

Fully Computer-Assisted Proofs in Extremal Combinatorics

Olaf Parczyk¹, Sebastian Pokutta^{2,3}, Christoph Spiegel³, Tibor Szabó¹

¹ Freie Universität Berlin, Institute of Mathematics, Berlin

² Technische Universität Berlin, Institute of Mathematics, Berlin

³ Zuse Institute Berlin, Department for AI in Society, Science, and Technology, Germany
parczyk@mi.fu-berlin.de, pokutta@zib.de, spiegel@zib.de, szabo@mi.fu-berlin.de

Abstract

We present a fully computer-assisted proof system for solving a particular family of problems in Extremal Combinatorics. Existing techniques using Flag Algebras have proven powerful in the past, but have so far lacked a computational counterpart to derive matching constructive bounds. We demonstrate that common search heuristics are capable of finding constructions far beyond the reach of human intuition. Additionally, the most obvious downside of such heuristics, namely a missing guarantee of global optimality, can often be fully eliminated in this case through lower bounds and stability results coming from the Flag Algebra approach.

To illustrate the potential of this approach, we study two related and well-known problems in Extremal Graph Theory that go back to questions of Erdős from the 60s. Most notably, we present the first major improvement in the upper bound of the Ramsey multiplicity of K_4 in 25 years, precisely determine the first off-diagonal Ramsey multiplicity number, and settle the minimum number of independent sets of size four in graphs with clique number strictly less than five.

Introduction

Computers and Artificial Intelligence (AI) have always been an important tool for mathematicians, allowing one to gather data and extrapolate connections to formulate conjectures, most notably the conjecture of Swinnerton-Dyer and Birch (1965), exhaustively execute case analysis too large to be done by hand, for example to establish the four color theorem (Appel and Haken 1989), or solve convex programming problems, as was done to prove the Kepler conjecture (Hales 2005). Lately the rigorous formalization and verification of proofs has also come to the forefront, cf. (Buzzard 2022). This has opened up the possibility for computers to execute fully automated reasoning either through traditional heuristics, as was for example done to prove that Robbins algebras are boolean (McCune 1996), or even through Machine Learning (ML) approaches (Polu et al. 2022).

The use of computer assistance has been particularly fruitful in Extremal Combinatorics, where the primary interest is to study the behaviour of abstract discrete structures, such as graphs, with the goal of determining the constructions

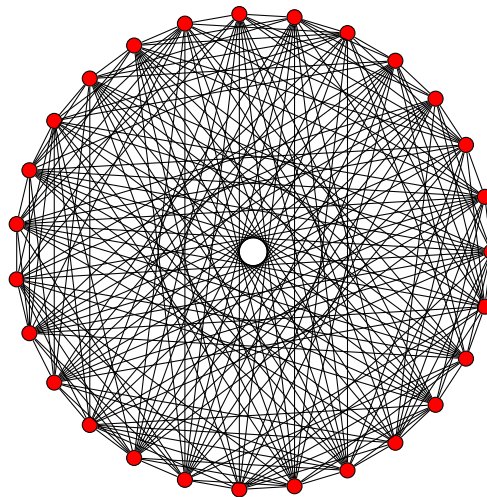


Figure 1: The blow-up sequence of the Schläfli graph is the unique solution to the off-diagonal $(3, 4)$ -Ramsey multiplicity problem and was found through search heuristics.

that minimize or maximize a given parameter under certain restrictions. When studying locally defined restrictions and parameters, the most formalized and successful computational approach in this field comes in the form of an application of Flag Algebras, allowing one to establish bounds through a double-counting argument by solving Semidefinite Programming (SDP) problems. Razborov (2007) introduced the notion of Flag Algebras, citing both Bondy’s work on the Caccetta-Häggkvist conjecture as well as Lovász and Szegedy’s work on graph limits as inspiration. Since then this approach has been for example applied to make progress on Turán’s conjecture (Razborov 2010; Falgas-Ravry and Vaughan 2013), determine the maximum number of five cycles in triangle-free graphs (Grzesik 2012), to establish that jumps exist for hypergraphs (Baber and Talbot 2011), and prove a conjecture of Brandt about the spectrum of triangle-free graphs (Balogh, Clemen, and Lidický 2022).

In many of these problems the counterpart to the bound given by the Flag Algebra approach is played by an explicit construction. Such constructions are commonly obtained by hand based on human combinatorial insights, but

there is a clear desire, as for example explicitly stated by Pikhurko, Sliáčan, and Tyros (2019), for a computer based approach that is easily applicable to a wide variety of problems and can supplement and surpass that human intuition. We show that for many cases the goal of finding constructive bounds can be explicitly stated as a simple discrete optimization problem. While a lot of approaches to find solutions to these problems exist and have been used in the past, we find that metaheuristics, most notably Simulated Annealing (Kirkpatrick, Gelatt Jr, and Vecchi 1983) and Tabu Search (Glover 1986, 1989, 1990), are particularly well suited to the type of problems described above; they are easy to implement and tune, allow one to quickly add additional insights into the problem formulation and solution process, scale far beyond exhaustive search methods and most importantly work very well in practice with limited computational resources. The biggest downside of using heuristics, namely the fact that they offer no guarantee that a found solution is in fact globally optimal, is often mitigated or even fully eliminated in this application; matching Flag Algebra lower bounds can not only establish that a found solution is a (not necessarily unique) global optimum, but one can in many cases also obtain stability results from them, establishing that anything that comes close to the optimal value must be close to the extremal construction, which in a strong sense shows the uniqueness of the solution found by the heuristics. While these search heuristics are well-established in other areas, they so far appear undervalued as a tool in Extremal Combinatorics despite a recent surge in interest in obtaining combinatorially relevant constructions through computational means (Lidický and Pfender 2017; Rowley 2022; Wagner 2020, 2021).

