Journal of the Operations Research
Society of Japan
2003, Vol. 46, No. 2, 178-188

© 2003 The Operations Research Society of Japan

NEW APPROXIMATION ALGORITHMS FOR MAX 2SAT AND MAX DICUT

Shiro Matuura Tomomi Matsui
University of Tokyo

(Received Received: June 7, 2002, Revised: February 21, 2003)

Abstract We propose a 0.935-approximation algorithm for MAX 2SAT and a 0.863-approximation algorithm for MAX DICUT. The approximation ratios improve upon the recent results of Zwick, which are equal to 0.93109 and 0.8596434254 respectively. Also proposed are derandomized versions of the same approximation ratios. We note that these approximation ratios are obtained by numerical computation rather than theoretical proof.

The algorithms are based on the SDP relaxation proposed by Goemans and Williamson but do not use the 'rotation' technique proposed by Feige and Goemans. The improvements in the approximation ratios are obtained by the technique of 'hyperplane separation with skewed distribution function on the sphere.'

Keywords: Algorithm, combinatorial optimization, nonlinear programming

1. Introduction

In this paper we propose approximation algorithms for the optimization problems called MAX 2SAT and MAX DICUT. The MAX 2SAT problem is described as follows. Given n boolean variables $x_1, x_2, \ldots, x_n$, m clauses $C_1, C_2, \ldots, C_m$ each consisting of two literals (either a boolean variable x_i or its negation $\neg x_i$), and nonnegative weights $w_1, w_2, \ldots, w_m$ associated with clauses, the MAX 2SAT is the problem of finding an assignment of true or false to x_i's that maximizes the total weight of satisfied clauses. The MAX 2SAT is formulated as follows:

$$\text{(ST)} \quad \text{maximize} \sum_{s:C_s \text{ is satisfied}} w_s \quad \text{subject to} \quad x_i \in \{\text{True, False}\}.$$

The MAX DICUT problem is described as follows. We are given a complete directed graph $D = (V, A)$ with vertex set $V = \{1, 2, \ldots, n\}$ and arc set $A = \{(i, j) \in V \times V \mid i \neq j\}$, and a nonnegative arc weight w_{ij} for each arc $(i, j) \in A$. An arc subset $A' \subseteq A$ is called a *dicut* if it can be represented as $A' = \{(i, j) \in A \mid i \in U \text{ and } j \in V \setminus U\}$ for some vertex subset $U \subseteq V$. The weight of a dicut A' is the sum of the weights of arcs contained in A'. The MAX DICUT is the problem of finding a dicut that maximizes its weight, and is formulated as follows:

$$\text{(DI)} \quad \text{maximize} \sum_{i \in U,\, j \in V \setminus U} w_{ij} \quad \text{subject to } U \subseteq V.$$

Both MAX 2SAT and MAX DICUT are NP-hard [4], and this fact motivates researches on algorithms for finding an approximate solution. As is well-known, Goemans and Williamson [5] proposed a novel algorithmic scheme of randomized polynomial-time algorithms that applies to these problems as well as MAX CUT problem. Their algorithmic scheme is based on a combination of Semi-Definite Programming (SDP) relaxation and random hyperplane separation. Their approximation ratios are 0.87856 for MAX 2SAT, and 0.79607 for MAX DICUT. Feige and Goemans [2] proposed approximation algorithms with the approximation ratios 0.93109 for MAX 2SAT and 0.859387 for MAX DICUT. Their algorithms are based on two ideas. First, some constraints introduced by Feige and Lovász in [3] are added to the SDP relaxation. Second, the solution obtained by SDP relaxation is modified by the 'rotation' technique. The approximation ratios of their algorithms are obtained through numerical computation. With a refinement of the rotation technique, Zwick [10] improved the approximation ratios to 0.9310900680 for MAX 2SAT and 0.8596434254 for MAX DICUT. It is also shown that those approximation ratios are nearly the best attainable by any rotation technique.

In this paper, we propose randomized/derandomized approximation algorithms without rotation technique whose approximation ratios are 0.935 for MAX 2SAT and 0.863 for MAX DICUT, where the latter was reported in a conference paper [8]. Our algorithms solve the SDP relaxation problems proposed by Goemans and Williamson with the constraints used in Feige and Goemans' algorithms. The improved approximation ratios are obtained by using a skewed distribution function on the sphere. We note that these approximation ratios are obtained by numerical computation rather than theoretical proof. The possibility of using hyperplane separation technique with skewed distribution is suggested by Feige and Goemans [2], and the contribution of the present paper lies in the following:

- pointing out the difficulty that the probability distribution resulting from the suggested method of skewed distribution depends on the dimension n,
- identifying a class of distribution functions for which the dependence on the dimension n can be given in explicit formulae,
- selecting a good distribution function within that class to achieve good approximation ratio,
- constructing derandomized algorithms by applying the derandomization technique of Mahajan and Ramesh [7].

The derandomized algorithms thus constructed coincide with the derandomized algorithms obtained by Mahajan and Ramesh on the basis of the randomized algorithms of Goemans and Williamson. As an important implication of this fact, we obtain improved bounds on the approximation ratios of Mahajan and Ramesh's derandomized algorithms, i.e., 0.935 for MAX 2SAT and 0.863 for MAX DICUT.

