

to appear in
J. Comb. Opt.

On approximate graph colouring and MAX- k -CUT algorithms based on the ϑ -function

E. de Klerk*[§]

D.V. Pasechnik[†]

J.P. Warners[‡]

July 2, 2002

Abstract

The problem of colouring a k -colourable graph is well-known to be NP-complete, for $k \geq 3$. The MAX- k -CUT approach to approximate k -colouring is to assign k colours to all of the vertices in polynomial time such that the fraction of ‘defect edges’ (with endpoints of the same colour) is provably small. The best known approximation was obtained by Frieze and Jerrum [9], using a semidefinite programming (SDP) relaxation which is related to the Lovász ϑ -function. In a related work, Karger *et al.* [18] devised approximation algorithms for colouring k -colourable graphs exactly in polynomial time with as few colours as possible. They also used an SDP relaxation related to the ϑ -function.

In this paper we further explore semidefinite programming relaxations where graph colouring is viewed as a satisfiability problem, as considered in De Klerk *et al.* [6]. We first show that the approximation to the chromatic number suggested in [6] is bounded from above by the Lovász ϑ -function. The underlying semidefinite programming relaxation in [6] involves a lifting of the approximation space, which in turn suggests a provably good MAX- k -CUT algorithm. We show that of our algorithm is closely related to that of Frieze and Jerrum; thus we can sharpen their approximation guarantees for MAX- k -CUT for small fixed values of k . For example, if $k = 3$ we can improve their bound from 0.832718 to 0.836008, and for $k = 4$ from 0.850301 to 0.857487. We also give a new asymptotic analysis of the Frieze-Jerrum rounding scheme, that provides a unifying proof of the main results of both Frieze and Jerrum [9] and Karger *et al.* [18] for $k \gg 0$.

Keywords: Graph colouring, approximation algorithms, satisfiability, semidefinite programming, Lovász ϑ -function, MAX- k -CUT

1 Introduction

The Lovász ϑ -function [23] of a graph $G = (V, E)$ forms the base for many semidefinite programming (SDP) relaxations of combinatorial optimization problems. The ‘sandwich’ theorem of Lovász ensures that $\omega(G) \leq \vartheta(\bar{G}) \leq \chi(G)$ (see [23, 14, 21]), where $\omega(G)$ and $\chi(G)$ denote the clique and chromatic numbers of $G = (V, E)$ respectively, and \bar{G} is the complementary graph of G . Karger *et al.* [18] devised approximation algorithms for colouring k -colourable graphs. Their so-called vector chromatic number is closely related to — and bounded from above by — the ϑ -function. The authors of [18] proved that k -colourable graphs can be coloured in polynomial time by using at most $\min \left\{ \tilde{O} \left(n^{1-3/(k+1)} \right), \tilde{O} \left(\Delta^{1-2/k} \right) \right\}$ colours, where $n = |V|$ and Δ is the valency of the graph; for a review of earlier results, see [18]. This work employs the ideas of semidefinite approximation algorithms and associated randomized rounding, as introduced in the seminal work of Goemans

*Faculty of Information Technology and Systems, Department of Technical Mathematics and Informatics, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands.

[†]Theoretische Informatik, Fachbereich Biologie und Informatik, J.W. Goethe-Universität, Robert-Mayer Str.11-15, Postfach 11 19 32, 60054 Frankfurt(Main), Germany. Partially supported by the DFG Grant SCHN-503/2-1.

[‡]KPN Research, PO Box 421, 2260 AK Leidschendam, The Netherlands.

[§]Corresponding author, e-mail: E.deKlerk@TWI.TUdelft.NL

and Williamson [12] on the MAX-CUT and other problems. The results in [18] cannot be improved much for general k , since approximating the chromatic number of a general graph to within a factor $n^{1-\delta}$ for some $\delta > 0$ would imply $P = RP$. There is still some room for improvement of the results of Karger *et al.* for fixed values of k — the best related hardness result states that all 3-colourable graphs cannot be coloured in polynomial time using 4 colours, unless $P = NP$ [20]. The best known (exponential) algorithm for exact 3-colouring runs in $O(1.3446^n)$ time [4].

The graph colouring problem for k -colourable graphs can be seen as a special case of the unweighted MAX- k -CUT problem.¹ Approximation algorithms for the MAX- k -CUT problem assign a colour from a set of k colours to each vertex in polynomial time so as to minimize the number of defect edges (see *e.g.* [5]). The approximation guarantee of a MAX- k -CUT approximation algorithm is the ratio of the number of non-defect edges in the approximate solution to the maximal number of non-defect edges. The fraction $1 - 1/k$ is achieved by a random assignment of colours to the vertices. This can be slightly improved to $(1 - \frac{1}{k} + 2 \ln(k)/k^2)$ for sufficiently large values of k , as shown by Frieze and Jerrum [9]; there is very little room for further improvement of this result, since the attainable approximation guarantee is upper bounded by $1 - 1/(34k)$, unless $P=NP$ [17]. For fixed values of k the approximation guarantee can be improved. For example, Frieze and Jerrum [9] showed that a guarantee of 0.832718 is attainable for MAX-3-CUT. The approach in [9] is closely linked to that by Karger *et al.* in the sense that both underlying semidefinite programming relaxations are related to the same formulation of the ϑ function. In another related work, Alon and Kahale [2] proposed an approximation algorithm for the independent set number of a graph based on the ϑ -function.

Recently, De Klerk *et al.* [6] proposed a SDP based approximation of $\chi(G)$ by casting the colouring problem as a logical formula (satisfiability (SAT) problem). This logical formula is an encoding of the problem: ‘Is $\chi(G) \leq k$?’ for a given integer k ; the SAT problem is subsequently relaxed to a semidefinite programming (SDP) feasibility problem, and the approximation to $\chi(G)$ is the smallest value of k for which the SDP feasibility problem is feasible. The authors showed that the resulting approximation to $\chi(G)$ satisfies the same sandwich theorem as does $\vartheta(\bar{G})$.

In the first part of this paper we show that a slightly tighter SDP relaxation of the SAT formula yields $[\vartheta(\bar{G})]$ as an approximation to $\chi(G)$. Thus we show that the approximation to $\chi(G)$ from [6] is bounded from above by $[\vartheta(\bar{G})]$. The SDP relaxation we employ is closely related to the types of MAX-SAT relaxations studied by Karloff and Zwick [19], Zwick [26], and Halperin and Zwick [15]. The relaxation problem involves a lifting of the usual $n \times n$ matrix variables to $kn \times kn$ matrix variables. The advantage is that the different choice of variables suggests a rounding strategy that is simple to analyze. In the second part of the paper we show that this rounding procedure is actually equivalent to the rounding procedure by Frieze and Jerrum. Thus we can refine their approximation guarantees MAX- k -CUT where k is fixed (small). For example, if $k = 3$ we can improve their bound from 0.832718 to 0.836008, and for $k = 4$ from 0.850301 to 0.857487. Since the first submission of this paper, the improved bound for MAX-3-CUT was also independently obtained by Goemans and Williamson [13]. They also used a semidefinite programming relaxation of MAX-3-CUT, but with the novel feature that complex (Hermitian) matrix variables are used. Nevertheless, they showed that their method is in fact equivalent to that by Frieze and Jerrum [9] for MAX-3-CUT as well as to the method by Anderson *et al.* [3] for solving linear equations mod p when applied to the special case of MAX-3-CUT. Goemans and Williamson [13] also show that the MAX-3-CUT guarantee can be expressed as the minimum of a function in closed form. However, there is no obvious way to extend their approach to MAX- k -CUT if $k > 3$.

We will also give a theorem describing the asymptotic behavior of the rounding scheme in the Frieze-Jerrum MAX- k -CUT algorithm for large k . We show how the main results of Frieze and Jerrum [9] and Karger *et al.* [18] for $k \gg 0$ follow from it, thus giving a unified view of these results.

Outline of this paper

We show how to formulate the graph colouring problem as a Boolean quadratic feasibility problem in Section 2, and thereafter review the standard procedure to relax such problems to semidefinite feasibility problems in Section 3. Subsequently we apply the general relaxation procedure to our reformulation of the graph colouring problem in Section 4, and thus define an approximation to the chromatic number which is an upper bound to the approximation defined by De Klerk *et al.* [6]. In Section 5 we show that the new approximation is simply

¹In the (unweighted) MAX- k -CUT problem the goal is to assign k colours to the vertices of a given graph so as to minimize the number of defect edges (edges where both endpoints have the same colour). Note that a k -colourable graph allows a k -cut where there are no defect edges.

$\lceil \vartheta(\bar{G}) \rceil$, thus giving an upper bound on the approximation in [6]. Based on this analysis we propose a MAX- k -CUT algorithm in Section 6 and prove that it is equivalent to the Frieze-Jerrum algorithm for graphs where $\vartheta(\bar{G}) = k$. We derive an expression for the performance guarantee of our algorithm in Section 7. In Sections 8 and 9 we give performance guarantees of our algorithm for the MAX- k -CUT problem; this is followed by a section on approximate graph colouring where we show how to derive the main result by Karger *et al.* [18] from the analysis in our paper.