To illustrate the potential of this approach, we study some related and well-known problems in Extremal Graph Theory that go back to questions of Erdős. The first concerns a generalisation of Ramsey’s well known theorem: knowing that any graph of large enough order must either contain a clique or an independent set of some fixed size t , the *Ramsey multiplicity problem* asks how few such cliques and independent sets an optimal graph must contain. Goodman (1959) gave an answer for triangles, that is $t = 3$, but the exact value still remains unknown even for $t = 4$. Using the approach discussed above, we present the first major improvement on the upper bound of this problem in over 25 years. We also study a natural off-diagonal variant of this problem, where one minimizes the number of cliques of size t and the number of independent sets of size s , and show that for $s = 4$ and $t = 3$ this version of the Ramsey multiplicity problem is minimized precisely by the sequence of unweighted blow-ups of the Schläfli graph. The final problem we address concerns the minimum number of independent sets of a given size in graphs of bounded independence number. In particular, we settle the minimum number of independent sets of size four in a graph with clique number strictly less than five, proving that it is uniquely attained by the sequence of unweighted blow-ups of the unique $(3, 5)$ -Ramsey graph on 13 vertices.

The upper bounds for all three problems were found using search heuristics by restating the search for these con-

structive bounds as discrete optimization problems. For the latter two problems the bounds were found in the space of all graphs of a given fixed order whereas for the traditional Ramsey multiplicity problem we biased the search by directly constructing the generating sets of Cayley graphs in predetermined groups. The matching lower bounds as well as stability result for the latter two problems were obtained using the Flag Algebra approach, establishing that the search heuristics found unique global minima in two of the three cases.

Contributions The major contributions of our work can be summarized as follows:

1. We present both significant improvements on the constructive upper bound as well as tight bounds with stability results for several long-standing and important open problems in Extremal Graph Theory.
2. We demonstrate that well-established search heuristics can construct combinatorial bounds that are far out of reach of human intuition without relying on significant prior knowledge. In our applications these heuristics also significantly surpass more recently proposed Machine Learning (ML) based approaches.
3. We improve the toolset used to derive stability results from Flag Algebra based proofs in order to show that several of the constructive bounds found using search heuristics in fact represent global optima.

Related Literature This research builds and improves upon a long list of previous work. The Flag Algebra based SDP approach was originally proposed by Razborov (2007) and has since been significantly further developed, most notably by Baber (2011), Vaughan (2013), and Pikhurko, Sliáčan, and Tyros (2019). Some of the most important uses of this methodology to derive results in Extremal Combinatorics have already been listed above.

The use of computers to exhaustively generate combinatorial structures of a prescribed size has a long history in Combinatorics, with some of the most important practical contributions by McKay (1998). The use of exhaustive bounded tree searches, often imbued with a tremendous amount of theoretical insights, to find optimal constructions for a particular parameter has also resulted in several significant achievements. The most notable works in this area concern the determination of Ramsey numbers, see for example McKay and Radziszowski (1995), as well as the more recent work of Heule (2018) on the number-theoretical equivalent Schur numbers, resulting in the largest computer-based proof to date.

The search for Ramsey numbers in fact has many strong parallels to the problems studied here and the successful use of Metaheuristics, in particular Simulated Annealing and Tabu Search, to obtain lower bounds on them, see for example the work of McKay and Radziszowski (1997), motivated our use of these methods. These heuristics for the most part have been used to directly construct graphs, but our approach to construct the generating sets of Cayley graphs does bear some similarities to the work of Exoo (1998).

Our methodology is also motivated by recent approaches to theoretical problems in mathematics based on ML. The work of Wagner (2021) is perhaps most closely related to this work, suggesting to find explicit counter examples to open conjectures in Extremal Combinatorics through a Search Heuristic based on Reinforcement Learning algorithms. This approach was previously proposed in a different context under the name ‘Active Learning’ by Bello et al. (2016) and was subsequently also further developed by Roucairol and Cazenave (2022). We experimented significantly with learning-based approaches as alternatives to the more traditional Metaheuristics, but while they also found improved bounds and solutions to the previously described problems, we found the traditional Metaheuristics not only drastically easier to implement and tune but also much more performant, giving better solutions often using orders of magnitude less compute time. Nevertheless, the use of ML remains interesting in this field and some recent works such as those of Davies et al. (2021) give hope that novel learning-based approaches can open up new possibilities for computer-assisted conjectures and proofs.

Outline We begin by formally stating the problems and results regarding the extremal behavior of graphs summarized in the introduction along with additional context in the next section. We then outline how to formulate the relevant discrete optimization problems along with the methods we used to solve them. Afterwards, we will outline Razborov’s Flag Algebra approach and how it can be applied to establish stability results, along with some specific improvements allowing us to derive stability results for our application. Lastly, we will discuss some additional problems as well as a broader context and outlook.

Problem Statements and Results

The type of problems that are amenable to our computational approach are those where one has at least a reasonable suspicion that one part of a solution consists of a finding a bound derived from a sequence of constructions that is finitely describable (though possibly involving randomness) and whose relevant properties are explicitly computable. The most well established way of obtaining such finitely describable sequences of constructions in Extremal Graph Theory, besides random graph models, is through blow-ups; here the *m-fold blow-up* of a graph C is the graph $C[m]$ obtained by replacing every vertex with an independent set of a size m and connecting two vertices from different independent sets if and only if the original vertices were adjacent. We will also refer to this as the *unweighted blow-up* of C since every vertex gets replaced by an independent set of equal size. Many generalizations and variants of this type of construction exist throughout all areas of Extremal Combinatorics.

The family described above covers a surprising amount of problems, for example covering many, if not most, problems in which the Flag Algebra approach has found application, i.e., those where the parameter to be optimized is describable in terms of densities of smaller structures. In fact the larger meta question for many of these problems is to decide

whether or not the optimum value is (uniquely) achieved by a sequence coming from some finite construction. Famous examples, where this is still undecided but current best bounds come from finitely describable construction, are the capset problem (Edel 2004; Ellenberg and Gijswijt 2017), the Sunflower conjecture by Erdős and Rado (Alweiss et al. 2020), Turán’s $(3, 4)$ -conjecture (Razborov 2010; Falgas-Ravry and Vaughan 2013), and the Shannon Capacity of odd cycles (Mathew and Östergård 2017). Other examples include a conjecture of Brandt on the sum of the first and last eigenvalue in triangle-free graphs (Csikvári 2022), the homomorphism threshold for odd cycles (Ebsen and Schacht 2020), and the equivalent problems from Additive Combinatorics to the problems studied here (Cameron, Cilleruelo, and Serra 2007; Wolf 2010; Lu and Peng 2012). Let us now formally introduce the problems we studied in order to illustrate the potential of our approach.

