

This is the accepted manuscript made available via CHORUS. The article has been published as:

# Quantum approximate optimization algorithm for MaxCut: A fermionic view

Zhihui Wang, Stuart Hadfield, Zhang Jiang, and Eleanor G. Rieffel

Phys. Rev. A **97**, 022304 — Published 5 February 2018

DOI: [10.1103/PhysRevA.97.022304](https://doi.org/10.1103/PhysRevA.97.022304)

# The Quantum Approximate Optimization Algorithm for MaxCut: A Fermionic View

Zhihui Wang,<sup>1,2</sup> Stuart Hadfield,<sup>3</sup> Zhang Jiang,<sup>1,4</sup> and Eleanor G. Rieffel<sup>1</sup>

<sup>1</sup> *Quantum Artificial Intelligence Laboratory (QuAIL),  
NASA Ames Research Center, California 94035*

<sup>2</sup> *Universities Space Research Association, Columbia 21046*

<sup>3</sup> *Department of Computer Science, Columbia University, New York 10027*

<sup>4</sup> *Stinger Ghaffarian Technologies Inc., Maryland 20770*

(Dated: January 11, 2018)

Farhi *et al.* recently proposed a class of quantum algorithms, the Quantum Approximate Optimization Algorithm (QAOA), for approximately solving combinatorial optimization problems. A level- $p$  QAOA circuit consists of  $p$  steps; in each step a classical Hamiltonian, derived from the cost function, is applied followed by a mixing Hamiltonian. The  $2p$  times for which these two Hamiltonians are applied are the parameters of the algorithm, which are to be optimized classically for the best performance. As  $p$  increases, parameter optimization becomes inefficient due to the curse of dimensionality. The success of the QAOA approach will depend, in part, on finding effective parameter-setting strategies. Here, we analytically and numerically study parameter setting for QAOA applied to MaxCut. For level-1 QAOA, we derive an analytical expression for a general graph. In principle, expressions for higher  $p$  could be derived, but the number of terms quickly becomes prohibitive. For a special case of MaxCut, the ring of disagrees, or the 1D antiferromagnetic ring, we provide an analysis for arbitrarily high level. Using a fermionic representation, the evolution of the system under QAOA translates into quantum control of an ensemble of independent spins. This treatment enables us to obtain analytical expressions for the performance of QAOA for any  $p$ . It also greatly simplifies numerical search for the optimal values of the parameters. By exploring symmetries, we identify a lower-dimensional sub-manifold of interest; the search effort can be accordingly reduced. This analysis also explains an observed symmetry in the optimal parameter values. Further, we numerically investigate the parameter landscape and show that it is a simple one in the sense of having no local optima.

PACS numbers:

## I. INTRODUCTION

Recently, Farhi *et al.* [1] proposed a new class of quantum algorithm, the Quantum Approximate Optimization Algorithm (QAOA), to tackle challenging approximate optimization problems on a gate model quantum computer. In QAOA, the problem Hamiltonian, which encodes the cost function of the optimization problem, and a mixing Hamiltonian are applied alternately. A handful of recent papers suggest the power of such circuits [2–5]. Once the problem and mixing Hamiltonians have been chosen, the parameters of the algorithm are the times for which each Hamiltonian is applied at each stage. With an optimized time sequence for each piece, the optimal output of the problem Hamiltonian is approximated.

The success of QAOA relies on being able to find a good time sequence. A level- $p$  algorithm has  $2p$  parameters, the times (angles) for which the problem Hamiltonian and the mixing Hamiltonian are applied at each iteration. For QAOA of a fixed level, straight-forward sampling of search space was proposed [1], but it is practical only for small  $p$ ; as the level increases the parameter optimization becomes inefficient due to the curse of dimensionality. Elegant analytical tools designed for specific problem class can provide parameter values for  $p \gg 1$  that give near optimal performance, e.g., search an unstructured database [5], but for a general problem, practical search strategies are needed. Here, we analytically

and numerically study the parameter setting problem, with a focus on the MaxCut problem. We demonstrate how analyzing parameter symmetries and the landscape of the expectation value over the space of the parameter values can aid in finding optimal parameter values.

In Ref. [1], Farhi *et al.* investigated MaxCut for specific (bounded-degree) graphs, and provided numerical results for a special case, termed *ring of disagrees*, which is a one-dimensional chain of spin-1/2's with nearest-neighbored antiferromagnetic couplings. We first extend the results of MaxCut to derive analytical expressions which can be solved to obtain the optimal parameters for level-one QAOA for MaxCut on arbitrary graphs. Direct analysis through operator reduction quickly becomes cumbersome as the level  $p$  of the algorithm increases. We focus on the ring of disagrees where we are able to advance the analysis to arbitrary levels.

Using a Fermionic representation, we show that the evolution of the system under QAOA translates into quantum optimal control of an ensemble of independent spins, significantly simplifying the analysis. In the new representation, the analytical expression for the expectation value as a trigonometric polynomial of the parameters can be efficiently derived for arbitrary level  $p$ . Furthermore, the reduction to independent spins simplifies the numerical search greatly because evaluation involves only  $\sim 2p$  matrix multiplications of 2-by-2 matrices and is linear in problem size, the number of spins in the orig-

inal problem,  $n$ . Further, by exploring symmetries, we identify a lower-dimensional sub-manifold whose critical points are also critical points in the full manifold. We numerically confirm for small  $p$  that all optimal parameters lie in this sub-manifold. The search effort can be accordingly reduced, and it also explains an observed symmetry in the optimal parameter values. Finally, a numerical investigation of the parameter landscape shows that it is a simple one in the sense of having only global optima.

In Sec. II we give a recap of the QAOA algorithm, and a literature review. In Sec. III we introduce QAOA for MaxCut and present an analytical expression for level-1 QAOA. From Sec. IV on, we focus on the anti-ferromagnetic chain (ring of disagrees). Sec. IV A reviews the formulation of QAOA on this problem. In Sec. IV B, we transform to a fermionic representation and reduce the problem to the control of non-interacting spins. Sec. IV C provides analysis on the size-dependence of the approximation ratio. We analyze the symmetry in the system in Sec. IV D, and identify criticality-constrained manifolds where the global optima of the parameters live. In Sec. IV E, QAOA of level-1 and level-2 are illustrated; In Sec. IV F, we discuss the landscape topography of the search space of the parameter values and its relation to known theory in quantum control. Sec. V summarizes the main results and conclusions of the paper.

## II. RECAP OF THE ALGORITHM

Given an objective function  $C : \{0, 1\}^n \rightarrow \mathbf{R}$  to maximize, the aim of an approximation algorithm is to find, upon specification of a desired approximation ratio  $r^*$ , a bit string  $\mathbf{x}$  such that  $C(\mathbf{x})$  is within a factor of  $r^*$  of the maximum:

$$\frac{C(\mathbf{x})}{C_{\max}} \geq r^*. \quad (1)$$

An algorithm is an  $r^*$ -approximation algorithm for problem, if for every instance of the problem, the algorithm finds a bit string with cost function within  $r^*$  of the maximum. QAOA is a quantum approximate optimization algorithm that iteratively alternates between applying a problem Hamiltonian  $H_C$  derived from the cost function and applying a mixing Hamiltonian  $H_B$ , which in the standard case is the transverse field  $H_B = \sum_j \sigma_j^x$ . For many problems, alternative mixing Hamiltonians that incorporate some problem constraints can reduce resource requirements and improve performance over the standard setup [6].

From a classical cost function that is a polynomial in binary variables  $x_1, \dots, x_n$ , we can construct a Hamiltonian  $H_C$  on  $n$  qubits by first rewriting the cost function in terms of variables  $z_i \in \{-1, 1\}$  where  $x_i = (1 - z_i)/2$  to obtain a polynomial  $f(\mathbf{z}) = \sum_{C \subset \{1, \dots, n\}} \alpha_C \prod_{j \in C} z_j$  and then replacing each occurrence of  $z_i$  with the Pauli operator  $\sigma_i^z$ . Thus,  $H_C$  is diagonal in the  $\sigma^z$ -basis and

takes the form

$$H_C = \sum_{C \subset \{1, \dots, n\}} \alpha_C \bigotimes_{j \in C} \sigma_j^z, \quad (2)$$

where  $C$  is a subset of all qubits, and  $\alpha_C$  is a real coefficient for the many-body coupling between qubits in the subset  $C$ .

