

A Simpler Analysis of Goemans and Williamson's LP-relaxation for MAX SAT

Takao Asano (浅野孝夫)
Chuo University (中央大学)
Kasuga, Bunkyo-ku, Tokyo 112-8551
asano@ise.chuo-u.ac.jp

Abstract. Asano and Williamson obtained two types of best approximation algorithms for MAX SAT: one with best *proven* performance guarantee 0.7846 and the other with performance guarantee 0.8331 if a conjectured performance guarantee of 0.7977 is true in the Zwick's algorithm. Both algorithms are based on their sharpened analysis of Goemans and Williamson's LP-relaxation for MAX SAT. In this paper, we present an improved analysis which is simpler than the previous analysis by Asano and Williamson. Furthermore, algorithms based on this analysis will play a role as a better building block in designing an improved approximation algorithm for MAX SAT. Actually we can show an example that algorithms based on this analysis lead to approximation algorithms with performance guarantee 0.7877 and conjectured performance guarantee 0.8353 which are slightly better than the best known corresponding performance guarantees 0.7846 and 0.8331 respectively.

1 Introduction

MAX SAT is one of the most popular NP-hard problems and stated as follows: given a set of clauses with weights, find a truth assignment that maximizes the sum of the weights of the satisfied clauses. More precisely, an instance of MAX SAT is defined by (\mathcal{C}, w) , where \mathcal{C} is a set of boolean clauses such that each clause $C \in \mathcal{C}$ is a disjunction of literals with a positive weight $w(C)$. Let $X = \{x_1, \dots, x_n\}$ be the set of boolean variables in the clauses of \mathcal{C} . A *literal* is a variable $x \in X$ or its negation \bar{x} . For simplicity we assume $x_{n+i} = \bar{x}_i$ ($x_i = \bar{x}_{n+i}$). Thus, $\bar{X} = \{\bar{x} \mid x \in X\} = \{x_{n+1}, x_{n+2}, \dots, x_{2n}\}$ and $X \cup \bar{X} = \{x_1, \dots, x_{2n}\}$. We assume that no literals with the same variable appear more than once in a clause in \mathcal{C} . For each $x_i \in X$, let $x_i = 1$ ($x_i = 0$, resp.) if x_i is true (false, resp.). Then, $x_{n+i} = \bar{x}_i = 1 - x_i$ and a clause $C_j = x_{j_1} \vee x_{j_2} \vee \dots \vee x_{j_{k_j}} \in \mathcal{C}$ can be considered to be a function $C_j = C_j(\mathbf{x}) = 1 - \prod_{i=1}^{k_j} (1 - x_{j_i})$ on $\mathbf{x} = (x_1, \dots, x_{2n})$. Thus, $C_j = C_j(\mathbf{x}) = 0$ or 1

for any truth assignment $\mathbf{x} \in \{0, 1\}^{2n}$ with $x_i + x_{n+i} = 1$ ($i = 1, 2, \dots, n$) and C_j is satisfied if $C_j(\mathbf{x}) = 1$. The value of a truth assignment \mathbf{x} is defined to be $F_{\mathcal{C}}(\mathbf{x}) = \sum_{C_j \in \mathcal{C}} w(C_j) C_j(\mathbf{x})$. That is, the value of \mathbf{x} is the sum of the weights of the clauses in \mathcal{C} satisfied by \mathbf{x} . Thus, the goal of MAX SAT is to find an optimal truth assignment (i.e., a truth assignment of maximum value). We will also use MAX k SAT, a restricted version of the problem in which each clause has at most k literals.

Goemans and Williamson considered the following LP relaxation (GW) of MAX SAT [3]:

$$\begin{aligned} \max \quad & \sum_{C_j \in \mathcal{C}} w(C_j) z_j \\ \text{s.t.} \quad & \sum_{i=1}^{k_j} y_{j_i} \geq z_j \quad \forall C_j = x_{j_1} \vee \dots \vee x_{j_{k_j}} \in \mathcal{C} \\ & y_i + y_{n+i} = 1 \quad \forall i \in \{1, 2, \dots, n\} \\ & 0 \leq y_i \leq 1 \quad \forall i \in \{1, 2, \dots, 2n\} \\ & 0 \leq z_j \leq 1 \quad \forall C_j \in \mathcal{C}. \end{aligned}$$

