Charles University in Prague
Faculty of Mathematics and Physics

MASTER THESIS

Lukáš Mach

Extremal properties of hypergraphs

Extremální vlastnosti hypergrafů

Department of Applied Mathematics

Supervisor of the master thesis: Doc. RNDr. Daniel Kráľ, Ph.D.
Study programme: Computer Science
Specialization: Discrete Models and Algorithms

Prague 2011

I would like to thank Daniel Kráľ for his patience and support. He has been a great advisor and I am grateful for all the help he provided. My thanks also belong to my parents for their love and encouragement during my studies.

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In Prague date 2. 8. 2011 Lukáš Mach

Název práce: Extremální vlastnosti hypergrafů

Autor: Lukáš Mach

Katedra: Katedra aplikované matematiky

Vedoucí diplomové práce: Doc. RNDr. Daniel Kráľ, Ph.D., Katedra aplikované matematiky

Abstrakt: V této práci podáme přehled o některých nedávných výsledcích o skocích v hypergrafech v oblasti extermální kombinatoriky. Číslo $\alpha \in [0,1)$ je skok pro r, pokud pro každé $\epsilon > 0$ a každé celé číslo $m \geq r$ jakýkoliv r-graf na $N > N(\epsilon, m)$ vrcholech a s alespoň $(\alpha + \epsilon)\binom{N}{r}$ hranami obsahuje podgraf na m vrcholech s alespoň $(\alpha + c)\binom{m}{r}$ hranami, kde $c := c(\alpha)$ závisí pouze na α. Baber a Talbot [1] nedávno ukázali první příklad existence skoku pro $r = 3$ v intervalu $[2/9, 1)$. Jejich výsledek používá kalkul flag algeber [10], který vede k rešení problému semidefinitní optimalizace. Součástí práce je softwarová implementace jejich metody.

Klíčová slova: extremální kombinatorika, flag algebra, skoky v hypergrafech

Title: Extremal properties of hypergraphs

Author: Lukáš Mach

Department: Department of Applied Mathematics

Supervisor: Doc. RNDr. Daniel Kráľ, Ph.D., Department of Applied Mathematics

Abstract: We give an overview of recent progress in the research of hypergraph jumps – a problem from extremal combinatorics. The number $\alpha \in [0,1)$ is a jump for r if for any $\epsilon > 0$ and any integer $m \geq r$ any r-graph with $N > N(\epsilon, m)$ vertices and at least $(\alpha + \epsilon)\binom{N}{r}$ edges contains a subgraph with m vertices and at least $(\alpha + c)\binom{m}{r}$ edges, where $c := c(\alpha)$ does depend only on α. Baber and Talbot [1] recently gave first examples of jumps for $r = 3$ in the interval $[2/9, 1)$. Their result uses the framework of flag algebras [10] and involves solving a semidefinite optimization problem. A software implementation of their method is a part of this work.

Keywords: extremal combinatorics, flag algebras, hypergraph jumps

Contents

Introduction **6**

1 Flag algebras **7**
- 1.1 Basic definitions 7
- 1.2 Operators 10
  - 1.2.1 Flag multiplication 10
  - 1.2.2 Unlabelling – the downward operator 12
  - 1.2.3 Upward operator 13
- 1.3 Semantics 14
  - 1.3.1 Graph limits 15

2 Estimating Turán densities **18**
- 2.1 Turán's theorem 18
- 2.2 Bounding Turán densities using flag algebras 21
- 2.3 Example – a bound on $\pi(\{K_4^-\})$ 23
- 2.4 Computational aspects 23

3 Hypergraph jumps and Lagrangians **27**
- 3.1 The Lagrange function of a hypergraph 27
- 3.2 Hypergraph jumps 29
- 3.3 The existence of jumps for 3-uniform hypergraphs 31

Conclusion **33**

A Semidefinite programming and the CSDP library **34**

Introduction

In this thesis, we present an overview of some recent developments in extremal combinatorics, particularly of results concerning hypergraph jumps.

We begin with an exposition of the framework of *flag algebras* developed by Razborov [10], building on previous work by Lovász and Szegedy [8] on graph limits. This calculus allows us to infer relations between limit densities of fixed subgraphs by providing an algebraic formalization of arguments typical in the field, such as averaging over all choices of some parameter. Applying this framework has several advantages. Its concise nature allows for very short proofs of many classical theorems. This is illustrated in Section 2.1, where we give a proof of Turán's theorem by a short computation in flag algebras $\mathcal{A}^0$ and $\mathcal{A}^1$. Furthermore, the formalism allows for a computerized search for the relations useful in proving conjectures or improving existing bounds. The latter property plays a key role in Section 2.2, where we discuss a technique used by [1,11] to bound the Turán density of a general finite family of r-uniform hypergraphs.

In Chapter 3, we discuss Erdős' question about hypergraph jumps, which can be roughly stated as follows: for what values of α are we guaranteed to find an r-uniform subgraph of density $\alpha + c$ (where c depends only on α and r) in every sufficiently large r-uniform graph with the density of *slightly more* than α? (The reader is referred to Definition 3.2.1 for a precise statement.) Both positive and negative results can be obtained using a characterization developed by Frankl and Rödl [7], discussed in Section 3.2. Their theorem reduces the question to finding a family of r-uniform graphs whose Turán density satisfies certain relation. Baber and Talbot [1] applied flag algebras to give sufficiently good bounds on this density and thus proved the existence of non-trivial jumps for $r = 3$. Their result, based on solving a semidefinite program of considerable size using a computer, is discussed in Section 3.3. We provide a software implementation of their method. The source code for this program is available on the attached DVD.

Chapter 1

Flag algebras

In this section, we review a framework developed by Razborov in [10], which formalizes certain typical arguments used in asymptotic extremal combinatorics. In line with [10], we present the definitions in their general forms using the language of finite model theory while also supplementing this with specific examples.

1.1 Basic definitions

We fix a universal first-order theory with equality and infinite models in a language containing only predicate symbols. Specifically, in this thesis, we are interested in the theory T_G of all undirected graphs and the theories T_r of r-uniform hypergraphs (with predicates representing edges). The ground set of a model M from this theory is denoted $V(M)$. This notation corresponds with the fact that our models are graphs and their ground sets are their vertex sets.

For $U \subseteq V(M)$ we define $M|_U$ to be the submodel induced by U. We let $M - U := M|_{V(M)\setminus U}$. A model embedding $\alpha : M \to N$ is an injective mapping $\alpha : V(M) \to V(N)$ inducing an isomorphism between M and $N|_{\mathrm{im}(\alpha)}$. The existence of an isomorphism between the models M and N is denoted by $M \approx N$ and $\mathcal{M}_n$ represents the set of all models (up to isomorphism) with a ground set of size n. Thus, in the theory T_G, $\mathcal{M}_n$ represents all non-isomorphic graphs of size n.

Definition 1.1.1. *A* ***type*** *σ is a model with ground set $[k] := \{1, \ldots, k\}$ for some non-negative integer k. A* ***σ-flag*** *is a pair (M, θ) where M is a model and $\theta : \sigma \to M$ is a model embedding. A* ***flag embedding*** *$\alpha : F \to F'$ where $F = (M, \theta)$ and $F' = (M', \theta')$ are σ-flags, is a model embedding such that $\theta' = \alpha\theta$. Two flags are* ***isomorphic*** *if there is a one-to-one flag embedding between them.*

The set of all σ-flags is denoted by $\mathcal{F}^\sigma$ and its restriction to flags of size l by $\mathcal{F}^\sigma_l$. In the context of our theory T_G, a type is a graph with vertices labeled by $[k]$, and a flag is a graph with several vertices distinguished using labels $[k]$.

We call the collection $V_1, \dots, V_t$ of finite sets satisfying the condition $V_i \cap V_j = C$ for all distinct $i, j \in [k]$ a *sunflower with the center C and petals $V_1, \dots, V_t$*. We consider the probability space of induced submodels of a fixed (large) model:

Definition 1.1.2. *For $M \in \mathcal{M}_l$ and $N \in \mathcal{M}_L$ with $l \leq L$ we let $p(M, N)$ be the probability that a uniformly randomly chosen subset U of $V(N)$ of size l satisfies $M \approx N|_U$.*

Definition 1.1.3. *For a fixed type σ of size k and integers $l, l_1, \dots, l_t \geq k$ satisfying*

$$l_1 + \ldots + l_t - k(t-1) \leq l,$$

and $F = (M, \theta) \in \mathcal{F}_l^\sigma$, $F_i \in \mathcal{F}_{l_i}^\sigma$, we denote by $p(F_1, \dots, F_t; F)$ the probability that uniformly randomly generated sunflower with center $im(\theta)$ and petals V_i of sizes l_i satisfies $F|_{V_i} \approx F_i$ for all $i \in [t]$. (For $t = 1$ this coincides with the previous definition.)

