Electronic Colloquium on Computational Complexity, Report No. 101 (2006)

Approximation Complexity of Nondense Instances of MAX-CUT

W. Fernandez de la Vega * Marek Karpinski †

Abstract

We prove existence of *approximation schemes* for instances of MAX-CUT with $\Omega(\frac{n^2}{\Delta})$ edges which work in $2^{O^{\sim}(\frac{\Delta}{\varepsilon^2})}n^{O(1)}$ time. This entails in particular existence of *quasi-polynomial* approximation schemes (QPTASs) for *mildly sparse* instances of MAX-CUT with $\Omega(\frac{n^2}{\operatorname{polylog} n})$ edges. The result depends on new sampling method for *smoothed* linear programs that approximate MAX-CUT.

On the other hand, we rule out existence of *polynomial time approximation schemes (PTASs)* for MAX-CUT instances with $\Omega(n^{2-\delta})$ edges for all $\delta > 0$, under the standard complexity theoretic assumptions.

1 Introduction

We study the problem of approximability of *nondense* instances of MAX-CUT. This complements the stream of results on existence of PTASs for *dense* or *subdense* instances of that problem (cf. [AKK95], [F96], [FK96], [GGR96], [FK00], [K01], [AFKK02], [FK02], [B05]). The paper aims at clarifying the inherent limits for existence of PTASs for instances of MAX-CUT which are parameterized by their densities. The algorithmic connections with Szemerédi's Regularity Lemma discovered in [FK96], raise also questions on their algorithmic nondense analogs. There were also recent breakthrough extensions of Szemerédi's Theorem to *nondense* settings dealing with prime numbers (cf. [GT06]). Back to the approximability of MAX-CUT on instances of arbitrary density. It is known that the (strongly) sparse instances of MAX-CUT with $\Theta(n)$ edges can be approximated with better factor ([FKL02]) than 1.1383 of Goemans-Williamson approximation algorithm ([GW95]). On the other hand, it is known that the

*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 done while visiting CSAIL, MIT and Microsoft Research, Redmond. Research partially supported by DFG grant, PROCOPE project, and by IST grant 14036 (RAND-APX). Email: marek@cs.uni-bonn.de

ISSN 1433-8092

general instances of MAX-CUT are *harder* to approximate, and 1.1383 is an approximation lower bound under the Unique Game Conjecture [KKMD04]. On the other end, dense instances are known to have PTASs.

In this paper we prove that the class of instances of MAX-CUT with $\Theta(n^{2-\delta})$ edges for any $\delta > 0$ is hard to approximate to within a constant. Furthermore, we prove existence of approximation schemes for instances of MAX-CUT with $\Omega(\frac{n^2}{\Delta})$ edges for arbitrary density $\frac{1}{\Delta}$, working in time $2^{O^{\sim}(\Delta/\varepsilon^2)} \cdot n^{O(1)}$. The main problem in designing approximation schemes for the classes of instances of arbitrary density is a possible wide distribution of their degrees. The method of solution is based on the smoothed analysis of linearized quadratic programs for MAX-CUT. The similar technique can be used for the classes of MAX-2CSP problems with $\Omega(\frac{n^2}{\Delta})$ constraints. Our results can be also generalized to MAX-rCSP problems with arbitrary $r \geq 2$, by using the recursive decomposition method of [AKK95] applied to the smooth r-degree programs.

2 Inapproximability Result

We prove the following approximation hardness result for MAX-CUT restricted to graphs having $\Omega(n^{2-\delta})$ edges.

Theorem 1. *For every $\delta > 0$, the class $\mathcal{G}$ of all graphs $G = (V, E)$ satisfying to*

$$|E| \in \Omega(|V|^{2-\delta})$$

is hard to approximate to within a constant.

Proof. We start with a cubic graph $G = (V(G), E(G))$, [BK99, BK03] proves explicit inapproximability bounds for that class of graphs. We introduce a parameter k, an exact value of which will be computed later. We associate with each $v \in V(G)$ a set U of k new vertices. Let $U = \bigcup_{v \in V(G)} U_v$. Map each cubic graph $G = (V(G), E(G))$ into a new graph $\Phi(G) = H = (V(H), E(H))$ with $V(H) = U$ and $\{x_i, x_j\} \in E(H)$ iff $x_i \in U_v$, $x_j \in U_w$ and $\{v, w\} \in E(G)$. Each V_i spans an independent set of H.