We will use  $\text{QAOA}_p$  to refer to a level- $p$  QAOA circuit consisting of  $p$  steps. In each step, we first apply the problem Hamiltonian  $H_C$ , and then a mixing Hamiltonian  $H_B$ . Once the mixing Hamiltonian and the problem Hamiltonian have been chosen, the parameters of a  $\text{QAOA}_p$  circuit are the  $2p$  real numbers  $(\gamma_i, \beta_i)$ , for  $1 \leq i \leq p$ , which determine how long each operator is applied in iteration  $i$ :

$$U_C(\gamma_i) = \exp[-i\gamma_i H_C] \quad (3)$$

$$U_B(\beta_i) = \exp[-i\beta_i H_B]. \quad (4)$$

Following Farhi *et al.* [2], we refer to these times as angles. The standard initial state  $|\psi_0\rangle$ , a superposition of all classical bit strings, is prepared as the ground state of  $-H_B$  with density matrix

$$\rho_0 = |\psi_0\rangle\langle\psi_0| = \bigotimes_j \frac{1}{2}(\mathbb{1} + \sigma_j^x). \quad (5)$$

The circuit

$$U = U_B(\beta_p)U_C(\gamma_p) \cdots U_B(\beta_2)U_C(\gamma_2)U_B(\beta_1)U_C(\gamma_1) \quad (6)$$

applied to the initial state creates a final state

$$|\gamma, \beta\rangle = U|\psi_0\rangle, \quad (7)$$

for which the expectation value of  $H_C$  is

$$F(\gamma, \beta) = \text{Tr}[H_C U \rho_0 U^\dagger]. \quad (8)$$

Let  $F^* = F(\gamma^*, \beta^*)$  be the optimal value of  $F$  over the value range of the parameter set  $\{(\gamma, \beta)\}$ . The approximation ratio for the QAOA circuit with parameters  $\{(\gamma, \beta)\}$  is

$$r \equiv F/C_{\max}. \quad (9)$$

The goal of the circuit  $U$  is to drive the system into a quantum state which, upon measuring in the computational basis, yields with high probability a classical bit string that is  $r^*$ -approximately optimal. This goal is achieved if the expectation value  $F^*$  in the final state is  $r^*$ -approximately optimal, i.e.,  $r \geq r^*$ , and the distribution of bit strings from measuring this state in the computational basis is concentrated on bit strings with costs close to this expectation value.

In Farhi *et al.* [2], a  $\text{QAOA}_1$  algorithm beat the existing best approximation bound for efficient classical algorithms for the problem E3Lin2, only to inspire a better

classical algorithm [7] that beats the approximation ratio for the QAOA<sub>1</sub> algorithm by a log factor. The performance of QAOA<sub>p</sub> for E3Lin2 with  $p > 1$  has yet to be determined. Circuits with the above alternating structure have been used for purposes other than approximate optimization. For example, such QAOA circuits have also been applied for exact optimization [5, 8] and sampling [3]. Wecker *et al.* [8] explores learning parameters for QAOA circuits on instances of MAX-2-SAT that result in high overlap with the optimal state. Jiang *et al.* [5] demonstrates that the class of QAOA circuits is powerful enough to obtain the  $\Theta(\sqrt{2^n})$  query complexity on Grover's problem, and also provides the first algorithm within the QAOA framework to show a quantum advantage for a number of iterations  $p$  in the intermediate range between  $p = 1$  and  $p \rightarrow \infty$ . Farhi and Harrow [3] proved that, under reasonable complexity assumptions, the output distribution of even QAOA<sub>1</sub> circuits cannot be efficiently sampled classically. QAOA circuits are therefore among the most promising candidates for early demonstrations of "quantum supremacy" [9, 10]. It remains an open question whether QAOA circuits provide a quantum advantage for approximate optimization.

QAOA has close connection with the Variational Quantum Algorithm (VQA), in which classical optimization of parameters for a quantum evolution is performed. The result of evaluation of the final state is fed back to the parameter optimization, forming a closed-loop learning process. Yang *et al.* [4] proved that for evolution under a Hamiltonian that is the weighted sum of Hamiltonian terms, with the weights allowed to vary in time, the optimal control is bang-bang, i.e. constant magnitude, of either the maximum or minimum allowed weight, for each of the terms in the Hamiltonian at any given time. Their work implies that QAOA circuits with the right parameters are optimal among Hamiltonians of the form  $H(s) = (1 - f(s))H_B + f(s)H_C$ , where  $f(s)$  is a real function in the range  $[0, 1]$ .

The ultimate success of the QAOA approach will depend on finding effective parameter-setting strategies. For fixed  $p$ , the optimal parameters can be computed in time polynomial in the number of qubits  $n$  [1]. With increasing  $p$ , however, exhaustive search of the QAOA parameters becomes inefficient due to the curse of dimensionality. If we discretize so that each parameter can take on  $m$  values, exhaustive search of the optimum takes exponential steps in  $p$  as  $m^{2p}$ . Here, we analytically and numerically study parameter setting for QAOA applied to MaxCut.

### III. QAOA<sub>1</sub> FOR MAXCUT

In this section, we derive an analytical expression for the expectation value  $F$  for QAOA<sub>1</sub> for MaxCut on general graphs, furthering the analysis in [1]. In principle, we could similarly derive expressions for higher  $p$ , but the workload quickly becomes prohibitive.

**MaxCut Problem:** Given a graph  $G = (V, E)$  with  $n = |V|$  vertices and  $|E|$  edges, the objective is to partition the graph vertices into two sets such that the number of edges connecting vertices in different sets is maximized.

The cost function for MaxCut is

$$C = \frac{1}{2} \sum_{(i,j) \in E} (1 - z_i z_j) \quad (10)$$

where  $z_i$  and  $z_j$  are binary variables associated to the vertices in  $V$  which assume value  $+1$  or  $-1$  depending on which of the two partitions defined by the cut are assigned. The Hamiltonian corresponding to this cost function is

$$H_C = \sum_{\langle uv \rangle \in E} C_{uv}, \quad C_{uv} = \frac{1}{2}(I - \sigma_u^z \sigma_v^z). \quad (11)$$

The expectation value of  $H_C$  in QAOA decomposes as

$$F(\gamma, \beta) = \sum_{\langle uv \rangle \in E} \langle C_{uv} \rangle \quad (12)$$

where  $\langle C_{uv} \rangle := \text{Tr}[C_{uv} U \rho_0 U^\dagger]$ . As  $\langle C_{uv} \rangle \leq 1$ , the expected approximation ratio is lower bounded as

$$r \geq \frac{F(\gamma, \beta)}{|E|}. \quad (13)$$

**Theorem 1.** For QAOA with  $p = 1$ , for each edge  $\langle uv \rangle$ ,

$$\begin{aligned} \langle C_{uv} \rangle &= \frac{1}{2} + \frac{1}{4}(\sin 4\beta \sin \gamma)(\cos^{d_u} \gamma + \cos^{d_v} \gamma) \\ &\quad - \frac{1}{4}(\sin^2 \beta \cos^{d_u+d_v-2\lambda_{uv}} \gamma)(1 - \cos^{\lambda_{uv}} 2\gamma), \end{aligned} \quad (14)$$

where  $d_u + 1$  and  $d_v + 1$  are the degrees of vertices  $u$  and  $v$ , respectively, and  $\lambda_{uv}$  is the number of triangles in the graph containing edge  $\langle uv \rangle$ .

See Appendix A for a proof. The theorem implies that for  $p = 1$  the expectation value of any edge  $\langle C_{uv} \rangle$  depends only on the parameters  $(d_u, d_v, \lambda_{uv})$ . Then, the overall expectation value is  $F(\gamma, \beta) = \sum_{(d_1, d_2, \lambda)} \langle C_{uv} \rangle \chi(d_1, d_2, \lambda)$ , where the summation is taken over distinct subgraphs  $(d_1, d_2, \lambda)$  and  $\chi$  is the multiplicity of the subgraph, i.e. the number of times the subgraph appears in  $G$ . Thus, for an arbitrary graph the expectation value  $F(\gamma, \beta)$  may be efficiently computed classically, while to find an actual bit string realizing an approximate solution, quantum computation resulting in the quantum state  $U|\psi_0\rangle$  followed by measurement is required.

**Corollary 1.** For a triangle-free  $(d + 1)$ -regular graph, the expectation value of QAOA<sub>1</sub> is

$$F(\gamma, \beta) = \frac{|E|}{2} (1 + \sin 4\beta \sin \gamma \cos^d \gamma) \quad (15)$$

with maximum

$$F^* = \frac{|E|}{2} \left( 1 + \frac{1}{\sqrt{d+1}} \left( \frac{d}{d+1} \right)^{\frac{d}{2}} \right) =: C_{\max}^{\text{reg}}(d). \quad (16)$$

For any such graph, one optimal pair of angles is  $(\gamma, \beta) = (\arctan(1/\sqrt{d}), \pi/8)$ .

Notice that  $\left(\frac{d}{d+1}\right)^{\frac{d}{2}} > \frac{1}{e}$ , the optimal approximation ratio is lower-bounded as

$$r > \frac{1}{2} \left( 1 + \frac{1}{\sqrt{e}} \frac{1}{\sqrt{d+1}} \right) \quad (17)$$

MaxCut in the case of a regular graph of degree 2, so the graph is a ring, is termed the ring of disagrees in Ref. [1]. In this case, for even  $n$ , the optimal partition is simply to include every other vertex into one set and the rest into the other set, hence  $C_{\max} = n$ . Equation (16) yields the approximation ratio 0.75 at  $(\beta, \gamma) = (\pi/8, \pi/4)$ , reproducing the results in [1]. For triangle-free 3-regular graph ( $d = 2$ ), the ratio is 0.692, also in agreement with the results of Ref. [1] for a general 3-regular graph.