2 The colouring problem as a Boolean quadratic feasibility problem

Consider a simple graph $G = (V, E)$ with $|V| := n$ and chromatic number $\chi(G)$. The colouring problem is stated as: Can G be coloured using k colours, *i.e.* is $\chi(G) \leq k$?

We define the following $\{-1, 1\}$ variables:

$$x_i^j = \begin{cases} 1 & \text{if vertex } i \text{ has colour } j \\ -1 & \text{otherwise.} \end{cases}$$

Adjacent vertices may not share a colour:

$$(x_i^p + x_j^p + 1)^2 = 1 \text{ if } (i, j) \in E \quad (1)$$

for each colour $p = 1, \dots, k$, and each vertex must be assigned exactly one colour:

$$x_i^1 + \dots + x_i^k = -k + 2 \quad (i = 1, \dots, n). \quad (2)$$

Squaring both sides of (2) and using the fact that the variables are $\{-1, 1\}$ one also has

$$\sum_{p=1}^k \sum_{q>p}^k x_i^p x_i^q = \frac{1}{2} ((2-k)^2 - k) = \frac{1}{2} (k^2 - 5k + 4). \quad (3)$$

Having chosen the Boolean quadratic representation the graph colouring problem, namely (1), (2) and (3), we can now use a standard procedure to relax it to an SDP feasibility problem.

3 SDP relaxations of Boolean quadratic equalities

Let us assume that

$$x^T A x + 2b^T x + c = 0, \quad x \in \{-1, 1\}^n \quad (4)$$

is a given Boolean quadratic equality. Equation (4) can be rewritten as

$$\text{Tr} \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \begin{bmatrix} x x^T & x \\ x^T & 1 \end{bmatrix} = 0, \quad x \in \{-1, 1\}^n, \quad (5)$$

where ‘Tr’ denotes the trace operator. We can relax (5) to

$$\text{Tr} \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \begin{bmatrix} X & y \\ y^T & 1 \end{bmatrix} = 0,$$

where X is now a symmetric $n \times n$ matrix which satisfies

$$X - y y^T \succeq 0, \quad (6)$$

and $X_{jj} = 1$ ($j = 1, \dots, n$) ($A \succeq 0$ means A is symmetric positive semidefinite). Note that the entries X_{ij} correspond to the products $x_i x_j$ and y_i corresponds to x_i . Also note that (6) is equivalent to

$$\begin{bmatrix} X & y \\ y^T & 1 \end{bmatrix} \succeq 0,$$

by the Schur complement theorem (see *e.g.* Theorem 7.7.6 in [16]).

4 An SDP relaxation of graph colouring

Recall that the variables of the SDP relaxation problem correspond to pairwise products of the $\{-1, 1\}$ variables and to the variables appearing in linear terms. The possible products which can occur and their corresponding relaxation variables are respectively listed and defined in Table 1.

Product	Corresponding relaxation variable
$x_i^p x_j^q$	X_{ij}^{pq}
$x_i^p x_j^p$	X_{ij}^p
$x_i^p x_i^q$	X_i^{pq}
$x_i^p \times 1$	X_{iT}^p

Table 1: The notation for variables appearing in the SDP relaxation

It is easy to check that (1) gives rise to the following constraint in the SDP relaxation:

$$X_{ij}^p + X_{iT}^p + X_{jT}^p = -1 \quad (7)$$

for each edge (i, j) and each colour p . Similarly (2) gives rise to

$$\sum_{p=1}^k X_{iT}^p = 2 - k \quad (8)$$

for each vertex i . Summing over p in (7) and using (8), we immediately derive

$$\sum_{p=1}^k X_{ij}^p = 4 - k, \quad (9)$$

if $(i, j) \in E$. Finally, (3) gives rise to

$$\sum_{p=1}^k \sum_{q < p}^k X_i^{pq} = \frac{1}{2} (k^2 - 5k + 4), \quad i = 1, \dots, n. \quad (10)$$

We now order our $(kn+1) \times (kn+1)$ SDP relaxation matrix X as follows: X consists of n^2 blocks of size $(k \times k)$ and is bordered by a kn -vector and has entry 1 in the lower right corner. Each $(k \times k)$ block corresponds to a pair of vertices (say (i, j)), and its entries correspond to the relaxation of the products $x_i^p x_j^q$, where p and q range over all the colours.

The bordering kn -vector is divided into n k -vectors, where the i th k -vector is given by $[X_{iT}^1 \dots X_{iT}^k]$ ($i = 1, \dots, n$). A schematic representation of the structure of X is given in Figure 1.

We formally define our SDP feasibility problem as follows:

$$\text{Find whether there exists an } X = X^T \in \mathbb{R}^{kn+1 \times kn+1} \quad (\text{SDFP})$$

such that²

$$X \succeq 0, \quad \text{diag}(X) = e, \quad i = 1, \dots, kn + 1,$$

and the constraints (7), (8) and (10) are satisfied.

We further define the number $\bar{k}(G)$ as the smallest value of k for which the SDP feasibility problem (SDFP) has a solution.

²Here, $\text{diag}(X)$ means the vector obtained by extracting the diagonal of X and e is the all-one vector.

1	X_1^{12}	...	X_1^{1k}	X_{1n}^1	X_{1n}^{12}	...	X_{1n}^{1k}	X_{1T}^1
	1		\vdots						X_{1n}^2		\vdots	\vdots
		\ddots	X_1^{k-1k}							\ddots	X_{1n}^{k-1k}	\vdots
			1								X_{1n}^k	X_{1T}^k
\vdots			\vdots	\ddots				\vdots			\vdots	\vdots
\vdots			\vdots		\ddots			\vdots			\vdots	\vdots
\vdots			\vdots			\ddots		\vdots			\vdots	\vdots
\vdots			\vdots				\ddots	\vdots			\vdots	\vdots
X_{1n}^1	X_{1n}^{12}	...	X_{1n}^{1k}	1	X_n^{12}	...	X_n^{1k}	X_{nT}^1
	X_{1n}^2		\vdots						1		\vdots	\vdots
		\ddots	X_{1n}^{k-1k}							\ddots	X_n^{k-1k}	\vdots
			X_{1n}^k								1	X_{nT}^k
X_{1T}^1	X_{1T}^k	X_{nT}^1	X_{nT}^k	1

Figure 1: Schematic representation of the block structure of the matrix X . The $k \times k$ blocks are symmetric and only the upper triangular parts are shown.

5 Relation with the Lovász ϑ -function

We show in this section that $\bar{k}(G) = \lceil \vartheta(\bar{G}) \rceil$. To quote Goemans [11]: ‘It seems all roads lead to ϑ ’!

The ϑ -function has many representations (see [14]); the representation we will use is by Karger *et al.* [18]:

$$\vartheta(\bar{G}) = \min_{U,k} k \tag{*}$$

subject to

$$\begin{aligned} U_{ij} &= \frac{-1}{k-1}, \text{ if } (i,j) \in E \\ U_{ii} &= 1, \quad i = 1, \dots, n \\ U &\succeq 0, \quad k \geq 2. \end{aligned}$$

We will show that if X is feasible in (SDFP) for some integer $k \geq 2$, then one can construct a feasible solution for problem (*). Conversely, if (U, k) is feasible in problem (*) for some integer k , then a feasible solution to (SDFP) can be constructed. This then implies that $\bar{k}(G) = \lceil \vartheta(\bar{G}) \rceil$.

Theorem 5.1 *Let $G = (V, E)$ be given. Then $\bar{k}(G) = \lceil \vartheta(\bar{G}) \rceil$.*

Proof: Let k be given and assume that $X \succeq 0$ is feasible in (SDFP) for k . We define the $(kn+1) \times (kn+1)$ matrix

$$\hat{Y} = \frac{1}{k!} \sum_{i=1}^{k!} Q_i X Q_i^T$$

where the $(kn+1) \times (kn+1)$ matrices Q_i ($i = 1, \dots, k!$) take the form

$$Q_i = \begin{bmatrix} P_i & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & P_i & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where P_i ($i = 1, \dots, k!$) are all the $k \times k$ permutation matrices. In other words, Q_i has n copies of the $k \times k$ permutation matrix P_i as diagonal blocks and 1 in the lower right corner. Note that $\hat{Y} \succeq 0$ since $X \succeq 0$.³

Note that \hat{Y} has the same block diagonal structure as X . For each square block of \hat{Y} there holds: the diagonal entries are constant and given by the average of the diagonal entries of the corresponding block of X ; the off-diagonal entries are likewise equal and given by the average of the off-diagonal entries of the corresponding blocks of X . The bordering vector of X is mapped to the vector of all ones times the average of the vector components. In other words, all components of the bordering vector of \hat{Y} equal $(2 - k)/k$, by (8).

The diagonal blocks of \hat{Y} have ones on their diagonals. The off diagonal elements are easily shown to equal $(k - 4)/k$ by using (10). Moreover, if block (i, j) corresponds to an edge $(i, j) \in E$, then all the diagonal entries of this block equal $(k - 4)/k$, by (9).

