A Quantum Approximate Optimization Algorithm Applied to a Bounded Occurrence Constraint Problem

Edward Farhi and Jeffrey Goldstone

Center for Theoretical Physics
Massachusetts Institute of Technology
Cambridge, MA 02139

Sam Gutmann

Abstract

We apply our recent Quantum Approximate Optimization Algorithm to the combinatorial problem of bounded occurrence Max E3LIN2. The input is a set of linear equations each of which contains exactly three boolean variables and each equation says that the sum of the variables mod 2 is 0 or is 1. Every variable is in no more than $D$ equations. A random string will satisfy $1/2$ of the equations. We show that the level one QAOA will efficiently produce a string that satisfies $\left(\frac{1}{2} + \frac{1}{22D^{3/4}}\right)$ times the number of equations. This beats the best known classical algorithm for this problem which has an approximation ratio of $\left(\frac{1}{2} + \frac{\text{constant}}{D}\right)$. We also show that if the hypergraph that describes the instance has no small loops then the quantum computer will output a string that satisfies $\left(\frac{1}{2} + \frac{1}{2\sqrt{3e}D^{1/2}}\right)$ times the number of equations.
I. INTRODUCTION

Recently we introduced a Quantum Approximate Optimization Algorithm[1], QAOA, which can be used to find an approximate solution for a combinatorial optimization problem. The algorithm depends on an integer parameter \( p \geq 1 \) and the approximation improves as \( p \) increases. Here we only use the \( p = 1 \) algorithm which we now restate. The input is an \( n \) bit instance of a combinatorial problem specified by an objective function \( C(z) \) where \( z \) is an \( n \) bit string and \( C(z) \) counts the number of constraints satisfied by the string \( z \). The algorithm works in the \( 2^n \) dimensional Hilbert space spanned by the computational basis states \(|z\rangle\). In this basis the objective function \( C \) can be viewed as a diagonal operator,

\[
C|z\rangle = C(z)|z\rangle. \tag{1}
\]

We also introduce the operator \( B \) which is the sum of the \( \sigma_x \) operators,

\[
B = X_1 + X_2 + \ldots + X_n. \tag{2}
\]

Take as the initial state

\[
|s\rangle = \frac{1}{2^{n/2}} \sum_z |z\rangle = |+\rangle_1 |+\rangle_2 \ldots |+\rangle_n
\tag{3}
\]

which we note is an eigenstate of each of the \( X_a \). Given parameters \( \gamma \) and \( \beta \), we define the state

\[
|\gamma, \beta\rangle = e^{-i\beta B} e^{-i\gamma C} |s\rangle. \tag{4}
\]

The quantum computer is used to produce the state \(|\gamma, \beta\rangle\) which is then measured in the computational basis to produce a string \( z \). The unitary operator \( e^{-i\beta B} \) is a product of \( n \) one qubit operators. The operator \( e^{-i\gamma C} \) can be written as a product of commuting unitaries each of which comes from a constraint in \( C \) and has the same locality as the corresponding constraint. So the number of gates required to produce \(|\gamma, \beta\rangle\) is no more than \( n \) plus the number of constraints.

Note that with \( \gamma \) and \( \beta \) both equal to zero we get a random string and the algorithm is equivalent to a classical algorithm which is pick a string \( z \) at random and evaluate \( C(z) \). For some non-zero \( \gamma \) and \( \beta \) we can do better. In fact for the problem MaxCut we were able to show that on any 3-regular graph the quantum algorithm improves the approximation
ratio from 1/2 (guessing) to .6924. We now look at another problem where we improve on guessing.

