On the Integrality Gap of the Maximum-Cut Semidefinite Programming Relaxation in Fixed Dimension

Fernando Mário de Oliveira Filho     Frank Vallentin*

Received 9 August 2018; Published 6 August 2020

Abstract: We describe a factor-revealing convex optimization problem for the integrality gap of the maximum-cut semidefinite programming relaxation: for each $n \geq 2$ we present a convex optimization problem whose optimal value is the largest possible ratio between the value of an optimal rank-$n$ solution to the relaxation and the value of an optimal cut. This problem is then used to compute lower bounds for the integrality gap.

Key words and phrases: maximum-cut problem, semidefinite programming, integrality gap

1 Introduction

For $x, y \in \mathbb{R}^n$, write $x \cdot y = x_1y_1 + \cdots + x_ny_n$ for the Euclidean inner product. Let $S^{n-1} = \{x \in \mathbb{R}^n : x \cdot x = 1\}$ be the $(n-1)$-dimensional unit sphere. Given a nonempty finite set $V$, a nonnegative matrix $A \in \mathbb{R}^{V \times V}$, and an integer $n \geq 1$, write

$$\text{SDP}_{n}(A) = \max \left\{ \sum_{x,y \in V} A(x,y)(1 - f(x) \cdot f(y)) : f : V \rightarrow S^{n-1} \right\}.$$ (1)

Replacing $S^{n-1}$ above by $S^\infty$, the set of all sequences $(a_k)$ such that $\sum_{k=0}^{\infty} a_k^2 = 1$, we obtain the definition of $\text{SDP}_{\infty}(A)$.

*This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie agreement No 764759. The second author is partially supported by the SFB/TRR 191 “Symplectic Structures in Geometry, Algebra and Dynamics” and by the project “Spectral bounds in extremal discrete geometry” (project number 414898050), both funded by the DFG.

© 2020 Fernando Mário de Oliveira Filho and Frank Vallentin
Licensed under a Creative Commons Attribution License (CC-BY)
Given a finite (loopless) graph $G = (V,E)$ and a nonnegative weight function $w: E \to \mathbb{R}_+$ on the edges of $G$, the maximum-cut problem asks for a set $S \subseteq V$ that maximizes the weight

$$\sum_{e \in \delta(S)} w(e) = \sum_{x \in S, y \in V \setminus S, xy \in E} w(xy)$$

of the cut $\delta(S) = \{ e \in E : |e \cap S| = 1 \}$. If $A: V \times V \to \mathbb{R}$ is the matrix such that $A(x,y) = w(xy)$ when $xy \in E$ and $A(x,y) = 0$ otherwise, then the weight of a maximum cut is $(1/4) \text{SDP}_1(A)$.

SDP$_n(A)$ is actually the optimal value of a semidefinite program with a rank constraint, namely

$$\max \sum_{x,y \in V} A(x,y)(1 - M(x,y))$$

$$M(x,x) = 1 \text{ for } x \in V,$$

$$M \in \mathbb{R}^{V \times V} \text{ is positive semidefinite and has rank at most } n.$$ (2)

In SDP$_\infty(A)$ the rank constraint is simply dropped. The optimization problem SDP$_\infty(A)$ is the semidefinite programming relaxation of the maximum-cut problem.

Obviously, SDP$_\infty(A) \geq \text{SDP}_1(A)$. In a fundamental paper, Goemans and Williamson [8] showed that, if $A$ is a nonnegative matrix, then

$$\text{SDP}_1(A) \geq \alpha_{GW} \text{SDP}_\infty(A),$$

where

$$\alpha_{GW} = \min_{t \in [-1,1]} \frac{1 - (2/\pi) \arcsin t}{1 - t} = 0.87856 \ldots.$$ 

The $n$-dimensional integrality gap of the semidefinite programming relaxation is

$$\gamma_n = \sup \left\{ \frac{\text{SDP}_n(A)}{\text{SDP}_1(A)} : A \text{ is a nonnegative matrix} \right\},$$

but it is often more natural to work with its reciprocal $\alpha_n = \gamma_n^{-1}$. Goemans and Williamson thus showed that $\alpha_n \geq \alpha_{GW}$; Feige and Schechtman [7] later showed that $\alpha_n = \alpha_{GW}$ (see also §8.3 in Gärtner and Matoušek [9]).

In dimension 2 it is known that

$$\alpha_2 = \frac{32}{25 + 5\sqrt{5}} = 0.88445 \ldots.$$ (3)

The ‘$\leq$’ direction was shown by Delorme and Poljak [4, 5]; the ‘$\geq$’ direction was shown by Goemans in an unpublished note (cf. Avidor and Zwick [2], who also provide another proof of this result). Avidor and Zwick [2] showed that $\alpha_3 \geq 0.8818$. Except for $n = 2$ and 3, it is an open problem whether $\alpha_n > \alpha_\infty = \alpha_{GW}$. 

**DISCRETE ANALYSIS, 2020:10, 17pp.**
1.1 A factor-revealing optimization problem

Theorem 1.1 below gives a factor-revealing optimization problem for \( \alpha_n \); an optimization problem defined for each \( n \geq 2 \) whose optimal value is \( \alpha_n \). Relaxations of it can be solved by computer to give upper bounds for \( \alpha_n \), as done in §4.

For a finite and nonempty set \( U \), write

\[
\text{CUT}_\square(U) = \text{conv}\{ f \otimes f^\ast : f : U \to \{-1, 1\}\}
\]

where \( f \otimes f^\ast \) is the external product of the vector \( f \), that is, the matrix whose entry \((x, y)\) is \( f(x)f(y) \). This set is known as the cut polytope and was extensively investigated [6].

A kernel is a square-integrable (with respect to the Lebesgue measure) real-valued function on \( S^{n-1} \times S^{n-1} \); the set of continuous kernels is denoted by \( C(S^{n-1} \times S^{n-1}) \). Write

\[
\text{CUT}_\square(S^{n-1}) = \{ K \in C(S^{n-1} \times S^{n-1}) : (K(x,y))_{x,y \in U} \in \text{CUT}_\square(U) \}
\]

for every finite and nonempty \( U \subseteq S^{n-1} \).