Note that \hat{Y} satisfies the constraints (7), (8) and (10). This means that X is feasible for (SDFP) if and only if \hat{Y} is feasible for (SDFP). The matrix \hat{Y} has the form:

$$\hat{Y} = \begin{bmatrix} Y_{kn \times kn} & \frac{2-k}{k} e_{kn} \\ \frac{2-k}{k} e_{kn}^T & 1 \end{bmatrix},$$

where e_t denotes the all-ones vector of length t . Now define the $(kn \times kn)$ matrix

$$\bar{Y} := Y_{kn \times kn} - \left(\frac{2-k}{k}\right)^2 e_{kn} e_{kn}^T \quad (11)$$

which is positive semidefinite by the Schur complement theorem. More precisely, $\bar{Y} \succeq 0$ if and only if $\hat{Y} \succeq 0$. All the diagonal blocks of \bar{Y} are identical and given by:

$$\left(\frac{4}{k}\right) I_k + \frac{k-4}{k} e_k e_k^T - \left(\frac{2-k}{k}\right)^2 e_k e_k^T = \frac{4}{k} \left(I_k - \frac{1}{k} e_k e_k^T \right),$$

where I_k denotes the identity matrix of size $k \times k$. The blocks of Y corresponding to edges have diagonal entries $-4/k^2$. The diagonal of \bar{Y} can now be normalized to one by defining

$$Y := \frac{k^2}{4(k-1)} \bar{Y}. \quad (12)$$

We now take a $(n \times n)$ principal submatrix of Y (say U) determined by the rows and columns with indices $t \times k + 1$ where $t = 0, 1, \dots, n - 1$. Straightforward calculations yield that

$$U_{ij} = \begin{cases} 1 & \text{if } i = j \\ \frac{-1}{k-1} & \text{if } (i, j) \in E. \end{cases}$$

This proves that $\bar{k}(G) \geq \lceil \vartheta(\bar{G}) \rceil$.

Conversely, assume that a feasible solution U, k is given for problem $(*)$, where k is an integer. We can now construct Y via

$$Y = U \otimes \frac{k}{k-1} \left(I_k - \frac{1}{k} e_k e_k^T \right), \quad (13)$$

where \otimes denotes the Kronecker product and subsequently obtain a feasible solution of (SDFP). This implies that $\lceil \vartheta(\bar{G}) \rceil \geq \bar{k}(G)$, which completes the proof. \square

Remark 5.1 Note that we can solve (SDFP) by solving problem $(*)$ to obtain $\vartheta(\bar{G})$. If $\vartheta(\bar{G}) > k$ then (SDFP) has no feasible solution. If $\vartheta(\bar{G}) \leq k$, one can find a feasible solution U of $(*)$ with objective value $\lceil \vartheta(\bar{G}) \rceil$ and subsequently construct a feasible solution of (SDFP) via (13).

³The above construction is well-known and was used *e.g.* by Lovász [23] in his study of the Shannon capacity of graphs.

Remark 5.2 The approximation to $\chi(G)$ from [6] (say $|C^*|$) was defined in a similar way as $\bar{k}(G)$. The only difference was that the requirement that each node should be coloured was modeled using the quadratic inequalities

$$\left(x_i^1 + \dots + x_i^k - 1\right)^2 \leq (k-1)^2, \quad (14)$$

instead of (2) and (3). Since (14) leads to a weaker restriction in the associated SDP feasibility problem than (2) and (3), we have

$$|C^*| \leq \bar{k}(G) = \lceil \vartheta(\bar{G}) \rceil. \quad (15)$$

Theorem 5.1 still holds if the constraint (10) is dropped in the formulation of (SDFP). The constraint (10) was only included to prove the inequality (15).

Remark 5.3 There are many possible valid inequalities for (SDFP); the well-known triangle inequalities may be added, for example. Feige and Goemans [7] derived improved approximation results for MAX-2-SAT by including the triangle inequalities associated with the last row of X . For (SDFP) the analogous set of inequalities is given by:

$$\pm X_{iT}^p \pm X_{jT}^q \pm X_{ij}^{pq} \geq -1, \quad i, j = 1, \dots, n, \quad p, q = 1, \dots, k, \quad (16)$$

where the \pm 's must contain a even number of minuses. If we add all triangle inequalities to (SDFP), we obtain a problem (SDFP') with associated $\bar{k}'(G) \geq \bar{k}(G) = \lceil \vartheta(\bar{G}) \rceil$. This is analogous to the definition of the ϑ' -function of Schrijver [25] which satisfies $\vartheta'(\bar{G}) \leq \vartheta(\bar{G})$.

The subclass of triangle inequalities where $p = q$ in (16), imply

$$U_{ij} \geq \frac{-1}{k-1}, \quad i, j = 1, \dots, n, \quad (17)$$

for problem (*). These valid inequalities for (*) are not mentioned explicitly in [18], but follow from arguments presented there; they *are* included in the MAX- k -CUT relaxation of Frieze and Jerrum [9].

The inclusion of all possible triangle inequalities in (SDFP) implies the addition of the valid inequalities (17) as well as

$$U_{ij} - U_{jl} - U_{il} \geq -1, \quad i, j, l = 1, \dots, n, \quad (18)$$

to problem (*).

6 A randomized MAX- k -CUT algorithm

We now propose a randomized MAX- k -CUT algorithm that will turn out to be closely related to that of Frieze and Jerrum [9], although it appears quite different at the first glance. The algorithm is inspired by the alternative formulation of the ϑ -function presented in the previous section.

We will work with vectors $[v_1^1, v_1^2, \dots, v_1^k, \dots, v_n^k]$, where each vector is associated with a particular vertex and a particular colour. We can then generate a random 'truth vector' on the unit ball in \mathbb{R}^{kn} and assign colour p to vertex i if v_i^p is the 'closest' to r of the k vectors associated with vertex i . In other words, we assign colour p to vertex i if $p = \arg \max_{q=1, \dots, k} \{r^T v_i^q\}$. Note that each vertex is assigned exactly one colour in this way, with probability one.

The algorithm we propose is as follows.

Algorithm 1: Randomized MAX- k -CUT Algorithm

1. Find an optimal solution $(U, \vartheta(\bar{G}))$ of problem (*), and let

$$Y = U \otimes \frac{k}{k-1} \left(I_k - \frac{1}{k} e_k e_k^T \right). \quad (19)$$

2. *Rounding scheme:*

Perform a factorization $Y = V^T V$ where $V = [v_1^1, v_1^2, \dots, v_1^k, \dots, v_n^k]$;

Choose a random unit vector $r \in \mathbb{R}^{kn}$ from the uniform distribution on the unit ball in \mathbb{R}^{kn} .⁴

Set $x_i^p = 1$ if $r^T v_i^p = \max_{q=1, \dots, k} \{r^T v_i^q\}$; Otherwise, set $x_i^p = -1$.

⁴This can be done by choosing each component of r independently from the standard normal distribution with mean zero and variance one, and subsequently normalizing.

We proceed to show that the rounding scheme of Algorithm 1 is equivalent to that used by the MAX- k -CUT algorithm of Frieze and Jerrum [9].

The MAX- k -CUT algorithm of Frieze and Jerrum works as follows.

Jerrum-Frieze MAX- k -CUT algorithm

1. Find $U \succeq 0$ such that $U_{ij} \geq \frac{-1}{k-1}$, ($i, j = 1, \dots, n$), and $U_{ii} = 1$ ($i = 1, \dots, n$) which maximizes

$$\frac{k-1}{k} \sum_{(i,j) \in E} (1 - U_{ij}).$$

2. *Rounding scheme:*

Take the factorization $U = \bar{V}^T \bar{V}$, and denote $\bar{V} = [v_1, \dots, v_n]$;

Generate k different random vectors $r^{(1)}, \dots, r^{(k)} \in \mathbb{R}^n$; (Random vectors here means that each component of each vector is drawn independently from the standard normal distribution with mean zero and variance one.) Each vector $r^{(i)}$ is associated with a colour i ;

Assign colour i to vertex j if $i = \arg \max_t v_j^T r^{(t)}$, *i.e.* assign the colour corresponding to the ‘closest’ random vector.

Our goal here is to prove the following.