The function p is the key quantity of our interest and the basis for the definition of flag algebras. We first give a simple observation [10, Lemma 2.2]:

Lemma 1.1.4 (Chain rule). *Consider σ-flags $F_i \in \mathcal{F}_{l_i}^\sigma$ for $i \in [t]$. Take $s \in [t]$, a flag $F \in \mathcal{F}_l^\sigma$ and $\tilde{l} \leq l$ such that:*

$$\tilde{l} + l_{s+1} + \cdots + l_t - |\sigma|(t-s) \leq l,$$
$$l_1 + \cdots + l_s - |\sigma|(s-1) \leq \tilde{l}.$$

Then

$$p(F_1, \dots, F_t; F) = \sum_{\tilde{F} \in \mathcal{F}_{\tilde{l}}^\sigma} p(F_1, \dots, F_s; \tilde{F}) p(\tilde{F}, F_{s+1}, \dots, F_t; F).$$

Proof. Let $F = (M, \theta)$. Consider another way of uniformly generating a random sunflower $(\boldsymbol{V_1}, \dots, \boldsymbol{V_t})$:

First generate uniformly and randomly the sunflower $(\tilde{\boldsymbol{V}}, \boldsymbol{V_{s+1}}, \dots, \boldsymbol{V_t})$ with the center $\mathrm{im}(\theta)$ and petals of sizes $\tilde{l}, l_{s+1}, \dots, l_t$. Then, within $\tilde{\boldsymbol{V}}$ generate uniformly and randomly a sunflower $(\boldsymbol{V_1}, \dots, \boldsymbol{V_s})$ with center $\mathrm{im}(\theta)$ and petals of sizes $l_1, \dots, l_s$. This also leads to a uniform distribution and the identity we are proving becomes the formula of total probability with the right-hand side corresponding to the partition of the probability space according to the isomorphism type of $F|_{\tilde{\boldsymbol{V}}}$. □

The following lemma motivates us to consider the vector space of all linear combinations of flags as a basis of our framework since it shows that, asymptotically, the function $p(F_1, \dots, F_t; F)$ is multiplicative in the parameters $F_1, \dots, F_t$.

Lemma 1.1.5. *[10, Lemma 2.3] Let $F_i \in \mathcal{F}^\sigma_{l_i}$ for $i \in [t]$ and $F \in \mathcal{F}^\sigma_l$. Then*

$$\Big|p(F_1, \dots, F_t; F) - p(F_1; F) \cdot \ldots \cdot p(F_t; F)\Big| \le \frac{(l_1 + \ldots + l_t)^{O(1)}}{l}. \tag{1.1}$$

Proof. Note that the difference between the two expressions corresponds to picking the elements differently. In the first expression, the randomly chosen sets forming the petals of the sunflower can't overlap, while the second case allows this.

Let $F = (M, \theta)$. Choose uniformly, randomly and independently the sets $\boldsymbol{V_i} \subseteq V(M)$ with sizes l_i satisfying $\mathrm{im}(\theta) \subseteq \boldsymbol{V_i}$ for all $i \in [t]$. Consider the following two events:

$$A \ldots \forall i \in [t](\mathcal{F}^\sigma|_{\boldsymbol{V_i}} \approx \mathcal{F}_i),$$
$$B \ldots \text{the sets } \boldsymbol{V_i} \setminus \mathrm{im}(\theta) \text{ are disjoint.}$$

The left-hand side of the inequality (1.1) now becomes $\big|P[A|B] - P[A]\big|$. This can be bounded from above by:

$$1 - P[B] \le \sum_{i \neq j} P[\mathrm{im}(\theta) \subsetneq (\boldsymbol{V_i} \cap \boldsymbol{V_j})] \le \frac{(l_1 + \ldots + l_t)^{O(1)}}{l},$$

and the lemma follows. □

Let $\mathbb{R}\mathcal{F}^\sigma$ be the space of all formal finite linear combinations of σ-flags and let $\mathcal{K}^\sigma$ be its linear subspace generated by the elements of form

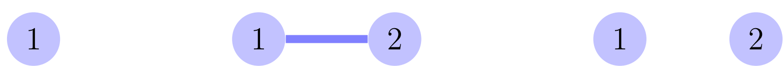
$$\tilde{F} - \sum_{F \in \mathcal{F}^\sigma_l} p(\tilde{F}, F)F, \tag{1.2}$$

for $\tilde{F} \in \mathcal{F}^\sigma_{\tilde{l}}$ and $|\sigma| \le \tilde{l} \le l$. The set $\mathcal{K}^\sigma$ represents the set of zeros in our algebra.

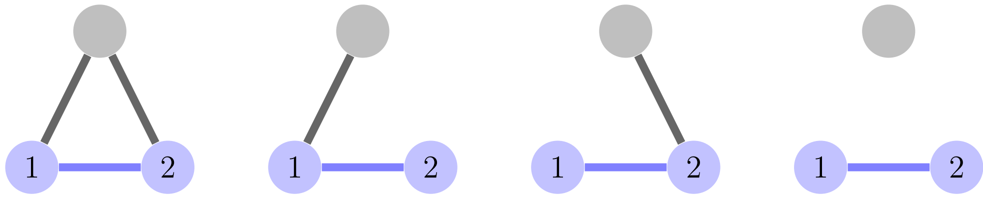
Definition 1.1.6 (Flag algebra $\mathcal{A}^\sigma$)**.**

$$\mathcal{A}^\sigma := (\mathbb{R}\mathcal{F}^\sigma)/\mathcal{K}^\sigma.$$

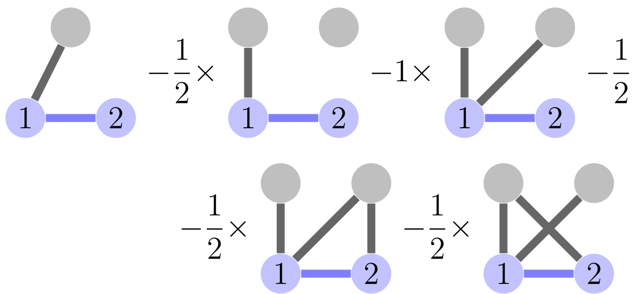
We conclude this section with several examples, which are also used to fix some notation. In the theory T_G of all simple graphs (without loops and parallel edges), there is a unique type of size 0 – when there is no risk of confusion, we also use 0 to represent this type. Note that $\mathcal{M}_l = \mathcal{F}^0_l$. There are three non-empty types of size at most 2, denoted by $1, E$ and $\overline{E}$:



The set $\mathcal{F}_3^E$ contains the following four (non-isomorphic) E-flags:



Below is an example of an element of $\mathcal{K}^E$ (and thus also of $\mathbb{R}\mathcal{F}^E$):



1.2 Operators

Typically, given a fixed set of models $M_1, \ldots, M_t$, our goal is to understand the relations between $p(M_i, N)$ that always hold for $N \in \mathcal{M}_L$ as $L \to \infty$. To achieve this, our plan is to encode the assumptions of our theory using the notation of *flag algebras*. We then use the operators defined in this section to take these assumptions and infer additional propositions from them.

1.2.1 Flag multiplication

We define the product of two flags $F_1 \in \mathcal{F}_{l_1}^\sigma, F_2 \in \mathcal{F}_{l_2}^\sigma$ as

$$F_1 \cdot F_2 := \sum_{F \in \mathcal{F}_l^\sigma} p(F_1, F_2; F)F, \tag{1.3}$$

where $l \geq l_1 + l_2 - |\sigma|$ is chosen arbitrarily. This mapping is extended onto $(\mathbb{R}\mathcal{F}^\sigma) \otimes (\mathbb{R}\mathcal{F}^\sigma)$ by linearity (where $\otimes$ denotes Cartesian product). The following Lemma shows the correctness of the definition:

Lemma 1.2.1. *Razborov [10, Lemma 2.4]*

(a) The right-hand side of (1.3) *is independed of the choice of* l *(modulo* $\mathcal{K}^\sigma$*).*

(b) The equation (1.3) *induces a bilinear mapping* $\mathcal{A}^\sigma \otimes \mathcal{A}^\sigma \to \mathcal{A}^\sigma$*.*

Proof. (a) Let $l \geq \tilde{l} \geq l_1 + l_2 - |\sigma|$. We have:

$$\begin{aligned}
\sum_{F \in \mathcal{F}_l^\sigma} p(F_1, F_2; F)F &= \sum_{F \in \mathcal{F}_l^\sigma} \sum_{\tilde{F} \in \mathcal{F}_{\tilde{l}}^\sigma} p(F_1, F_2; \tilde{F}) p(\tilde{F}, F)F \\
&= \sum_{\tilde{F} \in \mathcal{F}_{\tilde{l}}^\sigma} p(F_1, F_2; \tilde{F}) \sum_{F \in \mathcal{F}_l^\sigma} p(\tilde{F}, F)F \\
&= \sum_{\tilde{F} \in \mathcal{F}_{\tilde{l}}^\sigma} p(F_1, F_2; \tilde{F})\tilde{F} \mod \mathcal{K}^\sigma.
\end{aligned}$$

(b) We have to show that $\forall f_1 \in \mathcal{K}^\sigma \ \forall F' \in \mathcal{F}^\sigma \Rightarrow f_1 \cdot F' \in \mathcal{K}^\sigma$; the rest follows from symmetry of $(\cdot)$ and linearity of $\mathbb{R}\mathcal{F}^\sigma$. We may also assume that f_1 has the form (1.2). Therefore, we need to show:

$$\left(\tilde{F} - \sum_{F \in \mathcal{F}_l^\sigma} p(\tilde{F}, F)F\right) \cdot F' = 0.$$

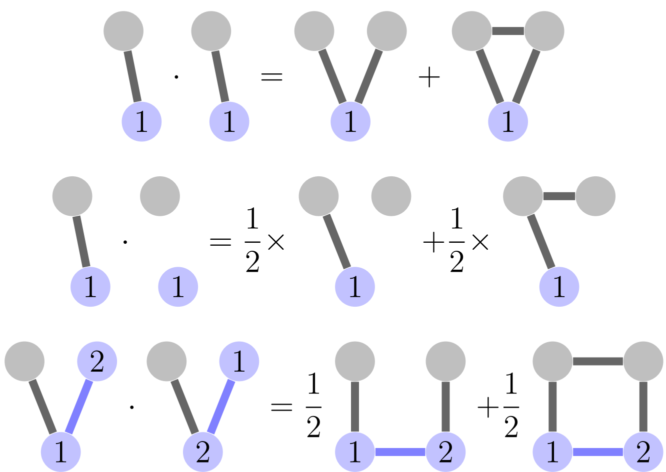
By the already proven part a, we may compare the following expressions:

$$\tilde{F} \cdot F' \qquad\qquad \sum_{F \in \mathcal{F}_l^\sigma} p(\tilde{F}, F)F \cdot F'$$

by expanding them as summations over $\mathcal{F}_L^\sigma$ for common L. Looking at a particular $\widehat{F} \in \mathcal{F}_L^\sigma$, its coefficient in the left expression is $p(\tilde{F}, F'; \widehat{F})$ while the coefficient in the expression on the right is $\sum_{F \in \mathcal{F}_l^\sigma} p(\tilde{F}, F) p(F, F'; \widehat{F})$. The equality follows from the *chain rule* of Lemma 1.1.4.

