An approximability-related parameter on graphs—properties and applications

Robert Engström, Tommy Färnqvist, Peter Jonsson, Johan Thapper

Discrete Mathematics and Theoretical Computer Science DMTCS vol. **17:1**, 2015, 33–66

*An Approximability-related Parameter on Graphs – Properties and Applications**

Robert Engström Tommy Färnqvist[1†] Peter Jonsson[1‡] Johan Thapper[2§]

[1] *Department of Computer and Information Science, Linköpings universitet, Sweden*
[2] *Laboratoire d'Informatique Gaspard-Monge, Université Paris-Est Marne-la-Vallée, France*

received 14th Apr. 2012, revised 21st Oct. 2014, accepted 22nd Jan. 2015.

We introduce a binary parameter on optimisation problems called *separation*. The parameter is used to relate the approximation ratios of different optimisation problems; in other words, we can convert approximability (and non-approximability) result for one problem into (non)-approximability results for other problems. Our main application is the problem *(weighted) maximum H-colourable subgraph* (MAX H-COL), which is a restriction of the general *maximum constraint satisfaction problem* (MAX CSP) to a single, binary, and symmetric relation. Using known approximation ratios for MAX k-CUT, we obtain general asymptotic approximability results for MAX H-COL for an arbitrary graph H. For several classes of graphs, we provide near-optimal results under the unique games conjecture. We also investigate separation as a graph parameter. In this vein, we study its properties on circular complete graphs. Furthermore, we establish a close connection to work by Šámal on *cubical colourings* of graphs. This connection shows that our parameter is closely related to a special type of chromatic number. We believe that this insight may turn out to be crucial for understanding the behaviour of the parameter, and in the longer term, for understanding the approximability of optimisation problems such as MAX H-COL.

Keywords: graph H-colouring, approximation, graph homomorphism, circular colouring, combinatorial optimisation, graph theory

*Preliminary versions of the results in this article appeared in the 4th International Computer Science Symposium in Russia (Färnqvist et al. (2009)) and in the 5th Latin-American Algorithms, Graphs, and Optimization Symposium (Engström et al. (2009b)).

†Email: `tommy.farnqvist@liu.se`. Supported by the *National Graduate School in Computer Science* (CUGS), Sweden.

‡Email: `peter.jonsson@liu.se`. Partially supported by the *Center for Industrial Information Technology* (CENIIT) under grant 04.01 and by the *Swedish Research Council* (VR) under grants 2006-4532 and 621-2009-4431.

§Email: `thapper@u-pem.fr`. Partially supported by the *European Community's Seventh Framework Programme* (FP7/2007-2013 Grant Agreement no. 257039) and by the *LIX-Qualcomm Postdoctoral Fellowship*.

1 Introduction

In this article we study an approximation-preserving reducibility called *continuous reduction* (Simon, 1989). A continuous reduction allows the transfer of constant ratio approximation results from one optimisation problem to another. We introduce a binary parameter on optimisation problems which measures the degradation in the approximation guarantee of a given continuous reduction. We call this parameter the *separation* of the two problems.

The separation parameter is then used to study a concrete family of optimisation problems called the maximum H-colourable subgraph problems, or MAX H-COL for short. This family includes the problems MAX k-CUT for which good approximation ratios are known (Frieze and Jerrum, 1997). Starting from these ratios, we use the notion of separation to obtain general approximation results for MAX H-COL. Our main contribution in relation to approximation is Theorem 3.10 which gives a constant approximation ratio of $1 - \frac{1}{r} + \frac{2 \ln k}{k^2}(1 - \frac{1}{r} + o(1))$ for MAX H-COL, when the graph H has clique number r and chromatic number k.

In the setting of MAX H-COL, we view separation as a binary graph parameter. While the initial motivation for introducing this parameter was to study the approximability of optimisation problems, it turns out that separation is a parameter of independent interest in graph theory. We investigate this aspect of separation in the second part of the article. Among the most striking results is the connection between separation and a (generalisation of) *cubical colourings* and *fractional covering by cuts* previously studied by Šámal (2005, 2006, 2012).

1.1 The Separation Parameter

Before we consider continuous reductions and the separation parameter, we formally define optimisation problems. An optimisation problem M is defined over a set of *instances* I_M; each instance $\mathcal{I} \in I_M$ has an associated finite set $\mathsf{Sol}_M(\mathcal{I})$ of *feasible solutions*. The objective is, given an instance $\mathcal{I}$, to find a feasible solution of optimum value, with respect to some *measure* (*objective function*) $m_M(\mathcal{I}, f)$, where $f \in \mathsf{Sol}_M(\mathcal{I})$. The optimum of $\mathcal{I}$ is denoted by $\mathsf{Opt}_M(\mathcal{I})$, and is defined as the largest measure of any solution to $\mathcal{I}$. (We will only consider maximisation problems in this article.) We will make the assumption that every instance of every problem considered has some feasible solution and that all feasible solutions have positive rational measure. Then, the following quantity is always defined, where $\mathcal{I} \in I_M$ and $f \in \mathsf{Sol}_M(\mathcal{I})$.

$$R_M(\mathcal{I}, f) = \frac{m_M(\mathcal{I}, f)}{\mathsf{Opt}_M(\mathcal{I})}.$$

A solution $f \in \mathsf{Sol}_M(\mathcal{I})$ to an instance $\mathcal{I}$ of a maximisation problem M is called *r-approximate* if it satisfies

$$R_M(\mathcal{I}, f) \geq r.$$

An approximation algorithm for M has *approximation ratio* $r(n)$ if, given any instance $\mathcal{I}$ of M, it outputs an $r(|\mathcal{I}|)$-approximate solution. We say that M *can be approximated within* $r(n)$ if there exists a polynomial-time algorithm for M with approximation ratio $r(n)$.

All optimisation problems that we consider belong to the class **NPO**; this class contains the problems for which the instances and solutions can be recognised in polynomial time, the solutions are polynomially bounded in the input size, and the objective function can be computed in polynomial time. An **NPO** problem is in the class **APX** if it can be approximated within a constant factor. If, in addition, for any rational value $0 < r < 1$, there exists an algorithm which, given an instance, returns an r-approximate solution in time polynomial in the size of the instance, then we say that the problem admits a *polynomial-time approximation scheme* (PTAS). Note that the dependence of the time on r may be arbitrary.

A reduction from an **NPO**-problem N to an **NPO**-problem M is specified by two polynomial-time computable functions, φ which maps instances of N to instances of M, and γ which takes an instance $\mathcal{I} \in I_N$ and a solution $f \in \mathsf{Sol}_M(\varphi(\mathcal{I}))$ and returns a solution to $\mathcal{I}$. Completeness in **APX** can be defined using appropriate reductions and it is known that an **APX**-complete problem does not admit a PTAS, unless **P** $=$ **NP**. For a more comprehensive introduction to these classes and their accompanying reductions, see Ausiello et al. (1999) and Crescenzi (1997). Our main focus will be on the following reducibility.

Definition 1.1 (Simon, 1989; Crescenzi, 1997) *A reduction φ, γ from N to M is called a* continuous reduction *if a positive constant α exists such that, for every $\mathcal{I} \in I_N$ and $f \in \mathsf{Sol}_N(\varphi(\mathcal{I}))$, it holds that*

$$R_N(\mathcal{I}, \gamma(\mathcal{I}, f)) \geq \alpha \cdot R_M(\varphi(\mathcal{I}), f). \tag{1}$$

Every continuous reduction preserves membership in **APX**. More specifically, we have the following.

Proposition 1.2 (Simon, 1989) *Assume that there is a continuous reduction from N to M with a constant α. If M can be approximated within a constant ratio r, then N can be approximated within $\alpha \cdot r$.*

Given a fixed continuous reduction φ, γ, we ask the following question:

Which is the largest constant α that satisfies (1)?

To answer this question we take the supremum of all positive constants satisfying (1) over all $\mathcal{I} \in I_N$ and $f \in \mathsf{Sol}_M(\varphi(\mathcal{I}))$.

Definition 1.3 *The* separation of M and N*, given a continuous reduction φ, γ from N to M, is defined as the following quantity.*

$$s(M, N) := \inf_{\substack{\mathcal{I} \in I_N \\ f \in \mathsf{Sol}_M(\varphi(\mathcal{I}))}} \frac{R_N(\mathcal{I}, \gamma(\mathcal{I}, f))}{R_M(\varphi(\mathcal{I}), f)}. \tag{2}$$

Needless to say, the separation is difficult to compute in the general case. Thus, we henceforth concentrate on one particular optimisation problem that is parameterised by graphs. It is, however, important to note that the parameter s can be defined over many different types of optimisation problems and it is by no means restricted to problems parameterised by graphs.

1.2 The Maximum H-Colourable Subgraph Problem

Denote by $\mathcal{G}$ the set of all non-empty, simple, and undirected graphs. For a graph $G \in \mathcal{G}$, let $\mathcal{W}(G)$ be the set of *weight functions* $w : E(G) \to \mathbb{Q}_{\geq 0}$ such that w is not identically 0. For $w \in \mathcal{W}(G)$, we let $\|w\|_1 = \sum_{e \in E(G)} w(e)$ denote the total weight of G with respect to w.

Let G and H be graphs in $\mathcal{G}$. A *graph homomorphism* from G to H is a vertex map which carries the edges in G to edges in H. The existence of such a map is denoted by $G \to H$. If both $G \to H$ and $H \to G$, then G and H are said to be *homomorphically equivalent*. This is denoted by $G \equiv H$.

We now consider the following collection of problems, parameterised by a graph $H \in \mathcal{G}$.

Problem 1 *The weighted maximum H-colourable subgraph problem, or* MAX H-COL *for short, is the maximisation problem with*

Instance: *An edge-weighted graph (G, w), where $G \in \mathcal{G}$ and $w \in \mathcal{W}(G)$.*

Solution: *A subgraph $G' \subseteq G$ such that $G' \to H$.*

Measure: *The weight of G' with respect to w.*

Example 1. Let G be a graph in $\mathcal{G}$. Given a subset of vertices $S \subseteq V(G)$, a *cut* in G with respect to S is the set of edges from a vertex in S to a vertex in $V(G) \setminus S$. The MAX CUT problem asks for the size of a largest cut in G.

More generally, for $k \geq 2$, a k-*cut* in G is the set of edges between S_i and S_j, $i \neq j$, where $S_1, \ldots, S_k$ is a partition of $V(G)$. The MAX k-CUT problem asks for the size of a largest k-cut. This problem is readily seen to be equivalent to finding a largest subgraph of G which has a homomorphism to the complete graph K_k, i.e. finding a largest k-colourable subgraph of (a uniformly edge-weighted graph) G. Hence, for each $k \geq 2$, the problem MAX k-CUT is included in the collection of MAX H-COL problems. It is well known that MAX k-CUT is **APX**-complete, for $k \geq 2$ (Ausiello et al., 1999).

Given an edge-weighted graph (G, w), denote by $mc_H(G, w)$ the measure of an optimal solution to the problem MAX H-COL. Denote by $mc_k(G, w)$ the (weighted) size of a largest k-cut in (G, w). This notation is justified by the equivalence of the problems MAX k-CUT and MAX K_k-COL. The decision version of MAX H-COL, the H-COLOURING problem, has been extensively studied (see the monograph by Hell and Nešetřil (2004) and its many references). Hell and Nešetřil (1990) were the first to show that this problem is in **P** if H contains a loop or is bipartite, and **NP**-complete otherwise.

Assuming that $M \to N$, we consider the reduction φ_1, γ_1 defined as follows. The function φ_1 maps an instance $(G, w) \in I_N$ to $(G, w) \in I_M$ and the function γ_1 maps a solution $G' \in \mathsf{Sol}_M(G, w)$ to the solution $G' \in \mathsf{Sol}_N(G, w)$. Here, $m_M(\varphi_1(\mathcal{I}), f) = m_N(\mathcal{I}, \gamma_1(\mathcal{I}, f))$, so the separation defined in (2) takes the following simplified form.

Definition 1.4 *Let $M, N \in \mathcal{G}$. The* separation of M and N *is defined as the following quantity.*

$$s(M, N) := \inf_{\substack{G \in \mathcal{G} \\ w \in \mathcal{W}(G)}} \frac{mc_M(G, w)}{mc_N(G, w)}. \tag{3}$$

Remark 1.5 *Note that we do not require $M \to N$ in Definition 1.4. The reason for this is that the definition makes sense independently of its connection to the reduction φ_1, γ_1. We can then study the separation parameter independently as a graph parameter.*

In Section 2 we show that the reduction φ_1, γ_1 is indeed a continuous reduction. Proposition 1.2 therefore implies the following.

Lemma 1.6 *Let $M \to N$ be two graphs in $\mathcal{G}$. If mc_M can be approximated within α, then mc_N can be approximated within $\alpha \cdot s(M, N)$. If it is* **NP**-*hard to approximate mc_N within β, then mc_M is not approximable within $\beta / s(M, N)$, unless* **P** = **NP**.

Example 2. Goemans and Williamson (1995) give an algorithm for MAX CUT which is a 0.87856-approximating algorithm for MAX K_2-COL. In Proposition 2.5 we will see that $s(K_2, C_{11}) = 10/11$.

We can now apply Lemma 1.6 with $M = K_2$ and $N = C_{11}$, and we find that this MAX CUT-algorithm approximates MAX C_{11}-COL within $0.87856 \cdot s(K_2, C_{11}) \approx 0.79869$.

1.3 Article Outline

The basic properties of separation for the MAX H-COL family are worked out in Section 2. The main result, Theorem 2.1, provides a simplification of (3) which we then use to obtain explicit values and bounds on separation. In particular, this shows that the reduction φ_1, γ_1 defined above is continuous. A linear programming formulation of separation is presented in Section 2.3.

In Section 3, we use separation to study the approximability of MAX H-COL and investigate optimality issues, for several classes of graphs. Comparisons are made to the bounds achieved by the general MAX 2-CSP-algorithm by Håstad (2005). Our investigation covers a spectrum of graphs, ranging from graphs with few edges and/or containing long shortest cycles to an asymptotic result, Theorem 3.10, for arbitrary graphs. We also look at graphs in the $\mathcal{G}(n, p)$ random graph model, pioneered by Erdős and Rényi (1960).

In order to apply our method successfully to the MAX H-COL problem but also to get a better understanding of the separation parameter, we want to compute some explicit values of $s(M, N)$ for various graphs M and N. To this end, we turn to the *circular complete graphs* in Section 4. We take a close look at 3-colourable circular complete graphs, and amongst other things, find that there are regions of such graphs on which separation is constant. The application of these results to MAX H-COL relies heavily on existing graph homomorphism results, and in this context we will see that a conjecture by Jaeger (1988) has precise and interesting implications (see Section 6.2).