Theorem 6.1 *For a given matrix $U \succeq 0$, the rounding schemes of the MAX- k -CUT algorithm of Jerrum-Frieze and Algorithm 1 are equivalent.*

Proof: Let $U \succeq 0$ and k be given. The first step of Algorithm 1 is to take the Kronecker product

$$Y = U \otimes c \left(I_k - \frac{1}{k} e_k e_k^T \right) = \bar{V}^T \bar{V} \otimes c \left(I_k - \frac{1}{k} e_k e_k^T \right)$$

where $c = \frac{k}{k-1}$, and subsequently obtaining the Choleski factor $[v_1^1 v_1^2 \dots v_n^k]$ of Y . Vertex i is then assigned colour p if

$$p = \arg \max_t v_i^{tT} r \tag{20}$$

for a random $r \in \mathbb{R}^{kn}$ (each component of r is independently drawn from the standard normal distribution with mean zero and variance one). The identity

$$\left(B^T B \right) \otimes \left(G^T G \right) = \left(B^T \otimes G^T \right) \left(B \otimes G \right) = \left(B \otimes G \right)^T \left(B \otimes G \right),$$

implies

$$\bar{V}^T \bar{V} \otimes c \left(I_k - \frac{1}{k} e_k e_k^T \right) = \left(\bar{V} \otimes \sqrt{c} \left(I_k - \frac{1}{k} e_k e_k^T \right) \right)^T \left(\bar{V} \otimes \sqrt{c} \left(I_k - \frac{1}{k} e_k e_k^T \right) \right), \tag{21}$$

where we have used the fact that the matrix $\left(I_k - \frac{1}{k} e_k e_k^T \right)$ is idempotent. We define the matrices:

$$X_i = \left[v_i^1 v_i^2 \dots v_i^k \right]^T, \quad i = 1, \dots, n. \tag{22}$$

Note that by (20), vertex i is assigned colour p iff

$$p = \arg \max_t (X_i r)_t.$$

By (21) and (22) we have

$$X_i = v_i^T \otimes \sqrt{c} I_k - v_i^T \otimes \frac{\sqrt{c}}{k} e_k e_k^T,$$

so that

$$X_i r = \left(v_i^T \otimes \sqrt{c} I_k \right) r - \left(v_i^T \otimes \frac{\sqrt{c}}{k} e_k e_k^T \right) r = \left(v_i^T \otimes \sqrt{c} I_k \right) r - c_i e_k$$

where c_i is a scalar depending on r and v_i . Note that we can ignore the last term when finding the largest component of $X_i r$. In other words,

$$\arg \max_p (X_i r)_p = \arg \max_p \left(v_i^T \otimes \sqrt{c} I_k r \right)_p. \tag{23}$$

Finally, we construct a set of k random vectors in \mathbb{R}^n from r as follows:

$$r^{(i)} := [r_i, r_{i+k}, r_{i+2k}, \dots, r_{i+(n-1)k}]^T, \quad i = 1, \dots, k.$$

Note that, by construction,

$$(v_i^T \otimes \sqrt{c} I_k) r = \sqrt{c} [v_i^T r^{(1)}, \dots, v_i^T r^{(k)}]^T, \quad i = 1, \dots, n,$$

so that

$$\arg \max_p (X_i r)_p = \arg \max_p v_i^T r^{(p)}, \quad i = 1, \dots, n,$$

by (23). This completes the proof. \square

Note that the matrix U used by the rounding scheme of Algorithm 1 may be different from that used in the Frieze-Jerrum rounding scheme. However, it is easy to verify that for graphs where $\vartheta(\bar{G}) = k$, the U -matrices will be the same, namely $U = \frac{k}{k-1}I - \frac{1}{k-1}ee^T$.

Corollary 6.1 *The MAX- k -CUT algorithm of Frieze-Jerrum and Algorithm 1 are equivalent for graphs $G = (V, E)$ where $\vartheta(\bar{G}) = k$.*

7 Analysis of Algorithm 1

We proceed to give a simple proof – using geometric arguments only – which establishes the probability that a given edge is defect after running Algorithm 1. In particular, we wish to know what the probability (say p_1) is that both endpoints of a given edge are assigned colour 1. The probability that the edge is defect is then simply kp_1 since the number of colours used equals k . The expected fraction of non-defect edges simply equals $(1 - kp_1)$, by the linearity of expectation.

Note that both endpoints of an edge (i, j) have been assigned colour 1 if and only if:

$$r^T v_i^1 \geq r^T v_i^q, \quad q = 2, \dots, k,$$

and

$$r^T v_j^1 \geq r^T v_j^q, \quad q = 2, \dots, k.$$

In other words, r must lie in the dual of the cone spanned by the vectors

$$(v_i^1 - v_i^2), \dots, (v_i^1 - v_i^k), (v_j^1 - v_j^2), \dots, (v_j^1 - v_j^k). \quad (24)$$

An alternative geometrical interpretation is that the half space with outward pointing normal vector $-r$ must contain the vectors (24), *i.e.* the vectors (24) must lie on a specific side of a random hyperplane with normal vector r (the same side as r). The probability that a given set of vectors lie on the same side of a random hyperplane has been investigated recently by Karloff and Zwick [19] (at most 4 vectors) and Zwick [26] (general case) in the context of MAX-SAT approximation algorithms. In what follows, we employ the same approach as these authors.

For convenience of notation we define the unit vectors

$$w_i^q = \frac{v_i^1 - v_i^q}{\|v_i^1 - v_i^q\|}, \quad q = 2, \dots, k, \quad i = 1, \dots, n.$$

The w vectors can be viewed as a set of $(2k - 2)$ points on the $(2k - 3)$ -dimensional unit hypersphere

$$\mathcal{S}^{(2k-3)} := \{x \in \mathbb{R}^{2k-2} \mid \|x\| = 1\},$$

and thus define a so-called *spherical simplex* (say \mathbf{S}) in the space $\mathcal{S}^{(2k-3)}$.

The *Gram matrix* of the w vectors (which has the inner products of the w -vectors – *i.e.* the cosines of the edge lengths of \mathbf{S} – as entries) is known explicitly, since the corresponding entries in the matrix Y in (19) are known. In particular, it is easy to show that the Gram matrix is given by:

$$\text{Gram}(\mathbf{S}) := \frac{1}{2} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \otimes (I_{k-1} + e_{k-1}e_{k-1}^T), \quad (25)$$

where $\rho = \frac{-1}{\vartheta(\bar{G})-1}$.

From a geometrical viewpoint, we are interested in the volume of the spherical simplex (say \mathbf{S}^*) which is dual to the spherical simplex \mathbf{S} , as a fraction of the total volume of the unit hypersphere $\mathcal{S}^{(2k-3)}$. This dual spherical simplex is given by:

$$\mathbf{S}^* = \left\{ x \in \mathcal{S}^{(2k-3)} \mid x^T z \geq 0 \quad \forall z \in \mathbf{S} \right\}.$$

The Gram matrix associated with \mathbf{S}^* is given by taking the inverse of $\text{Gram}(\mathbf{S})$ in (25) and subsequently normalizing its diagonal. Straightforward calculations show that this matrix takes the form

$$\text{Gram}(\mathbf{S}^*) := \frac{k}{k-1} \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix} \otimes \left(I_{k-1} - \frac{1}{k} e_{k-1} e_{k-1}^T \right). \quad (26)$$

The volume of a spherical simplex is completely determined by the off-diagonal entries of its Gram matrix. Unfortunately, there is no closed form expression available for the volume function, and it must be evaluated by (numerical) integration (see [19, 26, 1]). The integral which yields p_1 is given by [1]:

$$p_1 = \frac{\text{vol}(\mathbf{S}^*)}{\text{vol}(\mathcal{S}^{(2k-3)})} = \frac{1}{\sqrt{\det(\text{Gram}(\mathbf{S}))} \pi^{2k-2}} \int_0^\infty \dots \int_0^\infty e^{-y^T \text{Gram}(\mathbf{S})^{-1} y} dy_1 \dots dy_{2k-2}. \quad (27)$$

By Theorem 6.1, the expression for p_1 must be equivalent to that given by Frieze and Jerrum.⁵

$$p_1 = I(\rho) := \int_{-\infty}^\infty \int_{-\infty}^\infty f(x, y, \rho) F(x, y, \rho)^{k-1} dx dy, \quad (28)$$

where

$$f(x, y, \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right) \quad (29)$$

is the density function of the joint bivariate normal distribution in standard form and

$$F(x, y, \rho) = \int_{-\infty}^x \int_{-\infty}^y f(\xi, \eta, \rho) d\eta d\xi \quad (30)$$

the corresponding cumulative distribution function. (The notation $I(\rho)$ was used by Frieze and Jerrum [9].) The only difference is in the meaning of the parameter ρ : in Algorithm 1, ρ always corresponds to $\frac{-1}{\vartheta(\bar{G})-1}$; in the algorithm by Frieze–Jerrum there is no simple relation between ρ and the ϑ -function — one only knows that $\rho \in [-1/(k-1), 1]$ for each edge.

Both representations of p_1 are useful — (27) allows us to compute p_1 accurately for small, fixed values of k and (28) is more suitable for asymptotic analysis where $k \rightarrow \infty$.

8 On the MAX- k -CUT performance guarantees

We show here that the performance guarantee (say α_k) of Algorithm 1 for the MAX- k -CUT problem is bounded by

$$\alpha_k \geq \frac{k}{k-1} \left(\frac{1 - kI(\rho)}{1 - \rho} \right), \quad (31)$$

where $\rho = \frac{-1}{\vartheta(\bar{G})-1}$ as before, and $I(\rho)$ is defined in (28). Since $\vartheta(\bar{G}) \in [2, \infty)$, this bound becomes

$$\alpha_k \geq \frac{k}{k-1} \min_{-1 \leq \rho < 0} \frac{1 - kI(\rho)}{1 - \rho}.$$

⁵The equivalence of the expressions (28) and (27) can also be shown analytically as follows: Rewrite (30) as the double integral (say F_i) with limits from 0 to infinity by substituting $y_i = x - \xi$, $y_{i+k} = y - \eta$. Rewrite F^{k-1} in (28) as the product of the F_i 's ($i = 1, \dots, k-1$). Now change the order of integration and integrate with respect to x and y .

