic setwise stability, (m, w) an edge in A, ands z a player. Since (m, w) has negative externalities, $\Pi(m, w|z) < 0$, and so z's utility in A is suboptimal. Thus z will attempt to disrupt A by severing some of his matches and hoping that in response, $N \setminus \{z\}$ will organize themselves in the best possible way for z. The only case in which z cannot deviate is when z is unmatched. This holds for every $z \in N$, and so A must be the empty matching. ■

Theorem 22. *The only candidates for the optimistic setwise-stable set under positive externalities are the complete and the empty matchings.*

Proof. If $\Pi(m, w|m) \geq 0$ and $\Pi(m, w|w) \geq 0$ for all m, w, the complete matching is the only setwise-stable matching. Otherwise, assume $\Pi(m, w|m) < 0$ for some m, w, and let A be an optimistic setwise-stable matching. Any player $z \neq m, w$ satisfies $\Pi(m, w|z) > 0$. Thus z receives suboptimal utility in A in the absence of (m, w), and so z will block hoping the deviation will result in the formation of this match. Since A is stable, z is not deviating, hence it must be that z is unable to deviate, i.e. z is unmatched. Agent z was arbitrarily chosen from $N \setminus \{m, w\}$, thus A is the empty matching. ■

We note that many alternative definitions of optimism are possible. A natural and less extreme version of optimistic reasoning is for the players to not assume that others will initiate matches with them upon a deviation. In this case, any benefit from a deviation can only come from the removal of negative matches, or from a change in externalities.

3.5.3 Pessimistic Stability

Definition 30 (Pessimistic Setwise-Stable Set)**.** *A matching A belongs to the pessimistic setwise-stable set if there is no coalition B that can block under the assumption that $N \setminus B$ reacts in the worst possible way for B.*

The idea underlying pessimism is that the society punishes deviating behaviour maximally. Similarly to the optimistic case, every player z in a blocking coalition B assumes that $N \setminus B$ will organize themselves in the worst possible way for z. Again, there may be inconsistencies among the reactions that the members of B expect from $N \setminus B$, and thus the worst possible outcome may not happen for each member of B. Under pessimism, the players are willing to accept many more outcomes compared to the neutral or optimistic setwise-stable sets, since they are convinced the deviation can trigger the worst possible outcome for them. However, the pessimistic setwise-stable set can still be empty, as Example 3 illustrates.

Theorem 23. *Any fair matching satisfies pessimistic individual rationality.*

Proof. If a player z with $u(z) > 0$ deviates, the set $N \setminus \{z\}$ can always form the empty matching, which would decrease z's utility to zero. ■

Thus every fair matching is a candidate for the pessimistic setwise stable set. Under positive externalities, the following stronger condition holds.

Theorem 24. *Any matching satisfying pessimistic setwise-stability under positive externalities is fair.*

Proof. Let A be a matching in the pessimistic setwise-stable set and assume there exists player z with negative utility in A. Agent z cannot be unmatched, since all the unmatched players receive non-negative utility. Thus z is matched and can deviate by cutting all his matches. The worst matching that $N \setminus \{z\}$ can form is the empty matching. Thus, by deviating, z would receive a worst case utility of zero, which is better than his current utility in A. Since A is stable, z cannot exist. ■

In general, a deviating coalition B will assume that $N \setminus B$ forms the complete matching under negative externalities, and the empty matching under positive externalities.

Example 6. *Let $M = \{m_1, m_2\}, W = \{w_1, w_2\}$ with $\Pi(m_i, w_i|m_i) = \Pi(m_i, w_i|w_i) = \varepsilon, \Pi(m_i, w_i|m_j) = \Pi(m_i, w_i|w_j) = -\Delta$, where $i \neq j$ and $\Delta \gg \varepsilon > 0$. The pessimistic setwise-stable set is the empty matching, while no matching satisfies neutral setwise-stability.*

Unlike the neutral setwise-stable set, the pessimistic setwise-stable set of Example 4 is nonempty when $\delta \gg \varepsilon$. However, the pessimistic setwise-stable set can still be empty. The following example is based on the reduction in Theorem 16 and contains an unfeasible knapsack instance, with $U = \{u_1, u_2\}$, $s(u_1) = 3$, $s(u_2) = 5$, $v(u_1) = 2$, $v(u_2) = 10$, $B = 4$, and $K = 5$.

Example 7 (Empty Pessimistic Setwise-Stable Set)**.** *Let* $G = (M, W, \Pi)$ *with* $M = \{x_1, x_2, m_1, m_2\}$, $W = \{y_1, y_2, w_1, w_2\}$, *and* $\Pi(x_1, y_1|m_1) = -3$, $\Pi(x_2, y_2|m_1) = -5$, $\Pi(x_1, y_1 \mid m_2) = 2$, $\Pi(x_2, y_2|m_2) = 10$, $\Pi(m_1, w_1|m_1) = -4$, $\Pi(m_2, w_2|m_2) = -7$, *and* $\Pi(m_i, y_j|y_j) = -1$, $\Pi(x_i, w_j|x_i) = -1$, *for all* $i, j \in \{1, 2\}$*. If a matching does not contain edge* (x_2, y_2)*, then coalition* $\{x_1, y_1, x_2, y_2, m_2, w_2\}$ *can block by ensuring that matches* $\{(x_1, y_1), (x_2, y_2), (m_2, w_2)\}$ *are formed (possibly in addition to others). This would give player* m_2 *a utility of* 5*, which is strictly better than the maximum obtained by* m_2 *in the absence of match* (x_2, y_2)*. On the other hand, if a matching does contain edge* (x_2, y_2)*, then coalition* $\{x_1, y_1, x_2, y_2, m_1, w_1\}$ *can block by forming* (m_1, w_1) *and ensuring that none of the matches* $(x_1, y_1), (x_2, y_2)$ *occur. This deviation would give player* m_1 *a utility of* -4*, which is strictly better than the maximum obtained in the presence of match* (x_2, y_2)*.*

Next we consider the complexity of computing matchings in the pessimistic setwise-stable set.

Theorem 25. *Checking pessimistic setwise-stable set membership is coNP-complete.*

Proof. The reduction from Theorem 17 applies, by noting that in matching A of Theorem 17, a coalition is blocking under neutral reasoning if and only if it is blocking under pessimistic reasoning. ■

Theorem 26. *Checking nonemptiness of the pessimistic setwise-stable set is NP-hard.*

Proof. The same reduction as in Theorem 16 applies, by noting that a coalition is blocking for the constructed matching under pessimistic reasoning if and only if it is blocking under neutral reasoning. ■

Our results with respect to the complexity of many-to-many matchings are summarized in Table 3.5.3.

Table 3.1: Setwise Stability in Many-to-Many Matchings with Externalities

Setwise-Stable Set	*Optimistic*	*Neutral*	*Pessimistic*
Membership	P	*co*NP*-complete*	*co*NP*-complete*
Nonemptiness	NP*-complete*	NP*-hard*	NP*-hard*

3.5.4 Contractual Stability

From previous sections, the optimistic setwise-stable sets can often be characterized by extreme outcomes, such as the empty and complete matchings. These outcomes arise under optimistic reasoning because of the excessive expectations of the players, who refuse to settle for anything less than the best. The neutral setwise-stable set is still too permissive: a player never accepts to maintain a negative match, even when that match indirectly helps them obtain a very high utility (Example 4). On the other hand, under pessimistic reasoning, the players may accept inadequate matchings on the basis of them not being the worst possible. In this section we strive for a more balanced stability concept, that is not as permissive as the optimistic and neutral setwise-stable sets, yet not as negative as the pessimistic set.

Definition 31 (Contractual Setwise-Stable Set)**.** *A matching A belongs to the contractual setwise-stable set if there is no coalition B that can block under the assumption that*

(i) the players from $N \setminus B$ which are hurt by the deviation maximally punish B.

(ii) the players which are not harmed by the deviation do not react.

Every player in B evaluates independently their punishment from $N \setminus B$ (if any), when contemplating deviation $A^{'}$ from A. Denote by H the players that degrade in $A^{'}$ compared to A, and by I the players in $N \setminus B$ unharmed by $A^{'}$, where $H \cup I = N \setminus B$. Then every z in B assumes H will sever all the matches positive for z and will form all the matches negative for z. The players in I do not react to the deviation, and so they are not involved in the formation of new matches. However, I might also lose those matches that H decides to sever when punishing z. This results in matching $A^{''}_z$, which is the worst possible punishment that H can give z without any help from I or B. If z is still better off in $A^{''}_z$ compared to A, then z agrees to participate in blocking.