Another way to study separation is to relate it to known graph parameters. In Section 5 we show that our parameter is closely related to a fractional edge-covering problem and an associated "chromatic number", and that we can pass effortlessly between the two views, gaining insights into both. This part is highly inspired by work of Šámal (2005, 2006, 2012) on *cubical colourings* and *fractional covering by cuts*. In particular, our connection to Šámal's work brings about a new family of chromatic numbers that provides us with an alternative way of computing our parameter. We also use our knowledge of the behaviour of separation to decide two conjectures concerning cubical colourings.

Finally, we summarise the future prospects and open problems of the method in Section 6.

2 Basic Properties of the Separation Parameter

In this section we establish a basic theorem on separation, Theorem 2.1, and derive a number of results from it. It follows that the reduction φ_1, γ_1 is continuous. We also give a general bound on the separation parameter, exact values in some special cases, and a linear programming formulation.

Let $N \in \mathcal{G}$ be a graph. An *edge automorphism* of N is a permutation of the edges of N that sends edges with a common vertex to edges with a common vertex. The set of all edge automorphisms is called the *edge automorphism group of* N and its denoted by $\mathrm{Aut}^*(N)$. The graph N is called *edge-transitive* if $\mathrm{Aut}^*(N)$ acts transitively on $E(N)$. Let $\hat{\mathcal{W}}(N)$ be the set of weight functions $w \in \mathcal{W}(N)$ that satisfy $\|w\|_1 = 1$, and for which $w(e) = w(\pi \cdot e)$ for all $e \in E(N), \pi \in \mathrm{Aut}^*(N)$. That is, w is constant on the orbits of $\mathrm{Aut}^*(N)$.

Theorem 2.1 *Let $M, N \in \mathcal{G}$. Then,*

$$s(M, N) = \inf_{w \in \hat{\mathcal{W}}(N)} mc_M(N, w).$$

In particular, when N is edge-transitive,

$$s(M, N) = mc_M(N, 1/e(N)).$$

We postpone the proof of Theorem 2.1 to Section 2.2.

Corollary 2.2 *Let $M \to N$ be graphs in $\mathcal{G}$. The reduction φ_1, γ_1 is continuous with constant $s(M, N)$.*

Proof: For every instance $\mathcal{I} = (G, w)$ of MAX N-COL and solution $G' \to N$, we have, by definition, $R_N(\mathcal{I}, \gamma_1(\mathcal{I}, G')) \geq s(M, N) \cdot R_M(\varphi_1(\mathcal{I}), G')$. It remains to show that $s(M, N)$ is positive. For every edge-weighting w of N, there is at least one edge e with $w(e) \geq 1/|E(N)|$. Since M is non-empty, the subgraph consisting of only e maps homomorphically to M, so $mc_M(N, w) \geq 1/|E(N)|$. By Theorem 2.1, it follows that $s(M, N) \geq 1/|E(N)| > 0$. □

2.1 Exact Values and Bounds

Let $n(G)$ and $e(G)$ denote the number of vertices and edges in G, respectively. Let $\omega(G)$ denote the *clique number of* G; the greatest integer r such that $K_r \to G$. Let $\chi(G)$ denote the *chromatic number of* G; the least integer n such that $G \to K_n$. The *Turán graph* $T(n, r)$ is a graph formed by partitioning a set of n vertices into r subsets, with sizes as equal as possible, and connecting two vertices whenever they belong to different subsets. The properties of Turán graphs are given by the following theorem.

Theorem 2.3 (Turán, 1941) *The following holds:*

1. $e(T(n, r)) = \left\lfloor \left(1 - \frac{1}{r}\right) \cdot \frac{n^2}{2} \right\rfloor$;
2. $\omega(T(n, r)) = \chi(T(n, r)) = r$;
3. *if G is a graph such that $e(G) > e(T(n(G), r))$, then $\omega(G) > r$.*

Using Turán's theorem, we can determine the separation exactly when the second parameter is a complete graph.

Proposition 2.4 *Let H be a graph with $\omega(H) = r$ and let n be an integer such that $\chi(H) \leq n$. Then,*

$$s(H, K_n) = e(T(n, r))/e(K_n) = \left\lfloor \left(1 - \frac{1}{r}\right) \cdot \frac{n^2}{2} \right\rfloor \Big/ \binom{n}{2}.$$

Proof: The graph K_n is edge-transitive. Therefore, by the second part of Theorem 2.1, it suffices to show that $mc_H(K_n, 1/e(K_n)) = e(T(n, r))/e(K_n)$. By definition, $T(n, r)$ is an r-partite subgraph of K_n, so $T(n, r) \to H$. Hence, $mc_H(K_n, 1/e(K_n)) \geq e(T(n, r))/e(K_n)$. Conversely, any subgraph G of K_n such that $G \to H$ must satisfy $\omega(G) \leq \omega(H) = r$. Thus, by Theorem 2.3(3), $e(G) \leq e(T(n, r))$ which implies $mc_H(K_n, 1/e(K_n)) \leq e(T(n, r))/e(K_n)$. □

Next we consider separation between cycles. The even cycles are all bipartite and therefore homomorphically equivalent to K_2. The odd cycles, on the other hand, form a chain between K_2 and $C_3 = K_3$ in the following manner:

$$K_2 \to \cdots \to C_{2i+1} \to C_{2i-1} \to \cdots \to C_3 = K_3.$$

Note that the chain is infinite on the K_2-side. The following proposition gives the separation between pairs of graphs in this chain. The value depends only on the target graph.

Proposition 2.5 *Let $m > k$ be positive integers. Then,*

$$s(K_2, C_{2k+1}) = s(C_{2m+1}, C_{2k+1}) = \frac{2k}{2k+1}.$$

Proof: Since C_{2k+1} is edge-transitive, it suffices by Theorem 2.1 to show that $mc_2(C_{2k+1}) = 2k = mc_{C_{2m+1}}(C_{2k+1})$. Such cuts clearly exist since after removing one edge from C_{2k+1}, the remaining subgraph is isomorphic to a path, and therefore homomorphic to K_2 (and to C_{2m+1}). Moreover, these cuts are the largest possible: $C_{2k+1} \not\to K_2$ and $C_{2k+1} \not\to C_{2m+1}$. □

Given two graphs $M, H \in \mathcal{G}$, it may be difficult to determine $s(M, H)$ directly. However, if we know that H is "homomorphically sandwiched" between M and another graph N, so that $M \to H \to N$, then it turns out that we can use $s(M, N)$ as a lower bound for $s(M, H)$. More generally, we have the following lemma.

Lemma 2.6 *The following implications hold.*

$$\begin{aligned} H \to N \quad &\Longrightarrow \quad s(M, N) \leq s(M, H), \\ M \to K \quad &\Longrightarrow \quad s(M, N) \leq s(K, N). \end{aligned}$$

Proof: Assume that G' is a subgraph of G such that $G' \to H$, and $H \to N$. Then, $G' \to N$, so any solution $G' \subseteq G$ to (G, w) as an instance of MAX H-COL is also a solution to (G, w) as an instance of MAX N-COL of the same measure. It follows that $mc_H(G, w) \leq mc_N(G, w)$, so

$$s(M, H) = \inf_{\substack{G \in \mathcal{G} \\ w \in \mathcal{W}(G)}} \frac{mc_M(G, w)}{mc_H(G, w)} \geq \inf_{\substack{G \in \mathcal{G} \\ w \in \mathcal{W}(G)}} \frac{mc_M(G, w)}{mc_N(G, w)} = s(M, N).$$

The second part follows similarly. □

We will see several applications of Lemma 2.6 in the following sections, but first we will use it to get a general lower bound on the separation parameter.

Proposition 2.7 *Let $M \in \mathcal{G}$ be a fixed graph. Then, for any $N \in \mathcal{G}$,*

$$1 \geq s(M, N) > \sum_{\{u,v\} \in E(M)} \frac{\deg(u)\deg(v)}{2e(M)^2}.$$

Proof: The upper bound follows from Theorem 2.1. For the lower bound, let $n = \chi(N)$ be the chromatic number of N, so that $N \to K_n$. Then, $s(M, N) \geq s(M, K_n) = mc_M(K_n, 1/e(K_n))$ by Lemma 2.6 and the second part of Theorem 2.1, respectively.

To give a lower bound on $mc_M(K_n, 1/e(K_n))$ we apply a standard probabilistic argument. Let $f : V(K_n) \to V(M)$ be a function, chosen randomly as follows: for every $v_k \in V(K_n)$, and $v_m \in V(M)$, the probability that $f(v_k) = v_m$ is equal to $\deg(v_m)/2e(M)$. Every possible function f appears with non-zero probability and each function defines a subgraph of K_n by including those edges that are mapped by f to edges in M. We will show that there is at least one function that defines a subgraph $K \subseteq K_n$ with the right number of edges.

For $e \in E(K_n)$, and $e' = \{u, v\} \in E(M)$, let $Y_{e,e'} = 1$ if f maps e to e', and $Y_{e,e'} = 0$ otherwise. Then, $\mathbb{E}(Y_{e,e'}) = 2 \cdot \deg(u)\deg(v)/(2e(M))^2$. Let $X_e = 1$ if f maps e to *some* edge in $E(M)$, and

$X_e = 0$ otherwise, so that $X_e = \sum_{e' \in E(M)} Y_{e,e'}$. Then, the *total number of edges* in K is equal to $\sum_{e \in E(K_n)} X_e$, and by linearity of expectation,

$$\mathbb{E}(e(K)) = \sum_{e \in E(K_n)} \mathbb{E}(X_e) = e(K_n) \sum_{\{u,v\} \in E(M)} \frac{\deg(u)\deg(v)}{2e(M)^2}.$$

Finally, we note that for an arbitrary fixed vertex $v_m \in V(M)$, the function defined by $f(v) = v_m$ for all $v \in V(K_n)$ defines the empty subgraph, and has a non-zero probability. Since K_n and M are non-empty we have $\mathbb{E}(e(K)) > 0$, so there must exist at least one f which defines a K with strictly more than the expected total number of edges. □

From Proposition 2.7, we have $s(K_m, N) > \frac{m-1}{m}$. It may seem surprising that the value of $s(K_m, N)$ for *any* graph N can be bounded this close to 1, in particular since we can choose m as large as we like. For large m, it follows from Lemma 1.6 that when $K_m \to N$, an algorithm for MAX m-CUT will approximate MAX N-COL almost as well. This seems like a quite strong statement. However, there is a straightforward explanation:

$$s(M, N) = \inf_{\substack{G \in \mathcal{G} \\ w \in \mathcal{W}(G)}} \frac{mc_M(G, w/\|w\|_1)}{mc_N(G, w/\|w\|_1)} \geq \inf_{\substack{G \in \mathcal{G} \\ w \in \mathcal{W}(G)}} mc_M(G, w/\|w\|_1).$$

This implies that a lower bound for mc_M (mc_m in our case) immediately yields a lower bound for $s(M, N)$. Now, a probabilistic argument, analogous to the one in the proof of Proposition 2.7, shows that $mc_m(G, w/\|w\|_1) > \frac{m-1}{m}$, for every G and every (non-zero) w. Hence, the fact that $s(K_m, N)$ is close to 1 is merely a reflection of the fact that every (edge-weighted) graph has a large m-cut.

The *bipartite density* of a graph H is defined as $b(H) := mc_2(H, 1/e(H))$, and is a well-studied graph parameter (Alon, 1996; Berman and Zhang, 2003; Bondy and Locke, 1986; Hopkins and Staton, 1982; Locke, 1990). From Theorem 2.1, it follows that $b(H) = s(K_2, H)$ whenever H is edge-transitive. The following proposition shows that separation is invariant under homomorphic equivalence. Note that this is not true for bipartite density: Let T be a triangle and let H be the graph of two triangles sharing a common edge. In this case, $T \equiv H$, but $b(T) = 2/3 \neq 4/5 = b(H)$.

Proposition 2.8 *Let $M \equiv K$ and $N \equiv H$ be two pairs of homomorphically equivalent graphs. Then, $s(M, N) = s(K, H)$.*

Proof: Using Lemma 2.6, $H \to N$, and $M \to K$, we get $s(M, N) \leq s(M, H) \leq s(K, H)$. On the other hand, $N \to H$ and $K \to M$, so $s(K, H) \leq s(K, N) \leq s(M, N)$. □

2.2 Exploiting Symmetries (Proof of Theorem 2.1)

In this section we prove Theorem 2.1. First we need to establish a number of lemmas. The optimum $mc_H(G, w)$ is sub-linear with respect to the weight function, as is shown by the following result.

Lemma 2.9 *Let $G, H \in \mathcal{G}$, $\alpha \in \mathbb{Q}_{\geq 0}$ and let $w, w_1, \ldots, w_r \in \mathcal{W}(G)$ be weight functions on G. Then,*

- $mc_H(G, \alpha \cdot w) = \alpha \cdot mc_H(G, w)$,
- $mc_H(G, \sum_{i=1}^r w_i) \leq \sum_{i=1}^r mc_H(G, w_i)$.

Proof: The first part is trivial. For the second part, let G' be an optimal solution to the instance $(G, \sum_{i=1}^{r} w_i)$ of MAX H-COL. Then, the measure of this solution equals the sum of the measures of G' as a (possibly suboptimal) solution to each of the instances (G, w_i). $\square$

The solutions to a MAX H-COL instance have an alternative description, which is better suited for calculations: for any vertex map $f : V(G) \to V(H)$, let $f^{\#} : E(G) \to E(H)$ be the (partial) edge map induced by f (i.e. $f^{\#}$ maps an edge $\{u, v\}$ in $E(G)$ to $\{f(u), f(v)\}$ if the latter is in $E(H)$, and otherwise, $f^{\#}(\{u, v\})$ is undefined). Each vertex map f then determines a subgraph $G' = (f^{\#})^{-1}(H) \subseteq G$, but two distinct functions, f and g, do not necessarily determine distinct subgraphs. In this notation $h : V(G) \to V(H)$ is a graph homomorphism precisely when $(h^{\#})^{-1}(E(H)) = E(G)$ or, alternatively, when $h^{\#}$ is a total function. The set of solutions to an instance (G, w) of MAX H-COL can then be taken to be the set of vertex maps $f : V(G) \to V(H)$ with the measure

$$m_H(f) = \sum_{e \in (f^{\#})^{-1}(E(H))} w(e). \tag{4}$$

We will predominantly use this description of solutions.

Let $f : V(G) \to V(H)$ be a solution to the instance (G, w) of MAX H-COL, and define $w_f \in \mathcal{W}(H)$ as follows:

$$w_f(e) = \sum_{e' \in (f^{\#})^{-1}(e)} \frac{w(e')}{mc_H(G, w)}. \tag{5}$$

Note that $\|w_f\|_1 = 1$ iff f is optimal. The following lemma and its corollary establishes half of Theorem 2.1.

