Polyhedral Methods Applied to Extremal Combinatorics Problems

vorgelegt von
ANNIE RAYMOND
aus Montreal

von der Fakultät II - Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
-Dr. rer. nat.-

genehmigte Dissertation

Promotionausschuss:

Vorsitzende:	Prof. Dr. Gitta KUTYNIOK
Gutachter:	Prof. Dr. Martin GRÖTSCHEL
Gutachter:	Prof. Dr. Ralf BORNDÖRFER

Tag der wissenschaftlichen Aussprache: 23. Juni 2014

Berlin 2014
D83

Abstract

We study the polytopes describing two famous problems: the Turán hypergraph problem and the Frankl union-closed sets conjecture.

The Turán hypergraph problem asks to find the maximum number of r-edges in an r-uniform hypergraph on n vertices that does not contain a clique of size a. When $r = 2$, i.e. for graphs, the answer is well-known and can be found in Turán's theorem. However, when $r \geq 3$, the problem remains open. We model the problem as an integer program. We draw parallels between the stable set polytope and what we call the Turán polytope; we show that generalized and transformed versions of the clique, web, and wheel inequalities are also facet-defining for the Turán polytope. We also show that some other rank inequalities as well as doubling inequalities are facet-defining. These facets lead to a simple new polyhedral proof of Turán's theorem for graphs and yield some bounds in the case of the problem for hypergraphs.

The Frankl conjecture on the other hand states that there always exists an element contained in more than half of the sets of a non-empty union-closed family. We also model this problem as an integer program and discuss some helpful cuts and facets that can be added. The computations lead to a new conjecture: we claim that the maximum number of sets in a non-empty union-closed family where each element is present at most a times is independent of the number of elements n spanned by the sets if $n \geq \log_2(a) + 1$. We also present a minimization version of this conjecture and show that both conjectures are equivalent. We prove this new conjecture partially for $n \geq a$. By using a recent theorem of Balla, Bollobás and Eccles, we show that proving this new conjecture would imply that the original Frankl conjecture holds for all union-closed families of m sets when $\frac{2}{3}2^i \leq m \leq 2^i$ for some $i \in \mathbb{N}$. Moreover, proving the new conjecture would yield that any non-empty union-closed family of m sets contains an element in at least $\frac{6}{13}m$ of those sets, a much better bound than the current $\frac{m-1}{\log_2 m}$. Finally, we introduce the concept of twin sets and discuss its importance for the new conjecture.

The work on the Frankl conjecture was done jointly with Jonad Pulaj of the Konrad-Zuse-Zentrum and Dirk Theis of the University of Tartu.

Zusammenfassung

Wir untersuchen Polytope, die zwei bekannte Probleme beschreiben: das Hypergraphen-Problem von Turán und die Vermutung von Frankl.

Das Hypergraphen-Problem von Turán bestimmt die maximale Anzahl der r-Kanten in einem r-Hypergraph mit n Knoten, so dass der daraus entstandene r-Teil-Hypergraph keine Clique der Größe a enthält. Wenn $r = 2$ ist, also für Graphen, ist die Antwort bekannt und wird in Turáns Satz beschrieben. Jedoch, wenn $r \geq 3$ ist, bleibt das Problem offen. Wir modellieren das Problem als ein ganzzahliges lineares Programm. Wir ziehen eine Parallele zwischen dem Stabile-Mengen-Polytop und etwas, das wir Turán-Polytop nennen möchten. Damit zeigen wir, dass generalisierte und transformierte Versionen von den Cliquen-, Zirkulanten- und Rad-Ungleichungen auch facettendefinierend für das Turán-Polytop sind. Wir veranschaulichen, dass andere Rangungleichungen und Verdoppelungsungleichungen facettendefinierend sind. Diese Facetten führen zu einem einfachen neuen polyedrischen Beweis von Turáns Satz für Graphen und ergeben Schranken für das Problem für Hypergraphen.

Die Vermutung von Frankl besagt, dass es in jeder unter Vereinigung abgeschlossenen, endlichen Mengenfamilie ein Element gibt, das in mindestens der Hälfte der Mengen liegt. Wir modellieren dieses Problem auch als ein ganzzahliges lineares Programm und präsentieren hilfreiche Schnittebenen und Facetten, die hinzugefügt werden können. Die Rechenergebnisse führen zu einer neuen Vermutung: Wir behaupten, dass die maximale Anzahl der Mengen in einer Familie, in der jedes Element in maximal a Mengen enthalten ist, unabhängig von der Anzahl n der Elemente der Mengenfamilie ist, wenn $n \geq \log_2(a) + 1$ ist. Wir stellen auch eine Minimierungsversion dieser Vermutung vor und zeigen, dass beide Vermutungen äquivalent sind. Wir beweisen diese Aussage für $n \geq a$. Mit einem kürzlich vorgestellten Satz von Balla, Bollobás und Eccles zeigen wir, dass diese neue Vermutung impliziert, dass die Vermutung von Frankl für alle Familien mit m Mengen, $m \in [\frac{2}{3}2^i, 2^i]$, $i \in \mathbb{N}$, die unter Vereinigung abgeschlossen sind, wahr ist. Außerdem impliziert ein Beweis der neuen Vermutung, dass jede Familie die unter Vereinigung abgeschlossen ist, ein Element enthält, welches in mindestens $\frac{6}{13} \cdot m$ der m Mengen enthalten ist. Dies verbessert die aktuelle Schranke von $\frac{m-1}{\log_2 m}$ auf $\frac{6}{13} \cdot m$. Abschließend stellen wir das Konzept von Zwillings-Mengen vor, und diskutieren, welche wichtige Rolle es im Zusammenhang mit der neuen Vermutung spielt.

Die Ergebnisse bezüglich Frankls Vermutung wurden gemeinsam mit Jonad Pulaj von dem Konrad-Zuse-Zentrum und Dirk Theis von der Universität von Tartu erzielt.

Acknowledgements

First, I would like to thank my advisor, Martin Grötschel, who has given me as much freedom as possible to make my doctoral years as pleasant as possible. Every conversation I've had with him was a delight. I also want to thank my parents, Normand and Diane, who have supported me in everything I've ever undertaken. I also want to thank my brother for keeping me grounded throughout these years. My boyfriend Christoph Stüber has also been ever so patient while listening to me ranting about my research.

I also want to thank all of the incredible teachers and professors I've had the chance to meet along the way, but in particular, I want to thank: Paul Guertin who would put special extra questions for me in his exams and pushed me to go further than any teacher before; Margaret Readdy and Richard Ehrenborg who first introduced me to polytopes and gave me confidence to keep going with mathematics; Andreas Schulz who took me on as an undergraduate research assistant and who was so present; Rekha Thomas who opened her life to me and who remains my mentor; Michel Goemans whose class on combinatorial optimization was so beautiful; David Williamson who initiated me to approximation algorithms; Dirk Theis with whom I had the opportunity to work during a great week in Tartu; Ralf Borndörfer who was kind enough to read this thesis; and of course, Martin Grötschel, again, whose classes were so interesting and filled with historical anecdots.

I'm grateful to all of my fellow students and colleagues at MIT, the BMS and ZIB with whom I shared interesting conversations along the way, but in particular to: Doris Dobi, with whom I spent many late nights working on p-sets during my undergraduate days; Jean-Philippe Labbé, my old mathematician of a roommate; Christian Raack, Thomas Schlechte and Axel Werner who taught me a lot about soccer (sorry, football!); Jerry Gagelman and Moritz Schmitt, who were constant sources of encouragement; and, most of all, Fabio D'Andreagiovanni and Jonad Pulaj, without whom I couldn't have written this thesis.

Finally, I want to thank all of my friends all over the world who have always been there for me, no matter on which continent I was. Their love and support is what has driven me forward.

Contents

Introduction

The Turán Problem

One of the earliest theorems in combinatorics is Mantel's theorem from 1907 which states that the maximum number of edges in a graph on n vertices without any triangles is $\lfloor \frac{n^2}{4} \rfloor$, and that the maximum is attained only on complete bipartite graphs with parts of size $\lfloor \frac{n}{2} \rfloor$ and $\lceil \frac{n}{2} \rceil$.

This theorem was later generalized in 1941 by Turán who showed that the maximum number of edges in a graph on n vertices without any clique of size a is at most $(1 - \frac{1}{a-1})\frac{n^2}{2}$, and that this maximum is attained solely on complete $(a-1)$-partite graph with parts of size as equal as possible.

Since then, many different proofs of this theorem have been found using different techniques. We present a new polyhedral proof in theorem 1.2.5 by modeling the Turán problem as an integer program. Let

$$T(n,a,2) = \operatorname{conv}\left(\left\{x \in \{0,1\}^{\binom{n}{2}} \middle| \sum_{e \in E(Q^a)} x_e \le \binom{a}{2} - 1 \quad \forall Q^a \in \mathcal{Q}^a_{K_n}\right\}\right),$$

be the Turán polytope, where $\mathcal{Q}^a_{K_n}$ is the set of all cliques of size a in the complete graph on n vertices K_n. Then the goal of Turán's theorem is to find

$$ex(n,a,2) := \max\left\{\sum_{e \in E(K_n)} x_e \middle| x \in T(n,a,2)\right\}.$$

We prove in 1.2.5 that

$$ex(n,a,2) = \left\lfloor \frac{n}{n-2} \left\lfloor \frac{n-1}{n-3} \left\lfloor \frac{n-2}{n-4} \left\lfloor \cdots \left\lfloor \frac{a+2}{a} \left\lfloor \frac{a+1}{a-1} \cdot \left(\binom{a}{2} - 1 \right) \right\rfloor \right\rfloor \cdots \right\rfloor \right\rfloor \right\rfloor \right\rfloor$$

by using two types of valid inequalities of $T(n,a,2)$. Indeed, even though a complete linear description for $T(n,a,2)$ is unknown, it is possible to find the maximal value that $\sum_{e \in E(K_n)} x_e$ takes in this polytope.

Furthermore, we investigate some facet classes for $T(n,a,2)$. In theorems 1.2.23 and 1.2.24, we show that the clique inequality

$$\sum_{e \in E(Q^i)} x_e \leq ex(i,a,2)$$

is facet-defining for $T(n,a,2)$ for all $Q^i \in \mathcal{Q}^i_{K_n}$ with $i \neq 0 \mod (a-1)$ and $i \leq n$.

We then introduce the wheel graph W_l^a which consists of $l-1$ vertices forming a cycle and one vertex in its center with edges placed in such a way that every $a-1$ consecutive vertices on the cycle forms a clique of size a with the central vertex.

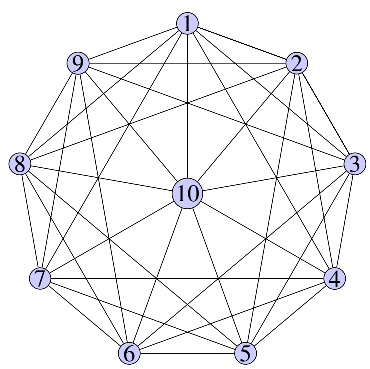

Figure 1: Wheel graph W_{10}^5.

In theorems 1.2.27 and 1.2.28, we show that the wheel inequality

$$\sum_{e \in E(W_l^a)} x_e \leq (a-1) \cdot (l-1) - \left\lceil \frac{l-1}{a-1} \right\rceil$$

defines a facet for $T(n,a,2)$ if $l-1=1 \mod (a-1)$ and $l \leq n$. Moreover, in theorem 1.2.33, we present a complete description for

$$T(W_l^3,3,2) := \text{conv}\left(\left\{\chi_S \in \{0,1\}^{|E(W_l^3)|} | S \in E(W_l^3) \text{ contains no clique of size } 3\right\}\right).$$

We then go on to consider web graphs $\overline{W}_{n'}^{a-1}$ which consist of n' vertices placed in a cycle such that every a consecutive vertices form a clique of size a, i.e. every vertex is connected to the next $a-1$ vertices on the cycle.

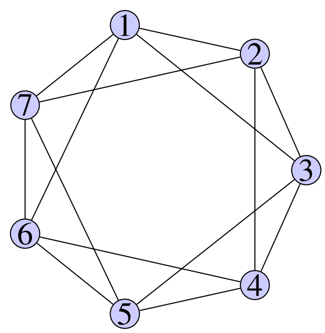

Figure 2: Web graph $\overline{W}_7^2$.

In theorems 1.2.36 and 1.2.37, we prove that the web inequality

$$\sum_{e \in E(\overline{W}_{n'}^{a-1})} x_e \leq (a-1) \cdot n' - \left\lceil \frac{n'}{a-1} \right\rceil$$

is facet-defining for $T(n,a,2)$ if $n'=1 \mod (a-1)$ and $n' \leq n$. In theorem 1.2.38, we also present a complete description for

$$T(\overline{W}_{n'}^2,3,2) := \text{conv}\left(\left\{\chi_S \in \{0,1\}^{|E(W_{n'}^2)|} | S \in E(W_{n'}^2) \text{ contains no clique of size } 3\right\}\right).$$

Finally, in theorems 1.2.39 and 1.2.40, we show that what we call the doubling inequality

$$\sum_{e \in \delta(v)} 2x_e + \sum_{e \in E(Q^i) \setminus \delta(v)} x_e \leq ex(i+1,a,2)$$

is facet-defining for $T(n,a,2)$ for any clique $Q^i \in \mathcal{Q}_{K_n}^i$ and for any $v \in V(Q^i)$ when $i=0 \mod (a-1)$, with $i \geq 3(a-1)$ and $n \geq i$.

Turán's theorem has since 1941 been generalized in two ways. First, by forbidding some other subgraph instead of a clique of a certain size; second, by considering the original problem applied to hypergraphs. We turn ourselves to this second generalization that Turán himself thought about after proving his theorem for graphs. The goal is to find the maximum number of r-hyperedges in a r-uniform hypergraph (i.e. a hypergraph for which every edge is composed of r vertices) on n vertices that does not contain any r-uniform hyperclique of size a.

Still, to this day, this problem remains unsolved. Even in the case when $a = 4$ and $r = 3$, the answer hasn't been found. Turán conjectured that in this case, the maximum number of edges is $\frac{4}{9}(1-o(1))\binom{n}{3}$. Many different attempts have been made and some bounds have been found, but a solution is still eluding researchers.

We decided to use the same approach that we used for graphs to attack the Turán hypergraph problem. We let

$$T(n,a,r) = \text{conv}\left(\left\{x \in \{0,1\}^{\binom{n}{r}} \middle| \sum_{e \in E(Q^a)} x_e \leq \binom{a}{r} - 1 \quad \forall Q^a \in \mathcal{Q}^a_{K^r_n}\right\}\right),$$

where $\mathcal{Q}^a_{K^r_n}$ is the set of all cliques of size a in the complete r-uniform hypergraph on n vertices K^r_n be the Turán polytope for hypergraphs. Then the goal is to find

$$ex(n,a,r) := \max\left\{\sum_{e \in E(K^r_n)} x_e \middle| x \in T(n,a,r)\right\}.$$

Unfortunately, the inequalities we used for the polyhedral proof of the graph problem are not enough to yield the conjectured solution. Already for $ex(7,4,3)$, we need to think of other techniques to obtain the optimal value; we present three different ways of doing so. We then turn again to facets of the polytope $T(n,a,r)$ to try to give us some more insights on the problem.

We prove in theorems 1.3.14, 1.3.15, 1.3.20 and 1.3.21 that the wheel and web facets generalize quite naturally to the hypergraph Turán polytope.

We also generalize some facets of $T(6,4,3)$ produced by the program Porta. All in all, we conclude that the Turán polytope is a very interesting polytope which deserves to be investigated in its own right, notwithstanding its connexion to the Turán hypergraph problem.

The Frankl Conjecture

A family of sets is said to be union-closed if the union of any two sets in the family is also a set in the family. In 1979, Frankl conjectured that there exists an element in at least half of the sets of any non-empty union-closed family.

Many mathematicians have tried to prove this, but were unsuccessful. Still, through their combined efforts and using methods coming from lattice theory, graph theory and probability, some results are known. Let n be the number of elements and m be the number of sets in a family. Then it is known that the conjecture holds for this family if $n \leq 12$, if $m \leq 50$, if $n \geq \frac{2}{3}2^m$, if $n \leq 2m$ and the family contains no blocks, if the family contains a singleton or a doubleton (i.e. a set containing only one or two elements respectively), or if the family can be represented as a lower semimodular lattice or by a subcubic graph. Moreover, it was proven that there always exists an element present in at least $\frac{m-1}{\log_2 m}$ sets of a family of m sets.

We decided to attack the problem from a new angle by formulating it as an integer program. Let $\mathcal{S}_n$ be the power set of the n elements $\{1,\dots,n\}$. We let

$$\begin{aligned}
f_n(a) := \max \sum_{S\in\mathcal{S}_n} x_S & \\
\text{s.t. } x_U + x_T \leq 1 + x_S & \qquad \forall T \cup U = S \in \mathcal{S}_n \\
\sum_{S\in\mathcal{S}_n : e\in S} x_S \leq a & \qquad \forall e \in [n] \\
x_S \in \{0,1\} & \qquad \forall S \in \mathcal{S}_n
\end{aligned}$$

and

$$\begin{aligned}
g_n(m) := \min \sum_{S\in\mathcal{S}_n : 1\in S} x_S & \\
\text{s.t. } x_U + x_T \leq 1 + x_S & \qquad \forall T \cup U = S \in \mathcal{S}_n \\
\sum_{S\in\mathcal{S}_n} x_S = m & \\
\sum_{S\in\mathcal{S}_n : i\in S} x_S - \sum_{S\in\mathcal{S}_n : j\in S} x_S \geq 0 & \qquad \forall 1 \leq i < j \leq n \\
x_S \in \{0,1\} & \qquad \forall S \in \mathcal{S}_n.
\end{aligned}$$

In other words, $f_n(a)$ is the maximum number of sets in a union-closed family on n elements in which every element is contained in at most a sets, and $g_n(m)$ is the minimum number of sets containing the most frequent element in a family of n elements on m sets. We then observe that the Frankl conjecture holds if and only if $f_n(a) \leq 2a$ for all $a, n \in \mathbb{N}$ if and only if $g_n(m) \geq \frac{m}{2}$ for all $n, m \in \mathbb{N}$.

We discuss some classes of valid inequalities for those programs and for two other formulations, which allow us to do some computations. These computations displayed a pattern that didn't seem to have been observed before and prompted us to formulate two new conjectures for union-closed families.

First, we conjecture in 2.7.1 that $f_n(a) = f_{n+1}(a)$ for all a and n such that $n \geq \lceil log_2\ a \rceil + 1$, that is, if there are enough elements for $f_n(a)$ not to be trivial. If $n < \lceil log_2\ a \rceil + 1$, then the power set of n contains at most a sets containing any given elements, and so the whole family will be present in an optimal solution.

Second, we conjecture in 2.7.2 that $g_n(m) = g_{n+1}(m)$ for all m and n such that $n \geq \lceil \log_2\ m \rceil$, that is, if there are enough elements for $g_n(m)$ to be feasible. If $n < \lceil \log_2\ m \rceil$, then the power set of n contains less than m sets, and so it is impossible to find a union-closed set family in this power set that contains m sets.

By proving some properties of the functions f and g, we show in theorem 2.7.14 that these two conjectures are equivalent.

We observe that these conjectures are different from the Frankl conjecture: neither do they imply the Frankl conjecture nor does the Frankl conjecture imply them. For example, if the f-conjecture is true, one would still need to show that $f_{\lceil \log_2 a \rceil + 1}(a) \leq 2a$ for all a to prove that the Frankl conjecture holds. Similarly, if the g-conjecture is true, one would still need to show that $g_{\lceil \log_2 m \rceil}(m) \geq \frac{m}{2}$ for all m to prove the Frankl conjecture. Conversely, if the Frankl conjecture holds, then certainly the value of $f_n(a)$ has to stabilize at some point as n increases, i.e. $\lim_{n \to \infty} f_n(a) = F_a$ for some $F_a \leq 2a$; however there is no reason why it should stabilize already at $n = \lceil \log_2 a \rceil + 1$. Equivalently, if the Frankl conjecture is true, then $g_n(m)$ must also stabilize as n increases, i.e. $\lim_{n \to \infty} g_n(m) = G_m$ where $G_m \geq \frac{m}{2}$; however, yet again, there is no reason why it should stabilize already at $n = \lceil \log_2 m \rceil$.

Still, those new conjectures would have a tremendous effect on the Frankl conjecture. If the f- and g-conjectures hold, then we could show that the Frankl conjecture holds for all families with m sets when $\frac{2}{3}2^i \leq m \leq 2^i$ for some $i \in \mathbb{N}$ (see corollary 2.7.4), thus proving that the Frankl conjecture holds in $\frac{2}{3}$ of all possible cases. Moreover, another consequence of proving the new conjectures would be that we could show that in any union-closed family on m sets, there exists an element in at least $\frac{6}{13} \cdot m$ sets of the family, a much better bound than the current $\frac{m-1}{\log_2 m}$ (see 2.7.5).

By introducing the notion of twin sets, i.e. sets that differ only in one element, we show in theorem 2.7.20 that $f_n(a) = f_{n+1}(a)$ when n is at least approximately $2a$. This result could potentially be improved by tweaking some lemmas. We obtain a better bound, $n \geq a$, by using a construction of Fargas-Ravry in theorem 2.7.21. We finally show in 2.7.22 that if $f_n(a) < f_{n+1}(a)$, then every element is the difference of at least two pairs of twin sets.

We conclude that more efforts must be put upon solving these two new conjectures, and that the concept of twins sets might be the best way to go.

Chapter 1

The Turán Problem

1.1 Preliminaries for the Turán Problem

1.1.1 Problem Statement

The typical goal in Turán-type problems is to find the maximum possible number of edges in a graph on n vertices that does not contain a copy of some graph H.

We consider a specific class of these problems, namely when the forbidden graph is a complete graph. From now on, we'll refer to this particular case of Turán-type problems as the *Turán problem*, which is stated as follows. Let $G = (V,E)$ be an undirected graph with vertex set V such that $|V| = n$ and edge set E such that it contains no clique of size a. What is the maximum size of E?

1.1.2 Some Notation

By $[n]$, we denote the set $\{1,2,\ldots,n\}$. In general, for a graph $G = (V,E)$ with $|V| = n$, we assume that $V = [n]$. For a graph G, $V(G)$ and $E(G)$ represent respectively the vertex set and edge set of G. All of the graphs we consider are undirected, and we use (v_1, v_2) to denote an edge formed by vertices v_1 and v_2, a notation sometimes used for directed edges only; we apologize for any confusion that ensues. For $G = (V,E)$, $\delta(v)$ for some vertex $v \in V$ corresponds to the set of edges in E that are adjacent to v and $d(v) := |\delta(v)|$ is the *degree* of v. We say a graph is *H-free* if it does not contain an induced subgraph isomorphic to H. We let K_n represent a complete graph on n vertices. We let $K_{c\times d}$ be the complete bipartite graph with d vertices on one side and d on the other. A vector $v \in \{0,1\}^n$ means that $v_i \in \{0,1\}$ for all $1 \leq i \leq n$. We let $\mathcal{Q}^i_G$ be the set of all cliques of size i in some graph G; a clique in this family is denoted by Q^i. Note that when we talk about cliques, we're mostly referring to their edge sets.

We let $ex(n,H,2)$ be the maximum number of edges in an H-free graph on n vertices. In the literature, the shorter notation $ex(n,H)$ is usually adopted; we add a "2" which refers to the the fact that we are now considering graphs, i.e. 2-uniform hypergraphs since we will later consider some general r-uniform hypergraphs, i.e. hypergraphs where every edge has cardinality r. We let $ex(n,K_a,2) = ex(n,a,2)$ for simplicity and call it the $(n,a,2)$-Turán number. We say an edge set on n vertices is *a-Turán* if it doesn't contain any clique of size a; we say it is *maximally a-Turán* if it is has as many edges as possible, i.e. if there are $ex(n,a,2)$ edges. We denote by $T(n,a,2)$ the convex hull of the characteristic vectors of all edge sets $F \subseteq E(K_n)$ that contain no clique of size a; we call $T(n,a,2)$ the *Turán polytope*.

Moreover, e_i will represent unit vectors with the ith component equal to one, and $\mathbb{0} = (0,0,\ldots,0)$.

1.1.3 Previous Work

An early result in extremal combinatorics states that a graph on n vertices that contains no triangles has at most $\lfloor \frac{n^2}{4} \rfloor$ edges. This was proven by Mantel in 1907. In our notation, this theorem says that $ex(n,3,2) = \lfloor \frac{n^2}{4} \rfloor$. We present a proof of this theorem for completeness sake.

Theorem 1.1.1 (Mantel,1907)**.** *The following equation holds true for all $n \geq 3$: $ex(n,3,2) = \lfloor \frac{n^2}{4} \rfloor$.*

Proof. Let $G = (V,E)$ be a triangle-free graph such that $|V| = n \geq 3$, and let v_1, v_2 be two vertices of V such that $(v_1,v_2) \in E$. Then

$$d(v_1) + d(v_2) \leq n$$

since v_1 and v_2 cannot both be adjacent to the same vertex in $V \backslash \{v_1,v_2\}$ without creating a triangle. Observe now that

$$\sum_{v\in V}((d(v))^2) = \sum_{(v_1,v_2)\in E}(d(v_1)+d(v_2))$$

since $d(v)$ for any $v \in V$ will be in the right-hand side summation $d(v)$ times, i.e. once for every edge it is adjacent to. Note also that $\sum_{v\in V} d(v) = 2|E|$ since each edge (v_1, v_2) is counted twice: once in $d(v_1)$ and once in $d(v_2)$. Moreover, we recall that the Cauchy-Schwarz inequality implies that

$$\sum_{v\in V}((d(v))^2) \geq \frac{(\sum_{v\in V} d(v))^2}{n}.$$

Stringing all of these easy observations together, we obtain

$$\frac{(2|E|)^2}{n} = \frac{(\sum_{v\in V} d(v))^2}{n} \leq \sum_{v\in V}((d(v))^2) = \sum_{(v_1,v_2)\in E}(d(v_1)+d(v_2)) \leq |E|n,$$

which can be simplified to

$$|E| \leq \frac{n^2}{4}.$$

Since the value of $|E|$ is integral, we can make this statement stronger: $|E| \leq \lfloor \frac{n^2}{4} \rfloor$. Finally, we note that this bound is tight since the graph $K_{\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil}$ is triangle free and has $\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil = \lfloor \frac{n^2}{4} \rfloor$ edges. $\square$

Pál Turán then generalized theorem 1.1.1 in [99] by considering the problem of finding out the value of $ex(n,a,2)$ for all $a \geq 3$. This is considered to be a ground-breaking theorem for the field of extremal graph theory. Turán thought of this problem in 1940 in an extreme situation: while building railways in a labor camp in Transylvania. This problem was of utter importance for him during those dark days:

> *...I immediately felt that here was the problem appropriate to the circumstances. I cannot properly describe my feelings during the next few days. The pleasure of dealing with a quite unusual type of*

problem, the beauty of it, the gradual nearing of the solution, and finally the complete solution made these days really ecstatic. The feeling of some intellectual freedom and being, to a certain extent, spiritually free of oppression only added to this ecstasy.

Many proofs have been found for Turán's theorem since Turán's original proof (see for example [47], [53], [70]), and we've also come up with a new polyhedral proof which we will present over the next few pages, but first, we present the most famous proof for this result.

Theorem 1.1.2. *The following inequality holds for all $n \geq a$:*

$$ex(n,a,2) \leq \left\lfloor (1-\frac{1}{a-1})\frac{n^2}{2} \right\rfloor.$$

Proof. Let $G=(V,E)$ be a graph such that E is maximally a-Turán. We first show that if $(v_1,v_2) \notin E$ and $(v_2,v_3) \notin E$ for some $v_1,v_2,v_3 \in V$, then $(v_1,v_3) \notin E$ either. Suppose not, suppose that $(v_1,v_2) \notin E$ and $(v_2,v_3) \notin E$, but $(v_1,v_3) \in E$.

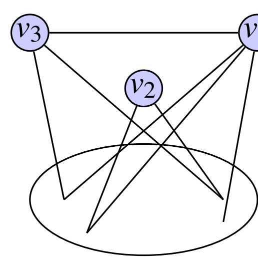

Figure 1.1: Graph $G=(V,E)$ with (v_1,v_2),$(v_2,v_3) \notin E$ and $(v_1,v_3) \in E$.

Then, if $d(v_2) < d(v_1)$, then construct a new graph G' by replacing v_2 by a copy of v_1.

Then G' is still K_a-free because any clique in G' contains at most one of v_1 and its copy since they don't form an edge. So any clique present is a clique or a copy of a clique that was already in G, and so none can be of size greater than $a-1$. But now notice that $|E(G')| = |E(G)| + d(v_1) - d(v_2) > |E(G)|$, thus contradicting that $E(G)$ was maximally a-Turán. We can reach the same conclusion if $d(v_2) < d(v_3)$.

Therefore, $d(v_2) \geq d(v_1)$ and $d(v_2) \geq d(v_3)$. We then construct a graph G'' by replacing v_1 and v_3 by two copies of v_2.

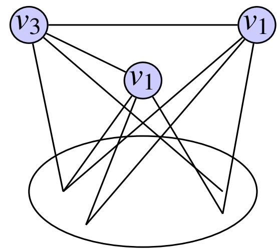

Figure 1.2: Replacing v_2 by v_1 in G.

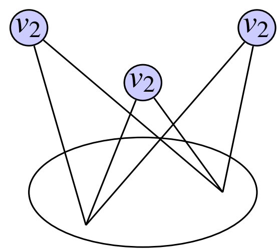

Figure 1.3: Replacing v_1 and v_3 by v_2 in G.

Then G'' is still K_a-free and $|E(G'')| = |E(G)| - (d(v_1) + d(v_3) - 1) + 2d(v_2) > |E(G)|$ where the "-1" comes from the fact that edge (v_1, v_3) gets counted twice. This is yet again a contradiction on the fact that $E(G)$ is maximally a-Turán.

Thus, we have that if $(v_1, v_2) \notin E$ and $(v_2, v_3) \notin E$ for some $v_1, v_2, v_3 \in V$, then $(v_1, v_3) \notin E$ as desired. Not being connected in G is thus an equivalence relation. For any two vertices v_1, v_2 that are in distinct maximal stable sets of the graph, (v_1, v_2) must be an edge, otherwise v_1 would not be connected either to any of the vertices in the stable set that contains v_2 and vice-versa, and actually the two stable sets considered would form one bigger stable set together, contradicting their maximality. This implies that G is a complete multipartite graph where the equivalence classes are the parts, i.e. the stable sets. Note that G can have at most $a-1$ parts, otherwise there would be a K_a in G. Moreover, the number of edges in G is maximized when all of the parts have cardinalities that differ by at most one. Indeed, if there were two parts V_1, V_2 such that $|V_1| > |V_2| + 1$, then moving a vertex from V_1 to V_2 would make the number of edges increase by $|V_1| - 1 - |V_2| > 0$ (note that this is also true for a "part" that contains no vertices). Thus, since $|E(G)|$ is maximal, G has exactly $a-1$ parts with sizes as equal as possible.

Since a maximally a-Turán graph is a multipartite graph with $a-1$ parts, namely with $a-1-(n \bmod (a-1))$ parts of size $\lfloor \frac{n}{a-1} \rfloor$ and $n \bmod (a-1)$ parts of size $\lceil \frac{n}{a-1} \rceil$, then $ex(n,a,2)$ is equal to

$$\binom{a-1-(n \text{ mod } (a-1))}{2} \cdot \left(\left\lfloor \frac{n}{a-1} \right\rfloor\right)^2 + \binom{n \text{ mod } (a-1)}{2} \cdot \left(\left\lceil \frac{n}{a-1} \right\rceil\right)^2 \\ + (a-1-(n \text{ mod } (a-1))) \cdot (n \text{ mod } (a-1)) \cdot \left\lfloor \frac{n}{a-1} \right\rfloor \cdot \left\lceil \frac{n}{a-1} \right\rceil$$

or, by removing the non-edges,

$$ex(n,a,2) = \binom{n}{2} - (a-1-(n \mod a-1)) \cdot \binom{\lfloor \frac{n}{a-1} \rfloor}{2} - (n \mod a-1) \cdot \binom{\lceil \frac{n}{a-1} \rceil}{2}.$$

Since neither of these formulas are very pretty, the nicer bound of $(1-\frac{1}{a-1})\frac{n^2}{2}$ is usually given. This bound represents the number of edges present in a complete $(a-1)$-partite graphs with parts of size $\frac{n}{a-1}$ (even if that is not an integer). $\square$

Listing all of the work done related to Turán-type problems since Turán's groundbreaking theorem would be nearly impossible since these problems went on to become a subfield of extremal graph theory in its own right. We refer the reader to two excellent surveys (who cite hundreds of articles!): [56] and [97]. We just list here a few more important results.

For example, what happens if we forbid some other subgraph G than K_i? Very quickly after Turán proved his eponymous theorem, Erdös and Stone proved the following in [34].

Theorem 1.1.3 (Erdös & Stone, 1946)**.** *Let $r = \chi_G$, i.e. let r be the chormatic number of G, then the maximum number of edges in a graph on n vertices that does not contain G is at most*

$$ex(n,r,2) + o(n^2).$$

Later on, Erdös and Simonovits found in [32] an asymptotic result as the number of vertices goes to infinity when several subgraphs are forbidden.

Theorem 1.1.4 (Erdös & Simonovits, 1966). *Let $ex(n, G_1, \ldots, G_k)$ be the maximum number of edges in a graph on n vertices that does not contain $G_1, \ldots, G_k$ as subgraphs. Then*

$$\lim_{n\to\infty} \frac{ex(n, G_1, \ldots, G_k)}{n^2} = \frac{1}{2}\left(1 - \frac{1}{r}\right),$$

where $r = \min_i \chi(G_i) - 1$ where $\chi(G_i)$ is the chromatic number of G_i.

There are also exact results known. For example, Füredi found the maximum number of edges in a graph without cycles of length 4, i.e. squares, on n vertices for certain n's (see [44] and [46]). In general, finding the Turán number for bipartite graphs (i.e. the maximum number of edges in an H-free graph on n vertices where H is bipartite) seems very hard.

Moreover, the problem was generalized to r-uniform hypergraphs where it turned out to be even more complicated. The maximum number of edges when we forbid some complete r-uniform hypergraph is still unknown for all $r \geq 3$ and n.

Still, there are a few success stories. Keevash and Sudakov found the Turán number for the Fano plane in [61] after the work of others ([25], [48]). Keevash also found the Turán number for a 3-uniform hypergraph that he calls $F_{3,3}$ in [57]. Other exact and asymptotical Turán numbers were found for hypergraphs in [39], [40], [41], [66], [77] and [95] among others.

Moreover, recently, Razborov successfully used flag algebras to prove the following theorem that is slightly weaker than the famous $(4,3)$-conjecture that we will see later on.

Theorem 1.1.5 (Razborov, 2010). *Let G be a 3-uniform hypergraph on four vertices with exactly one 3-edge present. Then the maximal number of 3-edges in a graph on n vertices without K_4^3 and without G is at most $\frac{4}{9}(1-o(1))\binom{n}{3}$, the number conjectured by Turán when only K_4^3 is forbidden.*

Moreover, still by using flag algebras in [80], Razborov found the best asymptotic bound for 3-hypergraphs that do not contain a clique of size 4:

$$\lim_{n\to\infty} \frac{ex(n,4,3)}{\binom{n}{3}} \leq 0.561666,$$

thus improving previous bounds found in [24], [66], [98]. In general, the best bound for the number of r-edges in a r-uniform hypergraph that is a-clique-free was given by de Caen in [23],

$$ex(n,a,r) \leq \binom{n}{r}\left(1 - \frac{n-a+1}{n-r+1} \cdot \frac{1}{\binom{a-1}{r-1}}\right),$$

which also yields the best asymptotic bound known.

1.2 Modeling the Turán Graph Problem

1.2.1 The Model

As we said before, we let $T(n,a,2)$ be the convex hull of the characteristic vectors of all edge sets $F \subseteq E(K_n)$ that contain no clique of size a. Our model will be based on the following proposition.

Proposition 1.2.1. *We have that*

$$T(n,a,2) = \operatorname{conv}\left(\left\{x \in \{0,1\}^{\binom{n}{2}} \mid \sum_{e \in E(Q^a)} x_e \leq \binom{a}{2} - 1 \quad \forall Q^a \in \mathcal{Q}^a_{K_n}\right\}\right).$$

Proof. Any vertex of $T(n,a,2)$ corresponds to some a-Turán edge set in K_n which means that for any induced subgraph on a vertices, there are at most $\binom{a}{2} - 1$ edges there, since otherwise there would be a clique of size a. Thus each vertex of $T(n,a,2)$ satisfies the conditions of points in the right-hand side set, that is,

$$T(n,a,2) \subseteq \operatorname{conv}\left(\left\{x \in \{0,1\}^{\binom{n}{2}} \mid \sum_{e \in E(Q^a)} x_e \leq \binom{a}{2} - 1 \quad \forall Q^a \in \mathcal{Q}^a_{K_n}\right\}\right).$$

Similarly, any vertex of the right-hand side is the characteristic vector of an a-Turán edge set, and must thus be in $T(n,a,2)$. Thus equality holds. □

We make the model a bit stronger by noticing the following easy fact.

Proposition 1.2.2. *For any a-Turán edge set F on n vertices, there is at most $ex(i,a,2)$ edges induced by any i vertices, for $a \leq i \leq n$.*

Proof. Any subset of edges induced by some i vertices of an a-Turán edge set F is also a-Turán since, if it contained a clique of size a, then so would F, a contradiction. Thus, the maximum number of edges induced by those i vertices is $ex(i,a,2)$. □

This easy observation is used to come up with the following linear model.

Definition 1.2.3. We let

$$Q(n,a,2) = \left\{ x \in \mathbb{R}^{\binom{n}{2}} \middle| \begin{array}{ll} \sum_{e \in E(Q^i)} x_e \leq ex(i,a,2) & \forall Q^i \in \mathbb{Q}^i_{K_n}, \forall a \leq i \leq n-1 \\ 0 \leq x_e \leq 1 & \forall e \in E(K_n) \end{array} \right\}$$

be the *clique-relaxation of the Turán polytope* $T(n,a,2)$ for $n > a$. Inequalities of type $\sum_{e \in E(Q^i)} x_e \leq ex(i,a,2)$ are called *clique inequalities*, $x_e \geq 0$, *non-negativity inequalities*, and $x_e \leq 1$, *edge inequalities*.

This linear program is the linear relaxation of $T(n,a,2)$ with clique inequalities added for any clique of size greater or equal to a and smaller than n. Note that we only consider the clique inequalities for $a \leq i \leq n-1$ since our strategy will be to calculate the value of $ex(n,a,2)$ by using this program inductively, and thus the program cannot include the value $ex(n,a,2)$.

Proposition 1.2.4. *We have that*

$$T(n,a,2) \subseteq Q(n,a,2)$$

and

$$T(n,a,2) = conv\left(Q(n,a,2) \cap \mathbb{Z}^{\binom{n}{2}}\right).$$

Proof. Any vertex of $T(n,a,2)$ corresponds to an a-Turán edge set, and by the previous proposition, we know such an edge set respects $\sum_{e \in E(Q^i)} x_e \leq ex(i,a,2)$ for all $Q^i \in Q^i_{K_n}$ and for all $a \leq i \leq n$. Moreover, any integral point of $Q(n,a,2)$ correspond to an a-Turán edge set. □

Remark. Since $T(n,a,2) = \text{conv}\left(Q(n,a,2) \cap \mathbb{Z}^{\binom{n}{2}}\right)$, any Gomory-Chvátal cut developed from inequalities in $Q(n,a,2)$ is valid for $Q(n,a,2) \cap \mathbb{Z}^{\binom{n}{2}}$ and thus is also valid for $T(n,a,2)$.

1.2.2 Polyhedral Proof of the Turán Theorem

We present in this section a new solution to the Turán problem for graphs using the previous model to calculate $ex(n,a,2)$.

Theorem 1.2.5. *There are exactly*

$$ex(n,a,2) = \left\lfloor \frac{n}{n-2} \left\lfloor \frac{n-1}{n-3} \left\lfloor \frac{n-2}{n-4} \left\lfloor \cdots \left\lfloor \frac{a+2}{a} \left\lfloor \frac{a+1}{a-1} \cdot \left(\binom{a}{2} - 1 \right) \right\rfloor \right\rfloor \cdots \right\rfloor \right\rfloor \right\rfloor \right\rfloor$$

edges in a maximally a-Turán edge set on n vertices.

Proof. We proceed by induction on n. For $n = a$, the theorem states that $ex(n,a,2) = ex(a,a,2) = \binom{a}{2} - 1$, which is trivially true. Now assume the theorem is true up to $n-1$ and let's prove that it still holds for n.

We generate the following Gomory-Chvátal cut valid for $T(n,a,2)$. Take the n inequalities $\sum_{e \in Q^{n-1}} x_e \leq ex(n-1,a,2)$ for all $Q^{n-1} \in \mathcal{Q}_{K_n}^{n-1}$, add them up with weight $\frac{1}{n-2}$ and round down. This yields the cut

$$\sum_{e \in E(K_n)} x_e \leq \left\lfloor \frac{n}{n-2} ex(n-1,a,2) \right\rfloor$$

since each edge of $E(K_n)$ is in $n-2$ of the n cliques of size $n-1$. Since $ex(n,a,2) = \max\{\sum_{e \in E(K_n)} x_e | x \in Q(n,a,2) \cap \mathbb{Z}^{\binom{n}{2}}\}$, we now have that

$$ex(n,a,2) \leq \left\lfloor \frac{n}{n-2} ex(n-1,a,2) \right\rfloor. \tag{1.1}$$

Applying this observation iteratively starting with $n = a+1$, we obtain that

$$ex(n,a,2) \leq \left\lfloor \frac{n}{n-2} \left\lfloor \frac{n-1}{n-3} \left\lfloor \frac{n-2}{n-4} \left\lfloor \cdots \left\lfloor \frac{a+2}{a} \left\lfloor \frac{a+1}{a-1} \cdot \left(\binom{a}{2} - 1 \right) \right\rfloor \right\rfloor \cdots \right\rfloor \right\rfloor \right\rfloor \right\rfloor, \tag{1.2}$$

for all $n > a \geq 3$. We claim that equality always holds in both (1.1) and (1.2). To prove this, we introduce another type of inequality to obtain more information about $ex(n,a,2)$.

The following type of inequality is valid for $T(n,a,2)$:

$$\sum_{e \in \delta(v)} 2 \cdot x_e + \sum_{e \in E(K_n) \setminus \delta(v)} x_e \leq ex(n+1,a,2) \tag{1.3}$$

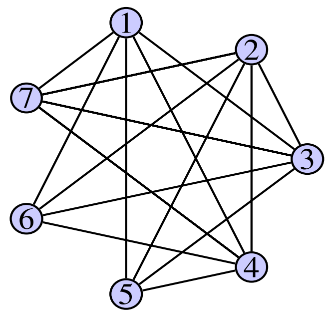

Figure 1.4: Graph without any 4-clique.

for any $v \in [n]$. Indeed, take any a-clique-free graph G on n vertices.

Fix a vertex v and consider the graph $G' = ([n+1], E')$ where $E' = E \cup \{(i, n+1) | (i, v) \in E\}$.

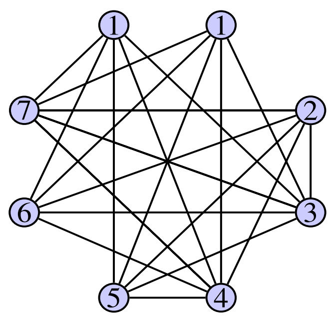

Figure 1.5: Previous graph with a copy of the first vertex.

It is easy to see that G' must also be a-clique-free since no clique contains both v and $n+1$. The maximum number of edges of an a-clique-free graph on $n+1$ vertices is thus at least the number of edges in G'.

From adding up the n inequalities of type (1.3) corresponding to each $v \in K_n$ with weight $\frac{1}{n+2}$, we get the following Gomory-Chvátal cut valid for $T(n, a, 2)$:

$$\sum_{e \in E(K_n)} x_e \leq \left\lfloor \frac{n}{n+2} \cdot ex(n+1, a, 2) \right\rfloor. \tag{1.4}$$

From this, we conclude that

$$ex(n, a, 2) \leq \left\lfloor \frac{n}{n+2} \cdot ex(n+1, a, 2) \right\rfloor. \tag{1.5}$$

Putting inequalities (1.1) and (1.5) together, we get that

$$ex(n+1,a,2) \leq \left\lfloor \frac{n+1}{n-1} \cdot ex(n,a,2) \right\rfloor \leq \left\lfloor \frac{n+1}{n-1} \left\lfloor \frac{n}{n+2} \cdot ex(n+1,a,2) \right\rfloor \right\rfloor. \quad (1.6)$$

We now show that

$$\left\lfloor \frac{n+1}{n-1} \left\lfloor \frac{n}{n+2} \cdot ex(n+1,a,2) \right\rfloor \right\rfloor \leq ex(n+1,a,2), \quad (1.7)$$

thus turning (1.6) into an equation and proving our claim.

