On the Sample Complexity of MAX-CUT

W. Fernandez de la Vega [*]   Marek Karpinski [†]

Abstract

We give a *simple* proof for the *sample complexity* bound $O^{\sim}(1/\varepsilon^4)$ of absolute approximation of MAX-CUT. The proof depends on a new analysis method for linear programs (LPs) underlying MAX-CUT which could be also of independent interest.

1 Introduction

The purpose of this paper is to give a simple proof based on linear programs of the following theorem proven in more generality in [AFKK02].

Theorem 1 (Main Theorem). *For any positive ε, there exists an integer $q \in O(\log^3(1/\varepsilon)/\varepsilon^4)$ such that for any graph $G = \{V, E\}$ if Q is a random subset of V of cardinality q and $G(Q)$ is the restriction of G to the sample Q, then with probability at least 9/10, we have,*

$$\left| \frac{n^2}{q^2} \mathrm{maxcut}(G(Q)) - \mathrm{maxcut}(G) \right| \leq \varepsilon n^2.$$

where $\mathrm{maxcut}(G)$ *is the maximum value of a cut of G.*

[*] LRI, CNRS, Université de Paris-Sud. Research partially supported by the IST grant 14036 (RAND-APX), and by the PROCOPE project. Email: lalo@lri.fr

[†] Dept. of Computer Science, University of Bonn. Research partially supported by DFG grant, PROCOPE project, and by IST grant 14036 (RAND-APX). Email: marek@cs.uni-bonn.de

It is apt to rephrase this theorem by saying that the sample complexity of MAX-CUT is in $O(\log^3(1/\varepsilon)/\varepsilon^4)$. A more general theorem was proved in [AFKK02]. [AFKK02] uses besides integer programs also special cut-norm and cut-array framework developed for that purpose (cf. also [AN04]). The sample bound $O^{\sim}(\frac{1}{\varepsilon^4})$ for MAX-CUT is the best known up to date (see [AKK95], [F96], [FK99] for early approximation algorithms for that problem), although there were several attempts to improve it. Any improvement of that bound would be of course an important contribution to the area (given the fact that $1/\varepsilon$ factors could be very large in various settings). [RV05] improved recently estimates for the *cut-norm* of random submatrices used in more generality in [AFKK02]. That however does not influence overall sampling bound $O^{\sim}(1/\varepsilon^4)$ for the MAX-CUT, see Errata of [RV05].

We will make essential use in our proof of the following theorem of [AFKK02]. It asserts that for a Linear Program P on n variables, each constrained to be between 0 and 1, we can make some assertion about the optimal value based on the optimal value of a small subprogram obtained by picking at random a small number of variables.

Theorem 2. *Suppose*

$$\boxed{\begin{array}{l} \mathsf{Max} \sum_{j=1}^{n} c_j x_j < \alpha \\ \\ \sum_{j=1}^{n} U_j x_j \leq v \qquad ; \qquad 0 \leq x_j \leq 1, \end{array}}$$

where each U_j is an $m-$vector. Suppose q is a positive integer and Q is a random subset of $\{1, 2, \ldots n\}$ of cardinality q. Then, for any positive real number λ, with probability at least $1 - 4e^{-\lambda^2/4}$, we have :

$$\begin{aligned} Max \sum_{j \in Q} c_j x_j &\leq \frac{q}{n}\alpha + \lambda\sqrt{q}||c||_\infty \\ \sum_{j \in Q} U_j x_j &\leq \frac{q}{n}v - \lambda\sqrt{q}||U||_\infty \qquad ; \qquad 0 \leq x_j \leq 1, j \in Q. \end{aligned}$$

Proof. See ([AFKK02]).

■

2 Preliminaries

Let $G = (V, E)$ be a graph. Suppose that (V_L, V_R), $V_R = V \backslash V_L$ is a bipartition of G. We use $x_i = 0$ (resp. $x_i = 1$) to indicate that the vertex v_i belongs to V_L (resp. that v_i belongs to V_R). We let e_{ij} be the indicator function of the edges of G:

$$e_{ij} = 1 \text{ if } v_i v_j \in E, \quad e_{ij} = 0 \text{ otherwise.}$$

Then, clearly, for each i, the number of neighbours of v_i in V_R, say ρ_i, satisfies