Lemma 2.10 *Let $M, N \in \mathcal{G}$ be two graphs. Then, for every $G \in \mathcal{G}$, every $w \in \mathcal{W}(G)$, and any solution f to (G, w) of* MAX N-COL*, it holds that*

$$\frac{mc_M(G, w)}{mc_N(G, w)} \geq mc_M(N, w_f).$$

Proof: Arbitrarily choose an optimal solution $g : V(N) \to V(M)$ to the instance (N, w_f) of MAX M-COL. Then, $g \circ f$ is a solution to (G, w) as an instance of MAX M-COL (Figure 1).

$$\begin{tikzcd}
(G, w) \arrow[rr, "f"] \arrow[dr, "g \circ f"'] & & (N, w_f) \arrow[dl, "g"] \\
& M &
\end{tikzcd}$$

Fig. 1: The relation between the graphs and solutions in the proof of Lemma 2.10.

Note that we have $((g \circ f)^{\#})^{-1} = (f^{\#})^{-1} \circ (g^{\#})^{-1}$. The measure of this solution is thus

$$\begin{aligned} m_M(g \circ f) &= \sum_{e \in ((g \circ f)^{\#})^{-1}(E(M))} w(e) = \sum_{e' \in (g^{\#})^{-1}(E(M))} \sum_{e \in (f^{\#})^{-1}(e')} w(e) \\ &= \sum_{e' \in (g^{\#})^{-1}(E(M))} w_f(e') \cdot mc_N(G, w) = mc_M(N, w_f) \cdot mc_N(G, w). \end{aligned}$$

Clearly, the measure of $g \circ f$ is at most $mc_M(G, w)$, so

$$mc_M(G, w) \geq mc_M(N, w_f) \cdot mc_N(G, w).$$

The result follows after division by $mc_N(G, w)$. □

From Lemma 2.10, we have the following corollary, which shows that it is possible to eliminate $G \in \mathcal{G}$ from the infimum in the definition of s.

Corollary 2.11 *Let $M, N \in \mathcal{G}$ be two graphs. Then,*

$$s(M, N) = \inf_{\substack{w \in \mathcal{W}(N) \\ \|w\|_1 = 1}} mc_M(N, w).$$

Proof: First, we fix some optimal solution $f = f(G, w)$ for each choice of (G, w). By taking infima over G and w on both sides of the inequality in Lemma 2.10, we then have

$$s(M, N) \geq \inf_{\substack{G \in \mathcal{G} \\ w \in \mathcal{W}(G)}} mc_M(N, w_f) \geq \inf_{\substack{w \in \mathcal{W}(N) \\ \|w\|_1 = 1}} mc_M(N, w),$$

where the second inequality holds since $\|w_f\|_1 = 1$ for any optimal f.

For the other direction, we specialise G to N, and restrict w to obtain:

$$s(M, N) \leq \inf_{\substack{w \in \mathcal{W}(N) \\ \|w\|_1 = 1}} \frac{mc_M(N, w)}{mc_N(N, w)} = \inf_{\substack{w \in \mathcal{W}(N) \\ \|w\|_1 = 1}} mc_M(N, w).$$

This concludes the proof. □

Proof of Theorem 2.1: From Corollary 2.11, we have that

$$s(M, N) = \inf_{\substack{w \in \mathcal{W}(N) \\ \|w\|_1 = 1}} mc_M(N, w) \leq \inf_{w \in \hat{\mathcal{W}}(N)} mc_M(N, w).$$

To complete the first part of the theorem, it will be sufficient to prove that for *any* graph $G \in \mathcal{G}$ and $w \in \mathcal{W}(G)$, there is a $w' \in \hat{\mathcal{W}}(N)$ such that the following inequality holds.

$$\frac{mc_M(G, w)}{mc_N(G, w)} \geq mc_M(N, w'). \tag{6}$$

Taking infima on both sides of this inequality then shows that

$$s(M, N) \geq \inf_{w' \in \hat{\mathcal{W}}(N)} mc_M(N, w').$$

Let $A = \mathrm{Aut}^*(N)$ be the edge automorphism group of N and let $\pi \in A$ be an arbitrary edge automorphism. If f is an optimal solution to (G, w) as an instance of MAX N-COL, then so is $\pi \circ f$. By Lemma 2.10, inequality (6) is satisfied by $w_{\pi \circ f}$. Summing π in this inequality over A gives

$$|A| \cdot \frac{mc_M(G, w)}{mc_N(G, w)} \geq \sum_{\pi \in A} mc_M(N, w_{\pi \circ f}) \geq mc_M(N, \sum_{\pi \in A} w_{\pi \circ f}),$$

where the last inequality follows from Lemma 2.9. The weight function $\sum_{\pi \in A} w_{\pi \circ f}$ is determined as follows:

$$\begin{aligned}\sum_{\pi \in A} w_{\pi \circ f}(e) &= \sum_{\pi \in A} \frac{\sum_{e' \in (f^{\#})^{-1}(\pi \cdot e)} w(e')}{mc_N(G,w)} \\ &= \frac{|A|}{|Ae|} \cdot \frac{\sum_{e' \in (f^{\#})^{-1}(Ae)} w(e')}{mc_N(G,w)},\end{aligned}$$

where Ae denotes the orbit of e under A. We have now shown that the inequality in (6) is satisfied by $w' = \sum_{\pi \in A} w_{\pi \circ f}/|A|$, and that w' is in $\hat{\mathcal{W}}(N)$. The first part of the theorem follows.

For the second part, note that when the edge automorphism group A acts transitively on $E(N)$, there is only one orbit $Ae = E(N)$ for all $e \in E(N)$. Then, the weight function w' is given by

$$w'(e) = \frac{1}{e(N)} \cdot \frac{\sum_{e' \in (f^{\#})^{-1}(E(N))} w(e')}{mc_N(G,w)} = \frac{1}{e(N)} \cdot \frac{mc_N(G,w)}{mc_N(G,w)},$$

since f is optimal. □

2.3 A Linear Programming Formulation

In this section, we first derive a linear program for the separation parameter based on Corollary 2.11. Later we will see how to reduce the size of this program, but it serves as a good first exercise, and it will also be used for comparison with the linear program studied in Section 5.

Each vertex map $f : V(N) \to V(M)$ induces an edge map $f^{\#}$, which provides a lower bound on the separation:

$$\sum_{e \in (f^{\#})^{-1}(E(M))} w(e) \le s(M,N). \tag{7}$$

By Corollary 2.11, we want to find the least s such that for *some* weight function $w \in \mathcal{W}(N)$, $\|w\|_1 = 1$, the inequalities (7) hold. Let the variables of the linear program be $\{w_e\}_{e \in E(N)}$ and s. We then have the following linear program for $s(M,N)$.

$$\begin{array}{lll} \text{Minimise} & s & \\ \text{subject to} & \displaystyle\sum_{e \in (f^{\#})^{-1}(E(M))} w_e \le s & \text{for each } f : V(N) \to V(M), \\ & \displaystyle\sum_{e \in E(M)} w_e = 1, & \\ & w_i, s \ge 0 & \end{array} \tag{8}$$

Given an optimal solution $\{w_e\}_{e \in E(N)}, s$ to (8), a weight function which minimises $mc_M(N,w)$ is given by $w(e) = w_e$ for $e \in E(N)$. The measure of this solution is $s = s(M,N)$. The program will clearly be very large with $|E(N)| + 1$ variables and $|V(M)|^{|V(N)|}$ inequalities. Fortunately it can be improved upon.

From Theorem 2.1 it follows that in order to determine $s(M,N)$, it is sufficient to minimise $mc_M(N,w)$ over $w \in \hat{\mathcal{W}}(N)$. We can use this to describe a smaller linear program for computing $s(M,N)$. Let $A_1, \ldots, A_r$ be the orbits of Aut*(N). The measure of a solution f when $w \in \hat{\mathcal{W}}(N)$ is equal to $\sum_{i=1}^{r} f_i \cdot w_i$, where w_i is the weight of an edge in A_i and f_i is the number of edges in A_i which are mapped to an edge in M by f. Note that given a w, the measure of a solution f depends only on the vector $(f_1, \ldots, f_r) \in \mathbb{N}^r$. Therefore, take the solution space to be the set of such vectors.

$$F = \{\, (f_1, \ldots, f_r) \mid f \text{ is a solution to } (N, w) \text{ of } \textsc{Max } M\text{-}\textsc{Col} \,\}.$$

Let the variables of the linear program be $w_1, \ldots, w_r$ and s, where w_i represents the weight of each element in the orbit A_i and s is an upper bound on the solutions.

$$\begin{array}{lll} \text{Minimise} & s & \\ \text{subject to} & \sum_{i=1}^{r} f_i \cdot w_i \leq s & \text{for each } (f_1, \ldots, f_r) \in F, \\ & \sum_{i=1}^{r} |A_i| \cdot w_i = 1, & \\ & w_i, s \geq 0 & \end{array} \tag{9}$$

Given an optimal solution w_i, s to this program, a weight function which minimises $mc_M(N, w)$ is given by $w(e) = w_i$ for $e \in A_i$. The measure of this solution is $s = s(M, N)$.

Example 3. The *wheel graph* on k vertices, W_k, is a graph that contains a cycle of length $k - 1$ plus a vertex v, which is not in the cycle, such that v is connected to every other vertex. We call the edges of the $k - 1$-cycle *outer edges* and the remaining $k - 1$ edges *spokes*. It is easy to see that the clique number of W_k is equal to 4 when $k = 4$ (W_4 is isomorphic to K_4) and that it is equal to 3 in all other cases. Furthermore, W_k is 3-colourable if and only if k is odd, and 4-colourable otherwise. This implies that for odd k, the wheel graphs are homomorphically equivalent to K_3.

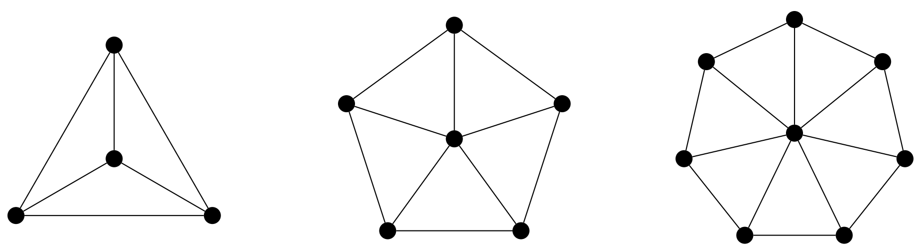

Fig. 2: The wheel graphs W_4, W_6, and W_8.

We will determine $s(K_3, W_k)$ for even $k \geq 6$ using the previously described construction of a linear program. The first three wheel graphs for even k are shown in Figure 2. Note that the group action of Aut*(W_k) on $E(W_k)$ has two orbits, one which consists of all outer edges and one which consists of all the spokes. If we remove one outer edge or one spoke from W_k, then the resulting graph can be mapped homomorphically to K_3. Therefore, it suffices to choose $F = \{f, g\}$ with $f = (k - 1, k - 2)$ and $g = (k - 2, k - 1)$ since all other solutions will have a smaller measure than at least one of these. The program for W_k looks as follows:

$$\begin{array}{ll} \text{Minimise} & s \\ \text{subject to} & (k-1)\cdot w_1 + (k-2)\cdot w_2 \leq s \\ & (k-2)\cdot w_1 + (k-1)\cdot w_2 \leq s \\ & (k-1)\cdot w_1 + (k-1)\cdot w_2 = 1 \\ & w_1, w_2, s \geq 0 \end{array}$$

The optimal solution to this program is given by $w_1 = w_2 = 1/(2k-2)$, with $s(K_3, W_k) = s = (2k-3)/(2k-2)$.

Example 4. In the previous example, the two weights in the optimal solution were equal. Here, we provide another example, where the weights turn out to be different for different orbits. The *circular complete graph* $K_{8/3}$ has vertex set $\{v_0, v_1, \ldots, v_7\}$, which is placed uniformly around a circle. There is an edge between any two vertices which are at a distance at least 3 from each other. Figure 3 depicts this graph.

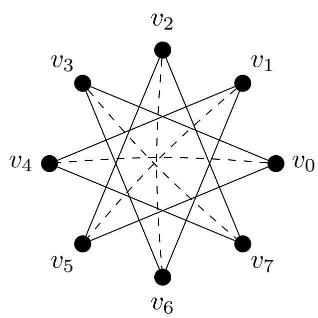

Fig. 3: The graph $K_{8/3}$.

We will now calculate $s(K_2, K_{8/3})$. Each vertex is at a distance 4 from exactly one other vertex, which means that there are 4 such edges. These edges, which are dashed in the figure, form one orbit under the action of $\text{Aut}^*(K_{8/3})$ on $E(K_{8/3})$. The remaining 8 edges (solid) form a second orbit. Let $V(K_2) = \{u_0, u_1\}$. We can obtain a solution f by mapping $f(v_i) = u_0$ if i is even, and $f(v_i) = u_1$ if i is odd. This solution will map all solid edges to K_2, but none of the dashed, hence $f = (0, 8)$. We obtain a second solution g by mapping $g(v_i) = u_0$ for $0 \leq i < 4$, and $g(v_i) = u_1$ for $4 \leq i < 8$. This solution will map all but two of the solid edges in $K_{8/3}$ to K_2, hence $g = (4, 6)$.

Let h be an arbitrary solution. If $h_1 = 0$, then $h_1 \cdot w_1 + h_2 \cdot w_2 \leq 0 \cdot w_1 + 8 \cdot w_2 = f_1 \cdot w_1 + f_2 \cdot w_2$. Otherwise, $h_1 > 0$, so it maps at least one dashed edge, say the edge between v_0 and v_4, to K_2. There are two edge-disjoint even-length paths using solid edges from v_0 to v_4, one via v_3 and one via v_5. The solution h sends at most 3 of these solid edges from each path to K_2. Hence, $h_2 \leq 6$, so $h_1 \cdot w_1 + h_2 \cdot w_2 \leq 4 \cdot w_1 + 6 \cdot w_2 = g_1 \cdot w_1 + g_2 \cdot w_2$. Therefore, the inequalities given by f and g imply the inequality given by any other solution, and we have the following program for $s(K_2, K_{8/3})$:

$$\begin{array}{ll} \text{Minimise} & s \\ \text{subject to} & 0 \cdot w_1 + 8 \cdot w_2 \leq s \\ & 4 \cdot w_1 + 6 \cdot w_2 \leq s \\ & 4 \cdot w_1 + 8 \cdot w_2 = 1 \\ & w_1, w_2, s \geq 0 \end{array}$$

The optimal solution to this program is given by $w_1 = 1/20$, $w_2 = 1/10$, and $s(K_2, K_{8/3}) = s = 4/5$.

3 Approximation Bounds for MAX H-COL

In this section we apply the reduction φ_1, γ_1 and use some of the explicit values obtained for s in Section 2 to bound the approximation ratio of MAX H-COL for various families of graphs. First, we would like to remind the reader of some earlier results and also give a hint of what to expect when we start studying the approximability of MAX H-COL.

