Linear Programming Relaxations of Maxcut

Wenceslas Fernandez de la Vega   Claire Kenyon-Mathieu

Abstract

It is well-known that the integrality gap of the usual linear programming relaxation for Maxcut is $2-\epsilon$. For general graphs, we prove that for any ϵ and any fixed bound$k$, adding linear constraints of support bounded by k does not reduce the gap below $2-\epsilon$. We generalize this to prove that for any ϵ and any fixed bound k, strengthening the usual linear programming relaxation by doing k rounds of Sherali-Adams lift-and-project does not reduce the gap below $2-\epsilon$. On the other hand, we prove that for dense graphs, this gap drops to $1+\epsilon$ after adding all linear constraints of support bounded by some constant depending on ϵ.

1 Introduction

Linear programming relaxations are a standard tool to design approximation algorithms for hard combinatorial optimization problems. This well-known algorithmic paradigm typically models the problem as an integer program, solves a linear programming relaxation, and then uses the optimal fractional solution to construct a feasible solution for the original problem by more or less clever rounding. The value of the rounded solution output by the algorithm is analyzed by comparing it to the value of the fractional solution, which is a solution of the linear program. Thus, the best approximation ratio one can hope to prove by such an approach is, at best, the maximum ratio between the value of the integer optimum and the value of the optimal solution of the linear program, *aka* the gap of the linear program. An approximation algorithm can sometimes be improved by designing better linear programs, enriching a given LP by putting in additional constraints so as to decrease the gap between the optimal fractional and optimal integral solutions. This paper studies attempts at doing such improvements for the Maxcut problem.

Maxcut, or finding the maximum cut in a graph, is a fundamental problem of combinatorial optimization. The standard linear programming relaxation has an integrality gap of $2-\epsilon$ [20]: there exist graphs for which the number of edges of the maximum cut is greater than the value of the linear program by a factor of $2-\epsilon$.

Much progress has been done on Maxcut by studying semi-definite relaxations instead of linear programming relaxations. The standard semi-definite programming relaxation has an integrality gap of at most 1/.878... [12]; in fact, the gap is exactly 1/.878... [8]. In efforts to improve the approximation by decreasding the integrality gap, people have considered adding triangular inequalities to the semi-definite program. Unfortunately, after several years, it was proved that this leaves a gap of at least 1/.94 [8], and in fact, that the gap is still 1/.878 [15]. (Indeed, if the unique games conjecture holds, then 1/.878 is best possible within P [16].)

This raises a natural question: How much could additional constraints help?

In this paper, we study this question, focusing exclusively on linear programs. This work was inspired by [3, 1] and can be seen as the analog of that work, in the context of the Maxcut problem.

We first study the strenghtening of linear programs for Maxcut obtained by adding all constraints of bounded support. Unfortunately, our first result is negative: we prove that the integrality gap of the Maxcut linear program remains at $2-\epsilon$, even after adding constraints of bounded support.

The proof technique is, in our view, at least as interesting as the result itself. Here is the high-level idea. In order to prove that a candidate vector x satisfies all the constraints of the LP, you focus on the ith constraint, $a_i^T x \leq b_i$; you construct a vector $x^{(i)}$ which is easily seen to satisfy that constraint (it is a convex combination of feasible integer solutions); you prove that it satisfies the constraint with some slack,

$a_i^T x^{(i)} - b_i \geq \epsilon > 0$; and you prove that $|x - x^{(i)}|$ is small, so small compared to the coefficients of a_i that $|a_i^T(x - x^{(i)})|$ must be less than ϵ. This implies that x satisfies the constraint. One nice feature of that proof is that it requires relatively few calculations: in particular, we never need to write the constraints explicitly.

We then go further and study the strenghtening of linear programs for Maxcut obtained by the lift-and-project technique of Sherali and Adams, a systematic way of decreasing linear programming integrality gap automatically; note that the Sherali-Adams variant is stronger than the variant of Lovasz and Shrijver [19], which has been the primary object of attention recently [3, 1, 6]. Unfortunately, we also prove that the integrality gap of the Maxcut linear program remains at $2 - \epsilon$ even after doing a small number of rounds of Sherali-Adams (and therefore, also of Lovasz-Shrijver) lift-and-project (and therefore, also after doing the same number of rounds of Lovasz-Shrijver lift-and-project). Note that this second result subsumes the first result, whose proof is only included for clarity.

To prove the second result, we apply the same proof technique. One difficulty is that we cannot necessarily, for every constraint, find a vector $x^{(i)}$ that satisfies the constraint with slack. Indeed, we work directly in the lifted LP (before the projection step), and that linear program is not full-dimensional: some constraints – the ones which reduce the dimension – are tight for every feasible solution, hence no slack is possible. In order to deal with that, we identify a subset of variables which are linearly independent define them in a natural manner using a probabilistic interpretation of lifted variables, and use them to define the rest of the variables.

Some special cases of Maxcut do have better approximations. In particular, in spite of its difficulty in general, Maxcut is easy to approximate on dense graphs. To complement our negative results, we prove that the integrality gap of the Maxcut linear program drops to $1 + \epsilon$ after a constant number of rounds of lift-and-project, in the special case when the underlying graph is dense.

