Obstacles to Variational Quantum Optimization from Symmetry Protection

Sergey Bravyi,[1] Alexander Kliesch,[2] Robert Koenig,[3] and Eugene Tang[4]

[1]*IBM Quantum, IBM T. J. Watson Research Center, Yorktown Heights, New York 10598, USA*
[2]*Zentrum Mathematik, Technical University of Munich, 85748 Garching, Germany*
[3]*Institute for Advanced Study and Zentrum Mathematik, Technical University of Munich, 85748 Garching, Germany*
[4]*Institute for Quantum Information and Matter, Caltech, Pasadena, California 91125, USA*

The quantum approximate optimization algorithm (QAOA) employs variational states generated by a parameterized quantum circuit to maximize the expected value of a Hamiltonian encoding a classical cost function. Whether or not the QAOA can outperform classical algorithms in some tasks is an actively debated question. Our work exposes fundamental limitations of the QAOA resulting from the symmetry and the locality of variational states. A surprising consequence of our results is that the classical Goemans-Williamson algorithm outperforms the QAOA for certain instances of MaxCut, at any constant level. To overcome these limitations, we propose a nonlocal version of the QAOA and give numerical evidence that it significantly outperforms the standard QAOA for frustrated Ising models.

Variational quantum optimization (VQO) has recently received significant attention as a candidate application of near-term quantum processors. The basic proposal is appealingly simple: The output state of a parameterized quantum circuit is used as a variational wave function to minimize the expected energy of a given Hamiltonian [1]. Depending on the envisioned application, the Hamiltonian may govern electronic interactions in a molecule or material of interest [2] or encode a classical cost function whose minimum is to be approximated [3]. Rotation angles that define individual gates in the state preparation circuit serve as variational parameters. These parameters are adjusted via a classical feedback loop that aims to minimize the expected energy.

The central question common to all VQO proposals is whether the chosen variational class of states is expressive enough to provide a good ground state approximation. Let us point out two factors that can limit the expressive power of VQO. First, the state preparation quantum circuit must have a small depth to enable reliable implementation on near-term noisy devices lacking error correction. This means that highly entangled states such as the ground state of Kitaev's toric code [4] may be out of scope for near-term VQO using gate-based devices and low-depth circuits [5,6]. Second, the number of variational parameters in the state preparation circuit must be small enough to enable efficient energy minimization. While this is not a serious concern for proof-of-principle experiments with a handful of qubits, it is anticipated that large-scale VQO with an extensive number of variational parameters may give rise to intractable optimization problems, for example, due to the barren plateau (vanishing gradient) effect [7].

In this Letter, we elaborate on the limitations of VQO and establish formal no-go results for the quantum approximate optimization algorithm (QAOA) [3]. Recall that the QAOA aims to approximate the maximum of a classical cost function $C(x)$ that depends on n binary variables, $x = (x_1, \ldots, x_n)$. The cost function is encoded into an n-qubit diagonal Hamiltonian $C = \sum_{x \in \{0,1\}^n} C(x)|x\rangle\langle x|$. The QAOA variationally maximizes the expected energy of C over n-qubit quantum states of the form [3]

$$|\psi(\beta, \gamma)\rangle = \prod_{a=1}^{p} e^{i\beta_a B} e^{i\gamma_a C} |+^n\rangle, \tag{1}$$

where β_a and γ_a are variational parameters, $|+^n\rangle$ is the tensor product of n single-qubit states $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, and $B = \sum_{j=1}^{n} X_j$ is the transverse magnetic field operator. The integer p, called the QAOA *level*, controls the expressive power of the variational ansatz. Finally, the QAOA outputs a bit string x obtained by preparing the optimal variational state $|\psi(\beta, \gamma)\rangle$ and measuring each qubit in the standard basis. The expected value of $C(x)$ coincides with the variational energy $\langle\psi(\beta, \gamma)|C|\psi(\beta, \gamma)\rangle$. The performance of the QAOA is commonly quantified by an *approximation ratio* defined as the ratio between the maximum variational energy and the maximum value of the cost function $C_{\max} = \max_x C(x)$.