In this formulation, variables $\mathbf{y} = (y_i)$ corre-

spond to the literals $\{x_1, \dots, x_{2n}\}$ and variables $z = (z_j)$ correspond to the clauses \mathcal{C} . Thus, variable $y_i = 1$ if and only if $x_i = 1$. Similarly, $z_j = 1$ if and only if C_j is satisfied. The first set of constraints implies that one of the literals in a clause is true if the clause is satisfied and thus IP formulation of this (GW) with $y_i \in \{0, 1\}$ ($\forall i \in \{1, 2, \dots, 2n\}$) and $z_j \in \{0, 1\}$ ($\forall C_j \in \mathcal{C}$) exactly corresponds to MAX SAT.

Using an optimal solution $(\mathbf{y}^*, \mathbf{z}^*)$ to this LP relaxation, Goemans and Williamson set each variable x_i to be true with probability y_i^* . Then they showed that a clause $C_j = x_{j_1} \vee x_{j_2} \vee \dots \vee x_{j_{k_j}}$ is satisfied by this random truth assignment $\mathbf{x}^p = \mathbf{y}^*$ with probability

$$\begin{aligned} C_j(\mathbf{y}^*) &= 1 - (1 - y_{j_1}^*)(1 - y_{j_2}^*) \cdots (1 - y_{j_{k_j}}^*) \\ &\geq \left(1 - \left(1 - \frac{1}{k}\right)^k\right) z_j^*. \end{aligned}$$

This implies that the expected value $F(\mathbf{y}^*)$ of the random truth assignment \mathbf{y}^* obtained in this way satisfies

$$\begin{aligned} F(\mathbf{y}^*) &= \sum_{C_j \in \mathcal{C}} w(C_j) C_j(\mathbf{y}^*) \\ &\geq \sum_{k \geq 1} \left(1 - \left(1 - \frac{1}{k}\right)^k\right) W_k^* \geq \left(1 - \frac{1}{e}\right) W^*, \end{aligned}$$

where e is the base of natural logarithm, $W^* = \sum_{C_j \in \mathcal{C}} w(C_j) z_j^*$ and $W_k^* = \sum_{C_j \in \mathcal{C}_k} w(C_j) z_j^*$ (\mathcal{C}_k is the set of clauses in \mathcal{C} with k literals and $W^* = \sum_{C_j \in \mathcal{C}} w(C_j) z_j^* \geq \hat{W} = \sum_{C_j \in \mathcal{C}} w(C_j) \hat{z}_j$ for an optimal solution $(\hat{\mathbf{y}}, \hat{\mathbf{z}})$ to the IP formulation of MAX SAT). Since $(1 - \frac{1}{e}) \approx 0.632$, this is a 0.632-approximation algorithm.

Goemans and Williamson [3] also considered three other non-linear randomized rounding algorithms. In these algorithms, each variable x_i is set to be true with probability $f_\ell(y_i^*)$ defined as follows ($\ell = 1, 2, 3$).

$$f_1(y) = \begin{cases} \frac{3}{4}y + \frac{1}{4} & \text{if } 0 \leq y \leq \frac{1}{3} \\ \frac{1}{2} & \text{if } \frac{1}{3} \leq y \leq \frac{2}{3} \\ \frac{3}{4}y & \text{if } \frac{2}{3} \leq y \leq 1 \end{cases}$$

$$f_2(y) = (2a - 1)y + 1 - a \quad \left(\frac{3}{4} \leq a \leq \frac{3}{\sqrt[3]{4}} - 1\right)$$

$$1 - 4^{-y} \leq f_3(y) \leq 4^{y-1}.$$

Note that $f_\ell(y_i^*) + f_\ell(y_{n+i}^*) = 1$ hold for $\ell = 1, 2$ and that $f_3(y_i^*)$ has to be chosen to satisfy $f_3(y_i^*) + f_3(y_{n+i}^*) = 1$. They then proved that all the random truth assignments $\mathbf{x}^p = f_\ell(\mathbf{y}^*) = (f_\ell(y_1^*), \dots, f_\ell(y_{2n}^*))$ obtained in this way have the expected values at least $\frac{3}{4}W^*$ and lead to $\frac{3}{4}$ -approximation algorithms.