Now take G to be a cubic graph on $2m$ vertices and fix k. Then $|V(H)| = 2km$ and

$$\begin{aligned} |E(H)| &= k^2|E(G)| = 3k^2 m \\ &= (2km)^{2-\delta} \\ &= |V(H)|^{2-\delta} \end{aligned}$$

by choosing $k = (1/6)3^{\epsilon} m^{\delta^{-1}-1}$.

Now we claim that MAX-CUT on cubic graphs G *approximation preserving E-reduces* (cf. [KMSV98]) to MAX-CUT on graphs H with $(2km)^{2-\delta}$ edges. Clearly, we can construct H from G in polynomial time. (Note however that the degree of the relevant polynomial is roughly δ^{-1}.) Assume now that we have a PTAS for MAX-CUT on H, running in time n^ℓ, say, $n^\ell = m^{\ell\delta^{-1}}$, which provides us with an approximate cut $(Q, V(H)\backslash Q)$ of H with value at least $(1-\epsilon)\mathrm{OPT}(H)$. Say that a cut of H is *pure* if it does not divide any of the sets V_i. By an argument of [FK01], [FK00], we can easily deduce in polynomial time from $(Q, \bar{Q})$ a pure cut of H, say $(P, \bar{P})$, with at least the same value. Let $(\Pi, \bar{\Pi})$ denote the cut induced by $(P, \bar{P})$ on $V(G)$ in the obvious way. We have then for a value of the cut $(\Pi, \bar{\Pi})$

$$\begin{aligned} \mathrm{Val}(\Pi, \bar{\Pi}) &= k^{-2}\mathrm{Val}(P, \bar{P}) \\ &\geq k^{-2}\mathrm{Val}(Q, \bar{Q}) \\ &\geq k^{-2}(1-\epsilon)\mathrm{OPT}(H) \\ &\geq (1-\epsilon)\mathrm{OPT}(G) \end{aligned}$$

since clearly $\mathrm{OPT}(H) \geq k^2\mathrm{OPT}(G)$. This concludes the proof of the theorem.

■

The construction used in the proof of Theorem 1 yields also the following.

Corollary 1. *MAX-CUT on r-regular graphs with $3 \leq r = \Theta(n^{1-\delta})$ is hard to approximate to within a constant for every $0 < \delta \leq 1$.*

Corollary 2. *MAX-CUT on $\Theta(n^{1-\delta})$-regular graphs for $0 < \delta < 1$ is polynomial time approximation preserving reducible (E-reducible) to MAX-CUT on $\Theta(n^{1-\delta'})$-regular graphs for any constant $\delta' < \delta$.*

Proof. Carry through the construction of Theorem 1 from $cn^{1-\delta}$-regular graphs to $c'n^{1-\delta'}$-regular graphs. Approximation properties of a construction are being preserved.

■

3 An Approximation Scheme for General MAX-CUT

We are going to prove now our main theorem. Recall that the density of graph $G = (V, E)$ is the ratio $|E|/\binom{|V|}{2}$. For a graph G on n vertices with density d we let

$$m = m(G, \epsilon) = \frac{10\log(1/\epsilon)}{\epsilon^2 d}. \tag{1}$$

Theorem 2. *There is an algorithm which when presented with a graph G on n vertices and density d, and an approximation accuracy $\epsilon > 0$, produces with probability at least $2/3$ a cut of G with value at least $(1-\epsilon)$ times the maximum value of a cut, in time*

$$2^{O(m)} \cdot n^{O(1)}$$

with m as defined above.

We formulate first several corollaries of the above theorem.