A paradigmatic test case for the QAOA is the maximum cut (MaxCut) problem [3]. Suppose $G = (V, E)$ is a graph with a set of n vertices V labeled by integers $j = 1, \ldots, n$ and a set of edges E. Given an n-bit string x, let $\text{cut}(x)$ be the set of edges $(j, k) \in E$ such that $x_j \neq x_k$. The cost function to be maximized is the cut size, $C(x) = |\text{cut}(x)|$. The corresponding n-qubit Hamiltonian is

$$C = \frac{1}{2} \sum_{(j,k) \in E} (I - Z_j Z_k). \qquad (2)$$

Here, Z_j is the Pauli Z operator applied to a qubit j, and I is the identity. The state preparation circuit defined in Eq. (1) has depth $\approx pD$, where D is the maximum vertex degree of the graph G and p is the QAOA level [8]. To keep the circuit depth and the number of variational parameters small, below, we focus on the regime when p and D are constants or slowly growing functions of n.

Our first result is an upper bound on the maximum variational energy attained by level-p states. Namely, we show that for any constant $D \geq 3$ and all large enough n there exists a degree-D graph G with n vertices such that

$$\frac{\langle \psi(\beta, \gamma) | C | \psi(\beta, \gamma) \rangle}{C_{\max}} \leq \frac{5}{6} + \frac{\sqrt{D-1}}{3D} \qquad (3)$$

for any $\beta, \gamma \in \mathbb{R}^p$ as long as $p < (1/3 \log_2 n - 4)(D+1)^{-1}$. This result severely limits the performance of the QAOA with any constant level p independent of n. Indeed, the right-hand side of Eq. (3) is approximately $5/6 \approx 0.833$ for large vertex degree D. For comparison, the best-known classical algorithm for MaxCut due to Goemans and Williamson [9] achieves the approximation ratio ≈ 0.878 on an arbitrary graph. Thus, the QAOA with a constant level p cannot outperform classical algorithms. We note that upper bounds on the QAOA approximation ratio were previously known only for $p = 1$ [3]. We refer to Ref. [10] for numerical studies of the QAOA applied to MaxCut.

Similar concerns about limitations of the QAOA have previously been voiced by Hastings [11], who showed analytically that certain local classical algorithms match the performance of the level-1 QAOA for Ising-like cost functions with multispin interactions. Hastings also gave numerical evidence for the same phenomenon for MaxCut with $p = 1$ and argued that this should extend to $p > 1$ [11].

QAOA states possess a certain symmetry that plays a crucial role in our analysis. Namely, the state $|\psi(\beta, \gamma)\rangle$ is invariant under a global spin flip:

$$X^{\otimes n} |\psi(\beta, \gamma)\rangle = |\psi(\beta, \gamma)\rangle.$$

Indeed, the Hamiltonians B and C commute with the symmetry operator $X^{\otimes n}$, while the initial state $|+^n\rangle$ is a $+1$ eigenvector of $X^{\otimes n}$. More generally, let us say that an n-qubit state $|\psi\rangle$ is $\mathbb{Z}_2$ symmetric if it is a $+1$ eigenvector of $X^{\otimes n}$. Our proof of Eq. (3) combines two observations: (i) The symmetry forces good variational states to be highly entangled, and (ii) low-depth circuits are not capable of preparing highly entangled states.

To elaborate on the role of the $\mathbb{Z}_2$ symmetry, suppose $x \in \{0, 1\}^n$ is an optimal cut such that $C_{\max} = C(x)$. Let $\bar{x}$ be the bitwise negation of x. Note that $C(\bar{x}) = C(x)$.

Although the state $|x\rangle$ has no entanglement whatsoever, its $\mathbb{Z}_2$-symmetric version $(|x\rangle + |\bar{x}\rangle)/\sqrt{2}$ is a highly entangled state, locally equivalent to the n-qubit Greenberger-Horne-Zeilinger (GHZ) state $(|0^n\rangle + |1^n\rangle)/\sqrt{2}$, which cannot be prepared by a low-depth circuit [5].

The fact that symmetry may prevent one from preparing ground states of certain Hamiltonians by low-depth circuits is well known in the theory of topological quantum order under the name *symmetry protection* [12–14]. The bound Eq. (3) asserts that the Hamiltonian C exhibits a strong form of symmetry protection that extends to all states with energy density above a certain constant threshold.