In Section 2, we briefly review the SDP relaxation and hyperplane separation. In Section 3, we describe the outline of our algorithm. In Section 4, we discuss some relations between the skewed distribution functions on the 2-dimensional sphere and on the n-dimensional sphere. In Section 5, we describe a numerical method used for finding a good distribution function defined on the 2-dimensional sphere. In Section 6, we show how to derandomize our algorithm.

2. Semi-Definite Programming Relaxation

Here we describe SDP relaxations of MAX 2SAT and MAX DICUT, and review the hyperplane separation technique.

First, we formulate the MAX 2SAT problem as an integer programming problem. Let v_i be a $\{-1,1\}$-variable associated with x_i and v_{i+n} be a $\{-1,1\}$-variable associated with $\neg x_i$. Let C be the set of index pairs of clauses, i.e., $C = \{(i,j) | \exists s : C_s = (x_i \vee x_j)\} \cup \{(i,j) | \exists s :$

$C_s = (x_i \vee \neg x_{j-n})\} \cup \{(i,j) | \exists s : C_s = (\neg x_{i-n} \vee x_j)\} \cup \{(i,j) | \exists s : C_s = (\neg x_{i-n} \vee \neg x_{j-n})\}$, and w_{ij} be the weight associated with the clause corresponding to $(i,j) \in C$. The following problem is equivalent to the original problem (ST).

$$\begin{array}{lll} \text{(ST')} & \text{maximize} & (1/4) \sum_{(i,j) \in C} w_{ij}(3 + v_0 v_i + v_0 v_j - v_i v_j), \\ & \text{subject to} & v_0 = 1, \ \ v_i + v_{i+n} = 0 \ \ (\forall i \in \{1, \ldots, n\}), \\ & & v_i \in \{-1, 1\} \quad (\forall i \in \{1, \ldots, n, n+1, \ldots, 2n\}). \end{array}$$

In the paper [5], Goemans and Williamson relaxed the above problem by replacing each variable $v_i \in \{-1, 1\}$ with a vector on the n-dimensional unit sphere $\boldsymbol{v}_i \in \boldsymbol{S}_n$, where $\boldsymbol{S}_n \stackrel{\text{def.}}{=} \{\boldsymbol{v} \in \mathbf{R}^{n+1} \mid ||\boldsymbol{v}|| = 1\}$. The idea of this relaxation can be traced back to Lovász [6]. With some additional valid constraints used in [2, 3], we obtain the following relaxation problem:

$$\begin{array}{lll} (\overline{\text{ST}}) & \text{maximize} & (1/4) \sum_{(i,j) \in C} w_{ij}(3 + \boldsymbol{v}_0 \cdot \boldsymbol{v}_i + \boldsymbol{v}_0 \cdot \boldsymbol{v}_j - \boldsymbol{v}_i \cdot \boldsymbol{v}_j), \\ & \text{subject to} & \boldsymbol{v}_0 = (1, 0, \ldots, 0)^\top, \quad \boldsymbol{v}_i + \boldsymbol{v}_{i+n} = \boldsymbol{0} \ \ (\forall i \in \{1, \ldots, n\}), \\ & & \boldsymbol{v}_i \in \boldsymbol{S}_n \quad (\forall i \in \{1, \ldots, n, n+1, \ldots, 2n\}), \\ & & \ \ \boldsymbol{v}_0 \cdot \boldsymbol{v}_i + \boldsymbol{v}_0 \cdot \boldsymbol{v}_j + \boldsymbol{v}_i \cdot \boldsymbol{v}_j \geq -1 \ \ (\forall (i,j)), \\ & & -\boldsymbol{v}_0 \cdot \boldsymbol{v}_i - \boldsymbol{v}_0 \cdot \boldsymbol{v}_j + \boldsymbol{v}_i \cdot \boldsymbol{v}_j \geq -1 \ \ (\forall (i,j)), \\ & & -\boldsymbol{v}_0 \cdot \boldsymbol{v}_i + \boldsymbol{v}_0 \cdot \boldsymbol{v}_j - \boldsymbol{v}_i \cdot \boldsymbol{v}_j \geq -1 \ \ (\forall (i,j)), \\ & & \ \ \boldsymbol{v}_0 \cdot \boldsymbol{v}_i - \boldsymbol{v}_0 \cdot \boldsymbol{v}_j - \boldsymbol{v}_i \cdot \boldsymbol{v}_j \geq -1 \ \ (\forall (i,j)). \end{array}$$

As is well-known, the above problem can be transformed to a semidefinite programming problem [5], which can be solved within any given gap ϵ in polynomial time by using an interior point method [1, 9].

The MAX DICUT problem can also be formulated as an integer programming problem:

$$\begin{array}{lll} \text{(DI')} & \text{maximize} & (1/4) \sum_{(i,j) \in A} w_{ij}(1 + v_0 v_i - v_0 v_j - v_i v_j), \\ & \text{subject to} & v_0 = 1, \ \ v_i \in \{-1, 1\} \quad (\forall i \in V). \end{array}$$