Corollary 3. *There exist approximation schemes for instances of MAX-CUT with* $\Omega(\frac{n^2}{\Delta})$ *edges working in* $2^{O^{\sim}(\frac{\Delta}{\varepsilon^2})} n^{O(1)}$ *time.*

By setting Δ to polylog n we have

Corollary 4. *There exist quasi-polynomial time approximation schemes (QPTASs) for* mildly sparse *instances of MAX-CUT with* $\Omega(\frac{n^2}{\text{polylog}\, n})$ *edges.*

We obtain also by setting Δ to $O(\log n)$ a result proven already by different method in [FK02] on existence of PTAS for *subdense instances* of MAX-CUT with $\Omega(\frac{n^2}{\log n})$ edges.

Corollary 5. *There exists a PTAS for subdense instances of MAX-CUT with* $\Omega(\frac{n^2}{\log n})$ *edges.*

In the subsequent sections we prove Theorem 2. Our approximation scheme for general MAX-CUT uses some ideas of smoothed analysis of quadratic programs of [FKK04]. There are several differences in our approach though, reflecting the fact that the degrees may be very widely distributed. In particular, we need a special *high discrepancy* Sampling Lemma to cope with those situations. Obviously, the simple pruning of small degrees does not work here anymore.

3.1 A New Sampling Lemma

We can suppose that $d = \omega(1/n)$ since in the case $d = O(1/n)$ we can use exhaustive search.

Lemma 1 (Sampling Lemma). *Let* (V_L, V_R) *be a bipartition of* V *with* $|V_R| \leq |V_L|$. (V_L, V_R) *is intended to be an optimum partition for MAXCUT on* G *so we assume that the inequality*

$$|\Gamma(x) \cap V_R| \geq (1/2)\deg(x) \tag{2}$$

holds for each x *in* V_L. *Let* S *be a random subset of* V *of size* m *as defined by 1 and let* $S_R = V_R \cap S$. *We have then*

$$\mathbf{E}\left(||\Gamma(x) \cap S_R| - \frac{\deg_R(x)|S_R|}{|V_R|}| \right) \leq \epsilon \sqrt{\frac{\deg(x)\log(1/\epsilon)}{nd}}, \tag{3}$$

and,

$$\mathbf{E}\left(\sum_{x \in S_R} ||\Gamma(x) \cap S_R| - \frac{\deg_R(x)|S_R|}{|V_R|}| \right) \leq m\epsilon(\log(1/\epsilon))^{1/2}. \tag{4}$$

Proof. Note that S_R is random in V_R, conditionally on $|S_R|$). The size of the intersection S_R satisfies $|S_R| \sim |V_R||S|/|V|$, so that by exhaustive sampling on the subsets of S, we will w.h.p. find one which we again call S_R which is contained in V_R and satisfies

$$0.9|S|/|V| \leq |S_R|/|V_R| \leq 1.1|S|/|V| \tag{5}$$

We assume that S_R is obtained by picking independently each point in V_R with probability $p = |S_R|/|V_R|$ (independent trials). $|\Gamma(x) \cap S_R|$ has the Binomial distribution with parameters $\deg_R(x) = |\Gamma(x) \cap V_R|$ and p. Assertion (3) follows from the fact that the average absolute deviation of a Binomial random variable from its expectation is bounded above by the standard deviation. To prove assertion (4), observe that the sum $\sum_{v \in V} \sqrt{\deg_v}$ is maximized for a fixed density by taking all the degrees equal.

■

3.2 The Algorithm and its Analysis

3.2.1 The Algorithm

For $u, v \in V$ we define $e_{uv} = 1$ if $\{u, v\} \in E(G)$. Otherwise, $e_{uv} = 0$. Consider the following *smoothed* linear program P (see also [FKK04]), depending on a set $S \in \binom{V}{m}$ and whose unknowns are the x_v, $v \in V$.

Maximize $\sum_{v \in V} x_v |\Gamma(x) \cap S|$

Subject to

$$\sum_{u \in V} (1 - x_u) e_{uv} \geq \frac{n}{m} |\Gamma(v) \cap S| - \delta_v \;\; \forall v \in V$$
$$\sum_{v \in V} \delta_v \leq (\epsilon/2)|E(G)|$$
$$\delta_v \geq 0 \;\; \forall v \in V$$
$$0 \leq x_v \leq 1 \;\; \forall v \in V$$

For a partition (V_L, V_R) of V, we intend $x_v = 0$ if $v \in V_L$. Otherwise, $x_v = 1$. The approximation scheme consists in picking a sufficiently large set of random subsets in $\binom{V}{m}$, for each of this subsets S solving P and then rounding randomly the solution by the usual device. The solution which gives the maximum cut value is then selected.

