Correctness of an application using QAOA

Hannu Reittu and Ville Kotovirta

***Abstract*— In this work, we analyze the Quantum Approximate Optimization Algorithm (QAOA) [1] to solve a particular type of problem related to a graph clustering [2]. We establish rigorous bounds for efficiency of QAOA in the limit of a large problem.**

***Keywords*— quantum approximate optimization algorithm, quantum community detection**

I. Introduction

In [2] it was demonstrated that the so-called regularity check can be formulated as a qubo (quadratic uncostrained optimization problem) which can be solved using a quantum annealing machine like D-Wave or on a quantum gate computer using QAOA. This work concentrates on QAOA for a very simple case which can be analyzed rigorously. Some extensions of the basic model are also introduced.

II. Regularity check as a qubo

The main ingredient of the problem is a cost function obtained from the Szemerédi's Regularity Lemma [3] and its version introduced by T. Tao [4], and corresponds to the so-called ϵ-regularity check. ϵ-regularity is a property of a bipartite graph. Denote a bipartite graph as $G(V_L, V_R, E)$ in which V_L and V_R are two finite and disjoint sets of nodes and edges E are only between these sets. The density of links in G is, by definition:

$$d := \frac{e(V_L, V_R)}{|V_L||V_R|}.$$

$e(V_L, V_R) = |E|$ is the number of links in G and $|\cdot|$ is number of elements in a set. G is ϵ-regular if and only if $\forall A \subseteq V_L$ and $\forall B \subseteq V_R$ we have:

$$L(A, B) := e(A, B) - d|A||B| = O(\epsilon|V_L||V_R|. \tag{1}$$

The problem of finding extremes of L-function can be formulated as a qubo, [2]. For each node $i \in V_L \cup V_R$ assing a binary variable q_i, which is used as an indicator variable of various sets. If $A \subseteq V_L$ is defined as $A = \{i \in V_L : q_i = 1\}$ and $B \subseteq V_R$ as $B = \{i \in V_R : q_i = 1\}$, then the cost function can be written as:

$$L(A, B) = \sum_{i \in A \subseteq V_L, j \in B \subseteq V_R} (a_{ij} - d) q_i q_j. \tag{2}$$

in which a_{ij} is an element of the adjacency matrix of G, with $a_{ij} = 1$ if $(i, j) \in E$ and $a_{ij} = 0$ otherwise. Our baseline qubo is:

$$\arg\max_{q_i \in \{0,1\}} \sum_{i \in \subseteq V_L, j \in \subseteq V_R} (a_{ij} - d) q_i q_j. \tag{3}$$

The corresponding sets that maximize the cost function (2) are densely connected subgraphs in the bipartite graph and can be used to find graph communities in a graph [2].

III. QAOA for the regularity check

We adapt the well-known Quantum Approximate Optimization Algorithm (QAOA), invented by Farhi et all. in [1]. QAOA can solve the optimization problem in Eq. (3).

The input bipartite graph is denoted as $G(V_L, V_R, E)$, Both sets have $n + m$ nodes. Assume that the optimal solution of the qubo corresponds to $2n$ nodes noted by blue color in Fig. 1. We use computational basis which corresponds to the same ordering of variables with basis states $|x_1, x_2, \cdots, x_{2n+2m}\rangle$. In QAOA we need a problem Hamiltonian $\mathcal{H}$, which we define through the following spectral decomposition. For any computational basis state $|X\rangle$:

$$\mathcal{H} |X\rangle = E(X) |X\rangle, \tag{4}$$

in which

$$E(X) = \sum_{1 \leq i \in V_L, j \in V_R} (a_{i,j} - d) x_i x_j.$$

We can write:

$$\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2, \tag{5}$$
$$\mathcal{H}_1 = (1 - d) \sum_{i \in V_L, j \in V_R} a_{i,j} s_i \otimes s_j,$$
$$\mathcal{H}_2 := d \sum_{i \in V_L, j \in V_R} (a_{i,j} - 1) s_i \otimes s_j.$$

with

$$s_i := \frac{1}{2}(1 - \sigma_{z,i}),$$

in which $\sigma_{z,i}$ Pauli $z-$ operator acting on qubit i.