Ramsey multiplicity We denote by $k_t(G)$ the number of cliques on t vertices in a graph G and let $k_t(n) = \min\{k_t(G) + k_t(\overline{G}) : |G| = n\}$. We know by Ramsey’s theorem that $k_t(n) > 0$ when n is sufficiently large depending on t and since $k_t(n)$ is increasing with n , we can study

$$c_t = \lim_{n \rightarrow \infty} k_t(n) / \binom{n}{t}.$$

The result of Goodman (1959) implies that $c_3 = 1/4$. Since this is the same value as given in expectation by the Erdős-Rényi random graph $G(n, 1/2)$, Erdős (1962) conjectured that $c_t = 2^{1-t(t-1)/2}$ holds for arbitrary t . This was soundly rejected by Thomason (1989) through explicit constructions based on unweighted blow-ups of for any $t \geq 4$. For $t \in \{4, 5\}$ in particular, Thomason (1997) later also found the significantly improved bounds of $c_4 < 0.03029$ and $c_5 < 0.001720$, the former of which was slightly improved by Even-Zohar and Linial (2015) to yield $c_4 < 0.03028$. Here we present the following improved upper bounds.

Theorem 1. *We have $c_4 < 0.03015$ and $c_5 < 0.001708$.*

These bounds are respectively established through the sequence of unweighted blow-ups of graphs on 768 and 192 vertices. Regarding lower bounds, Giraud (1979) proved that $c_4 \geq 1/46$ using traditional means, see also the updated version of this proof due to Wolf (2010), and Nieß (2012) and Sperfeld (2011) independently used the Flag Algebra approach to establish a lower bound of around $1/35 > 0.028571$. This was later improved by Grzesik et al. (2020) to 0.0296. To the best of our knowledge, no formal lower bound has been published for c_5 , but using the Flag Algebra approach, we obtained $c_5 \geq 0.001524$.

Off-diagonal Ramsey multiplicity There is a stark contrast between the difficulty of the Ramsey multiplicity problem for $t = 3$, which was solved in the 50s, and for $t = 4$, which is still open today. To get a grip on the latter, we propose to investigate a natural off-diagonal version of this problem by considering

$$c_{s,t} = \lim_{n \rightarrow \infty} \min \left\{ \frac{k_s(\overline{G})}{\binom{n}{s}} + \frac{k_t(G)}{\binom{n}{t}} : |G| = n \right\}$$

for arbitrary $s, t \geq 2$. A similar problem has also very recently been studied by Behague, Morrison, and Noel (2022). The value of $c_{2,t}$ follows from a result of Bollobás (2004) for every $t \geq 3$. Here we establish the first exact results when $t > s \geq 3$ and also show the stability of a unique construction on 27 vertices.

Theorem 2. *We have $c_{3,4} = 689 \cdot 3^{-8}$ and the problem is perfectly stable with respect to the Schläfli graph.*

The upper bound is given by the sequence of symmetric blow-ups of the Schläfli graph, a strongly regular graph on 27 vertices shown in Figure 1. The notion of perfect stability was introduced by Pikhurko, Slišan, and Tyros (2019) and strengthens the standard notion of stability. The lower bound and stability result are established using the Flag Algebra approach. Note that we can also show that $c_{5,3} = 24011 \cdot 3^{-12}$ where the upper bound is given by the sequence of symmetric blow-ups of the complement of the Schläfli graph, but we were unable to establish stability in this case.

As already alluded to at the beginning of this section, a central question underlying the Ramsey multiplicity problem is whether or not a tight upper bound can be attained by the sequence of (possibly weighted) blow-ups of some finite sized graph. For c_3 this holds true as the bound of Goodman is attained by, among other constructions, the sequence of blow-ups of K_2 . Already for c_4 this question remains unanswered.

Graphs with bounded clique number Another question related to the Ramsey multiplicity problem that goes back to Erdős concerns

$$g_{s,t} = \lim_{n \rightarrow \infty} \min \{k_s(\overline{G}) : |G| = n, k_t(G) = 0\} / \binom{n}{s}$$

for arbitrary $s, t \geq 2$. Note that $c_{s,t} \leq \min\{g_{s,t}, g_{t,s}\}$ for any $s, t \geq 2$. We trivially have $g_{s,2} = 1$ and the fact that $g_{2,t} = 1/(t-1)$ is implied by Turán’s theorem. In a clear parallel to the conjecture mentioned above, Erdős (1962) asked if the upper bound given by the Turán graphs $T_{t-1}(n)$ is asymptotically tight in general. This holds for triangles, that is when $s = t = 3$, as an easy consequence of the result of Goodman (1959), but Nikiforov (2001) showed that this upper bound can be sharp only for a finite number of pairs $s, t \geq 3$. Das et al. (2013) as well as Pikhurko and Vaughan (2013) established tight values for $g_{3,t}$ and $g_{s,3}$ when $4 \leq s, t \leq 7$, confirming Erdős’ intuition for the former and disproving it for the latter. We present one further tight value for $g_{s,t}$ that improves upon the bound given by Turán graphs and also show that stability holds with respect to a particular graph on 13 vertices.