Suppose not. Then

$$ex(n+1,a,2)+1 \leq \left\lfloor \frac{n+1}{n-1} \left\lfloor \frac{n}{n+2} \cdot ex(n+1,a,2) \right\rfloor \right\rfloor, \quad (1.8)$$

which implies that:

$$\begin{aligned} ex(n+1,a,2)+1 &\leq \left\lfloor \frac{n+1}{n-1} \cdot \frac{n}{n+2} \cdot ex(n+1,a,2) \right\rfloor \\ &= ex(n+1,a,2) + \left\lfloor \frac{2}{(n-1)\cdot(n+2)} \cdot ex(n+1,a,2) \right\rfloor. \quad (1.9) \end{aligned}$$

Certainly, this means that

$$\begin{aligned} 1 &\leq \frac{2}{(n-1)\cdot(n+2)} \cdot ex(n+1,a,2) \\ &\leq \frac{2}{(n-1)\cdot(n+2)} \cdot \left\lfloor \frac{n+1}{n-1} \left\lfloor \frac{n}{n-2} \left\lfloor \cdots \left\lfloor \frac{a+1}{a-1} \cdot \left(\binom{a}{2} - 1 \right) \right\rfloor \cdots \right\rfloor \right\rfloor \right\rfloor \\ &\leq \frac{2}{(n-1)\cdot(n+2)} \cdot \frac{n+1}{n-1} \cdot \frac{n}{n-2} \cdots \frac{a+1}{a-1} \cdot \left(\binom{a}{2} - 1 \right) \\ &= \frac{(n+1)\cdot n \cdot (a+1)\cdot(a-2)}{(n-1)\cdot(n+2)\cdot a\cdot(a-1)}. \quad (1.10) \end{aligned}$$

So we have that

$$\frac{(n-1)\cdot(n+2)}{(n+1)\cdot n} \leq \frac{(a+1)\cdot(a-2)}{a\cdot(a-1)}. \quad (1.11)$$

By simplifying this inequality, we obtain that

$$1 - \frac{2}{n\cdot(n+1)} \leq 1 - \frac{2}{a\cdot(a-1)}, \quad (1.12)$$

which implies that

$$n \cdot (n+1) \leq a \cdot (a-1), \tag{1.13}$$

which is impossible since $n > a \geq 3$. We have thus reached a contradiction, and so

$$\left\lfloor \frac{n+1}{n-1} \left\lfloor \frac{n}{n+2} \cdot ex(n+1,a,2) \right\rfloor \right\rfloor \leq ex(n+1,a,2),$$

which proves our claim. □

Corollary 1.2.6. *The integrality gap between $Q(n,a,2)$ and $T(n,a,2)$ is less than one in the $\mathbb{1}$-direction.*

Proof. In the proof of the last theorem, we saw that

$$\left\lfloor \max_{x \in Q(n,a,2)} \mathbb{1}x \right\rfloor = ex(n,a,2) = \max_{x \in T(n,a,2)} \mathbb{1}x.$$

The result follows. □

We note that

$$\left\lfloor \frac{n}{n-2} \left\lfloor \frac{n-1}{n-3} \left\lfloor \frac{n-2}{n-4} \left\lfloor \cdots \left\lfloor \frac{a+2}{a} \left\lfloor \frac{a+1}{a-1} \cdot \left(\binom{a}{2} - 1 \right) \right\rfloor \right\rfloor \cdots \right\rfloor \right\rfloor \right\rfloor \right\rfloor$$

is a new exact formula for $ex(n,a,2)$. Turán's bound calculates the number of edges in an complete $a-1$-partite graph where all parts have equal size. It is of course only possible to do so if $n = 0(\mod a-1)$, and so the Turán bound is equal to $ex(n,a,2)$ only if $n = 0(\mod a-1)$. For all other cases, $ex(n,a,2) < \left(1 - \frac{1}{a-1}\right)\frac{n^2}{2}$.

The values in black in table 1.1 are those for which the bound given by Turán and our formula for $ex(n,a,2)$ are less than one unit apart, i.e. $0 \leq \left(1 - \frac{1}{a-1}\right)\frac{n^2}{2} - ex(n,a,2) < 1$. This difference is at least one and less than two for those in red, i.e. $1 \leq \left(1 - \frac{1}{a-1}\right)\frac{n^2}{2} - ex(n,a,2) < 2$. As a increases, the possible gap between the Turán bound and $ex(n,a,2)$ also increases. Indeed, when $a = 3$, the Turán bound is exact for every other value, so for every value for which it is not exact, it cannot be too far from being exact. However, for a very big a, there is more space for the gap to increase and decrease again between two exact values.

Table 1.1: Values of $ex(n,a,2)$

$a\backslash n$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
3	2	4	6	9	12	16	20	25	30	36	42	49	56	64	72	81	90	100	110	121	132	144	156
4		5	8	12	16	21	27	33	40	48	56	65	75	85	96	108	120	133	147	161	176	192	208
5			9	13	18	24	30	37	45	54	63	73	84	96	108	121	135	150	165	181	198	216	234
6				14	19	25	32	40	48	57	67	78	90	102	115	129	144	160	176	193	211	230	250
7					20	26	33	41	50	60	70	81	93	106	120	135	150	166	183	201	220	240	260
8						27	34	42	51	61	72	84	96	109	123	138	154	171	189	207	226	246	267
9							35	43	52	62	73	85	98	112	126	141	157	174	192	211	231	252	273
10								44	53	63	74	86	99	113	128	144	160	177	195	214	234	255	277
11									54	64	75	87	100	114	129	145	162	180	198	217	237	258	280
12										65	76	88	101	115	130	146	163	181	200	220	240	261	283
13											77	89	102	116	131	147	164	182	201	221	242	264	286
14												90	103	117	132	148	165	183	202	222	243	265	288
15													104	118	133	149	166	184	203	223	244	266	289

Table 1.2: Smallest a for which there exists n such that $\left(1-\frac{1}{a-1}\right)\frac{n^2}{2} - ex(n,a,2) \geq d$

d	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
a	9	17	25	33	41	49	57	65	73	81	89	97	105	113	121	129	137	145	153	161	169	177	185

1.2.3 Different Objective Functions

Since we're considering the Turán problem as an optimization problem, the next obvious step is to see what can be said when we change the objective function from $\mathbb{1}$ to something else. We focus on objective functions with 0 or 1 coefficients.

We first note that the problem $\max c^T x$ such that $x \in T(n,a,2)$ and $c \in \{0,1\}^{\binom{n}{2}}$ is equivalent to maximizing the number of edges that can be chosen in G without forming any a-clique, where $G = (V,E)$ with $V = [n]$ and $e \in E$ if and only if $c_e = 1$. We call this maximum number the *Turán number for a subgraph G of K_n*, and we notate it by $ex(G,a,2)$. A natural concept emerges: that of the Turán density of a graph.

Definition 1.2.7. The *a-Turán density* of any graph $G = (V,E)$ is

$$\frac{ex(G,a,2)}{|E|}.$$

An instinctive idea is that the a-Turán density of a subgraph of K_n must be greater or equal than that of K_n, and we now prove that it is indeed so.

Lemma 1.2.8. *For any proper subgraph G of K_n, the a-Turán density of G is greater or equal than the a-Turán density of K_n, that is,*

$$\frac{ex(G,a,2)}{|E|} \geq \frac{ex(n,a,2)}{\binom{n}{2}}.$$

Proof. Suppose not: then

$$\frac{ex(G,a,2)}{|E|} < \frac{ex(n,a,2)}{\binom{n}{2}}$$

for some $G \subset K_n$. Then summing over all isomorphisms G' of G in K_n

$$\sum_{G':G'\sim G} \left(\sum_{e \in E(G')} x_e \leq ex(G,a,2) \right),$$

we get

$$\sum_{e \in E(K_n)} \frac{|E|}{\binom{n}{2}} \cdot |\{G' : G' \sim G\}| \cdot x_e \leq |\{G' : G' \sim G\}| \cdot ex(G,a,2)$$

since each edge will be in $\frac{|E|}{|E(K_n)|} = \frac{|E|}{\binom{n}{2}}$ of the isomorphisms. Thus we simplify and obtain

$$\sum_{e \in E(K_n)} x_e \leq \frac{ex(G,a,2) \cdot \binom{n}{2}}{|E|} < ex(n,a,2)$$

This is a contradiction since we know that the number of edges in an a-Turán edge set on K_n can be $ex(n,a,2)$, and so the Turán density of any subset of K_n must be greater or equal to that of K_n. □

So we know that $ex(G,a,2) \geq \frac{|E| \cdot ex(n,a,2)}{\binom{n}{2}}$.

Remark. Note that $\frac{ex(G,a,2)}{|E|} = \frac{ex(n,a,2)}{\binom{n}{2}}$ for some graph $G \subset K_n$ if and only if for each isomorphism of G in K_n, it is possible to have $ex(G,a,2)$ edges in a maximally a-Turán edge set of K_n.

For example, there exists an optimal 3-Turán edge set of $2n$ vertices such that, for every $(2n-1)$-clique in K_{2n}, we can put $ex(2n-1,3,2)$ edges in it, and so the Turán density of K_{2n-1} is the same as that of K_{2n}. The same isn't true for K_{2n+1} and its $2n$-cliques since any optimal 3-Turán edge set will be a bipartite graph with n vertices on one side and $n+1$ on the other, and a $2n$-clique missing a vertex in the smaller part will have $(n+1)(n-1) = n^2 - 1$ edges in the solution and not $ex(2n,3,2) = n^2$ edges. The density of K_{2n} is therefore greater than that of K_{2n+1}.

We now discuss how hard it might be to calculate what is $ex(G,a,2)$ in general for some $G = (V,E)$. Instead of considering the problem of maximizing the number of edges that can be chosen in E without forming an a-clique, we look instead at the complementary problem of minimizing the number of edges that need to be removed from E in order for the graph to become a-clique-free. It's easy to see that this problem can be rewritten as the vertex cover problem for some a-uniform hypergraph $G' = (V',E')$. Let's quickly recall the standard definition of a vertex cover for a uniform hypergraph.

Definition 1.2.9. A *vertex cover* of an a-uniform hypergraph $G' = (V',E')$ is a subset $S \subseteq V'$ such that every a-hyperedge $e' \in E'$ intersects S. A *minimum vertex cover* is one that minimizes the size of S for G'.

We transform the graph $G = (V,E)$ into an a-uniform hypergraph $G' = (V',E')$ as follows. Let the vertices V' correspond to the edges E of G, and let E' be the set of a-cliques of G. A vertex v' of V' is then adjacent to an edge e' if their corresponding edge e and clique Q in G are such that $e \in Q$. Thus a vertex cover in G' corresponds to a set of edges S in G such that every a-clique intersects S. Therefore, if we remove these edges from G, every a-clique will be missing at least one edge.

Example 1.2.10. For example, if we have the left-hand side graph $G = (V,E)$ of figure 1.6 for which we want to find $ex(G,3,2)$, then we can transform it into the adjacent hypergraph H.

A possible maximum Turán set in G could be

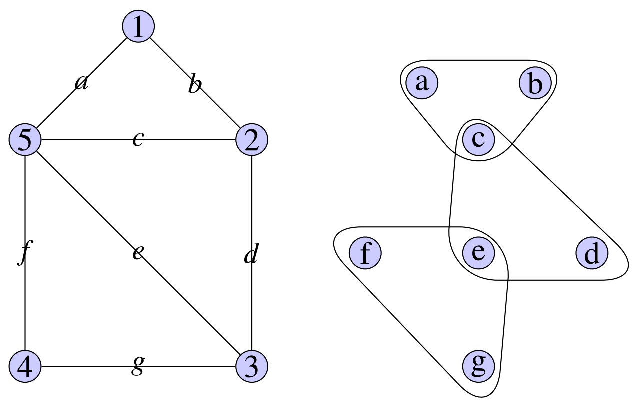

Figure 1.6: A graph and its corresponding 3-hypergraph.

$$E' := \{(1,2),(2,3),(3,4),(4,5),(1,5)\},$$

which corresponds to a minimum vertex cover in H, nameley the set of vertices $V' := \{c,e\}$. As expected, $ex(G,3,2) = |E'| = |E| - |V'|$.

The point of changing the Turán problem for G into a vertex cover problem for G' is that the complexity of the latter has been well-studied. Indeed, it is well-known that finding a minimum size vertex cover in an a-uniform hypergraph is NP-hard, even for $a = 2$ (in which case it is just the good old vertex cover for graphs). Recent results (such as [27]) have also shown that it is APX-hard, i.e. that it is even hard to get a good approximation for the vertex cover of uniform hypergraphs.

Theorem 1.2.11 (Dinur, Guruswami, Khot & Regev, 2005)**.** *For every integer $a \geq 3$ and every $\varepsilon > 0$, it is NP-hard to approximate the minimum size vertex cover of an a-uniform hypergraph within a factor of $a - 1 - \varepsilon$.*

This seems to indicate that there is not much hope of solving the Turán problem for some general G. However, we note that not any a-uniform hypergraph can represent the incidence of edges and a-cliques in G. Thus, succeeding in determining $ex(G,a,2)$ in general would not contradict the previous theorem.

Nevertheless, finding $ex(G,a,2)$ for $G \subset K_n$ is definitely not as easy as for K_n, since, in general, $T(G,a,2)$ (the convex hull of the incidence vectors of all the sets of edges in G that are a-clique-free) is not equal to

$$Q(G,a,2) := \Big\{x \in \mathbb{R}^{\binom{n}{2}} : \sum_{e \in E(Q^i)} x_e \leq ex(i,a,2) \qquad \forall a \leq i < |V|, \forall Q^i \in \mathcal{Q}^i_G$$
$$0 \leq x_e \leq 1 \qquad e \in E\Big\}$$

for $G = (V,E)$. Indeed, it is very easy to come up with an example where equality does not hold.

Example 1.2.12. Take for example G to be an odd wheel.

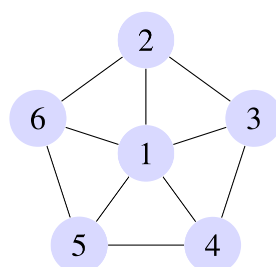

Figure 1.7: Odd 5-wheel.

Then it's easy to verify that the fractional point

$$\begin{array}{cccccccccccc} & 12 & 13 & 14 & 15 & 16 & 23 & 26 & 34 & 45 & 56 & \\ \hline (& \frac{1}{2}, & \frac{1}{2}, & \frac{1}{2}, & \frac{1}{2}, & \frac{1}{2}, & 1, & 1, & 1, & 1, & 1 &) \end{array}$$

is in $Q(G,3,2)$ by checking the validity of each inequality. Moreover, since the five 3-clique inequalities are tight with this point as well as the five edges inequalities $x_e \leq 1$ of the 5-cycle edges, and since those ten inequalities are linearly independent, we have that this point is actually a vertex of $Q(G,3,2)$, and so $Q(G,3,2) \neq T(G,3,2)$.

Thus, $Q(G,a,2)$ is not a sufficient description, and we will talk later on about how to make it better by adding extra facets. However, for now, we'd like to answer the following question: for which $G = (V,E)$ does

$$ex(G,a,2) = \left\lfloor \max \left\{ \sum_{e \in E} x_e | x \in Q(G,a,2) \right\} \right\rfloor ?$$

To study this question, it is useful to consider the dual of the program above. We now introduce the *dual Turán program* of the clique relaxation $Q(n,a,2)$. Let $\mathcal{Q}_G^i$ be the set of cliques of size i present in G.

$$Q_D(G,a,2) = \left\{ y \in \mathbb{R}^{|E|+|\mathcal{Q}_G^a|+\ldots+|\mathcal{Q}_G^{n-1}|} | y_e + \sum_{i=a}^{n-1} \sum_{\substack{Q^i \in \mathcal{Q}_G^i: \\ e \in E(Q^i)}} ex(i,a,2) y_{Q^i} \geq 1 \quad \forall e \in E \right.$$
$$y_e, y_{Q^i} \geq 0 \quad \forall Q^i \in \mathcal{Q}_G^i, e \in E$$
$$\left. 2 \leq i \leq n-1 \right\}$$

By the well-known strong duality theorem, the value given by minimizing

$$\sum_{e \in E} y_e + \sum_{i=a}^{n-1} \sum_{Q^i \in \mathcal{Q}_G^i} ex(i,a,2) y_{Q^i}$$

over $Q_D(G,a,2)$ is equal to the value given by maximizing $\sum_{e \in E} x_e$ over $Q(G,a,2)$. The binary version of the dual minimization problem is

$$\begin{aligned} \nu(G,a,2) := \min y_e + \sum_{i=a}^{n-1} \sum_{Q^i \in \mathcal{Q}_G^i} ex(i,a,2) y_{Q^i} & \\ \text{s.t. } y_e + \sum_{i=a}^{n-1} \sum_{\substack{Q^i \in \mathcal{Q}_G^i: \\ e \in E(Q^i)}} y_{Q^i} \geq 1 & \qquad \forall e \in E \\ y_e, y_{Q^i} \in \{0,1\} & \qquad \forall Q^i \in \mathcal{Q}_G^i, 2 \leq i \leq n-1, \forall e \in E \end{aligned}$$

where $\nu(G,a,2)$ is called the *Turán cover number* of G. Indeed, this problem corresponds to assigning weight $ex(i,a,2)$ to i-cliques and finding the minimum weight clique cover of the edges. In other words, we want to select the set of cliques of minimum weight such that each edge is in at least one clique in the set. From duality theory, we know that

$$
\begin{aligned}
ex(G,a,2) &\leq \max\left\{\sum_{e\in E} x_e | x \in Q(G,a,2)\right\} \\
&= \min\left\{\sum_{e\in E} y_e + \sum_{i=a}^{n-1} \sum_{Q^i \in \mathcal{Q}^i_G} ex(i,a,2) y_{Q^i} | y \in Q_D(G,a,2)\right\} \leq \nu(G,a,2).
\end{aligned}
$$

Definition 1.2.13. We call a graph $G = (V,E)$ for which

$$
\left\lfloor \max\left\{\sum_{e\in E} x_e | x \in Q(G,a,2)\right\}\right\rfloor = ex(G,a,2) < \max\left\{\sum_{e\in E} x_e | x \in Q(G,a,2)\right\}
$$

a *well-behaved graph*. Moreover, we call a graph $G = (V,E)$ for which

$$
ex(G,a,2) = \max\left\{\sum_{e\in E} x_e | x \in Q(G,a,2)\right\}
$$

a-perfect. Finally, we say a graph G is *badly-behaved* if the optimal solution in $Q(G,a,2)$ with additional rank constraints $\sum_{e\in E(G')} x_e \leq ex(G',a,2)$ for every subgraph $G' \subset G$ doesn't have value $ex(G,a,2)$ (without rounding).

We'd first like to find out which graphs are badly-behaved. A first logical assumption is to only consider graphs for which every edge is in an a-clique and which is connected. Indeed, we know that if there is an edge that is not in an a-clique, it will automatically be in all optimal solutions. Thus, we can remove them since the optimal solution we find on the rest of the graph will still be optimal after adding them again. We also only consider 2-node-connected graphs since, even if maximizing the number of edges in each component $G_C = (V_C, E_C)$ is equal to $\lfloor \max\left\{\sum_{e\in E_C} x_e | x_e \in Q(G_C,a,2)\right\}\rfloor$, that might not be true for the whole graph.

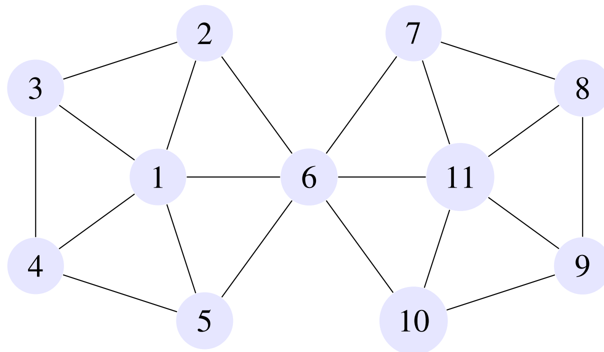

Figure 1.8: Two odd 5-wheels connected by one vertex.

Example 1.2.14. For example, if we have graph G in Figure 1.8,

and we let G_1 be the 5-wheel induced by vertices 1 through 6, and G_2 the one induced by vertices 6 through 11. From the previous example, we know $ex(G_1,3,2) = ex(G_2,3,2) \leq \lfloor \frac{15}{2} \rfloor = 7$, and since $ex(G,3,2) \leq ex(G_1,3,2) + ex(G_2,3,2)$, it is clear that $ex(G,3,2) \leq 7+7 = 14$. However, the maximum attained in $Q(G,3,2)$ will be $\lfloor \frac{15}{2} + \frac{15}{2} \rfloor = 15$ at vertex

$$\begin{array}{cccccccccccccccccccccc}
 & 12 & 13 & 14 & 15 & 16 & 23 & 26 & 34 & 45 & 56 & 67 & 78 & 89 & 910 & 610 & 611 & 711 & 811 & 911 & 1011 & \\
\hline
(& \frac{1}{2}, & \frac{1}{2}, & \frac{1}{2}, & \frac{1}{2}, & \frac{1}{2}, & 1, & 1, & 1, & 1, & 1, & 1, & 1, & 1, & 1, & 1, & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} &)
\end{array}$$

and so G is not well-behaved, but it is also not badly-behaved since by adding the rank inequalities associated to G_1 and G_2 to $Q(n,a,2)$, we find an optimal solution of value $ex(G,a,2)$..

Definition 1.2.15. We call an edge that is not in any a-clique an *a-outsider* edge.

Our question has thus become: which 2-connected a-outsider-free graph $G = (V,E)$ is not well-behaved? We know all complete graphs are such well-behaved, a-outsider-free, connected graphs. Is that a sufficient characterization? Unfortunately no.

Example 1.2.16. Another not well-behaved but also not badly-behaved example that is 2-connected and that has no 3-outsider edges is the graph G found in Figure 1.9.

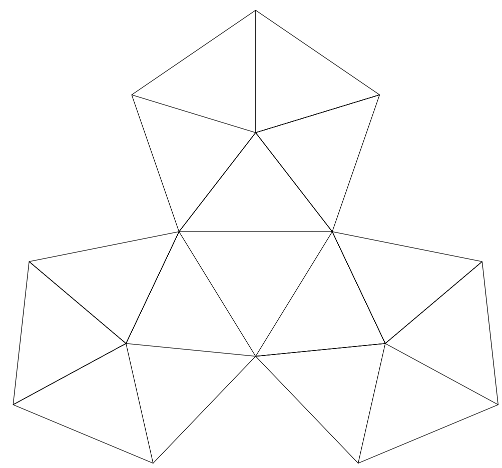

Figure 1.9: 2-connected 3-outsider-free not well-behaved graph.

The maximum on $Q(G,3,2)$ is 22, but $ex(G,3,2) = 21$ since $ex(G,3,2) \leq 3 \cdot ex(G',3,2)$ where G' is an odd 5-wheel.

We notice that the $\nu(G,3,2)$ here is equal to 23. So perhaps the question is more for which graph the cover is at most one bigger than $T(G,3,2)$.

Are there others? Can we classify them? More importantly, can we classify the badly-behaved graphs, which form a much smaller and more important class?

If we did, and if we could identify the badly-behaved graphs in polynomial-time, then we would be able to determine $ex(G,a,2)$ for any graph G in polynomial-time too. We'll show that, but first we need the following lemma.

Lemma 1.2.17. *Let $G = (V,E)$ such that $|V| \geq a+2$. If G' is a-clique-free for all $G' \subset G$, then G is a-clique-free too.*

Proof. Let $\overline{G}' = G - e$. Edge e is in all other $G' \subset G$, and does not form any a-clique then. We know all possible a-cliques containing e are present in the different $G' \subset G$ if $|V| \geq a+2$ since removing an edge leaves a subgraph on at least a vertices intact, so by looking at all $G' \subset G$, we will observe all possible a-cliques. Thus adding e to $\overline{G}'$ will not form any a-clique and G is thus a-clique-free. □

Theorem 1.2.18. *We have that*

$$ex(G,a,2) = \max\{\sum_{e\in E(G)} x_e | \sum_{e\in E(G')} x_e \le ex(G',a,2) \quad \forall G' \subset G,\ G' \text{ badly-behaved},$$
$$0 \le x_e \le 1 \qquad \forall e \in E(G)\}.$$

Proof. Let G' be a subgraph of G with $|E|-1$ edges. For any G', either it is badly-behaved, in which case there is an inequality associated to it, or it is not, meaning that there are inequalities present in the program which can be combined linearly to give a tight inequality for this subgraph. Thus, we know what is $ex(G',a,2)$ for each G'. We also know that

$$\max_{G'_1,G'_2}\{ex(G'_1,a,2) - ex(G'_2,a,2)\} \le 1$$

since $E(G'_1) = E(G'_2) - e_2 + e_1$ for some edges e_1, e_2. Adding up all these subgraphs with weight $\frac{1}{|E|-1}$, we get

$$ex(G,a,2) \le \left\lfloor \frac{\sum_{G'} ex(G',a,2)}{|E(G)|-1} \right\rfloor$$
$$= \begin{cases} |E(G)| & \text{if } ex(G',a,2) = |E(G)|-1\ \forall G' \\ \max_{G'} ex(G',a,2) & \text{otherwise.} \end{cases}$$

In the first case, all $G' \subset G$ are a-free and by the previous lemma, this means that G is a-free too, and so $ex(G,a,2) = E(G)$. In the second case, it is trivially tight since the set of edges tight with $\max_{G'} ex(G',a,2)$ is also valid for $ex(G,a,2)$.

Therefore, having all badly-behaved subgraphs of G is sufficient to find $ex(G,a,2)$. □

For example, maximizing $\sum_{e\in E} x_e$ over $Q(G,a,2)$ with all badly-behaed rank inequalities in the previous example would yield $ex(G,3,2) \le 21$ as desired.

What is of course problematic is that we haven't characterized fully the badly-behaved graphs. Of course, one could calculate all $ex(G',a,2)$, $G' \subset G$, recursively by increasing the number of edges in the subgraphs considered, thus making sure that all the badly-behaved subgraphs are listed. However, this would be time-consuming. One could also be a bit clever and only calculate $ex(G', a, 2)$, $G' \subset G$, when G' is connected and every edge is in an a-clique. This remains problematic, and so more work needs to be done to identify the badly-behaved subgraphs or at least to restrain the size of the set of possible subgraphs.

1.2.4 Some General Theorems for the Facet Analysis

As we saw, $T(G,a,2)$ and $Q(G,a,2)$ are not always equal, but we do know that $T(G,a,2) \subseteq Q(G,a,2)$. Adding more facets and valid inequalities of $T(G,a,2)$ to $Q(G,a,2)$ would thus help to close the gap between them. Certainly, we already know that one should add the badly-behaved rank inequalities that we discussed previously. We first state a few theorems that will help us in our facet analysis.

Proposition 1.2.19. *Any facet we will add to the existing ones for some $Q(G,a,2)$ will be of the form $\alpha^T x \leq b$ where $\alpha_e \geq 0$ for all $e \in E$ and $b > 0$.*

Proof. We simply need to show that $T(G,a,2)$ is down-monotone in $\mathbb{R}^n_+$. This is indeed the case since any y such that $0 \leq y \leq x$ for all $x \in T(G,a,2)$ will also be in $T(G,a,2)$ since y will respect all the constraints of $T(G,a,2)$. □

We now show that some facets of $T(G,a,2)$ can be lifted to $T(G',a,2)$ for any $G \subset G'$ because of the following theorem of Padberg.

Theorem 1.2.20 (Padberg, 1973)**.** *Let $S \subseteq \{0,1\}^n$ be monotone (i.e., $y \leq x \in S$ implies $y \in S$), $P_S := conv(S)$ be full-dimensional, $I \subseteq \{1,2,\ldots,n\}$, $P_S(I) := P_S \cap \{x \in \mathbb{R}^n | x_i = 0 \, \forall i \in I\}$ and $x^I \in \mathbb{R}^n$ denote a vector with $x^I_i = 0$ for all $i \in I$. Suppose $\sum_{j \notin I} \alpha_j x_j \leq \alpha_o$ with $\alpha_0 > 0$ defines a facet of $P_s(I)$ and $i \in I$. Define*

$$\alpha_i := \alpha_0 - \max\{\sum_{j \notin I} \alpha_j x^I_j | e_i + x^I \in S\}.$$

Here e_i is a unit vector in $\mathbb{R}^n$ with the ith component equal to one. Then

$$\alpha_i x_i + \sum_{j \notin I} \alpha_j x_j \leq a_0$$

defines a facet of $P_S(I \backslash \{i\})$.

Corollary 1.2.21. *Let G and H be two graphs such that $H \subseteq G$ and such that $ex(H_e,a,2) = ex(H,a,2)+1$ for every $e \in E(G)\backslash E(H)$ where $H_e = (V(H), E(H) \cup e)$. If $\sum_{e\in E(H)} x_e \leq ex(H,a,2)$ is a facet of $T(H,a,2)$, then it is also a facet of $T(G,a,2)$.*

Proof. We first show that $T(G,a,2)$ is a full-dimensional monotone polytope. We know that $\mathbb{0}$ as well as every unit vector is in $T(G,a,2)$ since the empty set and any single edge are a-clique-free; therefore, $T(G,a,2)$ is full-dimensional. Moreover, we know that if the characteristic vector of an edge set S is in $T(G,a,2)$, then it is a-clique-free and if we take a subset of S, then it is also a-clique-free, and so its characteristic vector will also be in $T(G,a,2)$. Thus, $T(G,a,2)$ is also monotone, and so the setup is similar to the one of the previous theorem.

We let $I := E(G)\backslash E(H)$, and we now show that $T(H,a,2) = T(G,a,2) \cap \{x \in \mathbb{R}^{|E(G)|} | x_e = 0 \quad \forall e \in I\}$. Indeed, $T(H,a,2)$ contains every point P that is in $T(G,a,2)$ for which $x_e = 0$ for all $e \in I$ since such a point must be a convex combination of vertices of $T(G,a,2)$ for which $x_e = 0$ for all $e \in I$ (if the combination contained a positive coefficient for some vertex of $T(G,a,2)$ for which there exists $x_{e'} = 1$ for some $e' \in I$, then P would have $x_{e'} > 0$, a contradiction). Since vertices of $T(G,a,2)$ for which $x_e = 0$ are also vertices of $T(H,a,2)$, P is also in $T(H,a,2)$. We can show similarly that every point in $T(H,a,2)$ is in $T(G,a,2) \cap \{x \in \mathbb{R}^{|E(G)|} | x_e = 0 \ \forall e \in I\}$.

Suppose $\sum_{e\in E(H)} x_e \leq ex(H,a,2)$ is a facet of $T(H,a,2)$. Take any $e' \in I$, and let $c_{e'} := ex(H,a,2) - \max\{\sum_{e\notin I} x_e | x \in T(H,a,2) \text{ and } e' \cup x \text{ is } a-\text{clique-free}\}$. Since $ex(H_{e'},a,2) = ex(H,a,2)+1$, we know there exists an a-clique-free set of edges S in $H_{e'}$ of size $ex(H,a,2)+1$ which must contain e' (else we would have $ex(H,a,2) = ex(H,a,2)+1!$), and $S\backslash\{e'\}$ is still a-clique-free and completely in H, so $c_{e'} = 0$ since $|S\backslash\{e'\}| = ex(H,a,2)$. Thus by the previous theorem, $\sum_{e\in E(H)} x_e \leq ex(H,a,2)$ is also a facet of $T(H_{e'},a,2)$.

Adding another edge $e'' \in E(G)\backslash(E(H) \cup \{e'\})$ will yield the same argument that $0 = c_0 - \max\{\sum_{e\notin I\backslash\{e'\}} x_e | x \in T(H_{e'},a,2) \text{ and } e'' \cup x \text{ is } a-\text{clique-free}\}$ since $ex(H_{e''},a,2) = ex(H,a,2)+1$, and so there is an edge set of size $ex(H,a,2)$ in $T(H,a,2)$ (and thus also in $T(H_{e'},a,2)$) that fulfills the requirements.

We can therefore add all of the edges in $E(G)\backslash E(H)$, and get that $\sum_{e\in E(H)} x_e \leq ex(H,a,2)$ is a facet of $T(G,a,2)$. $\square$

We'll see that the rank facets we have discussed so far all have the property that adding any edge to them increases their Turán number and can thus be lifted. We now generalize this last theorem to general facets.

Corollary 1.2.22. *Let G and H be two graphs such that $H \subseteq G$. If $\sum_{e\in E(H)} c_e x_e \leq c_0$ is a facet of $T(H,a,2)$ such that for every $e' \in E(G)\backslash E(H)$ there exists x^* such that $\sum_{e\in E(H)} c_e x_e^* = c_0$ for which $x^* \cup e'$ is a-clique-free, then it is also a facet of $T(G,a,2)$.*

Proof. As in the previous corollary, we have that the setup is similar to that of the Padberg lifting property. Here again, let $I := E(G)\backslash E(H)$ and $c_{e'} := c_0 - \max\{\sum_{e\notin I} x_e | x \in T(H,a,2) \text{ and } e' \cup x \text{ is } a-\text{clique-free}\}$ for $e' \in I$. Then, since there exists x^* such that $x^* \cup e'$ is a-clique-free and and such that $\sum_{e\in E(H)} c_e x_e^* = c_0$, we have that $c_{e'} = 0$ for any $e' \in I$. As in the previous corollary, we can add all of the edges in I one after the other without encountering problems, and so we get that $\sum_{e\in E(H)} c_e x_e \leq c_0$ is a facet for $T(G,a,2)$. □

This is good news, and bad news at the same time. Many facets we find will still be facets on higher-dimensional examples; our work on smaller graphs will often carry on to larger graphs. However, the fact that all those facets remain and don't get dominated by others when we add more edges to the graph means that higher-dimensional polytopes will have many, many, many facets, and thus makes it very unlikely that we'll find a complete description for them.

Note also that the Turán problem is very similar to many other combinatorial problems, and as such, some of its facets might be similar to the facets of the polytopes associated to those problems. One problem it resembles is that of finding the maximum stable (or independent) set in a graph. A stable set in a graph is a set of vertices such that none of them are adjacent. The usual integer program used to represent this problem is as follows:

$$\begin{aligned} \max \sum_{v\in V} & x_v \\ \text{s.t. } x_u + x_v \leq 1 & \quad \forall (u,v) \in E \\ x_v \in \{0,1\} & \quad \forall v \in V \end{aligned}$$

Though the Turán problem is edge-based, and the stable set problem, vertex-based, they show some similarities. In the Turán problem, in a k-clique, we want there to be at least one edge missing, and in the stable set problem, we want there to be at least one vertex missing in a 2-clique. The most trivial class of facets of the stable set polytope are the clique inequalities: $\sum_{v \in Q} x_v \leq 1$ for every clique Q. Such inequalities are facets of the stable set polytope for inclusionwise maximal cliques. Clique inequalities are also our most trivial class of facets.

1.2.5 Clique Facets

Theorem 1.2.23. *The clique inequality*

$$\sum_{e \in E(K_n)} x_e \leq ex(n,a,2)$$

is facet-defining for $T(n,a,2)$ if $n \neq 0 \mod (a-1)$.

Proof. It is well-known that any Turán graph with $ex(n,a,2)$ edges is an $(a-1)$-partite complete graph with parts of size as equal as possible (see Theorem 1.1.2). Suppose there are p_1 parts of size $\lfloor \frac{n}{a-1} \rfloor$ and p_2 parts of size $\lceil \frac{n}{a-1} \rceil$ such that $p_1 + p_2 = a-1$ and $p_1 \cdot \lfloor \frac{n}{a-1} \rfloor + p_2 \cdot \lceil \frac{n}{a-1} \rceil = n$.

Let $\alpha x \leq \beta$ be satisfied by all x in the Turán polytope with $\sum_{e \in E(K_n)} x_e = ex(n,a,2)$; then $\alpha(S) = \beta$ for each Turán edge set S with $|S| = ex(n,a,2)$, i.e. the optimal $(a-1)$-partite complete graphs we've just described.

Take two adjacent edges in K_n, without loss of generality, $(n,1)$ and $(1,2)$. We want to show that $\alpha_{(1,n)} = \alpha_{(1,2)}$. Since $n \neq 0 \mod (a-1)$, we know there exists optimal Turán edge sets in the clique such that vertex n is in a part of size $\lfloor \frac{n}{a-1} \rfloor$ and vertices 1 and 2 are together in a part of size $\lceil \frac{n}{a-1} \rceil$ such that $\lfloor \frac{n}{a-1} \rfloor < \lceil \frac{n}{a-1} \rceil$. For example, consider the optimal Turán solution S_1 where vertices $n - \lfloor \frac{n}{a-1} \rfloor + 1$ through n form one part, and vertices 1 through $\lceil \frac{n}{a-1} \rceil$ form another part. Fix the other vertices into a partition P that makes the whole solution optimal.

Now consider another optimal solution S_2 such that vertices $\lceil \frac{n}{a-1} \rceil + 1$ through $n - \lfloor \frac{n}{a-1} \rfloor$ are partitioned into P, vertices $n - \lfloor \frac{n}{a-1} \rfloor + 1$ through n with vertex 1 as well form a part of size $\lceil \frac{n}{a-1} \rceil$ and vertices 2 through $\lceil \frac{n}{a-1} \rceil$ form another part of size $\lfloor \frac{n}{a-1} \rfloor$. This is thus the previous solution but with vertex 1 moved to the other defined part. Note that we couldn't do that if $n = 0 \mod (a-1)$ and still have an optimal solution after the move.

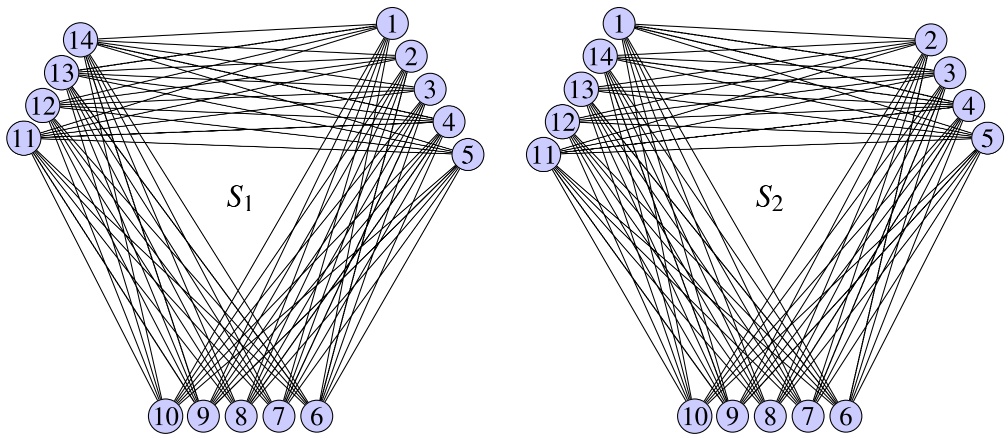

Figure 1.10: Example of S_1 and S_2 for $n = 14$ and $a = 4$.

Since $\alpha(S_1) = \beta = \alpha(S_2)$, this implies that

$$\alpha_{(1,n-\lfloor\frac{n}{a-1}\rfloor+1)} + \alpha_{(1,n-\lfloor\frac{n}{a-1}\rfloor+2)} + \ldots + \alpha_{(1,n)} \\ = \alpha_{(1,2)} + \alpha_{(1,3)} + \ldots + \alpha_{(1,\lceil\frac{n}{a-1}\rceil)}. \quad (1.14)$$

Now consider yet another optimal Turán edge set S_3 such that again vertices $\lceil\frac{n}{a-1}\rceil + 1$ through $n - \lfloor\frac{n}{a-1}\rfloor$ are partitioned into P, vertices $n - \lfloor\frac{n}{a-1}\rfloor + 1$ through $n-1$ with vertex 2 as well form a part of size $\lfloor\frac{n}{a-1}\rfloor$ and vertices 3 through $\lceil\frac{n}{a-1}\rceil$ with vertices 1 and n form another part of size $\lceil\frac{n}{a-1}\rceil$.

Finally, we consider one last optimal Turán solution S_4, again with vertices $\lceil\frac{n}{a-1}\rceil +$ 1 through $n - \lfloor\frac{n}{a-1}\rfloor$ are partitioned into P, vertices $n - \lfloor\frac{n}{a-1}\rfloor + 1$ through $n-1$ with vertices 1 and 2 as well form a part of size $\lceil\frac{n}{a-1}\rceil$ and vertices 3 through $\lceil\frac{n}{a-1}\rceil$ with vertex n form another part of size $\lfloor\frac{n}{a-1}\rfloor$ (thus the previous solution but with vertex 1 moved to the other defined part).

Again, since $\alpha(S_3) = \beta = \alpha(S_4)$, this implies that

$$\alpha_{(1,2)} + \alpha_{(1,n-\lfloor\frac{n}{a-1}\rfloor+1)} + \alpha_{(1,n-\lfloor\frac{n}{a-1}\rfloor+2)} + \ldots + \alpha_{(1,n-1)} \\ = \alpha_{(1,3)} + \alpha_{(1,4)} + \ldots + \alpha_{(1,\lceil\frac{n}{a-1}\rceil)} + \alpha_{(1,n)}. \quad (1.15)$$

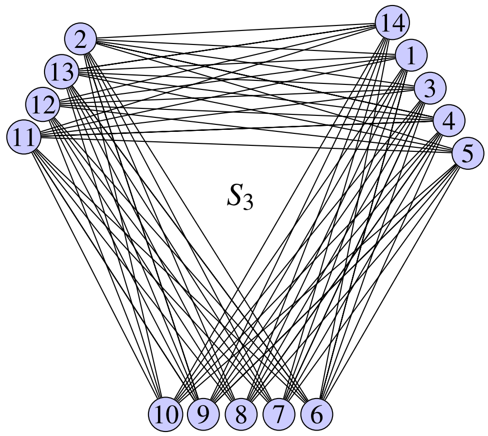

Figure 1.11: Example of S_3 and S_4 for $n = 14$ and $a = 4$.

By substracting equation 1.15 from equation 1.14, we get that

$$\alpha_{(1,n)} - \alpha_{(1,2)} = \alpha_{(1,2)} - \alpha_{(1,n)},$$

which implies that $\alpha_{(1,n)} = \alpha_{(1,2)}$. Since we chose those edges without loss of generality, we know that $\alpha_{e_1} = \alpha_{e_2}$ for any two adjacent edges e_1 and e_2. By applying this observation to all pairs of adjacent edges in the clique, we obtain that $\alpha_{e_3} = \alpha_{e_4}$ for any two edges e_3, e_4 of the K_n. Note that it is clear that $\alpha_e > 0$. Thus α is a positive scalar multiple of the left-hand side of the clique inequality, which is thus facet-defining when $n \neq 0 \mod (a-1)$. □

Corollary 1.2.24. *If G contains a clique of size i with $i \neq 0 \mod (a-1)$, say Q^i, then the corresponding clique inequality is a facet of $T(G,a,2)$.*

Proof. Note that $ex(Q^i + e, a, 2) = ex(i, a, 2) + 1$ for all $e \in E(G) \backslash E(Q^i)$ since any edge e in G but not in the clique can be added to any optimal edge set of the clique without forming an a-clique since other edges containing some vertices of both e and Q^i are missing. Thus, by the lifting theorem 1.2.21, these clique inequalities are facets of $T(G,a,2)$. In particular, they are facets of $T(n,a,2)$ for $n \geq i$. □

Let's try to adapt other classes of facets of the stable set polytope to the Turán problem.

1.2.6 Wheel Facets

An important class of inequalities for the stable set polytope is the odd wheel inequalities:

$$\sum_{v \in V \setminus v_c} x_v + \frac{|V|-2}{2} x_{v_c} \leq \frac{|V|-2}{2}$$

for any odd wheel with center v_c. As we saw previously in example 1.2.12, $Q(G,3,2)$ is also not sufficient to describe $T(G,3,2)$ completely when G is an odd wheel. We now generalize the concept of odd wheel for all a.

Definition 1.2.25. An *a-wheel on n vertices*, denoted by W_n^a, is a graph with a central vertex n and the other vertices 1 through $n-1$ placed cyclically around it. Every $a-1$ consecutive vertices on the cycle form an a clique with vertex n. This graph thus contains $n-1$ cliques of size a. Each *spoke*, i.e. every edge containing vertex n, will be present in $a-1$ cliques. Any other edge (i,j) with $j-i \mod (n-1) \leq a-2$, called *cycle edge*, will be present in $(a-1)-((j-i) \mod (n-1))$ cliques. We only consider wheels for which $n \geq 2a-1$ so that we do not have a complete graph.