We shall now argue that for a suitable family of graphs G all good variational states are qualitatively similar to the GHZ state. Specifically, the results of Refs. [15–17] show that for any constant $D \geq 3$ there exists an infinite family of bipartite degree-D graphs G such that

$$C(x) \equiv |\text{cut}(x)| \geq h \min \{|x|, n - |x|\} \qquad (4)$$

for any $x \in \{0, 1\}^n$, where h is a constant satisfying

$$h \geq \frac{D}{2} - \sqrt{D-1} \qquad (5)$$

and $|x|$ is the Hamming weight of x. Such graphs, known as Ramanujan expander graphs, maximize the spectral gap of their adjacency matrices among all D-regular graphs. Random D-regular bipartite graphs are known to approach the bound Eq. (5) with high probability [18].

Let G be a bipartite graph as above and $x_{\text{opt}} \in \{0, 1\}^n$ be an optimal solution of the MaxCut problem. For a bipartite graph, $C_{\max} = C(x_{\text{opt}}) = |E|$. Moreover, the optimal solution x_{opt} is unique up to the bitwise negation and

$$C(x) + C(x_{\text{opt}} \oplus x) = |E| \qquad (6)$$

for all $x \in \{0, 1\}^n$. Here, $\oplus$ denotes the bitwise XOR. Set $\epsilon = h/6$ and consider a level-p QAOA state such that

$$\langle \psi(\beta, \gamma) | C | \psi(\beta, \gamma) \rangle \geq |E| - \epsilon n. \qquad (7)$$

Let x be a random n-bit string sampled from the distribution $P(x) = |\langle x | \psi(\beta, \gamma) \rangle|^2$. Markov's inequality and Eq. (7) show that $C(x) \geq |E| - 2\epsilon n$ with a probability of at least $1/2$. From Eq. (6), one infers that $C(x_{\text{opt}} \oplus x) \leq 2\epsilon n$ with a probability of at least $1/2$. Let $\text{dist}(x, y) = |x \oplus y|$ be the Hamming distance between bit strings x and y. Combining Eq. (4) and the bound $C(x_{\text{opt}} \oplus x) \leq 2\epsilon n$, one gets

$$\min \{\text{dist}(x_{\text{opt}}, x), \text{dist}(\overline{x_{\text{opt}}}, x)\} \leq \frac{2\epsilon n}{h} = \frac{n}{3} \qquad (8)$$

with a probability of at least $1/2$. This shows that the state $|\psi(\beta, \gamma)\rangle$ has a non-negligible weight on bit strings close to

x_{opt} and on those close to $\overline{x_{\mathrm{opt}}}$. Here, closeness means being within a Hamming distance of at most $n/3$.

Finally, we employ a fascinating result by Eldar and Harrow stated as Corollary 43 in Ref. [19]. It asserts that the output distribution of a low-depth circuit cannot assign a non-negligible weight to subsets of bit strings that are far apart in Hamming distance. Namely, suppose $|\psi\rangle$ is an n-qubit state that can be prepared starting from a product state by a depth-d quantum circuit composed of one- and two-qubit gates. The state $|\psi\rangle$ does not have to be symmetric in any sense. Define a distribution $P(x) = |\langle x|\psi\rangle|^2$. Given a subset $S \subseteq \{0,1\}^n$, let $P(S) = \sum_{x\in S} P(x)$. Reference [19] showed that for any subsets $S, S' \subseteq \{0,1\}^n$ one has

$$\mathrm{dist}(S, S') \leq \frac{4n^{1/2}2^{3d/2}}{\min\{P(S), P(S')\}}. \tag{9}$$

Here, $\mathrm{dist}(S, S') = \min_{x\in S} \min_{y\in S'} \mathrm{dist}(x, y)$ is the minimum number of bit flips required to get from S to S'. Choose S and S' as the sets of n-bit strings x such that $\mathrm{dist}(x_{\mathrm{opt}}, x) \leq n/3$ and $\mathrm{dist}(x_{\mathrm{opt}}^-, x) \leq n/3$, respectively. Then $\mathrm{dist}(S, S') = n/3$. Choose $|\psi\rangle \equiv |\psi(\beta, \gamma)\rangle$. The $\mathbb{Z}_2$ symmetry of QAOA states gives $P(x) = P(\bar{x})$ and, thus, $P(S) = P(S')$. We have already shown that $P(S \cup S') \geq 1/2$ [see Eq. (8)]; that is, $P(S) = P(S') \geq 1/4$. Combining this and Eq. (9), one arrives at $1 \leq 48n^{-1/2}2^{3d/2}$. This gives a lower bound on the depth d required to approximate the maximum value $C_{\max} = |E|$ within a ratio

$$1 - \frac{\epsilon n}{|E|} = 1 - \frac{h}{3D} \leq \frac{5}{6} + \frac{\sqrt{D-1}}{3D}.$$