Asano and Williamson [1] sharpened the analysis of Goemans and Williamson to provide more precise bounds on the probability of a clause $C_j = x_{j_1} \vee x_{j_2} \vee \dots \vee x_{j_k}$ with k literals being satisfied (and thus on the expected weight of satisfied clauses in \mathcal{C}_k) by the random truth assignment $\mathbf{x}^p = f_\ell(\mathbf{y}^*)$ for each k (and $\ell = 1, 2$). From now on, we assume by symmetry, $x_{j_i} = x_i$ for each $i = 1, 2, \dots, k$ since $f_\ell(x) = 1 - f_\ell(\bar{x})$ and we can set $x := \bar{x}$ if necessary. They considered clause $C_j = x_1 \vee x_2 \vee \dots \vee x_k$ corresponding to the constraint $y_1 + y_2 + \dots + y_k \geq z_j$ in the LP relaxation (GW) of MAX SAT, and gave a bound on the ratio of the probability of clause C_j being satisfied by the random truth assignment $\mathbf{x}^p = f_\ell(\mathbf{y}^*)$ ($\ell = 1, 2$) to z_j^* . Actually, they analyzed parametrized functions f_1^a and f_2^a with $\frac{1}{2} \leq a \leq 1$ defined as follows:

$$f_1^a(y) = \begin{cases} ay + 1 - a & \text{if } 0 \leq y \leq 1 - \frac{1}{2a} \\ \frac{1}{2} & \text{if } 1 - \frac{1}{2a} \leq y \leq \frac{1}{2a} \\ ay & \text{if } \frac{1}{2a} \leq y \leq 1, \end{cases} \quad (1)$$

$$f_2^a(y) = (2a - 1)y + 1 - a. \quad (2)$$

Then their results are the following [1].

Theorem 1 *The probability of $C_j = x_1 \vee \dots \vee x_k \in \mathcal{C}$ satisfied by the random truth assignment $\mathbf{x}^p = f_1^a(\mathbf{y}^*) = (f_1^a(y_1^*), \dots, f_1^a(y_{2n}^*))$ is*

$$C_j(f_1^a(\mathbf{y}^*)) = 1 - \prod_{i=1}^k (1 - f_1^a(y_i^*)) \geq \gamma_k^a z_j^*,$$

for $\frac{1}{2} \leq a \leq 1$, where f_1^a is the function defined in Eq.(1) and

$$\gamma_k^a = \begin{cases} a & \text{if } k = 1 \\ \min\{\gamma_{k,1}^a, \gamma_{k,2}^a\} & \text{if } k \geq 2 \end{cases} \quad (3)$$

with

$$\gamma_{k,1}^a = 1 - \frac{1}{2}a^{k-1} \left(1 - \frac{1 - \frac{1}{2a}}{k-1}\right)^{k-1}, \quad (4)$$

$$\gamma_{k,2}^a = 1 - a^k \left(1 - \frac{1}{k}\right)^k. \quad (5)$$

Thus, the expected value $F(f_1^a(\mathbf{y}^*))$ of the random truth assignment $\mathbf{x}^p = f_1^a(\mathbf{y}^*)$ satisfies $F(f_1^a(\mathbf{y}^*)) \geq \sum_{k \geq 1} \gamma_k^a W_k^*$.

Theorem 2 The probability of $C_j = x_1 \vee \dots \vee x_k \in \mathcal{C}$ satisfied by the random truth assignment $\mathbf{x}^p = f_2^a(\mathbf{y}^*) = (f_2^a(y_1^*), \dots, f_2^a(y_{2n}^*))$ is

$$C_j(f_2^a(\mathbf{y}^*)) = 1 - \prod_{i=1}^k (1 - f_2^a(y_i^*)) \geq \delta_k^a z_j^*,$$

for $\frac{1}{2} \leq a \leq 1$, where f_2^a is the function defined in Eq.(2) and

$$\delta_k^a = 1 - a^k \left(1 - \frac{2 - \frac{1}{a}}{k}\right)^k. \quad (6)$$

Thus, the expected value $F(f_2^a(\mathbf{y}^*))$ of the random truth assignment $\mathbf{x}^p = f_2^a(\mathbf{y}^*)$ satisfies $F(f_2^a(\mathbf{y}^*)) \geq \sum_{k \geq 1} \delta_k^a W_k^*$.

