1 Background

As its name suggests, the quantum approximate optimization algorithm (QAOA) is a quantum algorithm for finding approximate solutions to optimization problems. Common examples include constraint satisfaction problems, for example, MaxCut. QAOA can be thought of as a discretization of the quantum adiabatic algorithm (QAA or QADI), which uses adiabatic quantum computing to solve optimization problems.

Given a problem Hamiltonian, \( C \), and a driving Hamiltonian, \( B \), and a specified number of layers or steps, \( p \), QAOA \( p \) can be described by the following circuit,

\[
|\gamma,\beta\rangle = e^{-i\beta_p B} e^{-i\gamma_p C} \cdots e^{-i\beta_1 B} e^{-i\gamma_1 C} |s\rangle
\]  

(1.1)

where \( |s\rangle \) is the ground state of \( B \), and where \( \gamma_i \) and \( \beta_i \) are user-defined parameters. In applying QAOA \( p \) to an optimization problem, we wish to maximize (or minimize) \( \langle \gamma,\beta | C | \gamma,\beta \rangle \), over all possible \( \gamma_i \) and \( \beta_i \).

The problem Hamiltonian encodes the optimization problem such that the ground state of the Hamiltonian is the optimal solution to the optimization problem. Given a problem specified by \( n \) bits and \( m \) clauses, the objective is defined as,

\[
C(z) = \sum_{\alpha=1}^{m} C_\alpha(z)
\]  

(1.2)

where \( z \) is an \( n \) bit string and \( C_\alpha(z) = 1 \) if \( z \) satisfies that clause, and is 0 otherwise.

For example in the case of MaxCut given a graph, \( G = (V,E) \) (\( n = |V|, m = |E| \)), the following problem Hamiltonian is used,

\[
C = \sum_{ij \in E} C_{ij} \quad \text{where} \quad C_{ij} = \frac{1}{2} (I - Z_i Z_j)
\]  

(1.3)

A commonly used driving Hamiltonian is,

\[
B = \sum_{j \in [n]} X_j
\]  

(1.4)

Thus, \( |s\rangle = |+\rangle^{\otimes n} \), the uniform superposition of all computational basis state. Intuitively, this driving Hamiltonian “maps” all computational basis states to all other computational basis states equally.

For optimizing the parameters, \( \gamma_i \) and \( \beta_i \), we consider the expectation of the Problem Hamiltonian in the output state of QAOA \( p \), which is given by,

\[
F_p(\gamma, \beta) = \langle \gamma, \beta | C | \gamma, \beta \rangle
\]  

(1.5)

Let \( M_p \) be the maximum (or minimum) of \( F_p \) over all angles,

\[
M_p = \max_{\gamma_i, \beta_i} F_p(\gamma, \beta)
\]  

(1.6)
Given $\gamma$ and $\beta$ vectors of $p - 1$ angles for $p - 1$ layers of QAOA. We can construct, $\gamma'$ and $\beta'$ vectors of $p$ angles, by setting the last angles, $\gamma_p$ and $\beta_p$, to be 0. This implies,

$$F_{p - 1}(\gamma, \beta) = F_p(\gamma', \beta') \implies M_p \geq M_{p - 1}$$

In other words with more layers, you have more degrees of freedom and thus, potentially better solutions.

This supposedly implies that,

$$\lim_{p \to \infty} M_p = \max_z C(z)$$

2 A Bounded Occurrence Constraint Problem

Max E3LIN2 is a combinatorial problem of bounded occurrences; given a set of linear equations containing exactly three boolean variables (E3) which sum to 0 or 1 mod 2 (LIN2), find a solution which maximizes the number of satisfied solutions. Each variable is guaranteed to be in no more than $D$ equations.

One can use Gaussian elimination to determine if the set of linear equations has a solution; however, the problem becomes hard, when we wish to maximize the number of equations satisfied. Specifically, for general instances (without bounded occurrences) it has been shown that there is no efficient $(1/2 + \epsilon)$ classical algorithm unless P=NP. Note, a random string is expected to satisfy half of the equations.