Consider the combinatorial problem Max E3LIN2 over \( n \) bits. The E3 means that each clause contains exactly 3 variables. The LIN2 means that each constraint is a linear equation mod 2 so for say bits \( x_1, x_2 \) and \( x_3 \) the constraint is either \( x_1 + x_2 + x_3 = 0 \) or \( x_1 + x_2 + x_3 = 1 \). It is possible using Gaussian elimination to determine if the set of linear equations has a solution. The computational task is to maximize the number of satisfied equations in the case when the equations do not have a solution. Guessing a random string will satisfy 1/2 of the equations. For general instances there is no efficient \( (\frac{1}{2} + \epsilon) \) classical approximation algorithm unless P=NP \[2\]. We now make the restriction that every bit is in no more than \( D + 1 \) equations. (The +1 is for later convenience.) There is a classical algorithm that achieves an approximation ratio of \( (\frac{1}{2} + \frac{\text{constant}}{D}) \) \[3\]. The existence of an efficient classical approximation algorithm that achieves \( (\frac{1}{2} + \frac{\text{constant}}{D^{1/2}}) \) for a sufficiently large constant would imply that P=NP \[4\]. We will show that the \( p = 1 \) QAOA will produce a string that satisfies \( (\frac{1}{2} + \frac{1}{22D^{1/2}}) \) times the number of clauses. This also means that for every instance at least \( (\frac{1}{2} + \frac{1}{22D^{1/2}}) \) of the clauses can be satisfied but you need the quantum computer to find the optimal string.

II. THE GENERAL CASE

For E3LIN2 we can write the objective operator for any three bits \( a, b, c \) as

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

(5)

where the \( Z \) operators are \( \sigma_z \)'s and the \( \pm \) in expression (5) corresponds to the two possible choices for the equation associated with bits \( a, b \) and \( c \). Dropping the constant 1/2 we write the objective operator as

\[
C = \sum_{a,b,c} d_{abc} Z_a Z_b Z_c
\]

(6)

where \( d_{abc} \) is 0 if there is no equation involving \( a, b \) and \( c \) and \( d_{abc} \) is \(+1/2\) or \(-1/2\) if there is an equation. We will evaluate

\[
\langle \gamma, \beta | C | \gamma, \beta \rangle
\]

(7)

for certain values of \( \gamma \) and \( \beta \). The expected number of satisfied equations is 1/2 times the number of equations plus (7). For \( \gamma \) and \( \beta \) equal to 0 the expression (7) is 0 corresponding
to just guessing a random string. One approach to running the algorithm is to search for good values of $\gamma$ and $\beta$ by using repeated calls to the quantum computer to produce output strings. However, below we will show how to pick $\gamma$ and $\beta$ in advance. We will pick $\beta = \pi/4$ because it simplifies the analysis. We will show that with $\gamma = \frac{3}{22D^{1/4}}$ we get the $(\frac{1}{2} + \frac{1}{22D^{1/4}})$ result. These values are not optimal but good enough to establish our result.

Consider one term in the quantum expectation (7) that comes from the clause involving say bits 1, 2 and 3. This is

$$\pm \frac{1}{2} \langle s | e^{i\gamma C} e^{i\beta B} Z_1 Z_2 Z_3 e^{-i\beta B} e^{-i\gamma C} | s \rangle.$$  

Now in the $B$ operator all terms except the $X_1 + X_2 + X_3$ commute through the three $Z$’s. We pick $\beta = \pi/4$ and get

$$\pm \frac{1}{2} \langle s | e^{i\gamma C} Y_1 Y_2 Y_3 e^{-i\gamma C} | s \rangle,$$

where the $Y$’s are $\sigma_y$’s. We separate out the clause involving bits 1,2 and 3 in $C$ and write

$$C = \bar{C} \pm \frac{1}{2} Z_1 Z_2 Z_3.$$  

Conjugating the $Y_1 Y_2 Y_3$ with the contribution from clause 123 we get

$$\pm \frac{1}{2} \langle s | e^{i\gamma \bar{C}} (\cos(\gamma) Y_1 Y_2 Y_3 \mp \sin(\gamma) X_1 X_2 X_3) e^{-i\gamma \bar{C}} | s \rangle.$$  

We will first evaluate the term with three $X$ operators whose coefficient is always $-\frac{1}{2} \sin(\gamma)$,

$$\langle s | e^{i\gamma \bar{C}} X_1 X_2 X_3 e^{-i\gamma \bar{C}} | s \rangle.$$  