Here, we recalled that $\epsilon = h/6$ and $|E| = Dn/2$ and used Eq. (5). In Appendix A in Supplemental Material [20], we show that the level-p QAOA circuit has depth $d \leq p(D+1)$ whenever G is a bipartite degree-D graph. Thus, $1 \leq 48n^{-1/2}2^{3d/2}$ implies $p \geq (1/3 \log_2 n - 4)(D+1)^{-1}$. This concludes the proof of Eq. (3).

The above arguments provide an upper bound on the variational energy for any $\mathbb{Z}_2$-symmetric state generated by a low-depth circuit. One may ask if stronger bounds can be derived by making use of the special structure of the QAOA ansatz. Indeed, so far this structure has been used only in establishing the $\mathbb{Z}_2$ symmetry and expressing the circuit depth d in terms of the degree D and the level p. A notable special feature of the QAOA ansatz is its *geometric locality*—the entangling operators $e^{i\gamma_a C}$ include interactions only between nearest-neighbor qubits with respect to the underlying graph G.

To elucidate implications of the geometric locality, consider a toy model known as the *ring of disagrees* [3]. It describes the MaxCut problem on the cycle graph. The latter has a set of vertices $V = \mathbb{Z}_n = \{0, 1, \dots, n-1\}$. An edge is drawn between any pair of vertices $j, k \in \mathbb{Z}_n$

such that $j = k \pm 1 \bmod n$. Quite recently, Ref. [21] proved that the optimal approximation ratio achieved by the level-p QAOA for the ring of disagrees is bounded above by $(2p+1)/(2p+2)$ for all p and conjectured that this bound is tight. Here, we prove a version of this conjecture for arbitrary $\mathbb{Z}_2$-symmetric geometrically local variational states. To quantify the geometric locality, let $\mathrm{dist}(j, k)$ be the distance between qubits j and k with respect to the cycle graph $\mathbb{Z}_n$. Define the R neighborhood of a qubit j as the set $\{i \in \mathbb{Z}_n : \mathrm{dist}(i, j) \leq R\}$. A unitary U acting on n qubits located at vertices of the cycle graph has range R if the operator $U^\dagger Z_j U$ has support on the R neighborhood of j for any qubit j. For example, the level-p QAOA circuit associated with the ring of disagrees has range $R = p$. A unitary U is said to be $\mathbb{Z}_2$ symmetric if $UX^{\otimes n} = X^{\otimes n}U$. Let C be the MaxCut Hamiltonian Eq. (2) associated with the cycle graph $\mathbb{Z}_n$, where n is even. Note that such a graph has a maximum cut of size n. We show that

$$\frac{1}{n}\langle +^n|U^\dagger CU|+^n\rangle \leq \frac{2R+1/2}{2R+1} \tag{10}$$

for any $\mathbb{Z}_2$-symmetric range-R unitary U with $R < n/4$. This bound is tight whenever n is an even multiple of $2R+1$. Since one can always round n to the nearest even multiple of $2R+1$, the bound Eq. (10) is tight for all n up to corrections $O(1/n)$, assuming that $R = O(1)$.

We shall now prove the upper bound Eq. (10). Let $\bar{X}$ be the operator that applies Pauli X to every second qubit, and let $W = \bar{X}U$. Note that W is a $\mathbb{Z}_2$-symmetric circuit with range R. Since $\bar{X}(Z_jZ_{j+1})\bar{X} = -Z_jZ_{j+1}$ for any qubit j, the MaxCut Hamiltonian Eq. (2) satisfies $C + \bar{X}C\bar{X} = nI$. Taking the expected value of this identity on the state $U|+^n\rangle$, one infers that Eq. (10) holds whenever

$$\frac{1}{n}\langle +^n|W^\dagger CW|+^n\rangle \geq 1 - \frac{2R+1/2}{2R+1} = \frac{1}{2(2R+1)}. \tag{11}$$