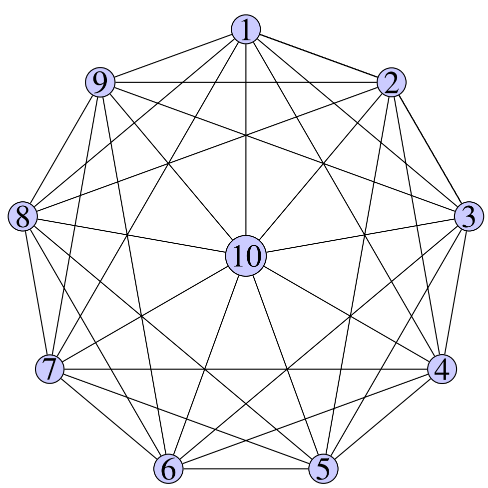

Figure 1.12: Wheel graph W_{10}^5.

Theorem 1.2.26. *The following inequalities are valid and tight for* $T(n,a,2)$

$$\sum_{e\in E(W_l^a)} x_e \leq (a-1)\cdot(l-1) - \left\lceil \frac{l-1}{a-1} \right\rceil,$$

for all wheels W_l^a in K_n with $l \leq n$.

Proof. We first show that the inequality is valid by determining that it is a Chvátal-Gomory cut. Add up the inequalities for the $l-1$ a-cliques present in W_l^a with weight $\frac{1}{a-1}$ and add up all the edge inequalities for cycle edges (i,j) with $(j-i) \mod (l-1) \leq a-2$ with weight $\frac{(j-i) \mod (l-1)}{a-1}$. This yields the inequality

$$\sum_{e\in E(W_l^a)} x_e \leq \left\lfloor \frac{(l-1)\left(\binom{a}{2}-1+\sum_{z=1}^{a-2} z\right)}{a-1} \right\rfloor$$

since spokes are present in $a-1$ cliques of weight $\frac{1}{a-1}$ and any other edge (i,j) is present in $(a-1)-((j-i) \mod (l-1))$ cliques of weight $\frac{1}{a-1}$ and in the edge inequality for (i,j) which has weight $\frac{((j-i) \mod (l-1))}{a-1}$. The right-hand side of the inequality comes from the fact that we're adding up $l-1$ clique inequalities which contain at most $\binom{a}{2}-1$ edges with weight $\frac{1}{a-1}$ and $l-1$ edge inequalies with weight $\frac{z}{a-1}$ for $1 \leq z \leq a-2$ since any cycle edge spans between 2 and $a-1$ vertices on the cycle. One can easily check that

$$\left\lfloor \frac{(l-1)\left(\binom{a}{2}-1+\sum_{z=1}^{a-2} z\right)}{a-1} \right\rfloor = (a-1)\cdot(l-1) - \left\lceil \frac{l-1}{a-1} \right\rceil.$$

We now show that this inequality is tight by producing an edge set of size $(a-1)\cdot(l-1)-\lceil\frac{l-1}{a-1}\rceil$ in $W_l^a \subset K_n$ which contains no clique of size a. Take all $(l-1)\cdot(a-2)$ cycle edges; these do not contain any a-clique. Then add as many spokes as possible without creating any clique of size a. To do so, take $a-2$ consecutive spokes, leave out the next spoke, then take the next $a-2$ consecutive spokes, then leave the next one, and so on until coming back to the first spoke; finally, make sure to leave a spoke out before the initial one.

This edge set is Turán since at least one spoke is missing in each clique, and such a solution contains all edges but $\lceil\frac{l-1}{a-1}\rceil$ spoke edges. Since there are $(a-1)\cdot(l-1)$ edges in W_l^a, this Turán construction thus has

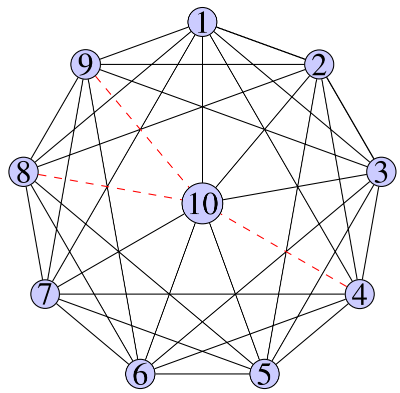

Figure 1.13: Removing three red edges in W_{10}^5 leaves it 5-clique-free.

$$(a-1)\cdot(l-1)-\lceil\frac{l-1}{a-1}\rceil$$

edges, as desired. Thus, these wheel inequalities are valid and tight for $T(n,a,2)$. □

Call the clique on the $a-1$ consecutive vertices on the cycle starting with i, i.e. vertices $i, i+1 \mod (l-1), \ldots, i+a-2 \mod (l-1)$, and central vertex l clique Q_i.

We saw one type of optimal Turán construction for the wheel in the previous proof; call such a construction a *type I construction for the wheel*. There exists a second type of optimal Turán edge set for the wheel, which is called a *type II edge set*. We more or less redo the same construction that we previously saw. First remove the edges $((a-1)\cdot i, l)$ for $1 \leq l \leq \lfloor\frac{l-1}{a-1}\rfloor$. Then the only cliques that are still full are $Q_{\lfloor\frac{l-1}{a-1}\rfloor\cdot(a-1)}$ through Q_{l-1}. Removing any edge contained in every one of these cliques, for example edge $(1, l-1)$, will lead to an optimal Turán construction.

Theorem 1.2.27. *Inequality*

$$\sum_{e\in E(W_l^a)} x_e \leq (a-1)\cdot(l-1)-\left\lceil\frac{l-1}{a-1}\right\rceil$$

is facet-defining for $T(W_l^a, a, 2)$ if $l-1=1 \mod (a-1)$.

Proof. Let $\alpha x \leq \beta$ be a facet of $T(W_l^a, a, 2)$ satisfied by all x in the Turán polytope with $\sum_{e \in E(W_l^a)} x_e = (a-1)\cdot(l-1) - \lceil \frac{l-1}{a-1} \rceil$; then $\alpha(S) = \beta$ for each Turán edge set S with $|S \cap W_l^a| = (a-1)\cdot(l-1) - \lceil \frac{l-1}{a-1} \rceil$.

Consider two distinct consecutive spoke edges, without loss of generality, say $(1,l)$ and $(l-1,l)$.Let

$$S_1 := E(W_l^a) \backslash \left\{ \left\{ (i\cdot(a-1), l) \,|1 \leq i \leq \left\lfloor \frac{l-1}{a-1} \right\rfloor \right\} \cup (1,l) \right\}$$

and

$$S_2 := E(W_l^a) \backslash \left\{ \left\{ (i\cdot(a-1), l) \,|1 \leq i \leq \left\lfloor \frac{l-1}{a-1} \right\rfloor \right\} \cup (l-1,l) \right\},$$

which are clearly two optimal type I Turán sets in W_l^a. This means that $\alpha(S_1) = \beta = \alpha(S_2)$ which implies that $\alpha_{(1,l)} = \alpha_{(l-1,l)}$. Since we can show this for any consecutive spoke edges, we obtain that $\alpha_{e_1} = \alpha_{e_2}$ for any two spoke edges e_1 and e_2.

Now consider a spoke edge and a cycle edge that are adjacent, without loss of generality $(l-1,l)$ and $(l-1,\gamma)$ where $1 \leq \gamma \leq a-2$. Again, if we let

$$S_3 := E(W_l^a) \backslash \left\{ \left\{ (i\cdot(a-1), l) \,|1 \leq i \leq \left\lfloor \frac{l-1}{a-1} \right\rfloor \right\} \cup (l-1,l) \right\}$$

and

$$S_4 := E(W_l^a) \backslash \left\{ \left\{ (i\cdot(a-1), l) \,|1 \leq i \leq \left\lfloor \frac{l-1}{a-1} \right\rfloor \right\} \cup (l-1,\gamma) \right\},$$

then S_3 and S_4 are both optimal Turán solutions, respectively of type I and II, since $l-1 = 1 \mod (a-1)$ and so there is only one clique that is still full after removing $\{(i\cdot(a-1), l) | 1 \leq i \leq \lfloor \frac{l-1}{a-1} \rfloor\}$, namely the one on vertices $1,2,\ldots,a-2, l-1, l$, and so removing any edge from that clique will make the graph a-clique-free. Thus, we have that $\alpha(S_3) = \beta = \alpha(S_4)$ which implies that $\alpha_{(l-1,l)} = \alpha_{(l-1,\gamma)}$ for any $1 \leq \gamma \leq a-2$. Since this is true for any spoke and adjacent cycle edge, we have that $\alpha_{e_1} = \alpha_{e_2}$ for any two edges in W_l^a. It is also clear that $\alpha_e > 0$ for all edges $e \in E(W_l^a)$.

We thus conclude that α is a positive scalar multiple of the left-hand side of the wheel inequality, which is thus facet-defining when $l-1 = 1 \mod (a-1)$. $\square$

Theorem 1.2.28. *The inequalities*

$$\sum_{e \in E(W_l^a)} x_e \leq (a-1)\cdot(l-1) - \left\lceil \frac{l-1}{a-1} \right\rceil$$

are facet-defining for $T(n,a,2)$ for all $W_l^a \subseteq K_n$ with $l-1 = 1 \mod a-1$.

Proof. By the lifting theorem 1.2.21, we simply need to show that there exists a Turán edge set of size $(a-1)\cdot(l-1) - \lceil \frac{l-1}{a-1} \rceil + 1$ in $E(W_l^a) \cup e$ for every $e \in E(K_n) \backslash E(W_l^a)$.

First consider an edge $e = (i,j) \in E(K_n) \backslash E(W_l^a)$ that is not an edge of the wheel. If $e = (i,j)$ such that $i \notin V(W_l^a)$ or $j \notin V(W_l^a)$ or both, then it is clear that we can add this edge to any optimal Turán edge set in W_l^a without creating a clique of size a. Now suppose $e = (i,j)$ such that $i,j \in V(W_l^a)$. Then we know that neither i nor j is the central vertex of the wheel (since all spoke edges are present in the wheel), and so e must be a cycle edge spanning at least a vertices since these are the only edges missing. If e spans more than a vertices, then we can add it to any optimal Turán edge set in W_l^a without creating an a-clique since any such clique containing e would have to contain also another edge spanning at least a vertices, which we know are absent from such an edge set given that they aren't in the wheel in the first place. So we just have to show that there exists an optimal Turán edge set S in W_l^a such that $S \cup e$ is still Turán when e spans a vertices. Fortunately, there exists an optimal solution where an edge $(b, b+1)$ is missing, namely the type II construction we discussed with, without loss of generality, let $e = (a-1, l-1)$ and remove edge $(1, l-1)$ and edges $((a-1)\cdot i, l)$ for $1 \leq l \leq \lfloor \frac{l-1}{a-1} \rfloor$ from the wheel. Adding e to this optimal type II Turán solution does not create an a-clique.

Thus, if $e \in E(K_n) \backslash E(W_l^a)$, then there always exists an optimal Turán edge set S in W_l^a such that $S \cup e$ is also Turán. Therefore, wheel inequalities on l vertices with $l-1 = 1 \mod (a-1)$ will still be facet-defining for $T(n,a,2)$, $n \geq l$. Actually, by

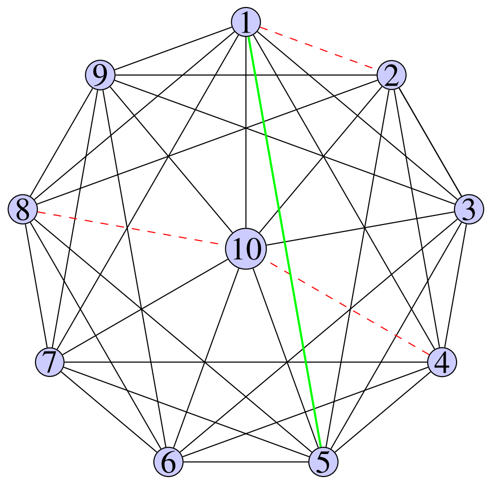

Figure 1.14: Removing three red edges from $W_{10}^5 \cup (1,5)$ leaves it 5-clique-free.

this argument, these wheel inequalities will be facet-inducing for any $T(G,a,2)$ for any graph G that contains such wheels as subgraphs.

□

We now show that in the case of $T(W_l^3,3,2)$, the wheel inequality was all that was missing from $Q(W_l^3,3,2)$ to have a complete description of $T(W_l^3,3,2)$. To do so, we use the well-known concept of total unimodularity which we recall here.

Definition 1.2.29. A matrix A is called *totally unimodular* if and only if each square submatrix has determinant equal to $-1,0$ or 1.

Total unimodularity plays a fundamental role in combinatorial optimization because of the following folkloric consequence. For reference, see [91].

Theorem 1.2.30 (Folklore)**.** *Let $A \in \{-1,0,1\}^{m\times n}$ be a totally unimodular matrix and let $b \in \mathbb{Z}^m$. Then the polyhedron $\{x|Ax \leq b\}$ is integral.*

We also need the following theorem of Ghouila-Houri about total unimodularity.

Theorem 1.2.31 (Ghouila-Houri, 1962)**.** *A matrix A is totally unimodular if and only if each collection R of rows of A can be partitioned into classes R_1 and R_2 such that the sum of the rows in R_1, minus the sum of the rows in R_2, is a vector with entries $0,\pm 1$ only.*

Definition 1.2.32. We say A is *locally totally unimodular* for some vertex v if there exists a totally unimodular submatrix A' of A that yields vertex v.

With those preliminaries out of the way, we can prove the following theorem.

Theorem 1.2.33. *We have that*

$$\begin{aligned}
T(W_l^3,3,2) = \{x \in \mathbb{R}^{|E(W_l^3)|} \mid \quad & x(Q^3) \leq 2 && \forall Q^3 \in \mathcal{Q}^3_{K_n} \\
& x(W_l^3) \leq 2 \cdot (l-1) - \left\lceil \frac{l-1}{2} \right\rceil && \\
& 0 \leq x_e \leq 1 && \forall e \in E(W_l^3)\} \\
=: \{x \in \mathbb{R}^{|E(W_l^3)|} \mid \quad & Ax \leq b\}. &&
\end{aligned}$$

Moreover, this linear description is irredundant if l is even.

Proof. Clearly, $T(W_l^3,3,2)$ is contained in $\{x \in \mathbb{R}^{|E(W_l^3)|} | Ax \leq b\}$. We now show that $\{x \in \mathbb{R}^{|E(W_l^3)|} | Ax \leq b\} \subseteq T(W_l^3,3,2)$ by showing that the system is locally totally unimodular for almost all vertices, and that any other vertex is also integral.

Let x^* be a vertex of $T(W_l^3,3,2)$. We want to show that the determinant of any square submatrix of the inequalities tight at x^* is in $\{-1,0,1\}$. We do so by induction on the size of the submatrix.

We know that a square submatrix of size one is simply an entry a_{ij} of the matrix A, and since our matrix contains only entries in $\{-1,0,1\}$, the hypothesis holds. Suppose it holds also for any submatrix of size at most $\lambda - 1$. Now look at a submatrix of size λ. Four cases arise.

- The submatrix contains a column made up only of zeroes. Then the determinant of this submatrix is zero, and the hypothesis holds.
- The submatrix contains a non-negativity inequality or an edge inequality, that is a row with only zeroes except for one entry that is either 1 or -1. The determinant is then equal to plus or minus the determinant of a $(\lambda-1)$-submatrix, which has by induction its determinant in $\{-1,0,1\}$. Thus the determinant of the λ-submatrix is also in $\{-1,0,1\}$.

- The submatrix contains the wheel inequality and some clique inequalities. Then let R_1 contain all of the clique inequalities present in the submatrix and R_2 contain the wheel inequality. Then $\sum_{i \in R_1} a_{ij} - \sum_{i \in R_2} a_{ij} \in \{-1, 0, 1\}$ for every j since any edge in the submatrix is present in 0, 1 or 2 cliques of the submatrix (so $\sum_{i \in R_1} a_{ij} \in \{0, 1, 2\}$), and each edge is in the wheel inequality (so $\sum_{i \in R_2} a_{ij} = 1$). By theorem 1.2.31, this submatrix is totally unimodular, and so its determinant is 1, 0 or -1.

- The submatrix contains only 3-clique inequalities (but not the wheel inequality), and at least one of them is missing, say the clique on vertices $1, 2, l$. Then put any 3-clique on vertices $i, i+1, l$ present in the submatrix in R_1 if i is odd and in R_2 if i is odd. Then $\sum_{i \in R_1} a_{ij} - \sum_{i \in R_2} a_{ij} \in \{-1, 0, 1\}$ since each sum gives either 0 or 1 because an edge is in at most two consecutive cliques, and no problem occurs when closing the cycle since the first clique is missing. By theorem 1.2.31, this submatrix is totally unimodular.

The only case left is if all clique inequalities are there. Then the submatrix spanning the spoke edges is an odd cycle whose determinant is bigger than one, and so a system that contains it is not totally unimodular. Consider $2(l-1)$ inequalities tight with x^* that are linearly independent and that contains the $l-1$ 3-clique inequalities. Note that we can assume that x^* is suboptimal since we can find a basis of the tight inequalities at x^* containing the wheel inequality (since linear independence forms a matroid), and so, by the third case above, we are certain that any optimal solution is integral.

- If all cycle edges have integral values in x^*, then, if one spoke has a fractional value, it follows that all of them do since no tight 3-clique inequality can contain only one fractional edge and since the 3-cliques are placed in a cycle. Thus, either all spokes have an integral value, which is fine, or they all have a fractional value. In that case, we know that the two spokes in any 3-clique must sum up to 1 and the cycle edge in it must have value 1. Thus, $\sum_{e \in E(W_l^3)} x_e^* = \frac{l-1}{2} + l - 1 > 2 \cdot (l-1) - \lceil \frac{l-1}{2} \rceil$, a contradiction to the wheel inequality, so then x^* cannot be a vertex of $T(W_l^3, 3, 2)$.

- If there exists a cycle edge that has a fractional value in x^*, say edge $(1, 2)$, then at least one other edge in the clique on vertices $1, 2, l$ must also have a fractional value in order for this clique inequality to be tight, say $(2, l)$. In order for the system at hand to be linearly independent, we know it must be must possible to associate each edge to a row of the matrix, i.e. to an inequality tight with x^*, in a specific way. What do we mean by that? Let A_{x^*} be the incidence matrix of the inequalities that are tight with x^*. Note that the rank of A_{x^*} is $2(l-1)$ since x^* is a vertex. The columns of A_{x^*} are thus linearly independent. Linear independence forms a matroid, and as such, if we look at a subset of s columns of A_{x^*}, they should be linearly independent. Note that this implies that the rank of the matrix formed by these s columns is s, and so at least s of these shortened rows must be nonzero.

Since $(1,2)$ is fractional, the system cannot include its non-negativity or edge inequality, and so the only row of A_{x^*} in which it is contained is the clique inequality for $1,2,l$. Let's now consider the submatrix of the columns associated to edges $(1,2)$ and $(2,l)$. There has to be at least two nonzero rows, else the rank of this submatrix would be one, and so A_{x^*} could not be full rank. This means that $(2,l)$, being also fractional, must be associated to the only other clique that contains it, $2,3,l$. Note now that edge $(2,3)$ cannot be fractional, else it would have to be associated to the clique inequality on $2,3,l$ which is already claimed (i.e. this is the only possibility so that the submatrix of the columns associated to edges $(1,2)$, $(2,l)$ and $(2,3)$ has at least three nonzero rows); since it it integral, we associate it to its non-negativity or edge inequality depending on whether it is 0 or 1. Actually, given that the clique on $2,3,l$ is tight with x^*, we know that $x^*_{(2,3)} = 1$. Now edge $(3,l)$ must be fractional in order for the clique on $2,3,l$ to be tight with x^*, and it must be associated to clique $3,4,l$, and so on, until we reach the clique on $l-2,l-1,l$ which will have fractional edges $(l-2,l)$ and $(l-1,l)$; edge $(l-2,l)$ will be associated to the clique on $l-2,l-1,l$ and so edge $(l-1,l)$ must be associated to clique $(1,l-1,l)$. In order for the clique on $1,l-1,l$ to be tight, either edge $(1,l)$ or $(l-1,l)$ will have to be factional. In both cases, there remains no clique for this fractional edge to be associated to, so the system is linearly dependent. Therefore, if there exists a cycle edge with a fractional value, then x^* is not a vertex, a contradiction.

Therefore, there are no fractional vertices for which all cliques are tight, and thus every vertex of the polytope is integral, as desired.

Furthermore, we can now show that each of these facets is essential for the description, that the system is irredundant if l is even.

If one of the 3-clique inequalities was missing, then the point with 1 for each edge in said clique and 0 everywhere else would be a vertex of the polytope, a

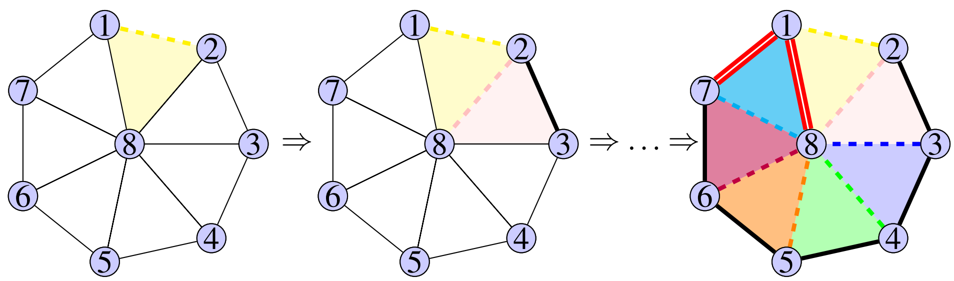

Figure 1.15: Assuming that edge $(1,2)$ is fractional leads to a linearly dependent solution (fractional edges are dashed and of the color of the clique that they are associated to, red double-edges are problematic).

contradiction.

If one of the non-negativity inequalities was missing, then the polytope would be a cone with a negative part, a contradiction.

If one of the edge inequalities was missing, say $x_e \leq 1$, then the point with 2 for e and zero everywhere else would be a vertex of the polytope, a contradiction.

And finally, if the wheel inequality was missing, then the point with 1 for cycle edges and $\frac{1}{2}$ for spoke edges would be a vertex if l is even since the inequalities tight with it form a linearly independent system. If l is odd, then the system is linearly dependent, and so the wheel inequality is redundant. $\square$

1.2.7 Web Facets

We now move on to web graphs, another class of graphs that is important for the stable set polytope.

Definition 1.2.34. A *web* $\overline{W}_{n'}^{a-1}$ is a graph with vertices $1,\ldots,n'$ placed cyclically in order where (i,j) is an edge if i and j differ by at most $a-1 \mod n$ and $i \neq j$. We'll only consider webs for which $n \geq 2a$, else we would have a complete graph.

We use the notation of [76]. Hopefully, it won't confuse the reader. We point out that the web $\overline{W}_{n'}^{a-1}$ is isomorphic to the wheel $W_{n'+1}^{a}$ with the central vertex and its adjacent edges removed.

Note also that every a consecutive vertices on the cycle form an a-clique. This graph thus contains n cliques of size a. Every edge (i,j) with $j-i \leq a \mod n'$ will be present in $(a-1)-(((j-i) \mod n')-1)$ of the a-cliques. Observe also that the facets in $Q(G,3,2)$ are also not sufficient to describe $T(G,3,2)$ completely when G is a web. Consider the following web on seven vertices.

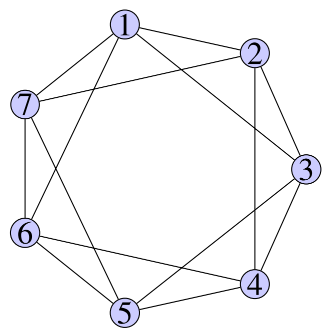

Figure 1.16: Web graph $\overline{W}_7^2$.

Then the following point is a vertex of $Q(\overline{W}_7^2,3,2)$.

	12	13	16	17	23	24	27	34	35	45	46	56	57	67	
(	$\frac{1}{2}$,	1,	1,	$\frac{1}{2}$,	$\frac{1}{2}$,	1,	1,	$\frac{1}{2}$	1,	$\frac{1}{2}$,	1,	$\frac{1}{2}$,	1,	$\frac{1}{2}$	)

We must thus add some cut to get rid of such points.

Theorem 1.2.35. *The following inequalities are valid and tight for $T(n,a,2)$*

$$\sum_{e \in E(\overline{W}_{n'}^{a-1})} x_e \leq (a-1)\cdot n' - \left\lceil \frac{n'}{a-1} \right\rceil,$$

for all webs $\overline{W}_{n'}^{a-1}$ in K_n with $n' \leq n$.

Proof. We first show that the inequality is valid by determining that it is a Chvátal-Gomory cut. Add up the n' a-clique inequalities with weight $\frac{1}{a-1}$ and add up all the edge inequalities for edges (i,j) with $2 \leq j-i \leq a-1 \mod n'$ with weight $\frac{((j-i) \mod n')-1}{a-1}$. This yields the inequality

$$\sum_{e \in E(\overline{W}_{n'}^{a-1})} x_e \leq \left\lfloor \frac{n'\left(\binom{a}{2} - 1 + \sum_{z=1}^{a-2} z\right)}{a-1} \right\rfloor$$

since edges (i, j) such that i and j are consecutive vertices on the cycle are present in $a-1$ cliques of weight $\frac{1}{a-1}$ and any other edge (i, j) is present in $(a-1) - (((j-i) \mod n') - 1)$ cliques of weight $\frac{1}{a-1}$ and in the edge inequality for (i, j) which has weight $\frac{((j-i) \mod n')-1}{a-1}$. The right-hand side of the inequality comes from the fact that we're adding up n' clique inequalities which contain at most $\binom{a}{2} - 1$ edges with weight $\frac{1}{a-1}$ and n' edge inequalies with weight $\frac{z}{a-1}$ for $1 \leq z \leq a-2$ since any other edge spans between 2 and $a-1$ vertices on the cycle. One can easily check that

$$\left\lfloor \frac{n'\left(\binom{a}{2} - 1 + \sum_{z=1}^{a-2} z\right)}{a-1} \right\rfloor = (a-1) \cdot n' - \left\lceil \frac{n'}{a-1} \right\rceil.$$

We now show that this inequality is tight by producing an edge set of size $(a-1) \cdot n' - \lceil \frac{n'}{a-1} \rceil$ in $\overline{W}_{n'}^{a-1} \subset K_n$ which contains no clique of size a. Take all $n' \cdot (a-2)$ edges (i, j) with i and j not consecutive vertices; this does not contain an clique of size a. Then add as many edges (i, j) with i, j consecutive vertices as possible without creating any a-clique. To do so, take $a-2$ consecutive such edges (e.g. $(i, i+1), (i+1, i+2), \ldots (i+a-3, i+a-2) \mod n'$), leave out the next edge $((i+a-2, i+a-1) \mod n'$ in our example), then take the next $a-2$ consecutive edges, then leave the next one, and so on until coming back to the first edge; finally, make sure to leave an edge out before the initial one.

This edge set is Turán since at least one edge is missing in each a-clique, and such a solution contains all edges of the web minus $\lceil \frac{n'}{a-1} \rceil$ edges (i, j) where i, j are consecutive vertices. Since the web graph has $n'(a-1)$ edges, this edge set has $n'(a-1) - \lceil \frac{n'}{a-1} \rceil$ edges, as desired.

Thus, these web inequalities are valid for $T(n, a, 2)$. □

Note that we saw one type of optimal Túran solution for webs in the previous proof; we call such solutions *type I optimal solutions* for web graphs. There exists another type of optimal solution which we call *type II*: without loss of generality, remove $(i \cdot (a-1), i \cdot (a-1) + 1)$ for all $1 \leq i \leq \left\lfloor \frac{n'}{a-1} \right\rfloor$. Then, removing an edge

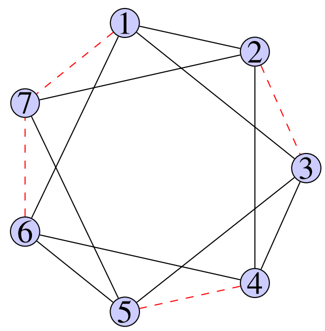

Figure 1.17: Removing four red edges from $\overline{W}_7^2$ leaves it 3-clique-free.

contained in all of the cliques left in the graph will make this a Turán optimal solution.

Theorem 1.2.36. *Inequality*

$$\sum_{e\in\overline{W}_{n'}^{a-1}} x_e \leq (a-1)\cdot n' - \left\lceil \frac{n'}{a-1} \right\rceil$$

is facet-defining for $T(\overline{W}_{n'}^{a-1}, a, 2)$ *if* $n' = 1 \mod (a-1)$.

Proof. Let $\alpha x \leq \beta$ be satisfied by all x in the Turán polytope with $\sum_{e\in E(\overline{W}_{n'}^{a-1})} x_e = (a-1)\cdot n' - \left\lceil \frac{n'}{a-1} \right\rceil$; then $\alpha(S) = \beta$ for each Turán edge set S with $|S \cap \overline{W}_{n'}^{a-1}| = (a-1)\cdot n' - \left\lceil \frac{n'}{a-1} \right\rceil$.

Consider two edges that both span two consecutive vertices and that are adjacent, without loss of generality, say $(1,2)$ and $(1,n')$. Let

$$S_1 := E(\overline{W}_{n'}^{a-1}) \backslash \left\{ \left\{ (i\cdot(a-1), i\cdot(a-1)+1) | 1 \leq i \leq \left\lfloor \frac{n'}{a-1} \right\rfloor \right\} \cup (1,2) \right\}$$

and

$$S_2 := E(\overline{W}_{n'}^{a-1}) \backslash \left\{ \left\{ (i\cdot(a-1), i\cdot(a-1)+1) | 1 \leq i \leq \left\lfloor \frac{n'}{a-1} \right\rfloor \right\} \cup (1,n') \right\},$$

which are clearly two optimal Turán sets of type I in $\overline{W}_{n'}^{a-1}$. This means that $\alpha(S_1) = \beta = \alpha(S_2)$ which implies that $\alpha_{(1,2)} = \alpha_{(1,n')}$. Since we can show this for any consecutive edges spanning two vertices, we obtain that $\alpha_{e_1} = \alpha_{e_2}$ for any two edges e_1 and e_2 spanning two consecutive vertices.

Now consider an edge spanning two vertices and an edge spanning more than two vertices such that they are adjacent, without loss of generality $(1,n')$ and (γ,n') where $2 \leq \gamma \leq a-1$. Again, if we let

$$S_3 := E(\overline{W}_{n'}^{a-1}) \setminus \left\{ \left\{ (i \cdot (a-1), i \cdot (a-1)+1) | 1 \leq i \leq \left\lfloor \frac{n'}{a-1} \right\rfloor \right\} \cup (1,n') \right\}$$

and

$$S_4 := E(\overline{W}_{n'}^{a-1}) \setminus \left\{ \left\{ (i \cdot (a-1), i \cdot (a-1)+1) | 1 \leq i \leq \left\lfloor \frac{n'}{a-1} \right\rfloor \right\} \cup (\gamma,n') \right\},$$

then S_3 and S_4 are both optimal Turán solutions respectively of type I and II since $n' = 1 \mod (a-1)$ and so there is only one clique that is still full after removing $\{(i \cdot (a-1), i \cdot (a-1)+1) | 1 \leq i \leq \lfloor \frac{n'}{a-1} \rfloor\}$, namely the one on vertices $1,2,\ldots,a-1,n'$, and so removing any edge from that clique will make the graph a-clique-free. Thus, we have that $\alpha(S_3) = \beta = \alpha(S_4)$ which implies that $\alpha_{(1,n')} = \alpha_{(\gamma,n')}$ for any $2 \leq \gamma \leq a-1$. Since this is true for any two edges that are adjacent, we have that $\alpha_{e_1} = \alpha_{e_2}$ for any two edges in $\overline{W}_{n'}^{a-1}$. It is also clear that $\alpha_e > 0$ for all edges $e \in E(\overline{W}_{n'}^{a-1})$.

We thus conclude that α is a positive scalar multiple of the left-hand side of the web inequality, which is thus facet-defining when $n' = 1 \mod (a-1)$. $\square$

Theorem 1.2.37. *The inequalities*

$$\sum_{e \in E(\overline{W}_{n'}^{a-1})} x_e \leq (a-1) \cdot n' - \left\lceil \frac{n'}{a-1} \right\rceil$$

are facet-inducing for $T(n,a,2)$ for $n' \leq n$ and $n' = 1 \mod (a-1)$.

Proof. By the lifting theorem 1.2.21, we simply need to prove that for any edge $e \in E(K_n) \backslash E(\overline{W}_{n'}^{a-1})$, there exists a Turán edge set in $E(\overline{W}_{n'}^{a-1}) \cup e$ of size $(a-1) \cdot n' - \left\lceil \frac{n'}{a-1} \right\rceil + 1$. First consider an edge $e \in E(K_n) \backslash E(\overline{W}_{n'}^{a-1})$ that is not an edge of the wheel. If $e = (i, j)$ such that $i \notin V(\overline{W}_{n'}^{a-1})$ or $j \notin V(\overline{W}_{n'}^{a-1})$ or both, then it is clear that we can add this edge to any optimal Turán edge set in $\overline{W}_{n'}^{a-1}$ without creating a clique of size a. Now suppose $e = (i, j)$ such that $i, j \in V(\overline{W}_{n'}^{a-1})$. Then we know that e must span at least $a+1$ vertices. If it spans at least $a+2$ vertices, then we can add it to any optimal Turán edge set in $\overline{W}_{n'}^{a-1}$ without creating an a clique since any such clique containing e would have to contain also another edge spanning at least $a+1$ vertices, which we know are absent from such an optimal solution given that they aren't in the web in the first place. If e spans exactly $a+1$ vertices, say $1, \ldots, a+1$, then we can use the optimal construction for webs we've seen before so that it contains all of the edges, except $\{(1+i \cdot (a-1), 2+i \cdot (a-1)) | 0 \leq i \leq \lfloor \frac{n'}{a-1} \rfloor\}$, and then adding e to this solution leaves it Turán if $n' \geq 2a+2$. If $n' = 7$ and $a = 3$ (the only case for which $n' = 2a$ or $n' = 2a+1$ and $n' = 1 \mod (a-1)$ and thus need to be more careful because adding an edge can create cliques on both side of the edge), then we remove $\{(1+i \cdot (a-1), 2+i \cdot (a-1)) | 0 \leq i \leq \lfloor \frac{n'}{a-1} \rfloor - 1\} \cup \{(1,6)\}$.

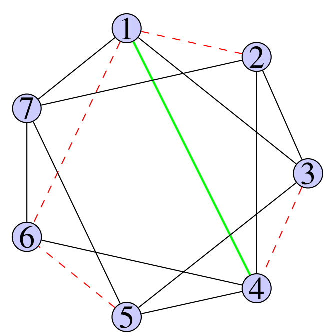

Figure 1.18: Removing four red edges from $\overline{W}_7^2 \cup (1,4)$ leaves it 3-clique-free.

Thus, if $e \in E(K_n) \backslash E(\overline{W}_{n'}^{a-1})$, then there always exists an optimal Turán edge set S in $\overline{W}_{n'}^{a-1}$ such that $S \cup e$ is also Turán. So the web inequality for $\overline{W}_{n'}^{a-1}$ with $n' = 1 \mod (a-1)$ is a facet for $T(n, a, 2)$ with $n \geq n'$. Moreover, by this argument, it is also a facet for $T(G, a, 2)$ for any G that contains $\overline{W}_{n'}^{a-1}$ as a subgraph. □

Theorem 1.2.38. *We have that*

$$\begin{aligned}
T(\overline{W}^2_{n'},3,2) = \{x \in \mathbb{R}^{|E(\overline{W}^2_{n'})|} \mid \quad & x(Q^3) \leq 2 && \forall Q^3 \in \mathcal{Q}^3_{K_n} \\
& x(\overline{W}^2_{n'}) \leq 2 \cdot n' - \left\lceil \frac{n'}{2} \right\rceil && \\
& 0 \leq x_e \leq 1 && \forall e \in E(\overline{W}^2_{n'})\} \\
=: \{x \in \mathbb{R}^{|E(\overline{W}^2_{n'})|} \mid \quad & Ax \leq b\}. &&
\end{aligned}$$

Moreover, none of these inequalities are redundant if n' is odd.

Proof. Clearly, $T(\overline{W}^2_{n'},3,2)$ is contained in $\{x \in \mathbb{R}^{|E(\overline{W}^2_{n'})|} | Ax \leq b\}$. We now show that $\{x \in \mathbb{R}^{|E(\overline{W}^2_{n'})|} | Ax \leq b\} \subseteq T(W^2_{n'},3,2)$ by showing that the system is locally totally unimodular for most vertices and that anhy other vertex is integral.

Let x^* be a vertex of $T(\overline{W}^2_{n'},3,2)$. We want to show that the determinant of any square submatrix of the inequalities tight at x^* is in $\{-1,0,1\}$, and thus that system is locally totally unimodular, meaning that x^* is integral. We do so by induction on the size of the submatrix.

We know that a square submatrix of size one is simply an entry a_{ij} of the matrix A, and since our matrix contains only entries in $\{-1,0,1\}$, the hypothesis holds. Suppose it holds also for any submatrix of size at most $\lambda - 1$. Now look at a submatrix of size λ. Four cases arise.

- The submatrix contains a column made up only of zeroes. Then the determinant of this submatrix is zero, and the hypothesis holds.
- The submatrix contains a non-negativity inequality or an edge inequality, that is, a row with only zeroes except for one entry that is either 1 or -1. The determinant is then equal to plus or minus the determinant of a $(\lambda - 1)$-submatrix, which has by induction determinant in $\{-1,0,1\}$. Thus the determinant of the λ-submatrix is also in $\{-1,0,1\}$.
- The submatrix contains the web inequality as well as some clique inequalities. Then put the clique inequalities in R_1 and the web inequality in R_2. We thus obtain that $\sum_{i \in R_1} a_{ij} - \sum_{i \in R_2} a_{ij} = \{-1,0,1\}$ since $\sum_{i \in R_1} a_{ij} \in \{0,1,2\}$ for every j since any edge is in at most two 3-cliques and $\sum_{i \in R_2} a_{ij} = 1$ for all j since only the web inequality is in R_2. By theorem 1.2.31, this submatrix is totally unimodular.

- The submatrix contains only 3-clique inequalities, and at least one of them is not present. Let Q_i be the clique going through vertices $i, i+1, \ldots, i+a-1$ mod (n'). Suppose Q^1 is missing. Then put an inequality Q_i that's present in R_1 if i is even and in R_2 if i is odd. Then $\sum_{i \in R_1} a_{ij} - \sum_{i \in R_2} a_{ij} = \{-1, 0, 1\}$ for all j since any edge spanning three vertices will be in at most one clique present and thus yield either -1,0 or 1, and any edge spanning two vertices will be in at most two consecutive which will not be in the same R_i. By theorem 1.2.31, this submatrix is totally unimodular.

Thus we have that any vertex for which there is at least one clique that is not tight with it is integral. Now consider a vertex for which all cliques are tight. Without loss of generality, we can imagine that x^* is suboptimal, otherwise there would exist a set of $E(\overline{W}^2_{n'})$ tight inequalities with x^* that form a linearly independent system containing the web inequality; we would then be in case 3 above. Consider $2n'$ inequalities tight with x^* that are linearly independent and that contains the n' 3-clique inequalities and n' edge or non-negativity inequalities. Two cases arise.

- Suppose that all edges spanning 3 vertices have integral values in x^*. If all edges spanning 2 vertices also have integral values, then that's fine, x^* is an integral point. However, if at least one edge spanning 2 vertices has a fractional value, then all of them do since no 3-clique inequality can contain only one fractional edge and the 3-cliques are placed in a cycle. In that case, we know that the two edges spanning 2 consecutive vertices in any 3-clique must sum up to 1 (else they wouldn't be fractional and/or the clique inequality would not be tight) and the edge spanning 3 vertices in it must have value 1. Thus, $\sum_{e \in E(\overline{W}^2_{n'})} x^*_e = \frac{n'}{2} + n' > 2 \cdot n' - \left\lceil \frac{n'}{2} \right\rceil$, meaning that x^* does not respect the web inequality, a contradiction, and so x^* cannot be a vertex of $T(\overline{W}^2_{n'}, 3, 2)$.
- If there exists an edge spanning 3 vertices that has a fractional value in x^*, say edge $(1,3)$, then at least one other edge in the clique on vertices $1,2,3$ must also have a fractional value in order for this clique inequality to be tight, say $(2,3)$. In order for the system at hand to be linearly independent, we know it must be must possible to associate each edge to a row of the matrix, i.e. to an inequality tight with x^*, in a specific way. Indeed, let A_{x^*} be the incidence matrix of the inequalities that are tight with x^*. Note that the rank of A_{x^*} is $2(l-1)$ since x^* is a vertex. The columns of A_{x^*} are thus linearly independent. Linear independence forms a matroid, and as such, if

we look at a subset of s columns of A_{x^*}, they should be linearly independent. Note that this implies that the rank of the matrix formed by these s columns is s, and so at least s of these shortened rows must be nonzero.

Since $(1,3)$ is fractional, the system cannot include its non-negativity or edge inequality, and so the only row in which it is contained is the clique inequality for $1,2,3$. This means that $(2,3)$, being also fractional, must be associated to clique $2,3,4$ so that, if we look at the submatrix of A_{x^*} with the columns associated to $(1,3)$ and $(2,3)$, there are at least two nonzero rows. Note now that edge $(2,4)$ cannot be fractional, else it would have to be associated to the clique inequality on $2,3,4$ which is already claimed (i.e. if $(2,4)$ were fractional, the submatrix with columns associated to $(1,3)$, $(2,3)$, and $(2,4)$ would have only two nonzero rows, and thus would not be linearly independent, meaning that A_{x^*} is also not full rank); since it it integral, we associate it to its non-negativity or edge inequality depending on whether it is 0 or 1. Actually, give that the clique on $2,3,4$ must be tight, and since edge $(2,3)$ is fractional, $x^*_{(2,4)} = 1$. Now edge $(3,4)$ must be fractional in order for the clique on $2,3,4$ to be tight with x^*, and it must be associated to clique $3,4,5$, and so on, until we reach the clique on $n'-1,n',1$ which will have fractional edges $(n'-1,n')$ and $(n',1)$; edge $(n'-1,n')$ will be associated to the clique on $n'-1,n',1$ and so edge $(n',1)$ must be associated to the clique on $1,2,n'$. Note that we hadn't decided at the beginning if edge $(1,2)$ is integral or fractional. Since the clique on vertices $1,2,n'$ is tight, then either $(2,n')$ or $(1,2)$ is fractional (in addition to edge $(1,n')$). In both cases, there remains no clique for this fractional edge to be associated with, and so the system is linearly dependent. Therefore, if there exists an edge spanning 3 vertices with a fractional value, then x^* is not a vertex, a contradiction.

Therefore, there are no fractional vertices for which all cliques are tight, and thus every vertex of the polytope is integral, as desired.

Furthermore, we can now show that each of these facets are essential for the description, that none are redundant if n' is odd.

If one of the a-clique inequalities was missing, then the point with 1 for each edge in said clique and 0 everywhere else would be a vertex of the polytope, a contradiction.

If one of the non-negativity inequalities was missing, then the polytope would be

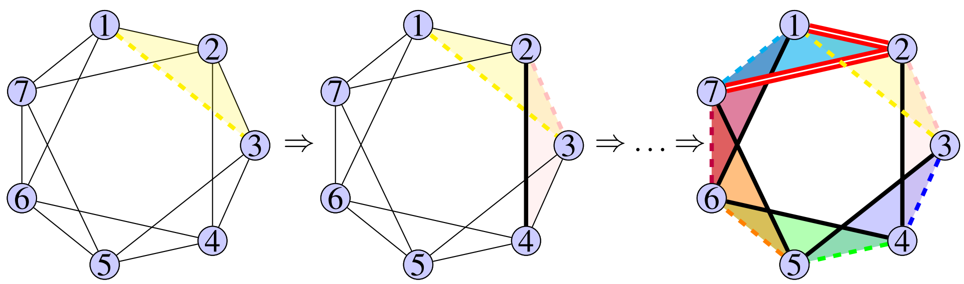

Figure 1.19: Assuming that edge $(1,3)$ is fractional leads to a linearly dependent solution (fractional edges are dashed and of the color of the clique that they are associated to, red double-edges are problematic).

a cone with a negative part, a contradiction.

If one of the edge inequalities was missing, say $x_e \leq 1$, then the point with 2 for e and zero everywhere else would be a vertex of the polytope, a contradiction.

And finally, if the web inequality was missing, then the point with 1 for edges spanning 3 vertices and $\frac{1}{2}$ for edges spanning 2 vertices would be a vertex of the polytope, since the clique and edge inequalities tight this point form a linearly independent system if n' is odd. If n' is even, this system is linearly dependent and so the web inequality is redundant.