2 Bounded support linear constraints

Let $\chi()$ denote the indicator function.

The integer cut polytope Cut_k on a set $U = \{x_1, x_2, \ldots, x_k\}$ of k vertices is defined as follows. Let F denote the set of all pairs of vertices, $|F| = m = k(k-1)/2$. Then Cut_k is the set of vectors $(x_e)_{e \in F}$ such that

$$\begin{cases} x_e = \sum_{S \subseteq U} \lambda_S \chi(e \text{ crosses } \mathrm{cut}(S, U \setminus S)) & \forall e \in F \\ \lambda_S \geq 0, \quad \sum_{S \subseteq U} \lambda_S = 1 \end{cases}$$

In other words, (λ_S) defines a distribution over cuts of $\{x_1, x_2, \ldots, x_k\}$, and x_e is the probability that e belongs to the cut under this distribution.

Let $V = \{1, 2, \ldots, n\}$. Given a graph with vertex set V, let K be the fractional polytope of the standard linear programming relaxation for MAXCUT:

$$\text{Maximise} \sum_{1 \leq i < j \leq n} X_{ij} e_{ij} \quad \text{subject to}:$$

$$X_{ij} \leq X_{ik} + X_{kj}$$

$$X_{ij} + X_{ik} + X_{kj} \leq 2$$

$$0 \leq X_{ij} \leq 1$$

Here, the X_{ij}'s are the variables, and (e_{ij}) is the indicator function of the edges of the graph.

It is known [20] that K has integrality gap $2 - \epsilon$. However, there exist many more valid cut inequalities than just the ones specified in K. For example, every cut satisfies $\sum_{\{u,v\} \subseteq \{i,j,k,l,m\}} X_{uv} \leq 6$. It is tempting to attempt to reduce the integrality gap by adding such inequalities. Given a set of k vertices, one could, at least in principle, write down every inequality which is satisfied by all cuts of those k vertices, i.e. every constraint of Cut_k. Each such constraint involves at most $\binom{k}{2}$ variables, hence has support bounded by $\binom{k}{2}$. For $k = 3$, it is known that the constraints already specified in the definition of K are all the constraints of Cut_3.

Theorem 2.1. *Let $\epsilon > 0$ and $k \geq 4$ be fixed. Consider the linear program obtained from K by adding all the constraints of Cut_k, for every k-subset of V. Then the integrality gap is at least $2 - \epsilon$.*

As in [20, 3], the proof relies on large girth graphs. We first define the vector $X = (X_{ij})$ which we will later prove to be feasible.

Definition 1. *Suppose that G is a graph with vertex set $V = \{1, 2, ...n\}$ and let $m = n(n-1)/2$. Let d_{ij}*

denote the shortest path distance between i and j. Let k be a fixed positive integer and let ϵ be a positive real with $\epsilon \le 1/4$. Let $s = \frac{1}{2}\epsilon^{-1}\log(1/\epsilon)k^4$. Define the following vector $X \in \mathbf{R}^m$.
If $d_{ij} < s$, then

$$X_{ij} = \frac{1}{2}(1-(1-2\epsilon)^{d_{ij}}) + (1-2\epsilon)^{d_{ij}}\chi(d_{ij} \ \textit{odd}).$$

And if $d_{ij} \ge s$, then

$$X_{ij} = \frac{1}{2}.$$

Intuitively, for $d_{ij} < s$, let P denote the path of length d_{ij} from i to j. Then X_{ij} is the probability that i and j are separated in the cut of P constructed by the following random process. Each edge of P is deleted independently with probability 2ϵ. For each resulting subpath of P, the vertices are placed in L and in R in an alternating manner as we go along the subpath, starting with a random choice of either L or R to start the subpath. This defines a random cut of P.

It is not hard to see by doing a short case-by-case analysis that if G has girth 4 or more, then X is in K. (A variant of this was used in [20]). The following proposition is the main new part of our proof.

PROPOSITION 2.1. *If the girth of G is at least $2(k-1)s$, then, for every k-subset of vertices, X satisfies all the constraints of* Cut_k.

Proof. Let $Q^T x \le a$ be a fixed constraint of $\mathrm{Cut}(k)$ in standard format, i.e., the coefficients and a are all co-prime integers. Let U be the set of vertices which are endpoints of some edge x_e occuring with non-zero coefficient in the constraint $Q^T x \le a$. By definition of constraints of Cut_k, set U consists of at most k vertices. To prove that X satisfies the constraint, first, we will define a certain vector $\tilde{X}^{(U)}$ on U; secondly, we will prove that $\tilde{X}^{(U)}$ satisfies the constraint with a certain slackness; thirdly, we will prove that the restriction of X to U and $\tilde{X}^{(U)}$ are very close to each other; combining will prove the Proposition.