□

We again give several examples:



1.2.2 Unlabelling – the downward operator

The downward operator is a linear operator acting between different flag algebras $\mathcal{A}^{\sigma}$ and $\mathcal{A}^{\sigma'}$ for $|\sigma| > |\sigma'|$. It allows us to unlabel some (or all) vertices of our flags by averaging and thus obtain statements about flags of a smaller type. Typically, we need to perform an unlabelling step at least at the end of our arguments, since we are mostly interested in statements about graphs, i.e., the elements of $\mathcal{F}^0$.

Definition 1.2.2. *For a type σ of size k and an injective mapping $\eta : [k'] \to [k]$, we define $\sigma|_{\eta}$ to be the type induced by $im(\eta)$, i.e. for any predicate symbol $P(x_1, \ldots, x_r)$ and any $i_1, \ldots, i_r \in [k']$ the following holds:*

$$\sigma|_{\eta} \models P(i_1, \ldots, i_r) \Leftrightarrow \sigma \models P(\eta(i_1), \ldots, \eta(i_r)).$$

Furthermore, given a σ-flag $F = (M, \theta)$, the $\sigma|_{\eta}$-flag $F|_{\eta}$ is defined as $F|_{\eta} := (M, \theta\eta)$.

Now, we can define the downward operator:

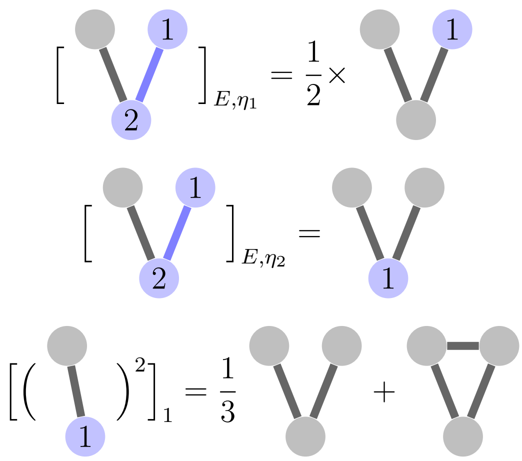
Definition 1.2.3. *For a type σ and η, k, k' as in the previous definition, we define a mapping $[\cdot]_{\sigma,\eta}$ from $\mathcal{F}^{\sigma}$ to $\mathcal{F}^{\sigma|_{\eta}}$ as follows:*

$$[F]_{\sigma,\eta} := q_{\sigma,\eta}(F) \cdot F|_{\eta},$$

where $q_{\sigma,\eta}(F)$ for $F = (M, \theta)$ is the probability that a uniformly randomly chosen injective mapping $\boldsymbol{\theta} : [k] \to V(M)$ satisfying $\theta\eta = \boldsymbol{\theta}\eta$ defines a model embedding $\sigma \to M$ and $(M, \boldsymbol{\theta}) \approx F$ holds.

Informally speaking, the coefficient q is equal to the probability that a random "extension" of the labeling of a partially unlabeled flag produces, again, the original flag (or an isomorphic one).

Let us illustrate this operator on a few examples. We let η_1 and η_2 be mappings $[1] \to [2]$, with $\eta_1(1) = 1$ and $\eta_2(1) = 2$. The empty mapping is denoted by 0 and dropped from notation by setting $[F]_{\sigma} := [F]_{\sigma,0}$.



1.2.3 Upward operator

The upward operator complements the one from the previous section. Having our example theory T_G in mind, the upward operator allows us to, e.g., take the propositions we know about graphs, or rather about $\mathcal{F}^0$, and translate them into propositions about $\mathcal{F}^\sigma$.

For this to work, we need an easy assumption of non-degeneracy of a type and also a variant of the operator defined in the previous section, which not only unlabels a subset of vertices, but completely removes them from the flag.

Definition 1.2.4. *A type σ is **non-degenerate** if $|\mathcal{F}_l^\sigma| > 0 \ \forall l \geq |\sigma|$.*

Definition 1.2.5. *Let σ be a non-degenerate type of size k and $\eta : [k'] \to [k]$ an injective mapping. For a σ-flag $F = (M, \theta)$, we define*

$$F \downarrow_\eta := F|_\eta - \theta([k] \setminus im(\eta)).$$

We can now define the upward operator:

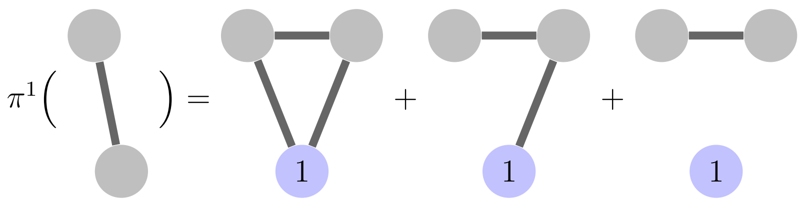
Definition 1.2.6. *Let σ, η, k, k' be as in the previous definition. The upward operator $\pi^{\sigma,\eta}(\cdot)$ is the following mapping from $\mathcal{A}^{\sigma|_\eta}$ to $\mathcal{A}^\sigma$:*

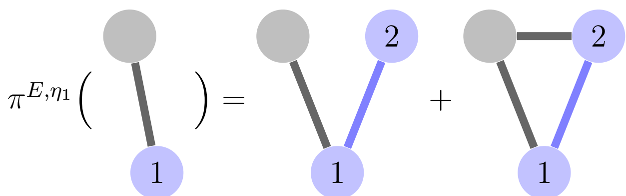
$$\pi^{\sigma,\eta}(F) := \sum \left\{ \widehat{F} : \widehat{F} \in \mathcal{F}_{l+d}^\sigma \wedge \widehat{F} \downarrow_\eta = F \right\},$$

where $d := k - k'$ is the number of vertices removed by $(\cdot) \downarrow_\eta$.

Informally, for a given type σ, vertex removal operation determined by η, and a $\sigma|_\eta$-flag F, the operator returns the sum of all σ-flags for which the vertex removal $(\cdot) \downarrow_\eta$ results in F.

Again, we illustrate the definition on some examples below. Recall the mappings η_1 and 0 from the example at the end of the Section 1.2.2. As usual, we drop the empty mapping from the notation by setting $\pi^\sigma := \pi^{\sigma,0}$.





1.3 Semantics

Recall that Lemma 1.1.5 says that the function p is "almost" multiplicative. In other words, the difference between $p(F_1, F) \cdot p(F_2, F)$ and $p(F_1 \cdot F_2, F)$ approaches 0 for $|V(F)| \to \infty$. Thus, the mapping $F \to p(F, \widehat{F})$ should converge to an algebra homomorphism for large flags $\widehat{F}$. In this section, we consider the set of all such limiting homomorphisms and determine which of them correspond to valid density assignments. Deciding a statement in extremal combinatorics is then equivalent to deciding whether the statement holds for all of these homomorphisms.

Definition 1.3.1. *Let σ be a non-degenerate type. We define $\mathrm{Hom}(\mathcal{A}^\sigma, \mathbb{R})$ to be the set of all algebra homomorphisms between $\mathcal{A}^\sigma$ and $\mathbb{R}$.*

Therefore, for every $\phi \in \mathrm{Hom}(\mathcal{A}^\sigma, \mathbb{R})$ we have $\phi(F_1) \cdot \phi(F_2) = \phi(F_1 \cdot F_2)$ for all σ-flags F_1, F_2 and $\phi(1_\sigma) = 1$.

Definition 1.3.2. *Let σ be a non-degenerate type. Denote by $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ the set of all $\phi \in \mathrm{Hom}(\mathcal{A}^\sigma, \mathbb{R})$ such that $\forall F \in F^\sigma : \phi(F) \geq 0$. We define **the semantic cone** $\mathcal{C}_{sem}(\mathcal{F}^\sigma)$ to be the set:*

$$\{f \in \mathcal{A}^\sigma : \forall \phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})(\phi(f) \geq 0)\}.$$

In the theory T_G of graphs, the elements of $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ correspond to limiting objects of convergent sequences of graphs, as discussed in [8] and [10]. The semantic cone $\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma)$ represents the (polynomial) statements that hold for all such objects.

For example, we can express Mantel's theorem as:

$$\max_\phi\{\phi(K_2) : \phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R}) \wedge \phi(K_3) = 0\} = \frac{1}{2},$$

where $K_r \in \mathcal{F}^0$ is a complete graph on r unlabeled vertices.

To make our proofs more readable, we introduce the following natural notation:

Definition 1.3.3. *For $f, g \in \mathcal{A}^\sigma$ the statement $f \leq_\sigma g$ is defined as:*

$$(g - f) \in \mathcal{C}_{sem}(\mathcal{F}^\sigma).$$

The subscript σ is dropped when the choice of the type is clear from the context.

We have already mentioned that the operators introduced in Section 1.2 correspond to inference rules of our calculus. This is formalized by the following easy proposition. (The definition of the operators is extended from flags to sets of flags in the straightforward way.)