$$\rho_i = |\Gamma(v_i) \cap V_R| = \sum_{j=1}^{n} x_j e_{ij}$$

The value of the cut defined by (V_L, V_R) is clearly

$$e(V_L, V_R) = \sum_{j=1}^{n} (1 - x_j) \rho_j.$$

For each j, let ρ_j^* denote an estimate of ρ_j. We can obtain these estimates by guessing a sample $S \subseteq V_R$ of a certain size m, say, and putting

$$\rho_j^* = \frac{|V_R|}{|S|} |\Gamma(v_j) \cap S|$$

Of course this will entail some error. For an appropriate size $|S|$ we may expect that the inequality

$$\rho_j^* - \varepsilon n \leq \rho_j \leq \rho_j^* + \varepsilon n \tag{1}$$

holds simultaneously for all, or nearly all vertices v_j. In [AKK95], the Set S has size $\Theta(\log n)$ and it is shown that then (1) holds with high probability for all v_i (if $S \subseteq V_R$). The following is shown

Proposition 1 ([AKK95]). *Assume that (V_L, V_R) is an optimum cut of G. Assume that S is a random sample of V_R with size $\Theta(\log(1/\varepsilon)\varepsilon^{-3} \log n)$. Then, with probability at least $2/3$, $\mathrm{maxcut}(G)$ is approximated with additive error $O(\varepsilon n^2)$ by applying randomized rounding to the solution of the LP*

Maximize $e(V_L, V_R) = \sum_{j=1}^{n}(1 - x_j)\rho_j^*$

Subject to

$$\begin{array}{l} \rho_j^* - \varepsilon n \leq \rho_j \leq \rho_j^* + \varepsilon n \\ 0 \leq x_j \leq 1, \quad 1 \leq j \leq n. \end{array}$$

Proof. See ([AKK95]).

■

(Technically the sample S is taken in turn to be each of the subsets of a fixed random sample T of appropriate size so that the inclusion $S \subseteq V_R$ need occur at least for one S.) We cannot use directly this result here since our intended sample size is only a constant (dependent on ε.) However, in this later case, we can still assert the following.

Claim 1. *Let (Q_L, Q_R) define an optimum cut of $G(Q)$. Take $m = \lambda \log(1/\varepsilon)\varepsilon^{-2}$ and assume that S is a random subset of Q_R of size m. Then, we have that, for any fixed j, with probability at least $1 - 2\varepsilon^{\lambda/2}$,*

$$\rho_j^* - \varepsilon n \leq \rho_j \leq \rho_j^* + \varepsilon n \tag{2}$$

Also, for $\lambda \geq 3$, the inequality (2) holds with probability at least $(1 - \varepsilon)$ for at least $q(1 - \varepsilon)$ vertices in Q.

If the above condition holds we shall say that S is *ε-representative* with respect to Q_R.

Proof. Let $q_R = |Q_R|$, $k = |\Gamma(v_j) \cap Q_R|$. Then $|\Gamma(v_j) \cap S|$ has the hypergeometric distribution with parameters Q_R, k, m which is tighter than the Binomial distribution $B(m, p)$ with parameters m and $p = k/q_R$. By a standard Hoeffding-Chernoff Bound we have

$$\begin{aligned} \Pr(||\Gamma(v_j) \cap S| - \frac{mk}{|Q_R|}| \geq \varepsilon m) &\leq 2\exp(-\frac{\varepsilon^2 q_k m}{2k}) \\ &\leq 2\exp(-\frac{\varepsilon^2 m}{2}) \\ &\leq 2\varepsilon^{\lambda/2} \end{aligned}$$

Using Markov inequality we get that the number of vertices for which (2) holds is at least $(1 - \varepsilon q)$ with probability at least $1 - 2\varepsilon^{\lambda/2-1} \geq 1 - \varepsilon$ for $\lambda = 4$.

■

The following proposition will be helpful

Proposition 2. *Let the ρ_j^* be any fixed constants (no relation between these ρ_j^* and the graph G is assumed here) and consider the LP*

Maximize $\sum_{j=1}^{n}(1-x_j)\rho_j^*$

Subject to

$$\rho_j^* - \varepsilon n \leq \rho_j \leq \rho_j^* + \varepsilon n$$
$$0 \leq x_j \leq 1,\ \ 1 \leq j \leq n.$$

Then, if this LP is feasible, the value $\mathrm{val}(P)^*$ *obtained by applying randomized rounding to the solution of this LP satisfies*

$$\mathrm{val}(P)^* \leq \mathit{maxcut}(G) + O(\varepsilon n^2)$$

Proof. The proof is implicit in [AKK95].