1. In order to define $\tilde{X}^{(U)}$, we first construct a tree $\tilde{T}$ with the following algorithm: initially we have a forest consisting of $|U|$ singleton trees, one for each vertex of U. Then, while there exist two trees in the forest within distance $< s$ of one another in G, connect them into a single tree with the shortest path. (Since $\mathrm{girth}(G) \ge 2(k-1)s$, that shortest path is unique.) We eventually get a forest $\{T_0, T_2, ...T_\ell\}$ covering U and such that in each tree T_i all the distances, and in particular the distances between the the points in $T_i \cap U$, are at most $(k-1)s$, and by the girth property the distances in T_i are the same as the distances in G. Furthermore, the distance between any two distinct trees is at least s. For $i = 0, 1, ...\ell$, we select one of the vertices of T_i, say r_i, as distinguished vertex. We link r_0 to $r_1, r_2, ...r_\ell$ by ℓ *fictitious paths* q_i each of length s. This defines tree $\tilde{T}$.

We then define a random process to label the vertices of $\tilde{T}$ with labels 0 or 1 as follows. Each edge of $\tilde{T}$ is deleted independently with probability 2ϵ. For each resulting subtree, some vertex is arbitrarily fixed as the root, and the vertices of the subtree are labeled as follows. The root is given the label 0 or 1 chosen uniformly at random. Proceeding by induction, if vertex u is a child of vertex v in the subtree, then u is given the label $\ell(u) = 1 - \ell(v)$. This defines the labeling. Let S be the (random) set of vertices in $V(\tilde{T})$ which are labeled 0, and let $\delta(S)$ be the corresponding cut vector of $V(\tilde{T})$. For $i, j \in U$, let $\tilde{X}_{ij} = E\delta(S)_{ij}$ be the probability that i and j are separated in the cut defined by S. This defines $\tilde{X}^{(U)}$.

2. The constraint $Q^T x \le a$ is (by definition of $\mathrm{Cut}(k)$) satisfied by every cut of U; but, by construction, $\tilde{X}^{(U)}$ is a convex combination of cuts of U. It follows that $\tilde{X}^{(U)}$ trivially satisfies the constraint: $Q^T \tilde{X}^{(U)} \le a$. Let us be more demanding and show that $\tilde{X}^{(U)}$ satisfies the constraint *with a certain slackness.*

LEMMA 2.1. *The cut polytope* $\mathrm{Cut}(k)$ *is full dimensional.*

Proof. See Appendix.

Appealing to Lemma 2.1, the cut polytope cannot lie entirely in the hyperplane of equation $Q^T x = a$, and so there must exist a cut $\delta(S^*)$ of U for which $Q^T\delta(S^*) \ne a$; since the coefficients are integer, that cut must be such that $Q^T\delta(S^*) \le a - 1$. Now, it is easy to see that the probability that cut $\delta(S^*)$ is the cut produced by our random labeling process is at least $(2\epsilon)^{k-1}(1/2)^{k-1} = \epsilon^{k-1}$. By linearity, this implies the desired slackness:

$$Q^T \tilde{X}^{(U)} \le \epsilon^{k-1}(a-1) + (1-\epsilon^{k-1})a = a - \epsilon^{k-1}.$$

3. It remains to relate $\tilde{X}$ to X.

LEMMA 2.2. *Let $x_i, x_j \in V(\tilde{T})$ and d be their distance in $\tilde{T}$. Then*

$$\tilde{X}_{ij} = \frac{1}{2}(1-(1-2\epsilon)^d) + (1-2\epsilon)^d \chi(d \text{ odd}).$$

Proof. If no edge of the path from x_i to x_j in $\tilde{T}$ is marked, then the labels alternate along the path, and $\delta(S)_{xz} = \chi(d \text{ odd})$. By independence of the marking, this event E has probability $(1-2\epsilon)^d$. Otherwise, at least one edge along the path is marked, the labels of x_i and of x_j are independent, and $E(\delta(S)_{x_i x_j} | \bar{E}) = 1/2$. That event has probability $1-(1-2\epsilon)^d$. Summing proves the lemma.

By Lemma 2.2 and the fact that small distances agree in G and in tree $\tilde{T}$, we have that $X_{ij} = \tilde{X}_{ij}^{(U)}$ whenever $d_{ij} < s$, and moreover:

$$\begin{aligned} &||X^{(U)} - \tilde{X}^{(U)}||_\infty \\ &= \max_{ij: d_{ij} \ge s} |X_{ij} - \tilde{X}_{ij}| \\ &\le |\frac{1}{2}(1-(1-2\epsilon)^s) + (1-2\epsilon)^s - \frac{1}{2}| \\ &= \frac{1}{2}(1-2\epsilon)^s \\ &\le \frac{1}{2}\exp(-2\epsilon s). \end{aligned}$$

LEMMA 2.3. *[7] Let ρ be the maximum value of the coefficients of the constraints of* Cut_k. *Then $\rho \le 2^{k^4}$.*

Putting everything together, we get

$$\begin{aligned} Q^T X^{(U)} &\le |Q^T X^{(U)} - Q^T \tilde{X}^{(U)}| + Q^T \tilde{X} \\ &\le \rho \binom{k}{2} ||X^{(U)} - \tilde{X}^{(U)}||_\infty + Q^T \tilde{X} \\ &\le 2^{k^4} k^2 \exp(-2\epsilon s) + a - (2\epsilon)^{k-1} \\ &\le a, \end{aligned}$$

where we used the definition of s and the fact that $k \ge 5$ and $\epsilon \le 1/4$ to infer $k^2 2^{k^4} \exp(-2\epsilon s) \le (2\epsilon)^{k-1}$. This concludes the proof of Proposition 2.1.

