Marcel K. de Carli Silva
Instituto de Matemática e Estatística
Universidade de São Paulo, Brazil
mksilva@ime.usp.br

Fernando Mário de Oliveira Filho
Instituto de Matemática e Estatística
Universidade de São Paulo, Brazil
fmario@ime.usp.br

Cristiane Maria Sato
Centro de Matemática, Computação e Cognição,
Universidade Federal do ABC, Brazil
c.sato@ufabc.edu.br

Flag algebras: a first glance

The theory of flag algebras, introduced by Razborov in 2007, has opened the way to a systematic approach to the development of computer-assisted proofs in extremal combinatorics. It makes it possible to derive bounds for parameters in extremal combinatorics with the help of a computer, in a semi-automated manner. In this article Marcel de Carli Silva, Fernando de Oliveira Filho and Cristiane Sato describe the main points of the theory in a complete way, using Mantel's theorem as a guiding example.

Mantel's theorem, perhaps the first result in extremal graph theory, was motivated by a problem proposed by W. Mantel in an issue of the journal *Wiskundige Opgaven*, published by the KWG [10]:

Vraagstuk XXVIII. K 13 a. Er zijn eenige punten gegeven waarvan geen vier in een zelfde vlak liggen. Hoeveel rechten kan men hoogstens tusschen die punten trekken zonder driehoeken te vormen? [W. Mantel]

(**Problem XXVIII.** K 13 a. Given are some points, no four of which lie on the same plane. How many lines at most can one draw between the points without forming triangles?)

In the language of graph theory, Mantel's problem asks for the maximum number of edges that a graph without triangles can have: the restriction that no four points lie on the same plane is there exactly to ensure that only triangles between the given points can be formed when lines are drawn.

A triangle-free graph on n vertices can be constructed as follows: divide the vertex set into two parts of $\lfloor n/2 \rfloor$ and $\lceil n/2 \rceil$ vertices each and add all edges between the parts. The resulting graph is bipartite, and hence in particular triangle-free, and has $\lfloor n^2/4 \rfloor$ edges. Mantel's theorem states that this is an *extremal example*, the best one can do: *every triangle-free graph on n vertices has at most $\lfloor n^2/4 \rfloor$ edges.*

This answer to Mantel's problem appeared in the same issue of *Wiskundige Opgaven*. There it is mentioned that solutions were provided by Mantel and several others; a proof by W.A. Wythoff (1865–1939), a former student of D.J. Korteweg (1848–1941), is included.

The theory of flag algebras allows us to computationally tackle extremal graph theory problems such as Mantel's problem and to obtain results such as Mantel's theorem. To understand how this is done, we first need to define exactly which extremal problems we consider.

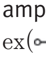
The *size* of a graph G is its number of vertices $|V(G)|$ and is denoted by $|G|$. For $U \subseteq V(G)$, we denote by $G[U]$ the subgraph of G *induced* by U, that is, the subgraph of G with vertex set U and all the edges of G between vertices of U. For graphs F and G, let $p(F;G)$ be the probability that a set $U \subseteq V(G)$ with $|U| = |F|$, chosen uniformly at random, is such that $G[U]$ is isomorphic to F. We say that $p(F;G)$ is the *density* of F in G. In other words, if $c(F;G)$ is the number of times F occurs as an induced subgraph of G, then

$$p(F;G) = c(F;G)\binom{|G|}{|F|}^{-1}.$$

Let $\mathcal{H}$ be a collection of graphs. A graph G is *$\mathcal{H}$-free* if no induced subgraph of G is isomorphic to a graph in $\mathcal{H}$. A fundamental problem in extremal graph theory is to determine, for a given graph C, the maximum asymptotic density of C in $\mathcal{H}$-free graphs

$$\mathrm{ex}(C, \mathcal{H}) = \sup_{(G_k)_{k\geq 0}} \limsup_{k \to \infty} p(C; G_k), \qquad (1)$$

where the supremum is taken over all sequences $(G_k)_{k\geq 0}$ of $\mathcal{H}$-free graphs that are *increasing*, i.e., with $(|G_k|)_{k\geq 0}$ strictly increasing.

Mantel's theorem shows that ex(, {}) $\leq \frac{1}{2}$. Together with the extremal example described above, we actually have ex(, {}) $= \frac{1}{2}$.

Let $\mathcal{G}$ be the set of all finite $\mathcal{H}$-free graphs taken up to isomorphism. An increasing sequence $(G_k)_{k\geq 0}$ is *convergent* if $\lim_{k\to\infty} p(F;G_k)$ exists for every $F \in \mathcal{G}$. Every increasing sequence of $\mathcal{H}$-free graphs has a convergent subsequence. Indeed, densities are numbers in $[0,1]$, so for $k \geq 0$ the function $F \mapsto p(F;G_k)$ can be identified with a point in $[0,1]^{\mathcal{G}}$, which is a compact space by Tychonoff's theorem.

In (1) we may therefore restrict ourselves to convergent sequences and this allows us to work with their limits. Call $\phi: \mathcal{G} \to \mathbb{R}$ a *limit functional* if there is a convergent sequence $(G_k)_{k\geq 0}$ of $\mathcal{H}$-free graphs such that

$$\phi(F) = \lim_{k\to\infty} p(F; G_k)$$

for all $F \in \mathcal{G}$ and let Φ denote the set of all limit functionals. *Then computing* $\mathrm{ex}(C, \mathcal{H})$ *is the same as solving an optimization problem over* Φ*:*

$$\mathrm{ex}(C, \mathcal{H}) = \sup\{\phi(C) : \phi \in \Phi\}. \qquad (2)$$

This is just a rewording of the original problem, but it emphasizes that the difficulty here lies in understanding Φ. This set may be very complex and computationally intractable, but to get an upper bound for $\mathrm{ex}(C, \mathcal{H})$ we do not need to work with Φ. Instead, we may look for a nice *relaxation* of Φ, that is, a set $\Phi' \supseteq \Phi$ for which we can solve the optimization problem. A first and obvious relaxation would be to take $\Phi' = [0,1]^{\mathcal{G}}$. Solving the optimization problem is then trivial, but we always get the bound $\mathrm{ex}(C, \mathcal{H}) \leq 1$. The difficulty lies in managing the trade-off between the quality of the relaxation and its tractability.