In the case we consider, where each variable is contained in most $D$ clauses, there is a classical algorithm [3] for a similar but more general problem, Max-$k$XOR, with an approximation ratio of,

$$\frac{1}{2} + \frac{\text{constant}}{D^{1/2}}$$

It can be shown [2] that one layer of QAOA will efficiently produce a string that satisfies,

$$\frac{1}{2} + \frac{1}{101D^{1/2}\ln D}$$

times the number of equations. However, in the typical case, the output string will satisfy

$$\frac{1}{2} + \frac{1}{2\sqrt{3e}D^{1/2}}$$

3 The Approximation Ratio for QAOA$_1$

In applying QAOA to this problem, we will use the standard driving hamiltonian,

$$B = \sum_{i=1}^{n} X_i$$

where $n$ is the number of variables. Thus, the initial state will be the ground state of $B$, $|s\rangle = |+\rangle^{\otimes n}$. The objective operator for each linear equation of E3LIN2 can be written as,

$$\frac{1}{2} (1 \pm Z_a Z_b Z_c)$$

where $a, b, c$ are the variables in the equation and the $\pm$ is determined by the sum of the equation.

Dropping the additive constant, which does not affect the solution to the problem, the problem Hamiltonian can be write as,

$$C = \frac{1}{2} \sum_{a < b < c} d_{abc} Z_a Z_b Z_c$$
where $d_{abc}$ is zero if there is no equation containing $a, b, c$ and $\pm 1$ if there is and depending on the sum.

Using these Hamiltonians and some parameters, $\gamma, \beta$, the expected number of satisfied equations will be,

$$
m \frac{1}{2} + \langle -\gamma, \beta | C | -\gamma, \beta \rangle \tag{3.4}
$$

where

$$
| -\gamma, \beta \rangle = e^{-i\beta B} e^{i\gamma C} | s \rangle \tag{3.5}
$$

For simplicity, we will set $\beta = \pi/4$. We will show that there exists $\gamma \in \left[ -\frac{1}{10D^{1/2}}, \frac{1}{10D^{1/2}} \right]$, such that the expectation is,

$$
\frac{1}{2} + \frac{1}{101D^{1/2} \ln D} \tag{3.6}
$$

Finding the actual value of $\gamma$ for a specific problem can be done using an efficient search.

### 3.1 Proof of Bound

Here we'll give the proof of the lower bound from [2]. The main idea in the proof is to break up the expectation into individual clauses (e.g. $Z_iZ_jZ_k$) and then see how the simplified observable interacts with the terms in the QAOA circuit. Specifically, only terms in $B$ and $C$ containing variables in the clause (e.g. $Z_iZ_jZ_k$) will contribute to the expectation. This will also us to simplify the general expectation into a trigonometric formula depending on fewer terms, which we will then be able to bound using some probability inequalities and facts about trigonometric functions.

Consider a specific clause, without loss of generality, we will say the clause contains variables 1, 2, and 3,

$$
\frac{1}{2} d_{123} \langle s | e^{-i\gamma C} e^{i\beta B} Z_1Z_2Z_3 e^{-i\beta B} e^{i\gamma C} | s \rangle \tag{3.7}
$$

All terms except for $X_1 + X_2 + X_3$ in $B$ will compute with $Z_1Z_2Z_3$; furthermore, since $\beta = \pi/4$,

$$
e^{i\beta(X_1+X_2+X_3)} Z_1Z_2Z_3 e^{-i\beta(X_1+X_2+X_3)} = \left( \prod_{i=1}^{3} \cos 2\beta + i \sin 2\beta X_1 \right) Z_1Z_2Z_3 = -Y_1Y_2Y_3 \tag{3.8}
$$