In principle, it is not clear that anything other than the constant 1 kernel is in \( \text{CUT}_\square(S^{n-1}) \). If \( f : S^{n-1} \to \{-1, 1\} \) is any nonconstant measurable function, then one could be tempted to say that \((x, y) \mapsto f(x)f(y) \) is in \( \text{CUT}_\square(S^{n-1}) \), but no such kernel is continuous, so to see that \( \text{CUT}_\square(S^{n-1}) \) is nontrivial requires a more complicated argument. Fix \( e \in S^{n-1} \) and let \( f(x) \) be 1 if \( e \cdot x \geq 0 \) and \(-1 \) otherwise; let \( K(x,y) = (2/\pi) \arcsin e \cdot y \) for every \( x, y \in S^{n-1} \). Grothendieck’s identity says that

\[
K(x,y) = \int_{O(n)} f(Tx)f(Ty) \, d\mu(T),
\]

where \( O(n) \) is the group of \( n \times n \) orthogonal matrices and \( \mu \) is the Haar measure on \( O(n) \) normalized so the total measure is 1. Then it is easy to see that \( K \) is continuous and that it belongs to \( \text{CUT}_\square(S^{n-1}) \). This kernel was first considered by McMillan [12], who was perhaps the first to use such an infinite-dimensional analogue of the cut polytope.

We say that a kernel \( K \) is invariant if \( K(Tx, Ty) = K(x, y) \) for all \( T \in O(n) \) and \( x, y \in S^{n-1} \). An invariant kernel is in fact a univariate function, since the value of \( K(x, y) \) depends only on the inner product \( x \cdot y \). Hence for \( t \in [-1, 1] \) we write \( K(t) \) for the common value taken by \( K \) on pairs \((x, y)\) with inner product \( t \).

**Theorem 1.1.** If \( n \geq 2 \), then \( \alpha_n \) is the optimal value of

\[
\sup \alpha \quad 1 - K(t) \geq \alpha(1-t) \quad \text{for all } t \in [-1, 1],
\]

(4)

\( K \in \text{CUT}_\square(S^{n-1}) \) is invariant.

This theorem is similar to the integral representation for the Grothendieck constant [14, Theorem 3.4]. The easy direction is to show that the optimal value of (4) is at most \( \alpha_n \).
Proof of the easy direction of Theorem 1.1. Let \((K, \alpha)\) be a feasible solution of (4) and let \(A \in \mathbb{R}^{V \times V}\) be any nonnegative matrix and \(f : V \to S^{n-1}\) be a function achieving the maximum in \(\text{SDP}_n(A)\). Note \((K(f(x), f(y)))_{x, y \in V} \in \text{CUT}_{\square}(V)\). This implies that there are nonnegative numbers \(\lambda_1, \ldots, \lambda_r\) that sum up to 1 and functions \(f_1, \ldots, f_r : V \to \{-1, 1\}\) such that
\[
K(f(x), f(y)) = K(f(x) \cdot f(y)) = \lambda_1 f_1(x)f_1(y) + \cdots + \lambda_r f_r(x)f_r(y)
\]
for all \(x, y \in V\). But then
\[
\text{SDP}_1(A) \geq \sum_{k=1}^{r} \lambda_k \sum_{x, y \in V} A(x, y)(1 - f_k(x)f_k(y))
\]
\[
= \sum_{x, y \in V} A(x, y)(1 - K(f(x) \cdot f(y)))
\]
\[
\geq \alpha \sum_{x, y \in V} A(x, y)(1 - f(x) \cdot f(x))
\]
\[
= \alpha \text{SDP}_n(A),
\]
so \(\alpha \leq \alpha_n\). \(\square\)

A proof that the optimal value of (4) is at least \(\alpha_n\) is given in §2, but it works only for \(n \geq 3\). For \(n = 2\) a direct proof can be given by showing a feasible solution of (4) with objective value \(\alpha_2\); this was done, in a different language, by Avidor and Zwick [2] and is outlined in §3, where a short discussion on how lower bounds for \(\alpha_n\) can be found is also presented.

Notice that the optimization problem (4) is infinite: the kernel \(K\) lies in an infinite-dimensional space and must satisfy infinitely many constraints, not to mention that the separation problem over \(\text{CUT}_{\square}(U)\) is NP-hard since the maximum-cut problem is NP-hard [10]. In §4 we will see how \(K\) can be parameterized and how the problem can be relaxed (by relaxing the constraint that \(K\) must be in \(\text{CUT}_{\square}(S^{n-1})\)) and effectively discretized so it can be solved by computer, providing us with upper bounds for \(\alpha_n\). From feasible solutions of this relaxation, instances with large integrality gap can be constructed, as shown in §4.1.

2 Proof of Theorem 1.1 for \(n \geq 3\)

The difficult part of the proof is to show that the optimal value of (4) is at least \(\alpha_n\). This is done here for \(n \geq 3\), and for this we need a few lemmas.

Let \(\mu\) be the Haar measure on the orthogonal group \(O(n)\), normalized so the total measure is 1. The Reynolds operator \(R\) projects a kernel \(K\) onto the space of invariant kernels by averaging:
\[
R(K)(x, y) = \int_{O(n)} K(Tx, Ty) d\mu(T)
\]
for all \(x, y \in S^{n-1}\). If \(K\) is a continuous kernel, then so is \(R(K)\) [3, Lemma 5.4], and if \(f \in L^2(S^{n-1})\), then \(R(f \otimes f^*)\) is continuous [3, Lemma 5.5], where \(f \otimes f^*\) is the kernel mapping \((x, y)\) to \(f(x)f(y)\).