The theory of flag algebras [12], developed by the Russian mathematician Alexander Razborov, winner of the Nevanlinna Prize in 1990 and the Gödel Prize in 2007, gives us computationally-tractable relaxations of Φ that have displayed good quality in practice. We may then use the computer to solve the corresponding optimization problems, thus obtaining upper bounds for $\mathrm{ex}(C, \mathcal{H})$ that are often tight. Perhaps the most attractive feature in the theory is that the whole process is more-or-less automatic: obtaining the relaxation and solving the corresponding problems is basically a computational matter. So the theory of flag algebras allows us to harness computational power and apply it to problems in extremal combinatorics; it can be understood as part of the growing trend for the use of computers in mathematics.

Razborov credits Bondy [3] with a predecessor of the theory of flag algebras. Bondy applies counting techniques to the Caccetta–Häggkvist conjecture and illustrates his idea on Mantel's theorem. (The Cacceta–Häggkvist conjecture states that every simple directed graph on n vertices with outdegree at least r has a cycle with length at most $\lceil n/r \rceil$.) Here is a proof that $\mathrm{ex}(\circ\!-\!\circ, \{K_3\}) \leq \frac{1}{2}$ that is a rewording of the proof by Bondy in terms of densities and limit functionals. This proof is a first glance into the theory of flag algebras; in it we will derive by hand some constraints on limit functionals of sequences of triangle-free graphs and then give an explicit simple relaxation of Φ from which Mantel's theorem will follow.

A triangle-free graph may have three different graphs on three vertices as induced subgraphs: the empty graph $\overline{K_3}$, the graph with one edge $K_2 \cup K_1$, and the graph with two edges P_3. (Nonedges are represented by dashed lines.) Let G be a triangle-free graph. Every edge of G belongs to $|G|-2$ induced subgraphs with three vertices, whence

$$p(K_2\cup K_1; G) + 2p(P_3; G) = 3p(\circ\!-\!\circ; G).$$

This is valid for every triangle-free graph G, hence also for a limit functional ϕ:

$$\phi(K_2\cup K_1) + 2\phi(P_3) = 3\phi(\circ\!-\!\circ).$$

We have our first constraint satisfied for all $\phi \in \Phi$.

A second constraint comes from the identity

$$p(P_3; G) = \binom{|G|}{3}^{-1} \sum_{v\in V(G)} \binom{d(v)}{2},$$

where $d(v)$ is the degree of vertex v. To rewrite the right-hand side above, we need to extend the definition of the density function p to partially-labeled graphs. Say F and G are graphs each having a special vertex labeled 1, and let x_1 be the vertex of G labeled 1. Let $p(F; G)$ be the probability that a set $U \subseteq V(G)\setminus\{x_1\}$ with $|U| = |F|-1$, chosen uniformly at random, is such that $G[U\cup\{x_1\}]$ is isomorphic to F via a label-preserving isomorphism, that is, an isomorphism that takes the labeled vertex of F to the labeled vertex of G.

For $v \in V(G)$, denote by G^v the labeled graph obtained from G by labeling vertex v with label 1. Let $P_3^{\bullet}$ denote the labeled graph obtained from P_3 by labeling the vertex of degree two with label 1; similarly for other graphs the solid vertex will be the labeled vertex. Then for a triangle-free graph G we have

$$\begin{aligned} p(P_3; G) &= \binom{|G|}{3}^{-1} \sum_{v\in V(G)} \binom{d(v)}{2} \\ &= \binom{|G|}{3}^{-1} \sum_{v\in V(G)} p(P_3^{\bullet}; G^v)\binom{|G|-1}{2} \\ &= \frac{3}{|G|} \sum_{v\in V(G)} p(P_3^{\bullet}; G^v). \end{aligned} \qquad (3)$$

Now comes a key observation. As the size of G goes to infinity, $p(P_3^{\bullet}; G^v)$ goes to $p(\bullet\!-\!\circ; G^v)^2$. This is not hard to prove (do it!), but the intuition should be clear: if G is very large, then choosing a subset of $V(G)\setminus\{v\}$ of size 2 uniformly at random is basically the same as choosing two vertices in $V(G)\setminus\{v\}$ independently — the probability of choosing the same vertex twice becomes negligible as $|G|$ grows larger.

So let ϕ be the limit functional of a convergent sequence $(G_k)_{k\geq 0}$ of triangle-free graphs. Then

$$\begin{aligned} \phi(P_3) &= \lim_{k\to\infty} p(P_3; G_k) \\ &= \lim_{k\to\infty} \frac{3}{|G_k|} \sum_{v\in V(G_k)} p(P_3^{\bullet}; G_k^v) \\ &= \lim_{k\to\infty} \frac{3}{|G_k|} \sum_{v\in V(G_k)} p(\bullet\!-\!\circ; G_k^v)^2. \end{aligned} \qquad (4)$$

Now, for any triangle-free graph G the Cauchy-Schwarz inequality gives

$$\sum_{v\in V(G)} p(\bullet\!-\!\circ; G^v)^2 \geq \frac{1}{|G|}\left(\sum_{v\in V(G)} p(\bullet\!-\!\circ; G^v)\right)^2.$$

Together with (4) and

$$\sum_{v\in V(G)} p(\bullet\!-\!\circ; G^v)(|G|-1) = 2p(\circ\!-\!\circ; G)\binom{|G|}{2}$$

we get

$$\phi(P_3) \geq \lim_{k\to\infty} 3p(\circ\!-\!\circ; G_k)^2 = 3\phi(\circ\!-\!\circ)^2.$$

So every limit functional ϕ satisfies the constraints

$$\phi(K_2\cup K_1) + 2\phi(P_3) = 3\phi(\circ\!-\!\circ),$$

$$\phi(P_3) \geq 3\phi(\circ\!-\!\circ)^2.$$

What do we get in (2) if we optimize over the set Φ' of all $\phi: \mathcal{G} \to [0,1]$ satisfying the constraints above? Well, suppose $\phi \in \Phi'$. Multiply the second constraint by 2 and subtract it from the first to get

$$\phi(K_2\cup K_1) \leq 3\phi(\circ\!-\!\circ) - 6\phi(\circ\!-\!\circ)^2.$$

Since $\phi(K_2\cup K_1) \geq 0$, we then have $\phi(\circ\!-\!\circ) \leq \frac{1}{2}$. So the optimal value of (2) with Φ' instead of Φ is at most $\frac{1}{2}$, hence $\mathrm{ex}(\circ\!-\!\circ, \{K_3\}) \leq \frac{1}{2}$.

In the following sections the main points of Razborov's theory of flag algebras are developed. Unless otherwise noted, every definition and result presented here can be found in Razborov's original paper [12].

Types and flags

In the introduction, we derived valid inequalities for Φ by combining densities of partially-labeled graphs as in (3). In the next few sections we will develop Razborov's theory of flag algebras, which automates this process. The discussion will be focused on families of graphs for concreteness, though one of the most attractive features of the theory is that it applies to a whole range of structures, including directed graphs, hypergraphs, and permutations.