Because  $C_{\max}^{\text{reg}}(d) > |E|/2$  holds for all  $d$ , and the values are concentrated around the expectation, QAOA<sub>1</sub> beats random guessing for arbitrary triangle-free regular graphs. For an arbitrary triangle-free graph with maximum vertex degree  $d+1$ , applying the value to the right-hand side of Eq. (16) gives a lower bound to  $F^*$ .

While it is straightforward to extend the analysis in the proof of Theorem 1 to QAOA of higher levels, the number of terms quickly becomes prohibitive for direct calculation; many more non-commuting terms coming from the  $U_C$ 's and  $U_B$ 's must be retained and carried through the calculation. The expectation value of a given edge will also depend on its local graph topology, which becomes difficult to succinctly characterize as  $p$  increases. (See Appendix C for the expression for the ring of disagrees for  $p = 2$ .)

#### IV. ANALYSIS OF THE PROBLEM OF RING OF DISAGREES (ANTI-FERROMAGNETIC CHAIN)

We now study in detail QAOA for the ring of disagrees. We show that analysis can be done for QAOA <sub>$p$</sub>  for arbitrary level  $p$ , extending numerical results for small  $p$  given in Ref. [1].

##### A. Formulation of the problem

The Hamiltonian for the ring of disagrees with  $n$  vertices i.e., a one-dimensional ring of spins of spin-1/2, is  $\tilde{H}_C = \frac{1}{2} \sum_{j=1}^n (1 - \sigma_j^z \sigma_{j+1}^z)$  where  $\sigma_{n+1}^z = \sigma_1^z$ . For convenience, we later consider only even  $n$ , in which case the

ground state of  $\tilde{H}_C$  is trivial with every pair of neighboring spins aligned in anti-parallel fashion, corresponding to  $C_{\max} = n$ . The approximation ratio is then  $r = F^*/n$ .

To simplify the derivation, and also to conform to the convention in physics to minimize instead of maximize, we drop the constant and rescale  $\tilde{H}_C$  to be

$$H_C = \sum_j \sigma_j^z \sigma_{j+1}^z, \quad (18)$$

which defines the operator  $U_C(\gamma)$  given in Eq. (3). The initial state of the system is prepared (Eq. (5)), and the algorithm is specified by the QAOA circuit of Eq. (6). Rewriting Eq. (8) taking into account our simplification, the approximation ratio for  $F$  with parameters  $(\beta, \gamma)$  is

$$r = \frac{1}{2} \left( 1 - \frac{F^*}{n} \right). \quad (19)$$

The problem is now to determine parameters  $(\beta, \gamma)$  that create a quantum state that approximately minimizes the expectation value of  $H_C$  (and thus maximizes  $r$ ). The relation between the angles and the expectation values used in the remainder of this paper and the ones in Ref. [1] (notations with tilde) is  $\gamma = -\tilde{\gamma}/2$ ,  $\beta = \tilde{\beta}$  and  $\tilde{F}(\tilde{\gamma}, \tilde{\beta}) = (n - F(\gamma, \beta))/2$ , while the approximation ratio is the same.

##### B. Fermionic representation

We show that using a fermionic representation, the parameter setting of QAOA reduces to finding the optimal quantum control of an ensemble of independent spins (spin-1/2).

Since spin operators do not obey canonical commutation relations, transforming them into bosonic or fermionic operators is often useful for analysis. Such transformations enable the application of standard techniques in condensed matter physics such as diagrammatic perturbation. The algebra of the original spin operators must be preserved in the mappings. The Jordan-Wigner transformation [11, 12] maps the spin operators to fermions with a long-range phase factor.

We apply the Jordan-Wigner transformation [11, 12],

$$a_j = S_j^- e^{-i\phi_j} \quad (20)$$

$$a_j^\dagger = S_j^+ e^{i\phi_j} \quad (21)$$

where  $S_j^+ = (\sigma_j^y + i\sigma_j^z)/2$ ,  $S_j^- = (\sigma_j^y - i\sigma_j^z)/2$ , and the phase factor  $\phi_j = \pi \sum_{j' < j} (\sigma_{j'}^x + 1)/2$  is long-ranged involving all operators for  $j' < j$ . The new operators  $a_j$ ,  $a_j^\dagger$  can be verified to obey the fermion anticommutation relations,  $\{a_j, a_{j'}^\dagger\} = a_j a_{j'}^\dagger + a_{j'}^\dagger a_j = \delta_{j,j'}$ , and  $\{a_j, a_{j'}\} = \{a_j^\dagger, a_{j'}^\dagger\} = 0$ . The inverse transformation

reads

$$S_j^+ = a_j^\dagger e^{-i\phi_j} \quad (22)$$

$$S_j^- = a_j e^{i\phi_j} \quad (23)$$

$$\sigma_j^x = 2a_j^\dagger a_j - 1 \quad (24)$$

and the phase factor in the fermionic representation is  $\phi_j = \pi \sum_{j' < j} a_{j'}^\dagger a_{j'}$ . The Jordan-Wigner transformation is a convenient tool for one-dimensional spin systems, particularly for nearest-neighbored couplings because in products of the neighboring spin operators like  $S_j^+ S_{j+1}^-$ , the phase factors drop out, leaving a concise expression with short-ranged coupling.

The Jordan-Wigner transformation for our problem works for both even and odd  $n$ . We will work on this general case in this section. Applying the transformation to the problem and mixing Hamiltonians, we get

$$H_B = \sum_{j=1}^n (2a_j^\dagger a_j - 1) \quad (25)$$

$$H_C = \sum_{j=1}^{n-1} a_j^\dagger a_{j+1} + a_j a_{j+1} - (a_N^\dagger a_1 + a_N a_1)G + \text{h.c.}, \quad (26)$$

where we introduce the gauge operator  $G = \exp[i\pi \sum_{l=1}^n a_l^\dagger a_l] = (-1)^n \prod_{j=1}^n \sigma_j^x$ , a necessary treatment for periodic boundary conditions. In the standard QAOA setting, the initial state is an eigenstate of  $G$  with eigenvalue 1 for even  $n$  and -1 for odd  $n$ . The operator  $G$  is a constant of motion since it commutes with both  $H_B$  and  $H_C$ , so the value of  $G$  remains constant throughout the evolution. Therefore for even  $n$ , the sign of the  $j = n$  term in  $H_C$  is different from the others and requires a special treatment.

We further introduce a phase factor to unify the expression,  $b_j = a_j e^{-ij\pi/n}$ . The Hamiltonians then read

$$H_B = \sum_{j=1}^n (2b_j^\dagger b_j - 1) \quad (27)$$

$$H_C = e^{i\pi/n} \sum_{j=1}^n (b_j^\dagger b_{j+1} + e^{2ij\pi/n} b_j b_{j+1}) + \text{h.c.} \quad (28)$$

Upon applying a Fourier transformation to  $b_j$  (to  $a_j$  for odd  $n$ ),

$$c_k = \frac{1}{\sqrt{n}} \sum_{j=1}^n e^{\omega j k} b_j, \quad \omega = 2i\pi/n, \quad (29)$$

the driver and the problem Hamiltonians in the momen-

tum space take the form

$$H_B = \sum_{k=0}^{n-1} (2c_k^\dagger c_k - 1) \quad (30)$$

$$H_C = 2 \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \cos \theta_k (c_k^\dagger c_k + c_{-k}^\dagger c_{-k}) + i \sin \theta_k (c_k c_{-k} + c_k^\dagger c_{-k}^\dagger) + H_{C,0} \quad (31)$$

where for even  $n$

$$\begin{cases} H_{C,0} = 0 \\ \theta_k = (2k+1)\pi/n \\ c_{-k} \equiv c_{n-1-k}, \end{cases} \quad (32)$$

and for odd  $n$

$$\begin{cases} H_{C,0} = -c_0^\dagger c_0 \\ \theta_k = 2k\pi/n \\ c_{-k} \equiv c_{n-k}. \end{cases} \quad (33)$$

Since in Eq. (31),  $c_k$  and  $c_k^\dagger$  are solely coupled to  $c_{-k}$  and  $c_{-k}^\dagger$ , we only need to solve a set of 2-fermion problems. Because both  $H_B$  and  $H_C$  preserve the parity of the fermionic excitations, we need to consider only the ground state and the double excited state of the two fermions. For each  $k$ , in this two-dimensional subspace the driver and the problem Hamiltonians become  $2\sigma^z$  and  $2\sigma^z \cos \theta_k + 2\sigma^x \sin \theta_k$ , respectively.