This is very similar to the bound given by Frieze and Jerrum for their algorithm, namely

$$\alpha_k^{JF} \geq \frac{k}{k-1} \min_{-1/(k-1) \leq \rho \leq 1} \frac{1 - kI(\rho)}{1 - \rho}.$$

In fact, Frieze and Jerrum mention that the minimum will be attained at a negative value of ρ . This means that the performance guarantee of their algorithm is bounded by

$$\alpha_k^{JF} \geq \frac{k}{k-1} \min_{-1/(k-1) \leq \rho < 0} \frac{1 - kI(\rho)}{1 - \rho} \geq \alpha_k. \quad (32)$$

The first step in deriving (31) is to give an upper bound on the cardinality of a maximum k cut in terms of the ϑ -function. To this end, we first give a reformulation of the MAX- k -CUT problem. Let r_1, \dots, r_k be a (fixed) set of vectors in \mathbb{R}^n ($n > k$) such that

$$r_i^T r_j = \begin{cases} -\frac{1}{k-1} & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases} \quad (33)$$

(Such a set of vectors always exists — one can take the columns of the Choleski factorization of the positive semidefinite matrix $\frac{k}{k-1}I - \frac{1}{k-1}ee^T$, for example.)

We will associate these vectors with k different colours. Similarly we will associate n unit vectors y_i ($i = 1, \dots, n$) with the set of vertices V . Thus we assign r_j to y_i if we wish to assign colour j to vertex i .

$$OPT := \max_{y_1, \dots, y_n} \frac{k-1}{k} \sum_{(i,j) \in E} (1 - y_i^T y_j), \quad (\text{MAX-}k\text{-CUT})$$

subject to

$$y_j \in \{r_1, \dots, r_k\} \quad (j = 1, \dots, n). \quad (34)$$

After an assignment of colours to the endpoints of the edge $(i, j) \in E$ we have

$$\frac{k-1}{k}(1 - y_i^T y_j) = \begin{cases} 1 & \text{if } (i, j) \text{ is non-defect} \\ 0 & \text{if } (i, j) \text{ is defect.} \end{cases}$$

Thus it is easy to see that we have given a valid mathematical formulation of MAX- k -CUT. We now use this formulation to prove the following.

Lemma 8.1 *Let a graph $G = (V, E)$ and an integer $k > 2$ be given, and let OPT denote the cardinality of the maximum k cut, as before. One has*

$$OPT \leq \frac{k-1}{k} |E| \left(\frac{\vartheta(\bar{G})}{\vartheta(\bar{G}) - 1} \right).$$

Proof: Let an optimal set of vectors $\bar{y}_1, \dots, \bar{y}_k$ be given for problem (MAX- k -CUT), and let \hat{t} be such that

$$\frac{-1}{\hat{t}-1} = \frac{1}{|E|} \sum_{(i,j) \in E} \bar{y}_i^T \bar{y}_j.$$

Note that $\frac{-1}{\hat{t}-1} \in \left[\frac{-1}{k-1}, 1 \right]$ and therefore there exist unit vectors $\hat{y}_1, \dots, \hat{y}_k$ such that

$$\frac{-1}{\hat{t}-1} = \hat{y}_i^T \hat{y}_j \quad \forall (i, j) \in E.$$

(For example, one can take the columns vectors of the Choleski factorization of the positive semidefinite matrix $\frac{\hat{t}}{\hat{t}-1}I - \frac{1}{\hat{t}-1}ee^T$).

It follows that

$$OPT = \frac{k-1}{k} |E| \frac{\hat{t}}{\hat{t}-1}.$$

Since the Gram matrix of $\hat{y}_1, \dots, \hat{y}_k$ (say U) together with \hat{t} give a feasible solution of problem $(*)$, we obtain $\hat{t} \geq \vartheta(\bar{G})$, which in turn implies the result of the lemma. \square

Let $\rho = -1/(\vartheta(\bar{G})-1)$, and recall that the expected size of a k -cut given by Algorithm 1 is simply $|E|(1-kI(\rho))$. The performance guarantee of Algorithm 1 is now bounded by

$$\begin{aligned} \alpha_k &= \frac{|E|(1-kI(\rho))}{OPT} \\ &\geq \frac{|E|(1-kI(\rho))}{\frac{k-1}{k}|E|(1-\rho)} \quad (\text{by Lemma 8.1}) \\ &= \frac{k}{k-1} \frac{1-kI(\rho)}{1-\rho}. \end{aligned}$$

In the next section we will compute this bound for $k = 3, \dots, 10$ by solving for $I(\rho)$ from (27). Moreover, we will look at the asymptotic behavior of α_k as $k \rightarrow \infty$.

9 Results for MAX- k -CUT

We will now calculate the MAX-3-CUT guarantee of Algorithm 1.

In this case ($k = 3$), the integral (28) can be solved analytically (see Appendix 1) to obtain:

$$p_1 = I(\rho) = \frac{1}{9} + \frac{\arccos(-\rho) - \arccos^2(\rho/2)}{4\pi^2}.$$

This function attains a minimum at $\rho = \frac{-1}{k-1} = -\frac{1}{2}$ on the interval $\rho \in [-1, 0)$, and we obtain

$$\alpha_3 \geq \frac{k}{k-1} \frac{1-kI(\rho)}{1-\rho} \geq \frac{3}{2} \left(\frac{1-3I(-\frac{1}{2})}{1-(-\frac{1}{2})} \right) \approx 0.836008.$$

The bound $\alpha_3 \geq 0.836008$ was first obtained via numerical integration in an earlier version of this paper. The same bound was independently obtained by Goemans and Williamson [13], who were the first to show (indirectly) that the integral (28) is analytically solvable for $k = 3$.

Note that the worst guarantee occurs where $\vartheta(\bar{G}) = k = 3$. We cannot prove that this is also true for all $k \geq 3$, but will show that it is true for $k = 4, \dots, 10$, and asymptotically (as $k \rightarrow \infty$). Based on this evidence we make the following conjecture.

Conjecture 9.1 *Let $k > 2$ be any integer. The worst-case performance of Algorithm 1 is attained for graphs where $\vartheta(\bar{G}) = k$.*

The conjecture would imply that the worst-case for the algorithm of Frieze–Jerrum is also attained when $\vartheta(\bar{G}) = k$, by (32).

In order to prove the conjecture for $k = 4, \dots, 10$, we evaluated the bound for α_k in (31) numerically on the interval $\rho \in (-1, 0)$ (i.e. for $\vartheta(\bar{G}) \in (2, \infty)$). The numerical integration of (27) (to obtain $I(\rho)$) was done using the software MVNDST (for calculating multivariate normal probabilities) by Genz [10]. In each case the minimum value of the bound in (31) was obtained when $\rho = \frac{-1}{k-1}$, i.e. when $\vartheta(\bar{G}) = k$. The corresponding approximate numerical values of the bound on α_k ($k = 4, \dots, 10$) are shown in Table 2, together with values listed in the paper by Frieze and Jerrum [9]. (The value for $k = 3$ is also included, for completeness.)

k :	3	4	5	6	7	8	9	10
This paper:	0.836008	0.857487	0.876610	0.891543	0.903259	0.912664	0.920367	0.926788
From [9]:	0.832718	0.850304	0.874243	–	–	–	–	0.926642

Table 2: MAX- k -CUT approximation guarantees for $3 \leq k \leq 10$

Asymptotic analysis

Now we investigate the asymptotic behavior of (28) for $k \rightarrow \infty$ and $\rho < 0$. We show the following in Appendix 2.

Theorem 9.1 *Let $I(\rho)$ be defined by (28) and let $\rho \in (-1, 0)$. One has*

$$I(\rho) \sim \frac{\Gamma(\frac{1}{1+\rho})^2 (4\pi \ln(k-1))^{\frac{-\rho}{1+\rho}}}{\sqrt{1-\rho^2} (k-1)^{\frac{2}{1+\rho}}}, \quad (35)$$

when $k \rightarrow \infty$, where Γ denotes the gamma function.

If we substitute (35) into (31) and differentiate with respect to ρ , we find that the minimum in (31) is attained when $\rho = -1/(k-1)$, i.e. for graphs G with $\vartheta(\bar{G}) = k$. In other words, we have shown that Conjecture 9.1 holds asymptotically as $k \rightarrow \infty$.

If $\rho = -1/(k-1)$ is easy to show that the performance guarantee becomes

$$\alpha_k \sim 1 - k^{\frac{-k}{k-2}} \sim 1 - \frac{1}{k} + \frac{2 \ln k}{k^2},$$

as $k \rightarrow \infty$. The $\alpha_k \sim 1 - \frac{1}{k} + \frac{2 \ln k}{k^2}$ result was first shown by Frieze and Jerrum [9].

10 Approximate colouring of κ -colourable graphs

We now turn our attention to the related problem of *approximate graph colouring*, i.e. how many colours are needed to give a legal colouring of a graph $G = (V, E)$ with $\chi(G) = \kappa$ in polynomial time. We will use Theorem 9.1 to derive the following result due to Karger *et al.* [18].

