Polynomial Time Approximability of Dense Weighted Instances of MAX-CUT

W. Fernandez de la Vega*   M. Karpinski†

Abstract

We give the first polynomial time approximability characterization of dense weighted instances of MAX-CUT, and some other dense weighted $\mathcal{NP}$-hard problems in terms of their empirical weight distributions.

1 Introduction

Significant results concerning Polynomial time approximation schemes (PTAS's) for "dense" instances of several $\mathcal{NP}$-hard problems such as MAX-CUT, MAX-k-SAT, BISECTION, DENSE-k-SUBGRAPH, and others have been obtained recently in Arora, Karger and Karpinski [AKK94], Vernandez de la Vega [FV96], Frieze and Kannan [AK97], Arora, Frieze and Kaplan [AFK95]. Still more recently, dense instances of approximation problems have been investigated from the point of view of the query complexity in Goldreich, Goldwasser and Ron [GGR96], Frieze and Kannan [AK97]. Recall that a PTAS for a given optimization problem is a family (A_ϵ) of algorithms

*LRI, CNRS, Université de Paris-Sud, 91405 Orsay. Research partially supported byt the ESPRIT BR Grants 7097 and EC-US 030, and by the PROCOPE grant. Email: lalo@lri.lri.fr.

†Dept. of Computer Science, University of Bonn, 53117 Bonn. Research partially supported by the International Computer Science Institute, Berkeley, California, by the DFG grant KA 673/4-1, by the ESPRIT BR Grants 7079 and EC-US 030, and by the Max–Planck Research Prize. Email: marek@cs.uni-bonn.de

indexed by a parameter $\epsilon \in (0, \infty)$ where each algorithm runs in polynomial time and, for each ϵ, the algorithm A_ϵ has approximation ratio $1 + \epsilon$. In most cases, the instances are graphs and a dense graph is defined as a graph with $\Theta(n^2)$ edges where n is the number of vertices. (In some cases, the algorithms apply only to graphs with minimum degree $\Theta(n)$.) Some of the problems considered in the papers listed above, such as MAX-CUT, are MAX-SNP-hard, and thus, if $\mathcal{P} \neq \mathcal{NP}$, have no PTAS's when the set of instances is not restricted.

Most algorithms in [AKK94] and those in [FV96] are based on the idea of *exhaustive sampling*: i) pick a small random set of vertices, ii) guess where they go in the optimum solution and, iii) use their placement to determine the placement of everything else. Phase iii) is essentially statistical: The decisions in phase iii) are made on the basis of a sample of the relevant information. (For instance, for MAX-CUT, the relevant information for the placement of a vertex x is the placement of everything else than x in the optimum solution.) The density requirement comes in to insure that these decisions are correct with high probability for a "small" sample size.

The natural instances of optimization problems (see the definitions given in [GJ79] involve weights while the results mentioned above deal mainly with the 0,1 case. We want to examine how these results can be extended to the weighted case. We want thus to define a concept of density for the weighted case, which ensures that our algorithms, possibly with minor modifications, work in the corresponding dense classes of instances and, hopefully since we want a genuine characterization of weighted polynomial time approximability, is such that the corresponding *non-dense* classes are not approximable under a standard intractability assumption. We have not succeeded completely in proving that our notion of density meets this later property, so that there would be some (small as we believe) gap in the general formulation. For the sake of simplicity, we concentrate here on MAX-CUT. In fact we start by considering MAX-BISECTION, which is MAX-CUT restricted to cuts with equal sides. (MAX-BISECTION is also called MAX-50/50-CUT or MAX-EQUI-CUT.) Our results extend easily to to other MAX-SNP problems such as Max-2SAT or Maximum Acyclic Subgraph. We note that weight problems have been briefly considered in [GGR96] and [AK97]. In both of these two papers, the authors evaluate the increase of the computation time of their algorithms when one allows weights belonging to some fixed interval $[0, a]$ instead of 0 1 weights. Weight problems are also considered in a recent paper of Crescenti, Silvestri and Trevisan [CST96]. His main result is that for weighted problems satisfying a certain "niceness" property (these problems include MAX-CUT) the approximability threshold for the unbounded version and that of a polynomially bounded version are equal.

The plan of this paper is as follows. We will first define dense classes of weighted instances via the corresponding classes of distribution functions (d.f.'s for short) of the weights which are, broadly speaking, classes in which the standard sampling problem: estimation of the mean of a distribution using the mean of a sample, can be solved efficiently. Next, we will give several natural examples of such density classes. In section 4, we shall give a general characterization of our dense classes of d.f.'s. In section 5, we prove that MAX-BISECTION has a PTAS in any dense class of weighted instances according to this definition. In section 6 we prove that approximating MAX-BISECTION on the instances to any fixed *non-dense* set of weight d.f.'s satisfying rather mild additional conditions, is NP-hard. In section 7, we extend these results to MAX-CUT. The last section contains a summary and open problems.

2 Definition of a Dense family

In as much as density requirements come in, any given instance is a set of non-negative real numbers (the weights) or rather a multi set. Let us associate to this instance the empirical distribution function of the weights:

$$F(x) = \frac{2}{n(n-1)} \sum_{x_i \leq x} m_i, \quad x \in \mathbb{R}^+$$

where m_i denotes the multiplicity of the weight x_i in the instance and n is the number of vertices.

We define our density classes in terms of families of weight distribution functions, i.e. instead of speaking of a multi set of real numbers, we speak of the corresponding d.f.. Clearly, such d.f.'s need have finite discrete support and rational individual probabilities. We call such d.f.'s *representable*. Conversely, the set $\mathcal{I}$ of instances corresponding to a representable d.f. with individual probabilities having smallest common denominator d, say, is given by

$$\mathcal{I} = \cup_{\{n:2d|n(n-1)\}} \mathcal{G}_n$$

where $\mathcal{G}_n$ is the set of weighted graphs on n vertices whose empirical weight distribution coincides with F. For convenience, in various occasions we state our theorems in terms of arbitrary d.f.'s. (not necessarily having finite or even discrete range). Moreover, we often assume in our proofs that our d.f.'s are continuous (i.e. we assume that the corresponding r.v.'s have densities). Whenever we make this assumption, it can be justified using the fact that any d.f. can be approximated arbitrarily closely by a continuous one.