Proof. (of Theorem 2.1.) We fix an arbitrary ϵ with $0 < \epsilon \le 1/4$ and fix $k \ge 5$. Let $h = k^5 \epsilon^{-1} \log(1/\epsilon)$.

Let $G = G(n,p)$ be the random graph on n vertices with edge probability $p = C/n$. Assume that $C \ge 32\epsilon^{-2}$ and $n \ge \ln(20)/(1-\ln(2))$. The following Lemma is a variant of results of Poljak and Tuza [20].

LEMMA 2.4. *With probability at least $9/10$, we have*

$$\mathrm{Maxcut}(G)) \le |E(G)|(\frac{1}{2} + 2\epsilon).$$

Proof. see Appendix.

As argued by Poljak and Tuza [20] and by Arora, Bollobas and Lovasz [3], the expectation of the number of cycles of length at most $h-1$ in G is less than

$$\sum_{\ell=3}^{h-1} \frac{n^\ell}{\ell} \left(\frac{C}{n}\right)^\ell = \sum_{\ell=3}^{h-1} \frac{C^\ell}{\ell} \le \frac{C^h}{h}$$

Thus, with probability at least $9/10$, G has at most C^h cycles of length at most $h-1$. Choosing n large enough that $h \le \frac{1}{2}\log_C n$, we can "kill" all these cycles by suppressing at most $\sqrt{n}$ edges. (Note that this requires $k \in O(\log^{1/5} n)$, where the O hides a factor depending only on ϵ.). Call the new graph G'. With Lemma 2.4, we have then that, with probability at least $8/10$,

$$\mathrm{maxcut}(G')) \le |E(G')|(\frac{1}{2} + 2\epsilon).$$

Moreover, by construction G' has girth at least h.

Consider the linear program $LP(G')$ which is the strenghtening of the standard linear relaxation of MAXCUT by including the constraints of Cut_k, and let *opt* be the value of this program on G'. By Proposition 2.1, vector X from Definition 1 is feasible; since $X_{ij} = 1-\epsilon$ for every adjacent pair of vertices of G', the objective function has value $(1-\epsilon)|E(G')|$ for X. Thus the integrality gap is bounded below by

$$\frac{\mathrm{maxcut}(G')}{opt} \ge \frac{(1-\epsilon)|E(G')|}{(\frac{1}{2}+2\epsilon)|E(G')|} \ge 2 - 6\epsilon,$$

concluding the proof of Theorem 2.1.

We remark that, thanks to the bound from Lemma 2.3, we can actually strenghten the Theorem to keep the integrality gap at $2-\epsilon$ for all $k \le O(\log^{1/5} n)$.

3 Lift-and-project linear constraints

3.1 The Sherali-Adams lift and project method

Integer polytopes. Recall the definition of the cut polytope at the beginning of section 2.

For $t \geq 0$, let us define the *lifted integer cut polytope* $R_t(P)$ as the set of vectors (x, y), $x \in \mathbf{R}^m$, $y \in \mathbf{R}^{\binom{m}{2}+\binom{m}{3}+\cdots+\binom{m}{t+1}}$ such that

$$\begin{cases} x_e = \sum_S \lambda_S \chi(e \text{ crosses cut } (S, V \setminus S)) \\ y_I = \sum_S \lambda_S \chi(\forall e \in I, e \text{ crosses cut } (S, V \setminus S)) \\ \lambda_S \geq 0, \quad \sum_S \lambda_S = 1 \end{cases}$$

(Here x_e is defined for every e in F and y_I is defined for every subset I of F of size at most $t+1$.) In other words, (λ_S) defines a distribution over cuts and y_I is the probability that every edge of I belongs to the cut under this distribution.

Fractional relaxations As in the previous section, consider again the usual fractional polytope K, $K \subseteq \mathbf{R}^m$, defined by the constraints

$$\begin{cases} 2 - x_{ij} - x_{jk} - x_{ki} \geq 0 & \forall i, j, k \text{ distinct vertices} \\ -x_{ij} + x_{ik} + x_{kj} \geq 0 & \forall i, j, k \text{ distinct vertices} \\ x_{ij} \geq 0 & \forall i, j \text{ distinct vertices} \\ 1 - x_{ij} \geq 0 & \forall i, j \text{ distinct vertices} \end{cases}$$

It is well known that $P = \text{Conv}(K \cap \{0,1\}^m)$.