The main ingredient of the QAOA algorithm is the following unitary transformation:

$$U(\tau, t_1, t_2) := \left(\prod_j^{\otimes} e^{-i\tau\sigma_{x,j}} \right) e^{-it_1\mathcal{H}_1 - it_2\mathcal{H}_2} \tag{6}$$

in which $\sigma_{x,j}$ is the σ_x Pauli operator acting on the j th qubit, $t - 1, t - 2$ and τ are real parameters (in a standard QAOA there are only two parameters (τ, t)). We have the following:

Hannu Reittu, VTT Technical research Centre of Finland e-mail: hannu.reittu@vtt.fi; Ville Kotovirta, QMill, Keilaranta 12 D, 02150 Espoo, Finland, e-mail: ville.kotovirta@qmill.org. This work was partially supported by Business Finland projects COHQCA and MatterQ, and CNPq/Brazil grants 407296/2021-2 and 306049/2025-2, and HORIZON-CL4-2022-QUANTUM01-SGA Project No. 101113946 OpenSuperQ-Plus100 of the EU Flagship on Quantum Technologies.

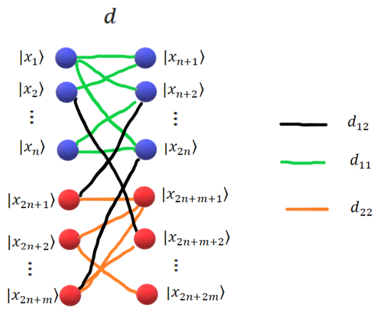

Fig. 1. Bipartite graph and assignments of qubits to nodes. Link weights between various subgraphs are shown as colour coded links. For instance link weight between blue nodes is d_{11} and so on.

Proposition 1:

$$U(\tau,t_1,t_2) = RX(2\tau)^{\otimes N}\prod_{a_{ij}=1} C_iP_j(-t_1(1-d))\times \prod_{a_{ij}=0} C_iP_j(t_2 d),$$

in which one-qubit operator RX is:

$$RX(2\tau) = \begin{pmatrix} \cos(\tau) & -i\sin(\tau) \\ -i\sin(\tau) & \cos(\tau)\end{pmatrix}$$

and where $C_iP_j(\phi)$ is controlled phase-shift operator with the control qubit i and phase-shift operator P_j acting on qubit number j as:

$$P(\phi) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi}\end{pmatrix},$$

in which ϕ is a real number. In QAOA the parameters $(\tau, t-1, t_2)$ are optimized by measuring the expected energy using some classical search method.

IV. Efficiency limit of QAOA

Consider a setting shown in Fig.(1). It is assumed that the maximum energy correspond to the configuration when the first $2n$ qubits are in state $|1\rangle$ and the rest are in the state $|0\rangle$. In Fig. 1 such a community corresponds to the blue nodes. The maximum energy has the value $E_{\max} = n^2(d_{11}-d)$, in which $d_{11} = \sum_{1\le i\le 2n} a_{ij}/n^2$ is the link density in a subgraph induced by the first $2n$ nodes (blue).

The state with maximum energy is denoted as $|\mathbf{1},\mathbf{0}\rangle$, and the state with all qubits in the state 0 as $|\mathbf{0},\mathbf{0}\rangle$.

A. A solvable case

First we consider a very simple model case, shown in Fig. (1), with $d_{11} = 1$ and all other densities are zero. Thus, the supposed solution is a full bipartite graph, K_{nn}, with the corresponding state denoted as $|\mathbf{1},\mathbf{0}\rangle$ (first $2nqubitsinstate1$. and all other in state 0) The state with all qubits in the state 0 is $|\mathbf{0},\mathbf{0}\rangle$. Define the amplitude:

$$a = \langle \mathbf{0},\mathbf{0}|\, H^{\otimes N} U(\tau,t_1,t_2)\,|\mathbf{1},\mathbf{0}\rangle\,, \tag{7}$$

then

$$p(\tau,t_1,t_2) = |a|^2$$

is the probability of finding the solution in one round of QAOA using fixed parameters (τ,t_1,t_2). Denote by $|s_1,s_2;f_1,f_2\rangle$ sum of all different computational basis states which fulfill a constrain: register has exactly $s_1+s_2+f_1+f_2$ qubits in state $|1\rangle$ and the rest are in state $|0\rangle$. And furthermore the $|1\rangle$'s are distributed in such a way that s_1 are among qubits $1,\cdots,n$, s_2 are among qubits $n+1,\cdots,2n$, f_1 are among qubits $2n+1,\cdots,2n+m$ and finally f_2 are among qubits $2n+m+1,\cdots,2n+2m$, see the numbering of qubits in Fig.1. We denote $|s_1,s_2;f_1,f_2\rangle := |\mathbf{sf}\rangle$ in which 4-dimensional vector $\mathbf{sf}$ has components: $\mathbf{sf} = (s_1,s_2,f_1,f_2)^T$. By $|\mathbf{x}\rangle, \mathbf{x}\sim\mathbf{sf}$ we denote a computational basis vector which meet the above condition with respect to a vector $\mathbf{sf}$. We denote:

$$\begin{aligned}|\phi\rangle &:= RX(-2\tau)^{\otimes N}|\mathbf{1},\mathbf{0}\rangle = (|1\rangle\cos\tau + i\sin\tau\,|0\rangle)^{2n}\times \\ &(|0\rangle\cos\tau + i\sin\tau\,|1\rangle)^{2m} = \\ &\sum_{0\le s_1,s_2\le n;0\le f_1,f_2\le m}(\cos\tau)^{2m+s_1+s_2-f_1-f_2}\times \\ &(i\sin\tau)^{2n-s_1-s_2+f_1+f_2}\,|s_1,s_2;f_1,f_2\rangle\,,\end{aligned} \tag{8}$$

For our graph we have for any computational basis vector $|\mathbf{x}\rangle$:

$$\begin{aligned}&\forall \mathbf{x}\sim(s_1,s_2;f_1,f_2) \\ &\mathcal{H}_1|\mathbf{x}\rangle = (1-d)s_1s_2\,|\mathbf{x}\rangle \\ &\mathcal{H}_2|\mathbf{x}\rangle = -d(s_1f_2+s_2f_1+f_1f_2)\,|\mathbf{x}\rangle\,.\end{aligned} \tag{9}$$

As a result we have:

$$\begin{aligned}&\forall \mathbf{x}\sim(s_1,s_2;f_1,f_2) \\ &e^{it_1\mathcal{H}_1+it_2\mathcal{H}_2|\mathbf{x}\rangle}\,|\mathbf{x}\rangle = e^{it_1(1-d)s_1s_2-it_2d(s_1f_2+s_2f_1+f_1f_2)}\,|\mathbf{x}\rangle\,.\end{aligned} \tag{10}$$

From Eqs. 8 and 10 we get

$$\begin{aligned}&e^{it_1\mathcal{H}_1+it_2\mathcal{H}_2}\,|\phi\rangle = \\ &\sum_{0\le s_1,s_2\le n;0\le f_1,f_2\le m}(\cos\tau)^{2m+s_1+s_2-f_1-f_2}\times \\ &(i\sin\tau)^{2n-s_1-s_2+f_1+f_2}e^{it_1\mathcal{H}_1+it_2\mathcal{H}_2}\,|s_1,s_2;f_1,f_2\rangle = \\ &\sum_{s_1,s_2,f_1,f_2}(\cos\tau)^{2m+s_1+s_2-f_1-f_2}(i\sin\tau)^{2n-s_1-s_2+f_1+f_2}\times \\ &e^{it_1(1-d)s_1s_2-it_2d(s_1f_2+s_2f_1+f_1f_2)}\,|s_1,s_2;f_1,f_2\rangle\,.\end{aligned} \tag{11}$$

Since:

$$H^{\otimes N}|\mathbf{0},\mathbf{0}\rangle = \frac{1}{2^{n+m}}\sum_{\mathbf{x}}|\mathbf{x}\rangle$$

and

$$\sum_{\mathbf{x}}\langle\mathbf{x}|\,|s_1,s_2;f_1,f_2\rangle = \binom{n}{s_1}\binom{n}{s_2}\binom{m}{f_1}\binom{m}{f_2},$$

we have

$$\begin{aligned}&a = \frac{1}{2^{n+m}}\times \\ &\sum_{s_1,s_2,f_1,f_2}(\cos\tau)^{2m+s_1+s_2-f_1-f_2}(i\sin\tau)^{2n-s_1-s_2+f_1+f_2}\times \\ &e^{it_1(1-d)s_1s_2-it_2d(s_1f_2+s_2f_1+f_1f_2)}\binom{n}{s_1}\binom{n}{s_2}\binom{m}{f_1}\binom{m}{f_2}.\end{aligned} \tag{12}$$

Using the Newton's binomial Theorem, Eq. 12 can be written as:

$$a = \frac{1}{2^{n+m}}\times \tag{13}$$
$$\sum_{s_2,f_2}(\cos\tau e^{it_1(1-d)s_2-it_2df_2}+i\sin\tau)^n\times$$
$$(i\sin\tau e^{-it_2d(s_2+f_2)}+\cos\tau)^m\times$$
$$(\cos\tau)^{m+s_2-f_2}(i\sin\tau)^{n-s_2+f_2}\binom{n}{s_2}\binom{m}{f_2}.$$