The probabilistic argument in Proposition 2.7 shows that MAX H-COL is in **APX**. Furthermore, Jonsson et al. (2009) have shown that whenever H is loop-free, MAX H-COL does not admit a PTAS, and otherwise the problem is trivial. Let us have a closer look at a concrete, well-known example: the MAX CUT problem. This problem was one of Karp's original 21 **NP**-complete problems (Karp, 1972) and has a trivial approximation ratio of 1/2, which is obtained by assigning each vertex randomly to either part of the partition. The trivial randomised algorithm is easy to derandomise; Sahni and Gonzalez (1976) gave the first such approximation algorithm. The 1/2 ratio was in fact essentially the best known ratio for MAX CUT until 1995, when Goemans and Williamson (1995), using *semidefinite programming* (SDP), achieved the ratio 0.87856 mentioned in Example 2. Frieze and Jerrum (1997) determined lower bounds on the approximation ratios for MAX k-CUT using similar SDP techniques. Sharpened results for small values of k have later been obtained by de Klerk et al. (2004). Håstad (2005) has shown that SDP is a universal tool for solving the general MAX 2-CSP problem (where every constraint only involves two variables) over any (finite) domain, in the sense that it establishes non-trivial approximation results for all of those problems. Until recently no other method than SDP was known to yield a non-trivial approximation ratio for MAX CUT. Trevisan (2009) broke this barrier by using techniques from algebraic graph theory to reach an approximation guarantee of 0.531. Soto (2009) later improved this bound to 0.6142 by a more refined analysis.

Khot (2002) suggested the *unique games conjecture* (UGC) as a possible direction for proving inapproximability properties of some important optimisation problems. The conjecture states the following (equivalent form from Khot et al. (2007)):

Conjecture 3.1 *Given any $\delta > 0$, there is a prime p such that given a set of linear equations $x_i - x_j = c_{ij} \pmod{p}$, it is* **NP***-hard to decide which one of the following is true:*

- *There is an assignment to the x_i's which satisfies at least a $1 - \delta$ fraction of the constraints.*
- *All assignments to the x_i's can satisfy at most a δ fraction of the constraints.*

Under the assumption that the UGC holds, Khot et al. (2007) proved the approximation ratio achieved by the Goemans and Williamson algorithm for MAX CUT to be essentially optimal and also provided upper bounds on the approximation ratio for MAX k-CUT, $k > 2$. The proof for the MAX CUT case crucially relies on Gaussian Analysis. In particular, it uses Borell's Theorem to answer the question of partitioning $\mathbb{R}^n$ into two sets of equal Gaussian measure so as to minimise the Gaussian noise-sensitivity, thereby transferring a Fourier analytic question to a geometric one.

Recently, Isaksson and Mossel (2012) showed that a similar geometric conjecture have further implications for the approximability of MAX k-CUT.

Conjecture 3.2 *The standard k-simplex partition is the most noise-stable balanced partition of $\mathbb{R}^n$ with $n \geq k-1$.*

A partition of $\mathbb{R}^n$ into k measurable sets $A_1, \ldots, A_k$ is called *balanced* if each A_i has Gaussian measure $1/k$. The ε-noise sensitivity is defined as the probability that two $(1-2\varepsilon)$-correlated n-dimensional standard Gaussian points $x, y \in \mathbb{R}^n$ belong to different sets in the partition. The standard k-simplex partition of $\mathbb{R}^n$ is obtained by letting $\mathbb{R}^n = \mathbb{R}^{k-1} \times \mathbb{R}^{n-k+1}$ and then partitioning $\mathbb{R}^{k-1}$ into k regular simplicial cones. Assuming this *standard simplex conjecture* (SSC) and the UGC, Isaksson and Mossel show that the Frieze and Jerrum SDP relaxation obtains the optimal approximation ratio for MAX k-CUT.

Every MAX H-COL problem can be viewed as a MAX CSP(Γ) problem, where Γ, the so called *constraint language*, is the set containing the single, binary, and symmetric relation given by the edge set of H. Raghavendra (2008) has presented approximation algorithms for every MAX CSP(Γ) problem based on semi-definite programming. Under the UGC, these algorithms optimally approximate MAX CSP(Γ) in polynomial-time, i.e. no other polynomial-time algorithm can approximate the problem substantially better. However, it seems notoriously difficult to determine the approximation ratio implied by this result, for a given constraint language: Raghavendra and Steurer (2009) show that this ratio can in principle be computed, but the algorithm is doubly exponential in the size of the domain. In combination with our results, such ratios could be used to confirm or disprove the UGC.

3.1 A General Reduction

Our main tool will be a generalisation of the reduction introduced in Section 1.2. Let M and N be (arbitrary) undirected graphs and consider the following reduction, φ_2, γ_2, from MAX N-COL to MAX M-COL: The function φ_2 maps an instance $(G, w) \in I_N$ to $(G, w) \in I_M$. Let $f : V(G) \to V(M)$ be a solution to (G, w). Let $g : V(M) \to V(N)$ be an optimal solution in $\mathsf{Sol}_N(M, w_f)$ (see (5)). The function γ_2 maps f to $g \circ f$.

Proposition 3.3 *The reduction φ_2, γ_2 from* MAX M-COL *to* MAX N-COL *is continuous with constant $s(M, N) \cdot s(N, M)$.*

Proof: First, we argue that γ_2 can be computed in polynomial time. We must show that g can be found in polynomial time. An optimal solution to (M, w_f) can be obtained by brute force. This takes $|V(M)|^{|V(N)|}$ times the time to evaluate a candidate solution g'. The measure of a solution depends on w_f, and thereby on f, but for a given candidate g', it can clearly be obtained in polynomial time. Since M and N are fixed, the total time is polynomial in the size of (G, w).

Next, we show that φ_2, γ_2 is continuous with constant $s(M, N) \cdot s(N, M)$.

$$
\begin{aligned}
m_N(g \circ f) &= m_M(f) \cdot mc_N(M, w_f) \\
&\geq m_M(f) \inf_{\substack{w \in \mathcal{W}(M) \\ \|w\|_1 = 1}} mc_N(M, w) \\
&= m_M(f) \cdot s(N, M),
\end{aligned}
$$

where the final equality follows from Corollary 2.11. From (3) we have the inequality $mc_N(G, w) \leq mc_M(G, w)/s(M, N)$ for all $G \in \mathcal{G}$ and $w \in \mathcal{W}(G)$. Consequently, with $\mathcal{I} = (G, w)$,

$$\begin{aligned} R_N(\mathcal{I}, \gamma_2(\mathcal{I}, f)) &= \frac{m_N(g \circ f)}{mc_N(G, w)} \\ &\geq s(M, N) \cdot s(N, M) \cdot \frac{m_M(f)}{mc_M(G, w)} \\ &= s(M, N) \cdot s(N, M) \cdot R_M(\varphi_2(\mathcal{I}), f). \end{aligned}$$

Since $s(M, N), s(N, M) > 0$, it follows that the reduction is continuous. □

As a direct consequence (using Proposition 1.2), we get the following generalisation of Lemma 1.6.

Lemma 3.4 *Let $M, N \in \mathcal{G}$. If mc_M can be approximated within α, then mc_N can be approximated within $\alpha \cdot s(M, N) \cdot s(N, M)$. If it is* **NP***-hard to approximate mc_N within β, then mc_M is not approximable within $\beta/(s(M, N) \cdot s(N, M))$, unless* **P** = **NP**.

The symmetric nature of this result has some interesting consequences. For example, It is possible to show that $1 - s(M, N) \cdot s(N, M)$ is a metric on the space of graphs taken modulo homomorphic equivalence; cf. Färnqvist et al. (2009).

Our main algorithmic tools will be the following two theorems.

Theorem 3.5 (Goemans and Williamson, 1995) MAX CUT *can be approximated within*

$$\alpha_{GW} = \min_{0<\theta<\pi} \frac{\theta/\pi}{(1 - \cos\theta)/2} \approx 0.87856.$$

A few logarithms will appear in the upcoming expressions. We fix the notation $\ln x$ for the *natural logarithm* of x, and $\log y$ for the *base-2 logarithm* of y.

Theorem 3.6 (Frieze and Jerrum, 1997) MAX k-CUT *can be approximated within*

$$\alpha_k = 1 - \frac{1}{k} + \frac{2\ln k}{k^2}(1 + o(1)).$$

We note that de Klerk et al. (2004) have presented the sharpest known bounds on α_k for small values of k. Table 1 lists α_{GW} together with the first of these lower bounds.

k	2	3	4	5	6
$\alpha_k \geq$	0.87856	0.836008	0.857487	0.876610	0.891543

Tab. 1: The best known lower bounds on α_k for $k = 2, \ldots, 6$.

Håstad (2005) has shown the following.

Theorem 3.7 (Håstad, 2005) *There is an absolute constant $c > 0$ such that mc_H can be approximated within*

$$1 - \frac{t(H)}{n^2} \cdot \Big(1 - \frac{c}{n^2 \log n}\Big),$$

where $n = n(H)$ and $t(H) = n^2 - 2e(H)$.

We will compare the performance of Håstad's algorithm on MAX H-COL with the performance of the algorithms in Theorems 3.5 and 3.6 when analysed using the reduction φ_2, γ_2 and estimates or exact values of the separation parameter. For this purpose, we introduce two functions, FJ_k and *Hå*, such that, if H is a graph, then $FJ_k(H)$ denotes the best bound on the approximation guarantee when Frieze and Jerrum's algorithm for MAX k-CUT is applied to the problem mc_H, while *Hå*(H) is the guarantee when Håstad's algorithm is used to approximate mc_H. This comparison is not entirely fair since Håstad's algorithm was not designed with the goal of providing optimal results; the goal was to beat random assignments. However, it is currently the best known algorithm for approximating MAX H-COL, for arbitrary $H \in \mathcal{G}$, which also provides an easily computable bound on the guaranteed approximation ratio; this is in contrast with the conjectured optimal algorithms of Raghavendra (2008) (see the discussion in Section 6).

Near-optimality of our approximation method will be investigated using results depending on Khot's unique games conjecture (Conjecture 3.1). Hence, we will henceforth assume that the UGC is true, which implies the following inapproximability results.

Theorem 3.8 (Khot et al., 2007) *Under the assumption that the UGC is true, the following holds:*

- *For every $\varepsilon > 0$, it is* **NP**-*hard to approximate mc_2 within $\alpha_{GW} + \varepsilon$.*
- *It is* **NP**-*hard to approximate mc_k within*

$$1 - \frac{1}{k} + \frac{2\ln k}{k^2} + \mathcal{O}\Big(\frac{\ln\ln k}{k^2}\Big).$$

Furthermore, assuming the standard simplex conjecture (Conjecture 3.2), we have the following strengthening of Theorem 3.8.

Theorem 3.9 (Isaksson and Mossel, 2012) *Under the assumption that the UGC and the SSC are true, the following holds:*

- *For every $\varepsilon > 0$, it is* **NP**-*hard to approximate mc_k within $\alpha_k + \varepsilon$.*

3.2 Asymptotic Performance

Next, we derive a general, asymptotic, result on the performance of our method.

Theorem 3.10 *Let $H \in \mathcal{G}$ be a graph with $\omega(H) = r$ and $\chi(H) = k$. Then,*

$$FJ_k(H) = 1 - \frac{1}{r} + \frac{2\ln k}{k^2}\Big(1 - \frac{1}{r} + o(1)\Big), \text{ where } o(1) \text{ is with respect to } k.$$

Furthermore, mc_H cannot be approximated within

$$1 - \frac{1}{k} + \frac{2\ln r}{r^2}\big(1 + o(1)\big), \text{ where } o(1) \text{ is with respect to } r.$$

Proof: By Proposition 2.4, we have

$$s(H, K_k) = \left(\left(1 - \frac{1}{r}\right) \cdot \frac{k^2}{2} + \mathcal{O}(1)\right) \Big/ \binom{k}{2} = \left(1 - \frac{1}{r}\right)\left(1 + \frac{1}{k-1}\right) + \mathcal{O}(k^{-2}).$$

By Lemma 3.4 and Theorem 3.6, we then have

$$\begin{aligned}
FJ_k(H) &= s(H, K_k) \cdot \alpha_k \\
&= \left(\left(1 - \frac{1}{r}\right)\left(1 + \frac{1}{k-1}\right) + \mathcal{O}(k^{-2})\right)\left(1 - \frac{1}{k} + \frac{2\ln k}{k^2}(1 + o(1))\right) \\
&= \left(1 - \frac{1}{r}\right)\left(1 - \frac{1}{k} + \frac{1}{k-1} - \frac{1}{k(k-1)} + \frac{2\ln k}{k^2}(1 + o(1))\right),
\end{aligned}$$

and the first inequality follows, since $-\dfrac{1}{k} + \dfrac{1}{k-1} - \dfrac{1}{k(k-1)} = 0$.

For the second part, we have $s(K_r, H) \geq s(K_r, K_k) = s(H, K_k)$ by Lemma 2.6 and Proposition 2.4, so

$$1/s(K_r, H) \leq \binom{k}{2} \Big/ \left(\left(1 - \frac{1}{r}\right) \cdot \frac{k^2}{2} + \mathcal{O}(1)\right) = \left(1 + \frac{1}{r-1}\right)\left(1 - \frac{1}{k}\right) + \mathcal{O}(k^{-2}).$$

Note the similarity between this upper bound and the expression for $s(H, K_k)$. In fact, without losing any precision in the following calculations, we could replace the $\mathcal{O}(k^{-2})$-term by $\mathcal{O}(r^{-2})$. By Lemma 3.4, mc_H cannot be approximated within $\alpha_r/s(K_r, H)$. An upper bound for $\alpha_r/s(K_r, H)$ can now be calculated as in the first part with r and k swapped. In the final expression we drop $-\frac{1}{k}$ from the last parenthesis since this is absorbed by $o(1)$. □

To give an upper bound on the performance of Håstad's algorithm, we can proceed as follows: Let $n = n(H)$ and $r = \omega(H)$, $k = \chi(H)$ as in the proposition. By Theorem 2.3, $e(H) \leq \left\lfloor \left(1 - \frac{1}{r}\right) \cdot \frac{n^2}{2} \right\rfloor$, hence $\dfrac{2e(H)}{n^2} \leq 1 - \dfrac{1}{r}$, and

$$\begin{aligned}
H\mathring{a}(H) &= 1 - \left(1 - \frac{2e(H)}{n^2}\right)\left(1 - \frac{c}{n^2 \log n}\right) \\
&= \frac{2e(H)}{n^2}\left(1 - \frac{c}{n^2 \log n}\right) + \frac{c}{n^2 \log n} \\
&\leq 1 - \frac{1}{r} + \frac{c}{k^2 \log k}.
\end{aligned}$$

We see that our algorithm performs asymptotically better.

