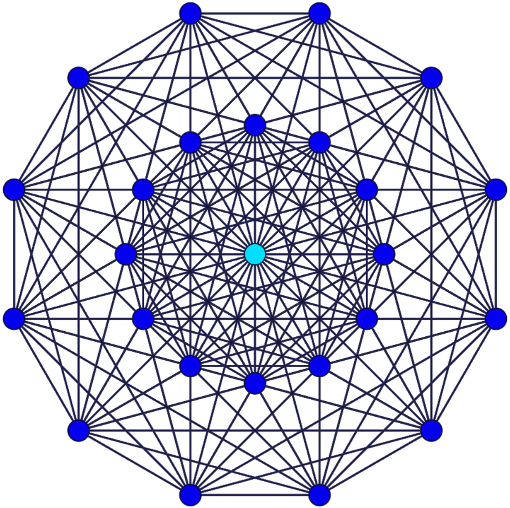
Computational challenges in Flag Algebra proofs

ICIAM 2023: Advances in Optimization I

Waseda University, Tokyo

Christoph Spiegel (Zuse Institute Berlin)

22nd of August 2023



Results are joint work with ...

Aldo Kiem
ZIB / TU Berlin

Sebastian Pokutta
ZIB / TU Berlin

Research partially funded through Math+ project EF1-21

Computational challenges in Flag Algebra proofs

The Ramsey Multiplicity Problem

Theorem (Ramsey 1930 – Multicolor Version)

For any $t \in \mathbb{N}$ and $c \geq 2$ there exists $R_c(t) \in \mathbb{N}$ such that any c-edge-coloring of the complete graph of order at least $R_c(t)$ contains a monochromatic clique of size t.

A well-known question

Can we determine $R_c(t)$?

A related question

How many cliques are required?

Theorem (Goodman 1959 – Asymptotic Version)

Asymptotically at least $1/4$ of all triangles are monochromatic in any 2-edge-coloring.

The Ramsey Multiplicity Problem

Theorem (Ramsey 1930 – Multicolor Version)

For any $t \in \mathbb{N}$ and $c \geq 2$ there exists $R_c(t) \in \mathbb{N}$ such that any c-edge-coloring of the complete graph of order at least $R_c(t)$ contains a monochromatic clique of size t.

A well-known question

Can we determine $R_c(t)$?

A related question

How many cliques are required?

Theorem (Goodman 1959 – Asymptotic Version)

Asymptotically at least $1/4$ of all triangles are monochromatic in any 2-edge-coloring.

The Ramsey Multiplicity Problem

Theorem (Ramsey 1930 – Multicolor Version)

For any $t \in \mathbb{N}$ and $c \geq 2$ there exists $R_c(t) \in \mathbb{N}$ such that any c-edge-coloring of the complete graph of order at least $R_c(t)$ contains a monochromatic clique of size t.

A well-known question

Can we determine $R_c(t)$?

A related question

How many cliques are required?

Theorem (Goodman 1959 – Asymptotic Version)

Asymptotically at least $1/4$ of all triangles are monochromatic in any 2-edge-coloring.

Beyond Goodman's Result

Notation. Let $\mathcal{G}_n = \{G : E(K_n) \to [c]\}$ denote all c-edge-colorings of K_n, G_i the subgraph of K_n given by color i and $k_t(G_i)$ the fraction of t-cliques in G_i.

Problem (Ramsey Multiplicity)

What is the value of $m_c(t) = \lim_n \min_{G \in \mathcal{G}_n} k_t(G_1) + \ldots + k_t(G_c)$?

The success of the binomial random graph for $m_2(3)$ lead to the following conjecture.

Conjecture (Erdős 1962)

$m_2(t) = 2^{1-\binom{t}{2}}$ for any $t \geq 2$. **False for $t \geq 4$ (Thomason 1989)**

The exact value of even $m_2(4)$ remains unknown with little progress!

Beyond Goodman's Result

Notation. Let $\mathcal{G}_n = \{G : E(K_n) \to [c]\}$ denote all c-edge-colorings of K_n, G_i the subgraph of K_n given by color i and $k_t(G_i)$ the fraction of t-cliques in G_i.