Theorem 3 For γ_k^a and δ_k^a defined in Eqs.(3) and (6), $\gamma_k^a > \delta_k^a$ hold for all $k \geq 3$ and for all a with $\frac{1}{2} < a < 1$. For $k = 1, 2$, $\gamma_k^a = \delta_k^a$ ($\gamma_1^a = \delta_1^a = a$, $\gamma_2^a = \delta_2^a = \frac{3}{4}$) hold.

2 Main Results

Asano and Williamson have not considered a parametrized function of f_3 . In this section we consider a parametrized function f_3^a of f_3 and show that it has better performance than f_1^a and f_2^a . Furthermore, its analysis (proof) is simpler. We also consider a generalization of both f_1^a and f_2^a and show that it has also better performance than f_1^a and f_2^a .

For $\frac{1}{2} \leq a \leq 1$, let f_3^a be defined as follows:

$$f_3^a(y) = \begin{cases} 1 - \frac{a}{(4a^2)^y} & \text{if } 0 \leq y \leq \frac{1}{2} \\ \frac{(4a^2)^y}{4a} & \text{if } \frac{1}{2} \leq y \leq 1. \end{cases} \quad (7)$$

Let

$$y_a = \frac{1}{a} - \frac{1}{2}. \quad (8)$$

Then the other parametrized function f_4^a is defined as follows:

$$f_4^a(y) = \begin{cases} ay + 1 - a & \text{if } 0 \leq y \leq 1 - y_a \\ \frac{a}{2}y + \frac{1}{2} - \frac{a}{4} & \text{if } 1 - y_a \leq y \leq y_a \\ ay & \text{if } y_a \leq y \leq 1. \end{cases} \quad (9)$$

Thus, $f_3^a(y) + f_3^a(1-y) = 1$ and $f_4^a(y) + f_4^a(1-y) = 1$ hold for $0 \leq y \leq 1$. Furthermore, f_3^a and f_4^a are both continuous functions which are increasing with y . Thus, $f_3^a(\frac{1}{2}) = f_4^a(\frac{1}{2}) = \frac{1}{2}$. We have the following theorems for the two parameterized functions f_3^a and f_4^a .

Theorem 4 For $\frac{1}{2} \leq a \leq \frac{\sqrt{e}}{2} = 0.82436$, the probability of $C_j = x_1 \vee x_2 \vee \dots \vee x_k \in \mathcal{C}$ being satisfied by the random truth assignment $\mathbf{x}^p = f_3^a(\mathbf{y}^*) = (f_3^a(y_1^*), \dots, f_3^a(y_{2n}^*))$ is

$$C_j(f_3^a(\mathbf{y}^*)) = 1 - \prod_{i=1}^k (1 - f_3^a(y_i^*)) \geq \zeta_k^a z_j^*, \quad (10)$$

where f_3^a is the function defined in Eq.(7) and

$$\zeta_k^a = \begin{cases} a & \text{if } k = 1 \\ 1 - \frac{1}{4}a^{k-2} & \text{if } k \geq 2. \end{cases} \quad (11)$$

Thus, the expected value $F(f_3^a(\mathbf{y}^*))$ of the random truth assignment $\mathbf{x}^p = f_3^a(\mathbf{y}^*)$ satisfies $F(f_3^a(\mathbf{y}^*)) \geq \sum_{k \geq 1} \zeta_k^a W_k^*$.

Theorem 5 For $\frac{\sqrt{e}}{2} = 0.82436 \leq a \leq 1$, the probability of $C_j = x_1 \vee x_2 \vee \dots \vee x_k \in \mathcal{C}$ being satisfied by the random truth assignment $\mathbf{x}^p = f_4^a(\mathbf{y}^*) = (f_4^a(y_1^*), \dots, f_4^a(y_{2n}^*))$ is

$$C_j(f_4^a(\mathbf{y}^*)) = 1 - \prod_{i=1}^k (1 - f_4^a(y_i^*)) \geq \eta_k^a z_j^*, \quad (12)$$

where f_4^a is the function defined in Eq.(9) and

$$\eta_k^a = \begin{cases} a & \text{if } k = 1 \\ \min\{\eta_{k,1}^a, \eta_{k,2}^a, \eta_{k,3}^a\} & \text{if } k \geq 2 \end{cases} \quad (13)$$