□

1.2.8 Doubling Facets

In the polyhedral proof of the Turán theorem, we introduced another type of valid inequality, which we call the *doubling inequality*. We showed there that

$$\sum_{e \in \delta(v)} 2x_e + \sum_{e \in E(K_n) \setminus \delta(v)} x_e \leq ex(n+1, a, 2)$$

is a valid inequality for $T(n,a,2)$ for any $v \in [n]$ since copying any vertex in a Turán edge set gives an edge set that is also Turán. We now show that this inequality is sometimes facet-defining.

Theorem 1.2.39. *The doubling inequality*

$$\sum_{e \in \delta(v)} 2x_e + \sum_{e \in E(K_n) \setminus \delta(v)} x_e \leq ex(n+1, a, 2)$$

is facet-defining for any $v \in [n]$ *for* $T(n,a,2)$ *when* $n = 0 \mod (a-1)$ *and with* $n \geq 3(a-1)$.

Proof. We first observe that there are two types of edge sets that are tight with the doubling inequality for $n = 0 \mod (a-1)$. The first type, which we call type I, is simply the optimal Turán edge set of a clique, that is, a complete $(a-1)$-partite graph with each part containing $\frac{n}{a-1}$ vertices. There are $\binom{\frac{n}{a-1}}{2} \cdot (\frac{n}{a-1})^2$ edges in such a solution, including $n - \frac{n}{a-1}$ edges that get doubled in the doubling inequality, thus the left-hand side of the inequality yield $\binom{\frac{n}{a-1}}{2} \cdot (\frac{n}{a-1})^2 + n - \frac{n}{a-1}$ which is equal to $ex(n+1, a, 2)$, as desired.

The second type of optimal Turán edge set, called type II, is given by the same construction, but with one vertex being moved to another part and the vertex v that gets doubled being in the part that lost a vertex. That is, an $(a-1)$-partite graph with all parts containing $\frac{n}{a-1}$ vertices except for two parts, one containing $\frac{n}{a-1} - 1$ vertices (including v) and one containing $\frac{n}{a-1} + 1$ vertices.

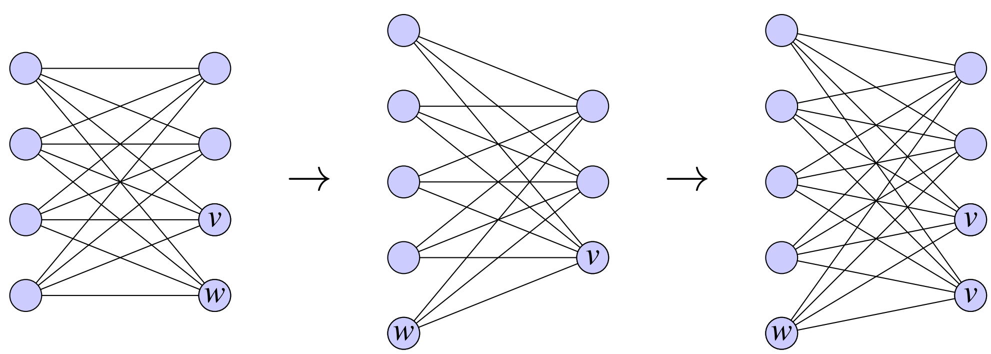

Figure 1.20: Constructing an optimal solution of type II for the doubling inequality for $T(8,3,2)$

We now show that the number of edges in a type II solution is the same as in a type I solution by calculating how many edges we lose and gain. So by moving a vertex from one part to another in the optimal Turán edge set for the clique to get the basis of a type II solution, the only edges that change are those adjacent to the vertex w we move. In the original clique solution, the vertex w we move was adjacent to $n - \frac{n}{a-1}$ edges, none of them adjacent to v, so forming a total of $n - \frac{n}{a-1}$

in the left-hand side of the doubling inequality. In the second type of solution, w is adjacent to only $n - \frac{n}{a-1} - 1$ edges, but one of them is adjacent to v, and so they form a total of $n - \frac{n}{a-1}$ in the left-hand side of the doubling inequality. So the total change is zero, and so the second type of solution is also tight with the doubling inequality.

Note that the type II construction is not always optimal for the doubling inequality when $n \neq 0 \mod a-1$, and observe that the type I constructions are already in the clique facet in these cases.

Let $\alpha x \leq \beta$ be a facet of $T(n,a,2)$ satisfied by all x in the Turán polytope with $\sum_{e \in \delta(v)} 2x_e + \sum_{e \in E(K_n) \setminus \delta(v)} x_e = ex(n+1,a,2)$; then $\alpha(S) = \beta$ for both types of Turán edge sets that we've just described.

Without loss of generality, let $v = \frac{n}{a-1}$. Take two adjacent edges that do not contain v, without loss of generality, $(n,1)$ and $(1,2)$. First consider the optimal solution S_1 of type I where vertices $i \cdot \frac{n}{a-1} + 1$ through $(i+1) \cdot \frac{n}{a-1}$ form a part for $0 \leq i \leq a-2$.

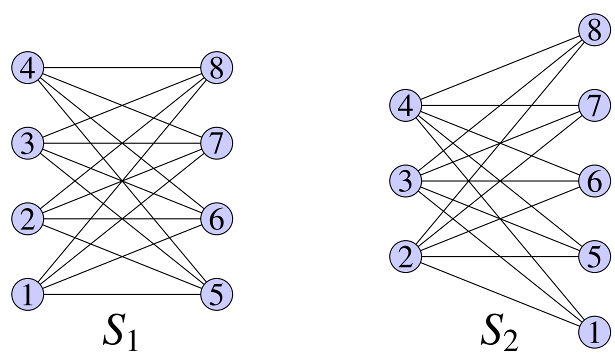

Figure 1.21: S_1 and S_2 for $T(8,3,2)$

Now consider the solution of type II S_2 which is the same as S_1 but with vertex 1 moved to the part containing vertices $n - \frac{n}{a-1} + 1$ through n.

Since $\alpha(S_1) = \beta = \alpha(S_2)$, we have that

$$\alpha_{(1,n-\frac{n}{a-1}+1)} + \alpha_{(1,n-\frac{n}{a-1}+2)} + \ldots + \alpha_{(1,n)} = \alpha_{(1,2)} + \alpha_{(1,3)} + \ldots + \alpha_{(1,\frac{n}{a-1})}.$$

Consider now two new optimal solutions. First, we consider one of type I, say S_3, where vertices $i \cdot \frac{n}{a-1}+1$ through $(i+1)\cdot\frac{n}{a-1}$ form a part for $1 \leq i \leq a-3$, vertices 3 through $\frac{n}{a-1}$ with vertices n and 1 form another part, and vertices $n-\frac{n}{a-1}+1$ through $n-1$ with vertex 2 form a final part.

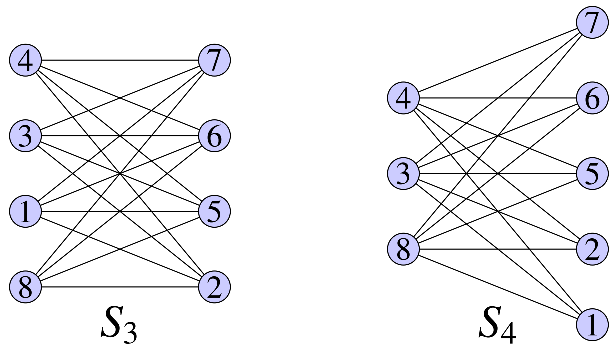

Figure 1.22: S_3 and S_4 for $T(8,3,2)$

Second, we consider a final solution S_4 of type II which is the same as S_3 but with vertex 1 moved to the part containing vertices $n-\frac{n}{a-1}+1$ through $n-1$ as well as vertex 2.

Since $\alpha(S_3)=\beta=\alpha(S_4)$, we obtain that

$$\alpha_{(1,2)}+\alpha_{(1,n-\frac{n}{a-1}+1)}+\alpha_{(1,n-\frac{n}{a-1}+2)}+\dots\alpha_{(1,n-1)} \\ =\alpha_{(1,3)}+\alpha_{(1,4)}+\dots+\alpha_{(1,\frac{n}{a-1})}+\alpha_{(1,n)}.$$

Substracting these two equations, we get that

$$\alpha_{(1,n)}-\alpha_{(1,2)}=\alpha_{(1,2)}-\alpha_{(1,n)},$$

meaning that $\alpha_{(1,n)}=\alpha_{(1,2)}$. Since edges $(1,2)$ and $(1,n)$ were chosen without loss of generality as adjacent edges not containing v, we get that $\alpha_{e_1}=\alpha_{e_2}$ for any two edges e_1,e_2 not containing v by applying this fact to pairs of adjacent edges not containing v repetitively. Let $\alpha_e=A$ for any edge not containing v which we fixed to be $\frac{n}{a-1}$ at the beginning. Then the first equation becomes

$$\left(n-\frac{n}{a-1}\right)\cdot A=\left(n-\frac{n}{a-1}-2\right)\cdot A+\alpha_{(1,\frac{n}{a-1})},$$

which yields that $\alpha_{(1,\frac{n}{a-1})} = 2A$. Since $v = \frac{n}{a-1}$ was chosen without loss of generality, we get that $\alpha_{e'} = 2A$ for any edge e' such that $v \in e'$. Note that it is clear that $\alpha_e > 0$ for any edge e, and thus we can conlude that α is a positive scalar multiple of the left-hand side of the doubling inequality. Since we know the doubling inequality is tight with $T(n,a,2)$ as we've seen through the two types of construcrtions, it is facet-defining. Finally, note that the sets S_1,S_2,S_3 and S_4 all exist only if each part contains at least three vertices, that is, if $n \geq 3(a-1)$. $\square$

Theorem 1.2.40. *The doubling inequality*

$$\sum_{e\in\delta(v)} 2x_e + \sum_{e\in E(K_i)\setminus\delta(v)} x_e \leq ex(i+1,a,2)$$

is facet-defining for any $v \in V(K_i)$ for $T(n,a,2)$ when $i = 0 \mod (a-1)$, with $i \geq 3(a-1)$ and $n \geq i$.

Proof. By the lifting theorem 1.2.22, we only need to show that for every $e \in E(K_n)\setminus E(K_i)$, there exists an optimal solution S of the doubling inequality in K_i such that $S \cup e$ is Turán. This is clear since, for any such edge e, $S \cup e$ is Turán for any optimal solution S of the doubling inequality because any a-clique containing the vertices of e as well as some vertices of K_i is missing some edges. $\square$

Clearly, $T(n,a,2)$ has many more facets than the ones we've spoken about. Since the optimal solutions of the Turán graph problem are already known, it's actually quite easy to produce more facets by using proofs like the ones we've seen so far. For example, instead of doubling just one vertex like in the last inequality, we could double two vertices, say v_1, v_2. Then

$$4 \cdot x_{(v_1,v_2)} + 2 \cdot \sum_{\substack{e\in E(K_n):\\ v_1 \text{ or } v_2\in e\\ \text{but not both}}} x_e + \sum_{\substack{e\in E(K_n):\\ v_1\notin e, v_2\notin e}} x_e \leq T(n+2,a,2)$$

is facet-defining if $n = 1 \mod (a-1)$, and we can keep playing this game by removing more vertices. Similarly, it is easy to come up with non-rank facets for the web and wheel graphs. However, these proofs all rely on the knowledge that we know what optimal solutions look like, knowledge that we are lacking in the next section.

1.3 Modeling the Turán Hypergraph Problem

1.3.1 The Model

We now discuss another generalization of the Turán problem: transferring the problem to hypergraphs. Some of the polyhedral study that we've done in the previous section can easily be extended to hypergraphs.

Even though the problem of finding the maximum number of edges in a K_a-free graph on n vertices has been solved a thousand (and now one) times, nobody has yet managed to solve the problem of finding the maximum number of r-hyperedges in a K_a^r-free r-uniform hypergraph on n vertices for any $a > r > 2$.

We now reintroduce the notation for hypergraphs. A r-uniform hypergraph is one where all edges have cardinality r. By K_a^r, we denote a r-uniform complete graph of size a, i.e. a hypergraph on a vertices for which the edges correspond to every set of r vertices. We let $ex(n,a,r)$ be the maximum number of r-hyperedges in a K_a^r-free r-uniform hypergraph on n vertices, and $ex(G,a,r)$ be the maximum number of r-hyperedges in a K_a^r-free subgraph of G. We denote by $T(n,a,r)$ the convex hull of the characteristic vectors of all r-edge sets $F \in E(K_n^r)$ that contain no r-hyperclique of size a; we call $T(n,a,r)$ the (n,a,r)-*Turán polytope* or simply *Turán polytope* when n,a,r are already clearly stated. Similarly, for a r-uniform graph G, we let $T(G,a,r)$ be the convex hull of the characteristic vectors of all r-edge sets $F \in E(G)$ that contain no r-hyperclique of size a. From now on, we might drop the prefix 'hyper-' to make reading easier when no confusion is possible.

As in the case for $r = 2$, we want to investigate the facet structure of $T(n,a,r)$.

Proposition 1.3.1. *We have that*

$$T(n,a,r) = conv\left(\left\{x \in \{0,1\}^{\binom{n}{r}} | \sum_{e \in E(Q^a)} x_e \leq \binom{a}{r} - 1 \quad \forall Q^a \in \mathcal{Q}^a_{K_n^r}\right\}\right).$$

Proof. Any vertex of $T(n,a,r)$ corresponds to some a-Turán edge set in K_n^r which means that for any induced subgraph on a vertices, there are at most $\binom{a}{r} - 1$ edges there, since otherwise there would be a clique of size a. Thus each vertex of

$T(n,a,r)$ satisfies is a $0/1$-vector that respects the constraints present in the right-hand side set, that is,

$$T(n,a,r) \subseteq \operatorname{conv}\left(\left\{x \in \{0,1\}^{\binom{n}{r}} \mid \sum_{e \in E(Q^a)} x_e \leq \binom{a}{r} - 1 \quad \forall Q^a \in \mathcal{Q}^a_{K^r_n}\right\}\right).$$

Similarly, any vertex of the right-hand side is the characteristic vector of an a-Turán edge set, and must thus be in $T(n,a,r)$. Thus equality holds. $\square$

Like we did for the Turán graph model, we make the model a bit stronger by noticing the following easy fact.

Proposition 1.3.2. *For any a-Turán edge set F on n vertices, there is at most $ex(i,a,r)$ hyperedges induced by any i vertices, for $a \leq i \leq n$.*

Proof. Any subset of hyperedges induced by some i vertices of an a-Turán edge set F is also a-Turán since, if it contained a r-uniform hyperclique of size a, then so would F, a contradiction. Thus, the maximum number of hyperedges induced by those i vertices is $ex(i,a,r)$. $\square$

This easy observation is used to come up with the following linear model.

Definition 1.3.3. We let

$$Q(n,a,r) = \left\{ x \in \mathbb{R}^{\binom{n}{r}} \mid \begin{array}{ll} \sum_{e \in E(Q^i)} x_e \leq ex(i,a,r) & \forall Q^i \in \mathcal{Q}^i_{K^r_n}, \forall a \leq i \leq n-1 \\ 0 \leq x_e \leq 1 & \forall e \in E(K^r_n) \end{array} \right\}$$

be the *clique-relaxation of the Turán polytope*.

This linear program is the linear relaxation of $T(n,a,r)$ with hyperclique inequalities added for any r-uniform hyperclique of size greater or equal to a and smaller than $n-1$. We do not include cliques of size n since our strategy will be again to try to calculate the value of $ex(n,a,r)$ by using this program inductively, and thus the program cannot include the value $ex(n,a,r)$.

Proposition 1.3.4. *$T(n,a,r) \subseteq Q(n,a,r)$ and $T(n,a,r) = conv\left(Q(n,a,r) \cap \mathbb{Z}^{\binom{n}{r}}\right)$.*

Proof. Any vertex of $T(n,a,r)$ corresponds to an a-Turán edge set, and by the previous proposition, we know such an edge set respects $\sum_{e\in E(Q^i)} x_e \leq ex(i,a,r)$ for all $Q^i \in Q^i_{K_n^r}$ and for all $a \leq i \leq n$. Moreover, any integral point of $Q(n,a,r)$ correspond to an a-Turán edge set. □

Note that we only consider $a \leq i \leq n-1$ since our strategy will be to calculate the value of $ex(n,a,r)$ by using this program inductively, and thus the program cannot include the value $ex(n,a,r)$.

Remark. Since $T(n,a,r) = \text{conv}\left(Q(n,a,r) \cap \mathbb{Z}^{\binom{n}{r}}\right)$, any Gomory-Chvátal cut developed from inequalities in $Q(n,a,r)$ is valid for $Q(n,a,r) \cap \mathbb{Z}^{\binom{n}{r}}$ and thus is also valid for $T(n,a,r)$.

1.3.2 The Turán Conjecture for $a = 4$ and $r = 3$

As we've previously mentioned, the value of $ex(n,a,r)$ is not known for $r \geq 3$. There is however a conjecture for $ex(n,4,3)$. Turán was the first to come up with it.

Conjecture 1.3.5 (Turán)**.** *The maximal number of edges in a* 3*-uniform hypergraph with n vertices containing no hyperclique of size* 4 *is*

$$ex(n,4,3) = \begin{cases} \frac{5m^3-3m^2}{2} & \text{if } n = 3m \\ \frac{5m^3+2m^2-m}{2} & \text{if } n = 3m+1 \\ \frac{5m^3+7m^2+2m}{2} & \text{if } n = 3m+2 \end{cases}$$

Different constructions yield those numbers (see for example [15], [43], [63]), so these are at least some lowerbounds for $ex(n,4,3)$. The first construction ever found is the following.

Example 1.3.6 (Construction of Turán for $ex(n,4,3)$)**.** Partition your n vertices into three sets of sizes as equal as possible, say V_1, V_2, V_3. Any set of three vertices such that two of them are in V_1 and one is in V_2, or two in V_2 and one in V_3, or two in V_3 and one in V_1 forms an edge.

Moreover, any set of three vertices such that each vertex comes from a different part of the partition also forms an edge.

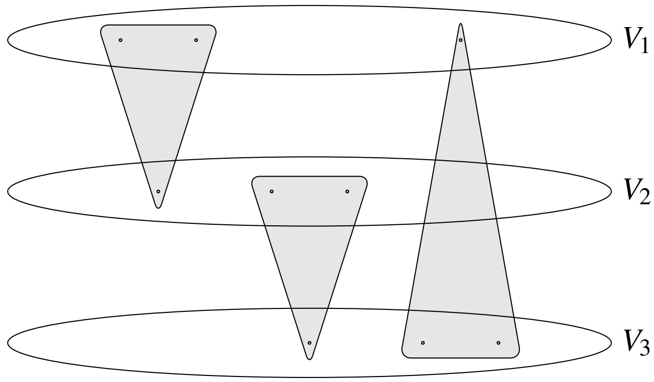

Figure 1.23: First type of edges of the construction.

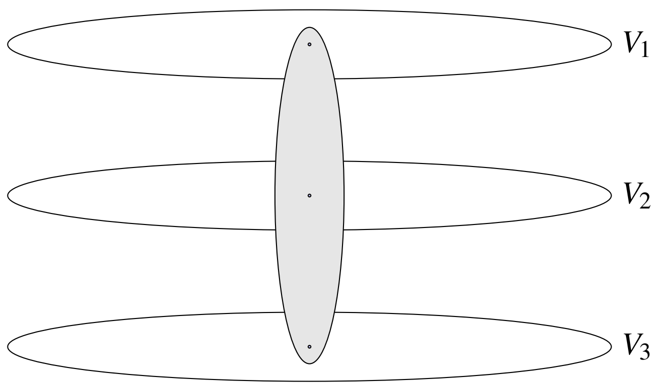

Figure 1.24: Second type of edges of the construction.

Then any set of four vertices that contains three or four vertices in the same part cannot be a clique since every part is a stable set. Any set of four vertices with two vertices in one part and two in another part will induce exactly two edges, and so these four vertices will once again not be a clique. Finally, if we have two vertices in one part, and one in each of the other two parts, then one of the two 3-sets that contain the two vertices from one part and one of the other two vertices will not be an edge. Thus, the construction is indeed K_4^3-free.

If $n = 3m$ for some $m \in \mathbb{N}$, then each part in this construction will contain exactly m vertices. There will be $\binom{m}{2} \cdot m$ edges with two vertices in V_1 and one in V_2, as well as for the two other similar types of edges, and m^3 edges with one vertex from each part. In total, there is thus $3 \cdot \binom{m}{2} \cdot m + m^3 = \frac{5m^3 - 3m^2}{2}$. Similar calculations show that the number of edges in the construction are those given in the conjecture.

For graphs, the integrality gap between $Q(n,a,2)$ and $T(n,a,2)$ was quite small, especially when maximizing in the $\mathbb{1}$ direction when it was always less than one. This is unfortunately not the case here if the conjecture is true.

Indeed, suppose now we try to find $ex(n,a,r)$ in an inductive way, i.e. imagine that we know that the conjecture holds true up to $ex(n-1,4,3)$ and that we want that it still holds for $ex(n,4,3)$. If $n = 3m$ for some $m \in \mathbb{N}$, then we can sum up the $3m$ cliques of size $3m-1$

$$\sum_{Q^{3m-1} \in \mathcal{Q}^{3m-1}_{K_n^3}} \left(\sum_{e \in E(Q^{3m-1})} x_e \leq ex(3m-1,4,3) \right)$$

which yields the cut

$$\sum_{e \in E(K_{3m}^3)} x_e \leq \left\lfloor \frac{3m}{3m-3} ex(3m-1,4,3) \right\rfloor .$$

Since we assume that the conjecture is true up to $3m-1$, then

$$\begin{aligned} ex(3m,4,3) \leq & \left\lfloor \frac{3m}{3m-3} ex(3m-1,4,3) \right\rfloor \\ & \leq \frac{3m}{3m-3} \cdot \frac{5m^3-8m^2+3m}{2} = \frac{5m^3-3m^2}{2} . \end{aligned}$$

But construction 1.3.6 ensures that $ex(3m,4,3) \geq \frac{5m^3-3m^2}{2}$, thus $ex(3m,4,3) = \frac{5m^3-3m^2}{2}$. So far, so good.

If we could also prove that the conjecture holds for $ex(3m+1,4,3)$ assuming that it is true up to $3m$ and that it also holds for $ex(3m+2,4,3)$ assuming that it is true up to $3m+1$, then we'd be done. However, the cut

$$ex(n,4,3) \leq \frac{n}{n-3} ex(n-1,4,3)$$

does not close the integrality gap in the $\mathbb{1}$-direction sufficiently for the other two cases. For $n = 3m+1$, this cut yields by induction

$$ex(3m+1,4,3) \leq \left\lfloor \frac{3m+1}{3m-2} \cdot \frac{5m^3-3m^2}{2} \right\rfloor = \frac{5m^3+2m^2-m}{2} + \left\lfloor \frac{2m}{3} \right\rfloor,$$

which is $\lfloor \frac{2m}{3} \rfloor$ units greater than the value given by construction 1.3.6. Thus, as m augments, the integrality gap in the $\mathbb{1}$ direction also gets wider for $n = 3m+1$.

Similarly, for $n = 3m+2$, the cut yields

$$ex(3m+2,4,3) \leq \left\lfloor \frac{3m+2}{3m-1} \cdot \frac{5m^3+2m^2-m}{2} \right\rfloor = \frac{5m^3+7m^3+2m}{2} + \left\lfloor \frac{m}{3} \right\rfloor$$

which is $\lfloor \frac{m}{3} \rfloor$ units greater than the value given by construction 1.3.6. Again, as m gets bigger, the integrality gap in the $\mathbb{1}$ direction also gets wider for $n = 3m+2$.

1.3.3 An Example: $ex(7,4,3)$

Very quickly, the previous cut is insufficient to obtain the value of $ex(n,4,3)$ by induction. This is already the case for $ex(7,4,3)$, since $ex(7,4,3) = 23$ and the cut generated by the 6-cliques gives $\lfloor 24.5 \rfloor$. One must already do more work to prove that $ex(7,4,3) = 23$. We present a two different ways to do so.

Example 1.3.7 ($ex(7,4,3)$, first method)**.** First, let's look carefully at the Chvátal-Gomory cut generated by adding up all the 6-clique inequalities:

$$ex(7,4,3) \leq \left\lfloor \frac{7 \cdot 14}{4} \right\rfloor = \lfloor 24.5 \rfloor = 24 = \frac{7 \cdot 14 - 2}{4}$$

which implies that there exists two 6-cliques, say Q_1^6 and Q_2^6 such that

$$\sum_{e \in E(Q_1^6)} x_e + \sum_{e \in E(Q_2^6)} x_e \leq 2 \cdot 14 - 2,$$

since adding up the number of edges in each 6-clique should give an integer. Without loss of generality, let $Q_1^6 := [2,3,4,5,6,7]$ and $Q_2^6 = [1,3,4,5,6,7]$, and add the previous inequality to $Q(7,4,3)$. Then this new smaller polytope still contains an optimal solution of $T(7,4,3)$. Then, we produce a Chvátal-Gomory cut by taking 1 of all 4-cliques that contain vertex 1 or 2 but not both, 4 of the 5-cliques that contain both of vertices 1 and 2, and 9 of the inequality we just added. Then each edge is present exactly 24 times, which yields

$$\sum_{e \in E(K_7^3)} x_e \leq \left\lfloor \frac{1 \cdot 20 \cdot 3 + 4 \cdot 10 \cdot 7 + 9 \cdot 1 \cdot 26}{24} \right\rfloor = \lfloor 23.91\bar{6} \rfloor = 23$$

in $Q(7,4,3)$ with the extra inequality, and thus $ex(7,4,3) \leq 23$ globally. Since Turán's construction yield that $ex(7,4,3) \geq 23$, we have that $ex(7,4,3) = 23$.

We now present a second way of calculating $ex(7,4,3)$ based on the idea that by fixing some non-edges, we'll reduce the value of the linear program optimal solution since most variables will be between 0 and 1 exclusively.

Example 1.3.8 ($ex(7,4,3)$, third method)**.** To make reading easier here, we denote a clique on vertices v_1, v_2, v_3, v_4 by $[v_1, v_2, v_3, v_4]$. Without loss of generality, suppose $(1,2,3) \notin E$. We know an edge must also be missing from $[4,5,6,7]$, say $(4,5,6)$ without loss of generality. We also know that an edge must be missing from $[1,2,4,5]$, say $(1,2,4)$ and this is again without loss of generality since any edge in $[1,2,4,5]$ has two vertices in one of the two non-edges already fixed, and one in the other, and those two edges are indistinguishable. Finally, an edge must be missing from $[3,5,6,7]$. Since $(3,5,7)$ is undistinguishable from $(3,6,7)$, we have three possible scenarios: either $(3,5,6) \notin E$, $(3,5,7) \notin E$ or $(5,6,7) \notin E$. At least one of these scenarios must still yield an optimal solution globally for $T(7,4,3)$.

- $(3,5,6) \notin E$**:** Let $E_1 = E(K_7^3) \setminus \{(1,2,3),(1,2,4),(3,5,6),(4,5,6)\}$. Produce a Chvátal-Gomory cut by taking four times the 4-clique inequality $[1,2,5,6]$ and twice the following 4-clique inequalities: $[1,3,4,5]$, $[1,3,4,6]$, $[1,3,5,7]$, $[2,3,4,5]$, $[2,3,4,6]$, $[2,4,6,7]$. Moreover, take one time the 5-clique inequalities for $[1,2,3,5,6]$, $[1,2,5,6,7]$, $[1,3,4,5,7]$, $[2,3,4,6,7]$; two times $[1,2,3,6,7]$, $[1,2,4,5,7]$, $[1,4,5,6,7]$, $[2,3,5,6,7]$, and three times $[1,3,4,6,7]$, $[2,3,4,5,7]$. Finally, give weight six to $x_e = 0$ for the four edges we have fixed, and three to $x_e \leq 1$ for edges $(1,2,7)$, $(2,4,6)$, $(5,6,7)$, two for edge $(1,3,5)$ and one for $(1,4,5)$, $(1,4,6)$, $(2,4,5)$. Then every single edge in K_7^3 is present exactly eight times. So this yields the cut:

$$\sum_{e \in E_1} x_e \leq \left\lfloor \frac{16 \cdot 3 + 18 \cdot 7 + 14 \cdot 1 + 24 \cdot 0}{8} \right\rfloor = \lfloor 23.5 \rfloor = 23$$

which is what we wanted.

- $(3,5,7) \notin E$**:** Let $E_2 = E(K_7^3) \setminus \{(1,2,3),(1,2,4),(3,5,7),(4,5,6)\}$. Produce a Chvátal-Gomory cut by taking once cliques $[1,3,4,5]$, $[1,4,5,7]$, $[2,3,4,5]$, $[2,4,5,7]$, $[1,2,3,5,6]$, $[1,2,5,6,7]$, $[1,3,4,6,7]$, $[2,3,4,6,7]$, and edges $(1,2,7)$, $(1,3,7)$, $(1,4,6)$, $(2,3,7)$, $(2,4,6)$, $(3,5,6)$, $(5,6,7)$, and non-edges $(1,2,3)$, twice each of the other non-edges that we have fixed. Each edge of K_7^3 is present exactly twice. This yields:

$$\sum_{e \in E_2} x_e \leq \left\lfloor \frac{4 \cdot 3 + 4 \cdot 7 + 7 \cdot 1 + 7 \cdot 0}{2} \right\rfloor = \lfloor 23.5 \rfloor = 23,$$

as desired.

- $(5,6,7) \notin E$**:** Let $E_3 = E(K_7^3) \setminus \{(1,2,3),(1,2,4),(4,5,6),(5,6,7)\}$. Produce a Chvátal-Gomory cut by taking once cliques $[1,3,4,5]$, $[1,3,5,7]$, $[1,4,5,7]$, $[2,3,4,6]$, $[2,3,6,7]$, $[2,4,6,7]$, twice the cliques $[1,2,4,5,6]$, $[1,2,5,7]$, $[1,2,6,7]$, $[1,3,5,6]$, $[2,3,5,6]$, and non-negativity inequalities for $(1,2,4)$ and $(4,5,6)$, three times cliques $[1,3,4,6]$, $[1,3,6,7]$, $[1,4,6,7]$, $[2,3,4,5]$, $[2,3,5,7]$, $[2,4,5,7]$, and four times the cliques for $[1,3,4,5,7]$, $[2,3,4,6,7]$, $[1,2,5,6]$, and edges $(1,2,7)$, $(3,5,6)$. Each edge of K_7^3 is present exactly eight times. This leads to the following cut:

$$\sum_{e \in E_3} x_e \leq \left\lfloor \frac{36 \cdot 3 + 5 \cdot 7 + 8 \cdot 1 + 4 \cdot 0}{8} \right\rfloor = \lfloor 23.25 \rfloor = 23,$$

as desired.

In all scenarios, we obtain that the Turán number of K_7^3 minus four non-edges fixed is at most 23. And since we know there exists an optimal Turán solution that misses one of these three possible four-edge sets, the Turán number $ex(7,4,3) \leq$

23. Since we know from Turán's construction that $ex(7,4,3) \geq 23$, we obtain that $ex(7,4,3) = 23$.

These two methods can all be extended for $n > 7$, however, they soon stop to be sufficient to find $ex(n,4,3)$. It makes therefore sense to investigate facets that are present for all n.

1.3.4 General Theorems for the Facet Analysis

Recall Padberg's lifting theorem 1.2.20. It is easy to see that it applies to $T(n,a,r)$ in the same way that it applied to $T(n,a,2)$.

Corollary 1.3.9. *Let G and H be two r-uniform hypergraphs such that $H \subseteq G$ and such that $ex(H_e,a,r) = ex(H,a,r)+1$ for every $e \in E(G)\backslash E(H)$ where $H_e = (V(H),E(H) \cup e)$. If $\sum_{e\in E(H)} x_e \leq ex(H,a,r)$ is a facet of $T(H,a,r)$, then it is also a facet of $T(G,a,r)$.*

Corollary 1.3.10. *Let G and H be two graphs such that $H \subseteq G$. If $\sum_{e\in E(H)} c_e x_e \leq c_0$ is a facet of $T(H,a,r)$ such that for every $e' \in E(G)\backslash E(H)$ there exists x^* such that $\sum_{e\in E(H)} c_e x^*_e = c_0$ for which $x^* \cup e'$ is a-clique-free, then it is also a facet of $T(G,a,r)$.*

Again, this is good news, and bad news at the same time. Most facets we find will still be facets on higher-dimensional examples unlike for the stable set polytope where facets get dominated by others; our work on smaller graphs will thus carry on to larger graphs. However, the fact that all those facets remain and don't get dominated by others when we add more edges to the graph means that higher-dimensional polytopes will have many, many, many facets, and thus that it is very unlikely that we will find a complete description.

Some of the facet classes that we've introduced previously for $T(n,a,2)$ can easily be generalized for $T(n,a,r)$.

1.3.5 Hyperwheel Facets

The wheel graph that induced facets for the Turán graph polytope generalizes very nicely to hypergraphs.

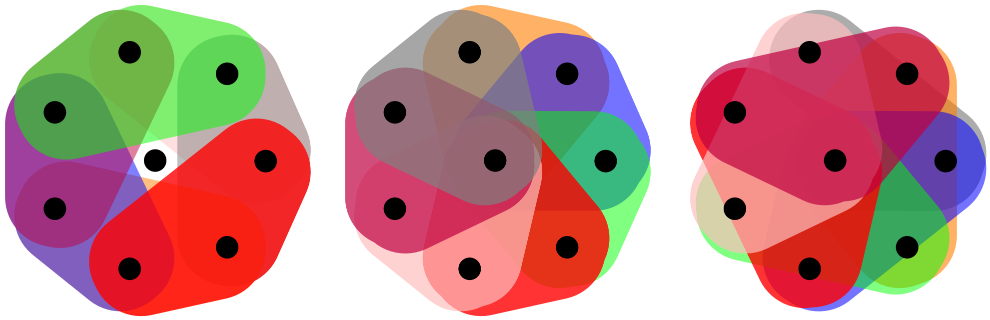

Figure 1.25: A 3-hyperwheel on 8 vertices.

Definition 1.3.11. A hyperwheel ${}^rW_l^a$ is a r-uniform hypergraph on l vertices with one vertex in the center, say l, and vertices $[l-1]$ placed in a cycle in increasing order around it. The r-edges present are such that every $a-1$ consecutive vertices form an a-clique with vertex l. We only consider wheels for which $n \geq 2a-1$, since otherwise we'd have a complete hypergraph.

Example 1.3.12. For example, Figure 1.25 represents ${}^3W_8^4$. As was the case for the graph version, 3-edge inequalities, non-negativity inequalities, and 3-uniform hypercliques of size 4 are not sufficient to describe $T({}^3W_8^4, 4, 3)$.

Indeed, by setting edges $(i, i+1 \mod 7, 8)$ to $\frac{1}{2}$ for $1 \leq i \leq 7$, and all other edges to 1, we obtain a point that satisfies all hyperclique inequalities as well as all other inequalities, and which is a vertex since the 24 tight inequalities are linearly independent. We must thus again devise an inequality for the hyperwheel to get rid of such points.

We first need to make a few key observations about ${}^rW_l^a$. Again, we call edges that contain the middle vertex, say l, *spoke edges* and those who don't, *cycle edges*. A r-edge spanning β vertices of the cycle is contained in $a-1-(\beta-1) = a-\beta$ of the hypercliques. Now label the hyperclique on vertices $1, \ldots, a-1, l$ as hyperclique 1, then the one on vertices $2, \ldots, a, l$ as hyperclique 2, and so on, so that hyperclique i is the hyperclique spanning vertices $i, i+1, \ldots, i+a-2, l \mod (l-1)$. Note that hypercliques containing any edge e are consecutive modulo $a-1$ in that labeling. A spoke edge spans between $r-1$ and $a-1$ vertices of the cycle and a cycle edge, between r and $a-1$ vertices.

To compute the number of edges in ${}^rW_l^a$, fix a vertex i on the cycle and the a-clique starting on the vertex. This clique contains $\binom{a-1}{r-1}$ edges containing vertex

i. Doing so for each vertex will give all edges, and no edge will be counted twice. The hyperwheel ${}^rW_l^a$ contains therefore $(l-1)\cdot\binom{a-1}{r-1}$ r-edges.

Theorem 1.3.13. *The following inequalities are valid and tight for $T(n,a,r)$*

$$\sum_{e\in E({}^rW_l^a)} x_e \leq |E({}^rW_l^a)| - \left\lceil \frac{l-1}{a-r+1} \right\rceil = \binom{a-1}{r-1}\cdot(l-1) - \left\lceil \frac{l-1}{a-r+1} \right\rceil,$$

for all wheels ${}^rW_l^a$ in K_n^r with $l \leq n$.

Proof. We first show that the inequality is valid by determining that it is a Chvátal-Gomory cut. Add up the $l-1$ inequalities of the r-hypercliques of size a with weight $\frac{1}{a-r+1}$. We now want to add edge inequalities with weights placed in a way so that each edge has total weight 1 when we add its edge inequality with the clique inequalities that contain it. Since we know that an edge spanning β vertices is in $a-\beta$ of the hypercliques, we'll give its edge inequality weight $\frac{\beta-r+1}{a-r+1}$.

Note that starting at any vertex of the cycle, there are $\binom{\beta-2}{\beta-r+1}$ spoke edges containing that vertex and spanning the next $\beta-1$ vertices as well for $r-1\leq\beta\leq a-1$; such edges are present $a-\beta$ times in the hypercliques, so we'll give its edge inequality weight $\frac{\beta-r+1}{a-r+1}$.

Observe also that starting at any vertex of the cycle, there are $\binom{\beta-2}{\beta-r}$ cycle edges containing that vertex as a first vertex and spanning the next $\beta-1$ vertices for $r\leq\beta\leq a-1$. Such edges are also present in $a-\beta$ cliques, and we therefore give their edge inequalities weight $\frac{\beta-r+1}{a-r+1}$.

This thus yields the following Chvátal-Gomory cut:

$$\sum_{e\in E({}^rW_l^a)} x_e \leq$$
$$\left\lfloor \frac{(l-1)\left(\binom{a}{r}-1+\sum_{\beta=r-1}^{a-1}\binom{\beta-2}{\beta-r+1}\cdot(\beta-r+1)+\sum_{\beta=r}^{a-1}\binom{\beta-2}{\beta-r}\cdot(\beta-r+1)\right)}{a-r+1} \right\rfloor.$$

We now show that

$$\left\lfloor \frac{(l-1)\left(\binom{a}{r}-1+\sum_{\beta=r-1}^{a-1}\binom{\beta-2}{\beta-r+1}\cdot(\beta-r+1)+\sum_{\beta=r}^{a-1}\binom{\beta-2}{\beta-r}\cdot(\beta-r+1)\right)}{a-r+1} \right\rfloor$$
$$= \binom{a-1}{r-1}\cdot(l-1)-\left\lceil \frac{l-1}{a-r+1} \right\rceil .$$

First, notice that in the first sum of the right-hand side of the Chvátal-Gomory cut, the term produced by $\beta = r-1$ is zero since spoke edges spanning $r-1$ vertices are already present in $a-r+1$ hypercliques. We can thus merge the two sums together, and by recalling that $\binom{\beta-2}{\beta-r+1}+\binom{\beta-2}{\beta-r}=\binom{\beta-1}{\beta-r+1}$, we obtain

$$\left\lfloor \frac{(l-1)\left(\binom{a}{r}-1+\sum_{\beta=r}^{a-1}\binom{\beta-1}{\beta-r+1}\cdot(\beta-r+1)\right)}{a-r+1} \right\rfloor .$$

We now switch the index to be going from 1 to $a-r$ to put it in a more familiar format:

$$\left\lfloor \frac{(l-1)\left(\binom{a}{r}-1+\sum_{\alpha=1}^{a-r}\binom{\alpha+r-2}{\alpha}\cdot\alpha\right)}{a-r+1} \right\rfloor .$$

Then one can easily check that this is equivalent to

$$\left\lfloor \frac{(l-1)\left(\binom{a}{r}-1+(r-1)\cdot\binom{a-1}{a-r-1}\right)}{a-r+1} \right\rfloor .$$

By rewriting the binomial coefficients as factorials, we obtain that

$$\left\lfloor \frac{(l-1)\left(\binom{a-1}{r-1}\cdot(a-r+1)-1\right)}{a-r+1} \right\rfloor ,$$

which is clearly equivalent to

$$\binom{a-1}{r-1} \cdot (l-1) - \left\lceil \frac{l-1}{a-r+1} \right\rceil,$$

as desired.

We now show that this inequality is tight by producing an edge set of size $\binom{a-1}{r-1} \cdot (l-1) - \lceil \frac{l-1}{a-r+1} \rceil$ in ${}^rW_l^a \subset K_n^r$ which contains no hyperclique of size a. The construction is similar to the one of the wheels for graphs: we'll want to take all of the edges except a minimum-size set of spoke edges which will ensure that no clique is full. To do so, we'll remove spoke edges that are contained in as many cliques as possible, namely spoke edges that span only $r-1$ vertices. So first remove such a spoke edge, say $(a-r+1, \ldots, a-1, l)$ without loss of generality, which ensures that the hypercliques starting on vertex 1 through $a-r+1$ are not full. Then remove spoke edge $(2a-2r+2, \ldots, 2a-r, l)$ which ensures that the hypercliques starting on vertices $a-r+2$ through $2a-2r+2$ are not full, and so on. When the full cycle has been explored that way, we only need to make sure we remove an edge before vertex 1. Thus, by removing $\lceil \frac{l-1}{a-r+1} \rceil$ edges, what remains is a-hyperclique-free since at least one spoke is missing in each hyperclique. Given that the hyperwheel contains $(l-1) \cdot \binom{a-1}{r-1}$ in the first place, it means that such a solution contains

$$\binom{a-1}{r-1} \cdot (l-1) - \left\lceil \frac{l-1}{a-r+1} \right\rceil$$

r-hyperedges and so hyperwheel inequalities of type ${}^rW_l^a$ are tight for $T(n,a,r)$ with $n \geq l$ $\square$

Note that in the previous proof, we have seen one type of optimal Turán edge set for the hyperwheel ${}^rW_l^a$ which we call a type I construction. Another type of optimal Turán construction is to remove, without loss of generality, $\lfloor \frac{l-1}{a-r+1} \rfloor$ (r-1)-spanning spoke edges, say edges $(i \cdot (a-r+1), \ldots, i \cdot (a-r+1) + r-2, l)$ for $1 \leq i \leq \lfloor \frac{l-1}{a-r+1} \rfloor$ which guarantees that the cliques starting on vertices 1 through $\lfloor \frac{l-1}{a-r+1} \rfloor$ and spanning the next $a-2$ vertices and central vertex l are all not full. Thus removing any edge contained in all cliques starting on $\lfloor \frac{l-1}{a-r+1} \rfloor$ through $l-1$ will yield an optimal Turán solution which we call of type II.

Theorem 1.3.14. *Inequality*

$$\sum_{e \in E({}^rW_l^a)} x_e \leq \binom{a-1}{r-1} \cdot (l-1) - \left\lceil \frac{l-1}{a-r+1} \right\rceil$$

is facet-defining for $T({}^rW_l^a, a, 2)$ *if* $l-1 = 1 \mod (a-r+1)$.

Proof. Let $\alpha x \leq \beta$ be satisfied by all x in the Turán polytope with $\sum_{e \in E({}^rW_l^a)} x_e = \binom{a-1}{r-1} \cdot (l-1) - \lceil \frac{l-1}{a-r+1} \rceil$; then $\alpha(S) = \beta$ for each Turán edge set S with $|S \cap {}^rW_l^a| = \binom{a-1}{r-1} \cdot (l-1) - \lceil \frac{l-1}{a-r+1} \rceil$.

Consider two distinct spoke edges, including one that spans $r-1$ vertices of the cycle, such that both start from the same vertex, without loss of generality, say $(1,2,\ldots,r-2,l-1,l)$ and $(i_1,\ldots,i_{r-2},l-1,l)$ where $1 \leq i_j \leq a-2$ for all $1 \leq j \leq r-2$. Let solution S_1 be the edge set

$$E({}^rW_l^a) \backslash \left\{ \left\{ (i \cdot (a-r+1), i \cdot (a-r+1)+1, \ldots, i \cdot (a-r+1)+r-2, l) | 1 \leq i \leq \left\lfloor \frac{l-1}{a-r+1} \right\rfloor \right\} \cup (1,2,\ldots,r-2,l-1,l) \right\}$$

and S_2 be

$$E({}^rW_l^a) \backslash \left\{ \left\{ (i \cdot (a-r+1), i \cdot (a-r+1)+1, \ldots, i \cdot (a-r+1)+r-2, l) | 1 \leq i \leq \left\lfloor \frac{l-1}{a-r+1} \right\rfloor \right\} \cup (i_1,\ldots,i_{r-2},l-1,l) \right\},$$

which are clearly two optimal Turán sets respectively of type I and II in ${}^rW_l^a$ since $l-1 = 1 \mod (a-r+1)$ and so the only clique that is still full after removing $\{(i \cdot (a-r+1), i \cdot (a-r+1)+1, \ldots, i \cdot (a-r+1)+r-2, l) | 1 \leq i \leq \lfloor \frac{l-1}{a-r+1} \rfloor\}$ is the clique $(l-1,l,1,2,\ldots,a-2)$ so removing any of those two spoke edges will make the graph a-clique-free. This means that $\alpha(S_1) = \beta = \alpha(S_2)$ which implies that $\alpha_{(1,2,\ldots,r-2,l-1,l)} = \alpha_{(i_1,\ldots,i_{r-2},l-1,l)}$. Since we can show this for any spoke edge starting on the same vertex as a spoke edge that spans $r-1$ vertices, we obtain that $\alpha_{e_1} = \alpha_{e_2}$ for any two spoke edges e_1 and e_2.