For $t \geq 0$, the Sherali-Adams lifted polytope $R_t(K)$ is the set of vectors (x, y), $x \in \mathbf{R}^m$, $y \in \mathbf{R}^{\binom{m}{2}+\binom{m}{3}+\cdots+\binom{m}{t+1}}$ such that the following constraints hold. Consider a constraint $a^T x - b \geq 0$ of P. Choose subsets I, J of V, with $|I \cup J| \leq t$. For every integer cut $(x_e) = (\delta(S)_e)$, we have:

$$(a^T x - b) \prod_{e \in I} x_e \prod_{f \in J} (1 - x_f) \geq 0.$$

This is a polynomial in (x_e) of degree at most $t+1$, which we can rewrite as

$$\sum_{U \subseteq V, |U| \leq t+1} \alpha_U \prod_{e \in U} x_e^{\beta_{e,U}} \geq 0, \quad \text{with } \beta_{e,U} \geq 1.$$

Equivalently, for every cut $\delta(S)$ we have

$$\sum_{U \subseteq V, |U| \leq t+1} \alpha_U \chi(\forall e \in U, e \text{ crosses cut } (S, V \backslash S)) \geq 0.$$

By linearity, for every $(x, y) \in R_t(P)$ we have

$$\sum_{U \subseteq V, |U| \leq t+1} \alpha_U y_U \geq 0. \tag{3.1}$$

These inequalities (3.1) are the constraints defining $R_t(K)$.

Projection $S_t(K)$, the polytope obtained from K by doing Sherali-Adams lift-and-project, is the set of vectors $x \in \mathbf{R}^m$ such that there exists a y with $(x, y) \in R_t(K)$.

Remark: The other types of lift-and-project can similarly be cast in a probabilistic framework. Instead of indexing y by a set of variables, index it by a *sequence* of variables, thus allowing the probability to depend on the order in which variables have been considered. To mitigate this effect, Lovasz and Shrijver add a constraint saying that the last two elements of the sequence must be able to commute.

Remark: In the semi-definite variant, one adds a constraint saying that when every element of I has been fixed except two, the resulting y-variables depend on the remaining two elements, and this two-dimensional matrix must be positive semi-definite.

3.2 Lift-and-Project Theorem

THEOREM 3.1. *Let K be the fractional polytope of the standard linear programming relaxation for MAX-CUT. For every ϵ and every fixed t, the integrality gap for the linear program obtained from K by doing the t rounds Sherali-Adams lift-and-project is at least $2 - \epsilon$.*

Proof. To prove Theorem 3.1, as in the proof of Theorem 2.1 we start with a high-girth graph G. We define $X = (X_{ij})$ as in definition 1, except that $s = 1 + (1/(2\epsilon)) \ln(2t^2(2/\epsilon)^{9t}) = \Theta((1/\epsilon) t \ln(1/\epsilon))$. However, in order to define Y, we must be more careful than in the previous section, because now $R_t(K)$ is not full dimensional. We will define Y_I in a iterative manner, by order of increasing value of $|I|$.

First, for each subset U of vertices of size at most $2t+3$, we define a distribution over cuts of U. Let i_0 be the element of U of smallest index. For each subset A of U containing i_0, let $p_{U,A}$ be the probability of the cut $(A, U \setminus A)$ according to the following inductive definition.Take $s = \Theta((1/\epsilon) t \ln(1/\epsilon))$ as defined above.

Initial case: $A = \{i_0\}, U = \{i_0, i_1\}$. Let d be the distance between i_0 and i_1 in G. We proceed as in Definition 1.
If $d < s$, then let

$$p_{U,A} = \frac{1}{2}(1 - (1-2\epsilon)^d) + (1-2\epsilon)^d \chi(d \text{ odd}),$$

and if $d \geq s$, then let $p_{U,A} = 1/2$.

We also set $p_{U,\{i_0,i_1\}} = 1 - p_{U,\{i_0\}}$ (empty cut).

Other base cases: $A = \{i_0\}, 3 \leq |U| \leq 2t+3$**.** We proceed as in Step 1 of the proof of Proposition 2.1. Starting from U, we define a tree T_U and a random labeling process on T_U (as in the proof of Proposition 2.1.) We define $p_{U,A}$ as the probability that $(A, U \setminus A)$ is the cut produced by the random process.

Inductive cases: $|A| \geq 2$**.** Take some element a of A, $a \neq i_0$. Set

$$p_{U,A} = p_{U\setminus\{a\},A\setminus\{a\}} - p_{U,A\setminus\{a\}}.$$

(Additionally, we replace $p_{U',A'}$ by $p_{U',U'\setminus A'}$ if A' does not contain the element of U' of minimum index.)

LEMMA 3.1. *For every set U of size at most $2t+3$, $(p_{U,A})_A$ defines a distribution over cuts of U.*

Proof. We need to prove that $p_{U,A}$ is non-negative and sums to 1.

We use a perturbation argument. Starting from U, consider the tree T_U and the random labeling process on T_U, as in the proof of Proposition 2.1. For every subset A of U containing i_0, define q_A to be the probability that $(A, U \setminus A)$ is the cut produced by the random process on T_U. By definition, $(q_A)_A$ is a distribution over cuts of U, and the minimum probability of any cut is at least ϵ^{2t+2}.

In the initial case $|U| = 2, |A| = 1$, as in the proof of Proposition 2.1 we see that $|p_{U,A} - q_A| \leq (1/2)\exp(-2\epsilon s)$.

In the general case $|A| \geq 2$, developing the recurrence formula for $p_{U,A}$ and seeing its similarity with the recurrence in Pascal's triangle, we see that $p_{U,A}$ is a linear combination of $\binom{|U|}{|A|}$ terms of the form $p_{U',A'}$ where $U' \subseteq U$ and A' is the singleton set containing the minimum element of U', with integer coefficients bounded by $t2^t$ (the t accounts for the possible t occurrences of the event when we replace $p_{U',A'}$ by $p_{U',U'\setminus A'}$.)