In summary, after transforming the problem to a fermionic representation, the original many-body Hamiltonian of a ring of  $n$  spins reduces to an ensemble of  $n/2$  non-interacting spins of spin-1/2, which we would refer to as *pseudospins* to distinguish them from spins in the original problem:

$$\begin{aligned} H_B &= \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} H_{B,k} \\ H_C &= \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} H_{C,k} \end{aligned} \quad (34)$$

each term taking the form

$$H_{B,k} = 2\sigma^z \quad (35)$$

$$H_{C,k} = 2(\cos \theta_k \sigma^z + \sin \theta_k \sigma^x) = 2\hat{k} \cdot \hat{\sigma}. \quad (36)$$

where the unit vector  $\hat{k} = (\sin \theta_k, 0, \cos \theta_k)$ .

The initial state for each pseudospin is the ground state of  $-H_{B,k}$ , i.e.,  $\rho_0 = |1\rangle\langle 1| = (\mathbb{1} + \sigma^z)/2$  and the optimization reduces to minimize

$$F(\gamma, \beta) = \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} F_k(\gamma, \beta), \quad (37)$$



where

$$F_k(\gamma, \beta) = \frac{1}{2} [H_{C,k} U_k \sigma^z U_k^\dagger] \quad (38)$$

$$= \text{Tr} [\hat{k} \cdot \hat{\sigma} U_k \sigma^z U_k^\dagger]. \quad (39)$$

Hereafter, for notation simplicity, we drop the subscript for  $U_k$  and use  $U$  to refer to the evolution operator for the single pseudospin.  $U = U_B(\beta_p) U_C(\gamma_p) \cdots U_B(\beta_1) U_C(\gamma_1)$  now consists of only single-spin operators

$$U_B(\beta_l) = \exp[-i2\beta_l \sigma^z] \quad (40)$$

$$U_C(\gamma_l) = \exp[-i2\gamma_l \hat{k} \cdot \hat{\sigma}] \quad (41)$$

for  $l = 1, 2, \dots, p$ .

### C. Size-dependence of the approximation ratio

For sufficiently large problem sizes, the approximation ratio of QAOA on the problem of ring of disagrees of even  $n$  is independent of the problem size. This property has been shown in Ref. [1] using an operator reduction argument. The specific value of the approximation ratio for QAOA <sub>$p$</sub>  was conjectured to be  $(2p+1)/(2p+2)$  therein. Here, we show that this size-independence feature comes naturally out of the picture of single spin rotations. Each  $U_C(\gamma) = \cos(2\gamma) - i \sin(2\gamma) \hat{k} \cdot \hat{\sigma}$  can contribute a trigonometric function of  $\theta_k$ . Thus,  $F_k$  takes a form

$$F_k = \sum_{\substack{l, l' = 0 \\ l + l' \leq 2p+1}}^{2p+1} f_{l,l'}(\gamma, \beta) \sin^l \theta_k \cos^{l'} \theta_k, \quad (42)$$

where  $(f_{l,l'})$ 's are real coefficients independent of  $\theta_k$ . Since each  $\sin \theta_k$  accompanies one  $\sigma^x$ , using properties of Pauli matrices,  $\text{Tr}[\sigma_\alpha \sigma_{\alpha'}] = 2\delta_{\alpha, \alpha'}$ , the coefficient  $f_{l,l'}(\gamma, \beta)$  is zero for odd  $l$ . Recall that  $\theta_k = (2k-1)\pi/n$ . When we consider  $F$ ,

$$F = \sum_{\substack{l, l' = 0 \\ l + l' \leq 2p+1}}^{2p+1} \left( f_{l,l'}(\gamma, \beta) \sum_{k=1}^{n/2} \sin^l \theta_k \cos^{l'} \theta_k \right), \quad (43)$$

for even  $l$ , we have  $\sum_{k=1}^{n/2} \sin^l \theta_k \cos^{l'} \theta_k = 0$  for odd  $l'$ . Therefore, we need to keep only terms with even  $l$  and even  $l'$ , reducing Eq. (42), to a trigonometric polynomial of  $2\theta_k$  of degree  $p$ ,

$$F_k = \sum_{s=0}^p d_{2s}(\gamma, \beta) \cos(2s\theta_k), \quad (44)$$

where  $d_{2s}(\gamma, \beta)$  is a coefficient independent of  $k$ . See the analysis for  $p = 1$  and  $2$  in Sec. IV E for example.

Eq. (44) takes the form of the Fourier transformation of series  $d_{2s}$  with a cutoff at order  $p$ . For any specific

parameter values  $(\gamma, \beta)$ , for  $n \geq 2p+2$ , the constant component  $d_0 = \sum_k F_k/n$ :

$$F = \frac{n}{2} \cdot d_0(\gamma, \beta). \quad (45)$$

Since the  $n$ -dependence of  $F_k$  lies in  $\theta_k$  and  $d_0$  is  $\theta_k$ -independent, the expectation value  $F$ , and furthermore the approximation ratio of QAOA, is independent of  $n$ . For an arbitrary level  $p$ , simplifying Eq. (39) to get the specific trigonometrical function form can be done easily. Finding parameters  $(\gamma, \beta)$  that optimize  $F$  appears to be highly non-trivial.

### D. Symmetry and criticality-constrained manifolds

In this section, we show that, based on symmetries in the pseudospin rotations, we can identify sub-manifolds in the search space that admit extrema. In later sections, we provide numerical evidence that the global minima always lie in these sub-manifolds. This evidence suggests that one can focus the search within the identified sub-manifolds and thus reduce the search effort.

#### 1. Physics: rotations of the Bloch vectors

For each pseudospin, Eq. (39) can be expressed as

$$F_k(\gamma, \beta) = 4\mathcal{F}_k - 2 \quad (46)$$

where

$$\mathcal{F}_k \equiv \text{Tr}[\rho_{\hat{k}} U \rho_z U^\dagger] \quad (47)$$

and  $\rho_{\hat{k}} = \frac{1}{2}(1 + \hat{k} \cdot \hat{\sigma})$  and  $\rho_z = \frac{1}{2}(1 + \sigma^z)$ .

On the Bloch sphere,  $\rho_{\hat{k}}$  and  $\rho_z$  can be interpreted as the density matrices for the Bloch vectors in the  $\hat{k}$ -direction and  $\hat{z}$ -direction, respectively. Equation (47) represents a single pseudospin, initialized along  $\hat{z}$ -direction, then rotated about  $\hat{k}$ -axis for angle  $4\gamma_1$ , rotated about  $\hat{z}$  for  $4\beta_1, \dots$ , rotated about  $\hat{k}$  for  $4\gamma_p$ , rotated about  $\hat{z}$  for  $4\beta_p$ , and measured along  $\hat{k}$ . The fidelity  $\mathcal{F}_k$  measures the overlap between the final state and the state  $\rho_{\hat{k}}$ , whose Bloch vector is along direction  $\hat{k}$ . Due to the periodicity in rotation,  $F(4\gamma + 2l\pi, 4\beta + 2l'\pi) = F(\gamma, \beta) \Rightarrow F(\gamma + l \cdot \pi/2, \beta + l' \cdot \pi/2) = F(\gamma, \beta)$ , where  $l, l' \in \mathbb{Z}^p$ . Hence the search space can be limited to  $\beta_k, \gamma_k \in [0, \pi/2]$  for  $k = 1, 2, \dots, p$ .

QAOA on the ring of disagrees thus corresponds to a physical picture in optimal quantum control, albeit with a specialized set of constraints. For the final average over  $k$  to get  $F$ , we can think of the system as an ensemble of pseudospins, each pseudospin  $k$  experiencing a constant magnetic field along  $\hat{k}$ , (the quantization field), and the system controlled by applying a strong uniform magnetic field along  $\hat{z}$  in the ‘‘bang-bang’’ style. Specifically, when

the field along  $\hat{z}$ ,  $\vec{B}_z$ , is on, the quantization field is negligible and all pseudospins are rotated about  $\hat{z}$  by the same angle  $4\gamma_p$ ; when  $\vec{B}_z$  is paused, each pseudospin evolves freely, i.e., rotates about its own quantization axis  $\hat{k}$  to pick up an angle  $4\beta_p$ . After the whole control sequence is applied, the overall magnetization along  $\hat{z}$ ,  $F = \sum_k \langle \sigma_k^z \rangle$ , is measured. The goal of the quantum control is to design a time sequence  $(\gamma, \beta)$  so that  $F$  is minimized.

## 2. Criticality-constrained sub-manifolds

Since the trace operator preserves cycling, and the role of  $\hat{z}$  and  $\hat{k}$  in Eq. (47) are interchangeable, after initializing the pseudospin along the  $\hat{k}$ -direction, the same result would be obtained by rotating about  $\hat{z}$ -axis for angle  $-4\beta_p$ , rotating about  $\hat{k}$  for  $-4\gamma_p, \dots$ , rotating about  $\hat{z}$  for  $-4\beta_1$ , rotating about  $\hat{k}$  for  $-4\gamma_1$ , and measuring along  $\hat{z}$ .