Insert two complete sets for qubits 1,2 and 3 to get

$$\sum_{z_1, z_2, z_3} \sum_{z'_1, z'_2, z'_3} \langle s | e^{i\gamma \bar{C}} | z_1, z_2, z_3 \rangle \langle z_1, z_2, z_3 | X_1 X_2 X_3 | z'_1, z'_2, z'_3 \rangle \langle z'_1, z'_2, z'_3 | e^{-i\gamma \bar{C}} | s \rangle.$$  

Now the $X$ operators are off diagonal so we get for (12)

$$\sum_{z_1, z_2, z_3} \langle s | e^{i\gamma \bar{C}} | z_1, z_2, z_3 \rangle \langle -z_1, -z_2, -z_3 | e^{-i\gamma \bar{C}} | s \rangle.$$  

We now need to look more carefully at $\bar{C}$ which includes all the clauses involving bits 1,2 and 3 except the 1,2,3 constraint. We assume that the instance has bounded degree, that is, that each bit is in no more than $D + 1$ clauses. So aside from the central clause each of bits 1,2 and 3 can be in at most $D$ clauses so there are at most $3D$ clauses in $\bar{C}$. Each of
these clauses involves two bits other than 1, 2 and 3 so there could be as many as 6D bits other than bits 1, 2 and 3 involved in the starting expression (8) as well as in (14) but there may be fewer. Let us write $\overline{C}$ as

$$\overline{C} = Z_1 C_1 + Z_2 C_2 + Z_3 C_3 + Z_1 Z_2 C_{12} + Z_1 Z_3 C_{13} + Z_2 Z_3 C_{23}$$  \hspace{1cm} (15)$$

where $C_1$ is a sum of terms of the form $\pm \frac{1}{2}Z_a Z_b$ where $a$ and $b$ are pairs of bits other than 2 and 3 that come from clauses with bit 1. Similarly for $C_2$ and $C_3$. $C_{12}$ is a sum of terms of the form $\pm \frac{1}{2}Z_a$ where bit $a$ is not 1, 2 or 3 and comes from being in a clause with bits 1 and 2. Similarly for $C_{13}$ and $C_{23}$. Note in expression (14) there is a $\overline{C}$ on the left where bits 1, 2 and 3 take the values $z_1$, $z_2$ and $z_3$ whereas on the right we have $-\overline{C}$ with the bits taking the values $-z_1$, $-z_2$ and $-z_3$. So expression (13) can be written as

$$\frac{1}{8} \sum_{z_1, z_2, z_3} \langle \overline{s} | e^{2 \gamma (z_1 C_1 + z_2 C_2 + z_3 C_3)} | \overline{s} \rangle$$  \hspace{1cm} (16)$$

where

$$| \overline{s} \rangle = \prod_{a \in Q} | + \rangle_a$$  \hspace{1cm} (17)$$

and $a$ is in the set $Q$ consisting of qubits that appear in $C_1$, $C_2$ and $C_3$. Note that the number of elements in $Q, Q$, can be as large as 6D. We can do the sum in (16) explicitly and we get

$$\frac{1}{4} \langle \overline{s} | \left[ \cos(2 \gamma (C_1 + C_2 + C_3)) + \cos(2 \gamma (C_1 - C_2 - C_3)) + \cos(2 \gamma (-C_1 - C_2 + C_3)) + \cos(2 \gamma (-C_1 + C_2 - C_3)) \right] | \overline{s} \rangle.$$  \hspace{1cm} (18)$$

Now we write

$$C_i = \frac{1}{2} \sum_{a < b} Z_a J_{ab}^{(i)} Z_b , \; i = 1, 2, 3.$$  \hspace{1cm} (19)$$

and $J^{(1)}$, $J^{(2)}$ and $J^{(3)}$ have entries of 1 or −1 if the pair $a, b$ comes from a + or − clause and the entry is 0 if the pair $a, b$ is not associated with a clause. We can write the quantum expectation in (18) as

$$\frac{1}{4 \cdot 2^Q} \prod_{a \in Q} \sum_{z_a} \left[ \cos(\gamma (c_1(z) + c_2(z) + c_3(z))) + \cos(\gamma (c_1(z) - c_2(z) - c_3(z))) + \cos(\gamma (-c_1(z) - c_2(z) + c_3(z))) + \cos(\gamma (-c_1(z) + c_2(z) - c_3(z))) \right]$$  \hspace{1cm} (20)$$
where
\[ c_i(z) = \sum_{a < b} z_a j_{ab}^{(i)} z_b, \quad i = 1, 2, 3. \] (21)