Recall that for $\delta \gg \varepsilon$, the setwise-stable set of Example 4 is empty. However, the contractual setwise-stable set consists of the complete matching, which is the outcome that one would intuitively expect. Thus under contractual reasoning players may accept to maintain negative-valued matches as long as long as those matches are valued by everyone else. Nevertheless, the contractual setwise-stable set can be empty (Example 3).

An Oceanic Game

In this section we consider an example where the contractual setwise-stable set is a natural stability concept. Consider a type of oceanic game [37] where the players are divided into two disjoint sets: big players and small players. The interaction between big players has a big influence on everyone, while the interaction between small players has a small influence. The interaction between a big player and a small player also has a small influence.

Definition 32 (Big-player small-player game)**.** *A big-player small-player game is defined as a matching game (M, W) such that B and S are disjoint sets of "big" players and "small" players respectively, where $B \cup S = N$. For the game to be nontrivial, we require that there are big players and small players on both sides of the market. For each $b, b^{'} \in B$, $\Pi(b, b^{'}|z) \in \{k\Delta \mid k \in \mathbb{Z}\}$, for all $z \in N$. For all other players $x, y \in N$ such that at least one of them is a small player, $\Pi(x, y|z) \in \{k\varepsilon \mid k \in \mathbb{Z}\}$, for all $z \in N$. In addition, the matches involving two big players have a much bigger influence than any other matches. Formally, we require that $\Delta > n^2\varepsilon$.*

Theorem 27. *Let (M, W) be a big-player small-player game. Then any matching that satisfies weak contractual setwise-stability in the original game (M, W) also satisfies weak contractual setwise-stability in the game reduced to the set of big players, $(M, W|B)$.*

Proof. Let A be a matching in the contractual setwise-stable set of (M, W) and A_B the same matching in the reduced game $(M, W|B)$. Assume by contradiction that A_B is unstable and let $D \subseteq B$ be a weakly blocking coalition, that can perform deviation $\mathcal{D}$ in A_B. From the Π function definition, it must be that each member of D improves with at least Δ as a result of $\mathcal{D}$. Consider now the following deviation, $\mathcal{D}^{'}$, that coalition D can perform in A:

- Set the matches among D identically in $\mathcal{D}^{'}$ and $\mathcal{D}$.
- Set the matches between D and $D \setminus B$ identically in $\mathcal{D}^{'}$ and $\mathcal{D}$.
- Leave the matches between D and S the same as in A.

Observe that the deviators D are not removing any matches with the small players in $\mathcal{D}^{'}$. This ensures that the big players hurt by deviation $\mathcal{D}^{'}$ are the same as those hurt by deviation $\mathcal{D}$. However, it is possible that at least one of the following happens:

(i) Some small players are hurt by deviation $\mathcal{D}'$.

(ii) Some deviators are hurt by maintaining the matches with the small players in $\mathcal{D}'$.

In each of the cases $(i), (ii)$, the additional loss in utility that each $d \in D$ incurs in $\mathcal{D}'$ compared to $\mathcal{D}$ can only come from small edges, and so is bounded by $n^2\varepsilon$. Since each $d \in D$ gained at least Δ from $\mathcal{D}$, it follows that each player $d \in D$ gains at least $\Delta - n^2\varepsilon > 0$ by performing $\mathcal{D}'$. Thus D is a blocking coalition for A, contradiction with A being stable. It follows that the assumption must have been false, and A_B is stable. ■

3.5.5 Relation between the Sets

Given a matching game, let $\mathcal{N}$-set(G) denote the neutral setwise-stable set, $\mathcal{P}$-set(G) the pessimistic setwise-stable set, $\mathcal{O}$-set(G) the optimistic setwise-stable set, and $\mathcal{C}$-set(G) the contractual setwise-stable set.

Theorem 28. *Given any matching game G, the following inclusions hold*

$$\mathcal{O}\text{-}set(G) \subseteq \mathcal{N}\text{-}set(G) \subseteq \mathcal{C}\text{-}set(G) \subseteq \mathcal{P}\text{-}set(G)$$

Proof. Let B denote a potentially blocking coalition. If a matching belongs to the optimistic set, then B cannot hope to improve even when the rest of the players organize themselves in the best possible way for B. Thus the optimistic set is included in all the other stable sets. If a matching belongs to the pessimistic set, then B cannot improve when the rest of the society will punish them maximally for the deviation. Under the other stability concepts, B assumes a potentially better reaction from $N \setminus B$. Thus if B's deviation is not profitable in the optimistic, neutral, or contractual scenarios, it is also not profitable in the pessimistic scenario. Hence the pessimistic set contains all the other sets. Finally, we show the neutral set is included in the contractual set. Let $A \in \mathcal{N}$-set(G) be a matching. Since A is setwise-stable, for every coalition B that contemplates deviation A', there exists a player $z \in B$ that does not improve in A'. Under contractual behaviour, the same player z either does not improve, or can expect to be hurt by $N \setminus B$'s reaction to A'. Thus if player z does not participate in blocking under neutral behaviour, z still does not participate in blocking under contractual behaviour. Hence $A \in \mathcal{C}$-set(G). ■

3.5.6 Social Welfare

Having discussed the issue of stability in many-to-many matching games, in the following section we discuss the social welfare and how it can be maximized. The social welfare of a matching A is defined as: $SW(A) = \sum_{z \in N} u(z, A)$.

We call a matching optimal if it maximizes social welfare among all the possible matchings. Finding the optimal matching is an important question in combinatorial optimization. The optimal one-to-one matching can be found in polynomial time using the Hungarian algorithm [13]. In our model, the optimal matching can be determined greedily.

Theorem 29 (Optimal Matching)**.** *In many-to-many matching games with additive externalities, the optimal matching can be computed in polynomial time.*

Proof. For all $(m, w) \in (M, W)$, let $\phi(m, w) = \sum_{z \in N} \Pi(m, w|z)$. The value $\phi(m, w)$ represents the influence of (m, w) on the welfare of any matching that includes it. Then the optimal matching can be determined greedily, by taking all the matches with non-negative influence on social welfare: $A^{OPT} = \{(m, w) | \phi(m, w) \geq 0\}$ ■

While the optimal matching is the best solution for the society overall, it is not necessarily stable:

Example 8 (Unstable Optimal Matching)**.** *Let $M = \{m\}$ and $W = \{w_1, \ldots, w_t\}$. Define $\Pi(m, w_i|m) = -\varepsilon$ and $\Pi(m, w_i|w_i) = \Delta$, with $\Delta > \varepsilon > 0$. The complete matching, A^{OPT} is optimal. However, A^{OPT} is not individually rational, since $u(m, A^{OPT}) = -n\varepsilon < 0$. Agent m can improve its utility to zero by severing all its matchings.*

There exist, however, games where the optimal matching is stable. When the influence of any edge (m, w) is always positive or always negative, that is $\Pi(m, w|z) \geq 0$, for all $z \in N$, or $\Pi(m, w|z) \leq 0$ for all $z \in N$, the optimal matching satisfies all the stability concepts discussed.

3.6 One-to-One Matchings

In this section we analyze compact externalities in the context of one-to-one matchings. In addition, we analyze the version of Γ-pairwise stability in which a pair can deviate if and only if both deviators improve *strictly* as a result of the deviation, which is consistent with the original definition of the stable marriage problem. In general, the requirement that the blocking coalition improves strictly with the deviation is known as *weak stability*, and examples of such solution concepts include the weak core and the weak pairwise stable set.

3.6.1 Pessimistic Stability

Definition 33 (Pessimistic Core)**.** *Matching A belongs to the pessimistic core if there is no coalition B that can block under the assumption that $N \setminus B$ reacts in the worst possible way for B.*

The core is a solution concept very similar to the setwise-stable set and is commonly used in one-to-one matchings. Unlike the setwise-stable set, the core does not allow deviators to maintain matches with the non-deviators.

Theorem 30. *Computing a maximal pessimistic pairwise stable maching is NP-hard.*

Proof. Recall that finding a maximal pairwise stable outcome in the stable marriage problem with ties and incomplete lists (SMTI) is known to be NP-complete [38]. Any SMTI instance can be expressed using the Π function, by setting all the externality influences to zero, $\Pi(m, w|m) = -\infty$, whenever player m prefers being alone to being matched with w, and $\Pi(m, w_1|m) = \Pi(m, w_2|m)$ whenever player m is indifferent between w_1 and w_2. ■

In addition, we consider a less extreme notion of pessimism. Namely, when Π is non-negative, any matching containing singletons (unmatched players) can weakly improve everyone's utility by pairing the singletons with each other. Based on this observation, the pessimistic deviators can have a less extreme attitude, and assume that while the rest of the players may punish them for the deviation, they will not stay unmatched.