Proposition 1.3.4. *Let σ be a non-degenerate type of size k and $\eta : [k'] \to [k]$ be an injective mapping. It holds that:*

(a) $[\mathcal{C}_{sem}(\mathcal{F}^\sigma)]_{\sigma,\eta} \subseteq \mathcal{C}_{sem}(\mathcal{F}^{\sigma|_\eta})$ *and*

(b) $\pi^{\sigma,\eta}(\mathcal{C}_{sem}(\mathcal{F}^{\sigma|_\eta})) \subseteq \mathcal{C}_{sem}(\mathcal{F}^\sigma)$

The next definition provides us with an initial set of true statements:

Definition 1.3.5. *The* ***ordinary cone*** $\mathcal{C}(\mathcal{F}^\sigma)$ *is defined as the set of elements of the form $f^2 F_1 F_2 \ldots F_t$, where $f \in \mathcal{A}^\sigma$ and $F_i \in \mathcal{F}^\sigma$ for $i \in [t]$.*

Clearly, $\mathcal{C}(\mathcal{F}^\sigma) \subseteq \mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma)$ and thus $\forall f \in \mathcal{C}(\mathcal{F}^\sigma) : [f]_\sigma \geq 0$. The following proposition will be particularly useful in the next chapter.

Proposition 1.3.6. *The following holds for any positive semidefinite quadratic form Q and $F_1, \ldots, F_t \in \mathcal{A}^\sigma$:*

$$[Q(F_1, \ldots, F_t)]_\sigma \geq 0.$$

Finally, we can easily obtain the Cauchy-Schwarz type inequalities:

Proposition 1.3.7. *It holds for any $f, g \in \mathcal{A}^\sigma$:*

$$[f^2]_{\sigma,\eta} \cdot [g^2]_{\sigma,\eta} \geq [f \cdot g]^2_{\sigma,\eta}.$$

1.3.1 Graph limits

We now proceed to prove that the definition $\mathcal{C}_{\mathrm{sem}}(\mathcal{F}^\sigma)$ indeed captures the intended semantics.

Definition 1.3.8. *Let σ be a non-degenerate type. An* ***increasing sequence*** *is a sequence of σ-flags*

$$F_1, F_2, \ldots, F_n, \ldots$$

such that $|F_1| < |F_2| < \ldots < |F_n| < \ldots$ We call an increasing sequence of σ-flags $\{F_n\}$ ***convergent*** *if the limit $\lim_{n\to\infty} p(F, F_n)$ exists for all $F \in \mathcal{F}^\sigma$.*

We assign each σ-flag $\widehat{F}$ a point $p^{\widehat{F}}$ in an infinitely dimensional space $[0,1]^{\mathcal{F}^\sigma}$:

$$p^{\widehat{F}}(F) := p(F, \widehat{F}).$$

The compactness of $[0,1]^{\mathcal{F}^\sigma}$ implies that every increasing sequence of σ-flags contains a convergent subsequence. Similarly, each $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ can be viewed as a point from this space and $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ can thus be interpreted as a (compact) subset of $[0,1]^{\mathcal{F}\sigma}$.

Theorem 1.3.9. *[10, Theorem 3.3]*

(a) If a sequence $\{F_n\}$ of σ-flags is convergent, then $\lim_{n\to\infty} p^{F_n} \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$.

(b) If $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$, then there exists a sequence $\{F_n\}$ of σ-flags such that $\lim_{n\to\infty} p^{F_n} = \phi$.

Proof. (a) We need to verify that the limiting point of the sequence $\{p^{F_n}\}$, when considered as a mapping from $\mathcal{F}^\sigma$ to $[0,1]$, is an algebra homomorphism between $\mathcal{A}^\sigma$ and $\mathbb{R}$. The expression (1.2) is mapped to zero for every F_n (provided that l in this expression is chosen so that $l \leq |F_n|$). The condition (1.3) on flag multiplication is satisfied by p^{F_n} in the limit by Lemma 1.1.5. We conclude that $\lim_{n\to\infty} p^{F_n} \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$.

(b) Consider $\phi \in \mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$. The fact $\phi(1_\sigma) = 1$ together with (1.2) implies:

$$\sum_{F \in \mathcal{F}_l^\sigma} \phi(F) = 1.$$

Thus, ϕ defines a probability measure on $\mathcal{F}_l^\sigma$ for every $l \geq |\sigma|$. We generate a random sequence $\boldsymbol{F_n}$ by choosing $\boldsymbol{F_n} \in \mathcal{F}_{n^2}^\sigma$, where the probability of generating a particular $F \in \mathcal{F}_{n^2}^\sigma$ is $\phi(F)$. It suffices to show:

$$P[\lim_{n\to\infty} p^{\boldsymbol{F_n}} = \phi] = 1.$$

This is equivalent to showing that $\forall F \in \mathcal{F}_l^\sigma$ and $\forall \varepsilon > 0$ the following holds:

$$P[\exists n_0 \text{ such that } \forall n \geq n_0 : |p(F, \boldsymbol{F_n}) - \phi(F)| \leq \varepsilon] = 1.$$

For $n^2 \geq |\sigma|$, using 1.2:

$$E[p(F, \boldsymbol{F_n})] = \sum_{F_n \in \mathcal{F}_{n^2}^\sigma} p(F, F_n)\phi(F_n) = \phi(F).$$

Next, calculate the variance:

$$\begin{aligned}
\mathrm{Var}[p(F, \boldsymbol{F_n})] &= E[p(F, \boldsymbol{F_n})^2] - \phi(F)^2 \\
&= E[p(F, \boldsymbol{F_n})^2] - p(F, F; F_n) \\
&= \sum_{F_n \in \mathcal{F}_{n^2}^\sigma} p(F, F_n)^2 \phi(F_n) - \sum_{F_n \in \mathcal{F}_{n^2}^\sigma} p(F, F; F_n)\phi(F_n) \\
&\in O(1/n^2),
\end{aligned}$$

since the difference between $p(F, F_n)^2$ and $p(F, F; F_n)$ is in $O(1/n^2)$ by Lemma 1.1.5. The statement of the theorem now follows from Chebyshev inequality and Borel-Cantelli Lemma.

□

We now formulate a corollary of Theorem 1.3.9 which links $\mathrm{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ and the study of polynomial relations in asymptotic extremal combinatorics:

Corollary 1.3.10. *[10, Corollary 3.4]*

Let $F_1, \ldots, F_t \in F_s$ be σ-flags and $f : \mathbb{R}^t \to \mathbb{R}$ be a polynomial. It holds that

$$f(F_1, \ldots, F_h) \in \mathcal{C}_{sem}(\mathcal{F}^\sigma)$$

if and only if

$$\liminf_{l \to \infty} \min_{F \in \mathcal{F}_l^\sigma} f(p(F_1, F), \ldots, p(F_h, F)) \geq 0.$$

Chapter 2

Estimating Turán densities

In this chapter, we apply flag algebras to the problem of bounding the Turán density of a fixed finite family. This is a fundamental task in dense extremal combinatorics. We shall see that the framework of flag algebras not only provides a concise way to express these arguments but it also allows to formulate the search for such arguments in terms of solving a semidefinite program.

2.1 Turán's theorem

The following theorem due to Turán is one of the most classical results in extremal combinatorics:

Theorem 2.1.1. *If a simple undirected graph $G = (V, E)$ on n vertices has no copy of K_p $(p \geq 2)$, then*

$$|E| \leq \Big(1 - \frac{1}{p-1}\Big)\frac{n^2}{2}.$$

This inequality is (asymptotically) tight, with the extremal example being a complete $(p-1)$-partite graph with partitions of (almost) equal sizes. We postpone the proof of this theorem to Section 3.1 and discuss only the special case $p = 3$, which is known as Mantel's theorem:

Theorem 2.1.2. *If a simple undirected graph $G = (V, E)$ on n vertices has no triangle, then*

$$|E| \leq \frac{n^2}{4} + O(n).$$

Here, we first present the proof of the theorem and identify its key parts. We then proceed to give an analogous proof using flag algebras. Our hope is to illustrate how the individual arguments translate to our framework while highlighting the systematic nature of the resulting proof.

Proof. As usual, for a vertex $v \in V$ the number of edges incident with v is denoted by $\deg(v)$. First, we estimate the number of edges by looking at 3-vertex configurations:

$$|E| \geq \frac{1}{n} \sum_{\{u,v\} \in E} \big(\deg(u) + \deg(v)\big).$$

This inequality holds because every edge contributes at most n to the sum – otherwise the vertices u and v would have a common neighbour which would result in a triangle. We can rewrite the sum as

$$\sum_{\{u,v\} \in E} \big(\deg(u) + \deg(v)\big) = \sum_{v \in V} \deg^2(v),$$

since each vertex v contributes $\deg(v)$ times to the sum on the left-hand side. Applying Cauchy-Schwarz inequality, we get:

$$\sum_{v \in V} \deg^2(v) \geq \frac{1}{n} \Big(\sum_{v} \deg(v) \Big)^2 = \frac{4|E|^2}{n}.$$

Therefore $|E| \geq \dfrac{4|E|^2}{n^2}$ which implies the theorem. □

Note the two key parts of the above proof:

1. inspecting possible subgraphs of size $l = 3$ and
2. using Cauchy-Schwarz inequality to exploit the information about 1-vertex overlaps between 2-vertex subgraphs obtained from the inspection of l-vertex subgraphs.