Similarly, we can obtain the following relaxation problem:

$$\begin{array}{lll} (\overline{\text{DI}}) & \text{maximize} & (1/4) \sum_{(i,j) \in A} w_{ij}(1 + \boldsymbol{v}_0 \cdot \boldsymbol{v}_i - \boldsymbol{v}_0 \cdot \boldsymbol{v}_j - \boldsymbol{v}_i \cdot \boldsymbol{v}_j), \\ & \text{subject to} & \boldsymbol{v}_0 = (1, 0, \ldots, 0)^\top, \quad \boldsymbol{v}_i \in \boldsymbol{S}_n \quad (\forall i \in V) \\ & & \ \ \boldsymbol{v}_0 \cdot \boldsymbol{v}_i + \boldsymbol{v}_0 \cdot \boldsymbol{v}_j + \boldsymbol{v}_i \cdot \boldsymbol{v}_j \geq -1 \ \ (\forall (i,j)), \\ & & -\boldsymbol{v}_0 \cdot \boldsymbol{v}_i - \boldsymbol{v}_0 \cdot \boldsymbol{v}_j + \boldsymbol{v}_i \cdot \boldsymbol{v}_j \geq -1 \ \ (\forall (i,j)), \\ & & -\boldsymbol{v}_0 \cdot \boldsymbol{v}_i + \boldsymbol{v}_0 \cdot \boldsymbol{v}_j - \boldsymbol{v}_i \cdot \boldsymbol{v}_j \geq -1 \ \ (\forall (i,j)), \\ & & \ \ \boldsymbol{v}_0 \cdot \boldsymbol{v}_i - \boldsymbol{v}_0 \cdot \boldsymbol{v}_j - \boldsymbol{v}_i \cdot \boldsymbol{v}_j \geq -1 \ \ (\forall (i,j)), \end{array}$$

which can be formulated again as a semidefinite programming problem.

Hereafter, we focus on the MAX DICUT problem while leaving to the reader the necessary adaptation to the MAX 2SAT problem.

The hyperplane separation technique by Goemans and Williamson may be described as follows. Let $(\overline{\boldsymbol{v}}_1, \overline{\boldsymbol{v}}_2, \ldots, \overline{\boldsymbol{v}}_n)$ be an optimal solution to $\overline{\mathrm{DI}}$. We generate a vector $\boldsymbol{r}$ uniformly on $\boldsymbol{S}_n$ and construct the vertex-subset

$$\overline{U} = \{i \in V \mid \mathrm{sign}(\boldsymbol{r} \cdot \boldsymbol{v}_0) = \mathrm{sign}(\boldsymbol{r} \cdot \overline{\boldsymbol{v}}_i)\}$$

and the corresponding dicut

$$\overline{A} = \{(i,j) \in A \mid i \in \overline{U} \text{ and } j \notin \overline{U}\}.$$

We denote the expected weight of the dicut $\overline{A}$ by $\mathrm{E}(\overline{U})$. Then the linearity of the expectation implies that

$$\begin{aligned}
\mathrm{E}(\overline{U}) &= \sum_{(i,j)\in A} w_{ij} \Pr[\mathrm{sign}(\boldsymbol{r} \cdot \boldsymbol{v}_0) = \mathrm{sign}(\boldsymbol{r} \cdot \bar{\boldsymbol{v}}_i) \wedge \mathrm{sign}(\boldsymbol{r} \cdot \boldsymbol{v}_0) \neq \mathrm{sign}(\boldsymbol{r} \cdot \bar{\boldsymbol{v}}_j)] \\
&= \sum_{(i,j)\in A} w_{ij} \left(\frac{1}{2} \Pr[\mathrm{sign}(\boldsymbol{r} \cdot \boldsymbol{v}_0) = \mathrm{sign}(\boldsymbol{r} \cdot \bar{\boldsymbol{v}}_i)] \right. \\
&\quad \left. + \frac{1}{2} \Pr[\mathrm{sign}(\boldsymbol{r} \cdot \boldsymbol{v}_0) \neq \mathrm{sign}(\boldsymbol{r} \cdot \bar{\boldsymbol{v}}_j)] - \frac{1}{2} \Pr[\mathrm{sign}(\boldsymbol{r} \cdot \bar{\boldsymbol{v}}_i) = \mathrm{sign}(\boldsymbol{r} \cdot \bar{\boldsymbol{v}}_j)] \right) \\
&= \sum_{(i,j)\in A} w_{ij} \left(\frac{1}{2}(1 - \frac{\arccos(\boldsymbol{v}_0 \cdot \bar{\boldsymbol{v}}_i)}{\pi}) \right. \\
&\quad \left. + \frac{1}{2} \frac{\arccos(\boldsymbol{v}_0 \cdot \bar{\boldsymbol{v}}_j)}{\pi} - \frac{1}{2}(1 - \frac{\arccos(\bar{\boldsymbol{v}}_i \cdot \bar{\boldsymbol{v}}_j)}{\pi}) \right) \\
&= \sum_{(i,j)\in A} w_{ij} \frac{\arccos(\overline{\boldsymbol{v}}_i \cdot \overline{\boldsymbol{v}}_j) + \arccos(\boldsymbol{v}_0 \cdot \overline{\boldsymbol{v}}_j) - \arccos(\boldsymbol{v}_0 \cdot \overline{\boldsymbol{v}}_i)}{2\pi}.
\end{aligned}$$