Definition 34 (Restricted Pessimistic Core)**.** *A matching A belongs to the* restricted pessimistic core *if there is no coalition B that can block under the assumption that $N \setminus B$ forms the worst possible matching for B, among all the matchings of size*

$$\min\left(|M \cap (N \setminus B)|, |W \cap (N \setminus B)|\right)$$

The restricted pessimistic core is only defined when $\Pi \geq 0$.

Restricted pessimistic pairwise stability is defined similarly, by limiting to 2 the size of the blocking coalition to at most two. Both the pessimistic and restricted pessimistic pairwise stable sets are always nonempty, and a stable solution can be computed in polynomial time when $\Pi \geq 0$ (Algorithm 3.6.1).

Theorem 31. *A restricted pessimistic pairwise stable matching can be computed in $O(n^5)$ when Π is non-negative.*

Proof. First note that the runtime of Algorithm 3.6.1 is $O(n^5)$. Let A be the matching returned by Algorithm 3.6.1 and assume by contradiction that A is not stable under pessimistic pairwise stability. Then there exists deviating pair $(m, w^{'})$, where the matches of the two players in A are (m, w) and $(m^{'}, w^{'})$, respectively. Then it must be the case that for any possible matching $A^{'}(m, w^{'})$ that includes the pair $(m, w^{'})$, both m and $w^{'}$ are better off in $A^{'}(m, w^{'})$ than in A. Equivalently:

$$E^{-}_{m}(w^{'}) > u(m, A) \geq E^{-}_{m}(w)$$

and

$$E^{-}_{w^{'}}(m) > u(w^{'}, A) \geq E^{-}_{w^{'}}(m^{'})(^{*})$$

However, matching A is stable under E^{-}, and so for any $(m, w^{'}) \notin A$, either $E^{-}_{m}(w^{'}) \leq E^{-}_{m}(w^{'})$ or $E^{-}_{w^{'}}(m^{'}) \leq E^{-}_{w^{'}}(m)$, contradiction with (*). Thus A must be stable. ■

Corollary 3. *A pessimistic pairwise stable matching can be computed in $O(n^5)$ when Π is non-negative.*

Proof. When $\Pi \geq 0$, the utility that pessimistic deviators expect upon a deviation is lower than the utility expected by restricted pessimistic deviators. Since the matching returned by Algorithm 3.6.1 is not blocked by restricted pessimistic deviators, it also cannot be blocked by pessimistic deviators. ■

The matching returned by Algorithm 3.6.1 also satisfies pessimistic pairwise stability.

A pessimistic pairwise stable matching can be computed in $O(n^2)$ by using Algorithm 3.6.1, which simply runs the Gale-Shapley algorithm by ignoring externalities. However, as the next example illustrates, the restricted pessimistic core can be empty.

Example 9. *(Empty Restricted Pessimistic Core) Consider the game $G = (M, W, \Pi)$, with $M = \{m_1, m_2, m_3\}$, $W = \{w_1, w_2, w_3\}$, and Π defined as follows:*

- *For $z_1 \in \{m_1, w_1\}$, let $\Pi(m_1, w_1|z_1) = \Pi(m_2, w_2|z_1) = \Pi(m_3, w_2|z_1) = 6$, $\Pi(m_1, w_3|z_1) = \Pi(m_3, w_1 \mid z_1) = 7$, $\Pi(m_1, w_2|z_1) = \Pi(m_2, w_1|z_1) = \Pi(m_3, w_3|z_1) = 9$, and $\Pi(m_2, w_3|z_1) = 0$.*

- *For $z_2 \in \{m_2, w_2\}$, let $\Pi(m_i, w_i|z_2) = 6$, $\forall i \in \{1, 2, 3\}$, $\Pi(m_1, w_2|z_2) = \Pi(m_2, w_1|z_2) = 7$, $\Pi(m_2, w_3|z_2) = \Pi(m_3, w_2|z_2) = 9$, $\Pi(m_1, w_3|z_2) = \Pi(m_3, w_1| \ z_2) = 0$.*

- *For $z_3 \in \{m_3, w_3\}$, let $\Pi(m_i, w_i|z_3) = 6$, $\forall i \in \{1, 2, 3\}$, $\Pi(m_2, w_1|z_3) = 0$, $\Pi(m_1, w_2|z_3) = 13$, $\Pi(m_2, w_3|z_3) = \Pi(m_3, w_2|z_3) = 7$, and $\Pi(m_1, w_3|z_3) = \Pi(m_3, w_1) = 9$.*

Algorithm 4 Restricted Pessimistic Pairwise Stability ($\Pi \geq 0$)

for all $z \in N$ **do**
    // If $z \in M$ *then* $P(z) = W$*, else* $P(z) = M$;
    for all $t \in P(z)$ **do**
        $M^{'} = M \setminus \{z, t\}$;
        $W^{'} = W \setminus \{z, t\}$;
        for all $(m, w) \in (M^{'}, W^{'})$ **do**
            $\psi(m, w) \leftarrow \Pi(m, w|z)$;
        end for
        // Compute A^-*, the worst case matching for* z *when paired with* t;
        A^- = Hungarian-Matching$(M^{'}, W^{'}, \psi)$;
        $E_z^-(t) \leftarrow \Pi(z, t|z)$;
        for all $(m, w) \in A^-$ **do**
            $E_z^-(t) \leftarrow E_z^-(t) + \psi(m, w)$;
        end for
    end for
end for
// Return a stable matching according to preferences given by E^-*, with ties broken arbitrarily;*
A = Gale-Shapley(M, W, E^-);
return A;

Algorithm 5 Pessimistic and Neutral Pairwise Stability ($\Pi \geq 0$)

for all $(m, w) \in (M, W)$ **do**
    $E_m(w) = \Pi(m, w|m)$;
end for
// Break ties in E *arbitrarily;*
A = Gale-Shapley(M, W, E);
return A;

The possible matchings are:

- $A_1 = \{(m_1, w_2), (m_2, w_1), (m_3, w_3)\}$
- $A_2 = \{(m_1, w_1), (m_2, w_3), (m_3, w_2)\}$
- $A_3 = \{(m_1, w_3), (m_2, w_2), (m_3, w_1)\}$
- $A_4 = \{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$
- $A_5 = \{(m_1, w_2), (m_2, w_3), (m_3, w_1)\}$
- $A_6 = \{(m_1, w_3), (m_2, w_1), (m_3, w_2)\}$

It can easily be verified that the players have the following preferences over the possible matchings:

- *Players* $z_1 \in \{m_1, w_1\} : (A_1 >_{z_1} A_6 >_{z_1} A_4 >_{z_1} A_3 >_{z_1} A_5 >_{z_1} A_2)$,
- *Players* $z_2 \in \{m_2, w_2\} : (A_2 >_{z_2} A_1 >_{z_2} A_4 >_{z_2} A_5 =_{z_2} A_6 >_{z_2} A_3)$,
- *Players* $z_3 \in \{m_3, w_3\} : (A_3 >_{z_3} A_2 >_{z_3} A_1 >_{z_3} A_4 >_{z_3} A_5 =_{z_3} A_6)$.

We obtain that coalition $\{(m_2, w_3), (m_3, w_2)\}$ *is blocking matching* A_1, $\{(m_1, w_3), (m_3, w_1)\}$ *is blocking* A_2, $\{(m_1, w_2), (m_2, w_1)\}$ *is blocking* A_3 *and* A_5, $\{(m_2, w_3), (m_3, w_2)\}$ *is blocking* A_4 *and* A_6, *and so* G *has an empty restricted pessimistic core.*

Note that any matching in the pessimistic core is pairwise stable under pessimism, and so the candidates for the pessimistic core are the solutions to the stable marriage problem given by E^-. However, determining whether a given matching belongs to the pessimistic core is coNP-complete.