Omitting the second step would leave us with a significantly worse bound: 3-vertex graphs without K_3 have $0, 1$ or 2 edges. Thus, using only the first argument, we would get the bound

$$|E| \leq \frac{2}{3} \cdot \frac{n^2}{2}.$$

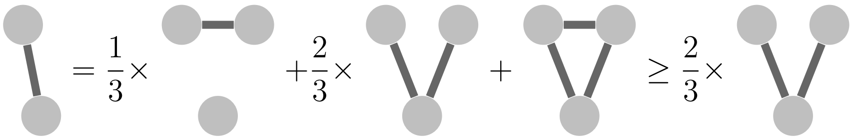
Increasing l to 5 would improve the multiplicative factor from $\frac{2}{3}$ to $\frac{3}{5}$, but the averaging argument of this kind is not strong enough to yield the tight bound.

We now reformulate the proof of the assymptotic version of the theorem in terms of flag algebras:

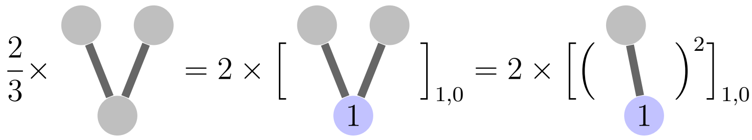
Proof. Our aim is to prove the following:

$$\max_{\phi} \{\phi(K_2) : \phi \in \mathrm{Hom}^+(\mathcal{A}^0, \mathbb{R}) \wedge \phi(K_3) = 0\} = \frac{1}{2}.$$

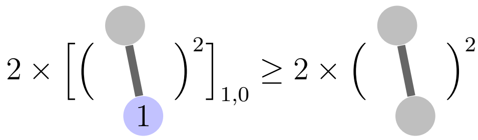
Analogously to the previous proof, we start by estimating the edge density $\phi(K_2)$ in terms of flags of size 3. This is made possible by the relation 1.2:



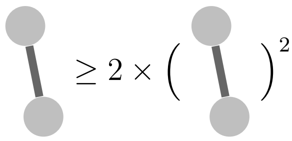
We omitted K_3, since $\phi(K_3) = 0$. We also discarded the flag with a single edge: since it is not present in our conjectured extremal example (a complete bipartite graph), our hope is that it does not play a significant role in proving the theorem. To apply Cauchy-Schwarz inequality, we distinguish one of the flag's vertices by labeling it (in the last equality, we again drop the flag containing a triangle):



The Cauchy-Schwarz inequality, as presented in Proposition 1.3.7 and applied with $g = 1$, now yields:



Thus, we arrive at the following:



Solving this inequality for $\phi(K_2)$ concludes the proof. □

In general, we are going to apply the two techniques discussed above for estimates on $|E|$ for families of prohibited graphs (or r-uniform hypergraphs) more general than $\{K_3\}$. Naturally, using larger values of l (i.e., using bigger flags) and/or considering overlaps of size larger than 1 (i.e., using larger type σ) result in better estimates. On the other hand, with bigger flags of larger types, it can be computationally prohibitive to find the right inequalities to use by computer search. This will be discussed in the next two sections. Finally, we note that while the above proof uses only a single trivial type K_1, using several types might be necessary to obtain good bounds for harder problems.

2.2 Bounding Turán densities using flag algebras

Recall that an r-uniform hypergraph (or shortly r-graph) is a tuple $H = (V, E)$ where $E \subseteq \binom{V}{3}$. We follow the standard graph theory terminology [3] while stressing that subgraphs do not necessarily have to be induced (unless stated otherwise).

Definition 2.2.1. *Let $\mathcal{H}$ be a finite family of r-uniform hypergraphs. We call an r-graph **$\mathcal{H}$-free** if it does not contain any graph $H \in \mathcal{H}$ as a subgraph. The set of all $\mathcal{H}$-free non-isomorphic r-graphs on n vertices is denoted by $\mathcal{H}_n$. The density $d(H)$ of an r-graph H is defined as:*

$$d(H) := \frac{|E(H)|}{\binom{|V|}{r}}.$$

We sometimes refer to $\mathcal{H}$-free graphs as *admissible* when the choice of the family of the prohibited subgraphs is clear from context.

Definition 2.2.2. *The **Turán density** $\pi(\mathcal{H})$ of a finite family of r-graphs $\mathcal{H}$ is defined as the following limit:*

$$\pi(\mathcal{H}) := \lim_{n\to\infty} \max\{d(H) : H \in \mathcal{H}_n\}. \tag{2.1}$$

In the following, we assume that the family of forbidden r-graphs does not contain a graph with no edges, since that results in a pathological situation where the maximum is taken over an empty set.

Using an averaging argument as in the previous section, we can obtain the following bound for any fixed l:

$$\pi(\mathcal{H}) \leq \max\{d(H) : H \in \mathcal{H}_l\}.$$

Combined with the compactness of the interval $[0, 1]$, this also shows the existence of the limit in (2.1) for any $\mathcal{H}$. In the remainder of the section, we describe a method applied in [11] and [1] to find a better upper bound on $\pi(\mathcal{H})$ using a calculation in flag algebras based on the theory of all admissible r-graphs.

Let us fix an ordering of all non-isomorphic admissible r-uniform hypergraphs $\mathcal{H}_l = \{H_1, H_2, \ldots, H_s\}$ where l is a fixed natural number. Also, let $\mathcal{H}_n^\sigma = \{F_1^\sigma, F_2^\sigma, \ldots, F_r^\sigma\}$ denote the set of all σ-flags of size n not containing any of the prohibited graphs as a subgraph. Consider t triples $(\sigma_i, m_i, \mathsf{Q}^i)$, where σ_i is a type, m_i is a natural number such that $2m_i - |V(\sigma_i)| \leq l$ (this ensures that two σ_i-flags of size m_i overlapping on σ_i fit into a graph with l vertices) and finally Q^i is a positive semidefinite matrix of type $|\mathcal{H}_{m_i}^{\sigma_i}| \times |\mathcal{H}_{m_i}^{\sigma_i}|$.

The matrices Q^i can be found using semidefinite programing while the exact choice of t and σ_i, m_i is a matter of some experimentation. Ideally, we would use all types of a certain size but this can be computationally intractable.

First, we again express the density as a sum of 0-flags of size l:

$$K_r = d(H_1) \times H_1 + \ldots + d(H_s) \times H_s, \tag{2.2}$$

where K_r denotes the flag consisting a single unlabeled hyperedge. Using Proposition 1.3.6 for quadratic forms defined by Q^i, we get the inequality:

$$0 \leq [\mathsf{Q}^i(F_1, \ldots, F_r)]_{\sigma_i} \text{ where } F_j \in \mathcal{H}^{\sigma_i}_{m_i}.$$

Now, we multiply the flags and group the coefficients of each H_i:

$$0 \leq \Big[\sum_{a,b} (\mathsf{Q}^i)_{a,b}(F_a \cdot F_b)\Big]_{\sigma_i} \tag{2.3}$$

$$= \Big[\sum_{a,b} (\mathsf{Q}^i)_{a,b} \sum_{H \in \mathcal{H}^{\sigma_i}_l} p(F_a, F_b; H) \cdot H\Big]_{\sigma_i} \tag{2.4}$$

$$= c^i_1 \times H_1 + \ldots + c^i_s \times H_s \tag{2.5}$$

Here, c^i_j is the sum of all coefficients in front of a 0-flag H_j. Note that this sum can be 0, e.g., if H_j does not contain a copy of the graph σ_i as an induced subgraph. Summing these inequalities for all $i \in [t]$ with (2.2) yields:

$$K_r \leq c_1 \times H_1 + \ldots + c_s \times H_s,$$

where $c_j = d(H_j) + \sum_i c^i_j$.

Now recall that in Section 1.3.1, we defined a certain set of homomorphisms from $\mathcal{A}^0$ to $\mathbb{R}$, denoted by $\text{Hom}^+(\mathcal{A}^0, \mathbb{R})$, and proved in Theorem 1.3.9 that these homomorphisms correspond to convergent sequences of graphs. We want to show that $\phi(K_r)$ is bounded by a universal constant for all $\phi \in \text{Hom}^+(\mathcal{A}^0, \mathbb{R})$, since this immediately implies a bound on the limit edge density in every convergent sequence. Because ϕ is a homomorphism to $\mathbb{R}$, we can use the simple fact that the weighted average of a set of real numbers is at most the maximum of the set. Thus, we get the bound:

$$\pi(\mathcal{H}) \leq \max_j \{c_j\}. \tag{2.6}$$

The positive semidefinite matrices Q_i minimizing this bound for a given choice of the types σ_i and the flag sizes m_i can be found using a semidefinite programming solver, e.g. CSDP [2].

2.3 Example – a bound on $\pi(\{K_4^-\})$

Baber and Talbot [1] applied flag algebras to give bounds on Turán density for several families of 3-graphs. We discuss their results in more detail in Section 3.3. Here, we present one of their bounds to illustrate the method.

The graph K_4^- is complete 3-uniform graph on 4 vertices without an edge:

$$K_4^- = \big\{\{1,2,3\},\{1,2,4\},\{1,3,4\}\big\}.$$

The value of Turán density $\pi(\{K_4^-\})$ is conjectured to be $2/7 \approx 0.2857$. Baber and Talbot [1] prove the following theorem:

Theorem 2.3.1. *The Turán density of K_4^- satisfies:*

$$\pi(\{K_4^-\}) \leq 0.2871.$$

This theorem is a result of a computation in flag algebras using the method described in the previous section. Specifically, they choose to analyze all 3-graphs of size $l = 7$ and use the following choice of $\sigma_i = ((V_i, E_i), \theta_i)$ and m_i:

$$\begin{aligned}
V_1 &= [3], & E_1 &= \emptyset, & m_1 &= 5,\\
V_2 &= [3], & E_2 &= \{123\}, & m_2 &= 5,\\
V_3 &= [4], & E_3 &= \{123\}, & m_3 &= 5,\\
V_4 &= [5], & E_4 &= \{123, 124, 125\}, & m_4 &= 6,
\end{aligned}$$

where we use the notation xyz to denote the set $\{x, y, z\}$ and θ_i is an identity on V_i for all $i \in [4]$. The subsequent optimization yields Theorem 2.3.1.