To estimate approximation ratio, we must pay attention to the arrangement of $\{\boldsymbol{v}_0, \boldsymbol{v}_i, \boldsymbol{v}_j\}$. So we think the 3-dimensional linear subspace including $\{\boldsymbol{v}_0, \boldsymbol{v}_i, \boldsymbol{v}_j\}$ and treat these vectors as 3-dimensional vectors in that subspace. Let α be defined by

$$\alpha \stackrel{\text{def.}}{=} \min_{(\boldsymbol{v}_i, \boldsymbol{v}_j) \in \Omega} \frac{(1/2\pi)(\arccos(\boldsymbol{v}_i \cdot \boldsymbol{v}_j) + \arccos(\boldsymbol{v}_0 \cdot \boldsymbol{v}_j) - \arccos(\boldsymbol{v}_0 \cdot \boldsymbol{v}_i))}{(1/4)(1 + \boldsymbol{v}_0 \cdot \boldsymbol{v}_i - \boldsymbol{v}_0 \cdot \boldsymbol{v}_j - \boldsymbol{v}_i \cdot \boldsymbol{v}_j)},$$

where

$$\Omega \stackrel{\text{def.}}{=} \left\{ (\boldsymbol{v}_i, \boldsymbol{v}_j) \in \boldsymbol{S}_2 \times \boldsymbol{S}_2 \;\middle|\; \begin{array}{r} \boldsymbol{v}_0 \cdot \boldsymbol{v}_i + \boldsymbol{v}_0 \cdot \boldsymbol{v}_j + \boldsymbol{v}_i \cdot \boldsymbol{v}_j \geq -1, \\ -\boldsymbol{v}_0 \cdot \boldsymbol{v}_i - \boldsymbol{v}_0 \cdot \boldsymbol{v}_j + \boldsymbol{v}_i \cdot \boldsymbol{v}_j \geq -1, \\ -\boldsymbol{v}_0 \cdot \boldsymbol{v}_i + \boldsymbol{v}_0 \cdot \boldsymbol{v}_j - \boldsymbol{v}_i \cdot \boldsymbol{v}_j \geq -1, \\ \boldsymbol{v}_0 \cdot \boldsymbol{v}_i - \boldsymbol{v}_0 \cdot \boldsymbol{v}_j - \boldsymbol{v}_i \cdot \boldsymbol{v}_j \geq -1 \end{array} \right\},$$

and $\boldsymbol{v}_0 = (1, 0, 0)^\top$. Then the approximation ratio of the algorithm can be estimated by

$$\mathrm{E}(\overline{U}) \geq \alpha \cdot (\text{optimal value of } (\overline{\mathrm{DI}})) \geq \alpha \cdot (\text{optimal value of (DI)}).$$

Namely, the expected weight $\mathrm{E}(\overline{U})$ of the dicut generated by the algorithm is greater than or equal to α times the optimal value of (DI). It is known [5] that $\alpha > 0.79607$.

3. Hyperplane Separation by Skewed Distribution on Sphere

Goemans and Williamson's algorithm generates a separating hyperplane at random. Our algorithm generates a separating hyperplane with respect to a distribution function defined on $\boldsymbol{S}_n$ which is skewed towards $\boldsymbol{v}_0$ but uniform in any direction orthogonal to $\boldsymbol{v}_0$.

Given the n-dimensional sphere $\boldsymbol{S}_n$, we define the class of skewed distribution function $\mathcal{F}_n$ by

$$\mathcal{F}_n \stackrel{\text{def.}}{=} \left\{ f : \boldsymbol{S}_n \to \mathbf{R}_+ \;\middle|\; \begin{array}{l} \int_{\boldsymbol{S}_n} f(\boldsymbol{v})\,\mathrm{d}s = 1,\;\; f(\boldsymbol{v}) = f(-\boldsymbol{v})\ (\forall \boldsymbol{v} \in \boldsymbol{S}_n), \\ [\boldsymbol{v}_0 \cdot \boldsymbol{v} = \boldsymbol{v}_0 \cdot \boldsymbol{v}' \Rightarrow f(\boldsymbol{v}) = f(\boldsymbol{v}')]\ (\forall \boldsymbol{v}, \forall \boldsymbol{v}' \in \boldsymbol{S}_n) \end{array} \right\}.$$

Let $f \in \mathcal{F}_n$ be a skewed distribution function defined on $\boldsymbol{S}_n$. For any pair $(\boldsymbol{v}_i, \boldsymbol{v}_j) \in \boldsymbol{S}_n$, we define

$$p(\boldsymbol{v}_i, \boldsymbol{v}_j \mid f) \stackrel{\text{def.}}{=} \int_{\boldsymbol{S}_n} \frac{1}{4}(1 + \mathrm{sign}(\boldsymbol{r} \cdot \boldsymbol{v}_0) \cdot \mathrm{sign}(\boldsymbol{r} \cdot \boldsymbol{v}_i))(1 - \mathrm{sign}(\boldsymbol{r} \cdot \boldsymbol{v}_0) \cdot \mathrm{sign}(\boldsymbol{r} \cdot \boldsymbol{v}_j)) f(\boldsymbol{r})\mathrm{d}s$$

which is equal to the probability that an arc (i, j) is contained in a dicut obtained by hyperplane separation technique based on f.

Then the expectation of the weight of the dicut with respect to a feasible solution $(\overline{\boldsymbol{v}}_1, \overline{\boldsymbol{v}}_2, \ldots, \overline{\boldsymbol{v}}_n)$ of $\overline{\mathrm{DI}}$ based on the distribution function f is $\sum_{(i,j)\in A} w_{ij}\, p(\overline{\boldsymbol{v}_i}, \overline{\boldsymbol{v}_j} \mid f)$.