We can assume that the mean of the weights in each instance is equal to 1. Indeed, when we divide all the weights by their mean, say m, in some instance, we also divide the values of the objective function by m so that the approximation ratios are unaffected. (We assume that the weights are not all 0.) Nevertheless, we shall not always impose this condition in the definition of our dense families. However, since our definitions are invariant under scaling we shall assume in the proofs that the d.f.'s have expectation 1. (Here and all along the paper, we speak with some abuse of language, of the expectation of a d.f. F meaning the expectation of a r.v. with d.f. F.)

We can now state our definition of a dense family of d.f.'s.

Definition 1 (Dense families of d.f.'s) *Let $\mathcal{F} = (F_j)_{j\in\mathcal{J}}$ be a family of integrable d.f.'s with supports contained in $\mathbb{R}^+$. Let μ_j denote the expectation of F_j. For each $j \in \mathcal{J}$ and each $k \in \mathbb{N}$, define*

$$M_{j,k} = \frac{1}{k\mu_j}\sum_{i=1}^{k} X_{j,i}$$

where the $X_{j,i}$ are independent r.v.'s each with d.f. F_j.

We say that the family $\mathcal{F}$ is dense if and only if, for each $j \in \mathcal{J}$, the sequence $(M_{j,k})_{k=1,2,..}$ converges in probability to 1, and moreover, this convergence is uniform for $j \in \mathcal{J}$.

In other words, $\mathcal{F}$ is dense iff there exists a function $n_\epsilon = n(\epsilon) :]0,1] \to \mathbb{N}$ such that the inequalities

$$\mathrm{P}[|M_{j,k} - 1| \le \epsilon] \ge (1-\epsilon),\ \ k \ge n_\epsilon \tag{1}$$

hold for each ϵ and simultaneously for all j, with an n_ϵ which depends only on ϵ (and not on j).

Definition 2 *A family $\mathcal{F}$ of d.f.'s which is not dense is called a non-dense family*

REMARK. As mentioned before, we do not restrict definition 1 to d.f.'s. representable by some graph (our characterization of dense families of d.f.'s does not require that the d.f.'s be representable). Representativity (or rather representativity by "small" graphs) will be considered when coming to inapproximability results.

An instant reflexion shows that definition 1 brings just what we want. If $\mathcal{F}$ is dense in the sense of definition 1, then we can estimate the mean of each $F \in \mathcal{F}$ with any desired relative accuracy by picking a sample whose size n_ϵ does not depend on F. In the next section we identify some natural dense families of d.f.'s.

3 Some Dense Families of d.f.'s

Recall the law of large numbers: If X has a finite mean EX, then the means of the partial sums of a sequence of independent r.v.'s each distributed as X converges in probability to EX. This implies immediately the next proposition.

Proposition 1 *Any singleton (and any finite set of integrable d.f.'s) with support in $\mathbb{R}^+$ is a dense family*

The following assertion can easily be checked.

Proposition 2 *The family of all integrable d.f.'s is not dense*

Our next dense family defines precisely, when restricted to 0 1 instances, the usual density classes. In the unweighted case (and after scaling), dense means "not too dispersed". We can thus define dense families based on some dispersion measure. The most common of these is the variance. and this leads to the following class of dense families.

Proposition 3 *For each $s \geq 0$ the family*

$$\mathcal{F}_s = \left\{ F_X : \frac{\mathrm{Var} X}{(EX)^2} \leq s \right\}$$

is dense.

For the proof, fix $s \geq 0$ and assume that F has mean 1 and variance $\leq s$. Using Tchebyshev's inequality one can immediately check that the mean of a sample of size $C\epsilon^{-3}$ from the distribution F, where C is a sufficiently large constant depending only on s, belongs to the interval $[1-\epsilon, 1+\epsilon]$ with probability at least $1-\epsilon$. □

We define next a family of dense sets in which the sample sizes required for a given approximation, though bigger than in the case of bounded variances, are bounded above by a polynomial in $1/\epsilon$.

Proposition 4 *For each pair (r, C) where $r \in]1, 2]$ and $C \in \mathbb{R}^+$, the family*

$$\mathcal{F} = \{F : \frac{1}{(EX)^r} \int_0^\infty x^r dF(x) \leq C\} \tag{2}$$

is dense.

Proof. The proof is very similar to the proof of the 'if' direction in Theorem 1 of the next section and is omitted. □

4 A Characterization of the Dense Families

The following theorem characterizes the dense families. Once again, this theorem, alike definition 1, is stated in terms of arbitrary (not necessarily representable) d.f.'s.

Theorem 1 *Let $\mathcal{F} = (F_j)_{j \in \mathcal{J}}$ be a family of non-negative integrable d.f.'s and assume all expectations equal to 1.*

The family $\mathcal{F}$ is dense in the sense of definition 1 if and only if one of the following conditions (i) and (ii) holds:

(i) For each j and each $x \in \mathbb{R}^+$, define $\tau_j(x) = x(1 - F_j(x))$. There is a function $\tau_o(x)$ tending to 0 as $x \to \infty$ and such that the inequalities

$$\tau_j(x) \leq \tau_o(x) \tag{3}$$

hold for each pair (j, x).

(ii) For each j and each $x \in \mathbb{R}^+$, define

$$s_j(x) = \int_x^\infty y dF_j(y). \tag{4}$$

There is a function $s_o(x)$ tending to 0 as $x \to \infty$ and such that the inequalities

$$s_j(x) \leq s_o(x) \tag{5}$$

hold for each pair (j, x).

Proof Notice that if we take $s_o = \tau_o$, then ii) is more stringent than i). Thus, it suffices to show that (ii) is necessary and (i) sufficient.