Thus, it suffices to prove that Eq. (11) holds for any $\mathbb{Z}_2$-symmetric range-R circuit W. For each $j, k \in \mathbb{Z}_n$ define

$$\epsilon_{j,k} = \frac{1}{2}\langle +^n|W^\dagger(I - Z_jZ_k)W|+^n\rangle.$$

We claim that

$$\epsilon_{j,k} = 1/2 \quad \text{if } \mathrm{dist}(j, k) > 2R. \tag{12}$$

Indeed, $\langle +^n|W^\dagger Z_j W|+^n\rangle = 0$ for any qubit j, since the states $W|+^n\rangle$ and $Z_jW|+^n\rangle$ are eigenvectors of $X^{\otimes n}$ with eigenvalues 1 and -1. Such eigenvectors have to be orthogonal. From $\mathrm{dist}(j, k) > 2R$, one infers that $W^\dagger Z_j W$ and $W^\dagger Z_k W$ have disjoint support. Thus,

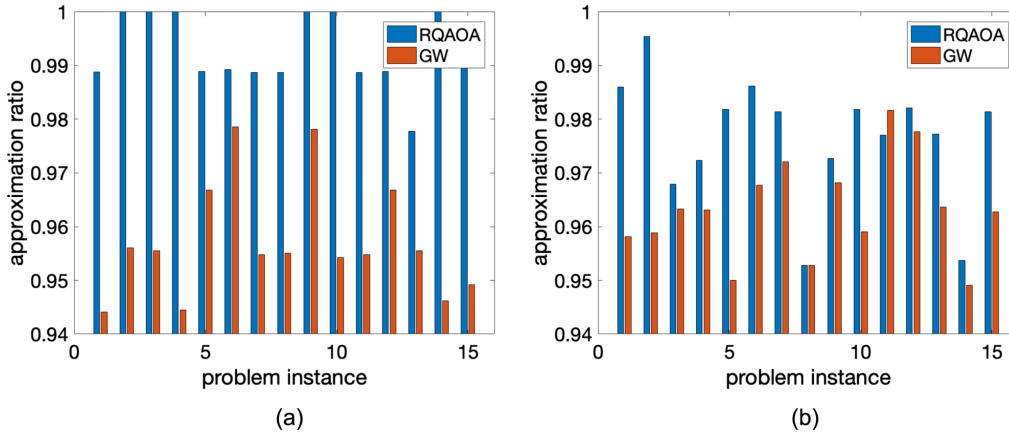(a)   (b)

FIG. 1.   (a) Approximation ratios achieved by the level-1 RQAOA (blue) and the Goemans-Williamson (GW) algorithm [9] (red) for 15 instances of the Ising cost function with random ± 1 couplings defined on the 2D toric grid of size 16×16. In case (b), the Ising Hamiltonian also includes random ± 1 external fields. The RQAOA threshold value is $n_c = 20$. We found that the standard level-1 QAOA achieves approximation ratios below $1/2$ for all considered instances (not shown). The GW algorithm was implemented with $n = 256$ rounding attempts, and the best found solution was selected. The exact maximum energy was computed using integer linear programming.

$$\langle +^n | W^\dagger Z_j Z_k W | +^n \rangle = \langle +^n | (W^\dagger Z_j W)(W^\dagger Z_k W) | +^n \rangle$$
$$= \langle +^n | W^\dagger Z_j W | +^n \rangle \cdot \langle +^n | W^\dagger Z_k W | +^n \rangle$$
$$= 0.$$

This proves Eq. (12).

Suppose one prepares the state $W|+^n\rangle$ and measures a pair of qubits $j < k$ in the standard basis. Then $\epsilon_{j,k}$ is the probability that the measured values on qubits j and k disagree. If qubits j and k disagree, at least one pair of qubits $(i, i+1)$ located in the interval $[j, k]$ must disagree. The probability of the latter event is $\epsilon_{i,i+1}$. Thus, the union bound gives

$$\epsilon_{j,k} \leq \sum_{i=j}^{k-1} \epsilon_{i,i+1}. \tag{13}$$

Set $k = j + 2R + 1$. Then $\epsilon_{j,k} = 1/2$ by Eq. (12). Take the expected value of Eq. (13) with respect to random uniform $j \in \mathbb{Z}_n$. This gives

$$\frac{1}{2} \leq \frac{2R+1}{n} \sum_{i \in \mathbb{Z}_n} \epsilon_{i,i+1} = \frac{2R+1}{n} \langle +^n | W^\dagger C W | +^n \rangle,$$

proving Eq. (11). As argued above, Eq. (11) is equivalent to Eq. (10). In Appendix B in Supplemental Material [20], we show that the bound Eq. (10) is tight by constructing a $\mathbb{Z}_2$-symmetric range-R circuit U such that $U|+^n\rangle$ is a tensor product of GHZ-like states on consecutive segments of $2R + 1$ qubits. This circuit is shown to saturate the upper bound Eq. (10).