Now consider a spoke edge spanning $r-1$ vertices of the cycle and any cycle edge that starts on the same vertex, without loss of generality $(1,2,\dots,r-2,l-1,l)$ and $(i_1,\dots,i_{r-1},l-1)$ where $1 \le i_j \le a-2$ for all $1 \le j \le r-1$. Again, if we let S_3 be the same edge set as S_1 and S_4 be

$$E({}^rW_l^a)\backslash\left\{\left\{(i\cdot(a-r+1), i\cdot(a-r+1)+1,\dots,i\cdot(a-r+1)+r-2,l) | 1 \le i \le \left\lfloor \frac{l-1}{a-r+1}\right\rfloor\right\}\right.$$
$$\left.\cup\,(i_1,\dots,i_{r-1},l)\right\},$$

then they are both optimal Turán solutions respectively of type I and II by the same argument as we've just seen. Thus, we have that $\alpha(S_3) = \beta = \alpha(S_4)$ which implies that $\alpha_{(1,2,\dots,r-2,l-1,l)} = \alpha_{(i_1,\dots,i_{r-1},l-1)}$. Since this is true for any spoke spanning $r-1$ vertices and any cycle edge starting on the same vertex, we have that $\alpha_{e_1} = \alpha_{e_2}$ for any two edges in ${}^rW_l^a$. It is also clear that $\alpha_e > 0$ for al edges $e \in E({}^rW_l^a)$.

We thus conclude that α is a positive scalar multiple of the left-hand side of the hyperwheel inequality, which is thus facet-defining when $l-1 = 1 \mod (a-r+1)$. $\square$

Theorem 1.3.15. *The inequalities*

$$\sum_{e\in E({}^rW_l^a)} x_e \le \binom{a-1}{r-1}\cdot(l-1) - \left\lceil \frac{l-1}{a-r+1}\right\rceil$$

are facet-defining for $T(n,a,r)$ for all ${}^rW_l^a \subseteq K_n^r$ with $l-1 = 1 \mod a-r+1$.

Proof. By the lifting theorem 1.3.9, we simply need to show that there exists a Turán edge set of size $\binom{a-1}{r-1}\cdot(l-1) - \left\lceil\frac{l-1}{a-r+1}\right\rceil + 1$ in $E({}^rW_l^a)\cup e$ for every $e \in E(K_n^r)\backslash E({}^rW_l^a)$. First consider an edge $e \in E\backslash E({}^rW_l^a)$ that is not an edge of the hyperwheel. If $e = (i_1,\dots,i_r)$ such that $i_j \notin V({}^rW_l^a)$ for some $1 \le j \le r$, then it is clear that we can add this edge to any optimal Turán edge set in ${}^rW_l^a$ without creating a clique of size a. Now suppose $e = (i_1,\dots,i_r)$ such that $i_j \in V({}^rW_l^a)$ for all $1 \leq j \leq r$. We know that e spans at least a vertices of the cycle since these are the only edges missing. If e spans more than a vertices, then we can add it to any optimal Turán edge set in ${}^rW_l^a$ without creating an a clique since any a-clique containing e would have to contain also another edge spanning at least a vertices, which we know are absent from such an edge set given that they aren't in the hyperwheel in the first place. So we just have to show that there exists an optimal Turán edge set S in ${}^rW_l^a$ such that $S \cup e$ is still Turán when e spans a vertices. If $r > 2$, then there exist other edges missing in those a vertices, and so they cannot form a clique.

Thus, if $e \in E(K_n^r) \backslash E({}^rW_l^a)$, then there always exists an optimal Turán edge set S in ${}^rW_l^a$ such that $S \cup e$ is also Turán. Therefore,hyperwheel inequalities on l vertices with $l-1 = 1 \mod (a-r+1)$ will still be facet-defining for $T(n,a,r)$, $n \geq l$. Actually, by this argument, these wheel inequalities will be facet-inducing for any $T(G,a,r)$ for any graph G that contains such hyperwheels as subgraphs.

$\square$

Theorem 1.3.16. *We have that*

$$\begin{aligned} T({}^rW_l^{r+1}, r+1, r) = \Big\{ x \in \mathbb{R}^{|E({}^rW_l^{r+1})|} \quad & x(Q^{r+1}) \leq r & \forall Q^{r+1} \in \mathcal{Q}_{K_n^r}^{r+1} \\ & x({}^rW_l^{r+1}) \leq r \cdot (l-1) - \left\lceil \frac{l-1}{2} \right\rceil & \\ & 0 \leq x_e \leq 1 & \forall e \in E({}^rW_l^{r+1}) \Big\} \\ = \{ x \in \mathbb{R}^{|E({}^rW_l^{r+1})|} \quad & Ax \leq b \} & \end{aligned}$$

and that none of these inequalities are redundant if l is even.

Proof. Clearly, $T({}^rW_l^{r+1}, r+1, r) \subseteq \{x \in \mathbb{R}^{|E({}^rW_l^{r+1})|} | Ax \leq b\}$. We now show that $\{x \in \mathbb{R}^{|E({}^rW_l^{r+1})|} | Ax \leq b\} \subseteq T({}^rW_l^{r+1}, r+1, r)$ by showing that the system is locally totally unimodular for almost all vertices, and those who aren't are also integral.

Let x^* be a vertex of $T({}^rW_l^{r+1}, r+1, r)$. We want to show that the determinant of any square submatrix of the inequalities tight at x^* is in $\{-1, 0, 1\}$. We do so by induction on the size of the submatrix.

We know that a square submatrix of size one is simply an entry a_{ij} of the matrix A, and since our matrix contains only entries in $\{-1,0,1\}$, the hypothesis holds. Suppose it also holds for any submatrix of size at most $\lambda-1$. Now look at a submatrix of size λ. Four cases arise.

- The submatrix contains a column made up only of zeroes. Then the determinant of this submatrix is zero, and the hypothesis holds.
- The submatrix contains a non-negativity inequality or an edge inequality, that is a row with only zeroes except for one entry that is either 1 or -1. The determinant is then equal to plus or minus the determinant of a $(\lambda-1)$-submatrix, which has by induction determinant in $\{-1,0,1\}$. Thus the determinant of the λ-submatrix is also in $\{-1,0,1\}$.
- The submatrix contains the hyperwheel inequality as well as clique inequalities. Put the clique inequalities in R_1 and the hyperwheel inequality in R_2. We thus obtain that $\sum_{i\in R_1} a_{ij} - \sum_{i\in R_2} a_{ij} = \{-1,0,1\}$ since $\sum_{i\in R_1} a_{ij} \in \{0,1,2\}$ for every j since any r-hyperedge is in at most two r-cliques and $\sum_{i\in R_2} a_{ij} = 1$ for all j since only the hyperwheel inequality is in R_2. By theorem 1.2.31, this submatrix is totally unimodular.
- The submatrix contains only $(r+1)$-clique inequalities, and at least one of them is not present. Let Q_i be the clique going through vertices $i, i+1, \ldots, i+r-1 \mod (l-1)$ and l. Suppose Q_1 is missing. Then put an inequality Q_i that's present in R_1 if i is even and in R_2 if i is odd. Then $\sum_{i\in R_1} a_{ij} - \sum_{i\in R_2} a_{ij} = \{-1,0,1\}$ for all j since any edge spanning $r+1$ vertices will be in at most one clique present and thus yield either -1,0 or 1, and any edge spanning r vertices will be in at most two consecutive cliques which will not be in the same R_i. By theorem 1.2.31, this submatrix is totally unimodular.

Thus we have that any vertex for which there is at least one clique that is not tight with it is integral. Now consider a vertex x^* for which all cliques are tight. Without loss of generality, we can assume that x^* is suboptimal, otherwise there would exist a set of $|E({}^rW_l^{r+1})|$ tight inequalities with x^* that forms a linearly independent system containing the hyperwheel inequality; we would then be in the third case above. Consider $|E({}^rW_l^{r+1})|$ inequalities tight with x^* that are linearly independent and that contains the $l-1$ $(r+1)$-clique inequalities and sufficiently many edge or non-negativity inequalities. Two cases arise.

- Suppose that all cycle edges and all spoke edges that span r vertices of the cycle have integral values in x^*. If all spoke edges spanning $r-1$ vertices of the cycle also have integral values, then that's fine, x^* is an integral point. However, if at least one such $(r-1)$-spanning spoke edge has a fractional value, then all of them do since no tight $(r+1)$-clique inequality can contain only one fractional edge and each $(r+1)$-clique contains two spoke edges spanning $r-1$ vertices which are placed in a cycle. In that case, we know that the two spoke edges spanning $r-1$ vertices in any $r+1$-clique must sum up to 1 (else they wouldn't be fractional and/or the clique inequality would not be tight) and the $r-1$ other edges in the clique must have value 1. Thus, $\sum_{e \in E({}^{r}W_{l}^{r+1})} x_e^* = \frac{l-1}{2} + (r-1)\cdot(l-1) > r\cdot(l-1) - \lceil \frac{l-1}{2} \rceil$, meaning that x^* does not respect the hyperwheel inequality, a contradiction, and so x^* cannot be a vertex of $T({}^{r}\overline{W}_{n'}^{r}, r+1, r)$.

- If there exists a cycle or spoke edge spanning r vertices of the cycle that has a fractional value in x^*, say edge e that is missing some vertex i for some $2 \leq i \leq r-1$ or $i = l$ in the clique on vertices $1, 2, \ldots, r, l$, i.e. Q_1, then at least one other edge in the clique Q_1 must also have a fractional value in order for this clique inequality to be tight. In order for the system at hand to be linearly independent, we know it must be must possible to associate each edge to a row of the matrix, i.e. to an inequality tight with x^*, in a specific way. Indeed, let A_{x^*} be the incidence matrix of the inequalities that are tight with x^*. Note that the rank of A_{x^*} is $2(l-1)$ since x^* is a vertex. The columns of A_{x^*} are thus linearly independent. Linear independence forms a matroid, and as such, if we look at a subset of s columns of A_{x^*}, they should be linearly independent. Note that this implies that the rank of the matrix formed by these s columns is s, and so at least s of these shortened rows must be nonzero.

  Since e is fractional, the system cannot include its non-negativity or edge inequality, and so the only row in which it is contained is the clique inequality for Q_1 since it is the only clique that contains it. Note that this means that no other spoke or cycle edge spanning r vertices in this clique can be fractional, since it would have to be associated to the same inequality which was already claimed by e, thus making the system linearly dependent if we looked at the submatrix of A_x^* of the two columns associated to those two edges. Since clique Q_1 is tight, we know that every other cycle or spoke edge spanning r vertices in this clique has value 1, and are thus associated to their edge inequality. So, without loss of generality, assume that edge

$(2,3,\ldots,r,l)$ is fractional. This means that edge $(2,3,\ldots,r,l)$ must be associated to the clique on Q_2. Note now that no other cycle or spoke edge spanning r vertices in the clique on Q_2 can be fractional, else such an edge would have to be associated to the clique inequality for Q_2 which is already claimed and so looking at the corresponding submatrix of columns associated to the edges in play, we would have enough nonzero rows to have linear independence; again they must have value 1 and be associated to their edge inequality. Now edge $(3,4,\ldots,r+1,l)$ must also be fractional in order for the clique on Q_2 to be tight with x^*, and it must be associated to clique Q_3, and so on, until we reach the clique Q_{l-2} which will have fractional edges $(l-2,l-1,l,1,\ldots,r-3)$ and $(l-1,l,1,2,\ldots,r-2)$; edge $(l-2,l-1,l,1,\ldots,r-3)$ will be associated to the clique Q_{l-2} and so edge $(l-1,l,1,2,\ldots,r-2)$ must be associated to the clique Q_{l-1}. Note that we hadn't decided at the beginning if edge $(1,2,\ldots,r-1,l)$ is integral or fractional. If it is integral, then some cycle or spoke edge spanning r vertices in the clique Q_1 will have to be fractional in order for the clique inequality to be tight, and so it will have to be associated to this clique as well, but it has already been claimed by e at the very beginning, and so the system cannot be linearly independent. If edge $(1,2,\ldots,r-1,l)$ is fractional, then it must also be associated to the clique Q_1 or Q_{l-1}, and so the system is again linearly dependent. Therefore, if there exists a cycle or spoke edge spanning r vertices with a fractional value, then x^* is not a vertex, a contradiction.

Thus any vertex for which all $(r+1)$-clique inequalities are tight is integral, and so the whole polytope is integral.

Furthermore, we can now show that each of these facets are essential for the description, that none are redundant if l is even.

If one of the $(r+1)$-hyperclique inequalities was missing, then the point with 1 for each r-edge in said hyperclique and 0 everywhere else would be a vertex of the polytope, a contradiction.

If one of the non-negativity inequalities was missing, then the polytope would be a cone with a negative part, a contradiction.

If one of the hyperedge inequalities was missing, say $x_e \leq 1$, then the point with r for e and zero everywhere else would be a vertex of the polytope, a contradiction.

And finally, if the hyperwheel inequality was missing, then the point with 1 everywhere except for edges of type $(i,i+1,\ldots,i+r-2,l) \mod (l-1)$ which we

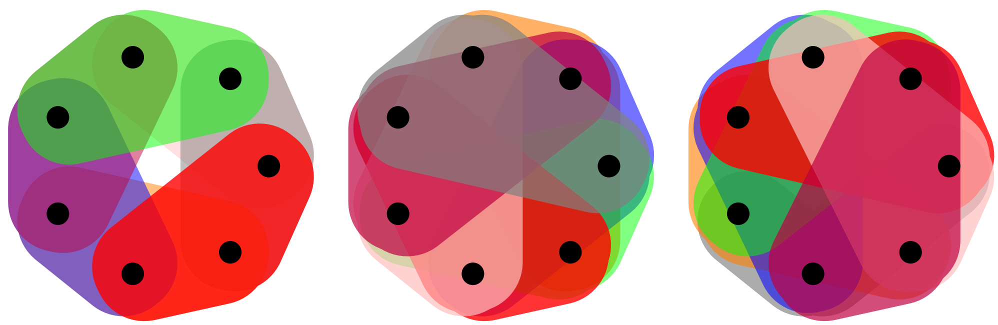

Figure 1.26: A 3-hyperweb ${}^3\overline{W}_7^3$ on 7 vertices.

set to weight $\frac{1}{2}$ would be a vertex of the polytope, since each clique contains two edges with weight $\frac{1}{2}$, and thus each clique would be tight with this point, as well as all edge inequalities for all other edges, which is an independent system if l is even. If l is odd, then they hyperwheel inequality is redundant since the right-hand side of the previously described Chvátal-Gomory cut for the hyperwheel is integral and tight.

$\square$

1.3.6 Hyperweb Facets

The web graph that induced facets for the Turán graph polytope also generalizes very nicely to hypergraphs.

Definition 1.3.17. A *hyperweb* ${}^r\overline{W}_{n'}^{a-1}$ is a r-uniform hypergraph on n' vertices placed in a cycle in increasing order, say 1 through n'. The r-edges present are such that every a consecutive vertices form an a-clique, i.e. for any given vertex, any edge starting with it and spanning at most the next $a-1$ vertices will be present. We only consider hyperwebs for which $n \geq 2a$, so that we do not have a complete hypergraph.

It is basically a hyperwheel with the central vertex removed.

Example 1.3.18. For example, Figure 1.26 represents the edges present in ${}^3\overline{W}_7^3$. As was the case for the graph version, $Q({}^3\overline{W}_7^3, 4, 3)$ is not sufficient to describe $T({}^3\overline{W}_7^3, 4, 3)$.

Indeed, by setting edges $(i, i+1, i+2 \mod 7)$ to $\frac{1}{2}$ for $1 \leq i \leq 7$, and all other edges to 1, we obtain a valid solution tight with all hyperclique inequalities as well as with edge inequalities for edges of type $(i, i+1, i+3 \mod 7)$ and $(i, i+2, i+3 \mod 7)$. Since those 21 tight inequalities are linearly independent, this solution is a vertex of $Q({}^3\overline{W}_7^3)$ and we must thus again devise an inequality for the hyperweb to get rid of it.

We first need to make a few key observations about ${}^r\overline{W}_{n'}^{a-1}$. A r-edge spanning β vertices of the cycle is contained in $a-(\beta-1) = a-\beta+1$ of the hypercliques. Now label the hyperclique on vertices $1, \ldots, a-1, a$ as hyperclique Q_1, then the one on vertices $2, \ldots, a, a+1$ as hyperclique Q_2, and so on, so that hyperclique Q_i is the hyperclique spanning vertices $i, i+1, \ldots, i+a-2, i+a-1 \mod n'$. Note that hypercliques containing any edge e are consecutive in this labeling. Moreover, note that any edge spans between r and a vertices.

To compute the number of edges in ${}^r\overline{W}_{n'}^{a-1}$, fix a vertex i on the cycle and the a-clique starting on the vertex. This clique contains $\binom{a-1}{r-1}$ edges containing vertex i. Doing so for each vertex will give all edges present in the hyperweb, and no edge will be counted twice. The hyperwheel ${}^r\overline{W}_{n'}^{a-1}$ contains therefore $n' \cdot \binom{a-1}{r-1}$ r-edges.

Theorem 1.3.19. *The following inequalities are valid and tight for $T(n,a,r)$*

$$\sum_{e \in E({}^r\overline{W}_{n'}^{a-1})} x_e \leq \binom{a-1}{r-1} \cdot n' - \left\lceil \frac{n'}{a-r+1} \right\rceil,$$

for all hyperwebs ${}^r\overline{W}_{n'}^{a-1}$ in K_n^r with $n' \leq n$.

Proof. We first show that the inequality is valid by determining that it is a Chvátal-Gomory cut. Add up the n' inequalities of the r-hypercliques of size a with weight $\frac{1}{a-r+1}$. We now want to add edge inequalities with weights placed in a way so that each edge has total weight 1 when we add its edge inequality with the clique inequalities that contain it. Since we know that an edge spanning β vertices is in $a-\beta+1$ of the hypercliques, we'll give its edge inequality weight $\frac{\beta-r}{a-r+1}$. Note that starting at any vertex of the cycle, there are $\binom{\beta-2}{\beta-r}$ edges containing that vertex and spanning the next $\beta-1$ vertices as well for $r \leq \beta \leq a$.

This yields thus yields the following Chvátal-Gomory cut:

$$\sum_{e \in E({}^{r}\overline{W}_{n'}^{a-1})} x_e \leq \left\lfloor \frac{n'\left(\binom{a}{r} - 1 + \sum_{\beta=r+1}^{a} \binom{\beta-2}{\beta-r} \cdot (\beta - r)\right)}{a-r+1} \right\rfloor .$$

We now show that

$$\left\lfloor \frac{n'\left(\binom{a}{r} - 1 + \sum_{\beta=r+1}^{a} \binom{\beta-2}{\beta-r} \cdot (\beta - r)\right)}{a-r+1} \right\rfloor = \binom{a-1}{r-1} \cdot n' - \left\lceil \frac{n'}{a-r+1} \right\rceil .$$

First, we switch the sum to be from 1 to $a-r$ to put it in a more familiar format:

$$\left\lfloor \frac{n'\left(\binom{a}{r} - 1 + \sum_{\alpha=1}^{a-r} \binom{\alpha+r-2}{\alpha} \cdot \alpha\right)}{a-r+1} \right\rfloor .$$

Then, again, one can easily check that this is equivalent to

$$\left\lfloor \frac{n'\left(\binom{a}{r} - 1 + (r-1) \cdot \binom{a-1}{a-r-1}\right)}{a-r+1} \right\rfloor .$$

By rewriting the binomial coefficients as factorials, we obtain that

$$\left\lfloor \frac{n'\left(\binom{a-1}{r-1} \cdot (a-r+1) - 1\right)}{a-r+1} \right\rfloor ,$$

which is clearly equivalent to

$$\binom{a-1}{r-1} \cdot n' - \left\lceil \frac{n'}{a-r+1} \right\rceil ,$$

as desired.

We now show that this inequality is tight by producing an edge set of size $\binom{a-1}{r-1} \cdot n' - \lceil \frac{n'}{a-r+1} \rceil$ in ${}^r\overline{W}_{n'}^{a-1} \subset K_n^r$ which contains no hyperclique of size a. The construction is similar to the one of the webs for graphs: we'll want to take all of the edges except a minimum-size set of edges spanning r vertices which will ensure that no clique is full. So first remove such an edge spanning r vertices, say $(a-r+1, \ldots, a-1, a)$ without loss of generality, which ensures that the hypercliques Q_1 through Q_{a-r+1} are not full. Then remove another edge spanning r vertices $(2a-2r+2, \ldots, 2a-r, 2a-r+1)$ which ensures that the hypercliques Q_{a-r+2} through $Q_{2a-2r+2}$ are not full, and so on. When the full cycle has been explored that way, we only need to make sure we remove such an edge before vertex 1, i.e. edge $(1, \ldots, r-1, n')$. Thus, by removing $\left\lceil \frac{n'}{a-r+1} \right\rceil$ edges, what remains is a-clique free since at least one spoke is missing in each clique. Given that the hyperweb contains $n' \cdot \binom{a-1}{r-1}$ in the first place, it means that such a solution contains

$$\binom{a-1}{r-1} \cdot n' - \left\lceil \frac{n'}{a-r+1} \right\rceil,$$

and so hyperweb inequalities of type ${}^r\overline{W}_{n'}^{a-1}$ are tight for $T(n,a,r)$ with $n \geq n'$ $\square$

The type of optimal Turán construction for hyperwebs seen in the previous proof is called type I. Instead of removing $(1, \ldots, r-1, n')$ at the wend, we could have removed any edge contained in all cliques $Q_{\lfloor \frac{n'}{a-r+1} \rfloor \cdot (a-r+1)}$ through $Q_{n'}$. We call this a type II construction.

Theorem 1.3.20. *Inequality*

$$\sum_{e \in E({}^r\overline{W}_{n'}^{a-1})} x_e \leq \binom{a-1}{r-1} \cdot n' - \left\lceil \frac{n'}{a-r+1} \right\rceil$$

is facet-defining for $T({}^r\overline{W}_{n'}^{a-1}, a, 2)$ *if* $n' = 1 \mod (a-r+1)$.

Proof. Let $\alpha x \leq \beta$ be satisfied by all x in the Turán polytope with $\sum_{e \in E({}^r\overline{W}_{n'}^{a-1})} x_e = \binom{a-1}{r-1} \cdot n' - \left\lceil \frac{n'}{a-r+1} \right\rceil$; then $\alpha(S) = \beta$ for each Turán edge set S with $|S \cap {}^r\overline{W}_{n'}^{a-1}| = \binom{a-1}{r-1} \cdot n' - \left\lceil \frac{n'}{a-r+1} \right\rceil$.

Consider two edges that both span only r vertices and that have $r-1$ vertices in common, without loss of generality, say $(1,2,\ldots,r)$ and $(1,2,\ldots,r-1,n')$. Let S_1 be the edge set

$$E({}^{r}\overline{W}_{n'}^{a-1})\backslash\left\{\left\{(i\cdot(a-r+1),i\cdot(a-r+1)+1,\ldots,i\cdot(a-r+1)+r-1)|1\leq i\leq\left\lfloor\frac{n'}{a-r+1}\right\rfloor\right\}\right.$$
$$\left.\cup(1,2,\ldots,r)\right\}$$

and S_2 be

$$E({}^{r}\overline{W}_{n'}^{a-1})\backslash\left\{\left\{(i\cdot(a-r+1),i\cdot(a-r+1)+1,\ldots,i\cdot(a-r+1)+r-1)|1\leq i\leq\left\lfloor\frac{n'}{a-r+1}\right\rfloor\right\}\right.$$
$$\left.\cup(1,2,\ldots,r-1,n')\right\},$$

which are clearly two optimal type I Turán sets in ${}^{r}W_{n'}^{a-1}$. This means that $\alpha(S_1)=\beta=\alpha(S_2)$ which implies that $\alpha_{(1,2,\ldots,r)}=\alpha_{(1,2,\ldots,r-1,n')}$. Since we can show this for any consecutive edges spanning r vertices and since those edges are placed in a cycle, we obtain that $\alpha_{e_1}=\alpha_{e_2}$ for any two edges e_1 and e_2 spanning r vertices.

Now consider an edge spanning r vertices and an edge spanning more than r vertices such that they start on the same vertex, without loss of generality $(1,2,\ldots,r-1,n')$ and $(i_1,\ldots,i_{r-1},n')$ where $1\leq i_j\leq a-1$ for all $1\leq j\leq r-1$ and all i_j's are distinct. Again, if we let S_3 be the edge set

$$E({}^{r}\overline{W}_{n'}^{a-1})\backslash\left\{\left\{(i\cdot(a-r+1),i\cdot(a-r+1)+1,\ldots,i\cdot(a-r+1)+r-1)|1\leq i\leq\left\lfloor\frac{n'}{a-r+1}\right\rfloor\right\}\right.$$
$$\left.\cup(1,2,\ldots,r-1,n')\right\}$$

and S_4 be

$$E({}^r\overline{W}_{n'}^{a-1})\backslash\left\{\left\{(i\cdot(a-r+1), i\cdot(a-r+1)+1,\dots,i\cdot(a-r+1)+r-1)|1\leq i\leq\left\lfloor\frac{n'}{a-r+1}\right\rfloor\right\}\right.$$
$$\left.\cup(i_1,\dots,i_{r-1},n')\right\},$$

then they are both optimal Turán solutions of type I and II since $n' = 1 \mod (a-r+1)$ and so there is only one clique that is still full after removing $\{(i\cdot(a-r+1), i\cdot(a-r+1)+1,\dots,i\cdot(a-r+1)+r-1)|1\leq i\leq\lfloor\frac{n'}{a-r+1}\rfloor\}$, namely $Q_{n'}$, and so removing any edge from that clique will make the graph a-clique-free; $(i_1,\dots,i_{r-1},n')$ is such an edge. Thus, we have that $\alpha(S_3)=\beta=\alpha(S_4)$ which implies that $\alpha_{(1,2,\dots,r-1,n')}=\alpha_{(i_1,\dots,i_{r-1},n')}$. Since any edge in the hyperweb starts on the same vertex as some edge spanning r vertices, we have that $\alpha_{e_1}=\alpha_{e_2}$ for any two edges in ${}^r\overline{W}_{n'}^{a-1}$. It is also clear that $\alpha_e>0$ for all edges $e\in E({}^r\overline{W}_{n'}^{a-1})$.

We thus conclude that α is a positive scalar multiple of the left-hand side of the web inequality, which is thus facet-defining when $n' = 1 \mod (a-1)$. □

Theorem 1.3.21. *The inequalities*

$$\sum_{e\in E({}^r\overline{W}_{n'}^{a-1})} x_e \leq \binom{a-1}{r-1}\cdot n' - \left\lceil\frac{l-1}{a-r+1}\right\rceil$$

are facet-defining for $T(n,a,r)$ for all ${}^r\overline{W}_{n'}^{a-1}\subseteq K_n^r$ with $n' = 1 \mod (a-r+1)$ and $n'\leq n$.

Proof. By the lifting theorem 1.3.9, we simply need to show that there exists a Turán edge set of size $\binom{a-1}{r-1}\cdot n' - \left\lceil\frac{n'}{a-r+1}\right\rceil+1$ in $E({}^r\overline{W}_{n'}^{a-1})\cup e$ for every $e\in E(K_n^r)\backslash E({}^r\overline{W}_{n'}^{a-1})$.

First consider an edge $e\in E(K_n^r)\backslash E({}^r\overline{W}_{n'}^{a-1})$ that is not an edge of the hyperweb. If $e=(i_1,i_2,\dots,i_r)$ such that $i_j\notin V({}^r\overline{W}_{n'}^{a-1})$ for some $1\leq j\leq r$, then it is clear that we can add this edge to any optimal Turán edge set in ${}^r\overline{W}_{n'}^{a-1}$ without creating a clique of size a. Now suppose $e=(i_1,i_2,\dots,i_r)$ such that $i_j\in V(\overline{W}_{n'}^{a-1})$ for all $1\leq j\leq r$. Then we know that e must span at least $a+1$ vertices since all r-hyperedges spanning at most a vertices are already present in the hyperweb. If it spans at least $a+2$ vertices, then we can add it to any optimal Turán edge set in ${}^r\overline{W}_{n'}^{a-1}$ without creating an a-clique since any such clique containing e would have to contain also another edge spanning at least $a+1$ vertices, which we know is absent from such an optimal solution given that edges spanning at least $a+1$ vertices aren't in the web. If e spans exactly $a+1$ vertices, say $1,\ldots,a+1$, then we can use the optimal construction for webs we've seen before so that it contains all of the edges, except $\{(1+i\cdot(a-r+1), 2+i\cdot(a-r+1))|0\leq i\leq\lfloor\frac{n'}{a-r+1}\rfloor\}$, and then adding e to this solution leaves it Turán.

Thus, if $e\in E(K_n^r)\backslash E({}^r\overline{W}_{n'}^{a-1})$, then there always exists an optimal Turán edge set S in ${}^r\overline{W}_{n'}^{a-1}$ such that $S\cup e$ is also Turán. $\square$

Theorem 1.3.22. *We have that*

$$\begin{aligned}
T({}^r\overline{W}_{n'}^r, r+1, r) = \{x\in\mathbb{R}^{|E({}^r\overline{W}_{n'}^r)|} \;|\; & x(Q^{r+1})\leq r && \forall Q^{r+1}\in\mathcal{Q}_{K_n^r}^{r+1}\\
& x({}^r\overline{W}_{n'}^r)\leq r\cdot n' - \left\lceil\frac{n'}{2}\right\rceil &&\\
& 0\leq x_e\leq 1 && \forall e\in E({}^r\overline{W}_{n'}^r)\}\\
= \{x\in\mathbb{R}^{|E({}^r\overline{W}_{n'}^r)|} \;|\; & Ax\leq b\},
\end{aligned}$$

and that none of these inequalities are redundant if n' is odd.

Proof. Clearly, $T({}^r\overline{W}_{n'}^r, r+1, r)\subseteq\{x\in\mathbb{R}^{|E({}^r\overline{W}_{n'}^r)|}|Ax\leq b\}$. We now show that $\{x\in\mathbb{R}^{|E({}^r\overline{W}_{n'}^r)|}|Ax\leq b\}\subseteq T({}^r\overline{W}_{n'}^r, r+1, r)$ by showing that the system is locally totally unimodular for most vertices and that other vertices are also integral.

Let x^* be a vertex of $T({}^r\overline{W}_{n'}^r, r+1, r)$. We want to show that the determinant of any square submatrix of the inequalities tight at x^* is in $\{-1,0,1\}$. We do so by induction on the size of the submatrix.

We know that a square submatrix of size one is simply an entry a_{ij} of the matrix A, and since our matrix contains only entries in $\{-1,0,1\}$, the hypothesis holds. Suppose it holds also for any submatrix of size at most $\lambda-1$. Now look at a submatrix of size λ. Four cases arise.

- The submatrix contains a column made up only of zeroes. Then the determinant of this submatrix is zero, and the hypothesis holds.

- The submatrix contains a non-negativity inequality or an edge inequality, that is a row with only zeroes except for one entry that is either 1 or -1. The determinant is then equal to plus or minus the determinant of a $(\lambda - 1)$-submatrix, which has by induction determinant in $\{-1,0,1\}$. Thus the determinant of the λ-submatrix is also in $\{-1,0,1\}$.
- The submatrix contains the hyperweb inequality as well as clique inequalities. Put the clique inequalities in R_1 and the hyperweb inequality in R_2. We thus obtain that $\sum_{i\in R_1} a_{ij} - \sum_{i\in R_2} a_{ij} = \{-1,0,1\}$ since $\sum_{i\in R_1} a_{ij} \in \{0,1,2\}$ for every j since any r-hyperedge is in at most two r-cliques and $\sum_{i\in R_2} a_{ij} = 1$ for all j since only the web inequality is in R_2. By theorem 1.2.31, this submatrix is totally unimodular.
- The submatrix contains only $(r+1)$-clique inequalities, and at least one of them is not present. Recall that Q_i be the clique going through vertices $i, i+1, \ldots, i+r \mod n'$. Suppose Q_1 is missing. Then put an inequality Q_i that's present in R_1 if i is even and in R_2 if i is odd. Then $\sum_{i\in R_1} a_{ij} - \sum_{i\in R_2} a_{ij} = \{-1,0,1\}$ for all j since any edge spanning $r+1$ vertices will be in at most one clique present and thus yield either -1,0 or 1, and any edge spanning r vertices will be in at most two consecutive cliques which will not be in the same R_i. By theorem 1.2.31, this submatrix is totally unimodular.

Thus we have that any vertex for which there is at least one clique that is not tight with it is integral. Now consider a vertex tight with all cliques. Without loss of generality, we can imagine that x^* is suboptimal then, otherwise, there would exist a set of $|E({}^r\overline{W}^r_{n'})|$ tight inequalities with x^* that form a linearly independent system containing the web inequality; we would then be in the third case above. Consider rn' inequalities tight with x^* that are linearly independent and that contains the n' $(r+1)$-clique inequalities and $(r-1)\cdot n'$ edge or non-negativity inequalities. Two cases arise.

- Suppose that all edges spanning $r+1$ vertices have integral values in x^*. If all edges spanning r vertices also have integral values, then that's fine, x^* is an integral point. However, if at least one edge spanning r vertices has a fractional value, then all of them do since no tight $r+1$-clique inequality can contain only one fractional edge and each clique contains two edges spanning r vertices which are placed in a cycle. In that case, we know that the two edges spanning r vertices in any $r+1$-clique must sum up to 1 (else they wouldn't be fractional and/or the clique inequality would not be tight) and the $r-1$ edges spanning $r+1$ vertices in it must have value 1. Thus, $\sum_{e \in E({}^r\overline{W}^r_{n'})} x^*_e = \frac{n'}{2} + (r-1) \cdot n' > r \cdot n' - \left\lceil \frac{n'}{2} \right\rceil$, meaning that x^* does not respect the hyperweb inequality, a contradiction, and so x^* cannot be a vertex of $T({}^r\overline{W}^r_{n'}, r+1, r)$.

- If there exists an edge spanning $r+1$ vertices that has a fractional value in x^*, say the edge e in the clique Q_1 that doesn't contain vertex i for some $2 \leq i \leq r$, then at least one other edge in the clique Q_1 must also have a fractional value in order for this clique inequality to be tight. In order for the system at hand to be linearly independent, we know it must be must possible to associate each edge to a row of the matrix, i.e. to an inequality tight with x^*, in a specific way. Indeed, let A_{x^*} be the incidence matrix of the inequalities that are tight with x^*. Note that the rank of A_{x^*} is $2(l-1)$ since x^* is a vertex. The columns of A_{x^*} are thus linearly independent. Linear independence forms a matroid, and as such, if we look at a subset of s columns of A_{x^*}, they should be linearly independent. Note that this implies that the rank of the matrix formed by these s columns is s, and so at least s of these shortened rows must be nonzero.

  Since e is fractional, the system cannot include its non-negativity or edge inequality, and so the only row in which it is contained is the clique inequality for Q_1. Note that this means that no other edge spanning $r+1$ vertices in this clique can be fractional, since it would have to be associated to the same inequality which was already claimed by e, thus making the system linearly dependent since the submatrix of the two columns associated to these two edges would have only one nonzero row. Actually, for Q_1 to be tight with x^*, they will all have to have value 1 in x^* and so will be associated with their edge inequality. So, without loss of generality, assume that edge $(2,3,\ldots,r+1)$ is fractional. This means that edge $(2,3,\ldots,r+1)$ must be associated to the clique Q_2. Note again that no edge spanning $r+1$ vertices in the clique Q_2 can be fractional, else such an edge would have to be associated to the clique inequality on $2,3,\ldots,r+2$ which is already claimed, and so the number of nonzero rows in the submatrix of A_{x^*} for the columns of the edges in play would not be sufficient to create linear indepedence. Now edge $(3,4,\ldots,r+2)$ must also be fractional in order for the clique Q_2 to be tight with x^*, and it must be associated to clique Q_3, and so on, until we reach the clique $Q_{n'-1}$ which will have fractional edges $(n'-1,n',1,\ldots,r-2)$ and $(n',1,2,\ldots,r-1)$; edge $(n'-1,n',1,\ldots,r-2)$

will be associated to the clique $Q_{n'-1}$ and so edge $(n', 1, 2, \ldots, r-1)$ must be associated to the clique $Q_{n'}$. Note that we hadn't decided at the beginning if edge $(1, 2, \ldots, r)$ is integral or fractional. If it is integral, then some edge spanning $r+1$ vertices in the clique$Q_{n'}$ will have to be fractional in order for the clique inequality to be tight since edge $(n', 1, 2, \ldots, r-1)$ is the only fractional edge it contains right now, and so it will have to be associated to this clique as well, but it has already been claimed, and so the system cannot be linearly independent. If edge $(1, 2, \ldots, r)$ is fractional, then it must also be associated to the clique $Q_{n'}$ or Q_1, both of which are already claimed, and so the system is again linearly dependent. Therefore, if there exists an edge spanning $r+1$ vertices with a fractional value, then x^* is not a vertex, a contradiction.

Thus any vertex for which all $(r+1)$-clique inequalities are tight is integral, and so the whole polytope is integral.

Furthermore, we can now show that each of these facets are essential for the description, that none are redundant if n' is odd.

If one of the $(r+1)$-hyperclique inequalities was missing, then the point with 1 for each r-edge in said hyperclique and 0 everywhere else would be a vertex of the polytope, a contradiction.

If one of the non-negativity inequalities was missing, then the polytope would be a cone with a negative part, a contradiction.

If one of the hyperedge inequalities was missing, say $x_e \leq 1$, then the point with r for e and zero everywhere else would be a vertex of the polytope, a contradiction.

And finally, if the wheel inequality was missing, then the point with 1 everywhere except for edges of type $(i, i+1, \ldots, i+r-2, i+a-2) \mod n'$ which we set to weight $\frac{1}{2}$ would be a vertex of the polytope, since each clique contains two edges with weight $\frac{1}{2}$, and thus each clique would be tight with this point, as well as all edge inequalities for all other edges, which is an independent system if n' is odd. If n' is even, then the hyperweb inequality is redundant. □

1.3.7 An Example: Rank Facets of $T(6,4,3)$

We now present some rank facets of $T(6,4,3)$. To do so, we used Porta [19]. Unfortunately, $T(6,4,3)$ was too big to be calculated, however, it was possible to calculate $T(G,4,3)$ for $G=(V,E)$ with $|V|=6$ and $E \subset E(K_6^3)$. We first present the rank facets we found, display how to get them as Chvátal cut, and show that they can be lifted. We call a rank facet $\sum_{e\in E} x_e \leq C$ associated to some graph $G=(V,E)$ a C-$\sum_{e\in E} 1$ rank facet.

The $11-14$ **rank facets of type I**
Let $G=(V,E)$ with $V=[6]$. Let E be generated by two 5-cliques, without loss of generality, say the cliques on vertices $1,2,3,4,5$ and $2,3,4,5,6$, each missing an edge, and such that these two missing edges do not intersect, without loss of generality, say that edge $(1,2,3)$ and $(4,5,6)$ are missing. Then $\sum_{e\in E} x_e \leq 11$ is a facet of $T(G,4,3)$ according to Porta. Here is how to obtain this inequality as a Chvátal-Gomory cut.

1 2 3	1 2 4	1 2 5	1 2 6	1 3 4	1 3 5	1 3 6	1 4 5	1 4 6	1 5 6	2 3 4	2 3 5	2 3 6	2 4 5	2 4 6	2 5 6	3 4 5	3 4 6	3 5 6	4 5 6	
0	1	1	0	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	0	≤ 11
0	1	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	≤ 3
0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	≤ 3
0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	1	0	0	0	≤ 3
0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	0	0	≤ 3
0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	1	0	≤ 3
0	1	1	0	1	1	0	0	0	0	0	0	0	0	1	1	0	1	1	0	≤ 8

By taking one half of all these inequalities, we obtain the rank facet.

Theorem 1.3.23. *The inequality* $\sum_{e\in E} x_e \leq 11$ *is facet-defining for* $T(n,4,3)$ *for all* $n \geq 6$.

Proof. By the lifting theorem 1.3.9, we only need to show that there exists for every $e \in E(K_n^r)\backslash E(G)$ an optimal Turán set S in G such that $S \cup e$ is also Turán. Clearly, if e contains vertices not in $[6]$, any optimal Turán set will do. If $e \in E(K_6^3)$, then, without loss of generality, it is either edge $(1,2,3)$ or $(1,2,6)$ in the example above. It is easy to find out if $ex(G'=([6],E(G)\cup e),4,3) = ex(G,4,3)+$

1 with integer programming solver such as SCIP [92]. For the sake of completion, we list here these solutions. If we add $(1,2,3)$ to G, then the following edge set is a solution of size 12, as desired:

$$\{(1,2,3),(1,2,4),(1,2,5),(1,3,4),(1,3,5),(2,3,6),(2,4,5),(2,4,6),(2,5,6),\\(3,4,5),(3,4,6),(3,5,6)\}.$$

Similarly, if we add $(1,2,6)$ instead, we can also find a solution of size 12:

$$\{(1,2,4),(1,2,5),(1,2,6),(1,3,4),(1,3,5),(2,3,4),(2,3,5),(2,4,5),(2,4,6),\\(2,5,6),(3,4,6),(3,5,6)\}.$$

Thus $\sum_{e\in E(G)} x_e \leq 11$ will be facet-defining for $T(n,4,3)$ for any $n \geq 6$. $\square$

Note, that there are $\binom{6}{2}\cdot 6\cdot 1 = 90$ subgraphs isomorphic to G in K_6^3, so $T(6,4,3)$ will have 90 inequalities of this type.

The $11-14$ rank facets of type II
Let $G=(V,E)$ with $V=[6]$. Let E be generated by two 5-cliques, say without loss of generality the cliques on vertices $1,2,3,4,5$ and $2,3,4,5,6$, each missing an edge, and such that those two missing edges intersect in one vertex, say without loss of generality $(1,2,3)$ and $(3,5,6)$. Then $\sum_{e\in E} x_e \leq 11$ is a facet of $T(G,4,3)$ according to Porta. Here is how to obtain this inequality as a Chvátal-Gomory cut.

1 2 3	1 2 4	1 2 5	1 2 6	1 3 4	1 3 5	1 3 6	1 4 5	1 4 6	1 5 6	2 3 4	2 3 5	2 3 6	2 4 5	2 4 6	2 5 6	3 4 5	3 4 6	3 5 6	4 5 6	
0	1	1	0	1	1	0	1	0	0	1	1	1	1	1	1	1	1	0	1	≤ 11
1	1	1	0	1	1	0	1	0	0	1	1	0	1	0	0	1	0	0	0	≤ 7
0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	≤ 3
0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	0	0	≤ 3
0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	≤ 3
0	1	1	0	0	0	0	0	0	0	0	1	1	0	0	1	0	1	0	1	≤ 7
-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	≤ 0

By taking one half of all these inequalities, we obtain the desired rank facet.

Theorem 1.3.24. *The inequality $\sum_{e \in E} x_e \leq 11$ is facet-defining for $T(n,4,3)$ for all $n \geq 6$.*

Proof. By the lifting theorem 1.3.9, we only need to show that there exists for every $e \in E(K_n^r) \backslash E(G)$ an optimal Turán set S in G such that $S \cup e$ is also Turán. Clearly, if e contains vertices not in $[6]$, any optimal Turán set will do. If $e \in E(K_6^3)$, then, without loss of generality, it is either edge $(1,2,3)$, $(1,2,6)$, $(1,3,6)$ or $(1,4,6)$ in the example above. It is easy to find out if $ex(G' = ([6], E(G) \cup e), 4, 3) = ex(G,4,3) + 1$ with SCIP. For the sake of completion, we list here these solutions. If we add $(1,2,3)$ to G, then the following edge set is a solution of size 12, as desired:

$$\{(1,2,3),(1,2,4),(1,2,5),(1,3,4),(1,4,5),(2,3,5),(2,3,6),(2,4,6),(2,5,6),\\ (3,4,5),(3,4,6),(4,5,6)\}.$$

Similarly, if we add $(1,2,6)$ instead, we can also find a solution of size 12:

$$\{(1,2,4),(1,2,5),(1,2,6),(1,3,4),(1,3,5),(2,3,5),(2,3,6),(2,4,5),(2,5,6),\\ (3,4,5),(3,4,6),(4,5,6)\}.$$

Again, if we add $(1,3,6)$ instead, we can find a solution of size 12:

$$\{(1,2,4),(1,2,5),(1,3,4),(1,3,5),(1,3,6),(2,3,5),(2,3,6),(2,4,5),(2,5,6),\\ (3,4,5),(3,4,6),(4,5,6)\}.$$

Finally, if we add $(1,4,6)$, there again exists a solution of size 12 that contains it in $G' = ([6], E(G) \cup (1,4,6))$:

$$\{(1,2,4),(1,2,5),(1,3,4),(1,3,5),(1,4,6),(2,3,5),(2,3,6),(2,4,5),(2,5,6),\\ (3,4,5),(3,4,6),(4,5,6)\}$$

Thus $\sum_{e \in E(G)} x_e \leq 11$ will be facet-defining for $T(n,4,3)$ for any $n \geq 6$. $\square$

Note, that there are $\binom{6}{2} \cdot \binom{4}{2} \cdot 2 \cdot 2 = 360$ subgraphs isomorphic to G in K_6^3, so $T(6,4,3)$ will have 360 inequalities of this type.