**Manifold 1:** Due to equivalence in the two views, it must hold that

$$F_k(\gamma, \beta) = F_k(-\beta', -\gamma') \quad (48)$$

where

$$\gamma = (\gamma_1, \gamma_2, \dots, \gamma_{p-1}, \gamma_p) \quad (49)$$

$$\beta = (\beta_1, \beta_2, \dots, \beta_{p-1}, \beta_p) \quad (50)$$

$$\gamma' = (\gamma_p, \gamma_{p-1}, \dots, \gamma_2, \gamma_1) \quad (51)$$

$$\beta' = (\beta_p, \beta_{p-1}, \dots, \beta_2, \beta_1). \quad (52)$$

This can be verified with the help of a  $\pi$ -rotation about the axis  $\hat{z} + \hat{k}$ , see Appendix B for a proof. Consider the manifold that satisfies

$$\gamma_i + \beta_{p+1-i} = 0 \text{ for } i = 1, 2, \dots, p. \quad (53)$$

It has a special property: the gradient of the function  $F_k(\gamma, \beta)$  is constrained to lie tangent to the manifold. Therefore, critical points in the manifold are critical points of the whole function.

For  $p = 1$ , the relation (53) can also be proven to be a necessary condition for the global minima. On the Bloch sphere for a pseudospin, the rotation trajectory has to switch at the intersection of the two circles on the Bloch sphere which are perpendicular to one axis and passes through the vector end of the other pseudospin, as illustrated in Fig. 1. Because the intersection lies in the plane spanned by  $\hat{y}$  and the bisector of  $-\hat{z}$  and  $\hat{k}$ , it is obvious that  $\gamma_1 + \beta_1 = 0$ , i.e., the relation (53) has to be observed.

**Manifold 2:** Equation (47) actually holds for  $\rho_{\hat{k}} = \frac{1}{2}(1 \pm \hat{k} \cdot \hat{\sigma})$  and  $\rho_z = \frac{1}{2}(1 \pm \sigma^z)$ . The  $+$  ( $-$ ) sign correspond to the picture when the initial and final states are parallel (anti-parallel) with respect to the rotation axes, respectively. Comparing these two pictures, since rotations by the same angle about any axis  $-\hat{v}$  and  $+\hat{v}$  by the same

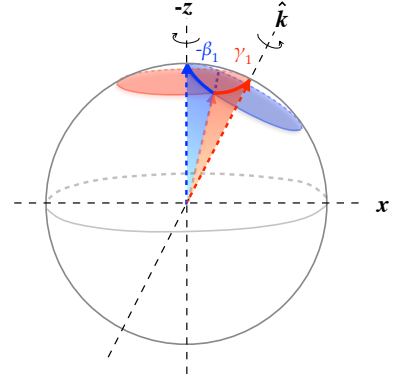


Fig. 1: (Color online). Schematic for the evolution trajectory of spin  $k$  on the Bloch sphere under QAOA<sub>1</sub> for arbitrary  $\theta_k > \theta_1^* = 2\pi/3$ . The trajectory of QAOA<sub>1</sub> is shown as arcs in solid lines.

angle are inverses to each other,  $\hat{R}_{(-\hat{v})}(\alpha) = \hat{R}_{(-\hat{v})}^\dagger(\alpha)$ , it must hold that

$$F_k(\gamma, \beta) = F_k(\beta', \gamma'). \quad (54)$$

Eq. (54) defines another manifold

$$\gamma_i - \beta_{p+1-i} = 0 \text{ for } i = 1, 2, \dots, p \quad (55)$$

with constrained gradient.

Eqs. (48) and (54) further indicate that  $F_k$  and is an even function of the angle sequence:  $F_k(\gamma, \beta) = F_k(-\gamma, -\beta)$ , and accordingly so is  $F$ ,

$$F(\gamma, \beta) = F(-\gamma, -\beta). \quad (56)$$

**Global extrema lie in the submanifolds** In our numerical search, the minima of  $F$  were always contained in the manifold defined by Eq. (53) while the maxima of  $F$  always lie in the manifold Eq. (55).

## E. Illustration of QAOA<sub>1</sub> and QAOA<sub>2</sub>

We use QAOA<sub>1</sub> and QAOA<sub>2</sub> to illustrate the results of symmetry and size-dependence of the optimization discussed above. Numerical results for higher levels ( $p > 2$ ) are shown in Appendix C.

For QAOA<sub>1</sub>, the unitary evolution operator is  $U = e^{-i2\beta_1\sigma^z} e^{-i2\gamma_1\hat{k}\cdot\hat{\sigma}}$ . Note that if a term  $f(k)$  in  $F_k$  satisfies  $f(n/2 + 1 - k) = -f(k)$ , then  $f(k)$  would vanish in  $F$  through the summation over  $k$ ; and note the properties of Pauli matrices,  $\text{Tr}[\sigma_\alpha \sigma_{\alpha'}] = 2\delta_{\alpha, \alpha'}$ , one comes to

$$F = 2 \sin(4\beta) \sin(4\gamma) \sum_k \sin^2 \theta_k \quad (57)$$

$$= \begin{cases} n \sin(4\beta) \sin(4\gamma) & \text{for } n = 2 \\ \frac{n}{2} \sin(4\beta) \sin(4\gamma) & \text{for } n > 2. \end{cases} \quad (58)$$



The optimal angles are  $(\gamma_1^*, \beta_1^*) = \pi \cdot (3/8, 1/8)$  or  $\pi \cdot (1/8, 3/8)$ .

For  $n = 2$ , the optimal angles correspond to  $F^* = -n$  while for larger problem size,  $F^* = -n/2$ . This reflects the property that QAOA<sub>p</sub> suffices to perfectly optimize the ring for  $n \leq 2p$  but for  $n \geq 2p + 2$  the optimization ratio is a fixed constant smaller than 1.

Eq. (58) is plotted in Fig. 2. Along the symmetry line  $\beta_1 + \gamma_1 = 0$ , the critical points are global minima and saddle points. While along the symmetry line  $\beta_1 - \gamma_1 = 0$ , the critical points are global maxima and saddle points.

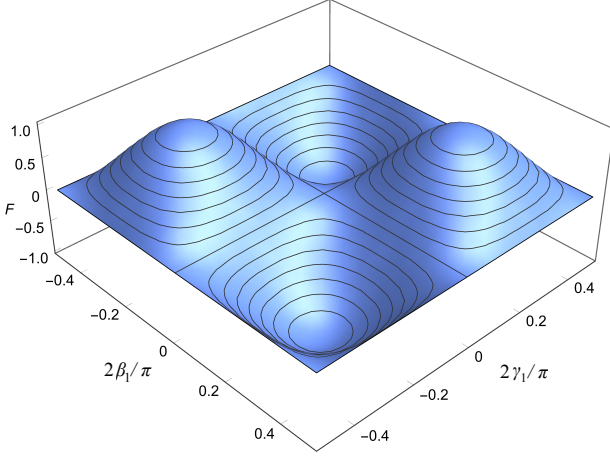


Fig. 2: (Color online). QAOA<sub>1</sub>. The expectation value  $F/n$  as a function of  $\gamma_1$  and  $\beta_1$ .

For level-2, the evolution operator reads

$$U = e^{-i2\beta_2\sigma^z} e^{-i2\gamma_2\hat{\sigma}\cdot\hat{k}} e^{-i2\beta_1\sigma^z} e^{-i2\gamma_1\hat{\sigma}\cdot\hat{k}}. \quad (59)$$

The expectation value  $F$  as a trigonometric function of  $(\gamma, \beta)$  is shown in Appendix C. Numerically found optimal angles are  $(\gamma_1^*, \beta_1^*, \gamma_2^*, \beta_2^*) = \pi \cdot (0.3956, 0.1978, 0.3022, 0.1044)$  or  $\pi \cdot (0.2052, 0.1026, 0.3974, 0.2948)$ . In both optimal angle sets,  $4(\gamma_1^* + \beta_2^*)$  and  $4(\gamma_2^* + \beta_1^*)$  are integer multipliers of  $2\pi$ , thus both optima lie in the manifold defined by Eq. (53).

## F. Discussion: controllability and optimality

The optima of  $F$  for QAOA level  $p = 1$  to 10 are tabulated in Appendix C. The optimal angles were obtained through numerical gradient descent search in the sub-manifold Eq. (53). The evaluation for  $F_k$  in each step is realized as Eq. (39), which only involves  $2p$  multiplications of 2-by-2 matrices, and sum over  $k$  gives  $F$ , so optimal angles for higher  $p$  could be computed easily if desired.