A function \(f : S^{n-1} \to \mathbb{R}\) respects a partition \(\mathcal{P}\) of \(S^{n-1}\) if \(f\) is constant on each \(X \in \mathcal{P}\); we write \(f(X)\) for the common value of \(f\) in \(X\).
Lemma 2.1. If \( n \geq 2 \), then for every \( \eta > 0 \) there is a partition \( \mathcal{P} \) of \( S^{n-1} \) into finitely many measurable sets such that for every finite set \( I \subseteq [-1,1] \) and every nonnegative function \( z : I \to \mathbb{R} \) there is a function \( f : S^{n-1} \to \{-1,1\} \) that respects \( \mathcal{P} \) and satisfies

\[
\sum_{t \in I} z(t)(1 - R(f \otimes f^*))(t) \geq \sum_{t \in I} z(t)(\alpha_n(1 - t) - \eta). \tag{5}
\]

Proof. Let \( \mathcal{P} \) be any partition of \( S^{n-1} \) into finitely many measurable sets of small enough diameter such that for all \( X, Y \in \mathcal{P}, x, x' \in X \), and \( y, y' \in Y \), we have \( |x \cdot y - x' \cdot y'| \leq \alpha_n^{-1}\eta \). Such a partition can be obtained by considering e.g. the Voronoi cell of each point of an \( \mathcal{E} \)-net for \( S^{n-1} \) for small enough \( \mathcal{E} \).

For \( u \in S^{n-1} \) and \( X \in \mathcal{P} \), write

\[
[u,X] = \{ T \in O(n) : Tu \in X \}.
\]

Then \( [u,X] \) is measurable [11, Theorem 3.7], so \( \{ [u,X] : X \in \mathcal{P} \} \) is a partition of \( O(n) \) into measurable sets, and hence for any \( u, v \in S^{n-1} \) so is the common refinement

\[
\{ [u,X] \cap [v,Y] : (X,Y) \in \mathcal{P} \times \mathcal{P} \text{ and } [u,X] \cap [v,Y] \neq \emptyset \}.
\]

Write \( u = (1,0,\ldots,0) \in S^{n-1} \) and for \( t \in [-1,1] \) let \( v_t = (t, (1 - t^2)^{1/2}, 0, \ldots, 0) \), so \( u \cdot v_t = t \). If \( f : S^{n-1} \to \mathbb{R} \) respects \( \mathcal{P} \), then for every finite \( I \subseteq [-1,1] \) and every nonnegative \( z : I \to \mathbb{R} \) we have

\[
\sum_{t \in I} z(t)(1 - R(f \otimes f^*))(t) = \sum_{t \in I} z(t) \int_{O(n)} 1 - f(Tu)f(Tv_t) d\mu(T)
= \sum_{t \in I} z(t) \sum_{X,Y \in \mathcal{P}} \int_{[u,X] \cap [v_t,Y]} 1 - f(Tu)f(Tv_t) d\mu(T)
= \sum_{t \in I} z(t) \sum_{X,Y \in \mathcal{P}} (1 - f(X)f(Y)) \mu(\{ [u,X] \cap [v_t,Y] \})
= \sum_{X,Y \in \mathcal{P}} (1 - f(X)f(Y)) \sum_{t \in I} z(t) \mu(\{ [u,X] \cap [v_t,Y] \}).
\]

By considering the matrix \( A_{\mathcal{E}} : \mathcal{P} \times \mathcal{P} \to \mathbb{R} \) such that

\[
A_{\mathcal{E}}(X,Y) = \sum_{t \in I} z(t) \mu(\{ [u,X] \cap [v_t,Y] \}), \tag{6}
\]

we see that finding a function \( f : S^{n-1} \to \{-1,1\} \) that respects \( \mathcal{P} \) and maximizes the left-hand side of (5) is the same as finding an optimal solution of \( \text{SDP}_1(A_{\mathcal{E}}) \), so there is such a function \( f \) satisfying

\[
\sum_{t \in I} z(t)(1 - R(f \otimes f^*))(t) = \text{SDP}_1(A_{\mathcal{E}}). \tag{7}
\]

Now let \( g : \mathcal{P} \to S^{n-1} \) be such that \( g(X) = x \) for some \( x \in X \) chosen arbitrarily. Recall that the sets
in $\mathcal{P}$ have small diameter, so

\[
\text{SDP}_n(A_z) \geq \sum_{X,Y \in \mathcal{P}} A_z(X,Y)(1 - g(X) \cdot g(Y))
\]

\[
= \sum_{t \in I} z(t) \sum_{X,Y \in \mathcal{P}} (1 - g(X) \cdot g(Y))\mu([u,X] \cap [v_t,Y])
\]

\[
= \sum_{t \in I} z(t) \int_{[u,X] \cap [v_t,Y]} 1 - g(X) \cdot g(Y) \, d\mu(T)
\]

\[
\geq \sum_{t \in I} z(t) \int_{[u,X] \cap [v_t,Y]} 1 - (Tu) \cdot (Tv_t) - \alpha_n^{-1} \eta \, d\mu(T)
\]

\[
= \sum_{t \in I} z(t) \int_{[u,X] \cap [v_t,Y]} 1 - \alpha_n^{-1} \eta \, d\mu(T)
\]

\[
= \sum_{t \in I} z(t)((1 - \alpha_n^{-1} \eta).
\]

Now take any finite $I \subseteq [-1,1]$ and any nonnegative $z: I \to \mathbb{R}$. If $f$ is a function that respects $\mathcal{P}$ and for which (7) holds, then

\[
\sum_{t \in I} z(t)(1 - R(f \otimes f^*)(t)) = \text{SDP}_1(A_z) \geq \alpha_n \text{SDP}_n(A_z) \geq \sum_{t \in I} z(t)(\alpha_n(1 - t) - \eta),
\]

as we wanted. □

Lemma 2.1 is enough to show the following weaker version of the difficult direction of Theorem 1.1:

**Lemma 2.2.** If $n \geq 2$ and $1 \geq \delta > 0$, then the optimal value of the optimization problem

\[
\sup \alpha \\
1 - K(t) \geq \alpha(1 - t) \quad \text{for all } t \in [-1,1 - \delta],
\]

\[
K \in \text{CUT}(S^{n-1}) \text{ is invariant}
\]

is at least $\alpha_n$.

**Proof.** Fix $\eta > 0$ and let $\mathcal{P}$ be a partition supplied by Lemma 2.1. Let $\mathcal{F}$ be the set of all functions $f: S^{n-1} \to \{-1,1\}$ that respect $\mathcal{P}$; note $\mathcal{F}$ is finite.