■

The following proposition extends Proposition 1 to the case where the approximation provided by the ρ_j^* fails on a small subset of vertices.

Proposition 3. *Assume that (V_L, V_R) is an optimum cut of G and assume that we have*

$$\rho_j^* - \varepsilon n \leq \rho_j \leq \rho_j^* + \varepsilon n$$

for each j in a set J, say, where $J \subseteq \{1, 2, ..n\}$ has size $|J| \geq (1-\varepsilon)n$. Then $\mathrm{maxcut}(G)$ *is approximated within $O(\varepsilon n^2)$ by applying randomized rounding to the solution of the LP*

Program LP0

Maximize $\sum_{j=1}^{n}(1-x_j)\rho_j^*$

Subject to

$$\rho_j^* - \varepsilon n \leq \rho_j \leq \rho_j^* + \varepsilon n,\ j \in J \text{ with } |J| \geq \varepsilon q$$
$$0 \leq x_j \leq 1,\ \ j = 1, \ldots, n.$$

We note that we do not know precisely the set J. We only have a lower bound on its cardinality.

Proof. Let us denote by $x = (x_j)$ an optimum solution of $LP0$, $\mathrm{val}^*(x) =$

$\sum_{j=1}^{n}(1-x_j)\rho_j^*$ its value. Define $\mathrm{val}(x) = \sum_{j=1}^{n}(1-x_j)\rho_j$. Then we have

$$\begin{aligned}|\mathrm{val}(x) - \mathrm{val}^*(x)| &\leq \sum_{j=1}^{n} |\rho_j - \rho_j^*| \\ &\leq \varepsilon(1-\varepsilon)n^2 + \varepsilon n^2 \\ &\leq 2\varepsilon n^2.\end{aligned}$$

Note that we have

$$\mathrm{maxcut}(G) \leq \max \sum_{j=1}^{n}(1-x_j)\rho_j$$

which implies

$$\mathrm{val}^*(x) \geq \mathrm{maxcut}(G) - 2\varepsilon n^2. \tag{3}$$

Let y be obtained from x by randomized rounding Then, $\mathrm{val}(y)$ is the value of the cut defined by y. From standard results we have that $|\mathrm{val}^*(x) - \mathrm{val}^*(y)| \in O(\varepsilon n^2)$ w.h.p. implies with (3)

$$|\mathrm{val}(y) - \mathrm{maxcut}(G)| \in O(\varepsilon^2 n).$$

■

3 End of the proof

Let Q be a random subset of vertices of size $q = \log^3(1/\varepsilon)\varepsilon^{-4}$. and let T be a random subset of Q of size $t = C\log(1/\varepsilon)\varepsilon^{-2}$ so that T is also a random subset of V. We can restrict ourselves to sets S ε representative for $G(Q)$ (since with high probability we have $\mathrm{cut}_S(G(Q)) = \mathrm{maxcut}(G(Q)) + O(\varepsilon q^2)$ for such a set S.) Define

$$\mathrm{Re}(G(Q)) = \{S \subseteq T : S \text{ is } \varepsilon - \text{representative for } G(Q)\}.$$

For each $S \in \mathrm{Re}(G(Q))$ we introduce the following LP's, LP1, LP2, LP3. So in each case, the estimates ρ_j^* are with reference to S. Also, we will consider in the sequel only the ordinary solutions to these programs (not the integer solutions) since randomized rounding plays a rather trivial role here.

Program LP1

Here the ρ_j are evaluated in Q:

$$\begin{aligned}\rho_j &= \sum_{i\in Q} x_i e_{ij}\\ \rho_j^* &= \frac{|V_R|}{|S|}|\Gamma(v_j)\cap S|\end{aligned}$$

Maximize $\sum_{j\in Q}^{n}(1-x_j)\rho_j^*$

Subject to

$\frac{q}{n}\rho_j^* - \frac{\varepsilon q}{2} \le \rho_j \le \frac{q}{n}\rho_j^* + \frac{\varepsilon q}{2},\quad j\in J^{(Q)}$

say

$0\le x_j\le 1,\quad j\in Q$

and we are assuming that $|J^{(Q)}|\ge(1-\varepsilon)q$.