The fact that condition (i) implies that the family $\mathcal{F}$ is dense can be established easily by adapting the proof of the law of large numbers in order to get an effective bound on the sample size. Actually, we will adapt a proof of Feller (see [Fe]) that he uses to show the convergence of the means of sums of independent r.v.'s to a not necessary constant specified function. The speed of convergence is governed by the function τ. Let us write

$$S_n = X_1 + ... + X_n$$

where the X_i are independent with the common d.f. F with expectation μ. Let us define new r.v.'s X_i' by truncation at level n:

$$X_i' = X_i \text{ if } X_i \leq n, \qquad X_i' = 0 \text{ if } X_i > n.$$

Put

$$S_n' = X_1' + ... + X_n', \qquad m_n' = E(S_n') = nE(X_1').$$

Then,

$$P[|S_n - m_n'| > t] \leq P[|S_n' - m_n'| > t] + P[S_n \neq S_n'].$$

Putting $t = n\epsilon$ and applying Tchebyshev's inequality to the first term on the right, we get

$$P[|S_n - m_n'| > t] \leq \frac{1}{n^2\epsilon^2} E(X_1'^2) + nP[X_1 > n] \tag{6}$$

Put

$$\sigma(t) = \int_0^t x^2 dF(x).$$

Then, an integration by parts gives

$$\begin{aligned} \sigma(n) &= -n\tau(n) + 2\int_0^n x\tau(x)dx \\ &\leq 2\int_0^n x\tau(x)dx. \end{aligned}$$

(Recall that $\tau(x) = x(1 - F(x))$.) We have thus, for each n,

$$P\left[\left|\frac{S_n}{n} - EX_1'\right| \geq \epsilon\right] \leq \frac{2}{n^2\epsilon^2}\int_0^n x\tau(x)dx + \tau(n)$$

Since EX_1' tends to EX_1 uniformly for $F \in \mathcal{F}$ as $n \to \infty$, this implies

$$P\left[\left|\frac{S_n}{n} - EX_1\right| \geq 2\epsilon\right] \leq \frac{2\int_0^n x\tau(x)dx}{n^2\epsilon^2} + \tau(n) \tag{7}$$

for sufficiently large n. Clearly, the right side tends to 0 again uniformly whenever $\tau(t) \leq \tau_o(t)$ with a $\tau_o(t) \to 0$. This concludes the proof of sufficiency of condition (i).

Let us prove now the necessity of condition (ii). Assume that all expectations are equal to 1 and suppose that $\mathcal{F}$ does not satisfy to condition 5. Then, there is an $\eta > 0$ such that, for any $y \in \mathbb{R}^+$, there is an $F \in \mathcal{F}$ with

$$\int_y^\infty x dF(x) \geq \eta. \tag{8}$$

(The contrary would state

$\forall \eta > 0 \ \exists y(\eta) \in \mathbb{R}^+$ s.t. $\int_{y(\eta)}^\infty x dF(x) < \eta$
for every $F \in \mathcal{F}$

Then, putting $y_k = y(2^{-k})$, we could define an s_o for $\mathcal{F}$ by $s_o(x) = 2^-k$ for $y_k \leq x < y_{k+1}$, which is a contradiction.)

Fix an arbitrarily large y such that (8) holds with equality and fix a sample size $n = y\eta^{-1}$. We claim that we cannot then estimate EX with accuracy $\epsilon = \eta/10$. The truth of the claim implies the necessity of (ii) since n is arbitrarily large. Note first that 8 implies $1 - F(y) \leq y\eta^{-1}$. Thus, with probability

$$(F(y))^n \geq \left(1 - \frac{1}{n}\right)^n \geq e^{-1.1}, \tag{9}$$

(assuming that y and thus n are sufficiently large), all the points in the sample lie on the left-side of y. Now let Z be distributed as X conditioned by $X \leq y$. Then,

$$EZ = \frac{1-\eta}{1-F(y)} \leq 1 - \frac{9\eta}{10}, \tag{10}$$

again for sufficiently large y. Let M denote the mean of the sample and let M_c denote the mean of n independent r.v.'s each distributed as Z. Thus $EM_c = EZ$ and Markov inequality gives, with (10),

$$\begin{aligned} P[M_c \geq 1 - 0.1\eta] &\leq \frac{EZ}{1-0.1\eta} \leq \frac{1-0.9\eta}{1-0.1\eta} \\ &\leq \frac{\eta}{5}. \end{aligned}$$

Setting $p = \mathrm{P}[M_c \leq 1 - \eta/10]$, we have $p \geq \frac{4\eta}{5}$. Then, using (9), we obtain

$$\mathrm{P}[M \leq 1 - \eta/10] \geq \frac{4\eta e^{-1.1}}{5} > \frac{\eta}{10},$$

showing, as we claimed, that M cannot be approximated with accuracy 0.1η.
□

5 A PTAS for Dense Weighted Instances of MAX-BISECTION and MAX-CUT

The purpose of this section is to prove the following theorem.

Theorem *Let the family of d.f.'s $\mathcal{F}$ be dense (i.e. $\mathcal{F}$ satisfies to the conditions of Proposition 4). Then, MAX-BISECTION and MAX-CUT both have PTAS when restricted to the instances corresponding to $\mathcal{F}$.*

PROOF We give first the proof for MAX-BISECTION and then state briefly the required additions needed for MAX-CUT.