Let $I_1 \subseteq I_2 \subseteq \cdots$ be a sequence of finite nonempty subsets of $[-1,1]$ whose union is the set of all rational numbers in $[-1,1]$. Suppose there is no $m_k: \mathcal{F} \to \mathbb{R}$ satisfying

\[
\sum_{f \in \mathcal{F}} (1 - R(f \otimes f^*)(t))m_k(f) \geq \alpha_n(1 - t) - \eta \quad \text{for all } t \in I_k,
\]

\[
\sum_{f \in \mathcal{F}} m_k(f) = 1,
\]

\[
m_k \geq 0.
\]
Farkas’s lemma [16, §7.3] says that, if this system has no solution, then there is \( z : I_k \to \mathbb{R}, \ z \geq 0, \) and \( \rho \in \mathbb{R} \) such that

\[
\rho + \sum_{t \in I_k} z(t)(1 - R(f \otimes f^*)(t)) \leq 0 \quad \text{for all } f \in \mathcal{F},
\]

\[
\rho + \sum_{t \in I_k} z(t)(\alpha_n(1 - t) - \eta) > 0.
\]

Together, these inequalities imply that for every \( f \in \mathcal{F} \) we have

\[
\sum_{t \in I_k} z(t)(1 - R(f \otimes f^*)(t)) < \sum_{t \in I_k} z(t)(\alpha_n(1 - t) - \eta),
\]

a contradiction to the choice of \( \mathcal{P} \).

Since all \( m_k \) lie in \([0, 1]^3\), which is a compact set, the sequence \((m_k)\) has a converging subsequence; say this subsequence converges to \( m : \mathcal{F} \to \mathbb{R} \). Then \( m \geq 0 \) and \( \sum_{f \in \mathcal{F}} m(f) = 1 \). Moreover,

\[
\sum_{f \in \mathcal{F}} (1 - R(f \otimes f^*)(t))m(f) \geq \alpha_n(1 - t) - \eta \quad \text{for all } t \in [-1, 1]. \tag{9}
\]

Indeed, the inequality holds for all \( t \in [-1, 1] \cap \mathbb{Q} \). But \( R(f \otimes f^*) \) is continuous for every \( f \), so the left-hand side above is a continuous function of \( t \), whence the inequality holds for every \( t \in [-1, 1] \).

Fix \( 1 \geq \delta > 0 \) and \( \varepsilon > 0 \) and set \( \eta = \alpha_n \varepsilon \delta \); let \( m \) be such that (9) holds. If \( t \leq 1 - \delta \), then \( 1 - t \geq \delta \) and

\[
(1 - \varepsilon)(1 - t) = (1 - t) - \varepsilon(1 - t) \leq (1 - t) - \varepsilon \delta.
\]

So, for \( t \in [-1, 1 - \delta] \), the left-hand side of (9) is at least

\[
\alpha_n(1 - t) - \eta = \alpha_n((1 - t) - \alpha_n^{-1} \eta) \geq \alpha_n(1 - \varepsilon)(1 - t).
\]

Now \( K_\varepsilon = \sum_{f \in \mathcal{F}} R(f \otimes f^*)m(f) \) is a continuous kernel that moreover belongs to \( \text{CUT}_{\mathbb{R}}(S^{n-1}) \). So for every \( \varepsilon > 0 \) there is \( K_\varepsilon \in \text{CUT}_{\mathbb{R}}(S^{n-1}) \) such that \( (K_\varepsilon, \alpha_n(1 - \varepsilon)) \) is a feasible solution of (8), and by letting \( \varepsilon \) approach 0 we are done. \( \square \)

For \( n \geq 3 \), Theorem 1.1 can be obtained from Lemma 2.2 by using the following lemma.

**Lemma 2.3.** For every \( n \geq 3 \), there is \( 1 \geq \delta > 0 \) such that if \( (K, \alpha) \) is any feasible solution of (8), then

\[
1 - K(t) \geq \alpha(1 - t) \quad \text{for all } t \in [1 - \delta, 1].
\]

The proof of this lemma uses some properties of Jacobi polynomials, and goes through only for \( n \geq 3 \). A proof of Theorem 1.1 for \( n = 2 \) is given in §3.

The Jacobi polynomials\(^1\) with parameters \((\alpha, \beta), \ \alpha, \ \beta > -1, \) are the orthogonal polynomials with respect to the weight function \((1 - t)^\alpha(1 + t)^\beta\) on the interval \([-1, 1]\). We denote the Jacobi polynomial with parameters \((\alpha, \beta)\) and degree \( k \) by \( P_k^{(\alpha, \beta)} \) and normalize it so \( P_k^{(\alpha, \beta)}(1) = 1. \)

A continuous kernel \( K : S^{n-1} \times S^{n-1} \to \mathbb{R} \) is positive if \((K(x, y))_{x,y \in U}\) is positive semidefinite for every finite and nonempty set \( U \subseteq S^{n-1} \). Schoenberg [15] characterizes continuous, positive, and invariant kernels via their expansions in terms of Jacobi polynomials:

\(^1\)See for example the book by Szegö [17] for background on orthogonal polynomials.
Theorem 2.4 (Schoenberg’s theorem). A kernel $K: S^{n-1} \times S^{n-1} \to \mathbb{R}$ is continuous, positive, and invariant if and only if there are numbers $a_k \geq 0$ satisfying $\sum_{k=0}^{\infty} a_k < \infty$ such that

$$K(x, y) = \sum_{k=0}^{\infty} a_k P_k^{(\nu, \nu)}(x \cdot y) \quad \text{for all } x, y \in S^{n-1}$$

with absolute and uniform convergence, where $\nu = (n - 3)/2$.

Schoenberg’s theorem is used in the proof of Lemma 2.3 and again in §§3 and 4.

Proof of Lemma 2.3. Fix $n \geq 3$ and set $\nu = (n - 3)/2$. Claim: there is $1 \geq \delta > 0$ such that $t = P_k^{(\nu, \nu)}(t) \geq P_k^{(\nu, \nu)}(t)$ for all $k \geq 2$ and $t \in [1 - \delta, 1]$.