We make the following assignment:

$$\cos\tau = \sin\tau = \frac{1}{\sqrt{2}}, \tag{14}$$

after which Eq. 13 takes the form:

$$a = \frac{i^n}{2^{2(n+m)}}\sum_{s_2,f_2}(e^{it_1(1-d)s_2-it_2df_2}+i)^n\times \tag{15}$$
$$(ie^{-it_2d(s_2+f_2)}+1)^m(i)^{f_2-s_2}\binom{n}{s_2}\binom{m}{f_2}.$$

For large m and n the binomial coefficients have a very sharp maximums around the centre $s_2 \approx n/2$ and $f_2 \approx m/2$. That is why, the main contribution to the amplitude a comes near the center and we should use the choice of the parameters t_1 and t_2 to maximize this contribution. The corresponding contribution in Eq. 15, which depends on parameters t_1 and t_2, is maximized if the following conditions hold:

$$e^{it_1(1-d)\frac{n}{2}-it_2d\frac{m}{2}}+i = 2i \tag{16}$$
$$ie^{-it_2d\frac{n+m}{2}}+1 = 2.$$

The solution of Eq. 16 lead to the following final assignment:

$$\tau^* = \frac{\pi}{4}, \tag{17}$$
$$t_1^* = \frac{\pi}{(1-d)n}(\frac{m}{n+m}+1),$$
$$t_2^* = \frac{\pi}{d(n+m)}.$$

This means that both parameters t_1^* and t_2^* are small in Eq. 17, since we are interested in asymptotic of the amplitude a as $n, m \to \infty$. We have the following strictly positive lower bound:

Lemma 1:

$$|a| \geq \sqrt{p_0} \to \frac{2}{\pi}e^{-\frac{\pi^2}{4}(2+mn/(m+n)^2)},$$
$$n\to\infty, n/m = const.$$

Proof: See Appendix A. ■

Let $K_{s,h}$ denote a fully connected bipartite graph, in which s and h are natural numbers indicating number of nodes in the two parts. We have the following result:

Theorem 1: Let G be a bipartite graph with N nodes and with parts each having $n+m$ nodes. Only links of G are in a subgraph $K_{n,n}$. Let $|\mathbf{0}\rangle$ denote a state of N-qubit quantum register with all qubits in the state $|0\rangle$ and the first $2n$. Using operators defined in Eqs. 5 and 6 and in which the adjacency matrix and the density parameter d corresponds to the graph G a quantum state $|\psi(\tau,t_1,t_2\rangle$ is prepared:

$$|\psi(\tau,t_1,t_2)\rangle := U(\tau,t_1,t_2)H^{\otimes N}|\mathbf{0}\rangle,$$

in which H is the Hadamar transformation. With the assignment $\tau=\tau^*, t1 = t1^*, t2 = t2^*$ according to Eq. 17, the measurement of all qubits of $|\psi(\tau^*,t_1^*,t_2^*)\rangle$ will yield subgraph $K_{n,n}$,in the sense that measured state will have value 1 in all positions in the quantum register assigned to the subgraph $K_{n,n}$ and 0's in all other positions, with probability at least p_0 and with

$$p_0 \to \frac{4}{\pi^2}e^{-\{\frac{\pi^2}{2}[2+mn/(m+n)^2]\}},$$
$$n\to\infty, n/m = const.$$

The qubits assigned to the subgraph $K_{n,n}$ are, according to definition in Eq. 5, all such positions i and j in the quantum register for which adjacency matrix of G has entry $a_{i,j}=1$, corresponding to the links in G.

Proof: Obviously, probability of finding K_{nn} when measuring the state $|\psi(\tau^*,t_1^*,t_2^*)\rangle$ is equal to $|a|^2$ in Eq.(7) and the claim follows from the Lemma 1. ■

B. Extension to a more generic case

Assume the same model as in previous section, described by the adjacency matrix A with elements a_{ij}. Next we allow a perturbation of the simple model by cutting some links and adding some links. This alternations are described by a sparse matrix with elements $\epsilon_{ij}\in\{0,1\}$ of the same size as A. By sparsity we mean that $\sum_{ij}\epsilon_{ij}/n \to 0, n\to\infty$. The value $\epsilon_{ij}=1$ means that the corresponding value of adjacency matrix element a_{ij} is reversed and if $\epsilon ij = 0$ the element is left unchanged. Obviously, the altered graph has adjacency matrix with elements:

$$a'_{ij} = a_{ij}(1-\epsilon_{ij})+(1-a_{ij})\epsilon_{ij} = a_{ij}+\epsilon_{ij}(1-2a_{ij}) \tag{18}$$