Since the formulas used in developing the recurrence hold for any distribution, using the same recurrence formulas for q_A, we can also rewrite q_A as the *same* linear combination of terms of the form $q_{A'}$.

For every subset U' of U, if A' is the element of U' of minimum index, then $p_{U',A'}$ is obtained by running the random process on tree $T_{U'}$. Since $T_{U'} \neq T_U$, this defines a probability which may be different from $q_{A'}$ in general. However, the two trees only differ in the tree paths which are longer than s, and so we have $|p_{U',A'} - q_{A'}| \leq (2t+2)(1/2)\exp(-2\epsilon s)$.

Plugging this into the linear combination, we deduce that, with $\binom{|U|}{|A|} \leq 2^{2t+3}$,

$$|p_{U,A} - q_A| \leq t2^{3t+3}(2t+2)(1/2)\exp(-\epsilon s) \leq$$
$$(1/2)\epsilon^{2t+2} \leq (1/2)q_A,$$

where we used our lower bound on s. This implies $p_{U,A} \geq 0$, as desired.

As for proving that $\sum_A p_{U,A} = 1$, we see that it is true for $|U| = 2$ and it is easy to verify it by induction on $|U|$.

Now we need to define variable Y_I, for I any set of at most $t+1$ pairs of vertices of V. Let U be the set of endpoints of edges in I. We define Y_I as the sum, over every cut $(A, U \setminus A)$ of U which contains all edges of I, of $p_{U,A}$. (Note that when $|I| = 1$, this is consistent with the definition of X.)

It only remains to verify that (X, Y) satisfies all the constraints of $R_t(K)$. Consider any constraint of $R_t(K)$, obtained from the polynomial inequality of the lifted polytope, for example

$$\prod_{e \in I} x_e \prod_{f \in J} x_f (x_{ij} + x_{jk} + x_{ki} - 2) \leq 0.$$

The endpoints of the edges appearing in this constraint span a set U of at most $2t+3$ vertices, hence is a set on which Lemma 3.1 holds, and so there exists a distribution of cuts on U associated to the restriction of (X, Y) to U, and so the constraint is valid for (X, Y).

By projection, this implies that $X \in S_t(K)$, and the rest of the proof goes along the same lines as for Theorem 2.1.

This ends the proof of Theorem 3.1.

4 Gap reduction in dense graphs

The fact that dense MAXCUT has a PTAS was discovered in the middle of the last decade [4, 9], see also [13, 10, 11]. The known PTASs for this problem are randomized, although they can be de-randomized, and are usually quite involved. We give here a new and very natural deterministic PTAS.

Let P denote the standard relaxation for MAXCUT: For each positive integer k, we denote by $N^k(P)$ the relaxation of the cut polytope obtained by applying k rounds of the lift-and-project operator of Lovàsz and Schriver to P. Recall that the density $d(G)$ of a graph $G = (V(G), E(G))$ is defined by $d(G) = |E(G)|\binom{|V(G)|}{2}^{-1}$. Then:

THEOREM 4.1. *For each $\epsilon > 0, d > 0$, there is an integer $g = g(\epsilon, d)$ with the property that the relaxation $N^g(P)$ has integrality gap at most $1+\epsilon$ on the set of graphs of density at least d.*

Proof The proof uses a result of [2]. We need some preparation to state this result. Recall that the input to a MAX-rCSP problem (for r fixed) consists of a set F of m distinct Boolean functions $f_1, f_2, \dots f_m$ of n Boolean variables $x_1, x_2, \dots x_n$, where each f_i is a function of only r of the n variables. The answer $\max(F)$ is the maximum number of functions which can be simultaneously set to 1 by a truth assignment to the variables. For a subset Q of the variables, we let F^Q denote the subset of F which are functions of only the variables in Q (and their negations). The following theorem is proved in [2].

THEOREM 4.2. *Let r, n be positive integers, with r fixed. Suppose ϵ is a positive real. There exists a positive integer $q \in O(\log(1/\epsilon)/\epsilon^4)$ such that for any F (as above), if Q is a random subset of $\{x_1, x_2, \dots x_n\}$ of cardinality q, then with probability at least 9/10, we have*

$$|\frac{n^r}{q^r}\max(F^Q) - \max(F)| \le \epsilon n^r.$$

We shall use this theorem with a set F corresponding to MAXCUT of a graph $G = (V, E)$, $(r = 2)$. (Here the f are indexed by the literal pairs u, v and we have $f_{u,v} = e_{\tilde{u},\tilde{v}}(u(1-v) + v(1-u))$ where $\tilde{u}, \tilde{v}$ are the variables corresponding to u and v, and $(e_{\tilde{u},\tilde{v}})$ is the indicator function of the edges of the graph.) The theorem says in this case that we have with probability at least 9/10,

$$(4.2)\quad |\frac{n^2}{q^2}\text{maxcut}(G(Q)) - \text{maxcut}(G)| \le \epsilon n^2.$$