Motivated by the limitations established above, we propose a nonlocal modification of QAOA which we call the recursive quantum approximate optimization algorithm

(RQAOA). This is a VQO-type algorithm based on the variational ansatz Eq. (1) with a constant level p. The key new feature of the RQAOA is a variable elimination step. The latter transforms a cost function with n variables to one with $n - 1$ variables by examining correlations present in the optimal variational state and identifying strongly correlated clusters of variables. To describe this formally, suppose the cost function $C(x)$ describes the Ising model on a graph $G = (V, E)$ with n vertices. The corresponding n-qubit Hamiltonian is

$$C = \sum_{(j,k) \in E} J_{j,k} Z_j Z_k. \tag{14}$$

Here, $J_{j,k}$ are arbitrary real coefficients. As before, our goal is to maximize $C(x) = \langle x | C | x \rangle$. The RQAOA consists of the following steps.

First, maximize the expected value $\langle \psi(\beta, \gamma) | C | \psi(\beta, \gamma) \rangle$ over $\beta, \gamma \in \mathbb{R}^p$. For every edge $(j, k) \in E$, compute the mean value $M_{j,k} = \langle \psi(\beta, \gamma) | Z_j Z_k | \psi(\beta, \gamma) \rangle$.

Next, find an edge $(i, j) \in E$ with the largest magnitude of $M_{i,j}$ (breaking ties arbitrarily). The corresponding variables Z_i and Z_j are correlated if $M_{i,j} > 0$ and anti-correlated if $M_{i,j} < 0$. Impose a parity constraint

$$Z_j = \text{sgn}(M_{i,j}) Z_i \tag{15}$$

and substitute it into the cost function C to eliminate the variable j. For example, a term $J_{j,k} Z_j Z_k$ with $k \notin \{i, j\}$ gets mapped to $J_{j,k} \text{sgn}(M_{i,j}) Z_i Z_k$. The term $J_{i,j} Z_i Z_j$ gets mapped to a constant energy shift $J_{i,j} \text{sgn}(M_{i,j})$. All other terms remain unchanged. This yields a new Ising cost function C' that depends on $n - 1$ variables. The underlying interaction graph G' with $n - 1$ vertices is obtained from G

by contracting the edge (i, j) [22]. The maximum energy of C' coincides with the maximum energy of C over the subset of assignments satisfying the constraint Eq. (15).

Next, call the RQAOA recursively to maximize the cost function C'. Each recursion step eliminates one variable from the cost function. The recursion stops when the number of variables reaches some specified threshold value $n_c \ll n$. The remaining instance of the problem with n_c variables is then solved by a purely classical algorithm (for example, by a brute force method). Thus, the value of n_c controls how the workload is distributed between quantum and classical computers.

Finally, assign a value to all eliminated variables Z_j by backtracking the steps of the algorithm and applying the parity constraints Eq. (15) imposed at each step. This results in a tentative solution $x \in \{0, 1\}^n$ of the original problem with n variables.

Importantly, the limitations established above for the QAOA with a constant level p on bounded degree graphs do not apply to its recursive version. Indeed, each variable elimination step performed by the RQAOA results in a contraction of some edge in the graph. The latter tends to increase the maximum vertex degree, thereby increasing the circuit depth of level-p variational states. In other words, the RQAOA overcomes the locality restriction of the standard QAOA without increasing the number of variational parameters that have to be optimized at each step.