3.3 Some Specific Graph Classes

Next, we investigate the performance of our method on sequences of graphs "tending to" K_2 (K_3) in the sense that the separation of K_2 (K_3) and a graph H_k from the sequence tends to 1 as k tends to infinity. In several cases, the *girth* of the graphs plays a central role. The girth of a graph G is the length of a shortest cycle in G. The *odd girth of* G is the length of a shortest odd cycle in G. Hence, if G has odd girth g, then $C_g \to G$, but $C_{2k+1} \not\to G$ for $3 \leq 2k+1 < g$.

Proposition 3.11 *We have the following bounds.*

1. *Let* $k \geq 1$. *Then,* $FJ_2(C_{2k+1}) \geq \frac{2k}{2k+1} \cdot \alpha_{GW}$ *and* $mc_{C_{2k+1}}$ *cannot be approximated within* $\frac{2k+1}{2k} \cdot \alpha_{GW} + \varepsilon$, *for any* $\varepsilon > 0$.

2. *Let $m > k \geq 4$ and let H be a graph on m vertices and with odd girth $g \geq 2k+1$ and minimum degree $\delta(H) \geq \frac{2m-1}{2(k+1)}$. Then, $FJ_2(H) \geq \frac{2k}{2k+1} \cdot \alpha_{GW}$ and mc_H cannot be approximated within $\frac{2k+1}{2k} \cdot \alpha_{GW} + \varepsilon$, for any $\varepsilon > 0$.*

3. *Let H be a planar graph with girth at least $g = \frac{20k-2}{3}$. Then, $FJ_2(H) \geq \frac{2k}{2k+1} \cdot \alpha_{GW}$ and mc_H cannot be approximated within $\frac{2k+1}{2k} \cdot \alpha_{GW} + \varepsilon$, for any $\varepsilon > 0$.*

4. *Let $k \geq 6$ be even. Then, $FJ_3(W_k) \geq \frac{2k-3}{2k-2} \cdot \alpha_3$ and mc_{W_k} cannot be approximated within $\frac{2k-2}{2k-3} \cdot \alpha_3 + \varepsilon$, for any $\varepsilon > 0$.*

Proof: (1) From Lemma 2.5 we see that $s(K_2, C_{2k+1}) = \frac{2k}{2k+1}$ which implies (using Lemma 3.4) that $FJ_2(C_k) \geq \frac{2k}{2k+1} \cdot \alpha_{GW}$. Furthermore, mc_2 cannot be approximated within $\alpha_{GW} + \varepsilon'$ for any $\varepsilon' > 0$. From the second part of Lemma 3.4, we get that $mc_{C_{2k+1}}$ is not approximable within $\frac{2k+1}{2k} \cdot (\alpha_{GW} + \varepsilon')$ for any ε'. With $\varepsilon' = \varepsilon \cdot \frac{2k}{2k+1}$ the result follows.

(2) Lai and Liu (2000) show that if H is a graph with the stated properties, then there exists a homomorphism from H to C_{2k+1}. Thus, $K_2 \to H \to C_{2k+1}$ which implies that $s(K_2, H) \geq s(K_2, C_{2k+1}) = \frac{2k}{2k+1}$. By Lemma 3.4, $FJ_2(H) \geq \frac{2k}{2k+1} \cdot \alpha_{GW}$, but mc_H cannot be approximated within $\frac{2k+1}{2k} \cdot \alpha_{GW} + \varepsilon$ for any $\varepsilon > 0$.

(3) Borodin et al. (2004) show that there exists a homomorphism from H to C_{2k+1}. The result follows as for case (2).

(4) We know from Example 3 that $K_3 \to W_k$ and $s(K_3, W_k) = \frac{2k-3}{2k-2}$. The result again follows by Lemma 3.4. □

We can compare the results of Proposition 3.11 to the performance of Håstad's algorithm as follows: Let $n = n(H)$. In (1), we have $e(H) = n$; for (2), Dutton and Brigham (1991) have given an upper bound on $e(H)$ of asymptotic order $n^{1+2/(g-1)}$; in (3), $e(H) \leq 3n - 6$, since H is planar; and finally in (4), we have $e(H) = 2(n-1)$. Now note that by ignoring lower-order terms in the expression for $\mathit{Hå}(H)$ in Theorem 3.7, we get $\mathit{Hå}(H) = \frac{2e(H)}{n^2}(1 + o(1))$. Hence, for each case (1)–(4), $\mathit{Hå}(H) \to 0$ as $n \to \infty$.

Proposition 3.11(3) can be strengthened and extended in several ways: For K_4-minor-free graphs, Pan and Zhu (2002) have given odd girth-restrictions for $2k+1$-colourability which is better than the girth-restriction in Proposition 3.11(3). Dvořák et al. (2008) have proved that every planar graph H of odd girth at least 9 is homomorphic to the Petersen graph, P. The Petersen graph is edge-transitive and its bipartite density is known to be $4/5$ (cf. Berman and Zhang (2003).) In other words, $s(K_2, P) = 4/5$. Consequently, mc_H can be approximated within $\frac{4}{5} \cdot \alpha_{GW}$ but not within $\frac{5}{4} \cdot \alpha_{GW} + \varepsilon$ for any $\varepsilon > 0$. This is an improvement on the bounds in Proposition 3.11(3) for planar graphs with girth strictly less than 13. We can also consider graphs embeddable on higher-genus surfaces. For instance, Proposition 3.11(3) is true for graphs embeddable on the projective plane, and it is also true for graphs of girth *strictly* greater than $\frac{20k-2}{3}$ whenever the graphs are embeddable on the torus or Klein bottle. These bounds are direct consequences of results in Borodin et al. (2004).

3.4 Random Graphs

Finally, we look at random graphs. Let $\mathcal{G}(n, p)$ denote the random graph on n vertices in which every edge is chosen uniformly at random, and independently, with probability $p = p(n)$. We say that $\mathcal{G}(n, p)$

has a property A *asymptotically almost surely* (a.a.s.) if the probability that it satisfies A tends to 1 as n tends to infinity. Here, we let $0 < p < 1$ be a fixed constant.

Proposition 3.12 *Let $H \in \mathcal{G}(n,p)$. Then, a.a.s.,*

$$FJ_k(H) = 1 - \frac{\ln(1/p)}{2\ln n}\big(1 + o(1)\big).$$

Proof: For $H \in \mathcal{G}(n,p)$ it is well known that a.a.s. $\omega(H)$ assumes one of at most two values around $\frac{2\ln n}{\ln(1/p)} + \Theta(\ln\ln n)$ (Matula, 1972; Bollobás and Erdős, 1976). Let $r = \omega(H)$. By Theorem 3.10,

$$FJ_k(H) = 1 - \frac{1}{r} + \frac{2\ln k}{k^2}\Big(1 - \frac{1}{r} + o(1)\Big) = 1 - \frac{\ln(1/p)}{2\ln n}\big(1 + o(1)\big).$$

$\square$

For a comparison to Håstad's algorithm, note that $e(H) = \frac{n^2 p}{2}(1 + o(1))$ a.a.s. for $H \in \mathcal{G}(n,p)$, so

$$H\mathring{a}(H) = \frac{2e(H)}{n^2}\big(1 + o(1)\big) = p + o(1).$$

The slow logarithmic growth of the clique number of $\mathcal{G}(n,p)$ works against our method in this case. However, we still manage to achieve an approximation ratio tending to 1 unlike Håstad's algorithm which ultimately is restricted by the density of the edges.

We conclude this section by looking at what happens for graphs $H \in \mathcal{G}(n,p)$ when p is no longer chosen to be a constant, but instead we let np tend to a constant $\varepsilon < 1$ as $n \to \infty$. The following theorem allows us to do this.

Theorem 3.13 (Erdős and Rényi, 1960) *Let c be a positive constant and $p = \frac{c}{n}$. If $c < 1$, then a.a.s. no component in $\mathcal{G}(n,p)$ contains more than one cycle, and no component has more than $\frac{\ln n}{c - 1 - \ln c}$ vertices.*

Now we see that if $np \to \varepsilon$ when $n \to \infty$, then $\mathcal{G}(n,p)$ almost surely consists of components with at most one cycle. Thus, each component resembles a cycle where, possibly, trees are attached to certain vertices of the cycle, and each component is homomorphically equivalent to the cycle it contains. Proposition 2.5 is therefore applicable in this part of the $\mathcal{G}(n,p)$ spectrum.

4 Circular Complete Graphs

The successful application of our method relies on the ability to compute $s(M,N)$ for various graphs M and N. In Section 2.3 we saw how this can be accomplished by the means of linear programming. This insight is put to use here in the context of *circular complete graphs*. We have already come across examples of such graphs in the form of (ordinary) complete graphs, cycles, and the graph $K_{8/3}$ in Figure 3. We will now take a closer look at them.

Definition 4.1 *Let p and q be positive integers such that $p \geq 2q$. The* circular complete graph*, $K_{p/q}$, has vertex set $\{v_0, v_1, \ldots, v_{p-1}\}$ and edge set $\{\{v_i, v_j\} \mid q \leq |i - j| \leq p - q\}$.*

The image to keep in mind is that of the vertices placed uniformly around a circle with an edge connecting two vertices if they are at a distance at least q from each other.

Example 5. Some well-known graphs are extreme cases of circular complete graphs:

- The complete graph $K_n, n \geq 2$ is a circular complete graph with $p = n$ and $q = 1$.

- The cycle graph $C_{2k+1}, k \geq 1$ is a circular complete graph with $p = 2k + 1$ and $q = k$.

These are the only examples of edge-transitive circular complete graphs.

A fundamental property of the circular complete graphs is given by the following theorem.

Theorem 4.2 (see Hell and Nešetřil, 2004) *For positive integers p, q, p′, and q′,*

$$K_{p/q} \to K_{p'/q'} \Leftrightarrow p/q \leq p'/q'$$

Due to this theorem, we may assume that whenever we write $K_{p/q}$, the positive integers p and q are relatively prime.

One of the main reasons for studying circular complete graphs is that they refine the notion of complete graphs. In particular, that they refine the notion of the *chromatic number* $\chi(G)$. Note that an alternative definition of $\chi(G)$ is given by $\chi(G) = \inf\{n \mid G \to K_n\}$. With this in mind, the following is a natural extension of proper graph colouring, and the chromatic number.

Definition 4.3 *The* circular chromatic number, $\chi_c(G)$, *of a graph G is defined as* $\inf\{p/q \mid G \to K_{p/q}\}$. *A homomorphism from G to* $K_{p/q}$ *is called a* (circular) p/q-colouring *of G.*

For more on the circular complete graphs and the circular chromatic number, see the book by Hell and Nešetřil (2004), and the survey by Zhu (2001).

We will investigate the separation parameter $s(K_r, K_t)$ for rational numbers $2 \leq r < t \leq 3$. In Section 4.1, we fix $r = 2$ and choose t so that Aut*(K_t) has few orbits. We find some interesting properties of these numbers which lead us to look at certain "constant regions" in Section 4.2, and at the case $r = 2 + 1/k$, in Section 4.3. Our method is based on solving a relaxation of the linear program (9) which was presented in Section 2.3, combined with arguments that the chosen relaxation in fact finds the optimum in the original program. Most of the calculations, which involve some rather lengthy ad hoc constructions of solutions, are left out. The complete proofs can be found in the technical report Engström et al. (2009a).

4.1 Maps to an Edge

We consider $s(K_2, K_t)$ for $t = 2 + n/k$ with $k > n \geq 1$, where n and k are integers. The number of orbits of Aut*(K_t) then equals $\lceil (n+1)/2 \rceil$. We will denote these orbits by

$$A_c = \{\{v_i, v_j\} \in E(K_{p/q}) \mid j - i \equiv q + c - 1 \pmod{p}\},$$

for $c = 1, \ldots, \lceil (n+1)/2 \rceil$. Since the number of orbits determine the number of variables of the linear program (9), we choose to begin our study of $s(K_2, K_t)$ using small values of n. For $n = 1$ we have seen that the graph $K_{2+1/k}$ is isomorphic to the cycle C_{2k+1}. For $n = 2$ we can assume that k is odd in order to have $2k + n$ and k relatively prime. We will write this number as $t = 2 + 2/(2k-1) = \frac{4k}{2k-1}$.

Proposition 4.4 *Let* $k \geq 1$ *be an integer. Then,* $s(K_2, K_{\frac{4k}{2k-1}}) = \frac{2k}{2k+1}$.

Proof: Let $V(K_{\frac{4k}{2k-1}}) = \{v_0, v_1, \ldots, v_{4k-1}\}$ and $V(K_2) = \{w_0, w_1\}$. We start by presenting two maps $f, g : V(K_{\frac{4k}{2k-1}}) \to V(K_2)$. The map f sends v_i to w_0 if i is even and to w_1 if i is odd. Then, f maps all of A_1 to K_2 but none of the edges in A_2, so $f = (4k, 0)$. The solution g sends a vertex v_i to w_0 if $0 \leq i < 2k$ and to w_1 if $2k \leq i < 4k$. It is not hard to see that $g = (4k-2, 2k)$. It remains to argue that these two solutions suffice to determine s. But we see that any map $h = (h_1, h_2)$ with $h_2 > 0$ must cut at least two edges in the even cycle A_1. Therefore, $h_1 \leq 4k-2$, so $h \leq g$, componentwise. The proposition now follows by solving the relaxation of (9) using only the two inequalities obtained from f and g. □

Note that the graph $K_{8/3}$ from Example 4 is covered by Proposition 4.4. The argument in that example is very similar to the proof of the general case.

For $n = 3$, $t = 2 + 3/k$, we see that if $k \equiv 0 \pmod 3$, then K_t is an odd cycle. If $k \equiv 2 \pmod 3$, then we can let $k = 3k' - 1$ and observe that $2 + 1/k' \leq t \leq 2 + 2/(2k' - 1)$. Hence, by Theorem 4.2, Lemma 2.6, and known values for s, we have

$$\frac{2k'}{2k'+1} = s(K_2, C_{2k'+1}) \geq s(K_2, K_t) \geq s(K_2, K_{\frac{4k'}{2k'-1}}) = \frac{2k'}{2k'+1}.$$

It follows that $s(K_2, K_t) = 2k'/(2k'+1) = (2k+2)/(2k+5)$ as well. Therefore we assume that t is of the form $2 + 3/(3k+1) = \frac{6k+5}{3k+1}$ for an integer $k \geq 1$.

Proposition 4.5 *Let $k \geq 1$ be an integer. Then,*

$$s(K_2, K_{\frac{6k+5}{3k+1}}) = \frac{6k^2+8k+3}{6k^2+11k+5} = 1 - \frac{3k+2}{(k+1)(6k+5)}.$$

For $t = 2 + 4/k$, we find that we only need to consider the case when $k \equiv 1 \pmod 4$. We then have graphs K_t with $t = 2 + 4/(4k+1) = \frac{8k+6}{4k+1}$ for integers $k \geq 1$.