$$\eta_{k,1}^a = 1 - a^k \left(1 - \frac{1}{k}\right)^k, \quad \eta_{k,2}^a = 1 - \frac{a^{k-2}}{4},$$

$$\eta_{k,3}^a = 1 - \frac{a^k}{2} \left(1 - \frac{1 - y_a}{k-1}\right)^{k-1}$$

($\eta_{k,1}^a = \gamma_{k,2}^a$, $\eta_{k,2}^a = \zeta_k^a$). The expected value $F(f_4^a(\mathbf{y}^*))$ of the random truth assignment $\mathbf{x}^P = f_4^a(\mathbf{y}^*)$ satisfies $F(f_4^a(\mathbf{y}^*)) \geq \sum_{k \geq 1} \eta_k^a W_k^*$.

Theorem 6 For γ_k^a , δ_k^a , ζ_k^a , and η_k^a defined in Eqs.(3), (6), (11), and (13), we have the following.

1. If $\frac{1}{2} \leq a \leq \frac{\sqrt{e}}{2} = 0.82436$, then $\zeta_k^a > \gamma_k^a > \delta_k^a$ hold for all $k \geq 3$.
2. If $\frac{\sqrt{e}}{2} = 0.82436 \leq a \leq 1$, then $\eta_k^a > \gamma_k^a > \delta_k^a$ hold for all $k \geq 3$.
3. For $k = 1, 2$, $\gamma_k^a = \delta_k^a = \zeta_k^a$ hold if $\frac{1}{2} \leq a \leq \frac{\sqrt{e}}{2} = 0.82436$, and $\gamma_k^a = \delta_k^a = \eta_k^a$ hold if $\frac{\sqrt{e}}{2} = 0.82436 \leq a \leq 1$ ($\gamma_1^a = \delta_1^a = \zeta_1^a = \eta_1^a = a$, $\gamma_2^a = \delta_2^a = \zeta_2^a = \eta_2^a = \frac{3}{4}$).

In this paper, we first give a proof of Theorem 4. It is very simple and we use only the following lemma.

Lemma 1 If $\frac{1}{2} \leq a \leq \frac{\sqrt{e}}{2} = 0.82436$, then $f_3^a(y) \geq ay$.

Proof. Let $g(y) \equiv \frac{(4a^2)^y}{4a} - ay$. Then its derivative is $g'(y) = \ln(4a^2) \frac{(4a^2)^y}{4a} - a$. Thus, $g'(y)$ is increasing with y and $g'(1) = a(\ln(4a^2) - 1) \leq 0$, since $\ln(4a^2) \leq \ln(4(\frac{\sqrt{e}}{2})^2) = 1$. This implies that $g'(y) \leq 0$ for all $0 \leq y \leq 1$ and that $g(y)$ is decreasing with $0 \leq y \leq 1$. Thus, $g(y)$ takes a minimum value at $y = 1$, i.e., $g(y) = \frac{(4a^2)^y}{4a} - ay \geq g(1) = \frac{4a^2}{4a} - a = 0$.

Now we are ready to prove the lemma. For $\frac{1}{2} \leq y \leq 1$, we have $f_3(y) - ay = g(y) = \frac{(4a^2)^y}{4a} - ay \geq 0$. For $0 \leq y \leq \frac{1}{2}$, we have

$$\begin{aligned} f_3(y) - ay &= 1 - \frac{a}{(4a^2)^y} - ay \\ &= -\frac{(4a^2)^{1-y}}{4a} + a(1-y) + 1 - a \\ &= -g(1-y) + 1 - a \\ &\geq -g\left(\frac{1}{2}\right) + 1 - a = \frac{1-a}{2} \geq 0 \end{aligned}$$

since $g(y)$ is decreasing and $g(1-y) \leq g(\frac{1}{2}) = \frac{1-a}{2}$ for $\frac{1}{2} \leq 1-y \leq 1$. ■

Proof of Theorem 4. Noting that clause $C_j = x_1 \vee x_2 \vee \dots \vee x_k$ corresponds to the constraint

$$y_1^* + y_2^* + \dots + y_k^* \geq z_j^* \quad (14)$$

in the LP relaxation (GW) of MAX SAT, we will show that

$$C_j(f_3^a(\mathbf{y}^*)) = 1 - \prod_{i=1}^k (1 - f_3^a(y_i^*)) \geq \zeta_k^a z_j^*.$$

If $k = 1$, then we have

$$C_j(f_3^a(\mathbf{y}^*)) = f_3^a(y_1^*) \geq ay_1^* \geq az_j^* = \zeta_1^a z_j^*$$

by Lemma 1 and inequality (14).