Starting with a random initial guess of  $(\gamma, \beta)$ , the search (with sufficiently small steps) always converges

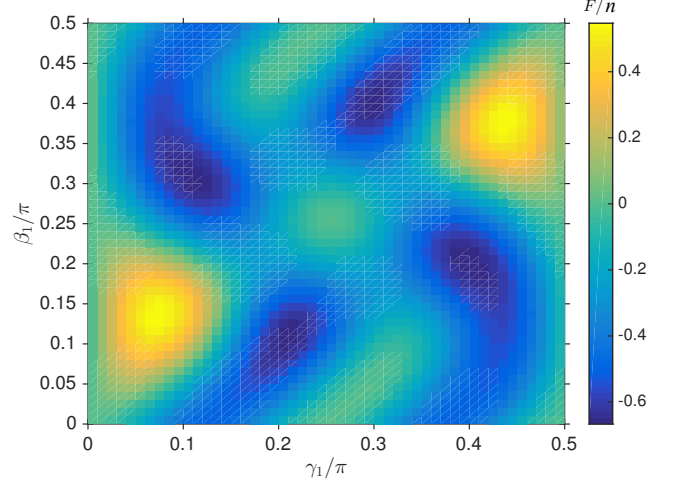


Fig. 3: (Color online). The landscape of  $F/n$  for QAOA<sub>2</sub>, in the sub-manifold Eq. (53). The four darkest spots indicate the global minima  $F^*/n = -2/3$ . The origin  $(0, 0)$  is a saddle point. No local minima are observed. The contour is symmetric w.r.t.  $(\gamma_1, \beta_1) = (\pi/4, \pi/4)$ , reflecting the symmetry in Eq. (56), (and the period  $\pi/2$ ).

to a global minimum. This behavior suggests that at least within the sub-manifold, all local minima are global minima. For example, for  $p = 2$ , there are two free parameters in the sub-manifold, which we choose to be  $\gamma_1$  and  $\beta_1$ . Figure 3 shows the landscape of the expectation value  $F$ . The four minima (darkest spots) observed in one period  $(\gamma_1, \beta_1) \in [0, \pi/2]$  are all global minima.

This result calls for extended understanding of landscapes of quantum control. In quantum control theory, it has been shown that assuming *controllability*, i.e., evolution between any two states is achievable via the set of controls given, the landscape of the infidelity  $F$  over the space of parameter values  $(\gamma, \beta)$  generically has only global minima [13–15]. Without controllability, the quantum control landscape in general is rugged and admits local minima [16].

In the case of QAOA, the controls are constrained in a specific way: if an infinite number of controls are allowed, i.e.,  $p \rightarrow \infty$ , then the system is controllable. The finite number of control steps dictated by the level  $p$  limits the controllability. For more general graphs, however, Eq. (14) shows that for QAOA on MaxCut, even for  $p = 1$ , there exists local optima in the space of parameter values. For the special case of ring of disagrees, the system is still not controllable, however, our numerical results indicate that, at least within the sub-manifold Eq. (53), all local minima are global minima.

## V. CONCLUSIONS

We studied parameter setting for QAOA on MaxCut. For QAOA<sub>1</sub>, we extended the results in Ref. [1], pro-

viding an analytical expression for general graphs. As a corollary, for triangle-free graphs with fixed vertex degree, the optimal angles for the driver Hamiltonian can be directly read off while the optimal angles for the problem Hamiltonian show a dependence on the vertex degree. For higher level  $p$ , direct analysis on the operator expansion becomes cumbersome, providing further evidence that more advanced parameter setting techniques need to be developed.

For a special case of MaxCut, the ring of disagrees, which corresponds to a one-dimensional anti-ferromagnetic spin ring, we analyze QAOA $_p$ , for arbitrary  $p$ , using a fermionic representation. Applying the Jordan-Wigner transformation transforms the evolution of an  $n$ -qubit system under QAOA to a set of  $n/2$  independent evolutions within two-dimensional subspaces. The parameter setting problem, thus, corresponds to finding optimal control parameters for an ensemble of  $n/2$  non-interacting pseudospins of spin- $\frac{1}{2}$ . From this treatment we obtain analytical expression for any  $p$ , and an easy numerical search for the corresponding optimal angles.

The fermionic picture also enables us to explain symmetries in the optimal parameters, suggesting a means to further reduce the effort required to find optimal parameters by restricting to manifolds defined by these symmetries. The specific symmetry in the problem of ring of disagrees has its roots in the equal footing of the action of the driver and the problem Hamiltonians – each corresponds to a single spin rotation. We observed numerically that within the parameter space, all minima are global minima. While such a “no-trap” character of a quantum control landscape can be explained given controllability, the structure of QAOA for finite  $p$  often does not guarantee controllability. Future research that reveals the underlying theory may shed further light on the control landscape and the structure of QAOA, and inspire ways to simplify or improve the algorithm.

In Ref. [1] it is conjectured that the best achievable approximation ratio for QAOA $_p$  on a ring of size  $n \geq 2p+2$  is  $(2p+1)/(2p+2)$ . The fermionic view we have presented, by simplifying the analysis, may be a useful step toward a proof, but for the moment the conjecture remains open. Further work will examine how realistic noise affects the performance. While for certain simple noise models, including control noise affecting the times for which the Hamiltonians are applied, can be analyzed within the model, for other cases a combination of more sophisticated analytical tools and experimentation on quantum hardware will be needed to evaluate performance under noise. The simplicity of QAOA on the ring makes it a promising target for implementation on early quantum processors.

## VI. ACKNOWLEDGEMENTS

We thank Sergey Knysh, Jason Dominy, Salvatore Mandrà, Bryan O’Gorman, Davide Venturelli, Nick Rubin, Will Zeng and Robert Smith for enlightening discussions. S.H. was supported by NASA under award NNX12AK33A. He would like to thank the Universities Space Research Association (USRA) and the NASA Ames Research Center for the opportunity to participate in the program that enabled this research. The authors would like to acknowledge support from the NASA Advanced Exploration Systems program and the NASA Ames Research Center. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purpose notwithstanding any copyright annotation thereon.

## Appendix A: Proof of Theorem 1

*Proof of Theorem 1.* For  $p = 1$ , only terms corresponding to neighbors of  $u$  and  $v$  can contribute to the expectation of  $C_{uv}$  [1]. We thus partition our objective function as

$$C = \frac{1}{2}(I - \sigma_u^z \sigma_v^z) + C_u + C_v + \bar{C},$$

where  $C_u$  is the  $d$ -many constraints involving only vertex  $u$  but not  $v$ , and  $C_v$  is the  $e$ -many constraints involving only  $v$ . The remaining constraints  $\bar{C}$  do not contribute to the expectation value  $\langle C_{uv} \rangle$ . For simplicity, we write  $(d, e, f) = (d_u, d_v, \lambda_{uv})$ .

Let  $c = \cos 2\beta$  and  $s = \sin 2\beta$ . We have

$$\begin{aligned} & e^{i\beta B} \sigma_u^z \sigma_v^z e^{-i\beta B} \\ &= c^2 \sigma_u^z \sigma_v^z + sc(\sigma_u^y \sigma_v^z + \sigma_u^z \sigma_v^y) + s^2 \sigma_u^y \sigma_v^y. \end{aligned} \quad (\text{A1})$$

The first term  $\sigma_u^z \sigma_v^z$  commutes with  $C$  and does not contribute to  $\langle C_{uv} \rangle$ . We conjugate each remaining term separately by  $e^{i\gamma C}$ . Let  $c' = \cos \gamma$  and  $s' = \sin \gamma$ . We have

$$\begin{aligned} & \text{Tr}[\rho_0 e^{i\gamma C} \sigma_u^y \sigma_v^z e^{-i\gamma C}] \\ &= \text{Tr}[\rho_0 (Ic' - is' \sigma_u^z \sigma_v^z) \prod_{i=1}^d (Ic' - is' \sigma_u^z \sigma_{w_i}^z) \sigma_u^y \sigma_v^z]. \end{aligned} \quad (\text{A2})$$

Expanding the product on the right gives a sum of tensor products of Pauli operators. Clearly, the only term that can contribute is proportional to  $\sigma_u^z \sigma_v^z * I^{\otimes d} * \sigma_u^y \sigma_v^z = -i\sigma_u^x$ . Thus we have

$$\text{Tr}[\rho_0 e^{i\gamma C} \sigma_u^y \sigma_v^z e^{-i\gamma C}] = \text{Tr}[\rho_0 (-i)s' c'^d (-i\sigma_u^x)] = -s' c'^d \quad (\text{A3})$$

By symmetry, we have  $\text{Tr}[\rho_0 e^{i\gamma C} \sigma_u^z \sigma_v^y e^{-i\gamma C}] = -s' c'^e$ . Observe that these terms are independent of the number of mutual neighbours (triangles) of  $u$  and  $v$ . The next term is

$$\begin{aligned} & \text{Tr}[\rho_0 e^{i\gamma C} \sigma_u^y \sigma_v^y e^{-i\gamma C}] = \text{Tr}[e^{2i\gamma C_u} e^{2i\gamma C_v} \sigma_u^y \sigma_v^y] \\ &= \text{Tr}\left[\prod_{i=1}^d (c'I - is' \sigma_u^z \sigma_{w_i}^z) \prod_{j=1}^e (c'I - is' \sigma_v^z \sigma_{w_j}^z) \sigma_u^y \sigma_v^y\right] \end{aligned} \quad (\text{A4})$$

The simplest term that contributes in this case is  $\text{Tr}[\rho_0 f c'^{d+e-2} (-is')^2 (-i\sigma_u^x) (-i\sigma_v^x)] = f c'^{d+e-2} s'^2$ . Corresponding to the triangles of  $\langle uv \rangle$ , in the above product we have  $f$ -many distinct values  $i$  such that  $w_i = w_j$ . As  $\sigma_u^z \sigma_{w_i}^z * \sigma_u^z \sigma_{w_i}^z = I$ , if  $f > 2$  then higher order terms depending on the number of triangles  $f$  will contribute. For example, the next order terms will result from three pairs of  $(\sigma_u^z \sigma_{w_i}^z, \sigma_u^z \sigma_{w_i}^z)$  and hence be proportional to  $s'^6$ .