Problem (Ramsey Multiplicity)

What is the value of $m_c(t) = \lim_n \min_{G \in \mathcal{G}_n} k_t(G_1) + \ldots + k_t(G_c)$?

The success of the binomial random graph for $m_2(3)$ lead to the following conjecture.

Conjecture (Erdős 1962)

$m_2(t) = 2^{1-\binom{t}{2}}$ for any $t \geq 2$. **False for $t \geq 4$ (Thomason 1989)**

The exact value of even $m_2(4)$ remains unknown with little progress!

Beyond Goodman's Result

Notation. Let $\mathcal{G}_n = \{G : E(K_n) \to [c]\}$ denote all c-edge-colorings of K_n, G_i the subgraph of K_n given by color i and $k_t(G_i)$ the fraction of t-cliques in G_i.

Problem (Ramsey Multiplicity)

What is the value of $m_c(t) = \lim_n \min_{G \in \mathcal{G}_n} k_t(G_1) + \ldots + k_t(G_c)$?

The success of the binomial random graph for $m_2(3)$ lead to the following conjecture.

Conjecture (Erdős 1962)

$m_2(t) = 2^{1-\binom{t}{2}}$ for any $t \geq 2$. **False for $t \geq 4$ (Thomason 1989)**

The exact value of even $m_2(4)$ remains unknown with little progress!

More than two colors

Goodman also asked about more than two colors and Fox noted the upper bound

$$m_c(3) \leq 1/R_{c-1}(3)^2.$$

Theorem (Cummings et al. 2013 – Asymptotic Version)

Asymptotically at least 1/25 of all triangles are monochromatic in any 3-edge-coloring.

Proofs for such statements are computational, relying on *Flag Algebras*. Getting an answer for larger t or larger c means solving a more challenging optimization problem.

Theorem (Kiem, Pokutta, S. 2023)

Asymptotically at least 1/256 of all triangles are monoc. in any 4-edge-coloring.

More than two colors

Goodman also asked about more than two colors and Fox noted the upper bound

$$m_c(3) \leq 1/R_{c-1}(3)^2.$$

Theorem (Cummings et al. 2013 – Asymptotic Version)

Asymptotically at least 1/25 of all triangles are monochromatic in any 3-edge-coloring.

Proofs for such statements are computational, relying on *Flag Algebras*. Getting an answer for larger t or larger c means solving a more challenging optimization problem.

Theorem (Kiem, Pokutta, S. 2023)

Asymptotically at least 1/256 of all triangles are monoc. in any 4-edge-coloring.

Computational challenges in Flag Algebra proofs

A trivial computational lower bound

Let $k(G) = k_t(G_1) + \ldots + k_t(G_c)$. Through double counting, we have

$$k(G) = \sum_{H \in \mathcal{G}_\ell} p(H; G)\, k(H) \tag{1}$$

as long as $\ell \leq v(G)$, where $p(H; G)$ is the *density* of H in G. This implies

$$k(G) \geq \min_{H \in \mathcal{G}_\ell} k(H). \tag{2}$$

If we had a_H satisfying $\sum_{H \in \mathcal{G}_\ell} p(H; G')\, a_H \leq o(1)$ for all G', this would imply

$$k(G) \geq \min_{H \in \mathcal{G}_\ell} k(H) - a_H. \tag{3}$$

Flag Algebras allow us to find exactly such coefficients a_H!

A trivial computational lower bound

Let $k(G) = k_t(G_1) + \ldots + k_t(G_c)$. Through double counting, we have

$$k(G) = \sum_{H \in \mathcal{G}_\ell} p(H; G)\, k(H) \tag{1}$$

as long as $\ell \leq v(G)$, where $p(H; G)$ is the *density* of H in G. This implies

$$k(G) \geq \min_{H \in \mathcal{G}_\ell} k(H). \tag{2}$$

If we had a_H satisfying $\sum_{H \in \mathcal{G}_\ell} p(H; G')\, a_H \leq o(1)$ for all G', this would imply

$$k(G) \geq \min_{H \in \mathcal{G}_\ell} k(H) - a_H. \tag{3}$$

Flag Algebras allow us to find exactly such coefficients a_H!