Next suppose $k \geq 2$. By symmetry, we assume $y_1^* \leq y_2^* \leq \dots \leq y_k^*$ and consider two cases as follows: Case 1: $0 \leq y_k^* \leq \frac{1}{2}$; and Case 2: $\frac{1}{2} < y_k^* \leq 1$.

Case 1: $0 \leq y_k^* \leq \frac{1}{2}$. Since all $y_i^* \leq \frac{1}{2}$ ($i = 1, 2, \dots, k$), we have $f_3^a(y_i^*) = 1 - \frac{a}{(4a^2)^{y_i^*}}$ and $1 - f_3^a(y_i^*) = \frac{a}{(4a^2)^{y_i^*}}$. Thus, we have

$$\begin{aligned} C_j(f_3^a(\mathbf{y}^*)) &= 1 - \prod_{i=1}^k (1 - f_3^a(y_i^*)) \\ &= 1 - \prod_{i=1}^k \frac{a}{(4a^2)^{y_i^*}} = 1 - \frac{a^k}{(4a^2)^{\sum_{i=1}^k y_i^*}} \\ &\geq 1 - \frac{a^k}{(4a^2)^{z_j^*}} \geq \left(1 - \frac{a^k}{4a^2}\right) z_j^* = \zeta_k^a z_j^*, \end{aligned}$$

where the first inequality follows by inequality (14), and the second inequality follows from the fact that $1 - \frac{a^k}{(4a^2)^{z_j^*}}$ is a concave function in $0 \leq z_j^* \leq 1$.

Case 2: $\frac{1}{2} < y_k^* \leq 1$. Suppose $z_j^* < y_k^*$. Then $1 - f_3^a(y_i^*) \leq a$ ($i = 1, 2, \dots, k-1$) and we have

$$\begin{aligned} C_j(f_3^a(\mathbf{y}^*)) &= 1 - \prod_{i=1}^k (1 - f_3^a(y_i^*)) \\ &\geq 1 - a^{k-1} \left(1 - \frac{(4a^2)^{y_k^*}}{4a}\right) \\ &\geq 1 - a^{k-1} \left(1 - \frac{(4a^2)^{z_j^*}}{4a}\right) \end{aligned}$$

$$\begin{aligned}
&\geq \left(1 - a^{k-1} \left(1 - \frac{(4a^2)}{4a}\right)\right) z_j^* \\
&= \left(1 - a^{k-1}(1-a)\right) z_j^* \\
&\geq \left(1 - \frac{a^{k-2}}{4}\right) z_j^* = \zeta_k^a z_j^*,
\end{aligned}$$

where the second inequality follows by $z_j^* < y_k^*$, the third inequality by the fact that $1 - a^{k-1} \left(1 - \frac{(4a^2)^{z_j^*}}{4a}\right)$ is a concave function in $0 \leq z_j^* \leq 1$, and the fourth inequality by $a(1-a) \leq \frac{1}{4}$.

Thus, we can assume $z_j^* \geq y_k^*$. If $y_{k-1}^* > \frac{1}{2}$, then $1 - f_3^a(y_i^*) \leq a$ ($i = 1, 2, \dots, k-2$), $1 - f_3^a(y_i^*) = 1 - \frac{(4a^2)^{y_i^*}}{4a} \leq \frac{1}{2}$ ($i = k-1, k$), and $z_j^* \leq 1$, and we have

$$\begin{aligned}
C_j(f_3^a(\mathbf{y}^*)) &= 1 - \prod_{i=1}^k (1 - f_3^a(y_i^*)) \\
&\geq 1 - a^{k-2} \left(\frac{1}{2}\right)^2 = 1 - \frac{a^{k-2}}{4} \\
&\geq \left(1 - \frac{a^{k-2}}{4}\right) z_j^* = \zeta_k^a z_j^*.
\end{aligned}$$