For an integer $k \geq 0$, write $[k] = \{1,\ldots,k\}$. Fix a family $\mathcal{H}$ of forbidden subgraphs. A *type* of *size* k is an $\mathcal{H}$-free graph σ with $V(\sigma) = [k]$. We can think of it as a graph with vertices labeled with $1,\ldots,k$, whereas we regard graphs as unlabeled. The empty type is denoted by $\varnothing$.

Let σ be a type of size k and F be a graph on at least k vertices. An *embedding* of σ into F is an injective function $\theta : [k] \to V(F)$ that defines an isomorphism between σ and the subgraph of F induced by $\operatorname{Im}\theta$.

A *σ-flag* is a pair (F,θ) where F is an $\mathcal{H}$-free graph and θ is an embedding of σ into F. So a σ-flag is a partially-labeled graph that avoids $\mathcal{H}$ and whose labeled part is a copy of σ. When the embedding itself is not important, we will drop it, speaking simply of the σ-flag F.

The *labeled vertices* of (F,θ) are the vertices in the image of θ. Note that an $\varnothing$-flag is just *an $\mathcal{H}$-free graph*. Any type σ of size k can also be seen as the σ-flag (σ,θ) where θ is the identity on $[k]$.

Isomorphism between σ-flags is defined just as for graphs, but now the labels should also be preserved by the bijection. More precisely, σ-flags (F,θ) and (G,η) are *isomorphic* if there is a graph isomorphism $\rho : V(F) \to V(G)$ between F and G such that $\rho(\theta(i)) = \eta(i)$ for $i = 1,\ldots,|\sigma|$. Write $(F,\theta) \simeq (G,\eta)$ when (F,θ) and (G,η) are isomorphic, or simply $F \simeq G$ when the embeddings are not important. In the introduction, this notion was used only for σ-flags where σ is the type of size 1. Figure 1 shows some flags of different types.

For $n \geq |\sigma|$, denote by $\mathcal{F}_n^\sigma$ the set of all σ-flags of size n, taken up to isomorphism; denote by $\mathcal{F}^\sigma$ the set of all σ-flags taken up to isomorphism. Note that the set $\mathcal{G}$ of all $\mathcal{H}$-free graphs is simply $\mathcal{F}^\varnothing$. A type σ is *degenerate* if $\mathcal{F}^\sigma$ is finite. If σ is nondegenerate, then $\mathcal{F}_n^\sigma \neq \emptyset$ for all $n \geq |\sigma|$. It is easy to construct a family $\mathcal{H}$ for which there are degenerate types: take for instance $\mathcal{H}$ as the set of all graphs with 1000 vertices containing at least one triangle. Then the triangle itself is $\mathcal{H}$-free, and

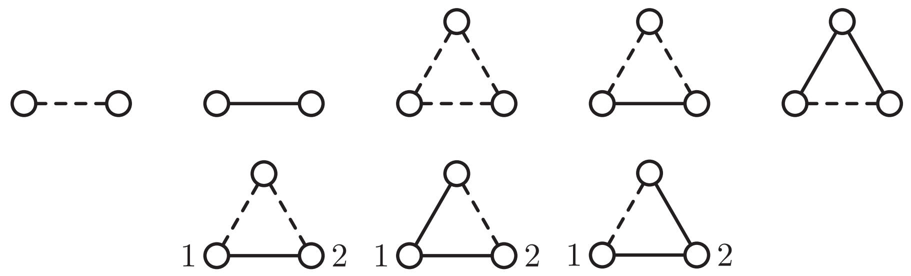


Figure 1 Let $\mathcal{H}=\{\ \}$. On the top row we have all $\varnothing$-flags of sizes 2 and 3, up to isomorphism (nonedges are shown as dashed lines); notice that the triangle itself is not a flag. On the bottom row we have all flags of type $\sigma = 1\!-\!2$; notice that the last two of these flags are not isomorphic, since the isomorphism has to preserve the labels.

hence is a type, but there are no -flags of size ≥ 1000.

From now on, we assume that all types are nondegenerate. In particular, every time a result about σ-flags is stated, it is implicitly assumed that σ is nondegenerate.

Density

The definition of density given in the introduction can be extended to σ-flags as follows. We say that σ-flags $F_1,\ldots,F_t$ *fit* in a σ-flag G if

$$|G| - |\sigma| \geq (|F_1| - |\sigma|) + \cdots + (|F_t| - |\sigma|).$$

Let $F_1,\ldots,F_t$ and (G,θ) be σ-flags such that $F_1,\ldots,F_t$ fit in G. Consider the following experiment: choose pairwise-disjoint sets $U_1,\ldots,U_t \subseteq V(G) \setminus \operatorname{Im}\theta$ of unlabeled vertices of G with $|U_i| = |F_i| - |\sigma|$ uniformly at random. Let $p(F_1,\ldots,F_t;G)$ be the probability that the σ-flag $(G[U_i \cup \operatorname{Im}\theta],\theta)$ is isomorphic to F_i for $i = 1,\ldots,t$. This is the *density* of $F_1,\ldots,F_t$ in G. For $\varnothing$-flags and $t = 1$, this definition coincides with the usual notion of density for graphs. In the introduction we also extended the definition of density to graphs with one labeled vertex; this corresponds to taking $t = 1$ and the only type of size 1 as σ.

Say $|F| \leq n \leq |G|$. To embed F into G, we may first try to embed F into a σ-flag F' of size n and then embed F' into G. This gives us another way to compute $p(F;G)$:

$$p(F;G) = \sum_{F' \in \mathcal{F}_n^\sigma} p(F;F')\,p(F';G). \tag{5}$$

This identity can be generalized, giving us the *chain rule*:

Theorem 1. *If $F_1,\ldots,F_t$, and G are σ-flags such that $F_1,\ldots,F_t$ fit in G, then for every $1 \leq s \leq t$ and every n such that $F_1,\ldots,F_s$ fit in a σ-flag of size n and a σ-flag of size n together with $F_{s+1},\ldots,F_t$ fit in G, the identity*

$$p(F_1,\ldots,F_t;G) = \sum_{F \in \mathcal{F}_n^\sigma} p(F_1,\ldots,F_s;F)\,p(F,F_{s+1},\ldots,F_t;G)$$

holds.