The lemma quickly follows from this claim. Indeed, say $(K, \alpha)$ is feasible for (8). Since every matrix in CUT(U) for finite $U$ is positive semidefinite, every kernel in CUT(S^{n-1}) is positive. Hence using Schoenberg’s theorem we write

$$K(t) = \sum_{k=0}^{\infty} a_k P_k^{(\nu, \nu)}(t) \quad \text{for all } t \in [-1, 1].$$

Since $K \in$ CUT(S^{n-1}), we have $K(1) = 1$, so $\sum_{k=0}^{\infty} a_k = 1$.

As $(K, \alpha)$ is a feasible solution of (8), we know that

$$1 - K(-1) \geq \alpha(1 - (-1)) = 2\alpha.$$

Now $|P_k^{(\nu, \nu)}(t)| \leq 1$ for all $k$ and all $t \in [-1, 1]$, so $K(-1) \geq a_0 - (1 - a_0)$, whence $a_0 \leq 1 - \alpha$. The claim implies that, if $t \in [1 - \delta, 1]$, then

$$K(t) \leq a_0 + (1 - a_0)t,$$

so for $t \in [1 - \delta, 1]$ we have

$$1 - K(t) \geq 1 - a_0 - (1 - a_0)t = 1 - t - a_0(1 - t) \geq 1 - t - (1 - \alpha)(1 - t) = \alpha(1 - t),$$

as we wanted.

To prove the claim, we use the following integral representation of Feldheim and Vilenkin for the Jacobi polynomials: for $\nu \geq 0$,

$$P_k^{(\nu, \nu)}(\cos \theta) = \frac{2\Gamma(\nu + 1)}{\Gamma(1/2)\Gamma(\nu + 1/2)} \int_0^{\pi/2} \cos^{2\nu} \phi (1 - \sin^2 \theta \cos^2 \phi)^{k/2} \cdot P_k^{(-1/2,-1/2)}(\cos \theta (1 - \sin^2 \theta \cos^2 \phi)^{-1/2}) d\phi. \quad (10)$$

This formula is adapted to our normalization of the Jacobi polynomials from Corollary 6.7.3 in the book by Andrews, Askey, and Roy [1]; see also equation (3.23) in the thesis by Oliveira [13].

For fixed $\theta$ and $\phi$, the function $k \mapsto (1 - \sin^2 \theta \cos^2 \phi)^{k/2}$ is monotonically decreasing. Write $t = \cos \theta$ and recall that the Jacobi polynomials are bounded by 1 in $[-1, 1]$; plug $k = 2$ in the right-hand side of (10) to get

$$P_k^{(\nu, \nu)}(t) \leq \frac{2\Gamma(\nu + 1)}{\Gamma(1/2)\Gamma(\nu + 1/2)} \int_0^{\pi/2} \cos^{2\nu} \phi (1 - (1 - t^2) \cos^2 \phi) d\phi \quad (11)$$
for all \( t \in [0, 1] \) and \( k \geq 2 \). For \( v = (n - 3)/2 \) with \( n \geq 4 \), we show that there is \( \delta > 0 \) such that the right-hand side above is at most \( t \) for all \( t \in [1 - \delta, 1] \); the case \( n = 3 \) will be dealt with shortly.

Let \( m \geq 2 \) be an integer. Write \( \cos^m \phi = \cos^{m-1} \phi \cos \phi \) and use integration by parts to get
\[
m \int_0^{\pi/2} \cos^m \phi \, d\phi = (m-1) \int_0^{\pi/2} \cos^{m-2} \phi \, d\phi.
\]

It follows by induction on \( m \) that, if \( v = (n - 3)/2 \) with \( n \geq 3 \), then
\[
\int_0^{\pi/2} \cos^{2v} \phi \, d\phi = \frac{\Gamma(1/2)\Gamma(v+1/2)}{2\Gamma(v+1)}.
\]

The right-hand side of (11) is a degree-2 polynomial on \( t \); let us denote it by \( p_v \). Use (12) to get
\[
p_v(t) = \frac{2v + 1}{2(v+1)} t^2 + \frac{1}{2(v+1)}.
\]

It is then a simple matter to check that, for \( v = (n - 3)/2 \) with \( n \geq 4 \), there is \( \delta > 0 \) such that \( p_v(t) \leq t \) for all \( t \in [1 - \delta, 1] \).

For \( n = 3 \) and hence \( v = 0 \), we have \( p_v(t) \geq t \) for all \( t \in [0, 1] \). In this case, we may take \( k = 4 \) in (10) and follow the same reasoning, proving that the degree 4 polynomial obtained will have the desired property. It then only remains to show that \( P_2^{(0,0)} \) and \( P_3^{(0,0)} \) are below \( P_1^{(0,0)} \) for \( t \) close enough to 1, and this can be done directly.

All that is left to do is to put it all together.

Proof of Theorem 1.1 for \( n \geq 3 \). In §1.1 we have seen that the optimal value of (4) is at most \( \alpha_n \). The reverse inequality follows from Lemmas 2.2 and 2.3 put together.

3 Lower bounds for \( \alpha_n \) and a proof of Theorem 1.1 for \( n = 2 \)

To get a lower bound for \( \alpha_n \), one needs to show a feasible solution of (4). One such feasible solution, that shows that \( \alpha_n \geq \alpha_{GW} \), is \((K_{GW}, \alpha_{GW})\) with
\[
K_{GW}(x \cdot y) = (2/\pi) \arcsin x \cdot y.
\]

We encountered this kernel in the introduction. Fix \( e \in S^{n-1} \) and let \( f_{GW}: S^{n-1} \to \{-1, 1\} \) be such that \( f_{GW}(x) = 1 \) if \( e \cdot x \geq 0 \) and \(-1\) otherwise. Recall that Grothendieck’s identity is
\[
K_{GW}(x \cdot y) = R(f_{GW} \otimes f_{GW}^*)(x \cdot y),
\]
whence in particular \( K_{GW} \in \text{CUT\_G}(S^{n-1}) \).

Let \( t_{GW} \in [-1, 1] \) be such that \( \alpha_{GW} = (1 - K_{GW}(t_{GW}))/ (1 - t_{GW}) \); then \( t_{GW} = -0.68918 \ldots \). The easy direction of the following result is implicit in the work of Avidor and Zwick [2].
Theorem 3.1. If $n \geq 2$, then $\alpha_n > \alpha_{GW}$ if and only if there is an invariant kernel $K \in \text{CUT}_{\Box}(S^{n-1})$ such that

$$1 - K(t_{GW}) > 1 - K_{GW}(t_{GW}). \tag{14}$$

If, moreover, $\alpha_n > \alpha_{GW}$, then there is a measurable function $f : S^{n-1} \to \{-1, 1\}$ such that (14) holds for $K = R(f \otimes f^*)$.