The $11-14$ rank facets of type III
Let $G=(V,E)$ with $V=[6]$. Let E be generated by two 5-cliques, say without loss of generality the cliques on vertices $1,2,3,4,5$ and $2,3,4,5,6$, with one edge in their intersection missing, without loss of generality say $(2,3,4)$, and an edge in only one of the two cliques intersecting the other missing edge in two vertices, say $(1,2,3)$. Then $\sum_{e \in E} x_e \leq 11$ is a facet of $T(G,4,3)$ according to Porta. Here is how to obtain this inequality as a Chvátal-Gomory cut.

1 2 3	1 2 4	1 2 5	1 2 6	1 3 4	1 3 5	1 3 6	1 4 5	1 4 6	1 5 6	2 3 4	2 3 5	2 3 6	2 4 5	2 4 6	2 5 6	3 4 5	3 4 6	3 5 6	4 5 6	
0	1	1	0	1	1	0	1	0	0	0	1	1	1	1	1	1	1	1	1	≤ 11
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	≤ 7
0	1	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	≤ 3
0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	≤ 3
0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	1	0	≤ 3
0	1	1	0	1	1	0	0	0	0	0	0	0	0	1	0	0	1	0	1	≤ 7
0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	≤ 0

By taking one half of all these inequalities and rounding down, we obtain the desired rank facet.

Theorem 1.3.25. *The inequality $\sum_{e \in E} x_e \leq 11$ is facet-defining for $T(n,4,3)$ for all $n \geq 6$.*

Proof. By the lifting theorem 1.3.9, we only need to show that there exists for every $e \in E(K_n^r) \backslash E(G)$ an optimal Turán set S in G such that $S \cup e$ is also Turán. Clearly, if e contains vertices not in $[6]$, any optimal Turán set will do. If $e \in E(K_6^3)$, then, without loss of generality, it is either edge $(1,2,3)$, $(1,2,6)$, $(1,4,6)$, $(1,5,6)$ or $2,3,4$ in the example above. It is easy to find out if $ex(G' = ([6], E(G) \cup e), 4, 3) = ex(G,4,3)+1$ with SCIP. For the sake of completion, we list here these solutions. If we add $(1,2,3)$ to G, then the following edge set is a solution of size 12, as desired:

$$\{(1,2,3),(1,2,4),(1,2,5),(1,3,4),(1,3,5),(1,4,5),(2,3,6),(2,4,6),(2,5,6),\\ (3,4,6),(3,5,6),(4,5,6)\}.$$

Similarly, if we add $(1,2,6)$ instead, we can also find a solution of size 12:

$$\{(1,2,4),(1,2,5),(1,2,6),(1,3,4),(1,3,5),(1,4,5),(2,3,6),(2,4,6),(2,5,6),\\ (3,4,6),(3,5,6),(4,5,6)\}.$$

Again, if we add $(1,4,6)$, there again exists a solution of size 12 that contains it in $G' = ([6], E(G) \cup (1,4,6))$:

$$\{(1,2,4),(1,2,5),(1,3,4),(1,3,5),(1,4,5),(1,4,6),(2,3,6),(2,4,6),(2,5,6),\\ (3,4,6),(3,5,6),(4,5,6)\}$$

One more time, if we add $(1,5,6)$ instead, we can find a solution of size 12:

$$\{(1,2,4),(1,2,5),(1,3,4),(1,3,5),(1,4,5),(1,5,6),(2,3,6),(2,4,6),(2,5,6),\\ (3,4,6),(3,5,6),(4,5,6)\}.$$

Finally, if we add $(2,3,4)$ instead, we again find a solution of size 12:

$$\{(1,2,4),(1,2,5),(1,3,4),(1,3,5),(1,4,5),(2,3,4),(2,3,5),(2,4,6),(2,5,6),\\ (3,4,6),(3,5,6),(4,5,6)\}.$$

Thus $\sum_{e \in E(G)} x_e \leq 11$ will be facet-defining for $T(n,4,3)$ for any $n \geq 6$. $\square$

Note, that there are $\binom{6}{2} \cdot \binom{4}{3} \cdot \binom{3}{2} \cdot 2 = 360$ subgraphs isomorphic to G in K_6^3, so $T(6,4,3)$ will have 360 inequalities of this type.

The $12-15$ **rank facets**

The $12-15$ rank facets are the ${}^3W_6^4$-hyperwheel inequelaties, which we've already discussed in 1.3.13. There are $6\cdot\frac{5!}{5\cdot 2}=72$ subgraphs isomorphic to ${}^3W_6^4$ in K_6^3, so $T(6,4,3)$ will have 72 inequalities of this type.

The $13-17$ **rank facets**

Let $G=(V,E)$ with $V=[6]$. Let E consist of all of the edges of K_6^3 except for a chain of three edges, say (a,b,c), (b,c,d), (c,d,e). For example, take all of the edges except $(1,2,3)$, $(2,3,4)$ and $(3,4,5)$. Then $\sum_{e\in E}x_e\leq 13$ is a facet of $T(G,4,3)$ according to Porta. Here is how to obtain this inequality as a Chvátal-Gomory cut.

1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	4	
2	2	2	2	3	3	3	4	4	5	3	3	3	4	4	5	4	4	5	5	
3	4	5	6	4	5	6	5	6	6	4	5	6	5	6	6	5	6	6	6	
0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	≤ 13
0	1	1	1	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	1	≤ 7
0	1	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	≤ 3
0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	≤ 3
0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	1	0	≤ 3
0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	1	0	≤ 3
0	0	0	1	1	1	0	0	0	0	0	1	1	0	1	0	0	1	0	1	≤ 8

By taking one half of all these inequalities and rounding down, we obtain the desired facet.

Theorem 1.3.26. *The inequality* $\sum_{e\in E}x_e\leq 13$ *is facet-defining for* $T(n,4,3)$ *for all* $n\geq 6$.

Proof. By the lifting theorem 1.3.9, we only need to show that there exists for every $e\in E(K_n^r)\backslash E(G)$ an optimal Turán set S in G such that $S\cup e$ is also Turán. Clearly, if e contains vertices not in $[6]$, any optimal Turán set will do. If $e\in E(K_6^3)$, then, without loss of generality, it is either edge $(1,2,3)$ or $(2,3,4)$ in the example above. It is easy to find out if $ex(G'=([6],E(G)\cup e),4,3)=ex(G,4,3)+1$ with SCIP. For the sake of completion, we list here these solutions. If we add $(1,2,3)$ to G, then the following edge set is a solution of size 14, as desired:

$$\{(1,2,3),(1,2,4),(1,2,6),(1,3,4),(1,3,5),(1,4,5),(1,5,6),(2,3,5),(2,3,6),\\ (2,4,5),(2,4,6),(3,4,6),(3,5,6),(4,5,6)\}.$$

Finally, if we add $(2,3,4)$ instead, we again find a solution of size 14:

$$\{(1,2,4),(1,2,5),(1,2,6),(1,3,4),(1,3,5),(1,3,6),(1,4,5),(2,3,4),(2,3,5),\\ (2,4,6),(2,5,6),(3,4,6),(3,5,6),(4,5,6)\}.$$

Thus $\sum_{e\in E(G)} x_e \leq 13$ will be facet-defining for $T(n,4,3)$ for any $n \geq 6$. $\square$

Note, that there are $6 \cdot 5!/2 = 360$ subgraphs isomorphic to G in K_6^3, so $T(6,4,3)$ will have 360 inequalities of this type.

The $13-18$ **rank facets**

Let $G = (V,E)$ with $V = [6]$. Let E consist of all of the edges of K_6^3 except for two disjoint edges, say $(1,2,3)$ and $(4,5,6)$. Then $\sum_{e\in E} x_e \leq 13$ is a facet of $T(G,4,3)$ according to Porta. Here is how to obtain this inequality as a Chvátal-Gomory cut.

1 2 3	1 2 4	1 2 5	1 2 6	1 3 4	1 3 5	1 3 6	1 4 5	1 4 6	1 5 6	2 3 4	2 3 5	2 3 6	2 4 5	2 4 6	2 5 6	3 4 5	3 4 6	3 5 6	4 5 6	
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	≤ 13
0	1	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	≤ 3
0	1	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	≤ 3
0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	≤ 3
0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	≤ 3
0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	1	0	0	≤ 3
0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	1	0	≤ 3
0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	1	0	0	0	≤ 3
0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	1	0	0	≤ 3
0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	1	0	≤ 3

By taking one half of all these inequalities and rounding down, we obtain the desired facet.

Theorem 1.3.27. *The inequality $\sum_{e \in E} x_e \leq 13$ is facet-defining for $T(n,4,3)$ for all $n \geq 6$.*

Proof. By the lifting theorem 1.3.9, we only need to show that there exists for every $e \in E(K_n^r) \backslash E(G)$ an optimal Turán set S in G such that $S \cup e$ is also Turán. Clearly, if e contains vertices not in $[6]$, any optimal Turán set will do. If $e \in E(K_6^3)$, then, without loss of generality, it is edge $(1,2,3)$ in the example above. It is easy to find out if $ex(G' = ([6], E(G) \cup e), 4, 3) = ex(G,4,3) + 1$ with SCIP. For the sake of completion, we list here these solutions. If we add $(1,2,3)$ to G, then the following edge set is a solution of size 14, as desired:

$$\{(1,2,3),(1,2,4),(1,2,6),(1,3,4),(1,3,6),(1,4,5),(1,5,6),(2,3,5),(2,4,5),\\ (2,4,6),(2,5,6),(3,4,5),(3,4,6),(3,5,6)\}.$$

Thus $\sum_{e \in E(G)} x_e \leq 13$ will be facet-defining for $T(n,4,3)$ for any $n \geq 6$. □

Note, that there are $6 \cdot 5!/(5 \cdot 2) = 72$ subgraphs isomorphic to G in K_6^3, so $T(6,4,3)$ will have 72 inequalities of this type.

Thus $T(6,4,3)$ already has at least 1314 rank facets other than the cliques. This gives a good idea of the size of the problem.

1.4 Conclusion

Obviously, there is still a lot of work to do on the Turán hypergraph problem and even just on its polytope. Hopefully, it is clear by now that the Turán polytope is interesting in itself, notwithstanding its connection to the famous problem. It is surprising that this polytope hasn't previously been researched since its structure is so rich and interesting, and even more so because many parallels can be drawn between its facets and those of the stable set polytope, one of the polytopes that has been studied the most. Indeed, we showed here that cliques, webs and wheels all play an important role, like for the stable set polytope. In the future, we'd love to see the rank facets of the stable set polytope for quasi-line graphs (see [29]) be transferred to the Turán polytope. This polytope is just so big and contains so many different type of combinatorial facets that it would be a shame not to study it more thoroughly.

Understanding some of the facet structure of the Turán polytope also allowed us to understand better why the Turán problem is so hard in general. Indeed, given that the facets we found do not get dominated as the number of vertices grows, this leads us to believe that the number of facets of $T(n,a,r)$ really explodes as n grows. For example, we proved that cliques of size i were facet-defining for $T(n,3,2)$ for all i odd and $n \geq i$. This means that $T(n,3,2)$ already has at least

$$\binom{n}{3} + \binom{n}{5} + \binom{n}{7} + \ldots + \binom{n}{n'} \approx 2^{n-1}$$

facets, where $n' = n$ if n is odd or $n' = n-1$ if n is even. And these are just the clique facets! Optimizing over a polytope with many facets in a random direction is generally hard, and so this might explain why the problem remains open in general.

Still, we saw that optimizing in the $\mathbb{1}$-direction over $T(n,a,2)$ was quite easy. Thus, even though the facets we found were not sufficient to yield the $(4,3)$-conjecture, it might well be that there exists a succession of cuts that would always yield the solution. This is even more likely because the Turán numbers are so well-defined in the conjecture. Moreover, when $n = 0 \mod 3$, we've already shown that this is the case. With more work focusing on facets of the Turán polytope, it might be possible to achieve similar results for $n = 1,2 \mod 3$. A deeper study of the Turán polytope is therefore a logical course of action.

Furthermore, there are many more Turán-type problems that would benefit from having their corresponding polytopes studied. We believe that studying the structure of the polytopes will go a long way in understanding what makes those problems so difficult to solve. Small instances that can be completely described with the software Porta already contain some non-trivial facets that are useful to describe the polytopes of larger instances and to understand the problem better in general.

This rich world of Turán-type polytopes is just waiting to be discovered.

Chapter 2

The Frankl Problem

The work in this chapter was done in collaboration with Jonad Pulaj and Dirk Theis.

2.1 Preliminaries for Frankl's Conjecture

2.1.1 Problem Statement

The Frankl conjecture, also known as the union-closed sets conjecture, was formulated in 1979 by, surprisingly enough, Péter Frankl. Before stating the conjecture, we need the following the definition.

Definition 2.1.1. A family of sets is said to be *union-closed* if and only if the union of two members of the family is also a member of the family.

The conjecture has been stated in many different ways over the years, but the simplest form is the following. It was first stated in Frankl's 1983 article ([38]) where he dates it back to 1979 (though some claim it was known even before).

Conjecture 2.1.2 (Frankl, 1979)**.** *In a union-closed family $\mathcal{F}$ such that $\mathcal{F} \neq \emptyset$, there exists an element that is in at least half of the sets of $\mathcal{F}$.*

Note that, throughout this section, we assume that any family has only distinct sets. We now restate the conjecture, but as two optimization problems.

Conjecture 2.1.3 (Maximization Version)**.** *Let $\mathcal{F} \neq \emptyset$ be a union-closed family such that each element is in at most a sets of $\mathcal{F}$. Then $|\mathcal{F}| \leq 2a$.*

For the minimization version, we also need a very natural definition.

Definition 2.1.4. For a family of sets, we call an element that is contained in the greatest number of sets a *most common element*.

Conjecture 2.1.5 (Minimization Version)**.** *Let $\mathcal{F} \neq \emptyset$ be a union-closed family composed of m sets. Then minimizing the number of sets that contain the most common element yields at least $\frac{m}{2}$.*

Remark. The three conjectures are equivalent since there exists a counterexample to Frankl's conjecture if and only if there exists a union-closed family $\mathcal{F} \neq \emptyset$ with b sets such that each element is contained in less than $b/2$ sets of $\mathcal{F}$.

2.1.2 Some Notation

We let $\mathcal{F}$ always represent a union-closed family that is not the empty set. We let m represent the number of sets in $\mathcal{F}$ and we'll usually let them run through $\{1,\ldots,m\} =: [m]$. We let n be the number of elements in $\mathcal{F}$ and we denote the set of elements by $E = \{1,\ldots,n\} =: [n]$. We let a be the maximum number of sets in $\mathcal{F}$ that can contain an element.

2.1.3 Previous Work

Though still unsolved in general, the conjecture has progressed a lot since 1979. It has attracted the attention of both lattice theorists as well as combinatorial probabilists; more recently, computer scientists have also joined the race to solve this conjecture. As far as we know, this is the first time that the problem is investigated through combinatorial optimization.

Two easy results were noticed early on. Their authorship is unclear.

Theorem 2.1.6 (Folklore)**.** *The Frankl conjecture holds true for any family containing a singleton, i.e. a set of one element.*

Theorem 2.1.7 (Folklore)**.** *The Frankl conjecture holds true for any family that has an average set size of at least* $\frac{n}{2}$.

The following result was proven many times by different people in different ways, but the first to do so were Sarvate and Renaud.

Theorem 2.1.8 (Sarvate and Renaud, 1989)**.** *The Frankl conjecture holds true for any family that contains a doubleton, i.e. a set of two elements.*

Another early result is the following:

Theorem 2.1.9 (Roberts, 1992)**.** *The Frankl conjecture holds true if* $m < 4n - 1$.

Those results and a few more that followed allowed to prove the conjecture for increasing values of m and n over time. The current status is as follows:

Theorem 2.1.10 (Roberts and Simpson, 2010)**.** *The Frankl conjecture is true if* $m \leq 46$.

Theorem 2.1.11 (Bošnjak and Marković, 2008)**.** *The Frankl conjecture is true if* $n \leq 11$.

These results follow others that became incrementally better over the years ($m \leq 11$ in [89], $m \leq 18$ in [90], $m \leq 24$ in [36], $m \leq 27$ in [78], $m \leq 32$ in [49], $m \leq 40$ in [87], $n \leq 7$ in [78], $n \leq 9$ in [36]).

More recently, Vučković and Živković announced the following result which is still unpublished.

Theorem 2.1.12 (Vučković and Živković, 2012)**.** *The Frankl conjecture is true if* $n \leq 12$ *and* $m \leq 50$.

Thus the conjecture is still open for $m \geq 51$ and $n \geq 13$.

Many more results have been obtained through the years, however we will introduce them later on as we need them.

2.2 Three models for the Frankl Conjecture

The work in this section is the result of a collaboration with Jonad Pulaj.

We model the Frankl conjecture in three different ways. First with integer programs and then with a quadratic program. We let $\mathcal{S}_n$ be the power set of n elements. We first propose the following model.

$$F(n,m) = \{x | x_T + x_U \leq 1 + x_S \qquad \forall S = T \cup U \in \mathcal{S}_n \tag{2.1}$$

$$\sum_{S \in \mathcal{S}_n} x_S = m \tag{2.2}$$

$$\sum_{\substack{S \ni e: \\ S \in \mathcal{S}_n}} x_S \leq \frac{m-1}{2} \qquad \forall e \in E \tag{2.3}$$

$$x_S \in \{0,1\} \qquad \forall S \in \mathcal{S}_n\} \tag{2.4}$$

Theorem 2.2.1. *The polytope $F(n,m)$ is non-empty if and only if there exists a counterexample to Frankl's conjecture with exactly m sets spanning n elements.*

Proof. In this formulation, $\mathcal{S}_n$ contains the 2^n possible subsets of $[n]$. We let x_S for $S \in \mathcal{S}_n$ be 1 if S is a part of the union-closed family in the solution and 0 otherwise. Constraint 2.1 ensures that the family in a solution is union-closed: if both T and U are in a family, then their union $S = T \cup U$ must be there as well. The equation 2.2 says that the families we consider contain m sets. Finally, the third inequality 2.3 states that we are looking for a family that is a counterexample to Frankl's conjecture: each element should be contained in at most $\frac{m-1}{2}$ sets of the family.

Thus, if there exists a solution in $F(n,m)$, it is a counterexample to the Frankl conjecture since it is a union-closed family for which each element is in less than half the sets present. $\square$

The previous model can be stated very simply, however it is quite large. The number of variables 2^n is "small" enough, however, there are $\binom{2^n}{2} + n + 1$ constraints, a number which grows very quickly with n. We thus propose the following extended formulation which doesn't grow too quickly in n, but rather in m. In the previous model, all sets of $\mathcal{S}_n$ were present and we chose a union-closed subset of these sets with certain properties. In the next two models, we construct directly a subset of $\mathcal{S}_n$ starting from nothing that holds the same desired properties.

$$
\begin{aligned}
Fr(n,m) = \{x,y,z| \sum_{k\in[m]} y_k^{i,j} &= 1 && \forall i<j\in[m] && (2.5)\\
x_e^k &\le x_e^i + x_e^j + 1 - y_k^{i,j} && \forall e\in[n], k, i<j\in[m] && (2.6)\\
x_e^k &\ge \frac{x_e^i + x_e^j}{2} - 1 + y_k^{i,j} && \forall e\in[n], k, i<j\in[m] && (2.7)\\
\sum_{k\in[m]} x_e^k &\le \frac{m-1}{2} && \forall e\in[n] && (2.8)\\
z_e^{i,j} &\le \frac{x_e^i - x_e^j + 1}{2} && \forall e\in[n], i\neq j\in[m] && (2.9)\\
z_e^{i,j} &\ge x_e^i - x_e^j && \forall e\in[n], i\neq j\in[m] && (2.10)\\
\sum_{e\in[n]} (z_e^{i,j} + z_e^{j,i}) &\ge 1 && \forall i<j\in[m] && (2.11)\\
x_e^k, y_k^{i,j}, z_e^{i,j} &\in \{0,1\} && \forall e\in[n], i,j,k\in[m]\}
\end{aligned}
$$

Theorem 2.2.2. *The polytope $Fr(n,m)$ is non-empty if and only if there exists a counterexample to Frankl's conjecture with m sets spanning n elements.*

Proof. Let $\mathcal{F} = \{S_1, \ldots, S_m\}$ be our family of m sets and $E = \{e_1, \ldots, e_n\}$ be our n elements. We let

$$x_e^i = \begin{cases} 1 & \text{if } e\in S_i \text{, and} \\ 0 & \text{otherwise,} \end{cases}$$

to decide whether element e is in set i, and

$$y_k^{i,j} = \begin{cases} 1 & \text{if } S_k = S_i \cup S_j \text{, and} \\ 0 & \text{otherwise,} \end{cases}$$

to determine if set k is the union of sets i and j. We observe that constraints 2.9 and 2.10 thus imply that

$$z_e^{i,j} = \begin{cases} 1 & \text{if } e\in S_i \backslash S_j \text{, and} \\ 0 & \text{otherwise,} \end{cases}$$

that is, whether e is in set i but not in j. We assume that $S_i \neq S_j$ for all $i \neq j$. Constraint 2.11 ensures that this assumption holds by forcing the symmetric difference $\sum_{e\in[n]}(z_e^{i,j} + z_e^{j,i})$ between any two sets S_i and S_j with $i \neq j$ to be at least one, thus implying that S_i and S_j cannot be equal.

Furthermore, for any $i \neq j$, there exists exactly one $k \in [m]$ such that $S_k = S_i \cup S_j$ by constraints 2.5, 2.6 and 2.7. Constraint 2.5 enforces that there is exactly one set that is the union of S_i and S_j. Constraints 2.6 and 2.7 say that if $y_k^{i,j} = 1$, that is if $S_k = S_i \cup S_j$, then $x_e^k = 1$ if $e \in S_i \cup S_j$ and $x_e^k = 0$ otherwise. Note that constraints 2.6 and 2.7 are trivial if $y_k^{i,j} = 0$, i.e. they have no impact on variables x_e^k.

This shows that the program does represent a union-closed family without repetion. Constraint 2.8 states that we are looking for a counterexample: no element can be in half of the sets in the family or more. Therefore, there exists a solution in this polytope if and only if there exists a counterexample to Frankl's conjecture with m sets spanning n elements. $\square$

Finally, we propose a quadratic program as well.

$$
\begin{aligned}
FQ(n,m) = \{x,y| \sum_{k\in[m]} y_k^{ij} &= 1 && \forall i < j \in [m] && (2.12)\\
y_k^{ij} \cdot \frac{x_e^i + x_e^j}{2} &\leq y_k^{ij} x_e^k && \forall e \in [n], k, i < j \in [m] && (2.13)\\
y_k^{ij} x_e^k &\leq y_k^{ij} \cdot (x_e^i + x_e^j) && \forall e \in [n], k, i < j \in [m] && (2.14)\\
\sum_{e\in E} (x_e^i - x_e^j)^2 &\geq 1 && \forall i < j \in [m] && (2.15)\\
\sum_{i\in[m]} x_e^i &\leq \frac{m-1}{2} && \forall e \in [n] && (2.16)\\
x_e^i, y_k^{ij} &\in \{0,1\} && \forall e \in [n] \forall i,j,k \in [m]\} && (2.17)
\end{aligned}
$$

Theorem 2.2.3. *The Frankl conjecture is true if and only if $FQ(n,m)$ is empty.*

Proof. Let x_e^i and y_k^{ij} be defined as in the second model. The first equation states as before that for any two sets, there is exactly one set that is their union. The inequalities 2.13 and 2.14 ensure that if $S_k = S_i \cup S_j$, i.e. if $y_k^{ij} = 1$, then this is reflected in the elements in the sets i, j and k. Inequality 2.15 enforces that no two sets contain exactly the same elements. Finally, inequality 2.16 once again states that we are looking for a counterexample. Thus, any solution in $FQ(n,m)$ must be a union-closed family that is a counterexample to Frankl's conjecture. $\square$

We wanted to list all models we found for completion, but we'll mostly focus on the two integer programs.

2.3 Analysis of the First Model

We first want to say a few words about the integrality gap of the first model. Of course, talking about an integrality gap when we are hoping that the model is always unfeasible doesn't quite make sense. However, it is easy to change the model into an optimization problem, either a maximization or a minimization, as can be seen in definition 2.3.1 for $f_n(a)$ and $g_n(m)$.

Definition 2.3.1. We let

$$\begin{aligned} f_n(a) := \max \sum_{S \in \mathcal{S}_n} x_S & & \\ \text{s.t. } x_U + x_T \leq 1 + x_S & & \forall T \cup U = S \in \mathcal{S}_n \\ \sum_{S \in \mathcal{S}_n : e \in S} x_S \leq a & & \forall e \in [n] \\ x_S \in \{0,1\} & & \forall S \in \mathcal{S}_n \end{aligned}$$

and we let

$$\begin{aligned} g_n(m) := \min \max_{e \in [n]} \sum_{S \in \mathcal{S}_n : e \in S} x_S & & \\ \text{s.t. } x_U + x_T \leq 1 + x_S & & \forall T \cup U = S \in \mathcal{S}_n \\ \sum_{S \in \mathcal{S}_n} x_S = m & & \\ x_S \in \{0,1\} & & \forall S \in \mathcal{S}_n. \end{aligned}$$

We can modify these problems a bit so that they are even more closely related by introducing a constraint that states that element i appears in at least as many sets as element j for $1 \leq i < j \leq n$.

$$
\begin{aligned}
f_n^*(a) := \max \sum_{S\in\mathcal{S}_n} x_S & & \\
\text{s.t. } x_U + x_T \le 1 + x_S & & \forall T\cup U = S \in \mathcal{S}_n \\
\sum_{S\in\mathcal{S}_n : i\in S} x_S \ge \sum_{S\in\mathcal{S}_n : j\in S} x_S & & \forall 1\le i\le j\le n \\
\sum_{S\in\mathcal{S}_n : 1\in S} x_S \le a & & \\
x_S \in \{0,1\} & & \forall S\in\mathcal{S}_n
\end{aligned}
$$

and we let

$$
\begin{aligned}
g_n^*(m) := \min \sum_{S\in\mathcal{S}_n : 1\in S} x_S & & \\
\text{s.t. } x_U + x_T \le 1 + x_S & & \forall T\cup U = S \in \mathcal{S}_n \\
\sum_{S\in\mathcal{S}_n : i\in S} x_S \ge \sum_{S\in\mathcal{S}_n : j\in S} x_S & & \forall 1\le i\le j\le n \\
\sum_{S\in\mathcal{S}_n} x_S = m & & \\
x_S \in \{0,1\} & & \forall S\in\mathcal{S}_n.
\end{aligned}
$$

Clearly $f_n^*(a) = f_n(a)$ and $g_n^*(m) = g_n(m)$. The star notation is interesting though because the underlying polytopes are exactly the same except that the extra constraint of one is the objective function of the other and vice-versa.

Proposition 2.3.2. *The three following statements are equivalent:*

1. *The Frankl conjecture holds,*
2. $f_n(a) \le 2a$*, and*
3. $g_n(m) \ge \frac{m}{2}$.

Proof. $1 \Rightarrow 2$**:** If the Frankl conjecture holds, and $f_{n'}(a') = m' \ge 2a'+1$ for some n', a', then there exists a union-closed family of m' sets such that each element is present in at most $a' < \frac{m'}{2}$ sets, a contradiction. $2 \Rightarrow 3$**:** If $f_n(a) \le 2a$ for all n and

a, and $g_{n'}(m') = a' < \frac{m'}{2}$ for some m', n', then there exists a family of m' sets where every element is in at most a' sets of the family, and so $f'_n(a') \geq m' \geq 2a' + 1$, a contradiction. $3 \Rightarrow 1$**:** If $g_n(m) \geq \frac{m}{2}$ for all n and m, then for any union-closed family with m sets, there exists an element in at least $\frac{m}{2}$ of these sets, and thus the Frankl conjecture holds. □

We call the polytope behind the program $f_n(a)$, $F_n(a)$, and the one behind $g_n(m)$, $G_n(m)$.

Now that we have optimization problems, talking about the integrality gap makes sense. Our computations reveal that the inequalities ensuring that the family is union-closed do little to close the integrality gap. Indeed, the optimal value of the linear relaxation of F is not too far from the relaxation without the union-closed inequalities. Given that number of constraints is what makes integer program solvers run out of memory when n grows, it might be useful in the future to consider programs without the union-closed constraints and find a few constraints they imply that help close the gap.

Second of all, we want to talk about the fact that the integrality gap between the integral value of $f_n(a)$ and its relaxed value increases as n increases, but decreases in a. We want to point out that when n is as small as possible for the problem to be non-trivial, i.e. when $2^{n-1} > a > 2^{n-2}$, then the integrality gap is not too bad when a is big enough. In fact, if we add that $x_S = 0$ for all $S \in \mathcal{S}_n$ such that $|S| = 1$ or $|S| = 2$ (otherwise we know the family would be Frankl by theorem 2.1.6), then the linear relaxation already tells us that union-closed families are Frankl for many values, see table 2.6 in the Appendix.

With all of the different considerations below, we were able to compute that the Frankl conjecture holds for any union-closed family with $n \leq 10$. Given that this does not beat the current best of $n \leq 12$, we do not give the tedious details of which cuts were used for which values of a and n. Up to $n = 8$, computing $f_n(a)$ (see table 2.4) was fairly straightforward and just a matter of waiting (for quite a while in the case of $n = 8$). For $n = 9$, we had to use many cuts and concepts found in this section and the next ones to compute $f_9(a)$ to optimality, as well as lots of patience. For $n = 10$, we couldn't calculate $f_{10}(a)$ to optimality for all a's, however it is not necessary to get optimality to get that the conjecture holds, only that $f_{10}(a) \leq 2a$, which was possible to achieve. Computing $g_n(m)$ took longer, so we calculated fewer values (see table 2.5).

We now look at some useful cuts for the first model.

Example 2.3.3. We first use Porta [19] to investigate the polytope $F_3(3)$. We order the sets of $\mathcal{S}_3$ lexicographically:

S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
1	1	1	1	0	0	0	0
1	1	0	0	1	1	0	0
1	0	1	0	1	0	1	0

For example, set 5 is the set composed of elements 2 and 3, i.e. $S_5 = (0,1,1)$. We set $x_1 = 1$ since we are looking for a minimal counterexample. Finally, we choose to allow each element to be in at most three sets, i.e. $a = 3$. The complete linear description of the integral hull of this polytope includes the nonnegativity inequalities as well as the inequalities $x_S \leq 1$ for each set $S \in [8]$, as well as the original union inequalities and element containment inequalities. Moreover, the following inequalities are necessary:

$$
\begin{array}{rrrrrrrr}
-x_2 & -x_3 & +x_4 & -x_5 & +x_6 & +x_7 & \leq 1 & (2.18) \\
 & & x_4 & +x_5 & & +x_7 & \leq 2 & (2.19) \\
 & & x_4 & +x_5 & +x_6 & & \leq 2 & (2.20) \\
 & x_3 & & & +x_6 & +x_7 & \leq 2 & (2.21) \\
 & x_3 & +x_4 & & +x_6 & & \leq 2 & (2.22) \\
x_2 & & & & +x_6 & +x_7 & \leq 2 & (2.23) \\
x_2 & & +x_4 & & & +x_7 & \leq 2 & (2.24) \\
 & & 2x_4 & +x_5 & +x_6 & +x_7 & \leq 3 & (2.25) \\
 & x_3 & +x_4 & & +2x_6 & +x_7 & \leq 3 & (2.26) \\
x_2 & & +x_4 & & +x_6 & +2x_7 & \leq 3 & (2.27) \\
x_2 & +x_3 & +x_4 & +x_5 & +x_6 & +x_7 & \leq 3 & (2.28)
\end{array}
$$

These new inequalities can be explained and generalized as follows.

Lemma 2.3.4. *The following inequalities are valid for $F(n,m)$ and thus also for $F_n(a)$ and $F_n(m)$ for all $\mathcal{S}' \subseteq \mathcal{S}_n$:*

$$\sum_{S \in \mathcal{S}'} x_S \leq 1 + \sum_{\substack{T := S_i \cup S_j: \\ S_i, S_j \in \mathcal{S}', i<j}} c_T x_T,$$

where c_T is the number of distinct pairwise unions in $\mathcal{S}'$ that yield set T. We call those inequalities clique inequalities*. Facet 2.18 in example 2.3.3 is a clique inequality.*

Proof. It suffices to notice that the union inequalities in the first model are related to the basic inequalities of the stable set polytope. Indeed, the union equality for $S = T \cup V$

$$x_T + x_V \leq 1 + x_S$$

is trivially redundant if $x_S = 1$ (since $x_T \leq 1$ and $x_V \leq 1$, so $x_T + x_V \leq 2$) and is a stable set inequality if $x_S = 0$. Indeed, we could imagine the problem to be a kind of stable set problem in a dynamic graph where each vertex is one of the 2^n sets and where there is an edge between two sets T, V if and only if $S := T \cup V$ is not chosen.

With this analogy in mind, it is easy to see that the inequality above is valid. If the right handside of the inequality is equal to 1, this means that there are edges between all the sets in $\mathcal{S}'$, i.e. they form a clique, and thus, indeed, at most one of them can be chosen.

If the left handside is equal to 1, then since $x_S \geq 0$ for all S, the right handside is at least 1, and so the inequality holds.

If the left handside is equal to $l \geq 2$, then we know that the $\binom{l}{2}$ unions of these sets (with repetition) must be present, and so the inequality holds since $l \leq 1 + \binom{l}{2}$. □

We now look at a few general theorems about inequalities. Note that whenever we mention cuts for $F_n(a)$ from now on, they are usually also valid for $F(n,m)$ and $G_n(m)$ by making a few simple changes in the notation. We omit to write this every single time to make the reading lighter.

Lemma 2.3.5. *Cuts valid for $F_{n-1}(a-1)$ are also valid for sets $2^{n-1}+1$ to 2^n for $F_n(a)$. In other words, if $\sum_{i \in I} c_i x_i \leq c_0$ where $I \subset \{1, \ldots, 2^{n-1}\}$ is a valid inequality for $F_{n-1}(a-1)$, then $\sum_{i \in I} c_i x_{i+2^{n-1}} \leq c_0$ is a valid inequality for $F_n(a)$.*

Proof. We know that, without loss of generality, the set $\mathbb{1}$ containing all of the n elements is part of the optimal solution of $F_n(a)$. Thus, every element is in at most $a-1$ sets from $2^{n-1}+1$ to 2^n, i.e. sets that do not contain element 1. The same is true for the set of sets not containing some element e since there are only $n-1$ elements in play within those sets. Therefore, any cut for $F_{n-1}(a-1)$ is valid for those sets. □

We need the natural definion of the closure of a family in the union-closed context.

Definition 2.3.6. Let $\mathcal{S}'$ be a family of sets on n elements. Its *closure* is the smallest union-closed family on the same n elements that contains $\mathcal{S}'$ as a subfamily.

Lemma 2.3.7. *The inequality*

$$\sum_{S\in\mathcal{S}'} x_S \leq |\mathcal{S}'|-1$$

is valid for $F_n(a)$ for every $\mathcal{S}' \subseteq \mathcal{S}_n$ such that the closure of $\mathcal{S}'$ contains an element that is there more than a times (or $a-1$ times if $\mathbb{1}$ is not in the closure). We call such an inequality a closure inequality. *Moreover, inequalities 2.19-2.24 are closure inequalities.*

Proof. If a family contains all of the sets in $\mathcal{S}'$, then the closure of these sets must also be in the family since it is union-closed, and thus there would be an element present more than a times, and so this family would not be a solution contained in $F_n(a)$. □

A useful tactic is to assume that there exists a counterexample, and thus that there exists a counterexample of minimal size. Inequalities valid for minimal counterexample might cut off some valid union-closed families including some that might be bigger counterexamples, but that's fine since if we find that $f_n(a) \leq 2a$ with these cuts, then we reach a contradiction and know that there is no counterexample for n and a in the first place.

Lemma 2.3.8. *The inequality*

$$(a-1)x_{S^*} + \sum_{S:S\cap S^*=\emptyset, S\neq\emptyset} x_S \leq 2a-2$$

is valid for a minimal counterexample of $F_n(a)$ for every $\emptyset \neq S^* \in \mathcal{S}_n$.

Proof. If S^* is present in a solution, then at most $a-1$ sets in $\{S : |S \cap S^*| = 0\}$ can also be present since the union of each of those sets with S^* will yield a distinct set that is in the closure and that contains all of the elements in S^*, which we know can be there at most a times. The left hand side is then at most $a-1+a-1 = 2a-2$ and the inequality holds.

If S^* is not present, since $|\{S : S \cap S^* = \emptyset, S \neq \emptyset\}| \leq f_{n-|S|}(a-1)$ because $\mathbb{1}$ is without loss of generality in the solution (since the counterexample is minimal) and so each elements can be in at most $a-1$ sets of $\{S : S \cap S^* = \emptyset, S \neq \emptyset\}$. Moreover, since the optimal solution of $F_n(a)$ is a minimal counterexample, $F_{n-|S|}(a-1)$ contains no counterexamples and every solution in it, including the optimal one, is thus Frankl, meaning that $|\{S : S \cap S^* = \emptyset, S \neq \emptyset\}| \leq 2(a-1)$. Thus the inequality holds in this case too. □

Lemma 2.3.9. *The inequality*

$$\sum_{S:S^* \subset S} x_S + (2a-2)(1 - x_{S^*}) \geq \sum_{S:S \cap S^* = \emptyset} x_S$$

is valid for a minimal counterexample in $F_n(a)$ for every $S^* \in \mathcal{S}_n$.

Proof. If S^* is present in a solution, then, as was stated in the previous proof, its union with each of the sets in $\{S : S \cap S^* = \emptyset\}$ will be distinct and of course present in the solution, and so $\sum_{S:S^* \subseteq S} x_S \geq \sum_{S:S \cap S^* = \emptyset} x_S$ and since $1 - x_{S^*} = 0$, the inequality holds.

If S^* is not present, then as was stated before, $|S : S \cap S^* = \emptyset| \leq 2(a-1)$, and the left hand side is greater or equal to $2a-2$. Thus the inequality holds again. □

It is useful to look at the dual of the relaxation of our model for $f_n(a)$ especially after fixing some variables. It is widely known that any set that contains a singleton or a doubleton is Frankl. We can thus fix the variables of such sets to be zero if we're looking for a counterexample. Moreover, if we're looking for a smallest counterexample, we can force the full set to be there, since if it isn't, then there exists another counterexample with a smaller n. We thus consider the following relaxation of our program.

$$
\begin{aligned}
\max \sum_{S \in \mathcal{S}_n} & x_S \\
\text{s.t. } & x_T + x_U \leq 1 + x_S && \forall S = T \cup U \\
& \sum_{S \ni e} x_S \leq a && \forall e \in E \\
& x_S = 0 && \forall S \in \mathcal{S}_n \text{ with } 1 \leq |S| \leq 2 \\
& x_1 = 1 \\
& x_S \geq 0 && \forall S \in \mathcal{S}_n
\end{aligned}
$$

We now consider the dual of this model:

$$
\begin{aligned}
\min\ & w_1 + a \cdot \sum_e z_e + \sum_{S = T \cup U} y_{STU} \\
\text{s.t. } & \sum_{e : S' \ni e} z_e - \sum_{\substack{S' = T \cup U \\ T \neq S' \\ U \neq S'}} y_{S'TU} + \sum_{U = S' \cup T} y_{US'T} \geq 1 && \forall S' \in \mathcal{S}_n, |S'| \neq 1,2 \text{ or } n \\
& w_{S'} + \sum_{e : S' \ni e} z_e - \sum_{\substack{S' = T \cup U \\ T \neq S' \\ U \neq S'}} y_{S'TU} + \sum_{U = S' \cup T} y_{US'T} \geq 1 && \forall S' \in \mathcal{S}_n, |S'| = 1,2 \text{ or } n \\
& w_S, z_e, y_{STU} \geq 0 && \forall e, S = T \cup U
\end{aligned}
$$

Theorem 2.3.10. *If there exists a counterexample to Frankl's conjecture for some n, then $m \leq 2^{n-1} - \frac{n^2}{2} + \frac{n}{2} + a + 1$.*

Proof. One can easily check that setting $z_e = \frac{1}{n}$ for all e, $y_{11S} = 1 - \frac{|S|}{n}$ for all S such that $|S| \neq 1,2$ or n, and $w_1 = 1$ as well as $w_S = 1$ for all S such that $|S| = 1,2$ or n consists in a valid solution for the dual. We observe that its objective value is

$$
1 + a + \sum_{\substack{S \in \mathcal{S}: \\ |S| \neq 1,2, \text{ or } n}} \left(1 - \frac{|S|}{n}\right)
$$

which is equal to

$$1 + a + 2^n - \binom{n}{1} - \binom{n}{2} - 1 - \frac{1}{n} \cdot \sum_{\substack{S \in \mathcal{S}: \\ |S| \neq 1,2, \text{ or } n}} |S|$$

which is itself equal to

$$a + 2^n - \binom{n}{1} - \binom{n}{2} - \frac{1}{n} \cdot \sum_{i=3}^{n-1} \left(i \cdot \binom{n}{i} \right)$$

which evaluates to

$$1 + a + 2^{n-1} - \frac{n^2}{2} + \frac{n}{2}.$$

Since the valued of $f_n(a)$ for a and n for which there exists a counterexample will be at most the value of the relaxation of $f_n(a)$ with the above fixings, and since that is smaller or equal to the value of its dual, the number of sets in a counterexample is at most $1 + a + 2^{n-1} - \frac{n^2}{2} + \frac{n}{2}$. $\square$

2.4 Analysis of the Second Model

The work in this section was done in collaboration with Jonad Pulaj.

We now consider the second model. We first discuss how to make the solution space smaller by symmetry breaking.

In order to have not too many solutions that are isometric by permutation of rows or columns, we first order the rows, i.e., the elements, in decreasing order of number of sets the element is contained in. In other words, if we let $\mathscr{S}_e = \{S \in \mathscr{S} | e \in S\}$, then $\mathscr{S}_1 \geq \mathscr{S}_2 \geq \ldots \geq \mathscr{S}_n$. This order yields the following obvious cuts for all $e, e+1 \in [n]$.

$$\sum_{i \in [m]} x_e^i \geq \sum_{i \in [m]} x_{e+1}^i$$

Then, we order the columns, i.e. the sets, in lexicographic order as follows.

Theorem 2.4.1. *The inequalities*

$$\sum_{e=1}^{l}\left((z_e^{i,i+1} - z_e^{i+1,i}) \cdot 2^{l-e}\right) \geq 0 \quad i, i+1 \in [m], l \in [n]$$

are valid cuts for $Fr(m,n)$*.*

Proof. These cuts are valid since they force the sets to be in lexicographic order and there exists an optimal solution whose sets are in lexicographic order. Fix some $l \in [n]$, and $i, i+1 \in [m]$. If S_i and S_{i+1} have exactly the same elements among $\{e_1, \ldots, e_l\}$, then $z_e^{i,i+1} = z_e^{i+1,i} = 0$ for all $1 \leq e \leq l$ and so the inequality is valid since $0 \geq 0$.

Otherwise, there exists e such that $1 \leq e \leq l$ and $x_e^i \neq x_e^{i+1}$. Let e^* be the smallest such element. Two cases can arise.

1. If $x_{e^*}^i = 1$ and $x_{e^*}^{i+1} = 0$, then $z_{e^*}^{i,i+1} = 1$ and $z_{e^*}^{i+1,i} = 0$. We also know that $z_e^{i,i+1} = z_e^{i+1,i} = 0$ for all $1 \leq e < e^*$. The minimum value that the left hand-side of the inequality can take is thus

$$2^{l-e^*} - \sum_{t=e^*+1}^{l} 2^{l-t} = 2^{l-e^*} - (2^{l-e^*} - 1) = 1$$

   which is greater than 0. Thus, a solution whose sets are ordered lexicographically is not cut by these inequalities.