Recall from the introduction that $p(\ ;G^v) \to p(\ ;G^v)^2$ as $|G| \to \infty$. The argument to see this can be rephrased in two steps as follows. First, since G is triangle-free, then $p(\ ;G^v) = p(\ ,\ ;G^v)$. This can be seen directly, but is also a consequence of the chain rule. Indeed, let $\mathcal{H} = \{\ \}$ and let $\bullet$ denote the only type of size 1. Then $\bullet$-flags , fit in a $\bullet$-flag of size 3. Since $\mathcal{F}_3^\bullet = \{\ ,\ ,\ ,\ ,\ \}$, the chain rule gives

$$\begin{aligned} p(\ ,\ ;G) &= \sum_{F' \in \mathcal{F}_3^\bullet} p(\ ,\ ;F')\,p(F';G) \\ &= p(\ ;G). \end{aligned} \tag{6}$$

Second, $p(\ ,\ ;G^v) \to p(\ ;G^v)^2$ as $|G| \to \infty$, that is, density exhibits multiplicative behavior in the limit:

Theorem 2. *If F_1, F_2 are fixed σ-flags, then there exists a function $f(n) = O(1/n)$ such that if F_1, F_2 fit in a σ-flag G, then*

$$|p(F_1,F_2;G) - p(F_1;G)\,p(F_2;G)| \leq f(|G|).$$

Identity (6), that comes from an application of the chain rule, suggests that there is a relation between the pair (,) and . In the next section, we will use the chain rule to define a product operation on σ-flags, and under this product it will hold that $\cdot$ = . This product will also commute with the density function in the limit: for σ-flags F_1 and F_2 we will have $p(F_1 \cdot F_2;G) \to p(F_1;G)\,p(F_2;G)$ as $|G| \to \infty$.

Flag algebras

In the introduction, we derived the constraint

$$\phi(\;) + 2\phi(\;) = 3\phi(\;),$$

valid for every $\phi \in \Phi$. If we see $\phi \in [0,1]^{\mathcal{G}}$ as a vector, then this is a linear constraint on the components of ϕ.

To enable the use of tools from optimization, mainly duality, we need to embed our domain into a vector space. We do so by extending ϕ linearly to the space $\mathbb{R}\mathcal{G}$ of formal real linear combinations of graphs in $\mathcal{G}$. We could then rewrite the latter constraint as

$$\phi(\;+2\;) = \phi(3\;),$$

or even

$$\phi(\;+2\;-3\;) = 0.$$

One of our main goals is to characterize the linear functionals on $\mathbb{R}\mathcal{G}$ that are limit functionals. Instead of describing all the constraints that characterize limit functionals, it is convenient to encode some of them algebraically, that is, by modifying the algebraic structure of $\mathbb{R}\mathcal{G}$. The resulting algebraic object will be the flag algebra, which we construct now for the more general case of σ-flags.

Let $\mathbb{R}\mathcal{F}^\sigma$ be the free vector space over the reals generated by all σ-flags, i.e., $\mathbb{R}\mathcal{F}^\sigma$ is the space of all formal real linear combinations of σ-flags. Let $(A_k)_{k\geq 0}$ be a convergent sequence in $\mathcal{F}^\sigma$ and let

$$\phi(F) = \lim_{k\to\infty} p(F; A_k)$$

be the pointwise limit of the functions $p(\cdot\,; A_k)$. Extend ϕ linearly to $\mathbb{R}\mathcal{F}^\sigma$, obtaining a linear functional. We say that ϕ is the *limit functional* of the convergent sequence $(A_k)_{k\geq 0}$ or, when the sequence itself is not relevant, that it is a *limit functional*.

For any limit functional ϕ, the chain rule in its form (5) implies that for every σ-flag F and $n \geq |F|$ we have

$$\phi(F) = \phi\Bigg(\sum_{F' \in \mathcal{F}^\sigma_n} p(F; F')F'\Bigg),$$

that is,

$$F - \sum_{F' \in \mathcal{F}^\sigma_n} p(F; F')F' \tag{7}$$

is in the kernel of ϕ. Instead of enforcing these infinitely many relations, we might as well just quotient them out. So let $\mathcal{K}^\sigma$ be the linear span of vectors of form (7) and define $\mathcal{A}^\sigma = \mathbb{R}\mathcal{F}^\sigma/\mathcal{K}^\sigma$. This is a nontrivial vector space, since for every σ-flag F we have $p(\sigma; F) = 1$, and hence σ is itself not in $\mathcal{K}^\sigma$. Since $\mathcal{K}^\sigma$ is contained in the kernel of every limit functional, every limit functional is also a linear functional of $\mathcal{A}^\sigma$.

The main advantage of working with $\mathcal{A}^\sigma$ instead of $\mathbb{R}\mathcal{F}^\sigma$ is that it is possible to define a product on $\mathcal{A}^\sigma$, turning it into an algebra. This product will conveniently encode the asymptotic multiplicative behavior of densities described in Theorem 2: for every limit functional ϕ and $f, g \in \mathcal{A}^\sigma$we will have $\phi(f\cdot g) = \phi(f)\phi(g)$.

For σ-flags F and G, let n be any integer such that F, G fit in a σ-flag of size n and set

$$F\cdot G = \Bigg(\sum_{H\in\mathcal{F}^\sigma_n} p(F, G; H)H\Bigg) + \mathcal{K}^\sigma. \tag{8}$$

This defines a function from $\mathcal{F}^\sigma \times \mathcal{F}^\sigma$ to $\mathcal{A}^\sigma$ and one may show that the definition is independent of the choice of n for each pair (F, G) of σ-flags. Now, extend this function bilinearly to $\mathbb{R}\mathcal{F}^\sigma \times \mathbb{R}\mathcal{F}^\sigma$. It is possible to prove that if $f \in \mathcal{K}^\sigma$ and $g \in \mathbb{R}\mathcal{F}^\sigma$, then $f\cdot g = \mathcal{K}^\sigma$, whence the bilinear extension is constant on cosets, and therefore defines a symmetric bilinear form on $\mathcal{A}^\sigma$, that is, a commutative product.

This turns $\mathcal{A}^\sigma$ into an algebra, the *flag algebra of type* σ. The product on $\mathcal{A}^\sigma$ is now defined, and we will use henceforth the natural correspondence $f \mapsto f + \mathcal{K}^\sigma$ between $\mathbb{R}\mathcal{F}^\sigma$ and $\mathcal{A}^\sigma$ without further notice, i.e., we will omit $\mathcal{K}^\sigma$ and write f instead of $f + \mathcal{K}^\sigma$ for an element of $\mathcal{A}^\sigma$. Sometimes, namely in the last section, it is important to work with explicit representatives of each coset; in such cases we will clearly distinguish between cosets and their representatives.