Theorem 32. *Checking pessimistic core membership is coNP-complete.*

Proof. The reduction is similar to that of Theorem 17, except for the addition of several dummy players to ensure that the grand coalition can always block in the one-to-one setting when the *KNAPSACK* instance has a solution. Given $I = \langle U, s, v, B, K \rangle$, let $G = (M, W, \Pi)$ such that $M = \{x_1, \ldots, x_{2n}, m_1, m_2\}$, $W = \{y_1, \ldots, y_{2n}, w_1, w_2\}$, and Π with non-zero entries:

- $\Pi(x_i, y_i|m_1) = \Pi(x_i, y_{n+i}|m_1) = -s(u_i)$ and $\Pi(x_i, y_i|w_1) = \Pi(x_i, y_{n+i}|w_1) = v(u_i)$, $\forall i \in \{1, \ldots, n\}$
- $\Pi(m_2, w_2|m_1) = -B - \varepsilon$ and $\Pi(m_2, w_2|w_1) = K - \varepsilon$, for some $0 < \varepsilon < 1$
- $\Pi(m_1, w|w) = \Pi(m, w_1|m) = -1$, for all $w \in W \setminus \{w_1\}$ and $m \in M \setminus \{m_1\}$

Let matching $A = \{(m_2, w_2), (m_1), (w_1), (x_1), \ldots, (x_{2n}), (y_1), \ldots, (y_{2n})\}$. It can be easily verified that if I has a solution $U' \subseteq U$, the grand coalition can block through matching:

$$\begin{aligned} A_B &= \left(\bigcup_{u_i \in U'} \{(x_i, y_i), (x_{n+i}, y_{n+i})\}\right) \cup \left(\bigcup_{u_i \notin U'} \{(x_i, y_{n+i+1}), (x_{n+i+1}, y_i)\}\right) \\ &\qquad \cup \{(m_1, w_1), (m_2), (w_2)\}, \end{aligned}$$

where $x_{2n+1} = x_1$ and $y_{2n+1} = y_1$.

Conversely, if A is blocked by a coalition B through some matching A_B, the only players that could improve their utility compared to A are m_1 and w_1. It can be shown that the matches with non-zero externalities in A_B can be used to give a solution to the I instance. ∎

Theorem 33. *Checking nonemptiness of the pessimistic core is NP-hard.*

Proof. The reduction is similar to that of Theorem 16, except for the addition of several dummy players to ensure that the grand coalition can always block in the one-to-one setting when the *KNAPSACK* instance has a solution. Given $I = \langle U, s, v, B, K\rangle$, let $G = (M, W, \Pi)$ such that $M = \{x_1, \ldots, x_{2n}, m_1, m_2\}$, $W = \{y_1, \ldots, y_{2n}, w\}$, and Π with non-zero entries:

- $\Pi(x_i, y_i|m_1) = \Pi(x_i, y_{n+i}|m_1) = -s(u_i)$ and $\Pi(x_i, y_i|w) = \Pi(x_i, y_{n+i}|w) = v(u_i)$, $\forall i \in \{1, \ldots, n\}$
- $\Pi(m_1, w|m_1) = -B$
- $\Pi(m_2, w|m_2) = K - \sum_{u_i \in U} v(u_i)$
- $\Pi(x_j, w|x_j) = -1$ and $\Pi(m_i, y_j|y_j) = -1$, $\forall i \in \{1, 2\}$ and $\forall j \in \{1, \ldots, 2n\}$

Let matching $A = \{(m_1), (m_2), (w), (x_1), \ldots, (x_{2n}), (y_1), \ldots, (y_{2n})\}$. Similarly to Theorem 16, A belongs to the pessimistic core if and only if the *KNAPSACK* instance has a solution. ∎

Checking membership to the restricted pessimistic core can also be shown to be coNP-complete, by using a reduction from the following NP-complete problem.

EXACT COVER BY 3-SETS (X3C)

***Input*:** Pair (R, S), where $R = \{1, \ldots, n\}$ and $S = \bigcup_{r,s,t}(r, s, t)$ is a collection of 3-element subsets of R.

***Question*:** Does S contain an exact cover of R, that is, a collection $S' \subseteq S$ such that every element of R occurs in exactly one member of S'?

Theorem 34. *Checking restricted pessimistic core membership is coNP-complete.*

Proof. Given instance (R, S) of X3C, we construct a pair (G, A), where G is a matching game with non-negative Π and A a matching, such that A has a blocking coalition under pessimism if and only if (R, S) admits an exact cover. First note that given a coalition B, it can be verified in $O(n^3)$ if B is blocking for A.

For each $r \in R$, construct two players r and d_r, such that $\Pi(r, d_r|r) = 2n + 2$ and $\Pi(r, d_r|d_r) = 1$. For each unordered triple $(r, s, t) \in S$, construct six players:

$$\{X^r_{r,s,t}, X^s_{r,s,t}, X^t_{r,s,t}, Y^r_{r,s,t}, Y^s_{r,s,t}, Y^t_{r,s,t}\}$$

such that $\forall z \in \{r, s, t\}$, Π has the following values:

- $\Pi(X^z_{r,s,t},\ Y^z_{r,s,t} \mid X^z_{r,s,t}) = 5$
- $\Pi(X^z_{r,s,t}, Y^z_{r,s,t}|Y^z_{r,s,t}) = 1$
- $\Pi(z, X^z_{r,s,t}|z) = 5$
- $\Pi(z, X^z_{r,s,t}|X^z_{r,s,t}) = 2$
- For all $v \neq z, v \in \{r, s, t\}$: $\Pi(v, X^v_{r,s,t}|X^z_{r,s,t}) = 2$
- For all $w \in R \setminus \{z\}$: $\Pi(z, X^z_{r,s,t}|w) = 2$

The unspecified Π entries are set to zero. Note that since the triple (r, s, t) is unordered, $X^r_{r,s,t} = X^r_{s,r,t}$, for example. Let $G = (M, W, \Pi)$ be a matching game given by

- $M = R \cup \left(\bigcup_{(r,s,t)\in S}\{Y^r_{r,s,t}\ ,\ Y^s_{r,s,t}\ ,\ Y^t_{r,s,t}\}\right)$,

- $W = \left(\bigcup_{r\in R} d_r\right) \cup \left(\bigcup_{(r,s,t)\in S}\{X^r_{r,s,t}, X^s_{r,s,t}, X^t_{r,s,t}\}\right)$,
- Π as specified above.

Consider the following matching:

$$A = \left(\bigcup_{r\in R}(r, d_r)\right) \cup \left(\bigcup_{(r,s,t)\in S}\{(X^r_{r,s,t}, Y^r_{r,s,t}), (X^s_{r,s,t}, Y^s_{r,s,t})(X^t_{r,s,t}, Y^t_{r,s,t})\}\right)$$

We claim that A has a blocking coalition if and only if the X3C instance has an exact cover. If (R, S) has an exact cover $T \subseteq S$, consider coalition

$$B^* = \left(\bigcup_{(r,s,t)\in T}\{X^r_{r,s,t}X^s_{r,s,t}, X^t_{r,s,t}\}\right) \cup R \tag{3.5}$$

For each $r \in R$, let $(r, s^r, t^r) \in T$ be the triplet that r belongs to in the T cover. Note that player r's utility in A is:

$$u(r, A) = \Pi(r, d_r|r) = 2n + 2$$

while player X_{r,s^r,t^r} has utility:

$$u(X_{r,s^r,t^r}, A) = \Pi(X_{r,s^r,t^r}, Y_{r,s^r,t^r}|X_{r,s^r,t^r}) = 5$$

In addition, $u(d_r, A) = \Pi(r, d_r|d_r) = 1$ and $u(Y_{r,s^r,t^r}, A) = \Pi(X_{r,s^r,t^r}, Y_{r,s^r,t^r}|Y_{r,s^r,t^r}) = 1$.

If deviating with B^*, player r obtains utility:

$$u(r, B^*) = \Pi(r, X^r_{r,s^r,t^r}|r) + \sum_{z\in R\setminus\{r\}} \Pi(z, X^z_{z,s^z,t^z}|r) = 5 + 2(n-1) = 2n + 3$$

For each $z \in \{r, s, t\}$, player $X^z_{r,s,t}$ receives the following utility if deviating with B^*:

$$u(X^z_{r,s,t}, B^*) = \Pi(z, X^z_{r,s,t}|X^z_{r,s,t}) + \sum_{v\in\{r,s,t\}\setminus\{z\}} \Pi(v, X^v_{r,s,t}|X^z_{r,s,t}) = 2 + 2 + 2 = 6$$

Thus each member of B^* improves with the deviation, and so B^* is a blocking coalition for A.

Conversely, assume that A has a blocking coalition. Observe that each of d_r and $Y^r_{r,s,t}$ achieve their maximum possible utility in A, and so these players cannot be involved in blocking. Thus any blocking coalition, B, for A, must contain players from the following set:

$$R \cup \left(\bigcup_{(r,s,t)\in T} \{X^r_{r,s,t}, X^s_{r,s,t}, X^t_{r,s,t}\} \right)$$

Since a blocking coalition must contain players from both M and W, $R \subset M$, and no $X^z_{r,s,t}$ is a member of M, it follows that B contains at least one player r from R. Let $|B \cap R| = k$ and denote by $X^r_{r,s,t}$ the player that r is matched with in B, for some $s, t \in R$. Player r is pessimistic, so when deviating, will assume that all of $R \setminus B$ are matched in the worst possible way for him. That is, r assumes that each $v \in R \setminus B$ remains matched with d_v, and r cannot expect any utility from externalities formed outside B. Then player r's estimated utility in B is:

$$u(r, B) = \Pi(r, X^r_{r,s,t}|r) + \sum_{z\in(B\cap R)\setminus\{r\}} \Pi(z, X^z_{z,s^z,t^z}|r) = 5 + 2(k-1) = 2k + 3$$

where X^z_{z,s^z,t^z} is player z's match in B. For r to be blocking, it must be that $u(r, B) > u(r, A)$, which is equivalent to $2k + 3 > 2n + 2$, i.e. $k \geq n$.