Thus, we can assume $y_{k-1}^* \leq \frac{1}{2}$. Then, since $1 - f_3^a(y_i^*) = \frac{a}{(4a^2)^{y_i^*}}$ ($i = 1, 2, \dots, k-1$), we have

$$\begin{aligned}
C_j(f_3^a(\mathbf{y}^*)) &= 1 - \prod_{i=1}^k (1 - f_3^a(y_i^*)) \\
&= 1 - \frac{a^{k-1}}{(4a^2)^{\sum_{i=1}^{k-1} y_i^*}} \left(1 - \frac{(4a^2)^{y_k^*}}{4a}\right) \\
&\geq 1 - \frac{a^{k-1}}{(4a^2)^{z_j^* - y_k^*}} \left(1 - \frac{(4a^2)^{y_k^*}}{4a}\right) \\
&= 1 - \frac{a^{k-1}}{(4a^2)^{z_j^*}} (4a^2)^{y_k^*} \left(1 - \frac{(4a^2)^{y_k^*}}{4a}\right) \\
&\geq 1 - \frac{a^{k-1}}{(4a^2)^{z_j^*}} a = 1 - \frac{a^k}{(4a^2)^{z_j^*}} \\
&\geq \left(1 - \frac{a^{k-2}}{4}\right) z_j^* = \zeta_k^a z_j^*,
\end{aligned}$$

by inequality (14), $y_k^* \leq z_j^*$, $(4a^2)^{y_k^*} \left(1 - \frac{(4a^2)^{y_k^*}}{4a}\right) = u \left(1 - \frac{u}{4a}\right) \leq a$ with $u = (4a^2)^{y_k^*}$, and the fact

that $1 - \frac{a^k}{(4a^2)^{z_j^*}}$ is a concave function in $0 \leq z_j^* \leq 1$. ■

Proofs of Theorems 5 and 6. Proofs of Theorems 5 and 6 are almost similar to ones in Asano and Williamson [1]. In this sense, proofs may be a little complicated, however, they can be done in a systematic way. We omit details. ■

3 Improved Approximation Algorithms

In this section, we briefly outline our improved approximation algorithms for MAX SAT based on a hybrid approach which is described in detail in Asano and Williamson [1]. We use a semidefinite programming relaxation of MAX SAT which is a combination of ones given by Goemans and Williamson [4], Feige and Goemans [2], Karloff and Zwick [6], Halperin and Zwick [5], and Zwick [7]. Our algorithms pick the best solution returned by the four algorithms corresponding to (1) f_3^a in Goemans and Williamson [3], (2) MAX 2SAT algorithm of Feige and Goemans [2] or of Halperin and Zwick [5], (3) MAX 3SAT algorithm of Karloff and Zwick [6] or of Halperin and Zwick [5], and (4) Zwick's MAX SAT algorithm with a conjectured performance guarantee 0.7977 [7]. The expected value of the solution is at least as good as the expected value of an algorithm that uses Algorithm (i) with probability p_i , where $p_1 + p_2 + p_3 + p_4 = 1$.

Our first algorithm pick the best solution returned by the three algorithms corresponding to (1) f_3^a in Goemans and Williamson [3], (2) Feige and Goemans's MAX 2SAT algorithm [2], and (3) Karloff and Zwick's MAX 3SAT algorithm [6] (this implies that $p_4 = 0$). From the arguments in Section 3, the probability that a clause $C_j \in \mathcal{C}_k$ is satisfied by Algorithm (1) is at least $\zeta_k^a z_j^*$, where ζ_k^a is defined in Eq.(11). Similarly, from the arguments in [4, 2], the probability that a clause $C_j \in \mathcal{C}_k$ is satisfied by Algorithm (2) is

$$\text{at least } 0.93109 \cdot \frac{2}{k} z_j^* \text{ for } k \geq 2,$$

and at least $0.97653z_j^*$ for $k = 1$.