Theorem 3. *We have $g_{4,5} = 29 \cdot 13^{-3}$ with perfect stability with respect to the unique $(3, 5)$ -Ramsey graph of order 13.*

The unique $(3, 5)$ -Ramsey graph of order 13 can be constructed as the Cayley graph on \mathbb{Z}_{13} whose edge relations are given by the cubic-non-residues. The upper bound is given by the sequence of unweighted blow-ups of that graph. The lower bound as well as the stability results was established using the Flag Algebra approach. Note that Pikhurko and Vaughan (2013) also found a construction using weighted

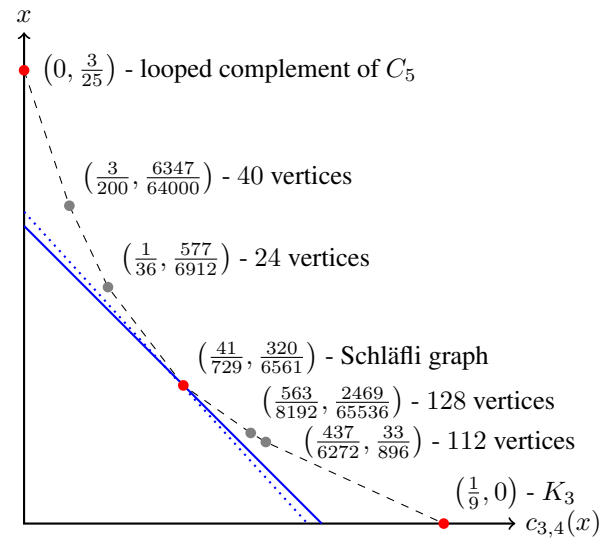


Figure 2: The known lower bounds and constructions for $c_{3,4}(x)$. The blue line is the lower bound $c_{3,4} \geq 689 \cdot 3^{-8}$ and the blue dotted line is an additional lower bound, both obtained by Flag Algebra. The red dots represent optimal constructions, where the ones on the axis were known before (Das et al. 2013). The grey dots are additional constructions that we found, where the dashed grey lines merely indicate that these are in convex position.

blow-ups that bounds $g_{4,4}$ away from the value given by $T_3(n)$ and that is conjectured to be tight.

Understanding the full spectrum Both the study of $c_{s,t}$ and $g_{s,t}$ are part of a broader question in which one would like to understand the full extremal tradeoff between the number of K_t in a graph and the number of K_s in its complement. More precisely, we are interested in

$$c_{s,t}(x) = \lim_{n \rightarrow \infty} \min \left\{ \frac{k_s(\overline{G})}{\binom{n}{s}} : |G| = n, \frac{k_t(G)}{\binom{n}{t}} \leq x \right\},$$

for $s, t \geq 2$ and $x \in [0, 1]$. Clearly $c_{s,t}(0) = g_{s,t}$ and $c_{s,t}(x) = 0$ if and only if $x \geq g_{t,s}$. This function was completely determined for $s = 2$ and arbitrary t by Razborov (2008), Nikiforov (2001), and a famous result of Reiher (2016). A simple construction also shows that the lower bound by Goodman for $s = t = 3$ is tight in this more general setting, establishing that $c_{3,3}(x) = 1/4 - x$ for all $x \in [0, 1/4]$.

Very little is known about the shape of $c_{s,t}(x)$ when $s \neq t \geq 3$, though we can show that it is decreasing and differentiable almost everywhere. For $s = 3$ and $t = 4$, the newly determined value of $c_{3,4}$ implies $c_{3,4}(41 \cdot 3^{-6}) = 320 \cdot 3^{-8}$ and the previously known values of $g_{3,4}$ and $g_{4,3}$ imply that $c_{3,4}(0) = 1/9$ and $c_{3,4}(3/25) = 0$. We can also show that $c_{3,4}(x)$ is not differentiable at $41 \cdot 3^{-6}$ using an additional lower bound using the Flag Algebra approach. This indicates that the Schläfli graph determines one of a number of ‘ex-

tremal points' of the curve $c_{3,4}(x)$.¹ We determined several additional constructions establishing upper bounds on $c_{3,4}$ at particular values of x and summarize our findings in Figure 2.

Constructive Bounds and Metaheuristics

As discussed previously, we are interested in studying problems where the goal of determining constructive bounds can be restated as an appropriate optimization problem. The problems introduced in the previous section fall into this category because of the following result of Thomason (1989).

Lemma 4. *For any graph C of order n and for $m \in \mathbb{N}$ going to infinity, we have*

$$k_t(C[m]) = \frac{t! k_t(C)}{n^t} \binom{mn}{t} (1 + o(1))$$

as well as

$$k_t(\overline{C[m]}) = \frac{\sum_{j=1}^t j! S(t, j) k_j(\overline{C})}{n^t} \binom{mn}{t} (1 + o(1)).$$

Here $S(t, j)$ is the Stirling number of the second kind. This lemma is an immediate consequence of the fact that the fraction of not necessarily injective homomorphic copies of cliques of size t in a graph stays the same under blow-up. The number of homomorphic copies asymptotically matches the number of injective copies, accounting for the $1 + o(1)$ term. This statement also readily generalizes to arbitrary graphs and not just cliques, as well as to weighted blow-ups and even quantum graphs.

From a more practical perspective, Lemma 4 means that we not only have a source of finitely describable constructions in the form of blow-ups of graphs, but we also have an efficient way of computing the necessary properties of these constructions. This means we can derive an upper bound on $c_{s,t}$ from any graph C , as well as an upper bound on $g_{s,t}$ from any graph C with $\omega(C) \leq t - 1$.

Given some fixed order n , we therefore associate the binary vector $\mathbf{s} = (s_1, \dots, s_N) \in \{0, 1\}^N$ where $N = \binom{n}{2}$ with the graph $C_{\mathbf{s}}$ given by the vertex set $\{1, \dots, n\}$ and edge set $\{ij : i < j, s_{\binom{j-1}{2}+i} = 1\}$ and are interested in the optimization problem

$$\min_{\mathbf{s} \in \{0,1\}^N} \sum_{j=1}^s \frac{S(t, j) n^j k_j(\overline{C_{\mathbf{s}}})}{n^t} + \lambda \frac{n^t k_t(C_{\mathbf{s}})}{n^t}. \quad (1)$$

Here $n^{\underline{j}}$ is the falling factorial. We set $\lambda = 1$ when we are interested in bounds for $c_{s,t}$ and when deriving bounds for $g_{s,t}$ the hard constraint $k_t(C_{\mathbf{s}}) = 0$ is addressed through a Lagrangian multiplier, i.e., some large enough $\lambda \gg 1$.