Thus we have

$$\begin{aligned} & \text{Tr}[\rho_0 (e^{i\gamma C} \sigma_u^y \sigma_v^y e^{-i\gamma C})] \\ &= \binom{f}{1} c'^{d+e-2} s'^2 + \binom{f}{3} c'^{d+e-6} s'^6 \\ &+ \binom{f}{5} c'^{d+e-10} s'^{10} + \dots \\ &= c'^{d+e-2f} \sum_{i=1,3,5,\dots}^f \binom{f}{i} (c'^2)^{f-i} (s'^2)^i. \end{aligned} \quad (\text{A5})$$

To sum this series, recall the binomial theorem, which we may split into even and odd sums as

$$\begin{aligned} & \sum_{i=0,2,\dots}^f \binom{f}{i} a^{f-i} b^i + \sum_{i=1,3,\dots}^f \binom{f}{i} a^{f-i} b^i \\ &= \sum_{i=0}^f \binom{f}{i} a^{f-i} b^i = (a+b)^f \end{aligned} \quad (\text{A6})$$

which also gives

$$\begin{aligned} & \sum_{i=0,2,\dots}^f \binom{f}{i} a^{f-i} b^i - \sum_{i=1,3,\dots}^f \binom{f}{i} a^{f-i} b^i \\ &= \sum_{i=0}^f (-1)^i \binom{f}{i} a^{f-i} b^i = (a-b)^f, \end{aligned} \quad (\text{A7})$$

and hence

$$\sum_{i=1,3,\dots}^f \binom{f}{i} a^{f-i} b^i = \frac{1}{2}((a+b)^f - (a-b)^f). \quad (\text{A8})$$

Thus the above sum becomes

$$\sum_{i=1,3,\dots}^f \binom{f}{i} (c'^2)^{f-i} (s'^2)^i = \frac{1}{2}(1 - \cos^f 2\gamma) \quad (\text{A9})$$

which yields

$$\text{Tr}[\rho_0 (e^{i\gamma C} \sigma_u^y \sigma_v^y e^{-i\gamma C})] = \frac{1}{2} c'^{d+e-2f} (1 - \cos^f 2\gamma) \quad (\text{A10})$$

Putting this all together, we have

$$\begin{aligned} \langle C_{uv} \rangle &= \text{Tr}[\rho_0 e^{i\gamma C} e^{i\beta B} C_{uv} e^{-i\beta B} e^{-i\gamma C}] \\ &= \frac{1}{2} + \frac{sc}{2} \text{Tr}[\rho_0 e^{i\gamma C} (\sigma_u^y \sigma_v^z + \sigma_u^z \sigma_v^y) e^{-i\gamma C}] \\ &\quad - \frac{s^2}{2} \text{Tr}[\rho_0 e^{i\gamma C} \sigma_u^y \sigma_v^y e^{-i\gamma C}] \\ &= \frac{1}{2} + \frac{1}{2} scs' (c'^d + c'^e) - \frac{1}{4} s^2 c'^{d+e-2f} (1 - \cos^f 2\gamma). \end{aligned} \quad (\text{A11})$$

□

**Appendix B: Proof of symmetry relation Eq. (48)**where  $U' \equiv U(-\beta', -\gamma')$ .

We prove Eq. (48)

Insert  $R^2 = 1$  to  $F_k$  we get

$$F_k(\gamma, \beta) = F_k(-\beta', -\gamma') \quad (\text{B1})$$

where

$$\gamma = (\gamma_1, \gamma_2, \dots, \gamma_{p-1}, \gamma_p) \quad (\text{B2})$$

$$\beta = (\beta_1, \beta_2, \dots, \beta_{p-1}, \beta_p) \quad (\text{B3})$$

$$\gamma' = (\gamma_p, \gamma_{p-1}, \dots, \gamma_2, \gamma_1) \quad (\text{B4})$$

$$\beta' = (\beta_p, \beta_{p-1}, \dots, \beta_2, \beta_1) . \quad (\text{B5})$$

*Proof.* We consider a unitary operator  $R = \cos \frac{\theta}{2} \sigma^z + \sin \frac{\theta}{2} \sigma^x$  which rotates a Bloch vector about axis  $(\hat{k} + \hat{z})$  by  $\pi$ . Note that  $R^\dagger = R$ ,  $R^2 = 1$  and

$$\begin{aligned} R\sigma^z R &= \hat{k} \cdot \hat{\sigma} \\ R\hat{k} \cdot \hat{\sigma} R &= \sigma^z \\ RU_B(\beta)R &= U_C(\beta) \\ RU_C(\gamma)R &= U_B(\gamma) , \end{aligned} \quad (\text{B6})$$

we have

$$\begin{aligned} RUR &= RU_B(\beta_p)RRU_C(\gamma_p)R \dots RU_B(\beta_1)RRU_C(\gamma_1)R \\ &= U_C(\beta_p)U_B(\gamma_p) \dots U_C(\beta_1)U_B(\gamma_1) \\ &= [U_B(-\gamma_1)U_C(-\beta_1) \dots U_B(-\gamma_p)U_C(-\beta_p)]^\dagger \\ &= U'^\dagger \end{aligned} \quad (\text{B7})$$

$$\begin{aligned} F_k(\gamma, \beta) &= \text{Tr}[(\hat{k} \cdot \hat{\sigma})U\sigma^z U^\dagger] \\ &= \text{Tr}[R(\hat{k} \cdot \hat{\sigma})RRURR\sigma^z RRU^\dagger R] \\ &= \text{Tr}[\sigma^z RU R(\hat{k} \cdot \hat{\sigma})RU^\dagger R] \\ &= \text{Tr}[\sigma^z U'^\dagger (\hat{k} \cdot \hat{\sigma})U'] \\ &= \text{Tr}[(\hat{k} \cdot \hat{\sigma})U'\sigma^z U'^\dagger] \\ &= F_k(-\beta', -\gamma') \end{aligned} \quad (\text{B8})$$

□

**Appendix C: Detailed results for  $p = 2$  and higher**For  $p = 2$ , terms in that are non-vanishing to  $F$  is

$$\begin{aligned} \frac{F}{n} &= \frac{1}{64} [ -7 \cos(4\beta_1 + 4\beta_2 + 4\gamma_1 + 4\gamma_2) - 6 \cos(4\beta_1 + 4\beta_2 + 4\gamma_1) \\ &\quad + 3 \cos(4\beta_1 + 4\beta_2 - 4\gamma_1 + 4\gamma_2) + 4 \cos(4\beta_1 + 4\beta_2 + 4\gamma_2) \\ &\quad + 3 \cos(4\beta_1 - 4\beta_2 + 4\gamma_1 + 4\gamma_2) - 6 \cos(4\beta_1 - 4\beta_2 + 4\gamma_1) - \\ &\quad 3 \cos(4\beta_1 - 4\beta_2 - 4\gamma_1 + 4\gamma_2) + 4 \cos(4\beta_1 + 4\gamma_1 + 4\gamma_2) - \\ &\quad 4 \cos(4\beta_1 + 4\gamma_1) - 4 \cos(4\beta_1 + 4\gamma_2) - 3 \cos(-4\beta_1 + 4\beta_2 + 4\gamma_1 + 4\gamma_2) \\ &\quad + 6 \cos(-4\beta_1 + 4\beta_2 + 4\gamma_1) + 3 \cos(-4\beta_1 + 4\beta_2 - 4\gamma_1 + 4\gamma_2) + \\ &\quad 7 \cos(-4\beta_1 - 4\beta_2 + 4\gamma_1 + 4\gamma_2) + 6 \cos(-4\beta_1 - 4\beta_2 + 4\gamma_1) - \\ &\quad 3 \cos(-4\beta_1 - 4\beta_2 - 4\gamma_1 + 4\gamma_2) - 4 \cos(-4\beta_1 - 4\beta_2 + 4\gamma_2) - \\ &\quad 4 \cos(-4\beta_1 + 4\gamma_1 + 4\gamma_2) + 4 \cos(-4\beta_1 + 4\gamma_1) + 4 \cos(-4\beta_1 + 4\gamma_2) - \\ &\quad 6 \cos(4\beta_2 + 4\gamma_1 + 4\gamma_2) - 6 \cos(4\beta_2 - 4\gamma_1 + 4\gamma_2) - 4 \cos(4\beta_2 + 4\gamma_2) \\ &\quad + 6 \cos(-4\beta_2 + 4\gamma_1 + 4\gamma_2) + 6 \cos(-4\beta_2 - 4\gamma_1 + 4\gamma_2) \\ &\quad + 4 \cos(-4\beta_2 + 4\gamma_2) ] . \end{aligned} \quad (\text{C1})$$