We report a numerical comparison between the level-1 RQAOA and the Goemans-Williamson algorithm [9] for the Ising cost function Eq. (14) with random coefficients $J_{j,k} = \pm 1$. Two graphs are considered: (a) the 2D grid and (b) the 2D grid with one extra vertex connected to all grid points. The latter is equivalent to the 2D Ising model with random ± 1 external fields. As shown in Ref. [23], the problem of maximizing the energy $C(x)$ admits an efficient algorithm in case (a), while case (b) is NP hard. To compute the mean values $\langle \psi(\beta, \gamma) | Z_j Z_k | \psi(\beta, \gamma) \rangle$, we used a version of the algorithm by Wang *et al.* [24], as detailed in Appendix C in Supplemental Material [20]. Figure 1 shows approximation ratios achieved by each algorithm for 15 problem instances with the grid size 16×16. It can be seen that the RQAOA outperforms the Goemans-Williamson algorithm for all except for one instance. We found that the standard level-1 QAOA achieves an approximation ratio below $1/2$ for all considered instances. Finally, we show analytically that the RQAOA with the level $p = 1$ finds the optimal solution for the ring of disagrees model; see Appendix D in Supplemental Material [20]. Meanwhile, the standard level-p QAOA achieves an approximation ratio of at most $(2p + 1)/(2p + 2)$ for this model [21]. This proves that in certain cases the RQAOA is strictly more powerful than the QAOA.

Note added.—Recently, analogous limitations were established for random regular graphs by exploiting the locality and spatial uniformity of the QAOA [25,26]. We focus on $\mathbb{Z}_2$ symmetry and locality, and our statements also apply to nonuniform local algorithms.

[1] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Nat. Commun. **5**, 4213 (2014).

[2] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature (London) **549**, 242 (2017).

[3] E. Farhi, J. Goldstone, and S. Gutmann, arXiv:1411.4028.

[4] A. Y. Kitaev, Ann. Phys. (Amsterdam) **303**, 2 (2003).

[5] S. Bravyi, M. B. Hastings, and F. Verstraete, Phys. Rev. Lett. **97**, 050401 (2006).

[6] D. Aharonov and Y. Touati, arXiv:1810.03912.

[7] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Nat. Commun. **9**, 4812 (2018).

[8] E. Farhi and A. W. Harrow, arXiv:1602.07674.

[9] M. X. Goemans and D. P. Williamson, J. ACM **42**, 1115 (1995).

[10] R. Shaydulin and Y. Alexeev, *Evaluating quantum approximate optimization algorithm: A case study,* in *2019 Tenth International Green and Sustainable Computing Conference (IGSC)* (IEEE, New York, 2019), pp. 1–6.

[11] M. B. Hastings, arXiv:1905.07047.

[12] Z.-C. Gu and X.-G. Wen, Phys. Rev. B **80**, 155131 (2009).

[13] N. Schuch, D. Pérez-García, and I. Cirac, Phys. Rev. B **84**, 165139 (2011).

[14] F. Pollmann, A. M. Turner, E. Berg, and M. Oshikawa, Phys. Rev. B **81**, 064439 (2010).

[15] A. W. Marcus, D. A. Spielman, and N. Srivastava, Interlacing families iv: Bipartite Ramanujan graphs of all sizes, in *Proceedings of the 2015 IEEE 56th Annual Symposium on Foundations of Computer Science* (IEEE, New York, 2015), pp. 1358–1377.

[16] A. W. Marcus, D. A. Spielman, and N. Srivastava, Ann. Math. **182**, 307 (2015).

[17] M. Morgenstern, J. Comb. Theory Ser. B **62**, 44 (1994).

[18] J. Friedman, arXiv:cs/0405020.

[19] L. Eldar and A. W. Harrow, Local Hamiltonians whose ground states are hard to approximate, in *Proceedings of the 2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, New York, 2017), pp. 427–438.

[20] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.125.260505 for further details on state preparation and implementation.

[21] G. Mbeng, R. Fazio, and G. Santoro, arXiv:1906.08948.

[22] Recall that contraction of an edge (i, j) in a graph G is an operation that removes the edge (i, j) from G and identifies the vertices i and j.

[23] F. Barahona, J. Phys. A **15**, 3241 (1982).

[24] Z. Wang, S. Hadfield, Z. Jiang, and E. G. Rieffel, Phys. Rev. A **97**, 022304 (2018).

[25] E. Farhi, D. Gamarnik, and S. Gutmann, arXiv:2004.09002.

[26] E. Farhi, D. Gamarnik, and S. Gutmann, arXiv:2005.08747.