Under the product just defined for $\mathcal{A}^\sigma$, the type σ, taken as a σ-flag, is the identity element. The identity σ can be decomposed in many different ways using relations (7). Indeed, for any $n \geq |\sigma|$, we have

$$\sigma = \sum_{F\in\mathcal{F}^\sigma_n} p(\sigma; F)F = \sum_{F\in\mathcal{F}^\sigma_n} F.$$

It now follows from Theorem 2 that limit functionals are multiplicative, i.e.,

$$\phi(f\cdot g) = \phi(f)\cdot\phi(g)$$

for $f, g \in \mathcal{A}^\sigma$. Since by construction $\phi(\sigma) = 1$, every limit functional ϕ is an algebra homomorphism between $\mathcal{A}^\sigma$ and $\mathbb{R}$. We denote the set of all algebra homomorphisms between $\mathcal{A}^\sigma$ and $\mathbb{R}$ by $\operatorname{Hom}(\mathcal{A}^\sigma, \mathbb{R})$.

As an example, recall the discussion at the end of the previous section. When $\mathcal{H} = \{\;\}$, if we expand the product $\;\cdot\;$ as a linear combination of $\bullet$-flags of size 3, then $\;\cdot\;=\;$. Hence every limit functional ϕ satisfies $\phi(\;) = \phi(\;\cdot\;) = \phi(\;)^2$.

Every limit functional ϕ lies in $\operatorname{Hom}(\mathcal{A}^\sigma, \mathbb{R})$. Another obvious constraint that every limit functional ϕ must satisfy is $\phi(F) \geq 0$ for every σ-flag F, which is not necessarily true of all homomorphisms. Call $\phi \in \operatorname{Hom}(\mathcal{A}^\sigma, \mathbb{R})$ *positive* if $\phi(F) \geq 0$ for every σ-flag F, and let $\operatorname{Hom}^+(\mathcal{A}^\sigma, \mathbb{R})$ denote the set of all positive homomorphisms.

It turns out that these are all the essential properties of a limit functional. It is clear that every limit functional is a positive homomorphism. The following theorem of Razborov [12] establishes the converse, and so positive homomorphisms are precisely the limit objects of convergent sequences of flags. In particular, the linear extension of the set Φ is precisely $\operatorname{Hom}^+(\mathcal{A}^\varnothing, \mathbb{R})$.

Theorem 3. *Every limit functional is a positive homomorphism and every positive homomorphism is a limit functional.*

Finally, notice that types and flags are defined in terms of the family $\mathcal{H}$ of forbidden subgraphs, so this family is encoded in the construction of the flag algebra $\mathcal{A}^\sigma$.

Downward operator

We are really interested in working with $\varnothing$-flags, that is, unlabeled graphs, so why consider other types altogether? Most times, in order to obtain results for $\varnothing$-flags, it is necessary to use other types. In the introduction, to obtain Mantel's theorem, it was not enough to work with unlabeled graphs: at some point, we had to introduce labeled graphs, namely to get (3).

The downward operator maps σ-flags into $\varnothing$-flags, in such a way that we can derive valid inequalities for densities of $\varnothing$-flags from valid inequalities for densities of σ-flags. If types can be seen as a form of lifting, then the downward operator is a projection back to our space of interest.

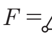
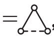
If F is a σ-flag, then $\downarrow F$ is the $\varnothing$-flag obtained from F simply by forgetting the embedding, that is, by forgetting the vertex labels. For a σ-flag F, let $q_\sigma(F)$ be the probability that an injective map $\theta:[k]\to V(F)$ taken uniformly at random is such that $(\downarrow F,\theta)$ is a σ-flag isomorphic to F and set

$$[\![F_\sigma]\!] = q_\sigma(F)\downarrow F,$$

then extend $[\![\cdot]\!]_\sigma$ linearly to $\mathbb{R}\mathcal{F}^\sigma$ to obtain a linear map from $\mathbb{R}\mathcal{F}^\sigma$ to $\mathbb{R}\mathcal{F}^\varnothing$. One key property of this map is that $[\![\mathcal{K}^\sigma]\!]_\sigma \subseteq \mathcal{K}^\varnothing$, and hence $[\![\cdot]\!]_\sigma$ gives a linear map from $\mathcal{A}^\sigma$ to $\mathcal{A}^\varnothing$, which we call *downward operator*. The main tool used in the proof of this result is the following lemma, which relates densities in the labeled and in the unlabeled cases by taking an average.

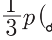
Lemma 4. *Let F be a σ-flag and G be an $\varnothing$-flag with $|G|\geq|F|$ and $p(\downarrow\sigma;G)>0$. If θ is an embedding of σ into G chosen uniformly at random, then* $p(F;(G,\theta))$ is a random variable and

$$\mathbb{E}[p(F;(G,\theta))] = \frac{q_\sigma(F)\,p(\downarrow F;G)}{q_\sigma(\sigma)\,p(\downarrow\sigma;G)}.$$

Note that equation (3) in the introduction follows trivially from this lemma. Indeed, take $\sigma = \bullet$ as the type of size 1 and let $F =$ . Then $\downarrow F =$ , $q_\sigma(F)=\frac{1}{3}$, $q_\sigma(\sigma)=1$, and $p(\downarrow\sigma;G)=1$ for any graph G. Thus, by Lemma 4,

$\frac{1}{|G|}\sum_{v\in V(G)} p($  $;G^v) = \mathbb{E}[p(F;(G,\theta))]$

$$= \frac{q_\sigma(F)\,p(\downarrow F;G)}{q_\sigma(\sigma)\,p(\downarrow\sigma;G)}$$

$= \frac{1}{3}p($  $;G).$

Conic programming

For $f\in\mathcal{A}^\sigma$ and a linear functional ϕ in the dual space $(\mathcal{A}^\sigma)^*$ of $\mathcal{A}^\sigma$, write $(\phi,f)=\phi(f)$. The *semantic cone* of type σ is the set

$$\mathcal{S}^\sigma = \{f\in\mathcal{A}^\sigma : (\phi,f)\geq 0 \text{ for all } \phi\in\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})\}.$$

This is a convex cone and its dual cone

$$(\mathcal{S}^\sigma)^* = \{\phi\in(\mathcal{A}^\sigma)^* : (\phi,f)\geq 0 \text{ for all } f\in\mathcal{S}^\sigma\}$$

contains every nonnegative multiple of functionals in $\mathrm{Hom}^+(\mathcal{A}^\sigma,\mathbb{R})$. So, given a graph C,

$$\begin{aligned}&\max\{(\phi,C):\phi\in\mathrm{Hom}^+(\mathcal{A}^\varnothing,\mathbb{R})\}\\ &\leq\max\{(\phi,C):\phi\in(\mathcal{S}^\varnothing)^* \text{ and } (\phi,\varnothing)=1\}.\end{aligned} \quad (9)$$

(Here we may write '$\max$' instead of '$\sup$' because $\mathrm{Hom}^+(\mathcal{A}^\varnothing,\mathbb{R})$ is compact. Actually, equality holds by the bipolar theorem.)