By an analysis obtained by Karloff and Zwick [6] and an argument similar to one in [4], the probability that a clause $C_j \in \mathcal{C}_k$ is satisfied by Algorithm (3) is at least

$$\text{at least } \frac{3}{k} \frac{7}{8} z_j^* \quad \text{for } k \geq 3,$$

$$\text{and at least } 0.87856z_j^* \quad \text{for } k = 1, 2.$$

Suppose that we set $a = 0.74054$, $p_1 = 0.7861$, $p_2 = 0.1637$, and $p_3 = 0.0502$ ($p_4 = 0$). Then

$$ap_1 + 0.97653p_2 + 0.87856p_3 \geq 0.7860$$

for $k = 1$,

$$\frac{3}{4}p_1 + 0.93109p_2 + 0.87856p_3 \geq 0.7860$$

for $k = 2$,

$$\zeta_k^a p_1 + \frac{2 \times 0.93109}{k} p_2 + \frac{3}{k} \frac{7}{8} p_3 \geq 0.7860$$

for $k \geq 3$.

Thus this is a 0.7860-approximation algorithm. Note that, under same conditions in Asano and Williamson [1], the algorithm picking the best solution returned by the three algorithms corresponding to (1) f_1^a with $a = \frac{3}{4}$ in Goemans and Williamson [3], (2) Feige and Goemans [2], and (3) Karloff and Zwick [6] only achieves the performance guarantee 0.7846.

Suppose next that we use three algorithms (1) f_3^a in Goemans and Williamson [3], (2) Halperin and Zwick's MAX 2SAT algorithm [5], and (3) Halperin and Zwick's MAX 3SAT algorithm [5] instead of Feige and Goemans [2] and Karloff and Zwick [6]. If we set $a = 0.739634$, $p_1 = 0.787777$, $p_2 = 0.157346$, and $p_3 = 0.054877$, then we have

$$ap_1 + 0.9828p_2 + 0.9197p_3 \geq 0.7877$$

for $k = 1$,

$$\frac{3}{4}p_1 + 0.9309p_2 + 0.9197p_3 \geq 0.7877$$

for $k = 2$,

$$\zeta_k^a p_1 + \frac{2 \times 0.9309}{k} p_2 + \frac{3}{k} \frac{7}{8} p_3 \geq 0.7877$$

for $k \geq 3$.

Thus we have a 0.7877-approximation algorithm for MAX SAT (note that the performance guarantees of Halperin and Zwick's MAX 2SAT and

MAX 3SAT algorithms are based on the numerical evidence [5]).

Suppose finally that we use two algorithms (1) f_4^a in Goemans and Williamson [3] and (4) Zwick's MAX SAT algorithm with a conjectured performance guarantee 0.7977 [7]. If we set $a = 0.907180$, $p_1 = 0.343137$ and $p_4 = 0.656863$ ($p_2 = p_3 = 0$), then the probability of clause C_j with k literals being satisfied can be shown to be at least $0.8353z_j^*$ for each $k \geq 1$. Thus, we can obtain a 0.8353-approximation algorithm for MAX SAT if a conjectured performance guarantee 0.7977 is true in Zwick's MAX SAT algorithm [7].

References

- [1] T. Asano and D.P. Williamson, Improved approximation algorithms for MAX SAT, *Journal of Algorithms* 42, pp.173–202, 2002.
- [2] U. Feige and M. X. Goemans, Approximating the value of two prover proof systems, with applications to MAX 2SAT and MAX DICUT, In *Proc. 3rd Israel Symposium on Theory of Computing and Systems*, pp. 182–189, 1995.
- [3] M. X. Goemans and D. P. Williamson, New 3/4-approximation algorithms for the maximum satisfiability problem, *SIAM Journal on Discrete Mathematics* 7, pp. 656–666, 1994.
- [4] M. X. Goemans and D. P. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *Journal of the ACM* 42, pp. 1115–1145, 1995.
- [5] E. Halperin and U. Zwick, Approximation algorithms for MAX 4-SAT and rounding procedures for semidefinite programs, In *Proc. 7th MPS Conference on Integer Programming and Combinatorial Optimization, LNCS 1610*, pp. 202–217, 1999.
- [6] H. Karloff and U. Zwick, A 7/8-approximation algorithm for MAX 3SAT?, In *Proc. 38th IEEE Symposium on the Foundations of Computer Science*, pp. 406–415, 1997.
- [7] U. Zwick, Outward rotations: a tool for rounding solutions of semidefinite programming relaxations, with applications to MAX CUT and other problems, In *Proc. 31st ACM Symposium on the Theory of Computing*, pp. 679–687, 1999.