It is convenient for us to view \( c_1, c_2 \) and \( c_3 \) as random variables coming from an underlying distribution of \( Q \) binary variables so we write our \( X_1X_2X_3 \) expectation in (12) as
\[
\frac{1}{4} E[\cos(\gamma(c_1 + c_2 + c_3)) + \cos(\gamma(c_1 - c_2 - c_3)) + \cos(\gamma(-c_1 - c_2 + c_3)) + \cos(\gamma(-c_1 + c_2 - c_3))]. \quad (22)
\]

The full contribution to (11) from the \( X_1X_2X_3 \) terms is expression (22) times \(-\frac{1}{2} \sin(\gamma)\). We are going to pick \( \gamma \) to be negative and small. We are first going to show that for small enough \( \gamma \), expression (22) is positive and bounded away from zero so the net contribution to (11) from the \( X_1X_2X_3 \) term is positive, regardless of the sign in (8).

For any random variable \( W \),
\[
E[\cos(W)] \geq 1 - \frac{1}{2} E[W^2] \quad (23)
\]

so we have that expression (22) is greater than or equal to
\[
1 - \frac{1}{2} \gamma^2 (E[c_1^2] + E[c_2^2] + E[c_3^2]) \quad (24)
\]

Now for each \( i \),
\[
E[c_i^2] = \sum_{a < b} J_{ab}^{(i)} j_{ab}^{(i)} \leq D, \quad (25)
\]

so expression (22) is greater than or equal to
\[
1 - \frac{3}{2} D \gamma^2. \quad (26)
\]

In order to bound the \( Y_1Y_2Y_3 \) term we are going to require that \( \gamma \) go to zero faster than \( 1/D^{1/2} \). Let us write
\[
\gamma = -\frac{f}{D^r} \quad (r > 1/2) \quad (27)
\]

with \( f \) a positive constant. We have that for large \( D \), the total contribution from the \( X_1X_2X_3 \) term in (11) is greater than
\[
\frac{1}{2} \sin \left( \frac{f}{D^r} \right) \left( 1 - \frac{3f^2}{2D^{2r-1}} \right) \quad (28)
\]
which for large $D$ goes like $\frac{f}{2D}$. We will determine the required values of $f$ and $r$ by looking at the $Y_1 Y_2 Y_3$ term.

The $Y_1 Y_2 Y_3$ term in (11) varies in sign depending on the sign of the clause so we are going to want to bound its absolute value to be below what we have in (28). From (11) we see that the $Y_1 Y_2 Y_3$ term comes with a coefficient of $\pm \frac{1}{2} \cos(\gamma)$ times the expectation of the operator. For the expectation we follow the same steps that brought us to (22) but now we get

$$\frac{1}{4} \mathbb{E}[\sin(\gamma(c_1 + c_2 + c_3)) + \sin(\gamma(c_1 - c_2 - c_3)) + \sin(\gamma(-c_1 - c_2 + c_3)) + \sin(\gamma(-c_1 + c_2 - c_3))]. \quad (29)$$

Each term is of the form

$$\mathbb{E}[\sin(\gamma c)] \quad (30)$$

where $c$ is one of the four combinations of $c_1$, $c_2$ and $c_3$ appearing in (29). Now

$$\mathbb{E}[\sin(\gamma c)] = \mathbb{E}[\sin(\gamma c) - \gamma c] \quad (31)$$

because the $\mathbb{E}[c_i] = 0$ for $i = 1, 2, 3$. We have

$$|\mathbb{E}[\sin(\gamma c)]| = |\mathbb{E}[\sin(\gamma c) - \gamma c]| \leq \mathbb{E}[|\sin(\gamma c) - \gamma c|] \leq \frac{1}{6} |\gamma|^3 \mathbb{E}[|c|^3] \quad (32)$$

where the last inequality comes because $|\sin(\theta) - \theta| \leq \frac{1}{6} |\theta|^3$ for all $\theta$. Using the Cauchy-Schwarz inequality we get

$$|\mathbb{E}[\sin(\gamma c)]| \leq \frac{1}{6} |\gamma|^3 \left( \mathbb{E}[c^2] \mathbb{E}[c^4] \right)^{1/2}. \quad (33)$$