2. If $x_{e^*}^i = 0$ and $x_{e^*}^{i+1} = 1$, then $z_{e^*}^{i,i+1} = 0$ and $z_{e^*}^{i+1,i} = 1$. We also know that $z_e^{i,i+1} = z_e^{i+1,i} = 0$ for all $1 \leq e < e^*$. The maximum value that the left handside of the inequality can take is thus

$$-2^{l-e^*} + \sum_{t=e^*+1}^{l} 2^{l-t} = -2^{l-e^*} + (2^{l-e^*} - 1) = -1$$

   which is smaller than 0. Thus, a solution whose sets are not ordered lexicographically is cut away by these inequalities.

□

Corollary 2.4.2. *The lexicographic order also yields the following cut for all* $2 \leq e \in [n]$ *and all* $i < j \in [m]$*:*

$$z_e^{ji} \leq \sum_{f=1}^{e-1} z_f^{ij}$$

Proof. If $z_e^{ji} = 1$, then $x_e^j = 1$ and $x_e^i = 0$. If $x_f^i = x_f^j$ for all $f < e$, then this would be a contradiction to the lexicographic order. There is at least one $f < e$ such that $x_f^i = 1$ and $x_f^j = 0$, for which $z_f^{ij} = 1$. □

We now consider some further cuts for the second model.

Theorem 2.4.3. *The following two cuts are valid for $Fr(m,n)$ for all $e \in [n]$:*

$$\frac{1}{m-1} \sum_{i \neq j} z_e^{ij} \leq \sum_i x_e^i,$$

$$\sum_{i \neq j} z_e^{ij} \leq \frac{m^2}{4}.$$

Proof. Let $\mathscr{S}$ be a union-closed family in $Fr(m,n)$ and let $\mathscr{S}_e := \{S \in \mathscr{S} | e \in S\}$, clearly $\sum_i x_e^i = |\mathscr{S}_e|$. We observe that $\sum_{i \neq j} z_e^{ij} = |\mathscr{S}_e| \cdot (m - |\mathscr{S}_e|)$. We know $\mathscr{S}_e \geq 1$ since if there exists an element that is in no sets, we can remove it and look for a smaller counterexample. So

$$\frac{1}{m-1} \sum_{i \neq j} z_e^{ij} = |\mathscr{S}_e| \cdot (m - |\mathscr{S}_e|) \frac{1}{m-1} \leq |\mathscr{S}_e| = \sum_i x_e^i,$$

which yields the first cut. The second cut is a stronger version based on the fact that we know that $\sum_{i \neq j} z_e^{ij} = |\mathscr{S}_e| \cdot (m - |\mathscr{S}_e|)$ increases as $|\mathscr{S}_e|$ increases, until $|\mathscr{S}_e| = \frac{m}{2}$, and on the fact that in a counterexample, $\mathscr{S}_e \leq \frac{m}{2}$ for all $e \in E$. □

Theorem 2.4.4. *The following cuts are valid for $Fr(m,n)$ for all $e \in [n]$ and $i, j, k \in [m]$.*

$$z_e^{ij} + 1 \geq z_e^{kj} + y_k^{ij},$$

$$z_e^{ij} + y_k^{ij} \leq z_e^{kj} + 1,$$

$$z_e^{ji} + 1 \geq z_e^{ki} + y_k^{ij},$$

$$z_e^{ji} + y_k^{ij} \le z_e^{ki} + 1,$$
$$z_e^{ik} + 1 \ge z_e^{jk} + y_k^{ij}, \text{ and}$$
$$z_e^{ik} + y_k^{ij} \le z_e^{jk} + 1.$$

Proof. If $y_k^{ij} = 1$, that is if $S_i \cup S_j = S_k$, then x_e^i and x_e^j determine the values of x_e^k as well as all z_e^{ab} where $\{a,b\} \in \{i,j,k\}$.

x_e^i	x_e^j	x_e^k	z_e^{ij}	z_e^{ji}	z_e^{ik}	z_e^{ki}	z_e^{jk}	z_e^{kj}
0	0	0	0	0	0	0	0	0
1	0	1	1	0	0	0	0	1
0	1	1	0	1	0	1	0	0
1	1	1	0	0	0	0	0	0

We thus observe that, if $y_k^{ij} = 1$, then $z_e^{ij} = z_e^{kj}$, $z_e^{ji} = z_e^{ki}$ and $z_e^{ik} = z_e^{jk}$. It's easy to see that, if $y_k^{ij} = 1$, the cuts enforce exactly that. If $y_k^{ij} = 0$, then the cuts do not force the z_e's to be anything. □

Theorem 2.4.5. *The following cuts are valid for $Fr(m,n)$ for all $i, j, k \in [m]$ with $i < j < k$.*

$$y_i^{ik} \ge y_j^{jk} + y_i^{ij} - 1$$

Proof. These cuts are based on the simple observation that if $S_i \cup S_j = S_i$ (i.e. $S_j \subset S_i$) and $S_j \cup S_k = S_j$ (i.e. $S_k \subset S_j$), that is if $y_i^{ij} = y_j^{jk} = 1$, then $S_k \subset S_i$, that is $S_k \cup S_i = S_i$, and so $y_i^{ik} = 1$. Thus, the cut says that if $y_i^{ij} = y_j^{jk} = 1$, then $y_i^{ik} \ge 1$, and since $y_i^{ik} \in \{0,1\}$, we have that $y_i^{ik} = 1$ as desired. Otherwise, the cut says that $y_i^{ik} \ge 0$ or -1, which is trivial. □

Theorem 2.4.6. *The following cuts are valid for $Fr(m,n)$ for all $i, j, k \in [m]$ with $i < j < k$:*

$$2 \cdot y_i^{jk} \le y_i^{ij} + y_i^{ik}$$

Proof. These cuts are based on another simple observation: if $S_j \cup S_k = S_i$, then $S_j \subset S_i$ and $S_k \subset S_i$. Thus, if $y_i^{jk} = 1$, so must y_i^{ij} and y_i^{ik}, which is what is enforced by the previous cut. Note that if $y_i^{jk} = 0$, then nothing is cut away. $\square$

Theorem 2.4.7. *The following cuts are valid for $Fr(m,n)$ for all $i, j, k \in [m]$ with $i < j < k$:*

$$\sum_{l=i}^{j} y_l^{jk} \geq y_i^{ij} + y_i^{ik} - 1$$

Proof. The following simple observation is this time necessary: if $S_j \subset S_i$ and $S_k \subset S_i$, then $S_j \cup S_k \subset S_i$. Thus, we know that the union of $S_j \cup S_k =: S_a$ is such that $i \leq a \leq j$. $\square$

Theorem 2.4.8. *The following cuts are valid for $Fr(m,n)$ for all $i, j, k, l \in [m]$ with $i < j < k < l$:*

$$y_i^{ij} \geq y_j^{kl} + y_i^{ik} + y_i^{il} - 2.$$

Proof. This cut is based on the following observation that if $S_k \subset S_i$, $S_l \subset S_i$, and $S_k \cup S_l = S_j$ (in which case the righthand side sums up to one), then $S_j \subset S_i$ (thus the left handside must also be one). Otherwise, the right handside is smaller or equal to zero and the inequality is trivial. $\square$

Theorem 2.4.9. *The following is a valid cut for $Fr(m,n)$:*

$$\sum_{i=1}^{m-1} z_1^{i,i+1} = 1.$$

Moreover, for $i < j \in [m]$, the following is also valid:

$$z_1^{ij} = \sum_{l=i}^{j-1} z_1^{l,l+1}.$$

Furthermore, for $i < j \in [m]$,

$$z_1^{ji} = 0.$$

Finally, we have that for all $e \in [n-1]$ and $i \in [m-1]$

$$\sum_{i=1}^{m-1} z_e^{i,i+1} \leq 2^{e-1}$$

as well as

$$\sum_{i=1}^{m-1} z_e^{i+1,i} = \sum_{i=1}^{m-1} z_e^{i,i+1} - 1$$

are valid inequalities for $Fr(m,n)$.

Proof. Since the sets are ordered lexicographically, we know the first row must be of the shape $1,1,\ldots,1,0,0,\ldots,0$, i.e. we go down from 1 to 0 only once and never come back up: a set containing the first element will always be ordered before than a set that doesn't contain the first element. This implies the two first cuts as well of the fixings of $z_1^{ji} = 0$.

The next cut is based on the following observation. For $e \in [n]$, we can partition the sets into sections such that the sets in a section contain the same elements among $\{1,\ldots,e\}$. Then, for $e+1$, because of the lexicographic order, the incidence vector of the sets in a same section, in which $e+1$ is contained must once again be of the shape $1,1,\ldots,1,0,0,\ldots,0$, i.e. it cannot go from 0 to 1 without breaking the order. Thus the number of times we encounter a 1 followed by a 0 in the incidence vector for $e+1$, $\sum_{i=1}^{m-1} z_{e+1}^{i,i+1}$, is at most the number of sections in e. The maximum number of sections in e is 2^{e-1}.

Finally, the last cut is based on the simple observation that, since the first set is fixed to contain all the elements and the last to contain none, for any element we look at, its containment incidence vector will be a succession of ones and zeroes that start with a one and ends with a zero. Therefore, the number of times we go from a zero to a one in this sequence will be one less than the number of times we go from a one to a zero. □

Theorem 2.4.10. *The following is a valid cut for $Fr(m,n)$ for all $i \in [m]$ and $e \in [n]$:*

$$z_e^{1,i} + x_e^i = 1$$

Proof. We know we can fix the first set to be all ones. Thus z_e^{1i} is 1 if x_e^i is 0 and 0 if x_e^i is 1. $\square$

We now add some cuts that aim at removing some fractional solutions of the linear relaxation of the model.

Theorem 2.4.11. *The following cuts are valid for $Fr(m,n)$ for all $e \in [n]$, $i \neq j \in [m]$:*

$$x_e^i + x_e^j + z_e^{ij} + z_e^{ji} \leq 2,$$
$$z_e^{ij} \leq x_e^i + x_e^j.$$

Moreoever, if we let $0 \leq z_e^{ij} \leq 1$ for all e,i,j be linear and $x_e^i \in \{0,1\}$ for all e,i, then $z_e^{ij} \in \{0,1\}$ for all e,i,j.

Proof. If x_e^i and z_e^{ij} are both binary, then for all triple e,i,j, there are only four scenarios possible:

	x_e^i	x_e^j	z_e^{ij}	z_e^{ji}
Scenario 1	1	1	0	0
Scenario 2	1	0	1	0
Scenario 3	0	1	0	1
Scenario 4	0	0	0	0

The two new cuts do not cut away any integer solution and are thus valid. They do cut away fractional solutions that contain components such as

x_e^i	x_e^j	z_e^{ij}	z_e^{ji}
1	1	0.5	0.5
0	0	0.5	0.5

Moreover, with these additional constraints, as well as $0 \leq z_e^{ij} \leq 1$, we have that if x_e^i for all e,i is binary, then so is z_e^{ij} for all e,i,j. Indeed, let's look again at the four possible scenarios of what x_e^i and x_e^j can be. Suppose that they are $(1,1)$. Then $x_e^i + x_e^j + z_e^{ij} + z_e^{ji} \leq 2$ implies that $z_e^{ij} + z_e^{ji} \leq 0$, and since $z_e^{ij}, z_e^{ji} \geq 0$, we have that $z_e^{ij} = z_e^{ji} = 0$ which is integral.

If $(x_e^i, x_e^j) = (1,0)$, then by the original inequality $z_e^{ij} \geq x_e^i - x_e^j$, we have that $z_e^{ij} \geq 1$ and since $z_e^{ij} \leq 1$, then $z_e^{ij} = 1$. Then, by the other original inequality $z_e^{ji} \leq \frac{x_e^j - x_e^i + 1}{2}$, we have that $z_e^{ji} \leq 0$, and since we also know that $z_e^{ji} \geq 0$, then $z_e^{ji} = 0$. So z_e^{ij} and z_e^{ji} are both integral. A similar argument holds for $(x_e^i, x_e^j) = (0,1)$.

Finally, if $(x_e^i, x_e^j) = (0,0)$, then $z_e^{ij} \leq x_e^i + x_e^j$ implies that $z_e^{ij} \leq 0$, and since $z_e^{ij} \geq 0$, we have that $z_e^{ij} = 0$. Similarly for z_e^{ji}. Thus, if x_e^i is binary, so is z_e^{ij} with these cuts. □

2.5 Cuts Derived From the Literature

The work in this section was the result of a collaboration with Jonad Pulaj.

Many authors discuss different cases for which the Frankl conjecture holds true. We can thus cut away these cases from all models since they only contain solutions for which the Frankl conjecture holds. In the optimization case for maximization, we already know that if $f_n(a) \leq 2a$ for the solutions we cut away, and if $f_n(a) \leq 2a$ afterwards, then all the solutions were Frankl. Again, most cuts valid for $F_n(a)$ are also valid for $F(n,m)$ and $G_n(m)$ by making some slight modifications to the notation. We leave those out to make the reading easier.

Theorem 2.5.1 (Folklore, see [18])**.** *Any family that contains a set that contains a single element is Frankl.*

Theorem 2.5.2 (Sarvate and Renaud, 1989)**.** *Any family that contains a set containing exactly two elements is Frankl.*

Corollary 2.5.3. *The following cuts are respectively valid for $Fr(m,n)$ for all $i \in [m]$ and for $F_n(a)$:*

$$\sum_{e \in E} x_e^i \geq 3, \text{ and}$$

$$x_S = 0 \quad \forall S \in \mathcal{S}_n \text{ such that } |S| = 1 \text{ or } 2.$$

□

Theorem 2.5.4 (Folklore, see [18]). *Any family for which the average cardinality of the sets is at least $n/2$ is Frankl.*

Corollary 2.5.5. *The following cuts are respectively valid for $Fr(m,n)$ and $F_n(a)$:*

$$\sum_{i\in[m]}\sum_{e\in[n]} x_e^S < \frac{n\cdot m}{2}, \text{ and}$$

$$\sum_{S\in\mathcal{S}_n} (|S| - \frac{n}{2})\cdot x_S < 0.$$

□

Remark. The cut $\sum_{i\in[m]}\sum_{e\in[n]} x_e^S < \frac{n\cdot m}{2}$ is redundant since we can recover it by summing up the inequalities $\sum_{i\in[m]} x_e^i < \frac{m}{2}$ over all elements $e \in E$.

Theorem 2.5.6 (Roberts, 1992). *A minimal counterexample contains a set of cardinality at most $\frac{m+1}{4}$.*

Corollary 2.5.7. *The following cut is valid for $Fr(m,n)$ simply by ordering the elements so that the elements contained in the mth set are the last elements:*

$$\sum_{e\in[n]} x_e^m \leq \frac{m+1}{4}.$$

For $F_n(a)$, we notice that $m+1 = 2a+2$ for a minimum counterexample, and so we have the following cut:

$$\sum_{0\leq|S|\leq\frac{a+1}{2}} x_s \geq 1.$$

□

Theorem 2.5.8 (Lo Faro, 1994). *A minimal counterexample contains a set of cardinality at least 9.*

Corollary 2.5.9. *The following cuts are valid for $Fr(m,n)$ and $F_n(a)$ respectively for $n \geq 9$:*

$$\sum_{e\in[n]} x_e^1 \geq 9, \text{ and}$$

$$\sum_{9\leq|S|\leq n-1} x_S \geq 1.$$

Proof. Since the family is union-closed and the sets are ordered lexicographically, the first set must be of maximum cardinality. Suppose not: suppose there exists $i > 1$ such that $|S_i| > |S_1|$. Then two cases arise. Either $S_1 \subseteq S_i$, which implies that S_i comes lexicographically before S_1, a contradiction; or either $S_1 \not\subseteq S_i$ (and $S_i \not\subseteq S_1$ either since $|S_i| > |S_1|$), and so $S_i \cup S_1 \neq S_1$ comes lexicographically before S_1, which is also a contradiction. □

This last proof can easily be extended to prove the following stronger cut.

Lemma 2.5.10. *The first set of the family contains all of the elements, i.e.*

$$\sum_{e\in[n]} x_e^1 = n$$

for $Fr(m,n)$ and

$$x_1 = 1$$

for $F_n(a)$.

Proof. Assume without loss of generality that each element is contained in at least one set; otherwise, we could remove that element and have a smaller counterexample. The union of the sets containing each element must be in the family since their pairwise union are. This union yields a set containing all the elements. Because our sets are lexicographically ordered, this set is the first one. □

Theorem 2.5.11 (Poonen, 1992)**.** *When looking for a minimal counterexample, it suffices to consider families $\mathscr{S}$ that contain the empty set.*

Corollary 2.5.12. *We have respectively for $Fr(m,n)$ and $F_n(a)$ that*

$$\sum_{e\in[n]} x_e^m = 0, \text{ and}$$

$$x_{2^n} = 1.$$

□

Theorem 2.5.13 (Poonen, 1992)**.** *When looking for a minimal counterexample, it suffices to consider families for which there does not exist $e, f \in E$ such that e and f are elements of exactly the same sets, i.e. blocks should be singletons.*

Corollary 2.5.14. *For $Fr(m,n)$, we introduce the variables $w_{e,f}^i \in \{0,1\}$ for all $e, f \in E$ with $e \neq f$, and $i \in [m]$, where $w_{e,f}^i$ is 1 if e is in set i but not f, and 0 otherwise. We add the following constraints for all $e, f \in E$ with $e \neq f$, and $i \in [m]$:*

$$\frac{x_e^i - x_f^i}{2} \leq w_{e,f}^i \leq \frac{x_e^i - x_f^i + 1}{2}.$$

The following cuts are then valid for $Fr(m,n)$ for all $e, f \in E$:

$$\sum_{i\in[m]} (w_{e,f}^i + w_{f,e}^i) \geq 1.$$

For $F_n(a)$, the following cuts are valid:

$$\sum_{\substack{S\in\mathcal{S}_n:\\ e\in S,\\ f\notin S}} x_S + \sum_{\substack{S\in\mathcal{S}_n:\\ e\notin S,\\ f\in S}} x_S \geq 1 \quad \forall e \neq f \in [n].$$

Proof. For $Fr(m,n)$, the variable w is very similar to z. The constraints force $w_{e,f}^i$ to be one if and only if $e \in S_i$ but $f \notin S_i$. And so the cut states that the symmetric difference between the sets containing e and/or f should be greater or equal to 1, i.e. e and f cannot be exactly in the same sets.

For $F_n(a)$, there exists a set in the family that contains one but not the other of any two distinct elements, meaning that they do not form a block. □

Corollary 2.5.15. *There exists a set of cardinality $n-1$ in a minimal counterexample, and so the following cuts are valid for $Fr(m,n)$ and $F_n(a)$:*

$$\sum_{e\in[n]} x_e^2 = n-1, \text{ and}$$

$$\sum_{S:|S|=n-1} x_S \geq 1.$$

Proof. This follows from Poonen's theorem on blocks. We know that $\mathbb{1}$ is in the solution. If there are no sets of size $n-1$, the second greatest set, say S, must have cardinality at most $n-2$. The elements missing from that set form a block: in any set, either they are all there or they are all missing, otherwise, there would exist a set of cardinality greater than S and smaller than $\mathbb{1}$, namely the union of S with such a set. Since there's more than one element missing, this is a block of size greater than one, which means that this family cannot be a minimal counterexample. □

Theorem 2.5.16 (Marić, Živković, Vučković, 2012)**.** *Frankl's conjecture holds true if the there exist three distinct sets S_i, S_j, S_k in the family such that $|S_i| = |S_j| = |S_k| = 3$ and $|S_i \cup S_j \cup S_k| \leq 5$.*

Corollary 2.5.17. *The following cut is valid for $Fr(m,n)$ for all distinct $i, j, k \in [m]$ and for $F_n(a)$ for all distinct sets of five elements $\{e_1, e_2, \dots, e_5\} \subset [n]$:*

$$\sum_{e\in E}(x_e^i + x_e^j + x_e^k) + \frac{1}{9}\cdot\sum_{e\in E}(z_e^{i,j} + z_e^{j,i} + z_e^{i,k} + z_e^{k,i} + z_e^{j,k} + z_e^{k,j} \geq 10, \text{ and}$$

$$\sum_{\substack{S\subset\{e_1,\dots,e_5\}:\\ |S|=3}} x_S \leq 2.$$

Proof. We already know that there are no sets of cardinality one or two in a counterexample, $\sum_{e\in E}(x_e^i + x_e^j + x_e^k)$ will already be at least 9. If one of the set has cardinality four or more, then the inequality holds trivially. The only solutions that get cut are those that contain some distinct sets S_i, S_j, S_k of cardinality three such that $\sum_{e\in E}(z_e^{i,j} + z_e^{j,i} + z_e^{i,k} + z_e^{k,i} + z_e^{j,k} + z_e^{k,j}) \leq 8$. Observe that if $U := |S_i \cup S_j \cup S_k| \geq 6$, then there can be at most one element that is in $S_i \cap S_j \cap S_k$ and that for any element $e^* \in E$ such that e^* is in one or two (but not three) of S_i, S_j, S_k, we have that $z^{i,j}_{e^*} + z^{j,i}_{e^*} + z^{i,k}_{e^*} + z^{k,i}_{e^*} + z^{j,k}_{e^*} + z^{k,j}_{e^*} = 2$. Thus if $U \geq 6$, then the sum of the symmetric difference between S_i and S_j, S_i and S_k as well as S_j and S_k yields $\sum_{e\in E}(z^{i,j}_e + z^{j,i}_e + z^{i,k}_e + z^{k,i}_e + z^{j,k}_e + z^{k,j}_e) \geq (U-1)\cdot 2 \geq 5 \cdot 2 = 10$. Therefore, if a solution contains S_i, S_j, S_k all of cardinality three such that $U \geq 6$, it does not get cut by these inequalities. However, if $U \leq 5$, then there is at least one element in $S_i \cap S_j \cap S_k$, and so $\sum_{e\in E}(z^{i,j}_e + z^{j,i}_e + z^{i,k}_e + z^{k,i}_e + z^{j,k}_e + z^{k,j}_e) \leq (U-1)\cdot 2 \leq 4 \cdot 2 = 8$. Thus the inequalities cut away any solution containing three sets of cardinality three that span at most five elements, solutions that are known to be Frankl. □

Theorem 2.5.18 (Marić, Živković, Vučković, 2012)**.** *Frankl's conjecture holds true if the there exist four distinct sets S_i, S_j, S_k, S_l in the family such that $|S_i| = |S_j| = |S_k| = |S_l| = 3$ and $|S_i \cup S_j \cup S_k| \leq 7$.*

Corollary 2.5.19. *For $F_n(a)$,*

$$\sum_{\substack{S\subset\{e_1,\ldots,e_7\}:\\ |S|=3}} x_S \leq 3 \quad \forall\{e_1,\ldots,e_7\} \subset [n].$$

For $Fr(m,n)$, we didn't find a way to include that cut without introducing even more variables.

2.6 Additional Theorems and Cuts

Definition 2.6.1. We call a set S of a union-closed family $\mathcal{F}$ *minimal* if and only if no other set in $\mathcal{F}$ is a proper subset of S except for the empty set.

Definition 2.6.2. We call a set S of a union-closed family $\mathcal{F}$ *quasi-minimal* if and only if there doesn't exist two proper subsets of S whose union is S.

Lemma 2.6.3. *Removing a minimal or a quasi-minimal set from a union-closed family leaves the family union-closed.*

Proof. The union of any two sets still remaining after the removal will still be there; the family is thus still union-closed. □

Theorem 2.6.4. *In a minimal counterexample family $\mathcal{F}$ with $f_n(a) \geq 2a+1$, there are no two disjoint non-empty minimal or quasi-minimal sets that span all n elements.*

Proof. Suppose not: remove from $\mathcal{F}$ the two disjoint minimal or quasi-minimal sets that span the n elements. Then $\mathcal{F}$ is a union-closed family for which each element is in at most $a-1$ sets. Thus, $\mathcal{F}$ now contains at most $f_n(a-1)$ sets. Since $F_n(a)$ was a minimal counterexample, we know $f_n(a-1) \leq 2(a-1)$. So the original family before the removal of the two sets had at most $2(a-1)+2 = 2a$ sets, and was thus Frankl, a contradiction. □

Theorem 2.6.5. *A minimal counterexample must satisfy*

$$\sum_{S: n-2 \leq |S| \leq n-1} x_S \geq 2.$$

Proof. Suppose not: then $\sum_{S: n-2 \leq |S| \leq n-1} x_S \leq 1$, and from the previous theorem, we know this means there's exactly one set of size $n-1$ and none of size $n-2$. Thus the third greatest set, say S, has cardinality at most $n-3$. This set is missing at least two elements that were in the $(n-1)$-set. Indeed, without loss of generality, we are in one of these two cases:

$$\begin{array}{ccc} 1 & 1 & 1 \\ 1 & 1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{array} \quad \text{or} \quad \begin{array}{ccc} 1 & 1 & 1 \\ 1 & 1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{array}$$

In the first case, if there is a set that contains exactly one of the two (or more) elements that are not in the third set, but are in the second set, then there would exist another set of cardinality greater than S and smaller than n that is not the already present $(n-1)$-set. So those two (or more elements) form a block, and so this cannot be a minimal counterexample.

In the second case, if any set with one or two elements from the elements that are in the second set but not in the third set is present, then the union with the third set must also be present, and this set will have cardinality greater than S and smaller than n and will not be the $(n-1)$-set already present since the last element will be present. So these three (or more) elements form a block, and so this cannot be a counterexample. $\square$

Theorem 2.6.6. *A minimal counterexample must satisfy*

$$\sum_{S: n-3 \leq |S| \leq n-1} x_S \geq 3.$$

Proof. Suppose not: then $\sum_{S: n-3 \leq |S| \leq n-1} x_S \leq 2$, and from the previous theorem, we know this means that there is one set of size $n-1$ and one of size $n-1$ or $n-2$. This leads to three different cases.

Case 1 We have two sets of size $n-1$, say S_1 and S_2. The next greatest set, say S, has size at most $n-4$, and it's easy to see that no matter what that set is, there will be at least two elements that are not in S but that are in S_1, S_2 and $\mathbb{1}$. These elements will form a block, since if there exists a set with only a fraction of these elements, then the union of that set with S would create a set of cardinality greater than S but smaller than $\mathbb{1}$ which isn't already present, since at least one of the block elements wouldn't be in the union, but are in S_1 and S_2.

Case 2 There is one set of size $n-1$, say S_1, and one of size $n-2$, say S_2, and $|S_1 \cap S_2| = n-2$. Once again, the next greatest set, say S, has size at most $n-4$ and there is at least two elements that are not in S, but that are in S_1, S_2, and $\mathbb{1}$, and so the previous argument holds.

Case 3 There is one set of size $n-1$, say S_1, and one of size $n-2$, say S_2, and $|S_1 \cap S_2| = n-2$, i.e. without loss of generality, we have

$$\begin{array}{ccc}
1 & 1 & 1 \\
1 & 1 & 1 \\
\vdots & \vdots & \vdots \\
1 & 1 & 1 \\
1 & 1 & 0 \\
1 & 1 & 0 \\
1 & 0 & 1
\end{array}$$

Without even looking at the next greatest set, we know there's already a block of two elements, namely the two elements before the last one. Indeed, if any set contained only one of those two elements, then its union with S_2 would produce a set of size $n-1$ that is not S_1.

Thus, in every case there is a block of size greater or equal to two, and so this cannot be a minimal counterexample. $\square$

We now generalize this theorem even more.

Theorem 2.6.7. *The following inequalities holds for any minimal counterexample:*

$$\sum_{\substack{S: \\ n-1-j \leq |S| \leq n-1}} x_S \geq 1 + j$$

for all $0 \leq j \leq n-2$.

Proof. We show it by induction. From the previous theorem, we know it is true for $j = 0, 1, 2$.

Suppose it holds for $j'-1$, but not for j'. Let $\mathscr{S}' := \{S : n-1-j' \leq |S| \leq n-1\}$. Then, we have $|\mathscr{S}'| \leq j'$, and since $|\{S : n-1-(j'-1) \leq |S| \leq n-1\}| \geq j'$, it's clear that $|\mathscr{S}'| = j'$. Order the sets in $\mathscr{S}'$ first by cardinality and then lexicographically. Look at the number b of elements that are spanned by zeroes of $\mathscr{S}'$. We consider two cases.

First, suppose $b \leq j'$. Then, since any set $S \neq \mathbb{1}$ part of the solution but which is not in $\mathscr{S}'$ contains at most $n-2-j'$ elements, and S has at least at least $j'+2$ zeroes. This means there exists at least two elements that were in all sets of $\mathscr{S}'$ but that are not in S. These elements form a block since if there exists some other set T which contains only some but not all these elements, then $S \cup T$ would be of cardinality greater than S, and thus should be in $\mathscr{S}'$. However, $S \cup T$ is still missing at least one of those elements, and thus cannot be any of the sets that are in $\mathscr{S}'$. Thus, these elements form a block of size at least two, and so the solution cannot be a minimal counterexample.

Now, suppose instead that $b \geq j+1$. This means there exists a set $U \in \mathscr{S}'$ such that U adds at least two elements more to the set of elements spanned by zeroes by its predecessors. In other words, there are at least two elements that are not in

U but that are in all the sets before U in $\mathscr{S}'$. We can once again show that these elements must form a block by the same argument as before. Indeed, if there exists a set V in the solution which contains some but not all of these elements, then its union with U should be listed in the sets ahead of U in $\mathscr{S}'$, but it clearly is not since it would still be missing at least one of these elements.

Thus, we have if $|\mathscr{S}'| \leq j'$, then the solution is not a minimal counterexample, which is a contradiction. Therefore, $|\mathscr{S}'| \geq 1 + j'$. □

Theorem 2.6.8. *For a minimal counterexample, there cannot be exactly one set of size $n-i$, $n-i+1$, …, and $n-1$ for any $2 \leq i \leq n-1$.*

Proof. Suppose not: suppose there is exactly one set of size $n-i$, say S_{n-i}, one of size $n-i+1$, say $S_{n-i+1}, \ldots$, and one of size $n-1$, say S_{n-1}. Then these $i+1$ sets of the solution must be nested inside each other like matryoshka dolls in order to avoid blocks of size two or more: $S_{n-j} \subset S_{n-j+1}$, and that, for all $2 \leq j \leq i$.

We now show that any sets of cardinality $(n-i-1)$ present in the solution must be contained in the set S_{n-i}. Suppose not: suppose there exists an $(n-i-1)$-set S that is not contained in S_{n-i}. If there are many such sets, take the smallest one lexicographically. Then, S is missing at least two elements that are in S_{n-i}, and thus in all greater sets. In order for those elements not to be a block (which would mean that the solution was not a minimal counterexample), there must be another set in the rest of the solution, say T, that is lexicographically smaller (and thus contained in S_{n-i}) and that contains some of these elements (but not all). If such a T exists, then $S \cup T$ is a set of cardinality greater or equal to $(n-i)$ that is missing at least one element present in S_{n-i}, meaning that it is not one of the sets already listed, a contradiction. Thus, every $(n-i-1)$-set present is contained in S_{n-i}.

We now show by induction on k that any $(n-k)$-sets present in the solution must be contained in S_{n-i}. We just showed that it was true for $k = i+1$, and we'd like to show it's true up to $n-1$. Assume it is true for some $k'-1$, and let's show it is true for k'.

Suppose it's not true for $(n-k')$. Then there exists a lexicographically smallest $(n-k')$ set U that is not contained in S_{n-i}. Then two possibilities arise:

- there exists a set $V \subseteq S_{n-i}$ of cardinality between $(n-i-1)$ and $(n-k')$ which has an element missing that is also missing in U but that is in S_{n-i}. Then $U \cup V$ form a set of cardinality greater than V which contains an element not in S_{n-i} (and thus cannot be one of the sets of cardinality $(n-i-1)$

to $(n-k'+1)$ since they are all contained in S_{i-1}) but which is also missing an element that is in S_{n-i} (and thus cannot be one of the sets of cardinality $n-i$ to n). Since $U \cup V$ is not in the family, the solution is not union-closed, which is a contradiction.

- there does not exist such a set V. In that case, every set listed before U contains the elements that are missing in U but that are in S_{n-i}. As seen previously, these elements form a block, and since there are at least two such elements, the solution cannot be a minimal counterexample.

Thus, every $(n-k)$-set present is contained in S_{n-i} for $i+1 \leq k \leq n-1$. This means that $f_n(a) \leq f_{n-i}(a-i)+i$. Since this family is supposed to be a minimal counterexample, $f_{n-i}(a-i) \leq 2(a-i)$ sets, and so $f_n(a) \leq 2a-i$, which is also Frankl, a contradiction. □

2.7 New Conjectures

2.7.1 The Conjectures and Their Consequences

The work in this subsection was done with Jonad Pulaj.

The point of studying these polytopes better was to see whether we could prove computationally that if $n = 13$, then the Frankl conjecture holds. Unfortunately, even with all the cuts we've added for the different models, we couldn't reach this result (see Appendix for computations). However, these computations were not completely worthless: they allowed us to notice a pattern that was, to the best of our knowledge, left unmentioned in the literature. We state this newfound pattern in the following two conjectures.

Conjecture 2.7.1 (f-conjecture)**.** *For every a, there exists F_a such that $f_n(a) = F_a$ for all n such that $n \geq \lceil \log_2 a \rceil + 1$, i.e. for all n for which it makes sense to calculate the number $f_n(a)$.*

Conjecture 2.7.2 (g-conjecture)**.** *For every m, there exists G_m such that $g_n(m) = G_m$ for all n such that $n \geq \lceil \log_2 m \rceil$, i.e. for all n for which it is possible to calculate the number $g_n(m)$.*

We checked these conjectures computationally up to $n = 9$ (see Appendix). Note that these conjectures are different than the Frankl conjecture. For one thing, even if the f- and g-conjectures hold, one would still need to show that $f_{\lceil \log_2 a \rceil + 1}(a) \leq 2a$ for every a or that $g_{\lceil \log_2 m \rceil}(m) \geq \frac{m}{2}$ for all m to prove the Frankl conjecture; therefore the f- and g-conjectures do not imply the Frankl conjecture. Moreover, the Frankl conjecture does not imply the f- and g-conjectures. Certainly, if the Frankl conjecture is true, then $\lim_{n\to\infty} f_n(a) = F'_a$ for some $F'_a \leq 2a$, else the Frankl conjecture wouldn't be true, and so the f function has to stabilize at some point as n increases since we show in theorem 2.7.8 that $f_n(a) \leq f_{n+1}(a)$, and we also know that $f_n(a) \in \mathbb{N}$. However, the Frankl conjecture does not imply that the function f should be stable immediately as $n \geq \lceil \log_2 a \rceil + 1$, i.e. as soon as there are enough elements for there to be at least a sets containing an element e. Similarly, if the Frankl conjecture is true, then $\lim_{n\to\infty} g_n(m) = G'_m$ for some $G'_m \geq \frac{m}{2}$, and so the g function has to stabilize at some point as n increases since we show in theorem 2.7.9 that $g_n(m) \geq g_{n+1}(m)$. But again, the Frankl conjecture does not imply that the function g has to stabilize immediately as $n \geq \lceil \log_2 m \rceil$, i.e. as soon as there are enough elements for there to be at least m sets in the power set of n.

Hopefully it is now clear that the f- and g-conjectures are different from the Frankl conjecture, however it might not yet be evident why one should care about these two new conjectures. The fact is that if we prove them to be true, this would have a tremendous impact on the Frankl conjecture. One consequence would come from combining them with a recent result of Balla, Bollobás and Eccles.

Theorem 2.7.3 (Balla, Bollobás & Eccles, 2013)**.** *The union-closed conjecture holds for any family based on n elements with at least $\frac{2}{3}2^n$ sets.*

Corollary 2.7.4. *If the g-conjectures holds, then the union-closed conjecture holds for any family with m sets where $\frac{2}{3}2^i \leq m \leq 2^i$ for some $i \in \mathbb{N}$.*

Proof. By theorem 2.7.3 of Balla, Bollobás and Eccles, we know that $g_{\lceil \log_2 m \rceil}(m) \geq \frac{m}{2}$ for $\frac{2}{3}2^{\lceil \log_2 m \rceil} \leq m \leq 2^{\lceil \log_2 m \rceil}$, and if the g-conjecture holds, then $g_n(m) \geq \frac{m}{2}$ for all $n \geq \lceil \log_2 m \rceil$ and $\frac{2}{3}2^{\lceil \log_2 m \rceil} \leq m \leq 2^{\lceil \log_2 m \rceil}$. So the Frankl conjecture would hold for any family consisting of m sets when $\frac{2}{3}2^i \leq m \leq 2^i$ for some $i \in \mathbb{N}$.

In other words, in the first table of 2.1 of values of $g_n(m)$, the theorem of Balla, Bollobás and Eccles says that the parts in red have to be at least $\frac{m}{2}$.

The g-conjecture states that for any row m' of the table 2.1, $g_n(m')$ is equal to some $G_{m'}$ for all sensible n. So if we know the first value in a row of the table is

Table 2.1: Values of $g_n(m)$ known to be at least $\frac{m}{2}$ respectively without and with the g-conjecture

$m \backslash n$	1	2	3	4	5	6	7	8
2	1	1	1	1	1	1	1	1
3		2	2	2	2	2	2	2
4		2	2	2	2	2	2	2
5			3	3	3	3	3	3
6			4	4	4	4	4	4
7			4	4	4	4	4	4
8			4	4	4	4	4	4
9				5	5	5	5	5
10				6	6	6	6	6
11				7	7	7	7	7
12				7	7	7	7	7
13				8	8	8	8	8
14				8	8	8	8	8
15				8	8	8	8	8
16				8	8	8	8	8

$m \backslash n$	1	2	3	4	5	6	7	8
2	1	1	1	1	1	1	1	1
3		2	2	2	2	2	2	2
4		2	2	2	2	2	2	2
5			3	3	3	3	3	3
6			4	4	4	4	4	4
7			4	4	4	4	4	4
8			4	4	4	4	4	4
9				5	5	5	5	5
10				6	6	6	6	6
11				7	7	7	7	7
12				7	7	7	7	7
13				8	8	8	8	8
14				8	8	8	8	8
15				8	8	8	8	8
16				8	8	8	8	8

at least $\frac{m}{2}$, then we know the other entries in that row will be exactly the same as the first one, and so will also be at least $\frac{m}{2}$.

That is to say, we'd know that all of these red numbers in the second table of 2.1 are at least $\frac{m}{2}$. These red values correspond to about $\frac{2}{3}$ of all rows, so the Frankl conjecture would thus hold for about $\frac{2}{3}$ of all possible m. $\square$

In other words, if the g-conjecture holds and a counterexample to the Frankl conjecture with m sets still exists, then m would have to be such that $2^i + 1 \leq m \leq \left\lfloor \frac{2}{3} 2^{i+1} \right\rfloor$ for some $i \in \mathbb{N}$.

Another fantastic consequence of the f- and g-conjectures would be that there would finally be a known constant percentage of sets containing the most frequent element in a union-closed family. Currently, the best known result on this matter is the following:

Theorem 2.7.5 (Knill, 1994)**.** *In a union-closed family of m sets, there always exists an element present in at least $\frac{m-1}{\log_2 m}$ sets.*

Wójcik improved this result a bit in [104], but still no constant is known. However, the f- and g-conjectures along with the following observation of Bruhn and

Schaudt would yield a constant frequency; moreover, that frequency would be very close to the desired $\frac{1}{2}$.

Theorem 2.7.6 (Bruhn & Schaudt, 2013)**.** *Any union-closed family on m sets and n elements with $2^{n-1} < m \leq 2^n$ contains an element contained in at least $\frac{6}{13} \cdot m$ sets of the family.*

Proof. We refer the reader to [18] for the proof, and simply mention that this observation is based on the theorems of [82] and [102]. $\square$

Corollary 2.7.7. *If the f- and g-conjectures hold, then any union-closed family on m sets contains an element in at least $\frac{6}{13} \cdot m$ sets of the family.*

Proof. By the observation 2.7.6, we know that $g_{\lceil \log_2 m \rceil}(m) \geq \frac{6}{13} \cdot m$, and by the g-conjecture, we know then that $g_n(m) \geq \frac{6}{13} \cdot m$ for all $n \geq \lceil \log_2 m \rceil$. Therefore, we know that any family on m sets contains an element in at least $\frac{6}{13} \cdot m$ sets of the family.

In other words, we're again using the trick of pushing values to the right in the g-table. The Bruhn and Schaudt theorem states the values in red in the first table of 2.2 (basically any first value in a row) are at least $\frac{6}{13}m$.

Since the g-conjecture states that all the values in a row of the g-table must be the same, then every other value in each row must also be at least $\frac{6}{13}m$ as seen in the second table of 2.2. $\square$

Of course, $\frac{6}{13}m$ is not the same as $\frac{m}{2}$, however it is quite close, and it is a constant, thus improving dramatically Knill's bound in 2.7.5.

Thus, the f- and g-conjectures deserve to be studied as their consequences would be quite spectacular. So first, let's study properties of $f_n(a)$ and $g_n(m)$ better.

2.7.2 Properties of the f- and g-functions

The work in this subsection is the result of a collaboration with Dirk Theis.

Certainly, the two following statements are true.

Theorem 2.7.8. *For every a and n such that $n \geq \lceil \log_2 a \rceil + 1$, $f_n(a) \leq f_{n+1}(a)$, that is that the function f is non-decreasing in n.*

Table 2.2: Values of $g_n(m)$ known to be at least $\frac{6}{13}m$ respectively without and with the g-conjecture

$m\backslash n$	1	2	3	4	5	6	7	8
2	1	1	1	1	1	1	1	1
3		2	2	2	2	2	2	2
4		2	2	2	2	2	2	2
5			3	3	3	3	3	3
6			4	4	4	4	4	4
7			4	4	4	4	4	4
8			4	4	4	4	4	4
9				5	5	5	5	5
10				6	6	6	6	6
11				7	7	7	7	7
12				7	7	7	7	7
13				8	8	8	8	8
14				8	8	8	8	8
15				8	8	8	8	8
16				8	8	8	8	8

$m\backslash n$	1	2	3	4	5	6	7	8
2	1	1	1	1	1	1	1	1
3		2	2	2	2	2	2	2
4		2	2	2	2	2	2	2
5			3	3	3	3	3	3
6			4	4	4	4	4	4
7			4	4	4	4	4	4
8			4	4	4	4	4	4
9				5	5	5	5	5
10				6	6	6	6	6
11				7	7	7	7	7
12				7	7	7	7	7
13				8	8	8	8	8
14				8	8	8	8	8
15				8	8	8	8	8
16				8	8	8	8	8

Proof. Suppose $f'_n(a') = m'$. Take a family $\mathcal{F}$ that is optimal. Add an element $n'+1$ to the family such that this element is exactly in the same sets as some other element e of the family. Then this augmented family is still union-closed, and every element of it is still in at most a' sets. Thus, $f_{n'+1}(a') \geq m' = f_{n'}(a')$. □

Theorem 2.7.9. *For every m and n such that $n \geq \lceil \log_2 m \rceil$, $g_n(m) \geq g_{n+1}(m)$, that is that the function g is non-increasing in n.*

Proof. Suppose $g_{n'}(m') = a'$ for some n', m'. Take a family $\mathcal{F}$ that is optimal. Add an element $n'+1$ to the family such that this element is exactly in the same sets as some other element e of the family. Then this augmented family is still union-closed, and every element of it is still in at most a' sets. Thus, $g_{n'+1}(m') \leq a' = g_{n'}(m')$. □

Theorem 2.7.10. *The function f is strictly increasing in a, that is, $f_n(a) < f_n(a+1)$ for every n and a such that $n \geq \lceil \log_2 a \rceil + 1$.*

Proof. Suppose $f_n(a) = m$. Take a family $\mathcal{F}$ that is optimal. Then add to this family one of the greatest set, i.e. a set containing as many elements as possible, that is not already present in the family. This is always possible if $f_n(a) < 2^n$. If $f_n(a) = 2^n$, then every set possible is present and so each element is in exactly 2^{n-1} sets, and so $a \geq 2^{n-1}$. Since we assume that $a \leq 2^{n-1}$, this means that $a = 2^{n-1}$, and that $a+1 = 2^{n-1}+1$, and so we do not consider that extreme case since $f_n(a') = 2^{n-1}$ for all $a' \geq 2^{n-1}$.

Otherwise, the new family we built has $m+1$ sets and is still union-closed since taking the union of any other set with the added set will give either the new set itself or a greater set (which are all present in the family). Moreover, every element is present in at most $a+1$ sets. Thus, $f_n(a+1) \geq m+1 = f_n(a)+1$. □

Theorem 2.7.11. *The function g is non-decreasing in m, that is* $g_n(m) \leq g_n(m+1)$ *for every m and n such that* $n \geq \lceil \log_2 m \rceil$.