2.4 Computational aspects

In the previous section, we have seen how the problem of bounding $\pi(\mathcal{H})$ can be turned into a semidefinite program minimizing (2.6). However, since the number of non-isomorphic admissible r-graphs grows exponentially, the sizes of the matrices Q^i increase rapidly, making the solution of the semidefinite program difficult to obtain. Razborov [11] has shown several methods to reduce the computational complexity of this task.

The method applied in [1] exploits the fact that semidefinite solvers run in time bounded by a polynomial in the sum of the sizes of the blocks in the block-diagonal structure of the matrices involved. Thus, for each i, type σ_i and flag size m_i, we consider the space of all linear combinations of *admissible* flags $\mathcal{H}_{m_i}^{\sigma_i}$, denoted by $\mathbb{R}\mathcal{H}_{m_i}^{\sigma_i}$, and find a basis with respect to which the quadratic forms to be found have a block-diagonal structure. (We drop i from the notation, since the argument is carried out for each i separately.)

Specifically, we find a base $\mathcal{B}$ of $\mathbb{R}\mathcal{H}_m^\sigma$ with the following two properties:

1. $\mathcal{B} = \mathcal{B}^+ \cup \mathcal{B}^-$, where $\mathcal{B}^+ \cap \mathcal{B}^- = \emptyset$,

2. for all $B^+ \in \mathcal{B}^+, B^- \in \mathcal{B}^-$ we have $[B^+ \cdot B^-]_\sigma = 0$.

The second property ensures that we can (after changing the coordinate frame) restrict the search to just those positive semidefinite matrices Q that have a block-diagonal structure with one block of size $|\mathcal{B}^+|$ and one of size $|\mathcal{B}^-|$. Even more formally, let T be the $|\mathcal{H}_m^\sigma| \times |\mathcal{H}_m^\sigma|$ matrix expressing the change of basis from $\mathcal{B}$ to the trivial basis $\mathcal{H}_m^\sigma$. In the subsequent optimization, we enforce Q to have the stated block-diagonal structure and replace Q in (2.3) by $\mathsf{T}^T\mathsf{Q}\mathsf{T}$. This has a potential of significantly reducing the number of variables to be determined, especially if the blocks have comparable sizes.

In the rest of this Section, we describe the construction of such a base $\mathcal{B}$. Let us fix a type $\sigma = (H, \theta)$ and a flag size m. We begin by considering the automorphism group Γ of all bijections $\alpha : \mathrm{V}(\sigma) \to V(\sigma)$ satisfying $(H, \theta) \approx (H\alpha, \theta\alpha)$. For $\alpha \in \Gamma$ and a flag $F = (H_F, \theta_F)$, we introduce the notation $F\alpha$ for the σ-flag $(H_F, \theta_F\alpha)$. First, we find bases for the following two subspaces of $\mathbb{R}\mathcal{H}_m^\sigma$:

$$\mathbb{R}\mathcal{H}_m^{\sigma+} = \{L : L \in \mathbb{R}\mathcal{H}_m^\sigma \wedge L\alpha = L \forall \alpha \in \Gamma\},$$

$$\mathbb{R}\mathcal{H}_m^{\sigma-} = \{L : L \in \mathbb{R}\mathcal{H}_m^\sigma \wedge \sum_{\alpha \in \Gamma} L\alpha = 0\}.$$

Note that $\mathbb{R}\mathcal{H}_m^{\sigma+}$ is the Γ-invariant subspace of $\mathbb{R}\mathcal{H}_m^\sigma$, while $\mathbb{R}\mathcal{H}_m^{\sigma-}$ is the anti-invariant one. Since every element of $\mathbb{R}\mathcal{H}_m^\sigma$ can be expressed as a sum of invariant and anti-invariant elements, the union of the two subspaces spans the whole $\mathbb{R}\mathcal{H}_m^\sigma$. Therefore, our aim is to first find the bases $\mathcal{B}^+$ and $\mathcal{B}^-$ for the subspaces $\mathbb{R}\mathcal{H}_m^{\sigma+}$ and $\mathbb{R}\mathcal{H}_m^{\sigma-}$, respectively. We then show that they indeed have the desired properties.

We begin the construction by taking the set $\mathcal{H}_m^\sigma$ of all admissible σ-flags of size m along with the following equivalence relation:

$$F_a \sim F_b \text{ if and only if } \exists \alpha \in \Gamma \text{ such that } F_a\alpha = F_b.$$

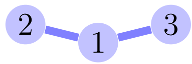
We partition the set $\mathcal{H}_m^\sigma$ into classes of equivalence relation $\sim$. These (distinct) classes are denoted by $Q_1, Q_2, \ldots, Q_u$ and called *orbits*. The base of $\mathbb{R}\mathcal{H}_m^{\sigma+}$ is formed by sums of individual orbits:

$$\mathcal{B}^+ = \{\sum_{\mathcal{F} \in O_i} F : i \in [u]\}.$$

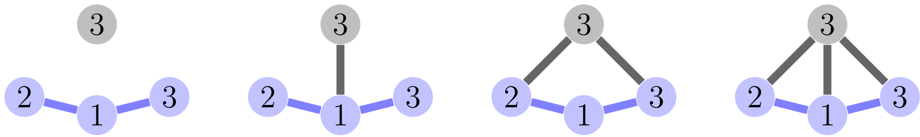
To construct the base of $\mathbb{R}\mathcal{H}_m^{\sigma-}$, we pick one arbitrary flag $\widehat{F_i}$ in each orbit O_i and set

$$\mathcal{B}^- = \{\widehat{F_i} - F' : F' \in O_i \wedge F' \neq \widehat{F_i}\}.$$

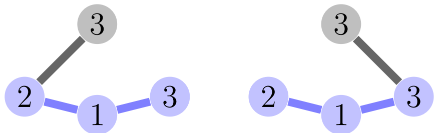
Let us provide an example to illustrate the situation in the theory T_G of all graphs. Set $m = 4$ and consider the following type (labeled path on 3 vertices):



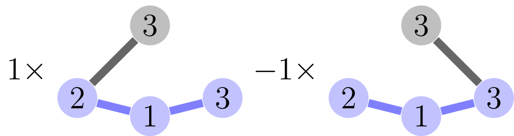
The group of automorphisms Γ consists of the identity mapping and the bijection switching the elements 2 and 3. The set $\mathcal{H}_4^\sigma$ consists of 8 graphs. The following four of them are in orbits of size 1:



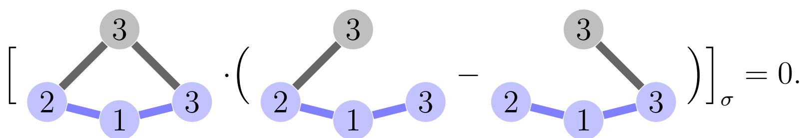
Each of these flags is an example of an element of $\mathcal{B}^+$. There are two orbits of size 2, one of them formed by the following pair of graphs:



Again, the sum of these two flags is an element of $\mathcal{B}^+$. One possible element of $\mathcal{B}^-$ is below:



We have already mentioned that the product of an element of $\mathcal{B}^+$ and an element of $\mathcal{B}^-$ should yield 0 after unlabeling. As an exampel of this property, it can be easily verified that:



This is established in full generality by the following lemma from [11]. This lemma shows that the use of this coordinate frame indeed results in a block-diagonal structure of the matrix Q.

Lemma 2.4.1. *$\forall B^+ \in \mathcal{B}^+, \forall B^- \in \mathcal{B}^-$ we have $[B^+ \cdot B^-]_\sigma = 0$.*

Proof. We have $B^- = F_b\alpha - F_b$ for a choice of $F_b \in \mathcal{H}_m^\sigma$ and $\alpha \in \Gamma$ as in the construction above. Also, $B^+\alpha = B^+$ for all $B^+ \in \mathcal{B}^+$, $\alpha \in \Gamma$. Therefore:

$$\begin{aligned}
[B^+ \cdot (F_b\alpha - F_b)]_\sigma &= [(B^+\alpha) \cdot (F_b\alpha - F_b)]_\sigma \\
&= [(B^+\alpha)(F_b\alpha) - (B^+ \cdot F_b)]_\sigma \\
&= [(B^+\alpha)(F_b\alpha)]_\sigma - [(B^+ \cdot F_b)]_\sigma \\
&= [(B^+ \cdot F_b)]_\sigma - [(B^+ \cdot F_b)]_\sigma
\end{aligned}$$

The last equality follows from $[F]_\sigma = [F\alpha]_\sigma$ for any σ-flag F and a permutation α of $V(\sigma)$. $\square$

Chapter 3

Hypergraph jumps and Lagrangians

In this chapter, we discuss the concept of *hypergraph jumps* and a related question of Erdős. This question can be reduced by a theorem of Frankl and Rödl [7] to finding a suitable family $\mathcal{H}$ of hypergraphs satisfying a certain inequality between two of its properties – the Turán density of the family and its minimal Lagrangian.

The former of these two properties has already been discussed in Chapter 2. The latter is the subject of Section 3.1. We then move on to describe the current state of knowledge about hypergraph jumps and prove a theorem of Frankl and Rödl relating them with Turán densities and Lagrangians. Finally, in Section 3.3 we discuss recent progress made by Baber and Talbot [1] for the 3-uniform case.