Formalisation through Algebras

With $\mathcal{G} = \bigcup_n \mathcal{G}_n$ and $\mathcal{K} = \{k(G) - \sum_{H \in \mathcal{G}_\ell} p(H; G)\, k(H) \mid G \in \mathcal{G}_n, \ell \leq n\} \subset \mathbb{R}\mathcal{G}$, let

$$\mathcal{A} = \mathbb{R}\mathcal{G} / \operatorname{span} \mathcal{K} \tag{4}$$

be the *Flag Algebra* (of the empty type) with product $H \cdot H' = \sum_G p(H, H'; G)\, G$.

Theorem (Razborov 2007; convergent sequences ↔ positive homomorphisms)

Sequences of graphs in which all $p(H; G_n)$ converge correspond one-to-one with $\varphi \in \operatorname{Hom}(\mathcal{A}, \mathbb{R})$ satisfying $\varphi(H) \geq 0$ for all $H \in \mathcal{G}$ through $p(H; G_n) = \varphi(H)$.

This means that any expression (such as  +  $- 1/4$) that is in the *semantic cone*

$$\mathcal{S} = \{f \in \mathcal{A} : \varphi(f) \geq 0 \text{ for all positive homomorphisms } \varphi\} \tag{5}$$

is 'true' in the world of combinatorics. **How do we prove something is in $\mathcal{S}$?**

Formalisation through Algebras

With $\mathcal{G} = \bigcup_n \mathcal{G}_n$ and $\mathcal{K} = \{k(G) - \sum_{H \in \mathcal{G}_\ell} p(H; G)\, k(H) \mid G \in \mathcal{G}_n, \ell \leq n\} \subset \mathbb{R}\mathcal{G}$, let

$$\mathcal{A} = \mathbb{R}\mathcal{G} / \operatorname{span} \mathcal{K} \tag{4}$$

be the *Flag Algebra* (of the empty type) with product $H \cdot H' = \sum_G p(H, H'; G)\, G$.

Theorem (Razborov 2007; convergent sequences $\leftrightarrow$ positive homomorphisms)

Sequences of graphs in which all $p(H; G_n)$ converge correspond one-to-one with $\varphi \in \operatorname{Hom}(\mathcal{A}, \mathbb{R})$ satisfying $\varphi(H) \geq 0$ for all $H \in \mathcal{G}$ through $p(H; G_n) = \varphi(H)$.

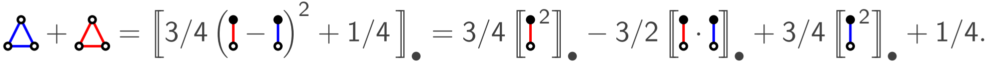
This means that any expression (such as  $+$  $- 1/4$) that is in the *semantic cone*

$$\mathcal{S} = \{f \in \mathcal{A} : \varphi(f) \geq 0 \text{ for all positive homomorphisms } \varphi\} \tag{5}$$

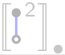
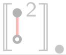
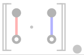
is 'true' in the world of combinatorics. **How do we prove something is in $\mathcal{S}$?**

SOS please someone help me

One approach to establishing the positivity in a *partially ordered algebra* are **sum-of-squares (SOS)** expressions (giving the a_H). Goodman's argument for example states



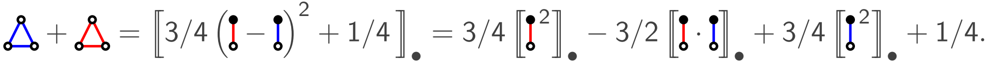
This is a *certificate*, that is verified over 'large enough'graphs ($\ell = 3$ for Goodman):

						a_H
	1	0	1	0	0	3/4
	0	0	1/3	0	1/3	-1/4
	0	0	0	1/3	1/3	-1/4
	0	1	0	1	0	3/4

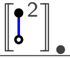
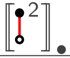
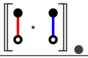
The SOS expression can be found using **Semidefinite Programming (SDP)**.

SOS please someone help me

One approach to establishing the positivity in a *partially ordered algebra* are **sum-of-squares (SOS)** expressions (giving the a_H). Goodman's argument for example states



This is a *certificate*, that is verified over 'large enough'graphs ($\ell = 3$ for Goodman):

						a_H
	1	0	1	0	0	3/4
	0	0	1/3	0	1/3	-1/4
	0	0	0	1/3	1/3	-1/4
	0	1	0	1	0	3/4

The SOS expression can be found using **Semidefinite Programming (SDP)**.