The optimization problem on the right-hand side above is a *conic programming problem*. It asks us to maximize a linear function $\phi\mapsto(\phi,C)$ over the intersection of a cone, namely $(\mathcal{S}^\varnothing)^*$, and an affine subspace, in our case determined by the linear equation $(\phi,\varnothing)=1$.

This conic programming problem has a *dual problem*, namely

$$\min\{\lambda : \lambda\varnothing - C\in\mathcal{S}^\varnothing \text{ and } \lambda\in\mathbb{R}\}, \quad (10)$$

where the optimization variable is λ. (We may write '$\min$' instead of '$\inf$' because the feasible region is a closed half-line in $\mathbb{R}$.)

Weak duality holds: any feasible solution of the dual has larger or equal objective value than any feasible solution of the primal. Indeed, if $\phi\in(\mathcal{S}^\varnothing)^*$ is such that $(\phi,\varnothing)=1$ and $\lambda\in\mathbb{R}$ is such that $\lambda\varnothing - C\in\mathcal{S}^\varnothing$, then

$$0\leq(\phi,\lambda\varnothing - C)=\lambda-(\phi,C).$$

Actually, it is easy to show that there is no duality gap, that is, that primal and dual have the same optimal value. Even more: the problem on the left-hand side of (9) has the same optimal value of the dual problem (10), and so all three optimization problems in (9) and (10) have the same optimal value. Indeed, notice that the maximum on the left-hand side of (9) is equal to

$$\min\{\lambda:(\phi,C)\leq\lambda \text{ for all } \phi\in\mathrm{Hom}^+(\mathcal{A}^\varnothing,\mathbb{R})\}.$$

Now, $\lambda\geq(\phi,C)$ for all $\phi\in\mathrm{Hom}^+(\mathcal{A}^\varnothing,\mathbb{R})$ if and only if $(\phi,\lambda\varnothing - C)\geq 0$ for all $\phi\in\mathrm{Hom}^+(\mathcal{A}^\varnothing,\mathbb{R})$ if and only if $\lambda\varnothing - C\in\mathcal{S}^\varnothing$, as we wanted.

To find an upper bound for $\mathrm{ex}(C,\mathcal{H})$ we work with the dual problem (10). One advantage is that we do not need to solve this problem to optimality to find an upper bound, since any feasible solution provides an upper bound. Solving (10) to optimality is the same as solving the primal problem to optimality, which is the same as computing $\mathrm{ex}(C,\mathcal{H})$.

One way to simplify the dual problem (10) is to replace $\mathcal{S}^\varnothing$ with a cone $C\subseteq\mathcal{S}^\varnothing$ for which it is easier to solve the resulting problem. Obviously, we still get a valid upper bound. We seem to have taken a tortuous path since the introduction, where we stated our goal of finding a relaxation of Φ, of which $\mathrm{Hom}^+(\mathcal{A}^\varnothing,\mathbb{R})$ is the linear extension, but that is exactly what we achieved, albeit via the dual:

$$\mathrm{Hom}^+(\mathcal{A}^\varnothing,\mathbb{R})\subseteq\{\phi\in(\mathcal{A}^\varnothing)^*:(\phi,f)\geq 0 \text{ for all } f\in C \text{ and } (\phi,\varnothing)=1\}.$$

What are some $f\in\mathcal{A}^\sigma$ that belong to the semantic cone $\mathcal{S}^\sigma$? Since a positive homomorphism ϕ is by definition nonnegative on every σ-flag F, then any conic combination of σ-flags is in the semantic cone. Another class of vectors in the semantic cone is the class of vectors that are sums of squares. We say that $f\in\mathcal{A}^\sigma$ is a *sum of squares* if there are $g_1,\ldots,g_t\in\mathcal{A}^\sigma$ such that $f=g_1^2+\cdots+g_t^2$. Then for any positive homomorphism ϕ (actually, for any homomorphism) we have $\phi(f)=\phi(g_1)^2+\cdots+\phi(g_t)^2\geq 0$. The class of sum-of-squares vectors is particularly interesting because it is computationally tractable, as we will soon see. Finally, the downward operator maps the semantic cone $\mathcal{S}^\sigma$of type σ into the semantic cone $\mathcal{S}^\varnothing$ of type $\varnothing$:

Theorem 5. *The image of $\mathcal{S}^\sigma$ under $[\![\cdot]\!]_\sigma$ is a subset of $\mathcal{S}^\varnothing$.*

This gives yet another way to obtain vectors in $\mathcal{S}^\varnothing$, by first considering a type σ, then obtaining a vector in $\mathcal{A}^\sigma$ (a sum-of-squares vector, for instance), and then using the downward operator.

The semidefinite programming method

Semidefinite programming is conic programming over the cone of positive semidefinite matrices. Using sum-of-squares vectors in $\mathcal{A}^\sigma$ and the downward operator, we may define a family of tractable cones contained in $\mathcal{S}^\varnothing$. Then using semidefinite programming it is possible to write down optimization problems that provide upper bounds to (10). This approach is known as the *semidefinite programming method*. Its main advantages are that writing down the semidefinite programming problems is mostly a mechanical affair, that can even be automated (and has been; see for instance flagmatic [5]), and solving the resulting problems can be done with a computer.

There is a well-known relation between sums-of-squares polynomials and positive semidefinite matrices (see e.g. the exposition by Laurent [9]). We now establish the analogous relation between sums-of-squares vectors in $\mathcal{A}^\sigma$ and positive semidefinite matrices. The *degree* of a vector $f \in \mathbb{R}\mathcal{F}^\sigma$ is the largest size of a flag appearing with a nonzero coefficient in the expansion of f; by convention, the degree of 0 is -1. The notion of degree can be extended to $\mathcal{A}^\sigma$, by setting the degree of $f + \mathcal{K}^\sigma \in \mathcal{A}^\sigma$ to be the smallest degree of any $g \in f + \mathcal{K}^\sigma$. For a type σ and $n \geq |\sigma|$, let $v_{\sigma,n} \colon \mathcal{F}_n^\sigma \to \mathcal{A}^\sigma$ be the canonical embedding, i.e., $v_{\sigma,n}(F) = F$ for all $F \in \mathcal{F}_n^\sigma$.