3.2.2 Analysis

We prove now that, with probability at least 2/3, the above algorithm gives a cut with value within $1 - \epsilon$ of the optimum.

Lemma 2. *Let (x_v) be a solution to* P. *Then, with probability at least 9/10 we have that*

$$\sum_{v,u \in V} x_v (1 - x_u) e_{vu} \geq maxcut(G) - (\epsilon/2)|E(G)|.$$

Proof. We just have to check that a vector x_v defining a maximum cut (V_L, V_R) satisfy w.h.p. the constraints of P. Let us define for each v

$$\delta_v^* = \left| \sum_{u \in V} (1 - x_u) e_{uv} - \frac{n}{m} |\Gamma(v) \cap S| \right|$$

By the Sampling Lemma the expectation of the sum $\sum_{v \in V} |\delta_v^*|$ is bounded above by $n\epsilon \log(1/\epsilon)^{1/2}$ while we are assuming that the ratio $|E(G)|/n$ tends to infinity. By Markov inequality, we get that the constraints of P are met by (x_v) with probability $1 - o(1)$.

■

We have now to examine the effect of the rounding. Let (x_v) be the solution of P. We define the rounded (randomized) solution x_v^* by

$$\Pr(x_v^* = 0) = 1 - x_v \qquad \Pr(x_v^* = 1) = x_v.$$

Therefore, with $Q = \sum_{u,v} x_u (1 - x_v) e_{uv}$, $Q^* = \sum_{u,v} x_u^* (1 - x_v^*) e_{uv}$, we have $E(Q^*) = Q$, and, since the x_v^* are independent,

$$\begin{aligned}
\mathrm{Var}(Q^*) &= \sum_{u,v,t} \mathrm{cov}(x_u^*(1 - x_v^*) e_{uv}, x_u^*(1 - x_t^*) e_{ut}) \\
&= \sum_{u,v,t} e_{uv} e_{ut} \mathrm{cov}(x_u^*(1 - x_v^*), x_u^*(1 - x_t^*)) \\
&\leq \sum_{u,v,t} e_{uv} e_{ut} \\
&\leq \sum_{u} \deg(u)^2 \\
&\in o(|E(G)|^2)
\end{aligned}$$

In order to justify the last deduction, observe that for given $|E(G)|$, the highest value of the sum of the squares of the degrees occurs when the number of vertices with maximum degree $n - 1$ is maximized. Noting that the number of vertices with this degree cannot exceed $|E(G)|d/n$, we have crudely

$$\begin{aligned}
\sum_u \deg(u)^2 &\leq 2nd|E(G)| \\
\frac{1}{|E(G)|^2} . \sum_u \deg(u)^2 &\leq \frac{2nd}{|E(G)|} \\
&\leq \frac{3}{n},
\end{aligned}$$

the last by the definition of the density. We have then by Chebyshev's inequality that the error due to the rounding is $o(|E(G)|)$.

From this and the fact that the value of the (not necessarily integer) solution (x_v) from which we start is by Lemma 2 at least the maximum of a cut minus $(\epsilon/2)|E(G)|$, and thus the correctness of the algorithm follows.

■

4 Extensions

The methods of this paper can be extended directly to the whole MAX-2CSP class of problems using appropriate smoothed linearized programs, and using recursive decompositions techniques of [AKK95] to the whole MAX-rCSP classes for all $r \geq 2$. This entails existence of approximation schemes for instances of MAX-rCSP problems with $\Omega(\frac{n^r}{\Delta})$ constraints working in time $2^{O^{\sim}(\frac{\Delta}{\varepsilon^2})} n^{O(1)}$. Again, *mildly sparse* instances of MAX-rCSP with $\Omega(\frac{n^r}{\operatorname{polylog} n})$ constraints have *quasi-polynomial approximation schemes* (QPTASs).