Computational challenges in Flag Algebra proofs

Computational barriers

Increasing ℓ improves bounds and makes the SDP harder:

ℓ	*value*	*time*	*memory*
6	0.02875	0.2s ±0.0	81.2MB ±24.7
7	0.02918	4.9s ±0.1	126.9MB ±26.3
8	0.02942	1.8h ±0.1	1.8GB ±0.0

Table: Difficulty of SDP problem formulations for $m_2(4)$ using `CSDP`

The number of colorings grows like $c^{\binom{n}{2}}/n!$, i.e., more colors means more colorings.
However, they also give us more symmetries in the problem!

How can we use combinatorial information to reduce these SDP formulations?

Computational barriers

Increasing ℓ improves bounds and makes the SDP harder:

ℓ	*value*	*time*	*memory*
6	0.02875	0.2s ± 0.0	81.2MB ± 24.7
7	0.02918	4.9s ± 0.1	126.9MB ± 26.3
8	0.02942	1.8h ± 0.1	1.8GB ± 0.0

Table: Difficulty of SDP problem formulations for $m_2(4)$ using `CSDP`

The number of colorings grows like $c^{\binom{n}{2}}/n!$, i.e., more colors means more colorings. However, they also give us more symmetries in the problem!

How can we use combinatorial information to reduce these SDP formulations?

Leveraging Symmetries

Graph isomorphisms only consider (partial) vertex permutations, but many relevant parameters are invariant under (partial) *color permutations*! In flag algebra terminology, each such permutation is a an *order- and unit-preserving linear map* $L : \mathcal{A} \to \mathcal{A}$, e.g. satisfying $L(\triangle + \triangle) = \triangle + \triangle$.

Method 1 Combine and therefore reduce the number of constraints by up to $c!$. This is *strictly stronger than considering edge partitions* (Balogh et al. 2017).

Method 2 Reduce the number of variables by block diagonalization. Many existing approaches like Gatermann and Parrilo (2004), Murota, Kanno, Kojima, Kojima (2010) and Bachoc et al. (2012) exist (applying Schur's Lemma).

(i) All symmetries are easily determined, (ii) need real or even rational matrices, (iii) data matrices are sparse, (iv) only deal with specific groups. We therefore split groups into isotypic components, strengthening the anti-invariant split of Razborov (2010).

Leveraging Symmetries

Graph isomorphisms only consider (partial) vertex permutations, but many relevant parameters are invariant under (partial) *color permutations*! In flag algebra terminology, each such permutation is a an *order- and unit-preserving linear map* $L : \mathcal{A} \to \mathcal{A}$, e.g. satisfying $L(\triangle + \triangle) = \triangle + \triangle$.

Method 1 Combine and therefore reduce the number of constraints by up to $c!$. This is *strictly stronger than considering edge partitions* (Balogh et al. 2017).

Method 2 Reduce the number of variables by block diagonalization. Many existing approaches like Gatermann and Parrilo (2004), Murota, Kanno, Kojima, Kojima (2010) and Bachoc et al. (2012) exist (applying Schur's Lemma).

(i) All symmetries are easily determined, (ii) need real or even rational matrices, (iii) data matrices are sparse, (iv) only deal with specific groups. We therefore split groups into isotypic components, strengthening the anti-invariant split of Razborov (2010).

Leveraging Symmetries

Graph isomorphisms only consider (partial) vertex permutations, but many relevant parameters are invariant under (partial) *color permutations*! In flag algebra terminology, each such permutation is a an *order- and unit-preserving linear map* $L : \mathcal{A} \to \mathcal{A}$, e.g. satisfying $L($$+$$) =$ $+$.

Method 1 Combine and therefore reduce the number of constraints by up to $c!$. This is *strictly stronger than considering edge partitions* (Balogh et al. 2017).

Method 2 Reduce the number of variables by block diagonalization. Many existing approaches like Gatermann and Parrilo (2004), Murota, Kanno, Kojima, Kojima (2010) and Bachoc et al. (2012) exist (applying Schur's Lemma).