While the difficulty of finding a good solution to Equation (1) is primarily governed by the number N of variables used to construct the graph, the quality of any solution,

¹We are being intentionally vague here with our notion of extremal points. Most likely we would expect there to be a, possibly infinite, set of points x in $[0, 1]$ at which $c_{3,4}(x)$ is not differentiable and which correspond to a change in the underlying graph construction.

at least for the motivating Ramsey multiplicity problem of K_4 problem, seems to primarily depend on the number n of vertices. Since in the graph space N is quadratic in n , we also considered a variant of this optimization problem in which we bias the search space towards Cayley graphs, which are easily constructed and where the number of variables only grows linearly with the order of the graph. That means for a fixed finite and abelian group \mathbf{G} and a set $S \subseteq \mathbf{G} \setminus \{1\}$ satisfying $S = S^{-1}$, the corresponding Cayley graph $C = C(\mathbf{G}, S)$ has vertex set \mathbf{G} and two vertices g_1, g_2 are connected by an edge if $g_1 s = g_2$ for some $s \in S$. A state $\mathbf{s} \in \{0, 1\}^N$ represents the generating set $\bigcup_{s_A=1} A$ where each subset $A = \{g, g^{-1}\}$ for $g \in \mathbf{G}$ corresponds to an entry s_A in \mathbf{s} . This means that N is between $|\mathbf{G}|/2$ and $|\mathbf{G}| - 1$. Denoting the Cayley graph constructed this way by $C = C(\mathbf{s})$, the relevant cost function is again given by Equation (1).

How to solve the optimization problem Regardless of whether a vector \mathbf{s} represents a graph of a fixed order n or a Cayley graph of a fixed group \mathbf{G} , the cost function implied by Equation (1) is efficiently computable due to Lemma 4. Many approaches to solve this optimization problem exist, though methods with guarantees of finding an optimal solution quickly hit a limit once the order n or group \mathbf{G} becomes too large. However, since the early 80s a plethora of heuristically motivated methods have been suggested to provide good feasible solutions to NP-hard optimization problems. In particular for combinatorial optimization problems with discrete search spaces, many Metaheuristics have been proposed that require little to no domain-specific knowledge. We will briefly introduce perhaps the two most well-established methods and refer the interested reader to the handbook of Gendreau, Potvin et al. (2010) for further information on this subject. We also note that we already briefly discussed the efficacy of more recent ML-based approaches for our particular application in the introduction.

Simulated Annealing (SA) was originally proposed by Kirkpatrick, Gelatt Jr, and Vecchi (1983) and is a probabilistic technique that can be interpreted as a modified local search, that attempts to avoid getting trapped in a local minimum early on while still ultimately being driven to better and better states. To more precisely describe SA, let $c : \{0, 1\}^N \rightarrow \mathbb{R}$ denote some cost function we are trying to minimize and $\mathcal{N} : \{0, 1\}^N \rightarrow \mathcal{P}(\{0, 1\}^N)$ some notion of neighborhoods of the states. We restricted ourselves to considering states as neighboring if their Hamming distance is 1. The algorithm starts with a randomly initialized state \mathbf{s}_0 and then executes a fixed number of iterations I where in each iteration $1 \leq i \leq I$ it picks a candidate state \mathbf{s}_c uniformly at random from $\mathcal{N}(\mathbf{s}_{i-1})$ and accepts it as \mathbf{s}_i with probability

$$\min \left(\exp \left(\frac{c(\mathbf{s}_{i-1}) - c(\mathbf{s}_c)}{t_i} \right), 1 \right), \quad (2)$$

where $(t_i)_{1 \leq i \leq I}$ is a sequence of decreasing temperatures in $\mathbb{R}_{>0}$. If a state is rejected we reuse \mathbf{s}_{i-1} as \mathbf{s}_i . A common choice for $(t_i)_{1 \leq i \leq I}$ consists of linearly decaying it from

some tuned initial value to 0 over the runtime of the algorithm.

We note that it is unclear how critical the individual design decisions of SA are for its efficacy; Dueck and Scheuer (1990) for example proposed a variant in the form of Threshold Accepting (TA) that completely avoids the probabilistic nature of SA along with Metropolis' criterion, that is Equation (2). For our particular purposes it is also computationally relatively cheap to determine $c(s')$ for any $s' \in \mathcal{N}(s)$ compared to determining the cost of a single state. This allows us to implement a version of SA that avoids rejecting states by directly sampling the candidates from an appropriate distribution dictated by Metropolis' criterion. This has previously been suggested by Greene and Supowit (1986).

Tabu Search is an arguably even simpler search heuristic than SA that was originally suggested by Glover (1986, 1989, 1990). Like SA it can be seen as a modified local search that tries to avoid getting stuck in local optima and cycles. We again start with a randomly initialized state s_0 and require a notion of neighborhood \mathcal{N} . We execute I iterations where in each iteration $1 \leq i \leq I$ we pick the neighboring state $s \in \mathcal{N}(s_{i-1})$ that has the lowest associated cost and has not recently been visited, regardless of whether it improves upon $c(s_{i-1})$. To make this more precise, if we let $\ell \in \mathbb{N}$ denote a fixed history length, we choose $s_i = \arg \min_{s \in \mathcal{N}(s_{i-1}) \setminus \{s_{i-\ell}, \dots, s_{i-1}\}} c(s)$.

There are of course again many degrees of freedom when implementing this algorithm. One liberty we took in this case was to modify the way the history is implemented: rather than storing a history of the last ℓ states and excluding those from the update, we simply store a list of the last ℓ modified bits from a state and exclude any state that differs from the current one in one of those bits. This slightly increases the number of states excluded but drastically decreases the computational and implementation effort. We should also note that this method, like SA, is commonly executed in parallel for several different initial states.