Using the adjacency matrix A' with elements a'_{ij}, we can rewrite Eq.5 as

$$\mathcal{H}' = \mathcal{H}'_1+\mathcal{H}'_2, \tag{19}$$
$$\mathcal{H}'_1 = (1-d')\sum_{i\in V_L, j\in V_R}(a_{i,j}+\epsilon_{ij}(1-2a_{ij}))s_i\otimes s_j,$$
$$\mathcal{H}_2 := d'\sum_{i\in V_L, j\in V_R}(a_{i,j}-1+\epsilon_{ij}(1-2a_{ij}))s_i\otimes s_j,$$

in which d' is the link density of the graph with adjacency matrix A'. Obviously $d'\to d, n\to\infty$.

$$\forall\mathbf{x}\sim(s_1,s_2;f_1,f_2) \tag{20}$$
$$e^{it_1\mathcal{H}'_1+it_2\mathcal{H}'_2|\mathbf{x}\rangle}|\mathbf{x}\rangle =$$
$$e^{it_1(1-d')s_1s_2-it_2d'(s_1f_2+s_2f_1+f_1f_2+it_1m_1(\mathbf{x})+it_2m_2(\mathbf{x})}|\mathbf{x}\rangle,$$

in which

$$m_1(\mathbf{x}) = (1-d')\sum_{x_i=1, j\in[2(n+m)]}\epsilon_{ij}(1-2a_{ij}) \tag{21}$$
$$m_2(\mathbf{x}) = d'\sum_{x_i=1, j\in[2(n+m)]}\epsilon_{ij}(1-2a_{ij}),$$
$$[k] := \{1,2,\cdots,k\}, k\in\mathbb{N} := \{1,2,\cdots\}.$$

Using Eq. 17 we have

$$\forall \mathbf{x} \sim (s_1, s_2; f_1, f_2) \tag{22}$$
$$e^{it_1^* \mathcal{H}_1' + it_2^* \mathcal{H}_2' |\mathbf{x}\rangle} |\mathbf{x}\rangle \to$$
$$e^{it_1^*(1-d)s_1 s_2 - it_2^* d(s_1 f_2 + s_2 f_1 + f_1 f_2)}(1 + o(1)) |\mathbf{x}\rangle .$$

To prove Eq.22, note that both t_1^* and t_2^* are of the order of $1/n$ and since $|m_1(\mathbf{x})|/n \leq (1-d')\sum_{ij}\epsilon_{ij}/n \to 0$ and $|m_2(\mathbf{x})|/n \leq d'\sum_{ij}\epsilon_{ij}/n \to 0$, we have $t_1^* m_1(\mathbf{x}) \to 0$ and $t_2^* m_2(\mathbf{x}) \to 0$. As a result we have in Eq.20:

$$e^{i\phi(\mathbf{x}) + it_1 m_1(\mathbf{x}) + it_2 m_2(\mathbf{x})} = e^{i\phi(\mathbf{x})}(1 + o_x(1)), n \to \infty,$$

in which we denoted by $o_x(1) = \sum_{k\geq 1}(it_1 m_1(\mathbf{x}) + it_2 m_2(\mathbf{x}))^n/n! \to 0$ and thus $o_x(1) = o(1)$, and $\phi(\mathbf{x}) = t_1^*(1-d')s_1 s_2 - t_2^* d'(s_1 f_2 + s_2 f_1 + f_1 f_2)$. As a result, we get

$$\langle \mathbf{0}, \mathbf{0}| H^{\otimes N} e^{it_1^* \mathcal{H}_1' + it_2^* \mathcal{H}_2'} \left(\prod_j^{\otimes} e^{i\tau^* \sigma_{x,j}} \right) |\mathbf{1}, \mathbf{0}\rangle \to$$
$$a + \frac{i^n}{2^{2(n+m)}} \sum_{\mathbf{x}} e^{i\phi(\mathbf{x})} o_x(1),$$

in which a is the amplitude with absolute value bounded away from zero according to Eq.1. The second term has zero limit because

$$|\frac{i^n}{2^{2(n+m)}} \sum_{\mathbf{x}} e^{i\phi(\mathbf{x})} o_x(1)| \leq \frac{1}{2^{2(n+m)}} \sum_{\mathbf{x}} |o_x(1)| \leq$$
$$\frac{|o_x^*(1)|}{2^{2(n+m)}} \sum_{\mathbf{x}} 1 = |o^*(1)| = o(1),$$

in which $|o^*(1)|$ is largest among $|o_x(1)|$.