Proposition 4.6 *Let $k \geq 1$ be an integer. Then,*

$$s(K_2, K_{\frac{8k+6}{4k+1}}) = \frac{8k^2+6k+2}{8k^2+10k+3} = 1 - \frac{4k+1}{(k+1/2)(8k+6)}.$$

The expressions for s in Propositions 4.5 and 4.6 have some interesting regularities, but for $n \geq 5$ it becomes much harder to choose a suitable set of solutions. Using brute force computer calculations, we have determined the first two values (for $k = 1, 2$) in each of the cases $t = 2 + 5/(5k+1)$ ($t = 17/6, 27/11$) and $t = 2 + 6/(6k+1)$ ($t = 20/7, 32/13$). These values are summarised in Table 2.

$s(K_2, K_t)$	$t = 2 + 5/(5k+1)$	$t = 2 + 6/(6k+1)$
$k = 1$	$322/425 \approx 0.7576$	$67/89 \approx 0.7528$
$k = 2$	$5/6 \approx 0.8333$	$94/113 \approx 0.8319$

Tab. 2: Some parameter values determined by brute force computer calculations

4.2 Constant Regions

In the previous section we saw that $s(K_2, C_{2k+1}) = s(K_2, K_{\frac{4k}{2k-1}})$ and used it to prove that $s(K_2, K_t)$ is constant in the interval $t \in [2 + 1/k, 2 + 2/(2k-1)]$. This is a special case of a phenomenon described more generally in the following proposition.

Proposition 4.7 *Let $k \geq 1$, and let r and t be rational numbers such that $2 \leq r < \frac{2k+1}{k} \leq t \leq \frac{4k}{2k-1}$. Then,*

$$s(K_r, K_t) = \frac{2k}{2k+1}.$$

Proof: From Theorem 4.2, we have the following chain of homomorphisms.

$$K_2 \to K_r \to K_{\frac{2k+1}{k}} \to K_t \to K_{\frac{4k}{2k-1}}.$$

By Lemma 2.6, this implies

$$s(K_r, K_{\frac{2k+1}{k}}) \geq s(K_2, K_{\frac{2k+1}{k}}) = \frac{2k}{2k+1},$$

but since $K_{\frac{2k+1}{k}} \not\to K_r$, and $K_{\frac{2k+1}{k}}$ is edge-transitive with $2k+1$ edges, $s(K_r, K_{\frac{2k+1}{k}}) \leq \frac{2k}{2k+1}$ and therefore $s(K_r, K_{\frac{2k+1}{k}}) = \frac{2k}{2k+1}$. Two more applications of Lemma 2.6 show that

$$\frac{2k}{2k+1} = s(K_r, K_{\frac{2k+1}{k}}) \geq s(K_r, K_t) \geq s(K_2, K_{\frac{4k}{2k-1}}) = \frac{2k}{2k+1},$$

which proves the proposition. □

We find that there are intervals $I_k = \{t \in \mathbb{Q} \mid 2 + 1/k \leq t \leq 2 + 2/(2k-1)\}$ where the function $s_r(t) = s(K_r, K_t)$ is constant for any $2 \leq r < (2k+1)/k$. In Figure 4 these intervals are shown for the first few values of k. The intervals I_k form an infinite sequence with endpoints tending to 2.

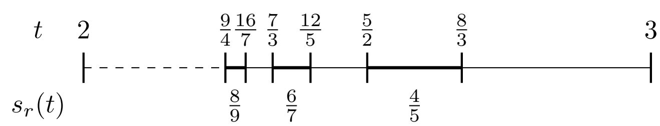

Fig. 4: The intervals I_k marked for $k = 2, 3, 4$.

It turns out that similar intervals appear throughout the space of circular complete graphs. Indeed, it follows from Proposition 2.4 that for a positive integer n and a rational number r such that $2 \leq r \leq n$, we have

$$s(K_r, K_n) = \left\lfloor \left(1 - \frac{1}{\lfloor r \rfloor}\right) \cdot \frac{n^2}{2} \right\rfloor \Big/ \binom{n}{2}. \tag{10}$$

From (10) we see that $s(K_r, K_n)$ remains constant for rational numbers r in the interval $k \leq r < k+1$, where k is any fixed integer $k < n$. Furthermore, for positive integers k and m, we have

$$\begin{aligned} e(T(km-1,k)) &= \left\lfloor \left(1 - \frac{1}{k}\right) \cdot \frac{(km-1)^2}{2} \right\rfloor \\ &= \left\lfloor \frac{(k-1)km^2}{2} - (k-1)m + \frac{k-1}{2k} \right\rfloor = \frac{(k-1)km^2}{2} - (k-1)m \\ &= \binom{k}{2} m^2 \cdot \left(1 - \frac{2}{km}\right) = e(T(km,k)) \cdot \binom{km-1}{2} \Big/ \binom{km}{2}. \end{aligned}$$

Thus, $s(K_k, K_{km-1}) = s(K_k, K_{km})$. When we combine this fact with (10) and Lemma 2.6, we find that $s(K_r, K_t)$ is constant on each region $(r, t) \in [k, k+1) \times [km-1, km]$.

4.3 Maps to Odd Cycles

It was seen in Proposition 4.7 that $s(K_r, K_t)$ is constant on the region $(r,t) \in [2, 2+1/k) \times I_k$. In this section, we will study what happens when t remains in I_k, but r assumes the value $2 + 1/k$. A first observation is that the absolute jump of the function $s(K_r, K_t)$ when r goes from being less than $2 + 1/k$ to $r = 2 + 1/k$ must be largest for $t = 2 + 2/(2k-1)$. Let $V(K_{2+2/(2k-1)}) = \{v_0, \ldots, v_{4k-1}\}$ and $V(K_{2+1/k}) = \{w_0, \ldots, w_{2k}\}$, and let the function f map v_i to w_j, with $j = \lfloor \frac{2k+1}{4k} \cdot i \rfloor$. Then, f maps all edges except $\{v_0, v_{2k-1}\}$ from the orbit A_1 to some edge in K_r. Since the subgraph A_1 is isomorphic to C_{4k}, any map to an odd cycle must exclude at least one edge from A_1. It follows that f alone determines s, and we can solve the linear program (9) to obtain $s(K_{2+1/k}, K_{2+2/(2k-1)}) = (4k-1)/4k$. Thus, for $r < 2 + 1/k$, we have

$$s(K_{2+1/k}, K_{2+2/(2k-1)}) - s(K_r, K_{2+2/(2k-1)}) = \frac{2k-1}{4k(2k+1)}.$$

Smaller $t \in I_k$ can be expressed as $t = 2 + 1/(k-x)$, where $0 \le x < 1/2$. We will write $x = m/n$ for positive integers m and n which implies the form $t = 2 + n/(kn-m)$, with $m < n/2$. For $m = 1$, it turns out to be sufficient to keep two inequalities from (9) to get an optimal value of s. From this we get the following result.

Proposition 4.8 *Let $k, n \ge 2$ be integers. Then,*

$$s(C_{2k+1}, K_{\frac{2(kn-1)+n}{kn-1}}) = \frac{(2(kn-1)+n)(4k-1)}{(2(kn-1)+n)(4k-1)+4k-2}.$$

There is still a non-zero jump of $s(K_r, K_t)$ when we move from $K_r < 2 + 1/k$ to $K_r = 2 + 1/k$, but it is smaller, and tends to 0 as n increases. For $m = 2$, we have $2(kn-m)+n$ and $kn-m$ relatively prime only when n is odd. In this case, it turns out that we need to include an increasing number of inequalities to obtain a good relaxation. Furthermore, we are not able to ensure that the obtained value is the optimum of the original (9). We will therefore have to settle for a lower bound on s. Brute force calculations have shown that, for small values of k and n, equality holds in Proposition 4.9. We conjecture this to be true in general.

Proposition 4.9 *Let $k \ge 2$ be an integer and $n \ge 3$ be an odd integer. Then,*

$$s(C_{2k+1}, K_{\frac{2(kn-2)+n}{kn-2}}) \ge \frac{(2(kn-2)+n)(\xi_n(4k-1)+(2k-1))}{(2(kn-2)+n)(\xi_n(4k-1)+(2k-1))+(4k-2)(1-\xi_n)},$$

where $\xi_n = \frac{1}{4}\left(z_1^{\frac{n-1}{2}} + z_2^{\frac{n-1}{2}}\right)$, and z_1^{-1}, z_2^{-1} are the roots of $\frac{2k-3}{4k-2}z^2 - 2z + 1$. Asymptotically, for a fixed k, one has $\xi_n \sim A_k \left(1 + \sqrt{\frac{2k+1}{4k-2}}\right)^{n/2}$, where A_k is a constant that does not depend on n.

5 Fractional Covering by H-cuts

In the following, we generalise the work of Šámal (2005, 2006, 2012) on fractional covering by cuts. We obtain a complete correspondence between a family of chromatic numbers, $\chi_H(G)$, and $s(H, G)$. These chromatic numbers are generalisations of Šámal's cubical chromatic number $\chi_q(G)$; the latter corresponds to the case when $H = K_2$. Two more expressions for $\chi_H(G)$ are given in Section 5.2. We believe that these alternative views on the separation parameter can provide great benefits to the understanding of its properties. We transfer a result in the other direction, in Section 5.3, disproving a conjecture by Šámal on χ_q, and settle another conjecture by him in the positive, in Section 5.4.

5.1 Separation as a Chromatic Number

We start by recalling the notion of a *fractional colouring* of a hypergraph. Let G be a (hyper-) graph with vertex set $V(G)$ and edge set $E(G) \subseteq 2^{V(G)}$. A subset J of $V(G)$ is called *independent* in G if no edge $e \in E(G)$ is a subset of J. Let $\mathcal{J}$ denote the set of all independent sets of G and for a vertex $v \in V(G)$, let $\mathcal{J}(v)$ denote all independent sets which contain v. Let $J_1, \dots, J_n \in \mathcal{J}$ be a collection of independent sets.

Definition 5.1 *An n/k* independent set cover *is a collection $J_1, \dots, J_n$ of independent sets in $\mathcal{J}$ such that every vertex of G is in at least k of them.*

The fractional chromatic number $\chi_f(G)$ *of G is given by the following expression.*

$$\chi_f(G) = \inf\{\frac{n}{k} \mid \textit{there exists an } n/k \textit{ independent set cover of } G\}.$$

The definition of fractional covering by cuts mimics that of fractional covering by independent sets, but replaces vertices with edges and independent sets with certain cut sets of the edges. Let G and H be undirected simple graphs and f be an arbitrary vertex map from G to H. Recall that the map f induces a partial edge map $f^{\#} : E(G) \to E(H)$. We will call the preimage of $E(H)$ under $f^{\#}$ an H*-cut* in G. When H is a complete graph K_k, this is precisely the standard notion of a k-cut in G. Let $\mathcal{C}$ denote the set of H-cuts in G and for an edge $e \in E(G)$, let $\mathcal{C}(e)$ denote all H-cuts which contain e. The following definition is a generalisation of *cut n/k-covers* (Šámal, 2006) to arbitrary H-cuts.

Definition 5.2 *An* H-cut n/k-cover of G *is a collection $C_1, \dots, C_N \in \mathcal{C}$ such that every edge of G is in at least k of them.*

The graph parameter χ_H is defined as:

$$\chi_H(G) = \inf\{\frac{n}{k} \mid \textit{there exists an } H\textit{-cut } n/k\textit{-cover of } G\}.$$

Šámal (2006) called the parameter $\chi_{K_2}(G)$, the *cubical chromatic number of G*. Both the fractional chromatic number and the cubical chromatic number also have linear programming formulations. This, in particular, shows that the value in the infimum of the corresponding definition is obtained exactly for some n and k. For our generalisation of the cubical chromatic number, the linear program is the following:

$$\begin{array}{lll} \text{Minimise} & \sum_{C \in \mathcal{C}} f(C) & \\ \text{subject to} & \sum_{C \in \mathcal{C}(e)} f(C) \geq 1 & \text{for all } e \in E(G)\,, \\ & f : \mathcal{C} \to \mathbb{Q}_{\geq 0}. & \end{array} \tag{11}$$

Proposition 5.3 *The graph parameter $\chi_H(G)$ is given by the optimum of the linear program in (11).*

Proof: The proof is completely analogous to those for the corresponding statements for the fractional chromatic number (cf. Godsil and Royle, 2001) and for the cubical chromatic number (Lemma 5.1.3 in Šámal, 2006). Let $C_1, \dots, C_n$ be an H-cut n/k-cover of G. The solution $f(C) = 1/k$ if $C \in \{C_1, \dots, C_n\}$, and $f(C) = 0$ otherwise, has a measure of n/k in (11). Thus, the optimum of the linear program is at most $\chi_H(G)$.

For the other direction, note that the coefficients of the program (11) are integral. Hence, there is a rational optimal solution f^*. Let N be the least common multiple of the divisors of $f^*(C)$ for $C \in \mathcal{C}$. Assume that the measure of f^* is n/k. Construct a collection of H-cuts by including the cut C a total of $N \cdot f^*(C)$ times. This collection covers each edge at least N times using $\sum_{C \in \mathcal{C}} N \cdot f^*(C) = N \cdot n/k$ cuts, i.e. it is an H-cut n/k-cover, so $\chi_H(G)$ is at most equal to the optimum of (11). □

We are now ready to work out the correspondence to separation.

Proposition 5.4 *The identity $\chi_H(G) = 1/s(H,G)$ holds for all $G, H \in \mathcal{G}$.*

Proof: Consider the dual program of (11).

$$\begin{array}{lll} \text{Maximise} & \displaystyle\sum_{e \in E(G)} g(e) & \\ \text{subject to} & \displaystyle\sum_{e \in X} g(e) \leq 1 & \text{for all } H\text{-cuts } X \in \mathcal{C}\,, \\ & g : E(G) \to \mathbb{Q}_{\geq 0}. & \end{array} \tag{12}$$

In (12) let $1/s = \sum_{e \in E(G)} g(e)$ and make the variable substitution $w = g \cdot s$. This leaves the following program.

$$\begin{array}{lll} \text{Maximise} & s^{-1} & \\ \text{subject to} & \displaystyle\sum_{e \in X} w(e) \leq s & \text{for all } H\text{-cuts } X \in \mathcal{C}\,, \\ & \displaystyle\sum_{e \in E(G)} w(e) = 1 & \\ & w : E(G) \to \mathbb{Q}_{\geq 0}. & \end{array} \tag{13}$$

Since $\max s^{-1} = (\min s)^{-1}$, a comparison with (8) establishes the proposition. □

5.2 More Guises of Separation

For fractional colourings, it is well-known that an equivalent definition is obtained by taking $\chi_f(G) = \inf\{n/k \mid G \to K_{n,k}\}$, where $K_{n,k}$ denotes the *Kneser graph*, the vertex set of which is the k-subsets of $[n]$ and with an edge between u and v if $u \cap v = \varnothing$. For $H = K_2$, a corresponding definition of $\chi_H(G) = \chi_q(G)$ was obtained in Šámal (2006) by taking the infimum over n/k for n and k such that $G \to Q_{n/k}$. Here, $Q_{n/k}$ is the graph on vertex set $\{0,1\}^n$ with an edge between u and v if $d_H(u,v) \geq k$, where d_H denotes the Hamming distance.