When we use a skewed distribution function $f \in \mathcal{F}_n$ defined on $\boldsymbol{S}_n$, the approximation ratio can be estimated by the distribution function $\widehat{f}$ defined by projection of a vector on $\boldsymbol{S}_n$ to the linear subspace spanned by $\{\boldsymbol{v}_0, \boldsymbol{v}_i, \boldsymbol{v}_j\}$. To be concrete, let H be the 3-dimensional linear subspace including $\{\boldsymbol{v}_0, \boldsymbol{v}_i, \boldsymbol{v}_j\}$ and define the distribution function $\widehat{f} \in \mathcal{F}_2$ by

$$\widehat{f}(\boldsymbol{v}') \stackrel{\text{def.}}{=} \int_{T(\boldsymbol{v}')} f(\boldsymbol{v})\,\mathrm{d}s,$$

where

$$T(\boldsymbol{v}') \stackrel{\text{def.}}{=} \{\boldsymbol{v} \in \boldsymbol{S}_n \mid \text{ the projection of } \boldsymbol{v} \text{ to } H \text{ is parallel to } \boldsymbol{v}'\}.$$

Here we note that the distribution function $\widehat{f}$ is independent of the 3-dimensional subspace H, since H contains $\boldsymbol{v}_0$ and f is uniform in any directions orthogonal to $\boldsymbol{v}_0$. For any distribution function $g \in \mathcal{F}_2$ we define

$$\alpha_g \stackrel{\text{def.}}{=} \min_{(\boldsymbol{v}_i, \boldsymbol{v}_j) \in \Omega} \frac{p(\boldsymbol{v}_i, \boldsymbol{v}_j \mid g)}{(1/4)(1 + \boldsymbol{v}_0 \cdot \boldsymbol{v}_i - \boldsymbol{v}_0 \cdot \boldsymbol{v}_j - \boldsymbol{v}_i \cdot \boldsymbol{v}_j)},$$

where we note that $p(\boldsymbol{v}_i, \boldsymbol{v}_j \mid g)$ is defined on $\boldsymbol{S}_2 = H \cap \boldsymbol{S}_n$. Then the approximation ratio of the algorithm using skewed distribution function $f \in \mathcal{F}_n$ is bounded by $\alpha_{\widehat{f}}$ from below.

To construct an algorithm with a good approximation bound, we need to find $f \in \mathcal{F}_n$ such that the induced distribution $g = \widehat{f} \in \mathcal{F}_2$ has a large value of α_g. We decompose our task into two subtasks:

(i) to identify a subclass of $\mathcal{F}_2$ consisting of g that can be induced from some $f \in \mathcal{F}_n$,
(ii) to find a function $g \in \mathcal{F}_2$ with large α_g.

In connection to (i) above, we observe that, when $n > 2$, not every distribution function $g \in \mathcal{F}_2$ has a distribution function $f \in \mathcal{F}_n$ satisfying $\widehat{f} = g$. For example, it is easy to show that there does not exist any distribution function $f \in \mathcal{F}_3$ such that

$$\widehat{f}(\boldsymbol{v}) = \begin{cases} 1/(2\sqrt{2}\pi) & (-0.5 \le \boldsymbol{v}_0 \cdot \boldsymbol{v} \le 0.5), \\ 0 & (\text{otherwise}). \end{cases}$$

In Section 4, we propose a class of functions in $\mathcal{F}_2$ such that a corresponding skewed distribution function exists for any sphere $\boldsymbol{S}_n$ with $n \ge 3$. In Section 5, we describe a numerical method for finding a good skewed distribution function in $\mathcal{F}_2$.

4. Main Theorem

This section affords the main technical results of this paper.

For any function $f \in \mathcal{F}_n$, we can characterize f by the function $P_f : [0, \pi/2] \to \mathbf{R}_+$ defined by

$$P_f(\theta) \stackrel{\text{def.}}{=} f(\boldsymbol{v})|_{\cos\theta = |\boldsymbol{v}_0 \cdot \boldsymbol{v}|}.$$

The following theorem, along with the corollary, gives a convenient class of skewed distribution function in $\mathcal{F}_n$.

Theorem 1 *Let $f \in \mathcal{F}_n$ be a skewed distribution function with $n \geq 2$ such that P_f can be represented as*

$$P_f(\theta) = \frac{1}{a}\sum_{k=0}^{\infty} a_k \cos^k \theta.$$

Then the function $P_{\widehat{f}}(\phi)$ can be described as

$$P_{\widehat{f}}(\phi) = \frac{1}{a}\sum_{k=0}^{\infty} \frac{S^{(k+n)}(1)}{S^{(k+2)}(1)} a_k \cos^k \phi,$$

where a is a constant used for normalizing the total probability to 1 and $S^{(m)}(r)$ is the area of the m dimensional sphere whose radius is equal to r.

Proof. First, we recall well-known formulae:

$$\Gamma(0) = 1, \;\; \Gamma(\frac{1}{2}) = \sqrt{\pi}, \;\; \Gamma(x+1) = x\Gamma(x),$$

$$\int_0^{\frac{\pi}{2}} \sin^p x \cos^q x \,\mathrm{d}x = \frac{\Gamma(\frac{p+1}{2})\Gamma(\frac{q+1}{2})}{2\Gamma(\frac{p+q+2}{2})}, \quad S^{(n)}(r) = \frac{2\pi^{\frac{n+1}{2}}}{\Gamma(\frac{n+1}{2})} r^n.$$

When we fix ϕ and $\mathrm{d}\phi$, we have the following:

$$\begin{aligned} & 2\pi \sin\phi P_{\widehat{f}}(\phi)\mathrm{d}\phi \\ = \;& \int_0^1 P_f(\arccos(r\cos\phi))(2\pi r \sin\phi)\left(S^{(n-3)}\left(\sqrt{1-r^2}\right)\right) \\ & \left(\frac{\sqrt{1-r^2\cos^2\phi}}{\sqrt{1-r^2}} r \frac{\mathrm{d}\phi}{\cos\phi}\right)\left(\frac{\cos\phi}{\sqrt{1-r^2\cos^2\phi}}\right)\mathrm{d}r. \end{aligned}$$