Thus, the expectation (3.7) becomes,

$$
\frac{1}{2} d_{123} \langle s | e^{-i\gamma C} e^{i\beta B} Y_1Y_2Y_3 e^{-i\beta B} e^{i\gamma C} | s \rangle \tag{3.9}
$$

We can rewrite $C$, separating out the clause containing 1, 2, and 3,

$$
C = \overline{C} + \frac{1}{2} d_{123} Z_1Z_2Z_3 \tag{3.10}
$$

Now we can conjugate the $Y_1Y_2Y_3$ with the contribution from the clause 123, the expectation (3.7) becomes,

$$
\frac{1}{2} d_{123} \langle s | e^{-i\gamma \overline{C}} (\cos(\gamma d_{123}) Y_1Y_2Y_3 + \sin(\gamma d_{123}) X_1X_2X_3) e^{i\gamma \overline{C}} | s \rangle \tag{3.11}
$$

We will first evaluate,

$$
\langle s | e^{-i\gamma \overline{C}} X_1X_2X_3 e^{i\gamma \overline{C}} | s \rangle \tag{3.12}
$$
By inserting \( I = \sum_{x \in \{0,1\}} |x\rangle \langle x| \), we get
\[
\sum_{z_1,z_2,z_3} \sum_{z'_1,z'_2,z'_3} \langle s| e^{-i \gamma \overline{C}} |z_1,z_2,z_3\rangle \langle z_1,z_2,z_3| X_1 X_2 X_3 |z'_1,z'_2,z'_3\rangle \langle z'_1,z'_2,z'_3| e^{i \gamma \overline{C}} |s\rangle \] (3.13)
\[
= \sum_{z_1,z_2,z_3} \langle s| e^{-i \gamma \overline{C}} |z_1,z_2,z_3\rangle \langle -z_1,-z_2,-z_3| e^{i \gamma \overline{C}} |s\rangle \] (3.14)

Now, we will more closely consider \( \overline{C} \). We will separate it into terms involving only \( Z_1 \), etc and \( Z_1 Z_2 \), etc. Terms which involved none of 1, 2, and 3 will commute and cancel out. Thus, we can consider,
\[
\overline{C} = Z_1 C_1 + Z_2 C_2 + Z_3 C_3 + Z_1 Z_2 C_{12} + Z_1 Z_3 C_{13} + Z_2 Z_3 C_{23} \] (3.15)

Note, both terms of \( Z_1 Z_2 \), \( Z_1 Z_3 \), and \( Z_2 Z_3 \) will anticommute with \( X_1 X_2 X_3 \). Thus, they will also be cancelled out. Thus, only the first three terms of \( \overline{C} \) contribute to the expectation; and the \( X_1 X_2 X_3 \) term of the expectation \((3.12)\) becomes,
\[
\frac{1}{8} \sum_{z_1,z_2,z_3} \langle \overline{s}| e^{-2 \gamma (z_1 C_1 + z_2 C_2 + z_3 C_3)} |\overline{s}\rangle \] (3.16)
where \( |\overline{s}\rangle = \prod_{a \in Q} |+\rangle_a \), where \( Q \) contains the terms appearing in \( C_1 \), \( C_2 \), and \( C_3 \). Let \( q \) be the size of \( Q \), which can be as large as \( 6D \).

Writing out the above \((3.16)\) and using some trigonometry identities, we get,
\[
\frac{1}{4} \langle \overline{s}| \{ \cos(2 \gamma (C_1 + C_2 + C_3)) + \cos(2 \gamma (C_1 - C_2 - C_3)) + \cos(2 \gamma (-C_1 + C_2 - C_3)) + \cos(2 \gamma (-C_1 - C_2 + C_3)) \} |\overline{s}\rangle \] (3.17)
\[
= \cos(2 \gamma (C_1 + C_2 + C_3)) + \cos(2 \gamma (-C_1 + C_2 - C_3)) + \cos(2 \gamma (-C_1 - C_2 + C_3)) \] (3.18)