The sampling property which makes the algorithm in [FV96], see also [AKK94], work in the 0 1 case (altogether for MAX-BISECTION and MAX-CUT) is the following. Assume that $\mathcal{F}$ is a family of unweighted instances with density α, say, and let β satisfy $\alpha/4 \leq \beta \leq 1/2$. Assume moreover that $\{X, Y\}$ is a bipartition of the set of vertices V defining a cut with maximum value $M_\beta = M_\beta(G)$, say, within the set of cuts with $|X| = \beta n$. Then, for each $\epsilon > 0$, there is an integer $m = m(\epsilon)$ which depends only on α and ϵ but not on the instance size and such that, if W' (resp. W) is a random subset with cardinality m coming from X (resp. from Y), we have

$$\begin{aligned}
\mathrm{P}\Big(\frac{m^2(1-\epsilon)M_\beta}{\beta(1-\beta)n^2} &\leq \sum_{x\in W'} |\Gamma(x)\cap W| \\
&\leq \frac{m^2(1+\epsilon)M_\beta}{\beta(1-\beta)n^2}\Big) \\
&\geq 1-\epsilon.
\end{aligned}$$

(The intuitive content of this inequality is that, for most choices for W and W' the pair W, W' represents well the cut (X, Y) in the sense that the number of edges between W and W' is near from its expectation.) Now we are going to check that an analogous inequality holds in the weighted case for constant sizes $|W|$ and $|W'|$ when our density condition is satisfied.

Let us fix $\beta = 1/2$. Note first that we have to consider not the whole distribution F of the weights, but its restriction, say F', to the weights in the cut $\{X, Y\}$. (As a rule and whenever no confusion arises, we use F rather than F_j to denote a generic element of $\mathcal{F}$). However, it is easy to check that the family $\mathcal{F}' = \{F' : F \in \mathcal{F}\}$ is dense when $\mathcal{F}$ is. We can also assume that F' has expectation 1. Thus for simplicity, and with a slight abuse of notation, we shall write F for F'. We shall assume that the size of the instance is even and write $2n$ for this size. We let $B(X,Y)$ denote the complete bipartite graph with color classes X and Y, ($|X| = |Y| = n$).

For any subsets $U \subseteq X$ and $R \subseteq Y$, we put

$$w(U,R) = \sum_{x \in U} \sum_{y \in R} w(x,y)$$

In particular, $w(y,X)$ denotes the sum of the weights of the edges connecting $y \in Y$ to X. Let now W (resp. W') denote random subsets of X, (resp. Y), both of size m. Let $\mu = w(W,W')$. We can write, for a fixed W,

$$w(W,W') = \sum_{y \in W'} w(y,W)$$

Thus, for a fixed W, we can consider μ as the sum of a sample of size m from the empirical distribution, say $G(.) = G_W(.)$, of the $w(y,W)$, $y \in Y$. (We use sampling with replacement instead of exhaustive sampling but this approximation is harmless). This suggest the following two stages proof.

- First, we will prove that the distribution G has (almost) the right mean with high probability if m is large enough.

- Then, we will derive a density property for G (Proposition (7)), holding whenever the distribution of the weights comes from a dense family. However, this density property holds only in average, roughly speaking, and we will need a little more work to show (Proposition (8)) that this property implies an efficient sampling.

To begin this program, note that we have $w(W,Y) = \sum_{x \in W} w(x,Y)$, i.e. $w(W,Y)$ is the sum of a sample of size m from the set $w(x,Y), x \in X$. Let H denote the empirical d.f. of this set. We have, for each t,

$$1 - H(t) \leq 1 - F(u)$$

where u is given by the equation

$$\frac{1}{1 - F(u)} \int_u^\infty v F\{dv\} = t.$$

This gives, using the preceding inequality and (4),

$$t(1 - H(t)) \leq s_o(t).$$

Let $\mathcal{H}$ denote the family of distributions H when F ranges in $\mathcal{F}$. The preceding inequality shows that, if $\mathcal{F}$ is dense, then $\mathcal{H}$ is dense as well, as we wanted to show. This concludes the first stage of the proof.

Turning to the second stage, we would like to show that, with high probability, $G = G_W$, the d.f of the $w(y, W)$, $y \in Y$, satisfies uniformly for $F \in \mathcal{F}$ to (one of) the conditions of Theorem 1. Actually, we are able to prove only the following weakened form of condition (3).

Proposition 7. *Let F be fixed with mean 1 and let $G(.)$ denote the empirical distribution of the r.v.'s $(m)^{-1}w(y, W)$, $y \in Y$, for some assignment A of weights with empirical overall distribution F to the edges of $B(X, Y)$ (hence G depends on the choice of the assignment) and where W is a random subset of Y of size $|W| = m$. We have then,*

$$E(t(1 - G(t)) = o(1)$$

as $t \to \infty$, and where the $o(1)$ is uniform when A ranges over the distinct possible assignments and F ranges over any fixed dense set $\mathcal{F}$.

PROOF The proof which is technical and lengthy is omitted in this extended abstract. It can be found in Appendix 1. □.

We are now well prepared to prove the following lemma from which the approximability of MAX-BISECTION and MAX-CUT on dense weighted instances can easily be deduced (see [FV96]).

Lemma (Sampling lemma) *Let us assign to the edges of $B(X, Y) \sim K_{n,n}$, the complete equilibrated bipartite graph on $2n$ vertices, weights whose empirical distribution F has mean 1 and belongs to some dense set $\mathcal{F}$. Then, for every $\epsilon > 0$, there is an integer m_ϵ, which depends only on $\mathcal{F}$ and such that the following holds:*
For each $m \geq m_\epsilon$, if n is sufficiently large and W, (resp. W') is a random subset of X, (resp. Y) of size m, we have

$$P[|w(W, W') - m^2| \leq \epsilon m^2] \geq 1 - \epsilon$$

Proof Define for $y \in Y$, $w_y = m^{-1}w(y, W)$ and denote by λ the mean of the w_y. By the denseness of the set $\mathcal{H}$ we know that we have, for sufficiently large m,

$$P[1 - \epsilon \leq \lambda \leq 1 + \epsilon] \geq 1 - \epsilon. \tag{11}$$

We also know by proposition 7 that we have

$$E(\tau_G(t)) \leq \zeta(t) \tag{12}$$

where G is the empirical d.f. of the w_y and $\zeta(t)$ tends to 0 as $t \to \infty$. We want now to make use of the formula (7), with $n = m$, to prove the sampling lemma. First we have, by the additivity of the expectation,