Thus we have

$$P_{\widehat{f}}(\phi) = \int_0^1 P_f(\arccos(r\cos\phi)) S^{(n-3)}\left(\sqrt{1-r^2}\right) r^2 \frac{\mathrm{d}r}{\sqrt{1-r^2}}.$$

On replacing r by $\sin\alpha$ and $P_f(\theta)$ by $(1/a)\sum_{k=0}^{\infty} a_k \cos^k \theta$, we can describe $P_{\widehat{f}}(\phi)$ as

$$\begin{aligned} P_{\widehat{f}}(\phi) &= \int_0^{\frac{\pi}{2}} \left(\frac{1}{a}\sum_{k=0}^{\infty} a_k \sin^k \alpha \cos^k \phi\right) \frac{2\pi^{\frac{n-2}{2}}}{\Gamma(\frac{n-2}{2})} \cos^{n-3}\alpha \; \sin^2\alpha \,\mathrm{d}\alpha \\ &= \frac{1}{a}\sum_{k=0}^{\infty} \frac{2\pi^{\frac{n-2}{2}}}{\Gamma(\frac{n-2}{2})} a_k \cos^k \phi \int_0^{\frac{\pi}{2}} \sin^{k+2}\alpha \; \cos^{n-3}\alpha \,\mathrm{d}\alpha \\ &= \frac{1}{a}\sum_{k=0}^{\infty} \frac{2\pi^{\frac{n-2}{2}}}{\Gamma(\frac{n-2}{2})} a_k \cos^k \phi \frac{\Gamma(\frac{k+3}{2})\Gamma(\frac{n-2}{2})}{2\Gamma(\frac{n+k+1}{2})} \\ &= \frac{1}{a}\sum_{k=0}^{\infty} \frac{\Gamma(\frac{k+3}{2})}{2\pi^{\frac{k+3}{2}}} \frac{2\pi^{\frac{n+k+1}{2}}}{\Gamma(\frac{n+k+1}{2})} a_k \cos^k \phi \\ &= \frac{1}{a}\sum_{k=0}^{\infty} \frac{S^{(k+n)}(1)}{S^{(k+2)}(1)} a_k \cos^k \phi. \end{aligned}$$

∎

The above theorem immediately implies the following.

Corollary 1 *Let $g \in \mathcal{F}_2$ be a distribution function such that P_g can be represented as*

$$P_g(\phi) = \frac{1}{b}\sum_{k=0}^{\infty} b_k \cos^k \phi$$

with $b_k \geq 0$. Then, for any $n \geq 2$, there exists a distribution function $f \in \mathcal{F}_n$ satisfying $\widehat{f} = g$ and

$$P_f(\theta) = \frac{1}{b}\sum_{k=0}^{\infty} \frac{S^{(k+2)}(1)}{S^{(k+n)}(1)} b_k \cos^k \theta,$$

where b is a constant used for normalizing the total probability to 1.

The following theorem, along with the corollary, extends the class of tractable distribution functions.

Theorem 2 *Let $f \in \mathcal{F}_n$ be a distribution function such that P_f can be represented as*

$$P_f(\theta) = (1/a)\sum_{t \in T} a_t \cos^t \theta$$

with a finite set T of nonnegative real numbers. Then the distribution function $\widehat{f}$ satisfies

$$P_{\widehat{f}}(\phi) = (1/a)\sum_{t \in T} c_t a_t \cos^t \phi,$$

where a is a normalization constant and

$$c_t = \frac{2\pi^{\frac{n-2}{2}}}{\Gamma(\frac{n-2}{2})}\int_0^{\frac{\pi}{2}} \sin^{t+2}\alpha \cos^{n-3}\alpha \, \mathrm{d}\alpha$$

for each $t \in T$.

Proof. We can prove this in a similar way to the proof of Theorem 1. ∎

This theorem implies the following.

Corollary 2 *Let $g \in \mathcal{F}_2$ be a distribution function such that P_g can be represented as*

$$P_g(\phi) = (1/b)\sum_{t \in T} b_t \cos^t \phi$$

with a finite set T of positive real numbers and $b_t \geq 0$ for $t \in T$. Then there exists a distribution function $f \in \mathcal{F}_n$ satisfying $\widehat{f} = g$ and

$$P_f(\theta) = (1/b)\sum_{t \in T} d_t b_t \cos^t \theta$$

where b is a normalization constant and

$$d_t = \left(\frac{2\pi^{\frac{n-2}{2}}}{\Gamma(\frac{n-2}{2})}\int_0^{\frac{\pi}{2}} \sin^{t+2}\alpha \cos^{n-3}\alpha \, \mathrm{d}\alpha\right)^{-1}$$

for each $t \in T$.

The above corollaries imply that if we have a good distribution function $g \in \mathcal{F}_2$ satisfying that $P_g(\phi)$ is a finite sum of nonnegative power of $\cos\phi$, then we can construct an approximation algorithm for MAX DICUT whose approximation ratio is greater than or equal to α_g.

A method of generating a distribution with the property required in Corollary 2 is described in Section 6.