Theorem 10.1 (Karger *et al.* [18]) *A κ -colourable graph with maximum degree Δ can be legally coloured in polynomial time using $\tilde{O}(\Delta^{1-2/\kappa})$ colours.*

The first step is to find a sufficiently large value of k (number of colours) so that a semicolouring⁶ is obtained (with high probability) when running Algorithm 1.

As shown by Karger *et al.* using a simple bisection argument, $O(\log n)$ successive semi-colourings (each using k colours) suffice to give a legal colouring of the graph.

The total number of edges is upper bounded by $|E| \leq \frac{1}{2}n\Delta$, where Δ is the largest degree of any node (valency).

By Theorem 9.1, the expected number of defect edges is therefore bounded via:

$$kp_1|E| = kI(\rho)|E| \leq \frac{1}{2}kI(\rho)n\Delta \sim \frac{1}{2}k^{\frac{\rho-1}{\rho+1}}n\Delta.$$

A semi-colouring will be obtained with high probability if the number of defect edges is at most $\frac{1}{2}n$. This will be true if

$$k \geq \Delta^{\frac{1+\rho}{1-\rho}}.$$

For κ -colourable graphs we have $\vartheta(\bar{G}) \leq \kappa$ and consequently $\rho := \frac{-1}{1-\vartheta(\bar{G})} \leq -1/(\kappa-1)$, and consequently we obtain a semi-colouring using $O(\Delta^{1-2/\kappa})$ colours, and subsequently a complete colouring using $\tilde{O}(\Delta^{1-2/\kappa})$ colours. This was the main result proved by Karger *et al.* in [18].

⁶An assignment of colours to at least half the nodes of G without any defect edges is called a semi-colouring in [18]. An assignment of colours to all the nodes of G such that at most $\frac{1}{2}n$ edges are defect yields a semi-colouring, by simply removing the colour from one endpoint of each defect edge.

11 Conclusions

The analysis in this paper linked different approaches where semidefinite programming is used for approximate graph colouring and MAX- k -CUT algorithms. We have shown that the ϑ -function is the common denominator in all these approaches. In particular, we have shown how to compute the MAX- k -CUT approximation guarantees for the algorithm by Frieze and Jerrum [9] more precisely for (small) fixed values of k , by employing the lifting procedure of De Klerk *et al.* [6]. This lifting procedure arises naturally from the coding of graph colouring as a satisfiability problem. We have also shown that one can derive both the results of by Jerrum and Frieze [9] and the results by Karger *et al.* [18] in a unified way, by using a single theorem on the asymptotic behaviour of the rounding scheme in Jerrum-Frieze MAX- k -CUT algorithm for large k .

Acknowledgements

The authors thank Uri Zwick for making his code for calculating the volume of spherical tetrahedrons available to them and for valuable comments on a draft version of this paper, and thank Monique Laurent, Michel Goemans, David Williamson, Michel Dekking and Ludolf Meester for useful comments. The second author thanks Viktor L. Ginzburg for a helpful discussion. Finally, the authors are indebted to Vladimir Koulikov for verifying the result of Theorem 9.1, to Piet Groeneboom for generously providing two alternative proofs of it, and to Vladimir Rotar for kindly supplying a preprint version of [24].

A part of the research was completed while the second author held positions at the Dept. of Computer Science, Utrecht University, and at the Dept. of Technical Mathematics and Informatics at Delft University of Technology.

Appendix 1

Here we show that the integral in (28) can be solved analytically if $k = 3$ and $\rho < 0$.

In what follows we denote $\lambda = k - 1$, and

$$\phi(x) = \frac{\exp(-x^2/2)}{\sqrt{2\pi}}$$

denotes the density function of the standard normal univariate distribution,

$$\alpha = \frac{1 + \sqrt{1 - \rho^2}}{\rho},$$

and

$$J = \sqrt{\frac{\rho}{2\alpha}} \begin{pmatrix} 1 & \alpha \\ \alpha & 1 \end{pmatrix}, \quad |\det J| = \sqrt{1 - \rho^2}.$$

Setting

$$\begin{pmatrix} x \\ y \end{pmatrix} = J \begin{pmatrix} s \\ t \end{pmatrix}$$

leads to

$$f(x, y, \rho) = \frac{\phi(s)\phi(t)}{\sqrt{1 - \rho^2}},$$

where f is defined in (29). Denoting by C_ρ the image of the negative orthant of \mathbb{R}^2 under J , one has

$$I(\rho) = \frac{|J|^{\lambda+1}}{\sqrt{1 - \rho^2}^{\lambda+1}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(x)\phi(y) \left(\iint_{C_\rho} \phi(u+x)\phi(v+y) du dv \right)^\lambda dx dy,$$

and by Fubini theorem

$$I(\rho) = \underbrace{\iint_{C_\rho} \cdots \iint_{C_\rho}}_{\lambda \text{ times}} Q(x, u) Q(y, v) du_1 dv_1 \dots du_\lambda dv_\lambda,$$

where

$$\begin{aligned}
Q(w) &= Q(z, w) = \int_{-\infty}^{\infty} \phi(z) \prod_{i=1}^{\lambda} \phi(w_i + z) dz = \\
&= \exp\left(\frac{(\sum_{i=1}^{\lambda} w_i)^2}{2(\lambda+1)}\right) \prod_{i=1}^{\lambda} \phi(w_i) \int_{-\infty}^{\infty} \phi\left(z\sqrt{\lambda+1} + \frac{\sum_{i=1}^{\lambda} w_i}{\sqrt{\lambda+1}}\right) dz = \\
&= \frac{1}{\sqrt{\lambda+1}} \exp\left(\frac{(\sum_{i=1}^{\lambda} w_i)^2}{2(\lambda+1)}\right) \prod_{i=1}^{\lambda} \phi(w_i).
\end{aligned}$$

Changing to polar coordinates for each cone $C_\rho = C_\rho(u_i, v_i)$, as follows: $v_i = r_i \sin \omega_i$, $u_i = r_i \cos \omega_i$, one obtains

$$\begin{aligned}
I(\rho) &= \frac{1}{(\lambda+1)\sqrt{2\pi}^\lambda} \underbrace{\int_0^{\omega_t} \int_{\omega_0}^{\omega_t} \dots \int_0^{\omega_t} \int_{\omega_0}^{\omega_t}}_{\lambda \text{ times}} e^{\frac{\sum_{i=1}^{\lambda} \sum_{j=1}^{\lambda} r_i r_j (\cos \omega_i \cos \omega_j + \sin \omega_i \sin \omega_j)}{2(\lambda+1)}} \prod_{i=1}^{\lambda} r_i \phi(r_i) dr_1 d\omega_1 \dots dr_\lambda d\omega_\lambda \\
&= \frac{(\lambda+1)^{\lambda-1}}{\sqrt{2\pi}^\lambda \lambda^\lambda} \underbrace{\int_0^{\omega_t} \int_{\omega_0}^{\omega_t} \dots \int_0^{\omega_t} \int_{\omega_0}^{\omega_t}}_{\lambda \text{ times}} e^{\frac{\sum_{i=1}^{\lambda-1} \sum_{j=i+1}^{\lambda} r_i r_j \cos(\omega_i - \omega_j)}{\lambda}} \prod_{i=1}^{\lambda} r_i \phi(r_i) dr_1 d\omega_1 \dots dr_\lambda d\omega_\lambda,
\end{aligned}$$

where $\omega_0 = \pi + \omega$, $\omega_t = 3\pi/2 - \omega$, and $\omega = \arctan(-1/\alpha)$, as can be observed by applying $(J^{-1})^T$ to the extreme rays of the negative orthant. Substituting $\omega_i = \eta_i + \omega_0$ and observing that $\omega_t - \omega_0 = \pi/2 - 2\omega = \arccos(-\rho)$, one obtains

$$I(\rho) = \frac{(\lambda+1)^{\lambda-1}}{\sqrt{2\pi}^\lambda \lambda^\lambda} \underbrace{\int_0^{\arccos(-\rho)} \dots \int_0^{\arccos(-\rho)}}_{\lambda \text{ times}} \underbrace{\int_0^{\infty} \dots \int_0^{\infty}}_{\lambda \text{ times}} e^{\frac{\sum_{i=1}^{\lambda-1} \sum_{j=i+1}^{\lambda} r_i r_j \cos(\eta_i - \eta_j)}{\lambda}} \prod_{i=1}^{\lambda} r_i \phi(r_i) dr_1 \dots dr_\lambda d\eta_1 \dots d\eta_\lambda. \quad (36)$$

In particular, for $\lambda = 2$ one obtains

$$\begin{aligned}
I(\rho) &= \frac{3}{8\pi} \int_0^{\arccos(-\rho)} \int_0^{\arccos(-\rho)} K(\cos(\omega_1 - \omega_2)) d\omega_1 d\omega_2 = \\
&= \frac{3(G(\pi/2 - \omega) - G(0))}{4\pi},
\end{aligned}$$

where

$$K(z) = \int_0^{\infty} \int_0^{\infty} e^{\frac{r_1 r_2 z}{2}} r_1 r_2 \phi(r_1) \phi(r_2) dr_1 dr_2, \quad G''(t) = K(\cos t).$$