A parameterised graph family which determines a particular chromatic number in this way is sometimes referred to as a *scale*. In addition to the previously mentioned fractional chromatic number χ_f, where the scale is the set of Kneser graphs, and the cubical chromatic number χ_q, where the scale is $\{Q_{n/k}\}$, another prominent example is the circular chromatic number (Section 4) for which the scale is given by the family of circular complete graphs $K_{n/k}$.

We now generalise the family $\{Q_{n/k}\}$ to produce one scale for each χ_H. To this end, let H^n_k be the graph on vertex set $V(H)^n$ and an edge between $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ when $|\{i \mid \{u_i, v_i\} \in E(H)\}| \geq k$. The proof of the following proposition is straightforward, but instructive.

Proposition 5.5 *For $G, H \in \mathcal{G}$, we have*

$$\chi_H(G) = \inf\{\frac{n}{k} \mid G \to H_k^n\}. \tag{14}$$

Proof: Both sides are defined by infima over n/k. Therefore, it will suffice to show how to translate each of the parameterised objects (an H-cut n/k-cover on the left-hand side and a homomorphism from G to H_k^n on the right-hand side) into the other, for some given values of n and k.

Let $h : V(G) \to H_k^n$ be a homomorphism and denote by $\mathrm{pr}_i\, h$ the projection of h onto the ith coordinate. Then, for each edge $e \in E(G)$, at least k of the edge maps $(\mathrm{pr}_i\, h)^\#$ must map e to an edge in H. Hence, the H-cuts $C_i = \{e \in E(G) \mid (\mathrm{pr}_i\, h)^\#(e) \in E(H)\}$, for $1 \leq i \leq n$, constitute an H-cut n/k-cover of G.

For the other direction, note that each H-cut C_i can be defined by a vertex map $f_i : V(G) \to H$. For $v \in V(G)$, let $h'(v) = (f_1(v), f_2(v), \ldots, f_n(v))$. To verify that h' is a homomorphism, note that for every edge $\{u, v\} \in E(G)$, at least k of the f_i must include e in their corresponding cut C_i. Hence $|\{i \mid \{f_i(u), f_i(v)\} \in E(H)\}| \geq k$, so by definition $\{h'(u), h'(v)\} \in E(H_k^n)$. □

Šámal further notes that $\chi_q(G)$ is given by the fractional chromatic number of a certain hypergraph associated to G. Inspired by this, we provide a similar formulation in the general case.

Proposition 5.6 *Let G' be the hypergraph obtained from G on vertex set $V(G') = E(G)$ with edge set $E(G')$ taken to be the set of minimal subgraphs $K \subseteq G$ such that $K \not\to H$. Then,*

$$\chi_H(G) = \chi_f(G').$$

Proof: We will let $\mathcal{J}$ denote the set of independent sets in G' and $\mathcal{C}$ the set of H-cuts in G. The parameter $\chi_H(G)$ is the infimum of n/k over all n/k-covers of $E(G)$ by sets in $\mathcal{C}$. Similarly, the parameter $\chi_f(G')$ is the infimum of n/k over all n/k-covers of $V(G') = E(G)$ by sets in $\mathcal{J}$. By definition, the independent sets of G' correspond precisely to the edge sets $E(K)$ of those subgraphs $K \subseteq G$ such that $K \to H$. Hence, $\mathcal{C} \subseteq \mathcal{J}$, so $\chi_H(G) \geq \chi_f(G')$.

On the other hand, assume that $K \subseteq G$ is a subgraph such that $K \to H$, and that the independent set $E(K) \in \mathcal{J}$ is not in $\mathcal{C}$. Then, any homomorphism $h : V(K) \to V(H)$ induces an H-cut C of G, and clearly we must have $E(K) \subseteq C$. Thus, we can replace all edge sets in a cover by sets from $\mathcal{C}$, without violating the constraint that all edges are covered at least k times. The proposition now follows. □

5.3 An Upper Bound

In Section 4 we obtained lower bounds on s by relaxing the linear program (9). In most cases, the corresponding solution was proved feasible in the original program, and hence optimal. Now, we take a look at the only known source of general upper bounds for s.

Let $G, H \in \mathcal{G}$, with $H \to G$ and let S be such that $S \to G$. Then, applying Lemma 2.6 followed by Theorem 2.1 gives

$$s(H, G) \leq s(H, S) = \inf_{w \in \hat{\mathcal{W}}(S)} mc_H(S, w) \leq mc_H(S, 1/|E(S)|). \tag{15}$$

We can therefore upper bound $s(H, G)$ by the least maximal H-cut taken over all subgraphs of G. For $H = K_2$, we have

$$s(K_2, G) \leq \min_{S \subseteq G} b(S),$$

where $b(S)$ denotes the bipartite density of S. Conjecture 5.5.3 in Šámal (2006) suggested that this inequality, expressed on the form $\chi_q(S) \geq 1/(\min_{S \subseteq G} b(S))$, could be replaced by an equality. We answer this in the negative, using $K_{11/4}$ as our counterexample. Lemma 4.5 with $k = 1$ gives $s(K_2, K_{11/4}) = 17/22$. If $s(K_2, K_{11/4}) = b(S)$ for some $S \subseteq K_{11/4}$ it means that S must have at least 22 edges. Since $K_{11/4}$ has exactly 22 edges it follows that $S = K_{11/4}$. However, a cut in a cycle must contain an even number of edges. Since the edges of $K_{11/4}$ can be partitioned into two cycles, we have that the maximum cut in $K_{11/4}$ must be of even size, hence $|E(K_{11/4})| \cdot b(K_{11/4}) \neq 17$. This is a contradiction.

5.4 Confirmation of a Scale

As a part of his investigation of the cubical chromatic number, Šámal (2006) set out to determine the value of $\chi_q(Q_{n/k})$ for general n and k. For the fractional chromatic number and the circular chromatic number, results for such measuring of the scale exist and provide very appealing formulae: $\chi_f(K_{n,k}) = \chi_c(C_{n/k}) = n/k$. For $\chi_q(Q_{n/k}) = 1/s(K_2, Q_{n/k})$, we are immediately out of luck as $1/2 < s(K_2, G) \leq 1$, i.e. $1 \leq \chi_q(G) < 2$ for all non-empty graphs. For $1 \leq n/k < 2$, however, Šámal gave a conjecture (Conjecture 5.4.2 in Šámal, 2006). We complete the proof of his conjecture to obtain the following result.

Proposition 5.7 *Let k, n be integers such that $k \leq n < 2k$. Then,*

$$\chi_q(Q_{n/k}) = \begin{cases} n/k & \text{if } k \text{ is even, and} \\ (n+1)/(k+1) & \text{if } k \text{ is odd.} \end{cases}$$

Corollary 5.8 *Let k, n be integers such that $1/2 < k/n \leq 1$. Then, $s(K_2, Q_{n/k}) = k/n$ if k is even, and $(k+1)/(n+1)$ if k is odd.*

If we make sure that k is even, possibly by multiplying both k and n by a factor two, we get the following interesting corollary.

Corollary 5.9 *For every rational number r, $1/2 < r \leq 1$, there is a graph G such that $s(K_2, G) = r$.*

Šámal (2006) provides the upper bound for Proposition 5.7 and an approach to the lower bound of using the largest eigenvalue of the Laplacian of a subgraph of $Q_{n/k}$. The computation of this eigenvalue boils down to an inequality (Conjecture 5.4.6 in Šámal, 2006) involving some binomial coefficients. We first introduce the necessary notation and then prove the remaining inequality in Proposition 5.11, whose second part, for odd k, corresponds to one of the formulations of the conjecture. Proposition 5.7 then follows from Theorem 5.4.7 in Šámal (2006), conditioned on the result of this proposition.

Let k, n be positive integers such that $k \leq n$, and let x be an integer such that $1 \leq x \leq n$. For $k \leq n < 2k$, let $S_o(n, k, x)$ denote the set of all k-subsets of $[n]$ that have an odd-sized intersection with $[n] \setminus [n - x]$. Define $S_e(n, k, x)$ analogously as the k-subsets of $[n]$ that have an even-sized intersection with $[n] \setminus [n-x]$, i.e. $S_e(n, k, x) = \binom{[n]}{k} \setminus S_o(n, k, x)$. Let $N_o(n, k, x) = |S_o(n, k, x)|$ and $N_e(n, k, x) = |S_e(n, k, x)|$. Then,

$$N_o(n, k, x) = \sum_{odd\, t} \binom{x}{t}\binom{n-x}{k-t} \text{ and } N_e(n, k, x) = \sum_{even\, t} \binom{x}{t}\binom{n-x}{k-t}.$$

When x is odd, the function $f : S_o(2k, k, x) \to S_e(2k, k, x)$, given by the complement $f(\sigma) = [n] \setminus \sigma$, is a bijection. Since $N_o(n, k, x) + N_e(n, k, x) = \binom{n}{k}$, we have

$$N_o(2k, k, x) = N_e(2k, k, x) = \frac{1}{2}\binom{2k}{k}. \tag{16}$$

Lemma 5.10 *Assume that x is odd, with $1 \leq x < n = 2k - 1$. Then, $N_e(n, k, x) = N_e(n, k, x + 1)$ and $N_o(n, k, x) = N_o(n, k, x + 1)$.*

Proof: First, partition $S_e(n, k, x)$ into $A_1 = \{\sigma \in S_e(n, k, x) \mid n - x \notin \sigma\}$ and $A_2 = S_e(n, k, x) \setminus A_1$. Similarly, partition $S_e(n, k, x+1)$ into $B_1 = \{\sigma \in S_e(n, k, x+1) \mid n-x \notin \sigma\}$ and $B_2 = S_e(n, k, x+1) \setminus B_1$. Note that $A_1 = B_1$. We argue that $|A_2| = |B_2|$. To prove this, define the function $f : 2^{[n]} \to 2^{[n-1]}$ by

$$f(\sigma) = (\sigma \cap [n - x - 1]) \cup \{s - 1 \mid s \in \sigma, s > n - x\}.$$

That is, f acts on σ by ignoring the element $n-x$ and renumbering subsequent elements so that the image is a subset of $[n - 1]$. Note that $f(A_2) = S_e(2k - 2, k - 1, x)$ and $f(B_2) = S_o(2k - 2, k - 1, x)$. Since x is odd, it follows from (16) that $|f(A_2)| = |f(B_2)|$. The first part of the lemma now follows from the injectivity of the restrictions $f|_{A_2}$ and $f|_{B_2}$. The second equality is proved similarly. □

Proposition 5.11 *Choose k, n and x so that $k \leq n < 2k$ and $1 \leq x \leq n$. Then,*

$$N_e(n, k, x) \leq \binom{n-1}{k-1} \textit{ for odd } k, \textit{ and } N_o(n, k, x) \leq \binom{n-1}{k-1} \textit{ for even } k.$$

Proof: We will proceed by induction over n and x. The base cases are given by $x = 1$, $x = n$, and $n = k$. For $x = 1$,

$$N_o(n, k, x) = \binom{n-1}{k-1} \text{ and } N_e(n, k, x) = \binom{n-1}{k} \leq \binom{n-1}{k-1},$$

where the inequality holds for all $n < 2k$. For $x = n$ and odd k, we have $N_e(n, k, x) = 0$, and for even k, we have $N_o(n, k, x) = 0$. For $n = k$,

$$N_e(n, k, x) = 1 - N_o(n, k, x) = \begin{cases} 1 & \text{if } x \text{ is even,} \\ 0 & \text{otherwise.} \end{cases}$$

Let $x > 1$ and consider $N_e(n, k, x)$ for odd k and $k < n < 2k - 1$. Partition the sets $\sigma \in S_e(n, k, x)$ into those for which $n \in \sigma$ on the one hand and those for which $n \notin \sigma$ on the other hand. These parts contain $N_o(n - 1, k - 1, x - 1)$ and $N_e(n - 1, k, x - 1)$ sets, respectively. Since $k - 1$ is even, and since $k \leq n - 1 < 2(k - 1)$ when $k < n < 2k - 1$, it follows from the induction hypothesis that

$$\begin{aligned} N_e(n, k, x) &= N_o(n - 1, k - 1, x - 1) + N_e(n - 1, k, x - 1) \\ &\leq \binom{n-2}{k-2} + \binom{n-2}{k-1} = \binom{n-1}{k-1}. \end{aligned}$$

The case for $N_o(n, k, x)$ and even k is treated identically.

Finally, let $n = 2k - 1$. If x is odd, then Lemma 5.10 is applicable, so we can assume that x is even. Now, as before

$$\begin{aligned} N_e(2k-1,k,x) &= N_o(2k-2,k-1,x-1) + N_e(2k-2,k,x-1) \\ &\leq \frac{1}{2}\binom{2k-2}{k-1} + \binom{2k-3}{k-1} = \binom{n-1}{k-1}, \end{aligned}$$

where the first term is evaluated using (16). The same inequality can be shown for $N_o(2k-1,k,x)$ and even k, which completes the proof. □

6 Discussion and Open Problems

What started out as a very simple idea has diverged in a number of directions, with plenty of room in each for further investigation and improvements. We single out two main topics, and discuss their respective future prospects and interesting open problems. These two topics relate to the application of our approach to the approximability of the problem MAX H-COL (and more generally to MAX CSP(Γ)), and to the computation and interpretation of the separation parameter.

6.1 Separation and Approximability

Our initial idea of separation applied to the MAX H-COL problem lead us to a binary graph parameter that, in a sense, measures how close one graph is to be homomorphic to another. While not apparent from the original definition in (3), which involves taking an infimum over all possible instances, we have shown that the parameter can be computed effectively by means of linear programming. Given a graph H and known approximability properties for MAX H-COL, this parameter allows us to deduce bounds on the corresponding properties for MAX H'-COL for graphs H' that are "close" to H. Our approach can be characterised as local; the closer the separation of two graphs is to 1, the more precise are our bounds. We have shown that given a "base set" containing the complete graphs, our method can be used to derive good bounds on the approximability of MAX H-COL for *any* graph H.

For the applications in Section 3, we have used the complete graphs as our base set of known problems. We have shown that this set of graphs is sufficient for achieving new, non-trivial bounds on several different classes of graphs. That is, when we apply Frieze and Jerrum's algorithm (Frieze and Jerrum, 1997) to MAX H-COL, we obtain results comparable to, or better than those guaranteed by Håstad's MAX 2-CSP algorithm (Håstad, 2005), for the classes of graphs we have considered. This comparison should however be taken with a grain of salt. The analysis of Håstad's MAX 2-CSP algorithm only aims to prove it better than a random assignment, and may leave room for strengthening of the approximation guarantee. At the same time, we are overestimating the distance for most of the graphs under consideration. It is likely that both results can be improved, within their respective frameworks. When considering inapproximability, we have relied on the unique games conjecture. Weaker inapproximability results, independent of the UGC, exist for both MAX CUT (Håstad, 2001) and MAX k-CUT (Kann et al., 1997), and they are applicable in our setting. We emphasise that our method is not per se dependent on the truth of the UGC.