5. Numerical Search for Good Distribution

We consider distribution functions $t \in \mathcal{F}_2$ satisfying $P_g(\phi) = (1/b)\cos^{1/\beta}\phi$, where $\beta \in \{1.00, 1.05, \ldots, 2.00\}$, to find that the distribution function with

$$P_g(\phi) = \cos^{(1/1.3)}\phi$$

yields the approximation ratio greater than 0.935 for MAX 2SAT, and that with

$$P_g(\phi) = \cos^{(1/1.95)}\phi$$

gives the approximation ratio greater than 0.863 for MAX DICUT.

For each function $P_g(\phi)$, we calculate the approximation ratio α_g as follows. We discretize the 2-dimensional sphere $\boldsymbol{S}_2$, choose every pair of points $(\boldsymbol{v}_i, \boldsymbol{v}_j)$ from the set

$$\left\{ (x,y,z)^\top \in \boldsymbol{S}_2 \;\middle|\; \begin{array}{l} \exists\eta, \exists\xi \in \{-32\pi/64, -31\pi/64, \ldots, 32\pi/64\} \\ x = \cos\eta,\; y = \sin\eta\cos\xi,\; z = \sin\eta\sin\xi, \end{array} \right\},$$

and calculate the value

$$\frac{p(\boldsymbol{v}_i, \boldsymbol{v}_j \mid g)}{(1/4)(1 + \boldsymbol{v}_0\cdot\boldsymbol{v}_i - \boldsymbol{v}_0\cdot\boldsymbol{v}_j - \boldsymbol{v}_i\cdot\boldsymbol{v}_j)}.$$

Next, we choose the minimum, 2nd minimum and 3rd minimum pairs of points. For each pair $(\boldsymbol{v}_i^*, \boldsymbol{v}_j^*)$ of the chosen three pairs, we make the grid size finer and check every pair of points $(\boldsymbol{v}_i, \boldsymbol{v}_j)$ satisfying that

$$\boldsymbol{v}_i \in \left\{ (x,y,z) \in \boldsymbol{S}_2 \;\middle|\; \begin{array}{l} \exists\eta, \exists\xi \in \{-64\pi/4096, -63\pi/4096, \ldots, 64\pi/4096\} \\ x = \cos(\eta_i^* + \eta),\; y = \sin(\eta_i^* + \eta)\cos(\xi_i^* + \xi), \\ z = \sin(\eta_i^* + \eta)\sin(\xi_i^* + \xi) \end{array} \right\},$$

and

$$\boldsymbol{v}_j \in \left\{ (x,y,z) \in \boldsymbol{S}_2 \;\middle|\; \begin{array}{l} \exists\eta, \exists\xi \in \{-64\pi/4096, -63\pi/4096, \ldots, 64\pi/4096\} \\ x = \cos(\eta_j^* + \eta),\; y = \sin(\eta_j^* + \eta)\cos(\xi_j^* + \xi), \\ z = \sin(\eta_j^* + \eta)\sin(\xi_j^* + \xi) \end{array} \right\},$$

where

$$\boldsymbol{v}_i^* = (\cos\eta_i^*, \sin\eta_i^*\cos\xi_i^*, \sin\eta_i^*\sin\xi_i^*)^\top$$

and

$$\boldsymbol{v}_j^* = (\cos\eta_j^*, \sin\eta_j^*\cos\xi_j^*, \sin\eta_j^*\sin\xi_j^*)^\top.$$

For each pair of points $(\boldsymbol{v}_i, \boldsymbol{v}_j)$ we calculate the value $p(\boldsymbol{v}_i, \boldsymbol{v}_j | g)$ by numerical integration.

Here we note that we have determined some polynomials with up to three terms and found that the approximation ratios are less than previous results.

6. Derandomization

Our randomized algorithm is amenable to the derandomization technique of Mahajan and Ramesh. To explain this we first describe how to generate a random vector that follows a distribution of the form specified in Corollary 2.

Let $X_0, \dots, X_n$ be independent random variables, each being a standard normal variate with the density function $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Let Z be a random variable, independent of $X_0, \dots, X_n$, such that the density function is represented as

$$\frac{1}{a} \sum_{t \in T} a_t |z|^t \mathrm{e}^{-z^2/2}$$

with a finite set T of positive real numbers and coefficients $a_t \geq 0$ for $t \in T$.

Then the distribution function $f : \boldsymbol{S}_n \to \mathbf{R}$ of $\boldsymbol{r} = (Z, X_0, \dots, X_n)/\sqrt{Z^2 + X_0^2 + \cdots + X_n^2}$ satisfies the conditions of $\mathcal{F}_n$ described in Section 3. From the above definitions,

$$\int 2\pi \sin\phi P_{\widehat{f}}(\phi) \mathrm{d}\phi = \frac{1}{2\pi a} \sum_{t \in T} a_t \int\int\int |z|^t \mathrm{e}^{-z^2/2} \mathrm{e}^{-(x_i^2 + x_j^2)/2} \mathrm{d}z \mathrm{d}x_i \mathrm{d}x_j.$$