Applying the standard change to polar coordinates, one obtains

$$\begin{aligned}
K(z) &= \frac{1}{2\pi} \int_0^{\pi/2} \sin \gamma \cos \gamma \int_0^{\infty} R^3 \exp(-R^2(1 - z \sin \gamma \cos \gamma)/2) dR d\gamma = \\
&= \frac{1}{2\pi} \int_0^{\pi/2} \frac{\sin 2\gamma d\gamma}{(1 - \frac{z}{2} \sin 2\gamma)^2} = \\
&= (4 - z^2)^{-3/2} (2\sqrt{4 - z^2} - 2z \arctan \frac{-z}{\sqrt{4 - z^2}} + \pi z) / \pi = \\
&= \frac{2}{\pi} \left(\frac{z \arccos(-z/2)}{\sqrt{4 - z^2}^3} + \frac{1}{4 - z^2} \right).
\end{aligned}$$

Computing the double indefinite integral, one obtains

$$G(t) = \frac{t^2 - \arccos^2(-\cos t/2)}{3\pi}.$$

Thus from (36) one obtains

$$I(\rho) = \frac{1}{9} + \frac{\arccos^2(-\rho) - \arccos^2(\rho/2)}{4\pi^2}.$$

Appendix 2: Proof of Theorem 9.1

In this appendix we give a proof of Theorem 9.1. The method we use closely resembles the one of [24], where an asymptotic expression for a similar univariate integral is derived.

We denote the standard normal univariate cumulative distribution by

$$\Phi(x) = \int_{-\infty}^x \phi(t)dt, \quad (37)$$

where

$$\phi(x) = \frac{\exp(-x^2/2)}{\sqrt{2\pi}}$$

is its density function, as before.

Note that F in (30) and Φ are connected as follows (see e.g. [22, Sect. 46.4]).

$$F(\infty, y, \rho) = \Phi(y); \quad F(x, \infty, \rho) = \Phi(x). \quad (38)$$

It is well-known (see e.g. [8]) that as $x \rightarrow \infty$:

$$1 - \Phi(x) \sim \frac{\phi(x)}{x}. \quad (39)$$

Similarly, one has the following.

Lemma 11.1

$$\int_x^\infty \int_y^\infty f(\xi, \eta, \rho) d\eta d\xi \sim \frac{(1-\rho^2)^2 f(x, y, \rho)}{-\rho(x^2 + y^2) + (1+\rho^2)xy} \quad (40)$$

for $\rho < 0$ as $x \rightarrow \infty$ and $y \rightarrow \infty$.

Proof: Substituting $\xi = x + u/x$, $\eta = y + v/y$ gives

$$J = \int_x^\infty \int_y^\infty f(\xi, \eta, \rho) d\eta d\xi = \frac{f(x, y, \rho)}{xy} \int_0^\infty \int_0^\infty e^{-q(u, v)} du dv,$$

where $q(u, v) = \frac{u(1-\rho y/x) + v(1-\rho x/y)}{1-\rho^2} + \frac{u^2/x^2 + v^2/y^2 - 2\rho uv/(xy)}{2(1-\rho^2)}$.

As $q(u, v) \sim \frac{u(1-\rho y/x) + v(1-\rho x/y)}{1-\rho^2}$, one obtains

$$J \sim \frac{f(x, y, \rho)}{xy} \int_0^\infty \int_0^\infty e^{-\frac{u(1-\rho y/x) + v(1-\rho x/y)}{1-\rho^2}} du dv = \frac{f(x, y, \rho)(1-\rho^2)^2}{xy(1-\rho y/x)(1-\rho x/y)}.$$

Finally, note that the denominator in the last expression equals $-\rho(x^2 + y^2) + (1+\rho^2)xy$. \square

Lemma 11.2

$$-\ln F(x, y, \rho) \sim 1 - F(x, y, \rho) \sim \frac{\phi(x)}{x} + \frac{\phi(y)}{y}, \quad (41)$$

for $\rho < 0$ as $x \rightarrow \infty$ and $y \rightarrow \infty$.

Proof: The first equivalence follows immediately from the fact that

$$\lim_{t \rightarrow 1} \frac{-\ln t}{1-t} = 1.$$

To establish the second equivalence, we write

$$\begin{aligned} F(x, y, \rho) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f d\eta d\xi - \int_{-\infty}^{\infty} \int_y^{\infty} f d\eta d\xi - \int_x^{\infty} \int_{-\infty}^{\infty} f d\eta d\xi + \int_x^{\infty} \int_y^{\infty} f d\eta d\xi = \\ &= 1 - (1 - F(\infty, y, \rho)) - (1 - F(x, \infty, \rho)) + \int_x^{\infty} \int_y^{\infty} f d\eta d\xi. \end{aligned}$$

Using (38), (39), and (40), we derive

$$\begin{aligned} F(x, y, \rho) &= 1 - (1 - \Phi(y)) - (1 - \Phi(x)) + \int_x^{\infty} \int_y^{\infty} f d\eta d\xi \sim \\ &\sim 1 - \frac{\phi(x)}{x} - \frac{\phi(y)}{y} + \frac{(1 - \rho^2)^2 f(x, y, \rho)}{-\rho(x^2 + y^2) + (1 + \rho^2)xy} \sim \\ &\sim 1 - \frac{\phi(x)}{x} - \frac{\phi(y)}{y}, \end{aligned}$$

as required. \square

The following is straightforward to check.

Lemma 11.3 *Let $\lambda := k - 1$ and $\Lambda = \lambda(1 + \rho)/\sqrt{2\pi}$ and $\tilde{x} \sim \sqrt{2 \ln \Lambda - \ln(2 \ln \Lambda)}$. Then one has*

$$\frac{\lambda \phi(\tilde{x})}{\tilde{x}} \sim \frac{1}{1 + \rho} \quad (42)$$

as $\lambda \rightarrow \infty$. \square

We note that one can show that $(x^*, y^*) = (\tilde{x}, \tilde{x})$ approximates a maximum of $f(x, y, \rho)F(x, y, \rho)^\lambda$, although we do not use this fact.

Next, we compute the integral $I_1 = \int_{\tilde{x}-1}^{\infty} \int_{\tilde{x}-1}^{\infty} f(u, v, \rho)F(u, v, \rho)^\lambda dudv$, that turns out to be asymptotically the main contribution to $I(\rho)$. Substituting $u = \tilde{x} + x/\tilde{x}$ and $v = \tilde{y} + y/\tilde{x}$, one obtains

$$I_1 = \frac{\phi(\frac{\tilde{x}}{\sqrt{1+\rho}})^2}{\tilde{x}^2 \sqrt{1-\rho^2}} \int_{-\tilde{x}-\tilde{x}}^{\infty} \int_{-\tilde{x}-\tilde{x}}^{\infty} e^{\lambda \ln F(\tilde{x} + \frac{x}{\tilde{x}}, \tilde{y} + \frac{y}{\tilde{x}}, \rho) - \frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)\tilde{x}^2} - \frac{x+y}{1+\rho}} dx dy. \quad (43)$$

As λ grows, the integral in (43) converges to an integral that is not dependent on λ .

Lemma 11.4 *As $\lambda \rightarrow \infty$,*

$$I_1 \sim \frac{\phi(\frac{\tilde{x}}{\sqrt{1+\rho}})^2}{\tilde{x}^2 \sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(-\frac{e^{-x} + e^{-y} + x + y}{1 + \rho}\right) dx dy. \quad (44)$$

Proof: As $\lambda \rightarrow \infty$,

$$\begin{aligned} -\lambda \ln F\left(\tilde{x} + \frac{x}{\tilde{x}}, \tilde{y} + \frac{y}{\tilde{x}}, \rho\right) &\sim \lambda \left(\frac{\phi(\tilde{x} + \frac{x}{\tilde{x}})}{\tilde{x} + \frac{x}{\tilde{x}}} + \frac{\phi(\tilde{x} + \frac{y}{\tilde{x}})}{\tilde{x} + \frac{y}{\tilde{x}}} \right) \sim \\ &\sim \frac{\lambda \phi(\tilde{x})}{\tilde{x}} (e^{-x} + e^{-y}) \sim \frac{e^{-x} + e^{-y}}{1 + \rho}. \end{aligned}$$

Hence for $\lambda \gg 0$ the integrand in (43) is equivalent to the integrand in (44).

Next, we show that on the domain of integration the integrand in (43) is bounded from above by an integrable function for all $\lambda \geq \lambda_0$, for some constant λ_0 depending only on ρ .

For nonnegative x and y , integrand in (43) is bounded from above by $e^{-\frac{x+y}{1+\rho}}$.