3.1 The Lagrange function of a hypergraph

The Lagrangian of a graph has been first introduced by Motzkin and Straus [9] to provide another proof of the Turán's theorem. Informally speaking, the Lagrangian measures the maximal local density of a graph. Although there are several ways to extend this property from graphs to r-graphs, the most natural generalization will suffice for our purposes.

Definition 3.1.1. *Fix an r-graph H with $V(H) = [n]$. The* ***Lagrangian function of*** *H is a function $\lambda_H : \mathbb{R}^n \to \mathbb{R}$ defined as follows:*

$$\lambda_H(x) := r! \cdot \sum_{\{i_1,\dots,i_r\}\in E(H)} x_{i_1}x_{i_2}\cdots x_{i_r},$$

where x_i denotes the i-th coordinate of the vector x. The ***Lagrangian of*** *H, denoted by $\lambda(H)$, is the maximum of this function on the simplex $S_n \subset \mathbb{R}^n$ given by $x_i \geq 0, \sum x_i = 1$:*

$$\lambda(H) := \max_{x\in S_n} \lambda_H(x).$$

(The purpose of the multiplicative factor $r!$ is mainly to normalize the possible values to the $[0, 1]$ interval.) For 2-graphs, there is a simple characterization of the value $\lambda(G)$ provided by the following lemma:

Lemma 3.1.2. *Let G be a graph in which the largest clique has size p. Then:*

$$\lambda(G) = 1 - \frac{1}{p}.$$

It follows that for 2-graphs, the possible values of $\lambda(G)$ form a discrete set and also that the computation (or even a reasonable approximation) of $\lambda(G)$ is NP-hard. The easy proof of Lemma 3.1.2 can be found in [7] where it is derived from a lemma stated below as Lemma 3.1.3. This lemma gives some insight into how Lagrangians behave for general value of r. The situation for $r \geq 3$ is more complex.

Lemma 3.1.3. *Let H be an r-graph ($r \geq 2$) with $V(H) = [n]$. Among all vectors $x \in S_n$ (where S_n is the same simplex as in the previous definition) satisfying $\lambda_H(x) = \lambda(H)$, choose one for which the set $J := \{i : x_i > 0\}$ has minimal size. For $i, j \in J, i \neq j$ there exists $e \in E(H)$ such that*

$$\{i, j\} \subseteq e \subseteq J.$$

Proof. Suppose there exist indices i and j violating the statement of the lemma. We proceed by adjusting x so that it has fewer non-zero positions. By symmetry, we can assume:

$$\frac{\partial}{\partial x_i}\lambda_H(x) \leq \frac{\partial}{\partial x_j}\lambda_H(x).$$

Define $\delta := \min\{x_i, 1 - x_j\}$ and construct a vector z from x by setting:

$$z_k := \begin{cases} x_i - \delta & \text{for } k = i \\ x_j + \delta & \text{for } k = j \\ x_k & \text{otherwise.} \end{cases}$$

Clearly, $z \in S_n$ and has fewer non-zero positions. Furthermore, we have

$$\lambda_H(z) = \lambda_H(x) + \delta\Big(\frac{\partial}{\partial x_j}\lambda_H(x) - \frac{\partial}{\partial x_i}\lambda_h(x)\Big) \geq \lambda_H(x),$$

since λ_H is linear in each variable and since

$$\frac{\partial^2}{\partial x_i \partial x_j}\lambda_H(x) = 0,$$

which follows from i and j being non-neighbours. □

The following proof of Turán's theorem, already mentioned in Section 2.1, illustrates the usefullness of Lagrangians:

Theorem 3.1.4. *Let G be a graph. If p is the order of the largest clique of G, then*

$$d(G) \leq 1 - \frac{1}{p-1}.$$

Proof. By Lemma 3.1.2, we have:

$$\lambda(G) = 1 - \frac{1}{p-1}.$$

We also have:

$$2 \cdot |E(G)| \cdot \left(\frac{1}{n}\right)^2 \leq \lambda(G),$$

since the left side of the inequality is the value of the Lagrangian function for the vector with $1/n$ in each position. The statement of the theorem follows from the observation that this is equal to the density $d(G)$. □

3.2 Hypergraph jumps

Let us start with a definition:

Definition 3.2.1. *The number $\alpha, 0 \leq \alpha \leq 1$ is **a jump** for r if for any $\epsilon > 0$ and any integer $m, m \geq r$ any r-hypergraph with $N > N(\epsilon, m)$ vertices and at least $(\alpha+\epsilon)\binom{N}{r}$ edges contains a subgraph with m vertices and at least $(\alpha+c)\binom{m}{r}$ edges, where $c := c(\alpha)$ does depend only on α.*

Erdős-Stone-Simonovitz Theorem [5,6] implies that for $r = 2$, every $\alpha \in [0, 1)$ is a jump. Erdős asked whether this also holds for all $r \geq 3$ and in particular whether $\alpha = r!/r^r$ is a jump for every r. The first question has been answered negatively in [7] by proving that $1 - l^{1-r}$ is not a jump for $r \geq 3$ and $l > 2r$. However, whether $r!/r^r$ is a jump remains open even for $r = 3$. Baber and Talbot [1] have found two intervals of jumps for $r = 3$, specifically $[0.2299, 0.2316)$ and $[0.2871, 8/27)$. Their proof is discussed in Section 3.3 and uses the following characterization developed by Frankl and Rödl [7]:

Theorem 3.2.2. *The following statements are equivalent:*

1. *A real number α is a jump for r.*
2. *There exists a family of r-graphs $\mathcal{H}$ such that:*

$$\min_{H \in \mathcal{H}} \lambda(H) > \alpha \geq \pi(\mathcal{H}).$$

This theorem immediately implies the result that $[0, r!/r^r)$ is an interval of jumps for every $r \geq 3$, since $\lambda(K_r) = r!/r^r$ and $pi(\{K_r\}) = 0$, where K_r denotes the r-graph on r vertices with a single edge.

For our purposes, it suffices to prove the implication (2) ⇒ (1). We need the following theorem of Erdős. We state it without a proof, which can be found in [4]:

Theorem 3.2.3. *Let $\{G_i\}$ be an infinite sequence of r-graphs such that $|V(G_i)| \to \infty$ as $i \to \infty$. If $\lim_{i\to\infty} d(G_i)$ exists and is non-zero, then there exists a sequence $\{H_i\}$ of complete r-partite r-graphs with each color class of the same cardinality such that $H_i \subseteq G_i$, $|V(H_i)| \to \infty$ as $i \to \infty$ and $\lim_{i\to\infty} d(H_i) = r!/r^r$.*

We can now prove the implication (2) ⇒ (1) in Theorem 3.2.2.

Proof. Let $\mathcal{H} = \{H_1, H_2, \ldots, H_t\}$ be a family of r-graphs as in the statement of the theorem and let ϵ and m be as in the Definition 3.2.1. Our aim is to find $N = N(\epsilon, m)$. The proof proceeds in the following steps: First, we show that all sufficiently large r-graphs with density $\alpha + \epsilon$ contain a large number of copies of one of the graphs of $\mathcal{H}$. In fact, their *density* in the r-graph can be bounded from below by a constant. This allows us to apply Theorem 3.2.3 to find a configuration of these copies isomorphic to any fixed "blow-up version" of this graph (provided that the r-graph is big enough). Since the density of a blow-up version of an r-graph approaches the value of its Lagrangian, we can conclude the proof by setting $c := \min_{H\in\mathcal{H}} \lambda(H) - \alpha$.

Note that without loss of generality we can assume that all graphs in $\mathcal{H}$ have the same number of vertices, denoted by l. We can add isolated vertices to graphs of smaller size without affecting their Turán density or their Lagrangian.

Let n_1 be such that for all $n \geq n_1$ the following holds:

$$\frac{n \cdot (n-1) \cdot \ldots \cdot (n-r+1)}{n^r} \geq \frac{\alpha + \epsilon/2}{\alpha + 2\epsilon/3}.$$

Let $n_T(\epsilon, \alpha, \mathcal{H})$ be a number with the property that every graph on at least $n_T(\epsilon, \alpha, \mathcal{H})$ vertices with density of at least $\alpha + \epsilon$ contains a copy of a graph from $\mathcal{H}$. Now, set $n_2 := \max\{n_1, n_T(\epsilon, \alpha, \mathcal{H})\}$ and consider a graph G on at least n_2 vertices. The expected density of a uniformly randomly chosen n_1-vertex induced subgraph of G is at least $\alpha + \frac{\epsilon}{2}$. This implies that at least

$$\frac{\epsilon}{4}\binom{|V(G)|}{n_2}$$

n_2-subsets of $V(G)$ induce a subgraph with density at least $\alpha + \epsilon/4$. Due to the choice of n_2, we know that each of these subgraphs contains a copy of an r-graph from $\mathcal{H}$. On the other hand, a single copy of $H_i \in \mathcal{H}$ cannot be contained in more than

$$\binom{|V(G)| - |V(H_i)|}{n_2 - |V(H_i)|}$$

subgraphs. By pigeon-hole principle, there exists a number $\widehat{c} > 0$ such that G contains at least

$$\frac{\widehat{c}}{m}\binom{n}{|V(H_i)|}$$

copies of H_i for some choice of i provided that G has at least $n_3 = n_3(n_2, \widehat{c})$ vertices. Similarly, we can argue that there exists a partition $V(G) = V_1 \cup V_2 \cup \ldots \cup V_l$ (recall that l is the size of H_i) such that the density of the copies of H_i with vertices in the same position, say $v_j \in V_j$ is

$$\frac{1}{l^l}\frac{\widehat{c}}{m}\binom{n}{l} > c'\binom{n}{l}.$$

Now define an auxiliary l-graph based on the r-graph G in the following way. Its vertex set is the same as G. To define the (l-uniform) edges, consider the partition $V_1 \cup V_2 \cup \ldots \cup V_l$ and the fixed ordering of the vertices of H_i as above. The vertices $\{v_1, v_2, \ldots, v_l\}$ form an edge of the l-graph if and only if $v_j \in V_j$ for all i and the vertices $\{v_1, \ldots, v_l\}$ induce a copy of H_i respecting the fixed ordering of $V(H_i)$.