Theorem 6. *If $f \in \mathcal{A}^\sigma$ and $n \geq |\sigma|$, then there are vectors $g_1, \ldots, g_t \in \mathcal{A}^\sigma$ for some $t \geq 1$, each of degree at most n, such that $f = g_1^2 + \cdots + g_t^2$ if and only if there is a positive semidefinite matrix $Q \colon \mathcal{F}_n^\sigma \times \mathcal{F}_n^\sigma \to \mathbb{R}$ such that $f = v_{\sigma,n}^\top Q v_{\sigma,n}$.*

Proof. Suppose that there are vectors $g_1, \ldots, g_t$ as described. Modulo $\mathcal{K}^\sigma$, every σ-flag of size m can be written as a linear combination of σ-flags of any fixed size greater than m. So by hypothesis we can take from each coset $g_i + \mathcal{K}^\sigma$ a representative $\hat{g}_i \in \mathbb{R}\mathcal{F}^\sigma$ which is a linear combination of σ-flags of size n.

Let c_i be the vector of coefficients of $\hat{g}_i$, in such a way that $\hat{g}_i = c_i^\top v_{\sigma,n}$. Then

$$\hat{g}_1^2 + \cdots + \hat{g}_t^2 = \sum_{i=1}^t \left(c_i^\top v_{\sigma,n}\right)^2 = \sum_{i=1}^t v_{\sigma,n}^\top c_i c_i^\top v_{\sigma,n},$$

and we may take $Q = c_1 c_1^\top + \cdots + c_t c_t^\top$.

For the converse, say there is a positive semidefinite matrix Q as described. Then for some t there are vectors $c_1, \ldots, c_t$ such that $Q = c_1 c_1^\top + \cdots + c_t c_t^\top$. But then $g_i = c_i^\top v_{\sigma,n}$ has degree at most n in $\mathcal{A}^\sigma$. Moreover, $f = g_1^2 + \cdots + g_t^2$, as we wanted. □

Let us describe the semidefinite programming method by applying it to Mantel's theorem. Fix $\mathcal{H} = \{\ \}$. We have the following $\varnothing$-flags of sizes 2 and 3: , , , and . There is also only one type of size 1, namely the graph on one vertex, which we denote by $\bullet$. These are the $\bullet$-flags of sizes 2 and 3: , , , , , , and .

Write $v = v_{\bullet,2}$, so that in vector notation we have $v = (\ ,\ )$. From Theorem 6, if $Q \colon \mathcal{F}_2^\bullet \times \mathcal{F}_2^\bullet \to \mathbb{R}$ is a positive semidefinite matrix, then $v^\top Q v$ belongs to the semantic cone $\mathcal{S}^\bullet$ of type $\bullet$, and hence from Theorem 5 we have that $[\![v^\top Q v]\!]_\bullet$ belongs to the semantic cone $\mathcal{S}^\varnothing$ of type $\varnothing$. Since any conic combination r of $\varnothing$-flags belongs to the semantic cone $\mathcal{S}^\varnothing$, we have that

$$r + [\![v^\top Q v]\!]_\bullet \in \mathcal{S}^\varnothing$$

for every conic combination r of $\varnothing$-flags and every positive semidefinite matrix Q.

So, recalling (10), any feasible solution of the following optimization problem gives an upper bound to $\mathrm{ex}(\ , \{\ \})$:

$$\begin{aligned} &\min \lambda \\ &\quad \lambda\varnothing - \ = r + [\![v^\top Q v]\!]_\bullet, \\ &\quad r \text{ is a conic combination of } \varnothing\text{-flags}, \\ &\quad Q \colon \mathcal{F}_2^\bullet \times \mathcal{F}_2^\bullet \to \mathbb{R} \text{ is positive semidefinite.} \end{aligned} \tag{11}$$

This problem is not quite a semidefinite programming problem: the first identity above is an identity between vectors in $\mathcal{A}^\varnothing$, not a linear constraint on λ and the entries of Q. This identity can be translated, however, into several linear constraints, as follows.

If A and B are $n \times n$ matrices, write $\langle A, B \rangle = \operatorname{tr} A^\top B = \sum_{i,j=1}^n A_{ij} B_{ij}$. Then

$$[\![v^\top Q v]\!]_\bullet = [\![\langle vv^\top, Q \rangle]\!]_\bullet = \langle [\![vv^\top]\!]_\bullet, Q \rangle.$$

Here, notice that $vv^\top$ is a matrix. The downward operator, when applied to the matrix $vv^\top$, is applied entrywise and yields a matrix of the same dimensions as the result.

So the first constraint in (11) can be rewritten as

$$\lambda\varnothing - \ = r + \langle [\![vv^\top]\!]_\bullet, Q \rangle, \tag{12}$$

which is still an identity between elements of $\mathcal{A}^\varnothing$. To test the above identity, we may choose a large enough N and use the chain rule to expand both left and right-hand sides as linear combinations of $\varnothing$-flags of size N. If the coefficients coincide, then equality holds. This is only a sufficient condition however: for a fixed N, equality may hold in $\mathcal{A}^\varnothing$ even though the coefficients differ, but it is not hard to show that there is always some N for which equality holds if and only if the coefficients coincide.

To make things precise, we have to choose for , r, and every element of $\mathcal{A}^\varnothing$ in $vv^\top$ a representative in $\mathbb{R}\mathcal{F}^\varnothing$. As a representative of $\ \in \mathcal{A}^\varnothing$ we may choose $\ \in \mathbb{R}\mathcal{F}^\varnothing$. For $vv^\top$ proceed as follows: use the definition of product in $\mathcal{A}^\bullet$ to get

$$vv^\top = \begin{pmatrix} \ + \ & \frac{1}{2}(\ + \ ) \\ \frac{1}{2}(\ + \ ) & \ \end{pmatrix}$$

and then apply the downward operator to get

$$[\![vv^\top]\!]_\bullet = \begin{pmatrix} \ + \frac{1}{3}\ & \frac{1}{3}(\ + \ ) \\ \frac{1}{3}(\ + \ ) & \frac{1}{3}\ \end{pmatrix}.$$

We will deal with r below in a different way (actually, we will get rid of it). Notice we could have chosen different representatives. For instance, we could have expanded the products in $vv^\top$ using $\bullet$-flags of size 6, say. All that matters, however, is to choose representatives, and it is usually a good idea to choose representatives of smallest possible degree.