$$E\left(\int_0^m x\zeta(x)dx\right) = \int_0^m E(x\zeta(x))dx,$$

This gives, using 12,

$$E\left(\int_0^m x\zeta(x)dx\right) \leq m \int_0^m \zeta(t)dt.$$

This implies, by Markov inequality,

$$P\left[2\epsilon^{-2}m^{-2}\int_0^m x\zeta(x)dx) + \zeta(m) \leq \phi(m)\right] \geq 1 - \epsilon,$$

with

$$\phi(m) = 2\epsilon^{-1}m^{-1}\int_0^m \zeta(t)dt + \epsilon^{-1}\zeta(m).$$

Now, using (7), we obtain

$$P\left[\left|\frac{w(W, W')}{m^2} - \lambda\right| \geq \epsilon\right] \leq \phi(m) + \epsilon$$

and

$$P\left[\left|\frac{w(W, W')}{m^2} - 1\right| \geq 3\epsilon\right] \leq 2\epsilon,$$

for sufficiently large m, using (11) and the fact that $\phi(m)$ tends to 0 with m. Since ϵ is arbitrary, the lemma follows. $\square$.

As mentioned before, the Sampling Lemma settles the case of MAX-BISECTION. This lemma gives in fact a PTAS for dense weighted MAX-CUT restricted to cuts with sides $\beta n, (1 - \beta)n$ where β is any fixed number in $]0, 1/2]$ and this fact was used in [FV96] to deduce a PTAS for MAX-CUT in the 0 1 case since then a strictly positive lower bound for the values of β which to be considered can be obtained, depending only on the (usual) density and the required accuracy. A similar bound can be given here for the instances corresponding to any given dense family $\mathcal{F}$. We omit the details in this extended abstract.

6 Hardness of MAX-BISECTION on a Non-Dense Set of Instances

We would like to prove that MAX-BISECTION is MAX-SNP-hard on any non-dense set of instances. We must however retreat a little from this aim and introduce the following conditions on $\mathcal{F}$, the first of which strengthens the non-density condition.

Condition 1: *There exist a sequence of indices $(j(i))_{i=1,2,...}$ and a sequence of reals $(t_i)_{i=1,2,...}$ such that the sequence (δ_i) where $\delta_i = 1 - F_{j(i)}(t_i)$, $i = 1, 2, ...,$ tends to 0 with i and, moreover, we have $t_i\delta_i \geq \eta$, $i = 1, 2, ...,$ with*

$$\delta_{i+1} \geq \delta_i^h, \quad i = 1, 2, ... \tag{13}$$

Here, η is a fixed positive real and $h \geq 1$.

Condition 2:*There are integers r and s such that each $F \in \mathcal{F}$ has a support with size at most r and individual probabilities with denominators bounded above by s.*

REMARK 1 Since we are of course assuming non-denseness of $\mathcal{F}$, condition 1 is really not very stringent

REMARK 1 Condition 2 is purely technical. It insures that the $F_i's$ be representable by "not too big" graphs. It implies in fact that each $F_\in\mathcal{F}$ is representable by any graph whose number of edges is a multiple of sr. Condition 2 can be considerably relaxed.

REMARK 3 By choosing large enough r and s, we can approximate with any required accuracy an arbitrary d.f. by a d.f. satisfying to condition 2.

We shall prove the following theorem.

Theorem 3 *Let the family of d.f.'s $\mathcal{F}$ satisfy to conditions 1 and 2. Then, approximating MAX-BISECTION on the instances corresponding to $\mathcal{F}$ is MAX-SNP-hard*

PROOF The proof has two parts. First we show that approximating MAX-BISECTION on the instances with given size n and given empirical distribution of weights F, involves approximating MAX-BISECTION on the (unweighted) graphs of order n and having a certain density d, say. Now, when the distributions of the weights comes from a non-dense family $\mathcal{F}$, the density d achieves arbitrarily small values when F ranges over $\mathcal{F}$. Thus, bearing in mind that MAX-BISECTION is MAX-SNP hard on 0 1 instances (see Papadimitriou and Yannakakis [PY91]), we should be done at this point apart from the apparently innocuous fact that, for each involved d, we are not required to solve MAX-BISECTION on all possible instances of density d, but only for a set of sizes which may be quite thin. We have been able to circumvent this difficulty only by adding the (once again rather mild) condition (15) (see below) which is the analogue for the unweighted case of condition (13), We begin by taking care of this point.

6.1 Hardness of MAX-BISECTION on a Non-Dense Set of Unweighted Instances

We need two lemmas.

Lemma 3 *Assume that the sequence $(n_k)_{k=1,\ldots}$ satisfies for any sufficiently large k the inequality*

$$n_{k+1} \leq n_k^h \tag{14}$$

where h is a fixed number greater than 1. Then, MAX-BISECTION remains MAX-SNP-hard when restricted to graphs whose vertex set sizes belong to (n_k).

PROOF Assume for a contradiction that for any $\epsilon \geq 0$, there exists an integer k such that 0 1 MAX-BISECTION is $(1-\epsilon)$-approximable in time n^k for vertex set sizes belonging to (n_j) by some algorithm A. Set for each n, $m = m(n) = \min\{n_j : n_j \geq n\}$. Clearly we have $m \leq n^h$. Thus, the algorithm A can be used with trivial modifications to solve MAX-BISECTION for any instance size n in time n^{kh}. This contradicts the MAX-SNP-hardness of unweighted MAX-BISECTION. □

REMARK Assertions similar to Lemma 3 apply in fact to any problem with the property that any instance of size n can also be considered in a natural way as an instance of size m whenever $m \geq n$.

Lemma 4 *Let $(n_i, d_i)_{i=1,2,\ldots}$ be an infinite sequence of pairs of integers where $3 \leq d_i \leq \Delta$ for each i with a fixed Δ. Let $\mathcal{G}(n,d)$ denote the set of graphs with n vertices and average degree d. Then, approximating MAX-BISECTION on the set*

$$\cup_{i\in\mathbb{N}}\ \mathcal{G}(n_i, 3)$$

reduces to approximating MAX-BISECTION on the set

$$\cup_{i\in\mathbb{N}}\ \mathcal{G}(n_i, d_i).$$

PROOF The easy proof is omitted. □

Lemmas 3 and 4 give easily the following particular case of Theorem 3.