For the purpose of extending the applicability of our method, a possible direction to take is to find a larger base set of MAX H-COL problems. We suggest two candidates for further investigation: the circular complete graphs, for which we have obtained partial results for the parameter s in Section 4, and the Kneser graphs, see for example Hell and Nešetřil (2004). Both of these classes generalise the complete graphs, and have been subject to substantial previous research. The Kneser graphs contain many examples of graphs with low clique number, but high chromatic number. They could thus prove to be an ideal starting point for studying this phenomenon in relation to our parameter.

We conclude this part of the discussion by considering some possible extensions of our approximability results. We have already noted that MAX H-COL is a special case of the MAX CSP(Γ) problem, parameterised by a finite constraint language Γ. It should be relatively clear that we can define a generalised separation parameter on a pair of general constraint languages. This would constitute a novel method for studying the approximability of MAX CSP—a method that may cast some new light on the performance of Raghavendra's algorithm. As a way of circumventing the hardness result by Khot et al. (2007), Kaporis et al. (2006) show that mc_2 is approximable within 0.952 for any given average degree d, and asymptotically almost all random graphs G in $\mathcal{G}(n, m = \lfloor \frac{d}{2} n \rfloor)$. Here, $\mathcal{G}(n, m)$ is the probability space of random graphs on n vertices and m edges, selected uniformly at random. A different approach is taken by Coja-Oghlan et al. (2005) who give an algorithm that approximates mc_k within $1 - \mathcal{O}(1/\sqrt{np})$ in expected polynomial time, for graphs from $\mathcal{G}(n, p)$. Kim and Williams (2011) give an algorithm for finding a cut with value at least an additive constant k better than α_{GW} times the value of an optimal cut (provided such a cut exists) in a given graph, if you are willing to spend time exponential in k to do so. In a similar vein, Crowston et al. (2011) show how to use time exponential in k to find a cut better than the Edwards-Erdős bound, i.e., with value $e(G)/2 + (n(G) - 1)/4 + k$ in a given graph G or decide that no such cut exists. It would be interesting to study whether separation can be used to extend these results to improved approximability bounds on MAX H-COL.

6.2 Separation as a Graph Parameter

For a graph G with a circular chromatic number r close to 2 we can use Lemma 2.6 to bound $s(K_2, G) \geq s(K_2, K_r)$. Due to Proposition 4.4, we have also seen that with this method, we are unable to distinguish between the class of graphs with circular chromatic number $2 + 1/k$ and the (larger) class of graphs with circular chromatic number $2 + 2/(2k - 1)$. Nevertheless, the method is quite effective when applied to sequences of graph classes for which the circular chromatic number tends to 2, as was the case in Proposition 3.11(1)–(3). Much of the extensive study conducted in this direction was instigated by the restriction of a conjecture by Jaeger (1988) to planar graphs. This conjecture is equivalent to the claim that every planar graph of girth at least $4k$ has a circular chromatic number at most $2 + 1/k$, for $k \geq 1$. The case $k = 1$ is Grötzsch's theorem; that every triangle-free planar graph is 3-colourable. Currently, the best lower bound on the girth of a planar graph which implies a circular chromatic number of at most $2 + 1/k$ is $\frac{20k-2}{3}$, and is due to Borodin et al. (2004). We remark that Jaeger's conjecture implies a weaker statement in our setting. Namely, if G is a planar graph with girth greater than $4k$, then $G \to C_k$ implies $s(K_2, G) \geq s(K_2, C_k) = 2k/(2k + 1)$. Deciding this to be true would certainly provide support for the original conjecture, and would be an interesting result in its own right. Our starting observation shows that the slightly weaker condition $G \to K_{2+2/(2k-1)}$ implies the same result.

For edge-transitive graphs G, it is not surprising that the expression $s(K_r, G)$ assumes a finite number of values, as a function of r. Indeed, Theorem 2.1 states that $s(K_r, G) = mc_{K_r}(G, 1/e(G))$, which leaves at most $e(G)$ possible values for s. This produces a number of constant intervals that are partly responsible for the constant regions of Proposition 4.7, and the discussion in Section 4.2. More surprising are the constant intervals that arise from $s(K_r, K_{2+2/(2k-1)})$. They give some hope that the behaviour of the separation parameter can be characterised more generally. We propose investigating the existence of more constant regions, and possibly showing that they tile the entire space.

In Section 5 we generalised the notion of covering by cuts due to Šámal. In doing this, we found a different interpretation of the separation parameter as an entire family of chromatic numbers. It is our belief that these alternate viewpoints can benefit from each other. The refuted conjecture in Section 5.3 is an immediate example of this. It is tempting to look for a generalisation of Proposition 5.7 with K_2 replaced by an arbitrary graph H. A trivial upper bound of $s(H, H_k^n) \leq k/n$ is obtained from Proposition 5.5, but we have not identified anything corresponding to the parity criterion which appears in the case $H = K_2$. This leads us to believe that this bound can be improved upon. The approach of Šámal on the lower bound does not seem to generalise. The reason for this is that it uses bounds on maximal cuts obtained from the Laplacian of (a subgraph of) $Q_{n/k}$. We know of no such results for maximal k-cuts, with $k > 2$, much less for general H-cuts.

In recent work, Šámal and coauthors have shown that the cubical chromatic number χ_q can be approximated within α_{GW} (Šámal, 2012). This suggests the interesting possibility of a close connection between the approximability of mc_H and that of $s(H, G)$, with H fixed. Let M, N, and H be graphs. A function $g : E(M) \to E(N)$ is said to be *H-cut continuous* if, for any H-cut $C \subseteq E(N)$ in N, we have that $g^{-1}(C) \subseteq E(M)$ is an H-cut in M. For any homomorphism h, the edge map $h^{\#}$ is H-cut continuous for every H. Šámal (2005) used cut continuous maps ($H = K_2$) to show that certain non-homomorphic graphs have the same cubical chromatic number. Here we show how general H-cut continuous maps can be used to generalise the implication in Lemma 2.6.

Lemma 6.1 *Let M, N, and H be graphs in $\mathcal{G}$. If there exists an M-cut continuous map from H to N, then $s(M, H) \geq s(M, N)$.*

Proof: Let $f : E(H) \to E(N)$ be an M-cut continuous function. It suffices to show that for any graph $H \in \mathcal{G}$ and $w \in \mathcal{W}(H)$, we have

$$mc_M(H, w) \geq mc_M(N, w_N), \tag{17}$$

where $w_N(e) = \sum_{e' \in f^{-1}(e)} w(e')$. Let $g : V(N) \to V(M)$ be an optimal solution to (N, w_N). Then, $C = (g^{\#})^{-1}(E(M))$ is an M-cut in N, so $f^{-1}(C)$ is an M-cut in H. Hence, there exists a solution to (H, w) which contains precisely the edges in $f^{-1}(C)$. The measure of this solution is given by

$$\sum_{e \in f^{-1}(C)} w(e) = \sum_{e \in (g^{\#})^{-1}(E(M))} \sum_{e' \in f^{-1}(e)} w(e') = m_M(g).$$

Since g is optimal, $m_M(g) = mc_M(N, w_N)$, and inequality (17) holds. $\square$

The possibility of efficiently computing (bounds on) $s(M, N)$ have an immediate application: Lemma 2.6 and Lemma 6.1 give necessary conditions for the existence of a homomorphism $N \to M$. As noted by Šámal (2012), this can be used as a *no-homomorphism lemma*, proving the absence of a homomorphism between two given graphs. Needless to say, establishing such properties is often a non-trivial task.

Acknowledgements

The authors are grateful to David Cohen for insightful comments leading to improvements in the presentation of this article, to Robert Šámal for valuable comments and corrections on the proof of Proposition 5.7, and to the anonymous referee for helpful feedback and suggestions.

References

N. Alon. Bipartite subgraphs. *Combinatorica*, 16:301–311, 1996.

G. Ausiello, P. Crescenzi, G. Gambosi, and V. Kann. *Complexity and Approximation*. Springer, 1999.

A. Berman and X.-D. Zhang. Bipartite density of cubic graphs. *Discrete Mathematics*, 260:27–35, 2003.

B. Bollobás and P. Erdős. Cliques in random graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 80(419):419–427, 1976.

J. A. Bondy and S. C. Locke. Largest bipartite subgraphs in triangle-free graphs with maximum degree three. *Journal of Graph Theory*, 10:477–504, 1986.

O. V. Borodin, S.-J. Kim, A. V. Kostochka, and D. B. West. Homomorphisms from sparse graphs with large girth. *Journal of Combinatorial Theory, Series B*, 90:147–159, 2004.

A. Coja-Oghlan, C. Moore, and V. Sanwalani. MAX k-CUT and approximating the chromatic number of random graphs. *Random Structures and Algorithms*, 28:289–322, 2005.

P. Crescenzi. A short guide to approximation preserving reductions. In *IEEE Conference on Computational Complexity*, pages 262–273, 1997.

R. Crowston, M. Jones, and M. Mnich. Max-Cut parameterized above the Edwards-Erdős bound. arXiv:1112.3506v1 [cs.DS], 2011.

E. de Klerk, D. V. Pasechnik, and J. P. Warners. Approximate graph colouring and MAX-k-CUT algorithms based on the θ function. *Journal of Combinatorial Optimization*, 8:267–294, 2004.

R. D. Dutton and R. C. Brigham. Edges in graphs with large girth. *Graphs and Combinatorics*, 7(4):315–321, 1991.

Z. Dvořák, R. Škrekovski, and T. Valla. Planar graphs of odd-girth at least 9 are homomorphic to the Petersen graph. *SIAM Journal on Discrete Mathematics*, 22(2):568–591, 2008.

R. Engström, T. Färnqvist, P. Jonsson, and J. Thapper. Graph homomorphisms, circular colouring, and fractional covering by H-cuts. arXiv:0904.4600v3 [cs.DM], 2009a.

R. Engström, T. Färnqvist, P. Jonsson, and J. Thapper. Properties of an approximability-related parameter on circular complete graphs. In *Proceedings of the 5th Latin-American Algorithms, Graphs, and Optimization Symposium (LAGOS-2009)*, pages 115–120, 2009b.

P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.

A. Frieze and M. Jerrum. Improved approximation algorithms for MAX k-CUT and MAX BISECTION. *Algorithmica*, 18(1):67–81, 1997.

T. Färnqvist, P. Jonsson, and J. Thapper. Approximability distance in the space of H-colourability problems. In *Proceedings of the 4th International Computer Science Symposium in Russia*, pages 92–104, 2009.

C. Godsil and G. Royle. *Algebraic Graph Theory*. Graduate Texts in Mathematics. Springer, 2001.

M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42:1115–1145, 1995.

P. Hell and J. Nešetřil. *Graphs and Homomorphisms*. Oxford Lecture Series in Mathematics and Its Applications. Oxford University Press, 2004.

P. Hell and J. Nešetřil. On the complexity of H-coloring. *Journal of Combinatorial Theory, Series B*, 48:92–110, 1990.

G. Hopkins and W. Staton. Extremal bipartite subgraphs of cubic triangle-free graphs. *Journal of Graph Theory*, 6: 115–121, 1982.

J. Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48:798–869, 2001.

J. Håstad. Every 2-CSP allows nontrivial approximation. In *Proceedings of the 37th Annual ACM Symposium on*

Theory of Computing (STOC-2005), pages 740–746, 2005.

M. Isaksson and E. Mossel. Maximally stable Gaussian partitions with discrete applications. arXiv:0903.3362v3 [math.PR], 2012.

F. Jaeger. Nowhere-zero flow problems. *Selected Topics in Graph Theory*, 3:71–95, 1988.

P. Jonsson, A. Krokhin, and F. Kuivinen. Hard constraint satisfaction problems have hard gaps at location 1. *Theoretical Computer Science*, 410(38–40):3856–3874, 2009.

V. Kann, S. Khanna, J. Lagergren, and A. Panconesi. On the hardness of approximating MAX k-CUT and its dual. *Chicago Journal of Theoretical Computer Science*, 1997(2), 1997.

A. C. Kaporis, L. M. Kirousis, and E. C. Stavropoulos. Approximating almost all instances of MAX-CUT within a ratio above the Håstad threshold. In *Proceedings of the 14th Annual European Symposium on Algorithms (ESA-2006)*, pages 432–443, 2006.

R. M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103, 1972.

S. Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC-2002)*, pages 767–775, 2002.

S. Khot, G. Kindler, E. Mossel, and R. O'Donnel. Optimal inapproximability results for MAX-CUT and other two-variable CSPs? *SIAM Journal of Computing*, 37(1):319–357, 2007.

E. J. Kim and R. Williams. Improved parameterized algorithms for constraint satisfaction. In *Proceedings of the 6th International Symposium on Parameterized and Exact Computation (IPEC 2011)*, 2011.

H.-J. Lai and B. Liu. Graph homomorphism into an odd cycle. *Bulletin of the Institute of Combinatorics and its Applications*, 28:19–24, 2000.

S. C. Locke. A note on bipartite subgraphs of triangle-free regular graphs. *Journal of Graph Theory*, 14:181–185, 1990.

D. W. Matula. The employee party problem. *Notices of the American Mathematical Society*, 19, 1972. A–382.

Z. Pan and X. Zhu. The circular chromatic number of series-parallel graphs of large odd girth. *Discrete Mathematics*, 245(1-3):235–246, 2002.

P. Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing (STOC-2008)*, pages 245–254, 2008.

P. Raghavendra and D. Steurer. How to round any CSP. In *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS-2009)*, 2009.

S. Sahni and T. Gonzalez. P-complete approximation problems. *Journal of the ACM*, 23(3):555–565, 1976.

R. Šámal. Fractional covering by cuts. In *Proceedings of the 7th International Colloquium on Graph Theory (ICGT-2005)*, pages 455–459, 2005.

R. Šámal. *On XY mappings*. PhD thesis, Charles University in Prague, 2006.

R. Šámal. Cubical coloring – fractional covering by cuts and semidefinite programming. arXiv:0911.2589v2 [math.CO], 2012.

H. U. Simon. Continuous reductions among combinatorial optimization problems. *Acta Informatica*, 26:771–785, 1989.

J. Soto. Improved analysis of a Max Cut algorithm based on spectral partitioning. arXiv:0910.0504v1 [cs.DS], 2009.

L. Trevisan. Max cut and the smallest eigenvalue. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC-2009)*, pages 263–272, 2009.

P. Turán. On an extremal problem in graph theory. *Matematicko Fizicki Lapok*, 48:436–452, 1941.

X. Zhu. Circular chromatic number. *Discrete Mathematics*, 229(1-3):371–410, 2001.