Now we are working exclusively with representatives in $\mathbb{R}\mathcal{F}^\varnothing$. For a given $N > 0$ and fixed $G \in \mathcal{F}_N^\varnothing$, extend $F \mapsto p(F; G)$ linearly to $\mathcal{F}_N^\varnothing$. If for every $G \in \mathcal{F}_N^\varnothing$ we have

$$p(\lambda\varnothing - \ ; G) = p(r; G) + p(\langle [\![vv^\top]\!]_\bullet, Q \rangle; G), \tag{13}$$

then (12) holds. Conversely, if (12) holds, then for some $N > 0$ (13) holds for every $G \in \mathcal{F}_N^\varnothing$ (this requires a short argument though).

Now, $p(r; G)$ is the coefficient of G in r; then, since r is a conic combination, $p(r; G) \geq 0$ for every $G \in \mathcal{F}_N^\varnothing$. Together with linearity this implies that we may rewrite (13) equivalently as

$$\lambda - p(\ ; G) \geq \langle p([\![vv^\top]\!]_\bullet; G), Q \rangle, \tag{14}$$

where $p(\cdot\,; G)$ is applied entrywise to $vv^\top$. Notice that $p(\ ; G)$ is a number and $p([\![vv^\top]\!]_\bullet; G)$ is a matrix of numbers, so for each $G \in \mathcal{F}_N^\varnothing$ the above inequality is a linear constraint on λ and the entries of Q.

In our case, we may take $N = 3$. Then (14) gives rise to one linear constraint for each of the $\varnothing$-flags of size 3:

$\varnothing$-flag	constraint
	$\lambda \geq \left\langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, Q \right\rangle$,
	$\lambda - \frac{1}{3} \geq \left\langle \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 \end{pmatrix}, Q \right\rangle$,
	$\lambda - \frac{2}{3} \geq \left\langle \begin{pmatrix} 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix}, Q \right\rangle$.

In this way we may rewrite problem (11), obtaining a semidefinite programming problem that gives an upper bound to the optimal value of (11), and hence also to $\mathrm{ex}(\ , \{\ \})$. This problem is not necessarily equivalent to (11), since for a given N equality in the algebra may hold even though the linear constraints are not satisfied.

Now, it is easy to check that $\lambda = \frac{1}{2}$ and $Q = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ form a feasible solution of this semidefinite programming problem (and hence also of (11)), and so we have Mantel's theorem.

All the steps of the semidefinite programming method are contained in the example we worked out above. In general, however, one may choose a finite set $\mathcal{T}$ of types instead of only one type and consider the vectors in $\mathcal{S}^{\varnothing}$ given by

$$r + \sum_{\sigma \in \mathcal{T}} \left[\!\left[v_{\sigma,n_\sigma}^{\top} Q_\sigma v_{\sigma,n_\sigma} \right]\!\right]_\sigma,$$

where r is a conic combination of $\varnothing$-flags, $n_\sigma \geq |\sigma|$, and each Q_σ is a positive semidefinite matrix. Choosing more types makes the problem larger, but also potentially stronger.

Summary

The theory of flag algebras provides a powerful, unifying approach for extremal problems involving a host of combinatorial structures. Its novelty is that it allows the formulation of relaxations for such problems using conic programming, which can be further relaxed to semidefinite programming problems, thus enabling the use of a computer to obtain bounds. Most importantly, the computed bounds are often tight. Hence, the theory yields relaxations that achieve the desired trade-off of computational tractability and high-quality bounds.

We have only scratched the surface of the theory of flag algebras. Many optimization aspects of the semidefinite method, such as the use of complementary slackness to obtain further constraints on the optimal solutions for (9), were left out. Complementary slackness can be useful to show properties of all increasing sequences $(G_k)_{k\geq 0}$ that attain $\mathrm{ex}(C, \mathcal{H})$, an important issue in extremal combinatorics. Razborov [12] further developed other methods involving flag algebras, such as the differential method and the inductive method.

Techniques involving flag algebras have been used to obtain many significant new results such as: computing the minimal number of triangles in graphs with given edge density [11, 13], computing the maximum number of pentagons in triangle-free graphs [6, 8], and obtaining new advances towards the Cacceta-Häggkvist conjecture [14]. Besides being applied in the context of graphs and digraphs, flag algebras have also been successfully used in the setting of colored graphs [1, 4] and of permutations [2]. For many more references, see the thesis of Grzesik [7].

References

1 R. Baber and J. Talbot, A solution to the 2/3 conjecture, *SIAM Journal on Discrete Mathematics* 28 (2014), 756–766.

2 J. Balogh, P. Hu, B. Lidický, O. Pikhurko, B. Udvari and J. Volec, Minimum Number of Monotone Subsequences of Length 4 in Permutations, *Combinatorics, Probability and Computing* 24(4) (2015), 658–679.

3 J.A. Bondy, Counting subgraphs: a new approach to the Caccetta-Häggkvist conjecture, *Discrete Mathematics* 165/166 (1997), 71–80.

4 J. Cummings, D. Král', F. Pfender, K. Sperfeld, A. Treglown and M. Young, Monochromatic triangles in three-coloured graphs, *Journal of Combinatorial Theory, Series B* 103 (2013), 489–503.

5 V. Falgas-Ravry and E.R. Vaughan, Applications of the semi-definite method to the Turán density problem for 3-graphs, *Combinatorics, Probability and Computing* 22 (2013), 21–54.

6 A. Grzesik, On the maximum number of five-cycles in a triangle-free graph, *Journal of Combinatorial Theory, Series B* 102 (2012), 1061–1066.

7 A. Grzesik, *Flag Algebras in Extremal Graph Theory*, PhD Thesis, Jagiellonian University, 2014.

8 H. Hatami, J. Hladký, D. Král', S. Norine and A. Razborov, On the number of pentagons in triangle-free graphs, *Journal of Combinatorial Theory, Series A* 120 (2013), 722–732.

9 M. Laurent, Semidefinite programming in combinatorial and polynomial optimization, *Nieuw Archief voor Wiskunde* 5/9 (2008), 256–262.

10 W. Mantel, Vraagstuk XXVIII, *Wiskundige Opgaven* 10 (1910), 60–61.

11 O. Pikhurko and A. Razborov, Asymptotic structure of graphs with the minimum number of triangles, *Combinatorics, Probability & Computing*, First View (2016), 1–23.

12 A. Razborov, Flag algebras, *Journal of Symbolic Logic* 72 (2007), 1239–1282.

13 A. Razborov, On the minimal density of triangles in graphs, *Combinatorics, Probability & Computing* 17(4) (2008), 603–618.

14 A. Razborov, On the Caccetta-Häggkvist Conjecture with Forbidden Subgraphs, *Journal of Graph Theory* 74 (2013), 236–248.