By replacing (x_i, x_j) with $(r\cos\psi, r\sin\psi)$, we have

$$= \frac{1}{a} \sum_{t \in T} a_t \int\int |z|^t \mathrm{e}^{-z^2/2} r \mathrm{e}^{-r^2/2} \mathrm{d}z \mathrm{d}r$$

and by replacing r with $z\cos\phi$,

$$\begin{aligned}
&= \frac{1}{a} \sum_{t \in T} a_t \int\int |z|^t z^2 \tan\phi (1 + \tan^2\phi) \mathrm{e}^{-z^2 \tan^2\phi/2} \mathrm{d}z \mathrm{d}\phi \\
&= \frac{1}{a} \int \left\{ \sum_{t \in T} a_t \frac{\sin\phi}{\cos^3\phi} \mathrm{d}\phi \int_{-\infty}^{\infty} |z|^{t+2} \mathrm{e}^{-z^2/(2\cos^2\phi)} \mathrm{d}z \right\} \\
&= \frac{1}{a} \int \left\{ \sum_{t \in T} a_t \frac{\sin\phi}{\cos^3\phi} \cos^{t+3} \sqrt{2}^{t+3} \mathrm{d}\phi \right. \\
&\qquad \left. 2 \int_0^{\infty} \left(\frac{z}{\sqrt{2}\cos\phi}\right)^{t+2} \mathrm{e}^{-(z/\sqrt{2}\cos\phi)^2} \frac{\mathrm{d}z}{\sqrt{2}\cos\phi} \right\} \\
&= \frac{1}{a} \int \sin\phi \sum_{t \in T} c_t a_t \cos^t\phi \, \mathrm{d}\phi,
\end{aligned}$$

where

$$c_t = 2^{(t+5)/2} \int_0^{\infty} z^{t+2} \mathrm{e}^{-z^2} \mathrm{d}z$$

is a constant. From these equations, we obtain

$$P_{\widehat{f}}(\phi) = \frac{1}{2\pi a} \sum_{t \in T} c_t a_t \cos^t\phi.$$

It is emphasized that the index set T and the coefficients a_t with $t \in T$ can be chosen to yield a good destribution, and any choice of $\{a_t \mid t \in T\}$ leads a distribution function to which Corollary 2 is applicable.

Then it is not difficult to observe that the derandomization scheme of Mahajan and Ramesh, with some modifications, can be applicable to our randomized algorithms. Namely, we first fix Z and then fix each variable step by step by calculating expectation with conditional probabilities.

This derandomization procedure yields a derandomized algorithm with the approximation ratio bounded by $\alpha_{\widehat{f}}$. The bound $\alpha_{\widehat{f}}$ is naturally dependent on the chosen distribution specified by $\{a_t \mid t \in T\}$. An interesting finding here is that the resulting derandomized algorithm itself is independent of the chosen distribution, and coincides with the one proposed by Mahajan and Ramesh. This implies, in particular, that we have obtained improved bounds on the approximation ratios, 0.935 for MAX 2SAT and 0.863 for MAX DICUT, for the algorithms of Mahajan and Ramesh.

7. Conclusion

In this paper, we have proposed an approximation algorithm for MAX 2SAT problem whose approximation ratio is 0.935, and one for MAX DICUT problem whose approximation ratio is 0.863. Our algorithm solves the SDP relaxation problem proposed by Goemans and Williamson with additional valid constraints by using hyperplane separation technique based on skewed distribution functions $f \in \mathcal{F}_n$ satisfying that $P_{\widehat{f}}(\theta) = \cos^{(1/1.3)} \phi$ and $P_{\widehat{f}}(\theta) = \cos^{(1/1.95)} \phi$. We have derandomized our algorithms with the observation that the above distribution function is not needed in the derandomized algorithms.

The authors are grateful to Kazuo Murota for his help in preparing the manuscript.

This work is supported by a Grant-in-Aid of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] F. Alizadeh: Interior point methods in semidefinite programming with applications to combinatorial optimization, *SIAM Journal on Optimization*, **5** (1995) 13–51.

[2] U. Feige and M. X. Goemans: Approximating the value of two prover proof systems, with applications to MAX 2SAT and MAX DICUT, *Proc. of 3rd Israel Symposium on the Theory of Computing and Systems* (1995) 182–189.

[3] U. Feige and L. Lovász: Two-prover one-round proof systems: Their power and their problems, *Proc. of the 24th Annual ACM Symposium on the Theory of Computing* (1992) 733–744.

[4] M. R. Garey and D. S. Johnson: *Computers and Intractability, a Guide to the Theory of NP-Completeness* (W. H. Freeman and Company, 1979).

[5] M. X. Goemans and D. P. Williamson: Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *Journal of the ACM*, **42** (1995) 1115-1145.

[6] L. Lovász: On the Shannon capacity of a graph, *IEEE Transactions on Information Theory*, **25** (1979) 1–7.

[7] S. Mahajan and H. Ramesh: Derandomizing semidefinite programming based approximation algorithms, *Proc. of the 36th IEEE Conference on Foundations of Computer Science* (1995) 162–169.

[8] S. Matuura and T. Matsui: 0.863-approximation algorithm for MAX DICUT, *Proc. of APPROX+RANDOM2001*, LNCS 2129 (2001) 138–146.

[9] Y. Nesterov and A. Nemirovskii: *Interior Point Polynomial Methods in Convex Programming* (SIAM Publications, SIAM, Philadelphia, USA, 1994).

[10] U. Zwick: Analyzing the MAX 2-SAT and MAX DI-CUT approximation algorithms of Feige and Goemans, currently available from http://www.math.tau.ac.il/~zwick/my-online-papers.html.

Shiro Matuura
Department of Mathematical Informatics
University of Tokyo
7-3-1, Hongo, Bunkyo-ku,
Tokyo, Japan.
E-mail: shiro@misojiro.t.u-tokyo.ac.jp