That is, if A has a blocking coalition, it must involve all the members of R. Consider now the players of the form $X^r_{r,s,t}$ for some $(r, s, t) \in S$. Since $X^r_{r,s,t}$ is deviating, the player can no longer be matched with $Y^r_{r,s,t}$. Thus player $X^r_{r,s,t}$ can receive utility only from: (a) her own match with r, (b) the externality of $(s, X^s_{r,s,t})$, and (c) the externality of $(t, X^t_{r,s,t})$, in case these matches form. Thus

$$u(X^r_{r,s,t}, B) \leq \Pi(r, X^r_{r,s,t}|X^r_{r,s,t}) + \Pi(s, X^s_{r,s,t}|X^r_{r,s,t}) + \Pi(t, X^t_{r,s,t}|X^r_{r,s,t}) = 6$$

Since $u(X^r_{r,s,t}, A) = 5$, it must be the case that all of the matches (a), (b), and (c) form, and so for every player $X^r_{r,s,t}$ that participates in blocking, all three matches $(z, X^z_{r,s,t})$, $\forall z \in \{r, s, t\}$ form. For each player $r \in R$, denote by (r, s^r, t^r) the triplet that r is matched with. Then it must be that $B = B^*$ (*), and so the following set is an exact cover for (R, S): $T' = \bigcup_{r\in R}(r, s^r, t^r)$ ■

3.6.2 Neutral Stability

Definition 35 (Neutral Core)**.** *Matching A belongs to the neutral core if no coalition B can block under the assumption that $N \setminus B$ will not react to their deviation.*

We note that neutral stability may seem unreasonable in the context of one-to-one matchings, since at first sight it implies that the previous partners of a deviating pair will not form a match with each other. However, the important idea underlying neutral stability is that the deviating pair uses the neutral stability value to estimate the value they may be getting if they deviate. This does not necessarily imply that they assume the rest of the players will remain matched the same way as before.

Theorem 35. *A neutral pairwise stable matching can be computed in $O(n^2)$ when Π is non-negative.*

Proof. First note that Algorithm 3.6.1 runs in $O(n^2)$. Let A be the matching returned by Algorithm 3.6.1 and assume by contradiction that A is not stable under neutral pairwise stability. Then there exists some deviation $(m, w^{'})$, where the matches of the two players in A are (m, w) and $(m^{'}, w^{'})$, respectively. For $(m, w^{'})$ to be blocking, it must be the case that

$$u(m, A) = \Pi(m, w|m) + ext(m, A) < \Pi(m, w^{'}|m) + ext(m, A) - \Pi(m^{'}, w^{'}|m)$$

and

$$u(w^{'}, A) = \Pi(m^{'}, w^{'}|w^{'}) + ext(w^{'}, A) < \Pi(m, w^{'}|w^{'}) + ext(w^{'}, A) - \Pi(m, w|w^{'})$$

The two conditions are equivalent to

$$(1)\ \Pi(m, w|m) < \Pi(m, w^{'}|m) - \Pi(m^{'}, w^{'}|m) \leq \Pi(m, w^{'}|m)$$

and

$$(2)\ \Pi(m^{'}, w^{'}|w^{'}) < \Pi(m, w^{'}|w^{'}) - \Pi(m, w|w^{'}) \leq \Pi(m, w^{'}|w^{'})$$

By using the notation in Algorithm 3.6.1, Inequalities (1) and (2) imply that $E_m(w) < E_m(w^{'})$ and $E_{w^{'}}(m^{'}) < E_{w^{'}}(m)$, which is equivalent to the pair $(m, w^{'})$ being blocking in A under the preferences given by E. However A is the result of running the Gale-Shapley algorithm on (M, W, E), and so A must be stable. This is a contradiction, thus A is pairwise stable under neutrality. ■

However, as the next example illustrates, when Π can be negative the existence of a neutral pairwise-stable matching is no longer guaranteed.

Example 10 (Empty Neutral Pairwise-Stable Set)**.** *Let $G = (M, W, \Pi)$, where $M = \{m_1, m_2\}$, $W = \{w_1, w_2\}$, and Π as follows:*

- $\Pi(m_i, w_j|m_i) = \Pi(m_i, w_j|w_j) = 1$*, for all* $i, j \in \{1, 2\}$

- $\Pi(m_i, w_j|z) = -1$, *for all* $i, j \in \{1, 2\}$ *and* $z \neq m_i, w_j$

Since all the edges have positive values for their endpoints, the only candidates for neutral pairwise stability are:

- $A_1 = \{(m_1, w_1), (m_2, w_2)\}$
- $A_2 = \{(m_1, w_2), (m_2, w_1)\}$

However, matching A_1 *is blocked by the pair* (m_1, w_2)*, while* A_2 *is blocked by* (m_1, w_1)*.*

The next example illustrates the separation between the neutral pairwise-stable set and the neutral core.

Example 11 (Empty Neutral Core)**.** *Let* $G = (M, W, \Pi)$*, where* $M = \{m_1, m_2\}$*,* $W = \{w_1, w_2\}$*, and* Π *as follows:*

- $\Pi(m_i, w_j|m_i) = \Pi(m_i, w_j|w_j) = -\varepsilon$
- $\Pi(m_i, w_j|z) = W \gg \varepsilon > 0$*, where* $z \in N \setminus \{m_i, w_j\}$*,* $\forall i, j \in \{1, 2\}$*.*

All of the matchings in which at least one edge forms, such as:

- $A_1 = \{(m_1, w_1), (m_2, w_2)\}$
- $A_2 = \{(m_1, w_2), (m_2, w_1)\}$
- $A_3 = \{(m_1, w_1), \{m_2\}, \{w_2\}\}$

fail individual rationality, since both players m_1 *and* w_1 *can sever edge* (m_1, w_1)*, and increase their utility by doing so. The "empty" matching:*

$$A_0 = \{\{m_1\}, \{m_2\}, \{w_1\}, \{w_2\}\}$$

gives utility of zero to all the players, and is blocked by the grand coalition, which can agree to form matching A_1*, and obtain the following utility for all the players:*

$$u(z, A_1) = W - \varepsilon > 0 = u(z, A_0), \forall z \in N$$

Thus G *has an empty neutral core, while matching* A_0 *satisfies neutral pairwise stability. We contrast this with the pessimistic core of* G*, which consists of matchings* A_1 *and* A_2*, and with the neutral pairwise stable set, which consists of matching* A_0*.*

We have the following results with respect to the complexity of the neutral core.

Theorem 36. *Checking neutral core membership is coNP-complete.*

Proof. The same reduction as in Theorem 17 applies, by noting that the weights are set such that any feasible matching (from the perspective of the x_i and y_i players) is one-to-one. In other words, a coalition can block the matching constructed in Theorem 17 if and only if it is a one-to-one matching. ■

Theorem 37. *Checking nonemptiness of the neutral core is NP-hard.*

Proof. The same reduction as in Theorem 16 applies, by noting that the weights are set up such that any feasible matching (from the perspective of the x_i and y_i players) is one-to-one, and so a matching of the game in Theorem 16 is stable if and only if it is a one-to-one matching. ■

3.6.3 Optimistic Stability

Definition 36 (Optimistic Core)**.** *Matching A belongs to the optimistic core if there is no coalition B that can block under the assumption that the players in $N \setminus B$ react in the best possibly way for B.*

That is, when coalition B considers deviation A' from A, every player $z \in B$ evaluates the deviation assuming the players in $N \setminus B$ will organize themselves in the best possible way for z.

Similarly to the result for many-to-many matchings (Theorem 38), the optimistic pairwise stable set is equivalent to the optimistic core in the one-to-one setting.