Since the density of this l-graph is at least c', we can use Theorem 3.2.3 to obtain a complete l-partite subgraph of this l-graph. This implies that for sufficiently large $|V(G)|$, the original graph G contains any blow-up of the graph H_i constructed in the following way: each vertex $v_j \in V(H_i)$ is replaced by t_j copies of v_j denoted by $v_j^1, \ldots, v_j^{t_j}$. For every edge $\{v_{j_1}, v_{j_2}, \ldots, v_{j_r}\}$ of H_i, we insert into the graph the edge-set of a complete r-partite r-graph with partitions $\{v_{j_1}^1, \ldots, v_{j_1}^{t_{j_1}}\}, \ldots, \{v_{j_r}^1, \ldots, v_{j_r}^{t_{j_r}}\}$. It can be easily seen that an appropriate choice of coefficients t_j leads to a subgraph with density approaching $\lambda(H_i)$: just set t_j proportionally to the weight of the vertex v_j in the weighting maximizing the Lagrangian function. This in turn implies that we can set $c := \min_{H \in \mathcal{H}} \lambda(H) - \alpha$, which proves the theorem. □

3.3 The existence of jumps for 3-uniform hypergraphs

By Theorem 3.2.2, proving that α is a jump for r can be reduced to finding an appropriate family of r-graphs and estimating its Lagrangians and its Turán density. Although calculating the Lagrangian is NP-hard, it can be easily done for small r-graphs using numerical routines implemented in software packages like

Mathematica. Providing good bounds on Turán density of families of r-graphs is more problematic but the method presented in Chapter 2 gives some results.

In Section 2.3, we established the bound

$$\pi(\{K_4^-\}) \leq 0.2871,$$

where K_4^- is the 3-graph on vertices $\{1,2,3,4\}$ with edge set

$$\big\{\{1,2,3\},\{1,2,4\},\{1,3,4\}\big\}.$$

Since the value of $\lambda(K_4^-)$ is $8/27$, we get the following:

Theorem 3.3.1 ([1]). *Every $\alpha \in [0.2871, 8/27)$ is a jump for $r = 3$.*

Baber and Talbot [1] also established a bound on the Turán density of the family of 3-graphs $\mathcal{H}' = \{H_1, H_2, H_3, H_4, H_5\}$, where

$$\begin{aligned}
H_1 =& \{123, 124, 134\},\\
H_2 =& \{123, 124, 125, 345\},\\
H_3 =& \{123, 124, 235, 145, 345\},\\
H_4 =& \{123, 135, 145, 245, 126, 246, 346, 356, 237, 147, 347, 257, 167\},\\
H_5 =& \{123, 124, 135, 145, 236, 346, 256, 456, 247, 347, 257, 357, 167\}.
\end{aligned}$$

We again use the notation xyz to represent the set $\{x, y, z\}$. We employ a semidefinite program constructed using the method of Chapter 2 applied with the following choices of $\sigma_i = ((V_i, E_i), \theta_i)$ and m_i:

$$\begin{aligned}
&V_1 = [1], && E_1 = \emptyset, && m_1 = 4,\\
&V_2 = [3], && E_2 = \emptyset, && m_2 = 5,\\
&V_3 = [3], && E_3 = \{123\}, && m_3 = 5,\\
&V_4 = [5], && E_4 = \{123, 124, 135\} && m_4 = 6,\\
&V_5 = [5], && E_4 = \{123, 124, 345\}, && m_5 = 6,\\
&V_6 = [5], && E_4 = \{123, 124, 135, 245\}, && m_6 = 6.
\end{aligned}$$

The subsequent optimization yields the bound $\pi(\mathcal{H}') \leq 0.2299$. This and the fact that the minimal Lagrangian of $\mathcal{H}'$ is 0.2316 gives us the second interval of hypergraph jumps:

Theorem 3.3.2 ([1]). *Every $\alpha \in [0.2299, 0.2316)$ is a jump for $r = 3$.*

Conclusion

In this thesis, we have studied the framework of flag algebras and discussed how it can be applied to bound Turán densities of families of hypergraphs and to resolve a question about hypergraph jumps.

The bounds on $\pi(\{K_4^-\})$ and on $\pi(\mathcal{H}')$ from Section 3.3 are almost certainly not tight. A better bound could be obtained by using more types and/or flags of larger size. Even though we are able to generate semidefinite programs in some of these situations, their size quickly becomes prohibitively large. Alternatively, one could look at the inequality (2.6) and observe which c_i's achieve the maximal value. Subsequently, one could try adding an ad-hoc Cauchy-Schwarz inequality which has negative coefficients in front of the graphs H_i where the value of c_i is maximal. This has the potential of lowering the value of $\max_i\{c_i\}$. It should be noted that the method described in Section 2.2 is capable, at least in some cases, to obtain a tight bound on the Turán density using a semidefinite program of a reasonable size (see, e.g., [11]).

A better bound on $\pi(\{K_4^-\})$ and on $\pi(\mathcal{H}')$ would result in extending the interval of hypergraph jumps. However, the pivotal question of whether $2/9$ is a jump for $r = 3$ cannot be answered by establishing a better bound on $\pi(\mathcal{H}')$. This follows from the fact that there are 3-graphs H with $\lambda(H) > 2/9$ that do not contain any of the 3-graphs in $\mathcal{H}'$ as a subgraph. By taking blow-ups of such a graph H as in the proof of Theorem 3.2.2, we can obtain a graph with density strictly larger than $2/9$ without containing a 3-graph from $\mathcal{H}'$ as a subgraph, proving that $\pi(\mathcal{H}') > 2/9$.

Appendix A

Semidefinite programming and the CSDP library

Semidefinite programming refers to finding the optimal solution to the following convex optimization problem:

$$\begin{aligned} &\text{Maximize } \operatorname{tr}(C^T X) \\ &\text{subject to } A(X) = \mathbf{b}, \\ &X \succeq 0, \end{aligned}$$

where X and C are $n \times n$ matrices, $A(X)$ is a linear operator mapping an $n \times n$ matrix into $\mathbb{R}^m$ and $\mathbf{b}$ is an m-vector. The values of $C, \mathbf{b}$ and A are fixed and the search is done over all positive semidefinite matrices X, as expressed by $X \succeq 0$.

We use semidefinite programming to bound Turán densities by *minimizing the maximum* of a certain set of linear functions, each parametrized by a semidefinite matrix Q_i. This is only a minor difference from the optimization described above. The presence of more than one matrix X to be searched for can be solved by forming a block-diagonal matrix from all the matrices Q_i. The transition from maximization to minimization is trivial. Optimization of a maximum of a closed set of objective functions can be performed by introducing several extra variables: one for each objective function c_j and then one additional variable c, which is enforced to satisfy $c \geq c_j \forall j$.

We use the CSDP library [2] to solve semidefinite programs. Specifically, we apply its standalone solver that loads a prepared input file describing each problem. The input files corresponding to some of the problems discussed in this thesis can be generated using programs, which can be found on the attached DVD. Below, we give a brief overview of the sparse input file format used by the solver.

The input file consists of 5 sections. With the exception of the last one, all sections are placed on a single line.

1. The number m of constraints.
2. The number of blocks in the block-diagonal structure of the matrices.
3. The sizes of individual blocks. Negative number indicates that the corresponding block is actually diagonal.
4. Vector specifying the objective function. As mentioned above, in our case we maximize only the value of the variable c. This vector is thus a vector of all 0's with the exception of one position, which is equal to 1.
5. The last section specifies the constraint matrices. Each line has the format:

   `<matrix_no> <block_no> <i> <j> <entry>`

   Here, `matrix_no` denotes the index of the matrix (starting from 0), `block_no` identifies the block, `i` and `j` the position within the block and finally `entry` specifies the matrix value on that position. Since the matrices are symmetric, only the entries in the upper triangle are specified.

Bibliography

[1] Rahil Baber and John Talbot. Hypergraphs do jump. *Comb. Probab. Comput.*, 20:161–171, 2011.

[2] Brian Borchers. CSDP, a c library for semidefinite programming. *Optimization Methods and Software*, 11(1):613–623, 1999.

[3] Reinhard Diestel. *Graph Theory.* Springer Verlag, 2010.

[4] P. Erdős. On extremal problems of graphs and generalized graphs. *Israel Journal of Mathematics*, 2:183–190, 1964. 10.1007/BF02759942.

[5] P. Erdős and M. Simoinovits. A limit theorem in graph theory. *Studia Sci. Math. Hung*, pages 51–57, 1966.

[6] P. Erdős and A. H. Stone. On the structure of linear graphs. *Bull. Amer. Math. Soc.*, pages 1087–1091, 1946.

[7] Peter Frankl and Vojtech Rödl. Hypergraphs do not jump. *Combinatorica*, 4(2):149–159, 1984.

[8] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *J. Comb. Theory, Ser. B*, 96(6):933–957, 2006.

[9] T. S. Motzkin and E. G. Straus. Maxima for graphs and a new proof of a theorem of turán. *Canadian Journal of Mathematics*, pages 533–540, 1965.

[10] Alexander A. Razborov. Flag algebras. *J. Symb. Log.*, 72(4):1239–1282, 2007.

[11] Alexander A. Razborov. On 3-hypergraphs with forbidden 4-vertex configurations. *SIAM J. Discret. Math.*, 24:946–963, August 2010.