Using,
\[
C_i = \frac{1}{2} \sum_{a < b} d_{ab} Z_a Z_b \] (3.19)
we can rewrite it \((3.17)\) as,
\[
\frac{1}{4} \cdot 2^q \sum_{z_a \in Q} \{ \cos(\gamma (c_1(z) + c_2(z) + c_3(z))) + \cos(\gamma (c_1(z) - c_2(z) - c_3(z))) \} \] (3.20)
\[
+ \cos(\gamma (-c_1(z) + c_2(z) - c_3(z))) + \cos(\gamma (-c_1(z) - c_2(z) + c_3(z))) \] (3.21)

We can think of each \( c_i \) as random variables from a distribution of \( q \) binary variables, then the contribution of the \( X_1 X_2 X_3 \) term \((3.12)\) to the original expectation, is the following expectation,
\[
\frac{1}{8} \mathbb{E}_x \{ \cos(\gamma (c_1 + c_2 + c_3)) + \cos(\gamma (c_1 - c_2 - c_3)) \} \] (3.22)
\[
+ \cos(\gamma (-c_1 + c_2 - c_3)) + \cos(\gamma (-c_1 - c_2 + c_3)) \] (3.23)

The contribution from the \( Y_1 Y_2 Y_3 \) can be derived similarly and is,
\[
\frac{1}{8} \mathbb{E}_x \{ \sin(\gamma (c_1 + c_2 + c_3)) + \sin(\gamma (c_1 - c_2 - c_3)) \} \] (3.24)
\[
+ \sin(\gamma (-c_1 + c_2 - c_3)) + \sin(\gamma (-c_1 - c_2 + c_3)) \] (3.25)

Therefore, the expectation \((3.7)\) becomes,
\[
\frac{1}{8} d_{123} \mathbb{E}_x \{ \sin(\gamma (d_{123} + c_1 + c_2 + c_3)) + \sin(\gamma (d_{123} + c_1 - c_2 - c_3)) \} \] (3.26)
\[
+ \sin(\gamma (d_{123} - c_1 + c_2 - c_3)) + \sin(\gamma (d_{123} - c_1 - c_2 + c_3)) \] (3.27)
Now, we consider the Taylor expansion of the above in terms of $\gamma$, singling out the linear term,

$$\frac{1}{2}d_{123}^2\gamma + P_{123}^k(\gamma) + R_{123}^k(\gamma)$$

(3.28)

where

$$P_{123}^k(\gamma) = \frac{1}{8}d_{123} \sum_{j=3,5,...}^{k} \gamma^j(-1)^{(j-1)/2} \sum_{j=3,5,...}^{k} \frac{1}{j!} E_z[(d_{123} + c_1 + c_2 + c_3)^j + ...]$$

(3.29)

and

$$|R_{123}^k| \leq \frac{1}{8} \frac{|\gamma|^{k+2}}{(k+2)!} E_z[(d_{123} + c_1 + c_2 + c_3)^{k+1} + ...]$$

(3.30)

Note, that $d_{123} = 1$ and $P_{123}^k$ starts with a cubic term and has degree $k$, where $k$ will depend on $D$.

For any polynomial of degree 2, $c$, (e.g. $c_i$), we have the following fact,

$$E_z[c_i^{k+2}] \leq (k+1)^{k+2} (E_z[c_i^2])^{(k+2)/2}$$

(3.31)

Since $E_z[c_i] = 0$,

$$E_z[(d_{123} \pm c_1 \pm c_2 \pm c_3)^2] = 1 + E_z[(c_1 \pm c_2 \pm c_3)^2]$$

(3.32)

Since $E_z[c_i]^2 = \sum_{a<b} d_{iab} d_{iab} \leq D$,

$$E_z[(d_{123} \pm c_1 \pm c_2 \pm c_3)^2] \leq 1 + 9D$$

(3.33)