Actually, the probability can be pushed up to $1-\epsilon$ by adding an extra log factor to the sample size. We fix the density d and ϵ. We need 4.2 to hold with accuracy ϵd which will hold for $q \in O(\log^2(1/\epsilon)\epsilon^{-4}d^{-4})$. Suppose that $(U, V\backslash U)$ is a partition of V for which $|\delta(U)| = \text{maxcut}(G)$. For each subset $S \in \binom{V}{q}$ define $S^{(U)} = S \cap U$. We have then

$$\begin{aligned}\binom{n}{q}\frac{q^2}{n^2}\text{maxcut}(G) &\sim \sum_{S\in\binom{V}{q}} |\delta(S^{(U)})| \\ &\le \sum_{S\in\binom{V}{q}} \text{maxcut}(G(S))\end{aligned}$$

By theorem 4.2 we have that the proportion of sets S for which

$$\text{maxcut}(G(S)) \ge (1+\epsilon d)\frac{q^2}{n^2}\text{maxcut}(G)$$

does not exceed ϵ. This implies, using the trivial upper bound $\text{maxcut}(G(S)) \le q^2/2$,

$$\begin{aligned}\sum_{S\in\binom{V}{q}}\text{maxcut}(G(S)) &\le \binom{n}{q}\frac{q^2}{n^2}\text{maxcut(G)} + \epsilon d\binom{n}{q}\frac{q^2}{2} \\ &\le \binom{n}{q}\frac{q^2}{n^2}\text{maxcut(G)}(1+\epsilon)\end{aligned}$$

the last by using the density condition. Let us take g such that the inequalities with support of size at most q are satisfied by $N^g(P)$. ($g = q$ will do.) Then we have that for each S the value of the cut induced on S does not exceed $\text{maxcut}(G(S))$. The value alg, say, output by using the relaxation $N^g(P)$ satisfies thus

$$\begin{aligned}\text{alg} &\le \binom{n}{q}\frac{q^2}{n^2}\sum_{S\in\binom{V}{q}}\text{maxcut}(G(S)) \\ &\le (1+\epsilon)\text{maxcut(G)}\end{aligned}$$

using the preceding inequality. This concludes the proof.

5 Conclusion

We expect that with a similar proof, one might be able to show that the integrality gap of the Vertex Cover LP stays at $2 - \epsilon$ even after applying Sherali-Adams lift-and-project (note that the proof from [3] only applies to the weaker Lovasz-Shrijver lift-and-project). The corresponding random process on a tree would delete edges independently, label both endpoints of every deleted edge with a 1 and remove them, and alternate labels in each remaining connected component starting from a label chosen uniformly at random as 0 or 1. This could also naturally extend to hypergraph vertex cover.

In general, it seems likely that the methods developed in this paper extend to proving lower bounds for other problems as well. The most interesting open question, of course, would be to deal with the semidefinite programming variants of lift-and-project, but that appears much more challenging.

References

[1] Michael Alekhnovich, Sanjeev Arora, Iannis Tourlakis: Towards strong nonapproximability results in the Lovasz-Schrijver hierarchy. STOC 2005: 294-303

[2] N. Alon, W. Fernandez-de-la-Vega, R. Kannan and M. Karpinski, Random Sampling and approximation of MAX-CSP's, Journal of Computer and System Sciences, 67 (2003) 212-243.

[3] Arora, Bollobas, and Lovasz. Proving integrality gaps without knowing the linear program. In FOCS: IEEE Symposium on Foundations of Computer Science (FOCS), 2002.

[4] S.Arora, D.Karger and M.Karpinski, *Polynomial time approximation schemes for dense instances of NP-hard problems*, Proceedings of the 27th Annual ACM Symposium on Theory of Computing (1995) 284-293; the full paper appeared in J. Comp. System. Sci. **58** (1999), 193-210.

[5] Balas, E., Ceria, S. and Cornuejols, G. (1993), "A Lift-and-project Cutting Plane Algorithm for Mixed 0-1 Programs," Mathematical Programming 58: 245324.

[6] Josh Buresh-Oppenheim, Nicola Galesi, Shlomo Hoory, Avner Magen, Toniann Pitassi: Rank Bounds and Integrality Gaps for Cutting Planes Procedures Joshua. FOCS 2003: 318.

[7] Michel Deza and Monique Laurent Geometry of Cuts and Metrics. Springer-Verlag, Berlin, 1997.

[8] Uriel Feige, Gideon Schechtman: On the integrality ratio of semidefinite relaxations of MAX CUT. STOC 2001: 433-442

[9] W.Fernandez-de-la-Vega, *MAX-CUT has a Randomized Approximation Scheme in Dense Graphs*, Random Structures and Algorithms 8 (1996) 187-199.

[10] A.M.Frieze and R.Kannan, *The Regularity Lemma and approximation schemes for dense problems*, Proceedings of the 37th Annual IEEE Symposium on Foundations of Computing, (1996) 12-20.

[11] A. M. Frieze and R. Kannan *Quick Approximation to matrices and applications*, Combinatorica **19** (2) (1999) pp 175-200.