Proof. Supose $g_n(m) = a$. First note that there always exists an optimal solution for $g_n(m+1)$ that contains the empty set. If not, taking the smallest set in an optimal family without the empty set and replacing it by the empty set would leave the family union-closed. Moreover, the total number of sets wouldn't change, and the number of times the most frequent element is in the family could only go down, and since the original family was thought to be optimal, it doesn't.

Thus take an optimal solution for $g_n(m+1)$ that contains the empty set. Remove the empty set. This new family has m sets, is union-closed, and every element is there at most a times, so $g_n(m) \leq a$. □

Theorem 2.7.12. *We have that* $g_n(f_n(a)) = a$ *for all a and n.*

Proof. Suppose that $f_n(a) = m$. This means that the most sets that can be in a union-closed family such that every element is present in at most a sets is m. Certainly, this means that $g_n(m) = g_n(f_n(a)) \leq a$. Suppose that $g_n(m) = g_n(f_n(a)) = a' < a$. Then, it means that $f_n(a') \geq m$. Since we have already shown that the function f strictly increases in a, this is a contradiction. □

Theorem 2.7.13. *We have that* $f_n(g_n(m)) \geq m$.

Proof. Suppose that $g_n(m) = a$. This means that the minimum number of sets that the most frequent element of a union-closed family with m sets is in is a. Certainly, this means that $f_n(a) \geq m$. □

It will come as no surprise that the f- and g-conjectures are equivalent.

Theorem 2.7.14. *We have that $f_n(a) = f_{n+1}(a)$ for every a and n such that $n \geq \lceil \log_2 a \rceil + 1$ if and only if $g_{n'}(m) = g_{n'+1}(m)$ for every m and n' such that $n' \geq \lceil \log_2 m \rceil$.*

Proof. $\Leftarrow$: Pick an arbitrary $a' \in \mathbb{N}$. By theorem 2.7.12, we know that there exists m such that $g_{n'}(m) = a'$, namely $m := f_{n'}(a')$. Let m^* be the greatest number for which $g_{n'}(m^*) = a'$. It follows that $f_{n'}(a') \leq m^*$ since, otherwise, if $f_{n'}(a') = m^* + b$, then we would have $g_{n'}(m^* + b) = g_{n'}(f_{n'}(a')) = a'$ by theorem 2.7.12, thus contradicting the maximality of m^*. Moreover, $f_{n'}(a') \geq m^*$ since $f_{n'}(g_{n'}(m^*)) \geq m^*$ by theorem 2.7.13. Thus, $f_{n'}(a') = m^*$ for all n' and a'.

$\Rightarrow$: For all m' such that there exists an a such that $f_n(a) = m'$, we have that $g_n(m') = g_n(f_n(a)) = a$ for all n by theorem 2.7.12. For all m'' such that there exists no a' such that $f_n(a') = m''$, we have that $f_n(g_n(m'')) = m'' + b$ by theorem 2.7.13, and that $g_n(m'') = g_n(m'' + b)$. Then $g_n(m'' + b) =: a''$ for all n since $f_n(g_n(m'' + b)) = m'' + b$, and so we are back to the first case. By theorems 2.7.9 and 2.7.11, this implies that $g_n(m'') = g_n(m'' + b) = a''$ for all n. $\square$

The two conjectures are thus equivalent: proving one would prove the other. We thus focus on the f-conjecture from now on.

2.7.3 Twin Sets

The work in this subsection was done in collaboration with Dirk Theis.

We now need to introduce a new idea: twin sets.

Definition 2.7.15. We call two sets S_1, S_2 with $n > |S_1| > |S_2|$ *twins sets* if $|S_1 \triangle S_2| = 1$. We call S_1 the *big twin* and S_2 the *little twin*. Moreover, we say of an element e that it is a *twin difference* if there exists twin sets S_3 and S_4 such that $S_3 \backslash S_4 = \{e\}$.

Remark. Twin sets play an important role here. Suppose that for $f_{n+1}(k) = m$ there exists an optimal solution for which there exists an element that is not a twin difference. Then removing this element will leave all of the sets distinct, and the family will still be union-closed, but with n elements. In that case, we know that $f_n(k) \geq m = f_{n+1}(k)$, and since we've already shown that $f_n(k) \leq f_{n+1}(k)$, then $f_n(k) = f_{n+1}(k)$ as in the conjecture. So the only case we have to worry about to prove the f-conjecture is if every optimal solution of some $f_{n+1}(k) = m$ is such that every element is a twin difference.

Example 2.7.16. For example, the following union-closed family on 6 elements (labeled as e_1 through e_6) and 18 sets (labeled as T_1 through T_{18} with $a = 10$ has a twin for each element.

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}	T_{16}	T_{17}	T_{18}
e_1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
e_2	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
e_3	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
e_4	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0
e_5	1	1	0	0	1	1	0	0	1	1	1	1	0	0	1	1	0	0
e_6	1	0	1	0	1	0	1	0	1	1	1	0	1	0	1	0	1	0

If we remove element e_3, then we're left with

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}	T_{16}	T_{17}	T_{18}
e_1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
e_2	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
e_4	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0
e_5	1	1	0	0	1	1	0	0	1	1	1	1	0	0	1	1	0	0
e_6	1	0	1	0	1	0	1	0	1	1	1	0	1	0	1	0	1	0

and we now have two sets that are the same, and so this is not a solution to our problem. Similarly, if we remove element e_6, we obtain

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	T_{14}	T_{15}	T_{16}	T_{17}	T_{18}
e_1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
e_2	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
e_3	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
e_4	1	1	1	1	0	0	0	0	1	1	1	1	1	1	0	0	0	0
e_5	1	1	0	0	1	1	0	0	1	1	1	1	0	0	1	1	0	0

where many sets are doubled. And so on for every element.

Table 2.3: $f_n(a)$ with twins

$a \backslash n$	1	2	3	4	5	6	7	8
1	–	–	–	–	–	–	–	–
2		–	–	–	–	–	–	–
3			–	–	–	–	–	–
4			8	–	–	–	–	–
5				8	–	–	–	–
6				10	9	–	–	–
7				12	11	10	–	–
8				16	13	12	11	–
9					15	14	13	12
10					18	16	15	14
11					19	19	17	16
12					21	20	20	18
13					23	22	21	21
14					25	24	23	22
15					27	25	25	24
16					32	28	26	26

Moreover, note that we did not consider that the set S_1 with every element counts as a twin, because if S_1 is a bigger twin for some element e' and if we then remove its twin element e', then we can replace its smaller twin $S_1 \backslash \{e'\}$ (that is now identical to S_1 after removing e') with the biggest set missing from the family. This will leave the family union-closed by the same argument as in theorem 2.7.10, and each element will still be in at most a sets, and so here again we have that $f_{n+1}(a) = f_n(a)$.

We now present in table 2.3 a few computations done in order to discover the maximum number of sets in a union-closed family on n elements with each element present in at most a sets of the family and where each element is at least one non-trivial twin difference (i.e. not only the difference between the full set and an $(n-1)$-set).

Remark. To calculate these values, one only needs to add a few constraints to the program $f_n(a)$. First, add $y_S^e \leq x_S$ and $y_S^e \leq x_{S \cup e}$ for all $S \in \mathcal{S}_n, e \in E$ to ensure that y_S^e is positive if and only if both twin sets x_S and $x_{S \cup e}$ are in the family. Then simply enforce that there is at least one pair of non-trivial twin sets per element with $\sum_{\substack{S:e\notin S \\ |S|\leq n-2}} y_S^e \geq 1$ for all $e \in E$.

From Proposition 2.7.21, we know that everything past the main diagonal in Table 2.3 is infeasible. The fact that $f_n(a)$ with the twins constraint decreases as n increases is a bit counterintuitive at first, however it does make sense. If there are more elements, each forced to be a twin difference, then there will be more distinct unions of sets, and the risk of blowing up the a-limit gets greater and greater, and so the number of sets allowable decreases. If we could prove that, then the f-conjecture would be proven, and thus the g-conjecture as well.

Definition 2.7.17. Let $G(\mathcal{F})$ be a graph representing the twins of family $\mathcal{F}$. We let $V(G(\mathcal{F}))$ correspond to any set that is a twin, and $E(G(\mathcal{F}))$ are edges going from a big twin to a small twin.

Theorem 2.7.18. *Any subgraph G' of $G(\mathcal{F})$ restricted to n edges so that every twin difference is present exactly once will be a forest.*

Proof. Suppose there is a cycle. Then there are two ways to go from any vertex on the cycle to any other. This means that there are two ways of going from one set, say S_1, to another, say S_2, by successively adding or removing an element. Suppose without loss of generality that $|S_1| > |S_2|$. Then any element $e \in S_1 \backslash S_2$ will have to be removed in both ways from S_1, and thus there will be two edges representing the twin difference e, a contradiction. Thus there is no cycles, and G' must be a forest. $\square$

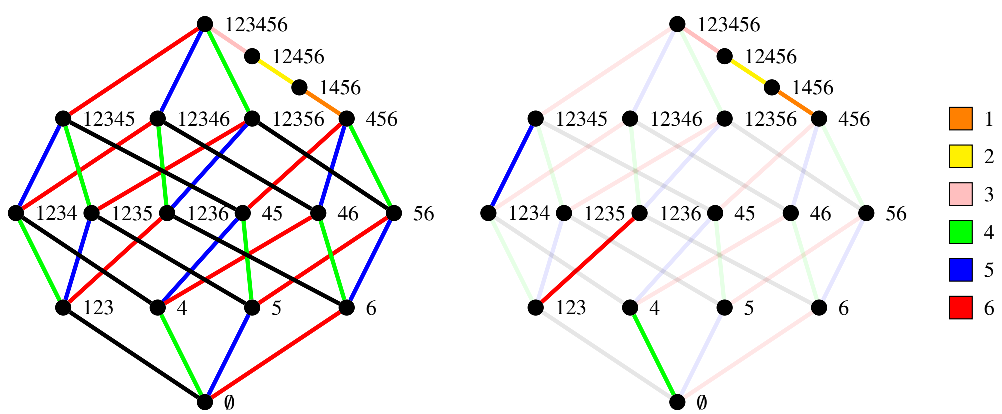

Figure 2.1: Example of $G(\mathcal{F})$ for a union-closed family $\mathcal{F}$ with each element being a twin difference, and its restriction to n edges representing different twin differences.

Theorem 2.7.19. *In an family with a ground set of n elements where each element is in at most a sets of $f_n(a)$, i.e. an optimal family for $f_n(a)$, there exists a set of size less or equal to $\frac{a \cdot n}{2a-b}$ where $a = 2^{\lfloor \log_2 a \rfloor} + b$.*

Proof. We know that in an optimal family for $f_n(a)$, if we add up all the elements in all the sets with repetition, we get at most $a \cdot n$. Moreover, we also know that $f_n(a) \geq 2^{\lfloor \log_2 a \rfloor + 1} + b = 2a - b$ since we know that $f_n(2^{\lfloor \log_2 a \rfloor}) \geq 2^{\lfloor \log_2 a \rfloor + 1}$ and since we also know that $f_n(a)$ is strictly increasing as a increases.

So, the average set in an optimal family for $f_n(a)$ has at most $\frac{a \cdot n}{2a-b}$ elements. This means that there exists a set in such a family with at most $\frac{a \cdot n}{2a-b}$ elements in it. $\square$

Theorem 2.7.20. *We have that $f_n(a) = f_{n+1}(a)$ for all $n > \frac{(a-2)\cdot(2a-b)}{a-b}$ where $a = 2^{\lfloor \log_2 a \rfloor} + b$.*

Proof. Let $\mathcal{F}$ be an optimal family for $f_n(a)$. If there exists an element that is not a twin difference, then $f_n(a) = f_{n-1}(a)$ trivially. So assume that each element is a twin difference.

Take a set S of minimal size in $\mathcal{F}$ that is not the empty set. We know such a set has at most $\frac{a \cdot n}{2a-b}$ elements in it by the previous theorem. Look at $G(\mathcal{F})$ and keep only one twin difference edge per element that is not in S. By theorem 2.7.18, this is a forest and so there must be a least $n - |S| + 1$ vertices induced by these edges. We now want to show that the union of S with the sets corresponding to these vertices are all distinct.

This is indeed true because all vertices are distinct even if we restrict the elements to $[n] \setminus S$. Then if we add S to all of these sets, they remain distinct sets. So there are $n - |S| + 2$ distinct sets that contain S. Every element in S is thus already present in $n - |S| + 2$ sets. Thus, $a \geq n - |S| + 2 \geq n - \frac{a \cdot n}{2a-b} + 2$, otherwise $\mathcal{F}$ could not be an optimal family for $f_n(a)$ since it would not respect the constraint that each element is in at most a sets.

So if $a \leq n - \frac{a \cdot n}{2a-b} + 2$, then the solution was not feasible and so there must have been an element that was not a twin difference meaning that $f_n(a) = f_{n+1}(a)$. In other words, if $n > \frac{(a-2)\cdot(2a-b)}{a-b}$, then $f_n(a) = f_{n+1}(a)$. $\square$

We can do even better than that by using the Falgas-Ravry construction in proposition 2.7.21. We included the previous proof since it can be improved upon by finding better upperbounds for the minimum size of a set in an optimal family and by analysing the lattice better.

2.7.4 Some Results for the New Conjectures

Proposition 2.7.21. *We have that $f_n(a) = f_{n+1}(a)$ for all $n \geq a$.*

Proof. First assume $f_{n-1}(a) < f_n(a)$ for some $n > a$. This means that any optimal solution for n elements contains no blocks since, if there was a block, there would be an element that has no twin. We can thus apply the same construction as in [35]. Order the elements $[n]$ by decreasing frequency. Observe now that for all $1 \leq i < j \leq n$, there exists S_{ij} such that $i \in S_{ij}$ and $j \notin S_{ij}$. For $2 \leq j \leq n$, let $S_j = \cup_{i=1}^{j-1} S_{ij}$, and let $S_{n+1} = [n]$. Note that the S_j's are all distinct since $[1, j-1] \subseteq S_j$ and $j \notin S_j$. Since the most frequent element, element 1, is in at least these n sets of the family, we have that $a \geq n$, which is a contradiction. $\square$

Note that having no block is a much weaker constraint than having each element being a non-trivial twin difference. Therefore, it should be possible to improve this.

Another direction worth investigating is if we could prove that if $f_n(a) < f_{n+1}(a)$, then every element is the difference of an increasingly large number of twins. At some point, this ceases to be possible, and so we would reach a contradiction.

Theorem 2.7.22. *If $f_n(a) < f_{n+1}(a)$, then every element is the difference of at least two pairs of twin sets.*

Proof. Suppose that $f_n(a) = m$ and $f_{n+1}(a) = m + k$ for some $k > 0$. Then we know that any optimal solution for $f_{n+1}(a)$ must be such that every element is a twin difference, otherwise we could remove that element and get an $m+k$ union-closed family spanning n elements such that none is in more than a sets, and so $f_n(a) \geq m + k$, a contradiction.

Now let

$$k' := \min_{e \in [n+1]} |\{S \in \mathcal{F} | e \notin S,\ e \cup S \in \mathcal{F} \text{ and } \mathcal{F} \text{ is an optimal family for } f_{n+1}(a)\}|,$$

i.e. k' is the minimum number of twin pairs for which an element is a twin difference in an optimal solution for $f_{n+1}(a)$.

We now show that $k \leq k'$. Suppose not. Let e' and $\mathcal{F}'$ be an element and a family that attain minimum k'. Remove e' from $\mathcal{F}'$, and remove the k' sets that are now duplicated. Call this new family $\mathcal{F}''$. What remains is a union-closed family of $m+k-k'$ sets on n elements. So $m = f_n(a) \geq m+k-k'$, which implies that $k \leq k'$.

Now suppose that $k = k'$. Notice then that there must exist e'' such that e'' is contained in a sets of $\mathcal{F}'$ that must still be contained in a sets of $\mathcal{F}''$. If not, each element of the new family $\mathcal{F}''$ would be contained in at most $a-1$ sets of $\mathcal{F}''$, and so we would have that $f_n(a-1) \geq m+k-k' = m$, which is a contradiction on the fact that $f_n(a-1) < f_n(a)$.

Thus if $k = k'$, then there exists e'' such that $|\{S \in \mathcal{F}' | S \ni e''\}| = a$ and such that e'' is never contained in a set of the family that does not contain e'. Indeed, if there existed $S' \in \mathcal{F}'$ such that $e' \notin S'$ and $e'' \in S'$, then for any set $S'' \in \mathcal{F}$ such that $S'' \cup e' \in \mathcal{F}$ as well, i.e. twin sets with difference e', then one of $S' \cup S''$ and $S' \cup (S'' \cup e')$ will disappear in $\mathcal{F}''$ and so e'' would be present $a-1$ times, a contradiction.

Thus $\{S \in \mathcal{F}' | S \ni e''\} \subseteq \{S \in \mathcal{F}' | S \ni e'\}$ and since $|\{S \in \mathcal{F}' | S \ni e''\}| = a$, then $|\{S \in \mathcal{F}' | S \ni e'\}| = a$ as well and $\{S \in \mathcal{F}' | S \ni e''\} = \{S \in \mathcal{F}' | S \ni e'\}$ which is a contradiction of the fact that e' and e'' are twin differences for some sets. Since they are copies of each other, we can remove either one of them without creating duplicate sets.

Thus, $k < k'$, and so if $k' = 1$, then $k \leq 0$, and then $f_n(a) \geq f_{n+1}(a)$.

$\square$

Note also that this means that if we remove any element from such a solution, and remove a copy of every duplicated set created, what remains is never an optimal solution for $f_n(a)$.

2.8 Conclusion

We saw that studying the Frankl problem from a polyhedral point of view gave us some new insight into the problem, as was also seen in the case of the Turán problem. Indeed, by modeling the Frankl conjecture as an integer program, we noticed a pattern at the heart of the problem that had remained hidden all these years. This allowed us to formulate two new equivalent conjectures, the f- and g-conjectures, which turned out to have great consequences on the Frankl conjecture. By proving the f- or g-conjecture, $\frac{2}{3}$ of all possible cases of the Frankl conjecture would be solved. Moreover, we would know that there exists an element in any union-closed family that is contained in at least $\frac{6}{13}$ of the sets. This is not the $\frac{1}{2}$ of the Frankl conjecture, but it is certainly not very far, especially considering that no constant fraction is known.

We were able to obtain partial results for the f-conjecture, namely that it holds if $n \geq a$ (the same can be done for g-conjecture). We think it would be possible to prove it for some smaller values of n, and maybe even for all sensible values of n. To do so, studying the structure of union-closed families with twin sets for each elements better, i.e., the families for which there exists two sets for every element such that the only difference between those two sets is that one contains that element and the other doesn't, should be helpful. Such families have a very particular structure and understanding it better might help to prove that the maximum number of sets decreases in such families when n increases for a fixed a.

So here are a few other open problems that would help to make some progress for the new conjectures.

- Show that $f_n(a) = f_{n+1}(a)$ for smaller values of n.
- Show that if $f_n(a) < f_{n+1}(a)$, each element is a twin difference for even more sets.
- Find a constant bound for $f_{n+1}(a) - f_n(a)$.
- Find a better upper bound on the size of a set of minimal cardinality in an optimal family.

Making a much more thorough analysis of the Frankl polytope would also be interesting. Time made fools of us, but we noticed many more interesting facet classes when we looked at complete examples with Porta, and proving that they are always facets would be interesting.

On the flip side, it seems that computers won't be able to make the Frankl conjecture progress much more... that is, until they get much faster or until more theory is developed to help them reduce the size of the problem. As of now, the number of variables and constraints grows too quickly for integer programming solvers after small values of n.

Therefore, focusing on theory for union-closed families with twins gives us the best hope for the future.

Appendix

In the following tables for $f_n(a)$ and $g_n(m)$, we remove unnecessary columns, i.e. the columns for which the values of $f_n(a)$ and $g_n(m)$ are trivial. By that, we mean that we keep the columns for $n \geq \log_2(a)+1$ for $f_n(a)$ and for $n \geq \log_2(m)$ for $g_n(m)$. Indeed, we know that $f_n(a) = 2^n$ for every $n < \log_2(a)+1$ and $g_n(m)$ doesn't exist if $n < \log_2(m)$.

Table 2.4: Values of $f_n(a)$

$a\backslash n$	1	2	3	4	5	6	7	8	9
1	2	2	2	2	2	2	2	2	2
2		4	4	4	4	4	4	4	4
3			5	5	5	5	5	5	5
4			8	8	8	8	8	8	8
5				9	9	9	9	9	9
6				10	10	10	10	10	10
7				12	12	12	12	12	12
8				16	16	16	16	16	16
9					17	17	17	17	17
10					18	18	18	18	18
11					19	19	19	19	19
12					21	21	21	21	21
13					23	23	23	23	23
14					25	25	25	25	25
15					27	27	27	27	27
16					32	32	32	32	32
17						33	33	33	33
18						34	34	34	34
19						35	35	35	35
20						36	36	36	36
21						38	38	38	38
22						40	40	40	40
23						41	41	41	41
24						43	42	42	42
25						45	45	45	45
26						47	47	47	47
27						49	49	49	49
28						52	52	52	52
29						53	53	53	53
30						56	56	56	56
31						58	58	58	58
32						64	64	64	64

$a\backslash n$	7	8	9
33	65	65	65
34	66	66	66
35	67	67	67
36	68	68	68
37	69	69	69
38	71	71	71
39	72	72	72
40	74	74	74
41	75	75	75
42	77	77	77
43	79	79	79
44	80	80	80
45	82	82	82
46	83	83	83
47	85	85	85
48	88	88	88
49	89	89	89
50	91	91	91
51	93	93	93
52	95	95	95
53	98	98	98
54	99	99	99
55	101	101	101
56	104	104	104
57	105	105	105
58	108	108	108
59	110	110	110
60	113	113	113
61	115	115	115
62	118	118	118
63	121	121	121
64	128	128	128

$a\backslash n$	8	9
65	129	129
66	130	130
67	131	131
68	132	132
69	133	133
70	134	134
71	136	136
72	137	137
73	139	139
74	140	140
75	142	142
76	144	144
77	145	145
78	146	146
79	147	147
80	149	149
81	150	150
82	152	152
83	154	154
84	156	156
85	157	157
86	158	158
87	160	160
88	162	162
89	164	164
90	166	166
91	168	168
92	170	170
93	171	171
94	173	173
95	175	175
96	176	176

$a\backslash n$	8	9
97	179	179
98	180	180
99	182	182
100	184	184
101	186	186
102	188	188
103	189	189
104	192	192
105	194	194
106	196	196
107	198	198
108	200	200
109	202	202
110	204	204
111	206	206
112	209	209
113	211	211
114	214	214
115	216	216
116	220	220
117	221	221
118	224	224
119	226	226
120	229	229
121	231	231
122	233	233
123	236	236
124	240	240
125	242	242
126	245	245
127	248	248
128	256	256

Table 2.5: Values of $g_n(m)$

$m\backslash n$	1	2	3	4	5	6	7	8
2	1	1	1	1	1	1	1	1
3		2	2	2	2	2	2	2
4		2	2	2	2	2	2	2
5			3	3	3	3	3	3
6			4	4	4	4	4	4
7			4	4	4	4	4	4
8			4	4	4	4	4	4
9				5	5	5	5	5
10				6	6	6	6	6
11				7	7	7	7	7
12				7	7	7	7	7
13				8	8	8	8	8
14				8	8	8	8	8
15				8	8	8	8	8
16				8	8	8	8	8
17					9	9	9	9
18					10	10	10	10
19					11	11	11	11
20					12	12	12	12
21					12	12	12	12
22					13	13	13	13
23					13	13	13	13
24					14	14	14	14
25					14	14	14	14
26					15	15	15	15
27					15	15	15	15
28					16	16	16	16
29					16	16	16	16
30					16	16	16	16
31					16	16	16	16
32					16	16	16	16

$m\backslash n$	6	7	8
33	17	17	17
34	18	18	18
35	19	19	19
36	20	20	20
37	21	21	21
38	21	21	21
39	22	22	22
40	22	22	22
41	23	23	23
42	24	24	24
43	24	24	24
44	25	25	25
45	25	25	25
46	26	26	26
47	26	26	26
48	27	27	27
49	27	27	27
50	28	28	28
51	28	28	28
52	28	28	28
53	29	29	29
54	30	30	30
55	30	30	30
56	30	30	30
57	31	31	31
58	31	31	31
59	32	32	32
60	32	32	32
61	32	32	32
62	32	32	32
63	32	32	32
64	32	32	32

$m\backslash n$	7	8
65	33	33
66	34	34
67	35	35
68	36	36
69	37	37
70	38	38
71	38	38
72	39	39
73	40	40
74	40	40
75	41	41
76	42	42
77	42	42
78	43	43
79	43	43
80	44	44
81	45	45
82	45	45
83	46	46
84	47	47
85	47	47
86	48	48
87	48	48
88	48	48
89	49	49
90	50	50
91	50	50
92	51	51
93	51	51
94	52	52
95	52	52
96	53	53

$m\backslash n$	7	8
97	53	53
98	53	53
99	54	54
100	55	55
101	55	55
102	56	56
103	56	56
104	56	56
105	57	57
106	58	58
107	58	58
108	58	58
109	59	59
110	59	59
111	60	60
112	60	60
113	60	60
114	61	61
115	61	61
116	62	62
117	62	62
118	62	62
119	63	63
120	63	63
121	63	63
122	64	64
123	64	64
124	64	64
125	64	64
126	64	64
127	64	64
128	64	64

Table 2.6: Values of the linear relaxation of $f_n(a)$ with singletons and doubletons fixed to be zero

$a\backslash n$	1	2	3	4	5	6	7	8	9
1	1	1	2	2	2	2	2	2	2
2		1	2	3	3	4	4	4	5
3			2	4	5	6	6	7	8
4			2	6	7	8	9	10	11
5				6	8	10	11	12	14
6				6	9	12	13	15	17
7				6	11	13	16	18	20
8				6	12	15	18	20	23
9					14	16	20	23	26
10					15	18	22	26	29
11					17	19	23	28	32
12					17	21	25	31	35
13					17	22	27	33	38
14					17	24	29	35	41
15					17	25	30	37	44
16					17	27	32	39	46
17						29	34	41	48
18						30	36	43	50
19						32	37	45	53
20						33	39	47	55
21						35	41	49	57
22						36	42	51	59
23						38	44	53	62
24						39	46	55	64
25						41	47	57	66
26						43	49	59	68
27						43	50	61	71
28						43	52	63	73
29						43	54	65	75
30						43	55	66	77
31						43	57	68	80
32						43	59	70	82

$a\backslash n$	7	8	9
33	60	71	84
34	62	73	86
35	64	75	89
36	65	76	91
37	67	78	93
38	68	80	95
39	70	82	98
40	72	83	100
41	73	85	102
42	75	87	104
43	77	88	107
44	78	90	108
45	80	92	110
46	82	93	112
47	83	95	114
48	85	97	116
49	86	99	117
50	88	100	119
51	90	102	121
52	91	104	123
53	93	105	125
54	95	107	126
55	96	109	128
56	98	110	130
57	100	112	132
58	100	114	134
59	100	116	135
60	100	117	137
61	100	119	139
62	100	121	141
63	100	122	143
64	100	124	144

$a\backslash n$	8	9
65	126	146
66	128	148
67	129	150
68	131	152
69	133	153
70	134	155
71	136	157
72	138	159
73	139	161
74	141	162
75	143	164
76	145	166
77	146	168
78	148	170
79	150	171
80	151	173
81	153	175
82	155	177
83	156	178
84	158	180
85	160	182
86	162	184
87	163	185
88	165	187
89	167	189
90	168	191
91	170	192
92	172	194
93	174	196
94	175	198
95	177	199
96	179	201

$a\backslash n$	8	9
97	180	203
98	182	205
99	184	206
100	185	208
101	187	210
102	189	212
103	191	213
104	192	215
105	194	217
106	196	219
107	197	220
108	199	222
109	201	224
110	202	226
111	204	227
112	206	229
113	208	231
114	209	233
115	211	235
116	213	236
117	214	238
118	216	240
119	218	242
120	220	243
121	220	245
122	220	247
123	220	249
124	220	250
125	220	252
126	220	254
127	220	256
128	220	257

$a\backslash n$	9
129	259
130	261
131	263
132	264
133	266
134	268
135	270
136	271
137	273
138	275
139	277
140	278
141	280
142	282
143	284
144	285
145	287
146	289
147	291
148	293
149	294
150	296
151	298
152	300
153	301
154	303
155	305
156	307
157	308
158	310
159	312
160	314

$a\backslash n$	9
161	315
162	317
163	319
164	321
165	322
166	324
167	326
168	328
169	329
170	331
171	333
172	335
173	336
174	338
175	340
176	342
177	343
178	345
179	437
180	349
181	351
182	352
183	354
184	356
185	358
186	359
187	361
188	363
189	365
190	366
191	368
192	370

$a\backslash n$	9
193	372
194	373
195	375
196	377
197	379
198	380
199	382
200	384
201	386
202	387
203	389
204	391
205	393
206	394
207	396
208	398
209	400
210	401
211	403
212	405
213	407
214	409
215	410
216	412
217	414
218	416
219	417
220	419
221	421
222	423
223	424
224	426

$a\backslash n$	9
225	428
226	430
227	431
228	433
229	435
230	437
231	438
232	440
233	442
234	444
235	445
236	447
237	449
238	451
239	452
240	454
241	456
242	458
243	459
244	461
245	463
246	465
247	467
248	467
249	467
250	467
251	467
252	467
253	467
254	467
255	467
256	467

Note that the values were rounded down to the previous integer. In red are the values for which this relaxation is not sufficient to show that the Frankl conjecture holds for these parameters a and n.

Bibliography

[1] T. Abe. Strong semimodular lattices and Frankl's conjecture. *Algebra Universalis*, 44:379–382, 2000.

[2] T. Abe and B. Nakano. Frankl's conjecture is true for modular lattices. *Graphs and Combinatorics*, 14:305–311, 1998.

[3] T. Abe and B. Nakano. Lower semimodular types of lattices: Frankl's conjecture holds for lower quasi-semimodular lattices. *Graphs and Combinatorics*, 16:1–16, 2000.

[4] N. Alon, A. Krech, and T. Szabó. Turán's theorem in the hypercube. *SIAM Journal on Discrete Mathematics*, 21:66–72, 2007.

[5] I. Balla. Minimum densities of union-closed families. e-print, 2011.

[6] I. Balla, B. Bollobás, and T. Eccles. Union-closed families of sets. *Journal of Combinatorial Theory (Series A)*, 120(3):531–544, 2013.

[7] J. Balogh. The Turán density of triple systems is not principal. *Journal of Combinatorial Theory (Series A)*, 100:176–180, 2002.

[8] J. Balogh, B. Bollobás, and M. Simonovits. On the number of graphs without forbidden subgraph. *Journal of Combinatorial Theory (Series B)*, 91:1–24, 2004.

[9] J. Balogh, B. Bollobás, and M. Simonovits. The typical structure of graphs without given excluded subgraphs. *Random Structures Algorithms*, 2014. to appear.

[10] J. Balogh and W. Samotij. The number of K_{st}-free graphs. *Journal of London Mathematical Society*, 2014. to appear.

[11] T. Bohman and P. Keevash. The early evolution of the H-free process. *Inventiones Mathematicae*, 181:291–336, 2010.

[12] B. Bollobás. *Extremal Graph Theory*. Dover Publications, Inc., Mineola, New York, 2004. reprint of the 1978 edition published by Academic Press, London.

[13] I. Bošnjak and P. Marković. The 11-element case of Frankl's conjecture. *The Electronic Journal of Combinatorics*, 15(1):R88, 2008.

[14] E. D. Boyer, D. L. Kreher, S. P. Radziszowski, and A. Sidorenko. On $(n,5,3)$-Turán systems. *Ars Combinatorica*, 37:13–31, 1994.

[15] W. G. Brown. On an open problem of Paul Turán concerning 3-graphs. *Studies in Pure Mathematics*, pages 91–93, 1983.

[16] H. Bruhn, P. Charbit, O. Schaudt, and J. A. Telle. The graph formulation of the union-closed sets conjecture. preprint, 2013.

[17] H. Bruhn and O. Schaudt. The union-closed sets conjecture almost holds for almost all random bipartite graphs. preprint, 2012.

[18] H. Bruhn and O. Schaudt. The journey of the union-closed sets conjecture. preprint, 2014.

[19] T. Christof and A. Löbel. Porta: polyhedron representation transformation algorithm (version 1.4.1).

[20] F. Chung and L. Lu. An upper bound for the Turán number $t_3(n,4)$. *Journal of Combinatorial Theory (Series A)*, 87:381–389, 1999.

[21] G. Czedli. On averaging Frankl's conjecture for large union-closed sets. *Journal of Combinatorial Theory (Series A)*, 116(3):724–729, 2009.

[22] G. Czedli, M. Maróti, and E. Schmidt. On the scope of averaging for Frankl's conjecture. *Order*, 26:31–48, 2009.

[23] D. de Caen. Extension of a theorem of Moon and Moser on complete subgraphs. *Ars Cominatorica*, 16:5–10, 1983.

[24] D. de Caen. On upper bounds for 3-graphs without tetrahedra. *Congressus Numerantium*, 63:193–202, 1988.

[25] D. de Caen and Z. Füredi. The maximum size of 3-uniform hypergraphs not containing a Fano plane. *Journal of Combinatorial Theory (Series B)*, 78:274–276, 2000.

[26] D. de Caen, D. L. Kreher, and J. Wiseman. On constructive upper bounds for the Turán numbers $T(n,2r+1,2r)$. In C. Numerantium, editor, *Nineteeth Southeastern Conference on Combinatorics, Graph Theory, and Computing (Baton Rouge, LA, 1988)*, pages 277–280, 1988.

[27] I. Dinur, V. Guruswami, S. Khot, and O. Regev. A new multilayered pcp and the hardness of hypergraph vertex cover. *SIAM Journal on Computing*, 34(5):1129–1146, 2005.

[28] K. Dohmen. A new perspective on the union-closed sets conjecture. *Ars Combinatorica*, 58:183–185, 2001.

[29] F. Eisenbrand, G. Oriolo, G. Stauffer, and P. Ventura. The stable set polytope of quasi-line graphs. *Combinatorica*, 28:45–67, 2008.

[30] M. El-Zahar. A graph-theoretic version of the union-closed sets conjecture. *Journal of Graph Theory*, 26:155–163, 1997.

[31] P. Erdös, P. Frankl, and V. Rödl. The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent. *Graphs and Combinatorics*, 2:113–121, 1986.

[32] P. Erdös and M. Simonovits. A limit theorem in graph theory. *Studia Scientiarum Mathematicarum Hungarica*, 1:51–57, 1966.

[33] P. Erdös and V. Sós. On Ramsey-Turán type theorems for hypergraphs. *Combinatorica*, 2:289–295, 1982.

[34] P. Erdös and A. H. Stone. On the structure of linear graphs. *Bulletin of the American Mathematical Society*, 52:1087–1091, 1946.

[35] V. Falgas-Ravry. Minimal weight in union-closed families. *Electronic Journal of Combinatorics*, 18(1):P95, 2011.

[36] G. L. Faro. A note on the union-closed conjecture. *Journal of the Australian Mathematical Society (Series A)*, 57:230–236, 1994.

[37] G. L. Faro. Union-closed sets conjecture: improved bounds. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 16:97–102, 1994.

[38] P. Frankl. On the trace of finite sets. *Journal of Combinatorial Theory (Series A)*, 34:41–45, 1983.

[39] P. Frankl. Asymptotic solution of a Turán-type problem. *Graphs and Combinatorics*, 6:223–227, 1990.

[40] P. Frankl and Z. Füredi. An exact result for 3-graphs. *Discrete Mathematics*, 50:323–328, 1984.

[41] P. Frankl and Z. Füredi. Exact solution of some Turán-type problems. *Journal of Combinatorial Theory (Series A)*, 45:226–262, 1987.

[42] P. Frankl and V. Rödl. Lower bounds for Turán's problem. *Graphs and Combinatorics*, 1:213–216, 1985.

[43] A. Frohmader. More constructions for Turán's $(3,4)$-conjecture. *Electronic Journal of Combinatorics*, 15:R137, 2008.

[44] Z. Füredi. Graphs without quadrilaterals. *Journal of Combinatorial Theory Series B*, 34:187–190, 1983.

[45] Z. Füredi. Turán type problems. *Surveys in combinatorics*, pages 253–300, 1991.

[46] Z. Füredi. On the number of edges of quadrilateral-free graphs. *Journal of Combinatorial Theory Series B*, 68:1–6, 1996.

[47] Z. Füredi and A. Kündgen. Turán problems for integer-weighted graphs. *Journal of Graph Theory*, 40:195–225, 2002.

[48] Z. Füredi and M. Simonovits. Triple systems not containing a Fano configuration. *Combinatorics Probability and Computing*, 14:467–484, 2005.

[49] W. Gao and H. Yu. Note on the union-closed sets conjecture. *Ars Combinatorica*, 49:280–288, 1998.

[50] A. Ghouila-Houri. Caractérisation des graphes non orientés dont on peut orienter les arrêtes de manière à obtenir le graphe d'une relation d'ordre. *Les Comptes rendus de l'Académie des sciences*, 254:1370–1371, 1962.

[51] R. Johnson and T. Vaughan. On union-closed families. *Journal of Combinatorial Theory (Series A)*, 85:112–119, 1999.

[52] G. Kalai. A new approach to Turán's conjecture. *Graphs and Combinatorics*, 1:107–109, 1985.

[53] G. Katona, T. Nemetz, and M. Simonovits. On a problem of Turán in the theory of graphs. *Matematikai Lapok*, 15:228–238, 1964.

[54] P. Keevash. The Turán problem for hypergraphs of fixed size. *Electronic Journal of Combinatorics*, 12:N11, 2005.

[55] P. Keevash. A hypergraph regularity method for generalised Turán problems. *Random Structures Algorithms*, 34:123–164, 2009.

[56] P. Keevash. Hypergraph Turán problems. *Surveys in Combinatorics*, pages 83–140, 2011.

[57] P. Keevash. The Turán number of F_{33}. *Combinatorics, Probability and Computing*, 21:451–456, 2012.

[58] P. Keevash, D. Mubayi, B. Sudakov, and J. Verstraete. Rainbow Turán problems. *Combinatorics, Probability and Computing*, 16:109–126, 2007.

[59] P. Keevash, M. Saks, B. Sudakov, and J. Verstraete. Multicolour Turán problems. *Advances in Applied Mathematics*, 33:238–262, 2004.

[60] P. Keevash and B. Sudakov. On a hypergraphs Turán problem of Frankl. *Combinatorica*, 25:673–706, 2005.

[61] P. Keevash and B. Sudakov. The Turán number of the Fano plane. *Combinatorica*, 25:561–574, 2005.

[62] E. Knill. Graph generated union-closed families of sets. e-print, 1994.

[63] A. Kostochka. A class of constructions for Turán's $(3,4)$ problem. *Combinatorica*, 2:187–192, 1982.

[64] U. Leck, I. Roberts, and J. Simpson. Minimizing the weight of the union-closure of families of two-sets. *Australasian Journal of Combinatorics*, 52:67–73, 2012.

[65] B. Llano, J. Montellano-Ballesteros, E. Rivera-Campo, and R. Strausz. On conjectures of Frankl and El-Zahar. *Journal of Graph Theory*, 57:344–352, 2008.

[66] L. Lu and Y. Zhao. An exact result for hypergraphs and upper bounds for the Turán density of K_{r+1}^r. *SIAM Journal of Discrete Mathematics*, 23:1324–1334, 2009.

[67] F. Marić, B. V. cković, and M. Živković. Formalizing Frankl's conjecture: *FC*-families. In J. Jeuring, J. Campbell, J. Carette, G. Reis, P. Sojka, M. Wenzel, and V. Sorge, editors, *Intelligent Computer Mathematics*, volume 7362 of *Lecture Notes in Computer Science*, pages 248–263. Springer Berlin Heidelberg, 2012.

[68] K. Markström. Extremal graphs and bounds for the Turán density of the 4-uniform K_5. *Discrete Mathematics*, 309:5231–5234, 2009.

[69] R. Morris. FC-families, and improved bounds for Frankl's conjecture. *European Journal of Combinatorics*, 27:269–282, 2006.

[70] T. Motzkin and E. Straus. Maxima for graphs and a new proof of a theorem of Turán. *Canadian Journal of Mathematics*, 17:533–540, 1965.

[71] D. Mubayi. On hypergraphs with every four points spanning at most two triples. *Electronic Journal of Combinatorics*, 10:N10, 2003.

[72] D. Mubayi. A hypergraph extension of Turán's theorem. *Journal of Combinatorial Theory (Series B)*, 96:122–134, 2006.

[73] D. Mubayi and O. Pikhurko. A new generalization of Mantel's theorem to k-graphs. *Journal of Combinatorial Theory (Series B)*, 97:669–678, 2007.

[74] D. Mubayi and V. Rödl. On the Turán number of triple systems. *Journal of Combinatorial Theory (Series A)*, 100:136–152, 2002.

[75] R. Norton and D. Sarvate. A note on the union-closed sets conjecture. *Journal of Australian Mathematical Society (Series A)*, 55:411–413, 1993.

[76] A. Pêcher and A. Wagler. Almost all webs are not rank-perfect. *Mathematical Programming*, 105(2-3):311–328, 2006.

[77] O. Pikhurko. An exact Turán result for the generalized triangle. *Combinatorica*, 28:187–208, 2008.

[78] B. Poonen. Union-closed families. *Journal of Combinatorial Theory (Series A)*, 59(1):253–268, 1992.

[79] A. Razborov. Flag algebras. *Journal of Symbolic Logic*, 72:1239–1282, 2007.

[80] A. Razborov. On 3-hypergraphs with forbidden 4-vertex configurations. *SIAM Journal of Discrete Mathematics*, 24:946–963, 2010.

[81] A. Razborov. On the Fon-der-Flaass interpretation of extremal examples for Turán's $(3,4)$-problem. e-print, 2010.

[82] D. Reimer. An average set size theorem. *Combinatorics, probability and computing*, 12(1):89–93, 2003.

[83] J. Reinhold. Frankl's conjecture is true for lower semimodular lattices. *Graphs and Combinatorics*, 16:115–116, 2000.

[84] J.-C. Renaud. Is the union-closed sets conjecture the best possible? *Journal of the Australian Mathematical Society (Series A)*, 51:276–283, 1991.

[85] J.-C. Renaud. A second approximation to the boundary function on union-closed collections. *Ars Combinatorica*, 41:177–188, 1995.

[86] J.-C. Renaud and L. Fitina. On union-closed sets and Conway's sequence. *Bulleting of the Australian Mathematical Society*, 47:321–332, 1993.

[87] I. Roberts. The union-closed conjecture. Technical Report Technical Report 2/92, Curtin University of Technology, 1992.

[88] I. Roberts and J. Simpson. A note on the union-closed families of finite sets. *Australasian Journal of Combinatorics*, 47:129–145, 2010.

[89] D. Sarvate and J.-C. Renaud. On the union-closed sets conjecture. *Ars Combinatoria*, 27:149–154, 1989.

[90] D. Sarvate and J.-C. Renaud. Improved bounds for the union-closed sets conjecture. *Ars Combinatoria*, 29:181–185, 1990.

[91] A. Schrijver. *Combinatorial Optimization - Polyhedra and Efficiency*. Springer-Verlag, Berlin, 2003.

[92] SCIP Optimization Suite. Constraint and mixed integer programming solver and modelling language, 2013. Free for academic use, download at scip.zib.de.

[93] A. Sidorenko. Method of quadratic forms in the Turán combinatorial problem. *Moscow University Mathematics Bulletin*, 37:3–6, 1982.

[94] A. Sidorenko. The maximal number of edges in a homogeneous hypergraph containing no prohibited subgraphs. *Math Notes*, 41:247–259, 1987.

[95] A. Sidorenko. Asymptotic solution for a new class of forbidden r-graphs. *Combinatorica*, 9:207–215, 1989.

[96] A. Sidorenko. Asymptotic solution of the Turán problem for some hypergraphs. *Graphs and Combinatorics*, 8:199–201, 1992.

[97] A. Sidorenko. What we know and what we do not know about Turán numbers. *Graph and Combinatorics*, 11:179–199, 1995.

[98] A. Sidorenko. Upper bounds for Turán numbers. *Journal of Combinatorial Theory (Series A)*, 77:133–147, 1997.

[99] P. Turán. On an extremal problem in graph theory (in hungarian). *Matematikai és Fizikai Lapok*, 48:436–452, 1941.

[100] T. Vaughan. Families implying the Frankl conjecture. *European Journal of Combinatorics*, 23:851–860, 2002.

[101] T. Vaughan. Three-sets in a union-closed family. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 49:73–84, 2004.

[102] B. Vučković and M. Živković. The 12 element case of Frankl's conjecture. preprint, 2012.

[103] P. Wójcik. Density of union-closed families. *Discrete Mathematics*, 105:259–267, 1992.

[104] P. Wójcik. Union-closed families of sets. *Discrete Mathematics*, 199:173–182, 1999.