Proof. First the easy direction. Suppose there is such a kernel $K$. Then

$$1 - K(t_{GW}) > 1 - K_{GW}(t_{GW}) = \alpha_{GW}(1 - t_{GW}). \tag{15}$$

Both functions

$$t \mapsto 1 - K(t) \quad \text{and} \quad t \mapsto 1 - K_{GW}(t)$$

are continuous in $[-1, 1]$. From (15), we see that there is $\varepsilon > 0$ such that the first function above is at least $(\alpha_{GW} + \varepsilon)(1 - t)$ in some interval $I$ around $t_{GW}$. The second function above is at least $\alpha_{GW}(1 - t)$ in $[-1, 1]$ and, if $\varepsilon$ is small enough, then it is at least $(\alpha_{GW} + \varepsilon)(1 - t)$ in $[-1, 1] \setminus I$ (recall from (13) that we know the second function explicitly). But then for some $\lambda \in [0, 1]$ and small enough $\varepsilon' > 0$ we will have that

$$K' = \lambda K + (1 - \lambda) K_{GW} \in \text{CUT}_{\Box}(S^{n-1})$$

is such that $1 - K'(t) \geq (\alpha_{GW} + \varepsilon')(1 - t)$ for all $t \in [-1, 1]$, so the optimal value of (4) is greater than $\alpha_{GW}$ and therefore $\alpha_n > \alpha_{GW}$ from the easy direction of Theorem 1.1 (proved in §1.1).

Now suppose $\alpha_n > \alpha_{GW}$. For every $\eta > 0$, Lemma 2.1 gives a measurable function $f : S^{n-1} \to \{-1, 1\}$ such that

$$1 - R(f \otimes f^*)(t_{GW}) \geq \alpha_n(1 - t_{GW}) - \eta$$

(take $I = \{t_{GW}\}$ and $z = 1$ in the lemma); set $K = R(f \otimes f^*)$. Then

$$1 - K(t_{GW}) \geq \alpha_n(\alpha_{GW}^{-1}(1 - K_{GW}(t_{GW}))) - \eta.$$ 

Since $\alpha_n / \alpha_{GW} > 1$, we finish by taking $\eta$ close enough to 0. \hfill \square

Theorem 3.1 shows that, to find a lower bound for $\alpha_n$, we need to find a better partition of the sphere $S^{n-1}$, and this can be done by finding a maximum cut in a graph defined on a discretization of the sphere (cf. the proof of Lemma 2.1). This can be tricky in general: Avidor and Zwick [2] present such a better partition for $n = 3$, but their construction is ad hoc. For $n = 2$, however, one may use the hyperplane rounding procedure to obtain such a better partition, in a curious application of the Goemans-Williamson algorithm to improve on itself.

We want to find an invariant kernel $K \in \text{CUT}_{\Box}(S^{n-1})$ satisfying (15), that is, we want to find a good solution of the following optimization problem:

$$\sup 1 - K(t_{GW})$$

implies $K \in \text{CUT}_{\Box}(S^{n-1})$ is invariant.
This seems to be a difficult problem, but we can relax the constraint that \( K \in \text{CUT}(S^{n-1}) \) by requiring only that \( K \) be positive. Then, using Schoenberg’s theorem to parameterize \( K \) as in §2, we get the following relaxation of our problem:

\[
\sup \left( 1 - \sum_{k=0}^{\infty} a_k P_k^{(v,v)}(t_{GW}) \right) \quad \text{subject to} \quad \sum_{k=0}^{\infty} a_k = 1, \quad a_k \geq 0 \quad \text{for all } k \geq 0. \tag{16}
\]

For \( n = 2 \) and hence \( v = -1/2 \), the optimal solution of (16) is \( a_k = 0 \) for all \( k \neq 4 \) and \( a_4 = 1 \), as may be proved, for instance, by showing a solution to the dual of (16) having the same objective value as the solution \( a \) (see §4 for a description of the dual problem of a problem related to (16)).

Using formula (5.1.1) from Andrews, Askey, and Roy [1], this means that the optimal kernel is

\[ K(\cos \theta) = P_4^{(-1/2,-1/2)}(\cos \theta) = \cos 4\theta. \]

If we identify the circle \( S^1 \) with the interval \([0, 2\pi]\), then the inner product between points \( \theta, \phi \in [0, 2\pi] \) is \( \arccos(\theta - \phi) \), so

\[ K(\theta, \phi) = \cos 4(\theta - \phi) = \cos 4\theta \cos 4\phi + \sin 4\theta \sin 4\phi. \]

Taking \( g: S^1 \rightarrow S^1 \) such that \( g(\theta) = (\cos 4\theta, \sin 4\theta) \), we have \( K(\theta, \phi) = g(\theta) \cdot g(\phi) \).

Now, let us round the rank-2 solution \( g \). Let \( e = (1, 0) \) and set \( f(\theta) = 1 \) if \( e \cdot g(\theta) \geq 0 \) and \( f(\theta) = -1 \) otherwise. The resulting partition is exactly the windmill partition that, combined with the partition \( f_{GW} \) of the sphere into two equal halves, shows that

\[ \alpha_2 = \frac{32}{25 + 5\sqrt{5}} \]

(cf. Avidor and Zwick [2]); see also Figure 1.

**Proof of Theorem 1.1 for \( n = 2 \).** In §1.1 we have seen that the optimal value of (4) is at most \( \alpha_2 \). The reverse inequality is proved by Avidor and Zwick [2]: they show how to pick \( \lambda \in [0, 1] \) such that, if \( f: S^1 \rightarrow \{-1, 1\} \) is the windmill partition of Figure 1 and \( f_{GW}: S^1 \rightarrow \{-1, 1\} \) is the partition into two equal halves, then \((K, \alpha)\) with

\[ K = \lambda R(f \otimes f^*) + (1 - \lambda) R(f_{GW} \otimes f_{GW}^*) \]

and

\[ \alpha = \frac{32}{25 + 5\sqrt{5}} \]

is a feasible solution of (4). Since \( \alpha_2 = \alpha \), we are then done.