Example 12. *(Empty Optimistic Core) Let $M = \{m_1, m_2, p_1\}$, $W = \{w_1, w_2, p_2\}$, and Π defined as follows:*

- $\Pi(m_i, w_j|m_i) = \Pi(m_i, w_j|w_j) = 1, \forall i, j \in \{1, 2\}$
- $\Pi(p_1, p_2|p_1) = \Pi(p_1, p_2|p_2) = 1$
- $\Pi(p_1, w_i|p_1) = \Pi(m_i, p_2) = \varepsilon$, *where* $0 < \varepsilon < 1$
- $\Pi(p_1, w_i|z) = \Pi(m_i, p_2|z) = +\infty$, $\forall z \in \{m_1, m_2, w_1, w_2\}$

All the remaining Π values are set to zero. Note that pair (p_1, p_2) forms in any stable structure, since both p_1 and p_2 receive their maximum utility when matched with each other, and they do not get any utility from externalities. Thus if the optimistic core is non-empty, it can only contain the following two matchings:

- $A_1 = \{(m_1, w_1), (m_2, w_2), (p_1, p_2)\}$
- $A_2 = \{(m_1, w_2), (m_2, w_1), (p_1, p_2)\}$

In both A_1 and A_2, each of the players m_1, m_2, w_1, w_2 receive utility of 1. The pair (m_1, w_2) is blocking A_1, since both m_1 and w_2 hope that matching

$$A^* = \{(m_1, w_2), (m_2, p_2), (p_1, w_1)\}$$

arises as a result of the deviation, giving both of them utility:

$$u(m_1, A^*) = \Pi(m_1, w_2|m_1) + \Pi(m_2, p_2|m_1) + \Pi(p_1, w_1|m_1) = 1 + \infty + \infty = +\infty$$

Similarly, A_2 is blocked by (m_1, w_1), who hope that

$$A^{**} = \{(m_1, w_1), (m_2, p_2), (p_1, w_2)\}$$

arises as a result of the deviation. Thus the game has an empty optimistic core. On the other hand, the pessimistic and neutral cores of G are non-empty, and consist of matchings: $\{A_1, A_2\}$.

Theorem 38. *The optimistic pairwise stable set is equivalent to the weak core.*

Proof. Let A be any matching. If a coalition C is blocking for A, then there must exist pair (m, w) that forms in C as a result of the deviation. The best value that (m, w) can hope for while deviating with C is bounded by the best value that (m, w) can obtain in general. It follows that (m, w) is also blocking for A. Thus if A is pairwise stable under optimistim, then A also belongs to the optimistic core. The reverse inclusion holds trivially, and so the core is equivalent to the pairwise stable set under optimism. ■

Theorem 39. *Checking nonemptiness of the optimistic core is NP-complete.*

Proof. First note that similarly to the optimistic setwise-stable set, checking nonemptiness of the optimistic core is in NP. Given a matching A, the utility of each player $z \in N$ in A can be computed in $O(n^2)$. In addition, for every pair $(m, w) \in (M, W)$, we can again compute in $O(n^2)$ the best case utilities of m and w when matched with each other. Verifying if A is stable can be done by iterating over all pairs (m, w) and checking in $O(1)$ if both m and w can (weakly) improve by deviating under optimism, compared to their current utility in A. The same reduction as in Theorem 16 applies, by noting that the weights are constructed such that any feasible matching (from the perspective of the x_i and y_i players) is one-to-one. ■

Moreover, the problem of checking nonemptiness of the optimistic pairwise stable set remains NP-complete even when Π is non-negative. We contrast this with Algorithm 3.6.1 and Algorithm 3.6.1 for computing a (restricted) pessimistic and neutral pairwise stable matching. We provide a proof using a reduction from the NP-complete problem *Exact Cover by 3-Sets*.

Theorem 40. *Checking nonemptiness of the optimistic core is NP-complete even when Π is non-negative.*

Proof. We provide a reduction from X3C. Given instance (R, S) of X3C, we construct matching game G, such that G has a non-empty optimistic core if and only if (R, S) admits an exact cover.

For each $r \in R$, construct a matching game $G_r = (M_r, W_r, \Pi_r)$, where:

- $M_r = \{m_1^r, m_2^r, p_1^r\}$
- $W_r = \{w_1^r, w_2^r, p_2^r\}$
- Π is set as in Example 12

We refer to p_1^r as the “representative” of r, and note that the game $(M_r \setminus \{p_1^r\}, W_r, \Pi_r)$ has a nonempty optimistic core. In particular, the following matching is stable:

$$A_r = \{(m_1^r, p_2^r), (m_2^r, w_2^r), \{w_1^r\}\}$$

In addition, for each unordered triple $(r, s, t) \in S$, let there be “committee members” $X_{r,s,t} = \{x_{r,s,t}^r, x_{r,s,t}^s, x_{r,s,t}^t\}$ with the role of bringing the triple (r, s, t) together, and dummy players $Y_{r,s,t} = \{y_{r,s,t}^r, y_{r,s,t}^s, y_{r,s,t}^t\}$, to be matched with $X_{r,s,t}$ in case (r, s, t) is not chosen in the cover.

Denote edge $(x_{r,s,t}^r, p_1^r)$ by $e_{r,s,t}$, and let its influence be the following:

$$\Pi(e_{r,s,t}^r | z) = 15, \forall z \in \{p_1^r\} \cup X_{r,s,t}$$

Similarly for $e^s_{r,s,t}$ and $e^t_{r,s,t}$.

Thus if a matching contains the set of edges:

$$A_{r,s,t} = \{(x^r_{r,s,t}, p^r_1), (x^s_{r,s,t}, p^s_1), (x^t_{r,s,t}, p^t_1)\},$$

then each player in $X_{r,s,t}$ receives utility 45. The non-zero Π values for $Y_{r,s,t}$ are set as follows, for all $i, j \in \{r, s, t\}, i \neq j$:

- $\Pi(x^i_{r,s,t}, y^i_{r,s,t} | x^i_{r,s,t}) = 14$
- $\Pi(x^j_{r,s,t}, y^j_{r,s,t} | x^i_{r,s,t}) = 12$
- $\Pi(x^i_{r,s,t}, y^i_{r,s,t} | y^i_{r,s,t}) = 1$

Thus when the following matching forms:

$$B_{r,s,t} = \{(x^r_{r,s,t}, y^r_{r,s,t}), (x^s_{r,s,t}, y^s_{r,s,t}), (x^t_{r,s,t}, y^t_{r,s,t})\},$$

we have that $u(x^i_{r,s,t}) = 38$ and $u(y^i_{r,s,t}) = 1$.

Let $G = (M, W, \Pi)$, where

- $M = \left(\bigcup_{r \in R} M_r\right) \cup \left(\bigcup_{(r,s,t) \in S} Y_{r,s,t}\right)$
- $W = \left(\bigcup_{r \in R} W_r\right) \cup \left(\bigcup_{(r,s,t) \in S} \ X_{r,s,t}\right)$
- Π as above, with all unspecified values set to zero.

We now show that the X3C instance has an exact cover if and only if G has a stable optimistic core. If (R, S) has an exact cover $T \subseteq S$, then the following matching belongs to the optimistic core of (M, W, Π):

$$\begin{aligned} A^*(T) \ &= \left(\bigcup_{r \in R} A_r\right) \cup \left(\bigcup_{(r,s,t) \in T} A_{r,s,t}\right) \cup \left(\bigcup_{(r,s,t) \in S \setminus T} B_{r,s,t}\right) \\ &\cup \left(\bigcup_{(r,s,t) \in T} \{y^r_{r,s,t}\}, \{y^s_{r,s,t}\}, \{y^t_{r,s,t}\}\right) \end{aligned}$$

Observe that everyone obtains their maximum possible utility in $A^*(T)$, except:

- The players in $X_{r,s,t}$, for all $(r,s,t) \in S \setminus T$, which form matching $B_{r,s,t}$. For each $z \in \{r,s,t\}$, player $X^z_{r,s,t}$ is matched with their "dummy" pair, $Y^z_{r,s,t}$, obtaining utility 38. The most that $X^z_{r,s,t}$ can hope for, if deviating from $B_{r,s,t}$, is a utility of 30, which can come from the externalities of $(X^s_{r,s,t}, p^s_1)$ and $(X^t_{r,s,t}, p^t_1)$.
- The players in $Y_{r,s,t}$, for all $(r,s,t) \in T$. Each player $Y^z_{r,s,t}$ is not influenced by any externality, and can only receive a utility of 1 when matched with $X^z_{r,s,t}$. Thus they have no incentive to block with any other player.

Conversely, assume that (M, W, Π) has a nonempty optimistic core, and let A be a stable matching. Note that A cannot contain any match of the form $(p^r_1, p^r_2), \forall r \in R$ (see Example 12). Thus it must be that every p^r_1 is matched with a $x^r_{r,s,t}$, for some $(r,s,t) \in S$.