How our constructions were found We implemented all search methods in Python 3.8 and logged the results using *Weights & Biases* (Biewald 2020). We also relied on the GAP (GAP) component in SageMath for the computations required for the Cayley graphs. We already stated the constructions used to establish the results regarding $c_{3,4}$ and $g_{4,5}$. Both of these constructions were found by running either of the two search heuristics in the graph space. We note that Tabu Search was significantly faster at finding these solutions compared to SA. They a posteriori turned out to be particular graphs that have previously already been observed to have particular properties of interest.

For an upper bound to c_4 , we ran heuristic searches to construct graphs on up to 40 vertices and the smallest graph we found whose blow-up sequence establishes a value below that original conjectured by Erdős was of order 33. The smallest previously known such graph was described by Thomason (1989) and was of order 36. We also found a graph on 32 vertices whose *weighted* blow-up gives a value slightly below $1/32$. Given that the blow-up sequences of small graphs seem to yield little of interest, we considered

Cayley graphs next. The best previous blow-up construction for c_4 can be described as a Cayley graph in $C_3^{\times 2} \times C_2^{\times 5}$ (order 288) and a Tabu Search in that particular group yielded no improvement. However, already in $C_3 \times C_2^{\times 6}$ (order 192) we were able to find a graph whose blow-up sequence improves upon the value found by Thomason (1997) and Even-Zohar and Linial (2015), giving an upper bound of 0.03027. Going up to $C_3 \times C_2^{\times 8}$ (order 768) produced the graph whose blow-up sequence gives the upper bound stated in Theorem 1.

We note that there seems to be no particular significance to the fact that we derived these constructions in groups defined only through direct products. In fact, running the Tabu Search on all groups of order at most 192 revealed (1) that in general good constructions seem to be found in groups of order $3 \cdot 2^n$ and (2) many groups of that order perform significantly better than the group $C_3 \times C_2^{\times n}$. For groups of order 192 for example the best value we found was around 0.03021. Unfortunately we were unable to determine any patterns indicating which groups might be preferable when going to groups of order 384 or 768. The sheer number of these groups and the amount of cliques to consider unfortunately makes it impossible to run a Tabu Search on anything more than a small selection of them

Regarding the value for c_5 , the previous best construction can be described as the blow-up sequence of a Cayley graph in $C_3 \times C_2^{\times 6}$ (order 192). A search run on Cayley graphs in this group yielded the slight improvement presented in Theorem 1. Due to the fact that we need to consider cliques of size 5, we were unable to go up Cayley graphs of order 384 or 768 as we did for c_4 .

Flag Algebra Proofs

Razborov (2007) suggested phrasing a particular subset of problems in Extremal Combinatorics that deals with the relation of asymptotic densities of combinatorial structures in the language of finite model theory in order capture and describe the rich algebraic structure underlying many of the techniques commonly used in this field. One particular application of his approach allows us to derive lower bounds for precisely the type of minimization problems of subgraph densities studied in this paper by solving an SDP. This method has found significant use over the last decade as discussed in the introduction. There exist not only several very good introductions to the topic (Falgas-Ravry and Vaughan 2013; Pikhurko and Vaughan 2013; Silva, Sato et al. 2016) but also a tool in the form of *flagmatic* (Vaughan 2013) to apply this method with relative ease.

Let us briefly describe the underlying optimization problem using the example of graphs. For any integer $m \in \mathbb{N}$, we let \mathcal{G}_m denote the family of graphs of order m , possibly avoiding a small set of forbidden subgraphs as for example when dealing with $g_{s,t}$. We are interested in solving

$$\max_{Q \geq 0} \min_{H \in \mathcal{G}_m} \frac{k_s(\overline{H})}{\binom{m}{s}} + \frac{k_t(H)}{\binom{m}{t}} - \langle Q, D_H \rangle, \quad (3)$$

where D_H a matrix representing particular pair densities of smaller subgraphs in $H \in \mathcal{G}_m$ and the maximum is over

positive semidefinite matrices Q . The rows and columns of Q and D_H are indexed by so-called flags, which are certain partially labelled graphs. While in theory this can potentially give the correct answer to many open problems, the main bottleneck for the quality of the bound the SDP solver can produce is predominantly determined by the size of the problem formulation, i.e., most significantly by the choice of m . Currently the largest formulations realistically solvable by standard solvers such as CSDP (Borchers 1999) and SDPA (Yamashita et al. 2012) are obtained for $m = 8$.

We employ this to obtain the lower bounds for our results, in particular, for $c_{3,4}$ in Theorem 2 and for $g_{4,5}$ in Theorem 3. As these two lower bounds match the upper bounds given by the constructions that we found it is natural to ask whether the construction is unique and stability holds, i.e., if anything that comes close to the optimal value must be close to the extremal construction. This is of particular interest, because this also implies that the constructions found by the heuristics are in fact unique global minima. Statements like this can usually be derived by extracting additional information from a Flag Algebra based proof and appealing to the Induced Removal Lemma of Alon et al. (2000). The process to do so has been formalized by Pikhurko, Slišan, and Tyros (2019), who formulated several sufficient but not necessary criteria to establish various forms of stability. Essentially, these require that the extremal construction is such that knowing all subgraph densities up to a certain size m is sufficient for forcing the whole graph to be a blow-up of it. We improved these criteria in order to apply them to the problems studied here and to obtain the stability results stated in Theorem 2 and Theorem 3.

Discussion and Outlook

The main goal of this paper was to use a computer driven search to obtain constructions of graphs with small density of cliques of order s and density of independent sets of order t . The most important question regarding the Ramsey multiplicity of K_4 is less its exact value, but rather whether it is given by the blow-up of a finite graph. We believe that our answer to the $c_{3,4}$ problem, where the density of monochromatic K_4 are counted in one color and the density of monochromatic K_3 is counted in the other color, provides some support to the possibility that the value of c_4 will also be determined by the blow-up sequence of a single graph. However, if this indeed was the case, then our efforts for Theorem 1 suggest that such a construction would have to be truly massive. Of course, we heavily relied on Cayley graph constructions in order to reduce our search space and there might be other types of constructions resulting in improved bounds and more compact descriptions.