If limited in the sub-manifold Eq. (53),

$p$	$r$	$F^*/n$	$\gamma_1$	$\beta_1$	$\gamma_2$	$\beta_2$	$\gamma_3$	$\beta_3$	$\gamma_4$	$\beta_4$	$\gamma_5$	$\beta_5$
1	3/4	-1/2	0.1250									
2	5/6	-2/3	0.2052	0.1026								
3	7/8	-3/4	0.2268	0.1888	0.0918							
4	9/10	-4/5	0.2357	0.2161	0.1791	0.0850						
5	11/12	-5/6	0.2403	0.2282	0.2094	0.1724	0.0802					
6	13/14	-6/7	0.3035	0.1639	0.2506	0.2835	0.0794	0.2409				
7	15/16	-7/8	0.2303	0.1623	0.3468	0.2690	0.1042	0.2397	0.1599			
8	17/18	-8/9	0.2445	0.1638	0.2839	0.3484	0.1539	0.1530	0.2581	0.1291		
9	19/20	-9/10	0.1929	0.1648	0.3307	0.3016	0.1551	0.2538	0.2174	0.1089	0.3117	
10	21/22	-10/11	0.2208	0.1374	0.3098	0.2974	0.2702	0.1205	0.3148	0.1904	0.1423	0.2572

TABLE I: Optimal angles for different levels of QAOA. Angles are in units of  $\pi$  radians. Gradient descend search is implemented with the optimal angles for level  $p$  set to be the initial guess for level  $p + 1$ . Arbitrary initial guess also always converges to a global minimum of  $F(\gamma, \beta)$  in the sub-manifold Eq. (53). Multiple sets of optimal angles exist for  $p \geq 2$ , only one of them is shown for each level.

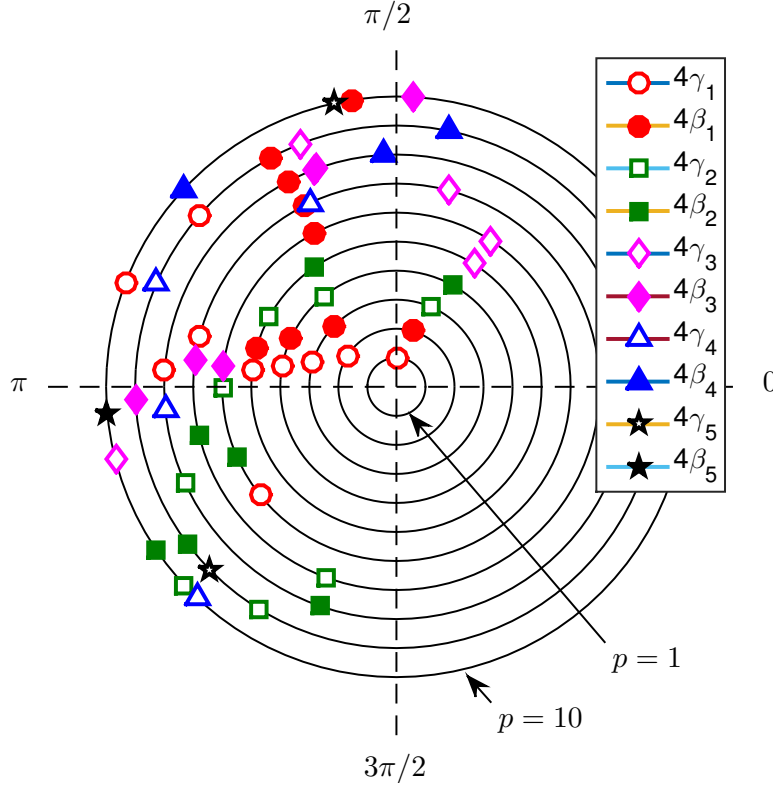


Fig. 4: (Color online). Optimal angles in the submanifold defined by Eq. (53). The optimal points are plotted on the complex plane with angles  $(4\gamma, 4\beta)$  as the argument and radius given by the level,  $p = 1$  to  $p = 10$  from the inner to outer circles.

$$\begin{aligned}
\frac{F}{n} = & \frac{1}{64} \left( -2 \cos(8\beta_1) + 3 \cos(8\beta_1 + 8\gamma_1) - 12 \cos(4\beta_1 + 8\gamma_1) \right. \\
& -8 \cos(4\beta_1 + 4\gamma_1) + 12 \cos(4\beta_1 - 8\gamma_1) + 8 \cos(4\beta_1 - 4\gamma_1) \\
& +7 \cos(8\beta_1 - 8\gamma_1) - 8 \cos(8\beta_1 - 4\gamma_1) + 6 \cos(8\gamma_1) \\
& \left. +8 \cos(4\gamma_1) - 14 \right). \tag{C2}
\end{aligned}$$



In Table. I we show numerical for optimal angles for higher QAOA levels in the manifold Eq. (53) (multiple optima were found for  $p \geq 2$ , we show only one for each

$p$ ). The same sets of angles are also plotted on the circles in Fig. 4.

- 
- [1] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A Quantum Approximate Optimization Algorithm. *arXiv:1411.4028 [quant-ph]*, November 2014.
  - [2] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A Quantum Approximate Optimization Algorithm Applied to a Bounded Occurrence Constraint Problem. *arXiv:1412.6062 [quant-ph]*, December 2014.
  - [3] Edward Farhi and Aram W. Harrow. Quantum Supremacy through the Quantum Approximate Optimization Algorithm. *arXiv:1602.07674 [quant-ph]*, February 2016.
  - [4] Zhi-Cheng Yang, Armin Rahmani, Alireza Shabani, Hartmut Neven, and Claudio Chamon. Optimizing variational quantum algorithms using Pontryagin’s minimum principle. *Phys. Rev. X*, 7:021027, May 2017.
  - [5] Zhang Jiang, Eleanor G. Rieffel, and Zhihui Wang. Near-optimal quantum circuit for Grover’s unstructured search using a transverse field. *Phys. Rev. A*, 95:062317, Jun 2017.
  - [6] Stuart Hadfield, Zhihui Wang, Bryan O’Gorman, Eleanor G. Rieffel, Davide Venturelli, and Rupak Biswas. From the Quantum Approximate Optimization Algorithm to a Quantum Alternating Operator Ansatz. *arXiv:1709.03489 [quant-ph]*, 2017.
  - [7] Boaz Barak, Ankur Moitra, Ryan O’Donnell, Prasad Raghavendra, Oded Regev, David Steurer, Luca Trevisan, Aravindan Vijayaraghavan, David Witmer, and John Wright. Beating the random assignment on constraint satisfaction problems of bounded degree. *arXiv:1505.03424*, 2015.
  - [8] Dave Wecker, Matthew B Hastings, and Matthias Troyer. Training a quantum optimizer. *Physical Review A*, 94(2):022309, 2016.
  - [9] John Preskill. Quantum computing and the entanglement frontier. *arXiv:1203.5813*, March 2012.
  - [10] Sergio Boixo, Sergei V. Isakov, Vadim N. Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, John M. Martinis, and Hartmut Neven. Characterizing Quantum Supremacy in Near-Term Devices. *arXiv:1608.00263*, July 2016.
  - [11] Elliott Lieb, Theodore Schultz, and Daniel Mattis. Two soluble models of an antiferromagnetic chain. *Annals of Physics*, 16(3):407–466, December 1961.
  - [12] Eytan Barouch, Barry M. McCoy, and Max Dresden. Statistical Mechanics of the XY Model. I. *Physical Review A*, 2(3):1075–1092, September 1970.
  - [13] Herschel A. Rabitz, Michael M. Hsieh, and Carey M. Rosenthal. Quantum optimally controlled transition landscapes. *Science*, 303(5666):1998–2001, 2004.
  - [14] Re-Bing Wu, Ruixing Long, Jason Dominy, Tak-San Ho, and Herschel Rabitz. Singularities of quantum control landscapes. *Phys. Rev. A*, 86:013405, Jul 2012.
  - [15] Benjamin Russell, Herschel Rabitz, and Rebing Wu. Quantum control landscapes are almost always trap free. *arXiv preprint arXiv:1608.06198*, 2016.
  - [16] Re-Bing Wu, Michael A. Hsieh, and Herschel Rabitz. Role of controllability in optimizing quantum dynamics. *Phys. Rev. A*, 83:062306, Jun 2011.