Using the bound for polynomials of degree 2, this implies,

$$E_z[(d_{123} \pm c_1 \pm c_2 \pm c_3)^{k+2}] \leq (k+1)^{k+2} (1 + 9D)^{(k+2)/2}$$

(3.34)

Therefore, using Stirling’s formula and $e < 3$,

$$|R_{123}^k| \leq \frac{1}{8} \frac{|\gamma|^{k+2}}{(k+2)!} \left((k+1)^{k+2} (1 + 9D)^{(k+2)/2}\right)$$

(3.35)

$$\leq \frac{1}{2} \left(e(1 + 9D)^{1/2} |\gamma|\right)^{k+2}$$

(3.36)

$$\leq (9D^{1/2}/|\gamma|)^{k+2}$$

(3.37)

Note, the original expectation can be written,

$$\langle -\gamma, \pi/4 | C | -\gamma, \pi/4 \rangle = \sum_{a<b<c} \frac{1}{2} \gamma + P_{abc}^k(\gamma) + R_{abc}^k(\gamma)$$

(3.38)

$$= \frac{m}{2} \gamma + P^k(\gamma) + \sum_{a<b<c} R_{abc}^k(\gamma)$$

(3.39)

where $P^k(\gamma)$ is the sum of the $m$ polynomials.

By the triangle inequality,

$$|\langle -\gamma, \pi/4 | C | -\gamma, \pi/4 \rangle| \geq \left| \frac{m}{2} \gamma + P^k(\gamma) \right| - \left| \sum_{a<b<c} R_{abc}^k(\gamma) \right|$$

(3.40)

$$\geq \left| \frac{m}{2} \gamma + P^k(\gamma) \right| - m(9D^{1/2}/|\gamma|)^{k+2}$$

(3.41)
To keep the negative term small, let,

$$|\gamma| \leq \frac{1}{10D^{1/2}}$$  \hspace{1cm} (3.42)

To lower bound the positive term, we use the following fact, for any constants $a_i$ and odd $k$,

$$\max r = 0, 1, ..., k \left| x_r + a_2x_r^2 + ... + a_kx_r^k \right| \geq \frac{1}{k}$$  \hspace{1cm} (3.43)

where $x_r = \cos(\pi r/k)$.

Thus, letting $\gamma_r = \frac{1}{10D^{1/2}} \cos(\pi r/k)$,

$$\max_{r=0,1,...,k} \left| \frac{m}{2} \gamma_r + P^k(\gamma_r) \right| \geq \frac{m}{20D^{1/2}k}$$  \hspace{1cm} (3.44)

Thus,

$$\max_{r=0,1,...,k} \left| \frac{m}{2} \gamma_r + P^k(\gamma_r) \right| - m(9D^{1/2}|\gamma_r|)^{k+2} \geq \frac{m}{20D^{1/2}k} - m \left( \frac{9}{10} \right)^{k+2}$$  \hspace{1cm} (3.45)

Letting $k = 5 \ln D$, the right hand side is greater than,

$$\frac{m}{101D^{1/2} \ln D}$$  \hspace{1cm} (3.46)

Therefore,

$$\max_{r=0,1,...,k} \langle -\gamma_r, \pi/4 | C | -\gamma_r, \pi/4 \rangle \geq \frac{m}{101D^{1/2} \ln D}$$  \hspace{1cm} (3.47)

where $\gamma_r \in \left[ -\frac{1}{10D^{1/2}} , \frac{1}{10D^{1/2}} \right]$, which completes the proof of the bound.

### 3.2 The Average Case

A common practice in bounding the performance of an algorithm is to bound the expected approximation ratio in the average case; in other words, for a random problem. In this case for a fixed selection of equations, we will consider a random assignment for the sums of the equations, i.e. $d_{abc} = (x_a + x_b + x_c) \mod 2$. In other words, if there are $m$ equations, there will be $2^m$ possible choices of all $d_{abc}$, and we will pick one uniformly at random.