Regarding future work that could build upon the presented tools, we believe there are three major points of interest:

1. Using different methodologies besides the mentioned search heuristics, the upper bounds derived from the optimization problems relating to c_4 and c_5 could be further improved. In fact, we believe that these problems could also serve as a possible source of benchmarks for these types of heuristics in the future.
2. Using different constructions, i.e., generalizing the notion of blow-ups or using other constructions besides Cayley graphs as the base, further improvements or even solutions to c_4 and c_5 could be obtained. It is quite likely that an immediate improvement can be gained from the found constructions using an iterative blow-up construction as done by Even-Zohar and Linial (2015). More generally, by incorporating such an iterative (and possibly weighted) construction into the optimization problem and using quantum graphs or graphs where edges are assigned probabilities rather than binary states as the base, further significant improvements could be found. It is also possible that the structure underlying the groups giving good constructions can be better understood, allowing one to more directly bias the search towards relevant constructions in this way.
3. As previously stated, there is a large number of important problems in Extremal Combinatorics besides the ones explicitly studied here, where the best current bounds are obtained by concrete and finitely describable constructions. It would be of great interest to see the methodologies applied here to the Ramsey Multiplicity problem and its variants also find application there.

We hope that, by making our results and tools available to the research community, this work will further stimulate interest both in these types of problems and more broadly to computational approaches to theoretical problems in mathematics.

Acknowledgements

This work was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – The Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID: 390685689).

References

- Alon, N.; Fischer, E.; Krivelevich, M.; and Szegedy, M. 2000. Efficient testing of large graphs. *Combinatorica*, 20(4): 451–476.
- Alweiss, R.; Lovett, S.; Wu, K.; and Zhang, J. 2020. Improved Bounds for the Sunflower Lemma. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2020, 624–630. New York, NY, USA: Association for Computing Machinery. ISBN 9781450369794.
- Appel, K. I.; and Haken, W. 1989. *Every planar map is four colorable*, volume 98. American Mathematical Soc.
- Baber, R. 2011. *Some results in extremal combinatorics*. Ph.D. thesis, UCL (University College London).
- Baber, R.; and Talbot, J. 2011. Hypergraphs do jump. *Combinatorics, Probability and Computing*, 20(2): 161–171.
- Balogh, J.; Clemen, F. C.; and Lidický, B. 2022. The spectrum of triangle-free graphs. *arXiv:2204.00093*.
- Behague, N.; Morrison, N.; and Noel, J. A. 2022. Common Pairs of Graphs. *arXiv:2208.02045*.

- Bello, I.; Pham, H.; Le, Q. V.; Norouzi, M.; and Bengio, S. 2016. Neural combinatorial optimization with reinforcement learning. *arXiv:1611.09940*.
- Biewald, L. 2020. Experiment Tracking with Weights and Biases. Software available from wandb.com. Accessed: 2021-07-21.
- Bollobás, B. 2004. *Extremal graph theory*. Courier Corporation.
- Borchers, B. 1999. CSDP, AC library for semidefinite programming. *Optimization methods and Software*, 11(1-4): 613–623.
- Buzzard, K. 2022. The rise of formalism in mathematics. International Congress of Mathematics Plenary Talk. Slides available at www.ma.imperial.ac.uk/~buzzard/xena/pdfs/BuzzardICM2022talk.pdf. Accessed: 2022-08-22.
- Cameron, P. J.; Cilleruelo, J.; and Serra, O. 2007. On monochromatic solutions of equations in groups. *Revista Matemática Iberoamericana*, 23(1): 385–395.
- Csikvári, P. 2022. Note on the sum of the smallest and largest eigenvalues of a triangle-free graph. *Linear Algebra and its Applications*.
- Das, S.; Huang, H.; Ma, J.; Naves, H.; and Sudakov, B. 2013. A problem of Erdős on the minimum number of k -cliques. *Journal of Combinatorial Theory, Series B*, 103(3): 344–373.
- Davies, A.; Veličković, P.; Buesing, L.; Blackwell, S.; Zheng, D.; Tomašev, N.; Tanburn, R.; Battaglia, P.; Blundell, C.; Juhász, A.; et al. 2021. Advancing mathematics by guiding human intuition with AI. *Nature*, 600(7887): 70–74.
- Dueck, G.; and Scheuer, T. 1990. Threshold accepting: A general purpose optimization algorithm appearing superior to simulated annealing. *Journal of computational physics*, 90(1): 161–175.
- Ebsen, O.; and Schacht, M. 2020. Homomorphism thresholds for odd cycles. *Combinatorica*, 40(1): 39–62.
- Edel, Y. 2004. Extensions of generalized product caps. *Designs, Codes and Cryptography*, 31(1): 5–14.
- Ellenberg, J. S.; and Gijswijt, D. 2017. On large subsets of with no three-term arithmetic progression. *Annals of Mathematics*, 339–343.
- Erdős, P. 1962. On the number of complete subgraphs contained in certain graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 7(3): 459–464.
- Even-Zohar, C.; and Linial, N. 2015. A Note on the Inducibility of 4-Vertex Graphs. *Graphs and Combinatorics*, 31(5): 1367–1380.
- Exoo, G. 1998. Some new Ramsey colorings. *the electronic journal of combinatorics*, R29–R29.
- Falgas-Ravry, V.; and Vaughan, E. R. 2013. Applications of the semi-definite method to the Turán density problem for 3-graphs. *Combinatorics, Probability and Computing*, 22(1): 21–54.
- GAP. 2021. *GAP – Groups, Algorithms, and Programming, Version 4.11.1*. The GAP Group.
- Gendreau, M.; Potvin, J.-Y.; et al. 2010. *Handbook of meta-heuristics*, volume 2. Springer.
- Giraud, G. 1979. Sur le probleme de Goodman pour les quadrangles et la majoration des nombres de Ramsey. *Journal of Combinatorial Theory, Series B*, 27(3): 237–253.
- Glover, F. 1986. Future paths for integer programming and links to artificial intelligence. *Computers & operations research*, 13(5): 533–549.
- Glover, F. 1989. Tabu search—part I. *ORSA Journal on computing*, 1(3): 190–206.
- Glover, F. 1990. Tabu search—part II. *ORSA Journal on computing*, 2(1): 4–32.
- Goodman, A. W. 1959. On sets of acquaintances and strangers at any party. *The American Mathematical Monthly*, 66(9): 778–783.
- Greene, J. W.; and Supowit, K. J. 1986. Simulated annealing without rejected moves. *IEEE Transactions on Computer-Aided Design*, 5(1): 221–228.
- Grzesik, A. 2012. On the maximum number of five-cycles in a triangle-free graph. *Journal of Combinatorial Theory, Series B*, 102(5): 1061–1066.
- Grzesik, A.; Lee, J.; Lidický, B.; and Volec, J. 2020. On tripartite common graphs. *arXiv:2012.02057*.
- Hales, T. C. 2005. A proof of the Kepler conjecture. *Annals of mathematics*, 162(3): 1065–1185.
- Heule, M. J. H. 2018. Schur Number Five. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence, AAAI’18/IAAI’18/EAAI’18*. AAAI Press. ISBN 978-1-57735-800-8.
- Kirkpatrick, S.; Gelatt Jr, C. D.; and Vecchi, M. P. 1983. Optimization by Simulated Annealing. *science*, 220(4598): 671–680.
- Lidický, B.; and Pfender, F. 2017. Semidefinite programming and Ramsey numbers. *arXiv:1704.03592*.
- Lu, L.; and Peng, X. 2012. Monochromatic 4-term arithmetic progressions in 2-colorings of \mathbb{Z}_n . *Journal of Combinatorial Theory, Series A*, 119(5): 1048–1065.
- Mathew, K. A.; and Östergård, P. R. 2017. New lower bounds for the Shannon capacity of odd cycles. *Designs, Codes and Cryptography*, 84(1): 13–22.
- McCune, W. 1996. Robbins algebras are boolean. *Association for Automated Reasoning Newsletter*, 35: 1–3.
- McKay, B. D. 1998. Isomorph-free exhaustive generation. *Journal of Algorithms*, 26(2): 306–324.
- McKay, B. D.; and Radziszowski, S. P. 1995. $R(4, 5) = 25$. *Journal of Graph Theory*, 19(3): 309–322.
- McKay, B. D.; and Radziszowski, S. P. 1997. Subgraph counting identities and Ramsey numbers. *journal of combinatorial theory, Series B*, 69(2): 193–209.
- Nieß, S. 2012. Counting monochromatic copies of K_4 : a new lower bound for the Ramsey multiplicity problem. *arXiv:1207.4714*.