Now the $c$ appearing in (33) is one of the sum of three $c_i$'s that appears in (29). It could be that $c_1 = c_2 = c_3$ and in this case the first term in (29) has $c = 3c_1$. Now using (25) we have that $\mathbb{E}[c^2] \leq 9D$. The same bound applies to other three terms in (29). We need to bound $\mathbb{E}[c^4]$. This is largest when $c_1 = c_2 = c_3$ and $c_1$ consists of $D$ terms all with positive coefficients. We then get that $\mathbb{E}[c^4]$ is less than or equal to $81 \cdot \mathbb{E}[c_1^4]$. We have

$$\mathbb{E}[c_1^4] = \sum_{a<b} \sum_{c<d} \sum_{e<f} \sum_{g<h} \mathbb{E}[z_a J_{ab}^{(1)} z_b z_c J_{cd}^{(1)} z_d z_e J_{ef}^{(1)} z_f z_g J_{gh}^{(1)} z_h]. \quad (34)$$
Regard $J^{(1)}$ as the adjacency matrix of a graph with $T$ edges, $T \leq D$. Terms in this sum will be 0 unless $ab$, $cd$, $ef$, and $gh$ correspond to one of the right combination of edges: either all the same edge, or two pairs of edges, or edges that form a square. There are fewer than $\binom{T}{2}$ squares in a graph with $T$ edges and this leads to the bound

$$E[c_4^4] \leq 15T^2 \leq 15D^2.$$  \hfill (35)

So we have that

$$|E[\sin(\gamma c)]| < \frac{1}{6} \cdot 27 \cdot \sqrt{15} \cdot |\gamma|^3D^{5/2}$$

and the absolute value of $\text{(29)}$ has the same bound. The $Y_1Y_2Y_3$ contribution to $\text{(11)}$ also has a factor of $\frac{1}{2}\cos(\gamma)$ so the absolute value of the $Y_1Y_2Y_3$ contribution to $\text{(11)}$ is less than or equal to

$$\frac{9}{4} \sqrt{15} \cdot \cos(\gamma) \cdot |\gamma|^3D^{3/2}.$$  \hfill (37)

With $\gamma$ of the form (27) and $D$ large we have that the absolute value of the total $Y_1Y_2Y_3$ contribution is less than

$$\frac{9}{4} \sqrt{15} \cdot \frac{f^3}{D^{3/2}}$$

whereas the $X_1X_2X_3$ contribution is greater than

$$\frac{f}{2D^{r}}.$$  \hfill (39)

Now let $r = 3/4$. The value of $f$ which makes $\text{(39)}$ minus $\text{(38)}$ as big as possible is 0.1383 but we will use the simpler fraction of 3/22. We then have that the net contribution is greater than

$$\frac{1}{22D^{3/4}}.$$  \hfill (40)

for large $D$.

Running the quantum computer repeatedly produces a sample of strings for which the expected number of equations satisfied is at least $\left(\frac{1}{2} + \frac{1}{22D^{3/4}}\right) \cdot m$ where $m$ is the total number of equations in the instance. A sample of size $m \log m$ will, with probability $1 - \frac{1}{m}$, include a string that satisfies at least $\left[\left(\frac{1}{2} + \frac{1}{22D^{3/4}}\right) \cdot m - 1\right]$ equations. We have used $(D + 1)$ as a bound on the number of equations that any variable can appear in. For any instance we can set $(D + 1)$ to be the maximum number of equations that any variable appears in and we see that $m$ must exceed $D$. For large $D$ the $-1$ that just appeared in our probability statement can be ignored.
III. THE CASE OF NO SMALL LOOPS

We can get an exact formula in the case that we make the additional restriction that the hypergraph that describes the instance has no small loops. Look at a clause involving say bits 1, 2 and 3. By no small loops we mean that the other clauses that each of bits 1, 2 and 3 are in involve distinct bits. In this case bits 1, 2 and 3 are involved in at most \(3D\) other clauses which would involve \(6D\) other bits.