Again considering the expectation for the term involving variables 1, 2, and 3, since the $c_i$ do not involve $d_{123}$, the expected value of $d_{123}$ is 0. Thus, the expectation can be written as,

$$\frac{1}{8} \sin \gamma \mathbb{E}_z \left[ \cos(\gamma(c_1 + c_2 + c_3)) + \cos(\gamma(c_1 + c_2 + c_3)) \right] + \cos(\gamma(c_1 + c_2 + c_3)) + \cos(\gamma(c_1 + c_2 + c_3))$$  \hspace{1cm} (3.48)

$$= \frac{1}{2} \sin \gamma \mathbb{E}_z [\cos(\gamma c_1) \cos(\gamma c_2) \cos(\gamma c_3)]$$  \hspace{1cm} (3.49)

Now looking at the expectations of each of these terms with respect to $d_{iab}$ for $i \in \{1, 2, 3\}$,

$$\mathbb{E}_d[\cos(\gamma c_i)] = \mathbb{E}_d \left[ \gamma \sum_{3 < a < b} d_{iab} z_az_b \right]$$  \hspace{1cm} (3.51)

$$= \frac{1}{2} \mathbb{E}_d \left[ \prod_{3 < a < b} \exp(i\gamma d_{iab} z_az_b) + \prod_{3 < a < b} \exp(-i\gamma d_{iab} z_az_b) \right]$$  \hspace{1cm} (3.52)

$$= \prod_{3 < a < b} \cos(\gamma z_az_b)$$  \hspace{1cm} (3.53)

$$= \prod_{3 < a < b} \cos \gamma$$  \hspace{1cm} (3.54)
Note, the term does not depending on $z$. Since each $c_i$ depend on distinct $d$'s, the expected value over all the $d$'s is,

$$
\frac{1}{2} \sin \gamma (\cos \gamma)^{D_1 + D_2 + D_3}
$$

(3.55)

where $D_i$ is the number of terms in $c_i$, which implies $0 \leq D_1 + D_2 + D_3 \leq 3D$. Therefore, the contribution from every clause is,

$$
\frac{m}{2} \sin \gamma \cos^{3D} \gamma \leq E_d [\langle -\gamma, \pi/4 \mid C \mid -\gamma, \pi/4 \rangle] \leq \frac{m}{2} \sin \gamma
$$

(3.56)

Choosing $\gamma$ to maximize the lower bound gives,

$$
\gamma = \frac{g}{D^{1/2}}
$$

(3.57)

For large $D$ the lower bound is,

$$
\frac{m}{2} \frac{g}{D^{1/2}} \exp \left( -\frac{3}{2} g^2 \right)
$$

(3.58)

where setting $g = 1/\sqrt{3}$ gives,

$$
\frac{m}{2 \sqrt{3cD^{1/2}}}
$$

(3.59)

as desired.

It can be shown that for problem sizes of order $m$, the standard deviation of the expectation over all $d$'s is of order $\sqrt{m}$. Thus, one can argue that with high probability, a “typical” instance will have at least the above approximation ratio minus a term of order $\sqrt{m}$.

4 A Classical Approximation Algorithm and Bound

In response to the first paper on QAOA’s application to E3LIN2, i.e. Max-3XOR, which gave slightly worse bounds than those showed above, Barak et. al. [3] gave a classical algorithm, which improved upon these bounds (and is still better than the above improved bounds for QAOA).

They showed that for any odd $k$, any instance of Max-$k$XOR, there is an efficient algorithm that finds an assignment satisfying,

$$
\frac{1}{2} + \Omega \left( \frac{1}{\sqrt{D}} \right)
$$

(4.1)

of the equations, where $D$ is the max number of equations any one variable can be in. In the “triangle-free” case, they showed that there is an efficient algorithm which gives an assignment satisfying $\mu + \Omega(1/\sqrt{D})$, where $\mu$ is the number of equations that would be satisfied in a random assignment.