C. A solvable weighted graph problem

Now consider a model which follows from a random bipartite graph with three parameters in the sense that number of links is taken from the average number of links in the corresponding random graph. This is similar to the case in Fig. 1. We assume that for any configuration (s_1, s_2, f_1, f_2) number of links is replaced by expectation $d_{11}s_1 s_2 + d_{12}(s_1 f_2 + f_1 s_2) + d_{22} f_1 f_2$ and number of no-links is replaced by $(1-d_{11})s_1 s_2 + (1-d_{12})(s_1 f_2 + f_1 s_2) + (1-d_{22}) f_1 f_2$.

The corresponding problem Hamiltonian is:

$$\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2, \tag{23}$$
$$V_L = \{1, 2, \cdots, n, 2n+1, 2n+2, \cdots, 2n+m\},$$
$$V_R = \{n+1, n+2, \cdots, 2n,$$
$$2n+m+1, 2n+m+2, \cdots, 2n+2m\}$$
$$\mathcal{H}_1 := (1-d) \sum_{i \in V_L, j \in V_R} D_{ij} s_i \otimes s_j,$$
$$\mathcal{H}_2 := d \sum_{i \in V_L, j \in V_R} (D_{ij} - 1) s_i \otimes s_j$$
$$d = \frac{\sum D_{ij}}{(n+m)^2}$$

and in which matrix elements D_{ij} can have three possible values, d_{11}, d_{12} and d_{22}. Those values are determined according to which part of the graph the corresponding nodes belong. See the corresponding assignments in Fig. 1. For instance if $i \leq n$ and $n+1 \leq j \leq 2n$, then $D_{ij} = d_{11}$ and so on.

Formally we have:

Definition 1:

$$G(V_{L,i}, V_{R,j}, d_{ij}, 1 \leq i, j \leq 2)$$

is a weighted bipartite graph with bipartition of nodes in two sets $V_{L,1} \cup V_{L,2}$ and $V_{R,1} \cup V_{R,2}$ and with weight matrix D between the nodes in two parts: for all nodes $\alpha \in V_{L,i}$, and $\beta \in V_{R,j}, D_{\alpha\beta} = d_{i,j}$ and in which $d_{ij} = d_{ji}$ are three real numbers.

$$\forall \mathbf{x} \sim (s_1, s_2; f_1, f_2) \tag{24}$$
$$\mathcal{H}_1 |\mathbf{x}\rangle = (1-d)\{d_{11}s_1 s_2 + d_{12}(s_1 f_2 + s_2 f_1) +$$
$$d_{22} f_1 f_2\} |\mathbf{x}\rangle$$
$$\mathcal{H}_2 |\mathbf{x}\rangle = d\{(d_{11}-1)s_1 s_2 + (d_{12}-1)(s_1 f_2 + s_2 f_1) +$$
$$(d_{22}-1) f_1 f_2\} |\mathbf{x}\rangle .$$

Similarly to the analysis of the previous section we can prove the following:

Theorem 2: Assume a weighted bipartite graph:

$$G(V_{L,i}, V_{R,j}, d_{ij}, 1 \leq i, j \leq 2),$$

according to the Definition 1, and a state

$$|\psi(\tau, t_1, t_2)\rangle := U(\tau, t_1, t_2) H^{\otimes N} |\mathbf{0}\rangle .$$

With

$$B = \tag{25}$$
$$\begin{pmatrix} (1-d)(nd_{11} + md_{12}), & -d(n(1-d_{11}) + m(1-d_{12})) \\ (1-d)(nd_{12} + md_{22}), & -d(n(1-d_{12}) + m(1-d_{22})). \end{pmatrix}$$

provided the determinant $\det B \neq 0$, set τ^*, t_1^* and t_2^*

$$t_1^* = \frac{\det \begin{pmatrix} \pi, & -d(n(1-d_{11}) + m(1-d_{12})) \\ -\pi, & -d(n(1-d_{12}) + m(1-d_{22})) \end{pmatrix}}{\det B}, \tag{26}$$
$$t_2^* = \frac{\det \begin{pmatrix} (1-d)(nd_{11} + md_{12}), & \pi \\ (1-d)(nd_{12} + md_{22}), & -\pi \end{pmatrix}}{\det B},$$
$$\tau^* = \frac{\pi}{4}$$

the solution can be found with a probability at least:

$$p_0 \to \frac{4}{\pi^2} e^{-\{\frac{\pi^2}{2}[2 + mn/(m+n)^2]\}},$$
$$n \to \infty, n/m = const$$
$$n = |V_{L,1}| = |V_{R,1}| \quad m = |V_{L,2}| = |V_{R,2}|.$$