Nikiforov, V. 2001. On the minimum number of k -cliques in graphs with restricted independence number. *Combinatorics, Probability and Computing*, 10(4): 361–366.

Pikhurko, O.; Sliáčan, J.; and Tyros, K. 2019. Strong forms of stability from flag algebra calculations. *Journal of Combinatorial Theory, Series B*, 135: 129–178.

Pikhurko, O.; and Vaughan, E. R. 2013. Minimum number of k -cliques in graphs with bounded independence number. *Combinatorics, Probability and Computing*, 22(6): 910–934.

Polu, S.; Han, J. M.; Zheng, K.; Baksys, M.; Babuschkin, I.; and Sutskever, I. 2022. Formal mathematics statement curriculum learning. *arXiv:2202.01344*.

Razborov, A. A. 2007. Flag algebras. *The Journal of Symbolic Logic*, 72(4): 1239–1282.

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

Razborov, A. A. 2010. On 3-hypergraphs with forbidden 4-vertex configurations. *SIAM Journal on Discrete Mathematics*, 24(3): 946–963.

Reiher, C. 2016. The clique density theorem. *Annals of Mathematics*, 683–707.

Roucairol, M.; and Cazenave, T. 2022. Refutation of Spectral Graph Theory Conjectures with Monte Carlo Search. *arXiv:2207.03343*.

Rowley, F. 2022. Improved Lower Bounds for Multicolour Ramsey Numbers using SAT-Solvers. *arXiv:2203.13476*.

Silva, M. K.; Sato, C. M.; et al. 2016. Flag algebras: A first glance. *arXiv:1607.04741*.

Sperfeld, K. 2011. On the minimal monochromatic K_4 -density. *arXiv:1106.1030*.

Swinnerton-Dyer, H.; and Birch, B. 1965. Notes on elliptic curves. II. *Journal für die reine und angewandte Mathematik*, 218: 79–108.

Thomason, A. 1989. A disproof of a conjecture of Erdős in Ramsey theory. *Journal of the London Mathematical Society*, 2(2): 246–255.

Thomason, A. 1997. Graph products and monochromatic multiplicities. *Combinatorica*, 17(1): 125–134.

Vaughan, E. 2013. Flagmatic: A tool for researchers in extremal graph theory. Version 2.0. Accessed: 2022-04-16.

Wagner, A. Z. 2020. Refuting conjectures in extremal combinatorics via linear programming. *Journal of Combinatorial Theory, Series A*, 169: 105130.

Wagner, A. Z. 2021. Constructions in combinatorics via neural networks. *arXiv:2104.14516*.

Wolf, J. 2010. The minimum number of monochromatic 4-term progressions in \mathbb{Z}_p . *Journal of Combinatorics*, 1(1): 53–68.

Yamashita, M.; Fujisawa, K.; Fukuda, M.; Kobayashi, K.; Nakata, K.; and Nakata, M. 2012. Latest developments in the SDPA family for solving large-scale SDPs. In *Handbook on semidefinite, conic and polynomial optimization*, 687–713. Springer.