[12] M.X. Goemans and D.P. Williamson, Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming, J. ACM, 42, 1115–1145, 1995.

[13] O.Goldreich, S.Goldwasser and D.Ron, *Property testing and its connection to learning and approximation*, Proc. 37th IEEE FOCS 96, 339-348; the full paper appeared in J. ACM **45** (1998), 653-750.

[14] Howard Karloff. How good is the Goemans-Williamson MAX CUT algorithm? In Proceedings of the 28th Annual ACM Symposium on Theory of Computing, 1996, 427–434.

[15] Subhash Khot, Nisheeth K. Vishnoi. The Unique Games Conjecture, Integrality Gap for Cut Problems and the Embeddability of Negative Type Metrics into l1, FOCS 2005.

[16] Subhash Khot, Guy Kindler, Elchanan Mossel, Ryan O'Donnell. Optimal Inapproximability Results for Max-Cut and Other 2-Variable CSPs? 45th Annual IEEE Symposium on Foundations of Computer Science (FOCS'04) pp. 146-154

[17] Monique Laurent. A comparison of the Sherali-Adams, Lovasz-Schrijver, and Lasserre relaxations for 0-1 programming. Mathematics of Operations Research, August, 2003.

[18] B Lasserre, An Explicit Exact. SDP Relaxation for Nonlinear 0-1 Programs. SIAM J. Optim. 12, pp. 756-769, 2002.

[19] L. Lovasz and A. Schrijver, Cones of matrices and setfunctions, and 0-1 optimization, SIAM J. of Optimization 1 (1990) 166-190.

[20] S Poljak and Zs. Tuza. On the expected relative error of the polyhedral approximation of the max-cut, Oper. Res. Letters 16 (1994) 191-198.

[21] H. Sherali and W. Adams, A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems, SIAM J. Disc. Math. 3:411-430, 1990.

A Proof of Lemma 2.1

We prove in this appendix that the cut polytope Cut_k is full dimensional for $k \geq 5$. We prove in fact that the set of cut vectors

$$B = \left\{ \delta(\{i,j\}) : \{i,j\} \in \binom{[k]}{2} \right\}$$

i.e., the set of cuts defined by the $\binom{k}{2}$ distinct pairs of vertices, is a basis of Cut_k. Actually, we will prove simultaneously that B is free and generating. We fix an arbitrary unit vector, say e_{ij} and express it as a function of the $\delta(i,j)$ We write

$$u = \alpha\delta(\{i,j\}) + \beta\sum_{k\neq i}\delta(\{i,k\}) + \beta\sum_{k\neq j}\delta(\{j,k\}) + \gamma\sum_{\{k,\ell\}\cap\{i,j\}=\emptyset}\delta(\{k,l\})$$

and compute α, β, γ to get $u = e_{ij}$. Inspection gives

$$\begin{aligned} u_{ij} &= 2\beta(n-2) \\ u_{ik} &= \alpha + \beta(n-2) + \gamma(n-3) \quad k \neq i \\ u_{kl} &= 4\beta + 2\gamma \quad k, \ell \neq i, j. \end{aligned}$$

Thus we need $\beta = \frac{1}{2(n-2)}$ to get $u_{ij} = 1$, and then successively $\gamma = -2\beta = -\frac{1}{n-2}$ and $\alpha = -\beta(n-2) - \gamma(n-3) = -\frac{1}{2} + \frac{n-3}{n-2}$ to annihilate the other components of u. Thus B generates indeed. Since the coefficient α of $\delta(i,j)$ is forced to be positive, we see that B is free.

B Proof of Lemma 2.4

Let $G = G(n,p)$ be the random graph on n vertices with edge probability $p = C/n$ and let $S \subseteq V(G)$ with $s = |S| \leq n/2$. The value $|\delta(S)|$ of the cut defined by S is distributed as a Binomial random variable $B(m,p)$ with $m = s(n-s)$. Therefore, using Hoeffding-Chernoff,

$$\begin{aligned} &\Pr\left(|\delta(S)| \geq \frac{n^2p}{4}(1+\epsilon)\right) \\ &\leq \exp\left(-\frac{1}{2}\frac{(\frac{n^2p}{4}(1+\epsilon) - s(n-s)p)^2}{s(n-s)p}\right) \\ &\leq \exp\left(-\frac{\epsilon^2 n^2 p}{32}\right) \\ &\leq \exp\left(-\frac{\epsilon^2 C n}{32}\right) \end{aligned}$$

This implies

$$\begin{aligned} \Pr(\mathrm{maxcut}(G) \geq \frac{n^2p}{4}(1+\epsilon)) &\leq 2^n \exp\left(-n(\frac{\epsilon^2 C}{32})\right) \\ &\leq 1/20 \end{aligned}$$

for $C \geq 32\epsilon^{-2}$ and $n \geq \ln(20)/(1-\ln(2))$. We also have then

$$\Pr(|E(G)| \leq \binom{n}{2}p(1+\epsilon)) \geq 19/20$$

so that, with the previous inequality, and with probability at least 9/10,

$$\mathrm{maxcut}(G)) \leq |E(G)|(\frac{1}{2} + 2\epsilon), \tag{2.3}$$

as desired.