For \( n \geq 3 \), the approach outlined above does not work. The optimal solution of the relaxation (16) is always \( a_k = 0 \) for all \( k \neq 1 \) and \( a_1 = 1 \). The hyperplane rounding then gives the partition \( f_{GW} \) into two equal halves, therefore not providing a lower bound for \( \alpha_n \) better than \( \alpha_{GW} \).
Figure 1: On the left we have the unit circle $S^1$ and, supported on a point $\theta$, the vector $f(\theta) = (\cos 4\theta, \sin 4\theta)$ — in blue if $\cos 4\theta \geq 0$ and in red otherwise. On the right, we have in gray the segments of the circle where $\cos 4\theta \geq 0$; this is the windmill partition.

4 Upper bounds for $\alpha_n$ and bad instances

Let us see how to solve a relaxation of (4) in order to get upper bounds for $\alpha_n$. The first order of business is to use Schoenberg’s theorem (Theorem 2.4) to parameterize $K$ as

$$K(t) = \sum_{k=0}^{\infty} a_k P_k^{(v)}(t) \quad \text{for all } t \in [-1, 1],$$

where $v = (n-3)/2$, $a_k \geq 0$ for all $k$, and $\sum_k a_k < \infty$.

Say now that $U \subseteq S^{n-1}$ is a nonempty finite set and $Z: U \times U \to \mathbb{R}$ and $\beta \in \mathbb{R}$ are such that

$$\sum_{x,y \in U} Z(x,y)A(x,y) \geq \beta$$

for all $A \in \text{CUT}_{\square}(U)$, so $Z$ and $\beta$ give a valid constraint for $\text{CUT}_{\square}(U)$. If $K \in \text{CUT}_{\square}(S^{n-1})$, then

$$\sum_{x,y \in U} Z(x,y)K(x,y) \geq \beta.$$

Rewriting this inequality using the parametrization of $K$ we see that the variables $a_k$ satisfy the constraint

$$\sum_{k=0}^{\infty} a_k r_k \geq \beta,$$

where $r = (r_k)$ is the sequence such that

$$r_k = \sum_{x,y \in U} Z(x,y) P_k^{(v)}(x \cdot y).$$
Let $\mathcal{R}$ be a finite collection of pairs $(r, \beta)$, each one associated with a valid constraint of $\text{CUT}_{\square}(U)$ for some finite set $U \subseteq S^n - 1$, as described above. Recall that, if $K \in \text{CUT}_{\square}(S^n)$, then $K(1) = 1$, and that $P_k^{(v,v)}(1) = 1$ in our normalization. Choose a finite nonempty set $I \subseteq [-1, 1]$. Then the following linear program with infinitely many variables but finitely many constraints is a relaxation of (4); its optimal value thus provides an upper bound for $\alpha_n$:

$$
\begin{align*}
\sup \alpha & \quad \sum_{k=0}^{\infty} a_k = 1, \\
\alpha(1-t) + \sum_{k=0}^{\infty} a_k P_k^{(v,v)}(t) & \leq 1 \quad \text{for all } t \in I, \\
\sum_{k=0}^{\infty} a_k r_k & \geq \beta \quad \text{for all } (r, \beta) \in \mathcal{R}, \\
a_k & \geq 0 \quad \text{for all } k \geq 0.
\end{align*}
$$

(18)

A dual problem for (18) is

$$
\begin{align*}
\inf \lambda & + \sum_{i \in I} z(i) - \sum_{(r, \beta) \in \mathcal{R}} y(r, \beta) \beta \\
\sum_{i \in I} z(i)(1-t) & = 1, \\
\lambda + \sum_{i \in I} z(i) P_k^{(v,v)}(t) - \sum_{(r, \beta) \in \mathcal{R}} y(r, \beta) r_k & \geq 0 \quad \text{for all } k \geq 0, \\
z, y & \geq 0.
\end{align*}
$$

(19)

It is routine to show that weak duality holds between the two problems: if $(a, \alpha)$ is a feasible solution of (18) and $(\lambda, z, y)$ is a feasible solution of (19), then

$$
\alpha \leq \lambda + \sum_{i \in I} z(i) - \sum_{(r, \beta) \in \mathcal{R}} y(r, \beta) \beta.
$$

So to find an upper bound for $\alpha_n$ it suffices to find a feasible solution of (19).

To find such a feasible dual solution we follow the same approach presented by DeCorte, Oliveira, and Vallentin [3, §8] for a very similar problem. We start by choosing a large enough value $d$ (say $d = 2000$) and truncating the series in (17) at degree $d$, setting $a_k = 0$ for all $k > d$. Then, for finite sets $I$ and $\mathcal{R}$, problem (18) becomes a finite linear program. We solve it and from its dual we obtain a candidate solution $(\lambda, z, y)$ for the original, infinite-dimensional dual. All that is left to do is check that this is indeed a feasible solution, or else that it can be turned into a feasible solution by slightly increasing $\lambda$. This verification procedure is also detailed by DeCorte, Oliveira, and Vallentin (ibid., §8.3).

Finding a good set $I \subseteq [-1, 1]$ is easy: one simply takes a finely spaced sample of points. Finding a good set $\mathcal{R}$ of constraints is another issue. The approach is, again, detailed by DeCorte, Oliveira, and Vallentin (ibid., §8.3); here is an outline. We start by setting $\mathcal{R} = \emptyset$. Then, having a solution of (18), and having access to a list of facets of $\text{CUT}_{\square}(U)$ for a set $U$ of 7 elements, numerical methods for unconstrained optimization are used to find points on the sphere for which a given inequality is violated. These violated inequalities are then added to (18) and the process is repeated.