Theorem 4 *Let $\mathcal{F} = (F_i)_{i=1,2,\ldots}$ be a non-dense set of 0 1 instances, i.e. there is an infinite vanishing sequence of numbers (d_i) such that, for each i, F_i gives probability d_i to the point d_i^{-1} and $1 - d_i$ to the origin. (d_i is the density common to the instances corresponding to F_i.) For each i, let $\frac{p_i}{q_i} = d_i$ be the shortest fraction representing d_i. Assume that the p_i's are bounded above by a fixed integer p_o and that the d_i's verify the inequality*

$$d_{i+1} \geq d_i^h \tag{15}$$

for some $h > 1$. MAX-BISECTION is MAX-SNP-hard on the instances corresponding to $\mathcal{F}$.

PROOF Because of lemma 3 and the fact that MAX-BISECTION is MAX-SNP-hard on the graphs with average degree 3 we need only exhibit an infinite sequence of sizes (n_k) satisfying to condition 14 and such that, for each k, the graphs on n_k vertices and with density d_k in our family of instances belonging to some fixed interval $[3, \Delta]$ (actually we shall take $\Delta = 3p_o$). We have

$$\begin{aligned} \frac{n_k\Delta_k}{2} &= \frac{n_k(n_k-1)d_k}{2} \\ \Delta_k &= (n_k-1)d_k = (n_k-1).\frac{p_k}{q_k}. \end{aligned}$$

Thus, choosing $n_k = 3q_k + 1$, we get $\Delta_k = 3p_k$ implying $3 \leq \Delta_k \leq 3p_o$. The proof is concluded by observing that condition 15 implies condition 14. □

The next lemma uses the following well known Martingale inequality (see a standard probability book or Alon and Spencer ([AS90]).

A Martingale inequality *If $Y_1, \ldots, Y_n$ are independent, satisfy $EY_i = 0$ and $|Y_i| \leq M$, a.s., then, with $S_n = \sum_1^n Y_i$, we have for any $a \geq 0$,*

$$P[|S_n| \geq a] \leq \exp\left(-\frac{a^2}{2nM^2}\right)$$

The following lemma asserts broadly speaking that putting random weights with mean 1 on the edges of a (not too sparse) graph G does not change significantly the maximum value of a cut of G.

Lemma (Averaging lemma) *Let $(G_n)_{n=1,\ldots}$ be a sequence of graphs where G_n has n vertices and $m = m(n)$ edges and $n = o(m)$. Assume that for each n the edges of G_n are given random non-negative weights picked from a fixed distribution F with mean 1. Let G'_n denote this weighted graph. We denote by $\delta_G(S)$ the value of the cut of G defined by the set of vertices S*
The quantity

$$\frac{1}{m} \max_S |\delta_{G'_n}(S) - \delta_{G_n}(S)|$$

where S ranges over all subsets of $V(G)$, tends to 0 in probability when $n \to \infty$.

PROOF The proof uses standard counting and is omitted in this extended abstract. □

6.2 End of the Proof

Because of Theorem 4, in order to conclude the proof of theorem 3 it suffices to prove the following.

Theorem 5 *Assume that $\mathcal{F}$ satisfies to conditions 1 and 2 of theorem 4 with parameters eta, h, r and s, say. Approximating MAX-BISECTION on $\mathcal{F}$ reduces then to approximating MAX-BISECTION on a non-dense set of 0 1 instances.*

PROOF Let the sequences (d_i) and (t_i) be defined as in condition 1 of theorem 1. Set $F \equiv F_{j(i)}$, $t = t_i$ and define

$$\alpha = \alpha_i = \frac{1}{1 - F(t)} \int_t^{\infty} s dF(s)$$

and

$$\beta = \beta_i = \frac{1}{F(t)} \int_0^t s dF(s).$$

For $n = \lambda rs$, $\lambda \in \mathbb{N}$, set $m = d_i\binom{n}{2}$ and use k to index the $\binom{\binom{n}{2}}{m}$ distinct subgraphs G_k of K_n having m edges.

We define for each k a partial instance J_k by giving to the edges of G_k random weights empirically distributed according to the d.f.

$$G(s) = \frac{F(s) - F(t)}{1 - F(t)}, \ s \geq t.$$

We define also a partial instance L_k by putting on the edges of $K_n \backslash G_k$ random weights on $[0, t]$ empirically distributed according to the d.f. $H(s) = F(s)/F(t),\ s \leq t$. We denote by I_k the instance obtained by sticking together J_k and L_k. Clearly, by the choice of $G(.)$ and $H(.)$, the empirical distribution of the weights in I_k coincides with F.

Let Opt denote the value of the optimal bisection of I_k and assume that we can find in polynomial time a bisection δ_{S_o} with value $\geq (1-\epsilon)Opt$ in the instance I_k. Let I'_k denote the instance obtained by replacing the weights in J_k (resp. in L_k) by their mean α, (resp. β).

By applying the averaging lemma separately to J_k and L_k, we see that the maximum value of a bisection in I'_k does not differ from $\delta_{S_o}(I'_k)$ by more than a $(1+\epsilon)$ factor.