(i) All symmetries are easily determined, (ii) need real or even rational matrices, (iii) data matrices are sparse, (iv) only deal with specific groups. We therefore split groups into isotypic components, strengthening the anti-invariant split of Razborov (2010).

Leveraging Symmetries

Graph isomorphisms only consider (partial) vertex permutations, but many relevant parameters are invariant under (partial) *color permutations*! In flag algebra terminology, each such permutation is a an *order- and unit-preserving linear map* $L : \mathcal{A} \to \mathcal{A}$, e.g. satisfying $L($ $+$ $) =$  $+$ .

Method 1 Combine and therefore reduce the number of constraints by up to $c!$. This is *strictly stronger than considering edge partitions* (Balogh et al. 2017).

Method 2 Reduce the number of variables by block diagonalization. Many existing approaches like Gatermann and Parrilo (2004), Murota, Kanno, Kojima, Kojima (2010) and Bachoc et al. (2012) exist (applying Schur's Lemma).

(i) All symmetries are easily determined, (ii) need real or even rational matrices, (iii) data matrices are sparse, (iv) only deal with specific groups. We therefore split groups into isotypic components, strengthening the anti-invariant split of Razborov (2010).

Concluding remarks and open problems

Theorem (Kiem, Pokutta, S. 2023)

Asymptotically at least 1/256 of all triangles are monoc. in any 4-edge-coloring.

- Rounding is another challenge (we used exact rational LP solver `SoPlex`.)
- Often one can extract stability results from Flag Algebra certificates.
- Other problems (e.g. rainbow cliques) have similar symmetries.
- Additional combinatorial information should further reduce the difficulty.
- We still need improved solvers in addition to improved problem formulations.

Open Problem: $m_{3,\ldots,3} = (R_{3,\ldots,3} - 1)^{-2}$ for all c?

Concluding remarks and open problems

Theorem (Kiem, Pokutta, S. 2023)

Asymptotically at least 1/256 of all triangles are monoc. in any 4-edge-coloring.

- Rounding is another challenge (we used exact rational LP solver `SoPlex`.)
- Often one can extract stability results from Flag Algebra certificates.
- Other problems (e.g. rainbow cliques) have similar symmetries.
- Additional combinatorial information should further reduce the difficulty.
- We still need improved solvers in addition to improved problem formulations.

Open Problem: $m_{3,\dots,3} = (R_{3,\dots,3} - 1)^{-2}$ for all c?

Concluding remarks and open problems

Theorem (Kiem, Pokutta, S. 2023)

Asymptotically at least $1/256$ of all triangles are monoc. in any 4-edge-coloring.

- Rounding is another challenge (we used exact rational LP solver `SoPlex`.)
- Often one can extract stability results from Flag Algebra certificates.
- Other problems (e.g. rainbow cliques) have similar symmetries.
- Additional combinatorial information should further reduce the difficulty.
- We still need improved solvers in addition to improved problem formulations.

Open Problem: $m_{3,\ldots,3} = (R_{3,\ldots,3} - 1)^{-2}$ for all c?

Thank you for your attention!

Selected related literature

- Thomason, A. "Graph products and monochromatic multiplicities." Combinatorica 17.1 (1997): 125-134.
- Razborov, A. "Flag algebras." The Journal of Symbolic Logic 72.4 (2007): 1239-1282.
- Razborov, A. "On 3-hypergraphs with forbidden 4-vertex configurations." SIAM Journal on Discrete Mathematics 24.3 (2010): 946-963.
- Cummings, J., et al. "Monochromatic triangles in three-coloured graphs." Journal of Combinatorial Theory, Series B 103.4 (2013): 489-503.
- Balogh, J., et al. "Rainbow triangles in three-colored graphs." Journal of Combinatorial Theory, Series B 126 (2017): 83-113.
- Gatermann, J., and Parrilo, P.. "Symmetry groups, semidefinite programs, and sums of squares." Journal of Pure and Applied Algebra 192.1-3 (2004): 95-128.
- Bachoc, C., et al. "Invariant semidefinite programs." Handbook on semidefinite, conic and polynomial optimization. Springer, Boston, MA, 2012. 219-269.