Now look at expression \((16)\). In this case \(C_1, C_2\) and \(C_3\) involve different qubits. \(C_1\) is a sum of at most \(D\) operators of the form \(\pm \frac{1}{2} Z_a Z_b\). Now

\[
e^{\pm i\gamma z_1 Z_a Z_b} = \cos(\gamma) \pm i \gamma \sin(\gamma) Z_a Z_b \tag{41}
\]

and the expectation in the state \(|+\rangle_a |+\rangle_b\) is \(\cos(\gamma)\). So the whole factor in equation \((16)\) is \(\cos(\gamma)\) of clauses. The full contribution to \((11)\) from the \(X_1 X_2 X_3\) term is, in the worst case,

\[-\frac{1}{2} \sin(\gamma) \cos(\gamma)^{3D}. \tag{42}\]

In fact the \(Y_1 Y_2 Y_3\) term contributes 0 so this is the whole answer. Now let

\[
\gamma = -\frac{g}{D^{1/2}}. \tag{43}
\]

For large \(D\) equation \((42)\) becomes

\[
\left(\frac{g}{2D^{1/2}}\right) \exp\left(-\frac{3}{2} g^2\right) \tag{44}
\]

which is maximized when \(g = \frac{1}{\sqrt{3}}\) and we satisfy \(\left(\frac{1}{2} + \frac{1}{2\sqrt{3}D^{1/2}}\right)\) times the number of equations. This also implies that for all instances with no small loops at least \(\left(\frac{1}{2} + \frac{1}{2\sqrt{3}D^{1/2}}\right)\) of the equations can be satisfied but to find the optimal solution you need to run the quantum computer.

IV. CONCLUSIONS

We applied the Quantum Approximate Optimization Algorithm at level \(p = 1\) with predetermined values of \(\gamma\) and \(\beta\) to the problem of Max E3LIN2 with each bit in no more than \(D + 1\) equations. We have shown that the quantum computer will output a string that satisfies \(\left(\frac{1}{2} + \frac{1}{22D^{1/4}}\right)\) of the equations. This means that all instances with bounded
occurrence $D$ have a solution that satisfies at least that many equations. The quantum computer is needed to find the solution.

The performance of the quantum algorithm can be improved. Here are a number of ways.

- We picked $\beta = \pi/4$ for ease of analysis and chose a value for $\gamma$ which sufficed for our results. Instead at $p = 1$ one could search for the optimal values of $\beta$ and $\gamma$ for each input instance. This could be done by classical preprocessing or by hunting for the best $\beta$ and $\gamma$ by making calls to the quantum computer.

- At $p = 1$ we could expand the parameter space. For example we could have a different angle $\gamma$ for each clause. This could only improve performance.

- Go to higher $p$. Perhaps at a higher value of $p$, the dependence of the approximation ratio on $D$ will be better than $constant/D^{3/4}$.

V. ACKNOWLEDGEMENTS

We thank Salil Vadhan, Madhu Sudan and Luca Trevisan for helpful correspondence. SG thanks Luca Trevisan for a stimulating coffee. We are indebted to Elizabeth Crosson for her great help in preparing the manuscript. This work was supported by the US Army Research Laboratory’s Army Research Office through grant number W911NF-12-1-0486, and the National Science Foundation through grant number CCF-121-8176.

[1] Edward Farhi, Jeffrey Goldstone, Sam Gutmann. A Quantum Approximate Optimization Algorithm, 2014. arXiv:1411.4028 [quant-ph]

[2] Johan Håstad. Some optimal inapproximability results, 2001. Journal of ACM, Vol. 48, pp 798–859.

[3] Johan Håstad. On bounded occurrence constraint satisfaction, 2000. Information Processing Letters, 74(1):1-6.

[4] Luca Trevisan. Non-approximability results for optimization problems on bounded degree instances, 2001. Proc. 30th STOC, Pages 453-461.