4.1 Proof of lower bound for $k = 3$

Here we will give the proof from [3] for the case of $k = 3$. The more general proof can be found in the paper. The paper uses a very common trick to lower bounding constraint satisfaction problems. First, they decouple the first coordinate. In other words instead of finding one solution (e.g $x_i$) for the $n$ variables; one finds two solutions, $y_i$ and $z_i$. The $y_i$ are used for the first term of each equation and the $z_i$ are used for the latter two terms. One can give a rounding scheme to convert, the $y_i$ and $z_i$, into a single solution $x_i$, which only reduces the number of satisfied equations by a constant number. Finally, one can show that a random assignment of $z_i$ (which then gives the $y_i$) gives the required lower bound.
To show the above, we will need some facts about the relationship between boolean functions and Fourier transformations. Any boolean function \( f : \{\pm 1\}^n \to \mathbb{R} \) can be represented by a multilinear polynomial, i.e. the following Fourier expansion,

\[
f(x) = \sum_{S \subseteq [n]} \hat{f}(S)x^S, \quad \text{where } x^S = \prod_{i \in S} x_i
\]

Note, the following,

\[
\mathbb{E}[f(x)] = \hat{f}(\emptyset) \tag{4.2}
\]

\[
\|f\|^2 = \mathbb{E}[f(x)^2] = \sum_S \hat{f}(S)^2 \implies \text{Var}[f(x)] = \sum_{S \neq \emptyset} \hat{f}(S)^2
\]

\[
\inf_l \sum_{S \supseteq l} \hat{f}(S)^2 = \mathbb{E}[(\partial_l f)(x)^2]
\]

Also, for any predicate \( P : \{\pm 1\}^r \to 0,1 \), \( r \geq 2 \), we have \( \text{Var}[(\partial_i P)(x)] \geq \Omega(2^{-r}) \) for all \( i \).

Finally, we will use the fact that low-degree polynomials often achieve their expectation. Let \( f : \{\pm 1\}^n \to \mathbb{R} \) be a multilinear polynomial of degree at most \( k \). Then \( \mathbb{P}[f(x) \geq \mathbb{E}[f]] \geq \frac{1}{4}\exp(-2k) \). In particular, for \( f^2 \),

\[
\mathbb{P}[|f(x)| \geq \|f\|_2] \geq \exp(-O(k))
\]

which implies,

\[
\mathbb{E}[|f(x)|] \geq \exp(-O(k)) \cdot \|f\|_2 \geq \exp(-O(k)) \cdot \text{stddev}[f(x)]
\]

In general, given an instance of the problem and an assignment \( x \in \{\pm 1\}^n \), the number of satisfied constraints is,

\[
\sum_{i=1}^{m} P_i(X_{S_i}) \tag{4.8}
\]

which can be thought of as a multilinear function with degree at most \( k \) (in our case, 3).

To simplify things, we will instead consider the fraction of satisfied constraints (i.e. divide the above by \( m \)). Furthermore, we will replace \( P_i \) with \( \bar{P}_i \),

\[
\bar{P}_i = P_i - \mathbb{E}[P_i] = P_i - \hat{P}_i(\emptyset)
\]

In this way, \( \bar{P}_i \) can be thought of as the advantage over a random assignment.

Therefore, given an instance, we will defined the associated polynomial \( \mathcal{B}(x) \) by,

\[
\mathcal{B}(x) = \frac{1}{m} \sum_{i=1}^{m} \bar{P}_i(x_{S_i}) \tag{4.10}
\]

Thus, in our case of \( k = 3 \), using the Fourier expansion,

\[
\mathcal{B}(x) = \sum_{|S|=3} \hat{\mathcal{B}}(S)x^S = \sum_{i,j,k \in [n]} a_{ijk}x_ix_jx_k \tag{4.11}
\]

where \( \hat{\mathcal{B}}(S) \in \{\pm \frac{1}{m}, 0\} \) depending on whether or not a constraint with those variables exists in the instance. Note, \( a_{ijk} = \frac{1}{8} \hat{\mathcal{B}}((i,j,k)) \).