Let now I_k" denote the instance obtained from I'_k by subtracting β from each weight. (Thus, I_k" has weights all equal to $\alpha - \beta$ on the edges of G_k and zero weights everywhere else.) For any bisection δ_S we have

$$\delta_S(I'_k) = \delta_S(I_k") + \frac{\beta n^2}{4}\ .$$

Note that we have $\alpha \geq \eta$ and since we have

$$\alpha(1 - F(t)) + \beta F(t) = 1,$$

while $F(t) = F_i(t_i)$ tends to 1 with i, we deduce that β has an upper bound strictly smaller than 1. Since the maximum value of a 50/50 cut of G is at least $n^2/4$, this implies for $S = S_o$ that the the ratio

$$\frac{\delta_{S_o}(I_k")}{\delta_{S_o}(I'_k)}$$

is bounded below by a strictly positive constant so that then δ_{S_o} is also an approximate solution for MAX-BISECTION on I_k". Thus, approximating MAX-BISECTION on $\mathcal{F}$ enables us to approximate 0 1 MAX-BISECTION on the instances corresponding to the non-dense sequence of densities (d_i) as we wanted to prove. □

7 MAX-CUT is MAX-SNP-hard on a non-dense set of weighted instances

To be filled

8 Summary and Conclusions

With the aim of separating as sharply as possible the approximable from the inapproximable families of weighted instances of MAX-CUT, we have introduced a notion of dense families of instances or, more precisely, a notion of dense families of weight distributions. We have shown that the corresponding families of instances have the (intended) polynomial time approximability property for MAX-CUT.

In the other direction, we have shown inapproximability only at the cost of a slight strengthening of our density condition and we believe that this is not necessary. This is our first question.

A second question is: Does our density definition capture the approximability of all MAX-SNP-hard problems instances in the weighted case? We know by [AK97] that all these problems are approximable in the dense *unweighted* case.

Acknowledgement. We thank Sanjeev Arora, Haine Kaplan and Uri Zwick for a number of helpful discussions.

References

[AS90] N. Alon and J.H. Spencer, *The Probabilistic Method*, Wiley 1992.

[AKK94] 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.*

[AFK95] S. Arora, A.M. Frieze and H. Kaplan, A New Rounding Procedure for the Assignment Problem with Applications to Dense Graph Arrangements, *Proc. 27th STOC (1995) pp.284-293*

[CST96] P. Crescensi, R. Silvestri and L. Trevisan, To weight or not to Weight: Where is the Question? *Proc. 26th STOCS (1996)*

[Fe] W. Feller, *An Introduction to Probability Theory and its Applications,* Vol 2. Wiley 1963

[FV96] W. Fernandez de la Vega, MAX-CUT has a Randomized Approximation Scheme in Dense Graphs, *Random Structures and Algorithms, Vol. 8. No.3 (1996)*

[AK96] A.M. Frieze and R. Kannan, The Regularity Lemma and Approximation Schemes for Dense Problems, *Proc. 37th FOCS (1996) 12-20.*

[AK97] A.M. Frieze and R. Kannan, Quick Approximation to Matrices and applications, *Manuscript (1997)*

[GJ79] M.R. Garey and D.S. Johnson, *Computers and Intractability, A Guide to the theory of NP-Completeness.* Freeman, San Francisco CA,1979.

[GGR96] O. Goldreich, S. Goldwasser and D. Ron, Property Testing and its Connection to Learning and Approximation, *Proc. 37th FOCS (1996)*

[PY91] C.H. Papadimitriou and M. Yannakakis, Optimization, Approximation, and Complexity Classes, *Journal of Computer and System Sciences, Vol 43, 425-440 (1991)*

Appendix 1: Proof of Proposition 7

Proposition 7. *Let F be fixed with mean 1 and let $G(.)$ denote the empirical distribution of the r.v.'s $(m)^{-1}w(y, W)$, $y \in Y$, for some assignment A of weights with empirical overall distribution F to the edges of $B(X, Y)$ (hence G depends on the choice of the assignment) and where W is a random subset of Y of size $|W| = m$. We have then,*

$$E(t(1 - G(t)) = o(1)$$

as $t \to \infty$, and where the $o(1)$ is uniform when A ranges over the distinct possible assignments and F ranges over any fixed dense set $\mathcal{F}$.

PROOF Let us fix some dense set $\mathcal{F}$ and an assignment of weights to $B(X, Y)$ with empirical distribution $F \in \mathcal{F}$. Recall that W is a random subset of X with $|W| = m$. For $y \in Y$, let $S_y = \sum_{x \in X} w(y, x)$. Define for each $t > 0$ the set

$$Y_1 = \{y \in Y : S_y \leq tn/2\},$$

and, for each $y \in Y_1$,

$$Z_y = \sum_{x \in W} w(y, x)$$

Let $T_y = \{x \in X : w(y, x) \geq mt\}$, $n_y = |T_y|$. We have of course

$$\begin{aligned} P[Z_y \geq tm] &= P[W \cap T_y \neq \emptyset] + \\ &+ P[Z_y \geq tm | W \cap T_y = \emptyset]. \end{aligned}$$

Let $\bar{Z}_y$ denote a r.v. distributed as Z_y conditioned by the event $W \cap T_y = \emptyset$. Since the r.v.'s $1_{x \in W}$ are negatively correlated for distinct x's, we have

$$\begin{aligned} \mathrm{Var}\bar{Z}_y &\leq \sum_{x \in X - T_y} \mathrm{Var}(w(y, x) 1_{x \in W}) \\ &\leq \frac{m}{n} \sum_{x \in X - T_y} w(y, x)^2. \end{aligned}$$

(Recall that we are assuming a fixed assignment of weights). This gives, using Tchebicheff's inequality,

$$\begin{aligned} P[\bar{Z}_y \geq tm] &\leq \frac{\mathrm{Var}\bar{Z}_y}{(tm - E\bar{Z}_y)^2} \\ &\leq \frac{4 \sum_{x \in X - T_y} w(y, x)^2}{mnt^2} \end{aligned}$$

using the inequality $E\bar{Z}_y \leq tm/2$. Thus, writing

$$N_1(t) = \sum_{y \in Y} 1_{w(y,W) \geq tm}$$

we obtain

$$E\left(\frac{N_1(t)}{n}\right) \leq T_1 + T_1'$$

with

$$\begin{aligned} T_1 &= \frac{\#\{(y,x) : w(y,x) \geq mt\}}{n} = 1 - F(tm) \\ &\leq \frac{\tau(tm)}{tm} \leq \frac{\tau_o(tm)}{tm}, \end{aligned}$$