Program LP2

Here the ρ_j^* are as above and the ρ_j are evaluated in G:

$$\rho_j = \sum_{i=1}^{n} x_i e_{ij}$$

Maximize $\frac{n}{q}\sum_{j=1}^{n}(1-x_j)\rho_j^*$

Subject to

$\frac{n}{q}\rho_j^* - \varepsilon n \le \rho_j \le \frac{n}{q}\rho_j^* + \varepsilon n,\ j\in J$

$0\le x_j\le 1.$

Here, $J\subseteq\{1,2,..n\}$ is a set of indices of size $|J|\ge(1{-}2\varepsilon)n$.

Program LP3

With ρ_j and ρ_j^* as in LP1 we set here

$$\rho_j = \sum_{i=1}^{n} x_i e_{ij}$$

Maximize $\frac{n}{q} \sum_{j \in Q}^{n} (1 - x_j)\rho_j^*$

Subject to

$$\rho_j^* - \tfrac{\varepsilon q}{2} \leq \rho_j \leq \rho_j^* + \tfrac{\varepsilon q}{2}, \quad j \in J^{(Q)}$$

say

$$0 \leq x_j \leq 1, \quad 1 \leq j \leq n$$

Clearly, the solution of LP3 is just equal to that of LP1 multiplied by $\frac{n}{q}$. Note that LP3 is a random subprogram of LP2. In view of a previous claim we need only compare, for each $S \in \mathrm{Re}(G(Q))$ the solutions of the programs LP3 and LP2. We use Theorem 2 in contrapositive form to assert that the value of LP2 is with sufficiently probability as big as the scaled value of LP1 and repeat this reasoning for each $S \in \mathrm{Re}(G(Q))$ ending with the conclusion that the assertion is true w.h.p. for all these S together and theorem (1) follows. (We also use Proposition 2 to assert that the maxcut(G) is at least nearly as big as the value of LP2.) We take $\lambda = 4\varepsilon^{-1}\sqrt{\log(1/\varepsilon)}$ in that theorem which implies $\lambda\sqrt{q} \leq \varepsilon q$. Thus we can use Theorem 2 with that λ to assert that with probability at least $1 - 4e^{-\varepsilon^{-2}\log(1/\varepsilon)}$,

$$\mathrm{val}(LP2) - \frac{q}{n}\mathrm{val}(LP3) \leq eqn$$

and, with $\mathrm{val}(LP3) = \frac{q}{n}\mathrm{val}(LP1)$, we get

$$\mathrm{val}(LP2) - \frac{q^2}{n^2}\mathrm{val}(LP3) \leq eq^2$$

Since there are at most $2^{|T|} = 2^{\log(1/\varepsilon)\varepsilon^{-2}}$ distinct possible choices for S, this will be true simultaneously for all these choices with probability at least 2/3 for each fixed sufficiently small ε. Theorem 1 follows.

■

Acknowledgments

We thank Ravi Kannan for stimulating remarks and discussions.

References

[AFKK02] N. Alon, W. Fernandez de la Vega, R. Kannan and M. Karpinski, *Random Sampling and Approximation of MAX-CSP Problems*, Proc. 34th ACM STOC (2002), pp. 232-239; journal version in J. Comput. System Sciences 67 (2003), pp. 212-243.

[AN04] N. Alon and A. Naor, *Approximating the Cut-Norm via Grothendieck's Inequality*, Proc. 36th ACM STOC (2004), pp. 72-80; journal version in SIAM J. Compu. 35 (2006), pp. 787-803.

[AKK95] S. Arora, D. Karger and M. Karpinski, *Polynomial Time Approximation Schemes for Dense Instances of NP-Hard Problems*, Proc. 27th ACM STOC (1995), pp. 284-293; journal version in J. Comput. System Sciences 58 (1999), pp. 193-210.

[F96] W. Fernandez de la Vega, *MAX-CUT has a Randomized Approximation Scheme in Dense Graphs*, Random Structures and Algorithms 8 (1996), pp. 187-198.

[FK99] A.M. Frieze and R. Kannan, *Quick Approximation to Matrices and Applications*, Combinatorica 19 (1999), pp. 175-200.

[RV05] M. Rudelson and R. Vershynin, *Sampling from Large Matrices: An Approach through Geometric Functional Analysis*, ArXiv: math.FA/0503442, 2005 (see also *Errata* ArXiv: math.FA/0503442v2, Aug. 2006).