We will use the trick of “decoupling” the first coordinate, i.e. the algorithm will consider \( \hat{\mathcal{B}}(y,z) = \sum_{i,j,k} a_{ijk}y_i z_j z_k \) for some \( y_i \) and \( z_i \). The algorithm will produce a good assignment for \( \mathcal{B} \) and round to produce as assignment for the original using the following randomized rounding scheme,

\[
w.p. \frac{4}{9}, x_i = \begin{cases} y_i & w.p. \frac{1}{2} \\ z_i & w.p. \frac{1}{2} \end{cases} \quad w.p. \frac{4}{9}, x_i = \begin{cases} y_i & w.p. \frac{1}{2} \\ -z_i & w.p. \frac{1}{2} \end{cases} \quad w.p. \frac{1}{9}, x_i = -y_i \tag{4.12}
\]
Thus, the expectation is equal to,

$$
\mathbb{E}[\mathcal{B}(x)] = \frac{4}{9} \sum_{i,j,k} a_{ijk} \left( \frac{y_i + z_i}{2} \right) \left( \frac{y_j + z_j}{2} \right) \left( \frac{y_z + z_z}{2} \right)
$$

(4.13)

$$
+ \frac{4}{9} \sum_{i,j,k} a_{ijk} \left( \frac{y_i - z_i}{2} \right) \left( \frac{y_j - z_j}{2} \right) \left( \frac{y_z - z_z}{2} \right)
+ \frac{1}{9} \sum_{i,j,k} a_{ijk} (-y_i)(-y_j)(-y_k)
$$

(4.14)

$$
= \frac{1}{9} \sum_{i,j,k} a_{ijk} (y_i z_j z_k + z_i y_j z_k + z_i z_j y_k)
$$

(4.15)

$$
= \frac{1}{3} \tilde{\mathcal{B}}(y, z)
$$

(4.16)

Now, we will write $\tilde{\mathcal{B}}(y, z) = \sum_i y_i G_i(z)$, where $G_i(z) = \sum_{j,k} a_{ijk} z_j z_k$. The algorithm just needs to find an assignment for $z$ such that $\sum_i |G_i(z)|$ is large; then, we can take $y_i = \text{sgn}(G_i(z))$.

The algorithm simply chooses a random $z \in \{0, 1\}^n$ uniformly. Thus, $\mathbb{E}[|G_i(z)|^2] = \sum_{j<k} (2a_{ijk})^2 = \frac{1}{9} \text{inf}_i[\mathcal{B}]$. Applying the fact about low-degree polynomials achieving their expectations gives, $\mathbb{E}[|G_i(z)|] \geq \Omega(1) \cdot \sqrt{\text{inf}_i[\mathcal{B}]}$. Finally, since $\text{inf}_i[\mathcal{B}] = \deg(i)/4m^2$,

$$
\mathbb{E} \left[ \sum_i |G_i(x)| \right] \geq \Omega(1) \cdot \sum_i \frac{\sqrt{\deg(i)}}{m} \geq \Omega(1) \cdot \sum_i \frac{\deg(i)}{m \sqrt{D}} = \frac{\Omega(1)}{\sqrt{D}}
$$

(4.17)

Since $\sum_i |G_i(x)|$ is bounded by $1/2$, by Markov’s inequality, the algorithm can find a $z$ achieving $\sum_i |G_i(x)| \geq \Omega(1/\sqrt{D})$ with high probability after $O(\sqrt{D})$ many trials of $z$. Then, as stated above, $y$ can be chosen to give $\tilde{\mathcal{B}}(y, z) \geq \Omega(1/\sqrt{D})$, which gives the correct lower bound for the expectation of $\mathcal{B}(x)$.

References