Now let $\tilde{x} \leq x \leq 0$. As $\ln F \leq F - 1$, using (41) (which is valid as $\tilde{x} + \frac{x}{\tilde{x}} \geq \tilde{x} - 1$ and $\tilde{x} + \frac{y}{\tilde{x}} \geq \tilde{x} - 1$) gives

$$\begin{aligned} \lambda \ln F\left(\tilde{x} + \frac{x}{\tilde{x}}, \tilde{y} + \frac{y}{\tilde{x}}, \rho\right) &\leq -\frac{\lambda \phi\left(\tilde{x} + \frac{x}{\tilde{x}}\right)}{2\left(\tilde{x} + \frac{x}{\tilde{x}}\right)} - \frac{\lambda \phi\left(\tilde{x} + \frac{y}{\tilde{x}}\right)}{2\left(\tilde{x} + \frac{y}{\tilde{x}}\right)} \leq \\ &\leq -\frac{\lambda \phi(\tilde{x})}{2\tilde{x}} e^{-x-x^2/(2\tilde{x}^2)} \leq \frac{-e^{-x}}{1+\rho}. \end{aligned}$$

Using this, one sees that, when $y \geq 0$, the integrand in (43) is bounded from above by the function $\exp\left(\frac{-e^{-x}-x-y}{1+\rho}\right)$ which is integrable on the upper half-plane.

Similarly, the bound $\lambda \ln F\left(\tilde{x} + \frac{x}{\tilde{x}}, \tilde{y} + \frac{y}{\tilde{x}}, \rho\right) \leq \frac{-e^{-y}}{1+\rho}$ that is valid for $\tilde{x} \leq y \leq 0$ implies that, when $x \geq 0$, the integrand in (43) is bounded from above by the function $\exp\left(\frac{-e^{-y}-x-y}{1+\rho}\right)$ which is integrable on the right halfplane (*i.e.* for $x \geq 0$).

Again, a similar argument shows that for $\tilde{x} \leq x \leq 0$ and $\tilde{x} \leq y \leq 0$ one has $\lambda \ln F\left(\tilde{x} + \frac{x}{\tilde{x}}, \tilde{y} + \frac{y}{\tilde{x}}, \rho\right) \leq -e^{-x} - e^{-y}$, thus bounding the integrand in (43) by $e^{(-e^{-x}-e^{-y}-x+y)/(1+\rho)}$ which is integrable everywhere.

Now an application of Lebesgue dominated convergence theorem establishes the equivalence of I_1 and the RHS of (44) with the lower limits of the integrals replaced by $-\tilde{x}$. It remains to analyze the ‘tail’ of the latter

integral. For instance, it suffices to show that $K = \int_{-\infty}^{\infty} \int_{-\infty}^{-\tilde{x}} e^{-\frac{e^{-x}+e^{-y}+x+y}{1+\rho}} dx dy$ is asymptotically negligibly small compared to $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{e^{-x}+e^{-y}+x+y}{1+\rho}} dx dy$. Indeed, $K = O(e^{-e^{\tilde{x}}})$, that converges to 0, as $\lambda \rightarrow \infty$. \square

Now let us show that I_1 provides the main contribution to I .

Lemma 11.5 *As $\lambda \rightarrow \infty$, one has $I \sim I_1$.*

Proof: Note that

$$\begin{aligned} -\lambda \ln F(\tilde{x} - 1, \infty, \rho) &= -\lambda \ln \Phi(\tilde{x} - 1) \sim \lambda \frac{\phi(\tilde{x} - 1)}{\tilde{x}} \sim \\ &\sim \sqrt{\frac{2\pi}{e}} e^{\tilde{x}} \frac{\lambda \phi(\tilde{x})}{\tilde{x}} \sim \sqrt{\frac{2\pi}{e}} e^{\tilde{x}} / (1 + \rho). \end{aligned} \tag{45}$$

We apply this estimate to obtain an upper bound on

$$I_2 = \int_{-\infty}^{\tilde{x}-1} \int_{-\infty}^{\infty} f(u, v, \rho) F(u, v, \rho)^\lambda dudv.$$

We write

$$I_2 \leq \int_{-\infty}^{\tilde{x}-1} \int_{-\infty}^{\infty} f(u, v, \rho) F(\tilde{x} - 1, \infty, \rho)^\lambda dudv = \exp(\lambda \ln \Phi(\tilde{x} - 1)) C',$$

where C' is a function of ρ that does not depend on λ . In view of (45),

$$I_2 \leq \exp\left(-\sqrt{\frac{2\pi}{e}} e^{\tilde{x}}\right) C'$$

for sufficiently large λ . Now (44) implies that $\lim_{\lambda \rightarrow \infty} I_2/I_1 = 0$.

The lemma follows from the observation (look at the limits of integration) that $I \leq I_1 + 2I_2$. \square

To complete the proof of the theorem, it remains to observe that the integral in (44) is the square of the integral $\int_{-\infty}^{\infty} \exp\left(-\frac{e^{-t}+t}{1+\rho}\right) dt$ that can be shown to be equal to $\Gamma(1/(1+\rho))(1+\rho)^{1/(1+\rho)}$ (for instance, by substituting $e^{-t} = z$), and make straightforward substitutions and asymptotic cancellations in (44).

References

- [1] D.V. Alekseevskij, E.B. Vinberg, and A.S. Solodovnikov. Geometry of spaces of constant curvature. In E.B. Vinberg, editor, *Geometry II*, volume 29 of *Encyclopedia of Mathematical Sciences*. Springer-Verlag, 1993.
- [2] N. Alon and N. Kahale. Approximating the independence number via the θ -function. *Mathematical Programming*, 80:253–264, 1998.
- [3] G. Andersson, L. Englbjergsen, and J. Hastad. A new way of using semidefinite programming with applications to linear equations mod p . *Journal of Algorithms*, 39:162–204, 2001.
- [4] R. Beigel and D. Eppstein. 3-Coloring in time $O(1.3446^n)$: a no-MIS algorithm. In *Proc. 36th IEEE Symp. Foundations of Comp. Sci.*, pages 444–453, 1995.
- [5] J.-D. Cho and M. Sarrafzadeh. Fast approximation algorithms on Maxcut, k -Coloring and k -Color ordering for VLSI applications. *IEEE Transactions on Computers*, 47(11):1253–1266, 1998.
- [6] E. de Klerk, H. van Maaren, and J. P. Warners. Relaxations of the satisfiability problem using semidefinite programming. *Journal of automated reasoning*, 24:37–65, 2000.
- [7] U. Feige and M. Goemans. Approximating the value of two prover proof systems with applications to MAX 2SAT and MAX DICUT. In *Proc. Third Israel Symposium on Theory of Computing and Systems*, pages 182–189, 1995.
- [8] William Feller. *An introduction to probability theory and its applications. Vol. I*. John Wiley & Sons Inc., New York, third edition, 1968.
- [9] A. Frieze and M. Jerrum. Improved approximation algorithms for MAX k -cut and MAX BISECTION. *Algorithmica*, 18:61–77, 1997.
- [10] A. Genz. Numerical computation of multivariate normal probabilities. *J. Comp. Graph. Stat.*, 1:141–149, 1992.
- [11] Michel X. Goemans. Semidefinite programming in combinatorial optimization. *Math. Programming*, 79(1-3, Ser. B):143–161, 1997. Lectures on mathematical programming (ismp97) (Lausanne, 1997).
- [12] M.X. Goemans and D.P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995.
- [13] M.X. Goemans and D.P. Williamson. Approximation algorithms for MAX 3-CUT and other problems via complex semidefinite programming. In *Proc. 33rd STOC*, pages 443–452. ACM, 2001.
- [14] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Springer-Verlag, Berlin, 1988.
- [15] E. Halperin and U. Zwick. Approximation algorithms for MAX 4-SAT and rounding procedures for semidefinite programs. In *Proc. 7th IPCO*, pages 202–217, 1999.
- [16] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.
- [17] 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), June 1997.
- [18] D. Karger, R. Motwani, and M. Sudan. Approximate graph coloring by semidefinite programming. *J. ACM*, 45(2):246–265, 1998.
- [19] H. Karloff and U. Zwick. A $7/8$ -approximation algorithm for MAX 3SAT? In *Proc. 38th FOCS*, pages 406–415, 1997.
- [20] S. Khanna, N. Linial, and S. Safra. On the hardness of approximating the chromatic number. *Combinatorica*, 20:393–415, 2000.
- [21] D.E. Knuth. The sandwich theorem. *The Electronic Journal of Combinatorics*, 1:1–48, 1994.
- [22] Samuel Kotz, N. Balakrishnan, and Norman L. Johnson. *Continuous Multivariate Distributions*, volume 1: Models and Applications of *Wiley Series in Probability and Statistics*. John Wiley & Sons Inc., New York, 2nd edition, 2000.
- [23] L. Lovász. On the Shannon capacity of a graph. *IEEE Trans. on Information Theory*, 25:1–7, 1979.
- [24] Y. Rinott and V. Rotar. A remark on quadrant normal probabilities in high dimensions. *Statistics and Probability Letters*, 51(1):47–51, 2001.
- [25] Alexander Schrijver. A comparison of the Delsarte and Lovász bounds. *IEEE Trans. Inform. Theory*, 25(4):425–429, 1979.
- [26] U. Zwick. Outward rotations: a new tool for rounding solutions of semidefinite programming relaxations, with applications to MAX CUT and other problems. In *Proc. 31st STOC*, pages 679–687, 1999.