Consider any triple $(r,s,t) \in S$. We claim that either (a) all of $X_{r,s,t}$ are matched with p^r_1, p^s_1, p^t_1, respectively, or (b) all of $X_{r,s,t}$ are matched with $y^r_{r,s,t}, y^s_{r,s,t}, y^t_{r,s,t}$, respectively. We show this by considering the alternatives:

Case (c)**:** Two of the players in $X_{r,s,t}$, say $x^r_{r,s,t}$ and $x^s_{r,s,t}$, are matched with p^r_1 and p^s_1, respectively, while $x^t_{r,s,t}$ is matched with $y^t_{r,s,t}$. Then player $x^r_{r,s,t}$ obtains utility:

$$u(x^r_{r,s,t}) = \Pi(x^r_{r,s,t}, p^r_1 | x^r_{r,s,t}) + \Pi(x^s_{r,s,t}, p^s_1 | x^r_{r,s,t}) + \Pi(x^t_{r,s,t}, y^t_{r,s,t} | x^r_{r,s,t}) = 15+15+12 = 42,$$

while $y^r_{r,s,t}$ is unmatched, and so obtains zero. Then pair $(x^r_{r,s,t}, y^r_{r,s,t})$ is blocking. Player $x^r_{r,s,t}$ hopes that matching

$$\{(x^r_{r,s,t}, y^r_{r,s,t}), (x^s_{r,s,t}, p^s_1), (x^t_{r,s,t}, p^t_1)\}$$

arises after the deviation, giving her utility:

$$\begin{aligned} u^{'}(x^r_{r,s,t}) &= \Pi(x^r_{r,s,t}, y^r_{r,s,t} | x^r_{r,s,t}) + \Pi(x^s_{r,s,t}, p^s_1 | x^r_{r,s,t}) + \Pi(x^t_{r,s,t}, p^t_1 | x^r_{r,s,t}) = \\ & \qquad 14 + 15 + 15 = 44 > 42, \end{aligned}$$

while player $y^r_{r,s,t}$ would get:

$$u^{'}(y^r_{r,s,t}) = \Pi(x^r_{r,s,t}, y^r_{r,s,t} | y^r_{r,s,t}) = 1 > 0$$

Case (d)**:** One of the players in $X_{r,s,t}$, say $x^r_{r,s,t}$, is matched with p^r_1, while $x^s_{r,s,t}$ and $x^t_{r,s,t}$ are matched with $y^s_{r,s,t}$ and $y^t_{r,s,t}$, respectively. Player $x^r_{r,s,t}$ obtains utility:

$$u(x^r_{r,s,t}) = \Pi(x^r_{r,s,t}, p^r_1 | x^r_{r,s,t}) + \Pi(x^s_{r,s,t}, y^s_{r,s,t} | x^r_{r,s,t}) + \Pi(x^t_{r,s,t}, y^t_{r,s,t} | x^r_{r,s,t}) = 39,$$

while $y^r_{r,s,t}$ gets zero. Then pair $(x^r_{r,s,t}, y^r_{r,s,t})$ is blocking. Player $x^r_{r,s,t}$ hopes that the following matching arises after the deviation:

$$\{(x^r_{r,s,t}, y^r_1), (x^s_{r,s,t}, p^s_1), (x^t_{r,s,t}, p^t_1)\},$$

giving him utility:

$$\begin{aligned} u(x^r_{r,s,t}) &= \Pi(x^r_{r,s,t}, y^r_1 | x^r_{r,s,t}) + \Pi(x^s_{r,s,t}, p^s_1 | x^r_{r,s,t}) + \Pi(x^t_{r,s,t}, p^t_1 | x^r_{r,s,t}) = \\ & 14 + 15 + 15 = 44 > 39, \end{aligned}$$

while player $y^r_{r,s,t}$ would get:

$$u'(y^r_{r,s,t}) = \Pi(x^r_{r,s,t}, y^r_{r,s,t} | y^r_{r,s,t}) = 1 > 0$$

Thus cases (c) and (d) are not feasible in a stable matching, and so for every set $X_{r,s,t}$, one of (a), (b) must hold. Hence every representative p^r_1 is matched with some $x^r_{r,s,t}$, and whenever $x^r_{r,s,t}$ is matched with p^r_1, it follows that both p^s_1 and p^t_1 are also matched with $x^s_{r,s,t}$ and $x^t_{r,s,t}$, respectively.

For every representative p^r_1, denote by $O_A(r)$ the set $X_{r,s,t}$ containing p^r_1's pair in matching A, and let $O_A = \bigcup_{r \in R} O(r, A)$. Then it follows that $A = A^*(O_A)$, and so O_A is an exact cover for (R, S). ∎

We note that while determining non-emptiness of the optimistic core is NP-complete, one can implement a more efficient search using the following observation. Denote by $E^+_m(w)$ the optimal (best case) utility that player m can get while in a matching with w. The value $E^+_m(w)$ can be computed in polynomial time, similarly to $E^-_m(w)$ (Algorithm **??**). In the next theorem we show that any matching belonging to the optimistic core must satisfy pairwise stability according to E^+, and so the search for the optimistic core can be restricted to those matchings which are pairwise stable under $E^+_m(w)$.

Lemma 6. *Any matching in the optimistic core is pairwise stable under E^+.*

Proof. Let A be a matching belonging to the optimistic core and assume by contradiction that A is not stable under E^+. Then A has a deviating pair (m, w') under E^+, and so the following hold:

- $u(m, A) \leq E^+_m(w) < E^+_m(w')$
- $u(w', A) \leq E^+_{w'}(m') < E^+_{w'}(m)$

Table 3.2: Core Stability in One-to-One Matchings with Externalities

Core	*Optimistic*	*Neutral*	*Pessimistic*
Membership	P	*co*NP*-complete*	*co*NP*-complete*
Nonemptiness	NP*-complete*	NP*-hard*	NP*-hard*

Table 3.3: Pairwise Stability in One-to-One Matchings with Externalities ($\Pi \geq 0$)

Pairwise Stability	*Optimistic*	*Neutral*	*Restricted Pessimistic*	*Pessimistic*
Membership	P	P	P	P
Nonemptiness	NP*-complete*	P	P	P

These conditions imply that $(m, w^{'})$ is blocking A under optimism, contradiction with the stability of A. ■

Our results with respect to the complexity of one-to-one matchings are summarized in Table 3.2 and Table 3.3.

3.7 Related Work

Klaus *et al* [30] consider pairwise and setwise stability, with the weak and strong variants, in many-to-many matching markets. Echenique *et al* [18] study several solution concepts, such as the setwise-stable set, the core, and the bargaining set in many-to-many matchings. These models do not account for externalities or the boundedness of the players as they use exponential preference profiles. Externalities in the classical marriage problem have been introduced by Sasaki and Toda [44], and in one-to-many models by Dutta and Masso [16], in both cases for complete preference profiles. Certain parallels can be drawn between our gamma-stability concept and that considered in Sasaki and Toda for one-to-one matchings, where the deviating pairs have estimations on the reaction of the other players. Nevertheless, the crucial difference is that our general Γ-stability defines how society *will* react to the deviation, which may differ from the players' predictions. In other words, the Γ-response is always correct, but possibly hard to compute, while the estimations in Sasaki and Toda can be wrong and even completely unrelated to the preferences.

Weighted preferences have been introduced in matchings by Pini *et al.* [35], in which the players rank each other using a numerical value. However, they study solution concepts different from ours, such as α-stability and link-additive stability, and do not consider externalities. The work of Laszlo Koczy [32] considers games in partition function form and formulates a very natural definition of a recursive core for such games, in which the deviators only deviate if they can be better off, regardless of what equilibrium forms in the residual game, if any. If the residual game does not have any stable structure, the deviators have to ensure they are better off regardless of what configuration forms in the residual.

Chapter 4

Conclusions and Future Work

4.1 Models of Interaction in Social Networks

This work is a step in the direction of understanding network interactions from the perspective of coalitional game theory. We formulated an intuitive mathematical model, analysed its welfare and stability properties, gave an approximation of the optimal welfare, and showed that core stable structures have small world characteristics. We studied the efficiency of the core and proposed two solution concepts with improved welfare guarantees.

This work can be extended in several ways. We would like to look at power indices and see how the degree and position of a node in the network are correlated with the welfare of that node in the equilibrium. It would be interesting to characterize the extent to which a node contributes to social welfare or to stabilizing the game, and to identify stabilizers in existent networks. It also remains to be determined whether stable structures are small worlds under general utility functions that reflect homophily. More generally, we are interested in game theoretic models to simulate the formation of social networks.

Finally, this type of game can be used to formulate alternative solution concepts that apply to social networks in general. Given that social networks are typically very large, we are mainly concerned with tractability and universal existence. One such direction is to take into account the notion of distance: players are likely to try deviating with close neighbours, rather than with players that are far away in the network.