V. Conclusions

We considered a problem of finding range of a so called regularity check qubo. This problem is known to be in the classical complexity class co-NP-complete. We formulate this problem as QAOA program. We analyzed rigorously some particular instances which can be effectively solved with one-layer QAOA in the limit of large graphs. The solution is based on Newton's binomial theorem which allows a constructive interference of amplitudes when there is enough smoothness in the cost function. It was shown that in these cases QAOA can find solution with a probability lower bounded by a positive constant. Regularity check problem has application to a graph clustering. In future we hope to extend similar analysis to more complicated cases using multi-level QAOA.

Appendix

Proof of the Lemma 1

Proof: Representing variables as $s_2 = n/2 + k,\ f_2 = m/2 + j$ and using the assignment Eq.17 we get Eq. 15 in the form:

$$a = \frac{i^{2n+\frac{m-n}{2}}}{2^{2(n+m)}} \sum_{k,j} (e^{i\frac{\pi k}{n}(\frac{m}{n+m}+1)-i\frac{\pi j}{n+m}} + 1)^n \times \quad (27)$$

$$(e^{-i\frac{\pi}{n+m}(k+j)} + 1)^m (i)^{j-k} \binom{n}{\frac{n}{2}+k}\binom{m}{\frac{m}{2}+j},$$

assuming for simplicity of notation that both n and m are even. We will find the leading contribution to a in the limit $n, m \to \infty$ and $n/m = const$. For this we use expansions in powers of small parameter $1/n$:

$$(e^{i\frac{\pi k}{n}(\frac{m}{n+m}+1)-i\frac{\pi j}{n+m}} + 1)^n =$$

$$\Big(2 + \frac{i\pi}{n}\big(k(\frac{m}{n+m}+1) - \frac{j}{1+m/n}t\big) -$$

$$\frac{\pi^2}{2n^2}\big(k(\frac{m}{n+m}+1) - \frac{j}{1+m/n}\big)^2 + \cdots\Big)^n =$$

$$2^n \exp\Big(\frac{i\pi}{2}\big(k(\frac{m}{n+m}+1) - \frac{j}{1+m/n}\big)$$

$$-\frac{\pi^2}{4n}\big(k(\frac{m}{n+m}+1) - \frac{j}{1+m/n}\big)^2 + \cdots\Big)$$

and

$$\big(e^{-i\frac{\pi}{n+m}(k+j)} + 1\big)^m =$$

$$\big(2 - i\frac{\pi(k+j)}{n+m} - \frac{(\pi(k+j))^2}{2(n+m)^2} + \cdots\big)^m =$$

$$2^m \exp\big(-\frac{i\pi}{2}\frac{(k+j)}{n/m+1} - \frac{\pi^2(k+j)^2}{4(n+m)(n/m+1)} + \cdots\big).$$

By substituting these expansions to Eq. 15, and ignoring the insignificant constant phase factor, we get:

$$a = \frac{1}{2^{n+m}} \sum_{j,k} e^{\{-\frac{\pi^2}{4n}[(k(\frac{m}{n+m}+1)-\frac{j}{1+m/n})]^2\}} \times \quad (28)$$

$$e^{\{-\frac{\pi^2(k+j)^2}{4(n+m)(n/m+1)}+\cdots\}} \times$$

$$\binom{n}{\frac{n}{2}+k}\binom{m}{\frac{m}{2}+j},$$

in which by $\cdots$ we denote the further terms in expansion on the small parameter $1/n$. The first order term has opposite phase with the factor i^{j-k} and, as a result, their product equals to 1. Indeed:

$$\frac{i\pi}{2}\big(k(\frac{m}{n+m}+1) - \frac{j}{1+m/n}\big) - \frac{i\pi}{2}\frac{(k+j)}{n/m+1} =$$

$$\frac{i\pi}{2}\big(k + k\frac{m-m}{n+m} - j\frac{n+m}{n+m}\big) = \frac{i\pi}{2}\big(k-j\big).$$