and,

$$T_1' = \frac{4}{mnt^2} \times \sum_{\{(y,x): w(y,x) \leq mt\}} w^2(y,x).$$

T_1' satisfies

$$T_1' \leq \frac{4}{mt^2} \int_0^{tm} y^2 dF(y)$$

We have for each ϵ,

$$\begin{aligned} \int_0^{tm} y^2 dF(y) &\leq \epsilon tm + \int_{\epsilon tm}^{tm} y^2 dF(y) \\ &\leq tm(\epsilon + s(\epsilon tm)), \\ &\leq tm(\epsilon + s_o(\epsilon tm)) \end{aligned}$$

since $s(.)$ is bounded above by the function s_o corresponding to $\mathcal{F}$, (see (4)). Choosing (somewhat arbitrarily) $\epsilon = (tm)^{-1/2}$, we obtain

$$T_1' \leq \frac{4}{t}\left(\frac{1}{\sqrt{mt}} + s_o(\sqrt{mt})\right).$$

It remains to deal with the vertices y which have $S_y \geq nt/2$. Their number $N_2(t)$ satisfies

$$T_2 = \frac{N_2(t)}{n} \leq 1 - F(u)$$

where u is defined by

$$\frac{1}{1 - F(u)} \int_u^{\infty} y dF(y) = \frac{t}{2}. \tag{16}$$

(T_2 is maximized by making $S_y = nt/2$ for a maximum number of $y's$.)
(16) gives

$$\begin{aligned} 1 - F(u) &= \frac{2}{t}\int_u^\infty y dF(y) \\ &\leq \frac{2s_o(u)}{t}, \end{aligned}$$

using (4). Since $s_o(.)$ is upper-bounded, it follows that u tends to infinity as $t \to \infty$, uniformly for $F \in \mathcal{F}$. But then $s_o(u)$ tends to 0 and we have thus $1 - F(u) = o(t^{-1})$ also uniformly, say

$$1 - F(u) \leq \frac{h(t)}{t}$$

for some function $h(t)$ tending to 0 when $t \to \infty$.

Collecting our estimates, we get, for any $F \in \mathcal{F}$,

$$\begin{aligned} E(1 - G(t)) &= E\left(\frac{N_1(t)}{n}\right) + E\left(\frac{N_2(t)}{n}\right) \\ E(1 - G(t)) &\leq T_1 + T_1' + \frac{h(t)}{t} \\ E(t(1 - G(t)) &\leq \frac{\tau_o(mt)}{m} + \frac{4}{\sqrt{mt}} + 4s_o(\sqrt{mt}) + h(t). \end{aligned}$$

Since $\tau_o(t)$, $s_o(t)$ and $h(t)$ all tend to 0 as $t \to \infty$, it is clear that $E(t(1 - G(t)) \to 0$ with t uniformly for $F \in \mathcal{F}$, as was to be proved. □

Appendix 2: Proof of the averaging lemma

Lemma (Averaging lemma) *Let $(G_n)_{n=1,...}$ be a sequence of graphs where G_n has n vertices and $m = m(n)$ edges and $n = o(m)$. Assume that for each n the edges of G_n are given random non-negative weights picked from a fixed distribution F with mean 1. Let G'_n denote this weighted graph. We denote by $\delta_G(S)$ the value of the cut of G defined by the set of vertices S*
The quantity

$$\frac{1}{m}\max_S |\delta_{G'_n}(S) - \delta_{G_n}(S)|$$

where S ranges over all subsets of $V(G)$, tends to 0 in probability when $n \to \infty$.

PROOF The proof uses standard counting and is omitted in this extended abstract

We must prove that, for every $\epsilon > 0$, we have for sufficiently large n,

$$P\left[\max_S |\delta_{G'_n}(S) - \delta_{G_n}(S)| \leq \epsilon m\right] \geq 1 - \epsilon \ .$$

We shall first get rid of the extreme values of F. Define $\theta = \theta(\epsilon)$ by

$$\int_\theta^\infty s dF(s) = \frac{\epsilon^2}{2},$$

so that, by Markov inequality, with probability at least $1-\epsilon$, the edges with weights $\geq \theta$ carry a total weight $\leq m\epsilon/2$. We also define

$$\eta = \int_\theta^\infty dF(s),$$

and note that, again with probability at least $1-\epsilon$, the maximum value of a cut of size $\leq \epsilon\eta m$ does not exceed $m\epsilon/2$ (since the expectation of the maximum total weight of $\epsilon\eta m$ edges is equal to $\epsilon^2/3$). Thus it remains to consider the cuts with size $\geq \eta m$. Let us fix a cut $\delta(S)$ with $\delta_G(S) = \lambda m$, $\eta \leq \lambda \leq 1/2$. $\delta_{G'}(S)$ is the sum of λm independent r.v.'s with the common distribution F. Call F' the distribution obtained from F by cutting the values $\geq \theta$: $F'(s) = F(s)/(1-F(\theta))$, $s \leq \theta$. By what has just been proved, we have with high probability, simultaneously for all S,

$$\delta_{G'}(S) \leq \delta'(S) + \frac{m\epsilon}{2} \tag{17}$$

where $\delta'(S)$ is the sum of λm independent r.v.'s with the common distribution F'. Using the fact that these r.v.'s are bounded above by θ and a familiar martingale inequality, we obtain

$$\begin{aligned} P[|\delta'(S) - \lambda m| \geq \epsilon m/2] &\leq \exp\left(-\frac{\epsilon^2 m^2}{8\theta^2 \lambda m}\right) \\ &\leq \exp\left(-\frac{\epsilon^2 m}{8\theta^2 \eta}\right). \end{aligned}$$

Since the total number of cuts is bounded above by 2^n, we obtain

$$\begin{aligned} P[|\delta'(S) - \lambda m| &\leq \epsilon m/2, \text{ all } S] \\ &\geq 1 - 2^n \exp\left(-\frac{\epsilon^2 m}{8\theta^2 \eta}\right) \end{aligned}$$

whose r.h.s. tends to 1 for any ϵ by our assumption on $m = m(n)$. The lemma follows from the last inequality and (17).

□