4.2 Matchings with Compact Externalities

In this work we introduced the first compact model for matching problems with externalities. Our formulation can be easily used to model externalities one-to-one, one-to-many, and many-to-many matchings. We also defined a general stability concept, Γ stability, which provides a framework for reasoning about classes of problems and different stability solution concepts. We studied instantiations of Γ stability in the context of many-to-many matchings and one-to-one matchings. In many-to-many matchings, we considered optimistic, pessimistic, neutral, and contractual reasoning under a solution concept called setwise stability. In one-to-one matchings, we considered optimistic, pessimistic, restricted pessimistic, and neutral reasoning under the pairwise and core stability solution concepts. In both the many-to-many and one-to-one settings, we studied the computational complexity of finding stable outcomes and provided both hardness results and polynomial algorithms.

This work can be extended in several ways. A coherent theory of stability for matchings with externalities remains to be further developed. The existing work on the stability of matchings assumes complete preference profiles, which require exponential space and reasoning abilities on behalf of the players. Such modeling is not realistic when dealing with bounded rational players. Furthermore, externalities are important in many other domains, including network formation games. It has been noted that effects due to externalities do appear in networks [29], however we are not aware of a theory of stability for network formation games that explicitly takes into account externalities. Since such games usually involve very large number of players, a compact model of externalities is crucial. Another possible future direction is to formulate solution concepts that consider recursive reasoning in the context of matchings, such as the recursive core [32].

References

[1] K. Aadithya, B. Ravindran, T. Michalak, and N. Jennings. Efficient computation of the shapley value for centrality in networks. In *WINE*, 2010. 28

[2] Noga Alon, Michal Feldman, Ariel D. Procaccia, and Moshe Tennenholtz. A note on competitive diffusion through social networks. *Information Processing Letters*, 110(6):221–225, 2010. 28

[3] Itai Ashlagi, Piotr Krysta, and Moshe Tennenholtz. Social context games. In *4th International Workshop on Internet and Network Economics (WINE)*, pages 675–683, 2008.

[4] Haris Aziz, Felix Brandt, and Hans Georg Seedig. Stable partitions in additively separable hedonic games. In *Proceedings of the Eighth International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2011. 5

[5] A-L. Barabasi and Z. N. Oltvai. Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.*, 5:101–113, 2004. 17

[6] Francis Bloch and Matthew O. Jackson. The formation of networks with transfers among players. Technical report, J. Economic Theory, 2004. 28

[7] A. Bogomolnaia and M. O. Jackson. The stability of hedonic coalition structures. *Games Econ. Behav.*, 38(2):201–230, 2002. 28

[8] Anna Bogomolnaia and Matthew O. Jackson. The stability of hedonic coalition structures. *Games and Economic Behaviour*, 38(2):201–230, 2002. 5, 30

[9] Christian Borgs, Jennifer T. Chayes, Jian Ding, and Brendan Lucier. The hitchhiker's guide to affiliation networks: A game-theoretic approach. *CoRR*, abs/1008.1516, 2010.

[10] Simina Brânzei and Kate Larson. Coalitional affinity games and the stability gap. In *Proceedings of the 21st International Joint Conference on Artifical Intelligence (IJCAI)*, pages 79–84, 2009. 5, 6, 19

[11] Simina Brânzei and Kate Larson. Social distance games. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, 2011.

[12] Simina Brânzei, Tomasz Michalak, Talal Rahwan, Kate Larson, and Nicholas Jennings. Matching problems with compact externalities: A stability analysis. *Working Paper*, 2011.

[13] Rainer Burkard, Mauro Dell'Amico, and Silvano Martello. *Assignment Problems*. Society for Industrial and Applied Mathematics (SIAM), 2009. 29, 52

[14] Chavdar Dangalchev. Residual closeness in networks. *Physica A: Statistical Mechanics and its Applications*, 365(2):556 – 564, 2006.

[15] Reinhard Diestel. *Graph Theory*. Springer-Verlag, Heidelberg, 2005. 24

[16] Bhaskar Dutta and Jordi Masso. Stability of matchings when individuals have preferences over colleagues. *Journal of Economic Theory*, 75:464–475, 1997. 70

[17] David Easley and Jon Kleinberg. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, 2010. 1

[18] Federico Echenique and Jorge Oviedo. A theory of stability in many-to-many matching markets. *Theoretical Econonomics*, 1(2):233–273, 2006. 29, 32, 70

[19] Federico Echenique and Bumin M. Yenmez. A solution to matching with preferences over colleagues. *Games and Economic Behaviour*, 59:46–71, 2005.

[20] Edith Elkind and Michael Wooldridge. Hedonic coalition nets. In *Proceedings of the Eighth International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2009. 5

[21] David Gale and Lloyd S. Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, 69:9–16, 1969. 29, 31

[22] Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979. 38

[23] Daniel Gomez, Enrique Gonzalez-Aranguena, Conrado Manuel, Guillermo Owen, Monica del Pozo, and Juan Tejada. Centrality and power in social networks: a game theoretic approach. *Mathematical Social Sciences*, 46(1):27–54, August 2003. 10

[24] M. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380, 1973. 8

[25] Jean-Loup Guillaume and Matthieu Latapy. Bipartite graphs as models of complex networks. *Combinatorial and Algorithmic Aspects of Networking*, pages 127–139, 2005.

[26] I. E. Hafalir. Stability of marriage with externalities. *International Journal of Game Theory*, 2008.

[27] Jochen Harant and Ingo Schiermeyer. On the independence number of a graph in terms of order and size. *Discrete Mathematics*, 232(1-3):131–138, 2001.

[28] Matthew O. Jackson. The stability and efficiency of economic and social networks. Springer-Verlag, 2001.

[29] Matthew O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008. 1, 2, 72

[30] Betina Klaus and Markus Walzl. Stable many-to-many matchings with contracts. *Journal Mathematical Economics*, 45(7-8):422–434, 2009. 70

[31] Jon Kleinberg. The small-world phenomenon: An algorithmic perspective. Technical report, 1999. 2, 28

[32] Laszlo Koczy. *Solution concepts and outsider behaviour in coalition formation games*. PhD thesis, Catholic University Leuven, 2003. 35, 42, 70, 72

[33] V. Latora and M. Marchiori. Efficient behavior of small-world networks. *Physics Review Letters*, 87(19), 2001. 9

[34] D. F. Manlove, R. W. Irving, K. Iwama, S. Miyazaki, and Y. Morita. Hard variants of stable marriage. *Theor. Comput. Sci.*, 276(1-2):261–279, 2002.

[35] Brent Venable Maria Silvia Pini, Francesca Rossi and Toby Walsh. Stability in Matching Problems with Weighted Preferences. In *ICAART*, 2011. 70

[36] M. McPherson, L. Smith-Lovin, and J. Cook. Birds of a feather: Homophily in social networks. *Annu. Rev. Sociol.*, 27:415–444, 2001. 9

[37] J. Milnor and L. Shapley. Values of large games 2: Oceanic games. *Mathematics of Operations Research*, 3(4):290–307, 1978. 50

[38] Francesca Rossi K. Brent Venable Mirco Gelain, Maria Silvia Pini and Toby Walsh. Local search for stable marriage problems with ties and incomplete lists. In *Pacific Rim International Conference on Artificial Intelligence (PRICAI)*, 2010. 53

[39] M. E. J. Newman. The mathematics of networks. *The New Palgrave Encyclopedia of Economics*, pages 1–12, 2007.

[40] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay Vazirani, editors. *Algorithmic Game Theory*. Cambridge University Press, 2007. 1, 19

[41] Kazuya Okamoto, Wei Chen, and Xiang-Yang Li. Ranking of closeness centrality for large-scale social networks. In *Proceedings of the 2nd annual international workshop on Frontiers in Algorithmics*, pages 186 – 195, 2008.

[42] Alvin E. Roth and Marilda O. Sotomayor. *Two-sided matching: A study in game-theoretic modeling and analysis*. Cambridge University Press, 1990. 29, 30

[43] Tuomas Sandholm and Victor Lesser. Coalitions among computationally bounded agents. *Artificial Intelligence*, 94(1-2):99–137, 1997. 35

[44] Hiroo Sasaki and Manabu Toda. Two-sided marriage problems with externalities. *Journal of Economic Theory*, 70:93 – 108, 1996. 70

[45] Yoav Shoham and Kevin Leyton-Brown. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, 2009. 4

[46] Marilda O. Sotomayor. Three remarks on the many-to-many stable matching problem. *Mathematical Social Sciences*, 38:55–70, 1999. 36

[47] Shao Sung and Dink Dimitrov. Computational complexity in additive hedonic games. *European Journal of Operational Research*, 203(3):635–639, 2010.

[48] N. Rama Suri and Y. Narahari. Determining the top-k nodes in social networks using the shapley value. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1509–1512, 2008. 28