First we split the sum in Eq. 28 in two parts. The first part has sum over values: $|j|, |k| \le n^\alpha$, in which $\frac{1}{2} \le \alpha < \frac{2}{3}$. The corresponding contribution to a is denoted a_0. The reminding contribution to a is denoted as a_1 and by definition $a = a_0 + a_1$. We have in the limit:

$$a_0 = \frac{1}{2^{n+m}} \sum_{j,k:|j|,|k|\le n^\alpha} e^{\{-\frac{\pi^2}{4n}[(k(\frac{m}{n+m}+1)-\frac{j}{1+m/n})^2]\}} \times \quad (29)$$

$$e^{\{-\frac{\pi^2}{4n}\frac{(k+j)^2}{(1+m/n)(n/m+1)}]\}} \times$$

$$\binom{n}{\frac{n}{2}+k}\binom{m}{\frac{m}{2}+j},$$

indeed, the next term of the power expansion in Eq.28 has the order of $n^{3\alpha}/n^2 \to 0$ as $n \to \infty$. The binomial coefficients are monotonously decreasing functions of k and j as their absolute values increase. In area of summation for a_0 we can use the asymptotic:

$$\binom{n}{\frac{n}{2}+k} \sim \frac{2^n}{\sqrt{n\pi/2}} e^{-\frac{2k^2}{n}}, \quad n \to \infty, k = o(n^{2/3}). \quad (30)$$

From Eqs. 29 and 30 we have:

$$a_0 \ge \frac{e^{-\{\frac{\pi^2}{4}[(m/n+2)^2+m/n]\}}}{2^{n+m}} \times \quad (31)$$

$$\sum_{j,k:|j|,|k|\le\sqrt{n}} \binom{n}{\frac{n}{2}+k}\binom{m}{\frac{m}{2}+j} \sim$$

$$\frac{2e^{-\{\frac{\pi^2}{4}[(m/n+2)^2+m/n]\}}}{\sqrt{nm}\pi} \sum_{j,k:|j|,|k|\le\sqrt{n}} e^{-\frac{2k^2}{n}-\frac{2j^2}{m}} \ge$$

$$\frac{2e^{-\{\frac{\pi^2}{4}[(m/n+2)^2+m/n]+2(n/m+1)\}}}{\sqrt{nm}\pi} \sum_{j,k:|j|,|k|\le\sqrt{n}} 1 \ge$$

$$\frac{2}{\pi} e^{-\{\frac{\pi^2}{4}[(m/n+2)^2+m/n]+2(n/m+1)\}}.$$

As a result of Eq. 31, amplitude of a_0 is lower-bounded by a constant.

Finally, we show that amplitude of a_1 is exponentially small. From Eqs. 15 and 30 we get an asymptotic upper bound

$$|a_1| \le \frac{1}{2^{n+m}}\{2^m m \sum_{|k|\ge n^\alpha} \binom{n}{\frac{n}{2}+k} + \quad (32)$$

$$2^n n \sum_{|j|\ge n^\alpha} \binom{m}{\frac{m}{2}+j} + \sum_{|j|,|k|\ge n^\alpha} \binom{n}{\frac{n}{2}+k}\binom{m}{\frac{m}{2}+j}\} \le$$

$$me^{-2n^{2\alpha-1}} + ne^{-2\frac{n}{m}n^{2\alpha-1}} + nme^{-2n^{2\alpha-1}}e^{-2\frac{n}{m}n^{2\alpha-1}},$$

since $\alpha > \frac{1}{2}$ we have $2\alpha - 1 > 0$, which indicates that $|a_1| \to 0$ with exponential speed $\sim e^{-n^\beta}$, $\beta > 0$ and $n \to \infty$. ■

References

[1] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm, 2014.

[2] Hannu Reittu, Ville Kotovirta, Lasse Leskelä, Hannu Rummukainen, and Tomi Räty. Towards analyzing large graphs with quantum annealing. In *2019 IEEE International Conference on Big Data (Big Data)*, pages 2457–2464, 2019.

[3] E. Szemerédi. Regular partitions of graphs. *Problemes combinatoires et theorie des graphes (Colloq. Internat. CNRS, **260**, Univ. Orsay), pp. 399–401*, 1978.

[4] T. Tao. Szemerédi's regularity lemma revisited. *Contributions to Discrete Mathematics*, 1(1):8–28, 2006.