Table 1 shows a list of upper bounds for $\alpha_n$ found with the procedure described above. These bounds have been rigorously verified using the approach of DeCorte, Oliveira, and Vallentin.
Table 1: Upper bounds for $\alpha_n$ from a relaxation of problem (4). For $n = 3$, the relaxation gives an upper bound of 0.8854, not better than $\alpha_2$. These bounds have all been computed considering a same set $\mathcal{R}$ with 28 constraints from the cut polytope found heuristically for the case $n = 4$; improvements could possibly be obtained by trying to find better constraints for each dimension. The bound using $\mathcal{R}$ decreases more and more slowly after $n = 19$; for $n = 10000$ one obtains the upper bound 0.878695.

### 4.1 Constructing bad instances

A feasible solution of (19) gives an upper bound for $\alpha_n$, but this upper bound is not constructive, that is, we do not get an instance of the maximum-cut problem with large integrality gap. Let us see now how to extract bad instances for the maximum-cut problem from a solution of (19).

Let $I \subseteq [-1, 1)$ be a finite nonempty set of inner products and $\mathcal{R}$ be a finite set of constraints from the cut polytope. Say $(\lambda, z, y)$ is a feasible solution of (19) and let

$$\alpha = \lambda + \sum_{t \in I} z(t) - \sum_{(r, \beta) \in \mathcal{R}} y(r, \beta) \beta$$

be its objective value.

The intuition behind the construction is simple. We consider a graph on the sphere $S^{n-1}$, where $x, y \in S^{n-1}$ are adjacent if $x \cdot y \in I$ and the weight of an edge between $x$ and $y$ is $z(x \cdot y)$. Bad instances will arise from discretizations of this infinite graph.

Given a partition $\mathcal{P}$ of $S^{n-1}$ into finitely many sets, denote by $\delta(\mathcal{P})$ the maximum diameter of any set in $\mathcal{P}$. Let $(\mathcal{P}_m)$ be a sequence of partitions of $S^{n-1}$ into finitely many measurable sets such that $\mathcal{P}_{m+1}$ is a refinement of $\mathcal{P}_m$ and

$$\lim_{m \to \infty} \delta(\mathcal{P}_m) = 0.$$

For $m \geq 0$, let $A^m_z : \mathcal{P}_m \times \mathcal{P}_m \to \mathbb{R}$ be the matrix defined in (6) for the partition $\mathcal{P} = \mathcal{P}_m$ and the function $z$. Since $\mathcal{P}_{m+1}$ is a refinement of $\mathcal{P}_m$, both limits

$$\lim_{m \to \infty} \text{SDP}_1(A^m_z) \quad \text{and} \quad \lim_{m \to \infty} \text{SDP}_n(A^m_z)$$

exist, as the sequences of optimal values are monotonically increasing and bounded. As $z \neq 0$, both limits are positive, hence

$$\lim_{m \to \infty} \frac{\text{SDP}_1(A^m_z)}{\text{SDP}_n(A^m_z)} = (20)$$
exists. Claim: the limit above is at most \( \alpha \).

Once the claim is established, we are done: for every \( \epsilon > 0 \), by taking \( m \) large enough (that is, by taking a fine enough partition) we have

\[
\frac{\text{SDP}_1(A^m)}{\text{SDP}_n(A^m)} \leq \alpha + \epsilon,
\]

that is, we get a sequence of bad instances for the maximum-cut problem.

To prove the claim, suppose (20) is at least \( \alpha + \epsilon \) for some fixed \( \epsilon > 0 \). Then for all large enough \( m \) we have

\[
\text{SDP}_1(A^m) \geq \frac{\alpha + \epsilon}{\text{SDP}_n(A^m)} \cdot \text{SDP}_n(A^m).
\]

Following the proof of Lemma 2.1, this means that for every large enough \( m \) there is a function \( f_m : S^{n-1} \to \{ -1, 1 \} \) that respects \( \mathcal{P}_m \) and satisfies

\[
\sum_{t \in I} z(t) ((1 - R(f_m \otimes f_m^*)(t)) \geq \sum_{t \in I} z(t)((\alpha + \epsilon)(1 - t) - \eta_m),
\]

where \( \eta_m \geq 0 \) and \( \eta_m \to 0 \) as \( m \to \infty \).

Use the feasibility of \((\lambda, z, y)\) for (19) together with the definition of \( \alpha \) to get from the above inequality that

\[
\lambda + \sum_{t \in I} z(t) R(f_m \otimes f_m^*)(t) - \sum_{(r, \beta) \in \mathbb{R}} y(r, \beta) \beta \leq -\epsilon + \eta_m \sum_{t \in I} z(t). \tag{21}
\]

Next, note that \( R(f_m \otimes f_m^*) \in \text{CUT}_{\mathbb{R}}(S^{n-1}) \). Using Schoenberg’s theorem (Theorem 2.4), write

\[
R(f_m \otimes f_m^*)(t) = \sum_{k=0}^{\infty} a_k P_k^{(v, v)}(t),
\]

where \( v = (n - 3)/2 \), \( a_k \geq 0 \), and \( \sum_{k=0}^{\infty} a_k = 1 \). Use again the feasibility of \((\lambda, z, y)\) for (19) together with (21) to get

\[
0 \leq \sum_{k=0}^{\infty} a_k \left( \lambda + \sum_{t \in I} z(t) P_k^{(v, v)}(t) - \sum_{(r, \beta) \in \mathbb{R}} y(r, \beta) r_k \right) \\
\leq \lambda + \sum_{t \in I} z(t) R(f_m \otimes f_m^*)(t) - \sum_{(r, \beta) \in \mathbb{R}} y(r, \beta) \beta \\
\leq -\epsilon + \eta_m \sum_{t \in I} z(t).
\]

Since \( \epsilon > 0 \) and \( \eta_m \to 0 \) as \( m \to \infty \), by taking \( m \) large enough we get a contradiction, proving the claim.

Acknowledgements

We thank the referees for valuable suggestions that improved the paper.
References


The integrality gap of the maximum-cut semidefinite relaxation in fixed dimension


AUTHORS

F.M. de Oliveira Filho  
Delft Institute of Applied Mathematics 
Delft University of Technology  
Van Mourik Broekmanweg 6, 2628 XE Delft, The Netherlands. 
fmario@gmail.com

F. Vallentin  
Mathematisches Institut  
Universität zu Köln  
Weyertal 86–90, 50931 Köln, Germany. 
frank.vallentin@uni-koeln.de