$\blacksquare$

5 Conclusion

This paper indicates explicit borderlines for existence of PTASs or QPTASs depending solely on the instance densities. It would be very interesting to shed some light on the approximation hardness of instances with the parameter Δ between n^δ and $\operatorname{polylog} n$. We suspect that the running times of approximation schemes of Theorem 2 can be generally improved in function of the underlying density.

This paper raises also new open questions. How about the algorithmic *nondense* analogs for using Szemerédi's Regularity Lemma along the lines of [FK96], and also how about the limits of nondense (say *mildly sparse*) graph sequences (cf. [LS04] for limits of dense graphs sequences)?

Acknowledgment: We thank Johan Håstad, Ravi Kannan and Laci Lovász for stimulating remarks and discussions.

References

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

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

[BK99] P. Berman and M. Karpinski, *On Some Tighter Inapproximability Results*, Proc. 25th ICALP (1999), LNCS 1644, Springer, pp. 200-209.

[BK03] P. Berman and M. Karpinski, *Improved Approximation Lower Bounds on Small Occurence Optimization*, ECCC Technical Reports, No 8 (2003).

[B05] A. Björklund, *Approximating Integer Quadratic Programs and MAXCUT in Subdense Graphs*, Proc. 13th ESA (2005), LNCS 3669, Springer, 2005, pp. 839-849.

[FKL02] U. Feige, M. Karpinski, and M. Langberg, *Improved Approximation of MAX-CUT on Graphs of Bounded Degree*, J. of Algorithms 43 (2002), pp. 201-219.

[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.

[FK01] W. Fernandez de la Vega and Claire Kenyon, *A Randomized Approximation Scheme for Metric MAX-CUT*, J. Comput. System Sciences 63 (2001), pp. 531-541.

[FK00] W. Fernandez de la Vega and M. Karpinski, *Polynomial Time Approximation of Dense Weighted Instances of MAX-CUT*, Random Structures and Algorithms 16 (2000), pp. 314-332.

[FK02] W. Fernandez de la Vega and M. Karpinski, *A Polynomial Time Approximation Scheme for Subdense MAX-CUT*, ECCC Technical Reports, No 44 (2002).

[FKK04] W. Fernandez de la Vega, M. Karpinski, and C. Kenyon, *Approximation schemes for Metric Bisection and Partitioning*, SODA 2004, pp. 506-515.

[FK96] A.M. Frieze and R. Kannan, *The Regularity Lemma and Approximation Schemes for Dense Problems*, Proc. 37th IEEE FOCS (1996), pp. 12-20.

[FK97] A.M. Frieze and R. Kannan, *Quick Approximation to Matrices and Applications*, Combinatorica 19 (1999), pp. 175-120.

[GW95] M. Goemans and D. Williamson, *Improved Approximation Algorithms for Maximum Cut and Satifiability Problems using Semidefinite Programming*, J. ACM 42 (1995), pp. 1115-1145.

[GGR96] O. Goldreich, S. Goldwasser, and D. Ron, *Property Testing and its Connection to Learning and Approximation*, Proc. 37th IEEE FOCS (1996), pp. 339-348; journal version in J. ACM 45 (1998), pp. 653-750.

[GT06] B. Green and T. Tao, The Primes Contain Arbitrary Long Arithmetic Progressions, arXiv math.NT/0404188 (Feb. 2006).

[K01] M. Karpinski, *Polynomial Time Approximation Schemes for Some Dense Instances of NP-Hard Optimization Problems*, Algorithmica 30 (2001), pp. 386-397.

[KMSV98] S. Khanna, R. Motwani, M. Sudan, and U. Vazirani, *On Syntactic versus Computational Views of Approximability*, SIAM J. Computing 28 (1998), pp. 164-191.

[KKMD04] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell, *Optimal Inapproximability Results for Max-Cut and Other 2-Variable CSPs?*, 45th IEEE FOCS (2004), pp. 146-154.

[LS04] L. Lovász and B. Szegedy, *Limits of Dense Graph Sequences*, MSR Tech. Report #MSR-TR-2004-79, 2004; http://research.microsoft.com/users/lovasz/limits.pdf.

ECCC ISSN 1433-8092
http://eccc.hpi-web.de/